The CRISPRs, They Are A-Changin': How Prokaryotes Generate Adaptive Immunity

Edze R. Westra, Daan C. Swarts, Raymond H.J. Staals, Matthijs M. Jore, Stan J.J. Brouns, and John van der Oost

Laboratory of Microbiology, Department of Agrotechnology and Food Sciences, Wageningen University, 6703 HB Wageningen, The Netherlands; email: edze.westra@wur.nl, john.vanderoost@wur.nl

Annu. Rev. Genet. 2012. 46:311–39

The *Annual Review of Genetics* is online at genet.annualreviews.org

This article's doi: 10.1146/annurev-genet-110711-155447

0066-4197/12/1201-0311$20.00

Keywords

bacteria, CRISPR/Cas, small RNA, phage, defense, restriction modification, Abi

Abstract

All organisms need to continuously adapt to changes in their environment. Through horizontal gene transfer, bacteria and archaea can rapidly acquire new traits that may contribute to their survival. However, because new DNA may also cause damage, removal of imported DNA and protection against selfish invading DNA elements are also important. Hence, there should be a delicate balance between DNA uptake and DNA degradation. Here, we describe prokaryotic antiviral defense systems, such as receptor masking or mutagenesis, blocking of phage DNA injection, restriction/modification, and abortive infection. The main focus of this review is on CRISPR (clustered regularly interspaced short palindromic repeats)/Cas (CRISPR-associated), a prokaryotic adaptive immune system. Since its recent discovery, our biochemical understanding of this defense system has made a major leap forward. Three highly diverse CRISPR/Cas types exist that display major structural and functional differences in their mode of generating resistance against invading nucleic acids. Because several excellent recent reviews cover all CRISPR subtypes, we mainly focus on a detailed description of the type I-E CRISPR/Cas system of the model bacterium *Escherichia coli* K12.

INTRODUCTION

All viruses live in intimate association with their host cell, which they require for propagation. The symbiosis between a virus and its host is often thought of as a parasitic interaction. In this view, viruses exploit host resources for their own benefit at the expense of the host. For example, after intracellular proliferation, the new generation of viruses spreads through the population, leaving behind a heavily damaged or dead host cell. It is important to bear in mind, however, that despite their bad reputation, some viruses can have a conditionally beneficial effect on their host (134). Examples from the eukaryotic literature include virus-host relationships in which endogenous (para)retroviruses have advanced host evolution through the acquisition of viral genes (69) or in which they function in host defense against exogenic variants of the same virus or closely related viruses (53). Another more complex example of virus-host mutualism is the interaction between polydnaviruses and endoparasitoid wasps, where viral genes are required for successful development of the wasp egg in the lepidopteran insect host (48). Likewise, bacteria and phages can have a mutualistic relationship. In the case of some pathogenic bacteria, phages encode virulence factors required for bacterial invasion (reviewed in 22, 27). Intriguingly, a mutualistic aphid-bacterium interaction has been described that relies on a phage that is integrated in the bacterial genome; a toxin encoded by this prophage specifically kills invading parasites of the aphid host (123). Moreover, lysogenic phages encoding superinfection exclusion (Sie) systems can benefit a bacterial population because their rare excision from the host genome results in the release of phage particles that infect competing bacteria that do not contain a similar prophage, whereas the lysogen-carrying bacterial population is protected by Sie (see below) (21, 26, 92). In view of these examples of viruses conferring (conditional) advantages, the combination of host defense and recombination systems should be considered as a compromise between resistance against harmful DNA on the one hand and admission of useful DNA on the other.

To protect themselves against parasitic invaders, many species have developed multilayered antiviral defense systems. Mammals combat viruses by apoptosis of infected cells, clearance of infected cells by natural killer cells, antibody production by the cell-based adaptive immune system (reviewed in 6), and probably also by RNA interference (RNAi) (137). In plants and insects, infection by RNA viruses is mainly controlled by RNAi (reviewed in 96, 162). Plants also employ programmed cell death through the hypersensitive response (reviewed in 178), whereas insects use intracellular protein-mediated antiviral defense (29, 174). As described in more detail below, archaea and bacteria also contain numerous lines of defense against their viruses and phages, respectively (reviewed in 79, 91), such as receptor masking, blocking DNA injection (Sie), restriction/modification (R-M) and R-M-like systems, abortive infection (Abi), and the recently identified CRISPR (clustered regularly interspaced short palindromic repeats)/Cas (CRISPR-associated) system, which is the main focus of this review.

BACTERIAL AND ARCHAEAL DEFENSE SYSTEMS

Antiviral defense systems in prokaryotes are multilayered, acting at virtually all stages of the viral life cycle (**Figure 1**). The different types of prokaryotic defense systems evolve rapidly because of antagonistic coevolution between virus and host, which is best described as a combination of an arms race of defense and counter-defense and fluctuating selection on rare host and parasite genotypes (59, 67).

Preventing Viral Attachment

An effective and simple means of acquiring phage resistance is the blocking of phage adsorption to the host cell, either by masking the receptor (e.g., through the expression of

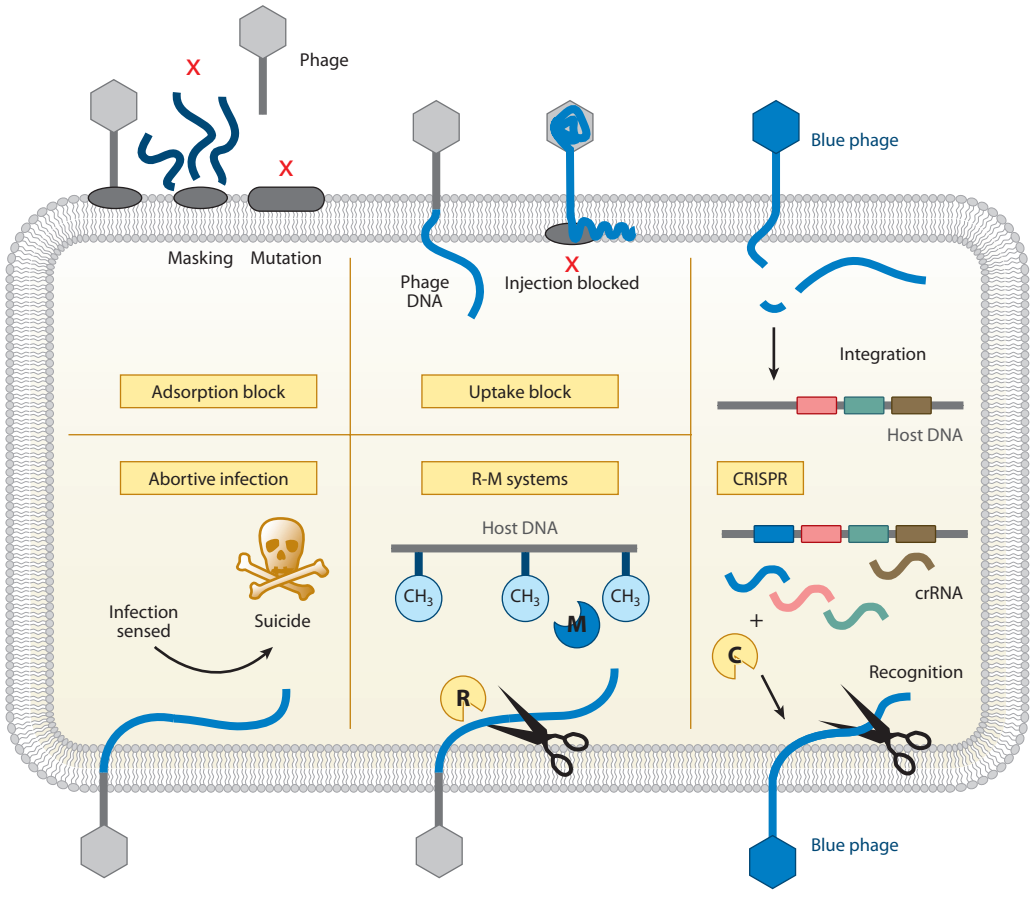

Figure 1

Schematic overview of prokaryotic defense systems. Bacterial cells contain several independent mechanisms to defend themselves against phage infection (or other invading DNA, such as conjugative plasmids). Defense mechanisms include blocking of phage adsorption or DNA injection. Other systems act directly on the invader DNA, such as restriction/modification and CRISPR (clustered regularly interspaced short palindromic repeats)/Cas (CRISPR-associated). Abortive infection systems are a form of altruistic defense causing cell death upon infection as a sacrifice to protect the rest of the population. These defense systems can act independently of each other, but cross talk between different systems may be possible. Note that the infecting blue phage contains a DNA fragment that has a sequence identical to the blue fragment in the CRISPR locus of the host cell. This coloring distinguishes it from the gray phages that are destroyed by other mechanisms. Abbreviations: C, Cas protein; M, methyltransferase; R, restriction endonuclease.

polysaccharides on the cell surface) or by downregulating or mutating receptor molecules (reviewed in 79). Examples of phage resistance by receptor masking have been described for *Aeromonas salmonicida* (81), *Staphylococcus aureus* (121, 122), *Escherichia coli* (132), and *Lactococcus* spp. (reviewed in 52). Some phages counteract receptor masking by carrying carbohydrate-degrading enzymes (138, 154). More direct alterations to phage receptors (mutagenesis, downregulation, and loss) usually come with (conditional) fitness costs or can make the host more susceptible to infection by other viruses (9). Nonetheless, this is a commonly encountered mechanism to escape phage predation, which phages can counter by mutations leading to recognition of the altered receptor or another receptor (reviewed in 73). This is nicely exemplified by a recent phage λ and *E. coli* coevolution experiment where, after *E. coli*

evolved to downregulate the natural receptor (LamB), the phage evolved through several point mutations in its receptor binding protein (J) to bind a new receptor (OmpF) (110).

Blocking DNA Injection

Obstruction of the entry of the viral genome into the host's cytoplasm is a second line of defense. The proteins blocking DNA injection are usually localized in association with or in close proximity to the membrane/cell wall and can be encoded by a plasmid (57) or by a prophage (107; 167; reviewed in 52, 97). In the latter case, the process is referred to as Sie. Sie was originally described for T-even phages and their *E. coli* hosts (45, 55) but has since also been found in other phages, such as those infecting lactic acid bacteria (52). Although Sie is often considered to be important at the level of phage-phage interactions, this system can in fact provide a selective advantage to the part of the bacterial population carrying the prophage (26). The bacteria can multiply for many generations while carrying the prophage. At some stage, upon specific (stress) stimuli, some prophages excise, multiply, and proliferate (killing the host cell). As such, the prophage is used as a weapon to infect and kill competing (prophage-free) bacteria, whereas the prophage-carrying fraction of the population remains intact because of Sie. Under these conditions, the host bacterium and Sie-encoding prophage have a mutualistic relationship (reviewed in 134).

Digestion of Nonself DNA

If phage adsorption and DNA injection are not prevented, intracellular defense systems may act directly on the viral DNA to neutralize the invader. The R-M system is a broad-range prokaryotic immune system that targets DNA. A typical R-M system consists of a DNA methyltransferase, which modifies specific DNA sequences, and a restriction endonuclease, which cleaves the same sequences when unmodified. The general principle of these systems is that the host's genomic DNA is methylated and hence protected against cleavage, whereas exogenic DNA is unmodified and restricted. Three major R-M types have been identified that follow this classical scheme. In addition, a fourth type has been identified that consists of restriction enzymes cleaving modified DNA only (reviewed in 133). More recently, an R-M-like system, known as phage growth limitation (Pgl), has been found in *Streptomyces coelicolor*. Pgl somehow modifies offspring phage DNA, such that it is targeted for destruction when these phages infect cells that also carry the Pgl system (153). The Pgl system is made up of at least four genes, one of which resembles a methyltransferase (78). The molecular mechanism of this interesting new system remains to be determined.

Type I R-M systems require *S*-adenosylmethionine (AdoMet) for DNA methylation and Mg^{2+} and ATP for endonuclease activity. These systems encode a nuclease (hsdR), a DNA methyltransferase (hsdM), and a recognition sequence–binding specificity subunit (hsdS) that together form a 400–500 kDa multiprotein complex with an $hsdR_2hsdM_2hsdS_1$ stoichiometry. Type I restriction sites are asymmetric and bipartite, with a separation into two fragments of 3–4 bp and 4–5 bp interspaced by 6–8 bp of nonspecific sequence (reviewed in 117). Although hemimethylated sequences trigger methyltransferase activity by the hsdM subunits (164), unmethylated sequences induce ATP-dependent DNA translocation by the hsdR subunits, and the hsdS subunit remains bound to the restriction site, leading to loop formation in the DNA. Subsequent cleavage is catalyzed by the hsdR subunits at sites remote from the restriction site when the R-M complex collides with a second R-M complex or when another collision block occurs (152; reviewed in 117).

Type II R-M systems are the best-known restriction/modification systems because they have revolutionized molecular cloning and are presently utilized in virtually all laboratories employing recombinant DNA techniques. In contrast to type I and type III R-M systems,

type II R-M systems are ATP-independent and cleave DNA at a well-defined position, most often within or very near to the restriction site (reviewed in 126). Restriction sites are usually 4–8 bp palindromic sequences. Often Mg^{2+}-dependent cleavage requires restriction endonucleases to form homodimers or homotetramers to generate symmetrical cleavage products with either blunt or sticky ends. Again, sequences are protected from DNA modification by a (generally) separately encoded DNA methyltransferase. In contrast to the other types, type II methyltransferases and restriction endonucleases work completely independently (reviewed in 126), therewith sharing certain properties with selfish addiction modules, such as toxin-antitoxin systems (118).

Type III R-M systems consist of a *mod* gene encoding a DNA methyltransferase and a *res* gene encoding a restriction endonuclease. Complexes with a Mod_2Res_2 stoichiometry are stimulated by AdoMet, and cleavage activity requires ATP, Mg^{2+}, and two inversely oriented unmethylated nonpalindromic recognition sequences (5–6 bp). The methyltransferase provides protection by catalyzing methylation of one strand of the recognition sequence (11, 23, 80). During DNA replication some sites become unmodified but remain protected, as these unmodified sites are never inversely oriented (109). Type III endonucleases cleave 25–27 bp downstream of the recognition site. Interestingly, expression of *mod* genes is phase variable in some organisms, which might be a way of downregulating host defense to allow for (potentially beneficial) DNA uptake (7, 43). Alternatively, these phase-variable type III R-M systems may have evolved into regulatory systems that control expression of target genes through their methylation status (reviewed in 54, 148).

