

Multiple nucleic acid cleavage modes in divergent type III CRISPR systems

Jing Zhang, Shirley Graham, Agnes Tello, Huanting Liu and Malcolm F. White*

Biomedical Sciences Research Complex, University of St Andrews, Fife KY16 9ST, UK

Received October 26, 2015; Revised December 14, 2015; Accepted January 07, 2016

ABSTRACT

CRISPR-Cas is an RNA-guided adaptive immune system that protects bacteria and archaea from invading nucleic acids. Type III systems (Cmr, Csm) have been shown to cleave RNA targets *in vitro* and some are capable of transcription-dependent DNA targeting. The crenarchaeon *Sulfolobus solfataricus* has two divergent subtypes of the type III system (Sso-IIIID and a Cmr7-containing variant of Sso-IIIB). Here, we report that both the Sso-IIIID and Sso-IIIB complexes cleave cognate RNA targets with a ruler mechanism and 6 or 12 nt spacing that relates to the organization of the Cas7 backbone. This backbone-mediated cleavage activity thus appears universal for the type III systems. The Sso-IIIB complex is also known to possess a distinct 'UA' cleavage mode. The predominant activity observed *in vitro* depends on the relative molar concentration of protein and target RNA. The Sso-IIIID complex can cleave plasmid DNA targets *in vitro*, generating linear DNA products with an activity that is dependent on both the cyclase and HD nuclease domains of the Cas10 subunit, suggesting a role for both nuclease active sites in the degradation of double-stranded DNA targets.

INTRODUCTION

One mechanism used by bacteria and archaea to protect themselves from invading nucleic acids is the adaptive immune system CRISPR-Cas, comprised of CRISPR arrays of repeat-spacer units and CRISPR-associated (Cas) genes (reviewed in (1)). The CRISPR-Cas system acts through three stages. First, short fragments of invading viruses and plasmids are captured and integrated into CRISPR arrays in a process known as acquisition or adaptation (2), generating a library of 'non-self' spacer sequences. Second, the CRISPR loci are transcribed and processed into short CRISPR RNAs (crRNAs) (reviewed in (3)). These unit length crRNAs are subsequently assembled with Cas proteins into large ribonucleoprotein effector or surveillance complexes. Guided by the crRNA sequence, the effector

complexes detect and degrade cognate nucleic acids (reviewed in (4)).

CRISPR-Cas systems have been classified into five types (5,6). Distant phylogenetic relationships have been reported between several proteins from type I and III systems (7), in particular the Cas7 subunit (8). This relationship was confirmed by electron microscopy studies of the type III complexes, which revealed a helical backbone formed by Cas7 family subunits binding to the crRNA (9–11), with a striking similarity to that observed in the type I Cascade structure (12). However, these two systems are mechanistically distinct. Type I systems target double-stranded DNA and recruit the Cas3 helicase:nuclease protein to unwind and degrade target DNA. In contrast, type III systems were originally shown to target single-stranded RNA using either a ruler-dependent cleavage with 6 nucleotide (nt) spacing, as observed for the type IIIB complex from *Pyrococcus furiosus* (13) or a sequence-dependent 'UA' cleavage mechanism for the Cmr7-containing type IIIB complex from *Sulfolobus solfataricus* (14). These activities have also been shown to occur *in vivo* in two *Sulfolobus* species (15,16).

In contrast, the type IIIA systems have long been known to possess a DNA targeting activity *in vivo* (17), but this proved difficult to demonstrate *in vitro*, although the HD nuclease domain of the large subunit (Csm1/Cas10) was reported to possess DNA nuclease activity (18). Recently, a ruler-dependent 'RNA shredding' activity was described for the type IIIA complex from *Streptococcus thermophilus* (19). This study demonstrated that the bound RNA target was cleaved at regular 6 nt intervals along the Cas7 backbone of the complex, an activity dependent on a conserved acidic residue in the Csm3 subunit (19). Subsequently, this mechanism of RNA degradation has been observed for several other type III systems (20–23). Most recently, the crystal structure of a chimeric type IIIB complex bound to crRNA and target RNA revealed regular nt flipping of crRNA with a 6 nt periodicity, caused by the insertion of a conserved beta-hairpin 'thumb' of the Cmr4 (Cas7) subunit into the RNA duplex (24). The thumb includes the conserved acidic residue important for target RNA cleavage, suggesting that the local distortion and chemical environment at these positions is key to the periodic, backbone-mediated target RNA cleavage observed in the type III sys-

*To whom correspondence should be addressed. Tel: +44 1334 463432; Fax: +44 1334 462595; Email: mfw2@st-and.ac.uk

tems. This structural arrangement is strikingly reminiscent of the Cas7-mediated distortion observed in the structure of *Escherichia coli* Cascade bound to a crRNA:DNA heteroduplex (25), emphasizing the fundamental conservation of nucleic acid binding across the type I and type III systems (26).

In parallel with these studies, genetic analyses of a *Sulfolobus islandicus* type III system suggested that the complex cleaved DNA in a transcription-dependent mechanism (27). This observation was confirmed by studies of the *Staphylococcus epidermidis* type IIIA system by the Marraffini lab, who showed that DNA targeting by the type IIIA system is dependent on the transcription of the DNA target—a mechanism that they proposed allows dormant lysogenic phage DNA in the genome to escape degradation (28). Most recently, this activity was reconstituted *in vitro*, demonstrating that DNA cleavage by the *S. epidermidis* type IIIA complex was dependent on transcription of the target DNA by RNA polymerase (29). The active site responsible for the DNA cleavage, which was observed on only one DNA strand, was identified in the conserved cyclase domain of the large subunit, Csm1 (29), consistent with recent biochemical studies of Csm1 in isolation (30).

Thus, there is an emerging consensus that type III systems have a backbone-mediated ‘RNA shredding’ activity (although the functional relevance of this activity remains uncertain) and some at least also possess a transcription-dependent DNA targeting activity. The type III systems present in *S. solfataricus* strain P2 are relatively unusual. The complex previously denoted as belonging to type IIIA is now classified as a type IIID system (6). The type IIIB complex was reported to have a more compact, ‘crab claw’ structure and has an extra subunit, Cmr7, found in only a few type IIIB systems (14). Here, we demonstrate that the *S. solfataricus* type IIID and IIIB systems (previously known as the Csm and Cmr complexes, hereafter Sso-IIID and Sso-IIIB) also possess a backbone-mediated RNA shredding activity *in vitro*, confirming that RNA shredding is a universal property of all type III systems. Furthermore, the Sso-IIID complex can cleave dsDNA targets *in vitro* and the Sso-IIIB system can switch between sequence-dependent and backbone-mediated cleavage modes, depending on the amount of RNA target available.

MATERIALS AND METHODS

Oligonucleotides for polymerase chain reaction (PCR), subcloning and site directed mutagenesis

The oligonucleotides were synthesized by Integrated DNA Technologies (IDT) and the sequences are available on request from the corresponding author.

Expression and purification of tagged Sso-IIID and Sso-IIIB complexes from *S. solfataricus*

The Sso-IIID complex used in this study was tagged on the large subunit, Csm1/Sso1428. The Sso-IIIB complex was tagged on the subunit Sso1990. The gene encoding Sso1428 or Sso1990 was PCR amplified and cloned into pMZ1 (31), then subcloned into expression vector pSVA9 that expresses the relevant subunit with a C-terminal strep-His tag (32).

