SUPPLEMENTAL EXPERIMENTAL PROCEDURES

Sample Preparation

The gene encoding full-length SaCas9 (residues 1–1053) was cloned between the *Nde*I and *Xho*I sites of the modified pE-SUMO vector (LifeSensors), and the N580A/C946A mutations were introduced by a PCR-based method. We mutated a non-conserved cysteine residue (Cys946) for crystallization, since solvent-exposed cysteine residues may hamper crystallization, and the mutation of two cysteine residues indeed improved the solution behavior of SpCas9, as previously described (Nishimasu et al., 2014). SaCas9 contains three cysteine residues (Cys237, Cys534 and Cys946), among which Cys237 and Cys534, but not Cys946, are conserved among close orthologs. The SaCas9 N580A/C946A mutant protein was expressed at 20°C in *Escherichia coli* Rosetta 2 (DE3) (Novagen), and was purified by chromatography on Ni-NTA Superflow resin (QIAGEN). The eluted protein was incubated overnight at 4°C with TEV protease to remove the His_6-SUMO-tag, and was further purified by chromatography on Ni-NTA, Mono S (GE Healthcare) and HiLoad Superdex 200 16/60 (GE Healthcare) columns. The SeMet-labeled SaCas9 N580A/C946A mutant protein was expressed in *E. coli* B834 (DE3) (Novagen), and was purified using a similar protocol to that for the native protein. The 73-nt sgRNA was transcribed *in vitro* with T7 RNA polymerase, using a double-stranded DNA template, which was generated using pairs of oligonucleotides. The transcribed RNA was purified by 8% denaturing (7 M urea) polyacrylamide gel electrophoresis. The 28-nt target and 8-nt non-target DNA strands were purchased from Sigma-Aldrich. The purified SaCas9 protein was mixed with the sgRNA, the target DNA strand and the non-target DNA strand (containing either the 5′-TTGAAT-3′ PAM or the 5′-TTGGGT-3′ PAM) (molar ratio, 1:1.5:2.3:3.4), and then the SaCas9–sgRNA–target DNA complex was purified by gel filtration chromatography on a Superdex 200 Increase column (GE Healthcare), in buffer consisting of 10 mM Tris-HCl, pH 8.0, 150 mM NaCl and 1 mM DTT. For *in vitro* cleavage assays, the wild-type SaCas9 protein was prepared using a similar protocol to that for the N580A/C946A mutant, except that the size-exclusion chromatography step was omitted.

Crystallography

The purified SaCas9–sgRNA–target DNA complex (containing either the 5′-TTGAAT-3′ PAM or the 5′-TTGGGT-3′ PAM) was crystallized at 20°C, by the hanging-drop vapor diffusion method.

Crystals were obtained by mixing 1 µl of complex solution ($A_{260\ nm}$ = 15) and 1 µl of reservoir solution (10–12% PEG 4,000, 0.75 M NaCl, 0.15 M Na_2HPO_4 and 0.15 M NaN_3). The SeMet-labeled complex (containing the 5′-TTGGGT-3′ PAM) was crystallized under similar conditions. X-ray diffraction data were collected at 100 K on beamlines BL32XU and BL41XU at SPring-8 (Hyogo, Japan). The crystals were cryoprotected in reservoir solution supplemented with 25% ethylene glycol. X-ray diffraction data were processed using XDS (Kabsch, 2010) and AIMLESS (Evans and Murshudov, 2013). The structure was determined by the Se-SAD method, using PHENIX AutoSol (Adams et al., 2010). The model was automatically built using Buccaneer (Cowtan, 2006), followed by manual model building using COOT (Emsley and Cowtan, 2004) and structural refinement using PHENIX (Adams et al., 2010). The final models of the 5′-TTGAAT-3′ PAM complex (2.6 Å resolution) and the 5′-TTGGGT-3′ PAM complex (2.7 Å resolution) were refined using native data sets.

