crRNA and tracrRNA guide Cas9-mediated DNA interference in *Streptococcus thermophilus*

Tautvydas Karvelis,[1,†] Giedrius Gasiunas,[1,†] Algirdas Miksys,[1] Rodolphe Barrangou,[2] Philippe Horvath[3] and Virginijus Siksnys[1,*]

[1]Institute of Biotechnology; Vilnius University; Vilnius, Lithuania; [2]Department of Food, Bioprocessing and Nutrition Sciences; North Carolina State University; Raleigh, NC USA; [3]DuPont Nutrition and Health; Dangé-Saint-Romain, France

[†]These authors contributed equally.

Abbreviations: Ap, ampicillin; bp, base pair; BSA, bovine serum albumin; cas, CRISPR-associated; Cm, chloramphenicol; CRISPR, clustered regularly interspaced short palindromic repeats; crRNA, CRISPR RNA; ds, double stranded; EDTA, ethylenediaminetetraacetic acid; IPTG, Isopropyl β-D-1-thiogalactopyranoside; nt, nucleotide; PAAG, polyacrylamide gel; PAGE, polyacrylamide gel electrophoresis; PAM, proto-spacer adjacent motif; PCR, polymerase chain reaction; RNP, ribonucleoprotein; SDS, sodium dodecyl sulfate; ss, single stranded; Str, streptomycin; PNK, T4 polynucleotide kinase; tracrRNA, trans-activating CRISPR RNA; Tris, tris(hydroxymethyl)aminomethane

The Cas9-crRNA complex of the *Streptococcus thermophilus* DGCC7710 CRISPR3-Cas system functions as an RNA-guided endonuclease with crRNA-directed target sequence recognition and protein-mediated DNA cleavage. We show here that an additional RNA molecule, tracrRNA (trans-activating CRISPR RNA), co-purifies with the Cas9 protein isolated from the heterologous *E. coli* strain carrying the *S. thermophilus* DGCC7710 CRISPR3-Cas system. We provide experimental evidence that tracrRNA is required for Cas9-mediated DNA interference both in vitro and in vivo. We show that Cas9 specifically promotes duplex formation between the precursor crRNA (pre-crRNA) transcript and tracrRNA, in vitro. Furthermore, the housekeeping RNase III contributes to primary pre-crRNA-tracrRNA duplex cleavage for mature crRNA biogenesis. RNase III, however, is not required in the processing of a short pre-crRNA transcribed from a minimal CRISPR array containing a single spacer. Finally, we show that an in vitro-assembled ternary Cas9-crRNA-tracrRNA complex cleaves DNA. This study further specifies the molecular basis for crRNA-based re-programming of Cas9 to specifically cleave any target DNA sequence for precise genome surgery. The processes for crRNA maturation and effector complex assembly established here will contribute to the further development of the Cas9 re-programmable system for genome editing applications.

Introduction

Clustered regularly interspaced short palindromic repeats (CRISPR) together with *cas* (CRISPR-associated) genes comprise an adaptive immune system that provides acquired resistance against invading foreign nucleic acids in bacteria and archaea.[1] CRISPR consists of arrays of short conserved repeat sequences interspaced by unique variable DNA sequences of similar size called spacers, which often originate from phage or plasmid DNA.[1-3] The CRISPR-Cas system functions by acquiring short pieces of foreign DNA (spacers), which are inserted into the CRISPR region and provide immunity against subsequent exposures to phages and plasmids that carry matching sequences.[1,4]

The highly diverse CRISPR-Cas systems are categorized into three major types, which are further subdivided into 10 subtypes, based on core element content and sequences.[5] Despite Cas diversity, the CRISPR-Cas immunity is generally performed through three stages, referred to as (1) adaptation (or immunization or spacer acquisition), (2) CRISPR expression (or crRNA biogenesis) and (3) interference (or immunity).[6-13]

During the adaptation stage, short pieces of foreign DNA (spacers) are acquired and inserted into the CRISPR locus in the host genome.[1] Although the detailed mechanism of spacer acquisition remains to be established, universal signature proteins Cas1 and Cas2 have been implicated in this step.[4,14,15] Indeed, *cas1* and *cas2* are the only *cas* genes required for new spacer selection and integration in vivo, in *Escherichia coli*.[16,17]

In the subsequent expression stage, CRISPR repeat-spacer arrays are transcribed into long primary transcripts, namely precursor crRNAs (pre-crRNAs), which are further processed into a set of short crRNAs containing a single spacer flanked by repeat fragments.[4,18-21] During the RNA maturation step in type I and type III CRISPR systems, pre-crRNA is cleaved within the repeat sequence by Cas6 endonucleases.[4,19,22-25] In contrast, type II CRISPR-Cas systems, as exemplified by *Streptococcus pyogenes* SF370, follow a different crRNA maturation pathway

Figure 1. tracrRNA in *S. pyogenes*, *S. thermophilus* LMD-9 and DGCC7710 strains. (**A**) tracrRNA-encoding regions, promoters and terminators. The tracrRNA length in *S. pyogenes* was previously determined by deep sequencing and in LMD-9 by northern blot analysis.[26] Putative promoters and terminators in DGCC7710 CRISPR3 are drawn as a dashed line. The distance in bp between the tracrRNA-encoding sequence and the *cas9* translation initiation codon is indicated above each fragment. (**B**) Alignment of tracrRNA-encoding loci in *S. thermophilus* LMD-9 CRISPR3 and DGCC7710 CRISPR3 systems. The anti-repeat regions complementary to the repeat sequences in crRNA are underlined, the putative promoters are shown in gray, the Rho-independent terminators are boxed. The RNase III cleavage site and transcription site in *S. pyogenes* is indicated by a gray triangle and arrow, respectively.

involving an additional trans-activating CRISPR RNA molecule (tracrRNA) and host RNase III.[26]

In the interference step, 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 and induces sequence-specific cleavage,[27] thereby preventing proliferation and propagation of foreign genetic elements. The structural organization and function of effector ribonucleoprotein (RNP) complexes involved in crRNA-mediated silencing of foreign nucleic acids differ between distinct CRISPR-Cas types.[11] In the type I-E systems of *E. coli*[4] and *Streptococcus thermophilus* CRISPR4-Cas,[28] crRNAs are incorporated into a multisubunit effector complex called Cascade (CRISPR-associated complex for antiviral defense),[4] which binds to the target DNA and triggers degradation by the signature Cas3 protein.[28-32] In type III CRISPR-Cas systems of *Sulfolobus solfataricus* and *Pyrococcus furiosus,* RNP complexes of Cas RAMP (Cmr) proteins and crRNA recognize and cleave synthetic RNA in vitro,[19,33,34] while the CRISPR-Cas system of *Staphylococcus epidermidis* targets DNA in vivo.[35] In type II CRISPR-Cas systems, as exemplified by *S. pyogenes*[36] and *S. thermophilus* CRISPR3-Cas (St-CRISPR3-Cas),[37-39] a single Cas9 protein, instead of a multisubunit Cascade or Cmr protein complex, provides DNA silencing.[37] In fact, the type II effector complex functions as an RNA-guided endonuclease which achieves target sequence recognition by crRNA and employs Cas9 for DNA cleavage within the target proto-spacer.

The *S. thermophilus* DGCC7710 strain contains four distinct CRISPR-Cas systems: CRISPR1, CRISPR2, CRISPR3 and CRISPR4.[6] Direct spacer acquisition activity has been demonstrated for the CRISPR1 and CRISPR3 systems, with the former being more active in this strain.[1,27,40] CRISPR1 and CRISPR3, which both belong to type II CRISPR/Cas systems,[5] share a similar architecture, with four *cas* and tracrRNA genes located upstream of the CRISPR spacer-repeat array (**Fig. 1A**). The proto-spacer adjacent motifs (PAM) characteristic for CRISPR1

and CRISPR3 systems differ and are NNAGAAW and NGGNG, respectively.[40] The CRISPR2 and CRISPR4[28,32] systems belong to the types III and I-E, respectively.[5,6]

The effector complex of the type II St-CRISPR3-Cas system contains a 42 nt crRNA combined with St-Cas9.[39] While tracrRNA involvement in the *S. pyogenes* effector complex has been documented,[36] the role of tracrRNA in DNA silencing provided by the St-CRISPR3-Cas effector complex remains to be established. Here, we show that the St-Cas9 protein preparation isolated from the heterologous *E. coli* strain contains both tracrRNA and crRNA. We also provide experimental evidence that tracrRNA is strictly required for Cas9-mediated DNA interference both in vitro and in vivo. Furthermore, we demonstrate that RNase III is required for processing of the primary pre-crRNA, but is not necessary for the maturation of a crRNA derived from a CRISPR array containing a single spacer. Finally, we show that Cas9 protein promotes annealing of the pre-crRNA:tracrRNA duplex in vitro and demonstrate that an in vitro-assembled ternary Cas9-crRNA-tracrRNA complex cleaves DNA.