The expression vector carrying the gene encoding Sso1428 or Sso1990 was transformed into the *S. solfataricus* PH1–16 expression strain. After transformation, cells were first cultivated in unselective Brock medium containing 0.1% (w/v) tryptone and 10 µg/ml uracil, and then transferred to selective media containing 0.2% (w/v) glucose and 0.1% (w/v) NZ-amine without uracil. Once the $A_{600\text{ nm}}$ reached 0.6, cells were transferred to expression media containing 0.2% (w/v) arabinose and 0.1% (w/v) NZ-amine to induce the expression of the tagged subunit and then collected at an $A_{600\text{ nm}}$ of 0.8 - 1.0. Cells were harvested by centrifugation and disrupted by sonication (Soniprep 150, MSE) in buffer A (20 mM HEPES pH 7.5, 250 mM NaCl, 30 mM imidazole). The supernatant was filtered and loaded onto a Histrap column (GE Healthcare) in buffer A. His-tagged proteins were eluted with a linear gradient of buffer B (20 mM HEPES pH 7.5, 250 mM NaCl, 500 mM imidazole). Fractions containing the tagged complex were pooled and exchanged into buffer C (20 mM Tris-HCl pH 8.0, 50 mM NaCl) then loaded onto a mono Q column (GE Healthcare) equilibrated in buffer C. Bound proteins were eluted with a linear gradient of buffer D (20 mM Tris-HCl pH 8.0, 1 M NaCl). Fractions containing the tagged complex were pooled, concentrated and loaded onto a Sephacryl 300 size exclusion column (GE Healthcare) equilibrated in buffer E (20 mM Tris-HCl pH 8.0, 250 mM NaCl). The purified proteins were concentrated and stored at 4°C.

To generate a variant Sso-IIIB complex with a mutated HD domain, the gene encoding the large subunit, Cmr2/Sso1991, was PCR amplified and cloned into a modified pMZ1 containing an N-terminal strep-His tag, resulting in pMZ1-Nhis-sso1991. Mutations of residues D181 and H182 to alanine were carried out by PCR mutagenesis using Phusion DNA polymerase (ThermoFisher Scientific) and pMZ1-Nhis-sso1991 as the template. Variants of the Sso-IIID complex with mutations targeted to the HD nuclease (H187A, D188A) and cyclase (D629A, D630A) domains were constructed in a similar manner. The procedures for subcloning, expression and purification of these variant proteins were as described above.

Cloning, expression and purification of Sso-IIID subunits

The gene encoding each subunit except *sso1428* was amplified by PCR and cloned into an expression vector with an N-terminal 6x His-tag. The *sso1424* gene was cloned into the pEHIS-Tev vector (33). The *sso1425*, *sso1426*, *sso1427*, *sso1430*, *sso1431* and *sso1432* genes were cloned into pDEST14 vector individually using the gateway recombination cloning (34). The constructs were expressed in C43 (DE3) *E. coli* grown in LB medium supplemented with appropriate antibiotics to an $A_{600\text{ nm}}$ of 0.6–0.8 at 37°C. Expression was induced by the addition of 0.4 mM IPTG and the cells were harvested after overnight incubation at 25°C. The cell pellets were resuspended in buffer F (50 mM Tris-HCl pH 7.5, 500 mM NaCl, 10 mM imidazole, 10% glycerol) and lysed by sonication. The lysate was clarified by centrifugation and filtered prior to loading onto a Histrap column (GE Healthcare) equilibrated in buffer G (50 mM Tris-HCl pH 7.5, 500 mM NaCl, 30 mM imidazole). The His-tagged proteins were eluted with a linear gradient

of buffer H (50 mM Tris-HCl pH 7.5, 500 mM NaCl, 500 mM imidazole). Fractions containing the His-tagged protein were pooled and treated with TEV protease overnight at room temperature for cleavage of the his-tag. The cleaved protein was passed through a HisTrap column in buffer G and collected in the flow through then further purified by size exclusion on a Superdex S200 column (GE Healthcare) equilibrated in buffer I (20 mM Tris-HCl pH 7.5, 250 mM NaCl). Fractions containing the proteins were concentrated and stored at -20°C .

The cell pellets for Sso1427 and Sso1431 were resuspended in buffer G containing 8 M urea and incubated overnight at room temperature. Samples were subject to sonication and ultracentrifugation at 40 000 rpm for 30 min prior to nickel affinity chromatography in buffer G containing 8 M urea. His-tagged proteins were eluted with a linear gradient of buffer H containing 8 M urea. Fractions containing the His-tagged proteins were stored at 4°C .

The *sso1428* gene was amplified and cloned into pMZ1 and then subcloned into pSVA9 and expressed from *S. solfataricus* as described above. The isolated Sso1428 subunit was purified using metal affinity, ion exchange, and size exclusion chromatography as for the intact complex (9). Fractions containing Sso1428 were concentrated and stored at 4°C .

Purification and end-labelling of RNA oligonucleotides

RNA oligonucleotide substrates were supplied by IDT and purified and end labelled as described previously (14).

Sso-IIID RNA cleavage assays

Purified Sso-IIID complex and $5'$ - ^{32}P end-labelled synthetic target RNA oligonucleotides were mixed in reaction buffer 1 (20 mM MES pH 6.0, 100 mM NaCl, 0.1 mg/ml BSA, 2 mM MgCl_2) and incubated at 75°C for 20 min. The reaction was stopped by adding $1\ \mu\text{l}$ of 20 mg/ml Proteinase K (Thermo scientific) and was further incubated at 37°C for 30 min. Then the reaction was phenol-chloroform (Ambion) extracted as described previously (35). Samples from the upper aqueous phase were mixed with formamide dye and heated at 95°C for 5 min before loading onto 20% polyacrylamide, 8 M urea, 1 x TBE gels. The gels were visualized by phosphor imaging. $5'$ end-labelled RNA size standards (Decade Markers, Ambion) were used to determine the sizes of the observed products.

Sso-IIIB RNA cleavage assays

Purified Sso-IIIB complex and unlabelled crRNA were mixed in reaction buffer 2 (20 mM MES pH 6.0, 100 mM potassium glutamate, 10 mM DTT, 10 mM MnCl_2) and pre-incubated at 75°C for 10 min prior to the addition of ^{32}P $5'$ end-labelled synthetic target RNA to the reaction mix. The reaction was further incubated at 75°C for 10 min and processed as described above.

Reconstitution of the Sso-IIID complex in vitro

Sso1424, Sso1425, Sso1426, Sso1427, Sso1428, Sso1430, Sso1431, Sso1432 and crRNA were mixed at a ratio of

3:1:4:1:1:1:1:1:1 in reaction buffer 1 and incubated at 45°C for 30 min prior to the addition of 4 nM radiolabelled target RNA. The reaction was further incubated at 75°C for 60 min then processed as described above.

