Cas9–crRNA ribonucleoprotein complex mediates specific DNA cleavage for adaptive immunity in bacteria

Giedrius Gasiunas[a], Rodolphe Barrangou[b], Philippe Horvath[c], and Virginijus Siksnys[a,1]

[a]Institute of Biotechnology, Vilnius University, LT-02241 Vilnius, Lithuania; [b]DuPont Nutrition and Health, Madison, WI 53716; and [c]DuPont Nutrition and Health, F-86220 Dangé-Saint-Romain, France

Edited by Arthur Landy, Brown University, Providence, RI, and approved August 1, 2012 (received for review May 21, 2012)

Clustered, regularly interspaced, short palindromic repeats (CRISPR)/CRISPR-associated (Cas) systems provide adaptive immunity against viruses and plasmids in bacteria and archaea. The silencing of invading nucleic acids is executed by ribonucleoprotein complexes preloaded with small, interfering CRISPR RNAs (crRNAs) that act as guides for targeting and degradation of foreign nucleic acid. Here, we demonstrate that the Cas9–crRNA complex of the *Streptococcus thermophilus* CRISPR3/Cas system introduces in vitro a double-strand break at a specific site in DNA containing a sequence complementary to crRNA. DNA cleavage is executed by Cas9, which uses two distinct active sites, RuvC and HNH, to generate site-specific nicks on opposite DNA strands. Results demonstrate that the Cas9–crRNA complex functions as an RNA-guided endonuclease with RNA-directed target sequence recognition and protein-mediated DNA cleavage. These findings pave the way for engineering of universal programmable RNA-guided DNA endonucleases.

nuclease | site-directed mutagenesis | RNA interference | DNA interference

Clustered, regularly interspaced, short palindromic repeats (CRISPR), together with CRISPR-associated (*cas*) genes constitute an adaptive microbial immune system which provides acquired resistance against viruses and plasmids via uptake of short fragments of invasive DNA (spacers) (1). *cas* genes typically are present in the vicinity of CRISPR arrays (2, 3). CRISPR repeat-spacer arrays are transcribed into long primary transcripts that are processed further into a set of short CRISPR RNAs (crRNAs) containing a conserved repeat fragment and a variable spacer sequence (guide) complementary to the invading nucleic acid (4–6). crRNAs combine with Cas proteins to form an effector complex which recognizes the target sequence in the invasive nucleic acid by base pairing to the complementary strand (7) and induces sequence-specific cleavage, thereby preventing proliferation and propagation of foreign genetic elements. CRISPR/Cas systems have been categorized into three main types, based on core elements content and sequences (8).

The structural organization and function of nucleoprotein complexes involved in crRNA-mediated silencing of foreign nucleic acids differ among distinct CRISPR/Cas types (9). In the Type I-E system, as exemplified by *Escherichia coli,* crRNAs are incorporated into a multisubunit effector complex called "Cascade" (from "CRISPR-associated complex for antiviral defense") (5), which binds to the target DNA and triggers degradation by the signature Cas3 protein (10, 11). In type III CRISPR/Cas systems of *Sulfolobus solfataricus* and *Pyrococcus furiosus,* the Cas RAMP module (Cmr) and the crRNA complex recognize and cleave synthetic RNA in vitro (12, 13), whereas the CRISPR/Cas system of *Staphylococcus epidermidis* targets DNA in vivo (14). The structure and composition of ribonucleoprotein complexes involved in DNA silencing by type II CRISPR/Cas systems remain to be established.

The CRISPR3/Cas system of *Streptococcus thermophilus* DGCC7710 (15) is a type II system that consists of four *cas* genes, *cas9, cas1, cas2,* and *csn2,* which are located upstream of 12 repeat-spacer units (Fig. 1*A*). *cas9* (formerly named "*cas5*" or "*csn1*") is the signature gene for type II systems (8). In the closely related *S. thermophilus* CRISPR1/Cas system, disruption of *cas9* abolishes crRNA-mediated DNA interference (DNAi) (1). We have shown recently that the *S. thermophilus* CRISPR3/Cas system can be transferred into *E. coli* and that this heterologous system provides protection against plasmid transformation and de novo phage infection (16). The interference against phage and plasmid DNA provided by *S. thermophilus* CRISPR3 requires the presence, within the target DNA, of a protospacer sequence complementary to the spacer-derived crRNA and a conserved protospacer-adjacent motif (PAM) sequence, NGGNG, located immediately downstream of the protospacer (17–19). Single point mutations in the PAM or defined protospacer positions allow the phages or plasmids to circumvent CRISPR-mediated immunity (16, 17, 20). We have established that in the heterologous system, *cas9* is the sole *cas* gene necessary for CRISPR-encoded interference (16), suggesting that this protein is involved in crRNA processing and/or crRNA-mediated silencing of invasive DNA.

Cas9 of *S. thermophilus* CRISPR3/Cas system is a large, multidomain protein comprising 1,409 amino acid residues (16). It contains two nuclease domains, a RuvC-like nuclease domain near the amino terminus and a HNH-like nuclease domain in the middle of the protein. Mutational analysis has established that interference provided in vivo by Cas9 requires both the RuvC and HNH motifs (16). Both active-site motifs presumably play an important role in the interference step; however, whether both these motifs contribute directly to the target DNA cleavage and/or crRNA maturation/processing remains to be established.

In this work we have isolated the Cas9–crRNA complex of the *S. thermophilus* CRISPR3/Cas system and demonstrate that it cleaves both synthetic oligodeoxynucleotide and plasmid DNA bearing a nucleotide sequence complementary to the crRNA, in a PAM-dependent manner. Furthermore, we provide experimental evidence that the PAM is recognized in the context of dsDNA and is critical for in vitro DNA binding and cleavage. Finally, we show that the Cas9 RuvC and HNH active sites are responsible for the cleavage of opposite DNA strands. Taken together, our data demonstrate that the Cas9–crRNA complex functions as an RNA-guided endonuclease that uses RNA for target site recognition and Cas9 for DNA cleavage.

Author contributions: G.G. and V.S. designed research; G.G. performed research; P.H. contributed new reagents/analytic tools; G.G., R.B., P.H., and V.S. analyzed data; and G.G. and V.S. wrote the paper.

Conflict of interest statement: R.B. and P.H. are employees of DuPont Nutrition & Health and G.G., R.B., P.H., and V.S. are inventors on patent applications related to CRISPR.

This article is a PNAS Direct Submission.

Freely available online through the PNAS open access option.

[1]To whom correspondence should be addressed. E-mail: siksnys@ibt.lt.

See Author Summary on page 15539 (volume 109, number 39).

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1208507109/-/DCSupplemental.

