
Supplementary information

Structural basis for pegRNA-guided reverse transcription by a prime editor

In the format provided by the
authors and unedited

Supplementary Information for

Structural basis for pegRNA-guided reverse transcription by prime editor

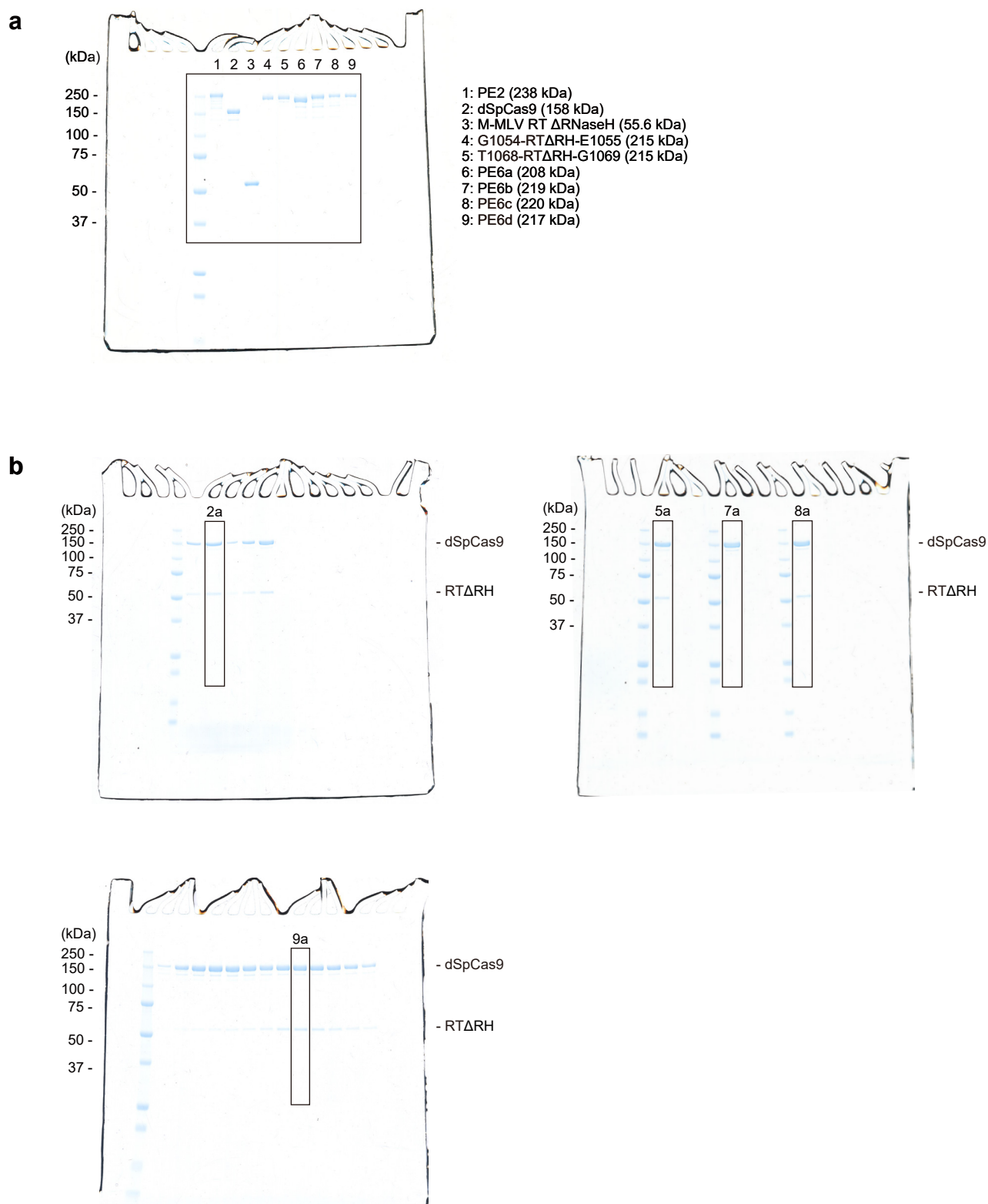
Yutaro Shuto^{1,7}, Ryoya Nakagawa^{1,7*}, Shiyong Zhu^{2,3,4,5,6}, Mizuki Hoki¹, Satoshi N. Omura¹,
Hisato Hirano¹, Yuzuru Itoh¹, Feng Zhang^{2,3,4,5,6}, and Osamu Nureki^{1*}

