

Efficient Single-Nucleotide Microbial Genome Editing Achieved Using CRISPR/Cpf1 with Maximally 3'-End-Truncated crRNAs

Ho Jung Lee, Hyun Ju Kim, Young-Jun Park, and Sang Jun Lee*

Cite This: *ACS Synth. Biol.* 2022, 11, 2134–2143

Read Online

ACCESS |



Metrics & More



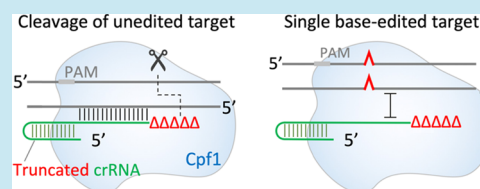
Article Recommendations



Supporting Information

ABSTRACT: Mismatch tolerance, a cause of the off-target effect, impedes accurate genome editing with the CRISPR/Cas system. Herein, we observed that oligonucleotide-directed single-base substitutions could be rarely introduced in the microbial genome using CRISPR/Cpf1-mediated negative selection. Because crRNAs have the ability to recognize and discriminate among specific target DNA sequences, we systematically compared the effects of modified crRNAs with 3'-end nucleotide truncations and a single mismatch on the genomic cleavage activity of FnCpf1 in *Escherichia coli*. Five nucleotides could be maximally truncated at the crRNA 3'-end for the efficient cleavage of the DNA targets of *galK* and *xylB* in the cells. However, target cleavage in the genome was inefficient when a single mismatch was simultaneously introduced in the maximally 3'-end-truncated crRNA. Based on these results, we assumed that the maximally truncated crRNA-Cpf1 complex can distinguish between single-base-edited and unedited targets in vivo. Compared to other crRNAs with shorter truncations, maximally 3'-end-truncated crRNAs showed highly efficient single-base substitutions (>80%) in the DNA targets of *galK* and *xylB*. Furthermore, the editing efficiency for the 24 bases in both *galK* and *xylB* showed success rates of 79 and 50%, respectively. We successfully introduced single-nucleotide indels in *galK* and *xylB* with editing efficiencies of 79 and 62%, respectively. Collectively, the maximally truncated crRNA-Cpf1 complex could perform efficient base and nucleotide editing regardless of the target base location or mutation type; this system is a simple and efficient tool for microbial genome editing, including indel correction, at the single-nucleotide resolution.

KEYWORDS: *FnCpf1*, crRNA, mismatch tolerance, single-base editing



INTRODUCTION

The potential of the CRISPR/Cas system as a genome editing tool was first reported a decade ago.¹ Various types of Cas nucleases have been discovered and applied in genome editing.² Among them, the CRISPR/Cpf1 system is a genome editing system functional in bacteria,³ plants,⁴ and mammals,⁵ including humans.⁶ Because Cpf1 exhibits nonspecific ssDNA trans-cleavage activity,⁷ it is used in a wide range of fields such as diagnostics.^{8,9} For a genome with a low average GC content, such as the human genome, Cpf1 with the 5'-TTN protospacer adjacent motif (PAM) helps edit more diverse targets than Cas9 with 5'-NGG PAM.¹⁰ Cas9 has been used to make point mutations in many bacterial cells,^{11,12} but in some microorganisms, increasing the intracellular concentration of Cas9 protein inhibits cell growth;^{13,14} therefore, the use of Cpf1 is absolutely necessary in certain cases of microbial genome editing. However, because Cpf1 and Cas9 exhibit mismatch tolerance⁶ and induce the off-target effect,¹⁵ accurate genome editing using Cpf1 has certain limitations.

To avoid the double-stranded cleavage of nontarget genomic regions, a base editor that induces chemical modifications of the base without introducing double-stranded cleavage has been developed.^{16,17} With the development of the glycosylase base editor, the introduction of transversion mutations has become feasible,¹⁸ and the precise editing of only the target base using a

sequence-specific deaminase has also become viable.¹⁹ In addition, all types of mutations can be introduced using a prime editor that synthesizes a new DNA strand to a desired target site using an engineered reverse transcriptase.²⁰ For highly efficient mutagenesis, including indel introduction, the use of gene scissors such as Cas9 and Cpf1 that can cut double strands is still needed.

Several studies have been performed to modify crRNA, a molecule that determines the target specificity of the CRISPR/Cas system to improve on-target specificity and editing efficiency. The gene knock-out efficiency was increased by the conjugation of a tRNA-like structure to the 5'-end of the crRNA.²¹ 5'-end extension in crRNA increased the editing efficiency, such as in gene knock-out and homology-directed repair,²² and 3'-end uridylation increased the indel efficiency.²³ Chemical base modification in crRNA, such as phosphorothioate backbone substitution in the target recognition sequence or methylation at the 3'-end carbon 2, improved

