Chimeric crRNAs with 19 DNA residues in the guide region show retained DNA cleavage activity of Cas9 with a potential to improve the specificity

Hyo Young Kim[a], Seong Jae Kang[a], Yongmoon Jeon[a], Jinsu An[a,b], Jihyun Park[a], Hee Jae Lee[a], Jeong-Eun Jang[a], JongSeong Ahn[c], Duhee Bang[c], Hak Suk Chung[a,b], Cherlhyun Jeong[a] and Dae-Ro Ahn[a,b,*]

[a]*Center for Theragnosis, Biomedical Research Institute, Korea Institute of Science and Technology (KIST), Hwarangno 14-gil 5, Seongbuk-gu, Seoul 02792, Republic of Korea,*
[b]*Division of Biomedical Science and Technology, Korea University of Science and Technology (UST), Hwarangno 14-gil 5, Seongbuk-gu, Seoul 02792, Korea.*
[c]*Department of Chemistry, College of Science, Yonsei University, Republic of Korea.*

MATERIALS AND METHODS

Oligonucleotides

DNA, crRNA, and chimeric oligonucleotides were purchased from Integrated DNA Technologies (USA) and Bioneer (Korea). Alt-R® CRISPR-Cas9 tracrRNA (Integrated DNA Technologies, USA) was used as tracrRNA.

Expression and purification of Cas9.

pET-NLS-Cas9-6xHis was purchased from Addgene (USA, plasmid #62934) and purified as previously described (1-2) with minor modifications. Pellets were harvested, resuspended in buffer A (50 mM Tris-HCl, pH 8.0, 0.5 M NaCl, 20 mM imidazole and 5 mM 2-Mercaptoethanol (Bio-rad) and lysed by sonication. After centrifugation at 10,000g for 40 minutes at 4°C, NLS-Cas9 was separated by Ni-NTA affinity chromatography. The eluted Cas9 was loaded onto a Hiprep SP HP 16/10 column (GE Health-care Life Sciences) and purified by a linear gradient of KCl from 0.15 M to 1M in buffer B (20 mM HEPES, pH 7.5, 10 % glycerol, and 2 mM DTT). Eluted protein was concentrated, flash-frozen in liquid

nitrogen and stored at -80°C. The final purity and concentration of Cas9 was determined by a SDS PAGE gel and Bradford protein assay (Bio-Rad, USA) using bovine serum albumin as a protein standard, respectively. Catalytically inactive mutant dCas9 (D10A/H840A) was generated by Quick change site-directed mutagenesis following manufacture's protocol (Agilent, USA) as well as previously described (3) and purified following the same procedure as for the wild type Cas9.

In vitro Cas9 activity assays

The pSMART-EGFP plasmid encoding a green fluorescent protein (GFP) gene (4) was kindly provided by Prof. Jong Bum Lee (University of Seoul) and used as the DNA substrate for the in vitro DNA cleavage assays (Table S1). The reaction mixtures (30 μL) containing Cas9 (33 nM), linearized pSMART-EGFP substrate (1 nM) crRNA (33 nM), and tracrRNA (33 nM) in the reaction buffer (20 mM HEPES pH6.5, 100 mM NaCl, 5 mM $MgCl_2$, 0.1 mM EDTA) were incubated at 37 °C for 10 or 60 min. The reactions were quenched by adding 6 ×gel loading buffer (19.8 mM Tris-HCl, pH 8.0, 66 mM EDTA, 0.017% SDS, 2.5% Ficoll®-400, 0.015% bromophenol blue,5 μL, New England Biolabs, USA) and analyzed by agarose gel (0.7%) electrophoresis. Gels were stained with SYBR gold (Life Technologies, USA) and imaged with ChemiDoc XRS^+ system (Bio-rad). Band intensities were quantified using ImageJ. The cleavage level (%) was determined by 100 × [sum of band intensities of cleavage products]/[sum of all band intensities]. The data show the averaged values with standard deviation of three independent experiments.

Single-molecule FRET experiments

The PEG-biotin coated quartz surface was prepared for single-molecule FRET imaging to prevent nonspecific binding. A sample chamber was constructed with prepared PEG-biotin

quartz as previously described (5). 5′-amine-NTS and 5′-biotin-TS-amine-3′DNA oligonucleotides (Integrated DNA technologies, USA) were conjugated with fluorescent dyes (Cy3 and Cy5) modified with NHS ester (GE Healthcare, USA) and purified using ethanol precipitation. The TS/NTS DNA duplex was annealed at 5 μM concentration (1:1.25 molar ratio of TS:NTS) in a buffer containing 10 mM Tris-HCl (pH 8.0) and 100 mM NaCl. The DNA duplex was immobilized on the PEG-coated surface via a streptavidin-biotin interaction. To prepare the Cas9/tracrRNA/crRNA complex, Cas9 was incubated with tracrRNA and crRNA that was annealed in advanced in a reaction buffer (50 mM Tris-HCl, pH 7.6, 100 mM KCl, 10 mM $MgCl_2$) for 10 min at room temperature. The oxygen scavenging system consisting of 3 mM Trolox (Merck, USA), 5 mM 3,4-dihydroxybenzoic acid (Sigma-Aldrich, USA), 1.64 U/ml recombinant protocatechuate 3,4-dioxygenase from bacteria (Oriental Yeast, Japan) was included in the reaction buffer to minimize photobleaching and suppress photoblinking. Fluorescence signals from single-molecules were collected in a prism-type total internal reflection microscope with 2 s or 0.4 s time resolution. To measure the FRET efficiency and the acceptor signal, the donor and acceptor molecules were alternately excited at 532 nm and 633 nm, respectively. Fluorescence signal from Cy3 and Cy5 were collected using a water immersion objective lens (UPlanSApo 60x, Olympus, Japan), filtered with emission filters (LP03-532RU and NF03-633E-25, Semrock, USA) to block scattered laser beams, separated with a dichroic mirror (635dcxr, Chroma, USA), and imaged onto an EMCCD (Andor, UK). Real-time single-molecule fluorescence intensity data were analyzed using Matlab 2016b (Mathworks, USA) and OriginPro 8 (OriginLab, USA) programs.