*Correspondence: ryoya.nakagawa@bs.s.u-tokyo.ac.jp, nureki@bs.s.u-tokyo.ac.jp

This PDF file includes:

Supplementary Figs. 1, 2, 3

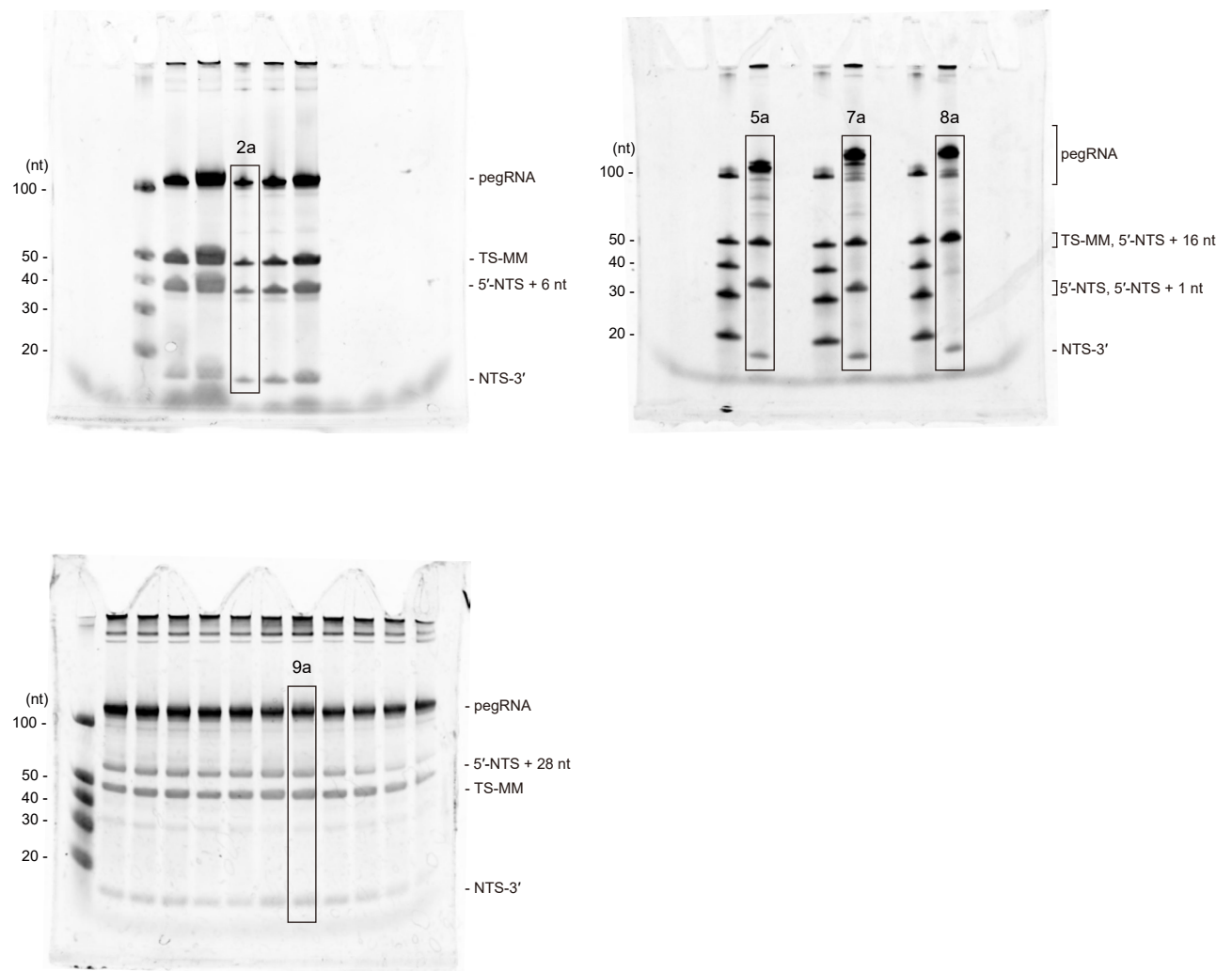
Supplementary Tables 1, 2, 3



Supplementary Fig. 1. Uncropped SDS-PAGE gel images in this study.

a, SDS-PAGE gel of purified proteins (used for Extended Data Fig. 1a).

b, SDS-PAGE gels of prime editor complex reconstitutions. After size-exclusion chromatography, the peak fraction was analysed (used for Extended Data Fig. 2a, 5a, 7a, 8a, and 9a).



