Supporting information

Rational Design of Minimum CRISPR Guide RNA by Site-Specific Cas9-RNA Conjugation

Xinyu Ling‡, Xiaoqin Gao‡, Liying Chang, Heqi Chen, Xiaomeng Shi, Tao Liu*

State Key Laboratory of Natural and Biomimetic Drugs, School of Pharmaceutical Sciences, Peking University, 38 Xueyuan Road, Haidian District, Beijing 100191, China

*Correspondence to: taoliupku@pku.edu.cn

‡These authors contributed equally to this work.

Material and Methods

General Information

For all chemical synthesis, solvents and reagents were purchased from Sigma-Aldrich, All the chemical modified RNA (DBCO or -NH2) were purchased from General Biosystems (Anhui) Co. Ltd. The tracrRNA was purchased from Integrated DNA Technologies, Inc. (IDT). U-tracrRNA was purchased from Suzhou Biosyntech Co. Ltd.

Plasmid construction

All plasmids were constructed using Gibson assembly (New England Biolabs). Mutagenesis PCR was conducted using KOD-One PCR master Mix (TOYOBO) followed by DpnI (New England Biolabs) digestion.

DBCO-RNA Modification

Chemical modified RNA (DBCO or -NH2) were purchased from General Biosystems (Anhui) Co. Ltd. Briefly, RNA was synthesized with the 3' NH2 C7 linker, DBCO-NHS(Sigma-Aldrich,761524) was treated 20 excess with the 3' NH2 C7 RNA under the boric acid buffer (PH 9.0) for 3 hours at room temperature. The final DBCO modified RNA was purified by HPLC and evaluated by mass spec.

Expression and purification of SpCas9. The pET28a vector containing 6xHis-tagged 2xNLS SpCas9 was gifted from Professor Caixia Gao at Institute of Genetics and Developmental Biology, Chinese Academy of Sciences. This construct was transformed into the *E.coli* BL21(DE3) strain for protein production. Expression and purification of SpCas9 was performed as described previously[1,2]. Briefly, cells were grown in 2YT medium at 37°C to OD600 of 0.6~0.8, 0.2mM IPTG (Sigma) was added and the temperature was lowered to 18°C. Cells were grown overnight and harvested by centrifugation at 4,000 rpm at 4°C. The protein was purified first by Ni^{2+} affinity chromatography, then by cation exchange (SP, GE healthcare) and finally by size-exclusion chromatography (Superdex 200, GE healthcare). Fractions containing Cas9 protein was collected and concentrated using Amicon Ultra 50K cellulose column (Millipore). Purity and concentration of Cas9 protein were analyzed by SDS-polyacrylamide gel

electrophoresis. Catalytic activities of purified Cas9 was comparable to those of purchased SpyCas9 (New England Biolabs, NEB) under the same optimize condition.

Expression and purification of AeF Cas9 mutants. The plasmid pET-28a-Cas9 containing TAG and plasmid pUltra-MjPolyRS were electroporated into BL21(DE3) strain for protein expression. Cells containing two plasmids were grown at 37°C to OD600 of 0.3, 1mM unnatural amino acid of AeF was added. At the point of OD600 reached 0.6~0.8, 0.2mM IPTG was added for protein induction. Expression and purification conditions were same to wide type Cas9. *In Vitro* catalytic activities of Cas9 variants were analyzed and compared with wide type Cas9 protein.

Nucleic acid preparation. sgRNA templates were amplified from a vector containing a T7 promoter, 20 nucleotide target sequence and sgRNA scaffold. The amplified PCR product was extracted using gel extraction kit (OMEGA) and severed as DNA template for sgRNA transcription reaction. Transcription reactions (20 μL) were conducted in 1x reaction buffer along with 2μL each of 100 mM ATP, GTP, CTP, and UTP, 2μL T7 RNA polymerase (HiScribe™ T7 High Yield RNA Synthesis Kit, New Langland Biolabs), and ∼1 μg DNA template. Reactions were incubated at 37 °C for 4 hours and subsequently purified with RNA Clean & Concentrator-25 Kit (Zymo Research). DNA oligonucleotides and DBCO modified RNAs (Supplementary Table 2) were synthesized by General Biosystems (Anhui) Co. Ltd and Integrated DNA Technologies.

