Published in final edited form as:

Nat Biomed Eng. 2017 May ; 1(5): . doi:10.1038/s41551-017-0066.

Engineering CRISPR-Cpf1 crRNAs and mRNAs to maximize genome editing efficiency

Bin Li¹, Weiyu Zhao¹, Xiao Luo¹, Xinfu Zhang¹, Chenglong Li², Chunxi Zeng¹, and Yizhou Dong^{1,3,4,5,6,7,*}

¹Division of Pharmaceutics and Pharmaceutical Chemistry, College of Pharmacy, The Ohio State University, Columbus, OH 43210, USA

²Department of Medicinal Chemistry, College of Pharmacy, University of Florida, Gainesville, FL 32610, USA

³Department of Biomedical Engineering, The Ohio State University, Columbus, OH 43210, USA

⁴The Center for Clinical and Translational Science, The Ohio State University, Columbus, OH 43210, USA

⁵The Comprehensive Cancer Center, The Ohio State University, Columbus, OH 43210, USA

⁶Dorothy M. Davis Heart & Lung Research Institute, The Ohio State University, Columbus, OH 43210, USA

⁷Department of Radiation Oncology, The Ohio State University, Columbus, OH 43210, USA

Abstract

Cpf1, a type-V CRISPR-Cas effector endonuclease, exhibits gene-editing activity in human cells through a single RNA-guided approach. Here, we report the design and assessment of an array of 42 types of engineered *Acidaminococcus* sp. Cpf1 (AsCpf1) CRISPR RNAs (crRNAs) and 5 types of AsCpf1 mRNAs, and show that the top-performing modified crRNA (cr3'5F, containing five 2'-fluoro ribose at the 3' termini) and AsCpf1 mRNA (full Ψ -modification) improved gene-cutting efficiency by, respectively, 127% and 177%, with respect to unmodified crRNA and plasmid-encoding AsCpf1. We also show that the combination of cr3'5F and Ψ -modified AsCpf1 or *Lachnospiraceae bacterium* Cpf1 (LbCpf1) mRNAs augmented gene-cutting efficiency by over 300% with respect to the same control, and discovered that 11 out of 16 crRNAs from Cpf1 orthologs enabled genome editing in the presence of AsCpf1. Engineered CRISPR-Cpf1 systems should facilitate a broad range of genome editing applications.

Users may view, print, copy, and download text and data-mine the content in such documents, for the purposes of academic research, subject always to the full Conditions of use: http://www.nature.com/authors/editorial_policies/license.html#terms

*Correspondence and requests for materials should be addressed to Y.D. dong.525@osu.edu.

Author contributions

B.L. and Y.D. conceived and designed the experiments; B.L., W.Z., X.L., C.Z. and X.Z. performed the experiments; B.L., W.Z., X.L., X.Z., C.L., C.Z. and Y.D. analyzed the data; B.L. and Y.D. wrote the paper with edits and comments from all authors.

Additional information

Supplementary information is available for this paper.

Competing interests

The authors declare no competing financial interests.

CRISPR (clustered regularly interspaced short palindromic repeats) and CRISPR-associated proteins are part of the adaptive immune system of bacteria and archaea^{1,2}. CRISPR-associated protein 9 (Cas9) induces double-stranded DNA breaks through complexation with two RNA molecules: CRISPR RNA (crRNA) and trans-activating crRNA (tracrRNA)^{3–5}. Recently, two Cpf1 (CRISPR from *Prevotella* and *Francisella* 1) proteins from *Acidaminococcus* sp. (AsCpf1) and *Lachnospiraceae* bacterium (LbCpf1) displayed comparable genome-editing capability to Cas9 (ref. ^{6–11}). Genome-wide analysis suggested that Cpf1 may cause fewer off-target cleavages in comparison to Cas9 (ref. ^{12,13}). To exert sequence-specific endonuclease activity, Cpf1 is functional through a single crRNA without an additional tracrRNA^{6,7,14}. This crRNA (43 nucleotides) consists of a 5'-handle (20 nucleotides) and a guide segment (23 nucleotides; Fig. 1a)⁶. The Cpf1 protein interacts with the pseudoknot structure formed by the 5'-handle of crRNA^{8,15}. A guide segment, composed of a seed region and 3' termini, possesses complementary binding sequences with the target DNA sequences. This protein-RNA complex recognizes a T-rich protospacer-adjacent motif and leads to a staggered DNA double-stranded break⁸.

To increase the genome editing efficiency of CRISPR-Cas systems, previous studies explored a variety of approaches^{16–19}. For example, chemical modifications of CRISPR-Cas9 led to enhanced activity in a number of human cells^{18,19}. Also, a chimeric single-guide RNA with three chemically modified nucleotides at both the 5' and 3' ends strongly improved Cas9-mediated genome editing in human primary T cells¹⁸. And the incorporation of chemically modified nucleotides in guide RNAs retained indel percentages in the CRISPR-Cas9 nuclease system (ref. ¹⁹). Furthermore, the structure of guide RNAs also plays important roles in gene cutting for the CRISPR-Cas9 system^{20,21}. Yet, to the best of our knowledge, the genome editing efficiency and off-target effects of engineered crRNAs and Cpf1 mRNAs have not been explored. Here, we report the systematic investigation of 42 chemically or structurally engineered crRNAs, and establish comprehensive structure-activity (genome editing efficiency) relationships. We also show that a ψ -modification is a favorable chemical alteration for AsCpf1 mRNA. Most importantly, we demonstrate that the combination of lead crRNA and ψ -mRNA significantly increased the gene-cutting efficiency by over 300% compared to the control group. This combination induced a more dramatic improvement of the gene-cutting efficiency when using LbCpf1. Also, we uncovered that the applicability of crRNAs from Cpf1 orthologs is different for AsCpf1 and LbCpf1; LbCpf1 is more conservative for the recognition of the loop structure at the 5'-handle. Our findings should enable a wide range of genome editing applications.