Supplementary Fig. 2. Uncropped urea-PAGE gel images of prime editor complex reconstitutions. After size-exclusion chromatography, the peak fraction was analysed. The gels were visualized by SYBR Gold staining (used for Extended Data Fig. 2a, 5a, 7a, 8a, and 9a).

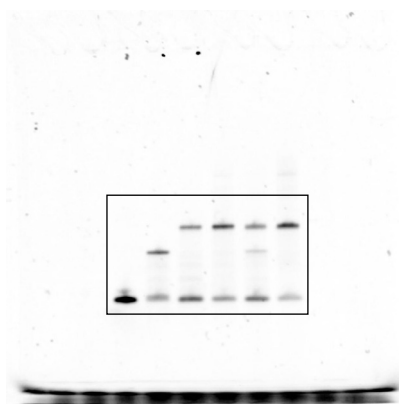


Fig. 1c

Products
- (RTT template)
- Products
(DNA template)
- Substrates

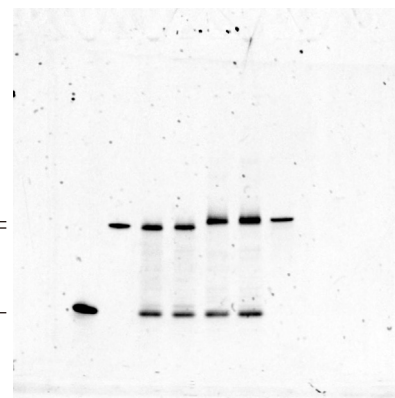


Fig. 3c

65-nt
62-nt
34-nt

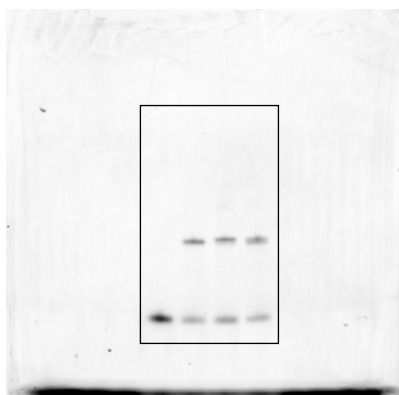
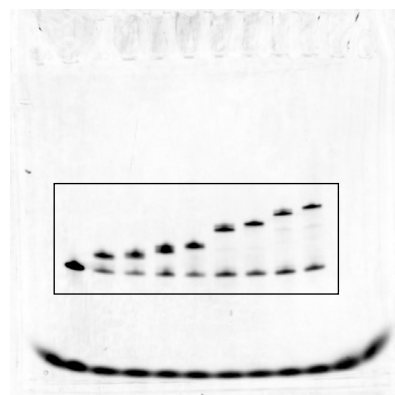


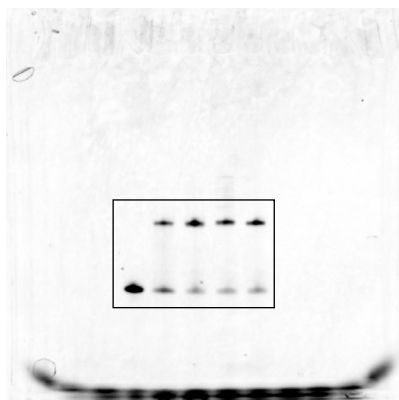
Fig. 4h

Products
- Substrates



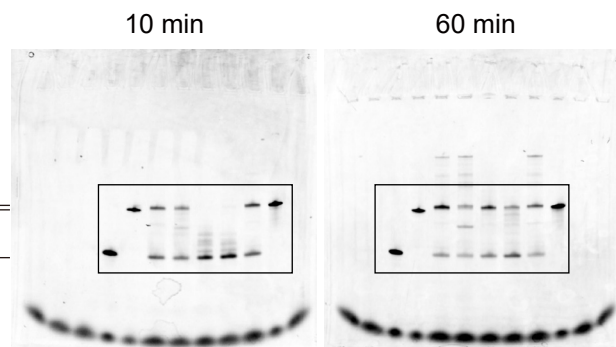
Extended Data Fig. 4a

Products
- Substrates



Extended Data Fig. 4d

Products
- Substrates



10 min

60 min

65-nt
62-nt
34-nt

Products
- Substrates

Extended Data Fig. 10f

Supplementary Fig. 3. Uncropped urea-PAGE gel images of *in vitro* prime editing assay.

After the reactions with Cy5-labeled substrates were stopped by the addition of quench buffer, the reaction products were separated on urea-PAGE gels. The gels were visualized by Cy5 fluorescence (used for Fig. 1c, 3c, and 4h and Extended Data Fig. 4a, 4d, and 10f).

