MICROBIOLOGY

RESEARCH ARTICLE

Shangguan and White, *Microbiology* 2023;169:001373

DOI 10.1099/mic.0.001373

MICROBIOLOGY SOCIETY

OPEN ACCESS

Repurposing the atypical type I-G CRISPR system for bacterial genome engineering

Qilin Shangguan and Malcolm F. White*

Abstract

The CRISPR-Cas system functions as a prokaryotic immune system and is highly diverse, with six major types and numerous subtypes. The most abundant are type I CRISPR systems, which utilize a multi-subunit effector, Cascade, and a CRISPR RNA (crRNA) to detect invading DNA species. Detection leads to DNA loading of the Cas3 helicase-nuclease, leading to long-range deletions in the targeted DNA, thus providing immunity against mobile genetic elements (MGE). Here, we focus on the type I-G system, a streamlined, 4-subunit complex with an atypical Cas3 enzyme. We demonstrate that Cas3 helicase activity is not essential for immunity against MGE *in vivo* and explore applications of the *Thioalkalivibrio sulfidiphilus* Cascade effector for genome engineering in *Escherichia coli*. Long-range, bidirectional deletions were observed when the *lacZ* gene was targeted. Deactivation of the Cas3 helicase activity dramatically altered the types of deletions observed, with small deletions flanked by direct repeats that are suggestive of microhomology mediated end joining. When donor DNA templates were present, both the wild-type and helicase-deficient systems promoted homology-directed repair (HDR), with the latter system providing improvements in editing efficiency, suggesting that a single nick in the target site may promote HDR in *E. coli* using the type I-G system. These findings open the way for further application of the type I-G CRISPR systems in genome engineering.

INTRODUCTION

CRISPR systems (classified into types I-VI) function in adaptive immune defence in prokaryotes [1, 2]. The programmable, sequence-specific nature of CRISPR effectors has led to widespread repurposing of CRISPR for genome editing [3]. In particular, the single subunit effectors Cas9 (type II) [4, 5] and Cas12a (type V) [6] have been intensively studied and applied in genome editing over the last decade. Type I CRISPR systems are more complex, with multi-subunit effectors, and have been less widely applied in gene editing applications, despite the fact that they are the most prevalent system in bacteria [7].

The signature enzyme of type I systems is Cas3, a helicase-nuclease fusion that degrades double-strand DNA once recruited to specific sites by the Cascade complex [7, 8]. The broad distribution of type I CRISPR systems in bacteria and archaea has potentiated prokaryotic genome editing applications using the species' own endogenous CRISPR apparatus. Examples include the type I-A system in *Saccharolobus islandicus* [9], type I-B in *Clostridia* [10–12] and *Haloarcula hispanica* [13], type I-C in *Pseudomonas aeruginosa* [14], types I-E [15–19] and I-F [20–23] in a range of species and type I-G in *Bifidobacteria* [24]. Examples of genome engineering in heterologous hosts are less common, but types I-C and I-F from *Pseudomonas* have been used in this manner [14, 25]. There are also successful attempts utilizing type I CRISPR in eukaryotic systems [26–33], showing promising potential for gene deletion and modulation of gene expression.

These genome engineering approaches typically result in long-range genomic deletions, due to the processive nature of Cas3. Cas3 is a dual helicase and nuclease protein that plays a key role in target DNA degradation [34–36]. It is either recruited by the Cascade complex upon binding target dsDNA [35, 37, 38], or pre-associated with Cascade and allosterically activated upon dsDNA recognition [26, 39]. Cas3 is a superfamily 2 helicase that unwinds target DNA in a 3′ to 5′ direction and degrades the resultant ssDNA in the HD nuclease site, leading to a 3′ to 5′ non-target strand degradation [36, 40–42]. Strand switching, or loading of a second Cas3 on the target strand, can lead to bi-directional DNA degradation [38, 43].

Received 24 April 2023; Accepted 18 July 2023; Published 01 August 2023
Author affiliations: [1]School of Biology, Biomedical Sciences Research Complex, University of St Andrews, St Andrews, UK.
***Correspondence:** Malcolm F. White, mfw2@st-andrews.ac.uk
Keywords: CRISPR; genome engineering; Type I-G Cascade; Cas3 helicase.
Abbreviations: CRISPR, clustered, regularly interspaced palindromic repeats; crRNA, CRISPR RNA; HDR, homology directed repair; MGE, mobile genetic element; MMEJ, microhomology mediated end joining; MOI, multiplicity of infection; PCR, polymerase chain reaction; SF2, superfamily 2.
Seven supplementary figures and four supplementary tables are available with the online version of this article.

Previously, we reported the structure and mechanism of the type I-G system from *Thioalkalivibrio sulfidiphilus* [39], demonstrating the constitutive association of Cascade with the Cas3 subunit. Here, we repurpose the *T. sulfidiphilus* system for gene disruption in *E. coli* via long-range, bidirectional DNA degradation, and show that abolition of the helicase activity of *T. sulfidiphilus* Cas3 results in efficient gene disruption by generation of small deletions in target genes.

METHODS

Cloning

For *E. coli* genome editing, the pM2 vector was constructed based on the previously described pACE-M1 [39]. The original T7 promoter was replaced by an araBAD promoter using overlap PCR extension [39] with overlap primers. Restriction sites (*NcoI* and *SalI*) were introduced to facilitate further construction. The *cas3* gene, present in the pEV5hisTEV expression vector as described previously [39], was digested with *Nco*I and *Sal*I (Thermo Scientific) and ligated into the promoter-swapped vector to generate the pM2 vector. Site-directed mutagenesis of the *cas3* gene in pM2 was carried out using standard protocols with Phusion enzyme (Thermo Scientific). Two *Bpi*I restriction sites with type I-G repeat sequence were introduced to pRAT-Duet MCS-I to get the spacer replaceable backbone of the pSPACER plasmid. 5′-phosphorylated oligos of CRISPR spacers were annealed and ligated into the *Bpi*I digested pSPACER backbone to obtain the pSPACER vector with target spacer [39].

For experiments that required a DNA donor, we constructed the pHR vector by introducing a homologous template into the pSPACER vector. Two 615 bp homologous arms (donor) for homologous directed repair (HDR) were PCR amplified from the *E. coli* MG1655 genome. The donor was incorporated into pSPACER MCS-II using restriction enzymes *Nde*I, *Xho*I and *Xho*I, *Avr*II for digestion, followed by ligation. All final constructs were verified by sequencing (GATC Biotech, Eurofins Genomics, DE). Primers and synthetic genes were obtained from Integrated DNA Technologies (Coralville, IA, USA), sequence can be found in Tables S1 and S2, available in the online version of this article. Plasmid information can be found in Table S4.