Engineering crRNAs to increase gene-cutting efficiency

To study the effects of engineered crRNAs on genome editing efficiency, we first used three types of chemically modified nucleotides: phosphorothioate (PS), 2'-O-Me-, and 2'-F-modifications²² (Supplementary Fig. 1). Modified crRNAs targeting the *DNMT1* locus⁶ were purified on denaturing polyacrylamide gels and validated by mass-spectrometry analysis (sequence-information and mass-spectrometry data are included in Supplementary Table 1). On the basis of previous findings for CRISPR-Cas9 (ref. ^{18,19}), we first treated human HEK293T cells with newly synthesized full-backbone PS-modified cr42PS (all 42 PO linkages were substituted with PS linkages) or cr5'&3'3F2PS (three 2'-F modifications

with two PS linkages at both 5' and 3' ends) and plasmid-encoding AsCpf1 complexed with Lipofectamine 3000 (Fig. 1b). Genome editing efficiency was quantified using the T7E1 assay, and normalized to the treatment of unmodified crRNA (crWT) and plasmid-encoding AsCpf1. Both crRNAs displayed dramatic reduction of gene-cutting capability (Fig. 2a), which suggests that chemical modifications of AsCpf1 crRNA are not of the same pattern as those of Cas9 guide RNAs (ref. ^{18,19}). Subsequently, we systematically examined the effects of chemical modifications at the 5'-handle [U(-20) to U(-1)], seed region [C(+1) to U(+8)], and 3'-termini [C(+9) to C(+23)] (Fig. 1a). We introduced five, ten, or twenty nucleotides with 2'-O-Me- or 2'-F-modifications at the 5'-handle of the crRNAs, which led to complete loss of activity. These data indicate that the 5'-handle is not favorable for chemical modifications. Then we installed different numbers of modifications at the seed region and synthesized crS3M3PS, crS3F, crS2F, and crS1F (Fig. 1b). Except crS1F, all other three crRNAs lowered the indel percentage (Fig. 2a), indicating that the seed region may tolerate slight modifications. Furthermore, at the 3'-termini, we incorporated five or ten nucleotides with 2'-O-Me- or 2'-F-modifications (Fig. 1b). Interestingly, 2'-F-modifications (cr3'5F and cr3'10F) were more favorable compared to their corresponding 2'-O-Me-modifications (cr3'5M and cr3'10M). cr3'5F improved the efficiency by 127% compared to the crWT ($p < 0.01$, Fig. 2a). We subsequently incorporated additional four PS linkage on the basis of five 2'-F-modifications (cr3'5F4PS), or replaced five 2'-F-modifications with five unlocked (cr3'5U) or locked (cr3'5L) nucleotides (Supplementary Fig. 1). cr3'5F4PS and cr3'5L reduced the potency, whereas cr3'5U completely lost function (Fig. 2a). These results imply that modifications at the 3'-termini are critical to gene editing activity. We next engineered the stem duplex [U(-15)-C(-11) and G(-6)-A(-2)] of the 5'-handle by split of the crRNAs (crSplit), or by deletion (crDel2-crDel8) or insertion of nucleotides (crIns4-crIns12) (Fig. 1c and Supplementary Table 1). Current crRNAs with split sequences and deletion of nucleotides exhibited no gene-cutting function. Insertion of four (crIns4') or six nucleotides (crIns6') retained the activity to some extent, which was dependent on the inserted base pairs (Fig. 2b).

Chemically modified AsCpf1 mRNA improved gene-cutting efficiency

We also investigated the effects of chemically modified AsCpf1 mRNAs on their gene-cutting efficiency. On the basis of previous results²³⁻²⁵, we designed pseudouridine- (ψ), N¹-methylpseudouridine- ($\text{me}^1\psi$), and 5-methoxyuridine- (5moU) modified AsCpf1 mRNA, which were produced via *in vitro* transcription (Fig. 3a and Supplementary Fig. 2). We then treated HEK293T cells with crWT and modified AsCpf1 mRNA. Unmodified AsCpf1 (WT AsCpf1 mRNA), ψ - and $\text{me}^1\psi$ -modified AsCpf1 mRNA showed higher activity (154%, 177% and 168%, respectively) than plasmid-encoding AsCpf1, whereas the activity of the 5moU-modified AsCpf1 mRNA resembled that of the AsCpf1 plasmid (Fig. 3b). In addition, it has been recently reported that plasmid-encoding AsCpf1 with a S1228A mutation may improve genome editing efficiency⁸. We therefore constructed AsCpf1 mRNA with both a S1228A mutation and ψ -modification (S1228A & ψ mRNA), which displayed comparable activity to the original ψ -modified mRNA (Fig. 3b).

Applicability of crRNAs from Cpf1 family orthologs

Inspired by crRNAs from other 15 members of the Cpf1 family⁶ (see Methods for a list of the Cpf1-family orthologs), we engineered the loop [U(−10)–C(−9)–U(−8)–U(−7)] of the 5′-handle by substituting the loop of wild-type AsCpf1 crRNA (crWT here is defined as AscrRNA to distinguish it from other crRNAs from the Cpf1-protein family) with those from crRNAs of other Cpf1-family orthologs to investigate the effects of the loop on gene-cutting activity (Fig. 3c and Supplementary Table 1; crRNAs were termed by combining the initials of the Cpf1 protein and crRNA; crRNAs of PbCpf1, PeCpf1 and LiCpf1 share the same loop, UUUU; crRNAs of Lb2Cpf1, PcCpf1 and PmCpf1 share the same loop, UAUU). Interestingly, co-delivery of ψ -modified AsCpf1 mRNA and a panel of loop-engineered crRNAs revealed that crRNAs with four-membered loop including FncrRNA, Pb/Pe/LicrRNA and Lb2/Pc/PmcrRNA induced evident cleavage on the target locus. Cytosine at position −9 tolerated nucleotide changes, with the order of potency being adenine (A) > cytosine (C) > guanine (G) > uridine (U). In addition, crRNAs with three-membered (EecrRNA) or five-membered loop (MbcrRNA and LbcrRNA) reduced the activity (Fig. 3d). Also, when the uridine at position −10 was replaced by G (Lb3crRNA and BpcrRNA) or A (SscrRNA) on the five-membered loop, crRNAs completely lost function (Fig. 3d). These results suggest that U(−10) is a critical position for gene editing, consistent with the findings from the crystal structure of AsCpf1 (ref. ⁸). After sequence alignment, we observed that the whole sequence of AscrRNA was fully matched with that of crRNAs from FnCpf1, Pb/LiCpf1 and Lb2/Pc/Pm Cpf1 except PeCpf1 (according to the crystal structure of AsCpf1–crRNA–target DNA complex, uridine at position −20 is not involved in the complexation; hence, U(−20) was not included for alignment)⁸. In other words, AsCpf1 is able to complex with these crRNAs, attaining effective genome editing. These data show cross-complexation between AsCpf1 and crRNAs in the Cpf1 family, expand the applicability of crRNAs, and are conducive to better understanding of the CRISPR-Cpf1 system and its potential biological and therapeutic applications.

