MINIREVIEW

CRISPR-Cas Systems in Prokaryotes

MICHAŁ BURMISTRZ[1] and KRZYSZTOF PYRĆ[1,2*]

[1] Microbiology Department, Faculty of Biochemistry Biophysics and Biotechnology,
Jagiellonian University, Krakow, Poland
[2] Malopolska Centre of Biotechnology, Jagiellonian University, Krakow, Poland

A b s t r a c t

Prokaryotic organisms possess numerous strategies that enable survival in hostile conditions. Among others, these conditions include the invasion of foreign nucleic acids such as bacteriophages and plasmids. The clustered regularly interspaced palindromic repeats-CRISPR-associated proteins (CRISPR-Cas) system provides the majority of bacteria and archaea with adaptive and hereditary immunity against this threat. This mechanism of immunity is based on short fragments of foreign DNA incorporated within the hosts genome. After transcription, these fragments guide protein complexes that target foreign nucleic acids and promote their degradation. The aim of this review is to summarize the current status of CRISPR-Cas research, including the mechanisms of action, the classification of different types and subtypes of these systems, and the development of new CRISPR-Cas-based molecular biology tools.

K e y   w o r d s: CRISPR-Cas, prokaryotes

Introduction

Prokaryotes inhabit many hostile environments that are often inaccessible to other organisms. Regardless of their environment, prokaryotes are able to adapt and adjust to the changing world by developing new tactics to promote survival. One of the systems constituting a natural defense mechanism is the ability to distinguish self from non-self nucleic acids. Nucleic acids may be inserted into prokaryotic cells by infection, transduction, conjugation, or transformation, and may have harmful effects. The defense mechanisms include the restriction-modification, abortive infection, and surface exclusion systems, all of which act in an innate, non-specific manner (Samson *et al.*, 2013). However, a unique defense system involving clustered regularly interspaced palindromic repeats-CRISPR-associated proteins (CRISPR-Cas) was reported recently. This system is present in most archaeal species and approximately half of all bacterial species. A unique feature of the CRISPR-Cas system is that it provides adaptive and hereditary immunity against foreign nucleic acids, mimicking in some ways the immune systems of eukaryotes.

The first reports of the CRISPR-Cas system originate from as early as 1987, but these repeat regions were not associated with a specific function (Ishino *et al.*, 1987).

In 2002, they were discovered in other bacteria and named CRISPR (Jansen *et al.*, 2002). Since then, numerous roles of CRISPR-Cas systems have been postulated, including chromosomal rearrangements, modulation of gene expression, replicon partitioning, and DNA repair (Babu *et al.*, 2011; Jansen *et al.*, 2002; Makarova *et al.*, 2002; Mojica *et al.*, 1995). Discovery of the similarity between some CRISPR spacer sequences and sequences from viruses and plasmids led to the suggestion that CRISPRs play roles in adaptive immunity against foreign nucleic acids (Makarova *et al.*, 2006). Later, spacers homologous to chromosomal genes of the host were discovered (Bolotin *et al.*, 2005; Mojica *et al.*, 2005), suggesting that, in addition to phage immunity, CRISPR-Cas systems may play a role in autoimmunity or as a regulatory mechanisms (Sorek *et al.*, 2008; Stern *et al.*, 2010).

Components of the CRISPR-Cas system

The CRISPR-Cas system comprises genomic (CRISPR) and proteomic (Cas) components. The genomic component is a DNA loci containing short fragments of targeted nucleic acid sequences (spacers) interspaced by short repeated sequences (repeats). These

---

* Corresponding author: K. Pyrć, Microbiology Department, Faculty of Biochemistry Biophysics and Biotechnology, Jagiellonian University, Krakow, Poland; **e-mail: k.a.pyrc@uj.edu.pl.**

spacer sequences can be of either foreign or self-origin (Stern *et al.*, 2010; Vercoe *et al.*, 2013). The proteomic component is responsible for the incorporation of new template sequences, processing them into a form that enables base pairing with target nucleic acids, as well as for scanning and cleavage of target DNA or RNA. Of note, not all CRISPR-Cas systems discovered to date are active (Haft *et al.*, 2005; van der Ploeg, 2009).

Structure of the CRISPR-Cas system

The genomic component of the CRISPR-Cas system is formed by a series of variable spacers, which in some cases share sequence similarity with viruses, plasmids, or bacteria. These regions are interspaced with repeat sequences that are identical or almost identical within a single CRISPR cassette. The length of the repeat sequences varies between 25 and 40 nt, whereas the length of the spacer sequences varies between 21 and 72 nt. As mentioned above, some spacers show high homology with foreign nucleic acids, but the origin of a significant percentage of spacers remains unknown. The sequence homology between the spacer and the target is the major determinant of nucleic acid degradation of the target. Some bacterial species contain more than one CRISPR locus within their genome (Louwen *et al.*, 2014). Depending on the specific bacterial species or strain, a CRISPR locus may contain from a few to several hundred repeat spacer units; however, most commonly, a single CRISPR locus contains approximately 50 units. The CRISPR repeat sequences play an

important role during both the acquisition of new spacers and the transcription and maturation of CRISPR RNA (crRNA). Based on the sequence similarity, the CRISPR repeats are assigned into groups (Kunin *et al.*, 2007). These groups were taken into account in the current classification of CRISPR-Cas systems (Makarova *et al.*, 2011). Although most CRISPR arrays are located on chromosomal DNA, there are examples of CRISPRs located on plasmids (Godde and Bickerton, 2006). In general, CRISPR loci are flanked by A/T rich leader sequences containing promoter elements and binding sites for regulatory proteins (Jansen *et al.*, 2002; Yosef *et al.*, 2012).

CRISPR arrays are usually located in close proximity to a set of genes encoding Cas proteins. Members of this large group of proteins are required for CRISPR activity and have multiple functions as nucleases, polymerases, helicases, and nucleic acid binders. The exact number and types of Cas proteins depend on the specific CRISPR-Cas class.

Similar to other immune systems, the CRISPR-Cas system requires differentiation between self and nonself species to avoid autoimmunity. The CRISPR array is the most susceptible to autoimmune events as it contains spacers used for target recognition. In most cases, the CRISPR-Cas system deals with this problem by utilizing protospacer adjacent motifs (PAMs) (Sashital *et al.*, 2012),which are present in the target DNA but not the CRISPR array. DNA cleavage occurs only if the correct PAM sequence is present (Jinek *et al.*, 2012; Semenova *et al.*, 2011; Westra *et al.*, 2013). However, in type III CRISPR-Cas systems the safety mechanism

Table I
Subtypes of CRISPR-Cas system.

