Primed CRISPR DNA uptake in *Pyrococcus furiosus*

Sandra Garrett[1],[†], Masami Shiimori[2],[†], Elizabeth A. Watts[2], Landon Clark[2], Brenton R. Graveley[1],* and Michael P. Terns [2],[3],[4],*

[1]Department of Genetics and Genome Sciences, Institute for Systems Genomics, UConn Stem Cell Institute, UConn Health, Farmington, CT 06030, USA, [2]Department of Biochemistry and Molecular Biology, University of Georgia, Athens, GA 30602, USA, [3]Department of Microbiology, University of Georgia, Athens, GA 30602, USA and [4]Department of Genetics, University of Georgia, Athens, GA 30602, USA

ABSTRACT

CRISPR-Cas adaptive immune systems are used by prokaryotes to defend against invaders like viruses and other mobile genetic elements. Immune memories are stored in the form of 'spacers' which are short DNA sequences that are captured from invaders and added to the CRISPR array during a process called 'adaptation'. Spacers are transcribed and the resulting CRISPR (cr)RNAs assemble with different Cas proteins to form effector complexes that recognize matching nucleic acid and destroy it ('interference'). Adaptation can be 'naïve', i.e. independent of any existing spacer matches, or it can be 'primed', i.e. spurred by the crRNA-mediated detection of a complete or partial match to an invader sequence. Here we show that primed adaptation occurs in *Pyrococcus furiosus*. Although *P. furiosus* has three distinct CRISPR-Cas interference systems (I-B, I-A and III-B), only the I-B system and Cas3 were necessary for priming. Cas4, which is important for selection and processing of new spacers in naïve adaptation, was also essential for priming. Loss of either the I-B effector proteins or Cas3 reduced naïve adaptation. However, when Cas3 and all crRNP genes were deleted, uptake of correctly processed spacers was observed, indicating that none of these interference proteins are necessary for naïve adaptation.

INTRODUCTION

CRISPR-Cas (clustered regularly interspaced short palindromic repeats and CRISPR-associated genes) systems provide adaptive immunity in bacteria and archaea. The systems store sequence information about potentially deleterious viruses and other mobile genetic elements in the CRISPR array (1) and use that stored information to carry out targeted, sequence-specific degradation of DNA or RNA, depending upon CRISPR type (2–8). CRISPR-Cas systems are diverse and have been classified into two classes, six distinct types (I–VI), and at least 30 subtypes (9), but certain characteristics are shared. All CRISPR arrays contain a series of direct repeats separated by short sequences called 'spacers' which match DNA from previously encountered invaders (10,11). An upstream leader sequence regulates transcription of the array and also mediates addition of new spacers (12–14). In addition to the CRISPR array, there are typically multiple nearby genes encoding CRISPR-associated (Cas) proteins, including effector nucleases capable of destroying target nucleic acid. After transcription, CRISPR array RNAs are processed into short guide RNAs (crRNAs) which associate with Cas nucleases to form a crRNA-guided effector complex (referred to as the crRNP effector complex here) (15,16). Base pairing between the crRNA and the target site (called the protospacer) allows for sequence-specific recognition of DNA or RNA (depending upon the CRISPR system type). For DNA-targeting CRISPR systems, if the target has an activating sequence motif present (called the Protospacer Adjacent Motif or PAM) then the complex degrades the target nucleic acid and silences the invader (i.e. carries out interference) (17–19).

New immune memories are formed when short fragments of DNA are taken from invading genetic elements, processed, and integrated into CRISPR arrays as new spacers (a process termed adaptation) (20–22). If no spacers match the invading genetic element, new spacer uptake is termed naïve adaptation. Adaptation can also be 'primed', which occurs when an existing spacer matches or partially matches the invader DNA. In this scenario, when the crRNP effector complex recognizes this match, it stimulates new spacer uptake using DNA in the vicinity of the protospacer target (23,24). Efficient interference usually requires a canonical PAM and high identity between the crRNA and the protospacer, particularly in the 'seed' region, which lies adjacent

to the PAM in type I and type II systems (25–27). However, primed adaptation can tolerate mismatches in the target or a non-consensus PAM (23–24,28–32) so mutations that might normally allow a target to escape CRISPR immune defence will still leave it vulnerable to interference once priming has updated the CRISPR array.

While mechanistic details are still emerging, some key components of adaptation have been identified, particularly for bacterial systems. Cas1 and Cas2 proteins, which are present in almost all active CRISPR-Cas systems described to date, are necessary for both naïve and primed adaptation. In *Escherichia coli*, the Cas1–Cas2 complex is required for recognizing PAMs, processing protospacers to the proper size, and integrating them into the CRISPR array (33–35). *In vitro*, Cas1 and Cas2 are sufficient to integrate a pre-spacer (DNA fragment with free 3′-OH ends) into plasmid DNA containing a CRISPR array (36). *In vivo*, numerous other proteins play a role in generating, processing and integrating new spacers. For example, IHF (Integration host factor) is essential for directing integration at the leader-proximal repeat for I-E and I-F systems (37–39). Additionally, spacer acquisition primarily occurs at sites of double-stranded breaks, like those that form at stalled DNA replication forks (40). As the RecBCD DNA repair complex resolves the stalled replication fork, it creates breakdown products, which are incorporated into the Cas1–Cas2 adaptation complex and converted into new spacers (40). Consequently, RecBCD influences which spacers are captured even though it is not an essential part of the adaptation complex.

As adaptation is investigated in a wider range of CRISPR-Cas systems and species, the list of factors involved in adaptation expands. In some Type I-F systems, Cas2 is fused with Cas3 (41–43), a nuclease recruited in *trans* by most type I Cas–crRNA complexes after target recognition (44–48). The Cas2–Cas3 fusion protein forms a complex with Cas1 and together they direct the recognition of protospacer PAMs, process spacers, and integrate them into the array (37). In the type I-E system in *E. coli*, Cas3 is not fused to Cas2, but was still found to play a key role in primed adaptation by providing Cas1 and Cas2 with partially double-stranded DNA fragments that are suitable for conversion into new spacers (23,49). In some systems, Cas4 is essential for PAM recognition, spacer processing and proper spacer integration; adaptation can occur without Cas4 in these systems, but new spacers are likely to be mis-sized, misoriented and derived from DNA fragments lacking a PAM and therefore non-functional (50–54). Several studies have also shown that the crRNP effector complexes can be involved in adaptation. With primed adaptation, it is not surprising that at least some elements of the crRNA effector complex are needed, since target recognition underlies priming. Evidence from some systems suggests that the effector complex also influences naïve adaptation. For example, the type I-F Csy effector complex in *Pseudomonas aeruginosa* and the type II-A effector nuclease Cas9 in *Streptococcus thermophilus* and *Streptococcus pyogenes* are essential for efficient adaptation (55–57). These various examples all suggest a complex interplay between adaptation, interference and other non-CRISPR cellular processes, but details and mechanism remain unclear.

Primed adaptation has been reported in at least four different type I systems: the type I-B in *Haloarcula hispanica*, the type I-C in *Legionella pneumophila*, the type I-E in *E. coli*, the type I-F of *Pectobacterium atrosepticum* (23–24,28–30,58) and was very recently reported in a Type II-A system (59). Much progress has been made in understanding how the crRNA effector complex couples with the adaptation machinery to produce priming for Type I systems. Priming requires the nuclease Cas3 and the immune effector complex, in addition to Cas1 and Cas2 (23,30,56). All of these components (effector complex, Cas3, Cas1/Cas2) can associate with one another in the presence of target DNA and can then translocate along the DNA together (or reel DNA toward the effector complex), potentially allowing for the simultaneous production and uptake of new spacers from DNA flanking a target (60,61). As noted above, there is evidence that DNA fragments produced by Cas3 during target interference can be preferentially captured by Cas1/Cas2 for conversion into new spacers (31,49).

We previously described robust spacer acquisition in the hyperthermophilic archean, *Pyrococcus furiosus* (53,62–63). Like many other organisms, *P. furiosus* harbors more than one CRISPR-Cas system. The *P. furiosus* CRISPR-Cas suite includes Type I-A (Csa) and Type I-B (Cst) systems that mediate crRNA-guided cleavage of DNA (44–45,64), and the Type III-B (Cmr) system that carries out cleavage of RNA and transcript-dependent DNA degradation (5,7,65). The type I-B system in *P. furiosus* was formerly referred to as Cst (66), later classified as I-G (67), but is now considered a I-B in the current classification (9). The *P. furiosus* genome contains seven shared CRISPR loci, a single CRISPR RNA processing enzyme gene and one set of adaptation genes (68). Genes of I-B and III-B effector complexes are encoded at one locus together with the *cas1*, *cas2* and *cas4-1* adaptation genes and the *cas6* gene encoding the crRNA processing endonuclease. The Type I-A gene locus is isolated with no associated adaptation genes. The other adaptation gene, *cas4-2*, is in an isolated locus and is not associated with any other cas genes. In *P. furiosus*, Cas1, Cas2, Cas4-1 and Cas4-2 proteins are required for acquisition of functional spacers and overexpression of these four proteins elevates adaptation (53,62–63). Here, we show that primed adaptation occurs in *P. furiosus* and is dependent on Cas3 and the Type I-B effector complex. Priming occurred both when there was a perfect protospacer target and when the target had a non-canonical PAM or mismatches in the seed region. In contrast to some other systems (55–57), we found that naïve adaptation of functional, correctly processed spacers in *P. furiosus* did not require any immune effector genes, but could be inhibited by loss of either Cas3 or the I-B effector complex, suggesting that these gene products interact, directly or indirectly, with Cas1 and Cas2 and influence adaptation in an interference-independent manner.

MATERIALS AND METHODS

P. furiosus strains and growth conditions

The strains used in this study are listed in Supplementary Table S1. *Pyrococcus furiosus* strains were grown anaerobically at 95°C in a defined medium with cellobiose as the

carbon source (69). Cultures were grown either overnight in anaerobic culture bottles or for 65 h on medium solidified with 1% wt/vol Gelrite (Research Product International). For growth of uracil auxotrophic strains, the defined medium contained 20 μM uracil. For growth of strains transformed with plasmids with *TrpAB*, the defined medium lacked tryptophan. Plasmid transformation was performed as described previously (69). All assays were performed with three replicates.

Plasmid construction

Plasmids were constructed by standard cloning techniques. The sequences of DNA oligonucleotides used in this study are shown in Supplementary Table S2 and plasmids are shown in Supplementary Table S3. To generate target plasmids, oligonucleotides were annealed and then digested by BamHI and NdeI. To construct plasmids with genome integration cassettes (pMS030 and pMS088), homologous regions were amplified from *P. furiosus* JFW02 genome and *gdh*-promoter-*pyrF* was amplified from pJFW18 plasmid. The amplified products were assembled by overlap polymerase chain reaction (PCR) and ligated into pHSG298 plasmid using GENEART seamless cloning kit (Thermo Fisher). To construct overexpression cassettes, Cas1, Cas2, Cas4-1, Cas4-2 and Cas3 coding regions and promoter regions were amplified from the *P. furiosus* wild-type (WT) genome and the *Thermococcus kodakarensis* (Tko) WT genome. The amplified products were assembled by overlap PCR and ligated with pJE47 plasmid (Cas1, Cas2, Cas4-1, Cas4-2 overexpression cassette: Tko *csg* promoter-Cas2 (PF1117), Tko *gdh* promoter-Cas4-2 (PF1793), *slp* promoter-Cas4-1 (PF1119), PRP synthetase promoter-Cas1 (PF1118)) or pJE64 (Cas3 overexpression cassette: Tko *csg* promoter-Cas3 (PF1120)), respectively. Overexpression cassettes were digested by NotI and EcoRV and ligated with pMS030 or pMS088 to yield genome integration plasmids. HD nuclease and helicase active sites of Cas3 (Supplementary Figure S1) were mutagenized via QuikChange PCR using pMS098 plasmid as the template. The plasmids were sequenced to confirm insert sequence.