Phages can escape R-M systems in a number of ways: (*a*) by mutagenesis of target sequences, (*b*) by carrying inhibitors of restriction endonucleases, (*c*) by carrying enzymes that hydrolyze R-M cofactors (e.g., AdoMet), (*d*) by incorporating modified nucleotides in their DNA (reviewed in 166), or (*e*) by encoding their own (multispecific) methyltransferases (reviewed in 18, 89). It is possible that type IV R-M systems, which specifically recognize methylated DNA or modified nucleotide-containing DNA (12, 133), evolved as a counter defense against these escape mechanisms. As type IV systems recognize and cleave modified DNA, these systems do not encode a modification enzyme.

Abortive Infection

Similar to the apoptosis mechanism in eukaryotic cells, the widespread Abi systems in prokaryotes induce cell death upon viral infection as a sacrifice to protect other cells of the same species. In *E. coli* several Abi systems have been described, either consisting of a single protein (e.g., LitA, PrrC, PifA) or, in the case of Rex, consisting of a two-component system (reviewed in 91, 115). Lactic acid bacteria also contain many Abi systems, with *Lactococcus lactis* alone containing as many as 20 Abi systems (AbiA to AbiU), most of which consist of single proteins (reviewed in 34). In general, the toxic effect of an Abi system is conferred by its cleavage of essential cellular components, therefore necessitating a tight regulation of these Abi systems. Interestingly, Fineran and coworkers (51) recently discovered a new class of toxin/antitoxin (TA) systems [protein toxin (ToxN)/RNA antitoxin (ToxI); type III TA systems] that functions as an Abi system. It has been proposed that the phage triggers disruption of ToxIN transcription or affects stability of the ToxIN complex, causing mobilization of the toxin, which allows it to exert its effect (19).

Adaptive Immunity

CRISPR/Cas, the latest addition to the list of known prokaryotic anti-invader systems, appears to be the most complex of these systems. The CRISPR/Cas system is unique in being the only adaptive and inheritable prokaryotic immune system known to date. Below, we briefly describe its discovery and summarize our current understanding of the system (for extensive reviews, see 5, 17, 76, 87, 104, 158, 160) before we elaborate in detail on the type I-E CRISPR/Cas system.

CRISPR/Cas DISCOVERY

In 1987, Ishino and coworkers (82) discovered an unusual structure of repetitive DNA downstream from the *E. coli iap* gene consisting of invariant direct repeats (29 nt) and variable spacing sequences (32 nt). Although similar repeat clusters were subsequently identified in *Haloferax mediterranei*, *Streptococcus pyogenes*, *Anabaena*, and *Mycobacterium tuberculosis* (61, 75, 106, 114, 161), their function remained unknown (113). Jansen and coworkers (83) coined the term CRISPR and reported that CRISPRs colocalize with specific *cas* genes. CRISPR/Cas systems are exclusively found in prokaryotes (84)—in approximately half of the bacteria and almost all archaea (60). Although different systems were proposed to classify the typical combinations of *cas* genes (63, 100), a recent classification defines three main types, each with two or more subtypes (101). The emergence of bacterial and archaeal genome sequences on the one hand and viral and plasmid sequences on the other resulted in the key discovery that CRISPR spacers resemble fragments of foreign genetic elements, suggesting that the spacers were derived from invading genomes (20, 112, 129). These observations, together with the detection of CRISPR locus transcripts with defined lengths of one or more spacer repeat units (93, 156, 157) and predicted nucleic acid–related activities for many of the *cas* genes, led Makarova and coworkers (100) to propose that CRISPR/Cas neutralizes invaders via a mechanism reminiscent of RNAi. Soon thereafter, Barrangou and coworkers (14) provided the first experimental evidence that the CRISPR/Cas system of the lactic acid bacterium *Streptococcus thermophilus* functions as an inheritable, adaptive prokaryotic immune system.

Given the long history of the study of phage biology, it may appear puzzling that the CRISPR/Cas immune system was not discovered earlier. To some extent, this is due to the rather tight regulation of CRISPR/Cas systems in model organisms, such as *E. coli* and *Salmonella enterica*, which causes their expression to be silenced under normal laboratory growth conditions. Genomics research since 1995 has resulted in the elucidation of the ubiquitous repetitive CRISPR loci and the associated gene clusters. Hence, the only adaptive and inheritable prokaryotic immune system known to date has remained hidden long enough to grant present-day researchers the opportunity to explore this exciting field. Indeed, during the last six years, CRISPR research has been flourishing, resulting in a rapid gain of insight in the molecular mechanisms of this highly diverse defense system.

CRISPR ARRAYS

A CRISPR array consists of host-derived repeating sequences of typically 30 bp that are interspaced by similar-sized acquired spacer sequences. A CRISPR locus may vary in length from just a few to several hundred repeat-spacer units. The repeat sequences have been classified in 12 different clusters, some of which are predicted to form a hairpin secondary structure, whereas others are predicted to be unstructured (90). Usually a CRISPR/Cas subtype associates with a single repeat cluster, although exceptions to this have been reported (see below). CRISPR loci are generally flanked by an adenine/thymine (AT)-rich leader sequence (83), which has been shown to contain promoter elements (65, 94, 128, 131) and binding sites for regulatory proteins (128, 131). Moreover, the observed polarized acquisition of new spacer sequences at the leader end of the CRISPR array gave rise to the speculation that leader sequences may contain binding sites for Cas proteins involved in CRISPR adaptation (5).

TYPE I, TYPE II, AND TYPE III CRISPR/CAS SYSTEMS

Careful reevaluation of the various sets of *cas* genes has revealed that CRISPR/Cas systems cluster into three basic types (Type I, Type II, and Type III) (**Figure 2**), which are further divided into at least ten subtypes (Types I-A–F, Types II-A–B, and Types III-A–B). The genes that are present in each CRISPR/Cas variant are *cas1* and *cas2* (101). The mechanism of all CRISPR/Cas systems is divided into

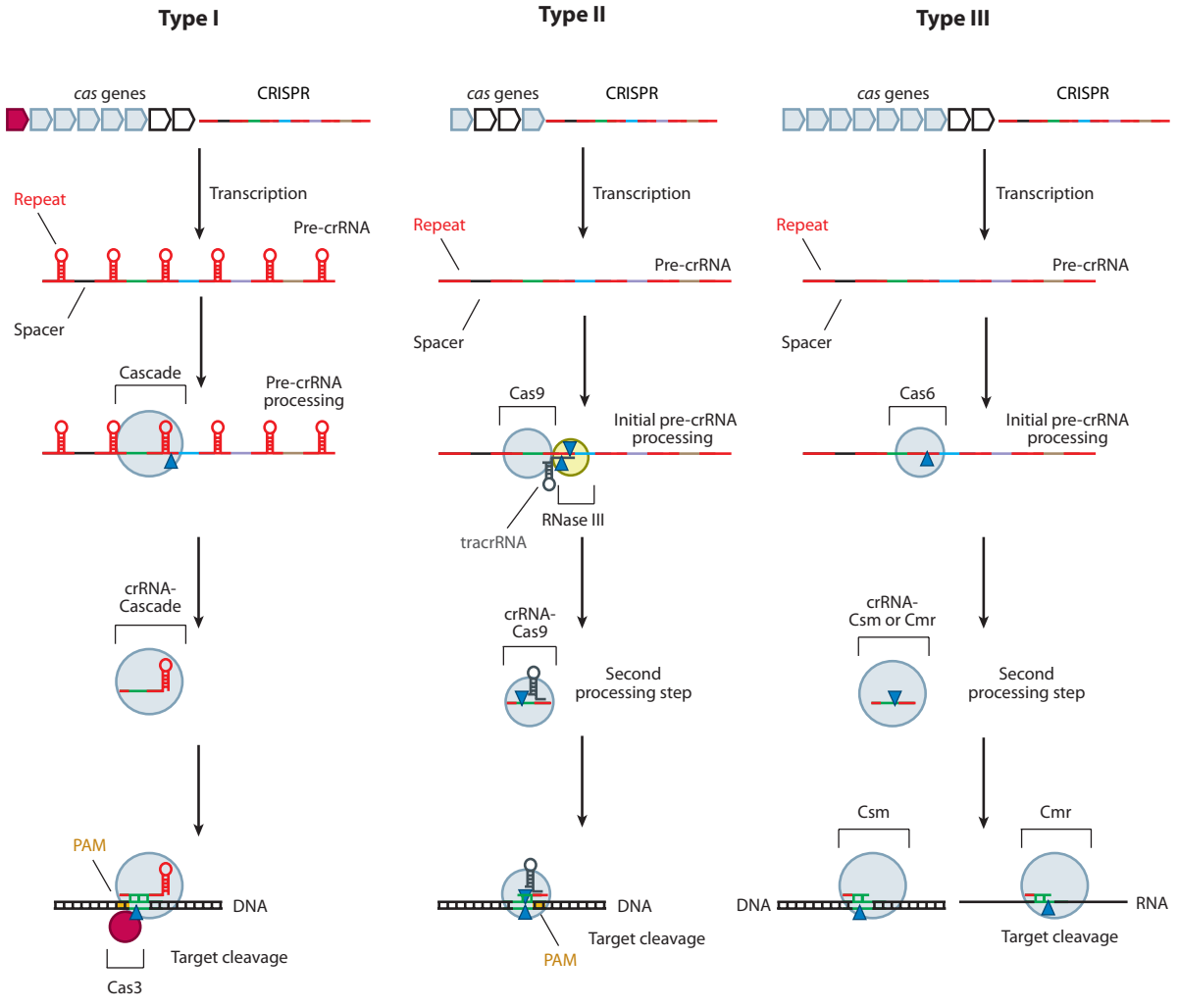

Figure 2

Schematic overview of the Type I, II, and III CRISPR (clustered regularly interspaced short palindromic repeats) expression and interference stages. Three main types of CRISPR/Cas (CRISPR-associated) systems are now known, and these display major mechanistic differences. In this schematic overview, the expression and interference stages are depicted for the three different types. Transcription of the CRISPR gives rise to a pre-crRNA molecule, which is subsequently cleaved in the repeat sequence by a Cas6 homolog in Type I (often a subunit of Cascade-like complexes) and Type III systems, and by RNase III and Cas9 in Type II systems. The generated crRNA molecules undergo a further processing event in Type II and Type III systems. Mature crRNA molecules are bound to a Cas protein or protein complex. In the case of Type III-B, the CRISPR RAMP Module (Cmr) complex binds and cleaves complementary RNA, whereas the other types bind and cleave dsDNA (double-stranded DNA). Targets contain a protospacer-adjacent motif (PAM) either downstream (Type I) or upstream (Type II) of the protospacer.

three stages: adaptation, expression, and interference. During the adaptation stage, resistance is acquired by integration of a new spacer sequence in a CRISPR array. During the expression stage, *cas* genes are transcribed and translated; in addition, CRISPRs are transcribed into precursor CRISPR RNAs (pre-crRNAs) that are subsequently cleaved by a Cas6 homolog in Type I and Type III systems (25, 31, 70, 71, 130) and by an RNase III in Type II systems (38). Mature crRNA contains (part of) a single spacer sequence, and hence

can recognize only a single target. During the interference stage, the crRNA guides one or more Cas proteins to cleave complementary nucleic acids.

Type I systems are characterized by the presence of Cas3, comprising a histidine/aspartate (HD)-nuclease domain and a DExH helicase domain, which can also be encoded in separate open reading frames or can be fused to other Cas proteins (100, 101) (see below). All Type I systems are thought to form a crRNA-guided surveillance ribonucleoprotein complex similar to the Cascade complex identified in Type I-E of *E. coli* K12 (25). Type I-A (archaeal Cascade, here referred to as IA-Cascade) and Type I-F (Csy complex, here referred to as IF-Cascade) Cascade-like complexes have been recently described (95, 171). Type I-A CRISPR/Cas systems have been studied in the hyperthermophilic crenarchaeon *Sulfolobus solfataricus* (62, 93, 94, 141). Although CRISPR transcript levels appear to be modulated by a repeat-binding protein, Cbp1 (39, 124), the system is constitutively expressed under laboratory growth conditions and protects against virus infection, probably by targeting DNA (102). Reverse transcripts of the CRISPR have also been detected, but their functional significance remains to be determined (94, 179). Lintner and colleagues (95) reported a *Sulfolobus* IA-Cascade complex consisting of at least four different Cas protein subunits (Csa5, Cas7, Cas5, and Cas6) loaded with crRNA of 60–70 nt derived from CRISPR loci belonging to any of three different CRISPR repeat clusters present in *S. solfataricus*. As in *E. coli* IE-Cascade, which has a Cas7 backbone of six copies (see below), the Cas7 subunit is most abundant in IA-Cascade. The Cas6 subunit cleaves pre-crRNA at a single position, yielding spacer sequences flanked by a repeat-derived 5′-handle of 8 nt and 3′-handle of 16–17 nt [identical to Cas6-mediated pre-crRNA cleavage in *Pyrococcus furiosus* (31)]. The *Sulfolobus* IA-Cascade binds ssDNA (single-stranded DNA) targets with a sequence complementary to the crRNA guide (95), but dsDNA (double-stranded DNA) binding remains to be determined. Recently, an IC-Cascade complex has been isolated from *Bacillus halodurans* (118b). A Cas5d subunit of this complex serves as an endoribonuclease that cleaves the pre-crRNA (118a). So far, very little literature exists on type I-B and type I-D systems. The Type I-E system has been extensively studied, mostly using *E. coli* K12 as a model organism (see below). The type I-F system has been studied in *Pseudomonas aeruginosa*; a CRISPR and *cas* gene–dependent effect on prophage-mediated inhibition of swarming and biofilm formation was reported (see section, Other Functions of CRISPR/Cas) (28, 177). More recently, CRISPR-based immunity of *P. aeruginosa* against phage infection was also demonstrated (27a). Type I-F Cas proteins have been studied biochemically and structurally. Co-crystal structures of crRNA and Cas6f, a metal-independent pre-crRNA endoribonuclease (71, 130), revealed specific protein-RNA interactions between Cas6f and the hairpin-forming crRNA repeat (71) and provided insight into the mechanism of pre-crRNA cleavage by Cas6f (72). In addition, a *Pseudomonas* IF-Cascade was recently isolated that consists of four different proteins ($Csy1_1$, $Csy2_1$, $Csy3_6$, and $Cas6f_1$) with a low-resolution structure and dsDNA-binding properties somewhat resembling those of *E. coli* IE-Cascade (171).

