CellPress

Molecular Cell
Article

Programmable RNA Shredding by the Type III-A CRISPR-Cas System of *Streptococcus thermophilus*

Gintautas Tamulaitis,[1] Migle Kazlauskiene,[1] Elena Manakova,[1] Česlovas Venclovas,[2] Alison O. Nwokeoji,[3] Mark J. Dickman,[3] Philippe Horvath,[4] and Virginijus Siksnys[1,*]
[1]Department of Protein–DNA Interactions
[2]Department of Bioinformatics
Institute of Biotechnology, Vilnius University, Graiciuno 8, Vilnius 02241, Lithuania
[3]Department of Chemical and Biological Engineering, ChELSI Institute, University of Sheffield, Mappin Street, Sheffield S1 3JD, UK
[4]DuPont Nutrition and Health, BP10, Dangé-Saint-Romain 86220, France
*Correspondence: siksnys@ibt.lt
http://dx.doi.org/10.1016/j.molcel.2014.09.027

SUMMARY

Immunity against viruses and plasmids provided by CRISPR-Cas systems relies on a ribonucleoprotein effector complex that triggers the degradation of invasive nucleic acids (NA). Effector complexes of type I (Cascade) and II (Cas9-dual RNA) target foreign DNA. Intriguingly, the genetic evidence suggests that the type III-A Csm complex targets DNA, whereas biochemical data show that the type III-B Cmr complex cleaves RNA. Here we aimed to investigate NA specificity and mechanism of CRISPR interference for the *Streptococcus thermophilus* Csm (III-A) complex (StCsm). When expressed in *Escherichia coli*, two complexes of different stoichiometry copurified with 40 and 72 nt crRNA species, respectively. Both complexes targeted RNA and generated multiple cuts at 6 nt intervals. The Csm3 protein, present in multiple copies in both Csm complexes, acts as endoribonuclease. In the heterologous *E. coli* host, StCsm restricts MS2 RNA phage in a Csm3 nuclease-dependent manner. Thus, our results demonstrate that the type III-A StCsm complex guided by crRNA targets RNA and not DNA.

INTRODUCTION

Clustered regularly interspaced short palindromic repeats (CRISPR) together with Cas (CRISPR-associated) proteins provide RNA-mediated adaptive immunity against viruses and plasmids in bacteria and archaea (Terns and Terns, 2014). Immunity is acquired through the integration of invader-derived nucleic acid (NA) sequences as "spacers" into the CRISPR locus of the host. CRISPR arrays are further transcribed and processed into small interfering CRISPR RNAs (crRNAs) that together with Cas proteins assemble into a ribonucleoprotein (RNP) complex which, using crRNA as a guide, locates and degrades the target NA. CRISPR-Cas systems have been categorized into three major types (I–III) that differ by the structural organization of RNPs and NA specificity (Makarova et al., 2011b).

Type I and II systems provide immunity against invading DNA. In type I-E systems, crRNAs are incorporated into a multisubunit RNP complex called Cascade (CRISPR-associated complex for antiviral defense) that binds to the matching invasive DNA and triggers degradation by the Cas3 nuclease/helicase (Brouns et al., 2008; Sinkunas et al., 2013; Westra et al., 2012). In type II systems, CRISPR-mediated immunity solely relies on the Cas9 protein. It binds a dual RNA into the RNP effector complex, which then specifically cuts the matching target DNA, introducing a double-strand break (Gasiunas et al., 2012; Jinek et al., 2012). In type I and II CRISPR-Cas systems, the target site binding and cleavage require a short nucleotide sequence (protospacer-adjacent motif, or PAM) near the target (Mojica et al., 2009). Target DNA strand separation, necessary for the crRNA binding, is initiated at PAM and propagates in a directional manner through the protospacer sequence to yield the R-loop intermediate, one strand of which is engaged into the heteroduplex with crRNA, while the other strand is displaced into solution (Sternberg et al., 2014; Szczelkun et al., 2014). Thus, despite differences in their architecture, type I and II RNP complexes share three major features: (1) they act on the invasive double-stranded DNA (dsDNA), e.g., viral DNA or plasmids; (2) they require the presence of a PAM sequence near the target site; and (3) they generate an R-loop as a reaction intermediate.

Intriguingly, type III CRISPR-Cas systems are believed to target either DNA (type III-A) or RNA (type III-B) (Makarova et al., 2011b). In the III-B systems, Cas RAMP proteins (Cmr) and crRNA assemble into a multisubunit RNP complex. Using crRNA as a guide, this complex in vitro binds single-stranded RNA (ssRNA) in a PAM-independent manner and triggers the degradation of target RNA (Hale et al., 2009; Staals et al., 2013; Zhang et al., 2012). The Cmr effector complex is composed of six Cmr proteins (Cmr1, Cas10, Cmr3-6) that are important for target RNA cleavage; however, roles of the individual Cmr proteins and the ribonuclease (RNase) component have yet to be identified. Cmr1, Cmr3, Cmr4, and Cmr6 are predicted RNA-binding proteins that share a ferredoxin-like fold and

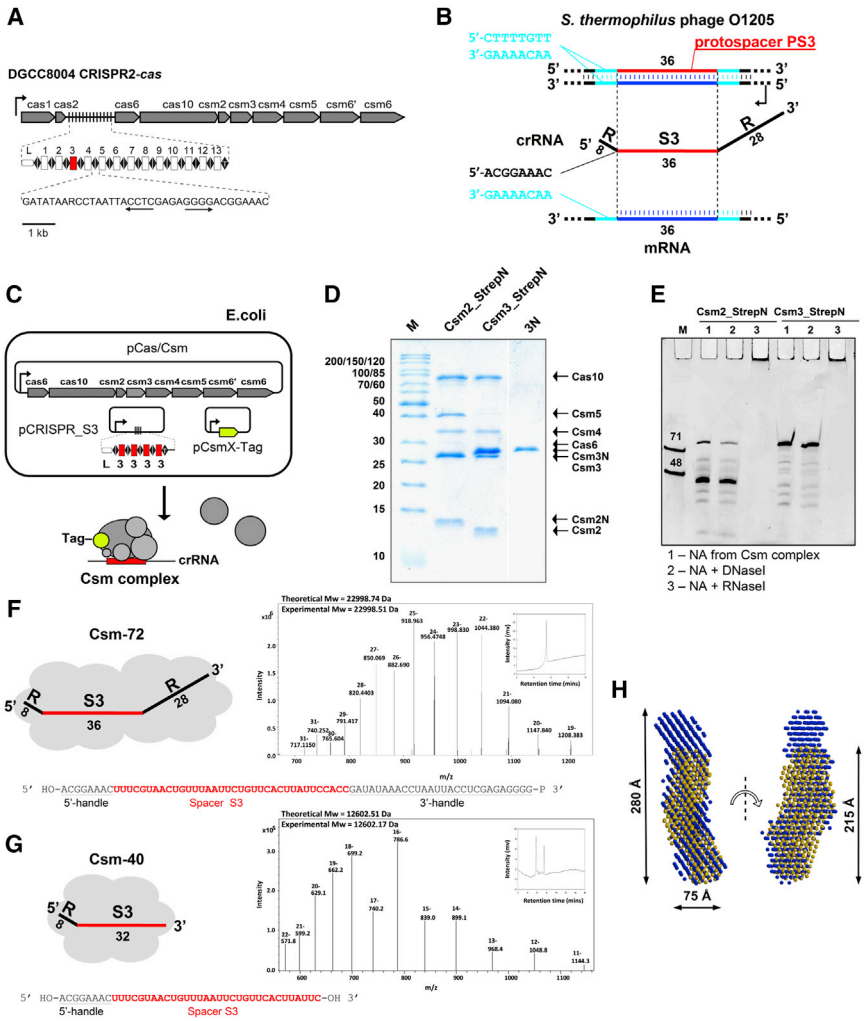

Figure 1. Cloning, Isolation, and Characterization of the Type III-A Csm Complex of *S. thermophilus* DGCC8004

(A) Schematic organization of the type III-A CRISPR2-*cas* locus (see also Figure S1). Repeats and spacers are indicated by diamonds and rectangles, respectively; T, terminal repeat; L, leader sequence; and the arrow indicates the promoter.

(B) Protospacer PS3 and the 5′-flanking sequence in the *S. thermophilus* phage O1205 genome.

(C) Strategy for expression and isolation of the StCsm complex. Four copies of the spacer S3 have been engineered into the pCRISPR_S3 plasmid to increase the yield of the Csm-crRNA complex.

(D) Coomassie blue-stained SDS-PAGE of Strep-tagged Csm2 and Csm3 pull-downs. 3N, Csm3_StrepN protein; M, protein mass marker.

(E) Denaturing PAGE analysis of NA copurifying with the Csm2_StrepN and Csm3_StrepN complexes. M, synthetic DNA marker.