Plasmid challenge assay

The method was described previously [39]. Briefly, pACE-M1 (Amp^R) encompassing the *cas7*, *cas8g* and *csb2* genes was co-transformed into *E. coli* C43 (DE3) strain with pCDF (Spc^R) vector containing a CRISPR array that targets tetracycline resistance gene (Tet^R). *cas3* or *cas3* mutant gene in pRAT-Duet vector with Tet^R was used to activate type I-G interference. Transformation reactions were then applied to three different selective plates in a 10-fold dilution series to investigate type I-G targeting. LB agar plates were supplemented with 100 μg ml^{-1} ampicillin and 50 μg ml^{-1} spectinomycin when selecting for recipients only; transformants were selected on LB agar containing 100 μg ml^{-1} ampicillin, 50 μg ml^{-1} spectinomycin, 25 μg ml^{-1} tetracycline, and further supplemented with 0.2% (w/v) D-lactose and 0.2% (w/v) L-arabinose for full induction of the type I-G system. Plates were incubated at 37 °C for 16–18 h. The experiment was performed with two biological replicates and at least two technical replicates.

Phage immunity assay

The method was described previously [39]. Briefly, the type I-G system was built in *E. coli* C43(DE3) by introducing pACE-M1, pCDF-*lpa* (CRISPR array targeting the phage P1 *lpa* gene) and pRAT-Cas3 plasmids. Cells were cultured overnight and infected with phage P1 at a range of m.o.i. in 96-well plates. The OD_{600} of the culture in the plate was measured using a FilterMax F5 Multi-Mode Microplate Reader (Molecular Devices) every 20 min over 20 h. The experiment was carried out with two biological replicates and three technical replicates. The OD_{600} was plotted against time using Graphpad Prism 8.

Genome targeting by the type I-G CRISPR system

pM2 was transformed into *E. coli* MG1655. Transformants were selected using 100 μg ml^{-1} ampicillin. Competent cells were prepared by diluting an overnight culture 50-fold into fresh, selective LB medium. The culture was incubated at 37 °C, 220 r.p.m. to reach OD_{600} 0.4 to 0.5. Cells were collected by centrifugation and the pellet resuspended in an equal volume of 100 mM $CaCl_2$, 40 mM $MgSO_4$. Following incubation on ice for 30 min, cells were collected and resuspended in 0.1 volumes of the same buffer containing 10% glycerol. Aliquots were stored at −80 °C. 60 ng pSPACER or pHR was transformed into 60 μl competent cells. In total, 400 μl LB medium was added after heat shock and cells incubated at 37 °C for 80 min. Then, 100 μl aliquots of cells were applied onto 10 cm petri dishes in a 10-fold serial dilution for colony number counting and the number was corrected for dilution and volume to obtain c.f.u. 0.1 ml^{-1}. The LB agar plates contained 100 μg ml^{-1} ampicillin, 12.5 μg ml^{-1} tetracycline, 1 mM IPTG, 0.2 mg/ml X-gal and 0.2% (w/v) L-arabinose for induced plates. A paired *t*-test was used for statistical analysis using Graphpad Prism 8. Further details are available in Table S3.

Tiling PCR

Transformants were submitted for colony PCR with sets of primers (Table S1). hen, 10 μl MyTaq Red Mix (Bioline, Meridian bioscience) was used with 2 μl 20 μM primer mix, colonies were added into the reaction. PCR products were analysed by separation on a 0.8% agarose gel, running in 1×TBE buffer.

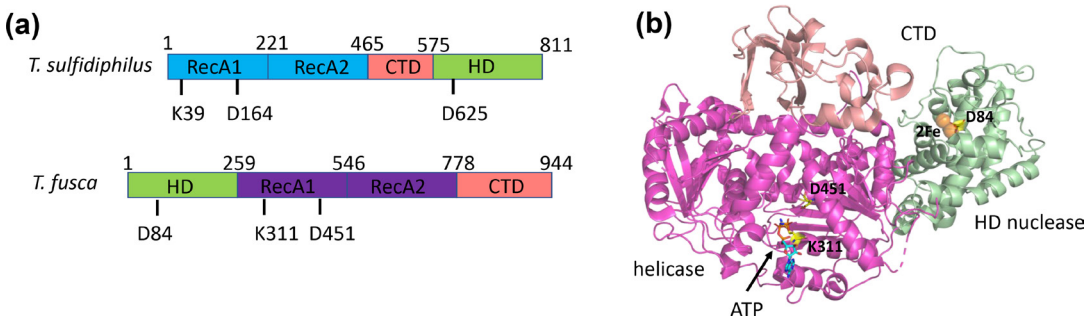

Fig. 1. Comparison of *T. fusca* and *T. sulfidiphilus* Cas3. (a) Domain organization and active site residues of *T. fusca* and *T. sulfidiphilus* Cas3. Each has two RecA-family motor domains containing active site residues that define the Walker A and Walker B boxes of the helicase motor. The HD nuclease domain is present at the N-terminus of *T. fusca* Cas3 but at the C-terminus of *T. sulfidiphilus* Cas3. Each has several conserved acidic residues that coordinate the catalytic metal ions in the active site, of which D625/D84 is labelled. (b) X-ray crystal structure of *T. fusca* Cas3 bound to the ATP analogue AMP-PNP (35), coloured to match the domains shown in (a). The positions of helicase active site residues K311 and D451, and of the HD nuclease active site residue D84, are indicated along with the two active site iron ions (green spheres).

RESULTS

Structure analysis of type I-G Cas3

The type I-G CRISPR system from *Thioalkalivibrio sulfidiphilus* consists of four protein-coding genes and a CRISPR locus. We previously reported a biochemical and structural study of this complex [39], demonstrating that Cas3 is an integral component of this complex rather than a separate subunit that is recruited on target DNA binding. Although the Cas3 structure was not observed in the cryo-EM density, bioinformatic predictions suggest that the domain organization is distinct from Cas3s from other type I CRISPR subtypes, with its HD nuclease domain located at the C-terminus instead of the N-terminus [7] (Fig. 1). By comparison with the crystal structure of Cas3 from *Thermobifida fusca*, we could predict the identity of key active site residues in the HD nuclease and helicase domains of the *T. sulfidiphilus* Cas3 protein (Fig. 1).