Received: January 29, 2022

Published: May 18, 2022



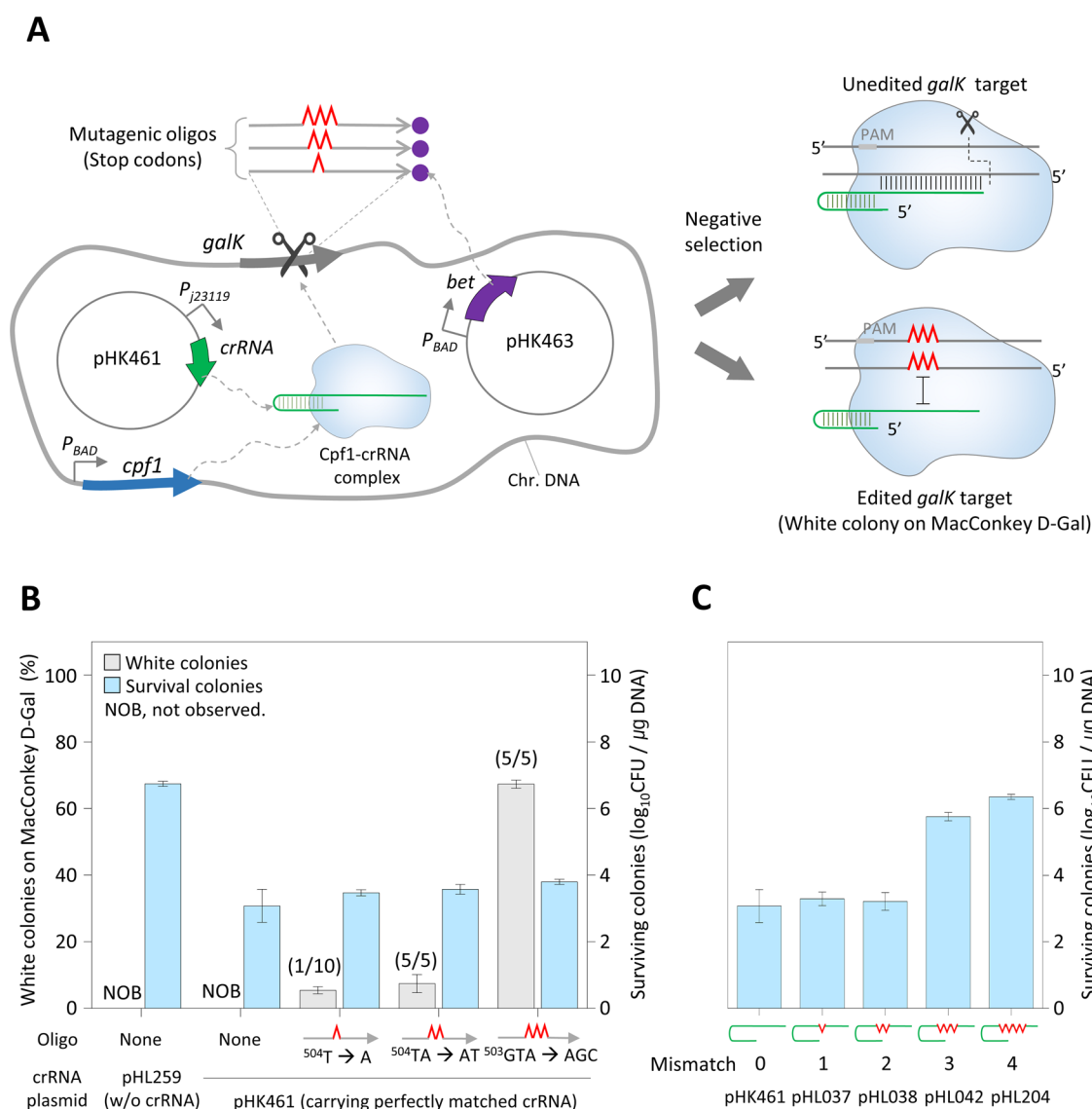


Figure 1. Comparison of single-base and multi-base genome editing using the CRISPR/Cpf1 system. (A) Negative selection of cells with genome editing using the crRNA-Cpf1 complex. The crRNA-Cpf1 complex can cleave the unedited DNA target, thereby leaving the edited DNA target uncut. The generation of stop codons in the *galk* gene in the cells with genome editing led to the formation of white colonies on D-galactose-containing MacConkey agar because of the premature termination of GalK protein translation. (B) Editing efficiencies in CRISPR/Cpf1-mediated one to three base(s) substitutions in *galk*. The genome editing efficiency was assessed as the ratio of the number of white colonies to the total number of colonies on MacConkey D-galactose agar. The parentheses indicate the number of colonies containing cells with correct editing among white colonies from which the cells were subjected to Sanger sequencing. (C) The target mismatch effect of CRISPR/Cpf1. crRNA plasmids carrying perfectly matched crRNAs or crRNAs with one to four bases mismatched were transformed into Cpf1-overexpressing cells. The number of transformants on LB agar containing spectinomycin indicates the genome-target-cleaving activity of the crRNA-Cpf1 complex.

the gene disruption activity.²⁴ When the 5'- and 3'-ends of crRNA were substituted with DNA, the target cleavage efficiency was improved and the off-target effect was reduced.²⁵ To overcome the mismatch tolerance of Cas9, we previously designed a sgRNA with a mismatch with the target in advance, which considerably enhanced the microbial single-base editing efficiency.²⁶ Target-mismatched crRNAs can also be used in the CRISPR/Cpf1 system, which aids accurate bacterial genome editing.³

Moreover, reportedly, the truncation of guide RNA reduces the off-target effect of Cas9.²⁷ It was recently reported that 2-nucleotide (nt) truncations at the 5'-end of a single molecular guide RNA (sgRNA) enhanced Cas9-mediated single-base substitution in the microbial genome.²⁸ As in the case of Cas9,

the 3'-end-truncated crRNA-Cpf1 complex can cause double-strand breaks in the target DNA even when 4–6 nt are truncated at the 3'-end of the canonical 23-nt spacer of the CRISPR/Cpf1 crRNA.^{29,30} In this study, we systematically investigated the relationship between single-base mismatch tolerance and 3'-end nucleotide truncation in the CRISPR/Cpf1 crRNA. We report a simple, accurate, and efficient method for single-nucleotide editing, including indel, using maximally 3'-end-truncated crRNAs in the CRISPR/FnCpf1 system, and also discuss the shorter target recognition sequences of crRNAs for better target specificity of the CRISPR/FnCpf1 system.

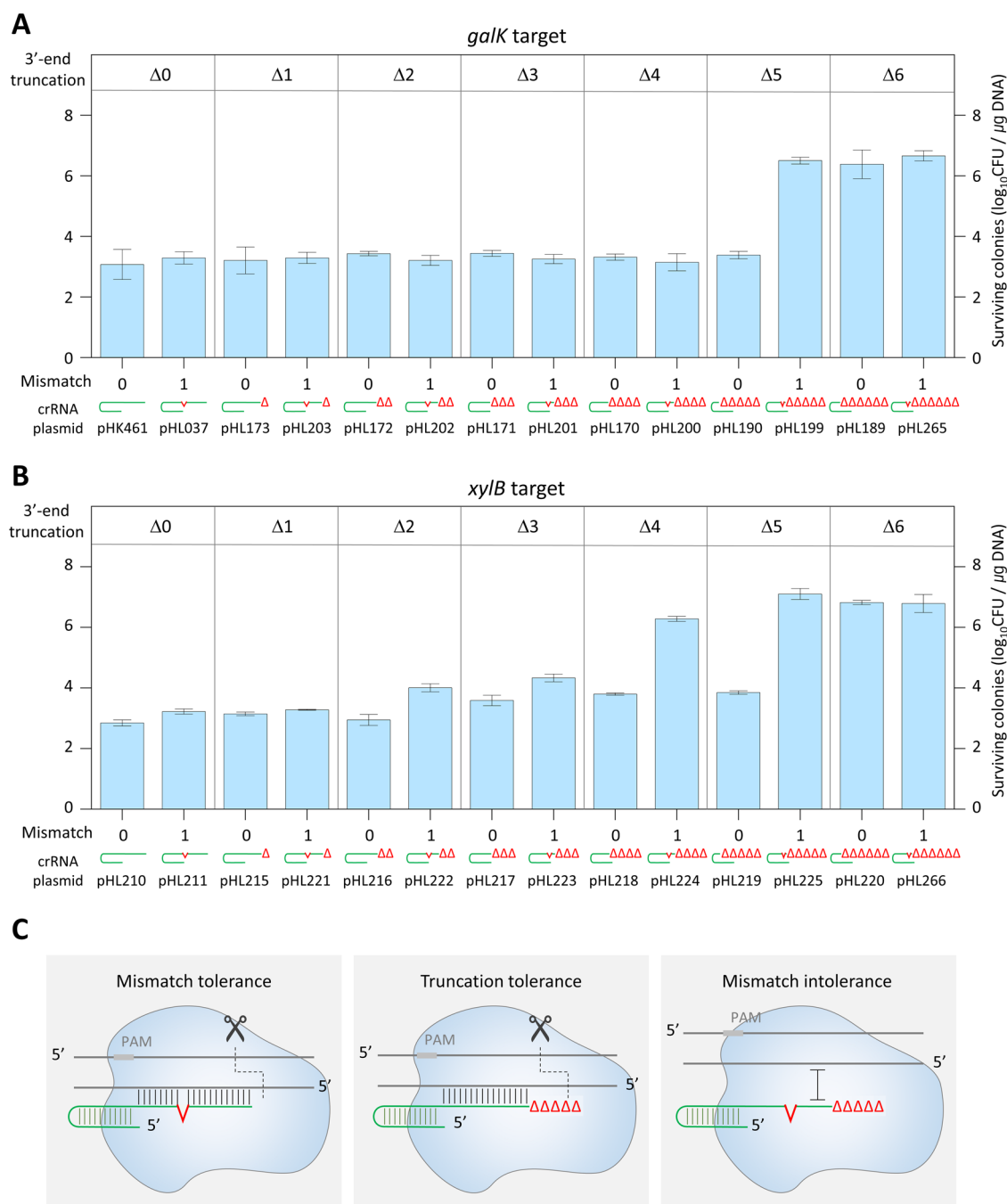


Figure 2. Effect of single-base mismatch in 3'-serially truncated crRNAs on Cpf1-mediated genomic DNA cleavage. (A) Genomic DNA cleavage efficiency of Cpf1 with *galK* target-mismatched and 3'-truncated crRNAs. The number of surviving colonies indicates the reversal of cleavage on the *galK* target by crRNA-Cpf1 complex. The large number ($>10^6$) of surviving cells represents the failure of target recognition and the cleavage of modified crRNAs by Cpf1. Each bar represents the mean value obtained from three independent experiments. (B) Genomic DNA cleavage efficiency of Cpf1 with *xylB* target-mismatched and 3'-truncated crRNAs. (C) Single-base mismatch intolerance of Cpf1 with 3'-5-nt-truncated crRNAs. Either a single-base mismatch or 5-nt truncation of crRNA against the target DNA is allowed for cleavage by the crRNA-Cpf1 complex.