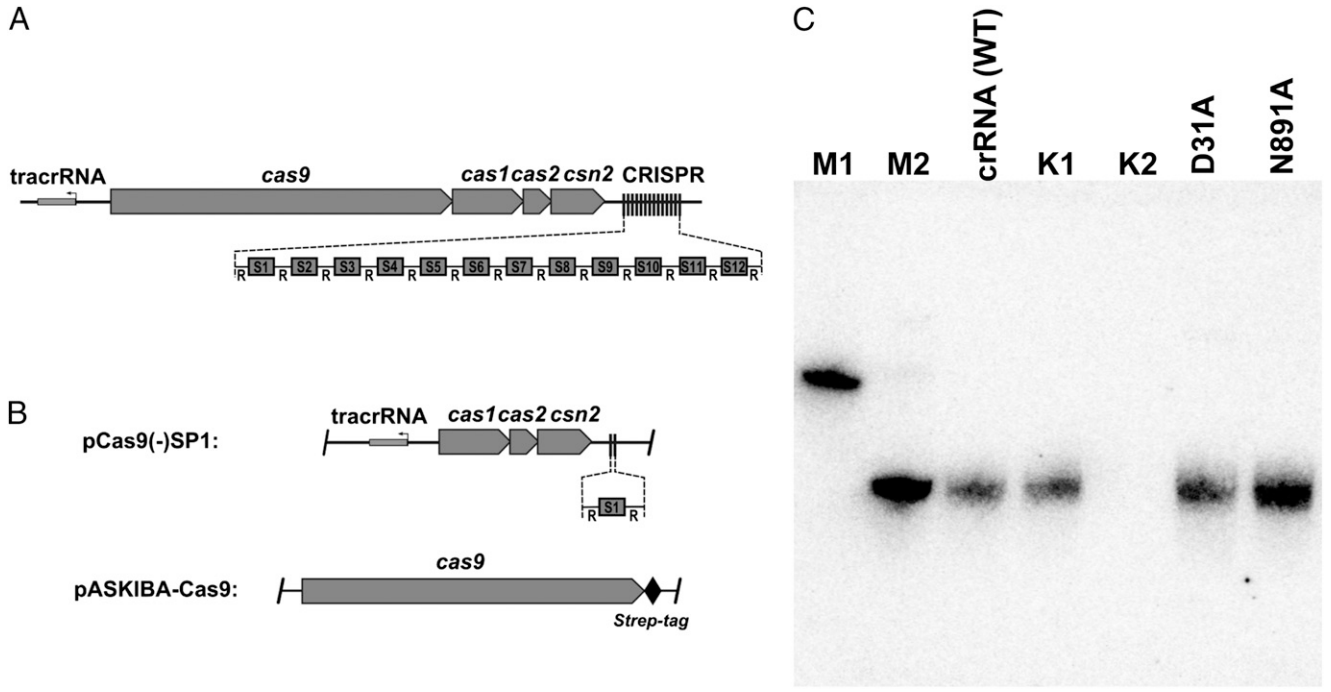

Fig. 1. The Cas9 protein copurifies with crRNA. (*A*) Schematic representation of the CRISPR3/Cas system of *S. thermophilus* DGCC7710. Four *cas* genes (*cas9, cas1, cas2, csn2*) are located upstream of the CRISPR repeat-spacer array, consisting of 13 repeat (R) sequences and 12 unique spacers (S1–S12). The tracrRNA, required for crRNA maturation in type II CRISPR/Cas systems (21), is located upstream of the *cas9* gene and is encoded on the opposite DNA strand (shown by an arrow) with respect to the other elements of this system. (*B*) Schematic representation of heterologous loci in two plasmids used for the coexpression of the Cas9–crRNA complex. *E. coli* RR1 contained pCas9(–)SP1 (encoding Cas1, Cas2, Csn2, SP1, and tracrRNA) and pASKIBA-Cas9 (encoding the Strep-tagged version of Cas9) plasmids. (*C*) Northern blot analysis of Cas9–crRNA complexes using anti-crDNA oligonucleotide as a probe. M1, 84-nt oligodeoxynucleotide corresponding to the spacer1-repeat unit; M2, 42-nt synthetic oligoribonucleotide corresponding to the predicted *S. thermophilus* CRISPR3 crRNA (Fig. S2); crRNA (WT), crRNA isolated from the WT Cas9–crRNA complex; K1, crRNA (WT) treated with DNase I for 15 min; K2, crRNA (WT) treated with RNase I for 15 min; D31A, crRNA purified from the Cas9 D31A mutant complex; N891A, crRNA purified from the Cas9 N891A mutant complex.

Results

Isolation and Analysis of the Cas9–crRNA Complex. The CRISPR3 system of *S. thermophilus* DGCC7710 comprises four *cas* genes that are located upstream of 12 repeat-spacer units (Fig. 1*A*). To isolate the Cas9–crRNA complex from the heterologous *E. coli* host, we first cloned the *cas9* gene into the pASK-IBA3 vector to generate a construct encoding a single Cas9 protein with a C-terminal Strep(II) tag (Fig. 1*B*). To obtain a homogeneous Cas9–crRNA complex, we next engineered a pCas9(−)SP1 plasmid with a single spacer (spacer1) inserted between two CRISPR repeats (Fig. 1*B*). A plasmid interference assay confirmed that the minimized CRISPR array carrying only spacer1 provides interference against plasmid pSP1 transformation in *E. coli* similar to that of the CRISPR3/Cas system carrying a complete 12-spacer array (Fig. S1*B*). To achieve simultaneous transcription of all target genes, we obtained *cas9* gene expression in two steps. First, we induced Cas9 expression in a small volume of *E. coli* culture and after 4 h transferred an aliquot of preinduced culture into a larger volume of fresh LB medium already containing the inductor and incubated the culture for 17 h. The Cas9–crRNA complex was isolated using a Strep-Tactin Sepharose column, and crRNA bound to Cas9 protein was analyzed.

The CRISPR3/Cas system of *S. thermophilus* belongs to type IIA (formerly "Nmeni") (8). In the homologous type IIA CRISPR/Cas system of *Streptococcus pyogenes,* transencoded small RNA (tracrRNA) and bacterial RNaseIII are involved in the generation of the 42-nt crRNAs that carry the 22-nt 3′ handle comprising the repeat sequence and a 20-nt spacer fragment (21). crRNAs of similar length also are generated in the *S. thermophilus* LMD-9 CRISPR3/Cas system (21), which is nearly identical to the CRISPR3/Cas system of DGCC7710 (Fig. S2 *A* and *B*). We assumed that crRNAs present in the Cas9–crRNA complex isolated from the heterologous *E. coli* strain might have the same length (Fig. S2*C*). Therefore, to probe nucleic acids extracted from the Strep-Tactin–purified Cas9 complex, we used a 42-nt anti-crRNA ssDNA oligonucleotide comprising 22 nt corresponding to the 3′ end of the repeat sequence and 20 nt at the 5′ end of the SP1 fragment. Nucleic acids present in the Cas9–crRNA complex hybridized with anti-crRNA oligonucleotide and were sensitive to RNase but not to DNase treatment (Fig. 1*C*). The size of extracted RNAs was identical to the 42-nt synthetic oligoribonucleotide corresponding to the putative crRNA of the CRISPR3 system of *S. thermophilus* DGCC7710 (Fig. 2*A* and Fig. S2*C*). Taken together, these data confirm that Cas9 Strep-tagged protein copurifies with the 42-nt crRNAs that are derived from the CRISPR3 array.