Type	Type hallmarks	Sub-type	Subtype hallmarks	crRNA processing*	Effector complex	Targeted nucleic acid	Exemplary organism
I	Cas3	I-A	Cas8a	Cas6 homologues, Cascade	Cascade	DNA	*Sulfolobus islandicus* (Held *et al.*, 2013)
		I-B	Cas8b				*Clostridium difficile* (Soutourina *et al.*, 2013)
		I-C	Cas8c, repeat group 3	Cas5d			*Xanthomonas oryzae* (Semenova *et al.*, 2009)
		I-D	Cas10	Cas6 homologues			*Thermofilum pendens* (Hrle *et al.*, 2014)
		I-E	Cse1, repeat group 2				*Escherichia coli* (Westra *et al.*, 2010)
		I-F	Csy1, repeat group 4				*Pseudomonas aeruginosa* (Bondy-Denomy *et al.*, 2013)
II	Cas9, repeat group 10	II-A	Csn2	RNase III, unknown nuclease	Cas9		*Streptococcus thermophilus* (Barrangou *et al.*, 2007)
		II-B	Cas4				*Francisella novicida* (Sampson *et al.*, 2013)
		II-C	lack of Csn2 and Cas4				*Neisseria meningitides* (Zhang *et al.*, 2013)
III	Cas10	III-A	Csm2	Cas6 homologues, Csm	Csm		*Staphylococcus epidermidis* (Marraffini and Sontheimer, 2008)
		III-B	Cmr5	Cas6 homologues, Cmr	Cmr	RNA	*Thermus thermophilus* (Staals *et al.*, 2013), *Sulfolubus solfataricus* (Zhang *et al.*, 2012b)

*An alternative crRNA biogenesis was described for *Neisseria* spp. of type II (details within main text).
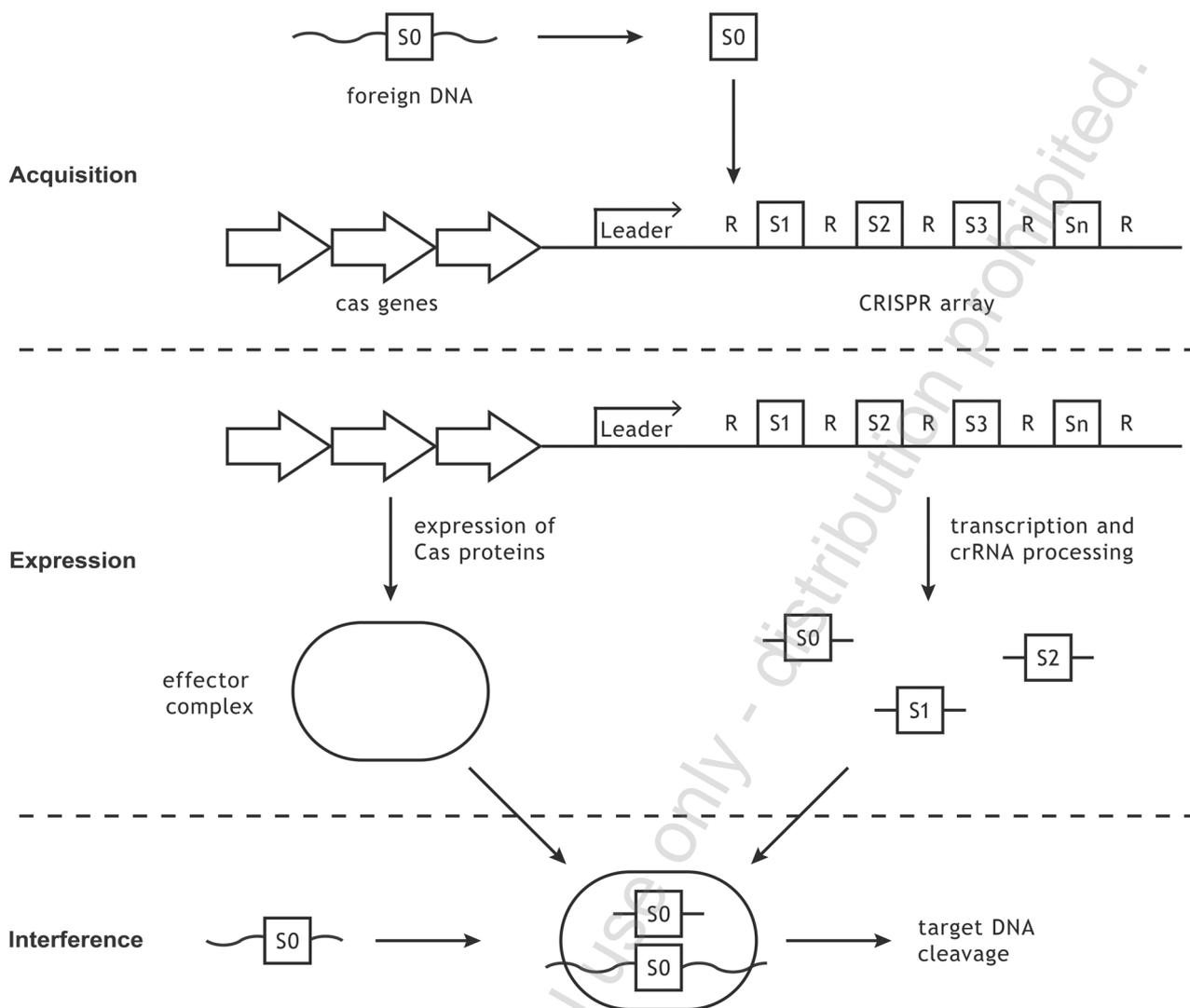
Fig. 1. Scheme of CRISPR-Cas system. S0, S1, S2, ..., Sn stands for spacer sequences. R stands for repeat sequence.

is PAM-independent and is based on the proximity of a repeat sequence to a spacer sequence in the CRISPR locus, which hinders its cleavage (Marraffini and Sontheimer, 2010).

The current classification of CRISPR-Cas systems is based on the sequences of the *cas* genes, the sequences of the repeats within the CRISPR arrays, and the organization of the *cas* operons (Makarova *et al.*, 2011). The classification distinguishes three main types of CRISPR-Cas systems (types I–III). Each type is also divided into subtypes, including I-A to I-F, II-A to II-C, III-A and III-B. The *cas1* and *cas2* genes are common to all three CRISPR-Cas types. The major criterion for classification includes the presence or absence of certain Cas proteins. For example, the Cas3, Cas9, and Cas10 proteins are hallmarks of CRISPR/Cas types I, II and III, respectively. Systems that do not have the specific hallmarks of types I–III are grouped as unclassified (type U). A description of the CRISPR-Cas subtypes is provided in Table I.

Mode of action

Profound studies of the mode of action of the CRISPR-Cas system revealed that the process can be divided into three distinct phases: (1) the acquisition of spacers following exposure to foreign nucleic acids, (2) the transcription of spacers and transcript processing to produce a crRNA that guides target recognition, and (3) target recognition and cleavage. Figure 1 shows a schematic overview of the CRISPR-Cas mode of action, and each phase is discussed in detail below.

Spacer acquisition

The most important feature of the CRISPR/Cas system is its capacity to respond dynamically to external stimuli. This adaptation relies on the ability to insert fragments of foreign nucleic acids into the CRISPR array, such that they serve as a template for the generation of crRNA, which mediates the subsequent

degradation of targeted nucleic acids. In this manner, bacteria are able to learn the patterns present in nucleic acids during invasion of foreign genetic elements. The mechanism responsible for the recognition, processing, and integration of new spacers has not been characterized fully; nevertheless, a number of factors involved in the process were identified.