Gel shift binding assay: dCas9/tracrRNA to crRNA

The chimeric crRNA (50 nM) was mixed with dCas9/tracrRNA complex at varying concentrations (0 – 200 nM) in the Cas9-reaction buffer (20 mM HEPES, pH 6.5, 100 mM

NaCl, 5 mM $MgCl_2$, 0.1 mM EDTA) and incubated at 37°C for 60 min. Free and complexed chimeric crRNA were analyzed on the non-denaturing 15% polyacrylamide gel electrophoresis using 0.5×TBS buffer supplemented with 5 mM $MgCl_2$ at 4 °C. The free and complexed crRNA bands were imaged in ChemiDoc XRS^+ system (Bio-Rad, USA) after staining with SYBR gold. The relative intensities of the bands quantified using ImageJ software, plotted in Prism (GraphPad), and fit to a one-site binding hyperbola equation to estimate the K_d values based on the non-linear regression.

Gel shift binding assay: dCas9/crRNA/tracrRNA to DNA duplex

FAM-labeled DNA duplex (GGGCGAGGAGCTGTTCACCGGGGTGGTGCC-FAM; GGCACCACCCCGGTGAACAGCTCCTCGCCC) (50 nM) was mixed with the dCas9/crRNA/tracrRNA complex at varying concentrations (0–500 nM) in the binding buffer (20 mM Tris-HCl, pH 7.5, 100 mM KCl, 5 mM $MgCl_2$, 1mM DTT) and incubated at 37°C for 60 min. After non-denaturing 15% polyacrylamide gel electrophoresis to separate free and complexed DNA duplex, K_d values were calculated by using the same manner employed for determination of K_d values for crRNA binding to Cas9/tracrRNA.

Preparation of off-targets

The mutant pSMART-EGFP plasmids containing off-targets were prepared by mutagenesis of the pSMART-EGFP plasmid substrate using Q5 site-directed mutagenesis kit (New England Biolabs, USA). The primers used for mutagenesis are listed in Table S2. The sequences of mutant plasmids were confirmed by DNA sequencing (Macrogene, Korea).

Cellular gene disruption

HeLa/EGFP cells (1×10^5) were seeded in 12-well plates and were cultured by following the

previously reported procedure (1) Cas9 (1 μg) were complexed with the mixture of 19-DNA chimeric crRNA (125 ng) and tracrRNA(125 ng) in one tube, while the Lipofectamine CRISPRMAX reagent (3 μL, Thermo Fisher Scientific, USA) was mixed with Opti-MEM medium (100 μL, USA) in another tube at room temperature for 5 min. Then, the solutions in the tubes were combined and incubated at room temperature for 10 min to form the ribonucleoprotein (RNP) lipoplex. After the treatment of the HeLa/EGFP cells treated with the lipoplex in serum-free DMEM medium (800 μL) for 48 h, the EGFP expression level in the cells was analyzed using flow cytometry (GUAVA, Millipore, USA).

To examine gene disruption level, T7 endonuclease I digestion (T7E1) assay was also performed using the PCR-amplified target gene in the cells treated with lipoplex by following the previously reported procedure (1). Briefly, the genomic DNA of the lipoplex-treated cells were prepared using DNA extraction kit (MagListo 5M genomic DNA extraction kit, Bioneer, Korea). The genomic DNA was PCR-amplified with Q5 Hot Start High-Fidelity 2× Master Mix (Engen mutation detection kit, New England Biolabs, USA) (forward primer: GATTCTAGCTGATCGTGGACCGG; reverse primer: TTGCCGGTGGTGCAGATGAA) by following the manufacturer's protocol. PCR products digested using T7E1 (New England Biolabs, USA) at 37 °C for 15 min and analyzed on 2 % agarose gels.

TIDE analysis

Hela-GFP (1×10^5) cells were transfected with 8 nM of crRNA/tracrRNA/cas9 complexes using CRISPRMax (ThermoFisher Scientific, USA)) reagent for 48 h at 37 ℃. Cells were washed twice with DPBS. The genomic DNA was extracted using MagListo™ 5M Genomic DNA extraction kit (Bioneer, Korea) according to manufacturer's instruction. On- and off-target region was amplified using polymerase chain reaction (PCR). DNA sequences were analyzed

by Sanger sequencing, and non-homologous end joining followed by CRISPR/Cas9 cleavage was validated using Tracking of Indels by DEcomposition (TIDE) (6).

Table S1. The sequence of the substrate plasmid (pSMART-EGFP). EGFP coding sequences are colored in blue. The Pvu I restriction enzyme site is colored in red. The target sequences used in this study are indicated with yellow boxes or underline.