Results

Genetic location of the tracrRNA-encoding sequence in the *S. thermophilus* DGCC7710 CRISPR3-Cas system. Sequence analysis of the DNA region upstream of the *cas9* gene in the St-CRISPR3-Cas system revealed that, similarly to the *S. pyogenes* CRISPR-Cas system,[26] a putative tracrRNA (St-tracrRNA) is encoded upstream of the *cas* operon (**Fig. 1A**). In *S. pyogenes*, deep sequencing data[26] revealed 171 nt and 89 nt tracrRNAs (Sp-tracrRNAs) that would result from transcription at two distinct promoters to a shared transcriptional terminator (**Fig. 1A**). Northern blot analysis of the *S. thermophilus* LMD-9 strain, which contains a CRISPR3-Cas system nearly identical to that of *S. thermophilus* DGCC7710 CRISPR3-Cas system,[39] revealed a ~100 nt tracrRNA transcript.[26]

Figure 2. The tracrRNA is required for interference. (**A**) Schematic representation of plasmids used for plasmid transformation interference assays. The pCRISPR3-Δt plasmid encodes a CRISPR3-Cas system without tracrRNA. In pCRISPR3-ΔtR, tracrRNA contains only the anti-repeat region and lacks its 3'-end. ptracrRNA plasmid was obtained by inserting a full-length tracrRNA-encoding sequence under the control of T7 RNA polymerase promoter in the pCDF-DUET plasmid. (**B**) The deletion or shortening of tracrRNA inactivates CRISPR3-Cas interference. (**C**) tracrRNA can be provided in trans on a separate plasmid. (**D**) Cas9 co-purifies with ~65 nt tracrRNA and 42 nt tracrRNA. Northern blot analysis of nucleic acids extracted from purified active St-Cas9 complex using anti-tracrRNA (left panel) and anti-crRNA (right panel) oligonucleotide probes. The estimated size of the tracrRNA is ~65 nt, albeit minor amounts of longer tracrRNAs intermediates are present. M, RNA size markers.

To identify a putative transcriptional start position for the DGCC7710 St-tracrRNA transcript, we compared DNA fragments located between the translation initiation codon of Cas9 and the anti-repeat region of the tracrRNA (**Fig. 1A**). The distance between the *cas9* start codon and tracrRNA is ~130 bp in DGCC7710 vs. ~200 bp in LMD-9 and ~300 bp in *S. pyogenes*. Within the shared frame, the corresponding DGCC7710 fragment is nearly identical to that of LMD-9. Therefore, we assumed that the transcription of St-tracrRNA (DGCC7710) could start at the same position as in LMD-9. The tracrRNA sequences (**Fig. 1B**) at the 3'-end were very similar (87% identical nucleotides) and contained Rho-independent transcription terminators. Based on this, we assumed that the estimated size of St-tracrRNA (DGCC7710) is ~100 nt, which would be consistent with the size of St-tracrRNA (LMD-9).[26]

tracrRNA is necessary for in vivo DNA interference by St-CRISPR3-Cas. The pCRISPR3 plasmid, which carries a complete St-CRISPR3-*cas* locus, including a CRISPR3 array comprised of 12 repeat-spacer units and a tracrRNA-encoding fragment located upstream of the *cas9* gene (**Fig. 2A**), provides interference against transformation of a donor pSP1 plasmid, which contains a proto-spacer identical to the SP1 spacer in the CRISPR3 array and the accompanying 5-NGGNG-3' PAM sequence.[37,39] To establish whether tracrRNA is required for St-CRISPR3-Cas-mediated plasmid interference in *E. coli*, we generated pCRISPR3 plasmid variants with a compromised tracrRNA-encoding sequence (**Fig. 2A**). In the pCRISPR3-Δt plasmid, the entire tracrRNA coding sequence is deleted, while in the pCRISPR3-ΔtR variant the tracrRNA-encoding sequence is truncated at the 3'-end, to leave only the region from the transcription start site up to the end of the anti-repeat sequence (**Fig. 2A**). Next, we analyzed transformation efficiency of two recipient *E. coli* strains carrying tracrRNA-deficient pCRISPR3-Δt or pCRISPR3-ΔtR variants by the pSP1 donor plasmid. The pUC18 plasmid, which lacks a proto-spacer but contains multiple PAMs was used as a control in the plasmid transformation

assay. We found that, in contrast to the pCRISPR3-carrying recipient strain, which was resistant to pSP1 transformation, the tracrRNA-deficient strains became permissive for transformation by pSP1 plasmid (**Fig. 2B**). Thus, the plasmid immunity provided by the heterologous St-CRISPR3-Cas system is compromised when either part of or the full tracrRNA-encoding sequence is eliminated, indicating that tracrRNA is necessary for Cas9-mediated interference.

To confirm that the St-CRISPR3-Cas-mediated plasmid interference is lost due to the tracrRNA gene deletion/truncation, we introduced a full-length tracrRNA in trans into the tracrRNA-deficient permissive recipient strain and evaluated pSP1 plasmid transformation efficiency. More specifically, the fragment encoding tracrRNA was cloned into the pCDF-DUET1 plasmid under T7 promoter control, the construct expressed in the pSP1 permissive *E. coli* BL21 (DE3) strain (**Table S1**) carrying the pCRISPR3-Δt plasmid (**Fig. 2C**) and transformation efficiency was evaluated by counting colonies on Ap, Str, Cm and IPTG-supplemented agar plates. Because the *E. coli* BL21 (DE3) strain is transformed less efficiently than RR1, 25 ng instead of 1 ng of pUC18 or pSP1 plasmid was used in the transformation assay. Under these conditions, we found that transformation by the pUC18 plasmid yielded ~1,000 colonies, while no colonies were obtained in the case of pSP1 plasmid (**Fig. 2C**). The same results were obtained in a recipient host, which carried the pCRISPR3 plasmid containing tracrRNA in cis (**Fig. 2C**). These results show that trans-complementation of tracrRNA converts a pSP1-permissive *E. coli* strain into a transformation-resistant (non-permissive) strain. Taken together, plasmid transformation assays demonstrate that tracrRNA is necessary for DNA interference provided by the St-CRISPR3-Cas effector complex, in vivo.

tracrRNA co-purifies with Cas9 protein. St-Cas9 is the sole Cas protein required for St-CRISPR3-Cas-mediated immunity.[37,39] St-Cas9 protein co-purifies with a 42 nt crRNA (**Fig. 2D**).[39] To probe whether tracrRNA also co-purifies with Cas9 and crRNA, we performed northern blot analysis using an anti-tracrRNA 36 nt oligodeoxynucleotide probe. Nucleic acids extracted from the Strep-Tactin-purified Cas9 preparation[39] hybridized with the anti-tracrRNA probe and were sensitive to RNase, but not to DNase treatment (**Fig. 2D**). The estimated size of the tracrRNA, which co-purified with Cas9 is ~65 nt. Minor amounts of longer tracrRNA intermediates were present (**Fig. 2D**). The northern blot revealed that Cas9 preparation expressed and purified in the absence of CRISPR3 repeat-spacer array sequence and tracrRNA, contains no crRNA or tracrRNA (**Fig. S1**).

Role of RNase III in pre-crRNA maturation. In *S. pyogenes,* crRNA maturation and effector complex assembly requires tracrRNA and a housekeeping RNase III.[26] During *S. pyogenes* crRNA maturation, the tracrRNA base pairs with pre-crRNA to form a double-stranded RNA that is recognized and cleaved by the host RNase III in the presence of Cas9.[26] Deletion of the *rnc* gene encoding RNase III in *S. pyogenes* abrogates DNA interference provided by the CRISPR-Cas system.