Sso-IIID plasmid cleavage assays

The pT6 plasmid (36) was used as the backbone for plasmid cleavage assays. A synthetic DNA fragment (gBlock, IDT) designed to incorporate sequences corresponding to three of the most abundant crRNA sequences identified in the Sso-IIID complex (corresponding to spacers A26 and B2) (9) and Sso-IIIB complex (spacer D63) (14) was subcloned into the BamHI site of the plasmid and the resultant plasmid, pSpacer, was checked by DNA sequencing. The sequence of the gBlock is presented in the supplementary data section. Purified wild type and mutant Sso-IIID complexes and plasmid DNA with or without the spacer insert were mixed in reaction buffer 3 (10 mM Tris pH 8.0, 50 mM KCl, 2 mM MgCl_2) and incubated at 40°C for 20 min. The reaction was stopped by adding $1\ \mu\text{l}$ of 20 mg/ml Proteinase K (Thermo scientific) and was further incubated at 37°C for 30 min. Then the reaction was phenol-chloroform (Ambion) extracted as described previously (32). Samples from the upper aqueous phase were mixed with formamide dye and resolved on 0.8% agarose gels. GeneRuler 1kb DNA ladder (Life Technologies) was used as a reference marker. Band intensities were quantified using ImageJ (37). Note that as EtBr may not bind supercoiled and relaxed DNA with equivalent binding densities, the absolute fraction of DNA cleaved is subject to some uncertainty. This does not affect the comparative analysis of wild-type and variant proteins.

RESULTS

Sso-IIID cleaves target RNA with 6 or 12 nt spacing

The type IIIA complexes from *Streptococcus thermophilus*, *S. epidermidis* and *Thermus thermophilus* cleave target RNAs at regular 6 nt intervals that correspond to the spacing of the Cas7 backbone subunits (19,21,29). To test the Sso-IIID complex for this activity we incubated the complex with a labelled target RNA, A26, which has sequence complementary to one of the most abundant crRNAs previously identified after extraction from the Sso-IIID complex, comprising $\approx 5\%$ of the total crRNA present (9). We observed three major cleavage products with a 12 or 6 nt spacing (Figure 1). Addition of further crRNA A26 reduced the activity, suggesting that the complex may not incorporate this crRNA productively *in vitro*. Deletion of 5 nt from the $5'$ end of the target RNA (A26S) yielded the same 12–6 cleavage pattern with the fragments now 5 nt shorter, suggesting that cleavage is directed by a ruler mechanism from the $5'$ end of the crRNA (Figure 1A), in common with other type III enzymes (19,23). The assay was repeated using a different target RNA, B2, which matches another abundant crRNA identified in the Sso-IIID complex (9). Again, a 12–6 nt spacing was observed and once again the addition of further cognate crRNA diminished the extent of target cleavage (Figure 1A). We proceeded to test Sso-IIID for backbone-mediated cleavage using a target RNA, A1, that

is not present in the purified complex (9). No cleavage was observed, regardless of whether crRNA complementary to A1 was provided in the assay (Figure 1C).

Initial assays were performed in the presence of 2 mM Mg^{2+} . Although no activity was detected in the presence of EDTA, a broad range of divalent metal ions supported backbone RNA cleavage by Sso-IIIID, including Mg^{2+} , Mn^{2+} , Ca^{2+} , Ni^{2+} , Co^{2+} and Zn^{2+} (Supplementary Figure S1B). A similar broad specificity was observed for the *S. thermophilus* IIIA complex (19). To determine the 3' end structure of the cleaved RNA species, we incubated the cleavage products with *E. coli* Polymerase A, which can extend RNA from a 3'-hydroxyl group. The cleavage products remained unchanged whereas the target RNA A26 was polyadenylated as expected (Supplementary Figure S1A). This is consistent with the presence of a 5' hydroxyl group and a 2',3'-cyclic phosphate moiety at the ends of the cleaved RNA products, as observed for other type III systems (13).

A time course of cleavage of the A26 target, performed under single turnover conditions where the Sso-IIIID complex was present in molar excess (500 nM, of which 5% or 25 nM is loaded with crA26) over the substrate (4 nM), demonstrated very rapid cleavage of the target at sites 1–3 with $\approx 50\%$ of substrate cleaved within the first 2 min. There was little evidence for interconversion between products, although the site 3 product built up more slowly over the course of the reaction, suggesting this site may be cleaved less rapidly than the others (Figure 1D). However, when a large excess (10 μM) of unlabelled target RNA A26 was added to the reaction, no substrate turnover was observed, even after 40 min. These observations suggest that backbone cleavage by Sso-IIIID may be a single-turnover process, with product generation limited by the concentration of active enzyme present.

Together, these data show that the Sso-IIIID complex shares the backbone-mediated RNA cleavage activity observed for several other type III systems, but with an unusual spacing that suggests one cleavage site is missed, which is explored further below. The Sso-IIIID system is not 'programmable' with crRNA added *in trans*, suggesting that crRNA co-purifying with the complex is tightly bound and cannot be exchanged *in vitro*. Furthermore, this activity seems to be limited to a single-turnover, at least *in vitro*.

Reconstitution of the Sso-IIIID complex from individually expressed subunits

Each subunit of Sso-IIIID was cloned and expressed individually, allowing reconstitution in recombinant form with a defined crRNA. The large subunit (Csm1, Sso1428) was expressed in *S. solfataricus* with a C-terminal polyhistidine tag, as described previously (9). Isolated Csm1 could be purified by metal affinity chromatography and separated from Csm1 in the Sso-IIIID complex during size exclusion chromatography (Supplementary Figure S2). All of the other subunits were expressed in *E. coli* with cleavable N-terminal polyhistidine tags. Subunits Sso1424, 1425, 1426, 1430 and 1432 were obtained in soluble form and purified by metal affinity chromatography and size exclusion. Subunits Sso1427 and 1431 were insoluble after expression in

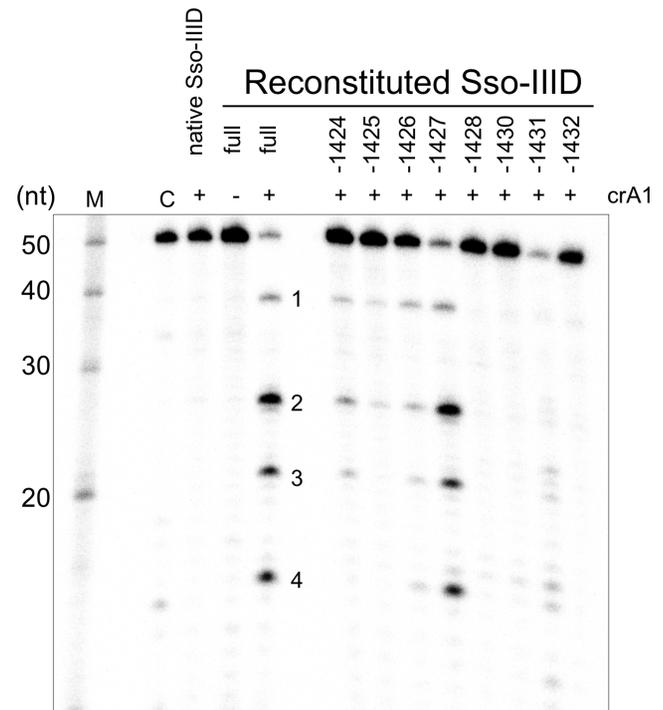


Figure 2. The Sso-IIIID complex can be reconstituted in active form *in vitro*. 5'- ^{32}P end-labelled target RNA A1 (4 nM) was incubated with purified native Sso-IIIID complex (500 nM) or recombinant, reconstituted Sso-IIIID (rSso-IIIID) in reaction buffer for 60 min at 75°C, followed by denaturing gel electrophoresis and phosphorimaging. Native Sso-IIIID lacks crRNA A1 and thus does not cleave this target. Sso-IIIID reconstituted from all the individual subunits (full complex) in the presence of crRNA A1 cleaves the A1 target RNA giving four products (1–4) with 12–6–6 spacing. Reconstitution experiments where one subunit was omitted (as indicated in the figure) showed varying levels of activity, with individual Cas7-like backbone subunits not essential for activity. Other lanes: M—RNA ladder; C—Control lane with no added Sso-IIIID.