Collectively, these results revealed the following structure–activity relationship and design criteria for crRNAs and AsCpf1 mRNA: (i) PS linkage replacement generally hampers genome editing efficiency; (ii) stem duplex in the 5′-handle is not a preferred region for structural engineering, as it does not tolerate the split, deletion or insertion of nucleotides; (iii) the seed region is sensitive to chemical modifications and can only incorporate minor modifications; (iv) the 3′ termini is favorable for certain chemical modifications, such as 2′-F modifications; (v) a combination of chemical modifications at both the 5′ and 3′ termini of crRNA reduce the activity to some extent; (vi) ψ - and me¹ ψ - modified AsCpf1 mRNAs are more active than plasmid-encoding AsCpf1 for gene cutting; (vii) AsCpf1 is capable of effectively complexing with the majority of crRNAs from the Cpf1 family to exert gene-cleavage activity.

Maximizing efficiency through the combination of crRNA and AsCpf1 mRNA

To study the combination effects of chemically modified crRNA and AsCpf1 mRNAs, we treated HEK293T cells with the top-performing modified crRNA and AsCpf1 mRNA (cr3'5F and ψ -modified AsCpf1 mRNA). Strikingly, this combination significantly enhanced the gene cutting efficiency over 250% compared to the treatment of crWT and plasmid encoding AsCpf1 ($p < 0.001$, Fig. 4a). We then analyzed the interaction among crRNA, AsCpf1 protein and target DNA. The crystal structure of the AsCpf1 complex indicates that 2'-OH of ribose on crRNA at the following positions plays a trivial role: -16, -15, -12, -11, -9, -8, -7, -5, -4, -3, -2, +1, +6, +8, +9, +10, +11, +12, +17, +19, +20, +21, +22 and +23 (ref. ⁸). On the basis of this analysis, we designed additional three type of crRNAs: crI21F, crI16M5F and crI16M10PS5F (Supplementary Fig. 6 and Supplementary Table 1), by introducing interspersed modifications at the ribose units without interaction with AsCpf1 and target DNA and by avoiding modifications at the seed region (+1 to +8). Interestingly, crI21F was comparable to cr3'5F, whereas crI16M5F and crI16M10PS5F dramatically weakened the activity of crRNA (Supplementary Fig. 6). Since cr3'5F possesses less modifications than crI21F, we further investigated the applicability of cr3'5F in Hep3B (a human hepatoma cell line) and U87 cells (a human glioblastoma cell line). Consistent with the results in HEK293T cells, combination of cr3'5F and ψ -modified AsCpf1 mRNA improved gene-cutting 329% for Hep3B ($p < 0.001$) and 293% for U87 cells ($p < 0.001$) (Fig. 4a) compared to the treatment of crWT and plasmid-encoding AsCpf1. In addition to *DNMT1* locus, we also examined gene-cutting efficiency for another two target genes, *AAVS1* and *FANCF-2* (ref. ^{12,13}). Regarding *AAVS1* locus (Fig. 4b), the order of potency in HEK293T cells is cr3'5F + ψ -modified AsCpf1 mRNA > crWT + ψ -modified AsCpf1 mRNA > cr3'5F + AsCpf1 plasmid > crWT + AsCpf1 plasmid. cr3'5F + ψ -modified AsCpf1 mRNA is 2.77-fold more efficient than crWT + AsCpf1 plasmid ($p < 0.001$). The enhanced efficiency was also observed in Hep3B and U87 cells (257% and 394% increase, respectively; Fig. 4b). For the *FANCF-2* locus, the trend is consistent with that for *AAVS1* locus (Fig. 4c). These results demonstrate the broad applicability of the chemically modified crRNAs and Cpf1 mRNAs.

Gene-cutting efficiency and applicability of the engineered crRNAs

As with AsCpf1, LbCpf1 is another important endonuclease in the Cpf1 family that also displayed genome editing ability in human cells^{6,9}. To investigate whether our strategy is applicable to LbCpf1, we used similar chemical modifications for LbcrRNA as well as its corresponding LbCpf1 mRNA. The combination of LbCpf1 crWT + LbCpf1 plasmid led to no detectable gene cutting (Fig. 5a), whereas the combination of LbCpf1 cr3'5F (Lbcr3'5F) + ψ -modified LbCpf1 mRNA induced remarkable gene-cutting activity in all three cell lines tested. Moreover, LbCpf1 cr3'5F was more efficient than LbCpf1 crWT ($p < 0.01$). These results further prove the concept of our chemical modifications to advance CRISPR-Cpf1-mediated genome editing. Lastly, we investigated the applicability of crRNAs for LbCpf1. We treated cells with loop-engineered crRNAs (Fig. 3c) in the presence of ψ -modified LbCpf1 mRNA. Except for its own LbcrRNA, LbCpf1 only led to reduced indel with

Lb2/Pc/PmcrRNA, AscrRNA, and MbcRNA, and no activity for other crRNAs (Fig. 5b), indicating different crRNA applicability between AsCpf1 and LbCpf1.