The *E. coli* RNase III is similar to St-RNaseIII (39.6% identical and 52.3% similar aa). To probe whether the *E. coli* RNase III is able to replace St-RNase III in the crRNA processing and contribute to the St-CRISPR3-Cas mediated immunity in the heterologous *E. coli* host, we performed pSP1 plasmid transformation assay in the *rnc+* (*E. coli* RR1) and *rnc-* (*E. coli* HT115) strains[41] (**Table S1**) carrying either a wild-type (with 12 distinct spacers in-between repeats) or a minimal (R-SP1-R) CRISPR locus (**Figs. 2B and 3A**, respectively). In-line with our previous findings,[37,39] the *rnc+* host, carrying either a wild-type or a minimal CRISPR region, provided interference against pSP1 transformation with nearly equal efficiency. On the other hand, in the *rnc-* host, the transformation efficiency radically differed depending on whether the recipient strain carried a wild-type or a minimal (R-SP1-R) CRISPR array. The *rnc-* strain carrying a wild-type CRISPR array became permissive for pSP1 plasmid transformation while the *rnc-* strain carrying a minimal CRISPR array was non-permissive to pSP1 transformation (**Fig. 3B**). This indicates that RNase III is not necessary for the maturation of pre-crRNA containing a single CRISPR spacer. This reduces the cloning needs and system requirements for a single-target Cas9-mediated programmable cleavage.

In vitro reconstitution of the St-CRISPR3 Cas9-crRNA-tracrRNA effector complex. Analysis of the protein and nucleic acid content of the St-CRISPR3-Cas effector complex isolated from the heterologous *E. coli* strain revealed the presence of Cas9, a ~42 nt crRNA[39] and a ~65 nt tracrRNA (see above). Next, we aimed to reconstitute an effector complex in vitro, by combining these three individual components. To assemble a ternary Cas9-crRNA-tracrRNA complex (St-Cas9t), Cas9 was pre-incubated with equimolar amounts of a synthetic 42 nt crRNA and 78 nt tracrRNA corresponding to a mature form, obtained by in vitro transcription. Extra nucleotides (78 nt vs. 65 nt) were introduced for in vitro transcription by a T7 RNA polymerase. The DNA cleavage activity of the reconstituted complex was monitored in vitro using the pSP1 plasmid, as described previously.[39] In the presence of a reconstituted ternary St-Cas9t complex, the pSP1 plasmid, which contains a proto-spacerSP1 sequence flanked by the 5'-NGGNG-3' PAM is converted into a linear form (**Fig. 4A**), indicating that both DNA strands are cleaved. If one of the complex components (Cas9, crRNA or tracrRNA) is missing, no cleavage of pSP1 is observed, indicating that all three components are necessary to form a functional St-Cas9t complex. On the other hand, the pUC18 plasmid which lacks the proto-spacer is not cleaved by the St-Cas9t. Hence, a catalytically competent Cas9-crRNA-tracrRNA ternary complex can be reconstituted in vitro by mixing individual components, similarly to the sp-Cas9 complex of *S. pyogenes*.[36]

According to the plasmid interference assay, RNase III is not required for DNA interference in the *E. coli* host carrying a plasmid with a minimal R-SP1-R unit (**Fig. 3**). Therefore, we next investigated whether RNase III is important for St-Cas9t complex assembly in vitro. To assemble the complex, we first pre-incubated the Cas9 protein with an equimolar mixture of a 105 nt tracrRNA (101 nt putative tracrRNA transcript + 4 nt for in vitro transcription by the T7 RNA polymerase) and a 150 nt pre-crRNA (including 20 nt of the leader and a R-SP1-R-SP2 fragment, which contains the targeting spacer SP1 and a truncated
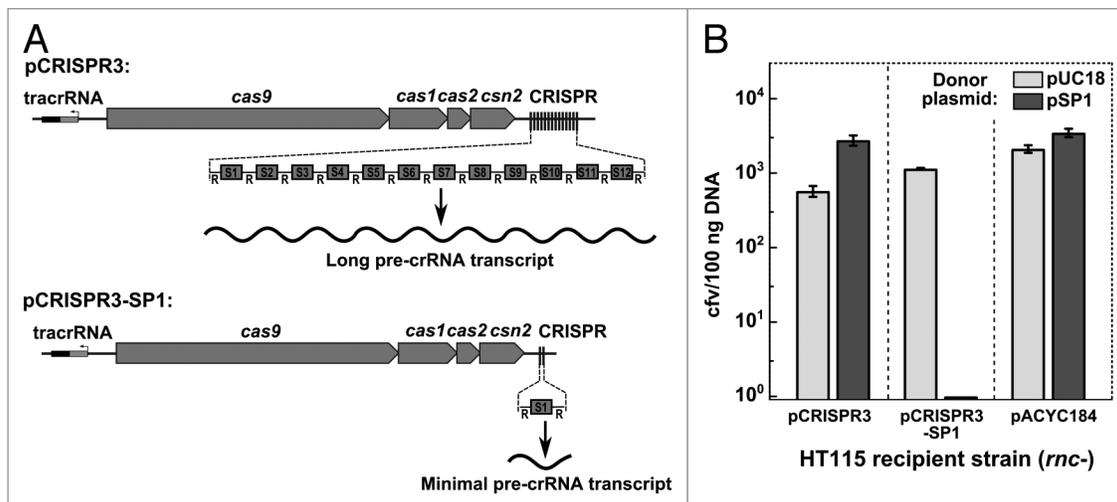
Figure 3. Involvement of tracrRNA and RNase III in plasmid interference. (**A**) Schematic representation of plasmids used in the plasmid transformation assay and corresponding pre-crRNA transcripts. The pCRISPR3 plasmid encodes a full-length pre-crRNA transcript, whereas pCRISPR3-SP1 encodes a minimal pre-crRNA transcript. (**B**) CRISPR3-Cas system plasmid interference assay in the *rnc-* (RNase III-deficient) *E. coli* HT115 strain.

SP2) transcripts, and then added RNase III. The pSP1 plasmid cleavage assay revealed that the complex assembled in vitro using pre-crRNA cleaved the pSP1 plasmid (**Fig. 4B**), generating a linear DNA form albeit at slower rate than a complex assembled using a 42 nt synthetic crRNA. Elimination of the RNase III from the reaction mix produced an active complex (**Fig. 4B**), consistent with in vivo data indicating that RNase III is not crucial for the assembly of the catalytically active complex when a short pre-crRNA (containing only one spacer) transcript is used.

Experiments with sp-Cas9 revealed that tracrRNA can be 3'-truncated in in vitro applications.[36] To define the minimal length of the 3'-end of tracrRNA required for crRNA-guided DNA cleavage by the St-Cas9t, we designed and generated a set of tracrRNA molecules truncated by stretches of 5 nt from the 3' terminus. The overall size of tracrRNA variants varied between 33 and 78 nt (**Fig. 4C**). The truncated tracrRNA variants were used for in vitro reconstitution of the ternary St-Cas9t complex, followed by analysis of the pSP1 plasmid cleavage. We found that 38 nt and 33 nt tracrRNA variants do not support DNA cleavage by St-Cas9 (**Fig. 4D**), whereas 43, 48 and 53 nt variants showed decreased cleavage activity. On the other hand, tracrRNAs longer than 58 nt support efficient St-Cas9 cleavage of the pSP1 plasmid. This again reduces cloning requirements for a synthetic system, and simplifies the elements necessary for Cas9-mediated programmable cleavage.

Cas9 role in formation of the pre-crRNA:tracrRNA duplex. In the *S. thermophilus* CRISPR3-Cas and *S. pyogenes* effector complexes, the target site recognition is achieved by the crRNA while the Cas9 protein provides two active sites for the cleavage of opposite DNA strands in the proto-spacer.[36,39] Indirect evidence obtained in vivo in *S. pyogenes* suggests that Cas9 is also an essential protein for crRNA maturation and may facilitate formation and stabilization of the pre-crRNA:tracrRNA duplex.[26] However, experimental evidence and molecular details remain to be established.