CCCGTGTAAA	ACGACGGCCA	GTTTATCTAG	TCAGCTTGAT	TCTAGCTGAT
CGTGGACCGG	AAGGTGAGCC	AGTGAGTTGA	TTGCAGTCCA	GTTACGCTGG
AGTCTGAGGC	TCGTCCTGAA	TGATATGCGA	CCGCCGGAGG	GTTGCGTTTG
AGACGGGCGA	CAGATCGACA	CTGCTCGATC	CGCTCGCACC	TAATACGACT
CACTATAGGG	ATGCCACCAT	GGATGGTGAG	CAAGGGCGAG	GAGCTGTTCA
CCGGGGTGGT	GCCCATCCTG	GTCGAGCTGG	ACGGCGACGT	AAACGGCCAC
AAGTTCAGCG	TGTCCGGCGA	GGGCGAGGGC	GATGCCACCT	ACGGCAAGCT
GACCCTGAAG	TTCATCTGCA	CCACCGGCAA	GCTGCCCGTG	CCCTGGCCCA
CCCTCGTGAC	CACCCTGACC	TACGGCGTGC	AGTGCTTCAG	CCGCTACCCC
GACCACATGA	AGCAGCACGA	CTTCTTCAAG	TCCGCCATGC	CCGAAGGCTA
CGTCCAGGAG	CGCACCATCT	TCTTCAAGGA	CGACGGCAAC	TACAAGACCC
GCGCCGAGGT	GAAGTTCGAG	GGCGACACCC	TGGTGAACCG	CATCGAGCTG
AAGGGCATCG	ACTTCAAGGA	GGACGGCAAC	ATCCTGGGGC	ACAAGCTGGA
GTACAACTAC	AACAGCCACA	ACGTCTATAT	CATGGCCGAC	AAGCAGAAGA
ACGGCATCAA	GGTGAACTTC	AAGATCCGCC	ACAACATCGA	GGACGGCAGC
GTGCAGCTCG	CCGACCACTA	CCAGCAGAAC	ACCCCCATCG	GCGACGGCCC
CGTGCTGCTG	CCCGACAACC	ACTACCTGAG	CACCCAGTCC	GCCCTGAGCA
AAGACCCCAA	CGAGAAGCGC	GATCACATGG	TCCTGCTGGA	GTTCGTGACC
GCCGCCGGGA	TCACTCTCGG	CATGGACGAG	CTGTACAAGT	AAGGATCGAC
GAGAGCAGCG	CGACTGGATC	AGTTCTGGAC	GAGCGAGCTG	TCGTCCGACC
CGTGATCTTA	CGGCATTATA	CGTATGATCG	GTCCACGATC	AGCTAGATTA
TCTAGTCAGC	TTGATGTCAT	AGCTGTTTCC	TGAGGCTCAA	TACTGACCAT
TTAAATCATA	CCTGACCTCC	ATAGCAGAAA	GTCAAAAGCC	TCCGACCGGA
GGCTTTTGAC	TTGATCGGCA	CGTAAGAGGT	TCCAACTTTC	ACCATAATGA
AATAAGATCA	CTACCGGGCG	TATTTTTTGA	GTTATCGAGA	TTTTCAGGAG
CTAAGGAAGC	TAAAATGAGT	ATTCAACATT	TCCGTGTCGC	CCTTATTCCC
TTTTTTGCGG	CATTTTGCCT	TCCTGTTTTT	GCTCACCCAG	AAACGCTGGT
GAAAGTAAAA	GATGCTGAAG	ATCAGTTGGG	TGCACGAGTG	GGTTACATCG
AACTGGATCT	CAACAGCGGT	AAGATCCTTG	AGAGTTTACG	CCCCGAAGAA
CGTTTTCCAA	TGATGAGCAC	TTTTAAAGTT	CTGCTATGTG	GCGCGGTATT
ATCCCGTATT	GACGCCGGGC	AAGAGCAACT	CGGTCGCCGC	ATACACTATT
CTCAGAATGA	CTTGGTTGAG	TACTCACCAG	TCACAGAAAA	GCATCTCACG
GATGGCATGA	CAGTAAGAGA	ATTATGCAGT	GCTGCCATAA	CCATGAGTGA
TAACACTGCG	GCCAACTTAC	TTCTGGCAAC	GATCGGAGGA	CCGAAGGAGC
TAACCGCTTT	TTTGCACAAC	ATGGGGGATC	ATGTAACTCG	CCTTGATCGT
TGGGAACCGG	AGCTGAATGA	AGCCATACCA	AACGACGAGC	GTGACACCAC
GATGCCTGTA	GCAATGGCAA	CAACGTTGCG	CAAACTATTA	ACTGGCGAAC
TACTTACTCT	AGCTTCCCGG	CAACAATTAA	TAGACTGGAT	GGAGGCGGAT

AAAGTTGCAG GATCACTTCT GCGCTCGGCC CTCCCGGCTG GCTGGTTTAT
TGCTGATAAA TCTGGAGCCG GTGAGCGTGG GTCTCGCGGT ATCATTGCAG
CACTGGGGCC AGATGGTAAG CCCTCCCGCA TCGTAGTTAT CTACACGACG
GGGAGTCAGG CAACTATGGA TGAACGAAAT AGACAGATCG CTGAGATAGG
TGCCTCACTG ATTAAGCATT GGTAATGAGG GCCCAAATGT AATCACCTGG
CTCACCTTCG GGTGGGCCTT TCTTGAGGAC CTAAATGTAA TCACCTGGCT
CACCTTCGGG TGGGCCTTTC TGCGTTGCTG GCGTTTTTCC ATAGGCTCCG
CCCCCCTGAC GAGCATCACA AAAATCGATG CTCAAGTCAG AGGTGGCGAA
ACCCGACAGG ACTATAAAGA TACCAGGCGT TTCCCCCTGG AAGCTCCCTC
GTGCGCTCTC CTGTTCCGAC CCTGCCGCTT ACCGGATACC TGTCCGCCTT
TCTCCCTTCG GGAAGCGTGG CGCTTTCTCA TAGCTCACGC TGTAGGTATC
TCAGTTCGGT GTAGGTCGTT CGCTCCAAGC TGGGCTGTGT GCACGAACCC
CCCGTTCAGC CCGACCGCTG CGCCTTATCC GGTAACTATC GTCTTGAGTC
CAACCCGGTA AGACACGACT TATCGCCACT GGCAGCAGCC ACTGGTAACA
GGATTAGCAG AGCGAGGTAT GTAGGCGGTG CTACAGAGTT CTTGAAGTGG
TGGCCTAACT ACGGCTACAC TAGAAGAACA GTATTTGGTA TCTGCGCTCT
GCTGAAGCCA GTTACCTCGG AAAAAGAGTT GGTAGCTCTT GATCCGGCAA
ACAAACCACC GCTGGTAGCG GTGGTTTTTT TGTTTGCAAG CAGCAGATTA
CGCGCAGAAA AAAAGGATCT CAAGAAGATC CTTTGATTTT CTACCGAAGA
AAGGCCCA