RESULTS

Failure of Single-Base Genomic Editing by crRNA-Cpf1 Negative Selection. We electroporated mutagenic oligonucleotides and crRNA plasmids into both λ Bet protein- and Cpf1-overexpressing *Escherichia coli* MG1655 cells. Intracellular λ Bet proteins in pHK463 are necessary for the stabilization of mutagenic oligonucleotides.^{26,31} When the target sequence in *galK* is properly altered by mutagenic oligonucleotides (via the introduction of stop codons), the crRNA-Cpf1 complex cannot

digest the target genomic DNA, and the surviving cells form white colonies in D-galactose-containing MacConkey agar. When the target was unedited, most cells could not survive owing to double-strand cleavages at the genomic DNA targets, that is, negative selection (Figure 1A). We designed three different nucleotides to introduce stop codons in the middle of the *galK* gene sequence. Edited targets that did not match the crRNA sequence were designed not to be recognized and cleaved by the crRNA-Cpf1 complex (Figure S1).

In the absence of oligonucleotides or crRNA plasmids during electroporation, the white colonies did not form on D-galactose-containing MacConkey agar (Figure 1B). When oligonucleotides were electroporated with a crRNA-deleted plasmid (pHL259), white colonies (Gal⁻) could not be obtained on the MacConkey agar containing D-galactose and spectinomycin, and the number of surviving transformants was found to increase ($\approx 10^7/\mu\text{g}$ of DNA) (Figure S2). These results indicated that crRNA-Cpf1 or mutagenic oligonucleotides cannot edit the *galk* target independently.

When we electroporated both mutagenic oligonucleotides and crRNA plasmids, white colonies with one or two bases edited were rarely obtained (5–7%; Figure 1B). Conversely, when three bases were edited, the proportion of white colonies among the total number of colonies increased to up to 67%. Moreover, white colonies were randomly selected, and the nucleotide sequences of the edited *galk* targets were amplified and analyzed using Sanger sequencing (Figure S3). In case of single-base editing ($^{504}\text{T} \rightarrow \text{A}$), only one among 10 sequences was correctly altered. In cases of double and triple base editing, cells from five among five colonies showed correct sequence editing. These data indicate that the actual accuracy of single-base editing was very low (1/10), even among the phenotypically selected 5% cells (1/10 in 5% = 0.5%).

We designed crRNAs carrying one to four base mismatch(es) against the *galk* DNA target and transformed the crRNA plasmids into MG1655 cells overexpressing Cpf1 nucleases (Figure 1C). The number of viable transformants indicates how each modified crRNA-Cpf1 complex can efficiently recognize and cleave the DNA target. When we used crRNAs with one or two bases mismatched, the number of transformants obtained (less than 10^4) was approximately equal to the number of transformants obtained with perfectly matched crRNAs. However, when we used crRNAs with three or four bases mismatched, the number of surviving transformants increased remarkably. These data indicate that one or two mismatch(es) tolerate recognition and are cleaved as a target, whereas more than two mismatches are not cleaved in the cells. Therefore, we demonstrated that the editing of single or double bases was not successful owing to mismatch tolerance.

Intolerance of Single-Base Mismatch between the DNA Target and Maximally 3'-End-Truncated crRNA. Reportedly, truncated sgRNAs reduce the off-target effect of Cas9,²⁷ and the truncated crRNA-Cpf1 complex can cleave DNA targets.^{29,30} Therefore, we investigated how accurately a truncated crRNA could help Cpf1 recognize and cleave DNA targets in vivo. We constructed various plasmids carrying crRNAs with 1 to 6 nt truncated that recognized 20 to 15 nt in the *galk* (497–517) gene as a target. After the transformation of crRNA plasmids, the number of viable transformants (less than 10^4) containing crRNAs with 1 to 5 nt truncated was approximately equal to the number of transformants containing perfectly matched crRNAs. This implies that even crRNAs with 1 to 5 nt truncated helped retain the in vivo target cleavage activity of Cpf1 nuclease. However, the genomic DNA cleavage activity of the Cpf1 nuclease was not retained when crRNA with 6 nt truncated was used. These data showed that the truncation of 5 nt is the upper limit for Fncpf1 nuclease activity retention in the cells (Figure 2A).

Next, we attempted to determine the number of base truncations that led to the loss of single-base mismatch tolerance. 3'-end truncations of up to 4 nt can ensure tolerance of single-base mismatch in the genomic DNA target cleavage.

However, when single-base mismatch and 5-nt truncation occurred simultaneously in the crRNA, the number of surviving transformants increased to up to $10^6/\mu\text{g}$ of DNA. These data indicate that a single-base mismatch was not tolerated in the recognition and cleavage of the *galk* target DNA when 5-nt-truncated crRNA was combined with Cpf1 nuclease in the cells.

In case of the *xylB* target, as in the case of *galk*, we observed the same limit of 5-nt truncation of crRNA for functional Cpf1 nuclease activity. However, serially 3'-end-truncated crRNAs with single-base mismatches progressively increased the number of viable transformed cells (Figure 2B). Of note, a single-base mismatch is almost never allowed in 4- or 5-nt-truncated crRNAs. These data showed that 3'-end truncation of the crRNA affected the in vivo cleavage activity of Cpf1, and a single-base mismatch was incompatible with substantial truncation at the 3'-end of the crRNA for the target cleavage by Cpf1 (i.e., mismatch intolerance; Figure 2C).

Single-Base Genome Editing with Maximally 3'-End-Truncated crRNAs. Based on the results of single-base mismatch intolerance of 3'-5-nt-truncated crRNAs (Figure 2C), it was hypothesized that a single-base-edited target could not be cleaved by 3'-5-nt-truncated crRNAs, whereas an unedited target could be cleaved for efficient negative selection (Figure 3A). We tested various lengths of truncated crRNAs for single-base genome editing. We electroporated single-mutagenic oligonucleotides ($^{510}\text{C} \rightarrow \text{A}$ at *galk*) with crRNA plasmids into both Cpf1- and Bet-overexpressing cells and spread the transformants on MacConkey agar containing D-galactose and spectinomycin. The proportions of white colonies obtained using 1- to 4-nt-truncated crRNAs (3–12%) and using untruncated crRNAs (5%) were similar (Figure 3B). When 3'-5-nt-truncated crRNAs were used, 88% of the transformants formed white colonies on the MacConkey (D-galactose) agar. We randomly selected the white colonies and analyzed the nucleotide sequences using Sanger sequencing (Figure S4). However, in the case of 3'-6-nt-truncated crRNAs, the number of viable cells increased to up to $10^7/\mu\text{g}$ of DNA, and no white colonies were observed. This implies that Cpf1 was not functional with 3'-6-nt-truncated crRNAs.

Another single-base alteration at a different genomic target was assessed using the same approach. We electroporated single-mutagenic oligonucleotides ($^{643}\text{G} \rightarrow \text{T}$ at *xylB*) with crRNA plasmids and spread the transformants on MacConkey agar containing D-xylose and spectinomycin. The proportion of white colonies gradually increased as the number of truncated bases increased (from $\Delta 0$ to $\Delta 5$; Figure 3C). We obtained the highest proportion (87%) of white colonies on the MacConkey (D-xylose) agar using 3'-5-nt-truncated crRNAs. However, when we used 3'-6-nt-truncated crRNAs, the number of viable cells increased to up to $10^7/\mu\text{g}$ of DNA, and we did not observe white colonies. These data indicated that Cpf1-mediated single-base genome editing could be performed efficiently using maximally 3'-end-truncated crRNAs.

