Science Advances

AAAS

advances.sciencemag.org/cgi/content/full/3/7/e1701620/DC1

Supplementary Materials for

Disabling Cas9 by an anti-CRISPR DNA mimic

Jiyung Shin, Fuguo Jiang, Jun-Jie Liu, Nicolas L. Bray, Benjamin J. Rauch, Seung Hyun Baik, Eva Nogales, Joseph Bondy-Denomy, Jacob E. Corn, Jennifer A. Doudna

Published 12 July 2017, *Sci. Adv.* **3**, e1701620 (2017)
DOI: 10.1126/sciadv.1701620

The PDF file includes:

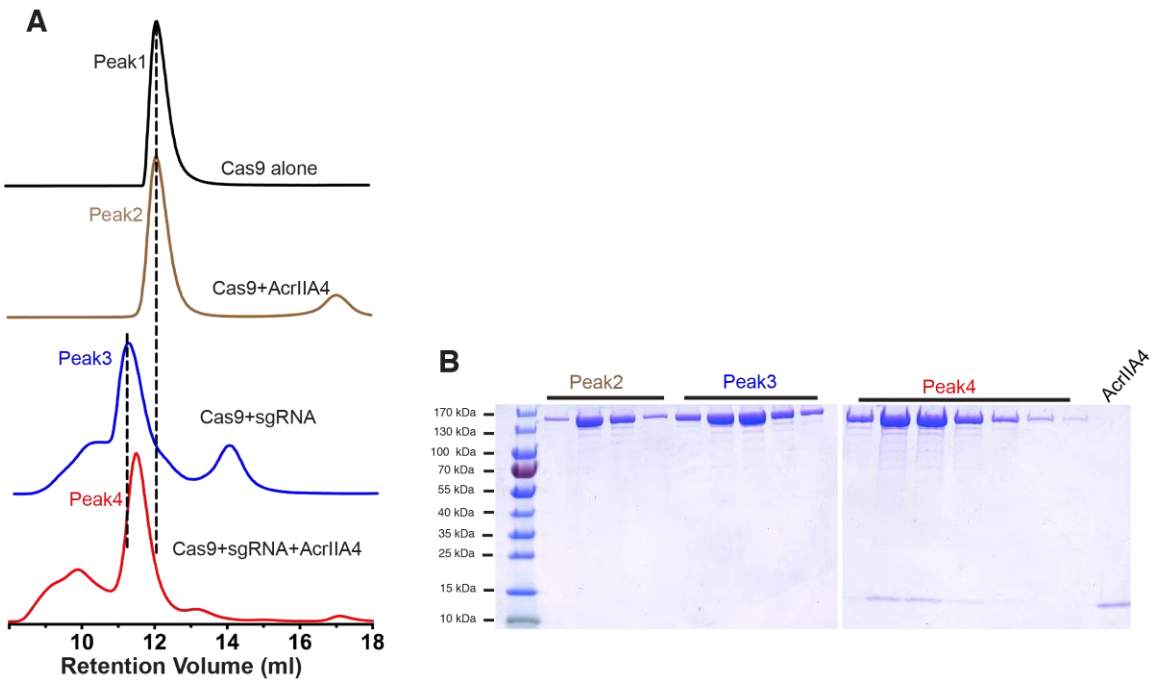

fig. S1. Gel filtration of Cas9 complexes with AcrIIA4. The size-exclusion chromatograph of Cas9 only (black), Cas9 with AcrIIA4 (brown), Cas9-sgRNA complex and Cas9-sgRNA-AcrIIA4 complex were aligned together based on the retention volume from the Supderdex200 column (GE Healthcare). The peaks were marked with dashed lines. (**B**). SDS-PAGE analysis of the fractions corresponding to major peaks in the panel (**A**).