Table S2. The primer sequences used for preparation of off-target substrate plasmids

Point mutation	Primer	Annealing (℃)	Reverse Primer
C1A	AAGGGAGAGGAGCTGTTC	59	GCTCACCATGGTGGCGA
C1T	AAGGGTGAGGAGCTGTTC	57	
C1G	AAGGGGGAGGAGCTGTTC	59	
T11A	GAGCTGATCACCGGGGTGGT	63	CTCGCCCTTGCTCACCAT
T11G	GAGCTGGTCACCGGGGTGGT	62	
T11C	GAGCTGCTCACCGGGGTGGT	62	
G19A	TCACCGGAGTGGTGCCCAT	68	ACAGCTCCTCGCCCTTGCTCACCA
G19T	TCACCGGTGTGGTGCCCAT	68	

G19C	TCACCGGCGTGGTGCCCAT	68	
G17A	CTGTTCACCAGGGTGGTGCC	55	GGCACCACCCTGGTGAACAG
G2C/A14T	GAGCAAGGGCCAGGAGCTGTTGTCCGGGGTGGT	67	ACCACCCCGGAGAACAGCTCCTGGCCCTTGCTG
C1A/G5C	GATGGTGAGCAAGGGAGAGCAGCTGTTCACCGGGG	67	CCCCGGTGAACAGCTGCTCTCCCTTGCTCACCATC
A6C/G7T	AAGGGCGAGGCTCTGTTCACCGG	55	CCGGTGAACAGAGCCTCGCCCTT
C8T/T9C	GGGCGAGGAGTCGTTCACCGGG	56	CCCGGTGAACGACTCCTCGCCC
C1T/A3T/G18A	TGAGCAAGGGTGTGGAGCTGTTCACCGAGGTGGTGCCC	68	GGGCACCACCTCGGTGAACAGCTCCACACCCTTGCTCA