The importance of Cas3 helicase and nuclease activities for defence against MGEs

Cas3 has two active sites – a 3′–5′ SF2 helicase and an HD nuclease (Fig. 1). To dissect the importance of these two activities for defence against MGE, we constructed three variant Cas3 proteins: one with a K39A mutation of the key Walker A motif, one with a D164A mutation of the Walker B motif and a third with a D625A mutation in the active site of the HD domain. Variant proteins equivalent to D164A and D625A were previously explored for *T. fusca* Cas3 [35]. We proceeded to investigate the phenotypes of these variants using our established plasmid challenge assay [39] (Fig. 2a, b). In this assay, *E. coli* cells with a functional *T. sulfidiphilus* Cascade were challenged with a plasmid containing a DNA sequence targeted by the crRNA. If the targeted plasmid is destroyed, cells do not become resistant to tetracycline and no colonies are observed. Consistent with previous studies, cells with wild-type Cas3 were fully resistant to plasmid challenge, regardless of whether full gene expression of Cascade was induced by lactose and arabinose (Fig. 2b). In contrast, cells with the K39A and D164A variant of Cas3, which lack helicase activity, were only resistant when protein expression was fully induced. These data suggest that when Cas3 cannot translocate on DNA, but can still cut it, type I-G CRISPR defence is weakened but not abolished.

When submitted to a phage P1 challenge assay, cells expressing the wild-type or helicase-deficient variants K39A and D164A all provided effective immunity at a range of m.o.i.s (Fig. 2c). Thus, as for the plasmid challenge assay, Cas3 helicase activity is not required for immune function. In previous *in vitro* experiments, we observed that *T. sulfidiphilus* Cascade in the absence of ATP (and thus helicase activity) only cleaved target DNA at the site of Cas3 loading [39]. This appears to be enough *in vivo* to prevent target plasmid and phage replication. At higher m.o.i.s, induction of the defence system resulted in higher cell counts compared to the uninduced cultures, suggesting that higher Cascade/Cas3 expression had a beneficial effect. In marked contrast, the D625A variant targeting the HD nuclease domain of Cas3 was indistinguishable from the ΔCas3 strain (Fig. 2b, c). To investigate this further, we expressed and purified the Cas3 D625A variant from *E. coli*. Unlike the wild-type protein, the D625A variant eluted as a large aggregate from a size exclusion column (Fig. S1). These data suggest that mutations disrupting the iron binding site of the Cas3 nuclease domain may disrupt protein folding rather than just inactivating the nuclease domain. This observation emphasizes the importance of checking the phenotypes of variant proteins *in vitro* as well as *in vivo*.

Genome engineering with the type I-G CRISPR system

We proceeded to explore the potential of the type I-G system in genome engineering. We first constructed the vector pM2, containing the type I-G operon under arabinose-inducible promoter control (Fig. 3). pM2 was transformed into *E. coli* MG1655

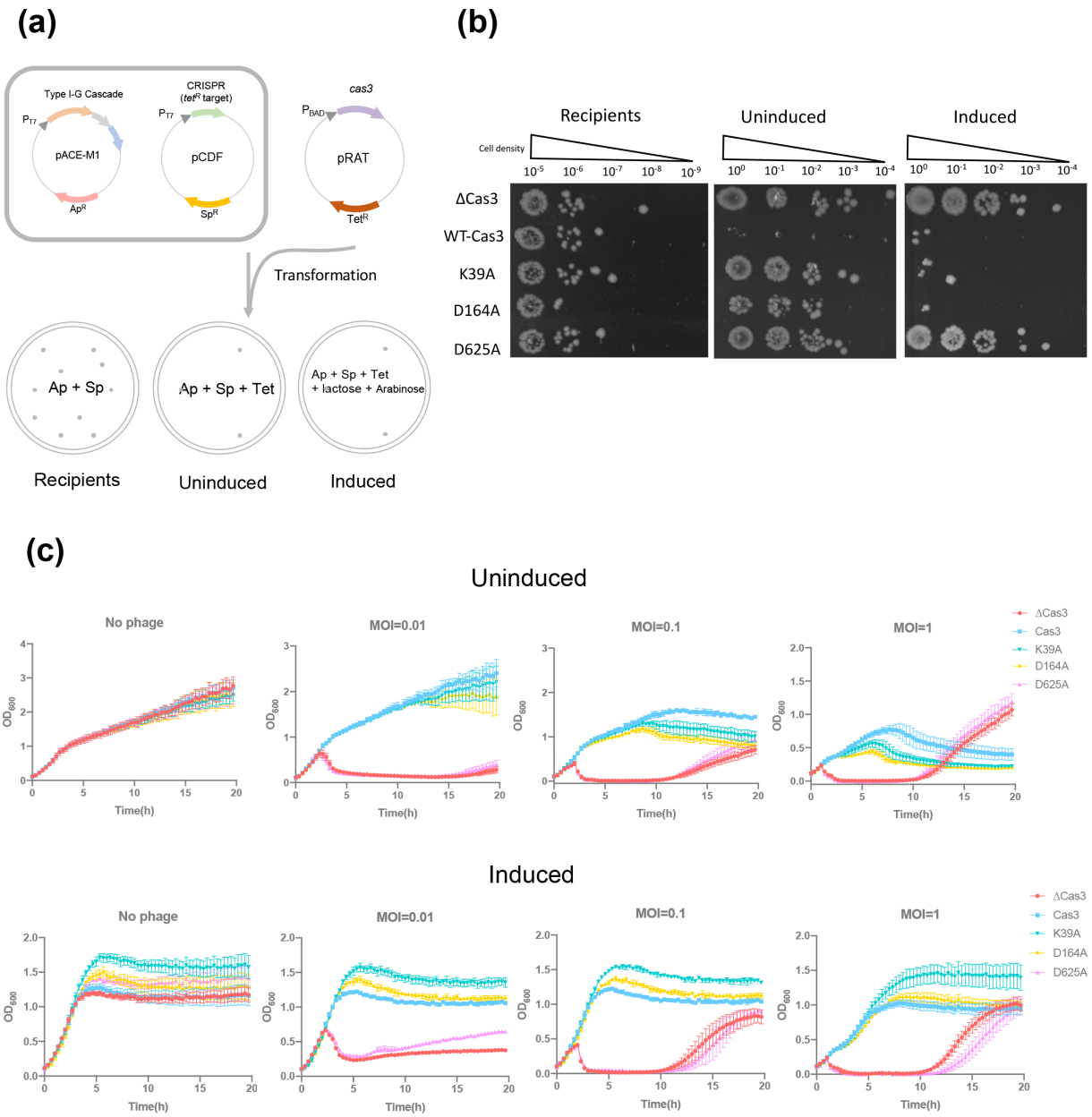

Fig. 2. Functional change in Cas3 mutants. (a, b) Plasmid challenge assay; type I-G Cascade genes (*cas7*, *csb2* and *cas8g*) were built in pACE-M1 vector under T7 promoter and *lac* operon control; CRISPR array targeting tetracycline resistance gene was introduced in pCDF vector under T7 promoter and *lac* operon control; the two vectors were co-transformed into *E. coli* following the pRAT plasmid challenge, cells were then spread on the plates in different conditions. Recipients, ampicillin and spectinomycin in plates; transformants, ampicillin, spectinomycin and tetracycline in plates. Induced, with all three antibiotics and lactose, arabinose for induction. ΔCas3, cells challenged with pRAT-Duet plasmid (Cas3 excluded). (c) Cell growth curve, no phage infection or phage infection (m.o.i.=0.01, 0.1 and 1). ΔCas3, cells lack of Cas3 gene. Data points represent the mean of six experimental replicates (two biological replicates and three technical replicates) with standard deviation shown.