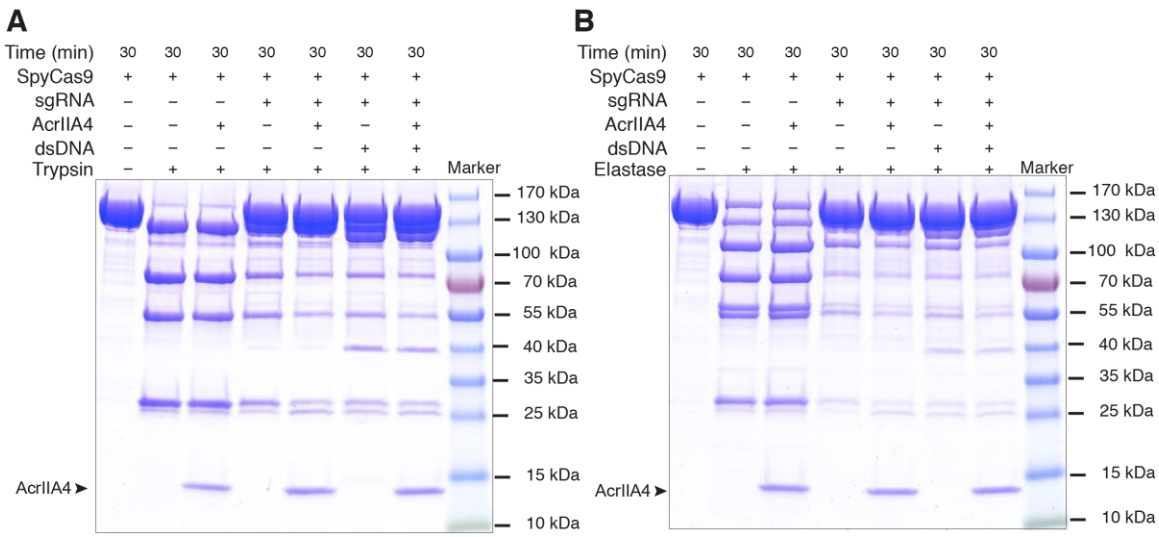

fig. S2. Exposed region analysis of SpyCas9 at AcrIIA4-free and AcrIIA4-bound states. Limited proteolysis of AcrIIA4-free and AcrIIA4-bound Cas9 by trypsin (**A**) and elastase (**B**) for 30 min. Undigested protein served as the zero time point.