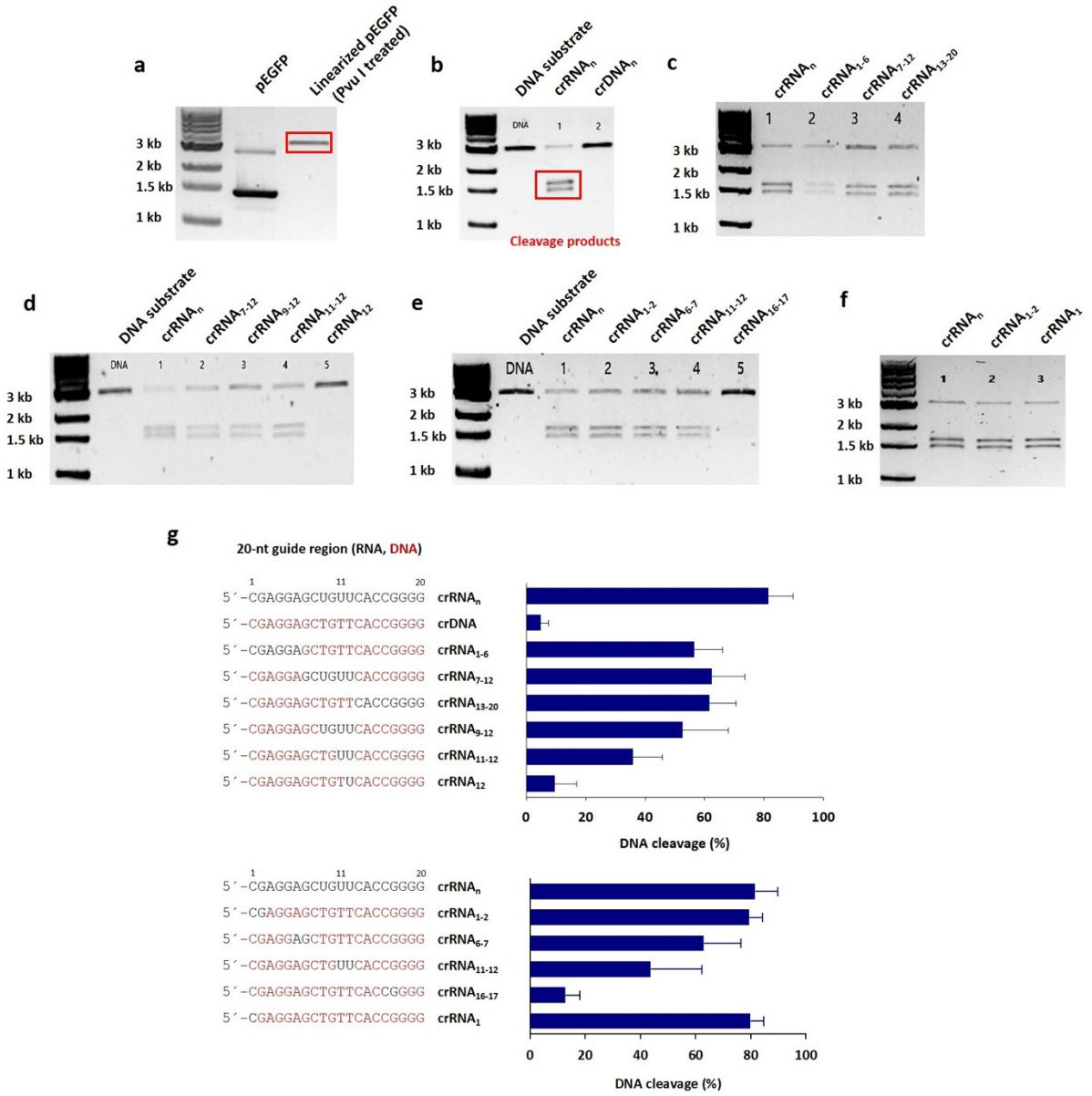

Figure S1. Cas9-mediated DNA cleavage assays using a linearized plasmid DNA substrate. (a) The linearized plasmid DNA substrate prepared by treatment a circular plasmid with Pvu I. (b) Activity of Cas9 guided by (b) $crDNA_n$ and (c-f) various chimeric crRNAs compared with that by $crRNA_n$. (g) Quantified activities of guides based on the intensity of gel bands were displayed in bar graphs.

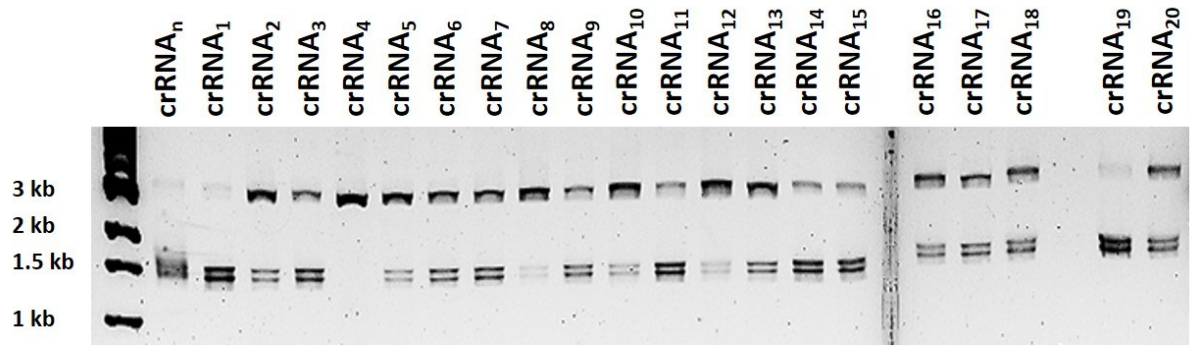

Figure S2. Activity of Cas9 guided by 19-DNA chimeric crRNAs.