E. coli. These proteins were purified by metal affinity chromatography in the presence of 8 M urea (Supplementary Figure S2). We reconstituted the Sso-IIIID complex by mixing the individual subunits in the presence of a defined crRNA, as described in the methods. The reconstituted complex was purified by size exclusion chromatography, with the intact complex migrating as a peak corresponding to the retention time of the native complex, while sub-complexes and individual subunits eluted later (Supplementary Figure S2). Although inefficient in terms of yield, this approach provided sufficient active reconstituted complex for biochemical studies.

The Sso-IIIID complex reconstituted with crRNA corresponding to the A1 spacer, which is not naturally present, catalysed cleavage of target RNA A1 with the same 12–6 nt spacing seen for the native complex (Figure 2). A further product (site 4) was observed, with a 6 nt spacing from site 3, suggesting that cleavage can occur at four different sites in the backbone. To investigate the subunit requirements of the complex, we reconstituted Sso-IIIID with individual subunits omitted, one at a time. For these experiments, the reconstituted complex was not subjected to gel filtration. Reconstitutions that omitted the large subunit (Csm1/Sso1428), Sso1430, Sso1431 or Sso1432 all resulted

in no detectable activity, suggesting that either complexes are not formed or are inactive when these subunits are absent. Sso1428 (Csm1/Cas10) and Sso1431 (Csm4/Cas5-like) form a close interaction at the base for all type III complexes (24,38,39). Native mass spectrometry places the Sso1430 subunit in contact with these two basal subunits in Sso-IIID (9), while Sso1432 is predicted to cap the Cas7-like backbone (9). When the small subunit (Cmr2/Sso1424) or the Cas7-like subunits Sso1425, or Sso1426 were missing, the cleavage activity was weakened but not eliminated, suggesting that these subunits play an important but not essential role. Omission of the Cas7-like subunit Sso1427 resulted in cleavage activity comparable to the full complex (Figure 2). For all these reconstituted complexes the cleavage pattern held the 12–6–6 pattern. These data suggest Sso-IIID has the potential to form a complex capable of backbone-mediated cleavage when one of the Cas7-like backbone subunits is absent, either by substituting the missing subunit with a paralogous one, or by forming sub-complexes that retain activity.

Dual modes of RNA cleavage by the Sso-IIIB Cmr7-containing complex

The type IIIB complexes from *P. furiosus* and *T. thermophilus* have been demonstrated to cleave the target RNA with a regular 6 nt spacing pattern (11,40). In contrast with these studies, we previously reported that the Sso-IIIB complex, which includes the unique Cmr7 subunit, cleaves target RNA selectively at UA sites, rather than using a 6-nt ruler-based mechanism (14). One difference between our original report and the work of other labs was that we had used a large excess of target RNA over enzyme in our assays. To investigate this further, we carried out the RNA cleavage reaction with 500 nM Sso-IIIB and a fixed concentration of radioactively labelled target RNA (4 nM), and varied the total target RNA concentration by adding unlabelled target RNA, resulting in a final substrate range from 4 to 1000 nM (Figure 3). Consistent with our previous report, Sso-IIIB cleaved target RNA at UA sites when target RNA was present in excess of the enzyme (Figure 3A, right hand side). As cold target RNA was reduced below the concentration of Sso-IIIB, the pattern changed to the canonical backbone-mediated cleavage with 6 nt spacing observed for other type III effector complexes. Both cleavage patterns could be detected simultaneously at intermediate ratios.

To examine the possibility that the HD nuclease domain of the Sso-IIIB complex, which is conserved in all type III systems, played a role in the UA cleavage mode, we constructed a variant complex where the conserved 'DH' dipeptide in the HD domain was changed to an 'AA' motif. This variant was expressed in *S. solfataricus* and purified and assayed as for the wild-type protein (Figure 3B). The complex gave robust backbone-mediated cleavage of the A1 target RNA, as expected. The UA cleavage mode was still apparent, suggesting that the HD nuclease domain is not essential for this activity. We were unable to create variants in subunits such as Cmr4 or the unique Cmr7 subunit, as these are present in multi-copy in the complex and our experimental system does not allow deletion of the wild-type genes in *S.*

solfataricus, with the result that complexes would contain mixtures of wild-type and variant subunits.

To further characterize the dual cleavage modes, we investigated the rate of substrate cleavage under the extreme conditions that result in only one cleavage mode (Figure 4). Under conditions of high enzyme excess, where only a single turnover is possible, backbone cleavage occurred very quickly, with $\approx 70\%$ of the target RNA cleaved by the first time point at 1 min. There was little further conversion over the following 20 min reaction time course. These properties are consistent with those of Sso-IIID under similar conditions (Figure 1). To elicit the UA cleavage mode, the reaction was carried out under conditions of high substrate excess by adding unlabelled target RNA. Target RNA was cleaved in multiple turnover conditions as observed previously (14), with around 50% cut over the 20 min reaction. Both experiments were carried out in triplicate and gels were quantified by phosphorimaging. Figure 4B shows a plot of the amount of target RNA substrate cleaved per nmole of protein versus time for the two conditions. It is apparent that the backbone-mediated cleavage occurs in rapid, single turnover mode whilst UA cleavage displays initially slower kinetics that nonetheless result in considerably higher levels of substrate turnover over the course of the 20 min reaction time.

DNA cleavage by the Sso-IIID complex

The activity of Sso-IIID wild type protein and three variants with amino acid replacements targeted to the HD domain, cyclase domain or both were tested for the ability to cleave plasmid DNA (Figure 5). We constructed a plasmid (pSpacer) containing sequences matching those of abundant crRNAs present in the Sso-IIID and -IIIB complexes. Supercoiled pSpacer plasmid DNA was cleaved by the wild type protein to generate nicked and linear DNA products (Figure 5A). This activity did not depend on the presence of a cognate target DNA sequence within the plasmid, as the parental plasmid was cleaved similarly (not shown). This suggests that the activity is unregulated compared to that expected *in vivo*. Single-strand nicking occurs first, with the generation of increasing quantities of linear DNA over time, suggesting that nicked products are further converted into linear ones by a second strand cleavage activity.