New spacers are incorporated into the leader end of the CRISPR array in an orderly manner (Barrangou et al., 2007; Pourcel et al., 2005). Such unilateral insertion of new spacers converts the CRISPR array into a timeline containing a chronological database of foreign nucleic acid encounters. Prior to the incorporation of a new spacer sequence into the CRISPR array, foreign nucleic acids must be recognized and fragmented. A study performed in Streptococcus thermophiles indicated that the restriction-modification system can supply the CRISPR-Cas system with potential substrates for integration (Dupuis et al., 2013). However, several factors are important for the incorporation of novel nucleic acids into the cell's genome. For example, the presence of a PAM (Deveau et al., 2008) sequence flanking the protospacer sequence, is necessary for the selection of spacers (Mojica et al., 2009). A positively selected protospacer sequence is then processed into a precursor (Swarts et al., 2012). To enable integration of this precursor into the CRISPR array, the repeat adjacent to the leader end sequence is nicked at the 3' ends of both of its strands, and the precursor is subsequently integrated between two single-stranded repeat sequences (Díez-Villaseñor et al., 2013). It has been hypothesized that an interaction between the leader sequence and Cas proteins ensures proper orientation of newly integrated spacers within the CRISPR array (Yosef et al., 2012). The last step of spacer integration is reconstruction of the second strands of the repeat; as a result, the repeat sequence is duplicated during the integration process (Yosef et al., 2012). In addition to PAM, other sequence motifs that affect the rate of spacer incorporation were also discovered in proto spacers. An example of such a sequence motif is the acquisition affecting motif, which comprises a dinucleotide AA and is located at the 3' end of a protospacer (Yosef et al., 2013).

In addition to genetic factors, the process of spacer acquisition also involves protein components. Among the many factors involved in spacer acquisition, the Cas1 and Cas2 proteins seem to be the most important. Both of these proteins are important for spacer integration (Yosef et al., 2012) and are conserved in all CRISPR-Cas systems (Haft et al., 2005; Makarova et al., 2011). A study in Escherichia coli showed that Cas1 and Cas2 form a protein complex that interacts with the CRISPR locus (Datsenko et al., 2012; Yosef et al., 2012). Despite the identification of a number of activities and structural motifs of these proteins (Babu et al., 2011; Beloglazova et al., 2008; Nam et al., 2012a; Reeks et al.,

2013; Wiedenheft et al., 2012; Wiedenheft et al., 2009), their exact function in spacer acquisition remains unknown. A side from Cas1 and Cas2, several other factors may also be involved in the CRISPR-Cas adaptation stage. In E. coli, the Cas1 protein interacts with housekeeping proteins involved in DNA maintenance, including the RecBCD and RuvB proteins (Babu et al., 2011). Several recent findings also indicate the involvement of additional Cas proteins, namely, Cas3, Cas4, Csa1, and Csn2 (Arslan et al., 2013; Jackson et al., 2014; Makarova et al., 2006; Makarova et al., 2011; van der Oost et al., 2009; Zhang et al., 2012a). Csn2 is thought to be responsible for the recruitment of additional factors and stabilization of DNA breaks in the CRISPR array during spacer acquisition (Arslan et al., 2013).

An interesting phenomenon linked with spacer acquisition has been described. Although one spacer is sufficient for target DNA degradation, a single CRISPR locus can contain multiple spacers originating from a single genetic entity. This finding led to the so-called hypothesis of "primed spacer acquisition", in which the presence of one spacer can trigger the acquisition of additional spacers from the same DNA target (Swarts et al., 2012). This primed acquisition can also occur in the presence of mismatches between the spacer and protospacer sequences that prevent the interference (Datsenko et al., 2012; Heler et al., 2014). It is thought that the role of primed spacer acquisition is to counter possible evasion. The hypothesis that the effector complex slides across the target DNA was proposed to explain the mechanism of primed acquisition (Datsenko et al., 2012; Fineran et al., 2014; Li et al., 2014). This hypothesis states that, after a failed interference attempt, the effector complex recruits Cas1 and Cas2, and then slides along the target strand to find suitable PAMs that enable the acquisition of new spacers. However, this proposed mechanism has not yet been confirmed.

Transcript processing

Spacers within the CRISPR array cannot interfere directly with foreign nucleic acids and need to be transcribed. In most cases, transcription is initiated by the leader sequence located upstream of the CRISPR array. As a result, a long precursor transcript called a pre-crRNA is created. Further processing of this transcript differs depending on the CRISPR-Cas system type; however, a notable exception has been described. In Neisseria spp., no pre-crRNA transcript is generated; instead, the crRNAs are transcribed separately from promoters embedded within repeat sequences, and their 5' ends are formed by transcription rather than maturation (Zhang et al., 2013). Notably, these crRNAs can undergo type II processing (see below), although it is not required for their functionality (Zhang et al., 2013).

In type I and III CRISPR-Cas systems, initial processing of the pre-crRNA is performed by Cas6 nuclease homologs carrying two repeat-associated mysterious protein domains (Reeks *et al.*, 2013). Cleavage of the pre-crRNA generates a product comprising a full spacer sequence flanked by repeat-derived handles on both ends. In most cases, the 5' handles are 8 nt (nucleotides) long (11nt in type I-C systems (Nam *et al.*, 2012b) and 13 nt in I-D systems (Scholz *et al.*, 2013), whereas the 3' handles have variable lengths depending on the specific CRISPR-Cas type. In some types, the 3' handles form hairpin structures (Carte *et al.*, 2010; Carte *et al.*, 2008; Niewoehner *et al.*, 2014; Wang *et al.*, 2011). In type I-C systems, initial processing is performed by Cas5 rather than Cas6 (Garside *et al.*, 2012; Nam *et al.*, 2012b), whereas in other CRISPR-Cas types, Cas5 is catalytically inactive and is thought to interact with the 5' handle of the crRNA (Wiedenheft *et al.*, 2011). In most type I and type III systems, after initial processing, the crRNA is transferred to CRISPR ribonucleoprotein (crRNP) complexes for secondary processing (Hatoum-Aslan *et al.*, 2011). In types I-A, I-B, I-C, and I-D, this processing results in shortening of the 3' handles. In both subtypes of type III, secondary processing generates two populations of products that differ in length by 6 nt (Hale *et al.*, 2012; Zhang *et al.*, 2012b). In contrast to the systems mentioned above, types I-E and I-F do not utilize secondary processing. After initial cleavage by Cas6 or its homolog, the Cas6 protein remains associated with the hairpin at the 3' handle of the crRNA. Subsequently, this sub-complex is incorporated into Cascade, the effector complex utilized by type I systems.

The mechanism of pre-crRNA processing in type II systems differs from those of type I and type III systems. First, an additional fragment encoding a trans activating crRNA (tracrRNA) is transcribed from the CRISPR locus. The tracrRNA contains a 25 nt sequence complementary to the repeat sequences within the pre-crRNA. Initial processing involves hybridization of the tracrRNA and pre-crRNA, and the hybrids are subsequently digested by RNase III associated with Cas9 (Deltcheva *et al.*, 2011; Jinek *et al.*, 2012). After cleavage, the hybrids attached to Cas9 undergo further trimming by an unidentified nuclease to yield functional, mature crRNA (Deltcheva *et al.*, 2011).

Target recognition and degradation

Mature crRNAs are incorporated into crRNP complexes. Depending on the type of CRISPR-Cas system, these effector complexes vary in their composition and mode of action; however, the general mechanism is common among types I–III. Initially, the crRNP complex scans nucleic acids for protospacer sequences.

