

Cas3/I-C mediated target DNA recognition and cleavage during CRISPR interference are independent of the composition and architecture of Cascade surveillance complex

Siddharth Nimkar and B. Anand *

Department of Biosciences and Bioengineering, Indian Institute of Technology Guwahati, Guwahati 781039, Assam, India

Received June 20, 2019; Revised December 17, 2019; Editorial Decision December 18, 2019; Accepted January 22, 2020

ABSTRACT

In type I CRISPR-Cas system, Cas3—a nuclease cum helicase—in cooperation with Cascade surveillance complex cleaves the target DNA. Unlike the Cascade/I-E, which is composed of five subunits, the Cascade/I-C is made of only three subunits lacking the CRISPR RNA processing enzyme Cas6, whose role is assumed by Cas5. How these differences in the composition and organization of Cascade subunits in type I-C influence the Cas3/I-C binding and its target cleavage mechanism is poorly understood. Here, we show that Cas3/I-C is intrinsically a single-strand specific promiscuous nuclease. Apart from the helicase domain, a constellation of highly conserved residues—which are unique to type I-C—located in the uncharacterized C-terminal domain appears to influence the nuclease activity. Recruited by Cascade/I-C, the HD nuclease of Cas3/I-C nicks the single-stranded region of the non-target strand and positions the helicase motor. Powered by ATP, the helicase motor reels in the target DNA, until it encounters the roadblock en route, which stimulates the HD nuclease. Remarkably, we show that Cas3/I-C supplants Cas3/I-E for CRISPR interference in type I-E *in vivo*, suggesting that the target cleavage mechanism is evolutionarily conserved between type I-C and type I-E despite the architectural difference exhibited by Cascade/I-C and Cascade/I-E.

INTRODUCTION

Archaea and Bacteria defend themselves from the assault of phages and plasmids by employing a genetically encoded RNA mediated adaptive immunity referred to

as CRISPR-Cas system. CRISPR constitutes an array of direct repetitive sequences (each of about 30–40 bp) that are intervened by similarly sized variable spacer sequences. These spacer sequences are preferentially captured from the invading foreign genetic element and get integrated into the CRISPR array. They serve as immunological memory—akin to antibodies in higher organisms—to specifically recognize and target the invader during recurrent infection (1–5). CRISPR based targeting can be operationally distinguished into three distinct stages: (i) adaptation, (ii) maturation and (iii) interference. Immunological memory is acquired during adaptation wherein short fragment of DNA from the foreign genetic element (protospacer) is acquired and incorporated into the CRISPR locus in a polarized fashion (6–11). This is followed by the transcription and processing of the pre-CRISPR transcript that give rise to the mature CRISPR RNA (crRNA) onto which either a single (Class II) or several Cas protein(s) (Class I) assemble(s) to form a ribonucleoprotein (RNP) surveillance complex. The crRNA within the RNP guides the target recognition by base complementarity whereas the protein component(s) facilitate(s) the cleavage of the target nucleic acid (1,3,4,12–14).

CRISPR-Cas system is highly diverse and it is broadly divided into two classes (I and II), which are further divided into six types (I–VI) and several subtypes (15–18). In Class II system, the RNP surveillance complex is comprised of a single Cas protein viz. Cas9/II-A (19,20) or Cpf1/V-A (21) and this effector complex participate in both target recognition and cleavage. On the contrary, the surveillance complex in Class I system is highly variable and comprises of several Cas proteins (3,22–24), which are involved in target recognition alone.

In Class I/type I system, the effector complex for CRISPR interference is referred to as Cascade—which comprises of crRNA, a large and a small Cas subunits, variable number of Cas5, Cas6 and Cas7—which

*To whom correspondence should be addressed. Tel: +91 361 2582223; Fax: +91 361 2582249; Email: banand@iitg.ac.in

recognize the target and recruit Cas3—a trans-acting motor protein—that cleaves the target DNA (22,25–27). Target identification by Cascade complex requires that the upstream of target sequence contains a short nucleotide sequence (2–4 nt) called proto-spacer adjacent motif (PAM), which is present only in the invading phages/plasmids (28). The primary function of the PAM sequence is to discriminate self versus non-self and thereby pre-empt targeting of the host genome (13,29). In line with this, mutations in the PAM sequence lead to target evasion (25,30,31). Target identification by Cascade leads to R-loop formation, wherein the target strand base pairs with the crRNA and the non-target strand becomes single-stranded (32–37). It appears that Cas3 recruitment is facilitated by a conformational change in the large subunit of Cascade complex—say Cas8E in type I-E—and this leads to Cas3-mediated preferential cleavage of the non-target strand (31,38–41). Cas3 acts on the invasive DNA by cooperative interaction between its ‘HD’ nuclease and Superfamily 2 (SF2) ‘DExD/H’ helicase domains (42,43). The HD domain shows divalent metal ion-dependent nuclease activity with a preference for single-stranded DNA (44–46). The DExD/H box helicase domain comprises of two juxtaposed RecA domains, and these are followed by a long linker and a C-terminal domain (CTD), whose function is poorly defined. After recruitment by Cascade, nuclease domain of Cas3 binds to non-target strand after making a single-stranded nick, followed by processive degradation of target DNA (36,47).

Befitting the enormous diversity exhibited by type I system, the composition and architecture of Cascade effector complex have undergone commensurate variation (15,16). In tune with this, the domain architecture of Cas3 in some sub-types has evolved in a very distinctive way such that HD and Helicase domains, which are fused in some sub-types, have undergone domain fission in others (45,48,49). The current mechanistic understanding on CRISPR interference in type I stems majorly from the prototypical type I-E system (25,31,34,37,40,42,43,47,50,51) with sporadic investigations from other type I systems such as I-A, I-F and I-G (48,49,52–59). In type I-C, the Cascade complex is composed of only three subunits, viz. Cas5, Cas7 and Cas8, as against five (CasA-E) in case of type I-E (24,26). Cas6—which is essential for crRNA maturation—is absent in type I-C and Cas5, which is inert in type I-E, has supplanted the role of Cas6 (24,60,61). Given the architectural differences among Cascade complex of type I-C and type I-E, it is reasonable to assume that this will be translated into functional differences. However, apart from type I-E, studies that systematically investigated the function of Cas3, as well as the functional interaction between Cascade and Cas3 in type I-C, are virtually non-existent hitherto. Arguably, this is partly due to the challenges associated with the reconstitution of Cascade complex and Cas3 in type I-C *in vitro*. Here, we have attempted to elucidate the functional mechanism of CRISPR interference in type I-C both by reconstituting the Cascade/I-C and Cas3/I-C *in vitro* and *in vivo*. This revealed the unique features of type I-C interference

machinery as well as the conserved functional mechanism with type I-E system.

MATERIALS AND METHODS

Strain construction

For studies related to the integration of genes encoding Cascade/I-C (Cas5, Cas7 and Cas8) from *B. halodurans*, *E. coli* IYB5101 (vide Supplementary Table S2) was used as a recipient strain. This strain harbors T7 polymerase, which is under the control of *araC* and hence it is inducible with arabinose. Genes encoding Cascade/I-C (Cas5, Cas7 and Cas8) were amplified as a polycistronic construct using *B. halodurans* C-125 genomic DNA as a template. This construct has T7 promoter sequence upstream of the coding region. Subsequently, this construct was integrated into *E. coli* IYB5101 using pOSIP-CT (addgene #45981) to create the strain IC-I through one-step cloning and chromosomal integration of DNA (clonetelegraph), which uses phage integrases (62). Double knockout of *cas3* ($\Delta cas3$) and *hns* (Δhns) was created in *E. coli* K-12 BW25113 (vide Supplementary Table S2) using λ Red recombineering (63) to create strain IC-2. The list of strains employed for this study is presented in Supplementary Table S2.

Molecular cloning

Bacillus halodurans C-125 strain was procured from The Microbial Type Culture Collection and Gene Bank (MTCC), India. Gene encoding Cas3/I-C (BH0336) and all its mutants were cloned in 1R (addgene #29664) at SspI restriction site by employing Gibson Assembly. Point mutations were created in Cas3/I-C using PCR-based mutagenesis. The 1R vector encodes Strep II tagged protein. Genes encoding Cascade/I-C (Cas5, Cas7 and Cas8) were amplified as a polycistronic construct and inserted into the 1R expression vector to produce pCascade/I-C. CRISPR array containing seven copies of identical repeat-spacer units that were preceded and succeeded by T7 promoter and terminator sequences, respectively, was commercially synthesized (GenScript) and inserted into 13S-R (addgene # 48328) to produce pCRISPR/I-C. 13S-R and 1R are compatible vectors and can be used for purification of Cascade/I-C complex. Target DNA sequence was inserted into the pUC19 vector using KpnI and HindIII restriction sites. All cloned constructs were verified by Sanger sequencing. The information on the oligonucleotides used in this study can be found in Supplementary Table S1.

Protein purification

E. coli BL21 (DE3) harboring pWT-1 (Cas3/I-C cloned in 1R) was grown in LB broth supplemented with 50 $\mu\text{g/ml}$ kanamycin at 37 °C till the OD₆₀₀ reached 0.6. At this point, Cas3/I-C expression was induced with 0.2 mM IPTG and cells were allowed to grow for 12 h at 18 °C. Cells were harvested and re-suspended in binding buffer (20 mM Tris-Cl (pH 8.0), 300 mM NaCl, 10% glycerol, 1 mM phenylmethylsulfonyl fluoride (PMSF)

and 6 mM β -mercaptoethanol). Cells were lysed using cell disruptor (20 kpsi), and cell debris was removed by centrifugation at 4 °C and 16000 $\times g$ for 30 min. After cell lysis, DNase I (50 μ g/ml) was added to reduce the viscosity due to the presence of intact genomic DNA. The clear supernatant was loaded onto pre-equilibrated 5 ml StrepTrap HP (GE Healthcare) column. After loading, the column was washed with 10 CV of binding buffer to remove unbound protein and Cas3/I-C was eluted in binding buffer containing 2.5 mM desthiobiotin. Cas3/I-C was further loaded onto HiTrap Heparin HP column (GE Healthcare) to remove impurities. Subsequently, it was eluted with a linear gradient of 0.15–2 M NaCl in binding buffer. Later, the protein was further purified through HiLoad 16/600 Superdex 200 prep grade column (GE Healthcare) and eluted in buffer containing 20 mM Tris-Cl (pH 7.5), 150 mM NaCl, 10% glycerol and 6 mM β -mercaptoethanol. After concentration, samples were flash frozen in liquid nitrogen and stored at –80 °C until use. Protein samples were run on SDS-PAGE to check the purity and concentration was measured by absorbance at 280 nm. Cas3/I-C mutants cloned in 1R (pD48A, pQ253A, pD395A, pK742A, pK743A, pQ745A, pQ746A and pY747A) were purified using similar protocol mentioned above.

E. coli BL21 (DE3) harboring pCascade/I-C and pCRISPR/I-C (vide Supplementary Table S3) was grown in LB broth supplemented with 50 μ g/ml kanamycin and 100 μ g/ml spectinomycin at 37 °C till the OD₆₀₀ was equal to 0.6. Cascade/I-C expression was induced with 0.2 mM IPTG and cells were allowed to grow overnight at 25 °C. Cells were harvested and re-suspended in binding buffer (20 mM Tris-Cl (pH 7.5), 150 mM NaCl, 10% glycerol, 1 mM PMSF and 6 mM β -mercaptoethanol). Cells were lysed using cell disruptor (20 kpsi) and cell debris was removed by centrifugation at 4 °C and 16000 $\times g$ for 30 min. After cell lysis, DNase I (50 μ g/ml) was added to reduce the viscosity. The clarified supernatant was passed through pre-equilibrated 5 ml StrepTrap HP (GE Healthcare) column. After washing with 10 CV of binding buffer, Cascade/I-C was eluted in binding buffer containing 2.5 mM desthiobiotin. Samples were pooled up and passed through HiTrap Heparin HP column (GE Healthcare) and eluted with a linear gradient of 0.15–2 M NaCl in the binding buffer to remove additional impurities. Later, proteins were further purified through HiLoad 16/600 Superdex 200 prep grade column (GE Healthcare) and eluted in buffer containing 20 mM Tris-Cl (pH 7.5), 150 mM NaCl, 10% glycerol and 6 mM β -mercaptoethanol. Concentrated samples were flash frozen in liquid nitrogen and stored at –80 °C until further use.