Type II systems are the most compact, encoding Cas1, Cas2, Cas9, and either Csn2 (Type II-A) or Cas4 (Type II-B) (**Figure 2**). The first evidence for CRISPR-dependent immunity came from the Type II-A system of *S. thermophilus*, where both Csn2-dependent CRISPR adaptation and Cas9-dependent interference were observed upon phage infection (14). The recently resolved structure of Csn2 shows that it forms a tetrameric ring-like structure that is 26 Å wide and that contains conserved lysine residues involved in binding dsDNA (49, 119). Until recently, the expression stage in Type II systems was poorly understood owing to the lack of a Cas6-like pre-crRNA endoribonuclease. A recent paper by Deltcheva and coworkers (38) resolved this issue, revealing an unexpected crRNA maturation

pathway that requires an 89- or 171-nt *trans*-acting CRISPR-associated RNA (named tracrRNA) that base-pairs with the pre-crRNA repeat sequences. Subsequent recruitment of a housekeeping RNase III catalyzes, together with Cas9, cleavage of the pre-crRNA in the repeat; this yields 66-nt crRNA molecules that are further trimmed at the 5′ end to produce 39–42-nt mature crRNA containing a 20-nt spacer sequence (38). During Type II CRISPR interference, target DNA is cleaved in the protospacer sequence (56). A recent study on the type II system of *S. pyogenes* demonstrated that this cleavage is carried out by a ribonucleoprotein complex consisting of Cas9, the mature crRNA, and the tracrRNA (84a). It was shown that the HNH nuclease domain of Cas9 cleaves the base-pairing strand, and the RuvC nuclease domain cleaves the displaced strand, resulting in a blunt-end cleavage product (84a).

Two Type III systems exist, Type III-A and Type III-B, which differ most notably in the nature of their target nucleic acid. Although Marraffini & Sontheimer (103) demonstrated that the Type III-A system of *Staphylococcus epidermidis* targets DNA, the Type III-B system of *P. furiosus* was shown to cleave RNA both in vitro (66) and in vivo (65), which is the first example of a prokaryotic immune system targeting RNA. The ribonucleoprotein effector complex consists of six different Cas proteins [Cmr (CRISPR RAMP Module 1), Cas10, Cmr3, Cmr4, Cmr5, and Cmr6] and mature crRNA of either 39 nt or 45 nt (66). Recently, an analogous crRNA-loaded Cmr complex, containing seven Cmr proteins (Cmr1–7), with manganese-dependent endoribonuclease activity on complementary RNA, was isolated from *S. solfataricus* (179). Detailed studies of the expression stage of type III-B demonstrated that Cas6 binds crRNA at the 5′ end of the unstructured repeat sequence (30) and wraps the crRNA around its surface to position the well-defined cleavage site in the Cas6 catalytic center (165), yielding 67-nt crRNAs through metal-independent endoribonuclease activity (31). These are further trimmed at the 3′ end through unknown mechanisms to 39-nt and 45-nt mature crRNA species that are complexed with Cas proteins (64). In the *S. epidermidis* Type III-A systems, pre-crRNA maturation also comprises a primary Cas6-mediated sequence-specific processing step at the base of a putative stem-loop structure in the pre-crRNA repeat (70, 103), followed by a ruler-based sequence-unspecific crRNA maturation step that involves crRNA trimming at the 3′ end, yielding mature crRNA of two defined lengths (37 nt and 43 nt) (70).

SELF/NONSELF DISCRIMINATION

Examination of protospacer-flanking sequences targeted by Type I and Type II CRISPR/Cas systems led to the identification of conserved sequence motifs, coined protospacer-adjacent motifs (PAMs) (20, 77, 111), that are essential for CRISPR interference (40). In contrast, Type III CRISPR interference appears to lack such a PAM requirement (101). Further analysis revealed a correlation between PAM sequences and specific repeat clusters (111) and hence with CRISPR/Cas subtypes (90) (**Table 1**).

Although Type I systems require PAM sequences of 2 or 3 nt at the 3′ end of the protospacer on the target DNA strand (see sidebar, A Modified Definition of Protospacer Sequences) (111), Type II systems require a PAM sequence of 4 or 5 nt at the 5′ end of the protospacer on the target DNA strand (20, 40, 56, 77). Some recent studies indicate that in the type I-B and I-E systems, some variation in PAM motifs is allowed (51a, 155, 169). For most CRISPR/Cas types, the sequence requirements, function, and recognition of PAMs are poorly understood. For type I-E systems, it has been shown that PAM recognition takes place only in the base-pairing strand (136, 169), through a Cascade subunit (Cse1) that specifically interacts with the PAM (136). However, type II systems recognize the PAM in the displaced strand (84a). PAMs could serve a dual role in selecting appropriate protospacers during CRISPR adaptation as well as in distinguishing CRISPR

Table 1 Protospacer-adjacent motif (PAM) sequences identified for CRISPR/Cas subtypes

Type	Species	References	PAM (5′–3′)[a]	Typical repeat	CRISPR cluster	
I-A	*Sulfolobus solfataricus* P2	(62, 94, 111)	Protospacer-NGG	GATAATCTCTTA TAGAATTGAAAG[b]	7	
	Metallosphaera sedula DSM5348	(111)	Protospacer-NGG	GTTAATCTTCTAT AGAGTTGAAAG	7	
			Unknown		11	
I-B	*Methanothermobacter thermautotrophicus* ΔH	(111)	Protospacer-NGG	GTTAAAATCAGA CCAAAATGGGA TTGAAAT	1	
	Listeria monocytogenes	(111)	Protospacer-NGG	GTTTTAACTACTT ATTATGAAATCT AAAT	1	
			Unknown		6,9	
I-C	*Streptococcus pyogenes*	(111)	Protospacer-GAA	GTCTCACCCTTC ATGGGTGAGTG GATTGAAAT	3	
	Xanthomonas oryzae	(111)	Protospacer-GAA	GTCGCGTCCTCA CGGGCGCGTGG ATTGAAAC	3	PAM downstream of protospacer
I-D			Unknown		Unknown	
I-E	*Escherichia coli* K12	(111, 139, 155, 169)	Protospacer-CTT Protospacer-CAT Protospacer-CCT Protospacer-CTC	GWGTTCCCCGCG CCAGCGGGGAT AAACCG[b]	2	
	Pseudomonas aeruginosa 2192	(111)	Protospacer-CTT	GTGTTCCCCACA TGCGTGGGGAT GAACCG	2	
I-F	*P. aeruginosa* PA14	(27a, 111)	Protospacer-GG	GTTCACTGCCGT GTAGGCAGCTA AGAAA[b]	4	
	Shewanella spp.	(111)	Protospacer-GG	GTTCACCGCCGC ACAGGCGGCTT AGAAA	4	
II-A	*Streptococcus thermophilus*	(77)	WTTCTNN - protospacer	GTTTTTGTACTCT CAAGATTTAAGT AACTGTACAAC	10	PAM upstream of protospacer
	S. thermophilus	(20)	TTTYRNNN - protospacer	GTTTTTGTACTCT CAAGATTTAAGT AACTGTACAAC	10	
II-B	*S. thermophilus*	(77)	CNCCN - protospacer	GTTTTAGAGCTG TGTTGTTTCGAA TGGTTCCAAAAC	10	
	S. pyogenes	(111)	CCN - protospacer	GTTTTAGAGCTA TGCTGTTTTGAA TGGTCCCAAAAC[b]	10	

(*Continued*)

Table 1 (*Continued*)

Type	Species	References	PAM (5′–3′)[a]	Typical repeat	CRISPR cluster	
	L. monocytogenes	(111)	CCN - protospacer	GTTTTAGAGCTA TGTTATTTTGAA TGCTACCAAAAC	10	
III-A	*Staphylococcus epidermidis*	(105)	No PAM	GATCGATACCCA CCCCGAAGAAA AGGGGACGAGAAC[b]	8	No PAM
III-B	*Pyrococcus furiosus*	(65)	No PAM	GTTCCAATAAGA CTAAAATAGAA TTGAAAG[b]	6	
	S. solfataricus	(179)	No PAM	GATTAATCCCAA AAGGAATTGAA AG[b]	7	

[a]PAM and protospacer correspond to sequence on target strand (i.e., strand that base-pairs with crRNA).
[b]Direction of CRISPR transcription verified.

spacers from their respective protospacers during CRISPR interference (85, 160).

On the one hand, the absence of a PAM requirement for the RNA-targeting type III-B CRISPR/Cas system may be explained by the fact that self-targeting of mRNA is less deleterious or may even serve a regulatory purpose (65). Type III-A systems, on the other hand, have a fundamentally different mechanism to discriminate self (the CRISPR DNA) from nonself (any other DNA) (105) that relies on base-pairing between the crRNA repeat sequence and the sequence flanking the protospacer. When three positions of the repeat fragment in the 5′ handle of the crRNA molecule base pair with a target nucleotide sequence, CRISPR interference is prohibited; mismatches at these positions trigger CRISPR interference (105). The III-A system therefore targets almost all randomly provided DNA sequences with a complementary protospacer, whereas PAM-dependent systems target only those protospacers having the PAM motif.

THE TYPE I-E CRISPR/Cas SYSTEM

Type I-E systems can be found in some actinobacteria, firmicutes, and methanogenic archaea as well as in many proteobacteria, including *E. coli* K12. All three stages of CRISPR defense have been studied for the type I-E system, placing it among the best understood CRISPR/Cas systems to date.

A MODIFIED DEFINITION OF PROTOSPACER SEQUENCES

Protospacer sequences were initially defined as sequences identical to the spacer sequences (40). We here refine this initial definition by proposing that the protospacer is the sequence complementary to the crRNA (CRISPR RNA) spacer (the reverse complement of the initially defined protospacer). This definition takes into account the direction of transcription of the CRISPR and thereby eliminates ambiguities of orientation and strand of the spacers. In addition, it has been demonstrated for several subtypes that CRISPR/Cas target recognition takes place by base-pairing between the crRNA spacer and the complementary nucleic acid sequence of a target (66, 86, 171). In the case of Type III-B systems, single-stranded nucleic acids (RNA) appear to be targeted (65, 66, 179). However, the established targeting of single-stranded nucleic acids imposes the need to define protospacer sequences as the sequence complementary to a crRNA spacer. Moreover, protospacer-adjacent motif (PAM) recognition in Type I-E systems takes place specifically in the base-pairing/target strand (169). This implies that the main interactions with target DNA involve the strand base-pairing with the crRNA, arguing for using the term protospacer to refer to this strand of the DNA.

Adaptation

CRISPR adaptation under laboratory conditions was first observed in the Type II system of *S. thermophilus* (14, 56, 77). Comparative genomics has revealed a variable spacer content in two CRISPR loci of *E. coli* strains (41, 159), suggesting that *E. coli* acquires new spacers over time. Indeed, very recently three groups have independently demonstrated CRISPR adaptation in the Type I-E system of *E. coli* (36, 155, 176).

Although CRISPR adaptation is dormant in wild-type *E. coli* K12 because of repression of *cas* gene expression (see below), spacer acquisition could be observed in cells in which *cas1* and *cas2* expression levels were elevated (36, 155). In addition, spacer acquisition was observed in *E. coli* BL21-Al cells in which *cas1* and *cas2* were overexpressed (176). In contrast, *E. coli* BL21-Al cells in which only *cas1* or only *cas2* was overexpressed were deficient in spacer integration, indicating that Cas1 and Cas2 together are responsible for CRISPR adaptation (176). Consistent with their role in CRISPR adaptation, *cas1* and *cas2* are invariably associated with CRISPR loci (101) and are not required for crRNA processing or CRISPR interference (25). Sometimes Cas1 and Cas2 are fused to each other or to other proteins (101, 127). In the *Thermoproteus tenax* type 1-A system, a complex consisting of the Cas1-Cas2 fusion protein, Cas4, and Csa1 has been reported and was termed Cascis (CRISPR-associated complex for spacer integration) (127).

So far, *E. coli* CRISPR-adaptation studies have been reported only for nonlytic bacteriophages and plasmids (36, 155, 176). Abedon (1) suggested that, in the case of a lytic phage infection, CRISPR adaptation may be preceded by a facilitation step. This step supports bacterial survival upon initial exposure to a new phage followed by degradation of the phage DNA, providing the cell with a DNA substrate from which new spacer sequences can be acquired.

As in *S. thermophilus* (14, 56, 77), new spacers are integrated in a polar fashion at the leader end of the CRISPR locus (36, 41, 155, 159, 176), suggesting that the integration machinery interacts directly with this region (5). Indeed, the first repeat and 60 bp of the leader sequence located immediately upstream of the first repeat are essential for integration of new spacers (176). Moreover, it has been demonstrated that the first repeat is duplicated during spacer integration (36, 176).

Integrated spacers were acquired from protospacers located on plasmid DNA (36, 155, 176) and phage DNA (36). Pre-spacers (5) of at least 33–34 bp are selected in a nonrandom process to target protospacers flanked by a CTT PAM, whereas the selection of the strand, locus, and nucleotide content of the protospacer appears to be random (36, 155, 176). Spacers are integrated in both CRISPR loci of *E. coli* K12 that have a leader sequence, and spacers from both loci actively contribute to CRISPR interference (36, 155); in the case of *E. coli* BL21-Al new spacers were integrated in only one of the two CRISPR loci (176), which is explained by the fact that the other CRISPR locus does not have an upstream leader sequence. The adaptation machinery appears to lack an intrinsic ability to distinguish between invading DNA and genomic DNA: Self-targeting spacers are readily acquired during phage infection, which causes autoimmunity and eventually cell death (149), hence imposing a need for tight regulation of CRISPR adaptation.

Often multiple spacers against the same target are integrated in a single clone (36, 155). Interestingly, the presence of the first targeting spacer has been found to accelerate acquisition of subsequent spacers from this target: a positive feedback loop mechanism referred to as priming (36, 155). Priming has been observed in response to escape mutagenesis of phages (36). The observation that all *E. coli* Cas proteins (Cas1, Cas2, Cascade and Cas3) are required for priming (36) is indicative of cross talk between the CRISPR interference and CRISPR adaptation pathways. Based on the priming phenomenon, it has been suggested that degradation of the invader DNA, mediated by the first integrated spacer, yields degraded DNA fragments suitable for acquisition of secondary spacers (155). Moreover, secondary

acquired spacers always target the same DNA strand as the primary integrated spacer, indicating that spacer acquisition after priming is a strand-specific process (36, 155). The acquisition of multiple spacers provides enhanced resistance and lowers the chance of invader escape by point mutations (25, 36, 155), hence providing an evolutionary benefit to the host.