Cas9–crRNA Complex Cleaves dsDNA Within the Protospacer. To test the in vitro activity of the purified Cas9–crRNA complex, we first used the SP1 oligoduplex (Table S1) containing a protospacer sequence (protospacer1) identical to spacer1 in the CRISPR3 array, the PAM sequence 5′-TGGTG-3′ (conserved nucleotides are underlined) downstream of the protospacer, and the 10-nt flanking sequences present in pSP1 plasmid (Fig. 2*A*) (16). The oligoduplex strand complementary to crRNA is named the "(+)strand," and the opposite duplex strand is called the "(−) strand." To monitor the cleavage reaction, either the (+)strand or the (–)strand of the SP1 oligoduplex was P^{33}-labeled at the 5′ terminus. Analysis of the reaction products by PAGE (Fig. 2*B*) revealed that the Cas9–crRNA complex cleaves both strands of the SP1 oligoduplex at a fixed position. Mapping of the cleavage position using synthetic oligonucleotides as size markers revealed that cleavage occurs within the protospacer, 3 nt away

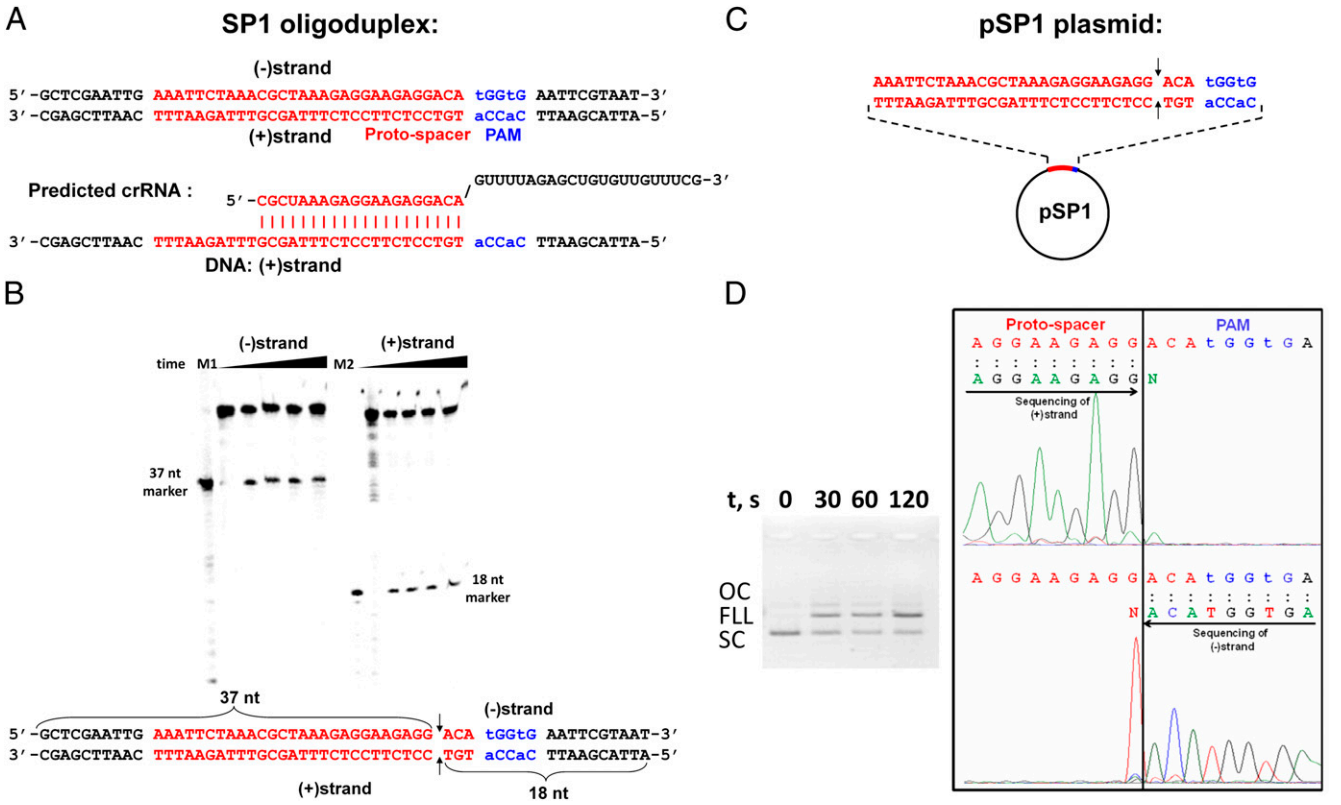

Fig. 2. The Cas9–crRNA complex cleaves dsDNA within the protospacer. (*A*) Oligoduplex substrate used in the cleavage assay. The 55-nt oligoduplex SP1 contains the protospacer1 (red letters), PAM (blue letters), and 10-nt flanking sequences on both sides identical to those in pSP1 plasmid. In the SP1 oligoduplex, the DNA strand complementary to the 5′-terminal fragment of crRNA (red letters) is termed the "(+)strand," and the opposite DNA strand is termed the "(–)strand." (*B*) Oligoduplex SP1 cleavage. The Cas9–crRNA complex (2.5 M) and 1 nM SP1 oligoduplex labeled with ^{33}P at the 5′ end of either the (+) or (–)strand were incubated in the reaction buffer [10 mM Tris·HCl (pH 7.5), 10 mM NaCl, 10 mM $MgCl_2$, 0.1 mg/mL BSA] at 37 °C for various time intervals (30 s to 10 min), and reaction products were analyzed by 20% PAGE. Lanes M1 and M2 contain chemically synthesized 5′-end, ^{33}P-labeled, 37-nt and 18-nt oligodeoxynucleotides corresponding to the cleavage products of (–) and (+) DNA strands, respectively. Cleavage positions are indicated by arrows. (*C*) Schematic representation of pSP1 plasmid (16) used in the plasmid cleavage assay. (*D*) pSP1 plasmid cleavage. (*Left*) Agarose gel analysis of pSP1 cleavage products. FLL, full-length linear DNA cut at both strands; OC, open circular DNA nicked at one of the strands; SC, supercoiled plasmid DNA. Final reaction mixtures at 37 °C contained 2.5 nM of the pSP1 plasmid and 2.5 nM of the Cas9–crRNA complex in the reaction buffer (see section B). (*Right*) Direct-sequencing electropherograms of (+) (*Upper*) and (–) (*Lower*) strands of the pSP1 plasmid cleavage product.

from the terminal end of the protospacer adjacent to the PAM, leaving blunt ends (Fig. 2*B*). Control experiments have shown that the crRNA guide is absolutely necessary for the SP1 oligoduplex cleavage, because no cleavage was produced by Cas9 in the absence of crRNA (Fig. S3*C*).

To test whether the Cas9–crRNA complex can recognize the protospacer in vitro and cut DNA in long DNA substrates mimicking in vivo invading foreign DNA, we analyzed cleavage of pSP1 (Fig. 2*C*) (16) carrying protospacer1 and the accompanying PAM. In the presence of the Cas9–crRNA complex, the supercoiled form of pSP1 plasmid was converted into a linear form (Fig. 2*D*), indicating DNA cleavage, but pUC18 plasmid lacking protospacer1 was not cleaved (Fig. 3*A*). Thus both strands of the pSP1 plasmid were cleaved specifically within the protospacer region. We used direct sequencing to determine the ends of the linear DNA molecule generated by Cas9–crRNA cleavage. Sequencing results confirmed that plasmid DNA cleavage occurred 3 nt away from the terminal end of the spacer adjacent to the PAM sequence, as seen in the SP1 oligoduplex cleavage (Fig. 2*D*). The cleavage position identified in the in vitro experiments (Fig. 2) for the CRISPR3/Cas system of *S. thermophilus* is identical to that determined in vivo for the CRISPR1/Cas system of *S. thermophilus* (20).