Strain construction

To create Cas1/Cas2/Cas4-1 overexpression strains and the Cas4-1 deletion strain, NruI-linearized pMS032 or pMS087 plasmids were transformed into TPF17. Plasmids are listed in Supplementary Table S3. Two rounds of colony purification were performed by plating 10^{-3} dilutions of transformant cultures onto selective plate medium (without uracil) and picking isolated colonies into selective liquid medium. Following marker replacement of the region of interest, 5-FOA, a toxic pyrF substrate, was used to select for cells that underwent pop-out of the pyrF marker by homologous recombination. The Cas1/Cas2/Cas4-1/Cas4-2 deletion strain (Δad) and ΔCas4-2 strain were created using the pop-out marker replacement strategy as described previously (70). The transformed PCR products were generated by overlap PCR. The sequence of DNA oligonucleotides used is shown in Supplementary Table S2.

Plasmid interference assays

Defined liquid media cultures of *P. furiosus* were allowed to grow overnight at 90°C for ∼16 h (mid-to-late log phase of growth). In aerobic conditions, 200 ng of either the no target control plasmid or target plasmid (containing a protospacer matching the 7.01 crRNA) were transformed into 100 μl of an overnight culture and incubated at room temperature for 15 min to 1 h. These transformations were split and plated onto three solid defined media plates lacking uracil (to select for the transformed plasmid). The plates were anaerobically sealed in chambers and incubated at 90°C for 3 days. Colony growth was observed and counted.

Adaptation assay and high throughput sequencing of expanded array PCR products

Transformations, colony growth and preparation of adaptation amplicon libraries were done as previously described (63). Briefly, genomic DNA was isolated from *P. furiosus* cells from 1 ml of overnight culture using the quick-gDNA miniprep kit (Zymo Research) and CRISPR arrays were amplified by PCR using a set of primers in which the forward primer annealed within the leader region of the CRISPR array and the reverse primer annealed to the existing spacer closest to the leader. Expanded PCR products were separated from unexpanded products by gel electrophoresis followed by DNA recovery (Zymo Research). PCR primers included an overhang corresponding to part of the adapter necessary for Illumina sequencing; additional PCRs were done to further select for expanded array products and to add Illumina barcodes. Final gel-purified amplicon libraries were pooled according to band intensity, normalized according to concentration and number of samples represented in the pool and sequenced on an Illumina MiSeq to yield 250 by 50 paired-end reads; the 250 base read 1 sequences were used in this study. After sequencing, samples were de-multiplexed by index and the sequence corresponding to a new spacer was extracted from each read. Spacer sequences were then aligned and analysed, as previously described, to characterize spacer size distribution, protospacer source, PAM and patterns of protospacer distribution along the genome and any plasmids used in the assay (63).

Spacer density tracks

To make spacer density tracks, total spacer alignments were used to generate base coverage files (bedtools, (71)), and these were then displayed on a custom track hub on the UCSC Genome Browser (https://genome.ucsc.edu/cgi-bin/hgHubConnect). For the PAM density track, an alignment file was generated *in silico* to include a single 37 bp protospacer adjacent to each NGG PAM in the genome and plasmid. This alignment file was converted to a base coverage file and displayed on the UCSC Genome Browser alongside spacer density tracks. For both the PAM and spacer density tracks, the 'mean' windowing option was applied to bin coverage values for display.

RESULTS

Adaptation is stimulated by plasmids with crRNA targets

We sought to determine if *P. furiosus* undergoes primed adaptation, and if so, what components mediate the process (Figure 1). First, we assessed interference using a negative control plasmid lacking a target sequence (no target) as well as a plasmid with a target that corresponds to a transcribed protospacer matching an endogenous crRNA (7.01; the first crRNA from CRISPR locus 7). Variations of the 7.01 target were made with one, two, three or four mismatches in the seed region (blue region highlighted in target, Figure 1A). These plasmids were transformed into *P. furiosus* JFW02 strain (referred to as WT/wild-type in this paper) (Figure 1A). Transformed strains were then assessed for both interference against the plasmid (Figure 1B) and adaptation (Figure 1C). A perfect target (i.e. one with a canonical GGG PAM and no mismatches between the 7.01 crRNA and the protospacer) on the plasmid reduced transformation efficiency by 1000-fold relative to the no target control plasmid, indicating robust interference (Figure 1B). One or two mismatches in the target seed region partially restored transformation efficiency while three and four consecutive mismatches returned transformation to the level observed for the no target plasmid, indicating a loss of interference (Figure 1B). Adaptation was observed by PCR amplification of the region between the leader and first spacer (63). If a new spacer was added to the CRISPR array (i.e. the array had been 'expanded'), PCR yielded a larger product, which included an additional spacer and repeat sequence (Figure 1C, * unexpanded, +1 expanded). Expanded CRISPR arrays were observed in all samples after two rounds of PCR, except the negative control strain (△ad) wherein the core adaptation genes, *cas1*, *cas2*, *cas4-1* and *cas4-2* had been deleted (Figure 1C). We noted that the intensity of the expanded band was relatively greater for the single and double mismatch plasmids than for the no target plasmid, suggesting the possibility of enhanced adaptation efficiency due to priming.

Type I-B effector crRNP complex is required for primed adaptation

Pyrococcus furiosus harbors three functional crRNP effector complexes (Type I-A (Csa), Type I-B (Cst) and Type III-B (Cmr)) that each utilize common crRNAs for their function (Figure 2A) (7,45,65,68). To examine whether these immune effectors are involved in naïve and primed adaptation in *P. furiosus*, we compared adaptation in the WT strain with adaptation in strains having only a single effector complex (I-A strain, I-B strain and III-B strain) or no effector complex (5,64). In each single effector strain, *cas3* (*Pfu* 1120) was left intact so as not to perturb downstream expression of *cas1, cas2 and cas4-1*. Plasmids with either no target (reflecting the case of naïve adaptation) or the two-mismatch target (exhibiting intermediate interference levels (Figure 1B) and potentially supporting primed adaptation) were transformed into each of these single effector strains, and interference and adaptation were assessed as in Figure 1. Individual effector strains had interference levels similar or identical to that observed for the WT strain con-

taining all three systems, showing that each system is capable of supporting robust anti-plasmid interference (Figure 2B). As expected, the strain lacking all three effector complexes (△crRNP) showed no reduction in transformation efficiency.

Expanded CRISPR arrays were observed in all strains after three rounds of PCR (Figure 2C), indicating that no one effector complex is strictly essential for naïve adaptation in *P. furiosus*. However, band intensity suggested that adaptation was reduced in the absence of the I-B effector complex and enhanced when this complex was present alone (Figure 2C), suggesting a role for the I-B module in naïve adaptation. To assess the relative adaptation efficiency among strains, we also examined the number of unique spacers in these samples. A single spacer uptake event in a given cell could potentially be duplicated many times during growth in culture and PCR amplification. We identified unique spacers by their unique alignments (position, length, strand), collapsed all duplicates, and, given sufficient sequencing depth, we used that unique count as a proxy for the number of initial spacer uptake events in the original culture. This approach also suggested lower relative levels of adaptation in strains lacking I-B (Supplementary Figure S2). We did not anticipate that loss of the I-B effector complex would reduce naïve adaptation since physical interactions between the I-B effector and the core adaptation proteins (Cas1, Cas2, Cas4-1 and Cas4-2) have not yet been detected (45). To determine whether naïve spacer acquisition could be rescued with by overexpression of *cas1*, *cas2*, *cas4-1* and *cas4-2*, these four genes were overexpressed by plasmid. Overexpression of *cas1*, *cas2*, *cas4-1* and *cas4-2* returned adaptation levels to that of WT (Supplementary Figure S2). The results suggest a potential role for the I-B effector complex in naïve adaptation.

In addition, the difference in band intensity between target and no target plasmid samples suggested that the I-B effector complex was required for priming (Figure 2C). We next characterized the source, distribution, size and flanking sequences of new spacers acquired into the CRISPR arrays for the strains shown in Figure 2. For all samples, including those in which bands were faint or absent, we cut from the gel at the position where expanded bands were present or expected and sequenced all samples in parallel. One hallmark of primed adaptation is an abundance of new spacers arising from the vicinity of the target protospacer (23–24,28–30,58). Since the protospacer target was located on a plasmid, we expected an increase in the proportion of plasmid-derived versus genome-derived new spacers for target samples, particularly for the I-B strain, which had shown a distinct increase in expanded band intensity when the target plasmid (with two seed region mismatches) was used (Figure 2C). When no target plasmids were used, fewer than 5% of new spacers were plasmid-derived, consistent with previous studies (53,63), but that rose to more than 80% when a target plasmid was used (Figure 3). New spacers were preferentially selected from DNA surrounding the target (Figure 3) and protospacers on the non-target strand (i.e. the strand of protospacer DNA opposite to the strand engaged in the crRNA base-paired interaction) were preferentially located 3′ of the primed protospacer (PAM side), while protospacers on the target strand were preferentially
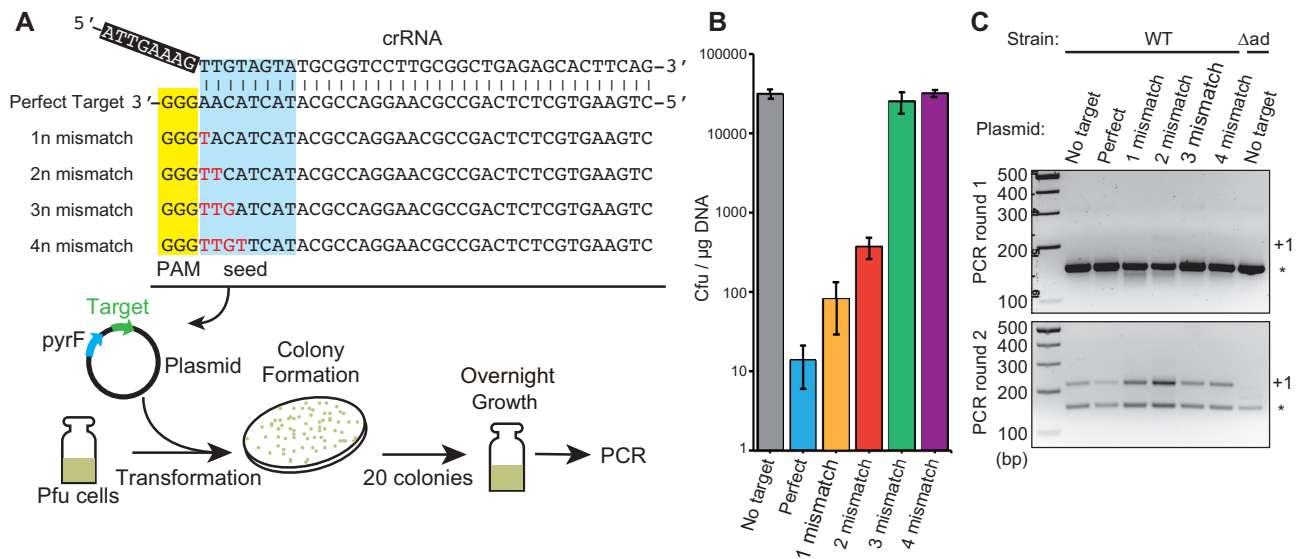
Figure 1. Mutations in the target DNA seed sequence affect interference and spacer uptake. (**A**) Experimental design for seed sequence mutants. A 'Perfect Target' plasmid was designed to contain a CRISPR target with 100% identity to spacer 7.01 of *Pyrococcus furiosus* and a canonical PAM (GGG). Additional plasmids were made to contain between one and four mismatches in the seed region (seed region highlighted in blue, mutated nucleotides indicated in red). (**B**) Transformation efficiency for target plasmids. A non-target plasmid and each of the five target plasmids were transformed into WT *P. furiosus* and then assessed for colony forming units. Plots show mean colony forming units per µg of input DNA ± SEM ($n = 8$), see Supplementary Table S4 for individual data. (**C**) Analysis of adaptation in WT *P. furiosus*. The leader-adjacent region of CRISPR7 array was amplified by PCR. PCR products corresponding to parental arrays (unexpanded) and arrays with one new repeat-spacer unit (expanded) are indicated with an asterisk and +1 respectively. Lane Δad (right-most lane) corresponds to the negative control strain wherein the adaptation genes for Cas1, Cas2, Cas4-1 and Cas4-2 have been deleted.

located 5′ of the primed protospacer (Figure 3). To confirm that the pattern of spacer distribution followed the protospacer target, the location and strand of the protospacer was changed (Target plasmid, position 2); a similar distribution was observed, again centered on the protospacer (Figure 3).