In vitro Cleavage Assay

In vitro plasmid DNA cleavage experiments were performed, essentially as described previously (Anders et al., 2014). The *Eco*RI-linearized pUC119 plasmid (150 ng, 7 nM), containing the 20-nt target sequence and the 5′-TTGAAT-3′ PAM, was incubated at 37°C for 1 h with the SaCas9–sgRNA complex (8, 16, 32 nM), in 10 µl of reaction buffer, containing 20 mM HEPES, pH 7.5, 100 mM KCl, 2 mM $MgCl_2$, 1 mM DTT and 5% glycerol. Reaction products were resolved on an ethidium bromide-stained 1% agarose gel, and then visualized using a Typhoon FLA 9500 scanner (GE Healthcare).

Cell Culture and Transfection

Human embryonic kidney 293FT (Life Technologies) cells were maintained in Dulbecco’s modified Eagle’s Medium (DMEM), supplemented with 10% FBS and 2 mM GlutaMAX, in incubators at 37°C with a 5% CO_2 supply.

Around 24 h prior to transfection, cells were seeded into 24-well plates (Corning) at a density of 2.5×10^5 cells/well, and transfected at 70–80% confluency using Lipofectamine 2000 (Life Technologies), according to the manufacturer’s recommended protocol. A total of 600 ng DNA was used for each well of the 24-well plate.

Detection and Quantification of Genomic Modification

About 72 h after transfection, genomic DNA was extracted using QuickExtract DNA Extraction Solution (Epicentre). Briefly, pelleted cells were resuspended in QuickExtract solution and incubated at 65°C for 15 min, 68°C for 15 min, and 98°C for 10 min (Cong et al., 2013).

Indel analyses by the SURVEYOR assay and targeted deep sequencing were performed as described previously (Cong et al., 2013). Briefly, the targeted genomic region was amplified using primers for the SURVEYOR assay or next-generation sequencing. For the SURVEYOR assay, the purified PCR product was re-annealed, subjected to SURVEYOR nuclease digestion, and then analyzed and quantified by polyacrylamide gel electrophoresis (4–20% TBE PAGE gel, Life Technologies). For deep sequencing, the purified PCR products were amplified again to attach Illumina sequencing adapters and barcodes, and then subjected to sequencing analysis using the Miseq sequencing system (Illumina).

Transcriptional Activation Assay

HEK 293FT cells were seeded into 24- or 96-well plates (Bio-coat, Corning) one day prior to transfection, at densities of 2.5×10^5 cells/well or 2.5×10^4 cells/well, respectively. Cells were transfected using Lipofectamine 2000 (Life Technologies), according to the manufacturer's recommended protocol. For 24-well plates, we used 600 ng total plasmid DNA, containing 50 ng of reporter plasmid, per well. For 96-well plates, we used 250 ng total plasmid DNA, containing 10 ng of reporter plasmid, per well. For the reporter assay, the transcriptional activation reporter was co-transfected into 293FT cells with plasmids carrying the SaCas9 activator system in 24-well plates. The MS2-p65-HSF1 plasmids were previously reported (Konermann et al., 2015). For endogenous gene transcriptional activation, a total of 600 ng of plasmid DNA, containing the SaCas9 activation system, was transfected in the 96-well format.

Flow Cytometry Analysis of Reporter Activation

Reporter activation levels were measured by flow cytometry, using an LSRFortessa cell analyzer (BD Biosciences). Cells were trypsinized from their culture plates approximately 48 h after transfection, and re-suspended in 200 µl of PBS for the flow cytometry analysis. The flow cytometry data were analyzed using FlowJo (FlowJo LLC). At least 10,000 events were analyzed for each transfection sample. The fold induction of reporter gene expression was determined

by a flow cytometry analysis of *mKate2* expression mean intensity over that from the control group, which expressed the sgRNA bearing a non-binding guide. The results were further normalized by the fluorescent intensity measured with a control plasmid that expresses GFP (spiked into the DNA mix for all experiments), to control for transfection differences between each experiment.