Interestingly, the complement of the protospacer-flanking nucleotide of the PAM is always conserved in the spacer-flanking nucleotide of the repeat (155). For example, integration of a spacer containing the prevalent CTT PAM co-occurs with a repeat with a CCG 3′-end, whereas deviations from this dominant PAM (e.g., ATT or TTT) co-occur with corresponding mutations in the last nucleotide of the proximal repeat (CCT or CCA, respectively). This observation indicates that the last nucleotide of the repeat is derived from the PAM of the invading DNA during spacer acquisition rather than being copied during repeat duplication (**Figure 3**). Polymorphisms at this last position of the repeat are not

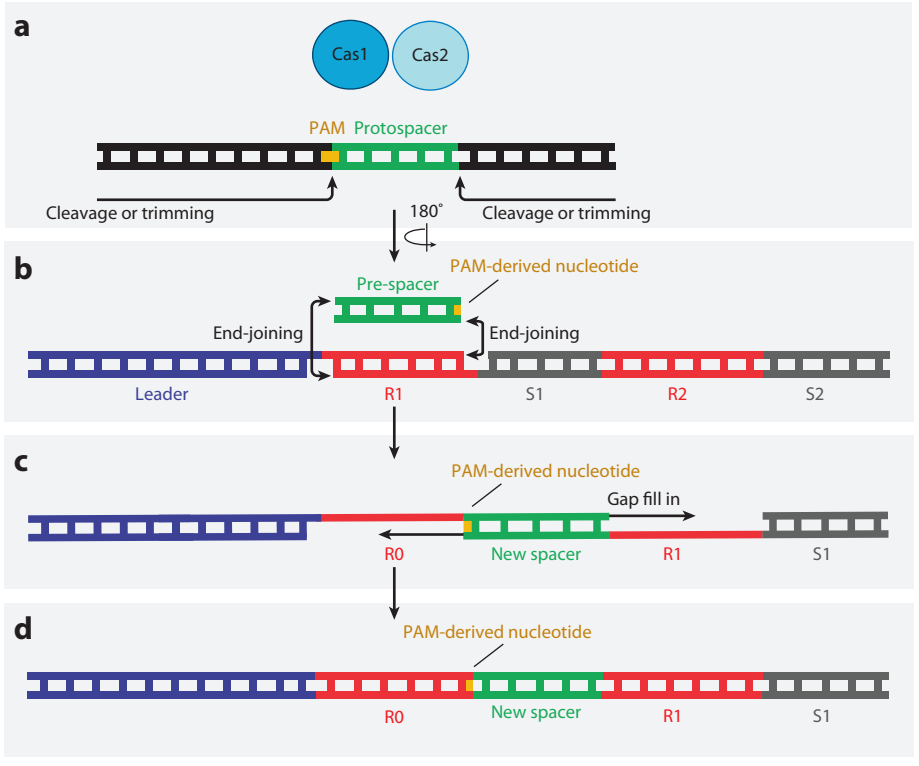

Figure 3

Hypothetical mechanism of CRISPR (clustered regularly interspaced short palindromic repeats) adaptation. (*a*) Spacer selection. DNA fragments containing a protospacer-adjacent motif (PAM) sequence are recognized. Cas1 and Cas2 are potentially involved in DNA cleavage or trimming, generating a pre-spacer containing at least the protospacer-proximal nucleotide of the PAM and the full protospacer sequence. The pre-spacer can either be single stranded or double stranded. (*b*) Spacer incorporation. The integration machinery (probably including Cas1 and/or Cas2) recognizes the leader strand and integrates the pre-spacer, most likely by end joining of either the 3′ or the 5′ ends of the repeat to the 5′ or 3′ ends of the pre-spacer, respectively, depending on which strand of the CRISPR is opened. One nucleotide of the PAM forms the last nucleotide of the repeat. This hypothetical scheme does not apply if pre-spacers are single stranded. (*c*) Repeat synthesis. After repeat unwinding, DNA polymerase synthesizes the complementary strand of the repeat, resulting in a newly incorporated spacer and a copied repeat, and (*d*) a ligase completes the adaptation stage by covalently linking the strands.

duplicated into newly synthesized repeats (36, 155), in contrast to polymorphisms in the next-to-last nucleotide of the repeat, which are duplicated (36). This provides further evidence that during spacer acquisition, only 28 bases of the repeat are copied and that 32–33bp of the protospacer plus the protospacer-adjacent nucleotide of the PAM are integrated between the copied repeats. This is likely to be of functional importance, as it ensures that new spacer sequences are integrated in the correct orientation. Indeed, space integration in *S. solfataricus*, which appears not to employ this mechanism, has been reported to sporadically and erroneously occur in inverted orientation (49a).

The studies described above show that, apart from potentially involved host factors, Cas1 and Cas2 together are sufficient and essential for integration of new spacers (176), but the molecular mechanism of action remains enigmatic. *E. coli* Cas1, like *P. aeruginosa* Cas1 (172), is a homodimeric endonuclease with in vitro metal-dependent nuclease activity on ssDNA, ssRNA, short dsDNA, and branched DNA (e.g., Holliday junctions) (10). The DNase activity of Cas1 is essential for spacer acquisition (176) and might be required for opening of the first repeat of a CRISPR locus during spacer integration or for pre-spacer processing. Cas1 was shown to be recruited to dsDNA breaks and to bind several housekeeping DNA repair proteins as well as the Cascade subunits Cas7 and Cas6e (10). Specific inhibition of Cas1-mediated Holliday junction cleavage by Cas6e suggests that Cas1 and Cascade may sometimes interact (10).

It was shown that Cas2 proteins from different organisms act as homodimeric endoribonucleases with specificity for U-rich regions (15). However, involvement of such an activity in CRISPR adaptation is not obvious. Only very recently was it shown that Cas2 proteins from *T. thermophilus* and *B. halodurans* possess pH-dependent endonucleolytic dsDNAse activity that generates dsDNA fragments of approximately 120 base pairs (118a). It is possible that this activity is important for either pre-spacer activity or CRISPR-locus sharing.

Proposed Mechanism of CRISPR Adaptation

With the recently gained insights into Type I-E CRISPR adaptation, it is tempting to speculate about a potential mechanism for primary spacer integration (**Figure 3**). We propose that the integration machinery, consisting of at least Cas1 and Cas2 and possibly other housekeeping proteins (but not Cascade or Cas3), recognizes and processes pre-spacers with a CTT PAM. The pre-spacer contains at least the protospacer-proximal nucleotide of the PAM (in most cases a C) and 32–33 bases that cover the entire protospacer. The integration machinery likely carries the pre-spacer and recognizes the leader strand. Integration of the pre-spacer takes place between the leader-proximal repeat and a newly synthesized repeat. Although the precise mechanism of repeat duplication and pre-spacer insertion remains a subject of speculation, the last nucleotide of the repeat is pre-spacer derived (155). It is possible that the mechanism of spacer integration is reminiscent of recombination by the HIV1 integrase homodimer (50), which binds and processes the ends of linear, to-be-integrated dsDNA (reviewed in 35). Next, HIV1 integrase end joins the 3′ end viral DNA strands to 5′ end host DNA, after which the gaps are filled by host DNA polymerase.

Expression-Regulation

The expression stage involves regulation of the *cas* gene and CRISPR expression, Cas complex formation, and pre-crRNA maturation (17, 160). The promoters of the two CRISPR loci in *E. coli* K12 are both located in their respective leader sequences (128, 131) and contain binding sites for regulatory proteins (131). Although the *cas3* gene is transcribed from its own constitutive promoter, transcription of the other *cas* genes initiates from a promoter (p_{cas}) upstream of *cse1*, giving rise to a polycistronic transcript encoding Cascade, Cas1, and Cas2 (**Figure 4**) (131, 168). The CRISPR/Cas system in *E. coli* K12 is tightly regulated with restricted transcription from the

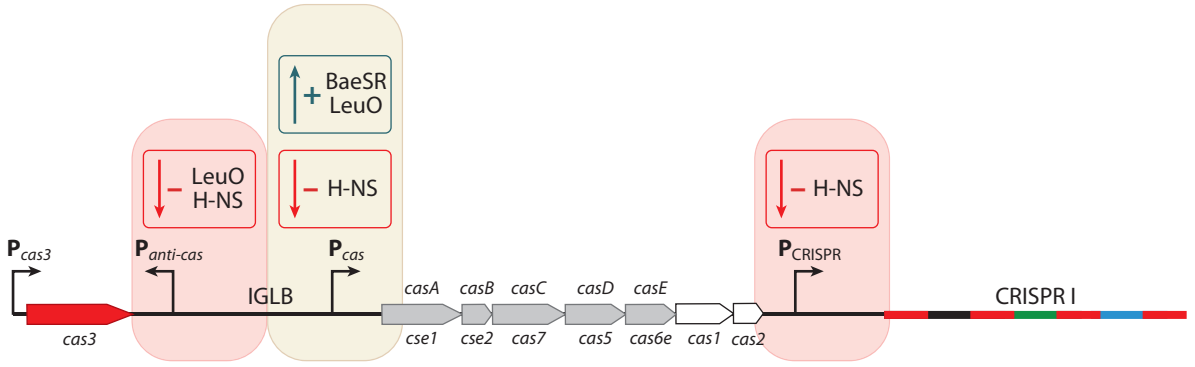

Figure 4

Cas operon organization with adjacent CRISPR (clusters of regularly interspaced short palindromic repeats) in *E. coli* K12. The Type I-E system of *E. coli* K12 comprises eight *cas* genes and an immediately adjacent CRISPR locus with Type 2 repeats. Promoters driving expression of the *cas* genes and the CRISPR are indicated by black arrows. Several regulatory proteins have been identified that directly act on the CRISPR and/or *cas* gene promoters in *E. coli* K12. Negative regulators (*depicted in red*) and positive regulators (*depicted in green*) are shown in proximity to the promoter they regulate.

p_{cas} promoter (128, 131, 168). The heat-stable nucleoid-structuring protein H-NS was identified as a key regulator of the CRISPR/Cas system of *E. coli* K12. Although H-NS binds both the p_{cas} and the CRISPR promoters, repression of p_{cas} is strongest, leading to complete blocking of transcription of the Cascade-encoding genes (131). Interestingly, phage-encoded H-NS and conjugative plasmid-encoded H-NS paralogs have been reported (44, 146), possibly counteracting induction of the CRISPR/Cas system (146).

CRISPR interference in *E. coli* has been observed when overexpressing *cas* genes and CRISPRs (25, 47), in H-NS deletion strains (128, 168, 175), and in cells overexpressing LeuO (168). LeuO is a LysR-type transcriptional activator that often serves as an H-NS antagonist (33, 37, 142, 143, 151). LeuO is a positive regulator of the *E. coli* CRISPR/Cas system that binds the *cse1* upstream region at positions flanking the H-NS binding sites, thereby blocking cooperative spreading of H-NS along the DNA (168). In *Salmonella enterica* serovar Typhi, the Type I-E CRISPR/Cas system was shown to be repressed by H-NS and LRP (leucine-responsive regulatory protein) and activated by LeuO (74, 108). Regulation of CRISPR/Cas in *E. coli* is complex because in addition to LeuO, envelope stress can also activate CRISPR defense through the two-component BaeSR pathway (13, 125). In *Thermus thermophilus*, expression of Type I-E *cas* genes was reported to be regulated by cyclic-AMP receptor protein (CRP) (144) and to be induced upon phage infection (2).

From a host-defense point of view, silencing of CRISPR/Cas seems puzzling. At least three possible explanations can be imagined. First, expressing host-defense systems is energy consuming and therefore disadvantageous when the cell is already well protected by other defense systems or when phage threats are absent. For this reason, CRISPR/Cas regulation by H-NS and LeuO could be tuned such that host defense is induced only when required. Under standard growth conditions, LeuO levels in *E. coli* are low because of H-NS-mediated repression of the *leuO* promoter (88, 98). H-NS is well known to bind AT-rich DNA (120), which may lead to H-NS titration from its own genome upon arrival of a mobile genetic element with AT-rich DNA. This releases repression of both the LeuO promoter (which is positively regulated by LeuO itself) and the p_{cas} promoter, hence instigating a CRISPR-based immune response. In fact, such an H-NS titration effect has been reported for an AT-rich plasmid in

Salmonella typhimurium (42). This regulatory mechanism may be particularly advantageous when a replicating nonlytic element, such as a high-copy plasmid, has invaded the cell. However, because preexisting Cascade surveillance complexes are absent, cells appear to be ill protected against aggressive invaders such as lytic phages. Innate defense systems could provide a first line of defense, potentially providing the cell with sufficient time to activate CRISPR/Cas while the infection is ongoing. Although speculative, it is also possible that under specific conditions (e.g., phage infection), bacterial cells communicate (e.g., through quorum sensing) to induce expression of defense systems. This is supported by the finding that spacer integration in *E. coli* took place simultaneously in many different cells in a culture (155).

A second possible reason to silence CRISPR/Cas is related to the fact that incoming DNA may be (conditionally) beneficial for the host; this is in agreement with the fact that some organisms evolved sophisticated DNA uptake systems. Although cells may be well protected against selfish DNA by constitutively expressing CRISPR/Cas and other defense systems, these systems will also block (conditionally) beneficial DNA (e.g., plasmids encoding antibiotic resistance or lysogenic phage encoding virulence factors). Elevated defense may be disadvantageous when harmless DNA enters the cell, and modest defense may be disadvantageous upon infection by a lytic phage. Presumably, fitness within a population of a bacterial species is increased by heterogeneity in expression of DNA-uptake and DNA-restriction systems. Interestingly, the aforementioned transcriptional regulators LeuO/LRP/CRP are associated to nutrient deprivation (24, 98, 147) when cells may tend to remove excess extrachromosomal genomic material. Moreover, under starvation conditions cells may depend on the uptake of foreign genetic material to acquire new (metabolic) traits. To control the resulting alien DNA accumulation, elevated expression levels of cellular defense systems may be required under these conditions.

A third reason to silence CRISPR/Cas could be to avoid autoimmunity problems. Approximately one in five CRISPR/Cas-containing organisms contains self-targeting spacers (4, 149). On the basis of a lack of conservation of these spacers across species and their co-occurrence with inactivating CRISPR or *cas* gene mutations, it could be ruled out that these spacers have regulatory functions (149). Their abundance indicates that self targeting may be a major threat imposed by CRISPR/Cas, necessitating tight regulation of this system to avoid cell death through erroneous incorporation of self DNA in CRISPR loci.