On-target and off-target effects of AsCpf1 and LbCpf1

Potential off-target effects are one of the major concerns for CRISPR-mediated gene editing and may limit its applications^{16,17,26,27}. To study off-target effects of chemically modified crRNAs and Cpf1 mRNAs, we selected the top four off-target sites defined previously^{12, 13} and performed the T7E1 assay at each genomic site. We observed no detectable indels in both treatment groups (cr3'5F + ψ -modified AsCpf1 mRNA and crWT + AsCpf1 plasmid; Supplementary Fig. 5). To further characterize on-target and off-target effects²⁸, we then conducted targeted deep sequencing using three biological samples per treatment group from the HEK293T cells used in Fig. 4a, and used CRISPResso²⁹ to analyze the deep sequencing data. Similar to the results from T7E1 assays, combination of cr3'5F and ψ -modified AsCpf1 mRNA induced significantly higher gene cutting (310%, $p < 0.001$) compared to the treatment of crWT and plasmid-encoding AsCpf1 (Fig. 6a). Compatible with the results in the literature¹³, Cpf1 mainly induced certain length of deletions (>95%) close to the predicted cleavage site rather than insertions (Fig 6b,c, Table 1, Supplementary Fig.7a,b and 8a,b and Supplementary Table 6). Also, we noticed that AsCpf1 showed low off-target effects at selected genomic sites, and that no significant difference was observed for off-target gene cutting between cr3'5F + ψ -modified AsCpf1 mRNA and crWT + AsCpf1 plasmid (Fig. 6a).

In addition to AsCpf1, we examined the on-target and off-target effects of LbCpf1 with three biological samples from the HEK293T cells used in Fig. 5a. As shown in Fig. 7a, Lbcr3'5F + ψ -modified LbCpf1 mRNA caused 46.7% indel compared to the minimal effects of LbcrWT + LbCpf1 plasmid. The indel pattern for LbCpf1 was consistent with that of AsCpf1, whereas LbCpf1 led to larger fragment deletion than AsCpf1 (Fig. 7b, Table 1, Supplementary Fig. 7c, 8c and Supplementary Table 6). Regarding off-target effects, LbCpf1 exhibited much higher mutagenesis at the position of off-target 4 (OT4) than AsCpf1 under the same condition (Fig. 7a). Similarly, Lbcr3'5F + ψ -modified LbCpf1 mRNA showed comparable off-target effects to that of LbcrWT + LbCpf1 plasmid (Fig. 7a). Taken together, the combination of engineered crRNA and Cpf1 mRNA enables the effective enhancement of genome editing efficiency without increasing off-target effects.

Discussion

Genome editing efficiency and off-targets effects are major challenges for the broad application of CRISPR systems^{1, 26}. To address these issues for the CRISPR-Cpf1 system, we designed an array of crRNA variants, including chemically modified crRNAs and structurally engineered crRNAs, and elucidated structure–activity relationships of crRNAs for improved genome editing activity. In contrast to CRISPR-Cas9 (ref. ^{18,19}), neither phosphorothioate substitutions nor dual-modification at both sides of crRNA enhanced genome editing efficiency. Deletion or insertion of nucleotides at the stem duplex reduced the activity. Moreover, modifications at the stem duplex of the 5'-handle or at the seed region severely hampered cleavage activity. Importantly, 2'-F modification at the 3' termini

(cr3'5F) exhibited higher potency compared with unmodified crRNA (crWT). In addition, the crystal structure of the Cpf1-crRNA-target DNA complex provides useful guidance to design new crRNAs. Regarding chemically modified AsCpf1 mRNA, pseudouridine (ψ) and N^1 -methylpseudouridine ($me^1\psi$) were favorable modifications compared to the mRNA with unmodified nucleotides. Notably, the combination of cr3'5F and AsCpf1 ψ -mRNA improved gene-cutting efficiency over 3-fold compared to crWT and plasmid-encoding AsCpf1. This phenomenon was confirmed in three cell lines and three target genome sites. In addition, the crystal structure of the Cpf1 complex provided useful guidance for the design of new crRNAs. Regarding LbCpf1, we observed 47% indel using the same combination of cr3'5F and LbCpf1 ψ -mRNA, and no detectable gene cutting with the treatment of crWT and plasmid-encoding LbCpf1. These results demonstrate the broad applicability of this engineering strategy for Cpf1-family-mediated genome editing. Importantly, we also explored the cross complexation of AsCpf1 and LbCpf1 with crRNAs from other members of the Cpf1 family. We elucidated that in addition to their own crRNAs, AsCpf1 was able to induce gene editing in the presence of other ten Cpf1-family crRNAs (FnCpf1, PbCpf1, LiCpf1, Lb2Cpf1, PcCpf1, PmCpf1, EeCpf1, MbCpf1, PdCpf1, LbcCpf1), whereas LbCpf1 only effectively complexed with Lb2/Pc/Pmc crRNA. These results suggest that LbCpf1 is more conservative than AsCpf1 when complexing with crRNAs. Targeted deep-sequencing data confirmed that the combination of cr3'5F and ψ -modified Cpf1 mRNA significantly enhanced the genome editing efficiency without increasing the level of off-target effects. Overall, our results offer new insights into CRISPR-Cpf1 systems and provide a set of useful design criteria for maximizing genome editing efficiency.

Methods

List of the Cpf1-family orthologs used

Acidaminococcus sp. *BV3L6* Cpf1 (AsCpf1); *Francisella tularensis* subsp. *Novicida* *U112* Cpf1 (FnCpf1); *Lachnospiraceae* bacterium *MC2017* Cpf1 (Lb3Cpf1); *Butyrivibrio proteoclasticus* Cpf1 (BpCpf1); *Parcubacteria* bacterium *GWC2011_GWC2_44_17* Cpf1 (PbCpf1); *Peregrinibacteria* bacterium *GW2011_GWA_33_10* Cpf1 (PcCpf1); *Leptospira inadai* Cpf1 (LiCpf1); *Smithella* sp. *SC_K08D17* Cpf1 (SsCpf1); *Lachnospiraceae* bacterium *MA2020* Cpf1 (Lb2Cpf1); *Porphyromonas crevioricanis* Cpf1 (PcCpf1); *Porphyromonas macacae* Cpf1 (PmCpf1); *Candidatus Methanoplasma termitum* Cpf1 (CmCpf1); *Eubacterium eligens* Cpf1 (EeCpf1); *Moraxella bovoculi* 237 Cpf1 (MbCpf1); *Prevotella disiens* Cpf1 (PdCpf1); *Lachnospiraceae* bacterium *ND2006* Cpf1 (LbCpf1).

Synthesis of crRNAs

The sequence of unmodified crRNA targeting *DNMT1-3* locus:

AsCpf1 crRNA: 5'-
UAAUUUCUACUCUUGUAGAUCUGAUGGUCCAUGUCUGUUACUC-3';
LbCpf1 crRNA: 5'-
AAUUUCUACUAAGUGUAGAUCUGAUGGUCCAUGUCUGUUACUC-3'.