Base pairing between the crRNA spacer and the proto spacer begins in a 7–8 nt seed region (Semenova *et al.*, 2011). Further hybridization to the protospacer region results in formation of an R-loop structure, in which the crRNA is paired with one of the DNA strands and the displaced DNA strand remains single-stranded (Sashital *et al.*, 2012; Sorek *et al.*, 2013). This structure triggers a conformational change in the crRNP complex, thereby initiating type-dependent nuclease activity (Jore *et al.*, 2011; Spilman *et al.*, 2013; Wiedenheft *et al.*, 2011). During the degradation phase, to prevent the effector complex from digesting its own CRISPR array, it is essential to distinguish between foreign and self DNA. In type I and type II CRISPR-Cas systems, this safety mechanism is based on the presence of PAMs (Gasiunas *et al.*, 2012; Jinek *et al.*, 2012; Semenova *et al.*, 2011; Westra *et al.*, 2013). Within the targeted nucleic acids, the PAM sequence is localized in close proximity to the protospacer sequence, whereas the spacer sequence in the CRISPR array lacks this element; consequently, the CRISPR region is not cleaved. In type III CRISPR-Cas systems, discrimination between self and foreign DNAs is PAM-independent. In these systems, the 5' handle of a crRNA interacts with a repeat sequence in the CRISPR locus. This interaction probably prevents nuclease recruitment, thereby preventing cleavage of self DNA (Marraffini and Sontheimer, 2010).

The effector complexes of type I and type III systems share a significant level of structural similarity. The small subunits of these complexes include Csa5 and Cse2 in type I systems, and Csm2 and Cmr5 in type III systems. The large subunits include Cas8 homologs and Cas10 homologs in types I and III, respectively. The effect or complexes of all subtypes of type I CRISPR-Cas systems share significant similarities and are collectively referred to as Cascade complexes (Reeks *et al.*, 2013). A typical Cascade complex consist of Cse1 (one subunit), Cse2 (two subunits), Cas5 (one subunit), Cas7 (six subunits), and Cas6 (one subunit) (Jore *et al.*, 2011; van Duijn *et al.*, 2012). The type I-E Cascade complex from *E. coli* is the most well-characterized of all CRISPR-Cas systems. The backbone of the complex consists of six Cas7 units, along which the crRNA is positioned. The 5' and 3' handles of the crRNA are anchored at opposite sides of the complex to Cas5 and Cas6, respectively. The Cse1 protein located at the 5' crRNA end of the complex is responsible for the non-specific interaction with DNA during initial target scanning (Jore *et al.*, 2011). This protein is also responsible for interaction with the PAM motif located at the 3' end of the proto spacer (Sternberg *et al.*, 2012), and this interaction is thought to trigger hybridization of the crRNA and target DNA by destabilization of the DNA duplex (Sashital *et al.*, 2012). In type I-E systems, the seed region, which requires perfect hybridization for target degradation, is located at the 5' end of the spacer (positions 1–5,

7, and 8) (Künne *et al.*, 2014; Semenova *et al.*, 2011). A conformational change in the Cascade complex, linked with formation of the R-loop, triggers recruitment of the Cas3 protein (Westra *et al.*, 2012). Cas3 possesses metal-dependent nuclease and ATP-dependent helicase activities (Beloglazova *et al.*, 2011; Jackson *et al.*, 2014; Mulepati and Bailey, 2011), and initial DNA cleavage occurs on the strand displaced during formation of the R-loop (Sinkunas *et al.*, 2013). The subsequent degradation of the target proceeds in an exonucleolytic manner from the 3' to the 5' end (Beloglazova *et al.*, 2011; Westra *et al.*, 2012). The second strand of the DNA is digested in both endonucleolytic and exonucleolytic manners (Beloglazova *et al.*, 2011; Mulepati and Bailey, 2011; Westra *et al.*, 2012).

The mechanism of interference used by type II CRISPR-Cas systems differs from that of type I systems. In type II systems, interference depends on the Cas9-RNP complex, the structure of which is relatively less intricate than those of other CRISPR types and includes Cas9 and a single guide RNA (sgRNA) (Jinek *et al.*, 2014; Nishimasu *et al.*, 2014). The sgRNA is a duplex of crRNA and tracrRNA (Jinek *et al.*, 2012). Cas9 has two domains, an α-helical recognition domain and a nuclease domain; the first of these domains is responsible for coordinating the guide RNA, and the second takes part in PAM recognition and target DNA cleavage. The nuclease domain actually comprises two individual nucleases, namely, RuvC and HNH (Nishimasu *et al.*, 2014). The interference process begins with the incorporation of a sgRNA into Cas9, which results in a conformational change of Cas9 that enables the complex to bind DNA (Jinek *et al.*, 2014). Subsequently, DNA scanning proceeds in a similar manner to that in type I systems. The PAM motif of type II systems is located at the 5' end of the proto spacer (Deveau *et al.*, 2008; Mojica *et al.*, 2009). Recognition of this motif is considered to lead to displacement of the DNA strand and formation of the R-loop. Base pairing between the spacer and proto spacer starts from a 12 nt seed region (Jiang *et al.*, 2013; Sternberg *et al.*, 2014). Each strand of the target DNA is digested by a different nuclease domain of Cas9 protein; specifically, the HNH domain cleaves the strand hybridized with crRNA, and the RuvC domain cleaves the displaced strand (Jinek *et al.*, 2012; Jinek *et al.*, 2014).

Type III CRISPR-Cas effector complexes resemble those of the type I Cascade complex. The backbone of the type III crRNP complex is composed of multiple Csm3 (type III-A) or Cmr4 (type III-B) proteins (Rouillon *et al.*, 2013; Spilman *et al.*, 2013; Staals *et al.*, 2013). Like type I CRISPR-Cas effector complexes, the crRNA is positioned along the backbone in type III complexes (Rouillon *et al.*, 2013). Parallel to the backbone, the small subunits (Csm2 or Cmr5) form a secondary string connecting the two ends of the complex. The Csm4 (type III-A) or Cmr3 (type III-B) protein is located at the 5' end of the crRNA and is bound to the large subunit (Csm1 or Cmr2, respectively). Together with the Cmr6 protein, Csm5 or Cmr1 is located at the 3' end of the crRNA and bound to a small subunit (Csm2 or Cmr5). Functioning of type III systems is not as well-characterized as that of other types, but some important differences have been described. As mentioned earlier, in type III-A systems, discrimination between self and non-self DNA proceeds in a PAM-independent manner (Rouillon *et al.*, 2013), and the large subunit (Csm1) is thought to take part in the discrimination process (Hatoum-Aslan *et al.*, 2014). The presence of a seed sequence within the spacer sequence was proposed (Hatoum-Aslan *et al.*, 2014) but not confirmed fully. Type III-B systems are unique among other CRISPR types because they target RNA rather than DNA (Hale *et al.*, 2009); however, there are examples of type III-B systems targeting plasmid DNA *in vivo* (Deng *et al.*, 2013). Most type III-B effector complexes comprise six proteins (Cmr1-6) (Gasiunas *et al.*, 2014); however, in *Sulfolobus solfataricus*, these effector complexes contain an additional protein (Cmr7) (Zhang *et al.*, 2012b). Recognized target RNA is cleaved at regular intervals in the 3' to 5' direction (Staals *et al.*, 2013). This digestion pattern indicates the presence of multiple active sites within the backbone of the Cmr complex (Staals *et al.*, 2013), which is composed of multiple subunits of Cmr4 and Cmr5.