(F and G) Characterization of crRNA in the isolated StCsm complexes. Cartoon models illustrate crRNA that copurifies with StCsm-72 and StCsm-40 complexes. Composition of the crRNA was determined using liquid chromatography ESI MS analysis (see also Figure S2). Ion pair reverse-phase HPLC analysis and ESI-MS spectra of Ion pair reverse-phase HPLC-purified crRNA from StCsm-40 and StCsm-72 are presented.

(H) Superimposed averaged dummy atom models obtained from SAXS data of StCsm-40 (yellow beads) and StCsm-72 (blue beads) (see also Figure S3).

RNA-recognition motif (RRM) identified in RNA-binding proteins (Terns and Terns, 2014).

The *cas* genes encoding the RNA-targeting type III-B (Cmr) and DNA-targeting type III-A (Csm) effector complexes share a partial synergy (Makarova et al., 2011a). In *Staphylococcus epidermidis*, the Csm complex (SeCsm) is composed of Cas10, Csm2, Csm3, Csm4, and Csm5 proteins; however, the function of individual Csm proteins is unknown. The evidence that type III-A systems target DNA remains indirect and relies on the experimental observation that type III-A RNP complex from *S. epidermidis* (SeCsm) limits plasmid conjugation and transformation in vivo, but the DNA degradation has not been demonstrated directly (Marraffini and Sontheimer, 2008, 2010). The Csm complex from the archaeon *Sulfolobus solfataricus* (SsCsm) binds dsDNA; however, it shows no crRNA-dependent nuclease activity in vitro (Rouillon et al., 2013). Thus, although the RNA cleavage activity of the Cmr complex has been characterized in vitro, the DNA degradation activity of the type III-A Csm complex has yet to be demonstrated. The Csm complex so far remains the only CRISPR-Cas effector complex, for which the function is not yet reconstituted in vitro. Here, we aimed to establish the composition and mechanism of the Csm complex for type III-A system *Streptococcus thermophilus* (St).

RESULTS

Type III-A CRISPR-Cas Loci in *S. thermophilus*

S. thermophilus strain DGCC8004 that was selected for experimental characterization carries 13 spacers in its type III-A CRISPR2 array (Figure 1A; Figure S1 available online). This strain also contains a type II CRISPR1 system that is ubiquitous in the *S. thermophilus* species. In the CRISPR2 locus of DGCC8004, the 36 nt repeat sequences, that are partially palindromic, are conserved with the exception of the two terminal repeats (Figure 1A). An A+T rich 100-base pair leader sequence is located upstream of the CRISPR2 array.

DGCC8004 CRISPR2 (type III-A) spacers range in size between 34 and 43 nucleotides (nt), but 36 nt spacers are the most abundant. In total, 38 unique spacers were identified among CRISPR2-positive *S. thermophilus* strains and a majority (20 of 38) of these spacer sequences have matches (protospacers) in *S. thermophilus* DNA phage sequences, although

phage interference for the *S. thermophilus* CRISPR2 locus has not yet been demonstrated. Analysis of the sequences located immediately upstream and downstream of these protospacers failed to identify any consensus sequence as a putative PAM, either due to the relatively small number of protospacers or targeting of RNA that is often PAM-independent (Hale et al., 2009). In DGCC8004, although no CRISPR2 spacer gives perfect identity with currently known sequences, six of 13 spacers (S3, S4, S6, S8, S12, and S13) show strong sequence similarity with *S. thermophilus* DNA phages (at least 94% identity over at least 80% of spacer length). Interestingly, all phage-matching protospacers appear to have been selected from the template strand. For instance, the 36 nt spacer S3 matches 34 nt of a protospacer in the *S. thermophilus* phage O1205 genome (Figure 1B). A corresponding crRNA would match the template DNA strand of the protospacer S3 and would pair with the target sequence on the coding strand of phage DNA or the respective mRNA sequence. If crRNA processing in the *S. thermophilus* type III-A locus is similar to that in *S. epidermidis* (Hatoum-Aslan et al., 2011, 2013, 2014), the resulting crRNA 5′-handle in the mature crRNA will be noncomplementary to the protospacer S3 3′-flank in the phage DNA coding strand or mRNA (Figure 1B). In the *S. epidermidis* type III-A system, which limits the spreading of plasmid DNA, the crRNA/target DNA noncomplementarity outside of the spacer sequence plays a key role in silencing of invading DNA and self versus nonself DNA discrimination (Marraffini and Sontheimer, 2010). Taking these elements into consideration, crRNA encoded by the spacer S3 was selected as the guide, and a complementary protospacer sequence as the NA target (DNA or RNA) (Figure 1B).

Cloning, Expression, and Isolation of the *S. thermophilus* DGCC8004 Type III-A Effector Complex

To isolate the type III-A RNP effector complex (StCsm) of the DGCC8004, we split the CRISPR2 locus into the three fragments and cloned them into three compatible vectors (Figure 1C). Plasmid pCas/Csm contained a cassette including all the *cas/csm* genes (except *cas1* and *cas2*), whereas plasmid pCRISPR_S3 carried four identical tandem copies of the repeat-spacer S3 unit flanked by the leader sequence and the terminal repeat. Plasmids pCsm2-Tag or pCsm3-Tag carried a StrepII-tagged variant of *csm2* or *csm3* genes, respectively. Next, all three plasmids were coexpressed in *Escherichia coli* BL21(DE3), and tagged Csm2 or Csm3 proteins were isolated by subsequent Strep-chelating affinity and size exclusion chromatography.

Strep-tagged Csm2 or Csm3 proteins pulled down from *E. coli* lysates copurified with other Csm/Cas proteins, suggesting the presence of a Csm complex (Figure 1D). Csm complexes isolated via N terminus Strep-tagged Csm2 (Csm2_StrepN) and the N terminus Strep-tagged Csm3 proteins (Csm3_StrepN) were subjected to further characterization. SDS-PAGE of these complexes revealed six bands that matched the individual Cas proteins Cas6, Cas10, Csm2, Csm3, Csm4, and Csm5 (Figure 1D). The identity of proteins in these Csm complexes was confirmed by mass spectrometry (MS) analysis (Tables S1 and S2).

We next examined the Csm complexes for the presence of NA using basic phenol-chloroform extraction followed by RNase I or DNase I digestion. Denaturing PAGE analysis revealed that ~70 nt and ~40 nt RNA molecules copurified with the Csm3_StrepN and Csm2_StrepN pulled-down Csm complexes, respectively (Figure 1E). The complex isolated via Csm2_StrepN subunit also contained ~10% of the ~70 nt RNA. When subjected to RNase I protection assay, the RNA in the complexes showed no visible degradation, indicating that the RNA is tightly bound and protected along its entire length (data not shown).

Characterization of the crRNA

We used denaturing RNA chromatography in conjunction with electrospray ionization mass spectrometry (ESI-MS) to analyze the crRNA sequence and determine the chemical nature of the 5′ and 3′ termini of crRNAs copurified with both Csm complexes. Denaturing ion pair reverse-phase chromatography was used to rapidly purify the crRNA directly from the Csm complexes. The crRNA isolated from the Csm3_StrepN pull-down complex revealed a single crRNA with a retention time consistent with an approximate length of 70 nt (Figure 1F). The crRNA isolated from Csm2_StrepN pull-down complex revealed the presence of an additional crRNA, with a retention time consistent with an approximate length of 40 nt (Figure 1G). Purified crRNAs were further analyzed using ESI-MS to obtain the accurate intact masses. A molecular weight of 22 998.5 Da was obtained for RNA isolated from Csm3 and 12 602.2 Da for RNA isolated from Csm2 pull-downs, respectively. Csm2 pull-down also contained a minor component, with a molecular weight of 12 907.3 Da (data not shown). In addition, ESI MS/MS was also used to analyze the oligoribonucleotide fragments generated from RNase A/T1 digestion of the crRNAs (Figure S2). In conjunction with the intact mass analysis, these results revealed a 72 nt crRNA in the complex isolated via Csm3 (further termed Csm-72 according to the length of crRNA) and a 40 nt crRNA in the complex isolated via Csm2 (further termed Csm-40 complex). The MS analysis of the 72 nt crRNA is consistent with the pre-crRNA cleavage at the base of the putative hairpin to yield a 8 nt 5′-handle, a 36 nt spacer and a 28 nt 3′-handle with 5′-OH and 3′-P, and could represent unmature crRNA intermediate (Figure 1F) similar to that of type III-A and III-B CRISPR-Cas systems (Hale et al., 2009; Hatoum-Aslan et al., 2013). Further verification of the 3′-P termini was obtained upon acid treatment of the 72 nt crRNA where no change in mass was observed using ESI-MS. Likewise, the MS analysis of the 40 nt crRNA in the Csm-40 complex revealed an 8 nt 5′-handle and a 32 nt spacer with 5′-OH and 3′-OH that would correspond to the mature crRNA (Figure 1G). The difference in the chemical nature of the 3′-end between intermediate and mature crRNAs suggests that primary processing and final maturation are achieved by distinct catalytic mechanisms as proposed by Hatoum-Aslan for the *S. epidermidis* model system (Hatoum-Aslan et al., 2011).