To investigate this further we compared the activities of variants with targeted changes in the HD and cyclase active sites, quantifying the proportions of supercoiled, nicked and linear DNA as a function of protein:DNA ratio in replicate experiments. The HD variant showed a modest but discernible defect in processing of supercoiled DNA compared to the wild type complex, with a small reduction in the amounts of nicked products generated (Figure 5B). The cyclase variant gave a stronger defect, with significantly more uncleaved plasmid apparent at the highest protein concentrations. The double variant yielded the most profound defect, with a significant proportion of the DNA uncleaved at the highest protein concentrations and strongly reduced generation of nicked or linear products, even when compared to the cyclase variant (Figure 5D, G). These data suggest that both the HD and cyclase domains play a role in DNA cleavage in Sso-IIID. It should be noted that, partic-

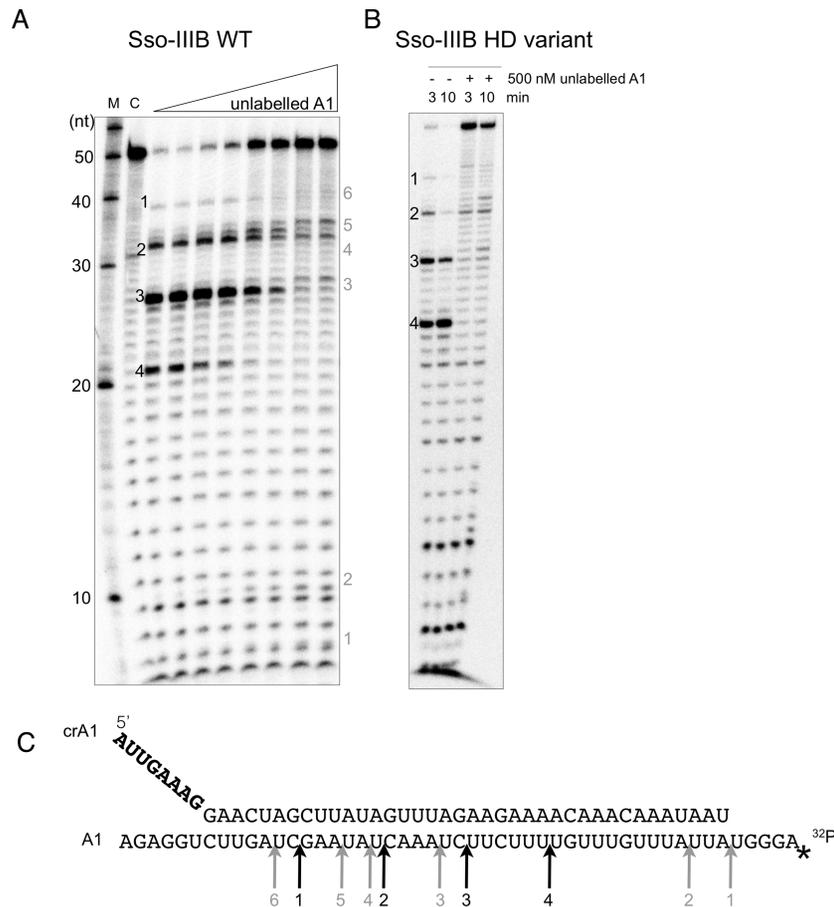


Figure 3. Dual modes of RNA cleavage by the Sso-III B complex. (A) 5'-³²P end-labelled target RNA A1 (4 nM) was incubated with purified endogenous Sso-III B complex (500 nM) and crRNA A1 (100 nM) in reaction buffer, together with a varied concentration of unlabelled A1 target RNA (0–1000 nM). Reactions were incubated at 75°C for 20 min, followed by denaturing gel electrophoresis and phosphorimaging. Backbone-mediated cleavage with a 6–6–6 spacing pattern predominated where target concentration was low, but as total substrate concentration increased the cleavage mode switched to the UA cleavage observed previously. Major cleavage sites are labelled 1–4 in black for backbone cleavage and 1–6 in grey for UA cleavage. Other lanes: M—RNA ladder; C—A1 RNA incubated in absence of Sso-III B. (B) Backbone and UA cleavage modes for an Sso-III B complex with a variant HD nuclease domain (D181A/H182A) in the large (Cmr2) subunit. The HD variant retained strong backbone-mediated cleavage and significant UA cleavage activity. (C) Sequences of the guide and target A1 RNAs used in the above assay. Arrows indicate the locations of the cleavage products, labelled 1–4 in black for backbone cleavage and 1–6 in grey for UA cleavage.

ularly for high protein:DNA concentrations, this assay may not necessarily discriminate between the independent nicking of both strands due to two distinct Sso-III D complexes and the coordinate nicking of both strands by the two active sites of a single complex. This may explain why the single HD and cyclase variants can still generate some linear DNA.

DISCUSSION

Backbone-mediated RNA cleavage by Sso-III D and III B

The observation that the Sso-III D and III B-Cmr7 complexes are both capable of target RNA cleavage with a backbone-mediated 5' ruler mechanism means that all type III systems tested *in vitro*, including these two divergent examples, possess this type of activity. Although the early EM studies of Sso-III B reported a 'crab claw' structure (14), the biochemical data presented here support the presence of a canonical Cas7 backbone found in all Type I and Type III complexes. For Sso-III B, the four RNA cleavage sites prob-

ably correspond to the positions of the Cmr4 (Sso1987) subunits that make up the backbone of many type III B complexes (22). As previously reported (19) the Sso1987 subunit has a conserved aspartate residue (D34) equivalent to D33 in the *S. thermophilus* type III A and D26 in the *P. furiosus* III B systems, both of which are essential for backbone cleavage activity (19,22). For technical reasons, we could not generate a variant form of Sso-III B uniformly mutated at this position, as gene knockouts are problematic in *S. solfataricus*.

For the Sso-III D effector the situation is more complex. Figure 6 shows a schematic representation of a generic type III complex, with details for Sso-III D shown below. The 12–6–6 nt spacing observed for three different target RNA species suggests that one specific position in the backbone does not support RNA cleavage. Sso-III D has two closely related Cas7-like subunits, Sso1425 and Sso1426, which are thought to be present at a stoichiometry of 1 and 4, respectively (9). A simple explanation for the missed cleavage site would therefore be that the single Sso1425 subunit does