Cas9–crRNA Cleavage Specificity Is Directed by the crRNA Sequence. To demonstrate directly that Cas9–crRNA complex specificity is preprogrammed by the crRNA guide sequence, we engineered a unique spacer sequence (spacerSN) into the CRISPR locus, purified the Cas9–crRNA complex, and monitored cleavage of a plasmid containing a protospacerSN complementary to the engineered spacer. More specifically, we first inserted spacerSN instead of spacer1 in CRISPR3, generating the pCas(−)SN plasmid containing only a minimal CRISPR array and the tracrRNA-encoding sequence (Fig. S4*A*). We then coexpressed this plasmid together with pASKIBA-Cas9 and purified the Cas9–crRNA complex on Strep-Tactin Sepharose. The cleavage specificity of the Cas9–crRNA complex was analyzed using pSP1 and pSP1+SPN (a plasmid carrying sequence corresponding to both spacer1 and spacerSN) plasmids. Only the latter plasmid, containing the protospacerSN matching the spacerSN in the CRISPR array, was linearized by the Cas9–crRNA complex, whereas the pSP1 plasmid, which lacks the complementary sequence, remained intact (Fig. S4*B*). Similar cleavage experiments using the SPN oligoduplex containing protospacerSN and the accompanying PAM revealed that cleavage occurred within protospacerSN 3 nt upstream of the PAM, identically to other Cas9–crRNA complexes (Fig. S4 *C* and *D*). Furthermore, to provide direct evidence that Cas9 can be programmed with cognate crRNAs to recognize any DNA target, we show that Cas9–crRNA complexes targeting two different protospacers in the pUC18 plasmid cleave DNA at positions specified by the complementary crRNA (Fig. S5 *B* and *C*).

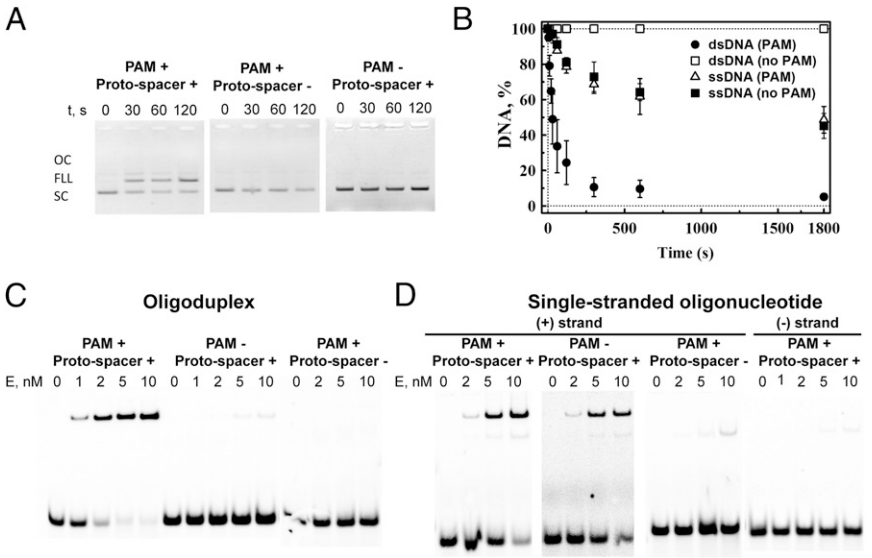

Fig. 3. A PAM is required for in vitro DNA binding and cleavage by the Cas9–crRNA complex. (*A*) Agarose gel analysis of plasmid DNA cleavage products. Three different plasmids—PAM+Protospacer+ (pSP1 containing both the protospacer and the PAM), PAM+Protospacer– (pUC18 containing multiple PAMs but no protospacer), and PAM-Protospacer+ [pSP1-pΔ (16) containing a protospacer without a PAM]—were incubated at 2.5-nM concentration with 2 nM of the Cas9–crRNA complex in the reaction buffer [10 mM Tris·HCl (pH 7.5), 10 mM NaCl, 10 mM $MgCl_2$, 0.1 mg/mL BSA] at 37 °C for various time intervals, and reaction products were analyzed by agarose gel electrophoresis. FLL, full- length linear DNA cut at both strands; OC, open circular DNA nicked at one of DNA strands; SC, supercoiled plasmid DNA. (*B*) Time courses of (+)strand hydrolysis in the single-stranded and double-stranded oligodeoxynucleotides. Reactions containing 2 nM Cas9–crRNA and 1 nM of oligodeoxynucleotide were conducted at 37 °C in the reaction buffer (see section A). SP1 (filled circles) and SP1-pΔ (open squares) oligoduplexes were used as dsDNA. s(+)SP1 (open triangles) and s(+)SP1-pΔ (filled squares) were used as ssDNA. (*C* and *D*) dsDNA and ssDNA (+)strand) binding by the Cas9–crRNA complex. The reactions contained 0.5 nM of the ^{33}P-labeled ssDNA or dsDNA oligonucleotide and the protein at the concentrations indicated above each lane. After 15 min at room temperature, the samples were subjected to PAGE for 2 h and analyzed as described in *Materials and Methods*.

The typical length of *S. thermophilus* CRISPR3 spacers is 30 nt (18). According to the data provided in Fig. 1*C*, the mature crRNA that copurified with the Cas9 protein has a length of 42 nt. Because 22 nt derive from the 3′ end of the repeat sequence, only 20 nt of the crRNA are complementary to the (+)strand of the protospacer. To assess whether the 5′ end of protospacers is important for plasmid interference, we engineered plasmids pSP1-27, pSP1-23, pSP1-19, pSP1-15, and pSP1-11 with a 5′-truncated protospacer1 and analyzed transformation efficiency of the recipient strain containing pCRISPR3 (Fig. S6*B*). Plasmids containing 4- or 7-bp truncations at the 5′ end of protospacer1 had no effect on the ability of the recipient strain to interfere with plasmid transformation. Shorter versions of protospacer (11, 15, or 19 bp) abolished the ability of the recipient strain to prevent plasmid transformation. These data show that the 10 nt at the 5′ end of protospacers, which have no complementarity to mature crRNAs, are not necessary for CRISPR3/Cas-mediated interference. In full support to the in vivo experiments, the SP1-20 oligoduplex containing only 20 nt of protospacer1 is cleaved efficiently by Cas9–crRNA (Fig. S6 *D* and *E*).

PAM Is Required for DNA Binding and Cleavage by the Cas9–crRNA Complex. Plasmids carrying a protospacer but no PAM (pSP1-pΔ) or multiple PAMs but no accompanying protospacer (pUC18) are resistant to Cas9–crRNA cleavage (Fig. 3*A*). Hence, in accordance with in vivo data, both PAM and protospacer are required for dsDNA cleavage by the Cas9–crRNA complex (16). To assess whether the PAM is recognized in the context of ds or ssDNA, we analyzed Cas9–crRNA binding to and cleavage of the oligodeoxynucleotides (*i*) SP1 (containing both protospacer and a PAM), (*ii*) SP1-Δp (containing only protospacer), and (*iii*) SP2 (containing only a PAM). The (+)strands of these oligodeoxynucleotides were used as ssDNA substrates s(+)SP1, s(+)SP1-Δp, and s(+)SP2, respectively (Table S1).

Consistent with the plasmid cleavage experiments, oligoduplexes that have only a protospacer but no associated PAM are not cut by Cas9–crRNA (Fig. 3*B*). On the other hand, the (+) strand in the single-stranded form is cut at a similar rate independently of whether it is associated with a PAM (Fig. 3*B*). In contrast, the (−)strand is cleaved only in the oligoduplex containing a PAM; in the single-stranded form, the (−)strand is resistant to Cas9–crRNA cleavage (Fig. S7). These data clearly show that a PAM is required only for dsDNA cleavage.

