

Review Type III CRISPR-Cas Immunity: Major Differences Brushed Aside

Gintautas Tamulaitis,¹ Česlovas Venclovas,¹ and Virginijus Siksnys^{1,*}

For a long time the mechanism of immunity provided by the Type III CRISPR-Cas systems appeared to be inconsistent: the Type III-A Csm complex of *Staphylococcus epidermidis* was first reported to target DNA while Type III-B Cmr complexes were shown to target RNA. This long-standing conundrum has now been resolved by finding that the Type III CRISPR-Cas systems are both RNases and target RNA-activated DNA nucleases. The immunity is achieved by coupling binding and cleavage of RNA transcripts to the degradation of invading DNA. The base-pairing potential between the target RNA and the CRISPR RNA (crRNA) 5'-handle seems to play an important role in discriminating self and non-self nucleic acids; however, the detailed mechanism remains to be uncovered.

CRISPR-Cas Immunity

Clustered regularly interspaced short palindromic repeats (CRISPR; see Glossary) together with **cas** (CRISPR-associated) genes provide adaptive immunity against viruses and plasmids in about 50% of bacteria and most of the archaea [1]. CRISPR-Cas systems highjack fragments of foreign DNA and insert them into the CRISPR array in the host genome to memorize the invader. An insert ('spacer') is then used to derive a **CRISPR RNA** (**crRNA**) molecule which, together with Cas proteins, is assembled into a ribonucleoprotein complex. During reinfection, immunity is provided by crRNA-guided recognition and degradation of invading nucleic acids by the CRISPR-Cas complex [2,3]. CRISPR-Cas systems are remarkably diverse. Depending on the composition of crRNA-**effector complexes** they are divided into two broad classes [4–6] (Table 1). Class 1 comprises multisubunit effector complexes, whereas Class 2 effector complexes consist of a single protein. The two classes of CRISPR-Cas systems are further subdivided into six types and multiple subtypes.

Class 1 systems include Type I, in which interference is achieved through the coordinated action of the multisubunit **Cascade** complex and the Cas3 nuclease/helicase [7–12], Type III, relying on either the **Csm** (Type III-A/D) or **Cmr** (Type III-B/C) effector complexes [13–18], both distantly related to the Cascade [19,20] and putative Type IV [4]. All Type III systems possess the signature gene *cas10*, which encodes a multidomain protein that is also the largest subunit of Type III crRNA–effector complexes. Type III-A and III-B systems are distinguished by the presence of distinct genes encoding **small subunits**, Csm2 (Type III-A) and Cmr5 (Type III-B). Moreover, subtype III-A loci usually contain *cas1*, *cas2*, and *cas6* genes, whereas most III-B loci lack these genes and therefore likely depend on other CRISPR–Cas systems present in the genome. The distinctive feature of subtype III-C is the apparent inactivation of the Palm domain of Cas10. Subtype III-D loci typically encode a Cas10 protein that lacks the HD domain [4].

Trends

Immunity against viruses and plasmids provided by CRISPR-Cas systems is mediated by a ribonucleoprotein effector complex. Csm (Type III-A) and Cmr (Type III-B) complexes function as RNA-activated single-stranded (ss) DNases that couple the target RNA binding/cleavage with ssDNA degradation.

Upon foreign DNA infection, the CRISPR RNA (crRNA)-guided binding of the Csm or Cmr complex to the emerging transcript recruits Cas10 DNase to the actively transcribed phage DNA, resulting in degradation of both the transcript and phage DNA, but not the host DNA.

The Cas10 HD-domain is responsible for the ssDNase activity, and Csm3/ Cmr4 subunits are responsible for the endoribonuclease activity of the Csm/ Cmr complex.

The 3'-flanking sequence of the target RNA is critical for the ssDNase activity of Csm/Cmr: the basepairing with the 5'-handle of crRNA protects host DNA from degradation.

¹Institute of Biotechnology, Vilnius University, Saulėtekio av. 7, Vilnius 10257, Lithuania

*Correspondence: siksnys@ibt.lt (V. Siksnys).

Table 1. Classification of CRISPR-Cas Systems^a

Class	Class 1 Multi-subunit	crRNA-effector com	plex	Class 2 Single-subunit crRNA-effector complex		
Туре	Туре І	Type III	Type IV	Type II	Type V	Type VI
Effector complex	Cascade	Csm and Cmr	n.d.	Cas9	Cpf1, C2c1, C2c3	C2c2
Target	dsDNA	ssRNA/ ssDNA	n.d.	dsDNA	dsDNA	ssRNA

^aAbbreviations: n.d., no data available; ss, single stranded; ds, double stranded.

Class 2 CRISPR systems comprise Type II, characterized by the single effector protein **Cas9**, containing RuvC and HNH nuclease domains [21–25], Type V systems which utilize single RuvC domain-containing effectors such as Cpf1 [26–29], C2c1, and C2c3 [6], and Type VI which relies on C2c2 protein, containing two HEPN nuclease domains [5].

So far, functionally characterized CRISPR-Cas systems, with the exception of Type III and recently discovered Type VI [5], have been reported to target double-stranded DNA (dsDNA). The type of nucleic acids targeted by Csm (Type III-A) and Cmr (Type III-B) effector complexes for a long time presented a puzzle. The Csm complex of S. epidermidis (SeCsm) was first reported to target DNA in vivo [30,31]. By contrast, Cmr complexes in vitro have been shown to degrade single-stranded (ss) RNA [17,18,32]. However, the idea that Csm targets DNA and Cmr targets RNA had to be gradually abandoned. The initial contribution to this change came from studies of the Sulfolobus islandicus Type III-B CRISPR-Cas system. The elimination of protospacercarrying plasmids by this system was shown to depend on directional transcription of the protospacer, prompting the proposal of transcription-dependent DNA targeting [33]. The concept of DNA targeting by the Csm complex has also evolved significantly. At first, it was thought that Type III-A CRISPR interference acts at the DNA level and that the transcription of protospacer is not required [30]. However, subsequent studies have shown the opposite - that the DNA targeting is transcription-dependent [34]. Further experiments revealed that DNA targeting by the SeCsm complex involves cleavage of both target DNA and its transcript, but at the time it was thought that these cleavage events are independent [35]. However, these two events were linked together when a common molecular mechanism of the target RNAactivated DNA interference by the Csm and Cmr effector complexes was established by three independent recent studies [36-38]. They found that the Csm/Cmr complexes, guided by crRNAs, bind target RNA to trigger two distinct enzymatic activities: (i) sequence-specific singlestranded ribonuclease (ssRNase) activity directed against the bound target RNA, and (ii) singlestranded deoxyribonuclease (ssDNase) activity, directed against the ssDNA and allosterically controlled by the target RNA binding. These findings have finally reconciled different functional activities of Type III-A and III-B CRISPR-Cas complexes.