The III-B only and I-A only strains do not appear to support primed adaptation. The proportion of plasmid-derived spacers did not increase using the target plasmid (Supplementary Figure S3, left panel). Since there were very few unique new spacers for these strains the spacer density tracks were too sparse to resolve distribution patterns around the protospacer target, we overexpressed the adaptation genes (*cas1*, *cas2*, *cas4-1*, *cas4-2*) to increase spacer uptake. We were then able to characterize the source (Supplementary Figure S3-A, right panel) and distribution of spacers for target and no target plasmids (Supplementary Figure S3B). These results also showed no evidence of priming. We noted a large cluster of protospacers on the right-hand end of the linearized plasmid for both the III-B and I-A strains that is independent of the target. This peak lies directly over the double stranded origin of replication for this rolling circle plasmid (Supplementary Figure S4). We had previously identified this protospacer cluster and proposed that it was due to DNA nicking by the Rep protein (63).

Noncanonical PAM sequences support priming

It was reported that mutations in PAM and/or protospacer sequence reduce interference but also stimulate primed adaptation for I-E and I-F systems (24,26,28–30). We gen-

erated two target plasmids wherein the canonical 5′-GGG-3′ PAM was mutated to either GAG and TGC, which were both previously found to reduce but not eliminate interference (64). These two plasmids were then transformed individually into the I-B strain and interference and adaptation were characterized as before. Consistent with previous findings, the two PAM mutants reduced interference (Figure 4A). The GAG and TGC PAMs both maintained or enhanced primed adaptation, as indicated by band intensity (Figure 4B), proportion of plasmid-derived spacers (Figure 4C) and the pattern of increased spacer density around the target site (Figure 4D). We found that a perfect target (i.e. a target with a canonical PAM (GGG) and no mismatches between the crRNA and the protospacer) was also able to stimulate priming (Figure 4B and D).

Cas3 is essential for primed adaptation

In Type I systems, protospacer recognition by the crRNP effector complex leads to recruitment of the Cas3 nuclease/helicase for degradation of target DNA (3,46,72–73). Cas3 can also be involved in adaptation, with reports demonstrating that it is required for priming and that Cas3 degradation products can be used as new spacers (23,30–31,49). To test whether Cas3 is also essential for primed adaptation in *P. furiosus*, we analyzed adaptation in a Cas3 deletion mutant in the I-B only background. We confirmed that Cas3 deletion in the I-B strain abolishes DNA interference (Figure 5A), as expected (64). When WT Cas3 was reconstituted by plasmid expression, interference was rescued (Figure 5A). Cas3 nuclease activity is localized to the HD domain and helicase activity to the SF2 helicase do-
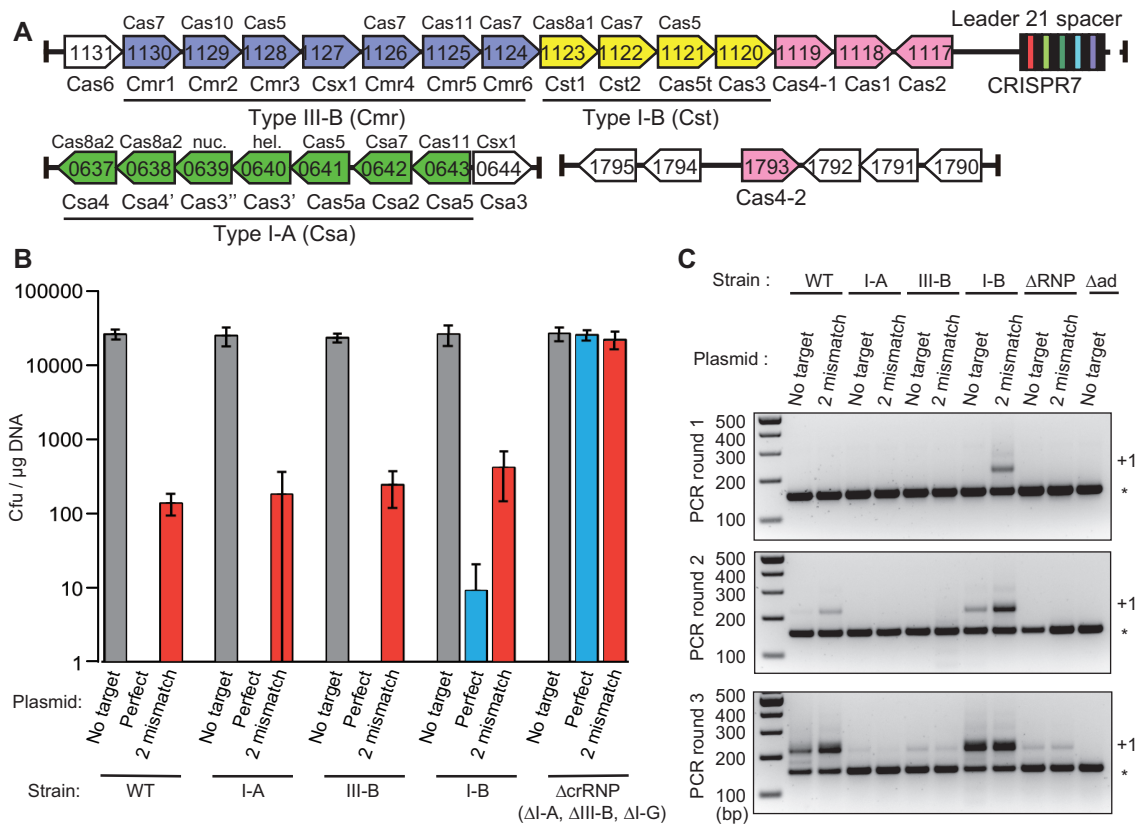
Figure 2. Primed adaptation occurs dependent on I-B crRNP effector complex. (**A**) Overview of cas gene loci of *Pyrococcus furiosus* strain JFW02 (WT). CRISPR adaptation genes are shown in pink. Genes for the three effector complexes (crRNPs) are color-coded as follows: blue for type III-B (Cmr), yellow for type I-B (Cst) and green for type I-A (Csa). Single effector strains were made by deletion of the other two effector complex genes, the ΔcrRNP strain was made by deletion of all three sets of effector genes. Note that Cas3 (Pfu 1120), which is upstream of the adaptation genes, was not deleted in the single effector or ΔcrRNP strains. (**B**) Plot of transformation efficiency. Mean colony forming units in WT, single effector strains and the ΔcrRNP strain are plotted ± SEM, (*n* = 4), see Supplementary Table S4 for individual data. Three plasmids were transformed for each strain: a plasmid with no target, a perfect target, or a target with two mismatches in the seed region (see Figure 1A). (**C**) Analysis of adaptation in *P. furiosus* strains with a no target plasmid or a target plasmid with two seed region mismatches. The leader-adjacent region of CRISPR7 was amplified by PCR and the gel images show unexpanded and expanded products with an asterisk and +1, respectively.

main (47,74). To test whether nuclease and helicase activities of Cas3 are both required for interference, we introduced active site mutations at the HD nuclease motif or the Walker B helicase motif, required for adenosine triphosphate binding and hydrolysis, into the plasmid-expressed Cas3 (Supplementary Figure S1). Complete rescue of interference required both intact nuclease and helicase domains, although some plasmid targeting was observed in the helicase mutant (Figure 5A). In naïve and primed adaptation assays, we noted that adaptation efficiency, as inferred from gel band intensity (Figure 5B), was dramatically reduced in the absence of Cas3 and there was no difference in intensity between target and no target plasmid samples, implying that priming did not occur without Cas3. Overall band intensity, along with the target/no target band intensity differences, could be rescued with plasmid-expressed Cas3. A nuclease-defective Cas3 mutant did not rescue adaptation, while a helicase mutant partially restored expanded band intensity but not the target/no target difference. We sequenced new spacers to look for the increase in plasmid-derived spacers for target samples that is indicative of prim-

ing and to determine unique spacer counts. The priming phenotype and normal unique spacer counts were only present when WT Cas3 was expressed (Figure 5C and Supplementary Figure S2). Taken together, these results indicate that Cas3 is essential for primed adaptation in the type I-B system, and additionally has a strong influence on naïve adaptation.

While a role for Cas3 in primed adaptation was not surprising, the effect on naïve adaptation was not anticipated. To help distinguish between a core role of Cas3 in spacer uptake and a secondary modulatory role, we tested adaptation in a ΔCas3 strain with the core adaptation overexpression background (OE Cas1, Cas2, Cas4-1, Cas4-2). As with the single immune effector strains, adaptation efficiency in the ΔCas3 strain was returned to WT-OE levels by Cas1, Cas2, Cas4-1 and Cas4-2 overexpression (Supplementary Table S2). These findings suggest that Cas3 and immune effectors are not central to the process of spacer processing and integration, but may influence availability of pre-spacers or may interact with core adaptation proteins to influence their activity.
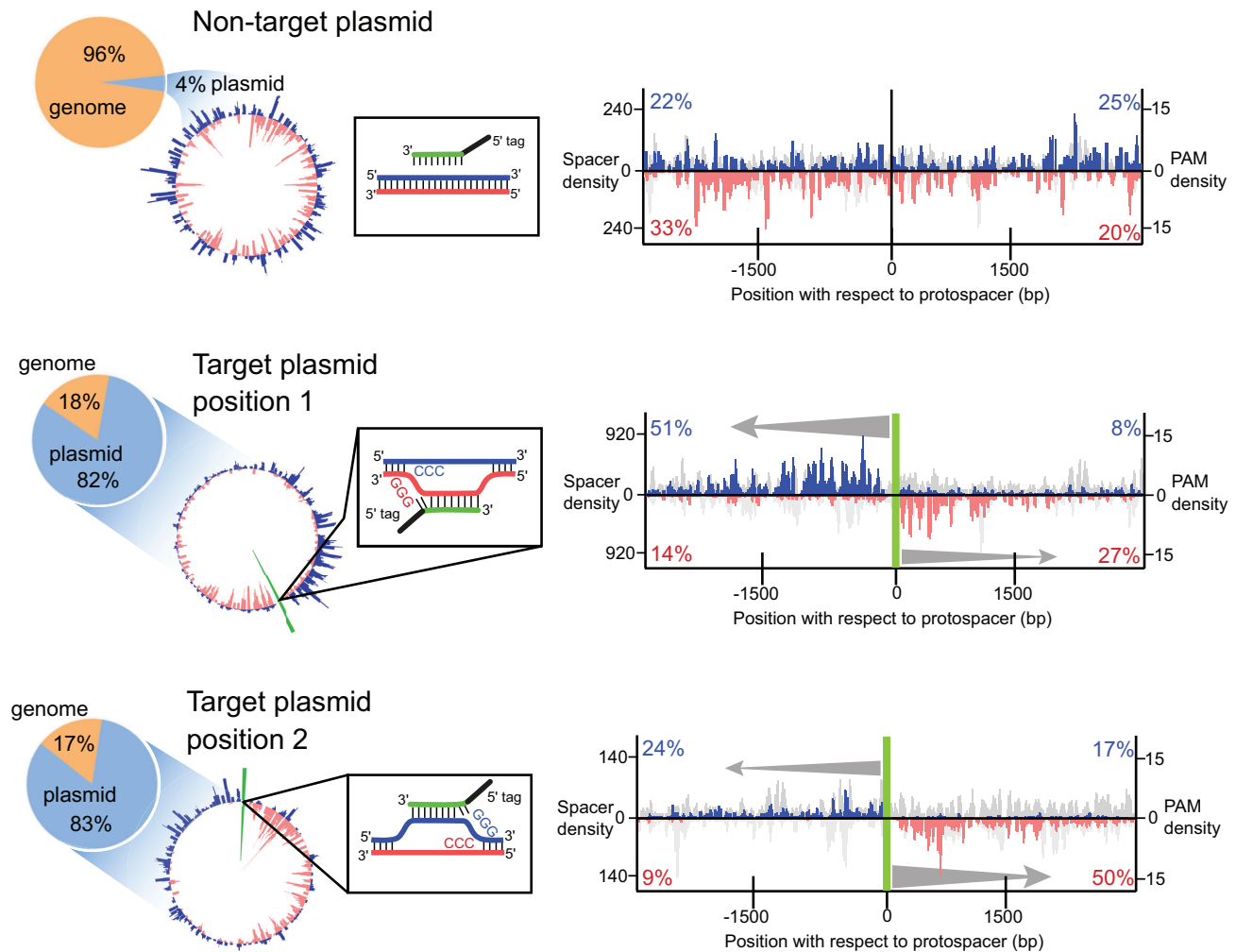
Figure 3. Patterns of spacer uptake correspond to target position. Expanded CRISPR7 and CRISPR5 arrays were amplified from the type I-B single effector strain and sequenced. New spacer sequences were extracted from each read and aligned to both the *Pyrococcus furiosus* genome and the plasmid used in the assay. The percentage of spacers aligning to the plasmid (blue) and the genome (orange) are indicated. Plasmid spacer density tracks for representative samples are shown; for target plasmids, the green bar indicates the location of the protospacer target. Spacers aligning to the plus and minus strand are indicated in blue and red, respectively. Spacer uptake patterns from CRISPR7 and CRISPR5 were highly similar; data from all replicates (11 for non-targeting, eight for position 1 and four for position 2) were pooled to calculate the percentage of new spacers originating from each quadrant of the linearized plasmid (plus strand 5′ of the target, minus strand 5′ of the target, plus strand 3′ of the target and minus strand 3′ of the target, numbers shown in red and blue in their respective quadrants). The same pooled data were also used to generate the spacer-origin pie charts.