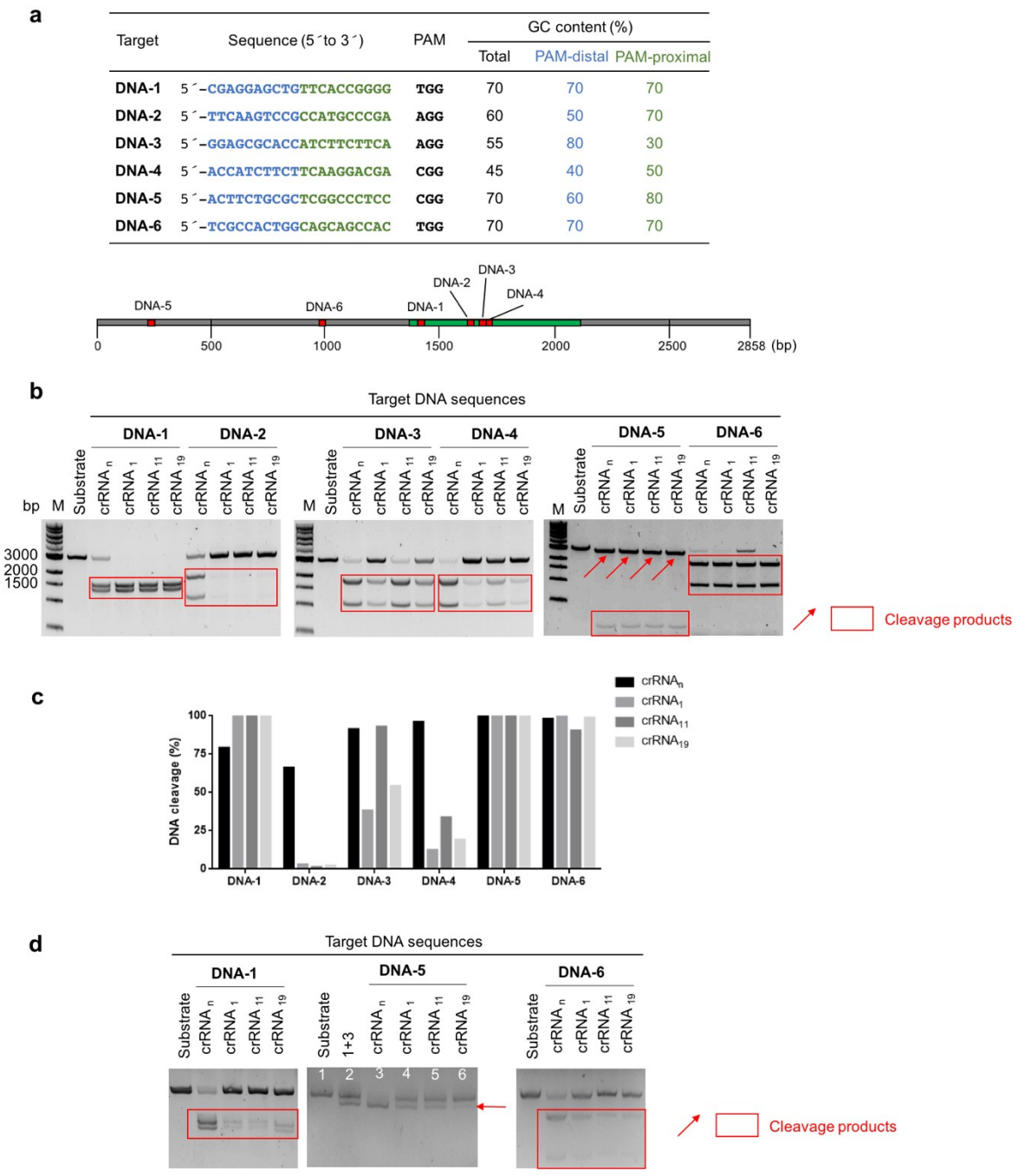

Figure S3. DNA target sequences with various GC-content in PAM-proximal and PAM-distal regions tested for the cleavage reaction with 19-DNA chimeric crRNAs (top). The target sites (red) in the linearized plasmid substrate containing EGFP gene (green) is schematically presented (bottom). (b) Gel analysis of in vitro cleavage of target DNAs (DNA-1 to -6) for 1 h using Cas9 guided by 19-DNA chimeric crRNAs and (c) a bar graph to show quantified DNA

cleavage yields estimated based on the band intensity. (d) Gel analysis of in vitro cleavage of target DNAs for 10 min.

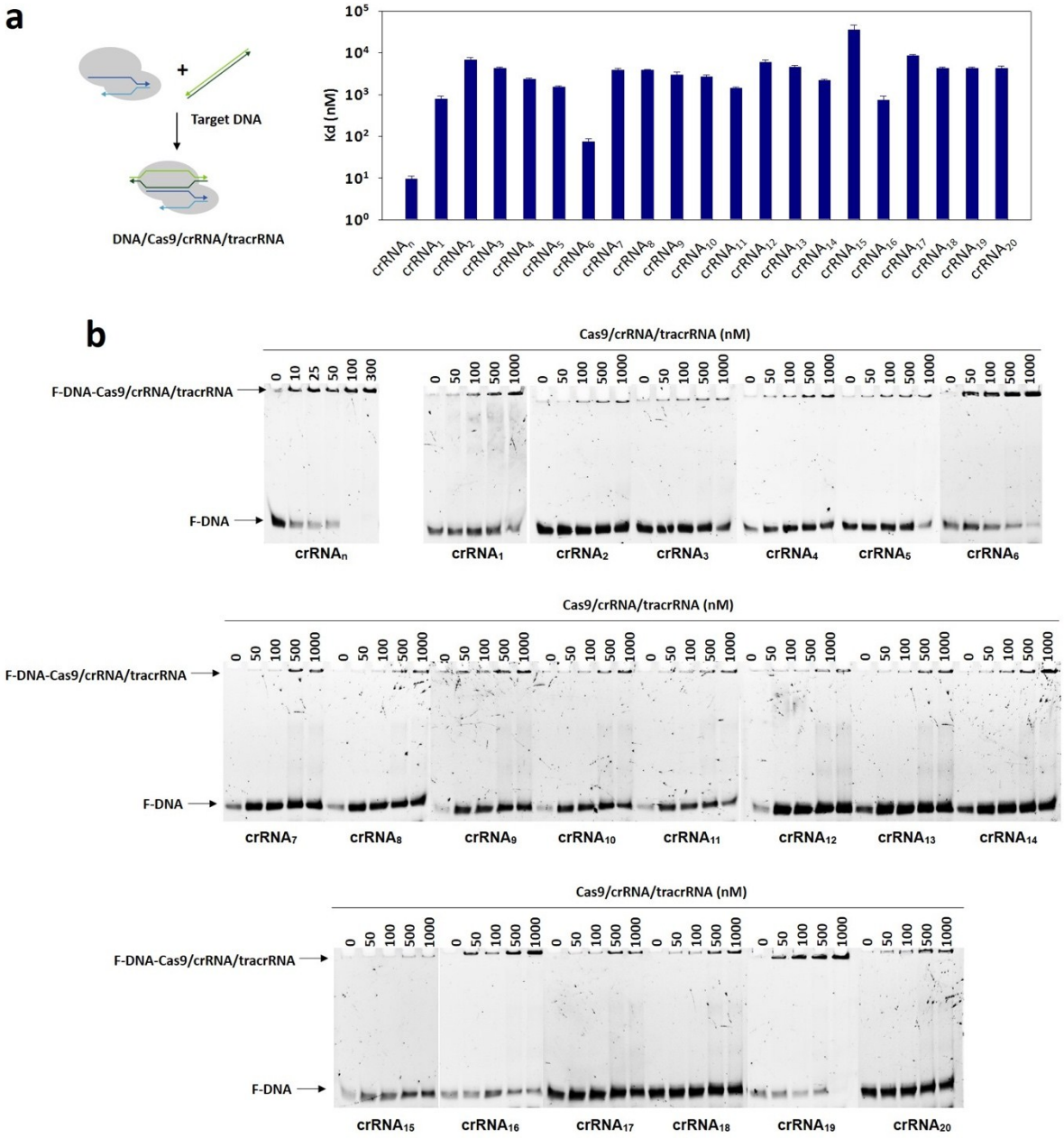

Figure S4. (a) K_d values of Cas9/crRNA/tracrRNA binding to the target DNA determined by (b) gel shift assays.