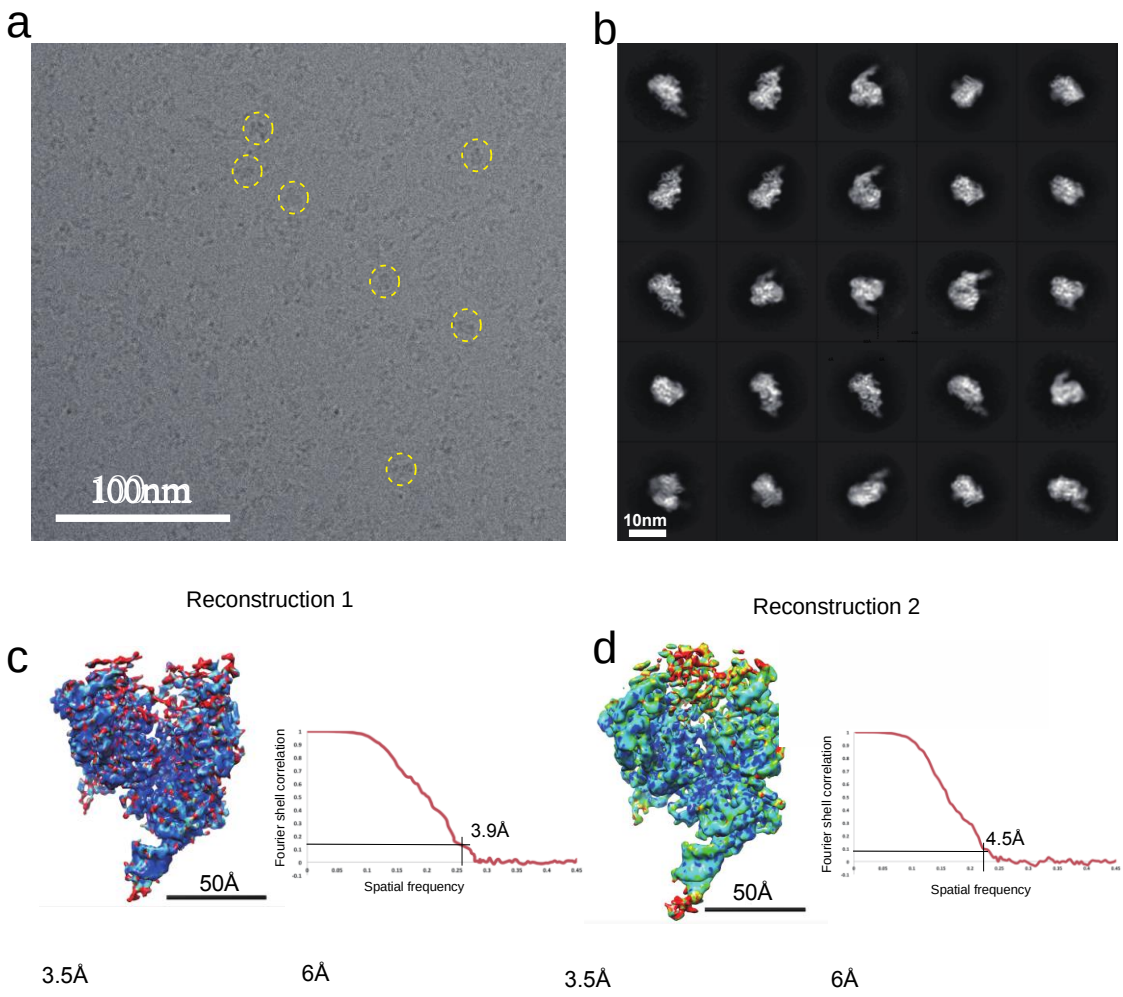

fig. S3. Cryo-EM of Cas9 ribonucleoprotein particles. (**a**) Example of a drift-corrected cryo-EM micrograph of SpyCas9-gRNA-AcrIIA4.Some representative particles are circled in yellow. (**b**) Reference-free 2D class averages showing Cas9 clam-like particles. (**c**) Cryo-EM structure of the SpyCas9-gRNA-AcrIIA4 used for model building shown colored by local resolution (the HNH domain is blurred out due to flexibility). Resolution ranges from 3.5 Å to 5.5 Å. The right panel shows the Fourier shell correlation (FSC) curve calculated using two independent half maps, indicating an overall resolution of 3.9Å (based on the 0.143 FSC criterion) (**d**) Cryo-EM structure of the SpyCas9-gRNA-AcrIIA4 obtained by sub-classification and showing a visible HNH domain (see EM_S2 figure for details), colored by local resolution ranges from 4 Å to 6 Å. Right panel, FSC curve indicating an overall resolution 4.5Å.