Supplementary Table 1. Nucleic-acid sequences used in prime editing assay.

In vitro prime editing assay	
Target DNA	
Name	Sequence
TS	CTAGTACTCTGCCATC <u>AC</u> GTGCTCAGTCTGGGCCGATTGCTCTAAGTGATC
TS-MM	CTAGTACTCTGCCATCAGAGCAAGCACTACGGCCGATTGCTCTAAGTGATC
5'-Cy5-NTS	GATCACTTAGAGCAATCGGCCAGACTGAGCAGC
NTS-3'	TGATGGCAGAGTACTAG
pegRNA	
Name	Sequence
pegRNA	<u>GGCCCAGACUGAGC</u> ACGUGAguuuuagagcuagaaauagcaaguuaaaauaaggcuaguccguuuaucaacuugaaaaaguggcaccgagucg gugcUGGAGGAAGCAGGGCAACCAACCACAGCGUGCUCAGUCUG
pegRNA-MM	<u>GGCCCGUAGUGCUUG</u> CGUCUGAguuuuagagcuagaaauagcaaguuaaaauaaggcuaguccguuuaucaacuugaaaaaguggcaccgagucg gugcUGGAGGAAGCAGGGCAACCAACCACAGCGUGCUCAGUCUG
pegRNA-6nt	<u>GGCCCAGACUGAGC</u> ACGUGAguuuuagagcuagaaauagcaaguuaaaauaaggcuaguccguuuaucaacuugaaaaaguggcaccgagucg gugcUCACAGCGUGCUCAGUCUG
pegRNA-10nt	<u>GGCCCAGACUGAGC</u> ACGUGAguuuuagagcuagaaauagcaaguuaaaauaaggcuaguccguuuaucaacuugaaaaaguggcaccgagucg gugcUGGAGGAAGCAGGGCAACCAACCACAGCGUGCUCAGUCUG
pegRNA-20nt	<u>GGCCCAGACUGAGC</u> ACGUGAguuuuagagcuagaaauagcaaguuaaaauaaggcuaguccguuuaucaacuugaaaaaguggcaccgagucg gugcUGGAGGAAGCAGGGCAACCAACCACAGCGUGCUCAGUCUG
pegRNA-ACG	<u>GGCCCAGACUGAGC</u> ACGUGAguuuuagagcuagaaauagcaaguuaaaauaaggcuaguccguuuaucaacuugaaaaaguggCGUccgaguc ggACGUGGAGGAAGCAGGGCAACCAACCACAGCGUGCUCAGUCUG
pegRNA-AAA	<u>GGCCCAGACUGAGC</u> ACGUGAguuuuagagcuagaaauagcaaguuaaaauaaggcuaguccguuuaucaacuugaaaaaguggUUUccgaguc ggAAAUGGAGGAAGCAGGGCAACCAACCACAGCGUGCUCAGUCUG
pegRNA-GGG	<u>GGCCCAGACUGAGC</u> ACGUGAguuuuagagcuagaaauagcaaguuaaaauaaggcuaguccguuuaucaacuugaaaaaguggCCCccgaguc ggGGGUGGAGGAAGCAGGGCAACCAACCACAGCGUGCUCAGUCUG

The guide sequence in the pegRNAs and the target sequence in the TS are underlined, and the scaffold region in the pegRNA is in lowercase. Moreover, in the lower three modified pegRNAs, the different sequences from the pegRNA are highlighted in bold or uppercase. The RTT sequences are colored yellow. The PBS sequences and their complementary sequences in the NTS are colored pink. The PAM sequences (TGG) in the NTS are colored purple.