Composition and Shape of the Csm Complex

Evaluation of the complex composition by densitometric analysis of the SDS gels suggests the $Cas10_1$:$Csm2_6$:$Csm3_{10}$:$Csm4_1$:$Csm5_{0.14}$ stoichiometry for Csm-72, and the $Cas6_{0.10}$:$Cas10_1$:$Csm2_3$:$Csm3_5$:$Csm4_1$:$Csm5_1$ stoichiometry for Csm-40.

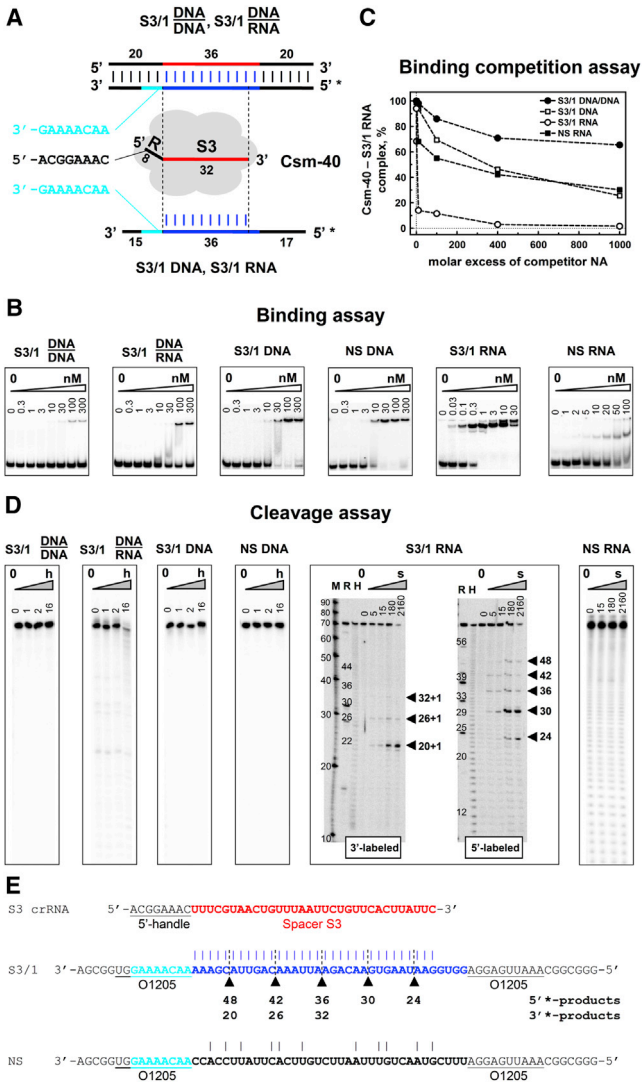

Figure 2. Nucleic Acid Binding and Cleavage by the Type III-A Csm Complex of *S. thermophilus*

(A) Schematic representation of DNA and RNA substrates used for in vitro binding and cleavage assays. NAs were 5′ or 3′ end-labeled with ^{32}P (indicated as *).

(B) EMSA analysis of DNA or RNA binding by StCsm-40. NS, nonspecific RNA.

(C) Binding competition assay: 0.5 nM of ^{32}P-labeled S3/1 RNA was mixed with increasing amounts of unlabeled competitor NAs and 0.3 nM StCsm-40 and analyzed by EMSA.

(D) StCsm-40 cleavage assay. Gel-purified DNA or RNA were used as substrates in the NA cleavage assay. Triangles with corresponding numbers indicate cleavage product length. M, RNA Decade marker; R, RNase A digest marker; H, alkaline hydrolysis marker.

(E) RNA cleavage products mapped on the S3/1 RNA substrate sequence. Triangles and dashed lines indicate cleavage positions. Short vertical lines above the sequence indicate nucleotides complementary to crRNA. crRNA (StCsm-40) sequence is depicted above the matching substrate fragments. See also Figure S4.

Fraction numbers for Cas6 and Csm5 proteins are presumably due to the weak transient interactions of these proteins in the respective complexes. Protein subunits that are involved in pre-crRNA processing, e.g., Cas6, would not necessarily occur in stoichiometric amounts in the purified effector complex.

We also performed small-angle X-ray scattering (SAXS) measurements to characterize the molecular mass/shape of both Csm-40 and Csm-72 effector complexes in solution. Molecular weight values obtained using SAXS are in agreement both with DLS and gel-filtration data (Table S3). Taken together, these data are consistent with the stoichiometry $Cas10_1$:$Csm2_6$:$Csm3_{10}$:$Csm4_1$:$crRNA_1$ (calculated molecular weight 486.2 kDa including the 72 nt crRNA) for Csm-72 and $Cas10_1$:$Csm2_3$:$Csm3_5$:$Csm4_1$:$Csm5_1$:$crRNA_1$ (calculated molecular weight 344.8 kDa including 40 nt crRNA) for Csm-40.

SAXS measurements revealed that the Csm-40 complex in solution has elongated and slightly twisted shape. The maximal interatomic distance (D_{max}) of the complex estimated from SAXS data is 215 Å, whereas its diameter is 75–80 Å (Table S4). The shape of this effector complex (Figure 1H) is very similar to the electron microscope structure of Cmr complexes from *Thermus thermophilus* (Staals et al., 2013), *Pyrococcus furiosus* (Spilman et al., 2013), and Cascade from *E. coli* (Wiedenheft et al., 2011) (Figure S3E). The Csm-72 complex with D_{max} of 280 Å (Table S4) is significantly more elongated than the Csm-40 complex (Figure 1H). The lowest normalized spatial discrepancy was obtained for the end-to-end superimposition of the Csm-40 and Csm-72 models (Figure 1H).

Nucleic Acid Specificity of the Type III-A StCsm Complex

In the CRISPR2 locus of DGCC8004, 34 of 36 nt of the spacer S3 match a sequence present in the genome of *S. thermophilus* phage O1205. Therefore, to probe the functional activity of the Csm-40 complex, we first designed DNA and RNA substrates that are fully complementary to the 32 nt crRNA encoded by spacer S3 and that carry phage O1205-flanking sequence. These flanking sequences lack complementarity to the 8 nt 5′-handle of the crRNA identified in the Csm-40/Csm-72 complexes (Figure 2A and Table S5). For binding analysis, DNA or RNA substrates were 5′-end radioactively labeled and the Csm-40 complex binding was evaluated by an electrophoretic mobility shift assay (EMSA) in the absence of any divalent metal (Me^{2+}) ions. Csm-40 showed weak affinity for oligoduplex S3/1 DNA/DNA and DNA/RNA substrates since binding was observed only at high (100–300 nM) complex concentrations. Single-stranded S3/1 DNA (ssDNA) was bound to Csm-40 with an intermediate affinity ($K_d \approx 30$ nM), whereas a single-stranded S3/1 RNA (ssRNA) showed high-affinity binding ($K_d \approx 0.3$ nM) (Figure 2B). Binding competition experiments with various nucleic acids further supported the single-stranded RNA specificity for the Csm-40 complex (Figure 2C). Cleavage data correlated with the binding affinity: S3/1 DNA/DNA, DNA/RNA, and ssDNA are refractory to cleavage, whereas S3/1 ssRNA complementary to the crRNA is cut by Csm-40 in the presence of Mg^{2+} ions (Figure 2D). RNase activity of Csm-40 complex requires Mg^{2+} or other divalent metal ions (Mn^{2+}, Ca^{2+}, Zn^{2+}, Ni^{2+} or Cu^{2+}) and is inhibited by EDTA (Figure S4E).

Notably, Csm-40 cuts the S3/1 RNA target at five sites regularly spaced by 6 nt intervals to produce 48, 42, 36, 30, and 24 nt products, respectively (Figures 2D and 2E). The sequence complementarity between the crRNA in the complex and the

RNA target is a key prerequisite for the cleavage: a nonspecific RNA (Figure 2E, bottom) was resistant to Csm-40. The Csm-40 cleavage pattern of the 3′-labeled S3/1 RNA substrate differs from that of the 5′-labeled variant. While the 5′-labeled substrate cleavage produces 48, 42, 36, 30, and 24 nt products, short degradation products of 21, 27, and 33 nt (1 nt shift is due to an additional nucleotide added during the 3′-labeling) are visible on the gel (Figures 2D and 2E). Taken together, cleavage data for the 5′ and 3′ end-labeled RNA substrates suggest that Csm-40 cuts the RNA molecule initially at one of its 3′ end internal sites and endonucleolytic degradation is further extended toward the 5′ end with 6 nt increments.