Cas4-1 is required for primed adaptation

Cas4 is one of the core Cas proteins and functions in spacer acquisition for diverse CRISPR-Cas systems (75). It was reported that Cas4 is required for primed adaptation in the *H. hispanica* Type I-B system (29) as well as a heterologous Type I-U system (76). *Pyrococcus furiosus* has two Cas4 proteins, Cas-associated Cas4-1 (PF1119) and isolated Cas4-2 (PF1793) (Figure 2A), and we recently reported that both forms are involved in spacer processing and integration (53). To test whether either Cas4-1 or Cas4-2 is required for primed adaptation, we generated Δcas4-1, Δcas4-2 and Δcas4-1/Δcas4-2 strains in the I-B only background. Since Cas4-1 and Cas4-2 deletion reduces adaptation efficiency (53), we used strong promoters upstream of the core adaptation genes (Cas1, Cas2 and Cas4-1 or Cas4-2 if not deleted

in that strain) to drive overexpression. Overexpression increases spacer uptake, allowing us to adequately characterize the effect of Cas4 deletion on priming and previous work demonstrated that adaptation characteristics are largely preserved in the Cas1/Cas2/Cas4-1/Cas4-2 overexpression background (63). As with the WT strain, the OE strain showed strong interference against a plasmid with a perfect target and intermediate interference with a target bearing two mismatches in the seed region (Figure 6A). Deletion of one or both of the *cas4* genes did not affect interference (Figure 6A). Regarding adaptation, we examined both the change in band intensity (Figure 6B) and the proportion of plasmid-derived spacers (Figure 6C) for target versus no target plasmid samples, and found that deletion of *cas4-1* eliminated the priming phenotype. We mapped new spacers to the plasmid reference sequence and com-
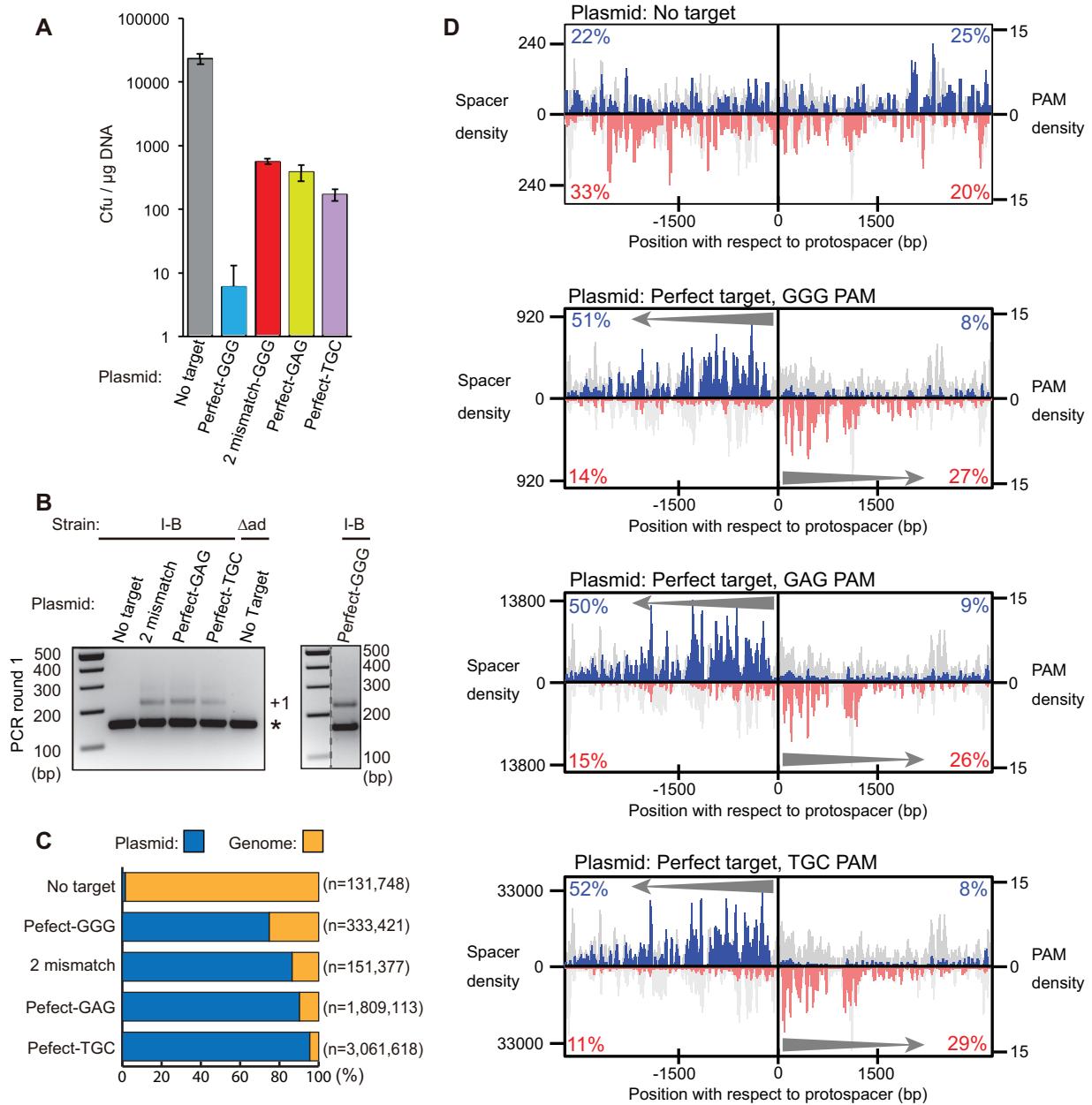
Figure 4. Adaptation and interference in targets with non-canonical PAMs. (**A**) Plot of transformation efficiency for different plasmids: no target, a perfect target with a canonical NGG PAM, a target with two mismatches in the seed region and a canonical NGG PAM (see Figure 1A), and two perfect targets with non-canonical PAMs, GAG and TGC. The two non-canonical PAMs were selected for testing because previous results suggested they yield intermediate levels of interference, a result that was replicated here. All transformations were into the type I-B single effector strain. Colony forming units are plotted ± SEM, ($n = 4$), see Supplementary Table S4 for individual data. (**B**) Expanded CRISPR7 arrays were amplified from the type I-B single effector strain transformed with the plasmids described for part A. PCR products corresponding to parental arrays (unexpanded) and arrays with one new repeat-spacer unit (expanded) are indicated with an asterisk and +1, respectively. Lane Δad corresponds to the negative control strain (adaptation genes deleted). (**C**) Expanded CRISPR7 and CRISPR5 array amplicons were sequenced; new spacer sequences were extracted and aligned to the *Pyrococcus furiosus* genome and appropriate plasmid. Bar graphs show the percentage of new spacers aligning to the genome versus plasmid. Data from CRISPR5 and CRISPR7 were highly similar; counts were pooled for two replicates from each locus (four experiments total). (**D**) Plasmid spacer density tracks for representative samples; blue and pink tracks indicate spacers aligning to the plus and minus strand, respectively. Data from the two replicates and the two loci (CRISPR7 and CRISPR5) were highly similar and thus were pooled to calculate the percentage of new spacers originating from each quadrant of the linearized plasmid (blue and red numbers located in respective quadrant of spacer density plots).

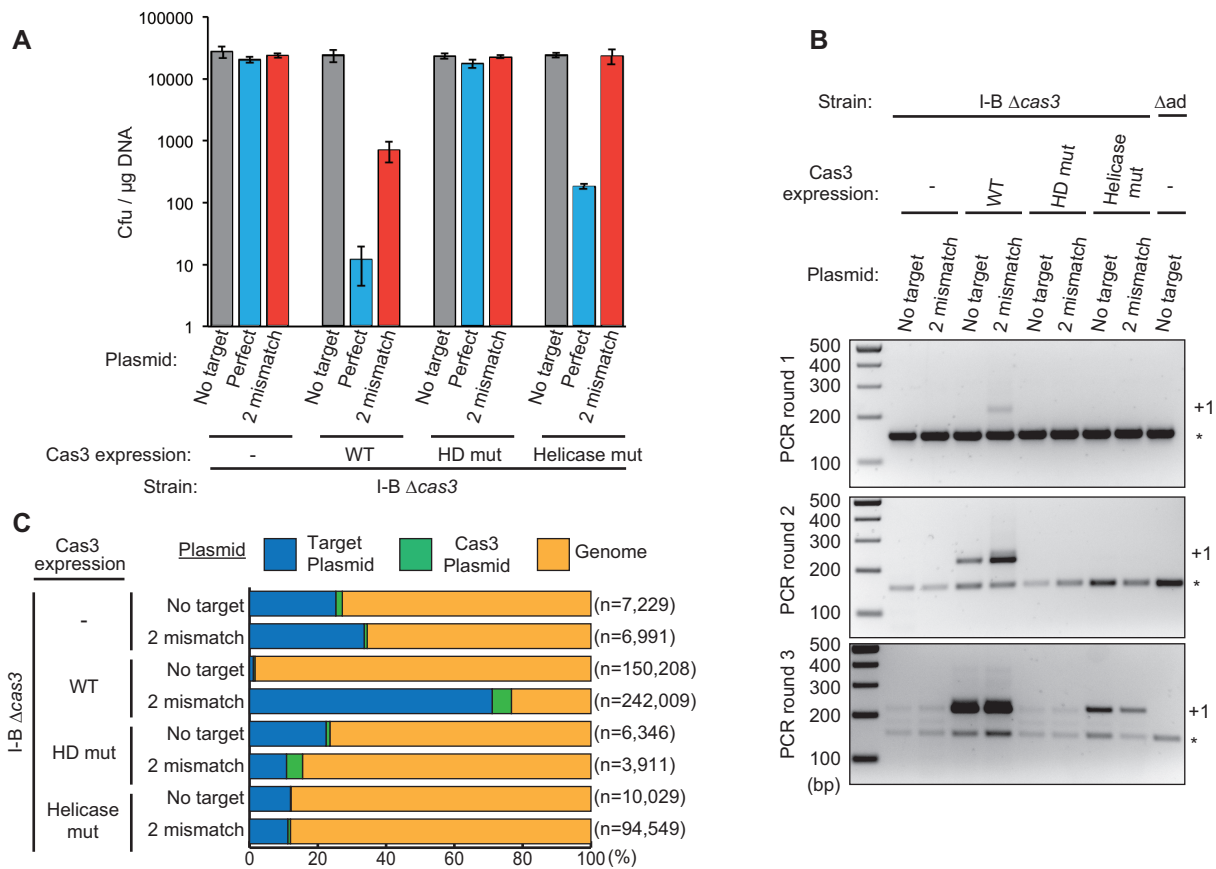**Figure 5.** Adaptation in Cas3 nuclease and helicase mutant strains. (**A**) Plot of transformation efficiency for target and no target plasmids into I-B ΔCas3 strains. Plasmid expression was used to reconstitute Cas3 in either WT or mutant form; substitutions from the literature were made to disrupt either the HD domain (H53A/D54A) or the helicase active site (D369A), see Supplementary Figure S1. Mean colony forming units are plotted ± SEM, ($n = 6$), see Supplementary Table S4 for individual data, (**B**) Analysis of adaptation in I-B ΔCas3 strains. The leader-adjacent region of CRISPR7 was amplified; PCR products corresponding to parental arrays and addition of one repeat-spacer unit are indicated with an asterisk and +1, respectively. (**C**) Spacers from expanded arrays were aligned to both the *Pyrococcus furiosus* genome and the plasmid used in the assay. The percentage of spacers aligning to the target plasmid (blue), OE plasmid (green) and the genome (orange) are shown. Pooled data from four experiments are presented (CRISPR5 and CRISPR7, two replicates each).