In vivo prime editing assay

pegRNA or nicking sgRNA		
Edit		Sequence
HEK3 +1 T to A	pegRNA-WT	<u>GGCCCAGACUGAGC</u> ACGUGAguuuuagagcuagaaauagcaaguuaaaauaaggcuaguccguuuaucaacuugaaaaaguggcaccgagucg gugcUCCUCUGCCAUUCGUGCUCAGUCUG
	pegRNA-mutant	<u>GGCCCAGACUGAGC</u> ACGUGAguuuuagagcuagaaauagcaaguuaaaauaaggcuaguccguuuaucaacuugaaaaaguggAAGccgaguc ggCUUUCUCUGCCAUUCGUGCUCAGUCUG
EMX1 +1 G to C	pegRNA-WT	<u>GAGUCCGAGCAG</u> AAGAAGAAguuuuagagcuagaaauagcaaguuaaaauaaggcuaguccguuuaucaacuugaaaaaguggcaccgagucg gugcGUGAUGGGAGCCCUUGUUCUUCUGCUCGG
	pegRNA-mutant	<u>GAGUCCGAGCAG</u> AAGAAGAAguuuuagagcuagaaauagcaaguuaaaauaaggcuaguccguuuaucaacuugaaaaaguggAUCccgaguc ggGAUUGAUGGGAGCCCUUGUUCUUCUGCUCGG
HEK3 +1 A insert	pegRNA-WT	<u>GGCCCAGACUGAGC</u> ACGUGAguuuuagagcuagaaauagcaaguuaaaauaaggcuaguccguuuaucaacuugaaaaaguggcaccgagucg gugcUCUGGCCAUACGUGCUCAGUCUG
	pegRNA-mutant	<u>GGCCCAGACUGAGC</u> ACGUGAguuuuagagcuagaaauagcaaguuaaaauaaggcuaguccguuuaucaacuugaaaaaguggGGAccgagu cggUCCUCUGGCCAUACGUGCUCAGUCUG
FANCF +5 G to C	pegRNA-WT	<u>GGAAUCCCUUCUG</u> CAGCACCguuuuagagcuagaaauagcaaguuaaaauaaggcuaguccguuuaucaacuugaaaaaguggcaccgagucg gugcGGAAAAGCGAUGCAGAGGUUCUGCAGAAAGGAU
	pegRNA-mutant	<u>GGAAUCCCUUCUG</u> CAGCACCguuuuagagcuagaaauagcaaguuaaaauaaggcuaguccguuuaucaacuugaaaaaguggGAGccgaguc ggCUCGGAAAAGCGAUGCAGAGGUUCUGCAGAAAGGAU
RNF2 +1 C to A	pegRNA-WT	<u>GUCAUCUUA</u> GUCAUUACCUGguuuuagagcuagaaauagcaaguuaaaauaaggcuaguccguuuaucaacuugaaaaaguggcaccgagucg gugcAACGAACACCCUCAGAGUAAUGACUAAGAUG
	pegRNA-mutant	<u>GUCAUCUUA</u> GUCAUUACCUGguuuuagagcuagaaauagcaaguuaaaauaaggcuaguccguuuaucaacuugaaaaaguggGUAccgaguc ggUACAACGAACACCCUCAGAGUAAUGACUAAGAUG
FANCF +5-7 GGA del	pegRNA-WT	<u>GGAAUCCCUUCUG</u> CAGCACCguuuuagagcuagaaauagcaaguuaaaauaaggcuaguccguuuaucaacuugaaaaaguggcaccgagucg gugcGGAAAAGCGAAGGUUCUGCAGAAAGGAU
	pegRNA-mutant	<u>GGAAUCCCUUCUG</u> CAGCACCguuuuagagcuagaaauagcaaguuaaaauaaggcuaguccguuuaucaacuugaaaaaguggGAGccgaguc ggCUCGGAAAAGCGAAGGUUCUGCAGAAAGGAU
	sgRNA	<u>GCCCUACUUC</u> CGCUUUCACCUguuuuagagcuagaaauagcaaguuaaaauaaggcuaguccguuuaucaacuugaaaaaguggcaccgaguc ggugc
FANCF +8 T to C	pegRNA-WT	<u>GGAAUCCCUUCUG</u> CAGCACCguuuuagagcuagaaauagcaaguuaaaauaaggcuaguccguuuaucaacuugaaaaaguggcaccgagucg gugcGGAAAAGCGGUCCAGGUUCUGCAGAAAGGAU
	pegRNA-mutant	<u>GGAAUCCCUUCUG</u> CAGCACCguuuuagagcuagaaauagcaaguuaaaauaaggcuaguccguuuaucaacuugaaaaaguggGAGccgaguc ggCUCGGAAAAGCGGUCCAGGUUCUGCAGAAAGGAU
	sgRNA	<u>GCCCUACUUC</u> CGCUUUCACCUguuuuagagcuagaaauagcaaguuaaaauaaggcuaguccguuuaucaacuugaaaaaguggcaccgaguc ggugc
RNF2 +1 T ins	pegRNA-WT	<u>GUCAUCUUA</u> GUCAUUACCUGguuuuagagcuagaaauagcaaguuaaaauaaggcuaguccguuuaucaacuugaaaaaguggcaccgagucg gugcAACGAACACCCUCAGAGUAAUGACUAAGAUG
	pegRNA-mutant	<u>GUCAUCUUA</u> GUCAUUACCUGguuuuagagcuagaaauagcaaguuaaaauaaggcuaguccguuuaucaacuugaaaaaguggGUAccgaguc ggUACAACGAACACCCUCAGAGUAAUGACUAAGAUG
	sgRNA	<u>UCAACCAUUA</u> AGCAAACAUguuuuagagcuagaaauagcaaguuaaaauaaggcuaguccguuuaucaacuugaaaaaguggcaccgagucg ugc
RNF2 +4 A del	pegRNA-WT	<u>GGAAUCCCUUCUG</u> CAGCACCguuuuagagcuagaaauagcaaguuaaaauaaggcuaguccguuuaucaacuugaaaaaguggcaccgagucg gugcAACGAACACCCAGGUAAUGACUAAGAUG
	pegRNA-mutant	<u>GUCAUCUUA</u> GUCAUUACCUGguuuuagagcuagaaauagcaaguuaaaauaaggcuaguccguuuaucaacuugaaaaaguggGUAccgaguc ggUACAACGAACACCCAGGUAAUGACUAAGAUG
	sgRNA	<u>UCAACCAUUA</u> AGCAAACAUguuuuagagcuagaaauagcaaguuaaaauaaggcuaguccguuuaucaacuugaaaaaguggcaccgagucg ugc