along with plasmid pSPACER to produce pre-crRNA for *lacZ* targeting. Cells were spread on X-gal plates for blue-white screening (Fig. 3c). In the absence of a targeting crRNA, large numbers of blue cells were observed. When a spacer targeting the *lacZ* gene on the bacterial genome was induced, colony counts were significantly reduced and both blue and white colonies were observed (Fig. 3d).

When we first introduced the type I-G system into *E. coli* for self-genome targeting, we noticed a significant decrease (around two orders of magnitude) in cell number compared to empty vector (no target spacer) control (Fig. 4a). Even in the absence of arabinose induction, a loss of cells was still observed, suggesting that the system was actively targeting the *lacZ* gene. On blue/

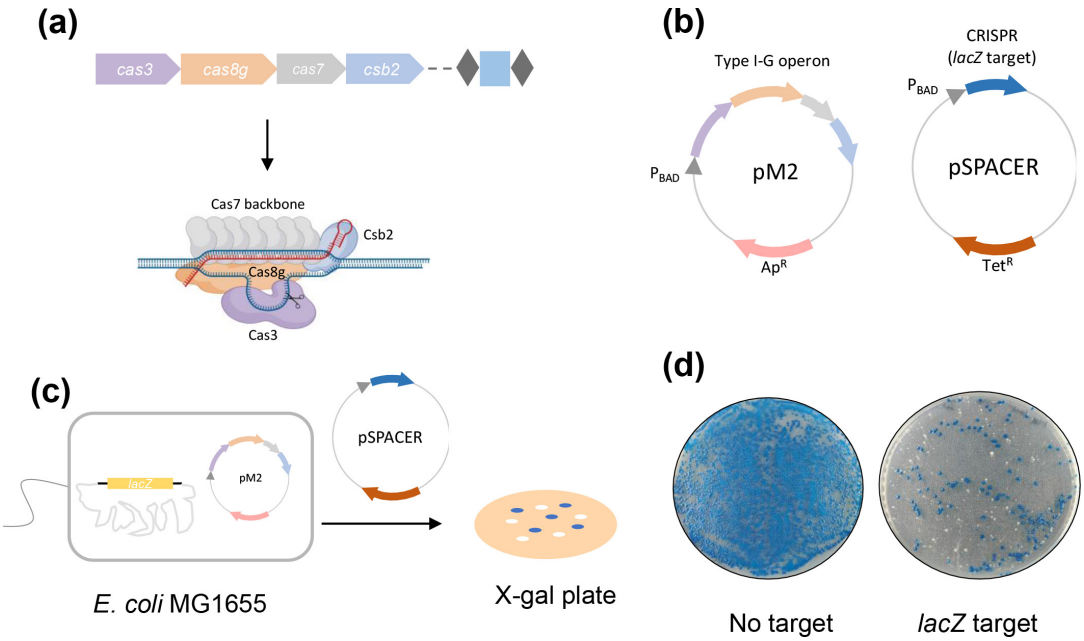

Fig. 3. Target editing by type I-G (a) An overview of type I-G CRISPR operon from *T. sulfidiphilus* HL-EbGr7. *cas* proteins (7 x Cas7, Csb2, Cas8g and Cas3) together with crRNA form the effector complex and target dsDNA. (b) pM2 vector contains type I-G operon (*cas3*, *cas8g*, *cas7* and *csb2*) under arabinose promoter control; pSPACER vector, with CRISPR repeat and spacer (target *lacZ*) under arabinose promoter control. (c) For *lacZ* disruption, pM2 was transformed into *E. coli* MG1655 strain (*lacZ* intact) first, and pSPACER was introduced to activate type I-G *lacZ* targeting, cells were spread on the X-gal plates for blue-white screening. (d) A representative blue-white screening assay with no target control and *lacZ* target spacer.

white screening, white colonies only appeared on the plates with a targeting spacer (Fig. 4b). The reduced cell counts resulting from experiments targeting the *lacZ* gene were most likely due to the generation of unrepaired double-strand breaks (DSB) and/or deletion of an essential flanking gene. Similar toxicity upon self-targeting has been seen in endogenous type I CRISPR and heterologous Cas9 chromosomal targeting in prokaryotic cells [16, 21, 44]. In particular, Cas3 of other type I CRISPR systems is responsible for long-range deletions of chromosomal DNA [14, 30].

Given the large decrease in viable cells and the extensive genomic deletions observed in survivors, we next explored *lacZ* gene disruption using the Cas3 K39A variant. Cells expressing the K39A variant of Cas3 as part of *T. sulfidiphilus* Cascade complex did not experience cell death in the absence of arabinose induction (Fig. 4a), but when fully induced we observed roughly two logs' reduction in viable cells, similar to cells with wild-type Cas3 (Fig. 4a), with survivors displaying a *ΔlacZ* phenotype in similar numbers to the wild-type system (Fig. 4b).

To test whether type I-G Cas3 yielded long-range deletion on genome targeting, white colony survivors were submitted for tiling PCR. In total, 17 out of 18 white survivors from induced plates had a *lacZ* target deletion (Figs 4c, S2 and S3). Out of 18 white colonies, 13 survivors had a deletion at least as far as 55 kb downstream of *lacZ*, 3 survivors experienced a deletion of 45 to 55 kb downstream and one survivor had a shorter 10 to 25 kb downstream deletion. The *lacZ* gene of one white colony (WT-18) remained intact with no detected deletion around the target site, suggestive of a spontaneous *lacZ* mutation. The essential *hemB* gene is located 20 kb upstream of *lacZ*. Unsurprisingly, all the survivors had an intact *hemB* gene locus (based on the presence of the tiling product 20 kb upstream of the *lacZ* gene), but 16 colonies showed 10–20 kb deletions of upstream DNA. In subsequent experiments, we targeted the *lacZ* adjacent genes *yahK* and *frmA*. When the downstream gene *yahK* was targeted, the yield of white survivors was comparable to *lacZ* target, while the upstream gene *frmA* target sharply lowered the number of white surviving colonies obtained, likely due to its proximity to the essential gene *hemB* (Fig. S4). These data suggest that the type I-G system yields long-range bidirectional degradation of the *E. coli* genome.