ATP hydrolysis assay

ATP hydrolysis reactions were conducted at 37 °C in the reaction buffer containing 20 mM Tris-Cl (pH 8.0), 50 mM NaCl, 10 mM Mg²⁺, 1 mM ATP, 6 mM β -Mercaptoethanol and 1 μ M Cas3/I-C. Hydrolysis reactions were performed in the presence of 10 ng of ssDNA, dsDNA and 16S rRNA. Malachite Green assay was used to determine the liberated inorganic phosphate. Reactions were initiated by

the addition of ATP and stopped by adding malachite green reaction mixture (ammonium molybdate: malachite green in the ratio 1:4). To this mixture, 34% sodium citrate was added to make a final concentration of 3.7%. The volume was made up to 1 ml using water and absorbance was measured at 630 nm. A standard curve between absorbance and inorganic phosphate concentration was established by using varying concentration of KH₂PO₄. All the reactions were performed in triplicates.

Electrophoretic mobility shift assay

Single-stranded DNA constructs (Supplementary Table S1) with or without 6-FAM labels were purchased from Integrated DNA Technologies, Inc (IDT) and gel purified to remove truncated DNA fragments. A 100 bp target DNA construct was generated by annealing the oligonucleotides having 3' complementarity and subjecting them to extension in a PCR using Pfu DNA polymerase. About 50 ng of target DNA was incubated with 500 nM of Cascade/I-C complex in the buffer containing 20 mM Tris-Cl (pH 8.0), 150 mM NaCl and 1 mM DTT at 37 °C for 30–60 min to allow R-loop formation. Post incubation, one set of the sample was treated with 1 mg/ml proteinase K to test the release of target DNA bound to Cascade/I-C. Samples were directly loaded onto 20% (w/v) native polyacrylamide gel and electrophoresed in 1 \times TBE at 4 °C. While 6-FAM labeled constructs were directly visualized under UV in gel documentation system (Bio-Rad), unlabeled constructs were visualized after staining with ethidium bromide (EtBr).

Assay for nuclease activity

In order to assess the nuclease activity of Cas3/I-C in the presence of Cascade/I-C, target DNA and Cascade/I-C were incubated with varied concentration of Cas3/I-C in the buffer containing 20 mM Tris-Cl (pH 8.0), 60 mM NaCl, 1 mM dithiothreitol (DTT), 10 mM MgCl₂ and 1 mM ATP. After the incubation, samples containing unlabeled DNA were directly loaded onto 20% (w/v) native polyacrylamide gel and electrophoresed in 1 \times TBE at 4 °C. DNA bands were visualized in the gel documentation system (Bio-Rad) after staining with EtBr. 6-FAM labeled DNA samples were analyzed on 20% (w/v) denaturing PAGE containing 8 M urea and directly visualized without any post-staining.

All other nuclease assays in the absence of Cascade/I-C were performed in cleavage buffer containing 20 mM Tris-Cl (pH 8.0), 60 mM NaCl, 1 mM dithiothreitol (DTT) and 10 mM MgCl₂. About 500 nM Cas3/I-C was incubated with 100 nM DNA or 50 nM RNA, unless specified otherwise. To check the cleavage in the presence of single-stranded DNA binding (SSB) protein, 5' 6-FAM labeled 60 nt ssDNA was incubated with 0.1–1 μ M of SSB and the reaction was incubated at room temperature to allow SSB to bind to DNA. Later, 500 nM Cas3/I-C along with 1 mM ATP was added in the reaction. To check metal ion dependency, 2 mM of various metal salts and 500 nM Cas3/I-C were used. The reactions were carried out at 37 °C for 10–60 min. Cleavage products were visualized by either 0.8% agarose gel electrophoresis or 10–15% (w/v) 8 M urea denaturing PAGE.

Nuclease activity on bubble DNA construct

Unlabeled single-stranded DNAs (100 nt) having complementary terminal ends were purchased from IDT and PAGE purified to remove truncated DNA fragments (Dloop-AAG-FP and Dloop-TTC-RP; refer Supplementary Table S1). These oligonucleotides (1 μ M) were annealed and incubated with increasing concentration of Cas3/I-C (0.2–1 μ M) in the reaction mixture containing 20 mM Tris-Cl (pH 8.0), 150 mM NaCl, 1 mM DTT, 10 mM MgCl₂ and 1 mM ATP at 37 °C for 120 min. Samples were directly loaded onto 20% (w/v) native polyacrylamide gel and electrophoresed in 1 \times TBE at 4 °C. Gels were post-stained with ethidium bromide (EtBr) and DNA bands were visualized in the gel documentation system.

Nuclease activity on biotin labeled DNA construct

Biotin labeled DNA construct with 5' 6-FAM label was designed and purchased from Bioserve Biotechnologies (India) Pvt. Ltd., and PAGE-purified to remove free 6-FAM and truncated oligonucleotides (ssDNA-biotin; refer Supplementary Table S1). To form double-stranded target DNA, both the strands were annealed by heating at 95 °C followed by gradual cooling. dsDNA or ssDNA was pre-incubated with streptavidin to form a roadblock in the molar ratio 2:3 (DNA: Streptavidin). Wherever mentioned, 500 nM of Cascade/I-C was added to form r-loop. Subsequently, Cas3/I-C was added in the reaction mixture containing 20 mM Tris-Cl (pH 8.0), 150 mM NaCl, 1 mM DTT, 10 mM MgCl₂ and 1 mM ATP/ADP/AMP-PNP. Samples were analyzed on 20% (w/v) denaturing PAGE containing 8 M urea and directly visualized in the gel documentation system.

Assay for helicase activity

Partial DNA duplexes were used to test the helicase activity of Cas3/I-C and its variants. 6-FAM labeled oligonucleotides were obtained from IDT. To generate the construct with 3' overhang, 3' 6-FAM labeled short oligonucleotide (36 nt) and unlabeled long oligonucleotide (70 nt) were annealed (refer Supplementary Table S1). Similarly, in order to generate the construct with 5' overhang, 5' 6-FAM labeled short oligonucleotide (34 nt) and unlabeled long oligonucleotide (70 nt) were annealed (refer Supplementary Table S1). In addition, an excess of an unlabeled short oligonucleotide having an identical sequence as that of 6-FAM labeled was used as trap DNA, in order to avoid the re-annealing of the separated strand. Helicase activity was assessed by mixing Cas3/I-C (0.5–1 μ M) and 6-FAM labeled DNA substrates (0.5 μ M) in the reaction buffer containing 20 mM Tris-Cl, 60 mM NaCl, 6 mM β -mercaptoethanol, 10 mM MgCl₂, 50 μ M Trap-DNA and 1 mM ATP (or other triphosphate variants). The reaction was carried out at 37 °C for 2 h. DNA bands were visualized on 12% (w/v) native PAGE.

Assay for CRISPR interference

E. coli IC-1 harboring pWT-1 and pCRISPR/I-C (vide Supplementary Table S3) was grown in LB broth

supplemented with 25 μ g/ml kanamycin, 50 μ g/ml spectinomycin, 25 μ g/ml chloramphenicol, 0.2% L-arabinose and 0.02 mM IPTG at 37 °C until OD₆₀₀ reached 0.3. Cells were harvested at 4 °C by centrifuging at 2700 \times g. Harvested cells were made chemically competent using 0.1 M calcium chloride and transformed with pT1/pNT1 (see Supplementary Table S3). Cells were grown overnight at 37 °C on LB agar supplemented with 25 μ g/ml kanamycin, 50 μ g/ml spectinomycin, 25 μ g/ml chloramphenicol, 50 μ g/ml ampicillin, 0.2% L-arabinose and 0.1 mM IPTG. The number of colonies obtained was counted, and transformation efficiency was calculated using the equation mentioned below.

$$\text{Transformation efficiency} = \frac{\text{No. of colonies} \times \text{Dilution factor}}{\mu\text{g of DNA used}}$$

Transformation efficiency for Cas3/I-C variants (pHD, pHL, pCas3 Δ CTD, pD48A, pQ253A, pD395A, pK742A, pK743A, pQ745A, pQ746A and pY747A; vide Supplementary Table S3) was calculated using similar protocol.

E. coli IC-2 harboring pWT-2 was cultured in LB broth supplemented with 50 μ g/ml ampicillin and 20 μ M IPTG at 37 °C. Cells were harvested when OD₆₀₀ reached 0.3. Subsequently, harvested cells were made chemically competent as mentioned above. These cells were transformed with pRSF-11/pST-KT (see Supplementary Table S3). Cells were grown overnight in LB agar with 25 μ g/ml kanamycin and 50 μ g/ml ampicillin. The transformation efficiencies for other variants of Cas3/I-C (pD48A, pQ253A, pD395A, pHD, pHL, pHD + pHL, pHD + pHL Δ CTD) and pCas3/I-E were determined as mentioned above.

Fluorescence quenching assay

DNA strand with 5' biotin and 6-FAM at 5th nucleotide (NTS) and Iowa Black FQ (IB) at 29th nucleotide was ordered from IDT. All the DNA constructs were PAGE purified to remove truncated fragments and annealed by heating at 95 °C followed by gradual cooling. About 100 nM of annealed dsDNA was pre-incubated with 200 nM streptavidin and 500 nM Cascade/I-C. Cas3/I-C at increasing concentration (0–5 μ M) was added in the reaction mixture containing 20 mM Tris-Cl (pH 8.0), 150 mM NaCl, 1 mM DTT, 10 mM MgCl₂ and 1 mM ATP/ADP/AMP-PNP, and fluorescence intensity was immediately measured using FluoroMax®-4 (Horiba Scientific). For time-dependent measurement, 500 nM of Cas3/I-C was added in the reaction mixture mentioned above and fluorescence intensity was measured for 120 min.

Anisotropy measurements

Biotin labeled ssDNA construct with 5' 6-FAM as mentioned previously (see Supplementary Table S1) was pre-incubated with streptavidin in the molar ratio 2:3 (DNA: Streptavidin). Cas3/I-C was added in the reaction mixture containing 20 mM Tris-Cl (pH 8.0), 150 mM NaCl, 1 mM DTT, 10 mM MgCl₂ and 1 mM ATP/ADP/AMP-PNP, and anisotropy values were measured using FluoroMax®-4 (Horiba Scientific).

To assess Cas3/I-C binding, anisotropy readings were taken immediately after adding Cas3/I-C (0–10 μ M). For time-dependent measurement of anisotropy, 500 nM of Cas3/I-C was added and readings were noted for 90 min.

Homology modeling

A model of Cas3/I-C was prepared using I-TASSER server (64). Cas3/I-C (BH0336) sequence was uploaded on I-TASSER server, and Cas3/I-C was modeled using Cas3/I-E from *Thermobifida fusca* [PDB ID: 4QQX and 6C66] as template structures. Cas3/I-C from *B. halodurans* and Cas3/I-E from *T. fusca* share 38% sequence identity. The model with the highest C-score was chosen for further analysis.

RESULTS

CRISPR interference requires coordination between nuclease and helicase domains of Cas3/I-C

To test the substrate preference for HD nuclease, different nucleic acid substrates were utilized in the presence of ATP and Mg^{2+} ions (Figure 1 and Supplementary Figure S1). This showed that Cas3/I-C cleaves both double-stranded (pQE2) and single-stranded (M13mp18) DNA proficiently (Figure 1B). Further, the nuclease activity does not differentiate between methylated (pQE2) and unmethylated (PCR amplicon) forms as substrates (Figure 1C). Prompted by the lack of apparent specificity for the target, we questioned whether Cas3/I-C shows any preference for the length of the target. We utilized DNA substrates of various lengths—70 bp, 400 bp and 2.5 kb. We found that while Cas3/I-C is active against 2.5 kb substrate, it is barely active against 400 bp and 70 bp substrates, respectively (Figure 1D). In line with this, though it acts on long RNA, no nuclease activity was discernable with small RNA fragment (150 nt) (Figure 1E and Supplementary Figure S1C,D and F). These suggest that despite the apparent preference for longer DNA, Cas3/I-C exists as a promiscuous nuclease with no apparent intrinsic specificity for the target.

Having characterized the HD nuclease activity, we set out to assess the DExD/H domain-mediated helicase activity using a non-specific DNA substrate. We utilized DNA containing 3' and 5' overhang substrates to assess the directionality of helicase activity (Figure 1F and G). We observed that Cas3/I-C unwinds only that DNA, which possesses a 3' overhang, suggesting that it translocates in 3'-5' direction in the presence of ATP and Mg^{2+} (Figure 1F and G; Supplementary Figure S2A,B). These suggest that Cas3/I-C is a generic 3'-5' helicase with no apparent sequence specificity.