To test whether the PAM is necessary for DNA binding by the Cas9–crRNA complex, EMSA were performed. To avoid cleavage, binding experiments were performed in the absence of Mg^{2+} ions, which are necessary for cleavage. Cas9–crRNA showed different binding patterns for double-stranded and single-stranded oligonucleotides. In the case of the SP1 oligoduplex a low-mobility complex is observed even at a concentration of 1 nM (Fig. 3*C*). On the other hand, no binding is observed under the same experimental conditions for oligoduplexes without a PAM (SP1-Δp) or without a protospacer (SP2). Moreover, no low-mobility complex is observed in the case of Cas9 protein without crRNA (Fig. S3*A*), confirming that crRNA is necessary for complex formation. Thus, taken together, the binding experiments show that Cas9 interaction with dsDNA requires both PAM and crRNA complementarity to the target.

On the other hand, single-stranded oligonucleotides [(+)strand] are bound by Cas9–crRNA with the same affinity independent of PAM presence (Fig. 3*D*). Again, no binding was observed for ssDNA oligonucleotide corresponding to the (−)strand (Fig. 3*D*) or for Cas9 protein lacking crRNA (Fig. S3*B*).

Because some type III CRISPR systems provide RNAi rather than DNAi, we studied RNA binding and cleavage by the Cas9–crRNA complex. Cas9–crRNA did not cleave specifically either ssRNA or dsRNA bearing a protospacer and a PAM (Fig. S8*B*). This finding is consistent with DNA being the primary target for type II system-mediated interference. The Cas9–crRNA complex binds a cRNA containing a protospacer, but bound RNA is not cleaved specifically by Cas9 within the protospacer.

Mutagenesis of RuvC and HNH Motifs in Cas9. Previous plasmid-transformation experiments have revealed that the RuvC and HNH motifs (Fig. 4*A*) are important for Cas9 function (16). To test whether these motifs are involved in target DNA cleavage, we constructed, expressed, and purified the Cas9 mutants D31A and N891A. Both mutants copurified with crRNA that appeared identical to the crRNA in the WT Cas9 complex (Fig. 1*C*). To determine whether mutant proteins retained cleavage activity, we monitored pSP1 plasmid cleavage by mutant Cas9–crRNA complexes. Surprisingly, both mutants generated nicked DNA (Fig. 4*B*), indicating that both active-site mutants cleave only one DNA strand of the plasmid substrate.

To determine whether mutant proteins exhibit a strand preference, we analyzed D31A and N891A mutant cleavage of the SP1 oligoduplex. The RuvC mutant (D31A) cut the (+)strand of the oligoduplex at the same position as the WT Cas9–crRNA protein, but the (−)strand remained intact (Fig. 4*C*). Conversely, the HNH mutant (N891A) cleaved only the (−)strand but not the (+)strand of the SP1 oligoduplex (Fig. 4*D*), indicating that each active site acts on opposite DNA strands to generate a double-strand break. To find out whether the same cleavage pattern is conserved during plasmid DNA cleavage, we sequenced the protospacer regions of nicked plasmids. Run-off sequence data confirmed that RuvC is implicated in DNA (+)strand cleavage, whereas HNH/McrA is involved in (−)strand cleavage (Fig. S9 *A* and *B*). Furthermore, we found that the RuvC mutant cleaved the (+)strand of ssDNA, but no such cleavage was detected for the HNH mutant (Fig. S9*D*).

To test whether mutations altered the DNA-binding affinity of protein–crRNA complexes, DNA binding was studied using EMSA. Both mutant protein–crRNA complexes bound oligoduplex SP1 with the same affinity as WT (Fig. S8*C*). Thus, mutations in the putative active sites of Cas9 have no significant effect on the dsDNA-binding properties of the Cas9–crRNA complex. Because 42 nt of crRNA was present in the mutant protein complexes (Fig. 1*C*), we conclude that mutant Cas9–crRNA complexes lost the ability to cut one of the target DNA strands. Because the Cas9-His–tagged protein (Cas9-Chis) is a monomer in solution (Fig. S10*A*), it is likely that Cas9 is functional as a monomer and uses two active sites for the cleavage of opposite DNA strands. A similar strategy is implemented by some restriction endonucleases (22).

Discussion

Cas9–crRNA Complex Guides dsDNA Cleavage. This work demonstrates that the Cas9–crRNA complex of type II systems is a crRNA-loaded endonuclease that cuts both DNA strands within the protospacer in the presence of Mg^{2+} ions, 3 nt upstream of the PAM sequence, to produce blunt-ended cleavage products. The sequence specificity of the Cas9–crRNA complex is dictated by the 42-nt crRNA, which includes a 20-nt fragment complementary to the protospacer sequence in the target DNA. Therefore, crRNA present in the Cas9–crRNA complex of the *S. thermophilus* CRISPR3/Cas system is complementary only to the part of the protospacer sequence adjacent to the PAM. Not surprisingly, truncation of the distal part of the protospacer sequence by 10 nt