The Csm-72 complex carrying a 72 nt crRNA (8 nt 5′-handle plus 36 nt of the spacer S3 and 28 nt of the 3′-handle, Figure S4A) showed ~30-fold weaker binding affinity ($K_d \approx 10$ nM) to S3/1 RNA compared to the Csm-40 (Figure S4B). Nevertheless, similar to the Csm-40 complex, in the presence of Mg^{2+} ions, Csm-72 cleaved S3/1 RNA, albeit at a decreased rate, which may correlate with its weaker binding affinity (Figure S4C). The 5′- and 3′-labeled S3/1 RNA cleavage pattern is identical to that of Csm-40 (Figures S4C and S4D and data not shown). Like the Csm-40 complex, Csm-72 showed no cleavage of S3/1 ssDNA, DNA/DNA, or DNA/RNA substrates (data not shown). The heterogeneous Csm complex isolated from the *E. coli* host carrying the wild-type CRISPR array containing 13 spacers produces RNA cleavage products identical to those of the homogenous StCsm (Figure S4F). Taken together, these data unambiguously demonstrate that Csm-40 and Csm-72 complexes in vitro target RNA but not DNA, and cut RNA at multiple sites regularly spaced by 6 nt intervals.

Reprogramming of the StCsm Complex

To demonstrate that the type III-A StCsm complex can be reprogrammed to cut a desired RNA sequence in vitro, we designed and isolated Csm complexes loaded with crRNA (+Tc) or crRNA (−Tc) targeting, respectively, the 68 nt sense (+) and antisense (−) mRNA fragments obtained by in vitro transcription of the tetracycline (Tc) resistance gene in the pBR322 plasmid (nt 851–886) (Figure S5A and Table S5). Both Csm-40 and Csm-72 complexes guided by the crRNA (+Tc) sliced the complementary sense RNA fragment but not the antisense RNA sequence (Figure S5B). In contrast, Csm-40 and Csm-72 complexes guided by the crRNA (−Tc) cleaved antisense RNA but not the sense Tc mRNA fragment (Figure S5B). In both cases, target RNA was cleaved at multiple sites regularly spaced by 6 nt intervals (Figure S5C).

Target RNA Determinants for Cleavage by crRNA-Guided Csm Complex

We examined further whether the nucleotide context downstream or upstream of the protospacer sequence modulates RNA cleavage by the Csm complexes. To this end, we designed the S3/2 RNA substrate in which the flanking regions originating from O1205 phage DNA in the S3/1 substrate are replaced by different nucleotide stretches that are noncomplementary to the 5′-handle of crRNA in the Csm-40 and Csm-72 complexes, and to the 3′-handle in the Csm-72 complex. RNA binding and cleavage data showed that despite differences in the nucleotide context of flanking sequences in the S3/1 and S3/2 substrates, cleavage patterns for the Csm-40 and Csm-72 complexes are nearly identical, except for an extra 18 nt product for the Csm-72 (Figure S4C).

Next, we addressed the question whether the base pairing between the flanking sequences of the RNA target and 5′- and 3′-handles of crRNA in the Csm-40 and Csm-72 complexes affect either the cleavage efficiency or pattern. We designed S3/3, S3/4, and S3/5 RNA substrates that contain flanking sequences complementary to the 5′-handle (40 or 72 nt crRNA), 3′-handle (72 nt crRNA) or both 5′- and 3′-handles in 72 nt scrRNA, respectively (Figure S6A and Table S5). The cleavage analysis revealed that base-pairing between the 8 nt 5′-handle of crRNA and the 3′-flanking sequence of the target had no effect on the cleavage pattern of the Csm-40 and Csm-72 complexes. Indeed, the S3/3 substrate is cleaved with the same 6 nt step by Csm-40, suggesting that the noncomplementarity of the flanking sequences is not a necessary prerequisite for cleavage by the Csm complex (Figure S6). Surprisingly, for the Csm-72 complex extension of the base-pairing between the 3′-handle of the 72 nt crRNA and the protospacer 5′-flanking sequence in S3/5 RNA substrate results in target RNA cleavage outside the protospacer, yielding 12 and 6 nt cleavage products (Figure 3). Moreover, the S3/6 substrate, which has extended complementarity between crRNA 3′-handle and 5′-flanking sequence was cleaved at multiple positions along the full length of RNA duplex, except for the region complementary to the crRNA 5′-handle (Figure 3). The cleavage at 18 and 12 nt outside the protospacer was also detected for the Csm-40 complex on S3/4 and S3/5 RNA substrates (Figure S6). The 40 nt crRNA present in the Csm-40 complex lacks the 3′-handle and therefore cannot form an RNA duplex with the 5′-flanking sequence in the S3/5 and S3/6 RNA substrates. However, the Csm-40 complex preparation still contains ~10% of unmatured 72 nt crRNA, and this heterogeneity results in the extra cleavage outside the protospacer (Figure S6C).

To interrogate the importance of base pairing within the protospacer region for target RNA cleavage, we designed a set of RNA substrates harboring two adjacent nucleotide mutations in the spacer region (substrates S3/7, S3/8, and S3/9, see Figure S6A and Table S5). Two nucleotide mismatches in these substrates did not compromise RNA cleavage by the Csm-40 (Figure S6) and Csm-72 complexes (data not shown), suggesting that the StCsm complex tolerates at least two contiguous mismatches in the protospacer region homologous to the crRNA.

To explore whether 3′ or 5′ ends of the target RNA are important for cleavage by the Csm-40 complex, we designed a set of truncated RNA substrates. In S3/10, S3/14, and S3/12 RNA substrates, unpaired flaps at the 3′, 5′, or both ends of the target RNA were truncated, whereas in S3/11 and S3/13 substrates, the truncations extend into the region complementary to crRNA (Figure 4A). Binding affinity for most of the substrates truncated outside the protospacer was not compromised (Figure 4B) and target RNA cleavage occurred at multiple sites spaced by 6 nt intervals at conserved protospacer positions (Figure 4C). Truncations extending into the protospacer region (S3/11 and S3/13) did not interfere with the target RNA cleavage albeit slightly reduced binding and cleavage rate presumably due to reduction of duplex stability. Importantly, for all

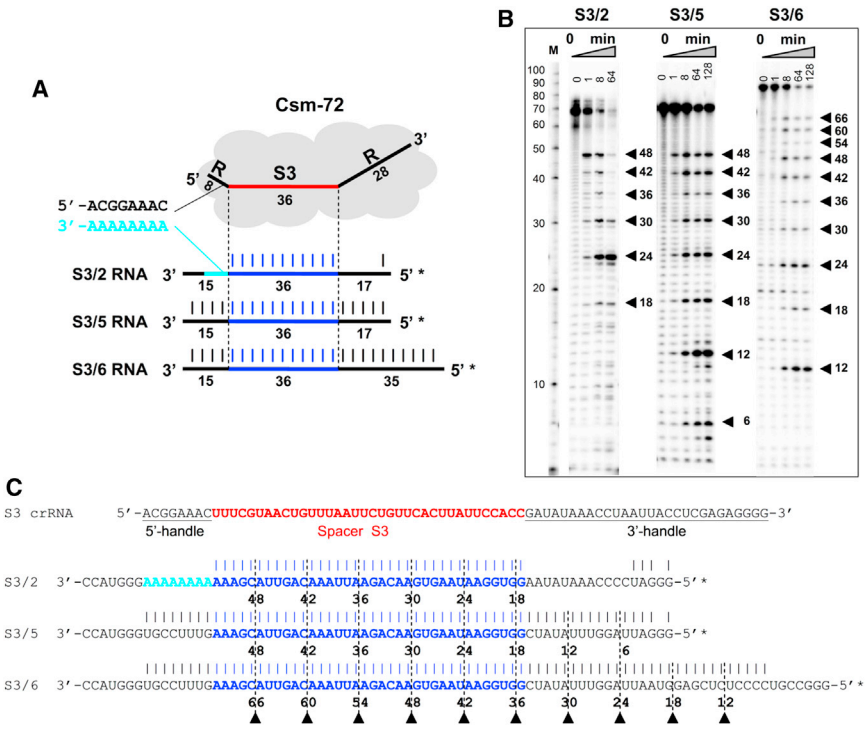

Figure 3. The Effect of the Sequence Complementarity outside the Spacer Region on the StCsm-72 Cleavage Pattern

(A) Schematic representation of the StCsm-72 complex and RNA substrates used in the cleavage assay. RNA substrates were 5′-end labeled with ^{32}P (indicated as *) and gel-purified.

(B) StCsm-72 cleavage assay. M, RNA Decade marker.

(C) RNA cleavage products mapped on the S3/2, S3/5, and S3/6 RNA substrates sequences. See also Figure S6.