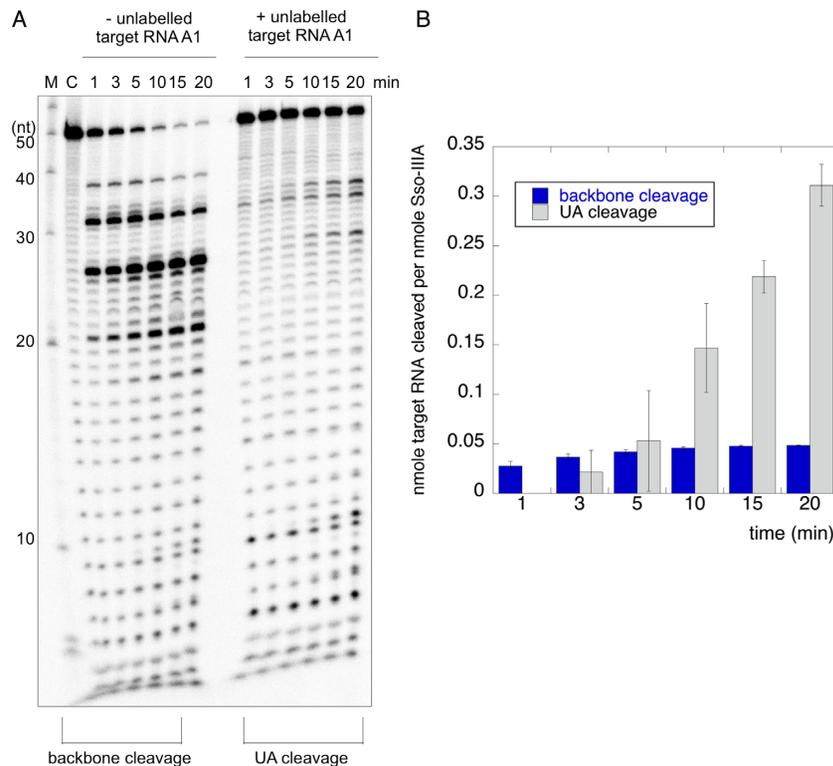


Figure 4. The two RNA cleavage modes of the Sso-III B complex possess distinct catalytic rates. (A) Time course reactions of backbone cleavage (left) and UA cleavage (right). Conditions were 40 nM Sso-III B complex with target RNA A1 at 2 nM (left), or 800 nM Sso-III B complex with total target RNA A1 at 500 nM (right). In both cases crRNA A1 was added to 100 nM final concentration. (B) Quantification of the extent of product formation for the two cleavage modes. The total nmoles of product formed per nmoles of protein is plotted against reaction time. Data points are the mean of triplicate experiments and standard errors are shown.

not support target RNA cleavage. However, reconstituted Sso-III D complexes lacking either Sso1425 or Sso1426 both support target RNA cleavage with the same 12–6–6 spacing as the native protein (Figure 2). This suggests that considerable plasticity in the backbone composition of Sso-III D is possible, and also rules out the simple explanation of one inactive subunit for the missed cleavage site. An alternative explanation is that the crRNA:target RNA duplex adopts a different conformation at the position corresponding to the missed site, which is at position 2 in the backbone (Figure 6). This would explain the conservation of the 12–6–6 spacing with different target RNAs and different subunit compositions. Given that the Sso-III D complex has more distinct subunits than any other type III system (eight, as opposed to seven in Sso-III B and six in most other type III complexes), it is perhaps not surprising that there may be subtle differences in RNA binding. Indeed, the Sso1426 subunit lacks the highly conserved ‘D33’ residue identified as responsible for backbone-mediated cleavage in the *S. thermophilus* type III A complex (19).

One possibility is that backbone-mediated RNA cleavage, which has largely been detected *in vitro* and which predominates under conditions of enzyme excess, may not be a relevant activity *in vivo*. In support of this, the abolition of backbone-mediated RNA cleavage in the *S. epidermidis* type III A system did not significantly affect plasmid immunity *in vivo*, whilst abrogation of DNA targeting abolished immunity (29). Backbone-mediated RNA cleavage by the

S. thermophilus type III A system has been shown to provide resistance against infection by the ssRNA phage MS2 *in vivo* (19), but it is open to question whether this phage is a relevant target for CRISPR systems as it is unclear how protospacers could be captured from such a virus.

Sulfolobus islandicus REY15A, encodes two type III systems. One resembles the Sso-III B Cmr7 complex and does not appear to target DNA *in vivo*. The other is a more canonical type III B system and has been demonstrated to possess both RNA and transcription-dependent DNA targeting that provides immunity against phage. The activities of these two complexes have been investigated *in vivo*. Both could target mRNA when programmed with a cognate crRNA sequence. The complex without Cmr7 was shown to cleave RNA with a 6 nt ruler mechanism, whilst the complex with Cmr7 cleaved at UA sites, and the latter complex provided more efficient RNA targeting *in vivo* (16). These data suggest that the UA cleavage mechanism of Sso-III B and related complexes, which predominates *in vitro* when RNA targets are present in excess, may be the dominant activity *in vivo*. Unfortunately, the active site responsible for UA cleavage remains to be determined. Mutation of the HD nuclease active site generated a variant protein that retained UA cleavage activity (Figure 3), whilst backbone-mediated cleavage activity was unaffected. The Cmr7 subunit is an attractive candidate for a catalytic role in UA cleavage. Unfortunately Cmr7 is present at a high stoichiometry in the Sso-III B complex (14) and we are unable to express and pu-

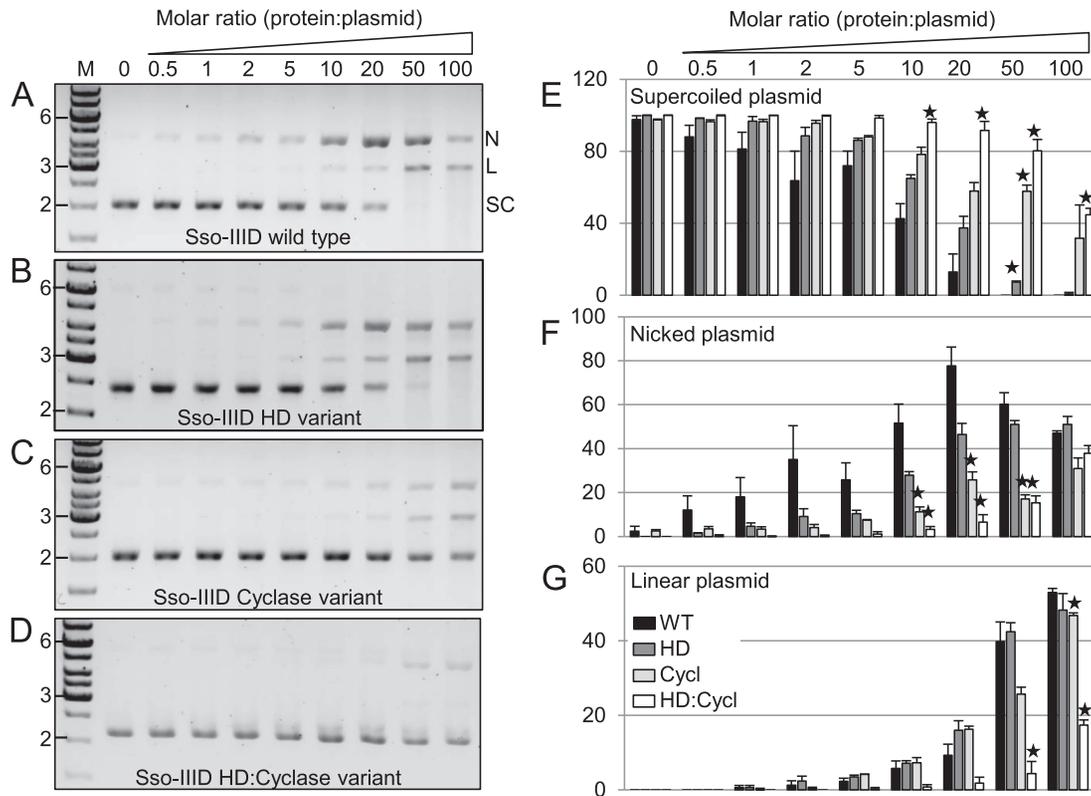


Figure 5. DNA cleavage by Sso-IIIID is dependent on the HD and cyclase domains of the large subunit (Sso1428). Comparison of DNA cleavage activity between wild type (A) and three variant (B-D) Sso-IIIID complexes. (A) The supercoiled DNA plasmid pSpacer (SC-bottom band) is nicked (N-top band) and linearized (L-middle band) by the wild type Sso-IIIID complex, with no apparent supercoiled plasmid remaining at a 50:1 protein:plasmid ratio after a 20 min incubation. (B-D) The efficiency of nicking and linearization is slightly reduced in the HD variant (B), more significantly reduced in the cyclase variant (C), and severely impaired in the HD:cyclase double variant (D). (E-G) Reduction in cleavage activity for the three variants was quantified. Bars represent the mean band intensity for each plasmid species (where SC+L+N = 100) from triplicate experiments, and error is displayed as standard deviation (SD). Differences between the wild type and each variant Sso-IIIID were determined using a T-test (one-tailed, unpaired), and significance is denoted by: $p < 0.001$. M- DNA marker (GeneRuler 1kb).