To probe experimentally whether the *S. thermophilus* CRISPR3 system Cas9 protein promotes pre-crRNA and tracrRNA annealing, we produced tracrRNA3 (105 nt) and pre-crRNA3 (94 nt containing a single 36 nt repeat sequence) from CRISPR3 by in vitro transcription and monitored tracrRNA3:pre-crRNA3 duplex formation in the presence of Cas9 (**Fig. 5**). In a control set of experiments, tracrRNA1 (105 nt) and pre-crRNA1 (94 nt) from the homologous *S. thermophilus* DGCC7710 CRISPR1 system were used (**Table S4**). To monitor duplex formation, radioactively labeled tracrRNAs were incubated for 10 min with pre-crRNAs in the presence of a non-specific 2 kb RNA transcript and different amounts of the Cas9 protein of the CRISPR3 system. The reaction was quenched by adding Proteinase K and an excess of unlabeled tracrRNA and samples were analyzed using non-denaturing PAGE (**Fig. 5**). Under these experimental conditions, in the absence of Cas9, no tracrRNA:pre-crRNA duplex is formed. Increasing amounts of Cas9 promote tracrRNA3:pre-crRNA3 formation but have no effect with tracrRNA1 and pre-crRNA1 (**Fig. 5**). Taken together, these results indicate that Cas9 specifically promotes pre-crRNA3 and tracrRNA3 annealing and facilitates formation of the pre-crRNA3:tracrRNA3 duplex.

Discussion

CRISPR-Cas systems provide adaptive immunity against viruses and plasmids in bacteria and archaea. The silencing of invading nucleic acids is executed by RNP complexes pre-loaded with small interfering crRNAs that act as guides for sequence-specific targeting and degradation of alien nucleic acids. To produce crRNA, CRISPR repeat-spacer arrays are transcribed into long primary transcripts (pre-crRNAs) that are further processed into a set of short CRISPR RNAs (crRNAs) containing a conserved repeat fragment ("handle" or "tag") and a variable spacer sequence ("guide"), which is complementary to the foreign nucleic acid.[4,22] Although this step has commonalities across different CRISPR-Cas types,

Figure 4. Plasmid DNA cleavage by the in vitro assembled effector pre-crRNA:tracrRNA:Cas9 complex. (**A**) Cas9, 42 nt crRNA and 78 nt tracrRNA were used for complex assembly. The Cas9-crRNA complex was incubated with pSP1 or pUC18 plasmids in a reaction buffer. pSP1 plasmid contained a proto-spacer SP1 sequence flanked by the 5'-NGGNG-3' PAM sequence. Proto-spacer SP1 sequence is absent in pUC18. Reaction products were analyzed by electrophoresis on agarose gel. SC, super-coiled plasmid DNA; OC, open circular DNA nicked on one of DNA strands; FLL, full-length linear DNA cut on both strands. Under the reaction conditions the pSP1 plasmid is converted into a linear form while the pUC18 plasmid lacking the proto-spacer SP1 sequence is resistant to cleavage. In the reaction mixes lacking one of the components (Cas9, crRNA or tracrRNA) pSP1 plasmid is not cleaved. (**B**) pSP1 plasmid cleavage by the effector complex assembled in vitro in presence and absence of RNase III. The effector complex was assembled by mixing st-Cas9, 150 nt pre-crRNA, 105 nt tracrRNA in the presence (left panel) or absence (right panel) of RNase III. The 150 nt pre-crRNA is arranged as follows (5'-3'): GGG-20 nt leader sequence 36 nt repeat(R1), 30 nt targeting spacer (S1), 36 nt repeat(R2), 25 nt spacer (truncated S2). (**C**) Schematic representation of 42 nt crRNA:78 nt crRNA duplex structure. The putative secondary structures of the non-complementary part of the tracrRNA identified by RNAfold[55] are shown. The positions of 5 nt 3'-truncations in the non-complementary region are indicated. (**D**) pSP1 plasmid cleavage by the effector complex assembled in vitro using tracrRNAs of different lengths.

Figure 5. tracrRNA and pre-crRNA annealing in the presence of Cas9. 105 nt tracrRNA and 94 nt pre-crRNA used in the annealing assay are represented above the gels. Complementary sequences are indicated in black. Non-denaturing PAGE analysis of duplex assembly between tracrRNA3 and pre-crRNA3 of CRISPR3 system (left panel) and tracrRNA1 and pre-crRNA1 from CRISPR1 (right panel) at varying concentrations of Cas9 from CRISPR3. C, control lanes, containing only labeled tracrRNA; D, tracrRNA/pre-crRNA duplex, formed by heating tracrRNA:pre-crRNA mixture (1:100 molar ratio) to 95°C and slowly cooling down to room temperature.

the crRNA maturation details differ. In the type I and type III systems, pre-crRNA is cleaved within the conserved repeat sequence by Cas6 endonucleases (Cas6, Cas6e or Cas6f).[4,19,22-25] All three endonucleases cleave pre-crRNA within the repeat region, 8 nt upstream of the spacer sequence to produce conserved 5'- and 3'-handles. After cleavage, Cas6e and Cas6f remain tightly bound to the crRNA and serve as docking points for type I effector complex assembly.[23,42] In the type III system, exemplified by *P. furiosus*, Cas6 does not become an integral part of the effector complex after precrRNA cleavage. In this case, the 8 nt 5' repeat handle of the mature crRNA serves as an anchor for the assembly of the RNP complex.[19] The resulting effector complexes that execute silencing in the type I and type III systems are large multisubunit RNPs, which include multiple proteins that form the CRISPR associated complex for antiviral defense, Cascade.[4,19,43]

The crRNA maturation and processing pathway in type II systems, as exemplified by *S. pyogenes*, differs significantly. It has been shown that the process requires an additional tracrRNA molecule and depends on the housekeeping RNase III.[26] The tracrRNA pairs with complementary repeat sequences within the pre-crRNA primary transcript and forms an RNA duplex, pre-crRNA:tracrRNA, which is recognized and cleaved by RNase III in the presence of Cas9 protein.[26] In the effector complex of *S. pyogenes*, Cas9 is bound to the crRNA and tracrRNA into a ternary complex which executes silencing.[36]

Here, we focused on the crRNA maturation pathway and effector complex assembly in the type II CRISPR3-Cas system of *S. thermophilus* DGCC7710. Sequence analysis revealed the presence of a putative tracrRNA-encoding sequence and estimated a tracrRNA transcript of ~100 nt (**Fig. 1**). Plasmid interference assays revealed that a DNA fragment encoding the putative tracrRNA is critical for DNA interference (**Fig. 2**). Moreover, northern blot analysis revealed that the ~65 nt tracrRNA co-purifies with the 42

nt crRNA and Cas9 protein (**Fig. 2**). Hence, the *S. thermophilus* DGCC7710 CRISPR3-Cas effector complex that provides interference against DNA consists of a ternary Cas9-crRNA-tracrRNA complex similarly to the *S. pyogenes* effector complex.[36] The length of the tracrRNA (~65 nt) co-purified with an effector complex isolated from the cell suggests that the tracrRNA transcript may undergo further processing before or after incorporation into the ternary effector complex. The nucleases that contribute to the tracrRNA transcript processing remain to be established.