The discovered intricate mechanism of Type III CRISPR-Cas systems is in sharp contrast to the Type I and II systems that rely strictly on DNA sequence recognition [7–9,22,24,39]. The RNA/ DNA targeting capabilities of Type III CRISPR-Cas systems which are widely spread in both bacteria and archaea (respectively, in 34% and 25% of the genomes encoding CRISPR-*cas* loci [1]) may provide a versatile immune response against many different viruses, plasmids, and other mobile genetic elements. By contrast, transcription-dependent DNA degradation provides an efficient fail-safe mechanism for degradation of both mRNA and transcriptionally coupled DNA of invading genetic elements while maintaining host genome integrity. This review attempts to summarize the current knowledge and to discuss open questions related to the nucleic acid interference by the effector complexes of Type III-A and Type III-B CRISPR-Cas systems.

Glossary

5'-'handle': 8 nt stretch derived from the CRISPR repeat and located at the 5'-end of the mature crRNA in the Type III-A (Csm) and Type III-B (Cmr) effector complexes. cas: CRISPR-associated genes which are located in the vicinity of the CRISPR array and are necessary for the CRISPR-Cas function. Cas9: an RNA-guided DNA endonuclease of the Type II CRISPR-Cas systems. It is comprised of a single Cas9 protein subunit bound to a dual crRNA tracrRNA molecule Cascade: a CRISPR-associated complex for antiviral defense of the Type I CRISPR-Cas systems. It is comprised of multiple Cas subunits and crRNA

Clustered regularly interspaced short palindromic repeat

(CRISPR) array: an array of short conserved repeat sequences interspersed between unique DNA sequences of similar size called spacers which often originate from phage or plasmid DNA. CRISPR array together with cas genes form the CRISPR–Cas system, which functions as an adaptive immune system in prokaryotes.

CRISPR RNA (crRNA): a small RNA molecule generated by transcription and processing of the CRISPR array. crRNA is composed of a conserved repeat fragment(s) and a spacer sequence, which guides the Cas protein(s) to the cognate invading nucleic acids for their destruction. Csm and Cmr (complexes): Type III CRISPR-associated multisubunit

effector complexes for antiviral defense. These complexes recognize and destroy phage RNA transcripts and phage DNA.

Effector complex: a

ribonucleoprotein complex containing crRNA bound to Cas (CRISPR associated) proteins. Guided by crRNA, the effector complex locates the nucleic acid target and triggers its degradation.

Ferredoxin-like fold: most Csm and Cmr proteins have a structurally similar ferredoxin fold core with different variations of the RNA recognition motif (RRM).

GGDD motif: the conserved GGDD active site motif present in the Cas10 Palm domain, which shares structural similarity with catalytic domains of DNA polymerases and nucleotide cyclases.



HD-domain: HD (histidine-aspartate) nuclease domain, found in Cas3 and Cas10 proteins. The HD-domain proteins represent a large group of enzymes which catalyse phosphomonoesterase or phosphodiesterase reactions within a broad range of substrates including nucleotides and nucleic acids and function primarily in nucleic acid metabolism and signal transduction. Large subunit: Cas10 protein representing the largest subunit of Type III systems, also known as Csm1 in Type III-A and Cmr2 in Type III-B.

Protospacer: sequence region in the target DNA or RNA molecule which is complementary to the crRNA spacer region of the effector complex.

Protospacer adjacent motif

(PAM): a conserved nucleotide motif that is found next to the target DNA sequences (protospacers). It is obligatory for the activity of Type I and II effector complexes. Small subunit: the smallest mostly α -helical subunit of Type III systems named Csm2 in Type III-A and Cmr5 in Type III-B systems.

Figure 1. Structural Arrangement of the Type III Complexes. (A) Schematic architecture of Csm (Type III-A) and Cmr (Type III-B) complexes. Homologous subunits are depicted by the same color. The 5'-handle of crRNA in the Csm and Cmr complexes is shown in red. (B) Structural model of the StCsm complex and the crystal structure of the chimeric Cmr complex lacking the HD-domain. The chimeric Cmr complex is fitted into the cryo-EM reconstruction of the intact PfCmr

complex bound to 45-mer crRNA (EMD-5740) [32,40]. The cryo-EM volume is represented by a black mesh.

Composition and Structure of Csm and Cmr Complexes

Csm and Cmr are ribonucleoprotein complexes composed of multiple Cas subunits and bound crRNA (Figure 1A). crRNA consists of a spacer region and a repeat-derived **5**'-**'handle'**, typically eight nucleotides long. Csm and Cmr complexes share similar architecture of two intertwined helical protein filaments with a **large subunit**, represented by the Cas10 family protein (Csm1 or Cmr2), at the base and another subunit capping the head (Csm5 or Cmr1/Cmr6) (Figure 2 and Table 2). The major filament, or the backbone, in both complexes is formed by a single copy of Csm4 or Cmr3 and multiple copies of Csm3 or Cmr4, a ribonuclease component of the corresponding complexes. The minor filament is formed by small subunits (Csm2 or Cmr5) together with the C-terminal domain of the large subunit.

The major helical filament is responsible for crRNA binding whereas minor filament mostly contributes to the binding of target strand. In different Type III systems the number of backbone and small subunits varies depending on the length of the crRNA, implying that it participates in the templated assembly of the effector complex core [32,40] (Figure 1B).

Trends in Microbiology

CellPress



Trends in Microbiology

Figure 2. Domain Organization of the PfCmr2 (PfCas10) Subunit. PfCmr2 lacking the HD-domain is shown together with Cmr3, crRNA, and a target analog as part of the chimeric Cmr complex (Figure 1B). The PfCmr2 structure is represented as solvent-excluded surface with individual domains denoted using different colors. Horizontal bar shows the PfCmr2 domain arrangement in sequence using the same coloring as in the structure. HD, HD-domain absent in the structure; Palm 1 and Palm 2, polymerase/nucleotide cyclase-like Palm domains; Zn, zinc-finger motif in Palm 1; D2, a small domain inserted between Palm domains; D4, the α-helical C-terminal domain; GGDD, GGDD motif.