Various Single-Base Editing Events in the Target Range Using Maximally 3'-End-Truncated crRNAs. We investigated whether 3'-end-truncated crRNAs could be used for the single-base editing of four types of nucleotides at various positions in the target (Figure 4). In the *galk* (497–512) target region, we selected two of each of G, A, T, and C (eight targets in all) as targets. Different oligonucleotides mutagenic at each of the eight positions were synthesized and electroporated with the untruncated crRNA plasmid (pHK461) and 3'-5-nt crRNA plasmid (pHL190), respectively. The transformed cells were

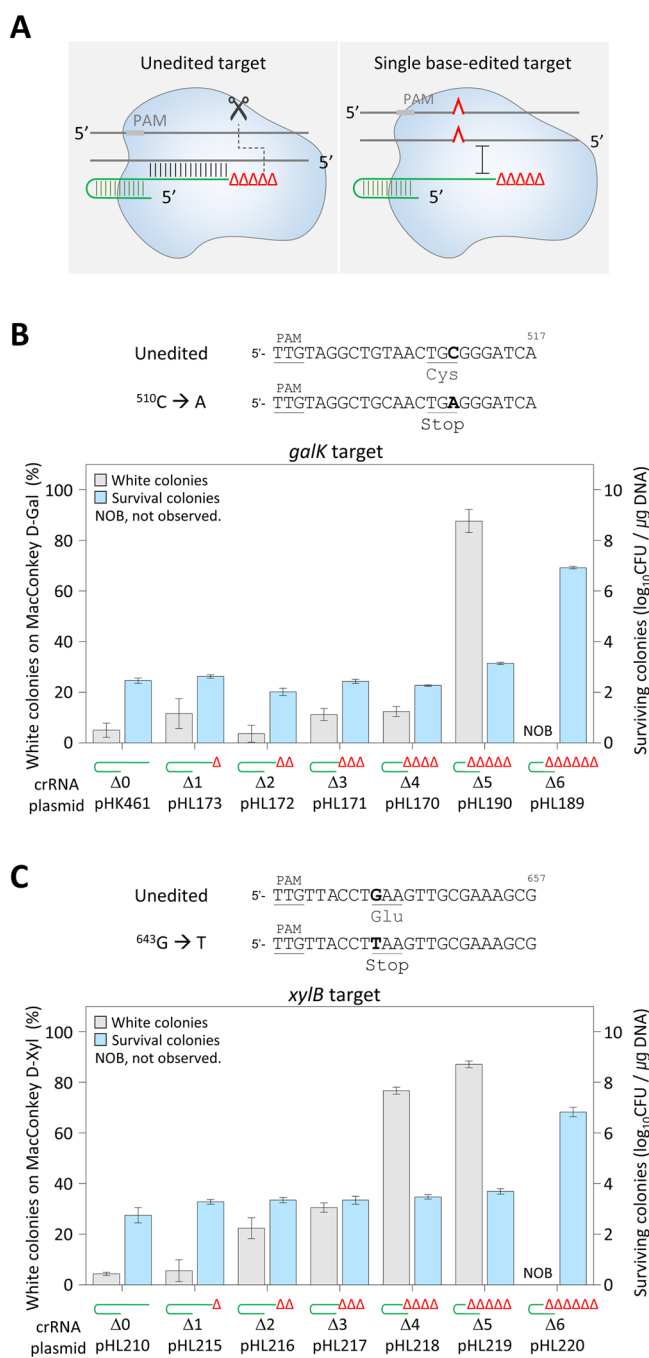


Figure 3. Efficiency of Cpf1-mediated single-base substitutions along the 3'-nt truncation lengths of crRNA. (A) Proposed method for the negative selection of the single-base-edited target facilitated by the mismatch intolerance of the truncated crRNA-Cpf1 complex. The unedited target could be cleaved by the 3'-5'-nt-truncated crRNA-Cpf1 complex (left). The single-base-edited target could not be recognized or cleaved by the truncated crRNA-Cpf1 complex (right). (B) and (C) Optimization of the 3'-end truncation length of crRNAs for Cpf1-mediated single-base alteration in the *galk* and *xylB* gene sequences, respectively. We tested six different 3'-end-truncated crRNAs for negative selection of single-base editing. Δ0 to Δ6 indicate the number of truncated nucleotides at the 3'-end of the crRNAs. Each bar represents the mean from three independent experiments.

spread on Luria–Bertani (LB) agar supplemented with spectinomycin instead of MacConkey agar, and four colonies were randomly selected in each case. Subsequently, the edited

nucleotide sequences were verified using Sanger sequencing (Figure S5). The success of base editing was indicated by at least one correct base editing among cells from the four selected colonies. Only one correct base editing was successful among 24 cases when untruncated crRNA (from pHK461) was used (Figure 4A). However, base editing was successful in 19 out of 24 cases (79.1%) when the 3'-5'-nt crRNA (pHL190) was used. At least one base editing was successful in all eight bases, and editing of all three different bases was successful for five bases (G_1 , A_1 , T_1 , A_2 , and T_2).

In the *xylB* (637–652) target region, editing was successful in 12 out of 24 cases (50%) using 3'-5'-nt-truncated crRNA (pHL219) (Figure 4B), the efficiency of which was slightly lower than that in the *galk* gene (79.1%). Based on the Sanger sequencing results (Figure S6), we observed at least one successful base editing at all sites, except at the A_2 position of the *xylB* gene. However, only one base editing out of the 24 cases was successful when the untruncated crRNA (from pHL210) was used. Therefore, for single-base editing, the use of maximally 3'-end-truncated crRNAs was more efficient than the use of untruncated crRNAs for both *galk* and *xylB* genes.

The success rate of base editing was analyzed for the mutation types (i.e., transition and transversion; Figure 4C). The success rate of transversion (pyrimidine to purine or vice versa) was slightly higher than that of transition (purine to purine or pyrimidine to pyrimidine). These data indicate that the truncated crRNA-Cpf1 complex can be used for both transitions and transversions in the DNA targets, regardless of the position of the target nucleotides.

Insertion and Deletion of Single Nucleotides by the 3'-End-Truncated crRNAs. For further investigating the 3'-end-truncated crRNA method, we evaluated how efficiently it can be applied to indels of single nucleotides during genome editing. Mutagenic oligonucleotides are designed to generate stop codons that disrupt the *galk* and *xylB* genes when the mutations are properly introduced. We tested an intact crRNA (Δ0) and three different truncated crRNAs (Δ4, Δ5, and Δ6) (Figure 5). For the 4-nt-truncated crRNA (Δ4), the proportion of white colonies formed upon the introduction of single indels was slightly higher than that achieved using the untruncated crRNA (Δ0). When the 5-nt-truncated crRNA (Δ5) was used, the proportions of white colonies formed upon the insertion (510G) and deletion (509ΔG) in the *galk* target increased markedly to 79 and 76%, respectively (Figure 5A). Similarly, when the 5-nt-truncated crRNA (Δ5) was used, the proportions of white colonies formed upon the insertion (643A) and deletion (643ΔG) in the *xylB* target increased considerably to 62 and 58%, respectively (Figure 5B). The single-nucleotide indel was verified using Sanger sequencing (Figure S7). When the 6-nt-truncated crRNA (Δ6) was used, we did not observe white colonies, and the number of transformants increased drastically ($\sim 10^7$ /μg of DNA), as observed in the abovementioned base substitution experiments.

In addition, we also tested the insertion and deletion of 2 nt in *galk* and *xylB* using 5-nt-truncated crRNAs. The efficiencies of the insertion (510AG) and deletion (509ΔGC) of 2 nt in the *galk* gene were observed to be 54, and 64%, respectively (Figure S8). In the case of *xylB*, the efficiencies of the insertion (643AG) and deletion (643ΔGA) of the 2 nt were observed to be 37 and 43%, respectively. These data indicate that the maximally 3'-end-truncated crRNA method also yields efficient results for the indel of single or double nucleotide(s) in the genomic DNA target.