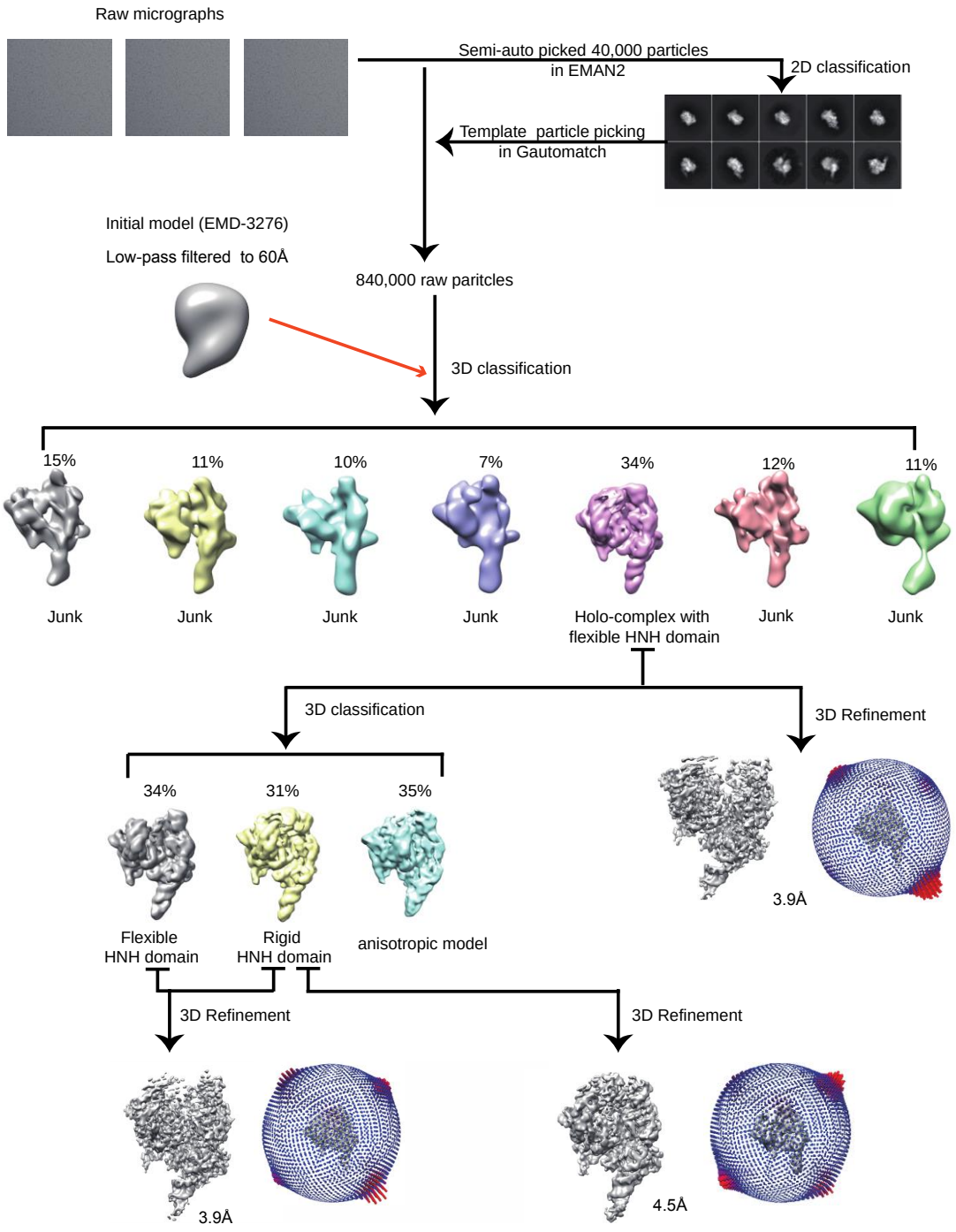

fig. S4. Classification and refinement workflow. A small stack of ~40,000 particles was subjected to 2D reference-free alignment. The resulting 2D class-averages were then used as templates for automatic particle picking. A total dataset of ~840,000 particles were ultimately picked for 3D classification into 7 classes, with our previous cryo-EM

Cas9 structure (EMD-3276) low-pass filtered to 60 Å used as a reference. The class with the most particles (~34%) was subsequently refined to 3.9Å resolution (reconstruction 1). To further improve the density of the HNH domain, which appears highly flexible, particles were subjected to further 3D classification using 3 classes. One of the subclasses, containing a strong HNH density was further refined and produced a structure at 4.5Å (reconstruction 2). 3D refinement by combining first and second subclasses produced a structure at 3.9 Å that was used for atomic modeling of the AcrIIA4.