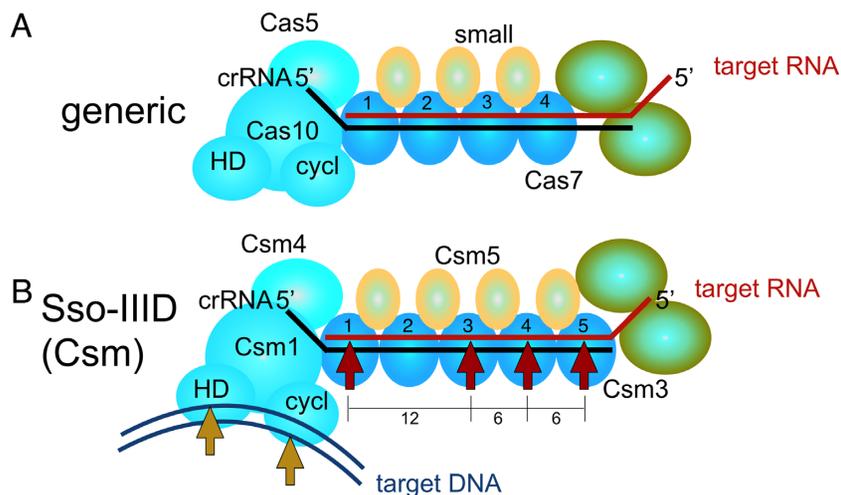


Figure 6. Schematic representation of the subunit organisation and target nucleic acid cleavage patterns for type III complexes. (A) Organisation of a generic type III complex. The Cas10-Cas5 sub-complex defines the ‘foot’ of the structure and binds the 5’ end of the crRNA. The Cas10 subunit has two nuclease active sites: HD nuclease (HD) and cyclase (cycl). The backbone is made up of a variable number of Cas7 and small subunits, whilst the ‘head’ has one or two subunits (green), which are not always essential. The crRNA is shown in black and target RNA in red. (B) The Sso-IIIID complex has a five-subunit backbone composed of four copies of Sso1426 and one of Sso1425. Cleavage of the target RNA (red arrows) is observed with a 12–6–6 spacing, consistent with cutting at sites 1, 3, 4 and 5. For clarity, subunits whose role and placement are unclear are omitted from this diagram. Target DNA (magenta) may be cleaved by the HD and cyclase nuclease sites of the Csm1 subunit on separate strands (orange arrows).

rify Sso-III_B variants lacking wild-type copies of the Cmr7 subunit as knockouts are technically very difficult in the *Sulfolobus* strain used for these experiments.

DNA targeting activities of type III complexes

A variety of nuclease activities have been ascribed to type III effector complexes (Figure 6). In addition to a presumably universal backbone-mediated RNA cleavage mechanism, distinct nuclease activities have been ascribed to the large subunit of type III systems. The *Thermococcus onnurineus* Cas10 subunit acts as a metal dependent nuclease *in vitro* with both circular and linear ssDNA substrates and the activity appears to be located in the HD domain (18). In contrast the DNA nuclease activity of *S. epidermidis* Cas10 *in vitro* was located in the cyclase domain (41), in keeping with recent studies showing cyclase-mediated cleavage of DNA during transcription (29). Here we have shown that the intact Sso-III_D complex also cleaves plasmid DNA *in vitro*, albeit in a sequence-independent manner that may represent a deregulated aspect of its activity. It appears that the cyclase domain is more important for plasmid nicking than the HD domain, however the most striking observation is that both the cyclase and HD domain active sites are required for efficient conversion of supercoiled DNA into linear form. This raises the prospect that these two nuclease domains in the Cas10 subunit each cut one strand of target DNA duplexes during interference *in vivo* (Figure 6). This would be analogous to the situation in Cas9, where there are two nuclease active sites, targeting opposing strands (42). The hypothesis that both the HD and cyclase domains, which are conserved and function as nucleases under certain circumstances, collaborate to cleave target DNA duplexes *in vivo*, is an attractive one. However, this awaits definitive evidence. Clearly, there is still much to learn about the activities and functions of the type III CRISPR effector complexes.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

Thanks to Christophe Rouillon and Clare Rollie for helpful discussions.

FUNDING

Biotechnology and Biological Sciences Research Council [BB/K000314/1, BB/M000400/1 to M.F.W.]. Funding for open access charge: Biotechnology and Biological Sciences Research Council [BB/K000314/1, BB/M000400/1 to M.F.W.].

Conflict of interest statement. None declared.

REFERENCES

1. Plagens, A., Richter, H., Charpentier, E. and Randau, L. (2015) DNA and RNA interference mechanisms by CRISPR-Cas surveillance complexes. *FEMS Microbiol. Rev.*, **39**, 442–463.