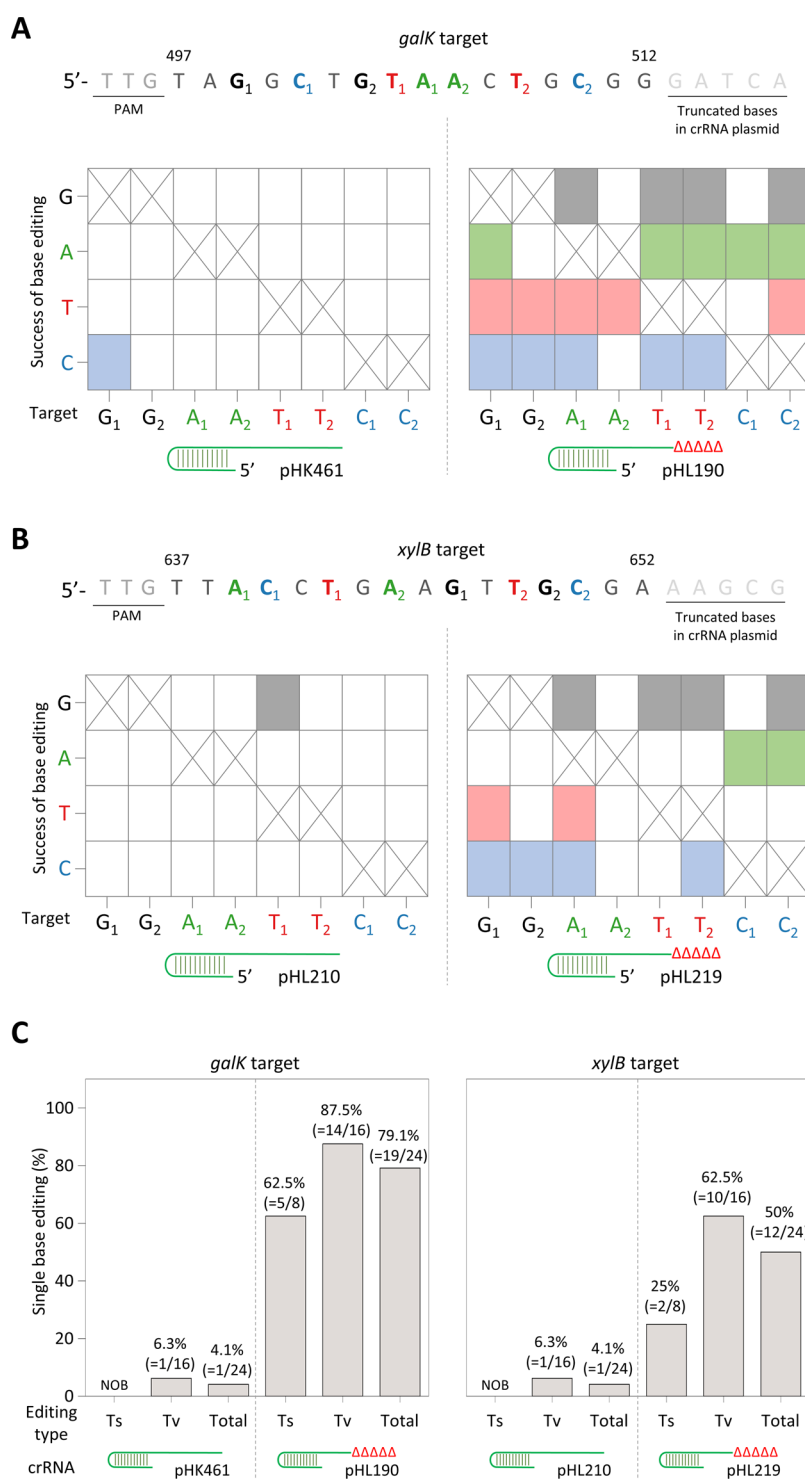


Figure 4. Single-base editing at various positions in the target DNA sequence achieved using the maximally 3'-end-truncated crRNA-Cpf1 complex. (A) Comparison of the single-base editing efficiency of the *galK* target using untruncated (left) and 3'-5-nt-truncated (right) crRNAs. The colored grids indicate successful single-base editing in at least one of the four cells selected randomly from the colonies on LB agar containing spectinomycin. The boxes with the X symbol were not applicable for base editing because they represent cases of base alterations to the same base. (B) Comparison of single-base editing efficiency in the *xyfB* target using untruncated (left) and 3'-5-nt-truncated (right) crRNAs. (C) Comparison of the efficiencies of successful single-base editing using untruncated and 3'-5-nt-truncated crRNAs. Two types of base editing—transitions (Ts) and transversions (Tv)—were separately analyzed in the DNA target.

DISCUSSION

The off-target effect, in which a target similar to the target of CRISPR/Cas is affected, is a major roadblock in accurate genome editing and has been widely reported in higher

organisms. Mismatch tolerance is considered to be one of the causes of off-target effects.³² As a defense system of the bacterial host, the CRISPR/Cas system is thought to effectively cleave foreign genes even when there is a minor change (~1 or 2 nt) in the recognition sequence caused by natural mutation. The

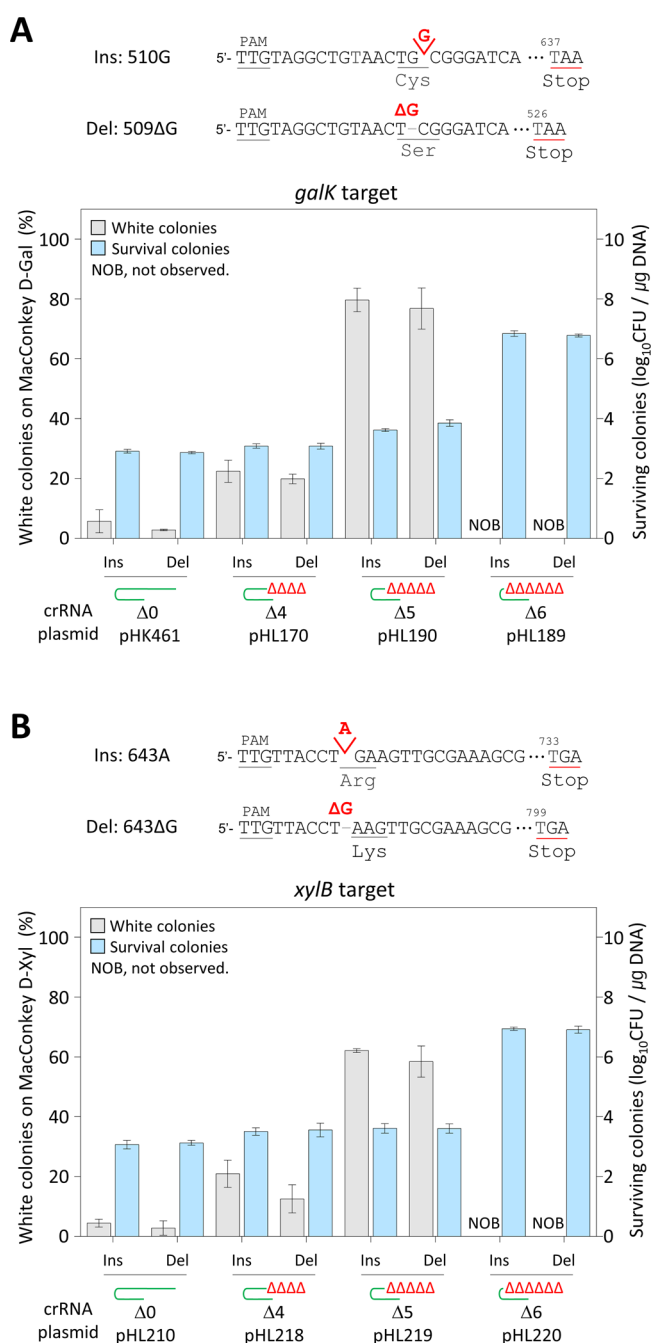


Figure 5. Cpf1-mediated insertion and deletion of nucleotides using 3'-end-truncated crRNAs. The efficiencies of insertion and deletion of single nucleotides in the *galk* (A) and *xylB* (B) genes were tested using 3'-end-truncated crRNAs. The editing efficiency was assessed by counting the white colonies on MacConkey agar containing D-galactose or D-xylose. The number of surviving colonies per unit crRNA plasmid added represents the genomic DNA breakage activity of the crRNA-Cpf1 complex. Δ0, Δ4, Δ5, and Δ6 represent the number of 3'-end-truncated nucleotides in each crRNA. Each bar represents the mean from three independent experiments.

mismatch tolerance of CRISPR/Cas may not pose a problem in bacterial defense because the size of the bacterial genome is considerably smaller than the number of possible permutations, with repetitions of approximately 20 nt ($\sim 4^{20}$) in the target recognition sequence.²⁶ Mismatch tolerance of the CRISPR/Cpf1 system for one or two base(s) was also observed in our study (Figure 1C).