Table 2. Components of Type III-A (Csm) and Type III-B (Cmr) Complexes

Large subunit	Csm1	Cmr2	Cas10 family/palm and HD domains
Small subunit	Csm2	Cmr5	small \propto -helical protein
Backbone	Csm3	Cmr4	Cas7 family/ferredoxin-like fold
	Csm4	Cmr3	Cas5 family/ferredoxin-like fold
Capping subunit	Csm5	Cmr1 and Cmr6	Cas7 family/ferredoxin-like fold

The large subunit (Cas10) is a signature protein of Type III systems (Figure 2). Cas10 contains two **ferredoxin-like fold** domains, homologous to the Palm domain of nucleic acid polymerases and nucleotide cyclases. The first Palm domain appears to be catalytically inactive, whereas the second one has a putative active site featuring the conserved **GGDD-motif** [41–44] (Figures 2 and 3). Cas10 is typically fused to the N-terminal HD (histidine-aspartate) nuclease domain, which features a circular permutation of the conserved motifs compared to the **HD-domain** of Cas3 in Type I CRISPR–Cas systems [1,45–47] (Figures 2 and 3). In addition, Cas10 has a C-terminal \propto -helical domain (D4) similar to the small subunits of Type III systems [44,48], a zinc-finger motif within the first Palm domain and a small structurally variable domain (D2) inserted between the Palm domains [41,46] (Figure 2).

The first atomic view of any Type III complex was provided by the crystal structure of the chimeric Cmr complex [19], which also allowed insight into the structural and functional features of a related Csm complex [38]. In general, the homologous relationship between most Csm and Cmr subunits was established previously from sequence-based studies [49]. Only the relationship for the corresponding small subunits, Csm2 and Cmr5, has remained uncertain. Remarkably, even with both structures at hand, their similarity was initially missed due to structure-swapping in a Csm2 dimer [50]. Only subsequent analysis revealed that



Trends in Microbiology

Figure 3. Comparison of Csm and Cmr Active Sites. Superimposed active sites of the HD-domain (left), the GGDD motif (center) and the ribonuclease (right) from the StCsm structural model and the chimeric Cmr structure (Figure 1B). StCsm active site residues are shown in gray with black labels. Cmr active site residues and their labels are shown in color. The HD-domain of Csm1 is shown complexed with a ssDNA fragment. Two metal ions bound in the HD active site are shown as orange and magenta balls for Csm1 and Cmr2, respectively. The GGDD nucleotide-binding site in the Palm domain of Csm1 is shown complexed with AMP. Active site motifs of Csm3 and Cmr4 are shown in the context of bound crRNA (orange) and target RNA (yellow); the scissile phosphate of target RNA is indicated by the red color.

Csm2 and Cmr5 display structural and functional similarity, implying that they have evolved from a common ancestor [48]. As a result of this finding, the homology can now be extended to all of the corresponding subunits in Csm and Cmr complexes, further supporting their structural and functional relatedness.

Functional Activities of the Csm and Cmr Complexes: Ribonuclease Activity

Type III CRISPR-Cas systems were initially believed to target either DNA (Type III-A) or RNA (Type III-B) [49]. Early in vivo genetic experiments suggested that Csm complexes target DNA. This evidence emerged from the observation that the Type III-A CRISPR-Cas system of S. epidermidis limits plasmid conjugation and transformation in vivo, although DNA degradation had not been demonstrated directly [30,31]. By contrast, early in vitro studies revealed that in the Type III-B systems, the Cmr complex guided by crRNA triggers the fragmentation of target RNA molecules [17,18,32] (Table 3). The RNAse component of the Cmr effector complex however was not identified. Later biochemical studies revealed that the Streptococcus thermophilus Csm (StCsm) complex binds the ssRNA target with high affinity and cuts it in a protospacer adjacent motif (PAM)-independent manner in the presence of Me²⁺ ions, producing a regular 6 nt cleavage pattern in the protospacer region [15]. The conserved D33 residue in S. thermophilus Csm3 was suggested to constitute the active site of the Csm complex [15]. Subsequently, a similar ssRNA cleavage pattern was demonstrated in vitro for other Csm complexes, namely, from Thermus thermophilus (TtCsm) and S. epidermidis [35,51] as well as for the Cmr complexes from Pyrococcus furiosus (PfCmr), Sulfolobus solfataricus (SsCmr), Thermotoga maritima (TmCmr), and T. thermophilus (TtCmr) [16,19,32,37,45,52–54] (Table 3). Csm3 and Cmr4 subunits, present in multiple copies within respective Csm and Cmr complexes, act as endoribonucleases [15,45,53,54] and account for multiple periodic cleavage sites on target RNA. The crystal structure of a chimeric Cmr complex bound to crRNA and a target analog revealed the structural basis for the periodic cleavage. The insertion of a conserved Cmr4 β-hairpin ('thumb') into the crRNA-target duplex causes nucleotide flipping from the duplex with the 6 nt periodicity [19]. In addition, distortion of the backbone likely helps position the target RNA scissile phosphates directly between the 2' hydroxyl of the adjacent ribose and the strictly conserved aspartate, which is critical for RNA cleavage [19]. RNA cleavage by both Csm and Cmr complexes was also confirmed by in vivo experiments [15,55,56].

	Organism	DNA interference		RNA interference		Target RNA-activated DNA degradation	
	0.94.10/11	in vivo	in vitro		in vitro	in vivo	in vitro
		111 0100		11 1100			
Csm (Type III-A)	Staphylococcus aureus					+ [68]	
	Staphylococcus epidermidis	+ [30,31]			+ [35]	+ [34]	+ [35]
	Streptococcus thermophilus			+ [15]	+ [15]		+ [38]
	Thermus thermophilus				+ [51]		
Cmr (Type III-B)	Pyrococcus furiosus				+ [17,19,45, 52–54,69]	+ [36]	+ [36]
	Sulfolobus islandicus			+ [55]		+ [33]	
	Sulfolobus solfataricus			+ [56]	+ [16,18]		
	Thermotoga maritima				+ [37]		+ [37]
	Thermus thermophilus				+ [32]		

Table 3. Characterized Type III-A and Type III-B Complexes

Functional Activities of the Csm and Cmr Complexes: DNAse Activity

The HD-domain of nuclease-helicase Cas3 is responsible for the ssDNA cleavage in Type I CRISPR-Cas systems [9,47,57,58]. Therefore, it could be expected that, in Type III complexes, the Cas10 HD-domain provides the same function. Indeed, several Cas10 proteins display weak ssDNase activity that is impaired by mutations in the HD-domain [36–38,46]. However, in the case of SeCsm1, the GGDD-motif of the Palm domain rather than the HD-domain was implicated in ssDNA cleavage [59,60].