It has been shown that in *S. pyogenes*, tracrRNA is required for the pre-crRNA maturation which also depends on the host RNase III.[26] It is thought that tracrRNA is annealed to the complementary regions of pre-crRNA forming a duplex that is cleaved by RNase III. We have found that the *S. thermophilus* CRISPR3-Cas long primary pre-crRNA maturation in the heterologous *E. coli* host depends on RNase III. Indeed, no plasmid interference is provided by the CRISPR3 system in the *rnc- E. coli* strain, which lacks RNase III (**Fig. 3**). On the other hand, RNase III is not required for the processing of a "minimal" crRNA transcript which includes a single spacer (**Fig. 3**). It seems that a short pre-crRNA is incorporated into the effector complex independently of host RNase III and does not require 3'-end processing by RNase III for interference against pSP1 plasmid transformation. These data suggest that RNase III in the type II systems is required for the dicing of a long pre-crRNA and generates the 3'-end of mature crRNAs when the CRISPR locus contains multiple repeat-spacer units. In this respect, host RNase III functions as an alternative for the Cas6 family nucleases, which are involved in the pre-crRNA processing in type I and III systems. However, it remains to be established whether other cellular nucleases contribute to the processing of the 5'-end of crRNA in type II systems. We found next that the catalytically active effector complex of *S. thermophilus* can be assembled in vitro by mixing Cas9 protein,

tracrRNA and crRNA (**Fig. 4**). Again, consistent with the in vivo data, *E. coli* RNase III was not absolutely required for the ternary complex assembly when a minimal pre-crRNA was used (**Fig. 4B**). We further used in vitro complex assembly to establish the tracrRNA minimal length for formation of the catalytically competent effector complex. We found that the tracrRNA can be shortened at the 3'-end and maintain activity, provided that at least 15 nt are maintained beyond the crRNA-complementary sequence. Further 3'-trimming of tracrRNA compromises DNA cleavage activity, suggesting that the tracrRNA fragment close to the duplex junction may be critical for interaction with the Cas9 protein. tracrRNAs of 53, 48 and 43 nt yielded effector complexes with compromised DNA cleavage activity (**Fig. 4C**). Interestingly, secondary structure analysis of tracrRNA suggests formation of three putative hairpin structures at its non-complementary 3' terminus (**Fig. 3C**). Our data indicate that the 3'-end proximal hairpin is not necessary for effector complex formation, while two other hairpins are required for Cas9 binding and/or cleavage. We were unable to detect specific tracrRNA binding by Cas9 protein using gel shift experiments, but we found that the in vitro assembled ternary complex is more active when the Cas9 protein is first pre-incubated with tracrRNA followed by addition of crRNA (data not shown).

It has been shown that the Sp-Cas9 protein is important for crRNA maturation.[26] We found here that the Cas9 protein of *S. thermophilus* CRISPR3 system specifically promotes crRNA:tracrRNA duplex formation. Of note is that Cas9 of the CRISPR3 system promoted formation of the crRNA3:tracrRNA3 duplex of the CRISPR3 system but had no effect on the crRNA1:tracrRNA1 duplex of the CRISPR1 system (**Fig. 5**). These data suggest that the Cas9 protein may specifically recognize structure and/or sequence determinants in the tracrRNA:crRNA duplex.

Taken together, experimental data provided here suggest the following mechanism for crRNA maturation and effector complex assembly in the *S. thermophilus* CRISPR3-Cas system and presumably in other type II systems (**Fig. 6**). The CRISPR array is transcribed into a long primary pre-crRNA that forms a duplex with tracrRNA containing a complementary sequence. Cas9 specifically promotes pre-crRNA:tracrRNA duplex annealing. It remains to be established whether the Cas9 contribution is due to the active role in the crRNA and tracrRNA annealing, or to the ability to remove secondary structures in the single-stranded crRNA and tracrRNA molecules. We cannot exclude that Cas9 remains bound to the duplex after the strand annealing, and the tracrRNA 3' terminus may contribute to Cas9 binding.

tracrRNA is not present in the effector complexes of the type I and III CRISPR systems. In the type I systems, five different Cas proteins form the Cascade complex, and assemble into a sea-horse-shaped structure which provides a scaffold for crRNA binding. Cascade-bound crRNA then locates a homologous DNA sequence, binds to the target site generating an R-loop and subsequently recruits the Cas3 helicase/nuclease to execute cleavage.[28,31,32] A similar mechanism is also found in type III systems.[19,33,34]

Type II systems in contrast contain a less complex CRISPR-*cas* locus encoding only three or four Cas proteins,[5] including the universal Cas1 and Cas2 proteins, as well as Cas4[44] or Csn2,[45-48] which are involved in new spacer acquisition. In type II systems, DNA binding and cleavage is achieved by an effector complex containing a sole Cas9 protein bound to the crRNA and tracrRNA.[36,39] It is tempting to speculate that in the type II systems Cas9-bound tracrRNA provides a scaffold for the crRNA binding and stabilization similarly to Cascade proteins in type I and III systems.[19,34,43,49-51] Once the complex is formed, RNase III cleaves double-stranded regions generating Cas9-crRNA complexes that undergo further maturation by unknown nuclease(s) that trim the 5'-end of the spacer[26] (**Fig. 6**) to generate a mature 42 nt crRNA containing 22 nt spacer and 20 nt repeat sequence.[26,39] The host RNase III functions as an equivalent of the Cas6 family nucleases, which are involved in the pre-crRNA processing and maturation in the type I and III systems. Interestingly, we show that RNase III is not required for the effector complex assembly since an active effector complex is assembled both in vivo and in vitro in the case of a minimal CRISPR locus, which contains a single spacer (**Fig. 6B**).

The demonstration that the DNA interference complex in type II CRISPR-Cas systems functions as an RNA-guided endonuclease that uses RNA for target site recognition and Cas9 for DNA cleavage[36,39] paved the way for the development of the novel tools for genome editing, notably in eukaryotes.[52,53] The mechanisms of crRNA maturation and effector complex assembly established here should contribute to the further development and fine-tuning of these molecular tools for re-programmable gene surgery and genome editing applications.

Materials and Methods

DNA manipulations. A DNA fragment encoding *cas9* was PCR amplified from *S. thermophilus* DGCC7710 strain genomic DNA using primers GG-384 and GG-385 (**Table S2**) and pre-cleaved with *Esp*3I and *Xho*I. The resulting *Esp*3I-*Xho*I fragment was cloned into a pBAD24-CHis expression vector using *Nco*I and *Xho*I sites to generate a pBAD-Cas9 plasmid, which was used for the expression of the C-terminal (His)$_6$-tagged Cas9 protein variant. Full sequencing of *cas9* in the pBAD-Cas9 plasmid revealed no difference with the original *cas9* sequence.

To generate pCRISPR3-Δt and pCRISPR3-ΔtR plasmid variants lacking an entire tracrRNA-encoding gene or its 5'-terminal fragment, respectively, DNA fragments were PCR amplified from pCRISPR3 plasmid[37] using primer pairs T1/T2 and T1/T3, and cloned into pCRISPR3 plasmid over *Xba*I and *Sna*BI sites. To obtain the ptracrRNA plasmid, encoding tracrRNA under the control of a T7 RNA polymerase promoter, a DNA fragment was PCR amplified using the T4/T5 primer pair and cloned into the pCDF-DUET (Invitrogen) expression vector using *Nco*I and *Pst*I sites.