Cleavage activity assay *in vitro*. Target DNA template used for cleavage activity was PCR amplified from plasmid pcDNA3.1-eGFP. 150 ng of DNA template was added to pre-formed RNP complexes (3 pmol) and incubate for 1 h at 37°C in NEB 1X reaction buffer. After incubation, Cas9 RNP was further heated to inactive at 65°C for 10min. The cleavage product was separated on a 1% agarose gel running at 150V. Relative intensities of full length and Cas9-cleavage DNA fragments were determined on a Tanon 1600R Gel Imaging System.

Covalent attachment of Cas9-H116-AeF protein and DBCO modified D-crRNA.

DBCO/NH2 modified 28nt crRNA purchased from by General Biosystems (Anhui) Co. Ltd and V1 tracrRNA were annealed at 95°C for 30 sec, and cooled to room temperature on bench top. For a 10ul

reaction system, 3 pmol Cas9-H116-AeF and 3 pmol annealed dual guide RNA were incubate at 4°C for 3 hours. The negative controls (wide type Cas9 + DBCO modified 28nt crRNA, or Cas9-H116-AeF + NH_2 modified 28nt crRNA) were treated in the same way. For confirming successful conjunction between protein and RNA, reaction products (1 μg RNP) were boiled and loaded on 10% SDS-PAGE gel, running at 150V for 90 min. Gels were stained with Coomassie blue. *In vitro* cleavage activity of these conjugated RNPs were processed as mentioned above.

Human cell culture. Both HEK293 cells and HEK293-EGFP-TetOn cells were cultured at 37 °C with 5% CO2 in DMEM containing 10% fetal bovine serum. Cell culture reagents were purchased from Thermo Fisher Scientific, and cell culture supernatant was tested twice a month for mycoplasma.

RNP-RNA conjugates electroporation. The electroporation machine (Catalog# CTX-1500A LE), the pressured electroporation tubes (Catalog# 20 μL: 12–0107), and the electroporation buffer (Catalog# 13–0104) were provided by Celetrix LLC, Manassas VA. For Cas9-RNA conjugates transfection, 30pmol Cas9 protein was pre-mixed with 30pmol annealed RNA complex at four degree for 3 hours before treating cells. Cells were resuspended in electroporation buffer to 1.5×10^6 cells/ml, then the formed RNP complex was mixed with the cells and transferred to 20 μl electroporation tube. The electroporation condition was 420V, 30ms. After electroporation the cells were immediately transferred back to 2ml warm medium.

Flow cytometry. Doxycycline (Sigma) was added to HEK293-EGFP-TetOn reporter cells 24h after transfection. 24h after induction, cells were trypsinized and resuspended in PBS containing 1% FBS. Data were acquired on a Beckman CytoFLEX Flow Cytometer and were further analyzed using CytExpert software. In all experiments, a minimum of 10,000 cells were analyzed. Live cells were gated based on FSC-A and SSC-A. Live-gated cells were further used to quantify the percentage of EGFP positive populations.

Endogenous Genome editing and off-target detection.

Previous reported Guide-seq was used for evaluation off-target[3]. Human EMX1 Gene was chose for genome editing and off-target detection. The target RNA and modified RNA sequences are listed in

Table S2. For genome editing, Cas9 mutant protein (30pmol), RNA complex (60pmol) and dsODN(90pmol) were pre-mixed at 4 degree for 3 hours before transfection. RNP-ssODN were elecporated into 293 cell by 4D-Nucleofector (Lonza) using program EO-115. Cells were harvested approximately 3 days post transfection and genomic DNA was extracted using TIANamp Genomic DNA Kit(DP304) Library construction ,Guide sequencing and data analysis was completed by Generulor Company Bio-X Lab, Guangzhou 510006, Guangdong, China.