HD-domains of purified StCsm and PfCmr complexes also degrade ssDNA in sequence nonspecific manner [36,38], whereas TmCmr cleaves ssDNA predominantly after every thymidine [37]. Interestingly, the HD-domain of *S. thermophilus* Cas3 displays a similar cleavage preference [8]. There are also reports regarding the HD nuclease activity that appear to be contradictory. Thus, in one recent study, the PfCmr complex was observed to degrade not only ssDNA, but also dsDNA [36]. At the same time the isolated large subunit (PfCmr2) of the same complex cleaved ssDNA but not dsDNA [36].

Based on transformation efficiencies, a transcription-dependent DNA interference mechanism has been proposed for the Type III-B CRISPR-Cmr systems in *S. islandicus* [33]. It has been also reported that the conditional tolerance of temperate phages in *S. epidermidis* is achieved through transcription-dependent DNA interference by the Type III-A CRISPR-Csm system [34]. Further genetic experiments in *S. epidermidis* revealed that transcription is required for antiplasmid immunity [35]. Samai *et al.* provided an indirect support for this mechanism, demonstrating that transcription across the target results in the dual cleavage of the target DNA and the RNA transcript by SeCsm *in vitro* [35]. The authors concluded that the crRNA-guided SeCsm independently targets either ssRNA or ssDNA through base-pairing between crRNA and target molecules [35]. Taken together, available data indicated that effector complexes of III-A and III-B systems degrade DNA targets, but the mechanism of degradation



remained obscure. Recently, three independent groups resolved this conundrum by proposing a mechanism of the target RNA-activated DNA degradation for StCsm, PfCmr, and TmCmr effector complexes [36–38]. This common mechanism finally unified Type III-A and III-B CRISPR-Cas subtypes.

Unified Mechanism for Csm/Cmr Complexes

In vitro experiments convincingly demonstrated that the binding of target RNA by either Csm or Cmr complex induces the DNase activity of Cas10 subunit (Csm1 or Cmr2), which then hydrolyses the ssDNA nonspecifically [36-38]. An important question is: how is this nuclease activity kept in check so that it does not degrade the host's own genome? Apparently, there are both spatial and temporal control mechanisms that keep Type III effector complexes tightly regulated (Box 1). When foreign DNA, such as phage DNA, enters the cell, transcription is initiated to establish phage infection. If transcription occurs through the region containing a protospacer, the nascent phage mRNA emerging from the transcription complex is targeted by the effector complex. The binding of crRNA to the complementary protospacer region in nascent mRNA will tether the complex to the transcribed DNA for in cis degradation of ssDNA available as part of the transcription bubble. In other words, the transcript serves as a 'leash' to keep the effector complex in the vicinity of RNA polymerase and to ensure that only the transcribed DNA, in this case phage DNA, will be cleaved. A number of observations are consistent with such spatial control of the Cas10 nuclease activity. In DNA cleavage assays designed to mimic the in vivo situation by including the RNA polymerase, it was observed that only the nontemplate DNA strand is cleaved [35]. This is in line with the observed susceptibility of the nontemplate strand to the nuclease treatment in the transcription elongation complex [61]. It is also consistent with the structure of the RNA polymerase transcription initiation complex, in which only a fragment of the nontemplate DNA strand within the transcription bubble is exposed [38,62]. In addition, this model agrees well with the in vitro experiments showing that the effector complex is able to cleave ssDNA even within short mismatched 'bubbled' regions of dsDNA [37,38].

However, whether tethered to the RNA polymerase or not, the long-lived RNA-activated complex would be extremely dangerous to the host cell. The activated complex potentially could cleave any ssDNA, for example DNA replication intermediates, causing damage to the host genome. It turns out that, in addition to the confinement in space, the DNase activity of Csm and Cmr complexes is under tight temporal control. Normally, the nuclease activity of the Cas10

Box 1. Mechanism of the Transcript-Activated Degradation of Foreign DNA by Type III CRISPR-Cas Systems

CRISPR refers to genomic loci consisting of DNA repeats interspaced by invader-derived DNA sequences termed 'spacers' that serve as memories of past invasions. CRISPR loci are often flanked by CRISPR-associated (*cas*) genes which encode the protein machinery of this adaptive immune system. CRISPR-Cas systems are subdivided into six major Types (I–VI) that differ by nucleic acid targets, the number and arrangement of *cas* genes, and the composition of silencing complexes [4–6].

In the Type III systems multiple Cas proteins and crRNA assemble into Csm (Type III-A) or Cmr (Type III-B) silencing complexes. Csm and Cmr function as target RNAactivated single-stranded (ss) DNases that couple the target RNA binding/cleavage with ssDNA degradation. The RNase and DNase activities of Csm/Cmr complexes are coordinated in space and time to ensure the destruction of foreign genetic elements while preventing the degradation of the host's own DNA [36–38]. When foreign DNA enters the cell, transcription is initiated. If transcription occurs through the region containing a protospacer, the nascent phage mRNA emerging from the transcription complex becomes a target of the Csm/Cmr complex. The RNA transcript binding by the Csm/Cmr complex activates the Cas10 protein for *in cis* degradation of the nontemplate ssDNA within the transcription bubble. Fast RNA cleavage by the Csm3/Cmr4 subunit in the Csm/Cmr complex provides a temporal control by suppressing the DNase activity of Cas10 and thereby preventing potential damage to the host genome.