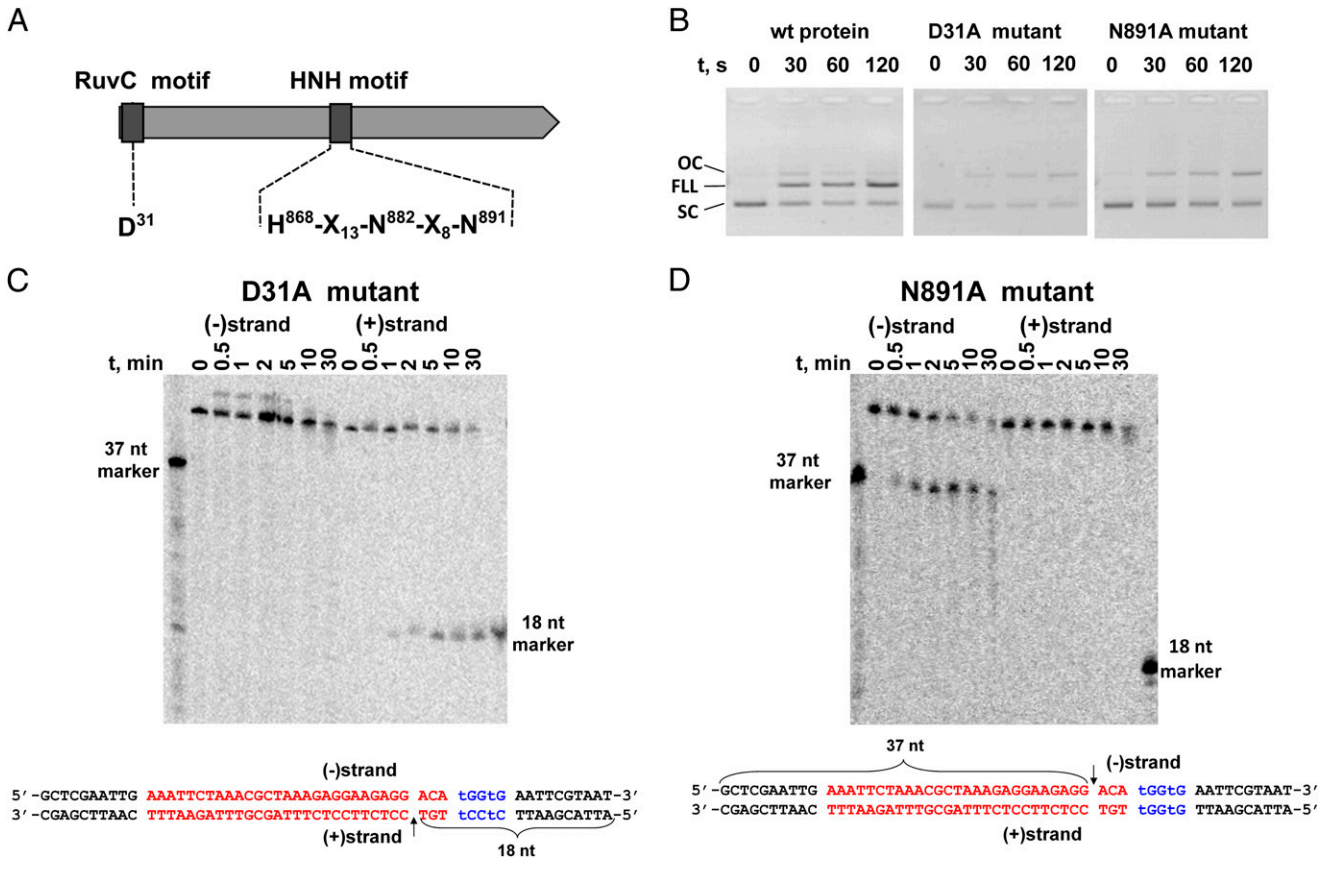

Fig. 4. RuvC and HNH active-site motifs of Cas9 contribute to the cleavage of opposite DNA strands. (*A*) Localization of the conserved active-site motifs within Cas9 protein. Amino acid residues identified as crucial for Cas9 in vivo activity (16) are indicated. (*B*) Agarose gel analysis of pSP1 plasmid cleavage by Cas9 and mutant proteins. Reactions were performed as described in *Materials and Methods*. (*C*) Strand preference of the D31A mutant. Reactions were performed as described in Fig. 2*A* and *Materials and Methods*. The D31 mutant cleaves only the (+)strand of the SP1 oligoduplex. (*D*) Strand preference of the N891A mutant. The N891 mutant cleaves only the (−)strand of the SP1 oligoduplex. Cleavage positions are indicated by arrows.

has no effect on Cas9–crRNA cleavage of synthetic oligoduplexes or plasmid DNA (Fig. S6).

The cleavage machinery of the Cas9–crRNA complex resides in the Cas9 protein, which provides two active sites for cleavage of the phosphodiester bond. The RuvC- and HNH-like active sites of Cas9 are located on different domains and act independently on individual DNA strands. Alanine replacement of the active-site residues in the RuvC and HNH motifs transforms the Cas9–crRNA complex into a strand-specific nicking endonuclease, similar to nicking enzymes (23). Consistent with in vivo studies, the functional activity of the Cas9–crRNA complex in vitro is strictly dependent on the presence of the PAM 5′-NGGNG-3′, in the immediate vicinity of the protospacer sequence. Data also show that the PAM is required for Cas9–crRNA binding to the dsDNA. If the PAM sequence is missing in dsDNA, the Cas9–crRNA complex does not bind such DNA, even if it contains a complementary protospacer sequence. On the other hand, Cas9–crRNA does not bind DNA containing PAM (or multiple PAMs) if the protospacer sequence is absent. Thus, consistent with the in vivo data, both PAM and protospacer sequences are required for dsDNA binding and subsequent cleavage. Unlike dsDNA, the PAM is not absolutely required for ssDNA binding and cleavage: A single-stranded oligodeoxynucleotide containing a protospacer with or without a PAM sequence is bound nearly as well as dsDNA. Moreover, in the presence of Mg^{2+} ions, Cas9 cuts ssDNA bound to the crRNA using its HNH active site independently of PAM.

Mechanism of DNAi in Type II Systems. Our results shed light on the mechanism of dsDNA cleavage by the Cas9–crRNA complex in the *S. thermophilus* CRISPR3/Cas system, a model type II CRISPR/Cas system (Fig. 5). Cas9–crRNA complexes, using a mechanism that yet has to be defined, locate and bind to a protospacer sequence within the dsDNA in a PAM-dependent process. The absolute requirement of the PAM for dsDNA binding by the Cas9–crRNA complex implies that the PAM serves as a priming site for strand separation or is essential for stabilization of the R-loop structure, because dsDNA lacking a PAM is not bound. Despite limited mechanistic details, our data clearly demonstrate that the PAM is recognized by Cas9–crRNA in the context of dsDNA. Cas9–crRNA binding to the target sequence in the dsDNA presumably results in a R-loop structure in which the (−)strand is displaced and the complementary (+) DNA strand is paired with the crRNA. In the presence of Mg^{2+} ions, phosphodiester bond cleavage occurs on both strands 3 nt upstream of the PAM sequence to generate blunt DNA ends. Further DNA degradation can be accomplished by host nucleases. DNA cleavage analysis by the RuvC- and HNH-motif mutants demonstrate that RuvC- and HNH-like active sites of Cas9 act on the (−) and (+)strands, respectively. Therefore, in the catalytically active Cas9–crRNA complex, the N-terminal domain containing the catalytic D31 residue of the RuvC motif is positioned at the displaced (−) DNA strand, whereas the central part of Cas9 containing the HNH motif is located in the vicinity of the scissile phosphodiester bond of the (+) DNA strand paired to loaded crRNA. Interestingly, after DNA cleavage, Cas9–crRNA remains bound to the reaction products (Fig. S11). Taken together, the data presented here provide a molecular basis for DNAi by type II CRISPR/Cas systems. Because *cas9* is a signature gene (8) for type IIA and type IIB systems, the cleavage mechanism proposed here is likely to be conserved in other type IIA and type IIB systems.

Comparison with Other RNAi Complexes. The mechanism proposed here for the cleavage of dsDNA by the Cas9–crRNA complex differs significantly from that for the type I-E (former "Ecoli") system (7). In the *E. coli* type I-E system crRNA and Cas proteins assemble into a large ribonucleoprotein complex, Cascade, that facilitates target recognition by enhancing sequence-specific hybridization between the crRNA and complementary target sequences (7). Target recognition is dependent on the PAM and governed by the seed crRNA sequence located at the 5′ end of the spacer region (24). However, although the Cascade–crRNA complex alone is able to bind dsDNA containing a PAM and a protospacer, it requires an accessory Cas3 protein for DNA cleavage. Cas3 is an ssDNA nuclease and helicase that is able to cleave ssDNA, producing multiple cuts (10). It has been demonstrated recently that Cas3 degrades *E. coli* plasmid DNA in vitro in the presence of the Cascade–crRNA complex (25). Thus, current data clearly show that the mechanistic details of the interference step for the type I-E system differ from those of type II systems, both in the catalytic machinery involved and the nature of the molecular mechanisms.

In type IIIB CRISPR/Cas systems, present in many archaea and some bacteria, Cmr proteins and cRNA assemble into an effector complex that targets RNA (6, 12). In *Pyrococcus furiosus* the RNA-silencing complex, comprising six proteins (Cmr1–Cmr6) and crRNA, binds to the target RNA and cleaves it at fixed distance from the 3′ end. The cleavage activity depends on Mg^{2+} ions; however, individual Cmr proteins responsible for target RNA cleavage have yet to be identified. The effector complex of *Sulfolobus solfataricus*, comprising seven proteins (Cmr1–Cmr7) and crRNA, cuts invading RNA in an endonucleolytic reaction at UA dinucleotides (13). Importantly, these two archaeal Cmr–crRNA complexes perform RNA cleavage in a PAM-independent manner.