Supplementary Tables

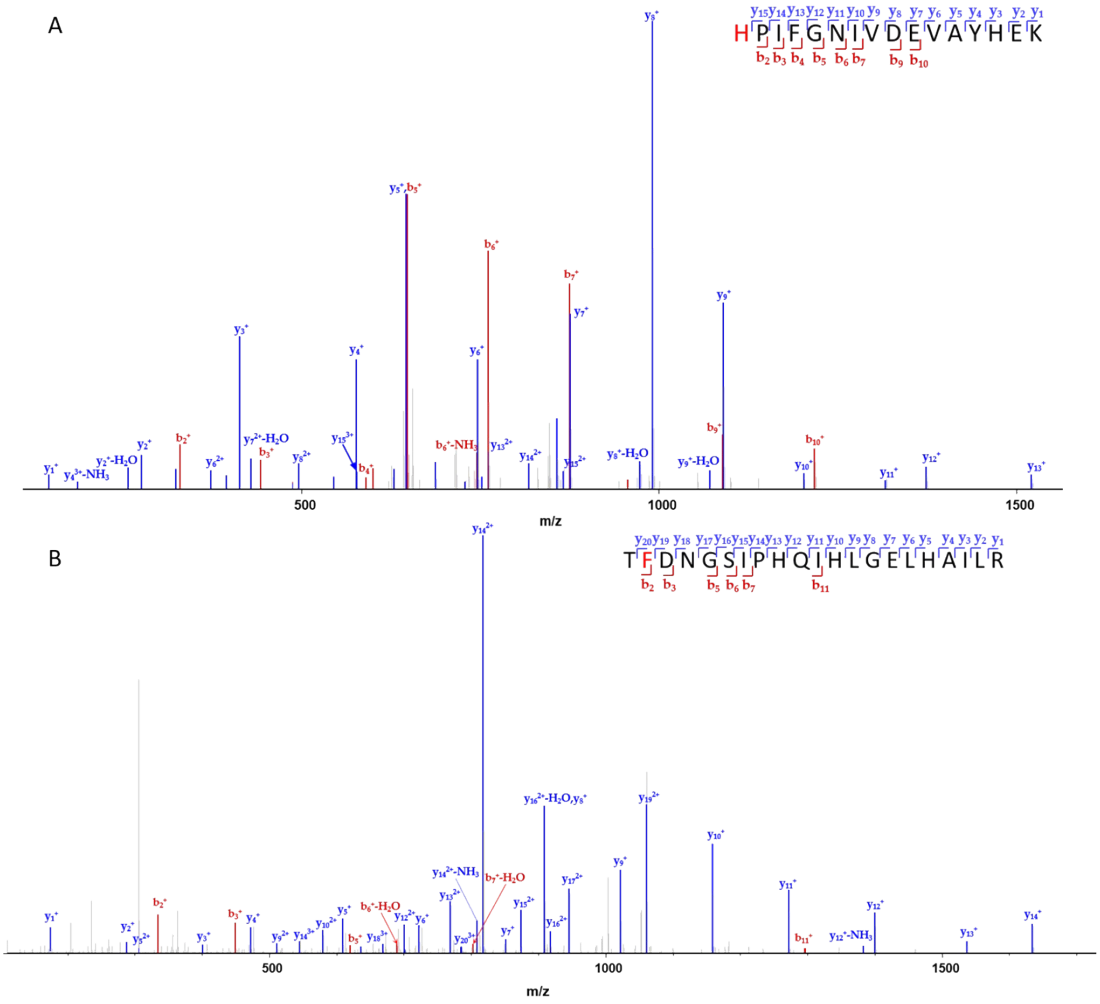
Name	**Sequence**
H116-FW	GAAGAAGACAAGAAGCATGAACGTTAGCCTATTTTTGGAAATATAGTAGATGAAGTTGC
H116-RV	GCAACTTCATCTACTATATTTCCAAAAATAGGCTAACGTTCATGCTTCTTGTCTTCTTC
F405-FW	GATTTGCTGCGCAAGCAACGGACCTAGGACAACGGCTCTATTCCCCATCAAATTCAC
F405-RV	GTGAATTTGATGGGGAATAGAGCCGTTGTCCTAGGTCCGTTGCTTGCGCAGCAAATC