A Cpf1-mediated negative selection system was constructed, in which microbial cells with nonmutated targets were removed, and only cells with mutated targets survived (Figure 1A). When one or two bases in the *galk* gene were edited, the formation of white colonies (the phenotype of cells with successful editing) on MacConkey agar containing D-galactose, a result of negative selection by Cpf1, was barely observed (Figure 1B). The results of the nucleotide sequence analysis of the target of cells from the white colonies was noteworthy. We observed various point mutations around the single-base-edited targets from the white colonies. However, in case of double-base editing, all white colony cells showed proper editing (Figure S3). This presumably resulted from the difference in the degree of mismatch tolerance for single-base-edited targets rather than from the impurity of synthesized oligonucleotides. It is speculated that additional mutations may be introduced into the edited target owing to errors in DNA repair or religation during repetitive double-strand break introduction in single-base-edited targets by CRISPR/Cpf1, which allows mismatches.

When only the crRNA plasmid was introduced into cells without using mutagenic oligonucleotides, it was anticipated that the cells transformed by crRNA plasmids would be killed by double-strand breakage caused by the crRNA-Cpf1 complex in the genome. However, the transformed cells consistently formed 10^3 – 10^4 colonies/μg of DNA on the agar (Figure 1B). When a crRNA-deleted plasmid was used for transformation, the number of transformants was approximately 10^6 – 10^7 /μg of DNA. Therefore, it appeared that negative selection by CRISPR/Cpf1 was unsuccessful in 1 cell per 10^3 – 10^4 cells. This was probably owing to subpopulation heterogeneity, in which case, Cpf1 would not work, or owing to occasional religation following double-strand breakage.

For assessing the effect of crRNA truncation on-target recognition and cleavage, we introduced single-base mismatches against the target into 3'-end-truncated crRNAs and compared the number of surviving transformants (Figure 2). It was found that even if 5 nt were truncated at the 3'-end, the DNA targets (*galk* and *xylB*) could be cleaved efficiently in vivo. However, 6-nt-truncated crRNAs (Δ6) could not recognize the targets (Figure 2A,B). Perhaps, because of the insufficient number of DNA-crRNA base pairings, the active nuclease form of the DNA-crRNA-Cpf1 ternary complex was not properly formed. In case of maximally 5-nt-truncated functional crRNAs (Δ5), a single-base mismatch could be clearly distinguished. As a crRNA with both single-base mismatch and maximally tolerated truncation (Δ5) cannot recognize the target, it can be considered that maximally 3'-end-truncated crRNAs exhibit mismatch intolerance (Figure 2C).

Therefore, it was expected that the maximally 3'-end-truncated crRNA could distinguish between the single-base-edited and unedited targets in the genome (Figure 3A). Single-base editing was performed using various lengths of 3'-end-truncated crRNAs. It was shown that the *galk* target with C510A single-base editing was successfully negatively selected with a high efficiency when the 5-nt-truncated crRNA was used (Figure 3B). In case of *xylB*, as the number of truncated nucleotides increased, the proportion of edited cells increased (Figure 3C). These results showed that the in vivo DNA cleavage activity of Cpf1 is affected by the length of the target recognition sequence of the crRNA, and maximally truncated crRNAs facilitate highly efficient single-nucleotide editing.

Single-base editing at various positions in the target DNA sequence was attempted in *galk* and *xylB* genes using the

maximally 3'-end-truncated crRNA-Cpf1 complex (Figure 4A,B). The success rates of 24 single-base editing events in the *galK* and *xylB* genes were 79.1 and 50%, respectively. As observed in the case of using target-mismatched sgRNA in Cas9,²⁶ transition was slightly less efficient than transversion in this study (Figure 4C). This was probably because the pairing of a pyrimidine to a pyrimidine and a purine to a purine between the target DNA and crRNA by transversion is thought to exert a more profound effect on-target recognition, which makes it easier to distinguish between edited and unedited sequences. Although the target-mismatched crRNA method requires a new crRNA construct for the editing of each base by Cpf1,³ the maximally 3'-end-truncated crRNA method can be applied for editing at multiple positions of the target using a single construct of maximally 3'-end-truncated crRNA.

The efficiency of nucleotide indel mutation showed an overall trend similar to that of base substitution (Figure 5). When a 4-nt-truncated crRNA ($\Delta 4$) was used, the white colony ratio increased considerably to 76% in single-base substitution in the *xylB* ⁶⁴³G target (Figure 3C), whereas the percentage of white colonies was as low as 17–22% in a single-nucleotide indel in *xylB* ⁶⁴³G (Figure 5B). In addition, the efficiency of the double nucleotide indel was lower than that of single-nucleotide indel. These results might be attributed to the differences in the action of the DNA repair systems on the mismatched bases and DNA bulge³³ during base substitution and indel mutagenesis, respectively.

In case of *xylB*, as the length of the target recognition sequence of 3'-end-truncated crRNAs reduced, the in vivo target cleavage activity progressively decreased, but the target specificity clearly increased (Figure 2B). The result was similar to that in vitro target cleavage activity was decreased gradually by the increase of 3'-end truncations.³⁴ It is assumed that the 3'-end of untruncated crRNAs is necessary for mismatch tolerance to improve the target recognition flexibility or Cpf1 activity for the cleavage of DNA targets modified by natural mutation (containing one or two base mutations, and 1 or 2 nt indels). The shortened target recognition sequence of crRNAs increases the mathematical probability of encountering the same sequence in the genome, which does not appear to be advantageous for increasing target specificity in the editing of the genomes of higher organisms. Because the number of cases with 16 nt target recognition sequences in 5-nt-truncated crRNA is 4^{16} ($\sim 4 \times 10^9$), the maximally 3'-end-truncated crRNA method can be applied to genome editing in higher organisms with large genome sizes as well as to microorganisms. As maximally 3'-end-truncated crRNAs can strictly recognize the target and effectively discriminate among single-base changes in the target, it can serve as a simple and efficient CRISPR/Cas-based method for accurate genome editing at a single-base/-nucleotide resolution, such as in indel correction.

METHODS

Strains and Culture Conditions. The *E. coli* strains used in this study are listed in Table S1; we cultured the strains in LB broth at 30 or 37 °C, depending on the *ori* sequences in the plasmids. *E. coli* DH5 α and MG1655 were used as cloning hosts and for genomic integration of the *cpf1*-KmR cassette, respectively. We cultured *E. coli* MG1655 cells in LB broth at 30 °C until the optical density at 600 nm (OD_{600}) approached 0.4. Subsequently, electrocompetent cells were harvested, washed twice, resuspended in 10% glycerol solution, aliquoted, and stored at -80 °C. The construction of an *E. coli* strain

carrying the *cpf1* gene in its chromosome is described below. As required, ampicillin, kanamycin, or spectinomycin at 50, 25, or $75 \mu\text{g mL}^{-1}$, respectively, was added to the culture medium. To overexpress Cpf1 in the chromosome or λ Bet in pHK463, L-arabinose (final concentration: 1 mM) was added to the culture flasks. After the OD_{600} reached 0.4, the cells were further cultured for 3 h.

Genomic Integration of the *cpf1* Gene. The primers used for the DNA amplification of the strains are listed in Table S2. The *cpf1* gene was PCR-amplified using the pJYS1Ptac plasmid (provided generously by Sheng Yang; Addgene plasmid # 85545) as a template and was fused with a kanamycin resistance marker using overlap PCR to generate a *cpf1*-KmR cassette. The *cpf1*-KmR cassette was amplified using primer pairs harboring homologous DNA sequences for recombineering in the arabinose operon. Subsequently, the purified *cpf1*-KmR PCR products were electroporated into *E. coli* MG1655 cells carrying the pKD46 plasmid after λ recombinases were fully induced by treatment with L-arabinose. After the pKD46 plasmid was cured at 42 °C, the strain was designated as *E. coli* HK1061, in which the *cpf1* gene was located downstream of the L-arabinose-inducible P_{BAD} promoter in the chromosome.