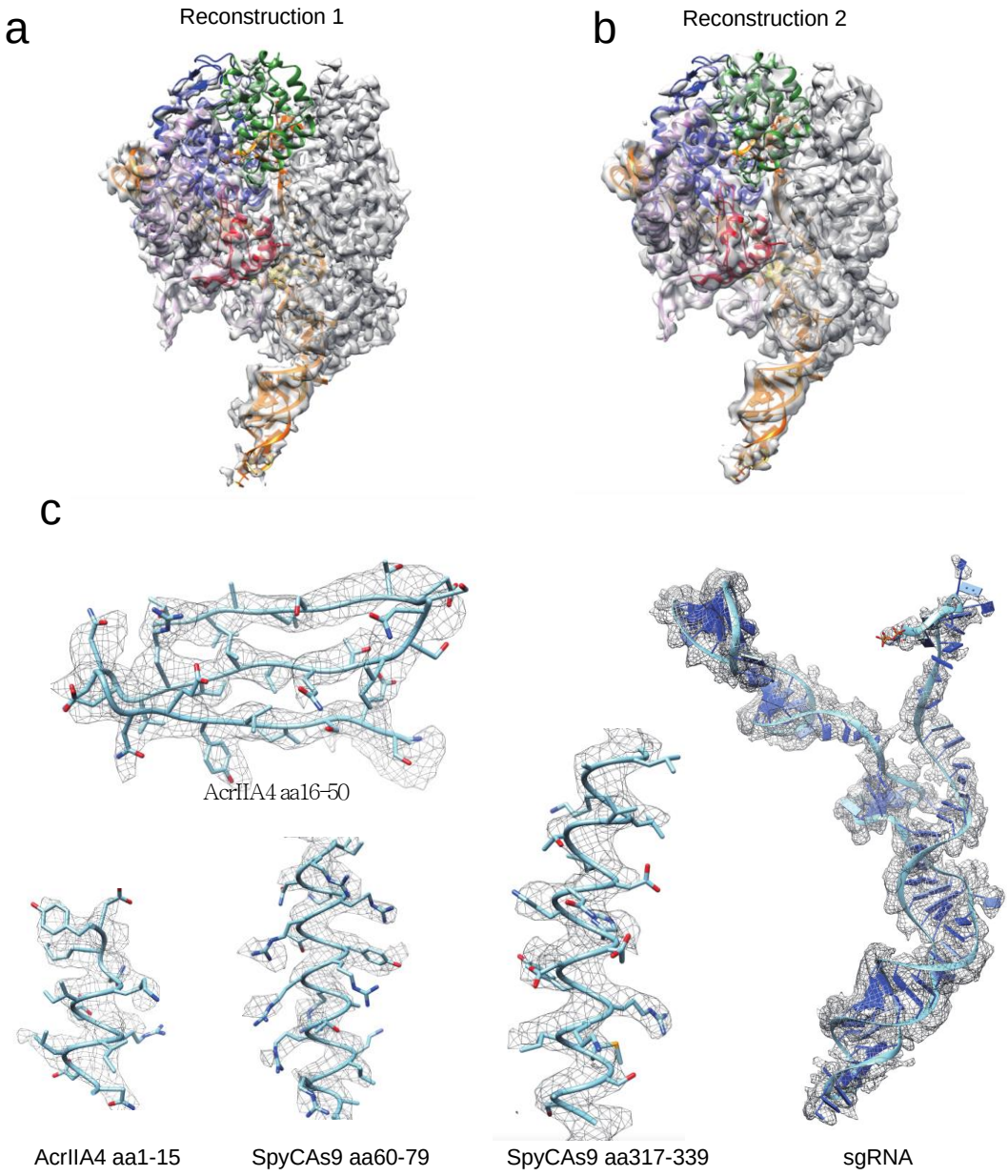

fig. S5. Atomic modeling. (**a**) Color coded atomic model of SpyCas9-sRNA-AcrIIA4 shown within the cryo-EM density of reconstruction 1 (3.9 Å resolution) (a) and reconstruction 2 (4.5 Å resolution) (**b**). The color scheme is the same as Figure 2. The map in (b) has better defined density for the HNH domain corresponding to the green ribbon. (**c**) Representative regions of the EM density map of SpyCas9-sRNA-AcrIIA4 (reconstruction 1), into which the atomic model was built. The subunit and sequence are marked under the corresponding regions.