Table S1 Mutagenesis PCR primers used in the study

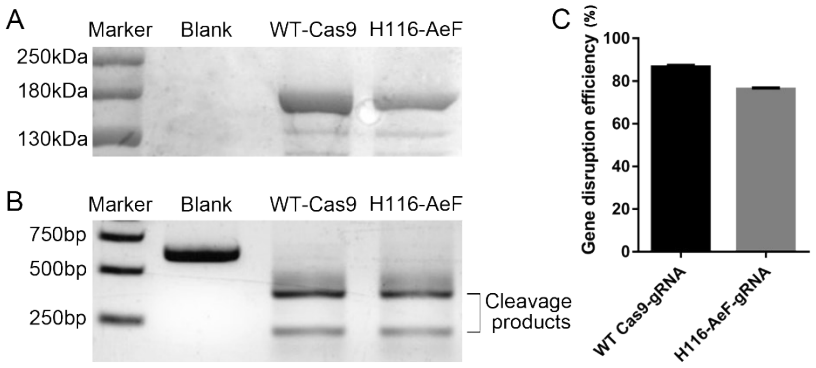
Name	Sequence
36nt crRNA-EGFP	GGGCGAGGAGCUGUUCACCGGUUUUAGAGCUAUGCU
36nt crRNA-EMX1	GAGUCCGAGCAGAAGAAGAAGUUUUAGAGCUAUGCU
67nt TracrRNA(V0)	AGCAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCUUU
28nt crRNA-EGFP	GGGCGAGGAGCUGUUCACCGGUUUUAGA
20nt crRNA-EGFP	GGGCGAGGAGCUGUUCACCG
28nt crRNA-EMX1	GAGUCCGAGCAGAAGAAGAAGUUUUAGA
20nt crRNA-EMX1	GAGUCCGAGCAGAAGAAGAA
tracrRNA V1	GCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGC
tracrRNA V2	AAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGC
tracrRNA V3	UAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGC
RNA V4(adapter RNA)	GCUAUGCU
hRNA1	GUUUUAGAGCUAUGCU
hRNA2	UUUUAGAGCUAUGCU
U-tracrRNA	UUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCUUUU

Table S2 RNA sequence used in the study

Supplementary Figures

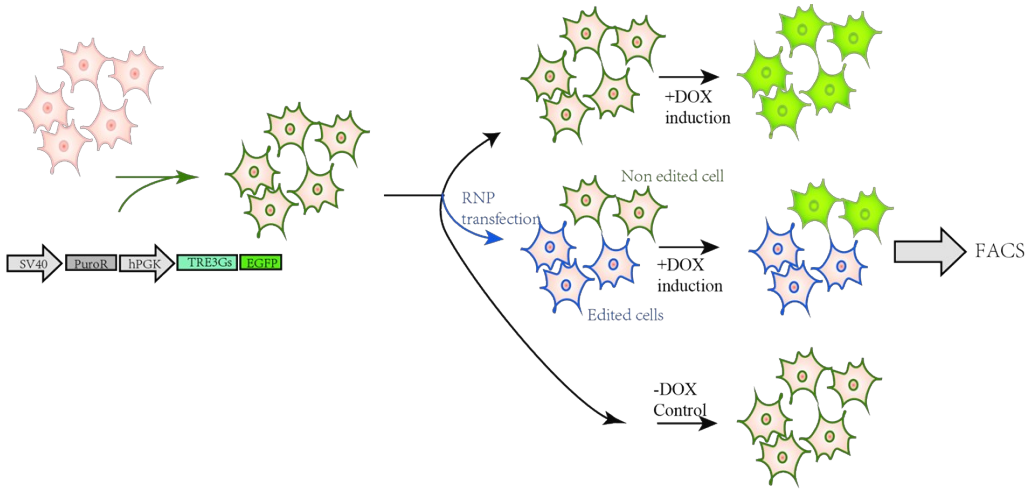


Supplementary Figure 1. Peptide mass spec of Cas9 mutants. (A) H116-AeF (B) F405-AeF