The sequence of unmodified crRNA targeting *AAVS1* locus:

AsCpf1 crRNA: 5'-
UAAUUUCUACUCUUGUAGAUCUUACGAUGGAGCCAGAGAGGAU-3'.

The sequence of unmodified crRNA targeting *FANCF-2* locus:

AsCpf1 crRNA: 5'-
UAAUUUCUACUCUUGUAGAUGUCGGCAUGGCCCCAUUCGCACG-3'.

Unmodified crRNA (crWT) and all other crRNA variants including chemically modified crRNAs, stem- and loop-engineered crRNAs were synthesized using an automated solid-phase DNA/RNA synthesizer. Chemically modified crRNAs consisted of partial or total chemically modified nucleotides including phosphate linkage (PS), 2'-O-Me, 2'-F modified, unlocked, locked nucleotides as well as their combinations (Fig. 1b and Supplementary Table 1). Stem-engineered crRNAs were designed by deleting certain number of Watson-Crick base pairs from the stem duplex or inserting additional paired or unpaired bases into the stem duplex of crWT (Fig. 1c and Supplementary Table 1). Split crRNA (crSplit) was generated by incubating equimolar of relevant RNA sequences listed in Supplementary Table 1 in TE buffer at 95 °C for 30 s, followed by gradient cooling (95-25 °C ramping at 0.1 °C/s). Loop-engineered crRNAs were designed by employing the loops from other Cpf1 orthologs (Fig. 3c and Supplementary Table 1). All crRNAs were purified on denaturing polyacrylamide gels and verified by mass spectrometry (Supplementary Table 1).

Production of Cpf1 mRNAs

Anti-Reverse Cap Analog (ARCA) capped and polyadenylated AsCpf1 mRNA and LbCpf1 mRNA transcripts were purchased from TriLink BioTechnologies (San Diego, CA, USA). For modified AsCpf1 mRNAs, uridines were fully substituted with pseudouridine (ψ), *N*¹-methylpseudouridine (me¹ ψ), or 5-methoxyuridine (5moU) (Fig. 2a). For S1228A & ψ modified AsCpf1 mRNA, the nucleotides that encode serine1228 (S1228) was substituted with that of alanine (A), and uridines were fully substituted with pseudouridine (ψ). For modified LbCpf1 mRNA, uridines were fully substituted with pseudouridine (ψ). These mRNAs were subjected to DNase and phosphatase treatment and silica-gel membrane spin column purification for further use. All mRNAs were verified by analytical agarose gels (Supplementary Fig. 2).

Co-delivery of crRNA and Cpf1

HEK-293T cells were cultured in Dulbecco's Modified Eagle's Medium without sodium pyruvate (Corning Incorporated). Hep3B cells were cultured in Eagle's Minimum Essential Medium with sodium pyruvate (Corning Incorporated). U87 cells were cultured in Dulbecco's Modified Eagle's Medium with sodium pyruvate, high glucose (Thermo Fisher Scientific). All medium supplement with 10% FBS, and all cell lines used in this study are from the American Type Culture Collection (ATCC) and are maintained at 37 °C with 5% CO₂. After overnight incubation (approximate 60~70% confluence), cells seeded on 24-well plates at an initial density of 100,000~150,000 cells per well were treated with either Cpf1 expression plasmid (provided by Dr. Feng Zhang) or Cpf1 mRNA (500 ng for *DNMT1-3* and *AAVS1* locus, and 1500 ng for *FANCF-2* locus). At the same time, crRNAs (38 pmol for *DNMT1-3* and *AAVS1* locus, and 114 pmol for *FANCF-2* locus) were added to each

well. All components were formulated with Lipofectamine 3000 (Life Technologies) in Opti-MEM I reduced serum medium (Life Technologies) following the manufacturer's recommended protocol.

Genomic DNA purification and PCR amplification

Two days post-treatment, cells were washed with PBS. The genomic DNA (gDNA) was then extracted and purified with a DNeasy Blood & Tissue Kit (QIAGEN) following the manufacturer's instructions. Concentrations of gDNA were determined on a Nanodrop 2000. Genomic regions flanking the on-target as well as previously predicted off-target sites off-target sites for T7E1 assay were amplified using 100 ng of purified gDNA template, Q5 high-fidelity DNA polymerase (New England Biolabs) and specific primers (Integrated DNA Technologies, Supplementary Table 2) on a T100 thermal cycler (Bio-Rad).

T7E1 Cleavage assay

The PCR products generated using Q5 high-fidelity DNA polymerase were hybridized in NEBuffer2 (50 mM NaCl, 10 mM Tris-HCl, 10 mM MgCl₂, 1 mM DTT, pH 7.9) (New England Biolabs) by heating to 98 °C for 10 minutes, followed by a 2 °C/s ramp down to 85 °C, 1 min at 85 °C, and a 0.1 °C/s ramp down to 25 °C on a T100 thermal cycler (Bio-Rad). Subsequently, the annealed samples were subjected to T7 Endonuclease I (New England Biolabs) digestion for 30 min, separated by a 2% agarose gel and quantified on ChemiDoc XRS (Bio-Rad) using Quantity One software. The mutation frequency (indel, %) was calculated with the following formula: $100 \times (1 - (1 - \text{fraction cleaved})^{1/2})$.

MiSeq library preparation and targeted deep sequencing

To further characterize on-target and off-target effects, genomic segment (200~300 bp, Supplementary Note 1) spanning the sites of interest were first amplified using sequencing primers with overhang adapter sequences (Supplementary Table 3) in the first round of PCR for 25 cycles. After purification, the second limited-cycle PCR amplification (10 cycles) was performed using the Nextera Index Kit (Illumina) to attach multiplexing indices and Illumina P5/P7 sequencing adapters (Supplementary Table 4, 5) to the first round PCR product. Next, libraries were normalized and pooled, and subjected to 2 × 300 paired-end sequencing on an Illumina MiSeq system. The raw deep sequencing data from MiSeq were analyzed by a bioinformatic tool, CRISPResso²⁹, with specific parameters (Supplementary Note 2).