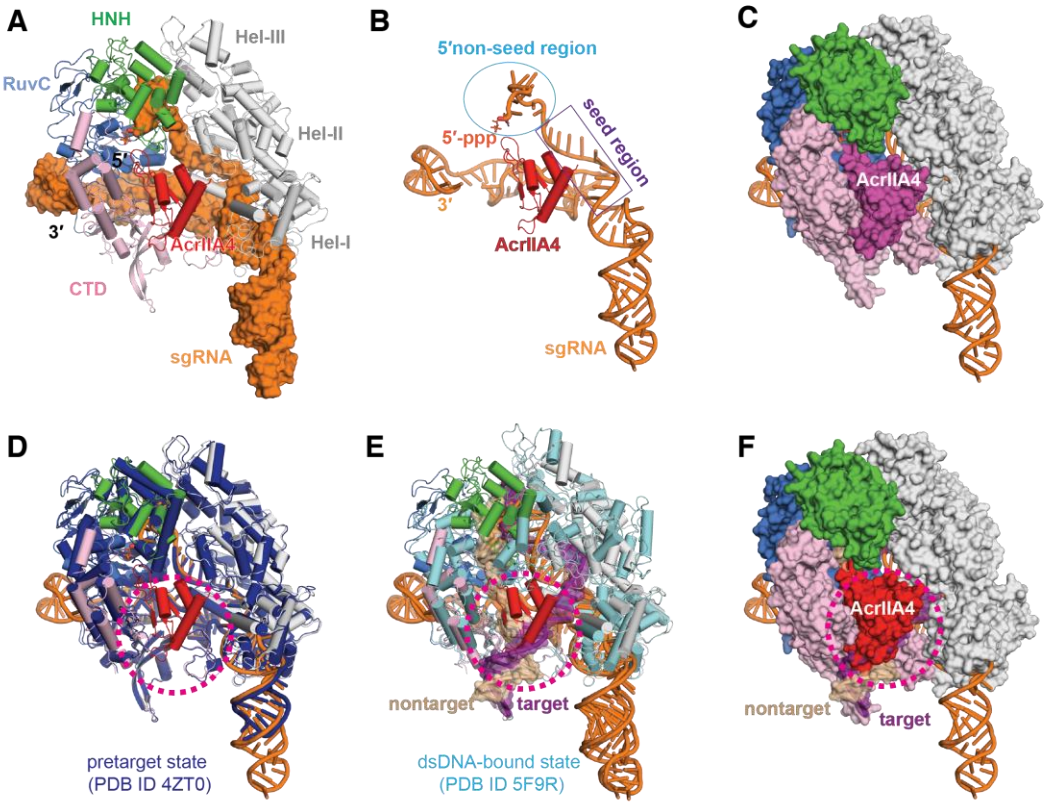

fig. S6. Model comparison between AcrIIA4-bound and DNA- bound SpyCas9-sgRNA complexes. (**A**) Atomic model of Spycas9-sgRNA-AcrIIA4. Domains were color as the same scheme in Fig2. The sgRNA is shown as surface representation and Cas9-AcrIIA4 is shown in ribbon diagram. (**B**) The model for AcrIIA4 and sgRNA were subtracted from the holo-model and shown independently. The seed-region and non-seed region are highlighted in the model. (**C**) The surface representation of Cas9 complexed with AcrIIA4. (**D**) The published atomic model for Cas9-sgRNA (PDB 4ZT0, colored by deep blue) was aligned with Cas9-sgRNA-AcrIIA4 model. (**E**) The published atomic model for dsDNA bound Cas9-sgRNA (PDB 5F9R, colored cyan for Cas9, orange for sgRNA, purple for target DNA strand and beige for nontarget strand) was aligned with Cas9-sgRNA-AcrIIA4 atomic model. Pink dash-lined circle represents the position of PAM-recognition cleft. (**F**) The published atomic model for dsDNA bound Cas9-sgRNA (PDB 5F9R) was aligned with the surface density of Cas9-sgRNA-AcrIIA4 atomic model. For clarity, only both DNA strand are shown with Cas9-sgRNA from 5F9R omitted. AcrIIA4 completely blocks PAM recognition cleft.