pared their distributions for the target versus no target plasmids; comparisons revealed that the distribution of spacers shifted toward the protospacer when Cas4-1 was present (I-B and Δ*cas4-2* strains) but not when it was absent (Δ*cas4-1* and Δ*cas4-1*/Δ*cas4-2* strains) (Supplementary Figure S5). As with the III-B and I-A only strains, a protospacer peak coinciding with the double stranded origin of rolling circle replication was observed for all strains, except when there was priming (Supplementary Figures S4 and 5).

We noted that the shift of spacers to the target site was not as clear in the Δ*cas4-2* strain as in the OE strain. Spacers were clustered around the protospacer position, but the strand bias that was normally associated with priming was absent, and this may be due to a defect in orientation during spacer integration. We previously showed that Cas4-1 defines the PAM for new spacers, Cas4-2 is critical for maintaining PAM-directional orientation during integration, and both Cas4-1 and Cas4-2 are important for processing new spacers to the correct length (53). In the I-B background with and without a priming target, we also found that Cas4-1 and Cas4-2 are necessary for PAM

definition, spacer orientation (Supplementary Figure S5) and spacer size (average size 44 and 43 bp for the Δ*cas4-1*/Δ*cas4-2* strain with and without the priming target, respectively). In the Δ*cas4-2* strain, strand information is lost when new spacers are integrated in random orientation, so strand biases would not be apparent, even if all other aspects of priming were preserved. Together the results indicate that only *P. furiosus* Cas4-1 is necessary for primed adaptation.

Naïve spacer uptake does not require Cas3 or any immune effector complex

Taken together, our experiments showed that Cas3 and the I-B effector complex were not only necessary for primed adaptation, but also appeared to positively influence naïve spacer uptake. Band intensity indicated that adaptation was reduced when Cas3 and I-B were deleted, and although bands were visible after three rounds of PCR, sequencing showed that those bands corresponded to a very small number of new, unique spacers, implying few spacer up-
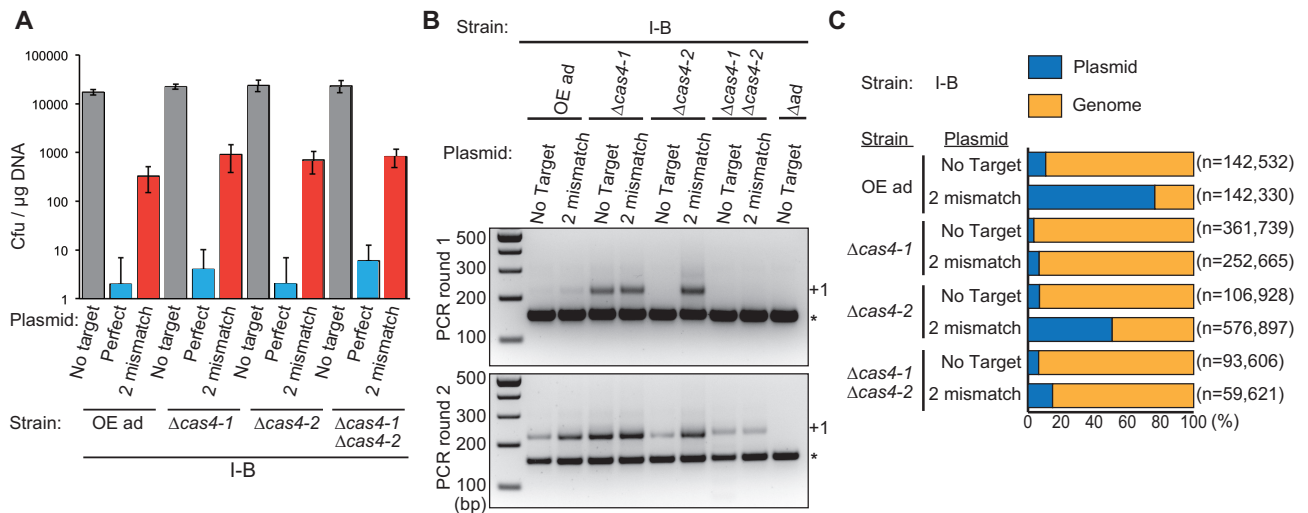
Figure 6. Cas4-1 is required for primed adaptation. (**A**) Plot of transformation efficiency for target and no target plasmids into the I-B strain. In these strains, adaptation proteins (Cas1, Cas2, Cas4-1 and Cas4-2) were over-expressed by strong promoters upstream of each gene; Cas4-1 and Cas4-2 were deleted individually or in combination in the adaptation over-expression strain. Mean colony forming units are plotted ± SEM, (*n* = 6), see Supplementary Table S4 for individual data. (**B**) Analysis of adaptation. The leader-adjacent region of CRISPR7 was amplified; PCR products corresponding to parental arrays and addition of one repeat-spacer unit are indicated with an asterisk and +1, respectively. (**C**) Spacers from expanded arrays were aligned to both the *Pyrococcus furiosus* genome and the plasmid used in the assay. The percentage of spacers aligning to the target plasmid (blue), and the genome (orange) are shown. Pooled data from four experiments are presented (CRISPR5 and CRISPR7, two replicates each).

take events (Supplementary Figure S2). However, when core adaptation proteins (Cas1, Cas2, Cas4-1, Cas4-2) were overexpressed in either the ΔCas3 strain or a strain lacking all immune effector genes, including I-B, (called the ΔcrRNP strain), adaptation efficiency was returned to a level similar to the WT OE samples (Supplementary Figure S2), suggesting that Cas3 and the I-B effector are not strictly necessary for adaptation. Instead, they may influence adaptation by interacting, directly or indirectly, with Cas1, Cas2, or Cas4-1 components.

We next evaluated adaptation in a strain in which *cas3* and genes of all the three immune effectors (I-A, I-B and III-B) were deleted (ΔcrRNP/ΔCas3) (Figure 7). Surprisingly, band intensity suggested that spacer uptake levels were higher in the ΔcrRNP/ΔCas3 strain as compared to the individual strains ΔCas3 and ΔcrRNP (Figure 7A). We characterized the new spacers in the ΔcrRNP/ΔCas3 strain and found that they were functionally similar to WT spacers in that they were typical in size (37 bp; Figure 7B) and had the correct, 5′-NGG-3′ PAM sequence (Figure 7C). The unique spacer counts for the ΔcrRNP/ΔCas3 strain were not statistically different from WT (Supplementary Figure S2), in agreement with the similar band intensities. These results implied that *cas3* and immune effector genes are dispensable for normal naïve adaptation. However, both Cas3 and crRNP immune effectors were essential for priming, consistent with our other results presented here. Without *cas3* and crRNP genes, the number of plasmid-derived spacers actually decreased for a target plasmid (Figure 7D). While new spacers from the ΔcrRNP/ΔCas3 strain closely resembled those of the WT strain (Figure 7B and C), we did note that the spacer distribution patterns differed. The ΔcrRNP/ΔCas3 strain appeared to lose most of the spacer-uptake hotspots (regions with a high density of protospacers that exceeds what would be expected based on PAM fre-

quency alone) that we had previously identified for *P. furiosus* (63), with the exception of transposons (Supplementary Figure S6) and rolling circle plasmid origin of replication (Supplementary Figure S4). Based on these observations, we speculate that Cas3 and the effector complexes play a role in generating, processing, or recruiting pre-spacer DNA.

DISCUSSION

Here, we have characterized both primed and interference-independent adaptation for *P. furiosus*. Like many other bacteria and archaea, *P. furiosus* has multiple CRISPR systems, but unlike other CRISPR-bearing organisms, there is functional overlap among the three CRISPR effector complexes (I-A, I-B and III-B) of this organism in that crRNAs from any of the seven CRISPR arrays can associate and function with any of the three effector complexes (45). Our results show that of the three crRNP effector complexes, only I-B participates in priming, with Cas3 and the adaptation proteins Cas1, Cas2 and Cas4-1 also playing essential roles. However, once these primed spacers are incorporated into an array, they would then be available for defense via I-A, I-B or III-B mediated interference.