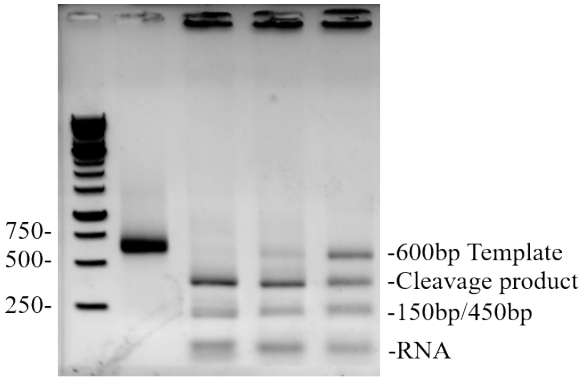


Supplementary Figure 2. Assessment of the activity of a chemically modified Cas9 mutant in vitro and in cells. (A) SDS-PAGE analysis of the purified H116-AeF Cas9 mutant. (B) Nuclease

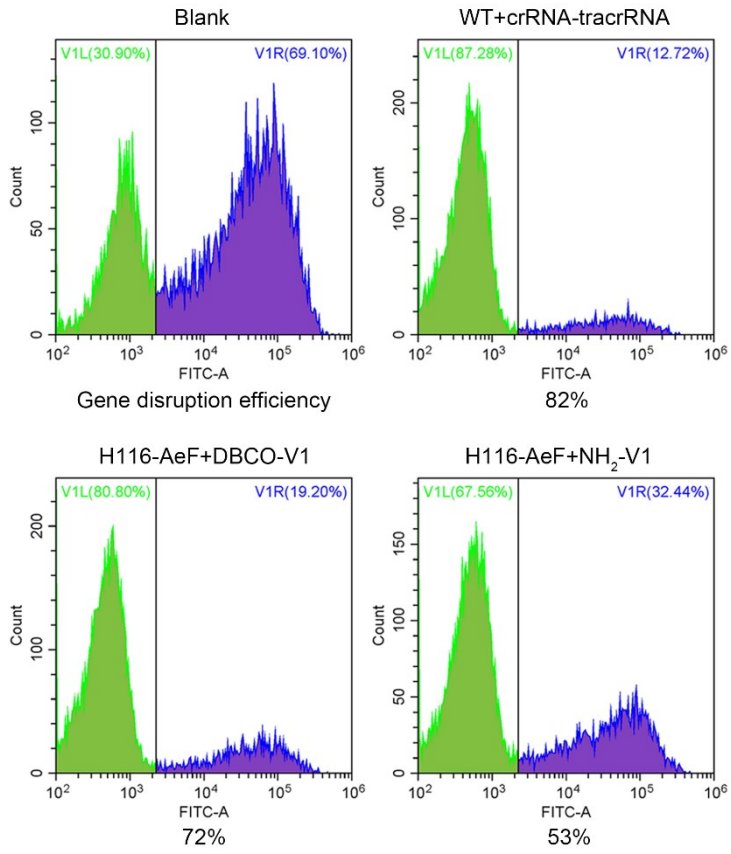
activities of WT and mutant Cas9 proteins in vitro. Double-stranded DNA was digested for 1 h at 37 °C in the presence of a pre-assembled RNP complex (3 pmol) comprising the protein and gRNA. (C) Gene disruption efficiencies of the WT and mutant Cas9 proteins. HEK293-EGFP-TetOn cells were transfected with 30 pmol mutant or WT Cas9 RNP by means of electroporation. Genome editing activity was quantified by flow cytometry 2 days after transfection. Error bars represent standard deviations (n = 3).