2. Heler, R., Marraffini, L.A. and Bikard, D. (2014) Adapting to new threats: the generation of memory by CRISPR-Cas immune systems. *Mol. Microbiol.*, **93**, 1–9.
3. Charpentier, E., Richter, H., van der Oost, J. and White, M.F. (2015) Biogenesis pathways of RNA guides in archaeal and bacterial CRISPR-Cas adaptive immunity. *FEMS Microbiol. Rev.*, **39**, 428–441.
4. Jiang, F. and Doudna, J.A. (2015) The structural biology of CRISPR-Cas systems. *Curr. Opin. Struct. Biol.*, **30**, 100–111.
5. 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.* (2011) Evolution and classification of the CRISPR-Cas systems. *Nat. Rev. Microbiol.*, **9**, 467–477.
6. Makarova, K.S., Wolf, Y.I., Alkhnbashi, O.S., Costa, F., Shah, S.A., Saunders, S.J., Barrangou, R., Brouns, S.J., Charpentier, E., Haft, D.H. *et al.* (2015) An updated evolutionary classification of CRISPR-Cas systems. *Nat. Rev. Microbiol.*, **13**, 722–736.
7. Makarova, K.S., Aravind, L., Wolf, Y.I. and Koonin, E.V. (2011) Unification of Cas protein families and a simple scenario for the origin and evolution of CRISPR-Cas systems. *Biology Direct*, **6**, 38.
8. Lintner, N.G., Kerou, M., Brumfield, S.K., Graham, S., Liu, H., Naismith, J.H., Sdano, M., Peng, N., She, Q., Copie, V. *et al.* (2011) Structural and Functional Characterization of an Archaeal Clustered Regularly Interspaced Short Palindromic Repeat (CRISPR)-associated Complex for Antiviral Defense (CASCADE). *J. Biol. Chem.*, **286**, 21643–21656.
9. Rouillon, C., Zhou, M., Zhang, J., Politis, A., Beilstein-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.
10. 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.
11. 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.
12. 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.
13. 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.
14. 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.
15. Zebec, Z., Manica, A., Zhang, J., White, M.F. and Schleper, C. (2014) CRISPR-mediated targeted mRNA degradation in the archaeon *Sulfolobus solfataricus*. *Nucleic Acids Res.*, **42**, 5280–5288.
16. Peng, W., Feng, M., Feng, X., Liang, Y.X. and She, Q. (2015) An archaeal CRISPR type III-B system exhibiting distinctive RNA targeting features and mediating dual RNA and DNA interference. *Nucleic Acids Res.*, **43**, 406–417.
17. Marraffini, L.A. and Sontheimer, E.J. (2008) CRISPR interference limits horizontal gene transfer in staphylococci by targeting DNA. *Science*, **322**, 1843–1845.
18. Jung, T.Y., An, Y., Park, K.H., Lee, M.H., Oh, B.H. and Woo, E. (2015) Crystal structure of the Csm1 subunit of the Csm complex and its single-stranded DNA-specific nuclease activity. *Structure*, **23**, 782–790.
19. Tamulaitis, G., Kazlauskienė, M., Manakova, E., Venclovas, C., Nwokeoji, A.O., Dickman, M.J., Horvath, P. and Siksnys, V. (2014) Programmable RNA shredding by the type III-A CRISPR-Cas system of *Streptococcus thermophilus*. *Mol. Cell*, **56**, 506–517.
20. Taylor, D.W., Zhu, Y.F., Staals, R.H.J., Kornfeld, J.E., Shinkai, A., van der Oost, J., Nogales, E. and Doudna, J.A. (2015) Structures of the CRISPR-Cmr complex reveal mode of RNA target positioning. *Science*, **348**, 581–585.
21. Staals, R.H., Zhu, Y., Taylor, D.W., Kornfeld, J.E., Sharma, K., Barendregt, A., Koehorst, J.J., Vlot, M., Neupane, N., Varossieau, K.

- et al.* (2014) RNA targeting by the type III-A CRISPR-Cas Csm complex of *Thermus thermophilus*. *Mol. Cell*, **56**, 518–530.
22. Ramia, N.F., Spilman, M., Tang, L., Shao, Y.M., Elmore, J., Hale, C., Coccozaki, A., Bhattacharya, N., Terns, R.M., Terns, M.P. *et al.* (2014) Essential structural and functional roles of the Cmr4 subunit in RNA cleavage by the Cmr CRISPR-Cas complex. *Cell Rep.*, **9**, 1610–1617.
 23. Zhu, X. and Ye, K.Q. (2015) Cmr4 is the slicer in the RNA-targeting Cmr CRISPR complex. *Nucleic Acids Res.*, **43**, 1257–1267.
 24. Osawa, T., Inanaga, H., Sato, C. and Numata, T. (2015) Crystal structure of the CRISPR-Cas RNA silencing Cmr complex bound to a target analog. *Mol. Cell*, **58**, 418–430.
 25. Mulepati, S., Heroux, A. and Bailey, S. (2014) Structural biology. Crystal structure of a CRISPR RNA-guided surveillance complex bound to a ssDNA target. *Science*, **345**, 1479–1484.
 26. Jackson, R.N. and Wiedenheft, B. (2015) A Conserved Structural Chassis for Mounting Versatile CRISPR RNA-Guided Immune Responses. *Mol. Cell*, **58**, 722–728.
 27. 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.
 28. Goldberg, G.W., Jiang, W.Y., Bikard, D. and Marraffini, L.A. (2014) Conditional tolerance of temperate phages via transcription-dependent CRISPR-Cas targeting. *Nature*, **514**, 633–637.
 29. Samai, P., Pyenson, N., Jiang, W., Goldberg, G.W., Hatoum-Aslan, A. and Marraffini, L.A. (2015) Co-transcriptional DNA and RNA Cleavage during Type III CRISPR-Cas Immunity. *Cell*, **161**, 1164–1174.
 30. Ramia, N.F., Tang, L., Coccozaki, A.I. and Li, H. (2014) *Staphylococcus epidermidis* Csm1 is a 3'-5' exonuclease. *Nucleic Acids Res.*, **42**, 1129–1138.
 31. Zolghadr, B., Weber, S., Szabo, Z., Driessen, A.J. and Albers, S.V. (2007) Identification of a system required for the functional surface localization of sugar binding proteins with class III signal peptides in *Sulfolobus solfataricus*. *Mol. Microbiol.*, **64**, 795–806.
 32. Albers, S.V., Jonuscheit, M., Dinkelaker, S., Urich, T., Kletzlin, A., Tampe, R., Driessen, A.J. and Schleper, C. (2006) Production of recombinant and tagged proteins in the hyperthermophilic archaeon *Sulfolobus solfataricus*. *Appl. Environ. Microbiol.*, **72**, 102–111.
 33. Liu, H. and Naismith, J.H. (2009) A simple and efficient expression and purification system using two newly constructed vectors. *Protein Expr. Purif.*, **63**, 102–111.
 34. Oke, M., Carter, L.G., Johnson, K.A., Liu, H., McMahon, S.A., Yan, X., Kerou, M., Weikart, N.D., Kadi, N., Sheikh, M.A. *et al.* (2010) The Scottish Structural Proteomics Facility: targets, methods and outputs. *J. Struct. Funct. Genomics*, **11**, 167–180.
 35. Sokolowski, R.D., Graham, S. and White, M.F. (2014) Cas6 specificity and CRISPR RNA loading in a complex CRISPR-Cas system. *Nucleic Acids Res.*, **42**, 6532–6541.
 36. Paytubi, S. and White, M.F. (2009) The crenarchaeal DNA damage-inducible transcription factor B paralogue TFB3 is a general activator of transcription. *Mol. Microbiol.*, **72**, 1487–1499.
 37. Schneider, C.A., Rasband, W.S. and Eliceiri, K.W. (2012) NIH Image to ImageJ: 25 years of image analysis. *Nat. Methods*, **9**, 671–675.
 38. Shao, Y., Coccozaki, A.I., Ramia, N.F., Terns, R.M., Terns, M.P. and Li, H. (2013) Structure of the Cmr2-Cmr3 subcomplex of the Cmr RNA silencing complex. *Structure*, **21**, 376–384.
 39. Benda, C., Ebert, J., Scheltema, R.A., Schiller, H.B., Baumgartner, M., Bonneau, F., Mann, M. and Conti, E. (2014) Structural model of a CRISPR RNA-silencing complex reveals the RNA-target cleavage activity in Cmr4. *Mol. Cell*, **56**, 43–54.
 40. Hale, C.R., Coccozaki, A., Li, H., Terns, R.M. and Terns, M.P. (2014) Target RNA capture and cleavage by the Cmr type III-B CRISPR-Cas effector complex. *Genes Dev.*, **28**, 2432–2443.
 41. Ramia, N.F., Tang, L., Coccozaki, A.I. and Li, H. (2013) *Staphylococcus epidermidis* Csm1 is a 3'-5' exonuclease. *Nucleic Acids Res.*, **42**, 1129–1138.
 42. 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. U.S.A.*, **109**, E2579–E2586.