One hypothesis for evolution of both primed adaptation and the existence of multiple CRISPR systems within a single host is that they provide backup in the face of viral counter-defenses. Mutations in either the PAM sequence or the protospacer can reduce or prevent interference and allow a virus to escape CRISPR-mediated targeting (Figures 1 and 3) and (17,23,25–27). By updating the CRISPR array when a partial match is detected, priming can overcome the viral escape mutations. Additionally, some CRISPR systems appear more tolerant of point mutations and can therefore limit escape (77). For example, horizontal trans-
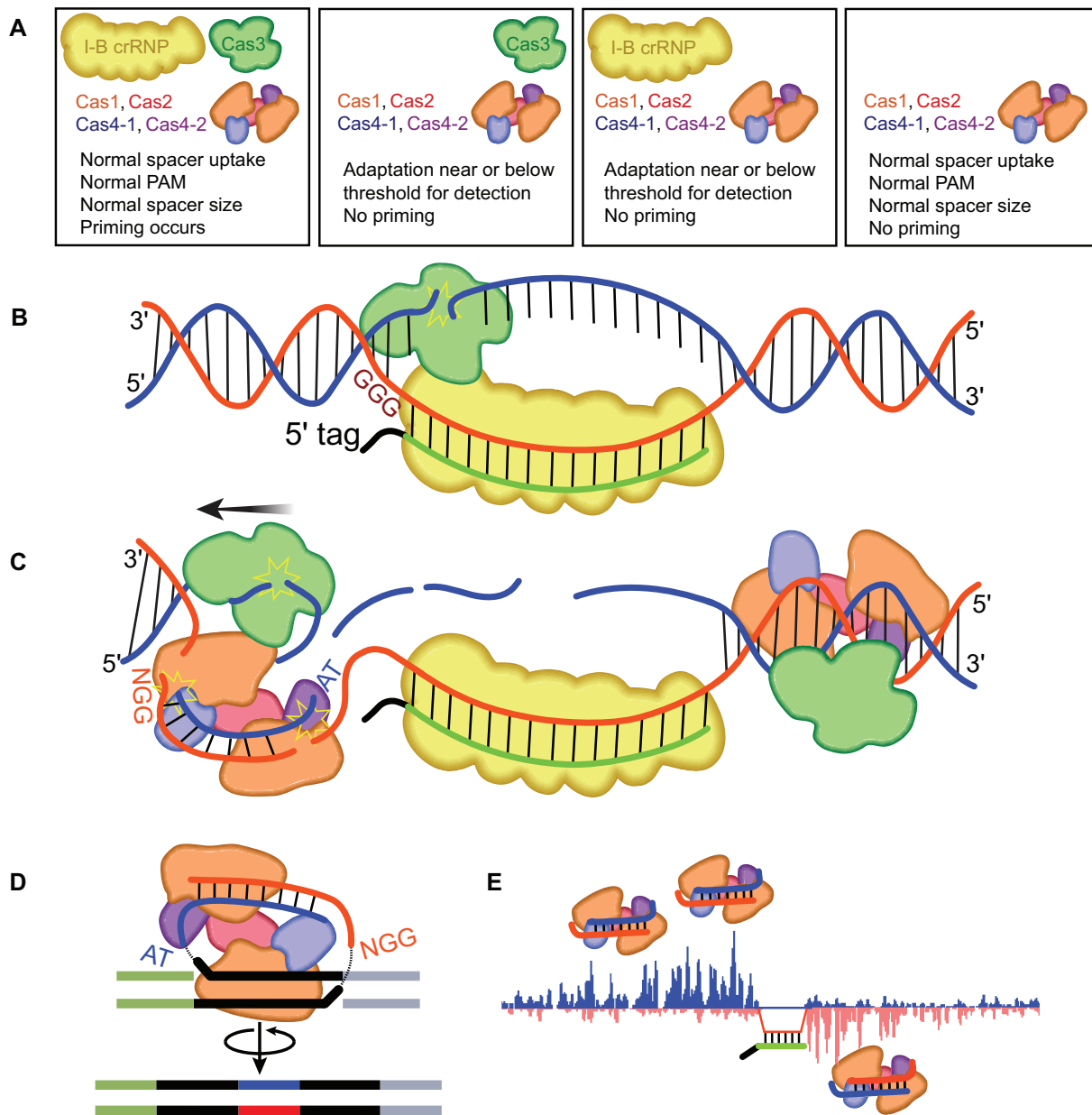
Figure 7. Adaptation in the absence of crRNPs and Cas3. (**A**) Analysis of adaptation. The leader-adjacent region of CRISPR5 was amplified; PCR products corresponding to unexpanded and expanded arrays are indicated with an asterisk and +1, respectively. Band intensity suggested that adaptation was equally efficient in the ΔcrRNP/ΔCas3 strain as in WT; expanded arrays were sequenced and new spacers were characterized. (**B**) Bubble plots show size distributions for all uniquely aligned new spacers from sequenced gel bands. Distributions show pooled data from CRISPR5 and CRISPR7 experiments, with a total of nine replicates for WT and 15 replicates for ΔcrRNP/ΔCas3. (**C**) Consensus upstream and downstream sequences for all uniquely aligned new spacers. A canonical upstream NGG PAM is observed for both WT and ΔcrRNP/ΔCas3. (**D**) Percentage of plasmid-aligned spacers in the WT and ΔcrRNP/ΔCas3 strains when transformed with either a no target plasmid or a plasmid with two mismatches in the target seed region. The percentage of spacers aligning to the target plasmid (blue), and the genome (orange) are shown. Data from CRISPR5 and CRISPR7 experiments are pooled; between six and 13 replicates were included for each treatment.

fer of type III-B systems into *Marinomonas mediterranea* strains that already contain a type I-F system appears to broaden targeting capabilities because the relatively promiscuous type III-B system can use I-F crRNAs in defense (78). Viruses can also resist interference through encoding anti-CRISPR proteins, which are usually small proteins that work by binding and disrupting activity of Cas proteins involved in one or more steps of immunity (79–82). Multiple co-existing CRISPR-Cas systems likely provide a way to resist anti-CRISPRs, since divergent *cas* genes and mechanisms make it unlikely that any single anti-CRISPR can inactivate all forms of interference. In *P. furiosus*, priming is carried out by the I-B effector complex and then shared with the functionally distinct I-A and III-B complexes, so the two avenues for backup defence against evolving viruses are essentially intermingled, potentially strengthening the whole suite to provide robust immunity.

The balance of intermingled defense systems may have some drawbacks. A given crRNA that is produced in *P. furiosus* is shunted into one of the three systems (45). Since only the I-B system can prime, crRNAs shunted to the I-A or III-B complexes are lost opportunities for priming. This dilution is likely why priming in our data is so much more apparent for the I-B only strain as compared to WT (Figure 2). Combining the unique capabilities from different systems, for example broad PAM tolerance in the III-B combined with priming in the I-B, may come at a cost to their individual contributions and may set an upper limit on the number of co-existing CRISPR systems.

Cas3 and I-B effector crRNP influence adaptation

When Cas3 or the I-B effector complex were deleted individually, we observed a dramatic decrease in band intensity for expanded array PCR products and a large drop in the number of unique new spacers relative to the WT *P. furiosus*, implying that there were very few adaptation events in the cultures (Figure 2 and Supplementary Figure S2). These findings suggested that both Cas3 and the I-B crRNP effector complex could be involved in adaptation. However, when we combined these individual deletions with overexpression of the core adaptation genes (*cas1*, *cas2*, *cas4-1*, *cas4-2*), spacer uptake returned to levels comparable to the WT strain with adaptation gene overexpression (Supplementary Figure S2). This showed that the effects of Cas3 and I-B crRNP on adaptation were experimentally conditional, and perhaps indirect. When both *cas3* and genes of the three (I-B, I-A and III-B) crRNP effector complexes were deleted, we saw that the intensity of the expanded array PCR band was comparable to WT, the counts of new unique spacers were similar to WT and the characteristics of the new unique spacers (PAM and size distribution) were also the same as in the WT strain (Figure 7 and Supplementary Figure S2). While the loss of either the effector complexes or Cas3 could inhibit adaptation, when both were lost, adaptation returned to normal (Figure 8). One possible explanation is that there are physical interactions among the adaptation proteins, Cas3 and the I-B complex and that these interactions can influence naive adaptation in addition to primed. This would be consistent with observations

Figure 8. Summary and model for effects of I-B effector crRNP and Cas3 on adaptation. (**A**) Summary of how deleting adaptation and interference components influences naïve and primed adaptation. (**B**) Model of initial steps of interference by Cas3 and the type I-B effector complex. (**C**) Model of interference-dependent pre-spacer generation by Cas3 and the adaptation complex (Cas1, Cas2, Cas4-1 and Cas4-2). (**D**) Model depicting integration of interference-generated pre-spacers. (**E**) Visual representation of new spacers acquired from a plasmid bearing a protospacer target; blue and red tracks show the density of protospacers acquired from the plus and minus strands, respectively. Half of all new spacers acquired from a target plasmid arose from the non-target strand on the PAM-proximal side of the plasmid, potentially reflecting the activity of Cas3 diagrammed in C.

from multiple recent studies. For example, the natural fusion of Cas2 and Cas3 into a single protein in the type I-F system implies that an important functional coupling exists for the two proteins, that are normally part of adaptation (Cas2) and downstream interference (Cas3) pathways. The Cas2/3 fusion protein forms a complex with Cas1 to carry out adaptation (37,83), and also interacts with the I-F (Csy) crRNP effector complex to carry out interference (43). *In vitro* experiments revealed that Cas1 and I-F crRNPs are opposing regulators of Cas2/3 nuclease activity, with Cas1 reducing the ability of Cas2/3 to degrade target DNA and I-

F crRNPs reversing that inhibition (43). Interference-driven adaptation is the primary source of new type I-F spacers detected in spacer uptake assays (84), and the patterns of naturally occurring spacers in type I-F systems suggests widespread priming (85). This is consistent with this fusion and close coupling between interference and adaptation proteins (43,84). However, naïve adaptation also occurs, albeit at a lower level (56,84), so functional coupling between adaptation and interference proteins may be important beyond a role in supplying Cas1 with interference-generated DNA fragments.

Evidence for interactions between crRNP effector complexes and adaptation complexes also appears in systems without fused Cas2/Cas3 proteins. In type I-E systems from both *Thermobifida fusca* and *E. coli*, Cas3, the I-E effector crRNP complex and the Cas1/Cas2 complex all co-assemble on target DNA and can translocate together along that DNA (60,61). Each component (Cas3, Cas1/Cas2 and crRNP effector complex) can modify target DNA binding affinity and activity of the others (60,61). Even in the absence of a protospacer target, which would be the context for naïve adaptation, the effector crRNP can interact with Cas1/Cas2, but this interaction is modulated by Cas3 (60). Extending these observations to our results, we speculate that, individually, Cas3 and the effector complexes can interact with Cas1/Cas2 (likely in complex with the two Cas4 proteins) in a way that inhibits Cas1/Cas2 spacer uptake, but when they are both present, the interaction between Cas3 and the crRNP with Cas1/Cas2 is modified and no longer inhibitory (Figure 8). Interactions between the effector complex and the adaptation complex could mean that adaptation is always poised and ready for priming, and could also act to constrain the adaptation complex and protect the cell from the potential for damage posed by free-roaming Cas4 nucleases.

Another potential explanation for the conditional effects of Cas3 and the I-B effector complex on adaptation is that these proteins influence the upstream process of pre-spacer generation. It is clear that Cas3 is essential for priming (Figure 5, (23,29,56)) and that the DNA fragments produced by Cas3 during interference are suitable material for generating functional pre-spacers (49). In *P. furiosus*, Cas3 may also help produce nicked or unwound DNA in a protospacer-independent manner. *In vitro*, the I-F Cas2/3 fusion protein could cleave DNA bubbles, which mimic R-loops, even in the absence of a crRNA (43). Cas1 largely, but not entirely, abolished this activity unless a target-bound I-F crRNP complex was also added (43). Cas3 has been demonstrated to possess both nuclease and helicase activity *in vitro* when partially duplex DNA or RNA and DNA are provided (47,86). Cas3 also appears to regulate replication of certain plasmids, potentially by affecting binding of the priming RNAs (87). We speculate that Cas3 may also cleave endogenous, crRNA-independent R-loops *in vivo* in *P. furiosus*, perhaps at a low level, and that these events may help Cas1/Cas2/Cas4-1/Cas4-2 to convert the nicked DNA into new spacers. We previously noted that highly sampled genomic or invader DNA protospacers (i.e. protospacer hotspots) in *P. furiosus* mapped to DNA loci where R-loops, single-stranded (ss) DNA intermediates, or nicked or broken DNA are expected (63). If Cas3 contributes to R-loop nicking and unwinding of nicked DNA, it would be consistent with our observation that some protospacer hotspots are lost for the ΔCas3/ΔcrRNP strain. For those protospacer hotspots that remain when Cas3 is absent, namely the relatively narrow hotspots around transposons and the double stranded origin of replication of a rolling circle plasmid (Supplementary Figures S4 and 6), the activity of Cas3 may be dispensable since transposases and replication proteins can carry out, or recruit factors to carry out, their own nicking and DNA unwinding/melting (88,89).