Expression-crRNA Maturation

In *E. coli* K12, five Cas proteins (Cse1, Cse2, Cas7, Cas5, and Cas6e) form a multiprotein complex called Cascade (25). The endoribonuclease Cas6e catalyzes metal-independent cleavage in the repeat sequence of the pre-crRNA, yielding mature 61-nt crRNA with an 8-nt repeat-derived 5′ handle and a 21-nt 3′ handle (25, 86). The crRNA remains Cascade bound, giving rise to a ribonucleoprotein complex with an unusual stoichiometry of $Cse1_1Cse2_2Cas7_6Cas5_1Cas6e_1crRNA_1$ and an asymmetrical seahorse-like shape (**Figure 5**) (86). Recently, an 8-Å resolution cryoelectron microscopy structure of the Cascade complex revealed the position of the subunits in the complex (170). The six Cas7 subunits form a backbone on which the other subunits assemble, with the crRNA running over the Cas7 subunits, spanning the entire complex. Although the structure of a Cas7 homolog from *S. solfataricus* was shown to contain an RNA recognition motif (RRM), it is unclear whether these motifs are functional in crRNA binding, as they lack some of the conserved sequence motifs generally involved in nucleic acid binding (95). The 5′ end of the crRNA is bound by Cse1 and/or Cas7 and/or Cas5, and the hairpin-forming 3′ end is bound by Cas6e (170). For Cas6f this binding was shown to be a very high affinity protein-RNA interaction with a Kd of 50 pM (150). The crystal

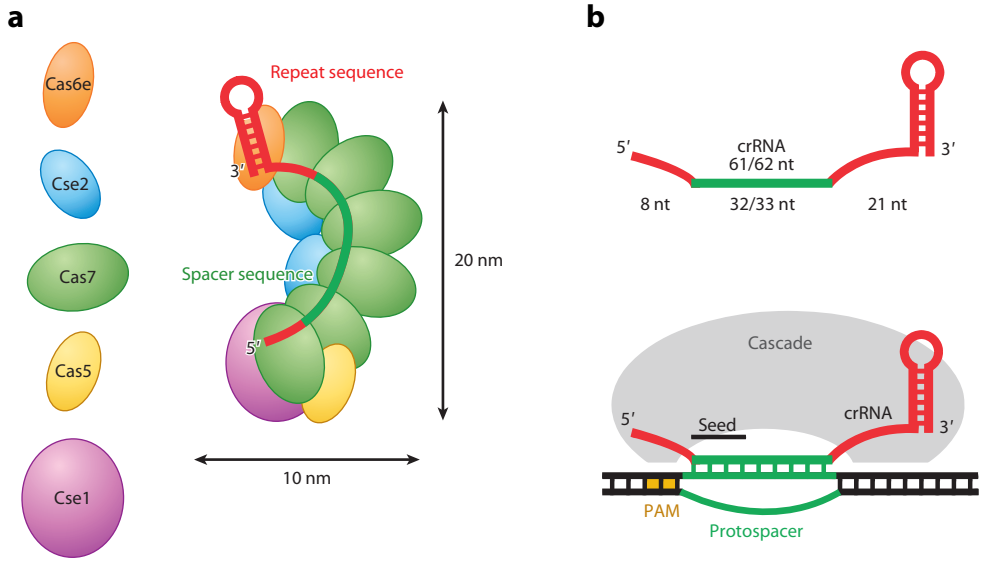

Figure 5

Schematic representation of Cascade and Cascade-mediated R-loop formation. (*a*) Cascade consists of five different Cas proteins that are present in an uneven stoichiometry. The Cas7 subunit forms the backbone of the complex. The 3′ end of the crRNA is bound by the Cas6e subunit, whereas binding of the 5′ end appears to involve a Cas7 subunit and/or Cas5 and/or Cse1. Spacer and repeat sequences of the crRNA are indicated in green and red, respectively. (*b*) Schematic of the R-loop formed by Cascade upon dsDNA binding. The protospacer-adjacent motif (PAM) sequence is indicated in yellow. The seed sequence corresponds to a noncontiguous sequence of 7 nt at the 5′ end of the spacer sequence.

structure of Cse2 from *T. thermophilus* reveals positively charged surface patches (3), which in the dimeric configuration of Cse2 in the complex might provide an interaction site for nucleic acids, such as the displaced strand after binding a target dsDNA. A high-resolution structure of Cas6e from *T. thermophilus* reveals a two-domain RAMP protein, with both domains adopting an RRM fold (a subclass of the ferredoxin fold) (46) structurally similar to *P. furiosus* Cas6 (31; reviewed in 160). More recently, high-resolution structures of Cas6e from *T. thermophilus* bound to part of the crRNA repeat reveal that the predicted crRNA hairpin structure (albeit with a stem of six instead of seven base pairs) runs over a basic groove on the Cas6e surface (58, 135). Most sequence-specific Cas6e-crRNA interactions occur between a Cas6e β-hairpin structure and the major groove of the crRNA, mainly at the 3′ end of the repeat sequence (58, 135). Although the positioning of the base of the crRNA hairpin relative to the N-terminal domain is similar in *P. aeruginosa* Cas6f (71), the overall structure of RNA-bound Cas6f is rather different from the tandem RRM-fold Cas6e-crRNA structure; Cas6f contains a single N-terminal ferredoxin fold domain that is connected to an α-helix-containing C-terminal domain through an extended linker, which is involved in sequence-specific interactions with the major groove of the crRNA repeat (71). *P. furiosus* Cas6 also has a tandem RRM-fold architecture but binds an unfolded repeat through wrapping of the crRNA around the protein, anchoring the 5′ end of the repeat in a groove between the RRM domains (165). Interestingly, despite the differences in structure and crRNA binding, pre-crRNA cleavage catalyzed by all Cas6 homologs yields mature crRNA with an 8-nt 5′ handle (25, 31, 71, 95, 103, 179).

Interference

The crRNA-loaded Cascade serves as a surveillance complex that functions in recognition of invader DNA but requires Cas3 for conferring CRISPR interference as a final step of the defense mechanism. The *E. coli* CRISPR/Cas system has been demonstrated to be functional in neutralizing distinct types of invasive DNA: phage infection (λ_{vir}, M13) (25, 139), plasmid transformation (25, 139), lysogenization, and prophage induction (λ) (47) as well as high-copy-number plasmid curing (155, 169). In addition to Cascade and Cas3, CRISPR interference requires the chaperone high-temperature protein G (HtpG) for Cas3 folding (175). The CRISPR-interference dsDNA target requires a PAM sequence, which is recognized by a loop structure of the Cse1 subunit of Cascade (136), and a fully complementary seed region (a noncontiguous 7-nt sequence at the 3′ end of the complementary DNA strand) (**Figure 5**), whereas up to five mismatches outside the seed region of the protospacer are tolerated (139). CRISPR interference is a multistep process, presumably starting with (*a*) Cascade scanning for PAM and protospacer seed regions on a dsDNA target, followed by (*b*) binding of the target sequence through base-pairing between the crRNA and the complementary DNA (initially limited to the seed, eventually resulting in R-loop formation over the entire crRNA spacer sequence) (**Figure 5**), (*c*) Cas3 recruitment by Cascade, (*d*) Cas3-mediated target DNA nicking, (*e*) reduced affinity of Cascade for the protospacer, presumably facilitating Cascade dissociation, and (*f*) full degradation of the dsDNA target by Cas3 (**Figure 6**). During the first step, the Cascade surveillance complex scans dsDNA for target sites. Although the scanning mechanism is poorly understood, PAM sites may function in this process by recruiting Cascade to potential target sites (136). Cascade may bind PAM sites and check for complementarity to the crRNA seed sequence. If the target DNA seed region is fully complementary to the crRNA seed sequence, base-pairing over the entire spacer/protospacer sequence follows (139). Upon DNA binding, both the protein complex (86, 170) and the target DNA (169) undergo conformational changes. Cascade does not require ATP for strand separation (86) but instead utilizes the energy stored in the negative supercoiled (nSC) DNA topology of the target (169). Because all circular DNA molecules in mesophiles have an nSC topology in vivo, it was initially suggested that this could be a general property of all mesophilic CRISPR/Cas systems that target dsDNA. However, recently Cas9-mediated R-loop formation was shown not to require negative supercoiling of the target DNA (84a). Cas3 recruitment takes place after Cascade has specifically bound its target (R-loop state), which is most likely triggered by the conformational changes of Cascade and/or the bending of the target DNA (169). In *E. coli* K12, the Cas3 protein consists of two domains: an N-terminal HD-nuclease domain and a C-terminal superfamily 2 DExD/H-box helicase domain (100). A number of recent structural and biochemical studies demonstrate that Cas3 has metal-dependent nuclease activities (16, 68, 116, 145). The HD domain of Cas3 from the Type I-E system of *S. thermophilus* displays magnesium-dependent endonuclease activity on ssDNA, and its helicase domain has ATP- and magnesium-dependent DNA/DNA and DNA/RNA unwinding activity in the 3′ to 5′ direction (145). In line with this, the Type I-A *Methanocaldococcus jannaschii* Cas3 HD domain has magnesium-dependent endonuclease and 3′-5′ exonuclease activity on ssDNA and ssRNA that is stimulated in the presence of ATP by the *M. jannaschii* Cas3 helicase domain (16). In accordance with these biochemical activities, it was recently demonstrated that the *E. coli* Cas3 HD-nuclease domain of a Cascade-Cas3 effector complex specifically nicks target DNA and subsequently degrades the target in the 3′ to 5′ direction through the combined action of the HD nuclease and DExD/H-box helicase domains (169). The generated cleavage products may be used for integration of new spacer sequences, which

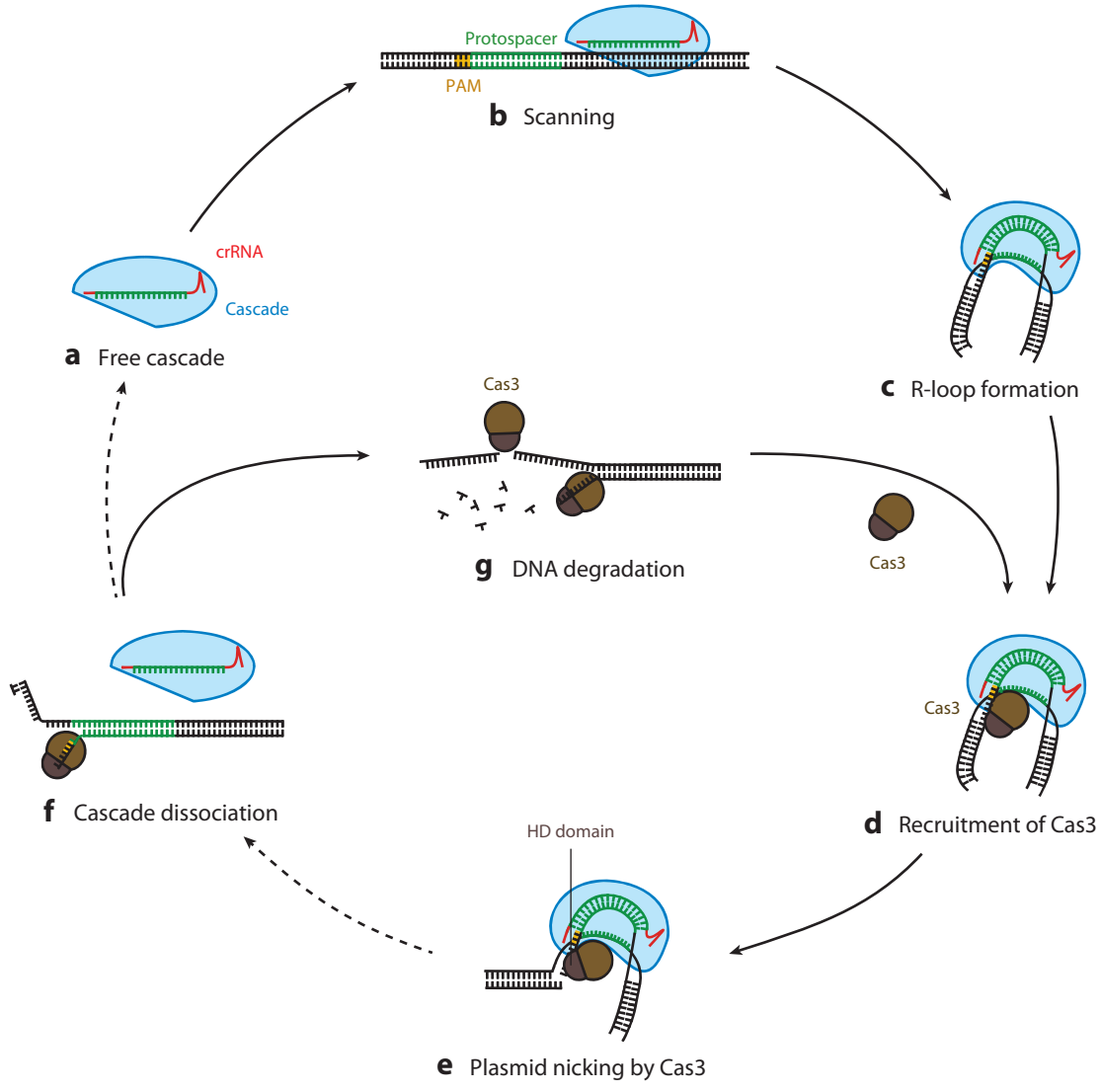

Figure 6

Proposed mechanism for CRISPR interference by Type I-E systems (from 169). Steps that are poorly understood are indicated by dotted arrows. (*a*) Cascade (*blue*) carrying a crRNA (*repeat in red*, *spacer in green*). (*b*) Cascade associates nonspecifically with the nSC (negative supercoiled) plasmid DNA and scans for a protospacer (*green*) with protospacer-adjacent motif (PAM) (*yellow*). (*c*) Sequence-specific binding to a protospacer is achieved through base-pairing between the crRNA and the complementary strand of the DNA, forming an R-loop. Upon binding, Cascade induces bending of the DNA, and Cascade itself undergoes conformational changes. (*d*) The Cse1 subunit of Cascade recruits the nuclease/helicase Cas3 (*brown*). This may be triggered by the conformational changes of Cascade and the target DNA. (*e*) The HD domain (*dark brown*) of Cas3 catalyzes Mg^{2+}-dependent nicking of the target DNA. (*f*) Plasmid nicking alters the topology of the target plasmid from nSC to relaxed open circular, causing lowered affinity of Cascade for the protospacer. (*g*) Cascade is presumably remobilized to locate new targets while Cas3 degrades the entire plasmid in an ATP-dependent manner as it progressively moves along, unwinds, and cleaves the target dsDNA. Exonucleolytic degradation takes place in the 3′-5′ direction. Reprinted from (169) with permission from Elsevier.

would explain the aforementioned priming effect during CRISPR adaptation (155).