CRISPR-Cas systems must destroy invading foreign nucleic acids, but prevent targeting the host's own DNA (Figure I). In the DNA-targeting Type I and II systems, PAM plays a key role in discriminating between self and non-self DNA. The discrimination between self and non-self DNA by Type III Csm and Cmr effector complexes is indirect as it depends on the 3'-flanking sequence of the target RNA. In the case of an anti-crRNA transcript emerging, for example, due to the bidirectional transcription of the CRISPR array, the transcript is complementary to both the spacer region and the 5'-handle of crRNA. The base-pairing with the crRNA 5'-handle represses the Cas10 ssDNase activity, thus protecting the host DNA. In the case of phage infection, the RNA transcript base-pairs only with the spacer region of crRNA, but not with its 5'-handle. This activates the Cas10 ssDNase leading to the phage DNA degradation. However, the features of the target RNA 3'-flanking sequence determining the balance between DNA degradation and protection are still poorly defined.



Figure I. Mechanism for Nucleic Acid Interference by the Csm (Type III-A) and Cmr (Type III-B) Complexes. Csm/Cmr complex assembly: the assembly stage of effector complex involves transcription of the CRISPR array into a long precursor CRISPR RNA (pre-crRNA). The latter is cleaved at the repeats by the endoribonuclease Cas6, encoded in the vicinity of the CRISPR array. The processed crRNA is bound by Csm or Cmr proteins to form a Csm/Cmr complex. The crRNA in the Csm/Cmr complex is further truncated by unknown nuclease. The mature crRNA guides the complex in the recognition of an invader transcript at the interference stage. Autoimunity avoidance: in the case of bidirectional transcription through the CRISPR array, anti-precrRNA (gray) transcribed from a putative promoter on the template strand will bind to the Csm/Cmr complex due to the complementarity to crRNA (orange). However, the complementarity between the crRNA 5'-handle (blue) and the 3'- flanking sequence (green) will repress the ssDNase activity, thus protecting the host DNA from cleavage. Immunity against invading DNA: in the case of phage infection, the RNA is transcribed, and crRNA (orange) will basepair to the protospacer region (magenta) but not to the 3'-flanking sequence of the protospacer. This will activate phage DNA degradation by the HD-domain (red) of Cas10.

HD-domain within the complex is attenuated. Only when the activating RNA is bound is the nuclease activity of the HD-domain switched on. However, the activating RNA is rapidly cleaved into 6 nt fragments that are released into solution [37,38]. This initiates a relaxation of the HD-domain into the inactive state. In this way the host cell is protected from the uncontrolled DNase activity that otherwise might be deleterious for the cell. Interestingly, only the binding but not the cleavage of the activating RNA is sufficient to unleash the HD-domain DNase. Moreover, if the bound activator RNA is not cleaved, the effector complex is locked in the nuclease-active state [37,38]. Consistent with this *in vitro* observation, the absence of target RNA cleavage by the SeCsm complex *in vivo* resulted in a hyperactive DNA-silencing phenotype [35]. Thus, tethering of the Csm/Cmr complex to the transcript confines the DNase activity within a limited space near nascent mRNA, while fast RNA degradation ensures a temporal regulation of DNase activity, avoiding the potential cleavage of the host genome.

In addition to the nuclease active site in the HD-domain, Cas10 also has an active site represented by the GGDD-motif in the Palm domain (Figure 3). Whereas by now there is consensus that the nuclease activity of the HD-domain is responsible for ssDNA cleavage



in vitro, the role of the GGDD-motif remains uncertain. There is also an unresolved question as to how these two active sites cooperate in silencing the invading DNA *in vivo*. *In vitro* experiments revealed that mutations of the Cas10 GGDD-motif do not significantly affect the DNA cleavage either by the isolated Cas10 or by the entire complex [36,38]. Interestingly, individual mutations of either the HD-domain or Palm domain of PfCmr2 did not disrupt plasmid silencing *in vivo* [36]. However, simultaneous mutations of both PfCmr2 domains did, suggesting that they both contribute in some way [36]. Therefore, the molecular function of the Palm domain and the nature of its cooperation with the HD-domain in CRISPR-Cas immunity provided by Type III systems remain important unanswered questions.

Avoiding Autoimmunity in Type III Systems

Complementarity to the spacer region of crRNA specifies invading nucleic acids targets. At the same time, to avoid autoimmunity, the CRISPR array has to be exempt from self-targeting. The main difference between the invading nucleic acid targets (protospacers) and spacers within the CRISPR array are the flanking sequences. To distinguish non-self from self DNA, Type I and II CRISPR-Cas systems rely on short (2–4 nt) PAMs [63–66]. PAMs are present in invading sequences but absent in CRISPR array repeats. By contrast, for Type III systems it was proposed that the base-pairing potential between the 5'-handle of crRNA and the 3'-flanking region of the DNA target, rather than PAM recognition, is responsible for distinguishing non-self from self [31]. According to this proposal, the noncomplementary 3'-flanking region of the target DNA specifies an invading sequence, whereas complementarity to the 5'-handle of crRNA indicates that the sequence derives from the CRISPR array. Recently, however, this view has been revised as it was shown that the 3'-flanking sequence of target RNA, not DNA, is involved in distinguishing self from non-self [36–38]. In other words, the 3'-flanking sequence of the transcript determines whether the DNase activity of Cas10 will be switched on or off.

Csm and Cmr complexes bind crRNAs that typically have the 8-nucleotide long 5'-handle followed by the spacer region (Figure 4A,B). Both the crystal structure of Cmr bound to the target analog and the Csm structural model show that not all positions in the handle are available for base-pairing [19,38] (Figure 4A,B). The first three nucleotides are bound into distinct pockets of Cmr3, whereas the 8th nucleotide is flipped-out like every 6th nucleotide within the spacer region of crRNA. These observations led to the suggestion that the complementary 3'-flanking sequence of target RNA could pair only with the four crRNA positions (4th to 7th) similarly as in the spacer region [38] (Figure 4B). Experimental data show that the full complementarity between the crRNA handle and the 3'-flank of the target RNA is indeed not necessary for protection. In different Type III systems the complementarity to the three (5–7th) [31,38] or even two (6–7th) (Figure S2 in [36]) crRNA positions was able to suppress DNA degradation. This agrees well with the finding that the introduction of two consecutive mismatches within the 5–7th positions of crRNA is sufficient to lose the protection of the CRISPR-Cas locus [31]. Collectively, these data indicate that the base-pairing potential at the three crRNA handle positions (5–7th) defines the balance between DNA degradation and protection.