Other functions of the CRISPR-Cas system

Recognition and degradation of invading nucleic acids is a major function of the CRISPR-Cas system; however, several studies showed that it may have additional roles. One of the intriguing features of the CRISPR-Cas system is the presence of spacers that are complementary to bacterial genes (Bolotin *et al.*, 2005; Horvath *et al.*, 2009; Mojica *et al.*, 2005; Shah *et al.*, 2009). This feature suggests the possibility of recognition and cleavage of self nucleic acids and, consequently, may be considered a form of autoimmunity (Stern *et al.*, 2010; van der Ploeg, 2009). However, it is worth noting that CRISPR-Cas systems containing self-targeting elements often lack some *cas* genes or are non-functional (Haft *et al.*, 2005; Stern *et al.*, 2010; van der Ploeg, 2009). None the less, the presence of autoimmunity within CRISPR-Cas systems does not exclude their regulatory roles, and alteration of *cas* gene expression in response to stress is a commonly observed phenomenon (Perez-Rodriguez *et al.*, 2011; Viswanathan *et al.*, 2007). There are also several examples of Cas proteins involved in the regulation of bacterial virulence. *Francisella novicida* utilizes a CRISPR-Cas system to evade recognition by host pattern recognition receptors (Sampson

et al., 2013); this is achieved by down-regulation of the mRNA coding for bacterial lipoprotein, a surface protein that is recognized by the host's Toll-like receptor-2. Down-regulation of the mRNA encoding bacterial lipoprotein weakens the Toll-like receptor-2 induced pro-inflammatory cytokine response. *Neisseria meningitides* and *Campylobacter jejuni* employ Cas9 for the modulation of virulence (Louwen *et al.*, 2013; Sampson *et al.*, 2013); however, in these species, the exact mechanism of action has not been deciphered fully. Furthermore, in *Legionella pneumophila*, Cas2 plays an important role in intracellular survival and replication in amoebae *via* an unidentified mechanism (Gunderson and Cianciotto, 2013).

CRISPR-Cas-based technology

The first practical application of CRISPR-Cas systems was typing of bacterial diversity (Groenen *et al.*, 1993; Kamerbeek *et al.*, 1997). Several typing methods were developed based on CRISPR locus size, spacer content of a given locus, or point mutations within specific spacer and/or repeat sequences (Shariat and Dudley, 2014). Increasing knowledge of the mechanisms of action of CRISPR-Cas opened new avenues for applications of these systems. It became clear that easily reprogrammable, precisely controllable, sequence-specific nucleic acid cleavage is useful for molecular biology-based studies. To date, the most well-developed methods are based on the type II CRISPR-Cas effector complex, mainly because of its simplicity, in that it consists of Cas9 and sgRNA only (Jinek *et al.*, 2012). In addition, DNA cleavage by Cas9 is very efficient (Gasiunas *et al.*, 2012), and target recognition can be programmed by exchanging the sgRNA guiding sequence (Jinek *et al.*, 2012). The limitation of the application of this effector complex for molecular biology studies is the requirement for a PAM sequence within the desired target.

Eukaryotic DNA repair mechanisms can be used in combination with the CRISPR-Cas9 system can be used as a genome editing tool. Introduction of a crRNA-tracr-RNA-Cas9 unit into eukaryotic cells results in sequence-specific DNA cleavage, and the cleaved strands undergo repair by an endogenous mechanism in the cell. Homologous recombination or non-homologous end joining of previously cleaved DNA enables genome editing (Cong *et al.*, 2013; Jinek *et al.*, 2013; Mali *et al.*, 2013b). In addition to dsDNA cleavage, Cas9-based systems can also be used for other purposes, such as nicking or binding DNA, which can be achieved by introducing mutations within the active sites of the Cas9 nuclease domains (HNH and RuvC) (Mali *et al.*, 2013a). The advantage of nicking DNA over blunt cutting is the possibility to create defined overhangs, thereby enabling direction of the specific recombination event (Mali *et al.*, 2013a).

Mutation of both the HNH and RuvC domains converts Cas9 into a RNA-guided dsDNA binding protein (dCas9), and programming dCas9 to bind to a specific promoter region can efficiently prevent transcription of a gene of interest (Bikard *et al.*, 2013).

The use of Cas9 fused with other proteins also enables its employment for other applications. When fused with an activator domain, dCas9 enhances transcription (Bikard *et al.*, 2013); for example, dCas9 fused with the ω subunit of bacterial RNA polymerase recruits other subunits of the polymerase (Opalka *et al.*, 2010). Furthermore, fusion of Cas9 with green fluorescent protein enables its use in live cell imaging (Chen *et al.*, 2013). It could be speculated that other CRISPR-Cas modules will be used in molecular biology applications in the future, such as targeting of RNA by type III-B systems.

Concluding remarks

Since its discovery, the CRISPR-Cas system has been an intriguing object of study. Before the identification of the functionality of this system, adaptive immunity was considered to be associated exclusively with eukaryotic organisms (Goren *et al.*, 2012). Adaptations of the CRISPR-Cas system in response to environmental conditions are hereditary; therefore, this system is a fine example of Lamarckian inheritance (Koonin and Wolf, 2009). Enhanced understanding of the mechanism of action of the CRISPR-Cas system enabled the development of new tools for genome engineering that are efficient alternatives to transcription activator-like effector nucleases (Boch *et al.*, 2009) and zinc-finger nucleases (Bibikova *et al.*, 2003). The CRISPR-Cas system was also used to build phage resistance in an industrial strain of *Streptococcus thermophiles* (Barrangou *et al.*, 2013). This application is of significant value to microorganism-based industries, such as the dairy and biopharmaceutical markets. Certainly, new CRISPR-Cas-based methods will be developed in the future. Despite intensive research during recent years, many questions concerning the mode of action and possible applications of the CRISPR-Cas system remain unanswered.

Literature

Arslan Z., R.Wurm, O. Brener, P. Ellinger, L. Nagel-Steger, F. Oesterhelt, l. Schmitt, D. Willbold, R. Wagner, H. Gohlke H. and others. 2013. Double-strand DNA end-binding and sliding of the toroidal CRISPR-associated protein Csn2. *Nucleic Acids Res.* 41: 6347–6359.

Babu M., N. Beloglazova, R. Flick, C. Graham, T. Skarina, B. Nocek, A. Gagarinova, O. Pogoutse, G. Brown, A. Binkowski and others. 2011. A dual function of the CRISPR-Cas system in bacterial antivirus immunity and DNA repair. *Mol. Microbiol.* 79: 484–502.

Barrangou R., C. Fremaux, H. Deveau, M. Richards, P. Boyaval, S. Moineau, D.A. Romero and P. Horvath. 2007. CRISPR provides acquired resistance against viruses in prokaryotes. *Science* 315: 1709–1712.

Barrangou R., A.C. Coûté-Monvoisin, B. Stahl, I. Chavichvily, F. Damange, D.A. Romero, P. Boyaval, C. Fremaux and P. Horvath. 2013. Genomic impact of CRISPR immunization against bacteriophages. *Biochem. Soc. Trans.* 41: 1383–1391.