Data availability

The authors declare that all data supporting the findings of this study are available within the paper and its Supplementary Information. Sequence reads used in this study are available at the NCBI Sequence Read Archive (SRP093361). Source data for the sequence analysis are available in figshare, with the identifier doi: 10.6084/m9.figshare.4696288 (ref. ³⁰).

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We acknowledge Dr. Feng Zhang and Mr. Bernd Zetsche for providing Cpf1 plasmids and technical assistance. This work was supported by the Early Career Investigator Award from the Bayer Hemophilia Awards Program, Research Awards from the National PKU Alliance, New Investigator Grant from the AAPS Foundation, R01HL136652 from the National Heart, Lung, and Blood Institute as well as the start-up fund from the College of Pharmacy at The Ohio State University.

References

1. Bosley KS, et al. CRISPR germline engineering--the community speaks. *Nat Biotechnol.* 2015; 33:478–486. [PubMed: 25965754]
2. Makarova KS, et al. Evolution and classification of the CRISPR-Cas systems. *Nat Rev Microbiol.* 2011; 9:467–477. [PubMed: 21552286]
3. Jinek M, et al. A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science.* 2012; 337:816–821. [PubMed: 22745249]
4. Cong L, et al. Multiplex genome engineering using CRISPR/Cas systems. *Science.* 2013; 339:819–823. [PubMed: 23287718]
5. Mali P, et al. RNA-guided human genome engineering via Cas9. *Science.* 2013; 339:823–826. [PubMed: 23287722]
6. Zetsche B, et al. Cpf1 Is a Single RNA-Guided Endonuclease of a Class 2 CRISPR-Cas System. *Cell.* 2015; 163:759–771. [PubMed: 26422227]
7. Fonfara I, Richter H, Bratovic M, Le Rhun A, Charpentier E. The CRISPR-associated DNA-cleaving enzyme Cpf1 also processes precursor CRISPR RNA. *Nature.* 2016; 532:517–521. [PubMed: 27096362]
8. Yamano T, et al. Crystal Structure of Cpf1 in Complex with Guide RNA and Target DNA. *Cell.* 2016
9. Dong D, et al. The crystal structure of Cpf1 in complex with CRISPR RNA. *Nature.* 2016; 532:522–526. [PubMed: 27096363]
10. Hur JK, et al. Targeted mutagenesis in mice by electroporation of Cpf1 ribonucleoproteins. *Nat Biotechnol.* 2016; 34:807–808. [PubMed: 27272385]
11. Kim Y, et al. Generation of knockout mice by Cpf1-mediated gene targeting. *Nat Biotechnol.* 2016; 34:808–810. [PubMed: 27272387]
12. Kleinstiver BP, et al. Genome-wide specificities of CRISPR-Cas Cpf1 nucleases in human cells. *Nat Biotechnol.* 2016; 34:869–874. [PubMed: 27347757]
13. Kim D, et al. Genome-wide analysis reveals specificities of Cpf1 endonucleases in human cells. *Nat Biotechnol.* 2016; 34:863–868. [PubMed: 27272384]
14. Zetsche B, et al. Multiplex gene editing by CRISPR-Cpf1 using a single crRNA array. *Nat Biotechnol.* 2017; 35:31–34. [PubMed: 27918548]
15. Gao P, Yang H, Rajashankar KR, Huang Z, Patel DJ. Type V CRISPR-Cas Cpf1 endonuclease employs a unique mechanism for crRNA-mediated target DNA recognition. *Cell Res.* 2016; 26:901–913. [PubMed: 27444870]
16. Slaymaker IM, et al. Rationally engineered Cas9 nucleases with improved specificity. *Science.* 2016; 351:84–88. [PubMed: 26628643]
17. Kleinstiver BP, et al. High-fidelity CRISPR-Cas9 nucleases with no detectable genome-wide off-target effects. *Nature.* 2016; 529:490–+. [PubMed: 26735016]
18. Hendel A, et al. Chemically modified guide RNAs enhance CRISPR-Cas genome editing in human primary cells. *Nat Biotechnol.* 2015; 33:985–989. [PubMed: 26121415]
19. Rahdar M, et al. Synthetic CRISPR RNA-Cas9-guided genome editing in human cells. *Proc Natl Acad Sci U S A.* 2015; 112:E7110–7117. [PubMed: 26589814]
20. Chen B, et al. Dynamic imaging of genomic loci in living human cells by an optimized CRISPR/Cas system. *Cell.* 2013; 155:1479–1491. [PubMed: 24360272]
21. Dang Y, et al. Optimizing sgRNA structure to improve CRISPR-Cas9 knockout efficiency. *Genome Biol.* 2015; 16:280. [PubMed: 26671237]

22. Watts JK, Deleavey GF, Damha MJ. Chemically modified siRNA: tools and applications. *Drug Discov Today*. 2008; 13:842–855. [PubMed: 18614389]
23. Kariko K, Muramatsu H, Keller JM, Weissman D. Increased erythropoiesis in mice injected with submicrogram quantities of pseudouridine-containing mRNA encoding erythropoietin. *Mol Ther*. 2012; 20:948–953. [PubMed: 22334017]
24. Andries O, et al. N(1)-methylpseudouridine-incorporated mRNA outperforms pseudouridine-incorporated mRNA by providing enhanced protein expression and reduced immunogenicity in mammalian cell lines and mice. *J Control Release*. 2015; 217:337–344. [PubMed: 26342664]
25. Li B, Luo X, Dong Y. Effects of Chemically Modified Messenger RNA on Protein Expression. *Bioconjug Chem*. 2016
26. Bolukbasi MF, Gupta A, Wolfe SA. Creating and evaluating accurate CRISPR-Cas9 scalpels for genomic surgery. *Nat Methods*. 2016; 13:41–50. [PubMed: 26716561]
27. Kim D, et al. Digenome-seq: genome-wide profiling of CRISPR-Cas9 off-target effects in human cells. *Nat Methods*. 2015; 12:237–243. 231–243. [PubMed: 25664545]
28. Lin YN, et al. CRISPR/Cas9 systems have off-target activity with insertions or deletions between target DNA and guide RNA sequences. *Nucleic Acids Research*. 2014; 42:7473–7485. [PubMed: 24838573]
29. Pinello L, et al. Analyzing CRISPR genome-editing experiments with CRISPResso. *Nat Biotechnol*. 2016; 34:695–697. [PubMed: 27404874]
30. Li, B., et al. Dataset for Engineering CRISPR-Cpf1 crRNAs and mRNAs to maximize genome editing efficiency. figshare. 2017. <http://dx.doi.org/10.6084/m9.figshare.4696288>