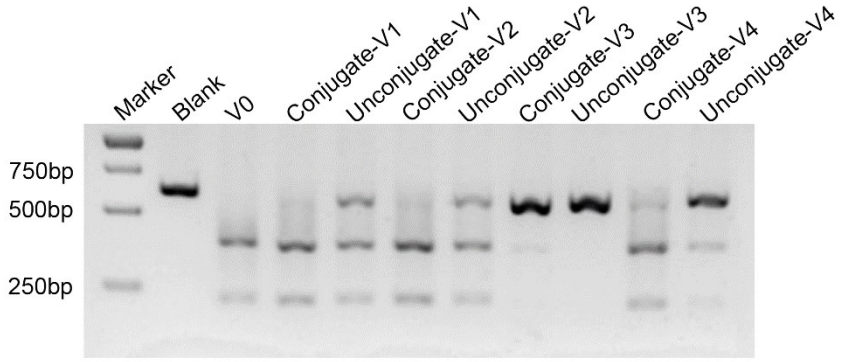
Supplementary Figure 3. Scheme of Inducible HEK293-EGFP-TetOn reporting cell line construction. Gene cluster containing Puromycin resistant gene driven by SV40 promoter and EGFP driven by hPGK promoter and TRE3Gs were packed in lentivirus. 293 cells were infected by lentivirus and sorting by flow cytometry to afford reporting cell line.



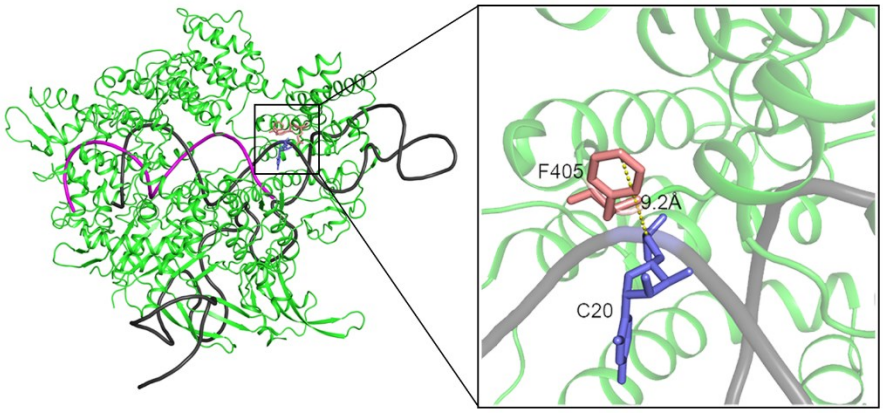
Supplementary Figure 4. In vitro evaluation of the effect of conjugation between D-crRNA and H116-AeF Cas9 on cleavage activity.



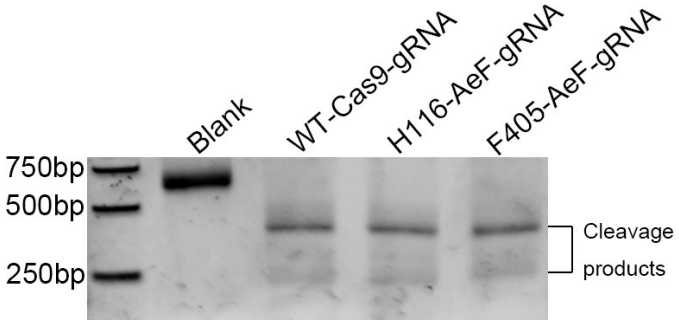
Supplementary Figure 5. Assessment of the activity of a H116-AeF 28nt D-RNA conjugates in cells. Flow cytometry analysis of HEK293-EGFP-TetOn cells transfected with mutant and WT RNPs (30 pmol) by electroporation (analysis was conducted 48h post-transduction).