Endogenous Gene Activation Assay

HEK 293FT cells were seeded in 96-well plates, and transfected as described in the previous section. Transfected cells were incubated for at least 48 h, and then harvested for RNA extraction, cDNA synthesis, and quantitative real-time PCR (qRT-PCR) analysis to measure the human *ASCL1* and *MYOD1* gene expression levels, using TaqMan Gene Expression Assays (Life Technologies), as described previously (Konermann et al., 2015).

Construction of Split and Inducible SaCas9

Split-SaCas9 fragments were constructed by Gibson Assembly (NEB). SaCas9 fragments and fragments for the inducible dimers FRB/FKBP and ABI/PYL were amplified as previously described (Konermann et al., 2015; Zetsche et al., 2015). DmC was amplified from the iDimerize plasmid system (Takara Clontech).

Genome Cleavage Test of Split and Inducible SaCas9

HEK 293FT cells were plated in 24-well plates, as described in the previous sections. Cells were transfected with 200 ng plasmid DNA for each SaCas9 fragment and 100 ng PCR amplified U6::sgRNA targeting *EMX1*. Dimerization was induced at the time of transfection with 100 nM rapamycin (Sigma) for SaCas9(N)FRB/SaCas9(C)FKBP, 100 nM A/C heterodimerizer (Takara Clontech) for SaCas9(N)DmC/SaCas9(C)FKBP, and 200 μM abscisic acid (Sigma) for SaCas9(N)ABI/SaCas9(C)PYL. DNA was harvested three days after transfection and analyzed with the SURVEYOR nuclease assay, as described previously (Cong et al., 2013).

Construct Sequences

1. NLS-dSaCas9 (D10A/N580A)-linker-NLS-VP64

Underlined: NLS sequences

Italic: linker

Underlined and Italic: VP64 activator domain

MAPKKKRKVGIHGVPAAKRNYILGLAIGITSVGYGIIDYETRDVIDAGVRLFKEANVENNEGRRSKRGAR
RLKRRRRHRIQRVKKLLFDYNLLTDHSELSGINPYEARVKGLSQKLSEEEFSAALLHLAKRRGVHNVNEV
EEDTGNELSTKEQISRNSKALEEKYVAELQLERLKKDGEVRGSINRFKTSDYVKEAKQLLKVQKAYHQLD
QSFIDTYIDLLETRRTYYEGPGEGSPFGWKDIKEWYEMLMGHCTYFPEELRSVKYAYNADLYNALNDLNN
LVITRDENEKLEYYEKFQIIENVFKQKKKPTLKQIAKEILVNEEDIKGYRVTSTGKPEFTNLKVYHDIKD
ITARKEIIENAELLDQIAKILTIYQSSEDIQEELTNLNSELTQEEIEQISNLKGYTGTHNLSLKAINLIL
DELWHTNDNQIAIFNRLKLVPKKVDLSQQKEIPTTLVDDFILSPVVKRSFIQSIKVINAIIKKYGLPNDI
IIELAREKNSKDAQKMINEMQKRNRQTNERIEEIIRTTGKENAKYLIEKIKLHDMQEGKCLYSLEAIPLE
DLLNNPFNYEVDHIIPRSVSFDNSFNNKVLVKQEEASKKGNRTPFQYLSSSDSKISYETFKKHILNLAKG
KGRISKTKKEYLLEERDINRFSVQKDFINRNLVDTRYATRGLMNLLRSYFRVNNLDVKVKSINGGFTSFL
RRKWKFKKERNKGYKHHAEDALIIANADFIFKEWKKLDKAKKVMENQMFEEKQAESMPEIETEQEYKEIF
ITPHQIKHIKDFKDYKYSHRVDKKPNRELINDTLYSTRKDDKGNTLIVNNLNGLYDKDNDKLKKLINKSP
EKLLMYHHDPQTYQKLKLIMEQYGDEKNPLYKYYEETGNYLTKYSKKDNGPVIKKIKYYGNKLNAHLDIT
DDYPNSRNKVVKLSLKPYRFDVYLDNGVYKFVTVKNLDVIKKENYYEVNSKCYEEAKKLKKISNQAEFIA
SFYNNDLIKINGELYRVIGVNNDLLNRIEVNMIDITYREYLENMNDKRPPRIIKTIASKTQSIKKYSTDI
LGNLYEVKSKKHPQIIKKGSA*GGGGSGGGGSGGGGSG*PKKKRKVAAA*GSGRADALDDFDLDMLGSDALDD*
FDLDMLGSDALDDFDLDMLGSDALDDFDLDMLINAS