The solved crystal structure of the Cmr complex is bound to the target analog totally lacking 3'-flanking sequence. Despite that, the route of the complementary 3'-flanking sequence can be inferred based on its expected pairing with the crRNA handle (Figure 4A,B) [38]. By contrast, the noncomplementary 3'-flanking sequence unable to pair with the crRNA handle is unlikely to fit into the same confined space and therefore is expected to take a different route. Although speculative, the most suggestive route is the deep groove between the zinc-finger and the D2 domain of the PfCmr2 subunit (Figure 4A).

The model whereby the complementarity between the 5'-handle of crRNA and the 3'-flank of the RNA target is the only factor determining the fate of target DNA has been challenged by a recent





Figure 4. Proposed Self Versus Non-Self Discrimination in the Type III CRISPR-Cas systems. (A) Possible alternative routes (yellow arrows) of the 3'-flanking region of target RNA on the PfCmr2 molecular surface. PfCmr2 surface is colored according to the electrostatic potential (blue, positive; red, negative). The 3'-flanking sequence of target RNA complementary to the 5'-handle of crRNA due to base-pairing is directed into the groove along the Cmr2–Cmr3 interface. The noncomplementary flanking sequence is prevented from entering this route and instead is likely directed through the deep groove formed by the zinc-finger and D2 domain of Cmr2. (B) A schematic representation of self target RNA recognition by crRNA, indicating that only four positions (4th–7th) of the crRNA handle (see panel A) are available for base-pairing. The scheme is adopted from [38] with permission. (C) Scheme illustrating the dependency between the route of the target flanking sequence of target RNA 5'-handle does not activate the HD-domain. Center panel: the alternative route switches on the ssDNAse activity. Right panel: the absence of the 3'-flanking sequence has a different outcome in different CRISPR-Cas systems – in StCsm the HD-domain is repressed, while in PfCmr and TmCmr the HD-domain is active. Abbreviation: N.D., no data available.

study [36]. The authors tested all 64 triplets immediately downstream of the protospacer region of the target RNA. Unfortunately, the experimental setup implied testing the base-pairing with only two (6th and 7th) positions of crRNA, because the 8th nucleotide is flipped-out and inaccessible for pairing [19]. Nonetheless, a surprising observation was that the PfCmr complex

did not target DNA *in vivo* even if both 6th and 7th positions within the crRNA handle contained mismatches. Drawing parallels with Type I and Type II CRISPR–Cas effector complexes that depend on the bipartite target DNA recognition (PAM and crRNA matching sequence) to activate target DNA cleavage [8,9,22,24,67], the authors concluded that the DNAse activation of PfCmr relies on the recognition of a specific RNA PAM motif (rPAM) within the flanking sequence rather than on noncomplementarity alone [36]. However, the proposed model of bipartite recognition is not without flaws. For example, the same authors showed that, even without any 3'-flanking sequence, the RNA target is able to activate the PfCmr2 HD nuclease (Figure 4C) [36]. The DNase activity of the TmCmr complex could also be switched on using an RNA target lacking any 3'-flanking sequence [37]. Since in the absence of the flanking sequence by definition there is no rPAM, it is unclear how to explain the Cmr activation. By contrast, the results of flanking sequence tests [36] perhaps may be reconciled with the base-pairing model assuming that not only Watson–Crick base-pairs are allowed at 4–7th positions of the crRNA handle.

Although the overall mechanism for distinguishing self from non-self might be expected to be the same for Cmr and Csm complexes, some differences are likely. PfCmr and TmCmr complexes on one hand and the StCsm complex on the other hand serve as a good illustration. As mentioned above, the HD-domain of these Cmr complexes is activated by the RNA target lacking any 3'-flanking sequence [36,37]. By contrast, the presence of 3'-flanking sequence is necessary for the activation of the StCsm HD-domain [38] (Figure 4C). One of the possible reasons for such a different behavior might be the significant structural difference between domain D2 in PfCmr2 [41,44] (Figure 2) and the corresponding domain in *Thermococcus onnurineus* Csm1 [46]. The observed variation in the stability of corresponding zinc-finger motifs might also be responsible for mechanistic differences between individual Type III complexes [46].

Concluding Remarks

Type III CRISPR-Cas systems were initially grouped together based on close evolutionary relationship between corresponding Cas proteins. Therefore, it seemed rather surprising when early experimental studies suggested that Csm (Type III-A) targets DNA, whereas Cmr (III-B) targets RNA. This functional division lingered for quite a while, perhaps mainly because of the complexity of these systems. However, recent developments have essentially reunited Csm and Cmr. Structural studies established homology between corresponding small subunits, Csm2 and Cmr5, the only subunits, for which the relationship has been uncertain. However, the most important breakthrough occurred in elucidating a common mechanism of RNA-dependent DNA degradation by both Csm and Cmr complexes. The mechanism has finally reconciled seemingly disparate activities of these complexes.

Although recent studies significantly advanced our knowledge regarding structure and function of Type III complexes, a number of intriguing questions remain (see Outstanding Questions). One of these is an important but still poorly understood role of the Cas10 Palm domain in Type III-mediated DNA silencing. Another open question is how Type III complexes distinguish self from non-self. At the moment two major models are proposed. One model is based on the positive recognition of two sequence elements, protospacer and rPAM, whereas the other model is based on the positive recognition of protospacer and the negative recognition (lack of complementarity) of its flanking sequence. Since differences between these models are well defined, their evaluation should be fairly straightforward. A more complicated question is how these recognition events lead to activation or repression of the Type III effector complexes. There is little doubt that further studies of Type III systems will bring about new exciting results and help us better understand CRISPR-Cas immunity.

Outstanding Questions

How do Type III ribonucleoprotein complexes assemble, and what is the ribonuclease which truncates the 3'end of crRNAs during the maturation of the Type III complexes?

What is the role of the GGDD-motif of the Cas10 Palm domain, and how does it cooperate with the HD-domain in the immunity provided by Csm and Cmr complexes?

What is the nature of the signal generated as a result of the target RNA binding to the spacer region of crRNA, and how is this signal transmitted to the HD-domain of Cas10?

How does the 3'-flanking sequence of target RNA control the DNA cleavage by the HD-domain of Cas10?

How do the two signals, one from the spacer region of crRNA and one from the 3'-flanking sequence, interact in regulating the activity of Cas10?