Figure 1. Design of chemically modified and structurally engineered crRNAs

a, Schematic of the AsCpf1-crRNA-target DNA complex. crRNA is composed of a 5'-handle (in a pseudoknot structure) and a guide segment (consists of a seed region and 3'-termini). The dotted line denotes the cleavage sites. **b**, Panel of the chemically modified crRNAs tested in this study. Unmodified nucleotides are shown in black. Full-length PS modifications are in pink. 2'-O-methyl modifications are highlighted in blue. 2'-F modifications are in red. 2'-O-methyl combined with PS modifications are in light blue. 2'-F combined with PS modifications are in green. **c**, Stem engineering of crRNAs. The length

of the stem duplex was altered by deletion or insertion (in light orange) of nucleotides. For crSplit, the pseudoknot-like structure was formed by hybridization of two arms of the direct repeat.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Figure 2. Gene-cutting efficiency of chemically modified and stem-engineered crRNAs

a. Gene-cutting efficiency of chemically modified crRNAs. Unlock nucleotides are shown in purple. Locked nucleotides are shown in orange. Other colors in sequence are the same as that in Fig. 1b. **b.** Gene-cutting efficiency of stem-engineered crRNAs. “T” in the left panel denotes the split site. The dash denotes deleted nucleotide. The lowercase letter denotes inserted nucleotides. Gene-cutting efficiencies in **a** and **b** were determined by the T7E1 cleavage assay, and normalized to that of the treatment with crWT and plasmid-encoding AsCpf1. N.D., Not detectable. Data were expressed as the mean $\pm$ s.d. from three biological replicates (**, $P < 0.01$, cr3'5F versus to crWT; t test, double-tailed).

Figure 3. Gene-cutting efficiency of chemically modified AsCpf1 mRNAs and loop-engineered crRNAs

a, Schematic diagram depicting chemical modifications applied to CRISPR-Cpf1 mRNAs. Modified Cpf1 mRNAs were generated by fully substituting natural uridines (blue box) of unmodified mRNA (WT mRNA) with pseudouridine (ψ), N¹-methylpseudouridine ($me^1\psi$), or 5-methoxyuridine (5moU). **b**, Gene-cutting efficiency of chemically modified AsCpf1 mRNAs. For S1228A & ψ modified AsCpf1 mRNA, the nucleotides that encode serine1228 (S1228) were substituted with that of alanine (A), and uridines were fully substituted with pseudouridine (ψ). Indel percentage was determined by the T7E1 cleavage assay, and normalized to that of the treatment with crWT and plasmid-encoding AsCpf1. N.D., Not detectable. (***, $P < 0.001$, ψ -modified AsCpf1 mRNA versus to AsCpf1 plasmid; t test, double-tailed). **c**, Loop engineering of crRNAs. Nucleotides of the loop were altered according to crRNAs from other 15 Cpf1-family orthologs. crRNAs were termed by

combining the initials of the Cpf1 protein followed by crRNA; crRNAs of PbCpf1, PeCpf1 and LiCpf1 share the same loop, UUUU; crRNAs of Lb2Cpf1, PcCpf1 and PmCpf1 share the same loop, UAUU. Differences in nucleotides between AscrRNA and other crRNAs in the loop are highlighted in red. **d**, Gene-cutting efficiency of loop-engineered crRNAs in the presence of Ψ -modified AsCpf1 mRNA. Red nucleotides denote sequence differences in the loop between AscrRNA and other crRNAs. Gene-cutting efficiency was determined by the T7E1 cleavage assay, and normalized to that of the treatment with AscrRNA and Ψ -modified AsCpf1 mRNA. N.D., Not detectable. Data were expressed as the mean $\pm$ s.d. from three biological replicates (*, $P < 0.05$, Lb2crRNA versus to AscrRNA; t test, double-tailed).

Figure 4. Maximizing genome editing efficiency through the combination of chemically modified crRNA and AsCpf1 mRNA

a, AsCpf1-mediated gene-cutting efficiency for the human *DNMT1-3* in HEK293T, Hep3B and U87 cells. **b**, AsCpf1-mediated gene-cutting efficiency for the human *AAVS1* in HEK293T, Hep3B and U87 cells. **c**, AsCpf1-mediated gene-cutting efficiency for the human *FANCF-2* in HEK293T cells. Indel percentage at each locus was determined using the T7E1 assay, and expressed as the mean $\pm$ s.d. from three biological replicates (*, $P < 0.05$; **, $P < 0.01$; ***, $P < 0.001$; t test, double-tailed). Gel images are in Supplementary Fig. 3.

Figure 5. Engineered LbCpf1 crRNAs and mRNAs, and crRNA applicability
a, LbCpf1-mediated gene-cutting efficiency for the human *DNMT1-3* in HEK293T, Hep3B and U87 cells. Indel percentage was determined using the T7E1 assay, and expressed as the mean $\pm$ s.d. from three biological replicates (**, $P < 0.01$; t test, double-tailed). Gel images are in Supplementary Fig. 4. **b**, Gene-cutting efficiency of loop-engineered crRNAs in the presence of ψ -modified LbCpf1 mRNA. Red nucleotides denote sequence difference in the loop between LbcrRNA and other crRNAs. Gene-cutting efficiency was determined by the T7E1 cleavage assay, normalized to that of the treatment with LbcrRNA and ψ -modified LbCpf1 mRNA, and expressed as the mean $\pm$ s.d. from three biological replicates. Gel images are in Supplementary Fig. 4.

Figure 6. Targeted-deep-sequencing analysis of on-target and off-target gene cutting for chemically modified crRNA and AsCpf1 mRNA

a, Indel percentage at on-target and predicted top four off-target sites analyzed by deep-sequencing data. Indel percentage was plotted as the mean $\pm$ s.d. from three biological replicates (***, $P < 0.001$; N.S.; Not significant; t test, double-tailed). **b**, Plot of representative indel-size distribution and position distribution of all reads for crWT + AsCpf1 plasmid. **c**, Representative plots of indel-size distribution and position distribution of all reads for cr3'5F+ ψ -modified AsCpf1 mRNA. Targeted-deep-sequencing analysis for biological replicates 2 and 3 are in Supplementary Fig. 7 and 8.