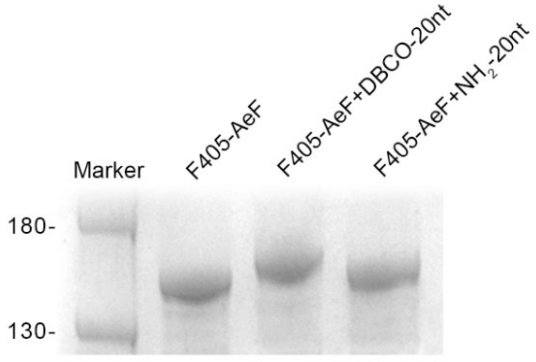
Supplementary Figure 6. In vitro assessment of nuclease activities of various RNA pairs. Double-stranded DNA was digested for 1 h at 37 °C in the presence of a pre-assembled RNP complex (3 pmol) consisting of Cas9 and pre-annealed RNA.



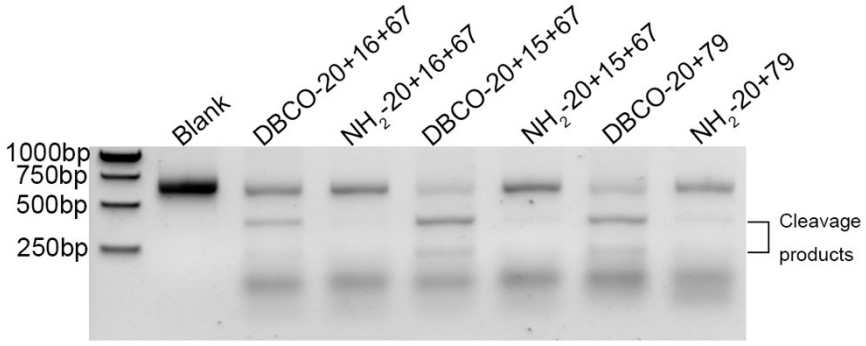
Supplementary Figure 7 Crystal structure of *Streptococcus pyrogenes* Cas9 with gRNA (PDB ID: 4OO8). Residue F405, which was selected for AeF mutagenesis, is labeled in red; the distance between F405 and base C20 was measured in PyMOL.



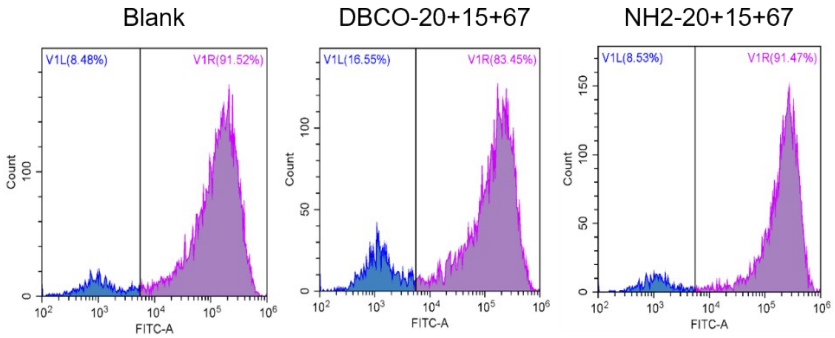
Supplementary Figure 8. In vitro cleavage activity of F405-AeF protein. Different Cas9 mutants were assembled into RNPs with gRNA targeting EGFP gene. The resulting complex were measured for cleavage activity, and the cleavage products were measured on a 1% agarose.



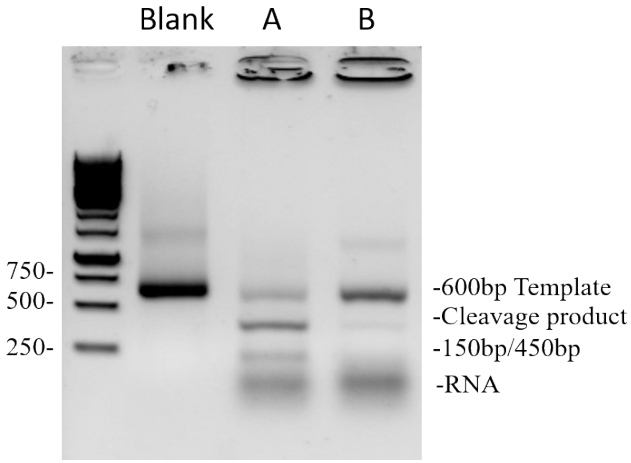
Supplementary Figure 9. SDS-PAGE analysis F405AeF-20nt RNA conjugates