RNA substrates, the cleavage sites were located at a fixed distance with respect to the conserved 5′-handle of crRNA (Figure 4C).

Identification of the Ribonuclease Subunit in the StCsm Complex

The regularly spaced cleavage pattern of the RNA target (Figures 2, 3, 4, and S4–S6) implies the presence of multiple cleavage modules in the Csm complex. According to the densitometric analysis, three Csm2 and five Csm3 subunits are identified in the Csm-40 complex, whereas six Csm2 and ten Csm3 subunits are present in the Csm-72 complex. Multiple copies of the Csm2 and Csm3 proteins in the Csm complexes make them prime candidates for catalytic subunits. StCsm2 is a small (121 amino acid [aa]) α-helical protein of unknown structure. StCsm3 (220 aa) contains a conserved RRM core and is fairly closely related (~35% sequence identity) to *Methanopyrus kandleri* Csm3, whose crystal structure has been solved recently (Hrle et al., 2013). We reasoned that, because the catalytic activity of the StCsm complex requires the presence of Me^{2+} ions, the active site is likely to contain one or more acidic residues. We inspected multiple sequence alignments of both Csm2 and Csm3 protein families for conserved aspartic or glutamic residues. We did not find any promising candidates in StCsm2, but identified several, including D33, D100, E119, E123, and E139, in StCsm3 (Figure 5A). To probe the role of these conserved negatively charged Csm3 residues, we constructed a single residue alanine replacement mutant. We also constructed the H19A mutant, because it was shown that the corresponding mutation (R21A) in *M. kandleri* Csm3 abolished binding of single-stranded RNA (Hrle et al., 2013). We expressed each mutant in the context of other StCsm/Cas proteins and analyzed the cleavage activity of the StCsm-40 complex containing mutant Csm3 subunits. StCsm3 H19A, E119A, E123A, and E139A mutants did not compromise the formation, RNA binding, or cleavage activity of Csm-40 complex (Figures 5B–5F). As in the *S. epidermidis* Csm3 (Hatoum-Aslan et al., 2013), the D100A mutation in StCsm3 affected crRNA length and compromised Csm complex stability; nevertheless, the StCsm complex still cleaved the target RNA with 6 nt periodicity (Figure 5E). Unlike other StCsm3 mutants, the D33A substitution impaired Csm-40 RNA cleavage (Figures 5E and 5F) without affecting RNA binding (Figure 5D) or complex assembly. Taken together, these data demonstrate that Csm3 is an RNase, producing multiple cleavage patterns spaced by regular 6 nt intervals, and that the D33 residue is part of the catalytic/metal-chelating site. The StCsm3 structural model based on the homologous structure of *M. kandleri* Csm3 is in good agreement with the identified role for this residue (Figure S7A). D33 belongs to the highly conserved surface patch that extends from the RRM core into the "lid" subdomain (Figure S7B). Part of this surface patch is positively charged, supporting the idea that it represents an RNA-binding site (Figure S7C).

In Vivo RNA Targeting by the StCsm Complex

To test whether the StCsm complex can target RNA in vivo, we used the MS2 phage restriction assay. MS2 is a lytic single-stranded RNA coliphage that infects *E. coli* via the fertility (F) pilus. The MS2 phage is a preferable model to investigate RNA targeting by the CRISPR-Cas system in vivo because no DNA intermediate is formed during the life cycle of this phage (Olsthoorn and van Duin, 2011). For the in vivo RNA-targeting experiment, the *E.coli* NovaBlue (DE3, F^+) strain was transformed with two compatible plasmids: (1) pCRISPR_MS2 plasmid bearing the synthetic CRISPR array of five repeats interspaced by four 36 nt spacers targeting correspondingly the mat, lys, cp, and rep MS2 RNA sequences; and (2) pCsm/Cas plasmid for the expression of Cas/Csm proteins (Figure 6A). The phage-targeting and control *E. coli* strains were plated and infected with series of dilutions of MS2 using the drop

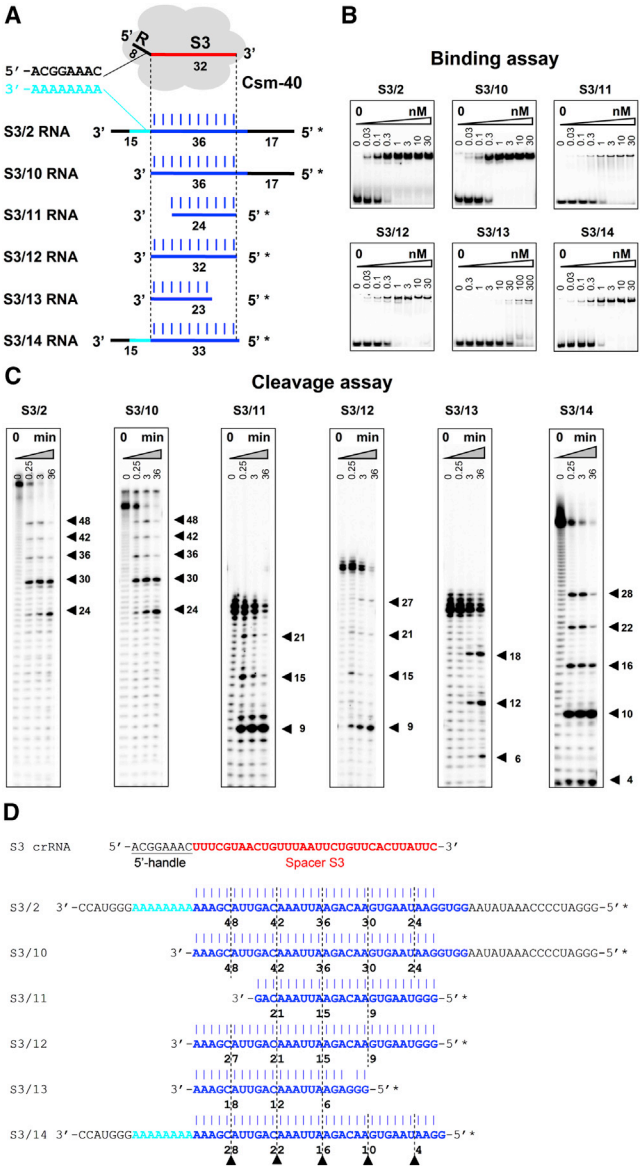

Figure 4. The Effect of Protospacer Truncations on the StCsm-40 Cleavage Pattern

(A) Schematic representation of the StCsm-40 complex and RNA substrates used in the binding and cleavage assays. RNA substrates were 5′-end labeled with ^{32}P (indicated as *) and gel-purified.

(B) EMSA analysis of RNA binding by StCsm-40.

(C) StCsm-40 cleavage assay. M, RNA Decade marker.

(D) RNA cleavage products mapped on the RNA substrates sequences.

plaque assay. The assay revealed that the *E. coli* strain expressing wild-type Csm and crRNAs that target MS2 induce a 3- to 4-log reduction of the plaquing efficiency with respect to the control cells (Figure 6). Importantly, no resistance to the MS2 phage infection was observed in the strain expressing either the nontargeting crRNA or the cleavage-deficient (D33A) Csm3 mutant. Taken together, these data demonstrate that the StCsm complex conveys in vivo resistance to RNA phage in the heterologous *E. coli* host.

DISCUSSION

We aimed to establish the NA specificity and mechanism for the type III-A CRISPR-Cas system of *S. thermophilus*. In sharp contrast to other CRISPR-Cas subtypes, the functional activity of the type III-A system so far has not been reconstituted in vitro. Cas/Csm proteins in the type III-A CRISPR locus of the *S. thermophilus* DGCC8004 are homologous to those of *S. thermophilus* DGCC7710 and LMD-9. They also show more distant but significant similarities to Cas/Csm proteins of *Lactococcus lactis*, *Enterococcus italicus*, and *S. epidermidis* (Marraffini and Sontheimer, 2008; Millen et al., 2012) (Figure S1).

Csm Complexes of *S. thermophilus*

We expressed the type III-A CRISPR-Cas locus of the DGCC8004 in *E. coli* and isolated two RNP complexes termed Csm-40 and Csm-72. Both complexes share a conserved set of Cas10, Csm2, Csm3, and Csm4 proteins. In addition to this core, the Csm-40 also contains the Csm5 protein. Two distinct crRNAs of 72 and 40 nt copurify with Csm-40 and Csm-72 complexes isolated from the heterologous *E. coli* host. The 72 nt crRNA composed of an 8 nt 5′-handle, a 36 nt spacer and a 28 nt 3′-handle would result from the pre-crRNA cleavage between 28 and 29 nt within the conserved repeat region, presumably by the Cas6 nuclease, similar to the III-B CRISPR-Cas system (Carte et al., 2008). The shorter 40 nt crRNA copurified with the Csm-40 complex of *S. thermophilus* contains the conserved 8 nt 5′-handle and 32 nt spacer, indicating that the 72 nt crRNA intermediate undergoes further 3′-end processing to produce a mature 40 nt crRNA that lacks the 3′-handle and 4 nt within the spacer region (Figure 7). The RNase involved in the maturation of 72 nt crRNA intermediate remains to be identified; however the Csm5 protein that is absent in Csm-72 but is present in Csm-40 could be a possible candidate. Indeed, *csm5* gene deletion in DGCC8004 produces only unmatured Csm-72 complexes (G.T., M.K., Irmantas Mogila; unpublished data).