Figure 7. Targeted-deep-sequencing analysis of on-target and off-target gene cutting for chemically modified Lbcr3

'5F and LbCpf1 mRNA.

a. Indel percentage at on-target and predicted top four off-target sites analyzed by deep-sequencing data. Indel percentage was plotted as the mean $\pm$ s.d. from three biological replicates (***, $P < 0.001$; N.S.; Not significant; t test, double-tailed). **b.** Representative plots of indel-size distribution and position distribution of all reads for Lbcr3'5F + ψ -modified LbCpf1 mRNA. Targeted-deep-sequencing analysis for biological replicates 2 and 3 are in Supplementary Fig. 7 and 8.

Table 1

Representative top ten high-frequent on-target mutagenesis aligned to the target site of *DNMT1-3*, induced by the CRISPR-Cpf1 system.

Indel pattern	On-target mutagenesis induced by the combination of AscrWT and AsCpf1 plasmid	Read (%)
WT	CAGTACGTTAATGTTCTGATGGTCCATGTCTGTACTCGCCTGTCAAGTGGCGTGACACCGGG	73.3
D6	CAGTACGTTAATGTTCTGATGGTCCATGT-----CTCGCCTGTCAAGTGGCGTGACACCGGG	2.2
D8	CAGTACGTTAATGTTCTGATGGTCCATG-----TCGCCTGTCAAGTGGCGTGACACCGGG	1.2
D7	CAGTACGTTAATGTTCTGATGGTCCA-----TACTCGCCTGTCAAGTGGCGTGACACCGGG	1.2
D6	CAGTACGTTAATGTTCTGATGGTCCA-----TTACTCGCCTGTCAAGTGGCGTGACACCGGG	1.0
D4	CAGTACGTTAATGTTCTGATGGTCCA---TGTTACTCGCCTGTCAAGTGGCGTGACACCGGG	0.8
D8	CAGTACGTTAATGTTCTGATGGTCCAT-----CTCGCCTGTCAAGTGGCGTGACACCGGG	0.8
D5	CAGTACGTTAATGTTCTGATGGTCCATG----TACTCGCCTGTCAAGTGGCGTGACACCGGG	0.7
D7	CAGTACGTTAATGTTCTGATGGTCCATGTCT-----GCCTGTCAAGTGGCGTGACACCGGG	0.6
D16	CAGTACGTTAATGTTCTGATGGT-----CCTGTCAAGTGGCGTGACACCGGG	0.6
D6	CAGTACGTTAATGTTCTGATGGTCCATG-----ACTCGCCTGTCAAGTGGCGTGACACCGGG	0.6

Indel pattern	On-target mutagenesis induced by the combination of Ascr3'5F and AsCpf1 mRNA (ψ)	Read (%)
WT	CAGTACGTTAATGTTCTGATGGTCCATGTCTGTACTCGCCTGTCAAGTGGCGTGACACCGGG	34.0
D6	CAGTACGTTAATGTTCTGATGGTCCATGT-----CTCGCCTGTCAAGTGGCGTGACACCGGG	6.7
D16	CAGTACGTTAATGTTCTGATGGT-----CCTGTCAAGTGGCGTGACACCGGG	4.4
D5	CAGTACGTTAATGTTCTGATGGTCCATG----TACTCGCCTGTCAAGTGGCGTGACACCGGG	2.8
D8	CAGTACGTTAATGTTCTGATGGTCCAT-----CTCGCCTGTCAAGTGGCGTGACACCGGG	2.5
D8	CAGTACGTTAATGTTCTGATGGTCCATG-----TCGCCTGTCAAGTGGCGTGACACCGGG	2.4
D7	CAGTACGTTAATGTTCTGATGGTCCA-----TACTCGCCTGTCAAGTGGCGTGACACCGGG	2.3
D6	CAGTACGTTAATGTTCTGATGGTCCATG-----ACTCGCCTGTCAAGTGGCGTGACACCGGG	2.1
D7	CAGTACGTTAATGTTCTGATGGTCCATG-----CTCGCCTGTCAAGTGGCGTGACACCGGG	1.8
D6	CAGTACGTTAATGTTCTGATGGTCCA-----TTACTCGCCTGTCAAGTGGCGTGACACCGGG	1.7
D25	CAGTACGTTAATGTTT-----CCTGTCAAGTGGCGTGACACCGGG	1.3

Indel pattern	On-target mutagenesis induced by the combination of Lbcr3'5F and LbCpf1 mRNA (ψ)	Read (%)
WT	CAGTACGTTAATGTTCTGATGGTCCATGTCTGTACTCGCCTGTCAAGTGGCGTGACACCGGG	55.5
D15	CAGTACGTTAATGTTCTGATGGTCCA-----TGTCAGTGGCGTGACACCGGG	8.4
D16	CAGTACGTTAATGTTCTGATGGT-----CCTGTCAAGTGGCGTGACACCGGG	5.6
D93	-----GTCAAGTGGCGTGACACCGGG	1.9
D25	----ACGTTAATGTTT-----CCTGTCAAGTGGCGTGACACCGGG	1.6
D85	-----CCTGTCAAGTGGCGTGACACCGGG	1.6
D22	CAGTACGTTAATGTTCTGA-----TGTCAGTGGCGTGACACCGGG	1.4
D25	CAGTACGTTAATGTTCTGATGGTCCAT-----GTGACACCGGG	0.7
D52	-----CTGTCAAGTGGCGTGACACCGGG	0.7
D6	CAGTACGTTAATGTTCTGATGGTCCATGT-----CTCGCCTGTCAAGTGGCGTGACACCGGG	0.7
D34	CAGTACGTTAATGTTCTGAGGG-----AACCAGGG	0.5

WT, unmodified sequence of DNMT1-3. The protospacer-adjacent motif is highlighted in red. Deletions are marked as dashes. Indel-pattern numbers refer to the size of the deletions. Read (%) refers to the read ratio of each mutated site.