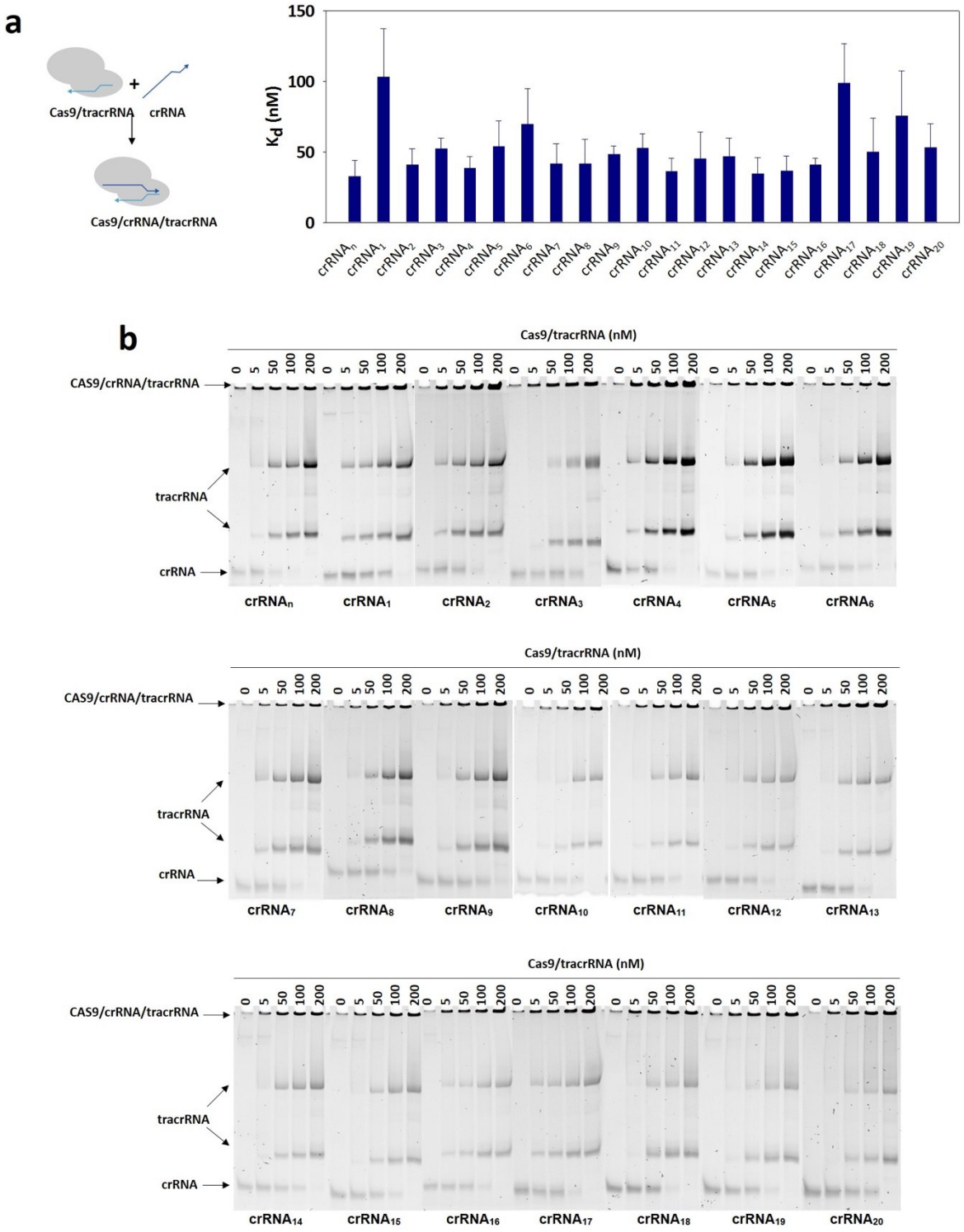

Figure S5. (a) K_d values of Cas9/tracrRNA binding to chimeric crRNAs determined by (b) gel shift assays.