The crRNA processing and maturation pathway in the *S. thermophilus* type III-A system (Figure 7) shows striking similarity to that in *S. epidermidis*. First, the SeCsm complex includes the same set of Cas10, Csm2, Csm3, Csm4, and Csm5 proteins as the StCsm-40. Furthermore, in *S. epidermidis*, the primary processing by Cas6 produces a 71 nt crRNA intermediate, that is subjected to further endonucleolytic processing at the 3′ end (Hatoum-Aslan et al., 2011, 2014).

StCsm Complex Cuts RNA, Producing a Regular Cleavage Pattern

The Csm complexes of *S. epidermidis* and *S. solfataricus* have been reconstituted and isolated; however, the NA cleavage activity has not been reported so far. In vivo studies in *S. epidermidis* suggested that the type III-A SeCsm RNP complex targets DNA (Marraffini and Sontheimer, 2008) in a PAM-independent manner and prevents autoimmunity by checking the complementarity between the crRNA 5′-handle and the 3′-flanking sequence in the vicinity of the protospacer (Marraffini and Sontheimer, 2010). In contrast to these data, we found that the

StCsm-40 and StCsm-72 complexes bind ssRNA with high affinity and cut a ssRNA target in a PAM-independent manner in the presence of Me^{2+} ions, producing a regular 6 nt cleavage pattern in the protospacer region (Figures 2D and S4C–S4E). In this respect, the type III-A StCsm complex resembles the RNA-targeting type III-B Cmr-complexes PfCmr, SsCmr, and TtCmr (Hale et al., 2009; Staals et al., 2013; Zhang et al., 2012) (Figure 7) rather than DNA targeting type I and II complexes. By targeting RNA rather than DNA, the StCsm complex avoids autoimmunity. We further show that the nucleotide context and noncomplementarity outside the protospacer have no effect on the target RNA cleavage, demonstrating that PAM or unpaired flanking sequences of the protospacer are not required for cleavage by the StCsm (Figure S6). The complementarity of the protospacer is the only prerequisite for the StCsm cleavage: nonmatching RNA is not cleaved; however, either two contiguous mismatches or end truncations in the complimentary protospacer S3 are tolerated (Figure S6). This suggests that StCsm may not rely on a seed sequence for target recognition. The differences in the cleavage patterns of the 5′- and 3′-labeled RNAs (Figure 2D) imply that cleavage first occurs at one of its 3′-end internal sites of the target RNA. It remains to be established whether the observed cleavage directionality is dictated by the structural features of the complex or RNA context-dependent cleavage rate differences.

Strikingly, we found that for the Csm-72 complex the target RNA is being cleaved at regular 6 nt intervals outside the protospacer if it retains base complementarity to the crRNA 3′ handle. Such regularly spaced cleavage pattern of the RNA target (Figures 2, 3, 4, S4, and S6) implies the presence of multiple cleavage modules in the Csm complex. The major difference between the Csm-40 and Csm-72 complexes is the number of Csm2 and Csm3 subunits. The Csm-40 contains 3 Csm2 and 5 Csm3 subunits, whereas Csm-72 contains six Csm2 and ten Csm3 subunits (Figure 1D). The size of the complexes determined by SAXS correlates with the different stoichiometry of Csm-40 and Csm-72. Indeed, both complexes show a slightly twisted elongated shape, but the Csm-72 is significantly more elongated than Csm-40 complex (Figure 1H). Taken together, these data suggest that the longer unmatured 72 nt crRNA intermediate in the Csm-72 complex binds additional copies of Csm2 and Csm3 subunits into a RNP filament (Figure 7).

Csm3 Is an RNase Subunit in the StCsm Complex

Computational analysis revealed that StCsm3 has a conserved RRM core and is fairly closely related (~35% sequence identity) to *M. kandleri* Csm3 (Hrle et al., 2013). StCsm3 displays close structural similarity to MkCsm3, in particular the RRM-core and insertions into RRM-core that form the "lid" subdomain (Figure S7A). In contrast, StCsm3 lacks both the N-terminal zinc-binding domain and the C-terminal helical domain, making its structure more compact compared to that of MkCsm3. Thus, StCsm3 may be considered as a trimmed-down version of MkCsm3. Guided by the multiple sequence alignment and homology model of StCsm3, we selected candidate active site/metal chelating residues of Csm3 and subjected them to alanine mutagenesis. We showed that the highly conserved D33 residue of the StCsm3 is critical for the RNA cleavage activity of the Csm complex, demonstrating that Csm3 is an RNase in the StCsm and other type III-A CRISPR-Cas systems (Figure 5).

Implications for Other RNA-Targeting CRISPR Systems

Taken together, our data indicate that the StCsm complex is specific for the RNA and cuts it in a PAM-independent manner, producing a regular 6 nt cleavage pattern. Furthermore, we demonstrate that the Csm3 protein, which is present in Csm-40 and Csm-72 complexes in multiple copies, acts as an RNase responsible for the target RNA cleavage. In this respect, the type III-A Csm complex of *S. thermophilus* closely resembles the RNA targeting type III-B Cmr complex of *T. thermophilus* (TtCmr complex) that also produces a regular 6 nt cleavage pattern (Staals et al., 2013). The RNA degrading subunit in the type III-B Cmr-module remains to be identified. Although there is currently no experimental evidence, Staals et al. suggested that Cmr4 could fulfill this role (Staals et al., 2013). Indeed, clustering of Csm3 and Cmr4 homologs by sequence similarity revealed that they form two related but separate groups (Figure S7D). On the other hand, neither Csm3 nor Cmr4 families are homogenous. They are comprised of sequence clusters of various sizes. StCsm3 is a member of a large representative group of Csm3 homologs that includes those from *S. epidermidis*, *L. lactis* and *M. kandleri*. Another large, but more loosely connected group does not have proteins from experimentally characterized systems, except for the Csm complex from *S. solfataricus*. Sso1425 and Sso1426, two of its Csm3-like proteins (Makarova et al., 2011a), are members of this group albeit they are nontypical. The Cmr4 family appears even more heterogeneous than Csm3. Cmr4 proteins of experimentally characterized III-B systems from *T. thermophilus* and *P. furiosus* represent one of the larger clusters, whereas Cmr4 from *S. solfataricus* is a nontypical outlier. Indeed, biochemical characterization revealed that PfCmr and TtCmr RNA cleavage mechanism are similar and follow a 3′- or 5′-ruler mechanism, respectively (Hale et al., 2009; Staals et al., 2013). Meanwhile, SsCmr endonucleolytically cleaves both target RNA and crRNA at UA dinucleotides (Zhang et al., 2012). Therefore, it would not be surprising if members of other, so far experimentally uncharacterized groups were part of Cmr complexes with somewhat different properties.

Naturally, we were curious if Csm3 and Cmr4 proteins may have similarly organized active site. The aligned sequences of Csm3 and Cmr4 subunits from characterized systems revealed that sequences of both families have Asp in the corresponding positions, suggesting similar active sites (Figure 5A). The exception is Sso1426. This is quite surprising, considering the composition of the *S. solfataricus* Csm complex. Four copies of Sso1426 were found to be present within the complex, suggesting that this subunit might play a role of the Csm3 (Rouillon et al., 2013). In contrast, another Csm3-like protein, Sso1425, does have the D33 counterpart, suggesting it can cleave ssRNA. However, only a single copy of Sso1425 was found in the *S. solfataricus* complex. Taken together, these data suggest that Csm-modules in *S. thermophilus* and *S. solfataricus* have different architectures and RNA cleavage mechanisms.