The guide sequence in the pegRNAs is underlined, and the scaffold region in the pegRNA is in lowercase. Moreover, the different sequences in the mutants from their WT are highlighted in bold and uppercase. The RTT sequences are colored yellow. The PBS sequences are colored pink.

Supplementary Table 2. Nucleic-acid sequences used in Cryo-EM analysis.

State	Name	Sequence
Termination	pegRNA-termination-MM	<u>GGCCGUAGUGCUUGCUCUCUGA</u> guuuuagagcuagaaa <u>uagcaagu</u> aaaa <u>uaaggcuaguccguu</u> aucaacuugaaaaaguggcaccgagucg gugc <u>U</u> <u>CACAG</u> <u>CGUGCUCAGUCUG</u>
	5'-NTS+3nt	CTAGTACTCTGCCAT <u>CAGAGCAAGCACTACGGCC</u> GATTGCTCTAAGTGATC
	TS-MM	GATCACTTAGAGCAATCGGCC <u>CAGACTGAGCACGCTG</u>
	NTS-3'	TGA <u>TGG</u> CAGAGTACTAG
Initiation	pegRNA-initiation-MM	<u>GGCCGUAGUGCUUGCUCUCUGA</u> guuuuagagcuagaaa <u>uagcaagu</u> aaaa <u>uaaggcuaguccguu</u> aucaacuugaaaaaguggcaccgagucg gugc <u>U</u> <u>CACAU</u> <u>CGUGCUCAGUCUG</u>
	TS-MM	CTAGTACTCTGCCAT <u>CAGAGCAAGCACTACGGCC</u> GATTGCTCTAAGTGATC
	5'-NTS	GATCACTTAGAGCAATCGGCC <u>CAGACTGAGCACG</u>
	NTS-3'	TGA <u>TGG</u> CAGAGTACTAG
Pre-initiation	pegRNA-pre-initiation-MM	<u>GGCCGUAGUGCUUGCUCUCUGA</u> guuuuagagcuagaaa <u>uagcaagu</u> aaaa <u>uaaggcuaguccguu</u> aucaacuugaaaaaguggcaccgagucg gugc <u>U</u> <u>GGAGGAAGCAGUGCAACCAAACCA</u> <u>CAGCGUGCUCAGUCUG</u>
	TS-MM	CTAGTACTCTGCCAT <u>CAGAGCAAGCACTACGGCC</u> GATTGCTCTAAGTGATC
	5'-NTS	GATCACTTAGAGCAATCGGCC <u>CAGACTGAGCACG</u>
	NTS-3'	TGA <u>TGG</u> CAGAGTACTAG
Elongation (16 nt)	pegRNA-elongation-16nt-MM	<u>GGCCGUAGUGCUUGCUCUCUGA</u> guuuuagagcuagaaa <u>uagcaagu</u> aaaa <u>uaaggcuaguccguu</u> aucaacuugaaaaaguggcaccgagucg gugc <u>U</u> <u>GGAGGAAGCAGUGCAACCAAACCA</u> <u>CAGCGUGCUCAGUCUG</u>
	TS-MM	CTAGTACTCTGCCAT <u>CAGAGCAAGCACTACGGCC</u> GATTGCTCTAAGTGATC
	5'-NTS+3nt	GATCACTTAGAGCAATCGGCC <u>CAGACTGAGCACGCTG</u>
	NTS-3'	TGA <u>TGG</u> CAGAGTACTAG
Elongation (28 nt)	pegRNA-elongation-28nt-MM	<u>GGCCGUAGUGCUUGCUCUCUGA</u> guuuuagagcuagaaa <u>uagcaagu</u> aaaa <u>uaaggcuaguccguu</u> aucaacuugaaaaaguggcaccgagucg gugc <u>U</u> <u>GGAGGAAGCAGGGCAACCAAACCA</u> <u>CAGCGUGCUCAGUCUG</u>
	TS-MM	CTAGTACTCTGCCAT <u>CAGAGCAAGCACTACGGCC</u> GATTGCTCTAAGTGATC
	5'-NTS+3nt	GATCACTTAGAGCAATCGGCC <u>CAGACTGAGCACGCTG</u>
	NTS-3'	TGA <u>TGG</u> CAGAGTACTAG