Plasmid transformation. Plasmid transformation assays were performed as described previously.[37] The *E. coli* RR1 strain was transformed with 1 ng, BL21 (DE3) strain, with 25 ng and HT115 strain, with 100 ng of pUC18 or pSP1 plasmid (**Table S1**). The transformants were plated on LB agar supplemented with appropriate antibiotics and 0.1 mM IPTG (for BL21 (DE3) strain). All
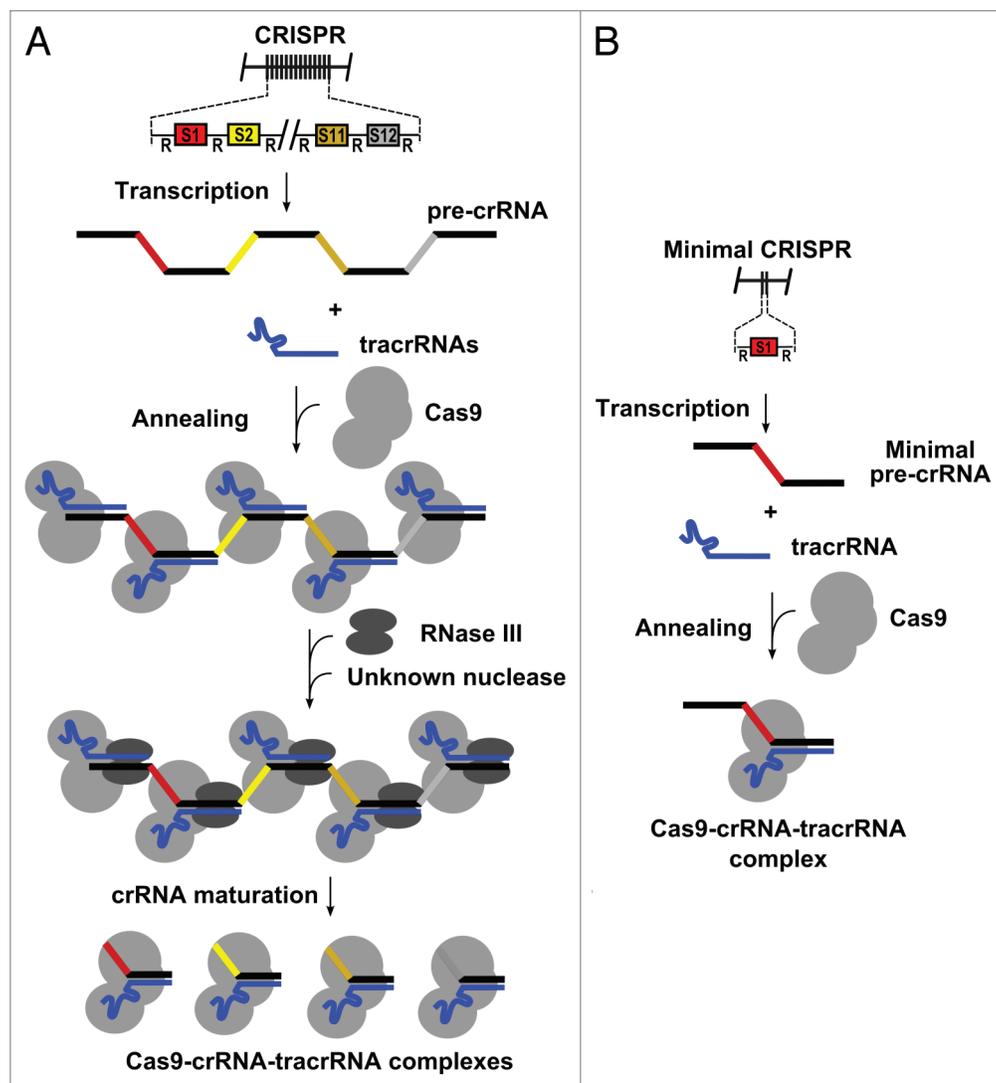
Figure 6. Schematic representation of crRNA maturation/processing and Cas9-crRNA-tracrRNA complex assembly pathways. (**A**) The wild-type CRISPR locus is transcribed as a long pre-crRNA molecule. Cas9 promotes pre-crRNA:tracrRNA duplex formation. The tracrRNA/pre-crRNA duplex regions are recognized and cleaved by the host RNase III to generate effector complexes that undergo further trimming at the 5' end by unknown nuclease(s) to produce mature Cas9-crRNA-tracrRNA complexes. (**B**) Short pre-crRNA transcripts produce functional effector complexes in the absence of RNase III.

transformation experiments were repeated at least three times. Bars in the graphs represent mean values from three or more independent experiments ± 1 SD.

Expression and purification of St-Cas9. Cas9 was expressed in the *E. coli* DH10B strain. Cells were grown in LB broth medium supplemented with Ap (100 μg/ml) at 37°C to OD_{600} of ~0.5. Cas9 expression was induced with 0.2% (w/v) arabinose for 5 h. Harvested cells were disrupted by sonication and cell debris removed by centrifugation. The supernatant was loaded onto the Ni^{2+}-charged 5 ml HiTrap chelating HP column (GE Healthcare) and eluted with a linear gradient of increasing imidazole concentration. The fractions containing the Cas9-$(His)_6$ protein were pooled and subsequently loaded onto a heparin column, and eluted using a linear gradient of increasing NaCl concentration. The fractions containing Cas9 were pooled and dialysed against 10 mM Tris-HCl (pH 7.0), 300 mM KCl, 1 mM

EDTA, 1 mM DTT, 50% (v/v) glycerol and stored at -20°C. The homogeneity of protein preparations was estimated by SDS-PAGE and western blotting against $(His)_6$-tag with anti-His antibodies (Novogen). Concentrations of Cas9-$(His)_6$ protein were determined by measuring absorbance at 280 nm using an extinction coefficient of 128,390 $M^{-1}cm^{-1}$.[54]

RNA production. RNAs were either purchased as synthetic oligoribonucleotides (Metabion) or synthesized by in vitro transcription using the TranscriptAid T7 High Yield Transcription Kit (Thermo Fisher Scientific). For in vitro transcription, fragments containing a T7 promoter at the proximal end of the RNA coding sequence were PCR-generated using primers and templates listed in the **Table S3**. Resulting RNA fragments were purified using RNeasy MinElute Cleanup Kit (Qiagen). The tracrRNAs, dephosphorylated with FastAP phosphatase (Thermo Fisher Scientific) were radiolabeled with (γ-^{33}P)ATP

(Hartmann Analytic) using T4 polynucleotide kinase (Thermo Fisher Scientific). Sequences of RNA fragments used in this study are provided in the **Table S4**.

In vitro assembly of the Cas9-RNA complexes. A ternary Cas9-crRNA-tracrRNA complex was assembled by mixing 100 nM of (His)$_6$-tagged Cas9 protein with a pre-annealed 42 nt crRNA and 78 nt tracrRNA duplex at 1:1 molar ratio and incubated in a buffer containing 10 mM TRIS-HCl (pH 7.5 at 37°C), 100 mM NaCl and 1 mM DTT at 37°C for 1 h.

To reconstitute crRNA maturation and complex assembly pathway in vitro, the (His)$_6$-tagged Cas9 protein (100 nM) was mixed with 150 nt pre-crRNA (100 nM) and 105 nt tracrRNA (200 nM) transcripts at 1:1:2 molar ratio and pre-incubated in a buffer containing 10 mM TRIS-HCl (pH 7.5 at 37°C), 100 mM NaCl and 1 mM DTT at 37°C for 30 min followed by addition of *E. coli* RNase III (0.1–2 μM; Ambion) and MgCl$_2$ (10 mM) and further incubated for additional 30 min at 37°C.

Reactions with plasmid substrates. Reactions on pSP1 and pUC18 plasmids were initiated by mixing plasmid DNA with a ternary Cas9-crRNA-tracrRNA complex (1:1 v/v ratio) and conducted at 37°C. Final reaction mixture contained 3 nM (500 ng) plasmid, 50 nM of the effector complex, 10 mM TRIS-HCl (pH 7.5 at 37°C), 100 mM NaCl, 1 mM DTT and 10 mM MgCl$_2$ in a 100 μl reaction volume. Aliquots were removed at timed intervals and quenched with phenol/chloroform. The aqueous phase was mixed with 3 × loading dye solution (0.01% bromphenol blue and 75 mM EDTA in 50% v/v glycerol) and reaction products analyzed by agarose gel electrophoresis and ethidium bromide staining.

RNA strand annealing activity assay. The annealing activity of Cas9 was assayed in 10 μl reaction volumes. Reactions were initiated by mixing 5 nM 94 nt pre-crRNA and varying concentrations of Cas9 with 5 nM 105 nt tracrRNA radiolabeled with (γ-^{33}P)ATP (Hartmann Analytic) and 1 μg nonspecific 2 kb RNA transcript in 10 mM TRIS-HCl (pH 7.5 at 37°C), 100 mM NaCl,

1 mM DTT, 0.05 mg/ml BSA reaction buffer and conducted for 10 min at 37°C. In order to inactivate Cas9, disrupt protein-nucleic acid complexes and prevent spontaneous pre-crRNA annealing to labeled tracrRNA, reactions were terminated by mixing 5 μl reaction aliquots with 20 μl of solution containing 0.625 mg/ml Proteinase K (Thermo Fisher Scientific), 0.625% SDS, 12.5% glycerol and RNA trap (187.5 nM solution of unlabeled tracrRNA). Reaction mixtures were incubated for additional 5 min at 37°C and analyzed in non-denaturing 8% PAGE. Gels were dried and visualized by a FLA-5100 phosphor-imager (Fujifilm).

Northern blot analysis. Northern blot analysis was performed as described previously.[39] Cas9-bound RNA was isolated from Strep-Tactin-purified Cas9-RNA complex[39] using the miRNeasy Mini kit (Qiagen). The RNA was probed with a (γ-^{32}P)ATP (Hartmann Analytic) labeled 36 nt oligodeoxynucleotide (GG-322) (**Table S2**) complementary to tracrRNA or 42 nt oligodeoxynucleotide (GG-321) (**Table S2**) complementary to crRNA. The size of tracrRNA was estimated by comparison with ^{33}P-labeled Decade RNA marker (Ambion) and RNA transcripts of different lengths.