What is the role of Csm6/Csx1 proteins in Type III immunity?

Acknowledgments

This work in part has been supported by the Research Council of Lithuania (grant MIP-40/2013 to G.T.).

References

- Makarova, K.S. et al. (2006) A putative RNA-interference-based immune system in prokaryotes: computational analysis of the predicted enzymatic machinery, functional analogies with eukaryotic RNAi, and hypothetical mechanisms of action. *Biol. Direct* 1, 7
- Amitai, G. and Sorek, R. (2016) CRISPR-Cas adaptation: insights into the mechanism of action. *Nat. Rev. Microbiol.* 14, 67–76
- Wright, A.V. et al. (2016) Biology and applications of CRISPR systems: harnessing Nature's toolbox for genome engineering. *Cell* 164, 29–44
- Makarova, K.S. et al. (2015) An updated evolutionary classification of CRISPR-Cas systems. Nat. Rev. Microbiol. 13, 722–736
- Abudayyeh, O.O. *et al.* (2016) C2c2 is a single-component programmable RNA-guided RNA-targeting CRISPR effector. *Science* 353, aaf5573
- Shmakov, S. et al. (2015) Discovery and functional characterization of diverse class 2 CRISPR-Cas systems. Mol. Cell 60, 385–397
- Brouns, S.J. et al. (2008) Small CRISPR RNAs guide antiviral defense in prokaryotes. Science 321, 960–964
- Sinkunas, T. et al. (2013) In vitro reconstitution of Cascade-mediated CRISPR immunity in Streptococcus thermophilus. EMBO J. 32, 385–394
- Westra, E.R. et al. (2012) CRISPR immunity relies on the consecutive binding and degradation of negatively supercoiled invader DNA by Cascade and Cas3. *Mol. Cell* 46, 595–605
- Jackson, R.N. et al. (2014) Structural biology. Crystal structure of the CRISPR RNA-guided surveillance complex from Escherichia coli. Science 345, 1473–1479
- Jackson, R.N. and Wiedenheft, B. (2015) A conserved structural chassis for Mounting versatile CRISPR RNA-guided immune responses. *Mol. Cell* 58, 722–728
- Mulepati, S. *et al.* (2014) Structural biology. Crystal structure of a CRISPR RNA-guided surveillance complex bound to a ssDNA target. *Science* 345, 1479–1484
- Hatoum-Aslan, A. *et al.* (2013) A ruler protein in a complex for antiviral defense determines the length of small interfering CRISPR RNAs. *J. Biol. Chem.* 288, 27888–27897
- Rouillon, C. et al. (2013) Structure of the CRISPR interference complex CSM reveals key similarities with Cascade. Mol. Cell 52, 124–134
- Tamulaitis, G. et al. (2014) Programmable RNA shredding by the type III-A CRISPR-Cas system of Streptococcus thermophilus. Mol. Cell 56, 506–517
- Zhang, J. *et al.* (2016) Multiple nucleic acid cleavage modes in divergent type III CRISPR systems. *Nucleic Acids Res.* 44, 1789–1799
- Hale, C.R. *et al.* (2009) RNA-guided RNA cleavage by a CRISPR RNA-Cas protein complex. *Cell* 139, 945–956
- Zhang, J. et al. (2012) Structure and mechanism of the CMR complex for CRISPR-mediated antiviral immunity. *Mol. Cell* 45, 303–313
- Osawa, T. et al. (2015) Crystal structure of the CRISPR-Cas RNA silencing Cmr complex bound to a target analog. *Mol. Cell* 58, 418–430
- Taylor, D.W. *et al.* (2015) Structural biology. Structures of the CRISPR-Cmr complex reveal mode of RNA target positioning. *Science* 348, 581–585
- Anders, C. *et al.* (2014) Structural basis of PAM-dependent target DNA recognition by the Cas9 endonuclease. *Nature* 513, 569–573
- Gasiunas, G. et al. (2012) Cas9-crRNA ribonucleoprotein complex mediates specific DNA cleavage for adaptive immunity in bacteria. Proc. Natl. Acad. Sci. U.S.A. 109, E2579–E2586
- Jiang, F. and Doudna, J.A. (2015) The structural biology of CRISPR-Cas systems. *Curr. Opin. Struct. Biol.* 30, 100–111
- 24. Jinek, M. et al. (2012) A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* 337, 816–821