Overall, we have shown that the Cas9–crRNA complex in type II CRISPR/Cas systems is a functional homolog of Cascade in type I systems and represents a minimal DNAi complex. The simple modular organization of the Cas9–crRNA complex, in which

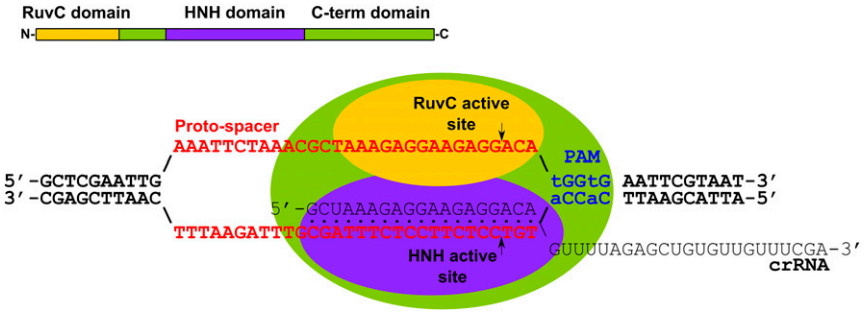

Fig. 5. Schematic arrangement and mechanism of crRNA-directed DNA cleavage by the Cas9–crRNA complex. Domain architecture of Cas9 is shown schematically at the top. The Cas9–crRNA complex binds to the dsDNA containing a PAM. crRNA binds to the complementary (+)strand, resulting in DNA strand separation and R-loop formation. In the ternary complex, the RuvC active site of Cas9 is positioned at the scissile phosphate on the unpaired (−)strand, and the HNH active site is located at the scissile phosphate on the DNA (+)strand bound to crRNA. Coordinated action of both active sites results in the double-strand break 3 nt upstream the PAM, generating blunt-ended DNA.

specificity for DNA targets is encoded by crRNAs and the cleavage enzymatic machinery is brought by a single, multidomain Cas protein, provides a versatile platform for engineering universal RNA-guided DNA endonucleases. Indeed, by altering the RNA sequence within the Cas9–crRNA complex, programmable endonucleases can be designed both for in vitro and in vivo applications. To provide proof of principle of such a strategy, we engineered de novo into a CRISPR locus a spacer targeted to a specific sequence on a plasmid and demonstrated that such a plasmid is cleaved by the Cas9–crRNA complex at a sequence specified by the designed crRNA. Experimental demonstration that RuvC and HNH active-site mutants of Cas9 are functional as strand-specific nicking enzymes opens the possibility of generating programmed DNA single-strand breaks de novo. Taken together, these findings pave the way for the development of unique molecular tools for RNA-directed DNA surgery.

Materials and Methods

DNA Manipulations. Genomic DNA of *S. thermophilus* DGCC7710 was used as a template in PCR reactions to clone *cas9*. To generate pASKIBA3-Cas9, which was used for the expression of the C-terminal Strep-tagged Cas9 protein variant, a PCR fragment was amplified with primers 5′-ACGTCTCAAATGTTGTTTAATAAGTGTATAATAATTTC-3′ and 5′-ACGTCTCCGCGCTACCCTCTCCTAGTTTG-3′. The product subsequently was cloned into the pASK-IBA3 expression vector using Esp3I sites. To generate the pBAD-Cas9 plasmid, which was used for the expression of the Cas9-Chis protein variant, PCR amplification was performed using primers 5′-ACGTCTCACATGACTAAGCCATACTCAATTGGAC-3′ and 5′-ACTCGAGACCCTCTCCTAGTTTGGCAA-3′. This product subsequently was cloned into the pBAD24-Chis expression vector using NcoI and XhoI sites. DNA sequencing of *cas9* in pASKIBA3-Cas9 and pBAD-Cas9 plasmids revealed no difference from the WT chromosomal *cas9* sequence.

To obtain plasmids pCas9(–)SP1 (Fig. 1*B*) and pCRISPR3-SP1 (Fig. S1*A*) bearing a single spacer1 (5′-AAATTCTAAACGCTAAAGAGGAAGAGGACA-3′), a PCR fragment amplified from pCRISPR3 plasmid with the primers 5′-GACCACTTATTGAGGTAAATGAG-3′ and 5′-CAAACCAGGATCCAAGCTAATACAGCAG-3′ (the BamHI site is underlined) was cloned into pCas9(–) and pCRISPR3 plasmids (16), respectively.

Expression and Purification of Cas9 Protein and Cas9–crRNA Complex. A His_6-tagged version of Cas9, Cas9-Chis, was expressed and purified using a scheme described for the Cas3 protein of the *S. thermophilus* CRISPR4/Cas system (10). For purification of the Cas9–crRNA complex, a Strep-tagged version of the Cas9 protein was expressed in *E. coli* RR1, bearing pCas9(–)SP1 plasmid (Fig. 1*B*). LB broth was supplemented with ampicillin (100 μg/mL) and chloramphenicol (10 μg/mL). *E. coli* cells for the Cas9–crRNA complex isolation were grown in two steps. First, 4 mL of cell culture was incubated at 37 °C to an OD_{600} of ∼0.5, and expression was induced by adding 0.2 μg/mL of anhydrotetracycline (AHT) (Sigma). After 4 h, 1/400th of the preinduced culture was inoculated into fresh LB medium supplemented with ampicillin (100 μg/mL), chloramphenicol (12 μg/mL), and AHT (0.2 μg/mL) and was grown at 37 °C for 17 h. Harvested cells were disrupted by sonication, and cell debris was removed by centrifugation. The supernatant was loaded onto a 1-mL Strep-Trap HP column (GE Healthcare) and eluted with 2.5 mM of desthiobiotin. Approximately 1.5 μg of the Cas9 protein was obtained in a single run from 1 L of *E. coli* culture. The fractions containing Cas9 were stored at 4 °C. The homogeneity of protein preparations was estimated by SDS/PAGE (Fig. S10*B*). Protein concentrations in the Cas9–crRNA complexes were determined by densitometric analysis of SDS/PAGE gels containing samples of Strep-Tactin–purified Cas9 proteins along with known amounts of Cas9-Chis.

Northern Blot Analysis. Cas9-bound RNA was isolated from Strep-Tactin–purified Cas9, coexpressed with pCas9(–)SP1 plasmid, using the miRNeasy Mini kit (Qiagen). Northern blots were performed by running RNA on a 10% polyacrylamide gel with 7 M urea in 20 mM Mops/NaOH (pH 8) buffer. The RNA was transferred to a SensiBlot Plus Nylon Membrane (Fermentas) by semidry blotting using a Transblot SD (Bio-Rad). RNA was cross-linked to the membrane with 0.16 M l-ethyl-3-(3-dimethylaminopropyl) carbodiimide (Pierce)/0.13 M 1-methylimidazole (Sigma), pH 8, at 60 °C for 1 h. The membrane was prehybridized with 2× SSC buffer containing 1% SDS and 0.1 mg/mL denatured DNA from fish testes (Ambion) for 1 h at 40 °C. Blots were probed for 12 h with a ^{32}P–5′-labeled 42-nt anti-crRNA DNA oligonucleotide containing 20 nt of spacer1 and 22 nt of the repeat sequence (5′-TCGAAACAACACAGCTCTAAAACTGTCCTCTTCCTCTTTAGC-3′). The blots were washed three times for 15 min with 0.2× SSC buffer containing 0.2% SDS and were visualized using phosphor imaging. A 42-nt synthetic oligoribonucleotide (5′-CGCUAAAGAGGAAGAGGACAGUUUUAGAGCUGUGUUGUUUCG-3′) and 84-nt DNA oligonucleotide (5′-AAC**AAATTCTAAACGCTAAAGAGGAAGAGG**-ACAGTTTTAGAGCTGTGTTGTTTCGAATGGTTCCAAAACTGAAGAGCTATTAG-3′(repeat sequence underlined, spacer1 sequence in bold) was used as size markers.