Supplemental Material

Supplemental material may be found here:
www.landesbioscience.com/journals/rnabiology/article/24203

References

1. Barrangou R, Fremaux C, Deveau H, Richards M, Boyaval P, Moineau S, et al. CRISPR provides acquired resistance against viruses in prokaryotes. Science 2007; 315:1709-12; PMID:17379808; http://dx.doi.org/10.1126/science.1138140.

2. Bolotin A, Quinquis B, Sorokin A, Ehrlich SD. Clustered regularly interspaced short palindrome repeats (CRISPRs) have spacers of extrachromosomal origin. Microbiology 2005; 151:2551-61; PMID:16079334; http://dx.doi.org/10.1099/mic.0.28048-0.

3. Mojica FJM, Díez-Villaseñor C, García-Martínez J, Soria E. Intervening sequences of regularly spaced prokaryotic repeats derive from foreign genetic elements. J Mol Evol 2005; 60:174-82; PMID:15791728; http://dx.doi.org/10.1007/s00239-004-0046-3.

4. Brouns SJJ, Jore MM, Lundgren M, Westra ER, Slijkhuis RJ, Snijders AP, et al. Small CRISPR RNAs guide antiviral defense in prokaryotes. Science 2008; 321:960-4; PMID:18703739; http://dx.doi.org/10.1126/science.1159689.

5. Makarova KS, Haft DH, Barrangou R, Brouns SJ, Charpentier E, Horvath P, et al. Evolution and classification of the CRISPR-Cas systems. Nat Rev Microbiol 2011; 9:467-77; PMID:21552286; http://dx.doi.org/10.1038/nrmicro2577.

6. Horvath P, Barrangou R. CRISPR/Cas, the immune system of bacteria and archaea. Science 2010; 327:167-70; PMID:20056882; http://dx.doi.org/10.1126/science.1179555.

7. Marraffini LA, Sontheimer EJ. CRISPR interference: RNA-directed adaptive immunity in bacteria and archaea. Nat Rev Genet 2010; 11:181-90; PMID:20125085; http://dx.doi.org/10.1038/nrg2749.

8. Deveau H, Garneau JE, Moineau S. CRISPR/Cas system and its role in phage-bacteria interactions. Annu Rev Microbiol 2010; 64:475-93; PMID:20528693; http://dx.doi.org/10.1146/annurev.micro.112408.134123.

9. Bhaya D, Davison M, Barrangou R. CRISPR-Cas systems in bacteria and archaea: versatile small RNAs for adaptive defense and regulation. Annu Rev Genet 2011; 45:273-97; PMID:22060043; http://dx.doi.org/10.1146/annurev-genet-110410-132430.

10. Terns MP, Terns RM. CRISPR-based adaptive immune systems. Curr Opin Microbiol 2011; 14:321-7; PMID:21531607; http://dx.doi.org/10.1016/j.mib.2011.03.005.

11. Wiedenheft B, Sternberg SH, Doudna JA. RNA-guided genetic silencing systems in bacteria and archaea. Nature 2012; 482:331-8; PMID:22337052; http://dx.doi.org/10.1038/nature10886.

12. Jore MM, Brouns SJJ, van der Oost J. RNA in defense: CRISPRs protect prokaryotes against mobile genetic elements. Cold Spring Harb Perspect Biol 2012; 4; PMID:21441598; http://dx.doi.org/10.1101/cshperspect.a003657.

13. Westra ER, Swarts DC, Staals RHJ, Jore MM, Brouns SJJ, van der Oost J. The CRISPRs, they are a-changin': how prokaryotes generate adaptive immunity. Annu Rev Genet 2012; 46:311-39; PMID:23145983; http://dx.doi.org/10.1146/annurev-genet-110711-155447.

14. van der Oost J, Jore MM, Westra ER, Lundgren M, Brouns SJJ. CRISPR-based adaptive and heritable immunity in prokaryotes. Trends Biochem Sci 2009; 34:401-7; PMID:19646880; http://dx.doi.org/10.1016/j.tibs.2009.05.002.

15. Babu M, Beloglazova N, Flick R, Graham C, Skarina T, Nocek B, et al. A dual function of the CRISPR-Cas system in bacterial antivirus immunity and DNA repair. Mol Microbiol 2011; 79:484-502; PMID:21219465; http://dx.doi.org/10.1111/j.1365-2958.2010.07465.x.

16. Yosef I, Goren MG, Qimron U. Proteins and DNA elements essential for the CRISPR adaptation process in *Escherichia coli*. Nucleic Acids Res 2012; 40:5569-76; PMID:22402487; http://dx.doi.org/10.1093/nar/gks216.

17. Datsenko KA, Pougach K, Tikhonov A, Wanner BL, Severinov K, Semenova E. Molecular memory of prior infections activates the CRISPR/Cas adaptive bacterial immunity system. Nat Commun 2012; 3:945; PMID:22781758; http://dx.doi.org/10.1038/ncomms1937.

18. Carte J, Pfister NT, Compton MM, Terns RM, Terns MP. Binding and cleavage of CRISPR RNA by Cas6. RNA 2010; 16:2181-8; PMID:20884784; http://dx.doi.org/10.1261/rna.2230110.

19. Hale CR, Zhao P, Olson S, Duff MO, Graveley BR, Wells L, et al. RNA-guided RNA cleavage by a CRISPR RNA-Cas protein complex. Cell 2009; 139:945-56; PMID:19945378; http://dx.doi.org/10.1016/j.cell.2009.07.040.

20. Pougach K, Semenova E, Bogdanova E, Datsenko KA, Djordjevic M, Wanner BL, et al. Transcription, processing and function of CRISPR cassettes in *Escherichia coli.* Mol Microbiol 2010; 77:1367-79; PMID:20624226; http://dx.doi.org/10.1111/j.1365-2958.2010.07265.x.

21. Djordjevic M, Djordjevic M, Severinov K. CRISPR transcript processing: a mechanism for generating a large number of small interfering RNAs. Biol Direct 2012; 7:24; PMID:22849651; http://dx.doi.org/10.1186/1745-6150-7-24.

22. Carte J, Wang R, Li H, Terns RM, Terns MP. Cas6 is an endoribonuclease that generates guide RNAs for invader defense in prokaryotes. Genes Dev 2008; 22:3489-96; PMID:19141480; http://dx.doi.org/10.1101/gad.1742908.

23. Sashital DG, Jinek M, Doudna JA. An RNA-induced conformational change required for CRISPR RNA cleavage by the endoribonuclease Cse3. Nat Struct Mol Biol 2011; 18:680-7; PMID:21572442; http://dx.doi.org/10.1038/nsmb.2043.

24. Gesner EM, Schellenberg MJ, Garside EL, George MM, Macmillan AM. Recognition and maturation of effector RNAs in a CRISPR interference pathway. Nat Struct Mol Biol 2011; 18:688-92; PMID:21572444; http://dx.doi.org/10.1038/nsmb.2042.

25. Richter H, Zoephel J, Schermuly J, Maticzka D, Backofen R, Randau L. Characterization of CRISPR RNA processing in *Clostridium thermocellum* and *Methanococcus maripaludis.* Nucleic Acids Res 2012; 40:9887-96; PMID:22879377; http://dx.doi.org/10.1093/nar/gks737.

26. Deltcheva E, Chylinski K, Sharma CM, Gonzales K, Chao Y, Pirzada ZA, et al. CRISPR RNA maturation by trans-encoded small RNA and host factor RNase III. Nature 2011; 471:602-7; PMID:21455174; http://dx.doi.org/10.1038/nature09886.

27. Garneau JE, Dupuis MÈ, Villion M, Romero DA, Barrangou R, Boyaval P, et al. The CRISPR/Cas bacterial immune system cleaves bacteriophage and plasmid DNA. Nature 2010; 468:67-71; PMID:21048762; http://dx.doi.org/10.1038/nature09523.

28. Sinkunas T, Gasiunas G, Fremaux C, Barrangou R, Horvath P, Siksnys V. Cas3 is a single-stranded DNA nuclease and ATP-dependent helicase in the CRISPR/Cas immune system. EMBO J 2011; 30:1335-42; PMID:21343909; http://dx.doi.org/10.1038/emboj.2011.41.