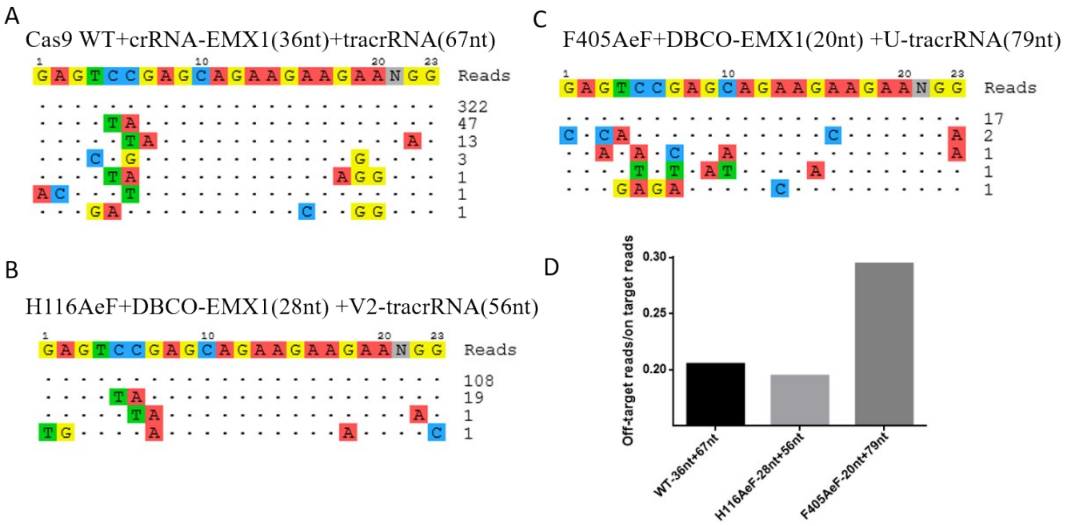
Supplementary Figure 10. In vitro cleavage activity of 20nt-crRNA with hRNAs and tracrRNA. RNPs were pre-incubated with NH2- or DBCO-modified annealed RNA, forming Cas9 mutants-RNA conjugate complex and analyzed for cleavage activity in vitro on agarose gel.



Supplementary Figure 11. In cell cleavage activity of 20nt-crRNA with 15nt hybridization RNA and tracrRNA. F405-AeF(18pmol) with DBCO- or NH2- modified 20nt M-crRNA, 15nt hRNA, 67nt tracrRNA were assembled in vitro and transfected in cell through electroporation. Editing efficiency was quantified by flow cytometry 48h post-transduction.



Supplementary Figure 12. In vitro activity of assemblies containing F405-AeF and 20 nt RNA conjugates: lane A, DBCO-20nt+79nt U-tracrRNA; lane B: NH2-20nt+79nt U-tracrRNA.



Supplementary Figure 13. Evaluate the target sequence specificity using Guide-seq. On target and off target reads are presented (A) General Cas9 RNP system (B) H116AeF+28nt RNA system (C)F405AeF+20nt RNA system. (D)The ratio of off-target and on target numbers

Notes and references

1 A. Mir, J. F. Alterman, M. R. Hassler, A. J. Debacker, E. Hudgens, D. Echeverria, M. H. Brodsky, A. Khvorova, J. K. Watts and E. J. Sontheimer, *Nature Communications*, 2018, **9**, 2641.

2 X. Ling, B. Xie, X. Gao, L. Chang, W. Zheng, H. Chen, Y. Huang, L. Tan, M. Li and T. Liu, *Science Advances*, 2020, **6**, eaaz0051.

3 S. Q. Tsai, Z. Zheng, N. T. Nguyen, M. Liebers, V. V. Topkar, V. Thapar, N. Wyvekens, C. Khayter, A. J. Iafrate, L. P. Le, M. J. Aryee and J. K. Joung, *Nature Biotechnology*, 2015, **33**, 187–197.