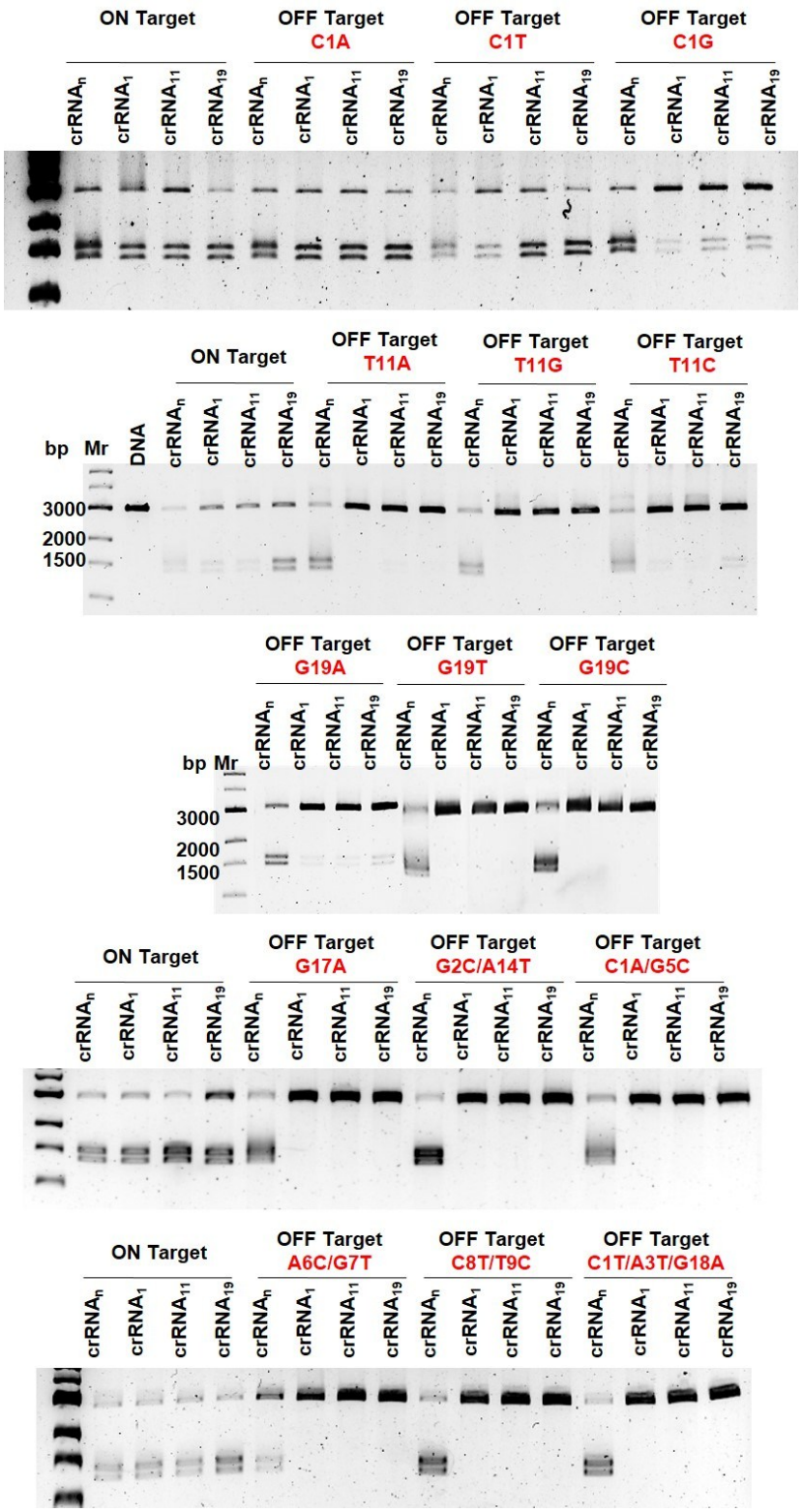

Figure S6. Off-target DNA cleavage with Cas9 guided by 19-DNA chimeric crRNAs.

a

	Sequence
Target DNA-1	CGAGGAGCTGTTCACCGGGG
Off-target in chromosome 4	AGACGAGCTGTTCACCGGGA
Target DNA-3	GGAGCGCACCATCTTCTTCA
Off-target in chromosome 10	GGAGCACCCCATCTTCTTCG

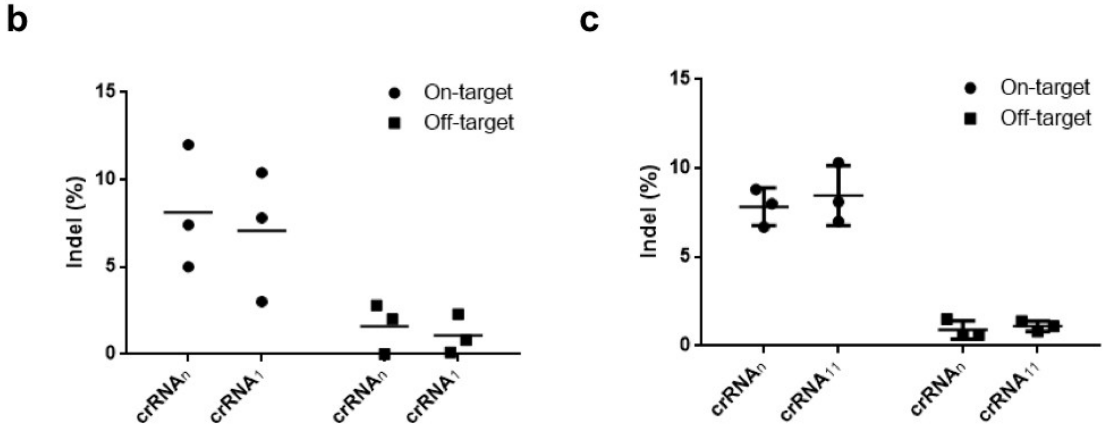

Figure S7. Cellular on-target and off-target cleavage with Cas9 guided by 19-DNA chimeric crRNA. (a) The target sequences and the corresponding off-target sequences found endogenously. (b) Indel levels in DNA-1 target and the off-target in chromosome 4 induced by treatment of cells with Cas9 guided by $crRNA_1$. The indel levels were estimated by Tracking of Indels by DEcomposition or TIDE (6). (c) Indel levels in DNA-3 target and the off-target in chromosome 10 induced by treatment of cells with Cas9 guided by $crRNA_{11}$. The indel levels were estimated by TIDE.

References

1. J. S. Ha, J. S. Lee, J. Jeong, H. Kim, J. Byun, S. A. Kim, H. J. Lee, H. S. Chung, J. B. Lee and D. R. Ahn, *J Control Release*, 2017, **250**, 27-35.
2. J. A. Zuris, D. B. Thompson, Y. Shu, J. P. Guilinger, J. L. Bessen, J. H. Hu, M. L. Maeder, J. K. Joung, Z. Y. Chen and D. R. Liu, *Nat Biotechnol*, 2015, **33**, 73-80.
3. M. Jinek, K. Chylinski, I. Fonfara, M. Hauer, J. A. Doudna and E. Charpentier, *Science*, 2012, **337**, 816-821.
4. H. Kim, Y. Park and J. B. Lee, *Sci Rep*, 2015, **5**, 12737.
5. Y. Jeon, D. Kim, J. V. Martin-Lopez, R. Lee, J. Oh, J. Hanne, R. Fishel and J. B. Lee, *Proc Natl Acad Sci U S A*, 2016, **113**, 3281-3286.
6. E. K. Brinkman, T. Chen, M. Amendola, B. van Steensel, *Nucleic Acids Res.* 2014, **42**, e168