2. MS2-linker-NLS-p65-linker-HSF1

Underlined: NLS sequences

Italic: linker

MASNFTQFVLVDNGGTGDVTVAPSNFANGVAEWISSNSRSQAYKVTCSVRQSSAQKRKYTIKVEVPKVAT
QTVGGVELPVAAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYSA*GGGGSGGG*
*GSGGGGSG*PKKKRKVAAAGSPSGQISNQALALAPSSAPVLAQTMVPSSAMVPLAQPPAPAPVLTPGPPQS
LSAPVPKSTQAGEGTLSEALLHLQFDADEDLGALLGNSTDPGVFTDLASVDNSEFQQLLNQGVSMSHSTA
EPMLMEYPEAITRLVTGSQRPPDPAPTPLGTSGLPNGLSGDEDFSSIADMDFSALLSQISSSGQ*GGGGS*G
FSVDTSALLDLFSPSVTVPDMSLPDLDSSLASIQELLSPQEPPRPPEAENSSPDSGKQLVHYTAQPLFLL
DPGSVDTGSNDLPVLFELGEGSYFSEGDGFAEDPTISLLTGSEPPKAKDPTVS

3. SaCas9-sgRNA MS2-fusion (TL) scaffold sequence (RNA)

GUUUUAGUACUCUGGGCCAACAUGAGGAUCACCCAUGUCUGCAGGGCCCAGAAUCUACUAAAACAAGGCA
AAAUGCCGUGUUUAUCUCGUCAACUUGUUGGCGAGAU

4. SaCas9-sgRNA MS2-fusion (SL1) scaffold sequence (RNA)

GUUUUAGUACUCUGGAAACAGAAUCUACUAAAACAAGGCAGGCCAACAUGAGGAUCACCCAUGUCUGCAG
GGCCUGCCGUGUUUAUCUCGUCAACUUGUUGGCGAGAU

5. SaCas9-sgRNA MS2-fusion (SL2) scaffold sequence (RNA)

GUUUUAGUACUCUGGAAACAGAAUCUACUAAAACAAGGCAAAAUGCCGUGUUUAUCUCGUCAAGGCCAAC
AUGAGGAUCACCCAUGUCUGCAGGGCCUUGGCGAGAU

6. SaCas9-sgRNA MS2-fusion (TL+SL1) scaffold sequence (RNA)

GUUUUAGUACUCUGGGCCAACAUGAGGAUCACCCAUGUCUGCAGGGCCCAGAAUCUACUAAAACAAGGCA
GGCCAACAUGAGGAUCACCCAUGUCUGCAGGGCCUGCCGUGUUUAUCUCGUCAACUUGUUGGCGAGAU

7. SaCas9-sgRNA MS2-fusion (TL+SL2) scaffold sequence (RNA)

GUUUUAGUACUCUGGGCCAACAUGAGGAUCACCCAUGUCUGCAGGGCCCAGAAUCUACUAAAACAAGGCA
AAAUGCCGUGUUUAUCUCGUCAAGGCCAACAUGAGGAUCACCCAUGUCUGCAGGGCCUUGGCGAGAU