Cas4-1 participates in primed adaptation

Cas4 is one of the core adaptation proteins in many CRISPR-Cas systems (75), a role supported by several lines of evidence: the frequent close proximity of the *cas1*, *cas2* and *cas4* genes, the existence of a Cas1/Cas4 fusion gene, and recent functional data (50–53,76,90). We previously showed that both of the *cas4* genes in *P. furiosus*, *cas4-1* and *cas4-2*, are necessary for integration of PAM-proximal, correctly sized and correctly oriented spacers (53). In our current study, we find that Cas4-1, but not Cas4-2, is also critical for priming. How might this occur? Structural work has shown that Cas4 forms a complex with Cas1 and Cas2 in the presence of duplex or partially duplex DNA (90). These structures, together with functional data, led to a model wherein Cas4, Cas1 and Cas2 come together around dsDNA in a 1:4:2 ratio, Cas4 sequesters 3′ overhangs on that DNA unless and until it detects a PAM in the overhang, Cas4 cleaves the ssDNA precisely at the PAM, then passes off the newly PAM-cleaved DNA end to the Cas1 integrase active site for integration into the CRISPR array (90). In the cleavage data used to generate this model, Cas4 only cleaved ssDNA, and so another, unidentified nuclease is presumed to unwind DNA to generate the ssDNA that Cas4 then processes. In our system, we speculate that Cas3 generates some of these Cas4 substrates in the course of interference and makes them readily available to Cas4 through their common interactions with Cas1 and Cas2 (Figure 8). In the Cas4-1 deletion strain, Cas3 may unwind dsDNA but Cas4-1 is not available to efficiently cleave the ssDNA and transfer the product to Cas1, and so a chain of events that promotes primed adaptation is disrupted.

Cas3 3′-5′ DNA reeling is consistent with the observed non-target strand bias

When the crRNP effector complex interacts with protospacer target DNA, this results in an R-loop structure wherein the crRNA is base-paired with the target strand while the other strand (non-target strand) is displaced and is unpaired (Figure 8B). We observed that about half of all new spacers arising from the target plasmids during primed adaptation, were from the non-target (displaced) strand of the plasmid, on the PAM-side of the protospacer (Figures 3, 4D and 8E). We believe this strand bias is consistent with data (61,91) on how Cas3 reels and unwinds DNA in the 3′ to 5′ direction (Figure 8B and C). Once recruited to a target-engaged effector complex, Cas3 carries out an endonucleolytic nick of the displaced DNA strand and can unwind DNA on the non-target strand, moving away from the PAM and producing stretches of single-stranded (ss)DNA (61). Cas1 and Cas2 can also assemble with Cas3 at the target-engaged effector complex in these systems, and may even promote Cas3 recruitment (61). Extending these observations to *P. furiosus*, we could expect that Cas4-1 and Cas4-2, together with Cas1 and Cas2, would localize to the target-engaged Cas3 and I-B crRNP effector complex (Figure 8C). We speculate that Cas3 unwinds DNA and periodically makes endonucleolytic cuts on the non-target DNA strand, thereby producing multiple ssDNA fragments, which then re-anneal with the target DNA strand (49). These products could act as suitable material for the adaptation complex,

with Cas4-1 and Cas4-2 (or possibly Cas1) carrying out endonucleolytic cleavage of the target strand and exonucleolytic trimming to yield a partially double stranded pre-spacer with NGG (PAM) and AT overhangs (Figure 8C). The pre-spacers would then be ready for integration into the CRISPR array (Figure 8D). We also noted a tapered peak of spacers on the PAM-distal target strand, though it was not as prominent as the PAM-adjacent peak (Figures 3, 4D and 8E). A previous report indicated that when Cas1 and Cas2 are recruited to a mutant PAM target, they permitted Cas3 to translocate in either direction away from the target-engaged protospacer (61). The secondary peak we observe may reflect a similar bi-directional movement during priming (Figure 8C and E) but the details of pre-spacer generation are unclear.

We observed the same pattern of spacer distribution for plasmids with a canonical versus non-canonical PAM by the protospacer, even though interference was reduced by about two orders of magnitude in the non-canonical PAM samples. If Cas3 underlies this protospacer pattern, it would appear that it is still recruited to the non-canonical PAM target. This would be in agreement with recent studies, which found that Cas3 was recruited to the interference complex both in the presence and absence of a typical PAM (32,61). These studies presented evidence for distinct PAM-dependent versus PAM-independent conformations in the interference complex; our data and recent work by others (92,93) suggest that if there are two conformations, they must both support Cas3-dependent priming.

Our work highlights the modular-yet-interconnected nature of CRISPR-Cas systems. Although the adaptation complex could function alone, i.e. functionally competent spacers could be efficiently acquired in the absence of Cas3 and all immune effector genes (Figure 8), spacer uptake was clearly influenced by the interference modules (Figures 2 and 5; Supplementary Figure S6), raising the interesting possibility of physical interaction between the adaptation and interference complexes. Moreover, while Cas3 has a well-established role in crRNA-mediated interference, its effect on naïve adaptation in *P. furiosus* suggested that it may also facilitate the supply of spacers from non-target DNA, through unknown mechanisms. Further work could help provide a mechanistic explanation for the effect of Cas3 deletion on naïve adaptation as well as lead to a detailed molecular understanding of I-B mediated, primed adaptation.

DATA AVAILABILITY

Sequence data were deposited in the NCBI Sequence Read Archive under the BioProject ID PRJNA603216.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

REFERENCES

1. 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.
2. Abudayyeh,O.O., Gootenberg,J.S., Konermann,S., Joung,J., Slaymaker,I.M., Cox,D.B., Shmakov,S., Makarova,K.S., Semenova,E., Minakhin,L. *et al.* (2016) C2c2 is a single-component programmable RNA-guided RNA-targeting CRISPR effector. *Science*, **353**, aaf5573.
3. 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.
4. 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.
5. Elmore,J.R., Sheppard,N.F., Ramia,N., Deighan,T., Li,H., Terns,R.M. and Terns,M.P. (2016) Bipartite recognition of target RNAs activates DNA cleavage by the Type III-B CRISPR-Cas system. *Genes Dev.*, **30**, 447–459.
6. 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.
7. 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.
8. Zetsche,B., Gootenberg,J.S., Abudayyeh,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.
9. Makarova,K.S., Wolf,Y.I., Iranzo,J., Shmakov,S.A., Alkhnbashi,O.S., Brouns,S.J.J., Charpentier,E., Cheng,D., Haft,D.H., Horvath,P. *et al.* (2019) Evolutionary classification of CRISPR-Cas systems: a burst of class 2 and derived variants. *Nat. Rev. Microbiol.*, **18**, 67–83.
10. Grissa,I., Vergnaud,G. and Pourcel,C. (2007) The CRISPRdb database and tools to display CRISPRs and to generate dictionaries of spacers and repeats. *BMC Bioinformatics*, **8**, 172.
11. Mojica,F.J., Diez-Villasenor,C., Garcia-Martinez,J. and Soria,E. (2005) Intervening sequences of regularly spaced prokaryotic repeats derive from foreign genetic elements. *J. Mol. Evol.*, **60**, 174–182.
12. Alkhnbashi,O.S., Shah,S.A., Garrett,R.A., Saunders,S.J., Costa,F. and Backofen,R. (2016) Characterizing leader sequences of CRISPR loci. *Bioinformatics*, **32**, i576–i585.
13. Wei,Y., Chesne,M.T., Terns,R.M. and Terns,M.P. (2015) Sequences spanning the leader-repeat junction mediate CRISPR adaptation to phage in Streptococcus thermophilus. *Nucleic Acids Res.*, **43**, 1749–1758.
14. 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.
15. Hille,F., Richter,H., Wong,S.P., Bratovic,M., Ressel,S. and Charpentier,E. (2018) The biology of CRISPR-Cas: Backward and forward. *Cell*, **172**, 1239–1259.
16. Jackson,R.N., van Erp,P.B., Sternberg,S.H. and Wiedenheft,B. (2017) Conformational regulation of CRISPR-associated nucleases. *Curr. Opin. Microbiol.*, **37**, 110–119.
17. Deveau,H., Barrangou,R., Garneau,J.E., Labonte,J., Fremaux,C., Boyaval,P., Romero,D.A., Horvath,P. and Moineau,S. (2008) Phage

response to CRISPR-encoded resistance in Streptococcus thermophilus. *J. Bacteriol.*, **190**, 1390–1400.

18. 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.

19. Shah,S.A., Erdmann,S., Mojica,F.J. and Garrett,R.A. (2013) Protospacer recognition motifs: mixed identities and functional diversity. *RNA Biol.*, **10**, 891–899.

20. Amitai,G. and Sorek,R. (2016) CRISPR-Cas adaptation: insights into the mechanism of action. *Nat. Rev. Microbiol.*, **14**, 67–76.

21. McGinn,J. and Marraffini,L.A. (2019) Molecular mechanisms of CRISPR-Cas spacer acquisition. *Nat. Rev. Microbiol.*, **17**, 7–12.

22. Sternberg,S.H., Richter,H., Charpentier,E. and Qimron,U. (2016) Adaptation in CRISPR-Cas systems. *Mol. Cell*, **61**, 797–808.

23. 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.

24. Swarts,D.C., Mosterd,C., van Passel,M.W. and Brouns,S.J. (2012) CRISPR interference directs strand specific spacer acquisition. *PLoS One*, **7**, e35888.

25. Jiang,F., Zhou,K., Ma,L., Gressel,S. and Doudna,J.A. (2015) STRUCTURAL BIOLOGY. A Cas9-guide RNA complex preorganized for target DNA recognition. *Science*, **348**, 1477–1481.

26. Semenova,E., Jore,M.M., Datsenko,K.A., Semenova,A., Westra,E.R., Wanner,B., van der Oost,J., Brouns,S.J. and Severinov,K. (2011) Interference by clustered regularly interspaced short palindromic repeat (CRISPR) RNA is governed by a seed sequence. *Proc. Natl. Acad. Sci. U.S.A.*, **108**, 10098–10103.

27. 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.

28. Fineran,P.C., Gerritzen,M.J., Suarez-Diez,M., Kunne,T., Boekhorst,J., van Hijum,S.A., Staals,R.H. and Brouns,S.J. (2014) Degenerate target sites mediate rapid primed CRISPR adaptation. *Proc. Natl. Acad. Sci. U.S.A.*, **111**, E1629–E1638.

29. Li,M., Wang,R., Zhao,D. and Xiang,H. (2014) Adaptation of the Haloarcula hispanica CRISPR-Cas system to a purified virus strictly requires a priming process. *Nucleic Acids Res.*, **42**, 2483–2492.

30. Richter,C., Dy,R.L., McKenzie,R.E., Watson,B.N., Taylor,C., Chang,J.T., McNeil,M.B., Staals,R.H. and Fineran,P.C. (2014) Priming in the Type I-F CRISPR-Cas system triggers strand-independent spacer acquisition, bi-directionally from the primed protospacer. *Nucleic Acids Res.*, **42**, 8516–8526.

31. Semenova,E., Savitskaya,E., Musharova,O., Strotskaya,A., Vorontsova,D., Datsenko,K.A., Logacheva,M.D. and Severinov,K. (2016) Highly efficient primed spacer acquisition from targets destroyed by the Escherichia coli type I-E CRISPR-Cas interfering complex. *Proc. Natl. Acad. Sci. U.S.A.*, **113**, 7626–7631.

32. Xue,C., Seetharam,A.S., Musharova,O., Severinov,K., Brouns,S.J., Severin,A.J. and Sashital,D.G. (2015) CRISPR interference and priming varies with individual spacer sequences. *Nucleic Acids Res.*, **43**, 10831–10847.

33. Arslan,Z., Hermanns,V., Wurm,R., Wagner,R. and Pul,U. (2014) Detection and characterization of spacer integration intermediates in type I-E CRISPR-Cas system. *Nucleic Acids Res.*, **42**, 7884–7893.

34. Nunez,J.K., Harrington,L.B., Kranzusch,P.J., Engelman,A.N. and Doudna,J.A. (2015) Foreign DNA capture during CRISPR-Cas adaptive immunity. *Nature*, **527**, 535–538.