Beloglazova N., G. Brown, M.D. Zimmerman, M. Proudfoot, K.S. Makarova, M. Kudritska, S. Kochinyan, S. Wang, M. Chruszcz, W. Minor and others. 2008. A novel family of sequence-specific endoribonucleases associated with the clustered regularly interspaced short palindromic repeats. *J. Biol. Chem.* 283: 20361–20371.

Beloglazova N., P. Petit, R. Flick, G. Brown, A. Savchenko and A.F. Yakunin. 2011. Structure and activity of the Cas3 HD nuclease MJ0384, an effector enzyme of the CRISPR interference. *EMBO J.* 30: 4616–4627.

Bibikova M., K. Beumer, J.K. Trautman and D. Carroll. 2003. Enhancing gene targeting with designed zinc finger nucleases. *Science* 300: 764.

Bikard D., W. Jiang, P. Samai, A. Hochschild, F. Zhang and L.A. Marraffini. 2013. Programmable repression and activation of bacterial gene expression using an engineered CRISPR-Cas system. *Nucleic Acids Res.* 41: 7429–7437.

Boch J., H. Scholze, S. Schornack, A. Landgraf, S. Hahn, S. Kay, T. Lahaye, A. Nickstadt and U. Bonas. 2009. Breaking the code of DNA binding specificity of TAL-type III effectors. *Science* 326: 1509–1512.

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

Bondy-Denomy J., A. Pawluk, K.L. Maxwell and A.R. Davidson. 2013. Bacteriophage genes that inactivate the CRISPR/Cas bacterial immune system. *Nature* 493: 429–432.

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

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

Chen B., L.A. Gilbert, B.A. Cimini, J. Schnitzbauer, W. Zhang, G.W. Li, J. Park, E.H. Blackburn, J.S. Weissman, L.S. Qi and others. 2013. Dynamic imaging of genomic loci in living human cells by an optimized CRISPR/Cas system. *Cell* 155: 1479–1491.

Cong L., F.A. Ran, D. Cox, S. Lin, R. Barretto, N. Habib, P.D. Hsu, X. Wu, W. Jiang, L.A. Marraffini and others. 2013. Multiplex genome engineering using CRISPR/Cas systems. *Science* 339: 819–823.

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

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

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

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

Díez-Villaseñor C., N.M. Guzmán, C. Almendros, J. García-Martínez and F.J. Mojica F.J. 2013. CRISPR-spacer integration reporter plasmids reveal distinct genuine acquisition specificities among CRISPR-Cas I-E variants of *Escherichia coli*. *RNA Biol.* 10: 792–802.

Dupuis M., M. Villion, A.H. Magadán and S. Moineau. 2013. CRISPR-Cas and restriction-modification systems are compatible and increase phage resistance. *Nat. Commun.* 4: 2087.

Fineran P.C., M.J. Gerritzen, M. Suárez-Diez, T. Künne, J. Boekhorst, S.A. van Hijum, R.H. Staals and S.J. Brouns. 2014. Degenerate target sites mediate rapid primed CRISPR adaptation. *Proc. Natl. Acad. Sci. USA* 111: E1629–1638.

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

Gasiunas G., R. Barrangou, P. Horvath and V. Siksnys. 2012. Cas9-crRNA ribonucleoprotein complex mediates specific DNA cleavage for adaptive immunity in bacteria. *Proc. Natl. Acad. Sci. USA* 109: E2579–2586.

Gasiunas G., T. Sinkunas and V. Siksnys. 2014. Molecular mechanisms of CRISPR-mediated microbial immunity. *Cell Mol. Life Sci.* 71: 449–465.

Godde J.S. and A. Bickerton. 2006. The repetitive DNA elements called CRISPRs and their associated genes: evidence of horizontal transfer among prokaryotes. *J. Mol. Evol.* 62: 718–729.

Goren M., I. Yosef, R. Edgar and U. Qimron. 2012. The bacterial CRISPR/Cas system as analog of the mammalian adaptive immune system. *RNA Biol.* 9: 549–554.

Groenen P.M., A.E. Bunschoten, D. van Soolingen and J.D. van Embden. 1993. Nature of DNA polymorphism in the direct repeat cluster of *Mycobacterium tuberculosis*; application for strain differentiation by a novel typing method. *Mol. Microbiol.* 10: 1057–1065.

Gunderson F.F. and N.P. Cianciotto. 2013. The CRISPR-associated gene cas2 of *Legionella pneumophila* is required for intracellular infection of amoebae. *MBio.* 4: e00074–00013.

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

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

Hale C.R., S. Majumdar, J. Elmore, N. Pfister, M. Compton, S. Olson, A.M. Resch, C.V. Glover, B.R. Graveley, R.M. Terns and others. 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.

Hatoum-Aslan A., I. Maniv and L.A. Marraffini. 2011. Mature clustered, regularly interspaced, short palindromic repeats RNA (crRNA) length is measured by a ruler mechanism anchored at the precursor processing site. *Proc. Natl. Acad. Sci. USA* 108: 21218–21222.

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

Held N.L., A. Herrera and R.J. Whitaker. 2013. Reassortment of CRISPR repeat-spacer loci in *Sulfolobus islandicus*. *Environ. Microbiol.* 15: 3065–3076.

Heler R., L.A. Marraffini and D. Bikard. 2014. Adapting to new threats: the generation of memory by CRISPR-Cas immune systems. *Mol. Microbiol.* 93: 1–9.

Horvath P., A.C. Coûté-Monvoisin, D.A. Romero, P. Boyaval, C. Fremaux and R. Barrangou. 2009. Comparative analysis of CRISPR loci in lactic acid bacteria genomes. *Int. J. Food Microbiol.* 131: 62–70.

Hrle A., L.K. Maier, K. Sharma, J. Ebert, C. Basquin, H. Urlaub, A. Marchfelder and E. Conti. 2014. Structural analyses of the CRISPR protein Csc2 reveal the RNA-binding interface of the type I-D Cas7 family. *RNA Biol.* 11(8): 1072–1082.

Ishino Y., H. Shinagawa, K. Makino, M. Amemura and A. Nakata. 1987. Nucleotide sequence of the *iap* gene, responsible for alkaline phosphatase isozyme conversion in *Escherichia coli*, and identification of the gene product. *J. Bacteriol.* 169: 5429–5433.

Jackson R.N., M. Lavin, J. Carter and B. Wiedenheft. 2014. Fitting CRISPR-associated Cas3 into the helicase family tree. *Curr. Opin. Struct. Biol.* 24: 106–114.

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

Jiang W., D.Bikard, D. Cox, F. Zhang and L.A. Marraffini. 2013. RNA-guided editing of bacterial genomes using CRISPR-Cas systems. *Nat. Biotechnol.* 31: 233–239.

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

Jinek M., A. East, A. Cheng, S. Lin, E. Ma and J. Doudna. 2013. RNA-programmed genome editing in human cells. *Elife* 2: e00471.

Jinek M., F. Jiang, D.W. Taylor, S.H. Sternberg, E. Kaya, E. Ma, C. Anders , M. Hauer, K. Zhou, S. Lin and others. 2014. Structures of Cas9 endonucleases reveal RNA-mediated conformational activation. *Science* 343: 1247997.