The guide sequence in the pegRNAs and the target sequence in the TS are underlined, and the scaffold region in the pegRNA is in lowercase. The RTT sequences are colored in yellow, with uracil in the RTT highlighted in bold except for the pegRNA-pre-initiation-MM. The PBS sequences and their complementary sequences in the NTS are colored pink. The reverse transcription sequences pre-added to the NTS are colored in blue. The PAM sequences (TGG) in the NTS are colored purple.

Supplementary Table 3. Oligonucleotides and constructs used in this study.

Mutants			
Mutation		Forward primer	Reverse primer
D10A		GCCATCGGCACCAACTCTGTGGGC	CAGGCCGATGCTGTACTTCTTGTACCTCCA
G1054-RTΔRH-E1055	insertion	ACCCTGGCCAAACGGCTCCGGGACAGGGGGTACC CTAAATATAGAAGATGAGTATCGGCTACATG	CCGCTTCCGGATCTCTCCGACCCCCGTTCCAAGG CAGTTGTGTTGCAGCCCTTC
	vector	GAGATCCGGAAGCGGCCTCTGATCGAGAC	GCCGTTGGCCAGGGTAATCTCGGTCTTG
T1068-RTΔRH-G1069	insertion	ACAAACGGCGAAACCTCCGGGACAGGGGGTACCC TAAATATAGAAGATGAGTATCGGCTACATG	CCACACGATCTCCCCTCCGACCCCCGTTCCAAGG CAGTTGTGTTGCAGCCCTTC
	vector	GGTTTCGCCGTTTGTCTCGATCAGAGGCCGCTTC	GCCGTTGGCCAGGGTAATCTCGGTCTTG
Constructs			
Plasmid			
His-SUMO-PE2	https://benchling.com/s/seq-NSua7gucllgrqjbUgTH?m=slm-f2zy3CFR9yplWc0x6aoS		
His-SUMO-SpCas9 D10A/H840A	https://benchling.com/s/seq-PKZ29WQG5SLvXMUFvTH4?m=slm-j9XhGzQ0hVi92MwJrd86		
His-SUMO-RTΔRH	https://benchling.com/s/seq-5spl3SO3h9lYxs3bsWFa?m=slm-YahBQRVNgYQS8mvqXRtr		
His-SUMO-G1054-RTΔRH-E1055	https://benchling.com/s/seq-OyRwnDSdsOuam6NtNa3G?m=slm-w3oQgNCixA32chdwWUgy		
His-SUMO-T1068-RTΔRH-G1069	https://benchling.com/s/seq-To7DIVFp4fkK7Jaa4z6P?m=slm-wW8hCIEjSSDEHseFwMil		
His-SUMO-PE6a	https://benchling.com/s/seq-ljRzTiRgBo3mwWaxiPrw?m=slm-q6276dfnucvtCsm1LVrE		
His-SUMO-PE6b	https://benchling.com/s/seq-lWyrqDfQLRmgT8x3tA?m=slm-1dsj1kE73NEwo2rv6S7G		
His-SUMO-PE6c	https://benchling.com/s/seq-7cXy6MwkHCLeGRG8U495?m=slm-c43MX30JwkzXID5TCUtQ		
His-SUMO-PE6d	https://benchling.com/s/seq-0pmF3oRQxiEX6avdostb?m=slm-FFHeOcVK2GMAUUAbhi4A		
In vitro transcription templates			
Forward template	Sequence		
T7promoter-Fw	GGATCCTAATACGACTCACTATA		
Reverse template	Sequence		
pegRNA-Rv	CAGACTGAGCACGCTGTGGTTGGTTGCCCTGCTTCCTCCAgcaccgactcggtgccacttttcaagttgataacggactagccttatttta acttgctatttctagctctaaaaacTCACGTGCTCAGTCTGGGCCCTATAGTGAGTCGTATTAGGATCC		