In contrast, PCR tiling revealed a marked difference in editing outcomes when using the *cas3* K39A mutant. Out of 18 white colony survivors analysed by tiling PCR, 14 gave a PCR product when using an internal *lacZ* primer, four had a localized *lacZ* deletion, and only one of them showed long-range deletion (Figs 4c, d, S3 and S5). These white survivors were further analysed using PCR primers that covered the target area. Intriguingly, PCR products from these survivors varied in size, indicative of small deletions (Fig. S5C). Out of 13 sequenced PCR products, one contained a point mutation, seven had a precise 180 bp deletion flanking the target site and the other five had no obvious edit (Fig. 4e). Close investigation of the sequences revealed the presence of 11 bp direct repeats flanking the deleted region (Fig. 4e). This suggests that the DNA break introduced by Cas3 was repaired by

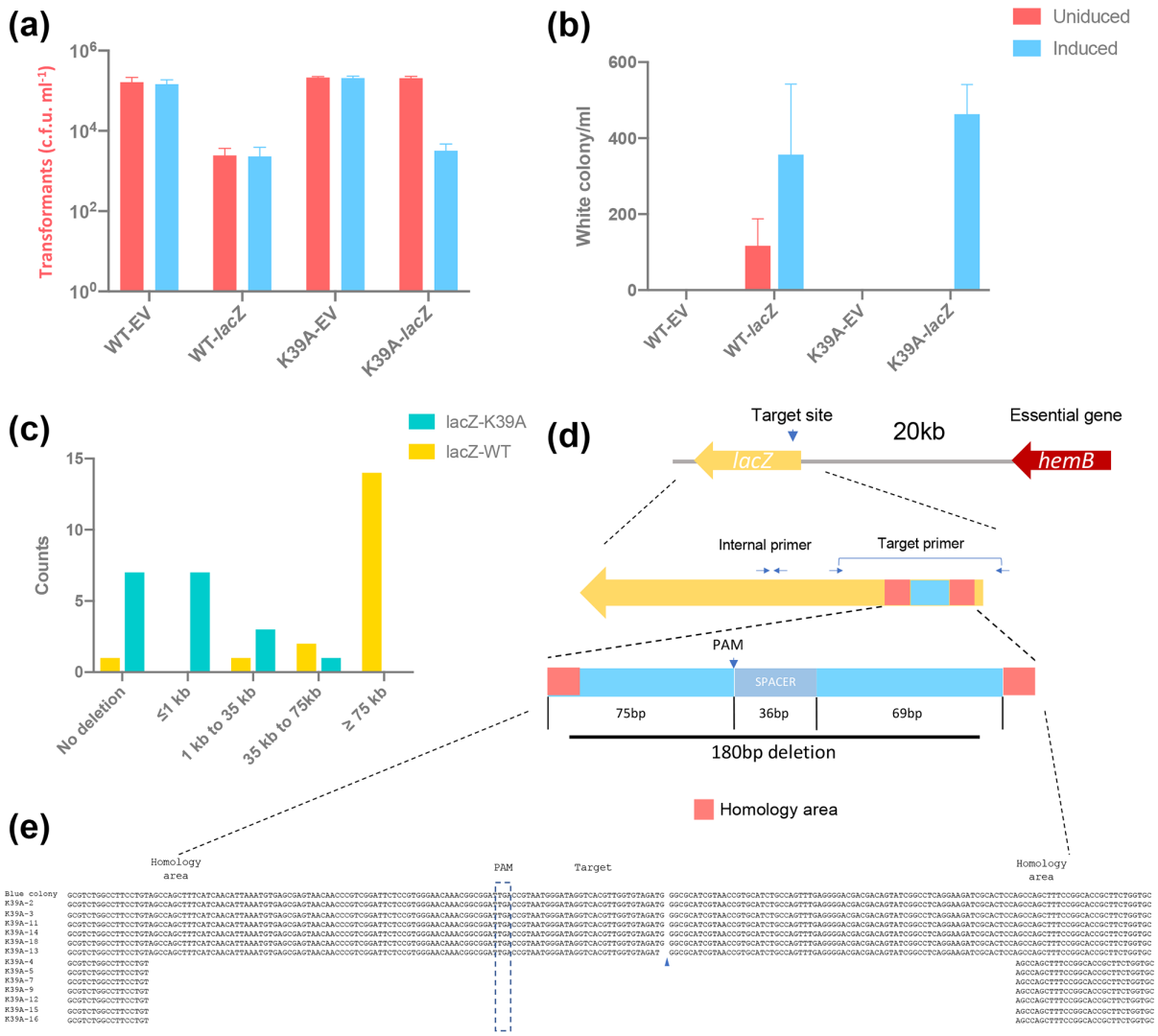

Fig. 4. Type I-G editing on *lacZ*. (a, b) Transformants' number and white colony number on the plate after transformation of the empty vector control or *lacZ* target spacer to wild-type Cas3 strain or K39A Cas3 strain with L-arabinose induction (blue) or without induction (red); Transformation efficiency was calculated as the number of transformants divided by the number of transformants for original plasmid without target (empty vector control). Values and error bars represent the mean of three biological replicates and standard deviation. (c) Editing outcome of *lacZ* targeting by wildtype type I-G (yellow) or Cas3 K39A (green); 18 white colonies generated by *lacZ* targeting were submitted for tiling PCR to determine the deletion range; counts of white colonies are plotted against deletion range. (d) A schematic of 180 bp deletion by Cas3 K39A targeting *lacZ*; blue, deleted sequence; red, homology area; PAM, protospacer adjacent motif; SPACER, spacer sequence for *lacZ* targeting; primers used for amplification of target area were indicated by arrows. (e) Sequence analysis of the 180 bp deletion by Cascade with Cas3 K39A; blue arrow, a point mutation was detected; direct repeats in the flanking sequence are shown in red.

limited strand resection and repair by microhomology-mediated end joining (MMEJ). Gene deletions flanked by short regions of microhomology were recently observed in a genome engineering study in *Bifidobacteria* [24].