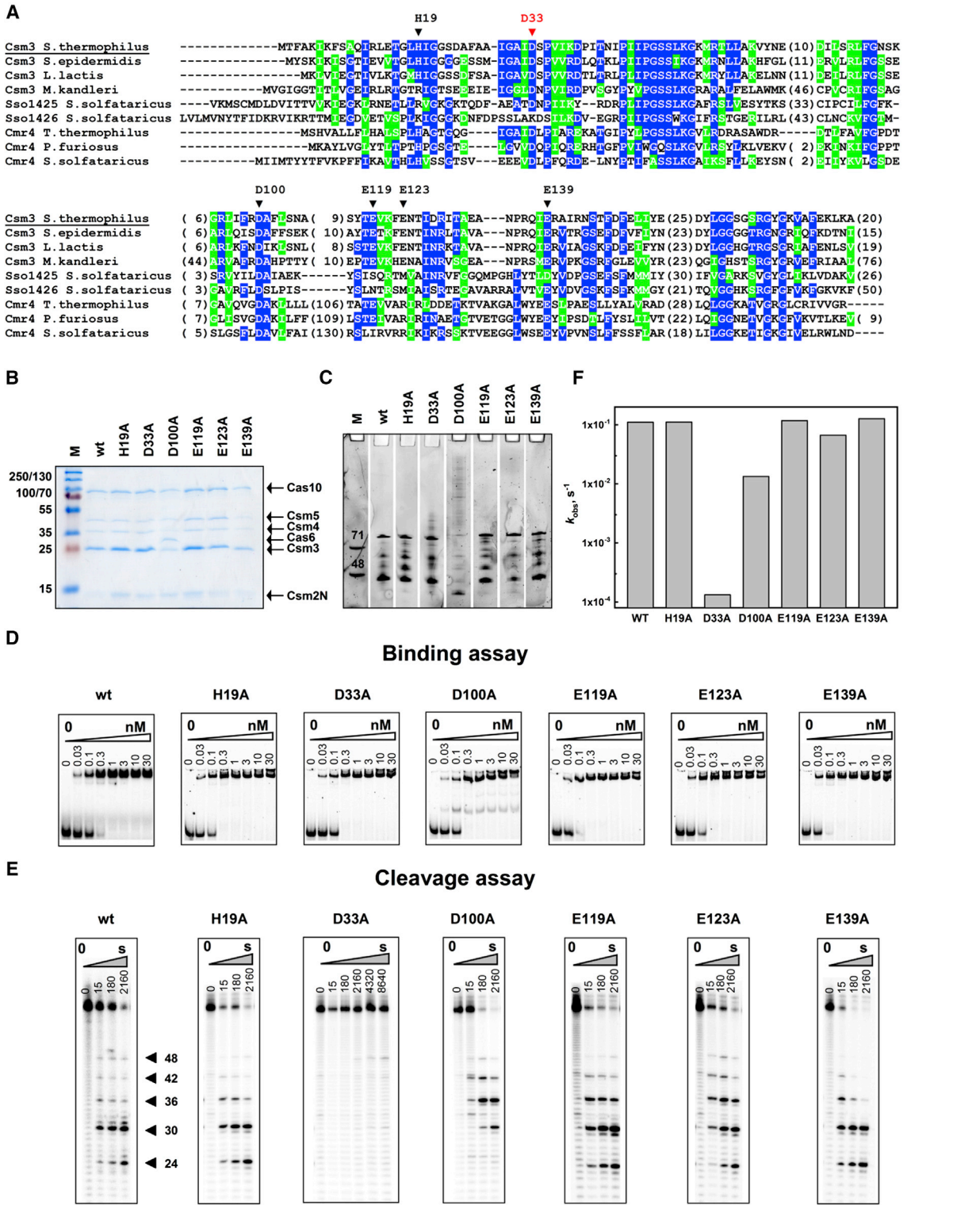

Figure 5. Computational and Mutational Analysis of Csm3

(A) Alignment of Csm3 and Cmr4 sequence representatives from experimentally characterized type III effector complexes. Identical and similar residues in more than half of sequences are shaded in blue and green correspondingly. StCsm3 positions subjected to site-directed mutations are indicated by triangles above the sequence (see also Figure S7).

(B) Coomassie blue-stained SDS-PAGE of StCsm complexes containing Csm3 mutants. M, protein marker.

(legend continued on next page)

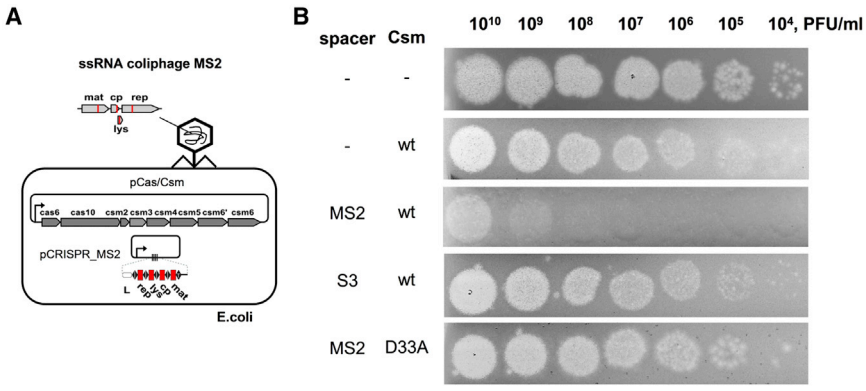

Figure 6. Restriction of ssRNA Phage MS2 in *E. coli* Cells Expressing StCsm Complex

(A) Schematic representation of the assay. The arrow indicates the promoter.

(B) Phage plaque analysis. Serial 10-fold dilutions of MS2 were transferred onto lawns of *E. coli* NovaBlue (DE3, F^+) strain expressing StCsm-crRNA complex targeting the MS2 genome or control cells.

Concluding Remarks

We show here that the Csm effector complex of the *S. thermophilus* type III-A system targets RNA and establish the mechanism of RNA cleavage. We demonstrate that in the type III-A effector complex, Cas/Csm proteins assemble into an RNP filament (Figure 7) that contains multiple copies of Csm2 and Csm3 proteins. Furthermore, we provide evidence that the Csm3 subunit acts as an RNase that cleaves target RNA at multiple sites spaced by regular 6 nt intervals (Figure 7). The number of cleavage sites correlates with the number of Csm3 subunits in the Csm effector complex. Easy programmability of the type III-A StCsm complex by custom crRNAs (Figure S5) paves the way for the development of molecular tools for RNA interference.

RNA cleavage specificity established here for the StCsm complex in vitro is supported by in vivo experiments of MS2 RNA phage interference in the heterologous *E.coli* host (Figure 6). It remains to be established whether RNA silencing by the StCsm complex can contribute to the DNA phage interference in the *S. thermophilus* host. Transcription-dependent DNA targeting mechanism has been proposed recently for the type III-B CRISPR-Cmr system (Deng et al., 2013); however, it has yet to be demonstrated for *S. thermophilus* and other type III-A systems.

EXPERIMENTAL PROCEDURES

Expression and Isolation of Csm Complexes

Heterologous *E. coli* BL21(DE3) cells producing the Strep-tagged Csm complexes were engineered and cultivated as described in the Supplemental Experimental Procedures. Csm-40 and Csm-72 complexes were isolated by subsequent Strep-chelating affinity and size exclusion chromatography steps (Supplemental Experimental Procedures).

Bioinformatic Analysis and Mutagenesis of Csm3

Putative active site residues of Csm3 were identified from multiple alignment of Csm3/Cmr4 (see Supplemental Experimental Procedures). Csm3 mutants were constructed using quick-change mutagenesis and purified as described in the Supplemental Experimental Procedures.

Extraction, HPLC Purification, and ESI-MS Analysis of crRNA

NAs copurified with Csm-40 and Csm-72 were isolated using phenol:chloroform:isoamylalcohol (25:24:1, v/v/v) extraction and precipitated with isopropanol. Purified NAs were incubated with 0.8 U DNase I or 8 U RNase I (Thermo Scientific) for 30 min at 37°C. NAs were separated on a denaturing 15% polyacrylamide gel (PAAG) and depicted with SybrGold (Invitrogen) staining.

Ion pair reversd-phase high-performance liquid chromatography (HPLC)-purified crRNA architecture was determined using denaturing RNA chromatography in conjunction with electrospray ionization mass spectrometry (ESI-MS) as described in (Sinkunas et al., 2013) and the Supplemental Experimental Procedures.

SAXS Experiments

SAXS data for Csm-40 and Csm-72 were collected at P12 EMBL beam-line at PETRAIII storage ring of DESY synchrotron in Hamburg (Germany). Csm-40 and Csm-72 complexes were measured in three different concentrations in buffer containing 20 mM Tris-HCl (pH 8.5 at 25°C), 0.5 M NaCl, 1mM EDTA, and 7 mM 2-mercaptoethanol. Data collection, processing, and ab initio shape modeling details are presented in Table S4 and Figure S3.

DNA and RNA Substrates

Synthetic oligodeoxynucleotides were purchased from Metabion. All RNA substrates were obtained by in vitro transcription using TranscriptAid T7 High Yield Transcription Kit (Thermo Scientific). A full description of all the DNA and RNA substrates is provided in Table S5. DNA and RNA substrates were either 5′-labeled with [γ^{32}P] ATP and PNK or 3′-labeled with [α^{32}P] cordycepin-5′-triphospate (PerkinElmer) and poly(A) polymerase (Life Technologies) followed by denaturing gel purification.