Plasmid Construction. The crRNA plasmids used are listed in Table S1, and the primers used are listed in Table S2. All crRNA plasmids were constructed as follows: the *galK* (497–517) and *xylB* (637–657) sequences in the *E. coli* genome were chosen as DNA target sequences recognized by crRNAs for Cpf1-mediated genome editing. Two overlapping DNA fragments containing the spectinomycin resistance gene and the crRNA gene were amplified using pJYS2_crtYf (generously provided by Sheng Yang; Addgene plasmid # 85544) as a template. Two fragments were assembled using Gibson Assembly Master Mix (NEB, Ipswich, MA, USA) to generate pHK461. All other crRNA plasmids and a crRNA-deleted plasmid were amplified using pHK461 as the template and subjected to Gibson assembly.

In Vivo Target Cleavage Assay Using Modified crRNAs.

Various crRNA plasmids that recognized the *galK* or *xylB* gene sequences as the DNA target were constructed. Each crRNA was modified to introduce 1 to 4 bp mismatches and 1 to 6 nt truncations at the 3'-end, or 1-bp mismatch and 1 to 6 nt truncations at the 3'-end. The modified crRNA plasmids (200 ng) were electroporated into Cpf1-overexpressing HK1061 cells. Electroporation was performed under the same conditions described above. The number of surviving colonies was determined to evaluate target recognition and cleavage by the modified crRNA-Cpf1 complex.

Oligonucleotide-Directed Genome Editing. The mutagenic oligonucleotides used for genome editing are listed in Table S3. Mutagenic oligonucleotides (100 pmol) and crRNA plasmids (200 ng) were electroporated simultaneously into the HK1061 cells carrying the pHK463 plasmid, in which both the λ Bet protein and Cpf1 were overexpressed in response to the addition of L-arabinose. Electroporation was performed at 25 μF , 200 Ω , and 1.8 kV, and a 0.1-cm electroporation cuvette was used. Afterward, the cells were transferred to 950 mL of SOC immediately and incubated for 1 h at 37 °C and 180 rpm for recovery. Thereafter, the recovered cells were spread on MacConkey agar containing D-galactose or D-xylose (0.5%) and spectinomycin ($75 \mu\text{g mL}^{-1}$) and incubated for 16 h at 37 °C. The numbers of surviving colonies and white colonies were counted for assessing the genome editing efficiencies. The white

colonies were randomly selected for Sanger sequencing of edited genomic DNA targets.

Base Substitution and Indel Editing. For single-base substitution experiments in *galK* and *xylB* targets, mutagenic oligonucleotides generating stop codons and each of the two types of crRNA plasmids (i.e., untruncated and 5-nt-truncated crRNA plasmids) were electroporated into the HK1061 cells carrying the pHK463 plasmid, as mentioned above. The recovered cells were spread on LB agar plates containing spectinomycin ($75 \mu\text{g mL}^{-1}$). After incubation for 16 h at 37°C , four colonies (per electroporation experiment) were randomly selected, and Sanger sequencing was carried out to verify the desired single-base editing in the *galK* and *xylB* targets. The primers used for PCR amplification and Sanger sequencing are listed in Table S2.

The indel mutagenic oligonucleotides could also generate stop codons that caused the premature termination of GalK or XylB protein translation. Electroporation was performed under the same conditions described above. The recovered cells were spread on MacConkey agar containing D-galactose or D-xylose (0.5%) and spectinomycin ($75 \mu\text{g mL}^{-1}$). The cells obtained from four white colonies from the MacConkey agar plate were subjected to Sanger sequencing for the confirmation of accurate single-base editing. The indel mutation efficiencies were calculated based on the ratio of the number of white colonies to the total number of colonies on the MacConkey agar plates.

■ ASSOCIATED CONTENT

SI Supporting Information

The Supporting Information is available free of charge at <https://pubs.acs.org/doi/10.1021/acssynbio.2c00054>.

Primers, plasmids, mutagenic oligonucleotides, base editing results, and Sanger sequences (PDF)

■ AUTHOR INFORMATION

Corresponding Author

Sang Jun Lee – Department of Systems Biotechnology and Institute of Microbiomics, Chung-Ang University, Anseong 17546, Republic of Korea; orcid.org/0000-0002-2803-753X; Phone: +82-31-670-3356; Email: sangjlee@cau.ac.kr

Authors

Ho Jung Lee – Department of Systems Biotechnology and Institute of Microbiomics, Chung-Ang University, Anseong 17546, Republic of Korea; orcid.org/0000-0002-4946-0230

Hyun Ju Kim – Department of Systems Biotechnology and Institute of Microbiomics, Chung-Ang University, Anseong 17546, Republic of Korea; orcid.org/0000-0002-5852-9578

Young-Jun Park – Environmental Disease Research Center, Korea Research Institute of Bioscience and Biotechnology (KRIBB), Daejeon 34141, Republic of Korea; orcid.org/0000-0002-9958-3976

Complete contact information is available at:

<https://pubs.acs.org/doi/10.1021/acssynbio.2c00054>

Author Contributions

H.J.L. and H.J.K. contributed equally to this work. H.J.L. contributed to the study design, performed experiments and data analysis, and wrote the manuscript; H.J.K. performed experiments, data analysis, and contributed to the manuscript

writing; Y.P. contributed to the funding acquisition and manuscript writing. S.J.L. contributed to the study design, funding acquisition, data analysis, and manuscript writing.

Notes

The authors declare the following competing financial interest(s): H.J.L., H.J.K., and S.J.L. have filed a patent application based on this work.

■ ACKNOWLEDGMENTS

This study was supported by the Korea Research Institute of Bioscience and Biotechnology (KRIBB) Research Initiative Program (KGM532214) and the National Research Foundation of Korea (2021R1A2C1013606), Republic of Korea.

■ ABBREVIATIONS

CRISPR/Cpf1, Clustered Regularly Interspaced Short Palindromic Repeats/CRISPR from *Prevotella* and *Francisella* 1; crRNA, CRISPR RNA; Cas, CRISPR-associated protein; FnCpf1, Cpf1 derived from *Francisella tularensis* subsp. *novicida* strain U112; *E. coli*, *Escherichia coli*; PAM, Protospacer adjacent motif; PCR, Polymerase chain reaction; Cas9, CRISPR-associated protein 9

■ REFERENCES

- (1) Jinek, M.; Chylinski, K.; Fonfara, I.; Hauer, M.; Doudna, J. A.; Charpentier, E. A programmable Dual-RNA-Guided DNA Endonuclease in Adaptive Bacterial Immunity. *Science* **2012**, *337*, 816–821.
- (2) Anzalone, A. V.; Koblan, L. W.; Liu, D. R. Genome editing with CRISPR-Cas nucleases, base editors, transposases and prime editors. *Nat. Biotechnol.* **2020**, *38*, 824–844.
- (3) Kim, H. J.; Oh, S. Y.; Lee, S. J. Single-base genome editing in *Corynebacterium glutamicum* with the help of negative selection by target-mismatched CRISPR/Cpf1. *J. Microbiol. Biotechnol.* **2020**, *30*, 1583–1591.
- (4) Kim, H.; Kim, S. T.; Ryu, J.; Kang, B. C.; Kim, J. S.; Kim, S. G. CRISPR/Cpf1-mediated DNA-free plant genome editing. *Nat. Commun.* **2017**, *8*, 14406.
- (5) Kim, Y.; Cheong, S. A.; Lee, J. G.; Lee, S. W.; Lee, M. S.; Baek, I. J.; Sung, Y. H. Generation of knockout mice by Cpf1-mediated gene targeting. *Nat. Biotechnol.* **2016**, *34*, 808–810.
- (6) Kim, D.; Kim, J.; Hur, J. K.; Been, K. W.; Yoon, S. H.; Kim, J. S. Genome-wide analysis reveals specificities of Cpf1 endonucleases in human cells. *Nat. Biotechnol.* **2016**, *34*, 863–868.
- (7) Swarts, D. C.; Jinek, M. Mechanistic Insights into the *cis*- and *trans*-Acting DNase Activities of Cas12a. *Mol. Cell* **2019**, *73*, 589–600.e4.
- (8) Li, S. Y.; Cheng, Q. X.; Wang, J. M.; Li, X. Y.; Zhang, Z. L.; Gao, S.; Cao, R. B.; Zhao, G. P.; Wang, J. CRISPR-Cas12a-assisted nucleic acid detection. *Cell Discov.* **2018**, *4*, 20.
- (9) Gootenberg, J. S.; Abudayyeh, O. O.; Kellner, M. J.; Joung, J.; Collins, J. J.; Zhang, F. Multiplexed and portable nucleic acid detection platform with Cas13, Cas12a, and Csm6. *Science* **2018**, *360*, 439–444.
- (10) Safari, F.; Zare, K.; Negahdaripour, M.; Barekati-Mowahed, M.; Ghasemi, Y. CRISPR Cpf1 proteins: structure, function and implications for genome editing. *Cell Biosci.* **2019**, *9*, 36.
- (11) Zhang, H.; Cheng, Q. X.; Liu, A. M.; Zhao, G. P.; Wang, J. A Novel and Efficient Method for Bacteria Genome Editing Employing both CRISPR/Cas9 and an Antibiotic Resistance Cassette. *Front. Microbiol.* **2017**, *8*, 812.
- (12) Arroyo-Olarte, R. D.; Bravo Rodríguez, R.; Morales-Ríos, E. Genome Editing in Bacteria: CRISPR-Cas and Beyond. *Microorganisms* **2021**, *9*, 844.
- (13) Rock, J. M.; Hopkins, F. F.; Chavez, A.; Diallo, M.; Chase, M. R.; Gerrick, E. R.; Pritchard, J. R.; Church, G. M.; Rubin, E. J.; Sasseti, C. M.; et al. Programmable transcriptional repression in mycobacteria using an orthogonal CRISPR interference platform. *Nat. Microbiol.* **2017**, *2*, 16274.