To investigate this phenomenon in more detail, we targeted a second site in the *lacZ* gene using a different crRNA with the K39A Cas3 variant Cascade and analysed it as before. DNA sequencing showed various editing outcomes: intact target site (8 out of 17 tested colonies), point mutation (2/17), 24 bp short deletion (5/17), 324 bp long deletion (1/17) and a 11 bp insertion (1/17) (Fig. S6). Microhomologies ranging from 2 to 8 bp were apparent flanking the deleted regions (Fig. S6). Overall, genome targeting via Cas3 variant K39A generated distinct editing outcomes compared with the wild-type editing.

Utilising the type I-G system in homology-directed repair

In the absence of any donor DNA to direct repair, the main products of gene disruption likely arise from error-prone end joining pathways. We proceeded to investigate whether the provision of donor DNA as a homology-directed repair (HDR) template would

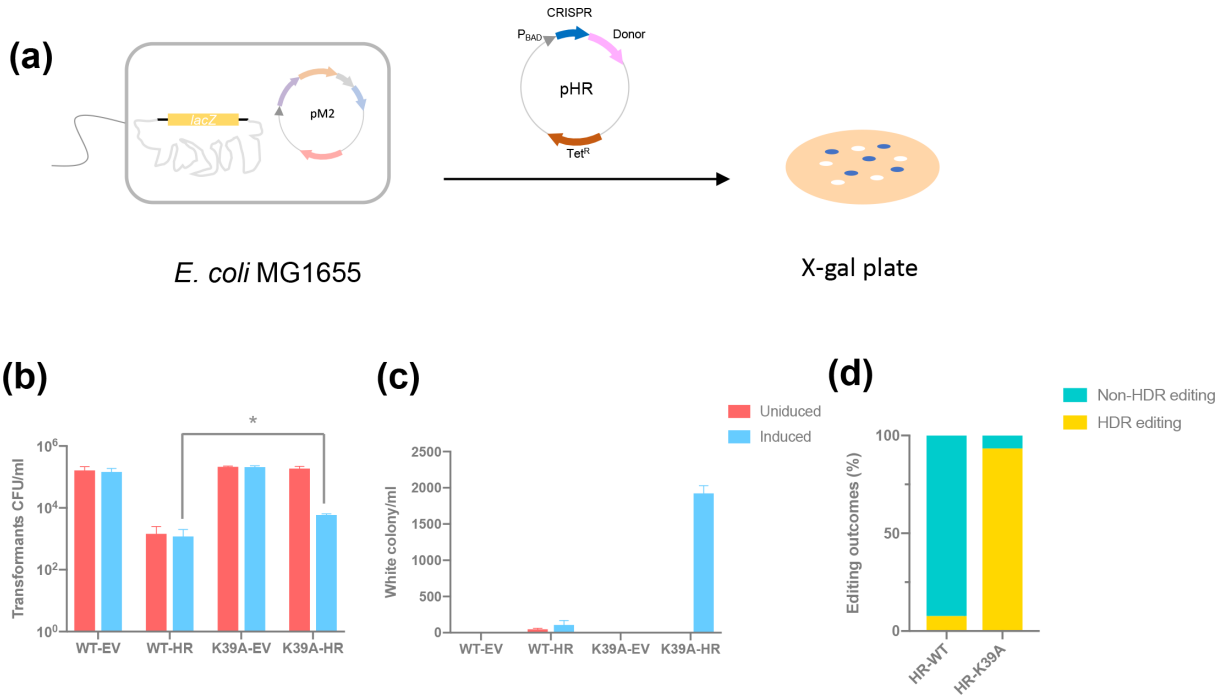

Fig. 5. Homology-directed repair. (a) A schematic of homology directed repair; pM2 was transformed into *E. coli* MG1655 strain (*lacZ* intact) first, and pHR with donor sequence and *lacZ* target spacer was introduced to activate type I-G *lacZ* targeting and desired HDR, cells were spread on the X-gal plates for blue-white screening. (b, c) Transformants' number and white colony number on the plate after transformation of pHR into WT *cas3* strain or K39A *cas3* mutant strain with L-arabinose induction (blue) or without induction (red); EV, empty vector control; transformation efficiency was calculated as the number of transformants divided by the number of transformants for original plasmid without target (empty vector control). Values and error bars represent the mean of three biological replicates and standard deviation; *P<0.05, paired t-test. (d) Editing outcomes with donor templates; percentage of white colonies with desired HDR editing (yellow), without HDR editing (green); 30 individual white colonies from HR-K39A, 13 from HR-WT were assayed.

result in different outcomes. To investigate this, the pHR vector was constructed with a donor sequence incorporated downstream of the CRISPR array (Figs 5A and S7A). If successful, the design of the HDR donor should result in a 13 kb deletion in the genome. For the type I-G system with wild-type Cas3, even with homologous recombination templates, the survivability was still low and target HDR efficiency was under 20% (one target deletion out of eight assayed white colonies) (Fig. 5). However, when targeting with Cas3 K39A, an increase in cell survivability and white colony number was observed. The target editing efficiency was over 90% (28 target deletions out of 30 white colonies) (Fig. 5d) and a 13 kb deletion was observed (Fig. S7). Overall, the introduction of donor DNA for HDR increased the cell survivability and yielded the desired long-range deletion in the bacterial genome when the K39A variant was used.

DISCUSSION

Although Cas3 translocates unidirectionally on target DNA *in vitro* [34, 37], type I CRISPR systems are capable of generating both unidirectional [27, 28, 30–33] and bi-directional [14, 26, 29] deletions in target genomes. Here, we have shown that type I-G Cascade-Cas3 specifically creates bi-directional long-range genome deletions in *E. coli*. Similar bidirectional degradation has been observed for the type I-C CRISPR in *Pseudomonas aeruginosa* [14], where an average deletion of around 90 kb was observed. However, when the type I-C system from *Neisseria lactamica* was repurposed for application in eukaryotes, it resulted in unidirectional degradation [28]. These observations suggest that editing outcomes are dependent on both the specific type I CRISPR system and the experimental species under study.

Cas3 is universally conserved across all type I CRISPR systems, despite the differences in Cascade composition [7]. In the type I-G system Cas3 is an integral subunit of Cascade rather than being recruited on target DNA binding [24, 39]. This association between Cas3 and Cascade has also been observed in a type I-A system where the authors proposed an allosteric activation mode, in contrast to the common trans-recruitment mode for Cas3 [26, 43]. The type I-G Cas3 shares common features with canonical Cas3 (a SF2-helicase domain and a HD-nuclease domain), but at the protein sequence level the HD nuclease domain is found at the C-terminus instead of N-terminus. In type I-A systems, Cas3 is encoded on two individual genes that separately express

the helicase and HD nuclease domains [7]. These differences in Cas3 domain organization seem to correlate with the existence of two different Cascade-Cas3 activation modes.