It is not clear whether the nuclease and helicase domains of Cas3/I-C cooperate during interference *in vivo*. To address this, we utilized *E. coli* IYB5101 as a surrogate host (IC-1, Supplementary Table S2) for porting the type I-C system (CRISPR/I-C and Cascade/I-C) from *B. halodurans* (Figure 2A). We introduced the protospacer sequence abutting the targeting PAM sequence (TTC) into pUC19 vector, which acted as target plasmid (T1). On the other

hand, an empty pUC19 vector lacking protospacer was used as the non-target plasmid (NT1). In the absence of Cas3/I-C, the transformation efficiency of both T1 and NT1 was comparable (Figure 2B). When Cas3/I-C was introduced via a plasmid-borne construct, the transformation efficiency of NT1 was retained; however, that of T1 was drastically reduced (Figure 2B). This suggests that the type I-C system is able to discriminate between the target and the non-target substrates *in vivo*. In addition, we observed that CRISPR interference is not active in those strains under the following conditions: (i) Cas3/I-C lacks C-terminal (CTD) domain; (ii) separate nuclease and helicase domains are produced; and (iii) a mutation in HD (D48A) or DExD/H (D395A) domain that inhibits nuclease and helicase activity, respectively (Figure 2B). However, co-expression of stand-alone HD and HL domains results in the functional assembly of interference complex as evidenced by the reduction in transformation efficiency (Figure 2B). This suggests that the interplay between these domains is crucial for CRISPR interference in type I-C system.

A constellation of highly conserved residues in the C-terminal domain of Cas3/I-C that regulates target cleavage

Apart from the two major functional domains—nuclease and helicase—Cas3/I-C has a CTD, whose function is poorly defined (Figure 2D). Intrigued by the fact that CTD is crucial for CRISPR interference (Figure 2B), we set out to probe its role in more detail. Previously determined crystal structure of Cas3/I-E with target DNA [PDB ID: 4QQX] shows that CTD caps the two juxtaposed RecA domains (that are part of the DExD/H domain) such that the target DNA is sandwiched between CTD and RecA domains (Figure 2D). Sequence alignment of Cas3/I-C orthologs showed high conservation of K742, Q745 and Y747 and these conserved residues were absent in Cas3/I-E, suggesting a functional significance for these residues in type I-C (Figure 2C and Supplementary Figure S3A). Since the structure of Cas3/I-C is not determined, we employed the Cas3/I-E structure [PDB ID: 4QQX] as template and made a homology model for Cas3/I-C, which suggested that K742, K743, Q745, Q746 and Y747 that are part of CTD could stabilize the interaction of the target DNA with the helicase core (Figure 2D). Based on this, we hypothesized that CTD may be involved in stabilizing the interaction with target DNA so that Cas3/I-C does not dissociate from DNA during its translocation (Supplementary Figure S4). To test this hypothesis, we performed alanine-scanning mutagenesis of the aforementioned residues and assessed for the nuclease activity. Toward this, we designed a 100 bp target substrate harboring a TTC PAM and performed the assay in the presence of Cascade/I-C and Cas3/I-C (Figure 2E and F). As compared to the wild-type, mutants harboring K743A, Q745A and Y747A were found to be inactive (Figure 2F). The effect of these CTD mutations on CRISPR interference *in vivo* was tested in *E. coli* IC-1. This strain was transformed with pCas3/I-C harboring the desired mutation in the CTD and the transformation efficiency was

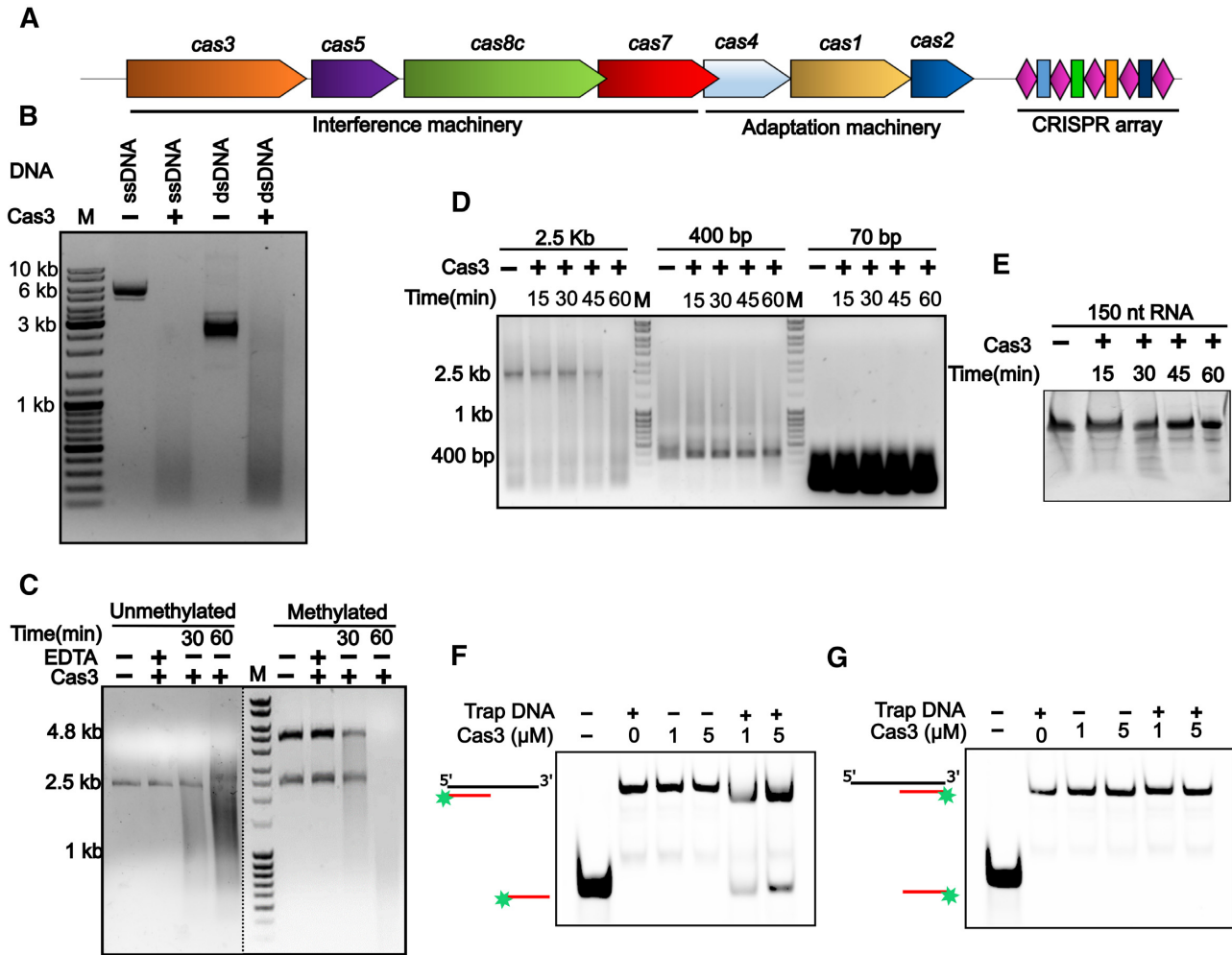


Figure 1. Cas3/I-C cleaves the target DNA indiscriminately. (A) A schematic overview of type I-C CRISPR-Cas locus from *B. halodurans* C-125. Genes are drawn to represent locus approximately to the scale. Effector nuclease (*cas3*), cascade (*cas5*, *cas7* and *cas8*) and adaptation (*cas1*, *cas2* and *cas4*) genes are shown. Repeats are shown as diamonds in magenta and spacers as rectangles in different colors. (B) ssDNA and dsDNA represent single-stranded and double-stranded DNA substrates, respectively. Cas3/I-C cleaves both single (M13mp18) and double (pQE2) stranded DNA in the presence of divalent metal ion (Mg^{2+}) and ATP. DNA marker (M) positions are shown on the left. (C) Methylated and unmethylated DNA fragments are of the same size (2.5 kb). Unmethylated DNA was generated using *Pfu* polymerase that generates blunt-end products. The methylated DNA fragment of 2.5 kb was released from plasmid using restriction enzymes NdeI and KpnI. An additional band of 4.8 kb corresponds to the linearized vector pQE2. DNase activity is not sensitive to methylation. The results were visualized using 0.8% agarose gel electrophoresis. Dotted line indicates a discontinuity in the gel for clarity and M denotes the position of DNA marker. (D and E) DNase/RNase activity depends on the size of the substrate. DNase and RNase activities were assessed on 0.8% agarose and 12% denaturing PAGE, respectively. To test the nuclease activity, 500 nM Cas3/I-C was incubated with 100 nM of 2.5 kb or 400 bp DNA. Similarly, 1.5 μM Cas3/I-C was used against 300 nM of 70 bp DNA. RNA concentration was 50 nM. Cas3/I-C preferred large DNA substrates (2.5 kb), whereas smaller DNA (400 and 70 bp) and RNA (150 nt) substrates were not completely cleaved even after 60 min of incubation. M denotes the position of DNA marker. (F and G) DNA unwinding was tested using fluorescently labeled partial DNA duplex. Displacement of 36 nt was observed when a substrate containing 3' overhang was used. About 15% native PAGE was used to analyze the results.

measured against pUC19 harboring the target sequence. We anticipated that the active Cas3/I-C would target the pUC19 resulting in reduced transformation efficiency whereas the inactive Cas3/I-C would result in target evasion leading to high transformation efficiency. We observed that K743A, Q745A and Y747A in CTD produced high transformants compared to WT, suggesting that these mutations indeed render the CRISPR interference inactive (Figure 2G and Supplementary Figure S7A). Intriguingly, though Q746A showed nuclease activity in the absence and presence of Cascade/I-C *in vitro*, there was no appreciable CRISPR interference *in vivo* (Figure 2F and G).

Resection of short double-stranded linear DNA by Cas3/I-C necessitates the presence of Cascade/I-C

Piqued by the observation that Cas3/I-C lacks intrinsic target specificity *in vitro* and retains one *in vivo*, we set out to resolve this apparent paradox. We hypothesized that the absence of 'intrinsic' target specificity could be compensated by a mechanism that introduces 'induced' target specificity. To test our hypothesis, we designed a 100 bp DNA substrate with 'TTC' PAM at the 5' end (Figure 2E). When Cas3/I-C alone was introduced, no apparent cleavage was noted (Figure 3A and B). This was surprising; however, it was consistent with our earlier observation

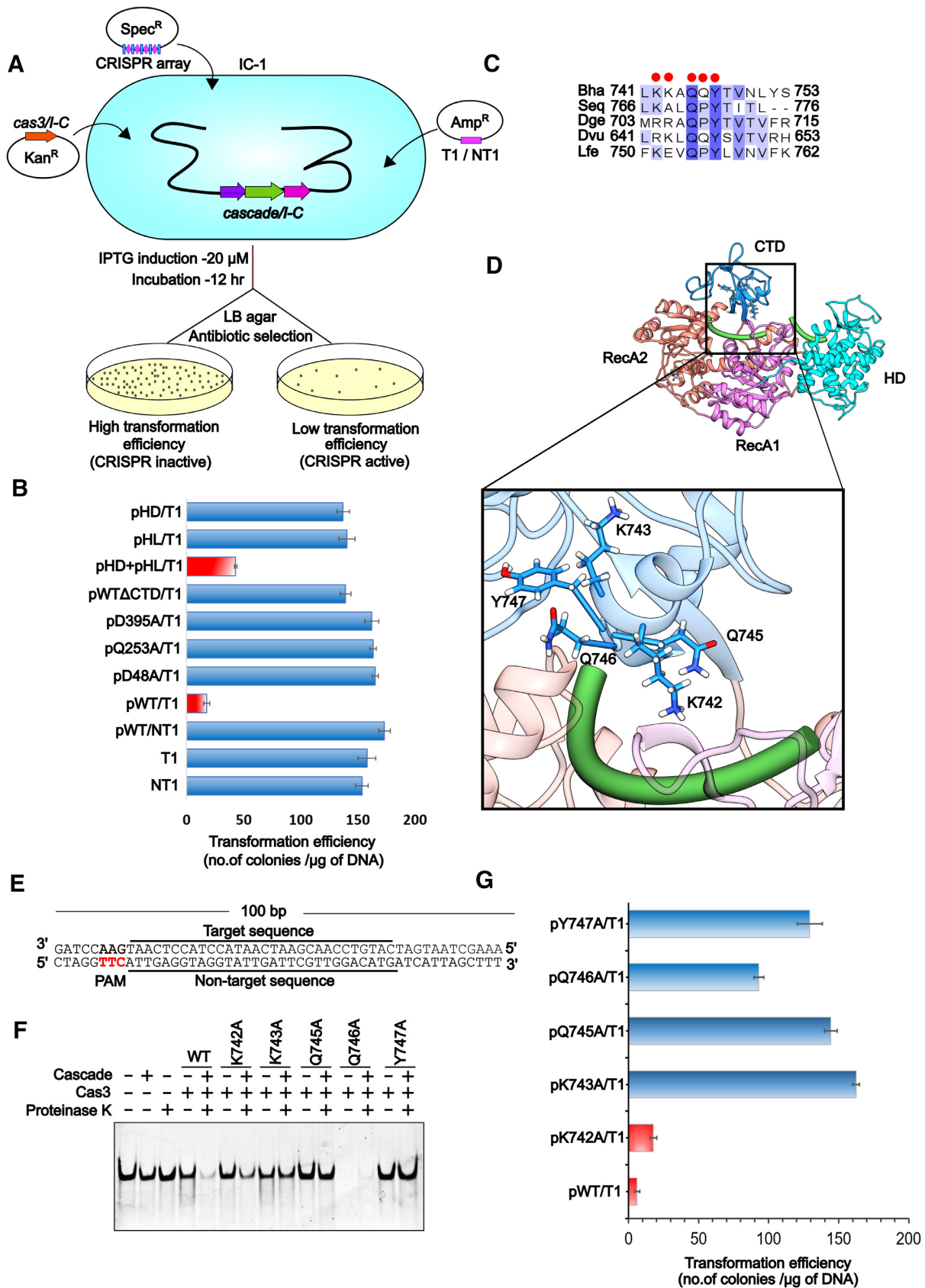


Figure 2. Inter-domain interaction in Cas3/I-C influences CRISPR interference. (A) An *E. coli* IYB5101 strain harboring genes encoding Cascade/I-C (IC-1) was used as a surrogate for *in vivo* experiments. Cas3/I-C and crRNA were expressed through IPTG inducible vectors having kanamycin and