CAS3 IN TYPE I CRISPR/Cas SYSTEMS

All Type I systems contain Cas3 but are not necessarily encoded as a single two-domain protein, as is the case in *E. coli* K12. So far, five variants of *cas3* gene arrangements have been recognized. The HD-nuclease and DExD/H helicase domains can be encoded as separate proteins as a fusion protein either with N-terminal HD and C-terminal helicase domains or in inverted orientation with N-terminal helicase and C-terminal HD-nuclease domains (8). This indicates some degree of flexibility with respect to the positioning of these domains relative to each other. Moreover, two examples of Cas3 fused to other Cas proteins have been described. In the *P. aeruginosa* I-F system, N-terminal *cas2* is fused to the *cas3* HD-nuclease domain, followed by the C-terminal helicase domain (100). In *Streptomyces griseus*, N-terminal *cas3* (with an N-terminal HD and C-terminal helicase domain) is fused to a C-terminal *cse1* gene. This Cas3-Cse1 fusion suggested that the separately encoded Cas3 and Cse1 directly interact in vivo, which was demonstrated to occur upon Cascade protospacer recognition (169). The observed truncation of HD-nuclease and helicase domains suggests that either these single-domain proteins are recruited separately to Cascade or a Cascade-like complex or they preassemble to form a Cas3 nuclease/helicase two-protein complex that functions similar to *E. coli* Cas3. The fusion of Cas2 to Cas3 provides another indication for cross talk between the interference stage and the adaptation stage because Cas2 is involved in CRISPR adaptation (176), whereas Cas3 is known to be involved in target DNA cleavage.

OTHER FUNCTIONS OF CRISPR/Cas

Many proteins have more than one function, such as methyltransferases of the type III R-M system that, besides their role in bacterial defense, have also been implicated in phase variation (reviewed in 148). Another example concerns the eukaryotic RNAi system that functions both in host defense and in gene regulation (reviewed in 32). Several studies have recently suggested that CRISPR/Cas has functions in addition to host defense. In *P. aeruginosa*, the type I-F CRISPR/Cas system has been implicated in affecting phage-mediated inhibition of swarming and biofilm formation (28, 177). It was reported that these effects are caused by a spacer with partial complementarity to the phage genome. It should be noted, however, that despite the five mismatches with the phage target, all mismatches are at positions in the spacer that, analogous to the type I-E system, would still allow for functional spacer/protospacer interaction (139, 170). Also in type I-E, a CRISPR-mediated effect based on partial complementarity between spacer and protospacer was reported (125). Type III-B systems may well have a regulatory role by cleaving complementary mRNAs (65). The possibility that CRISPR/Cas can exert gene regulatory functions is an interesting hypothesis that awaits further mechanistic analysis.

Another study has recently suggested a link between Cas1 and DNA repair (10), in line with the initial hypothesis of CRISPR/Cas being a DNA repair system (99). The authors describe interactions of Cas1 with DNA repair components and reveal defects in chromosome segregation and DNA repair in ΔCas1 strains. Further support for a role in DNA repair comes from a study on *P. furiosus*, which describes that *cas* gene expression is upregulated in response to gamma irradiation (173).

GENERAL ISSUES, CONCLUSIONS, AND OUTLOOK

Biochemical insights into the CRISPR/Cas system have increased tremendously over the past few years, but many mechanistic details remain to be determined. In particular, little is known about the I-B, I-C, and I-D systems. Another issue that has already received some

attention from a number of studies is how phages escape CRISPR-based immunity. Phages have evolved a number of strategies to escape R-M systems (see Introduction), ranging from removal of restriction sites by mutagenesis to protein-mediated inhibition of R-M enzymes. Similarly, all eukaryotic RNA plant viruses carry inhibitors of RNAi (reviewed in 163). Mechanisms to evade CRISPR interference include mutagenesis of the protospacer and/or PAM (40, 139, 140) and possibly phage-encoded H-NS (146), although the latter remains to be verified. Other strategies to overcome CRISPR/Cas can be imagined, including protein-mediated or RNA-mediated inhibition of essential CRISPR/Cas components (e.g., binding of surveillance/effector complexes or binding of crRNA). Furthermore, modification of phage DNA may potentially affect CRISPR interference through inhibition of R-loop formation or inhibition of Cas3-mediated cleavage.

It will also be highly interesting to see whether synergistic effects of multiple immune systems exist. As outlined in the introduction, bacteria can contain various defense systems, but their interactions are unknown. For example, DNA fragments generated by R-M systems could serve as spacer precursors to instigate a CRISPR-based response against an invader before it has escaped the R-M system. Also, many bacteria and in particular archaea contain multiple CRISPR/Cas systems belonging to different (sub)types. Whether cross talk between these systems occurs remains unknown. These and other open questions will undoubtedly receive more attention over the coming years.

DISCLOSURE STATEMENT

The authors are not aware of any affiliations, memberships, funding, or financial holdings that might be perceived as affecting the objectivity of this review.

ACKNOWLEDGMENTS

We would like to thank Konstantin Severinov and colleagues for sharing data prior to publication. This work was financially supported by an NWO Veni grant to S.J.J.B. (863.08.014) and an NWO-TOP grant to J.v.d.O. (854.10.003).

LITERATURE CITED

1. Abedon ST. 2011. Facilitation of CRISPR adaptation. *Bacteriophage* 1:179–81
2. Agari Y, Sakamoto K, Tamakoshi M, Oshima T, Kuramitsu S, Shinkai A. 2010. Transcription profile of *Thermus thermophilus* CRISPR systems after phage infection. *J. Mol. Biol.* 395:270–81
3. Agari Y, Yokoyama S, Kuramitsu S, Shinkai A. 2008. X-ray crystal structure of a CRISPR-associated protein, Cse2, from *Thermus thermophilus* HB8. *Proteins* 73:1063–67
4. Aklujkar M, Lovley DR. 2010. Interference with histidyl-tRNA synthetase by a CRISPR spacer sequence as a factor in the evolution of *Pelobacter carbinolicus*. *BMC Evol. Biol.* 10:230
5. Al-Attar S, Westra ER, van der Oost J, Brouns SJ. 2011. Clustered regularly interspaced short palindromic repeats (CRISPRs): the hallmark of an ingenious antiviral defense mechanism in prokaryotes. *Biol. Chem.* 392:277–89
6. Altfeld M, Fadda L, Frleta D, Bhardwaj N. 2011. DCs and NK cells: critical effectors in the immune response to HIV-1. *Nat. Rev. Immunol.* 11:176–86
7. Ando T, Xu Q, Torres M, Kusugami K, Israel DA, Blaser MJ. 2000. Restriction-modification system differences in *Helicobacter pylori* are a barrier to interstrain plasmid transfer. *Mol. Microbiol.* 37:1052–65

8. Aravind L, Koonin EV. 1998. The HD domain defines a new superfamily of metal-dependent phosphohydrolases. *Trends Biochem. Sci.* 23:469–72
9. Avrani S, Wurtzel O, Sharon I, Sorek R, Lindell D. 2011. Genomic island variability facilitates *Prochlorococcus*-virus coexistence. *Nature* 474:604–8
10. Babu M, Beloglazova N, Flick R, Graham C, Skarina T, et al. 2011. A dual function of the CRISPR-Cas system in bacterial antivirus immunity and DNA repair. *Mol. Microbiol.* 79:484–502
11. Bachi B, Reiser J, Pirrotta V. 1979. Methylation and cleavage sequences of the *Eco*P1 restriction-modification enzyme. *J. Mol. Biol.* 128:143–63
12. Bair CL, Black LW. 2007. A type IV modification dependent restriction nuclease that targets glucosylated hydroxymethyl cytosine modified DNAs. *J. Mol. Biol.* 366:768–78
13. Baranova N, Nikaido H. 2002. The baeSR two-component regulatory system activates transcription of the *yegMNOB* (*mdtABCD*) transporter gene cluster in *Escherichia coli* and increases its resistance to novobiocin and deoxycholate. *J. Bacteriol.* 184:4168–76
14. Barrangou R, Fremaux C, Deveau H, Richards M, Boyaval P, et al. 2007. CRISPR provides acquired resistance against viruses in prokaryotes. *Science* 315:1709–12
15. Beloglazova N, Brown G, Zimmerman MD, Proudfoot M, Makarova KS, et al. 2008. A novel family of sequence-specific endoribonucleases associated with the clustered regularly interspaced short palindromic repeats. *J. Biol. Chem.* 283:20361–71
16. Beloglazova N, Petit P, Flick R, Brown G, Savchenko A, Yakunin AF. 2011. Structure and activity of the Cas3 HD nuclease MJ0384, an effector enzyme of the CRISPR interference. *EMBO J.* 30:4616–27
17. Bhaya D, Davison M, Barrangou R. 2011. CRISPR-Cas systems in bacteria and archaea: versatile small RNAs for adaptive defense and regulation. *Annu. Rev. Genet.* 45:273–97
18. Bickle TA, Kruger DH. 1993. Biology of DNA restriction. *Microbiol. Rev.* 57:434–50
19. Blower TR, Pei XY, Short FL, Fineran PC, Humphreys DP, et al. 2011. A processed noncoding RNA regulates an altruistic bacterial antiviral system. *Nat. Struct. Mol. Biol.* 18:185–90
20. Bolotin A, Quinquis B, Sorokin A, Ehrlich SD. 2005. Clustered regularly interspaced short palindrome repeats (CRISPRs) have spacers of extrachromosomal origin. *Microbiology* 151:2551–61
21. Bossi L, Fuentes JA, Mora G, Figueroa-Bossi N. 2003. Prophage contribution to bacterial population dynamics. *J. Bacteriol.* 185:6467–71
22. Boyd EF, Brussow H. 2002. Common themes among bacteriophage-encoded virulence factors and diversity among the bacteriophages involved. *Trends Microbiol.* 10:521–29
23. Boyer HW. 1971. DNA restriction and modification mechanisms in bacteria. *Annu. Rev. Microbiol.* 25:153–76
24. Brinkman AB, Ettema TJ, de Vos WM, van der Oost J. 2003. The Lrp family of transcriptional regulators. *Mol. Microbiol.* 48:287–94
25. Brouns SJ, Jore MM, Lundgren M, Westra ER, Slijkhuis RJ, et al. 2008. Small CRISPR RNAs guide antiviral defense in prokaryotes. *Science* 321:960–64
26. Brown SP, Le Chat L, De Paepe M, Taddei F. 2006. Ecology of microbial invasions: amplification allows virus carriers to invade more rapidly when rare. *Curr. Biol.* 16:2048–52
27. Brussow H, Canchaya C, Hardt WD. 2004. Phages and the evolution of bacterial pathogens: from genomic rearrangements to lysogenic conversion. *Microbiol. Mol. Biol. Rev.* 68:560–602
27a. Cady KC, Bondy-Denomy J, Heussler GE, Davidson AR, O'Toole GA. 2012. The CRISPR/Cas adaptive immune system of *Pseudomonas aeruginosa* mediates resistance to naturally occurring and engineered phages. *J. Bacteriol.* doi:10.1128/JB.01184-12
28. Cady KC, O'Toole GA. 2011. Non-identity-mediated CRISPR-bacteriophage interaction mediated via the Csy and Cas3 proteins. *J. Bacteriol.* 193:3433–45
29. Carre-Mlouka A, Gaumer S, Gay P, Petitjean AM, Coulondre C, et al. 2007. Control of sigma virus multiplication by the *ref(2)P* gene of *Drosophila melanogaster*: an in vivo study of the PB1 domain of Ref(2)P. *Genetics* 176:409–19
30. Carte J, Pfister NT, Compton MM, Terns RM, Terns MP. 2010. Binding and cleavage of CRISPR RNA by Cas6. *RNA* 16:2181–88
31. Carte J, Wang R, Li H, Terns RM, Terns MP. 2008. Cas6 is an endoribonuclease that generates guide RNAs for invader defense in prokaryotes. *Genes Dev.* 22:3489–96

32. Carthew RW, Sontheimer EJ. 2009. Origins and mechanisms of miRNAs and siRNAs. *Cell* 136:642–55
33. Chen CC, Wu HY. 2005. LeuO protein delimits the transcriptionally active and repressive domains on the bacterial chromosome. *J. Biol. Chem.* 280:15111–21
34. Chopin MC, Chopin A, Bidnenko E. 2005. Phage abortive infection in lactococci: variations on a theme. *Curr. Opin. Microbiol.* 8:473–79
35. Craigie R. 2001. HIV integrase, a brief overview from chemistry to therapeutics. *J. Biol. Chem.* 276:23213–16
36. Datsenko KA, Pougach K, Tikhonov A, Wanner BL, Severinov K, Semenova E. 2012. Molecular memory of prior infections activates the CRISPR/Cas adaptive bacterial immunity system. *Nat. Commun.* 3:945
37. De la Cruz MA, Fernández-Mora M, Guadarrama C, Flores-Valdez MA, Bustamante VH, et al. 2007. LeuO antagonizes H-NS and StpA-dependent repression in *Salmonella enterica ompS1*. *Mol. Microbiol.* 66:727–43
38. Deltcheva E, Chylinski K, Sharma CM, Gonzales K, Chao Y, et al. 2011. CRISPR RNA maturation by *trans*-encoded small RNA and host factor RNase III. *Nature* 471:602–7
39. Deng L, Kenchappa CS, Peng X, She Q, Garrett RA. 2011. Modulation of CRISPR locus transcription by the repeat-binding protein Cbp1 in *Sulfolobus*. *Nucleic Acids Res.* 40:2470–80
40. Deveau H, Barrangou R, Garneau JE, Labonte J, Fremaux C, et al. 2008. Phage response to CRISPR-encoded resistance in *Streptococcus thermophilus*. *J. Bacteriol.* 190:1390–400
41. Diez-Villasenor C, Almendros C, Garcia-Martinez J, Mojica FJ. 2010. Diversity of CRISPR loci in *Escherichia coli*. *Microbiology* 156:1351–61
42. Dillon SC, Cameron AD, Hokamp K, Lucchini S, Hinton JC, Dorman CJ. 2010. Genome-wide analysis of the H-NS and Sfh regulatory networks in *Salmonella* Typhimurium identifies a plasmid-encoded transcription silencing mechanism. *Mol. Microbiol.* 76:1250–65
43. Donahue JP, Israel DA, Peek RM, Blaser MJ, Miller GG. 2000. Overcoming the restriction barrier to plasmid transformation of *Helicobacter pylori*. *Mol. Microbiol.* 37:1066–74
44. Doyle M, Fookes M, Ivens A, Mangan MW, Wain J, Dorman CJ. 2007. An H-NS-like stealth protein aids horizontal DNA transmission in bacteria. *Science* 315:251–52
45. Dulbecco R. 1952. Mutual exclusion between related phages. *J. Bacteriol.* 63:209–17
46. Ebihara A, Yao M, Masui R, Tanaka I, Yokoyama S, Kuramitsu S. 2006. Crystal structure of hypothetical protein TTHB192 from *Thermus thermophilus* HB8 reveals a new protein family with an RNA recognition motif-like domain. *Protein Sci.* 15:1494–99
47. Edgar R, Qimron U. 2010. The *Escherichia coli* CRISPR system protects from lambda lysogenization, lysogens, and prophage induction. *J. Bacteriol.* 192:6291–94
48. Edson KM, Vinson SB, Stoltz DB, Summers MD. 1981. Virus in a parasitoid wasp: suppression of the cellular immune response in the parasitoid's host. *Science* 211:582–83
49. Ellinger P, Arslan Z, Wurm R, Tschapek B, MacKenzie C, et al. 2012. The crystal structure of the CRISPR-associated protein Csn2 from *Streptococcus agalactiae*. *J. Struct. Biol.* 178:350–62
49a. Erdmann S, Garrett RA. 2012. Selective and hyperactive uptake of foreign DNA by adaptive immune systems of an archaeon via two distinct mechanisms. *Mol. Microbiol.* doi:10.1111/j.1365-2958.2012.08171.x
50. Esposito D, Craigie R. 1999. HIV integrase structure and function. *Adv. Virus Res.* 52:319–33
51. Fineran PC, Blower TR, Foulds IJ, Humphreys DP, Lilley KS, Salmond GP. 2009. The phage abortive infection system, ToxIN, functions as a protein-RNA toxin-antitoxin pair. *Proc. Natl. Acad. Sci. USA* 106:894–99
51a. Fischer S, Maier LK, Stoll B, Brendel J, Fischer E, et al. 2012. An archaeal immune system can detect multiple protospacer adjacent motifs (PAMs) to target invader DNA. *J. Biol. Chem.* doi:10.1074/jbc.M112.377002
52. Forde A, Fitzgerald GF. 1999. Bacteriophage defence systems in lactic acid bacteria. *Antonie Van Leeuwenhoek* 76:89–113
53. Tarlinton RE, Meers J, Young PR. 2006. Retroviral invasion of the koala genome. *Nature* 442:79–81
54. Fox KL, Srikhanta YN, Jennings MP. 2007. Phase variable type III restriction-modification systems of host-adapted bacterial pathogens. *Mol. Microbiol.* 65:1375–79