Guide and Primer Sequences

Mammalian sgRNA guide sequences

Gene	Guide target	PAM sequence
DYRK1A	ATGGTTCCTTAAATAAGAACTT	TAGGAT
EMX1-sg1	TGGCCAGGCTTTGGGGAGGCC	TGGAGT
EMX1-sg2	GGCCTCCCCAAAGCCTGGCCA	GGGAGT

SURVEYOR assay primers

Gene	Surveyor primer F	Surveyor primer R
DYRK1A	GGAGCTGGTCTGTTGGAGAA	TCCCAATCCATAATCCCACGTT
EMX1	CCATCCCCTTCTGTGAATGT	GGAGATTGGAGACACGGAGA

Targeted sequencing primers

Gene	NGS primer F (only priming sequence, without adapter)	NGS primer R (only priming sequence, without adapter)
DYRK1A	CTGTTGTGTTGAGTAACATATACCTG	TTGCATGCTGAAGTCTCTCC
EMX1	AAGAAGGGCTCCCATCACAT	AGTGGCCAGAGTCCAGCTT

SaCas9 self-assembly primer

Split 1	**Forward (N-term)** ggtaggcgtgtacggtgggagg	**Reverse (N-term)** cacagtcgaggctgatcagcgagctct aggaattcttaggacaggtccaccttc ttgggc
	Forward (C-term) agcagagctctctggctaactaccggt gccaccATGCAGCAGAAAGAGATCCCC ACCACC	**Reverse (C-term)** cagtcgaggctgatcagcgagc
Split 2	**Forward (N-term)** ggtaggcgtgtacggtgggagg	**Reverse (N-term)** gcacagtcgaggctgatcagcgagctc taggaattcttactcggcctgcttttc ctcgaa
	Forward (C-term) ataagcagagctctctggctaactacc ggtgccaccATGAGCATGCCCGAGATC GAAACC	**Reverse (C-term)** cagtcgaggctgatcagcgagc
Split 3	**Forward (N-term)** ggtaggcgtgtacggtgggagg	**Reverse (N-term)** cacagtcgaggctgatcagcgagctct aggaattcttagatcagggtgttgccc ttgtcg
	Forward (C-term) gcagagctctctggctaactaccggtg ccaccATGGTGAACAATCTGAACGGCC TGTACG	**Reverse (C-term)** cagtcgaggctgatcagcgagc

SaCas9 dimer fusion primer

Dimer fusion	**<u>Forward (N-term)</u>** taagcagagctctctggctaactaccggtgccaccATGAAGCGGAACTACATCCTGGGCC	**<u>Reverse (N-term)</u>** gcacagtcgaggctgatcagcgagctctaggaattcttactcggcctgcttttcctcgaa
	<u>Forward (C-term)</u> ataagcagagctctctggctaactaccggtgccaccATGAGCATGCCCGAGATCGAAACC	**<u>Reverse (C-term)</u>** cagtcgaggctgatcagcgagctctaggaattcttagccctttttgatgatctgagggtg

Table S1. Data Collection and Refinement Statistics, Related to Figure 1.