(Figure 1D) that nuclease activity is attenuated for short DNA constructs (<400 bp). Led by the requirement of Cascade/I-C for CRISPR interference *in vivo*, we introduced the Cascade/I-C into the reaction. Remarkably, the target DNA—which was initially refractory to nuclease activity of Cas3/I-C—was cleaved in the presence of Cascade/I-C (Figure 3A and Supplementary Figure S5E). Intrigued by this, we asked what was bestowed by the Cascade/I-C on Cas3/I-C to target the DNA? Based on previous reports in type I-E (22,33,34), we interpreted that the Cascade/I-C binding to target DNA ensues the generation of a DNA–RNA hybrid and a single-stranded DNA referred to as R-loop. Since Cas3/I-C requires 3'-overhang for loading (Figure 1F and Supplementary Figure S2), we hypothesized that the duplex nature of the target DNA perhaps abrogates the DNA binding. Therefore, it is likely that the ssDNA that was formed as a consequence of R-loop could then provide a platform for Cas3/I-C loading (Figure 3A and B). To test this, we designed a 100 bp bubble DNA construct that has no base pairing in the central region and that mimics the R-loop in the absence of Cascade/I-C (Figure 3C). Remarkably, we found that Cas3/I-C sliced the bubble DNA even in the absence of Cascade/I-C, suggesting that it is indeed the single-stranded region that is becoming the loading point for Cas3/I-C (Figure 3D and Supplementary Figure S5F). Additionally, we also found the presence of small DNA fragment (~75–85 nt) when ATP was not included in the reaction that hints at the ATP independent nicking of the target DNA (Supplementary Figure S5F)

Cas3/I-C cleaves target via a reeling mechanism

The interaction between Cascade/I-C and Cas3/I-C facilitates the target DNA cleavage. After recruitment by Cascade/I-C, the helicase activity of Cas3/I-C may allow it to translocate on the ssDNA, unwinding the target DNA in 3' to 5' direction followed by cleavage via nuclease domain, eventually leading to the dissociation from Cascade/I-C. An alternative possibility is to remain associated with Cascade/I-C and reel in the target DNA, which may result in the formation of a loop in the target strand. In order to test, how Cas3/I-C unwinds double-stranded target DNA upon recruitment by Cascade/I-C,

we designed a 100 bp dsDNA with a fluorophore (6-FAM) and a quencher (Iowa Black FQ). Non-target strand (NTS) was tagged with biotin at 5' end and 6-FAM at 5th nucleotide from the 5' end. On target strand (TS), Iowa Black FQ (IB) was introduced proximal to the PAM at 29th nucleotide from 3' end (Figure 4A). Biotin was introduced at 5' end with the presumption that Cas3/I-C would halt on encountering the biotin-streptavidin block thus sustaining 6-FAM quenching by Iowa Black FQ for a sufficiently long duration for measurements. Since 6-FAM and IB were far apart, the fluorescence of 6-FAM was not quenched. However, if Cas3/I-C reeled in the DNA, 6-FAM would come in proximity with IB, which in turn would quench the fluorescence. Target DNA was saturated with streptavidin and Cascade/I-C before adding Cas3/I-C, and fluorescence intensity was measured to estimate the extent of quenching. Interestingly, we observed significant fluorescence quenching with increasing Cas3/I-C concentrations (0–5 μ M) when Cascade/I-C was present. On the other hand, quenching was not observed when the reaction did not contain Cascade/I-C (dsDNA) or only non-target strand (ssDNA) was present (Figure 4B and Supplementary Figure S6F). Moreover, fluorescence quenching was observed only when ATP was supplied to Cas3/I-C in addition to Cascade/I-C (Figure 4C). Taken together, our result suggests that Cas3/I-C remain bound to Cascade/I-C and pulls in 6-FAM towards IB, which leads to fluorescence quenching.

Stalling the helicase motor of Cas3/I-C stimulates the nuclease activity

Bacterial cytoplasm holds a repertoire of DNA-binding proteins and therefore, Cas3/I-C during its processive nuclease–helicase activity is expected to encounter proteins such as single-stranded binding protein (SSB), polymerases etc. In order to decipher what would happen if Cas3/I-C encounters such roadblocks, we intentionally introduced a stalling site in the target DNA. The stalling site comprises a biotin labeled nucleotide that binds to Streptavidin, which is expected to block the movement of Cas3/I-C helicase motor. We labeled the 5' end of the non-target strand (NTS) with FAM and 12th nucleotide from 5' end with biotin (Figure 5A). Prior to the addition of Cas3/I-C, target DNA

spectinomycin antibiotic resistance markers, respectively. The target sequence was inserted into pUC19 vector (T1), whereas an empty pUC19 vector was used as non-target (NT1). Low transformation efficiency indicates functional CRISPR interference. (B) Transformation efficiencies of the strain IC-1 against the target (T1) and the non-target (NT1) for wild-type and variants of Cas3/I-C are shown. In the presence of Cas3/I-C wild-type (pWT) and co-expressed construct of HD and DExD/H domain (pHD+pHL), there was a significant reduction in transformation efficiency (indicated by a red bar). Interference was rendered ineffective when the HD domain mutant (pD48A), Helicase domain mutant (pQ253A and pD395A), CTD deletion mutant (pWT Δ CTD), stand-alone nuclease domain (pHD) and stand-alone helicase domain (pHL) were used. Error bar represents standard deviation measured from three independent trials. The boundary of the domain variants are as follows: HD domain (pHD) 1–248; DExD/H domain (pHL) 249–800; WT without CTD (pWT Δ CTD) 1–709. (C) Highly conserved residues in the CTD domain from type I-C system are shown. Red dots indicate those residues from *B. halodurans* that are chosen for alanine-scanning mutagenesis. (D) Homology model of Cas3/I-C from *B. halodurans* based on crystal structure of Cas3/I-E [PDB ID: 4QQX] shows the organization of nuclease (HD in cyan), DExD/H helicase (RecA1 in pink and RecA2 in orange) and CTD (in blue) domains. DNA is shown as a rod in green. Amino acids (K742, K743, Q745, Q746 and Y747) from CTD domain that are possibly interacting with target DNA are shown. (E) A 100 bp target DNA harboring functional PAM sequence TTC is depicted. Target sequence represents the strand, which is complementary to crRNA. (F) CTD point mutants were tested for their nuclease activity. Cascade/I-C was incubated with target DNA in the reaction mixture containing Mg²⁺ and ATP to form R-loop. Wild-type or CTD mutant was added to the reaction and the cleavage was monitored. Proteinase K was added to release residual Cascade bound DNA fraction. DNA was analyzed using 20% native PAGE. (G) CTD mutants were tested for their role in CRISPR interference using *E. coli* IC-1 strain as the surrogate host. Wild-type and CTD mutants were expressed through an IPTG inducible 1R vector (referred to as pWT, pK742A and so on). Transformation efficiency was measured against pUC19 harboring the target sequence with functional PAM TTC (referred to as pT1). Error bar represents standard deviation measured from three independent trials.

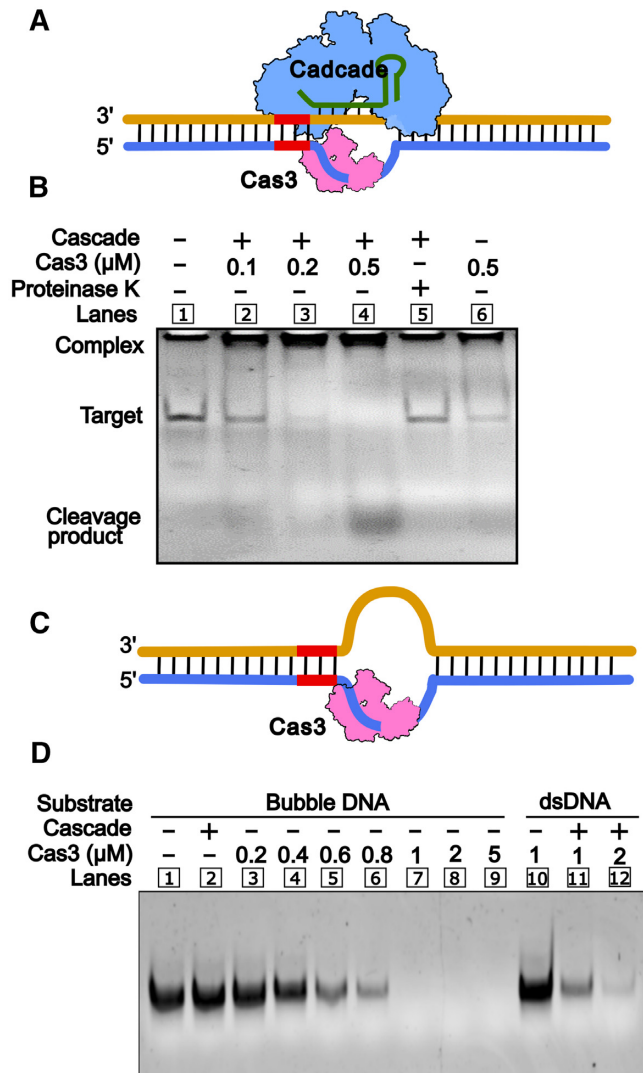


Figure 3. Cascade/I-C provides single-stranded DNA loop for Cas3/I-C binding. (A) A schematic model that suggests that Cascade/I-C binding to target DNA provides an ssDNA platform for Cas3 binding. The location of PAM (TTC in red) in the 100 bp construct is indicated. (B) DNA cleavage was tested on 20% native PAGE with increasing concentration of Cas3/I-C in the presence of Cascade (lanes 2–4). Proteinase K was added to release bound DNA (lane 5). No cleavage was noticed in the absence of Cascade (lane 6). (C) A 100 bp bubble DNA construct that lacks base pairing in the center was designed with PAM (TTC in red). The single-stranded DNA loop is suggested to facilitate Cas3/I-C binding in the absence of Cascade/I-C. (D) Bubble DNA was cleaved in the absence of Cascade/I-C (lanes 3–9), whereas dsDNA was cleaved only in the presence of Cascade/I-C (lanes 11–12). DNA was visualized using 20% native PAGE.

was incubated with Streptavidin and Cascade/I-C. Cas3/I-C was introduced with ATP/ADP/AMP-PNP and cleavage was monitored using denaturing PAGE (Figure 5B). We could discern the accumulation of a prominent ~ 55 – 60 nt band in the absence of ATP (lanes 10, 11, 13, 14 and 16 in Figure 5B). Since helicase motor is inactive in absence of ATP, we presume it to be a single-stranded nick generated by Cas3/I-C nuclease upon binding to NTS. Apart from the above observation, the presence of cleavage product (~ 55 – 60 nt) in the absence of nucleotide suggests that Cas3/I-C