- (14) Wendt, K. E.; Ungerer, J.; Cobb, R. E.; Zhao, H.; Pakrasi, H. B. CRISPR/Cas9 mediated targeted mutagenesis of the fast growing cyanobacterium *Synechococcus elongatus* UTEX 2973. *Microb. Cell Fact.* **2016**, *15*, 115.
- (15) Murugan, K.; Seetharam, A. S.; Severin, A. J.; Sashital, D. G. CRISPR-Cas12a has widespread off-target and dsDNA-nicking effects. *J. Biol. Chem.* **2020**, *295*, 5538–5553.
- (16) Kim, D.; Lim, K.; Kim, D. E.; Kim, J. S. Genome-wide specificity of dCpf1 cytidine base editors. *Nat. Commun.* **2020**, *11*, 4072.
- (17) Komor, A. C.; Kim, Y. B.; Packer, M. S.; Zuris, J. A.; Liu, D. R. Programmable editing of a target base in genomic DNA without double-stranded DNA cleavage. *Nature* **2016**, *533*, 420–424.
- (18) Zhao, D.; Li, J.; Li, S.; Xin, X.; Hu, M.; Price, M. A.; Rosser, S. J.; Bi, C.; Zhang, X. Glycosylase base editors enable C-to-A and C-to-G base changes. *Nat. Biotechnol.* **2021**, *39*, 35–40.
- (19) Lee, S. S.; Ding, N.; Sun, Y. D.; Yuan, T. L.; Li, J.; Yuan, Q. C.; Liu, L. Z.; Yang, J.; Wang, Q.; Kolomeisky, A. B.; Hilton, I. B.; Zuo, E.; Gao, X. Single C-to-T substitution using engineered APOBEC3G-nCas9 base editors with minimum genome- and transcriptome-wide off-target effects. *Sci. Adv.* **2020**, *6*, No. eaba1773.
- (20) Anzalone, A. V.; Randolph, P. B.; Davis, J. R.; Sousa, A. A.; Koblan, L. W.; Levy, J. M.; Chen, P. J.; Wilson, C.; Newby, G. A.; Raguram, A.; et al. Search-and-replace genome editing without double-strand breaks or donor DNA. *Nature* **2019**, *576*, 149–157.
- (21) Zhang, X.; Xu, L.; Fan, R.; Gao, Q.; Song, Y.; Lyu, X.; Ren, J.; Song, Y. Genetic editing and interrogation with Cpf1 and caged truncated pre-tRNA-like crRNA in mammalian cells. *Cell Discov.* **2018**, *4*, 36.
- (22) Park, H. M.; Liu, H.; Wu, J.; Chong, A.; Mackley, V.; Fellmann, C.; Rao, A.; Jiang, F.; Chu, H.; Murthy, N.; et al. Extension of the crRNA enhances Cpf1 gene editing *in vitro* and *in vivo*. *Nat. Commun.* **2018**, *9*, 3313.
- (23) Bin Moon, S.; Lee, J. M.; Kang, J. G.; Lee, N. E.; Ha, D. I.; Kim, D. Y.; Kim, S. H.; Yoo, K.; Kim, D.; Ko, J. H.; et al. Highly efficient genome editing by CRISPR-Cpf1 using CRISPR RNA with a uridinylate-rich 3'-overhang. *Nat. Commun.* **2018**, *9*, 3651.
- (24) McMahon, M. A.; Prakash, T. P.; Cleveland, D. W.; Bennett, C. F.; Rahdar, M. Chemically Modified Cpf1-CRISPR RNAs Mediate Efficient Genome Editing in Mammalian Cells. *Mol. Ther.* **2018**, *26*, 1228–1240.
- (25) Kim, H.; Lee, W. J.; Oh, Y.; Kang, S. H.; Hur, J. K.; Lee, H.; Song, W.; Lim, K. S.; Park, Y. H.; Song, B. S.; et al. Enhancement of target specificity of CRISPR-Cas12a by using a chimeric DNA-RNA guide. *Nucleic Acids Res.* **2020**, *48*, 8601–8616.
- (26) Lee, H. J.; Kim, H. J.; Lee, S. J. CRISPR-Cas9-mediated pinpoint microbial genome editing aided by target-mismatched sgRNAs. *Genome Res.* **2020**, *30*, 768–775.
- (27) Fu, Y. F.; Sander, J. D.; Reyon, D.; Cascio, V. M.; Joung, J. K. Improving CRISPR-Cas nuclease specificity using truncated guide RNAs. *Nat. Biotechnol.* **2014**, *32*, 279–284.
- (28) Lee, H. J.; Kim, H. J.; Lee, S. J. Mismatch Intolerance of 5'-Truncated sgRNAs in CRISPR/Cas9 Enables Efficient Microbial Single-Base Genome Editing. *Int. J. Mol. Sci.* **2021**, *22*, 6457.
- (29) Kim, H. K.; Song, M.; Lee, J.; Menon, A. V.; Jung, S.; Kang, Y. M.; Choi, J. W.; Woo, E.; Koh, H. C.; Nam, J. W.; et al. *In vivo* high-throughput profiling of CRISPR-Cpf1 activity. *Nat. Methods* **2017**, *14*, 153–159.
- (30) Kleinstiver, B. P.; Tsai, S. Q.; Prew, M. S.; Nguyen, N. T.; Welch, M. M.; Lopez, J. M.; McCaw, Z. R.; Aryee, M. J.; Joung, J. K. Genome-wide specificities of CRISPR-Cas Cpf1 nucleases in human cells. *Nat. Biotechnol.* **2016**, *34*, 869–874.
- (31) Costantino, N.; Court, D. L. Enhanced levels of λ Red-mediated recombinants in mismatch repair mutants. *Proc. Natl. Acad. Sci. U. S. A.* **2003**, *100*, 15748–15753.
- (32) Feng, H.; Guo, J.; Wang, T.; Zhang, C.; Xing, X. H. Guide-target mismatch effects on dCas9-sgRNA binding activity in living bacterial cells. *Nucleic Acids Res.* **2021**, *49*, 1263–1277.
- (33) Hsiao, Y. Y.; Fang, W. H.; Lee, C. C.; Chen, Y. P.; Yuan, H. S. Structural insights into DNA repair by RNase T—an exonuclease processing 3' end of structured DNA in repair pathways. *PLoS Biol.* **2014**, *12*, No. e1001803.
- (34) 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. Cpf1 Is a Single RNA-Guided Endonuclease of a Class 2 CRISPR-Cas System. *Cell* **2015**, *163*, 759–771.