	5′-TTGAAT-3′ PAM (Native)	5′-TTGGGT-3′ PAM (Native)	5′-TTGGGT-3′ PAM (SeMet)
Data collection			
Beamline	SPring-8 BL41XU	SPring-8 BL41XU	SPring-8 BL41XU
Wavelength (Å)	1.0000	1.0000	0.9791
Space group	$P2_12_12$	$P2_12_12$	$P2_12_12$
Cell dimensions			
a, *b*, *c* (Å)	67.6, 345.6, 98.1	67.6, 345.7, 98.1	67.9, 346.0, 98.3
α, *β*, *γ* (°)	90, 90, 90	90, 90, 90	90, 90, 90
Resolution (Å)*	49.4–2.6 (2.66–2.60)	49.7–2.7 (2.77–2.70)	49.4–3.0 (3.11–3.00)
R_{merge}	0.066 (0.806)	0.106 (1.14)	0.169 (2.00)
R_{pim}	0.028 (0.371)	0.044 (0.468)	0.046 (0.564)
$I/\sigma I$	14.6 (1.9)	10.6 (1.5)	17.8 (2.2)
Completeness (%)	99.4 (92.8)	99.9 (100)	99.9 (99.8)
Multiplicity	6.7 (5.6)	6.7 (6.8)	14.1 (13.4)
CC(1/2)	0.998 (0.833)	0.996 (0.823)	0.999 (0.786)
Wilson *B*-factor (A^2)	53.4	49.2	67.6
Refinement			
Resolution (Å)	49.0–2.6	49.7–2.7	
No. reflections	71,321	64,175	
R_{work}/ R_{free}	0.212 / 0.241	0.203 / 0.230	
No. atoms			
Protein	8,419	8,418	
Nucleic acid	2,312	2,312	
Ion	10	9	
Solvent	93	84	
B-factors ($Å^2$)			
Protein	92.0	81.8	
Nucleic acid	93.6	85.4	
Ion	114	114	
Solvent	60.3	53.8	
R.m.s. deviations			
Bond lengths (Å)	0.0020	0.0030	
Bond angles (°)	0.474	0.597	
Ramachandran plot (%)			
Favored region	96.9	98.1	
Allowed region	3.1	1.9	
Outlier region	0	0	

*Values in parentheses are for the highest resolution shell.

SUPPLEMENTAL REFERENCES

Adams, P.D., Afonine, P.V., Bunkoczi, G., Chen, V.B., Davis, I.W., Echols, N., Headd, J.J., Hung, L.W., Kapral, G.J., Grosse-Kunstleve, R.W., *et al.* (2010). PHENIX: a comprehensive Python-based system for macromolecular structure solution. Acta Crystallogr D Biol Crystallogr *66*, 213-221.

Anders, C., Niewoehner, O., Duerst, A., and Jinek, M. (2014). Structural basis of PAM-dependent target DNA recognition by the Cas9 endonuclease. Nature *513*, 569-573.

Cong, L., Ran, F.A., Cox, D., Lin, S., Barretto, R., Habib, N., Hsu, P.D., Wu, X., Jiang, W., Marraffini, L.A., *et al.* (2013). Multiplex genome engineering using CRISPR/Cas systems. Science *339*, 819-823.

Cowtan, K. (2006). The Buccaneer software for automated model building. Acta Crystallogr D Biol Crystallogr *62*, 1002-1011.

Emsley, P. and Cowtan, K. (2004). Coot: model-building tools for molecular graphics. Acta Crystallogr D Biol Crystallogr *60*, 2126-2132.

Evans, P.R., and Murshudov, G.N. (2013). How good are my data and what is the resolution? Acta Crystallogr D Biol Crystallogr *69*, 1204-1214.

Kabsch, W. (2010). Xds. Acta Crystallogr D Biol Crystallogr *66*, 125-132.

Konermann, S., Brigham, M.D., Trevino, A.E., Joung, J., Abudayyeh, O.O., Barcena, C., Hsu, P.D., Habib, N., Gootenberg, J.S., Nishimasu, H., *et al.* (2015). Genome-scale transcriptional activation by an engineered CRISPR-Cas9 complex. Nature *517*, 583-588.

Nishimasu, H., Ran, F.A., Hsu, P.D., Konermann, S., Shehata, S.I., Dohmae, N., Ishitani, R., Zhang, F., and Nureki, O. (2014). Crystal structure of Cas9 in complex with guide RNA and target DNA. Cell *156*, 935-949.

Zetsche, B., Volz, S.E., and Zhang, F. (2015). A split-Cas9 architecture for inducible genome editing and transcription modulation. Nat Biotechnol *33*, 139-142.