Jore M.M., M. Lundgren, E. van Duijn, J.B. Bultema, E.R. Westra, S.P. Waghmare, B. Wiedenheft, U. Pul, R. Wurm, R. Wagner and others. 2011. Structural basis for CRISPR RNA-guided DNA recognition by Cascade. *Nat. Struct. Mol. Biol.* 18: 529–536.

Kamerbeek J., L. Schouls, A. Kolk, M. van Agterveld, D. van Soolingen, S. Kuijper , A. Bunschoten, H. Molhuizen, R. Shaw, M. Goyal and others. 1997. Simultaneous detection and strain differentiation of *Mycobacterium tuberculosis* for diagnosis and epidemiology. *J. Clin. Microbiol.* 35: 907–914.

Koonin E.V. and Y.I. Wolf. 2009. Is evolution Darwinian or/and Lamarckian? *Biol. Direct.* 4: 42.

Kunin V., R.Sorek and P. Hugenholtz. 2007. Evolutionary conservation of sequence and secondary structures in CRISPR repeats. *Genome Biol.* 8: R61.

Künne T., D.C. Swarts and S.J. Brouns. 2014. Planting the seed: target recognition of short guide RNAs. *Trends Microbiol.* 22: 74–83.

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

Louwen R., D. Horst-Kreft, A.G. de Boer, L. van der Graaf, G. de Knegt, M. Hamersma, A.P. Heikema, A.R. Timms, B.C. Jacobs, J.A. Wagenaar and others. 2013. A novel link between *Campylobacter jejuni* bacteriophage defence, virulence and Guillain-Barré syndrome. *Eur. J. Clin. Microbiol. Infect. Dis* 32: 207–226.

Louwen R., R.H. Staals, H.P. Endtz, P. van Baarlen and J. van der Oost. 2014. The role of CRISPR-Cas systems in virulence of pathogenic bacteria. *Microbiol. Mol. Biol. Rev.* 78: 74–88.

Makarova K.S., L. Aravind, N.V. Grishin, I.B. Rogozin and E.V. Koonin. 2002. A DNA repair system specific for thermophilic Archaea and bacteria predicted by genomic context analysis. *Nucleic. Acids. Res.* 30: 482–496.

Makarova K.S.,N.V. Grishin, S.A. Shabalina, Y.I. Wolf and E.V. Koonin. 2006. A putative RNA-interference-based immune system in prokaryotes: computational analysis of the predicted enzymatic machinery, functional analogies with eukaryotic RNAi, and hypothetical mechanisms of action. *Biol. Direct.* 1: 7.

Makarova K.S., D.H. Haft, R. Barrangou, S.J. Brouns, E. Charpentier, P. Horvath, S. Moineau, F.J. Mojica, Y.I. Wolf, A.F. Yakunin and others. 2011. Evolution and classification of the CRISPR-Cas systems. *Nat. Rev. Microbiol.* 9: 467–477.

Mali P., K.M. Esvelt and G.M. Church. 2013a. Cas9 as a versatile tool for engineering biology. *Nat. Methods* 10: 957–963.

Mali P., L. Yang, K.M. Esvelt, J.Aach, M. Guell, J.E. DiCarlo, J.E. Norville and G.M. Church. 2013b. RNA-guided human genome engineering via Cas9. *Science* 339: 823–826.

Marraffini L.A. and E.J. Sontheimer. 2008. CRISPR interference limits horizontal gene transfer in staphylococci by targeting DNA. *Science* 322: 1843–1845.

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

Mojica F.J., C. Ferrer, G. Juez and F. Rodríguez-Valera. 1995. Long stretches of short tandem repeats are present in the largest replicons of the Archaea *Haloferax mediterranei* and *Haloferax volcanii* and could be involved in replicon partitioning. *Mol. Microbiol.* 17: 85–93.

Mojica F.J., C. Díez-Villaseñor, J. García-Martínez and E. Soria E. 2005. Intervening sequences of regularly spaced prokaryotic repeats derive from foreign genetic elements. *J. Mol. Evol.* 60: 174–182.

Mojica F.J., C. Díez-Villaseñor, J. García-Martínez and C. Almendros. 2009. Short motif sequences determine the targets of the prokaryotic CRISPR defence system. *Microbiology* 155: 733–740.

Mulepati S. and S. Bailey. 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.

Nam K.H., F. Ding, C. Haitjema, Q. Huang, M.P. DeLisa and A. Ke. 2012a. Double-stranded endonuclease activity in *Bacillus halodurans* clustered regularly interspaced short palindromic repeats (CRISPR)-associated Cas2 protein. *J. Biol. Chem.* 287: 35943–35952.

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

Niewoehner O., M. Jinek and J.A. Doudna. 2014. Evolution of CRISPR RNA recognition and processing by Cas6 endonucleases. *Nucleic Acids Res.* 42: 1341–1353.

Nishimasu H., F.A. Ran, P.D. Hsu, S. Konermann, S.I. Shehata, N. Dohmae, R. Ishitani, F. Zhang and O. Nureki. 2014. Crystal structure of Cas9 in complex with guide RNA and target DNA. *Cell* 156: 935–949.

Opalka N., J. Brown, W.J. Lane, K.A. Twist, R. Landick, F.J. Asturias and S.A. Darst. 2010. Complete structural model of *Escherichia coli* RNA polymerase from a hybrid approach. *PLoS Biol* 8: e1000483

Perez-Rodriguez R., C. Haitjema, Q.Huang, K.H. Nam, S. Bernardis, A. Ke and M.P. De Lisa. 2011. Envelope stress is a trigger of CRISPR RNA-mediated DNA silencing in *Escherichia coli*. *Mol. Microbiol.* 79: 584–599.

Pourcel C.,G. Salvignol and G. Vergnaud. 2005. CRISPR elements in *Yersinia pestis* acquire new repeats by preferential uptake of bacteriophage DNA, and provide additional tools for evolutionary studies. *Microbiology* 151: 653–663.

Reeks J., J. Naismith and M.F. White. 2013. CRISPR interference: a structural perspective. *Biochem. J.* 453: 155–166.

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

Sampson T.R., S.D. Saroj, A.C. Llewellyn, Y.L. Tzeng and D.S. Weiss. 2013. A CRISPR/Cas system mediates bacterial innate immune evasion and virulence. *Nature* 497: 254–257.

Samson J.E., A.H. Magadán, M. Sabri and S. Moineau. 2013. Revenge of the phages: defeating bacterial defences. *Nat. Rev. Microbiol.* 11: 675–687.

Sashital D.G., B. Wiedenheft and J.A. Doudna. 2012. Mechanism of foreign DNA selection in a bacterial adaptive immune system. *Mol. Cell* 46: 606–615.

Scholz I., S.J. Lange, S. Hein, W.R. Hess and R. Backofen. 2013. CRISPR-Cas systems in the cyanobacterium *Synechocystis* sp. PCC6803 exhibit distinct processing pathways involving at least two Cas6 and a Cmr2 protein. *PLoS One* 8: e56470.

Semenova E., M. Nagornykh, M. Pyatnitskiy, I.I. Artamonova and K. Severinov. 2009. Analysis of CRISPR system function in plant pathogen *Xanthomonas oryzae. FEMS Microbiol Lett* 296: 110–116.