Oligonucleotide Substrates. All oligonucleotide substrates used in this study are listed in Table S1. Oligodeoxyribonucleotides were purchased from Metabion. The 5′ ends of oligonucleotides were radiolabeled using T4 polynucleotide kinase (PNK; Fermentas) and [γ-^{33}P]ATP (Hartmann Analytic). Duplexes were made by annealing two oligonucleotides with complementary sequences (SP1, SP1-Δp, SP2). A radioactive label was introduced at the 5′ end of individual DNA strands before the annealing with unlabeled strands.

Reactions with Oligonucleotide Substrates. Typically, reactions were carried out by adding 2 nM of Cas9–crRNA complex to 1 nM labeled oligonucleotide in 10 mM Tris·HCl (pH 7.5), 10 mM NaCl, 0.1 mg/mL BSA, and 10 mM $MgCl_2$ at 37 °C. Aliquots were removed at timed intervals and quenched with loading dye [95% (vol/vol) formamide, 0.01% bromophenol blue, 25 mM EDTA, pH 9.0] and subjected to denaturing gel electrophoresis (20% polyacrylamide) followed by a FLA-5100 phosphorimager (Fujifilm) detection.

Reactions with Plasmid Substrates. Reactions on pUC18 plasmid and its derivatives (16) were conducted at 37 °C in the buffer used for reactions on oligonucleotide substrates. Reaction mixtures typically contained 2.5 nM supercoiled plasmid and 2 nM of Cas9–crRNA complex. The reactions were initiated by adding protein to the mixture of the other components. Aliquots were removed at timed intervals and quenched with phenol/chloroform. The aqueous phase was mixed with loading dye solution [0.01% bromophenol blue and 75 mM EDTA in 50% (vol/vol) glycerol] and analyzed by agarose electrophoresis.

Determination of Plasmid Cleavage Position. To achieve complete plasmid cleavage, 8 nM of the Cas9–crRNA complex was incubated with 2.5 nM of supercoiled plasmid in the reaction buffer at 37 °C for 10 min. Reaction products were purified and concentrated using the GeneJET PCR Purification Kit (Fermentas). Spacer1 regions in Cas9 linearized and nicked pSP1 plasmids were sequenced directly with the primers 5′-CCGCATCAGGCGCCATTCGCC-3′ [(+)strand] and 5′-GCGAGGAAGCGGAAGAGCGCCC-3′ [(–)strand].

Binding Assay. Increasing amounts of Cas9–crRNA complex were mixed with 0.5 nM of ^{33}P-labeled ds and ssDNA substrates (Table S1) in the binding buffer [40 mM Tris·acetate (pH 8.3), 0.1 EDTA, 0.1 mg/mL BSA, 10% (vol/vol) glycerol)] at 25 °C and incubated for 15 min at 22 °C. Free DNA and protein–DNA complexes were separated on the nondenaturing 8% polyacrylamide gel (ratio of acrylamide/N,N′-methylenebisacrylamide, 29:1) using 40 mM Tris·acetate (pH 8.3) supplemented with 0.1 mM EDTA as the running buffer. Electrophoresis was run at 22 °C for 3 h at 6 V/cm.

Mutagenesis. The mutants D31A and N891A were obtained by site-directed mutagenesis as previously described (26). Sequencing of the entire gene for each mutant confirmed that the sole difference consisted of the designed mutation.

ACKNOWLEDGMENTS. We thank Dr. Christophe Fremaux and members of the V.S. laboratory for helpful discussions. This work was funded by the European Social Fund under Global Grant measure Project R100.

1. Barrangou R, et al. (2007) CRISPR provides acquired resistance against viruses in prokaryotes. *Science* 315:1709–1712.
2. Jansen R, Embden JD, Gaastra W, Schouls LM (2002) Identification of genes that are associated with DNA repeats in prokaryotes. *Mol Microbiol* 43:1565–1575.
3. 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.
4. 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* 22:3489–3496.
5. Brouns SJJ, et al. (2008) Small CRISPR RNAs guide antiviral defense in prokaryotes. *Science* 321:960–964.

6. Hale CR, et al. (2009) RNA-guided RNA cleavage by a CRISPR RNA-Cas protein complex. *Cell* 139:945–956.
7. Jore MM, et al. (2011) Structural basis for CRISPR RNA-guided DNA recognition by Cascade. *Nat Struct Mol Biol* 18:529–536.
8. Makarova KS, et al. (2011) Evolution and classification of the CRISPR-Cas systems. *Nat Rev Microbiol* 9:467–477.
9. Wiedenheft B, Sternberg SH, Doudna JA (2012) RNA-guided genetic silencing systems in bacteria and archaea. *Nature* 482:331–338.
10. Sinkunas T, et al. (2011) Cas3 is a single-stranded DNA nuclease and ATP-dependent helicase in the CRISPR/Cas immune system. *EMBO J* 30:1335–1342.
11. Beloglazova N, et al. (2011) Structure and activity of the Cas3 HD nuclease MJ0384, an effector enzyme of the CRISPR interference. *EMBO J* 30:4616–4627.
12. Hale CR, 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.
13. Zhang J, et al. (2012) Structure and mechanism of the CMR complex for CRISPR-mediated antiviral immunity. *Mol Cell* 45:303–313.
14. Marraffini LA, Sontheimer EJ (2008) CRISPR interference limits horizontal gene transfer in staphylococci by targeting DNA. *Science* 322:1843–1845.
15. Horvath P, Barrangou R (2010) CRISPR/Cas, the immune system of bacteria and archaea. *Science* 327:167–170.
16. Sapranauskas R, et al. (2011) The *Streptococcus thermophilus* CRISPR/Cas system provides immunity in *Escherichia coli*. *Nucleic Acids Res* 39:9275–9282.
17. Deveau H, et al. (2008) Phage response to CRISPR-encoded resistance in *Streptococcus thermophilus*. *J Bacteriol* 190:1390–1400.
18. Horvath P, et al. (2008) Diversity, activity, and evolution of CRISPR loci in *Streptococcus thermophilus*. *J Bacteriol* 190:1401–1412.
19. Mojica FJM, Díez-Villaseñor C, García-Martínez J, Almendros C (2009) Short motif sequences determine the targets of the prokaryotic CRISPR defence system. *Microbiology* 155:733–740.
20. Garneau JE, et al. (2010) The CRISPR/Cas bacterial immune system cleaves bacteriophage and plasmid DNA. *Nature* 468:67–71.
21. Deltcheva E, et al. (2011) CRISPR RNA maturation by trans-encoded small RNA and host factor RNase III. *Nature* 471:602–607.
22. Armalyte E, et al. (2005) Mva1269I: A monomeric type IIS restriction endonuclease from *Micrococcus varians* with two EcoRI- and FokI-like catalytic domains. *J Biol Chem* 280:41584–41594.
23. Chan SH, Stoddard BL, Xu SY (2011) Natural and engineered nicking endonucleases—from cleavage mechanism to engineering of strand-specificity. *Nucleic Acids Res* 39: 1–18.
24. Semenova E, et al. (2011) Interference by clustered regularly interspaced short palindromic repeat (CRISPR) RNA is governed by a seed sequence. *Proc Natl Acad Sci USA* 108:10098–10103.
25. Westra ER, 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.
26. Tamulaitis G, Zaremba M, Szczepanowski RH, Bochtler M, Siksnys V (2007) Nucleotide flipping by restriction enzymes analyzed by 2-aminopurine steady-state fluorescence. *Nucleic Acids Res* 35:4792–4799.