While the type I-G system with wild-type Cas3 generates large bi-directional deletions, we have shown that the helicase activity of Cas3 is not essential for functional immunity against MGEs in *E. coli*, as long as the proteins are well expressed. This suggests that localized nicking of the target strand in targeted MGEs can be sufficient to prevent replication, possibly because replication forks collapse when encountering a nick in the leading strand template, resulting in a double-strand break. Likewise, the use of the helicase-deficient K39A variant of Cas3 did not abolish gene disruption in *E. coli*. However, the edited products were fundamentally different, with small deletions predominating. The recurrence of a specific 180 bp deletion was initially surprising, but may be explained by the presence of direct repeats flanking the deleted sequence, which could allow end joining following limited strand resection. Microhomology mediated end joining is observed when HDR is not an option in *E. coli* [45], and has been seen previously for Cas3-mediated deletions in *P. aeruginosa* [14]. Five non-edited white colonies were observed when targeting with Cas3 K39A. One possible explanation is that *lacZ* transcription was blocked by the type I-G targeting without target site cleavage, resulting in white colonies with an intact *lacZ* gene.

Provision of a donor template to enhance HDR during edits resulted in the expected HDR outcomes. In particular, the use of the Cas3 K39A variant in combination with a donor template enhanced editing efficiency. Recently, a Cas3 helicase variant of the *Zymomonas mobilis* type I-F system was applied endogenously to carry out genome editing with high efficiency [46]. In that study, crRNA was used to target each strand of the target, resulting in 'dual nicking' of the target gene. Our data suggest that targeting genes with a single guide RNA can still lead to efficient gene disruption and gene replacement, simplifying the procedure.

For obvious reasons, the compact Cas9 and Cas12 enzymes have been favoured for genome editing in a wide range of species. Nonetheless, type I systems have now been widely used to create larger edits in a range of cognate and heterologous cell types. Here, we have demonstrated that the relatively simple, 4-gene type I-G system can be used to generated gene deletions and replacements in *E. coli*. Surprisingly, abolishing the helicase activity of Cas3 does not prevent either effective immunity against MGE, and can result in enhanced genome editing, suggesting that the cellular responses to the DNA nicks produced can result in gene disruption with smaller deletions, while still supporting HDR when a donor template is provided.

Funding information
This work was supported by the Biotechnology and Biological Sciences Research Council (REF: BB/S000313/1 to MFW), and the China Scholarship Council (REF: 202008060345 to QS).

Conflicts of interest
The authors declare they are aware of no conflicts of interest.

References

1. Barrangou R, Fremaux C, Deveau H, Richards M, Boyaval P, *et al*. CRISPR provides acquired resistance against viruses in prokaryotes. *Science* 2007;315:1709–1712.
2. Makarova KS, Wolf YI, Iranzo J, Shmakov SA, Alkhnbashi OS, *et al*. Evolutionary classification of CRISPR-Cas systems: a burst of class 2 and derived variants. *Nat Rev Microbiol* 2020;18:67–83.
3. Barrangou R, Doudna JA. Applications of CRISPR technologies in research and beyond. *Nat Biotechnol* 2016;34:933–941.
4. Cong L, Ran FA, Cox D, Lin S, Barretto R, *et al*. Multiplex genome engineering using CRISPR/Cas systems. *Science* 2013;339:819–823.
5. Mali P, Yang L, Esvelt KM, Aach J, Guell M, *et al*. RNA-guided human genome engineering via Cas9. *Science* 2013;339:823–826.
6. Zetsche B, Gootenberg JS, Abudayyeh OO, Slaymaker IM, Makarova KS, *et al*. Cpf1 is a single RNA-guided endonuclease of a class 2 CRISPR-Cas system. *Cell* 2015;163:759–771.
7. Makarova KS, Wolf YI, Alkhnbashi OS, Costa F, Shah SA, *et al*. An updated evolutionary classification of CRISPR-Cas systems. *Nat Rev Microbiol* 2015;13:722–736.
8. Barrangou R. CRISPR-Cas systems and RNA-guided interference. *Wiley Interdiscip Rev RNA* 2013;4:267–278.
9. Li Y, Pan S, Zhang Y, Ren M, Feng M, *et al*. Harnessing type I and type III CRISPR-Cas systems for genome editing. *Nucleic Acids Res* 2016;44:e34.
10. Pyne ME, Bruder MR, Moo-Young M, Chung DA, Chou CP. Harnessing heterologous and endogenous CRISPR-Cas machineries for efficient markerless genome editing in *Clostridium*. *Sci Rep* 2016;6:25666.
11. Zhang J, Zong W, Hong W, Zhang Z-T, Wang Y. Exploiting endogenous CRISPR-Cas system for multiplex genome editing in *Clostridium tyrobutyricum* and engineer the strain for high-level butanol production. *Metab Eng* 2018;47:49–59.
12. Maikova A, Kreis V, Boutserin A, Severinov K, Soutourina O. Using an endogenous CRISPR-Cas system for genome editing in the human pathogen *Clostridium difficile*. *Appl Environ Microbiol* 2019;85:e01416-19.
13. Cheng F, Gong L, Zhao D, Yang H, Zhou J, *et al*. Harnessing the native type I-B CRISPR-Cas for genome editing in a polyploid archaeon. *J Genet Genomics* 2017;44:541–548.
14. Csörgő B, León LM, Chau-Ly IJ, Vasquez-Rifo A, Berry JD, *et al*. A compact Cascade-Cas3 system for targeted genome engineering. *Nat Methods* 2020;17:1183–1190.
15. Kiro R, Shitrit D, Qimron U. Efficient engineering of a bacteriophage genome using the type I-E CRISPR-Cas system. *RNA Biol* 2014;11:42–44.
16. Gomaa AA, Klumpe HE, Luo ML, Selle K, Barrangou R, *et al*. Programmable removal of bacterial strains by use of genome-targeting CRISPR-Cas systems. *mBio* 2014;5:e00928-13.
17. Hidalgo-Cantabrana C, Goh YJ, Pan M, Sanozky-Dawes R, Barrangou R. Genome editing using the endogenous type I CRISPR-Cas system in *Lactobacillus crispatus* . *Proc Natl Acad Sci* 2019;116:15774–15783.
18. Cañez C, Selle K, Goh YJ, Barrangou R. Outcomes and characterization of chromosomal self-targeting by native CRISPR-Cas systems in *Streptococcus thermophilus*. *FEMS Microbiol Lett* 2019;366:fnz105.