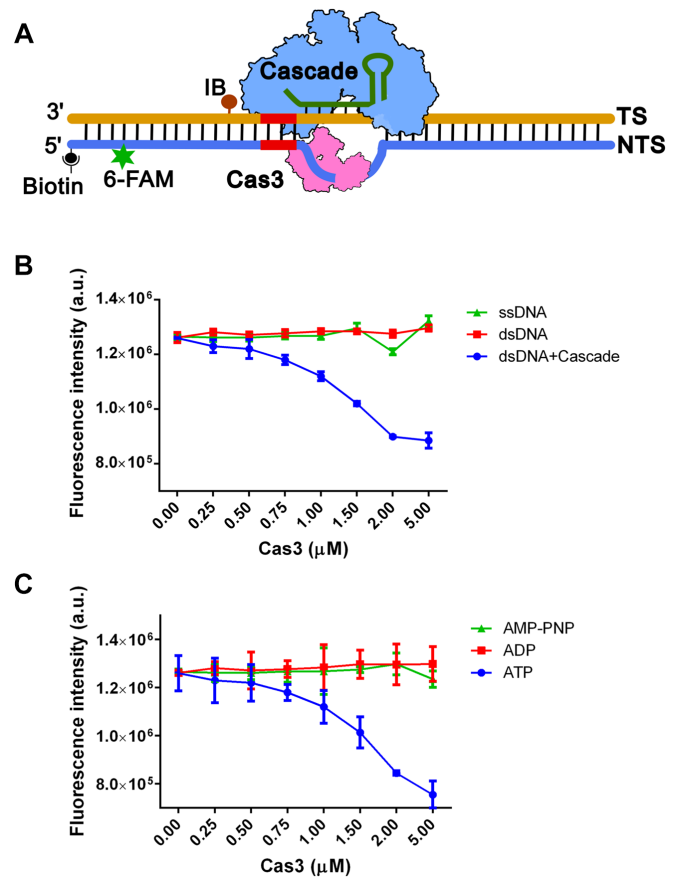


Figure 4. Cascade/I-C and Cas3/I-C form a stable complex during interference. (A) Schematic representation of 100 bp DNA substrates with biotin (black dot) at 5' end of the non-target strand (NTS), 6-FAM (shown as green star) at 5th nucleotide, Iowa Black[®] FQ (shown as brown dot) at 28th nucleotide on target strand (TS) and PAM sequence TTC (depicted in red). (B) Substrate mentioned above was incubated with or without Cascade/I-C (1 μM) and increasing concentration of Cas3/I-C (0–5 μM). A significant decline in fluorescence intensity is evident when both Cascade/I-C and Cas3/I-C were present. There was no apparent quenching when dsDNA and ssDNA were used in the absence of Cascade/I-C. (C) Fluorescence quenching was observed in the presence of ATP but not when ADP and AMP-PNP were used.

induced initial nick is independent of the nucleotide-bound state (lane 16 in Figure 5B). Interestingly, in the presence of ATP, we observed a prominent band at ~ 40 nt; however, the band corresponding to ~ 55 – 60 nt was not perceptible (lanes 7 and 8 in Figure 5B). On the contrary, ~ 40 nt band was not visible when the biotin-streptavidin block was absent, even in the presence of ATP; however, we could spot the ~ 55 – 60 nt band (lane 18 in Figure 5B). This suggests that Cas3/I-C introduces a nick in the single-stranded region of NTS irrespective of the nucleotide-bound state.

To account for the appearance of ~ 40 nt band specifically in the presence of ATP and biotin-streptavidin block, we hypothesized that the helicase motor of Cas3/I-C, in the presence of Cascade/I-C and ATP, would have reeled in the NTS from the loading point and a biotin-streptavidin roadblock could have triggered cleavage by the trailing nuclease domain. To test our hypothesis, we intentionally introduced similar stalling site in single-stranded DNA

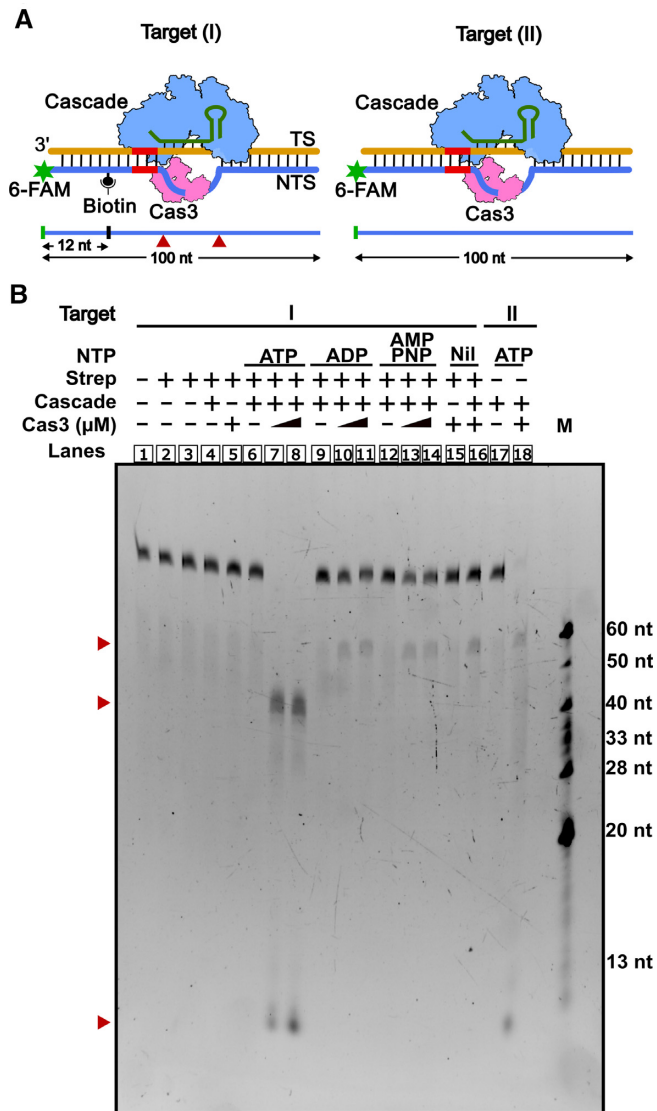


Figure 5. Stalling the translocation of helicase motor stimulates cleavage. (A) Schematic representation of target DNA (100 bp) with PAM (TTC in red) end-labeled with 6-FAM at 5' end of the non-target strand and biotinylated at 12th nucleotide (Target I). A similar target DNA without biotin (Target II) is also represented. Indicated in red triangle are the prominent cleavage sites. (B) Target (I) and (II) were incubated with streptavidin and Cascade/I-C to form interference complex. Cleavage was initiated by addition of Cas3 (0.2 and 0.5 μM) and 1 mM ATP/ADP/AMP-PNP. Prominent cleavages of the target (I) and (II) are indicated by a red triangle (~60, 40 and 10 nt). The size of the DNA markers is shown in the lane M. A 20% denaturing PAGE was used to assess the cleavage of DNA.

labeled with 6-FAM at 5'-end. Two constructs were made in which the biotin was labeled at 12th (Target A) and 20th (Target B) nucleotide positions, respectively (Figure 6A and D). We measured the change in fluorescence anisotropy of above-mentioned substrates upon addition of Cas3/I-C (Figure 6B, E and Supplementary Figure S6C–E). Initially, we observed a slight increase in anisotropy for biotin-labeled DNA, suggesting early binding of Cas3/I-C to ssDNA. With time, we could discern a significant decline in anisotropy value in the absence or when ATP/ADP/AMP-

PNP was present; however, the decay was steep when ATP was present. Remarkably, in the absence of biotin the anisotropy decay is not steep (Figure 6B and E). This decrease in anisotropy value with time suggests that DNA is fragmented possibly due to Cas3/I-C induced cleavage. In order to understand such cleavage, Cas3/I-C was mixed with ATP/ADP/AMP-PNP and bands were analyzed using denaturing PAGE (Figure 6C and F). In line with above observations, two prominent bands at the top corresponding to ~30 and 40 nt for Target A (lanes 4, 5, 7, 8 and 11 in Figure 6C) and ~45 and 55 nt for Target B (lanes 4, 5, 7, 8 and 11 in Figure 6F), respectively, were perceptible in presence of ATP/ADP/AMP-PNP. For Target A, we posit that the ~30 nt band could arise by cleavage due to biotin block and the other ~40 nt could emerge by cleavage when a second Cas3/I-C gets stalled by the preceding Cas3/I-C. A similar scenario can be envisaged for Target B as well. Whereas in the absence of a biotin-streptavidin block, we could not observe any cleavage (Supplementary Figure S6E). Interestingly, we also observed small sized bands (<10 nt), which suggests that few Cas3/I-C could trespass the biotin-streptavidin block and possibly get stalled by 6-FAM at the 5' end (Figure 6C and F). Altogether the data suggest that when the helicase motor of Cas3/I-C encounters roadblock en route, the nuclease activity associated with HD domain is stimulated.

Evolutionary conservation of target recognition and cleavage mechanism between type I-C and type I-E

Buoyed by the fact that it is the generation of single-stranded DNA as a by-product of R-loop that provides a platform for Cas3/I-C loading and target cleavage, we hypothesized that such mechanism could be pervasive among type I system irrespective of the differences in composition and architecture of Cascade. To test this, we chose Cascade/I-E—which is made up of five Cas subunits—as against Cascade/I-C, which comprises of three subunits only. Despite these differences, both Cascades facilitate the formation of R-loop and therefore, we anticipated that in such scenario it is possible to complement the Cascade/I-E with Cas3/I-C in order to promote the target cleavage *in vivo*. To test this, we created an *E. coli* strain (IC-2, Supplementary Table S2) that harbors a spacer that targets pRSF-11 (T2, Supplementary Table S3) and Cascade/I-E. In IC-2, the innate Cas3/I-E was deleted and therefore the intrinsic CRISPR/I-E machinery can recognize the target but it is deficient in target cleavage (Figure 7A). When IC-2 was challenged with T2, more number of transformants were observed (Figure 7B), suggesting that target cleavage is impaired in the absence of Cas3/I-E. When strain IC-2 was complemented with plasmid-borne Cas3/I-E (pCas3/I-E), there was a drastic reduction in transformation efficiency, which suggests restoration of functional CRISPR interference.

Interestingly, similar phenotype was noted when plasmid-borne Cas3/I-C (pCas3/I-C) complemented IC-2, suggesting that T2 is efficiently targeted by Cas3/I-C (Figure 7B and Supplementary Figure S7B). To probe this further, the HD (pHD) and DExD/H (pHL) domains