35. Wang,J., Li,J., Zhao,H., Sheng,G., Wang,M., Yin,M. and Wang,Y. (2015) Structural and mechanistic basis of PAM-Dependent spacer acquisition in CRISPR-Cas systems. *Cell*, **163**, 840–853.

36. Nunez,J.K., Lee,A.S., Engelman,A. and Doudna,J.A. (2015) Integrase-mediated spacer acquisition during CRISPR-Cas adaptive immunity. *Nature*, **519**, 193–198.

37. Fagerlund,R.D., Wilkinson,M.E., Klykov,O., Barendregt,A., Pearce,F.G., Kieper,S.N., Maxwell,H.W.R., Capolupo,A., Heck,A.J.R., Krause,K.L. *et al.* (2017) Spacer capture and integration by a type I-F Cas1-Cas2-3 CRISPR adaptation complex. *Proc. Natl. Acad. Sci. U.S.A.*, **114**, E5122–E5128.

38. Nunez,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.

39. Wright,A.V., Liu,J.J., Knott,G.J., Doxzen,K.W., Nogales,E. and Doudna,J.A. (2017) Structures of the CRISPR genome integration complex. *Science*, **357**, 1113–1118.

40. Levy,A., Goren,M.G., Yosef,I., Auster,O., Manor,M., Amitai,G., Edgar,R., Qimron,U. and Sorek,R. (2015) CRISPR adaptation biases explain preference for acquisition of foreign DNA. *Nature*, **520**, 505–510.

41. 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.

42. 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.

43. 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.

44. 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.

45. Majumdar,S., Zha,P., Pfister,N.T., Compton,M., Olson,S., Glover,C.V.C. III, Wells,L., Graveley,B.R., Terns,R.M. and Terns,M.P. (2015) Three CRISPR-Cas immune effector complexes coexist in Pyrococcus furiosus. *RNA*, **21**, 1147–1158.

46. 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.

47. 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.

48. 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.

49. Kunne,T., Kieper,S.N., Bannenberg,J.W., Vogel,A.I., Miellet,W.R., Klein,M., Depken,M., Suarez-Diez,M. and Brouns,S.J. (2016) Cas3-derived target DNA degradation fragments fuel primed CRISPR adaptation. *Mol. Cell*, **63**, 852–864.

50. Kieper,S.N., Almendros,C., Behler,J., McKenzie,R.E., Nobrega,F.L., Haagsma,A.C., Vink,J.N.A., Hess,W.R. and Brouns,S.J.J. (2018) Cas4 facilitates PAM-compatible spacer selection during CRISPR Adaptation. *Cell Rep.*, **22**, 3377–3384.

51. Lee,H., Zhou,Y., Taylor,D.W. and Sashital,D.G. (2018) Cas4-dependent prespacer processing ensures high-fidelity programming of CRISPR arrays. *Mol. Cell*, **70**, 48–59.

52. Rollie,C., Graham,S., Rouillon,C. and White,M.F. (2018) Prespacer processing and specific integration in a Type I-A CRISPR system. *Nucleic Acids Res.*, **46**, 1007–1020.

53. Shiimori,M., Garrett,S.C., Graveley,B.R. and Terns,M.P. (2018) Cas4 nucleases define the PAM, length, and orientation of DNA fragments integrated at CRISPR Loci. *Mol. Cell*, **70**, 814–824.

54. Zhang,Z., Pan,S., Liu,T., Li,Y. and Peng,N. (2019) Cas4 nucleases can effect specific integration of CRISPR spacers. *J. Bacteriol.*, **201**, e00747-18.

55. Heler,R., Samai,P., Modell,J.W., Weiner,C., Goldberg,G.W., Bikard,D. and Marraffini,L.A. (2015) Cas9 specifies functional viral targets during CRISPR-Cas adaptation. *Nature*, **519**, 199–202.

56. Vorontsova,D., Datsenko,K.A., Medvedeva,S., Bondy-Denomy,J., Savitskaya,E.E., Pougach,K., Logacheva,M., Wiedenheft,B., Davidson,A.R., Severinov,K. *et al.* (2015) Foreign DNA acquisition by the I-F CRISPR-Cas system requires all components of the interference machinery. *Nucleic Acids Res.*, **43**, 10848–10860.

57. Wei,Y., Terns,R.M. and Terns,M.P. (2015) Cas9 function and host genome sampling in Type II-A CRISPR-Cas adaptation. *Genes Dev.*, **29**, 356–361.

58. Rao,C., Chin,D. and Ensminger,A.W. (2017) Priming in a permissive type I-C CRISPR-Cas system reveals distinct dynamics of spacer acquisition and loss. *RNA*, **23**, 1525–1538.

59. Nussenzweig,P.M., McGinn,J. and Marraffini,L.A. (2019) Cas9 cleavage of viral genomes primes the acquisition of new immunological memories. *Cell Host Microbe*, **26**, 515–526.

60. 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.

61. 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.

62. Grainy,J., Garrett,S., Graveley,B.R. and Terns,M.P. (2019) CRISPR repeat sequences and relative spacing specify DNA integration by Pyrococcus furiosus Cas1 and Cas2. *Nucleic Acids Res.*, **47**, 7518–7531.

63. Shiimori,M., Garrett,S.C., Chambers,D.P., Glover,C.V.C. 3rd, Graveley,B.R. and Terns,M.P. (2017) Role of free DNA ends and protospacer adjacent motifs for CRISPR DNA uptake in Pyrococcus furiosus. *Nucleic Acids Res.*, **45**, 11281–11294.

64. 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.

65. Hale,C.R., Majumdar,S., Elmore,J., Pfister,N., Compton,M., Olson,S., Resch,A.M., Glover,C.V. 3rd, Graveley,B.R., Terns,R.M. *et al.* (2012) Essential features and rational design of CRISPR RNAs that function with the Cas RAMP module complex to cleave RNAs. *Mol. Cell*, **45**, 292–302.

66. Haft,D.H., Selengut,J., Mongodin,E.F. and Nelson,K.E. (2005) A guild of 45 CRISPR-associated (Cas) protein families and multiple CRISPR/Cas subtypes exist in prokaryotic genomes. *PLoS Comput. Biol.*, **1**, e60.

67. Vestergaard,G., Garrett,R.A. and Shah,S.A. (2014) CRISPR adaptive immune systems of Archaea. *RNA Biol.*, **11**, 156–167.

68. Terns,R.M. and Terns,M.P. (2013) The RNA- and DNA-targeting CRISPR-Cas immune systems of Pyrococcus furiosus. *Biochem. Soc. Trans.*, **41**, 1416–1421.

69. Lipscomb,G.L., Stirrett,K., Schut,G.J., Yang,F., Jenney,F.E. Jr, Scott,R.A., Adams,M.W. and Westpheling,J. (2011) Natural competence in the hyperthermophilic archaeon Pyrococcus furiosus facilitates genetic manipulation: construction of markerless deletions of genes encoding the two cytoplasmic hydrogenases. *Appl. Environ. Microbiol.*, **77**, 2232–2238.

70. Farkas,J., Stirrett,K., Lipscomb,G.L., Nixon,W., Scott,R.A., Adams,M.W. and Westpheling,J. (2012) Recombinogenic properties of Pyrococcus furiosus strain COM1 enable rapid selection of targeted mutants. *Appl. Environ. Microbiol.*, **78**, 4669–4676.

71. Quinlan,A.R. (2014) BEDTools: the swiss-army tool for genome feature analysis. *Curr. Protoc. Bioinformatics*, **47**, 11–34.

72. 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.

73. 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.

74. 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.

75. Hudaiberdiev,S., Shmakov,S., Wolf,Y.I., Terns,M.P., Makarova,K.S. and Koonin,E.V. (2017) Phylogenomics of Cas4 family nucleases. *BMC Evol. Biol.*, **17**, 232.

76. Almendros,C., Nobrega,F.L., McKenzie,R.E. and Brouns,S.J.J. (2019) Cas4-Cas1 fusions drive efficient PAM selection and control CRISPR adaptation. *Nucleic Acids Res.*, **47**, 5223–5230.

77. Pyenson,N.C., Gayvert,K., Varble,A., Elemento,O. and Marraffini,L.A. (2017) Broad targeting specificity during bacterial type III CRISPR-Cas immunity constrains viral escape. *Cell Host Microbe*, **22**, 343–353.

78. Silas,S., Lucas-Elio,P., Jackson,S.A., Aroca-Crevillen,A., Hansen,L.L., Fineran,P.C., Fire,A.Z. and Sanchez-Amat,A. (2017) Type III CRISPR-Cas systems can provide redundancy to counteract viral escape from type I systems. *Elife*, **6**, e27601.

79. Davidson,A.R., Lu,W.T., Stanley,S.Y., Wang,J., Mejdani,M., Trost,C.N., Hicks,B.T., Lee,J. and Sontheimer,E.J. (2020) Anti-CRISPRs: Protein inhibitors of CRISPR-Cas systems. *Annu. Rev. Biochem.*, doi:10.1146/annurev-biochem-011420-111224.

80. Borges,A.L., Davidson,A.R. and Bondy-Denomy,J. (2017) The discovery, mechanisms, and evolutionary impact of anti-CRISPRs. *Annu. Rev. Virol.*, **4**, 37–59.

81. Pawluk,A., Davidson,A.R. and Maxwell,K.L. (2018) Anti-CRISPR: discovery, mechanism and function. *Nat. Rev. Microbiol.*, **16**, 12–17.

82. Stanley,S.Y. and Maxwell,K.L. (2018) Phage-encoded anti-CRISPR defenses. *Annu. Rev. Genet.*, **52**, 445–464.

83. Richter,C., Chang,J.T. and Fineran,P.C. (2012) Function and regulation of clustered regularly interspaced short palindromic repeats (CRISPR) / CRISPR associated (Cas) systems. *Viruses*, **4**, 2291–2311.

84. Staals,R.H., Jackson,S.A., Biswas,A., Brouns,S.J., Brown,C.M. and Fineran,P.C. (2016) Interference-driven spacer acquisition is dominant over naive and primed adaptation in a native CRISPR-Cas system. *Nat. Commun.*, **7**, 12853.

85. Nicholson,T.J., Jackson,S.A., Croft,B.I., Staals,R.H.J., Fineran,P.C. and Brown,C.M. (2019) Bioinformatic evidence of widespread priming in type I and II CRISPR-Cas systems. *RNA Biol.*, **16**, 566–576.

86. 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.

87. He,L., St John James,M., Radovcic,M., Ivancic-Bace,I. and Bolt,E.L. (2020) Cas3 protein-a review of a multi-tasking machine. *Genes (Basel)*, **11**, 208.

88. Arinkin,V., Smyshlyaev,G. and Barabas,O. (2019) Jump ahead with a twist: DNA acrobatics drive transposition forward. *Curr. Opin. Struct. Biol.*, **59**, 168–177.

89. Ruiz-Maso,J.A., Macho,N.C., Bordanaba-Ruiseco,L., Espinosa,M., Coll,M. and Del Solar,G. (2015) Plasmid rolling-circle replication. *Microbiol. Spectr.*, **3**, PLAS-0035–2014.

90. Lee,H., Dhingra,Y. and Sashital,D.G. (2019) The Cas4-Cas1-Cas2 complex mediates precise prespacer processing during CRISPR adaptation. *Elife*, **8**, e44248.

91. 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.

92. Jackson,S.A., Birkholz,N., Malone,L.M. and Fineran,P.C. (2019) Imprecise spacer acquisition generates CRISPR-Cas immune diversity through primed adaptation. *Cell Host Microbe*, **25**, 250–260.

93. Krivoy,A., Rutkauskas,M., Kuznedelov,K., Musharova,O., Rouillon,C., Severinov,K. and Seidel,R. (2018) Primed CRISPR adaptation in Escherichia coli cells does not depend on conformational changes in the Cascade effector complex detected in vitro. *Nucleic Acids Res.*, **46**, 4087–4098.