19. Zhou Y, Yang Y, Li X, Tian D, Ai W, *et al*. Exploiting a conjugative endogenous CRISPR-Cas3 system to tackle multidrug-resistant *Klebsiella pneumoniae*. *EBioMedicine* 2023;88:104445.
20. Zheng Y, Han J, Wang B, Hu X, Li R, *et al*. Characterization and repurposing of the endogenous Type I-F CRISPR–Cas system of *Zymomonas mobilis* for genome engineering. *Nucleic Acids Res* 2019;47:11461–11475.
21. Vercoe RB, Chang JT, Dy RL, Taylor C, Gristwood T, *et al*. Cytotoxic chromosomal targeting by CRISPR/Cas systems can reshape bacterial genomes and expel or remodel pathogenicity islands. *PLoS Genet* 2013;9:e1003454.
22. Hampton HG, McNeil MB, Paterson TJ, Ney B, Williamson NR, *et al*. CRISPR-Cas gene-editing reveals RsmA and RsmC act through FlhDC to repress the SdhE flavinylation factor and control motility and prodigiosin production in *Serratia*. *Microbiology* 2016;162:1047–1058.
23. Xu Z, Li M, Li Y, Cao H, Miao L, *et al*. Native CRISPR-Cas-mediated genome editing enables dissecting and sensitizing clinical multidrug-resistant *P. aeruginosa*. *Cell Rep* 2019;29:1707–1717.
24. Pan M, Morovic W, Hidalgo-Cantabrana C, Roberts A, Walden KKO, *et al*. Genomic and epigenetic landscapes drive CRISPR-based genome editing in *Bifidobacterium*. *Proc Natl Acad Sci* 2022;119:30.
25. Xu Z, Li Y, Cao H, Si M, Zhang G, *et al*. A transferrable and integrative type I-F Cascade for heterologous genome editing and transcription modulation. *Nucleic Acids Res* 2021;49:16.
26. Hu C, Ni D, Nam KH, Majumdar S, McLean J, *et al*. Allosteric control of type I-A CRISPR-Cas3 complexes and establishment as effective nucleic acid detection and human genome editing tools. *Mol Cell* 2022;82:2754–2768.
27. Pickar-Oliver A, Black JB, Lewis MM, Mutchnick KJ, Klann TS, *et al*. Targeted transcriptional modulation with type I CRISPR-Cas systems in human cells. *Nat Biotechnol* 2019;37:1493–1501.
28. Tan R, Krueger RK, Gramelspacher MJ, Zhou X, Xiao Y, *et al*. Cas11 enables genome engineering in human cells with compact CRISPR-Cas3 systems. *Mol Cell* 2022;82:852–867.
29. Osakabe K, Wada N, Murakami E, Miyashita N, Osakabe Y. Genome editing in mammalian cells using the CRISPR type I-D nuclease. *Nucleic Acids Res* 2021;49:6347–6363.
30. Dolan AE, Hou Z, Xiao Y, Gramelspacher MJ, Heo J, *et al*. Introducing a spectrum of long-range genomic deletions in human embryonic stem cells using type I CRISPR-cas. *Mol Cell* 2019;74:936–950.
31. Morisaka H, Yoshimi K, Okuzaki Y, Gee P, Kunihiro Y, *et al*. CRISPR-Cas3 induces broad and unidirectional genome editing in human cells. *Nat Commun* 2019;10:5302.
32. Cameron P, Coons MM, Klompe SE, Lied AM, Smith SC, *et al*. Harnessing type I CRISPR-Cas systems for genome engineering in human cells. *Nat Biotechnol* 2019;37:1471–1477.
33. Chen Y, Liu J, Zhi S, Zheng Q, Ma W, *et al*. Repurposing type I–F CRISPR–Cas system as a transcriptional activation tool in human cells. *Nat Commun* 2020;11.
34. Mulepati S, Bailey S. *In vitro* reconstitution of an *Escherichia coli* RNA-guided immune system reveals unidirectional, ATP-dependent degradation of DNA target. *J Biol Chem* 2013;288:22184–22192.
35. Huo Y, Nam KH, Ding F, Lee H, Wu L, *et al*. Structures of CRISPR Cas3 offer mechanistic insights into Cascade-activated DNA unwinding and degradation. *Nat Struct Mol Biol* 2014;21:771–777.
36. Sinkunas T, Gasiunas G, Fremaux C, Barrangou R, Horvath P, *et al*. Cas3 is a single-stranded DNA nuclease and ATP-dependent helicase in the CRISPR/Cas immune system. *EMBO J* 2011;30:1335–1342.
37. Redding S, Sternberg SH, Marshall M, Gibb B, Bhat P, *et al*. Surveillance and processing of foreign DNA by the *Escherichia coli* CRISPR-Cas system. *Cell* 2015;163:854–865.
38. Dillard KE, Brown MW, Johnson NV, Xiao Y, Dolan A, *et al*. Assembly and translocation of a CRISPR-cas primed acquisition complex. *Cell* 2018;175:934–946.
39. Shangguan Q, Graham S, Sundaramoorthy R, White MF. Structure and mechanism of the type I-G CRISPR effector. *Nucleic Acids Res* 2022;50:11214–11228.
40. Westra ER, van Erp PBG, Künne T, Wong SP, Staals RHJ, *et al*. CRISPR immunity relies on the consecutive binding and degradation of negatively supercoiled invader DNA by Cascade and Cas3. *Mol Cell* 2012;46:595–605.
41. Gong B, Shin M, Sun J, Jung C-H, Bolt EL, *et al*. Molecular insights into DNA interference by CRISPR-associated nuclease-helicase Cas3. *Proc Natl Acad Sci* 2014;111:16359–16364.
42. Hochstrasser ML, Taylor DW, Bhat P, Guegler CK, Sternberg SH, *et al*. CasA mediates Cas3-catalyzed target degradation during CRISPR RNA-guided interference. *Proc Natl Acad Sci* 2014;111:6618–6623.
43. Xiao Y, Luo M, Dolan AE, Liao M, Ke A. Structure basis for RNA-guided DNA degradation by Cascade and Cas3. *Science* 2018;361:eaat0839.
44. Cui L, Bikard D. Consequences of Cas9 cleavage in the chromosome of *Escherichia coli*. *Nucleic Acids Res* 2016;44:4243–4251.
45. Chayot R, Montagne B, Mazel D, Ricchetti M. An end-joining repair mechanism in *Escherichia coli*. *Proc Natl Acad Sci* 2010;107:2141–2146.
46. Hao Y, Wang Q, Li J, Yang S, Zheng Y, *et al*. Double nicking by RNA-directed Cascade-nCas3 for high-efficiency large-scale genome engineering. *Open Biol* 2022;12.

Edited by: M. Brockhurst and S. van Houte