29. Beloglazova N, Petit P, Flick R, Brown G, Savchenko A, Yakunin AF. Structure and activity of the Cas3 HD nuclease MJ0384, an effector enzyme of the CRISPR interference. EMBO J 2011; 30:4616-27; PMID:22009198; http://dx.doi.org/10.1038/emboj.2011.377.

30. Mulepati S, Bailey S. Structural and biochemical analysis of nuclease domain of clustered regularly interspaced short palindromic repeat (CRISPR)-associated protein 3 (Cas3). J Biol Chem 2011; 286:31896-903; PMID:21775431; http://dx.doi.org/10.1074/jbc.M111.270017.

31. Westra ER, van Erp PBG, Künne T, Wong SP, Staals RH, Seegers CL, et al. CRISPR immunity relies on the consecutive binding and degradation of negatively supercoiled invader DNA by Cascade and Cas3. Mol Cell 2012; 46:595-605; PMID:22521689; http://dx.doi.org/10.1016/j.molcel.2012.03.018.

32. Sinkunas T, Gasiunas G, Waghmare SP, Dickman MJ, Barrangou R, Horvath P, et al. *In vitro* reconstitution of Cascade-mediated CRISPR immunity in *Streptococcus thermophilus.* EMBO J 2013; 32:385-94; PMID:23334296; http://dx.doi.org/10.1038/emboj.2012.352.

33. Hale CR, Majumdar S, Elmore J, Pfister N, Compton M, Olson S, et al. Essential features and rational design of CRISPR RNAs that function with the Cas RAMP module complex to cleave RNAs. Mol Cell 2012; 45:292-302; PMID:22227116; http://dx.doi.org/10.1016/j.molcel.2011.10.023.

34. Zhang J, Rouillon C, Kerou M, Reeks J, Brugger K, Graham S, et al. Structure and mechanism of the CMR complex for CRISPR-mediated antiviral immunity. Mol Cell 2012; 45:303-13; PMID:22227115; http://dx.doi.org/10.1016/j.molcel.2011.12.013.

35. Marraffini LA, Sontheimer EJ. CRISPR interference limits horizontal gene transfer in staphylococci by targeting DNA. Science 2008; 322:1843-5; PMID:19095942; http://dx.doi.org/10.1126/science.1165771.

36. Jinek M, Chylinski K, Fonfara I, Hauer M, Doudna JA, Charpentier E. A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. Science 2012; 337:816-21; PMID:22745249; http://dx.doi.org/10.1126/science.1225829.

37. Sapranauskas R, Gasiunas G, Fremaux C, Barrangou R, Horvath P, Siksnys V. The *Streptococcus thermophilus* CRISPR/Cas system provides immunity in *Escherichia coli.* Nucleic Acids Res 2011; 39:9275-82; PMID:21813460; http://dx.doi.org/10.1093/nar/gkr606.

38. Magadán AH, Dupuis MÈ, Villion M, Moineau S. Cleavage of phage DNA by the *Streptococcus thermophilus* CRISPR3-Cas system. PLoS One 2012; 7:e40913; PMID:22911717; http://dx.doi.org/10.1371/journal.pone.0040913.

39. Gasiunas G, Barrangou R, Horvath P, Siksnys V. Cas9-crRNA ribonucleoprotein complex mediates specific DNA cleavage for adaptive immunity in bacteria. Proc Natl Acad Sci USA 2012; 109:E2579-86; PMID:22949671; http://dx.doi.org/10.1073/pnas.1208507109.

40. Deveau H, Barrangou R, Garneau JE, Labonté J, Fremaux C, Boyaval P, et al. Phage response to CRISPR-encoded resistance in *Streptococcus thermophilus.* J Bacteriol 2008; 190:1390-400; PMID:18065545; http://dx.doi.org/10.1128/JB.01412-07.

41. Takiff HE, Chen SM, Court DL. Genetic analysis of the *rnc* operon of *Escherichia coli.* J Bacteriol 1989; 171:2581-90; PMID:2540151.

42. Sternberg SH, Haurwitz RE, Doudna JA. Mechanism of substrate selection by a highly specific CRISPR endoribonuclease. RNA 2012; 18:661-72; PMID:22345129; http://dx.doi.org/10.1261/rna.030882.111.

43. Jore MM, Lundgren M, van Duijn E, Bultema JB, Westra ER, Waghmare SP, et al. Structural basis for CRISPR RNA-guided DNA recognition by Cascade. Nat Struct Mol Biol 2011; 18:529-36; PMID:21460843; http://dx.doi.org/10.1038/nsmb.2019.

44. Zhang J, Kasciukovic T, White MF. The CRISPR Associated Protein Cas4 Is a 5' to 3' DNA Exonuclease with an Iron-Sulfur Cluster. Marinus MG, ed. *PLoS ONE.* 2012; 7(10):e47232.

45. Nam KH, Kurinov I, Ke A. Crystal structure of clustered regularly interspaced short palindromic repeats (CRISPR)-associated Csn2 protein revealed Ca2+-dependent double-stranded DNA binding activity. J Biol Chem 2011; 286:30759-68; PMID:21697083; http://dx.doi.org/10.1074/jbc.M111.256263.

46. Ellinger P, Arslan Z, Wurm R, Tschapek B, MacKenzie C, Pfeffer K, et al. The crystal structure of the CRISPR-associated protein Csn2 from *Streptococcus agalactiae.* J Struct Biol 2012; 178:350-62; PMID:22531577; http://dx.doi.org/10.1016/j.jsb.2012.04.006.

47. Lee KH, Lee SG, Eun Lee K, Jeon H, Robinson H, Oh BH. Identification, structural, and biochemical characterization of a group of large Csn2 proteins involved in CRISPR-mediated bacterial immunity. Proteins 2012; 80:2573-82; PMID:22753072; http://dx.doi.org/10.1002/prot.24138.

48. Koo Y, Jung DK, Bae E. Crystal structure of *Streptococcus pyogenes* Csn2 reveals calcium-dependent conformational changes in its tertiary and quaternary structure. PLoS One 2012; 7:e33401; PMID:22479393; http://dx.doi.org/10.1371/journal.pone.0033401.

49. Wiedenheft B, van Duijn E, Bultema JB, Waghmare SP, Zhou K, Barendregt A, et al. RNA-guided complex from a bacterial immune system enhances target recognition through seed sequence interactions. Proc Natl Acad Sci USA 2011; 108:10092-7; PMID:21536913; http://dx.doi.org/10.1073/pnas.1102716108.

50. Wiedenheft B, Lander GC, Zhou K, Jore MM, Brouns SJ, van der Oost J, et al. Structures of the RNA-guided surveillance complex from a bacterial immune system. Nature 2011; 477:486-9; PMID:21938068; http://dx.doi.org/10.1038/nature10402.

51. Lintner NG, Kerou M, Brumfield SK, Graham S, Liu H, Naismith JH, et al. Structural and functional characterization of an archaeal clustered regularly interspaced short palindromic repeat (CRISPR)-associated complex for antiviral defense (CASCADE). J Biol Chem 2011; 286:21643-56; PMID:21507944; http://dx.doi.org/10.1074/jbc.M111.238485.

52. Cong L, Ran FA, Cox D, Lin S, Barretto R, Habib N, et al. Multiplex genome engineering using CRISPR/Cas systems. Science 2013; 339:819-23; PMID:23287718; http://dx.doi.org/10.1126/science.1231143.

53. Mali P, Yang L, Esvelt KM, Aach J, Guell M, DiCarlo JE, et al. RNA-guided human genome engineering via Cas9. Science 2013; 339:823-6; PMID:23287722; http://dx.doi.org/10.1126/science.1232033.

54. Gill SC, von Hippel PH. Calculation of protein extinction coefficients from amino acid sequence data. Anal Biochem 1989; 182:319-26; PMID:2610349; http://dx.doi.org/10.1016/0003-2697(89)90602-7.

55. Hofacker IL, Fontana W, Stadler PF, Bonhoeffer LS, Tacker M, Schuster P. Fast folding and comparison of RNA secondary structures. Monatsh Chem 1994; 125:167-88; http://dx.doi.org/10.1007/BF00818163.