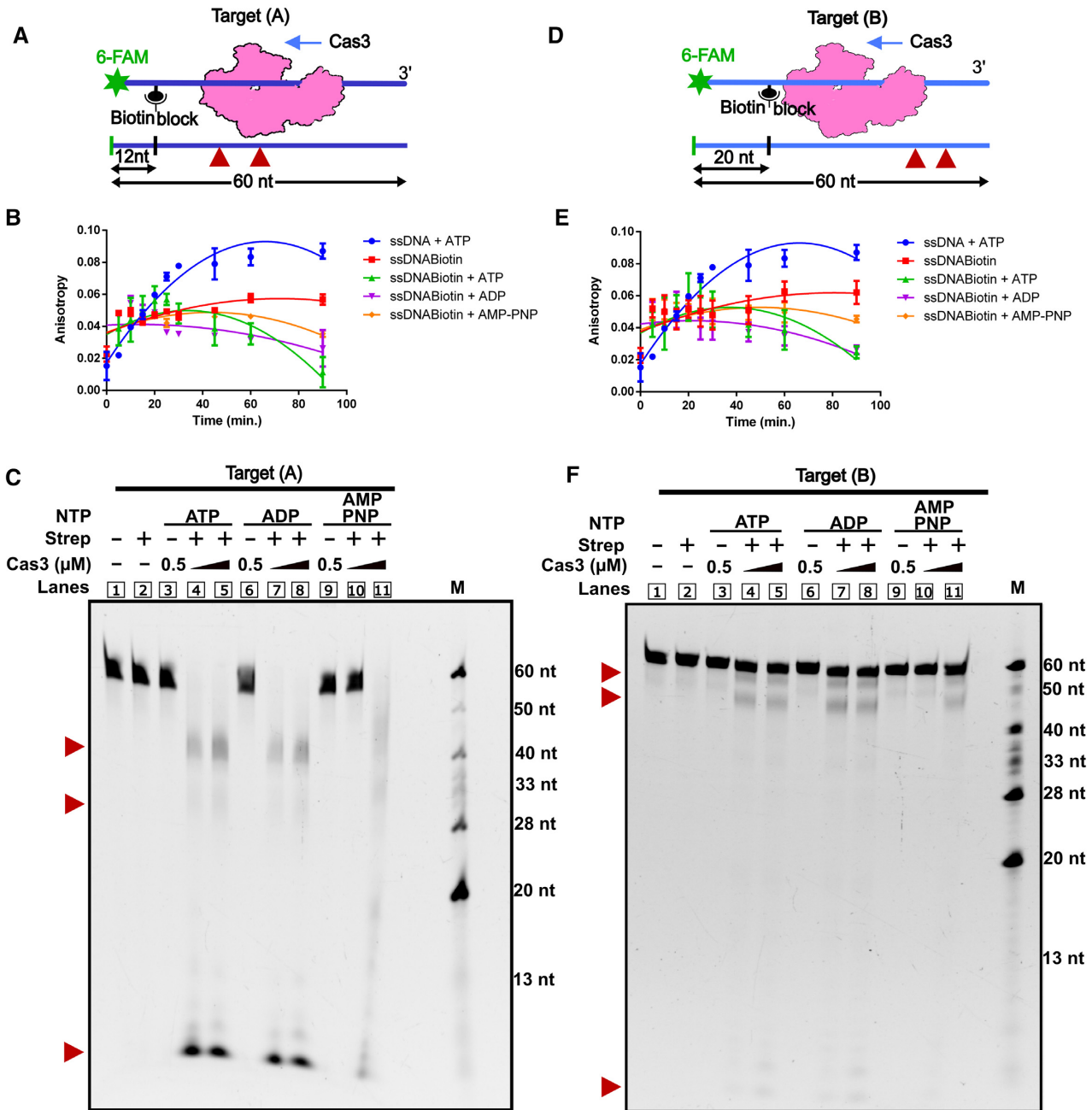


Figure 6. Roadblock in the translocation of Cas3/I-C stimulates cleavage (A). (D) A schematic representation of 60 nt 5' 6-FAM labeled ssDNA with biotin at 12th (Target A) and 20th nucleotide (Target B), respectively is shown. (B) and (E) ssDNA mentioned above was incubated with 200 nM of Cas3/I-C for several time points, and anisotropy measurements were recorded. With time, the decrease in anisotropy values was observed for ssDNA with a biotin roadblock. (C) and (F) ssDNA was pre-incubated with streptavidin before the addition of Cas3/I-C. Prominent DNA cleavage products were observed in the presence of ATP at ~40 nt in target A and ~50 nt in target B, indicated with a red triangle. In the presence of AMP-PNP, higher Cas3 concentration was required for the cleavage. The size of the DNA markers is shown in the lane M. A 20% denaturing PAGE was used to assess the cleavage of DNA.

of Cas3/I-C were co-expressed as independent domains for complementation *in vivo*. Remarkably, this strain showed less transformation efficiency that is comparable to Cas3/I-C WT, suggesting that these domains interact and produce functional Cas3/I-C *in vivo* during CRISPR interference. It is to be noted that HD and DExD/H domains, when expressed separately, failed to restore CRISPR interference and it was only when they were

co-expressed, T2 was targeted (Figure 7B). In line with this, pCas3/I-C, harboring point mutation in lieu of catalytic residue in HD (pD48A) or Helicase domain (pQ253A and pD395A) is found to be inactive (Figure 7B). This suggests that despite the difference in the architecture of Cascade/I-C and Cascade/I-E, the Cas3-mediated target cleavage mechanism seems to be conserved between type I-C and type I-E.

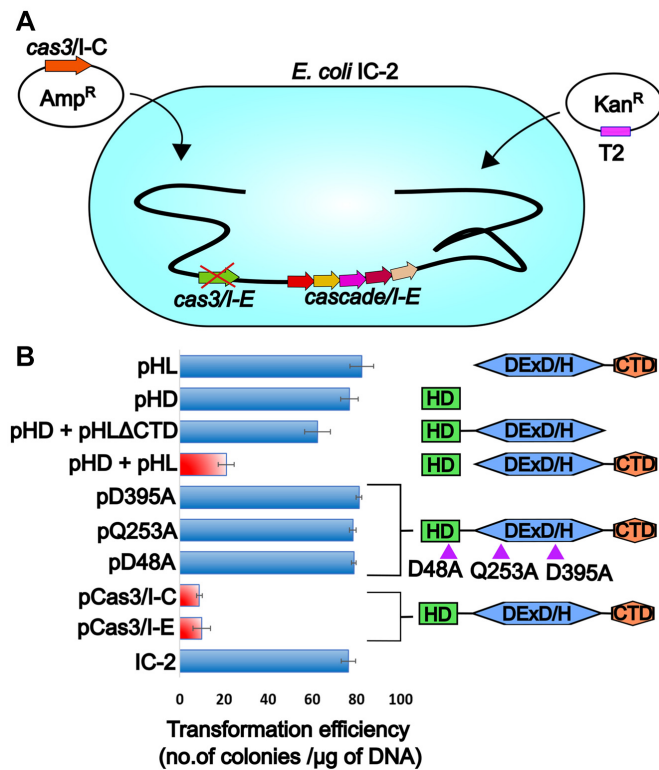


Figure 7. Cas3/I-C supplants Cas3/I-E in type I-E CRISPR-Cas system (A) *E. coli* K-12 harboring Cas3/I-E null mutant (IC-2) was used. IC-2 harbors spacer sequence targeting plasmid pRSF-11 (pT2). Cas3/I-C was heterologously expressed through IPTG inducible vector having ampicillin selection marker. Transformation efficiency was calculated after induction with 20 μM IPTG and 12 h incubation. (B) Cas3/I-C and its variants were tested for interference activity using *E. coli* IC-2. Heterologous expression of Cas3/I-C (pCas3/I-C) and Cas3/I-E (pCas3/I-E) showed reduced transformation efficiency (indicated as a red bar). When the HD (pHD) and helicase (pHL) domains of Cas3/I-C were co-expressed as two separate constructs, the transformation efficiency mirrored that of wild-type (indicated as a red bar). However, HD nuclease mutant (pD48A), helicase mutants (pQ253A and pD395A) and co-expressed constructs of HD and helicase without CTD (pHD + pHLΔCTD), stand-alone nuclease domain (pHD) and stand-alone helicase domain (pHL) showed high transformation efficiency. Error bar represents standard deviation measured from three independent trials. Purple triangles represent the location of the point mutations. The boundary of the domain variants are as follows: HD domain (pHD) 1–248; DExD/H domain (pHL) 249–800; pHD+pHLΔCTD 1–709.

DISCUSSION

During CRISPR interference, precise recognition of target forms the cornerstone of self versus non-self-recognition. In Class II, a multi-domain single protein, i.e. Cas9 in type II-A, executes both target recognition and annihilation. However, in Class I/type I, Cascade recognizes the target, whereas Cas3 cleaves the target. Unlike Class II, this division of labor between Cascade and Cas3 in Class I/type I suggests the possibility of additional regulatory control to the CRISPR-based genome defense and therefore alludes to the prevalence of type I among the prokaryotes. In this context, one could perceive that Cascade and Cas3 should have ‘intrinsic’ target specificity to obviate any collateral damage on the host. In the case of Cascade, this

is satisfied by complementary base pairing between crRNA and the target DNA. In line with this, our data suggest that the differences in the architecture and composition of Cascade among type I are unlikely to wield high influence in determining the specific recognition of the target DNA—a role that is exquisitely delegated to crRNA (Figure 7). On the contrary, we noted that the nuclease activity of Cas3 in type I is highly promiscuous (Figure 1)—which is counter-intuitive—and this has an intrinsic potential for indiscriminate host-targeting (44,45,65). Intriguingly, in type I variants such as type I-C and type I-E, though the composition and architecture of Cascade seem to differ, the overall architecture of Cas3 is broadly conserved, suggesting that Cas3 recognizes the target DNA via a ‘generic’ mechanism and it is likely that this is independent of the structure of Cascade. Piqued by this, we set out to unravel the underlying mechanism by which Cas3 selectively degrades the target DNA without inflicting collateral damage.

In Cas3/I-C, the requirement of HD, DExD/H and CTD domains during CRISPR interference suggests that the concerted domain–domain crosstalk plays a pivotal role in regulating the nuclease activity (Figures 2 and 7; Supplementary Figure S7). Strikingly, both HD and DExD/H domains require single-stranded DNA substrates, which are transiently formed during highly orchestrated events such as replication and transcription. Therefore, Cas3 is unlikely to have access to its preferred single-stranded substrates most part of the time and this may preclude any inadvertent self-targeting. The fact that unlike plasmid DNA, small double-stranded DNA (<400 bp) is not cleaved, strengthens this conjecture (Figure 1). Due to the typical *in vitro* setting by which plasmid DNA is isolated by employing denaturation and renaturation cycles as well as given its large size, it is likely that some regions retain single-stranded state even after renaturation. Moreover, pQE2 is found to be negatively supercoiled, which may also produce a single-stranded region. These single-stranded regions could become the substrate for Cas3 mediated nicking and its subsequent loading points. Whereas shorter DNA fragment is likely to have efficient renaturation and this can pre-empt Cas3 loading. Also, this scenario is likely to get reversed when the Cascade mediated recognition of the foreign DNA leads to the formation of R-loop and this sets the stage for Cas3 to access the single-stranded substrate. Our data suggest that the recruitment of Cas3 toward R-loop is not solely contingent on Cascade, which is in part driven by the single-stranded region (Figure 3C and D).

The R-loop has no free 3′ end for helicase loading and therefore threading the ssDNA into the helicase domain creates a topological problem. This can be overcome if free 3′ end is produced by the nuclease activity of Cas3/I-C. Nonetheless, the homology model of Cas3/I-C based on the crystal structure of Cas3/I-E suggests that the active site of HD nuclease is oriented toward the helicase domain and therefore it is unlikely to access the ssDNA directly. Consequently, it appears that the ssDNA has to enter the HD active site via the helicase domain only. However, our data suggest that upon recruitment to the target site, Cas3/I-C nicks non-target strand at ~30–35 nt