55. French RC, Lesley SM, Graham AF, van Rooyen CE. 1951. Studies on the relationship between virus and host cell. III. The breakdown of P32 labelled $T2r^+$ bacteriophage adsorbed to *E. coli* previously infected by other coliphages of the T group. *Can. J. Med. Sci.* 29:144–48
56. Garneau JE, Dupuis ME, Villion M, Romero DA, Barrangou R, et al. 2010. The CRISPR/Cas bacterial immune system cleaves bacteriophage and plasmid DNA. *Nature* 468:67–71
57. Garvey P, Hill C, Fitzgerald GF. 1996. The lactococcal plasmid pNP40 encodes a third bacteriophage resistance mechanism, one which affects phage DNA penetration. *Appl. Environ. Microbiol.* 62:676–79
58. Gesner EM, Schellenberg MJ, Garside EL, George MM, Macmillan AM. 2011. Recognition and maturation of effector RNAs in a CRISPR interference pathway. *Nat. Struct. Mol. Biol.* 18:688–92
59. Gomez P, Buckling A. 2011. Bacteria-phage antagonistic coevolution in soil. *Science* 332:106–9
60. Grissa I, Vergnaud G, Pourcel C. 2007. The CRISPRdb database and tools to display CRISPRs and to generate dictionaries of spacers and repeats. *BMC Bioinform.* 8:172
61. Groenen PM, Bunschoten AE, van Soolingen D, van Embden JD. 1993. Nature of DNA polymorphism in the direct repeat cluster of *Mycobacterium tuberculosis*: application for strain differentiation by a novel typing method. *Mol. Microbiol.* 10:1057–65
62. Gudbergsdottir S, Deng L, Chen Z, Jensen JV, Jensen LR, et al. 2011. Dynamic properties of the *Sulfolobus* CRISPR/Cas and CRISPR/Cmr systems when challenged with vector-borne viral and plasmid genes and protospacers. *Mol. Microbiol.* 79:35–49
63. Haft DH, Selengut J, Mongodin EF, Nelson KE. 2005. A guild of 45 CRISPR-associated (Cas) protein families and multiple CRISPR/Cas subtypes exist in prokaryotic genomes. *PLoS Comput. Biol.* 1:e60
64. Hale C, Kleppe K, Terns RM, Terns MP. 2008. Prokaryotic silencing (psi)RNAs in *Pyrococcus furiosus*. *RNA* 14:2572–79
65. Hale CR, Majumdar S, Elmore J, Pfister N, Compton M, 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
66. Hale CR, Zhao P, Olson S, Duff MO, Graveley BR, et al. 2009. RNA-guided RNA cleavage by a CRISPR RNA-Cas protein complex. *Cell* 139:945–56
67. Hall AR, Scanlan PD, Morgan AD, Buckling A. 2011. Host-parasite coevolutionary arms races give way to fluctuating selection. *Ecol. Lett.* 14:635–42
68. Han D, Krauss G. 2009. Characterization of the endonuclease SSO2001 from *Sulfolobus solfataricus* P2. *FEBS Lett.* 583:771–76
69. Harris JR. 1991. The evolution of placental mammals. *FEBS Lett.* 295:3–4
70. Hatoum-Aslan A, Maniv I, Marraffini LA. 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* 108:21218–22
71. Haurwitz RE, Jinek M, Wiedenheft B, Zhou K, Doudna JA. 2010. Sequence- and structure-specific RNA processing by a CRISPR endonuclease. *Science* 329:1355–58
72. Haurwitz RE, Sternberg SH, Doudna JA. 2012. Csy4 relies on an unusual catalytic dyad to position and cleave CRISPR RNA. *EMBO J.* 31:2824–32
73. Heller KJ. 1992. Molecular interaction between bacteriophage and the gram-negative cell envelope. *Arch. Microbiol.* 158:235–48
74. Hernandez-Lucas I, Gallego-Hernandez AL, Encarnacion S, Fernandez-Mora M, Martinez-Batallar AG, et al. 2008. The LysR-type transcriptional regulator LeuO controls expression of several genes in *Salmonella enterica* serovar Typhi. *J. Bacteriol.* 190:1658–70
75. Hoe N, Nakashima K, Grigsby D, Pan X, Dou SJ, et al. 1999. Rapid molecular genetic subtyping of serotype M1 group A *Streptococcus* strains. *Emerg. Infect. Dis.* 5:254–63
76. Horvath P, Barrangou R. 2010. CRISPR/Cas, the immune system of bacteria and archaea. *Science* 327:167–70
77. Horvath P, Romero DA, Coute-Monvoisin AC, Richards M, Deveau H, et al. 2008. Diversity, activity, and evolution of CRISPR loci in *Streptococcus thermophilus*. *J. Bacteriol.* 190:1401–12
78. Hoskisson PA, Smith MC. 2007. Hypervariation and phase variation in the bacteriophage "resistome." *Curr. Opin. Microbiol.* 10:396–400

79. Hyman P, Abedon ST. 2010. Bacteriophage host range and bacterial resistance. *Adv. Appl. Microbiol.* 70:217–48
80. Iida S, Meyer J, Bachi B, Stalhammar-Carlemalm M, Schrickel S, et al. 1983. DNA restriction-modification genes of phage P1 and plasmid p15B. Structure and in vitro transcription. *J. Mol. Biol.* 165:1–18
81. Ishiguro EE, Kay WW, Ainsworth T, Chamberlain JB, Austen RA, et al. 1981. Loss of virulence during culture of *Aeromonas salmonicida* at high temperature. *J. Bacteriol.* 148:333–40
82. Ishino Y, Shinagawa H, Makino K, Amemura M, Nakata A. 1987. Nucleotide sequence of the *iap* gene, responsible for alkaline phosphatase isozyme conversion in *Escherichia coli*, and identification of the gene product. *J. Bacteriol.* 169:5429–33
83. Jansen R, van Embden JD, Gaastra W, Schouls LM. 2002. Identification of genes that are associated with DNA repeats in prokaryotes. *Mol. Microbiol.* 43:1565–75
84. Jansen R, van Embden JD, Gaastra W, Schouls LM. 2002. Identification of a novel family of sequence repeats among prokaryotes. *OMICS J. Integr. Biol.* 6:23–33
84a. Jinek M, Chylinski K, Fonfara I, Hauer M, Doudna JA, Charpentier E. 2012. A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* 337:816–21
85. Jore MM, Brouns SJ, van der Oost J. 2012. RNA in defense: CRISPRs protect prokaryotes against mobile genetic elements. *Cold Spring Harb. Perspect. Biol.* 4:a003657
86. Jore MM, Lundgren M, van Duijn E, Bultema JB, Westra ER, et al. 2011. Structural basis for CRISPR RNA-guided DNA recognition by Cascade. *Nat. Struct. Mol. Biol.* 18:529–36
87. Karginov FV, Hannon GJ. 2010. The CRISPR system: small RNA-guided defense in bacteria and archaea. *Mol. Cell* 37:7–19
88. Klauck E, Böhringer J, Hengge-Aronis R. 1997. The LysR-like regulator LeuO in *Escherichia coli* is involved in the translational regulation of *rpoS* by affecting the expression of the small regulatory DsrA-RNA. *Mol. Microbiol.* 25:559–69
89. Kruger DH, Bickle TA. 1983. Bacteriophage survival: multiple mechanisms for avoiding the deoxyribonucleic acid restriction systems of their hosts. *Microbiol. Rev.* 47:345–60
90. Kunin V, Sorek R, Hugenholtz P. 2007. Evolutionary conservation of sequence and secondary structures in CRISPR repeats. *Genome Biol.* 8:R61
91. Labrie SJ, Samson JE, Moineau S. 2010. Bacteriophage resistance mechanisms. *Nat. Rev. Microbiol.* 8:317–27
92. Lehnherr H, Maguin E, Jafri S, Yarmolinsky MB. 1993. Plasmid addiction genes of bacteriophage P1: *doc*, which causes cell death on curing of prophage, and *phd*, which prevents host death when prophage is retained. *J. Mol. Biol.* 233:414–28
93. Lillestol RK, Redder P, Garrett RA, Brugger K. 2006. A putative viral defence mechanism in archaeal cells. *Archaea* 2:59–72
94. Lillestol RK, Shah SA, Brugger K, Redder P, Phan H, et al. 2009. CRISPR families of the crenarchaeal genus *Sulfolobus*: bidirectional transcription and dynamic properties. *Mol. Microbiol.* 72:259–72
95. Lintner NG, Kerou M, Brumfield SK, Graham S, Liu H, et al. 2011. Structural and functional characterization of an archaeal clustered regularly interspaced short palindromic repeat (CRISPR)-associated complex for antiviral defense (CASCADE). *J. Biol. Chem.* 286:21643–56
96. Llave C. 2010. Virus-derived small interfering RNAs at the core of plant-virus interactions. *Trends Plant Sci.* 15:701–7
97. Lu MJ, Henning U. 1994. Superinfection exclusion by T-even-type coliphages. *Trends Microbiol.* 2:137–39
98. Majumder A, Fang M, Tsai KJ, Ueguchi C, Mizuno T, Wu HY. 2001. LeuO expression in response to starvation for branched-chain amino acids. *J. Biol. Chem.* 276:19046–51
99. Makarova KS, Aravind L, Grishin NV, Rogozin IB, Koonin EV. 2002. A DNA repair system specific for thermophilic archaea and bacteria predicted by genomic context analysis. *Nucleic Acids Res.* 30:482–96
100. Makarova KS, Grishin NV, Shabalina SA, Wolf YI, Koonin EV. 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

101. Makarova KS, Haft DH, Barrangou R, Brouns SJ, Charpentier E, et al. 2011. Evolution and classification of the CRISPR-Cas systems. *Nat. Rev. Microbiol.* 9:467–77
102. Manica A, Zebec Z, Teichmann D, Schleper C. 2011. In vivo activity of CRISPR-mediated virus defence in a hyperthermophilic archaeon. *Mol. Microbiol.* 80:481–91
103. Marraffini LA, Sontheimer EJ. 2008. CRISPR interference limits horizontal gene transfer in staphylococci by targeting DNA. *Science* 322:1843–45
104. Marraffini LA, Sontheimer EJ. 2010. CRISPR interference: RNA-directed adaptive immunity in bacteria and archaea. *Nat. Rev. Genet.* 11:181–90
105. Marraffini LA, Sontheimer EJ. 2010. Self versus non-self discrimination during CRISPR RNA-directed immunity. *Nature* 463:568–71
106. Masepohl B, Gorlitz K, Bohme H. 1996. Long tandemly repeated repetitive (LTRR) sequences in the filamentous cyanobacterium *Anabaena* sp. PCC 7120. *Biochim. Biophys. Acta* 1307:26–30
107. McGrath S, Fitzgerald GF, van Sinderen D. 2002. Identification and characterization of phage-resistance genes in temperate lactococcal bacteriophages. *Mol. Microbiol.* 43:509–20
108. Medina-Aparicio L, Rebollar-Flores JE, Gallego-Hernandez AL, Vazquez A, Olvera L, et al. 2011. The CRISPR/Cas immune system is an operon regulated by LeuO, H-NS, and leucine-responsive regulatory protein in *Salmonella enterica* serovar Typhi. *J. Bacteriol.* 193:2396–407
109. Meisel A, Bickle TA, Kruger DH, Schroeder C. 1992. Type III restriction enzymes need two inversely oriented recognition sites for DNA cleavage. *Nature* 355:467–69
110. Meyer JR, Dobias DT, Weitz JS, Barrick JE, Quick RT, Lenski RE. 2012. Repeatability and contingency in the evolution of a key innovation in phage lambda. *Science* 335:428–32
111. Mojica FJ, Diez-Villasenor C, Garcia-Martinez J, Almendros C. 2009. Short motif sequences determine the targets of the prokaryotic CRISPR defence system. *Microbiology* 155:733–40
112. Mojica FJ, Diez-Villasenor C, Garcia-Martinez J, Soria E. 2005. Intervening sequences of regularly spaced prokaryotic repeats derive from foreign genetic elements. *J. Mol. Evol.* 60:174–82
113. Mojica FJ, Diez-Villasenor C, Soria E, Juez G. 2000. Biological significance of a family of regularly spaced repeats in the genomes of archaea, bacteria and mitochondria. *Mol. Microbiol.* 36:244–46
114. Mojica FJ, Ferrer C, Juez G, Rodriguez-Valera F. 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
115. Molineux IJ. 1991. Host-parasite interactions: recent developments in the genetics of abortive phage infections. *New Biol.* 3:230–36
116. Mulepati S, 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.* 286:31896–903
117. Murray NE. 2000. Type I restriction systems: sophisticated molecular machines (a legacy of Bertani and Weigle). *Microbiol. Mol. Biol. Rev.* 64:412–34
118. Naito T, Kusano K, Kobayashi I. 1995. Selfish behavior of restriction-modification systems. *Science* 267:897–99
118a. Nam KH, Ding F, Haitjema C, Huang Q, Delisa MP, Ke A. 2012. Double-stranded endonuclease activity in *B. halodurans* clustered regularly interspaced short palindromic repeats (CRISPR)-associated Cas2 protein. *J. Biol. Chem.* doi:10.1074/jbc.M112.382598
118b. Nam KH, Haitjema C, Liu X, Ding F, Wang H, et al. 2012. Cas5d protein processes pre-crRNA and assembles Cascade-like interference complex in subtype I-C/Dvulg CRISPR-Cas system. *Structure* 20:1574–84
119. Nam KH, Kurinov I, Ke A. 2011. Crystal structure of clustered regularly interspaced short palindromic repeats (CRISPR)-associated Csn2 protein revealed Ca^{2+}-dependent double-stranded DNA binding activity. *J. Biol. Chem.* 286:30759–68
120. Navarre WW, Porwollik S, Wang Y, McClelland M, Rosen H, et al. 2006. Selective silencing of foreign DNA with low GC content by the H-NS protein in *Salmonella*. *Science* 313:236–38
121. Nordström K, Forsgren A. 1974. Effect of protein A on adsorption of bacteriophages to *Staphylococcus aureus*. *J. Virol.* 14:198–202