- Nishimasu, H. et al. (2014) Crystal structure of Cas9 in complex with guide RNA and target DNA. Cell 156, 935–949
- 26. Dong, D. et al. (2016) The crystal structure of Cpf1 in complex with CRISPR RNA. Nature 532, 522–526
- Fonfara, I. et al. (2016) The CRISPR-associated DNA-cleaving enzyme Cpf1 also processes precursor CRISPR RNA. Nature 532, 517–521
- Yamano, T. et al. (2016) Crystal Structure of Cpf1 in Complex with Guide RNA and Target DNA. Cell 165, 949–962
- Zetsche, B. et al. (2015) Cpf1 is a single RNA-guided endonuclease of a class 2 CRISPR-Cas system. Cell 163, 759–771
- Marraffini, L.A. and Sontheimer, E.J. (2008) CRISPR interference limits horizontal gene transfer in staphylococci by targeting DNA. *Science* 322, 1843–1845
- Marraffini, L.A. and Sontheimer, E.J. (2010) Self versus non-self discrimination during CRISPR RNA-directed immunity. *Nature* 463, 568–571
- Staals, R.H. et al. (2013) Structure and activity of the RNA-targeting Type III-B CRISPR-Cas complex of Thermus thermophilus. Mol. Cell 52, 135–145
- Deng, L. et al. (2013) A novel interference mechanism by a type IIIB CRISPR-Cmr module in Sulfolobus. Mol. Microbiol. 87, 1088–1099
- Goldberg, G.W. et al. (2014) Conditional tolerance of temperate phages via transcription-dependent CRISPR-Cas targeting. *Nature* 514, 633–637
- Samai, P. et al. (2015) Co-transcriptional DNA and RNA Cleavage during Type III CRISPR-Cas Immunity. Cell 161, 1164–1174
- Elmore, J.R. et al. (2016) Bipartite recognition of target RNAs activates DNA cleavage by the Type III-B CRISPR-Cas system. *Genes Dev.* 30, 447–459
- Estrella, M.A. et al. (2016) RNA-activated DNA cleavage by the Type III-B CRISPR-Cas effector complex. Genes Dev. 30, 460–470
- Kazlauskiene, M. et al. (2016) Spatiotemporal control of Type III-A CRISPR-Cas immunity: coupling DNA degradation with the target RNA recognition. Mol. Cell 62, 295–306
- Garneau, J.E. et al. (2010) The CRISPR/Cas bacterial immune system cleaves bacteriophage and plasmid DNA. Nature 468, 67–71
- Spilman, M. et al. (2013) Structure of an RNA silencing complex of the CRISPR-Cas immune system. Mol. Cell 52, 146–152
- Cocozaki, A.I. et al. (2012) Structure of the Cmr2 subunit of the CRISPR-Cas RNA silencing complex. Structure 20, 545–553
- Osawa, T. et al. (2013) Crystal structure of the Cmr2-Cmr3 subcomplex in the CRISPR-Cas RNA silencing effector complex. J. Mol. Biol. 425, 3811–3823
- 43. Shao, Y. et al. (2013) Structure of the Cmr2-Cmr3 subcomplex of the Cmr RNA silencing complex. Structure 21, 376–384
- Zhu, X. and Ye, K. (2012) Crystal structure of Cmr2 suggests a nucleotide cyclase-related enzyme in type III CRISPR-Cas systems. FEBS Lett. 586, 939–945
- Benda, C. *et al.* (2014) Structural model of a CRISPR RNA-silencing complex reveals the RNA-target cleavage activity in Cmr4. *Mol. Cell* 56, 43–54
- Jung, T.Y. et al. (2015) Crystal structure of the Csm1 subunit of the Csm complex and its single-stranded DNA-specific nuclease activity. Structure 23, 782–790
- Mulepati, S. and Bailey, S. (2011) Structural and biochemical analysis of nuclease domain of clustered regularly interspaced short palindromic repeat (CRISPR)-associated protein 3 (Cas3), J. Biol. Chem. 286, 31896–31903
- Venclovas, Č. (2016) Structure of Csm2 elucidates the relationship between small subunits of CRISPR-Cas effector complexes. FEBS Lett. 590, 1521–1529
- Makarova, K.S. et al. (2011) Unification of Cas protein families and a simple scenario for the origin and evolution of CRISPR-Cas systems. *Biol. Direct* 6, 38

Trends in Microbiology

- 50. Gallo, G. *et al.* (2016) Structural basis for dimer formation of the CRISPR-associated protein Csm2 of *Thermotoga maritima*. *FEBS J.* 283, 694–703
- Staals, R.H. et al. (2014) RNA targeting by the type III-A CRISPR-Cas Csm complex of Thermus thermophilus. Mol. Cell 56, 518–530
- 52. Hale, C.R. et al. (2014) Target RNA capture and cleavage by the Cmr type III-B CRISPR-Cas effector complex. Genes Dev. 28, 2432–2443
- Ramia, N.F. et al. (2014) Essential structural and functional roles of the Cmr4 subunit in RNA cleavage by the Cmr CRISPR-Cas complex. Cell Rep. 9, 1610–1617
- 54. Zhu, X. and Ye, K. (2015) Cmr4 is the slicer in the RNA-targeting Cmr CRISPR complex. *Nucleic Acids Res.* 43, 1257–1267
- 55. Peng, W. et al. (2015) An archaeal CRISPR type III-B system exhibiting distinctive RNA targeting features and mediating dual RNA and DNA interference. *Nucleic Acids Res.* 43, 406–417
- Zebec, Z. et al. (2014) CRISPR-mediated targeted mRNA degradation in the archaeon Sulfolobus solfataricus. Nucleic Acids Res. 42, 5280–5288
- Mulepati, S. and Bailey, S. (2013) In vitro reconstitution of an Escherichia coli RNA-guided immune system reveals unidirectional, ATP-dependent degradation of DNA target. J. Biol. Chem. 288, 22184–22192
- Sinkunas, T. *et al.* (2011) Cas3 is a single-stranded DNA nuclease and ATP-dependent helicase in the CRISPR/Cas immune system. *EMBO J.* 30, 1335–1342
- Hatoum-Aslan, A. et al. (2014) Genetic characterization of antiplasmid immunity through a type III-A CRISPR-Cas system. J. Bacteriol. 196, 310–317
- 60. Ramia, N.F. et al. (2014) Staphylococcus epidermidis Csm1 is a 3'-5' exonuclease. Nucleic Acids Res. 42, 1129–1138

- Wang, D. and Landick, R. (1997) Nuclease cleavage of the upstream half of the nontemplate strand DNA in an *Escherichia coli* transcription elongation complex causes upstream translocation and transcriptional arrest. *J. Biol. Chem.* 272, 5989–5994
- Zuo, Y. and Steitz, T.A. (2015) Crystal structures of the *E. coli* transcription initiation complexes with a complete bubble. *Mol. Cell* 58, 534–540
- Bolotin, A. *et al.* (2005) Clustered regularly interspaced short palindrome repeats (CRISPRs) have spacers of extrachromosomal origin. *Microbiology* 151, 2551–2561
- Deveau, H. et al. (2008) Phage response to CRISPR-encoded resistance in Streptococcus thermophilus. J. Bacteriol. 190, 1390–1400
- 65. Mojica, F.J. et al. (2009) Short motif sequences determine the targets of the prokaryotic CRISPR defence system. *Microbiology* 155, 733–740
- Westra, E.R. et al. (2013) Type I-E CRISPR-Cas systems discriminate target from non-target DNA through base pairing-independent PAM recognition. PLoS Genet. 9, e1003742
- Plagens, A. et al. (2014) In vitro assembly and activity of an archaeal CRISPR-Cas type I-A Cascade interference complex. *Nucleic Acids Res.* 42, 5125–5138
- Cao, L. *et al.* (2016) Identification and functional study of type III-A CRISPR-Cas systems in clinical isolates of *Staphylococcus aureus. Int. J. Med. Microbiol.* Published online August 30, 2016. http://dx.doi.org/10.1016/j.ijmm.2016.08.005
- Hale, C.R. et al. (2012) Essential features and rational design of CRISPR RNAs that function with the Cas RAMP module complex to cleave RNAs. *Mol. Cell* 45, 292–302

CelPress