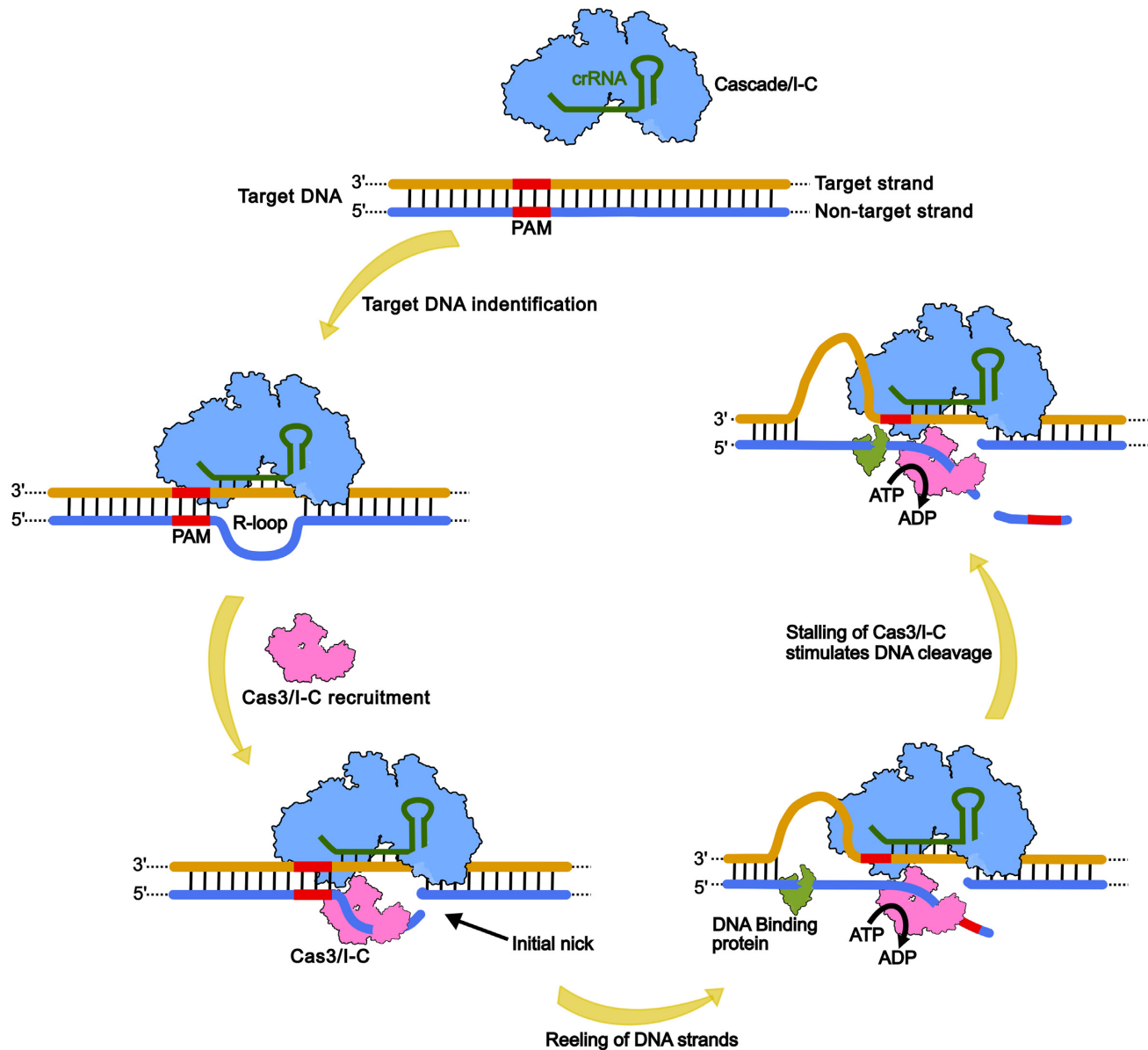


Figure 8. A schematic model representing the possible scenarios leading to Cas3/I-C mediated target DNA cleavage. Cascade/I-C (shown in blue) through base pairing between crRNA (indicated as stem and loop structure in green) and target DNA locates the invading foreign genetic element, thereby forming an R-loop. The ensuing conformational change in the large subunit of Cascade/I-C facilitates its binding to Cas3/I-C. This is followed by nucleotide independent single-stranded nick on the non-target strand by HD domain. This promotes the loading of the helicase domain of Cas3/I-C onto the ssDNA of the non-target strand. Based on the longevity of Cascade/I-C–Cas3/I-C interaction, we propose that Cas3/I-C remains attached to Cascade/I-C as a stable complex and reels in the non-target strand (this leads to DNA looping on target strand), utilizing energy from ATP. Subsequently, Cas3/I-C may hit a roadblock (e.g. a DNA-binding protein), which can trigger further DNA cleavage. Here, the cleavage site is likely to be in proximity to the Cascade/I-C.

from PAM (Figure 5). This initial nick in the absence of ATP points toward the direct binding of nuclease domain at PAM distal region of R-loop, which then renders a loading point for the helicase domain (50). This also concurs with the interpretation from the recent Cryo-EM structure of Cascade/I-E and Cas3/I-E (47). Our data further suggest that though Cas3/I-C lacks DNA unwinding activity in the absence of ATP, it can translocate on the ssDNA (Figure 6). In the presence of ATP, it seems that Cas3/I-C unwinds the target DNA as well as channelizes the displaced ssDNA toward the active site of the HD nuclease

through the helicase motor activity. This implies that in the absence of the helicase motor activity, the DNA is unlikely to reach the HD nuclease. This is in line with the current understanding on Cas3/I-E mediated cleavage and explains the requirement of the helicase domain for CRISPR interference (Figures 2 and 5) (36,47). We envisage another possible scenario wherein despite the helicase motor and HD nuclease being active, mutations in the channel forming residues are likely to derail the conduit so that the displaced ssDNA is unlikely to be threaded toward the active site of the HD nuclease resulting in the loss

of target cleavage. Previous studies on Yeast DEAD Box Protein Mss116p demonstrates the modulation of helicase core activity by CTD (66,67). Similarly, we identified a few critical residues in the CTD of Cas3/I-C—which appears to be unique with respect to type I-C—and that possibly stabilizes the DNA-bound helicase core thus bolting the channel for DNA translocation (Figure 2; Supplementary Figures S2E and S4).

Presumably, we posit that the incessant movement of DNA toward the HD nuclease would not provide enough residence time for the alignment of catalytic residues. This is likely to impede the nuclease activity as long as the helicase motor is active (Figures 4 and 5; Supplementary Figures S4B, S6E and S6G). Therefore, whenever there is a blockade for helicase movement, it is likely that the catalytic residues of HD nuclease have sufficient residence time to be suitably oriented toward the scissile bond (Figures 4 and 5; Supplementary Figure S4B). This appears to be an ingenious strategy to time the cleavage. Our data on Cas3/I-C are in agreement with the recent finding where stalling of Cas3/I-E translocation at roadblocks stimulates nuclease activity that is otherwise reported to be sparse (50,68). After the Cas3/I-C loading onto the single-stranded region, we envisage two modes of helicase motor activity: (i) translocation—Cas3/I-C is mobile with respect to the bound ssDNA and (ii) reeling—Cas3/I-C is stationary with respect to the bound ssDNA. However, in line with the previous report on Cas3/I-E (50), our data showed the cleavage pattern that concurs with a DNA reeling mechanism wherein Cas3/I-C has an intimate interaction with Cascade/I-C during DNA unwinding (Figure 4 and Supplementary Figure S6F). In this set up—where the mobility of the first Cas3/I-C molecule is locked due to interaction with Cascade/I-C—a caravan consisting of multiple Cas3/I-C that can load onto the R-loop is likely to collide with each other as well as with the one that interacts with Cascade/I-C to inflict multiple cleavages (Figure 6). While the rationale for adopting reeling over translocation is not abundantly clear, it is tempting to speculate that reeling allows Cas3/I-C to inflict cleavage that is proximal to the interference effector complex and endows with an apparent target site specificity (Figure 8).

The fact that Cas3/I-C—albeit exhibiting differences such as the presence of highly conserved residues in CTD—shows functional similarities with respect to Cas3/I-E has raised a question how this functional equivalence could exist in harmony with Cascade/I-C—which exhibits compositional and architectural differences with Cascade/I-E. Strikingly, the complementation of Cas3/I-C in lieu of Cas3/I-E in type I-E system *in vivo* shows that the target recognition and cleavage mechanism are conserved between type I-C and I-E (Figure 7). When we compared the homology model of Cas3/I-C from *B. halodurans* (BhaCas3/I-C) (Figure 2D and Supplementary figure S3C) and the homology model of Cas3/I-E from *E. coli* (EcoCas3/I-E)—both are modeled on the basis of the recent cryo-EM structure of Cas3/I-E from *Thermobifida fusca* (TfuCas3/I-E) (36)—we could discern that BhaCas3/I-C shows significant sequence and structural similarities with EcoCas3/I-E, if not complete,

in interfaces I–IV (Supplementary Figures S3C and S8). Based on this analysis, we attribute the reported incompatibility for hetero-assembly between TfuCas3/I-E and EcoCascade/I-E (47) to the following factors: (i) the reconstitution of hetero-assembly was attempted *in vitro*. Therefore, it is quite likely that the lack of cellular milieu could be a possible deterrent for hetero-assembly. (ii) TfuCas3/I-E and EcoCascade/I-E are not thermo-compatible. The former is thermophilic and the latter is mesophilic. They have undergone different adaptation attuned with the thermal conditions. (iii) In our study, types I-C and I-E are not co-existing as composite CRISPR-Cas system within the same organism. This precludes any selection pressure to avert cross-over Cascade–Cas3 interaction between type I-C and I-E. Thus, our data suggest that despite the compositional and architectural differences seen between Cascade/I-C and Cascade/I-E, Cas3 recruitment and target cleavage mechanism seem to be conserved, at least in I-C and I-E systems, if not all type I.

CONCLUDING REMARKS

It is quite remarkable to note that Cas3/I-C—albeit being a promiscuous nuclease—is induced toward achieving its target specificity after it associates with Cascade/I-C by segregating its nuclease and helicase activities in space and time. It seems that instead of producing a new effector nuclease with intrinsic target specificity, type I CRISPR-Cas system has adopted an existing generic helicase-cum-nuclease module and further evolved ways and means to nudge its unrestrained activity toward specifically targeting the foreign DNA.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

Vector pOSIP-CT (Addgene #45981) was a kind gift from Dr Drew Endy and Dr Keith Shearwin; pST-KT was gifted by Dr Vinay K. Nandicoori; p1R (Addgene #29664), and p13SR (Addgene #48328) were provided by Dr Scott Gradia; pKD46 (CGSC #7739) was a kind gift from Dr Barry L. Wanner; *E. coli* IYB5101 was a gift from Dr Udi Qimron. We acknowledge the geniality of the aforementioned scientists for sharing their plasmids and bacterial strains. We thank all members of the MAB lab for their critical comments and suggestions.

FUNDING

Department of Biotechnology (DBT) [BT/PR15925/NER/95/141/2015, BT/08/IYBA/2014/05, BT/406/NE/UEXCEL/2013, BT/PR5511/MED/29/631/2012, BT/341/NE/TBP/2012 (in part)]; Science and Engineering Research Board (SERB) [YSS/2014/000286]. The open access publication charge for this paper has been waived by Oxford University Press – NAR.

Conflict of interest statement. None declared.