Semenova E., M.M. Jore, K.A. Datsenko, A. Semenova, E.R. Westra, B. Wanner, J. van der Oost, S.J. Brouns and K. Severinov. 2011. Interference by clustered regularly interspaced short palindromic repeat (CRISPR) RNA is governed by a seed sequence. *Proc. Natl. Acad. Sci. USA* 108: 10098–10103.

Shah S.A., N.R. Hansen and R.A. Garrett. 2009. Distribution of CRISPR spacer matches in viruses and plasmids of crenarchaeal acidothermophiles and implications for their inhibitory mechanism. *Biochem. Soc. Trans.* 37: 23–28.

Shariat N. and E.G. Dudley. 2014. CRISPRs: molecular signatures used for pathogen subtyping. *Appl. Environ. Microbiol.* 80: 430–439.

Sinkunas T., G. Gasiunas, S.P. Waghmare, M.J. Dickman, R. Barrangou, P. Horvath and V. Siksnys. 2013. In vitro reconstitution of Cascade-mediated CRISPR immunity in *Streptococcus thermophilus. EMBO J.* 32: 385–394.

Sorek R., V. Kunin and P. Hugenholtz. 2008. CRISPR--a widespread system that provides acquired resistance against phages in bacteria and archaea. *Nat. Rev. Microbiol.* 6: 181–186.

Sorek R., C.M. Lawrence and B. Wiedenheft. 2013. CRISPR-mediated adaptive immune systems in bacteria and archaea. *Ann. Rev. Biochem.* 82: 237–266.

Soutourina O.A., M. Monot, P. Boudry, L. Saujet, C. Pichon, O. Sismeiro, E. Semenova, K. Severinov, C. Le Bouguenec, J.Y. Coppée and others. 2013. Genome-wide identification of regulatory RNAs in the human pathogen *Clostridium difficile. PLoS Genet* 9: e1003493.

Spilman M., A. Cocozaki, C. Hale, Y. Shao, N. Ramia, R. Terns, M. Terns, H. Li and S. Stagg. 2013. Structure of an RNA silencing complex of the CRISPR-Cas immune system.

Staals R.H., Y. Agari, S. Maki-Yonekura, Y. Zhu, D.W. Taylor, E. van Duijn, A. Barendregt, M. Vlot, J.J. Koehorst, K. Sakamoto and others. 2013. Structure and activity of the RNA-targeting Type III-B CRISPR-Cas complex of *Thermus thermophilus. Mol. Cell* 52: 135–145.

Stern A., L. Keren, O. Wurtzel, G. Amitai and R. Sorek. 2010. Self-targeting by CRISPR: gene regulation or autoimmunity? *Trends Genet.* 26: 335–340.

Sternberg S.H., R.E. Haurwitz and J.A. Doudna. 2012. Mechanism of substrate selection by a highly specific CRISPR endoribonuclease. *RNA* 18: 661–672.

Sternberg S.H., S. Redding, M. Jinek, E.C. Greene and J.A. Doudna. 2014. DNA interrogation by the CRISPR RNA-guided endonuclease Cas9. *Nature* 507: 62–67.

Swarts D.C., C. Mosterd, M.W. van Passel and S.J. Brouns. 2012. CRISPR interference directs strand specific spacer acquisition. *PLoS One* 7: e35888.

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

van der Ploeg J.R. 2009. Analysis of CRISPR in *Streptococcus mutans* suggests frequent occurrence of acquired immunity against infection by M102-like bacteriophages. *Microbiology* 155: 1966–1976.

van Duijn E., I.M. Barbu, A. Barendregt, M.M. Jore, B. Wiedenheft, M. Lundgren, E.R. Westra, S.J. Brouns, J.A. Doudna, J. van der Oost and others. 2012. Native tandem and ion mobility mass spectrometry highlight structural and modular similarities in clustered-regularly-interspaced shot-palindromic-repeats (CRISPR)-associated protein complexes from.

Vercoe R.B., J.T. Chang, R.L. Dy, C. Taylor, T. Gristwood, J.S. Clulow, C. Richter, R. Przybilski, A.R. Pitman and P.C. Fineran. 2013. Cytotoxic chromosomal targeting by CRISPR/Cas systems can reshape bacterial genomes and expel or remodel pathogenicity islands. *PLoS Genet* 9: e1003454.

Viswanathan P., K. Murphy, B. Julien, A.G. Garza and L. Kroos. 2007. Regulation of dev, an operon that includes genes essential for *Myxococcus xanthus* development and CRISPR-associated genes and repeats. *J. Bacteriol.* 189: 3738–3750.

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

Westra E.R., U. Pul, N. Heidrich, M.M. Jore, M. Lundgren, T. Stratmann, R. Wurm, A. Raine, M. Mescher, L. Van Heereveld and others. 2010. H-NS-mediated repression of CRISPR-based immunity in *Escherichia coli* K12 can be relieved by the transcription activator LeuO. *Mol. Microbiol.* 77: 1380–1393.

Westra E.R., P.B. van Erp, T. Künne, S.P. Wong, R.H. Staals, C.L. Seegers, S. Bollen, M.M. Jore, E. Semenova, K. Severinov and others. 2012. CRISPR immunity relies on the consecutive binding and degradation of negatively supercoiled invader DNA by Cascade and Cas3. *Mol. Cell* 46: 595–605.

Westra E.R., E. Semenova, K.A. Datsenko, R.N. Jackson, B. Wiedenheft, K. Severinov and S.J. Brouns. 2013. Type I-E CRISPR-cas systems discriminate target from non-target DNA through base pairing-independent PAM recognition. *PLoS Genet* 9: e1003742.

Wiedenheft B., G.C. Lander, K. Zhou, M.M. Jore, S.J. Brouns, J.van der Oost, J.A. Doudna and E. Nogales. 2011. Structures of the RNA-guided surveillance complex from a bacterial immune system. *Nature* 477: 486–489.

Wiedenheft B.,K. Zhou, M. Jinek, S.M. Coyle, W. Ma and J.A. Doudna. 2009. Structural basis for DNase activity of a conserved protein implicated in CRISPR-mediated genome defense. *Structure* 17: 904–912.

Wiedenheft B., S.H. Sternberg and J.A. Doudna. 2012. RNA-guided genetic silencing systems in bacteria and archaea. *Nature* 482: 331–338.

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

Yosef I., D. Shitrit, M.G. Goren, D. Burstein, T. Pupko and U. Qimron. 2013. DNA motifs determining the efficiency of adaptation into the *Escherichia coli* CRISPR array. *Proc. Natl. Acad. Sci. USA* 110: 14396–14401.

Zhang J., T. Kasciukovic and M.F. White. 2012a. The CRISPR associated protein Cas4 Is a 5' to 3' DNA exonuclease with an iron-sulfur cluster. *PLoS One* 7: e47232.

Zhang J., C. Rouillon, M. Kerou, J. Reeks, K. Brugger, S. Graham, J. Reimann, G. Cannone, H. Liu, S.V. Albers and others. 2012b. Structure and mechanism of the CMR complex for CRISPR-mediated antiviral immunity. *Mol. Cell* 45: 303–313.

Zhang Y., N. Heidrich, B.J. Ampattu, C.W. Gunderson, H.S. Seifert, C. Schoen, J. Vogel and E.J. Sontheimer. 2013. Processing-independent CRISPR RNAs limit natural transformation in *Neisseria meningitidis. Mol. Cell* 50: 488–503.