Electrophoretic Mobility Shift Assay

Binding assays were performed by incubating different amounts of Csm complexes with 0.5 nM of ^{32}P-5′-labeled NA in the binding buffer (40 mM Tris, 20 mM acetic acid [pH 8.4 at 25°C], 1 mM EDTA, 0.1 mg/ml BSA, 10% [v/v] glycerol). All reactions were incubated for 15 min at room temperature prior to electrophoresis on native 8% (w/v) PAAG. Electrophoresis was carried out at room temperature for 3 hr at 6 V/cm using 40 mM Tris, 20 mM acetic acid (pH 8.4 at 25°C), 0.1 mM EDTA as the running buffer. Gels were dried and depicted using an FLA-5100 phosphorimager (Fujifilm). The K_d for NA binding by Csm-72 and Csm-40 was evaluated assuming the complex concentration at which half of the substrate is bound as a rough estimate of K_d value. For binding competition assay, 0.5 nM ^{32}P-labeled S3/1 RNA was mixed with 0.5-5000 nM of unlabelled competitor NA and 0.3 nM StCsm-40, and analyzed by EMSA.

Cleavage Assay

The Csm-40 reactions were performed at 25°C and contained 20 nM of 5′- or 3′-radiolabeled NA (Table S5) and 62.5 nM (unless stated otherwise)

(C) Denaturing PAGE analysis of NA copurifying with mutant StCsm complexes. M, synthetic DNA marker.

(D) EMSA analysis of S3/2 RNA binding by mutant StCsm complexes.

(E) S3/2 RNA cleavage by the mutant StCsm complexes.

(F) The cleavage rate constant k_{obs} values for Csm3 mutant variants of StCsm-40.

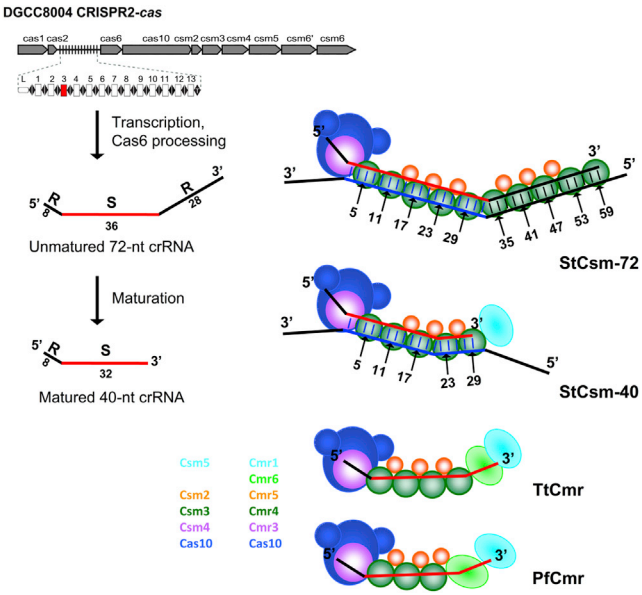

Figure 7. Structural and Cleavage Models of StCsm Complexes

The CRISPR2 transcript is first processed into 72 nt crRNA intermediates that undergo further maturation into 40 nt crRNA. Both crRNAs are incorporated into StCsm complexes that target RNA but differ by the number of Csm3 and Csm2 subunits. The number of RNA cleavage products correlates with the number of Csm3 nuclease subunits. Schematic models of StCsm complexes were generated based on similarity to TtCmr and PfCmr. Csm analogs of Cmr proteins according to (Makarova et al., 2011a) are colored identically.

complex in the reaction buffer (33 mM Tris-acetate [pH 7.9 at 25°C], 66 mM K-acetate, 0.1 mg/ml BSA, and 10 mM Mg-acetate). Csm-72 reactions were performed in the same reaction buffer at 37°C and contained 20 nM of radiolabeled NA and 125 nM of complex unless stated otherwise. Reactions were initiated by addition of the Csm complex. The samples were collected at timed intervals and quenched by mixing 10 μl of reaction mixture with 2× RNA loading buffer (Thermo Scientific) followed by incubation for 10 min at 85°C. The reaction products were separated on a denaturing 20% PAAG and depicted by autoradiography. ^{32}P-5′-labeled RNA Decade marker (Ambion) was used as a size marker. To map the cleavage products, oligoribonucleotide markers were generated by RNase A (Thermo Scientific, final concentration 10 ng/ml) treatment of RNA substrates for 8 min at 22°C or by alkaline hydrolysis in 50 mM $NaHCO_3$ (pH 9.5) at 95°C for 5 min.

Phage Drop Plaque Assay

Phage drop plaque assay was conducted using LGC Standards recommendations (see Supplemental Experimental Procedures).

ACCESSION NUMBERS

The GenBank accession number for the sequence of CRISPR2-*cas* locus of DGCC8004 reported in this work is KM222358.

SUPPLEMENTAL INFORMATION

Supplemental Information includes Supplemental Experimental Procedures, seven figures, and five tables and can be found with this article online at http://dx.doi.org/10.1016/j.molcel.2014.09.027.

ACKNOWLEDGMENTS

The authors thank Giedrius Gasiunas for suggestions and Tomas Sinkunas and Mindaugas Zaremba for plasmids. G.T. acknowledges support from Research Council of Lithuania (grant MIP-40/2013). The authors thank Dr. Andris Dislers (Latvian Biomedical Research and Study Center) for a kind gift of MS2 phage sample. M.J.D. acknowledges support from the Engineering and Physical Sciences Research Council (UK) and the Biotechnology and Biological Sciences Research Council (UK). P.H. is an employee of DuPont Nutrition and Health, Dangé-Saint-Romain, France.

Received: June 2, 2014
Revised: July 30, 2014
Accepted: September 25, 2014
Published: November 6, 2014

REFERENCES

Brouns, S.J., Jore, M.M., Lundgren, M., Westra, E.R., Slijkhuis, R.J., Snijders, A.P., Dickman, M.J., Makarova, K.S., Koonin, E.V., and van der Oost, J. (2008). Small CRISPR RNAs guide antiviral defense in prokaryotes. Science *321*, 960–964.

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

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

Gasiunas, G., Barrangou, R., Horvath, P., and Siksnys, V. (2012). Cas9-crRNA ribonucleoprotein complex mediates specific DNA cleavage for adaptive immunity in bacteria. Proc. Natl. Acad. Sci. USA *109*, E2579–E2586.

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

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

Hatoum-Aslan, A., Samai, P., Maniv, I., Jiang, W., and Marraffini, L.A. (2013). A ruler protein in a complex for antiviral defense determines the length of small interfering CRISPR RNAs. J. Biol. Chem. *288*, 27888–27897.

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

Hrle, A., Su, A.A., Ebert, J., Benda, C., Randau, L., and Conti, E. (2013). Structure and RNA-binding properties of the type III-A CRISPR-associated protein Csm3. RNA Biol. *10*, 1670–1678.

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

Makarova, K.S., Aravind, L., Wolf, Y.I., and Koonin, E.V. (2011a). Unification of Cas protein families and a simple scenario for the origin and evolution of CRISPR-Cas systems. Biol. Direct *6*, 38.

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

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

Marraffini, L.A., and Sontheimer, E.J. (2010). CRISPR interference: RNA-directed adaptive immunity in bacteria and archaea. Nat. Rev. Genet. *11*, 181–190.

Millen, A.M., Horvath, P., Boyaval, P., and Romero, D.A. (2012). Mobile CRISPR/Cas-mediated bacteriophage resistance in Lactococcus lactis. PLoS ONE *7*, e51663.

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

Olsthoorn, R., and van Duin, J. (2011). Bacteriophages with ssRNA. (Chichester: John Wiley & Sons Ltd).

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

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

Spilman, M., Cocozaki, A., Hale, C., Shao, Y., Ramia, N., Terns, R., Terns, M., Li, H., and Stagg, S. (2013). Structure of an RNA silencing complex of the CRISPR-Cas immune system. Mol. Cell *52*, 146–152.

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

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

Szczelkun, M.D., Tikhomirova, M.S., Sinkunas, T., Gasiunas, G., Karvelis, T., Pschera, P., Siksnys, V., and Seidel, R. (2014). Direct observation of R-loop formation by single RNA-guided Cas9 and Cascade effector complexes. Proc. Natl. Acad. Sci. USA *111*, 9798–9803.

Terns, R.M., and Terns, M.P. (2014). CRISPR-based technologies: prokaryotic defense weapons repurposed. Trends Genet. *30*, 111–118.

Westra, E.R., van Erp, P.B., Künne, T., Wong, S.P., Staals, R.H., Seegers, C.L., Bollen, S., Jore, M.M., Semenova, E., Severinov, K., 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.

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

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