122. Nordström K, Forsgren A, Cox P. 1974. Prevention of bacteriophage adsorption to *Staphylococcus aureus* by immunoglobulin G. *J. Virol.* 14:203–6
123. Oliver KM, Degnan PH, Hunter MS, Moran NA. 2009. Bacteriophages encode factors required for protection in a symbiotic mutualism. *Science* 325:992–94
124. Peng X, Brugger K, Shen B, Chen L, She Q, Garrett RA. 2003. Genus-specific protein binding to the large clusters of DNA repeats (short regularly spaced repeats) present in *Sulfolobus* genomes. *J. Bacteriol.* 185:2410–17
125. Perez-Rodriguez R, Haitjema C, Huang Q, Nam KH, Bernardis S, et al. 2011. Envelope stress is a trigger of CRISPR RNA-mediated DNA silencing in *Escherichia coli*. *Mol. Microbiol.* 79:584–99
126. Pingoud A, Fuxreiter M, Pingoud V, Wende W. 2005. Type II restriction endonucleases: structure and mechanism. *Cell. Mol. Life Sci.* 62:685–707
127. Plagens A, Tjaden B, Hagemann A, Randau L, Hensel R. 2012. Characterization of the CRISPR/Cas subtype I-A system of the hyperthermophilic crenarchaeon *Thermoproteus tenax*. *J. Bacteriol.* 194:2491–500
128. Pougach K, Semenova E, Bogdanova E, Datsenko KA, Djordjevic M, et al. 2010. Transcription, processing and function of CRISPR cassettes in *Escherichia coli*. *Mol. Microbiol.* 77:1367–79
129. Pourcel C, Salvignol G, Vergnaud G. 2005. CRISPR elements in *Yersinia pestis* acquire new repeats by preferential uptake of bacteriophage DNA, and provide additional tools for evolutionary studies. *Microbiology* 151:653–63
130. Przybilski R, Richter C, Gristwood T, Clulow JS, Vercoe RB, Fineran PC. 2011. Csy4 is responsible for CRISPR RNA processing in *Pectobacterium atrosepticum*. *RNA Biol.* 8:517–28
131. Pul U, Wurm R, Arslan Z, Geissen R, Hofmann N, Wagner R. 2010. Identification and characterization of *E. coli* CRISPR-Cas promoters and their silencing by H-NS. *Mol. Microbiol.* 75:1495–512
132. Riede I, Eschbach ML. 1986. Evidence that TraT interacts with OmpA of *Escherichia coli*. *FEBS Lett.* 205:241–45
133. Roberts RJ, Belfort M, Bestor T, Bhagwat AS, Bickle TA, et al. 2003. A nomenclature for restriction enzymes, DNA methyltransferases, homing endonucleases and their genes. *Nucleic Acids Res.* 31:1805–12
134. Roossinck MJ. 2011. The good viruses: viral mutualistic symbioses. *Nat. Rev. Microbiol.* 9:99–108
135. Sashital DG, Jinek M, Doudna JA. 2011. An RNA-induced conformational change required for CRISPR RNA cleavage by the endoribonuclease Cse3. *Nat. Struct. Mol. Biol.* 18:680–87
136. Sashital DG, Wiedenheft B, Doudna JA. 2012. Mechanism of foreign DNA selection in a bacterial adaptive immune system. *Mol. Cell* 46:606–15
137. Schnettler E, de Vries W, Hemmes H, Haasnoot J, Kormelink R, et al. 2009. The NS3 protein of rice hoja blanca virus complements the RNAi suppressor function of HIV-1 Tat. *EMBO Rep.* 10:258–63
138. Scholl D, Merril C. 2005. The genome of bacteriophage K1F, a T7-like phage that has acquired the ability to replicate on K1 strains of *Escherichia coli*. *J. Bacteriol.* 187:8499–503
139. Semenova E, Jore MM, Datsenko KA, Semenova A, Westra ER, et al. 2011. Interference by clustered regularly interspaced short palindromic repeat (CRISPR) RNA is governed by a seed sequence. *Proc. Natl. Acad. Sci. USA* 108:10098–103
140. Semenova E, Nagornykh M, Pyatnitskiy M, Artamonova II, Severinov K. 2009. Analysis of CRISPR system function in plant pathogen *Xanthomonas oryzae*. *FEMS Microbiol. Lett.* 296:110–16
141. Shah SA, Hansen NR, Garrett RA. 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
142. Shimada T, Bridier A, Briandet R, Ishihama A. 2011. Novel roles of LeuO in transcription regulation of *E. coli* genome: antagonistic interplay with the universal silencer H-NS. *Mol. Microbiol.* 82:378–97
143. Shimada T, Yamamoto K, Ishihama A. 2009. Involvement of the leucine response transcription factor LeuO in regulation of the genes for sulfa drug efflux. *J. Bacteriol.* 191:4562–71
144. Shinkai A, Kira S, Nakagawa N, Kashihara A, Kuramitsu S, Yokoyama S. 2007. Transcription activation mediated by a cyclic AMP receptor protein from *Thermus thermophilus* HB8. *J. Bacteriol.* 189:3891–901
145. Sinkunas T, Gasiunas G, Fremaux C, Barrangou R, Horvath P, Siksnys V. 2011. Cas3 is a single-stranded DNA nuclease and ATP-dependent helicase in the CRISPR/Cas immune system. *EMBO J.* 30:1335–42

146. Skennerton CT, Angly FE, Breitbart M, Bragg L, He S, et al. 2011. Phage encoded H-NS: a potential achilles heel in the bacterial defence system. *PLoS ONE* 6:e20095
147. Spector MP. 1998. The starvation-stress response (SSR) of *Salmonella*. *Adv. Microb. Physiol.* 40:233–79
148. Srikhanta YN, Fox KL, Jennings MP. 2010. The phasevarion: phase variation of type III DNA methyltransferases controls coordinated switching in multiple genes. *Nat. Rev. Microbiol.* 8:196–206
149. Stern A, Keren L, Wurtzel O, Amitai G, Sorek R. 2010. Self-targeting by CRISPR: gene regulation or autoimmunity? *Trends Genet.* 26:335–40
150. Sternberg SH, Haurwitz RE, Doudna JA. 2012. Mechanism of substrate selection by a highly specific CRISPR endoribonuclease. *RNA* 18:661–72
151. Stoebel DM, Free A, Dorman CJ. 2008. Anti-silencing: overcoming H-NS-mediated repression of transcription in gram-negative enteric bacteria. *Microbiology* 154:2533–45
152. Studier FW, Bandyopadhyay PK. 1988. Model for how type I restriction enzymes select cleavage sites in DNA. *Proc. Natl. Acad. Sci. USA* 85:4677–81
153. Sumby P, Smith MCM. 2002. Genetics of the phage growth limitation (Pgl) system of *Streptomyces coelicolor* A3(2). *Mol. Microbiol.* 44:489–500
154. Sutherland IW, Hughes KA, Skillman LC, Tait K. 2004. The interaction of phage and biofilms. *FEMS Microbiol. Lett.* 232:1–6
155. Swarts DC, Mosterd C, van Passel MW, Brouns SJ. 2012. CRISPR interference directs strand specific spacer acquisition. *PLoS ONE* 7:e35888
156. Tang TH, Bachellerie JP, Rozhdestvensky T, Bortolin ML, Huber H, et al. 2002. Identification of 86 candidates for small non-messenger RNAs from the archaeon *Archaeoglobus fulgidus*. *Proc. Natl. Acad. Sci. USA* 99:7536–41
157. Tang TH, Polacek N, Zywicki M, Huber H, Brugger K, et al. 2005. Identification of novel non-coding RNAs as potential antisense regulators in the archaeon *Sulfolobus solfataricus*. *Mol. Microbiol.* 55:469–81
158. Terns MP, Terns RM. 2011. CRISPR-based adaptive immune systems. *Curr. Opin. Microbiol.* 14:321–27
159. Touchon M, Rocha EP. 2010. The small, slow and specialized CRISPR and anti-CRISPR of *Escherichia* and *Salmonella*. *PLoS ONE* 5:e11126
160. van der Oost J, Jore MM, Westra ER, Lundgren M, Brouns SJ. 2009. CRISPR-based adaptive and heritable immunity in prokaryotes. *Trends Biochem. Sci.* 34:401–7
161. van Embden JD, van Gorkom T, Kremer K, Jansen R, van Der Zeijst BA, Schouls LM. 2000. Genetic variation and evolutionary origin of the direct repeat locus of *Mycobacterium tuberculosis* complex bacteria. *J. Bacteriol.* 182:2393–401
162. van Rij RP, Berezikov E. 2009. Small RNAs and the control of transposons and viruses in *Drosophila*. *Trends Microbiol.* 17:163–71
163. Voinnet O. 2005. Induction and suppression of RNA silencing: insights from viral infections. *Nat. Rev. Genet.* 6:206–20
164. Vovis GF, Horiuchi K, Zinder ND. 1974. Kinetics of methylation of DNA by a restriction endonuclease from *Escherichia coli* B. *Proc. Natl. Acad. Sci. USA* 71:3810–13
165. Wang R, Preamplume G, Terns MP, Terns RM, Li H. 2011. Interaction of the Cas6 riboendonuclease with CRISPR RNAs: recognition and cleavage. *Structure* 19:257–64
166. Warren RA. 1980. Modified bases in bacteriophage DNAs. *Annu. Rev. Microbiol.* 34:137–58
167. Watanabe K, Ishibashi K, Nakashima Y, Sakurai T. 1984. A phage-resistant mutant of *Lactobacillus casei* which permits phage adsorption but not genome injection. *J. Gen. Virol.* 65:981–86
168. Westra ER, Pul U, Heidrich N, Jore MM, Lundgren M, et al. 2010. H-NS-mediated repression of CRISPR-based immunity in *Escherichia coli* K12 can be relieved by the transcription activator LeuO. *Mol. Microbiol.* 77:1380–93
169. Westra ER, van Erp PBG, Künne T, Wong SP, Staals RHJ, 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
170. Wiedenheft B, Lander GC, Zhou K, Jore MM, Brouns SJ, et al. 2011. Structures of the RNA-guided surveillance complex from a bacterial immune system. *Nature* 477:486–89

171. Wiedenheft B, van Duijn E, Bultema JB, Waghmare SP, Zhou K, et al. 2011. RNA-guided complex from a bacterial immune system enhances target recognition through seed sequence interactions. *Proc. Natl. Acad. Sci. USA* 108:10092–97
172. Wiedenheft B, Zhou K, Jinek M, Coyle SM, Ma W, Doudna JA. 2009. Structural basis for DNase activity of a conserved protein implicated in CRISPR-mediated genome defense. *Structure* 17:904–12
173. Williams E, Lowe TM, Savas J, DiRuggiero J. 2007. Microarray analysis of the hyperthermophilic archaeon *Pyrococcus furiosus* exposed to gamma irradiation. *Extremophiles* 11:19–29
174. Wyers F, Petitjean AM, Dru P, Gay P, Contamine D. 1995. Localization of domains within the *Drosophila* Ref(2)P protein involved in the intracellular control of sigma rhabdovirus multiplication. *J. Virol.* 69:4463–70
175. Yosef I, Goren MG, Kiro R, Edgar R, Qimron U. 2011. High-temperature protein G is essential for activity of the *Escherichia coli* clustered regularly interspaced short palindromic repeats (CRISPR)/Cas system. *Proc. Natl. Acad. Sci. USA* 108:20136–41
176. Yosef I, Goren MG, Qimron U. 2012. Proteins and DNA elements essential for the CRISPR adaptation process in *Escherichia coli*. *Nucleic Acids Res.* 40:5569–76
177. Zegans ME, Wagner JC, Cady KC, Murphy DM, Hammond JH, O'Toole GA. 2009. Interaction between bacteriophage DMS3 and host CRISPR region inhibits group behaviors of *Pseudomonas aeruginosa*. *J. Bacteriol.* 191:210–19
178. Zhang H, Zheng X, Zhang Z. 2010. The role of vacuolar processing enzymes in plant immunity. *Plant Signal. Behav.* 5:1565–67
179. Zhang J, Rouillon C, Kerou M, Reeks J, Brugger K, et al. 2012. Structure and mechanism of the CMR complex for CRISPR-mediated antiviral immunity. *Mol. Cell* 3:303–13

Annual Review of Genetics

Volume 46, 2012

Contents

Errata

An online log of corrections to *Annual Review of Genetics* articles may be found at http://genet.annualreviews.org/errata.shtml