REFERENCES

- Barrangou, R., Fremaux, C., Deveau, H., Richards, M., Boyaval, P., Moineau, S., Romero, D.A. and Horvath, P. (2007) CRISPR provides acquired resistance against viruses in prokaryotes. *Science*, **315**, 1709–1712.
- Bolotin, A., Quinquis, B., Sorokin, A. and Ehrlich, S.D. (2005) Clustered regularly interspaced short palindrome repeats (CRISPRs) have spacers of extrachromosomal origin. *Microbiology*, **151**, 2551–2561.
- 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.
- 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.
- Hille, F., Richter, H., Wong, S.P., Bratovič, M., Ressel, S. and Charpentier, E. (2018) The biology of CRISPR-Cas: Backward and forward. *Cell*, **172**, 1239–1259.
- Datsenko, K.A., Pougach, K., Tikhonov, A., Wanner, B.L., Severinov, K. and Semenova, E. (2012) Molecular memory of prior infections activates the CRISPR/Cas adaptive bacterial immunity system. *Nat. Commun.*, **3**, 945.
- Nuñez, J.K., Bai, L., Harrington, L.B., Hinder, T.L. and Doudna, J.A. (2016) CRISPR Immunological memory requires a host factor for specificity. *Mol. Cell*, **62**, 824–833.
- Nunez, J.K., Kranzusch, P.J., Noeske, J., Wright, A.V., Davies, C.W. and Doudna, J.A. (2014) Cas1-Cas2 complex formation mediates spacer acquisition during CRISPR-Cas adaptive immunity. *Nat. Struct. Mol. Biol.*, **21**, 528–534.
- Yoganand, K.N., Sivathanu, R., Nimkar, S. and Anand, B. (2017) Asymmetric positioning of Cas1-2 complex and Integration Host Factor induced DNA bending guide the unidirectional homing of protospacer in CRISPR-Cas type I-E system. *Nucleic Acids Res.*, **45**, 367–381.
- Yosef, I., Goren, M.G. and Qimron, U. (2012) Proteins and DNA elements essential for the CRISPR adaptation process in *Escherichia coli*. *Nucleic Acids Res.*, **40**, 5569–5576.
- Yoganand, K.N., Muralidharan, M., Nimkar, S. and Anand, B. (2019) Fidelity of prespacer capture and processing is governed by the PAM mediated interactions of Cas1-2 adaptation complex in CRISPR-Cas type I-E system. *J. Biol. Chem.*, **294**, 20039–20053.
- Garneau, J.E., Dupuis, M.E., Villion, M., Romero, D.A., Barrangou, R., Boyaval, P., Fremaux, C., Horvath, P., Magadan, A.H. and Moineau, S. (2010) The CRISPR/Cas bacterial immune system cleaves bacteriophage and plasmid DNA. *Nature*, **468**, 67–71.
- Marraffini, L.A. and Sontheimer, E.J. (2010) Self versus non-self discrimination during CRISPR RNA-directed immunity. *Nature*, **463**, 568–571.
- 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.
- Koonin, E.V., Makarova, K.S. and Zhang, F. (2017) Diversity, classification and evolution of CRISPR-Cas systems. *Curr. Opin. Microbiol.*, **37**, 67–78.
- 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.
- Makarova, K.S., Wolf, Y.I. and Koonin, E.V. (2018) Classification and Nomenclature of CRISPR-Cas Systems: Where from Here? *CRISPR J.*, **1**, 325–336.
- Mohanraju, P., Makarova, K.S., Zetsche, B., Zhang, F., Koonin, E.V. and van der Oost, J. (2016) Diverse evolutionary roots and mechanistic variations of the CRISPR-Cas systems. *Science*, **353**, aad5147.
- 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.
- 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.
- Zetsche, B., Gootenberg, J.S., Abudayeh, O.O., Slaymaker, I.M., Makarova, K.S., Essletzbichler, P., Volz, S.E., Joung, J., van der Oost, J., Regev, A. *et al.* (2015) Cpf1 is a single RNA-guided endonuclease of a class 2 CRISPR-Cas system. *Cell*, **163**, 759–771.
- Jore, M.M., Lundgren, M., van Duijn, E., Bultema, J.B., Westra, E.R., Waghmare, S.P., Wiedenheft, B., Pul, U., Wurm, R., Wagner, R. *et al.* (2011) Structural basis for CRISPR RNA-guided DNA recognition by Cascade. *Nat. Struct. Mol. Biol.*, **18**, 529–536.
- Wiedenheft, B., van Duijn, E., Bultema, J.B., Waghmare, S.P., Zhou, K., Barendregt, A., Westphal, W., Heck, A.J., Boekema, E.J., Dickman, M.J. *et al.* (2011) RNA-guided complex from a bacterial immune system enhances target recognition through seed sequence interactions. *Proc. Natl. Acad. Sci. U.S.A.*, **108**, 10092–10097.
- Nam, K.H., Haitjema, C., Liu, X., Ding, F., Wang, H., DeLisa, M.P. and Ke, A. (2012) Cas5d protein processes pre-crRNA and assembles into a cascade-like interference complex in subtype I-C/Dvulg CRISPR-Cas system. *Structure*, **20**, 1574–1584.
- Hayes, R.P., Xiao, Y., Ding, F., van Erp, P.B., Rajashankar, K., Bailey, S., Wiedenheft, B. and Ke, A. (2016) Structural basis for promiscuous PAM recognition in type I-E Cascade from *E. coli*. *Nature*, **530**, 499–503.
- Hochstrasser, M.L., Taylor, D.W., Kornfeld, J.E., Nogales, E. and Doudna, J.A. (2016) DNA targeting by a minimal CRISPR RNA-Guided cascade. *Mol. Cell*, **63**, 840–851.
- Wiedenheft, B., Lander, G.C., Zhou, K., Jore, M.M., Brouns, S.J.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.
- Mojica, F.J., Diez-Villasenor, C., Garcia-Martinez, J. and Almendros, C. (2009) Short motif sequences determine the targets of the prokaryotic CRISPR defence system. *Microbiology*, **155**, 733–740.
- Westra, E.R., Semenova, E., Datsenko, K.A., Jackson, R.N., Wiedenheft, B., Severinov, K. and Brouns, S.J. (2013) Type I-E CRISPR-cas systems discriminate target from non-target DNA through base pairing-independent PAM recognition. *PLoS Genet.*, **9**, e1003742.
- Sapranaukas, R., Gasiunas, G., Fremaux, C., Barrangou, R., Horvath, P. and Siksnys, V. (2011) The *Streptococcus thermophilus* CRISPR/Cas system provides immunity in *Escherichia coli*. *Nucleic Acids Res.*, **39**, 9275–9282.
- Mulepati, S. and Bailey, S. (2013) In vitro reconstitution of an *Escherichia coli* RNA-guided immune system reveals unidirectional, ATP-dependent degradation of DNA target. *J. Biol. Chem.*, **288**, 22184–22192.
- Blosser, T.R., Loeff, L., Westra, E.R., Vlot, M., Kunne, T., Sobota, M., Dekker, C., Brouns, S.J.J. and Joo, C. (2015) Two distinct DNA binding modes guide dual roles of a CRISPR-Cas protein complex. *Mol. Cell*, **58**, 60–70.
- Rutkauskas, M., Sinkunas, T., Songailiene, I., Tikhomirova, M.S., Siksnys, V. and Seidel, R. (2015) Directional R-Loop formation by the CRISPR-Cas surveillance complex cascade provides efficient Off-Target site rejection. *Cell Rep.*, **10**, 1534–1543.
- 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.
- 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. U.S.A.*, **111**, 9798–9803.
- Xiao, Y., Luo, M., Hayes, R.P., Kim, J., Ng, S., Ding, F., Liao, M. and Ke, A. (2017) Structure basis for directional R-loop formation and substrate handover mechanisms in Type I CRISPR-Cas system. *Cell*, **170**, 48–60.
- Cooper, L.A., Stringer, A.M. and Wade, J.T. (2018) Determining the specificity of cascade binding, Interference, and primed adaptation in vivo in the *Escherichia coli* type I-E CRISPR-Cas system. *mBio*, **9**, e02100-17.
- Hochstrasser, M.L., Taylor, D.W., Bhat, P., Guegler, C.K., Sternberg, S.H., Nogales, E. and Doudna, J.A. (2014) CasA mediates

- Cas3-catalyzed target degradation during CRISPR RNA-guided interference. *Proc. Natl. Acad. Sci. U.S.A.*, **111**, 6618–6623.
39. Redding, S., Sternberg, S.H., Marshall, M., Gibb, B., Bhat, P., Guegler, C.K., Wiedenheft, B., Doudna, J.A. and Greene, E.C. (2015) Surveillance and processing of foreign DNA by the Escherichia coli CRISPR-Cas system. *Cell*, **163**, 854–865.
 40. van Erp, P.B.G., Patterson, A., Kant, R., Berry, L., Golden, S.M., Forsman, B.L., Carter, J., Jackson, R.N., Bothner, B. and Wiedenheft, B. (2017) Conformational dynamics of DNA binding and Cas3 recruitment by the CRISPR RNA-Guided cascade complex. *ACS Chem. Biol.*, **13**, 481–490.
 41. Westra, E.R., Nilges, B., van Erp, P.B., van der Oost, J., Dame, R.T. and Brouns, S.J. (2012) Cascade-mediated binding and bending of negatively supercoiled DNA. *RNA Biol.*, **9**, 1134–1138.
 42. Gong, B., Shin, M., Sun, J., Jung, C.H., Bolt, E.L., van der Oost, J. and Kim, J.S. (2014) Molecular insights into DNA interference by CRISPR-associated nuclease-helicase Cas3. *Proc. Natl. Acad. Sci. U.S.A.*, **111**, 16359–16364.
 43. Huo, Y., Nam, K.H., Ding, F., Lee, H., Wu, L., Xiao, Y., Farchione, M.D. Jr., Zhou, S., Rajashankar, K., Kurinov, I. et al. (2014) Structures of CRISPR Cas3 offer mechanistic insights into Cascade-activated DNA unwinding and degradation. *Nat. Struct. Mol. Biol.*, **21**, 771–777.
 44. Sinkunas, T., Gasiunas, G., Fremaux, C., Barrangou, R., Horvath, P. and Siksnys, V. (2011) Cas3 is a single-stranded DNA nuclease and ATP-dependent helicase in the CRISPR/Cas immune system. *EMBO J.*, **30**, 1335–1342.
 45. Beloglazova, N., Petit, P., Flick, R., Brown, G., Savchenko, A. and Yakunin, A.F. (2011) Structure and activity of the Cas3 HD nuclease MJ0384, an effector enzyme of the CRISPR interference. *EMBO J.*, **30**, 4616–4627.
 46. Mulepati, S. and Bailey, S. (2011) Structural and biochemical analysis of nuclease domain of clustered regularly interspaced short palindromic repeat (CRISPR)-associated protein 3 (Cas3). *J. Biol. Chem.*, **286**, 31896–31903.
 47. Xiao, Y., Luo, M., Dolan, A.E., Liao, M. and Ke, A. (2018) Structure basis for RNA-guided DNA degradation by Cascade and Cas3. *Science*, **361**, eaat0839.
 48. Plagens, A., Tjaden, B., Hagemann, A., Randau, L. and Hensel, R. (2012) Characterization of the CRISPR/Cas subtype I-A system of the hyperthermophilic crenarchaeon Thermoproteus tenax. *J. Bacteriol.*, **194**, 2491–2500.
 49. Majumdar, S. and Terns, M.P. (2019) CRISPR RNA-guided DNA cleavage by reconstituted Type I-A immune effector complexes. *Extremophiles*, **23**, 19–33.
 50. Loeff, L., Brouns, S.J.J. and Joo, C. (2018) Repetitive DNA reeling by the Cascade-Cas3 complex in nucleotide Unwinding steps. *Mol. Cell*, **70**, 385–394.
 51. Westra, E.R., van Erp, P.B., Kunne, 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.
 52. Majumdar, S., Ligon, M., Skinner, W.C., Terns, R.M. and Terns, M.P. (2017) Target DNA recognition and cleavage by a reconstituted Type I-G CRISPR-Cas immune effector complex. *Extremophiles*, **21**, 95–107.
 53. Dwarakanath, S., Brenzinger, S., Gleditsch, D., Plagens, A., Klingl, A., Thormann, K. and Randau, L. (2015) Interference activity of a minimal Type I CRISPR-Cas system from Shewanella putrefaciens. *Nucleic Acids Res.*, **43**, 8913–8923.
 54. Elmore, J., Deighan, T., Westpheling, J., Terns, R.M. and Terns, M.P. (2015) DNA targeting by the type I-G and type I-A CRISPR-Cas systems of Pyrococcus furiosus. *Nucleic Acids Res.*, **43**, 10353–10363.
 55. Plagens, A., Tripp, V., Daume, M., Sharma, K., Klingl, A., Hrle, A., Conti, E., Urlaub, H. and Randau, L. (2014) In vitro assembly and activity of an archaeal CRISPR-Cas type I-A Cascade interference complex. *Nucleic Acids Res.*, **42**, 5125–5138.
 56. Richter, C., Gristwood, T., Clulow, J.S. and Fineran, P.C. (2012) In vivo protein interactions and complex formation in the Pectobacterium atrosepticum subtype I-F CRISPR/Cas system. *PLoS One*, **7**, e49549.
 57. Rollins, M.F., Chowdhury, S., Carter, J., Golden, S.M., Wilkinson, R.A., Bondy-Denomy, J., Lander, G.C. and Wiedenheft, B. (2017) Cas1 and the Csy complex are opposing regulators of Cas2/3 nuclease activity. *Proc. Natl. Acad. Sci. U.S.A.*, **114**, E5113–E5121.
 58. Peng, W., Li, H., Hallstrom, S., Peng, N., Liang, Y.X. and She, Q. (2013) Genetic determinants of PAM-dependent DNA targeting and pre-crRNA processing in Sulfolobus islandicus. *RNA Biol.*, **10**, 738–748.
 59. Pausch, P., Müller-Esparza, H., Gleditsch, D., Altegoer, F., Randau, L. and Bange, G. (2017) Structural variation of Type I-F CRISPR RNA Guided DNA Surveillance. *Mol. Cell*, **67**, 622–632.
 60. Punetha, A., Sivathanu, R. and Anand, B. (2014) Active site plasticity enables metal-dependent tuning of Cas5d nuclease activity in CRISPR-Cas type I-C system. *Nucleic Acids Res.*, **42**, 3846–3856.
 61. Garside, E.L., Schellenberg, M.J., Gesner, E.M., Bonanno, J.B., Sauder, J.M., Burley, S.K., Almo, S.C., Mehta, G. and MacMillan, A.M. (2012) Cas5d processes pre-crRNA and is a member of a larger family of CRISPR RNA endonucleases. *RNA*, **18**, 2020–2028.
 62. St-Pierre, F., Cui, L., Priest, D.G., Endy, D., Dodd, I.B. and Shearwin, K.E. (2013) One-step cloning and chromosomal integration of DNA. *ACS Synth. Biol.*, **2**, 537–541.
 63. Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K.A., Tomita, M., Wanner, B.L. and Mori, H. (2006) Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. *Mol. Syst. Biol.*, **2**, 2006.0008.
 64. Zhang, Y. (2008) I-TASSER server for protein 3D structure prediction. *BMC Bioinformatics*, **9**, 40.
 65. Howard, J.A., Delmas, S., Ivancic-Bace, I. and Bolt, E.L. (2011) Helicase dissociation and annealing of RNA-DNA hybrids by Escherichia coli Cas3 protein. *Biochem. J.*, **439**, 85–95.
 66. Mohr, G., Del Campo, M., Mohr, S., Yang, Q., Jia, H., Jankowsky, E. and Lambowitz, A.M. (2008) Function of the C-terminal domain of the DEAD-box protein Mss116p analyzed in vivo and in vitro. *J. Mol. Biol.*, **375**, 1344–1364.
 67. Del Campo, M. and Lambowitz, A.M. (2009) Structure of the yeast DEAD box protein Mss116p reveals two wedges that crimp RNA. *Mol. Cell*, **35**, 598–609.
 68. Dillard, K.E., Brown, M.W., Johnson, N.V., Xiao, Y., Dolan, A., Hernandez, E., Dahlhauser, S.D., Kim, Y., Myler, L.R., Anslyn, E.V. et al. (2018) Assembly and translocation of a CRISPR-Cas primed acquisition complex. *Cell*, **175**, 934–946.