pegRNA-MM-Rv	CAGACTGAGCACGCTGTGGTTGGTTGCCCTGCTTCCTCCAgcaccgactcggtgccacttttcaagttgataacggactagccttatttta acttgctatttctagctctaaaaacTCAGAGCAAGCACTACGGCCCTATAGTGAGTCGTATTAGGATCC		
pegRNA-6nt-Rv	CAGACTGAGCACGCTGTGAgcaccgactcggtgccacttttcaagttgataacggactagccttattttaactgctatttctagctctaaaaacTCACGT GCTCAGTCTGGGCCCTATAGTGAGTCGTATTAGGATCC		
pegRNA-10nt-Rv	CAGACTGAGCACGGCTTCCTCCAgcaccgactcggtgccacttttcaagttgataacggactagccttattttaactgctatttctagctctaaaaacTC ACGTGCTCAGTCTGGGCCCTATAGTGAGTCGTATTAGGATCC		
pegRNA-20nt-Rv	CAGACTGAGCACGTGGTTGCCCTGCTTCCTCCAgcaccgactcggtgccacttttcaagttgataacggactagccttattttaactgctatttct agctctaaaaacTCACGTGCTCAGTCTGGGCCCTATAGTGAGTCGTATTAGGATCC		
pegRNA-ACG-Rv	CAGACTGAGCACGCAGAGGAAAGGAAGCCCTGCTTCCTCCAgctccgactcgggacgcacttttcaagttgataacggactagccttatttta aactgctatttctagctctaaaaacTCACGTGCTCAGTCTGGGCCCTATAGTGAGTCGTATTAGGATCC		
pegRNA-AAA-Rv	CAGACTGAGCACGCAGAGGAAAGGAAGCCCTGCTTCCTCCAtttccgactcgggaacacttttcaagttgataacggactagccttatttta actgctatttctagctctaaaaacTCACGTGCTCAGTCTGGGCCCTATAGTGAGTCGTATTAGGATCC		
pegRNA-GGG-Rv	CAGACTGAGCACGCAGAGGAAAGGAAGCCCTGCTTCCTCCAccccgactcgggggacacttttcaagttgataacggactagccttatttta taactgctatttctagctctaaaaacTCACGTGCTCAGTCTGGGCCCTATAGTGAGTCGTATTAGGATCC		
pegRNA-termination-MM-Rv	CAGACTGAGCACGCTGTGAgcaccgactcggtgccacttttcaagttgataacggactagccttattttaactgctatttctagctctaaaaacTCAGAG CAAGCACTACGGCCCTATAGTGAGTCGTATTAGGATCC		
pegRNA-initiation-MM-Rv	CAGACTGAGCACGATGTGcaccgactcggtgccacttttcaagttgataacggactagccttattttaactgctatttctagctctaaaaacTCAGAG CAAGCACTACGGCCCTATAGTGAGTCGTATTAGGATCC		
pegRNA-pre-initiation-MM-Rv	CAGACTGAGCACGCTGTGGTTGGTTGCACTGCTTCCTCCTGcaccgactcggtgccacttttcaagttgataacggactagccttatttta acttgctatttctagctctaaaaacTCAGAGCAAGCACTACGGCCCTATAGTGAGTCGTATTAGGATCC		
pegRNA-elongation-16nt-MM-Rv	CAGACTGAGCACGCTGTGGTTGGTTGCACTGCTTCCTCCTGcaccgactcggtgccacttttcaagttgataacggactagccttatttta acttgctatttctagctctaaaaacTCAGAGCAAGCACTACGGCCCTATAGTGAGTCGTATTAGGATCC		
pegRNA-elongation-28nt-MM-Rv	CAGACTGAGCACGCTGTGGTTGGTTGCCCTGCTTCCTCCAgcaccgactcggtgccacttttcaagttgataacggactagccttatttta acttgctatttctagctctaaaaacTCAGAGCAAGCACTACGGCCCTATAGTGAGTCGTATTAGGATCC		

The scaffold region in the pegRNA templates is in lowercase. The sequences complementary to T7promoter-Fw in the reverse templates are underlined.