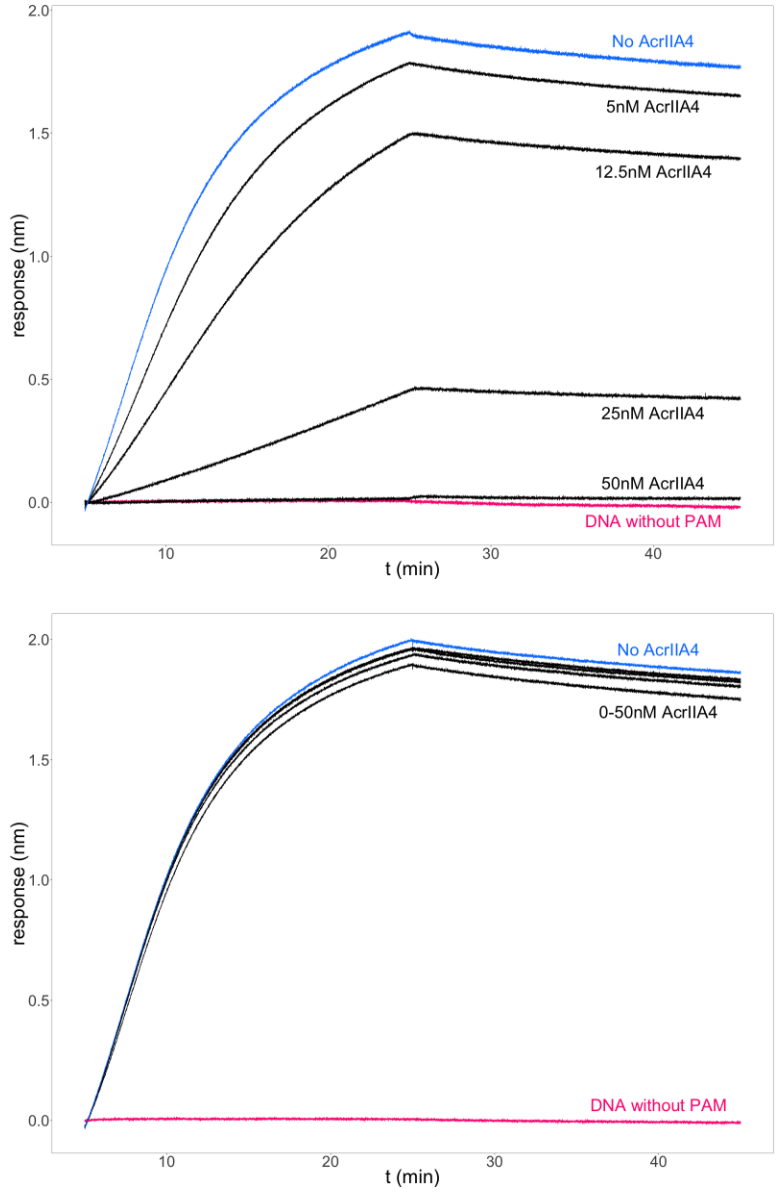

fig. S7. Biological replicate data for BLI data shown in Fig. 3C. (Top) Pre-incubation of dCas9-sgRNA with AcrIIA4 inhibits DNA binding. (Bottom) Addition of AcrIIA4 after allowing dCas9-sgRNA to bind DNA.

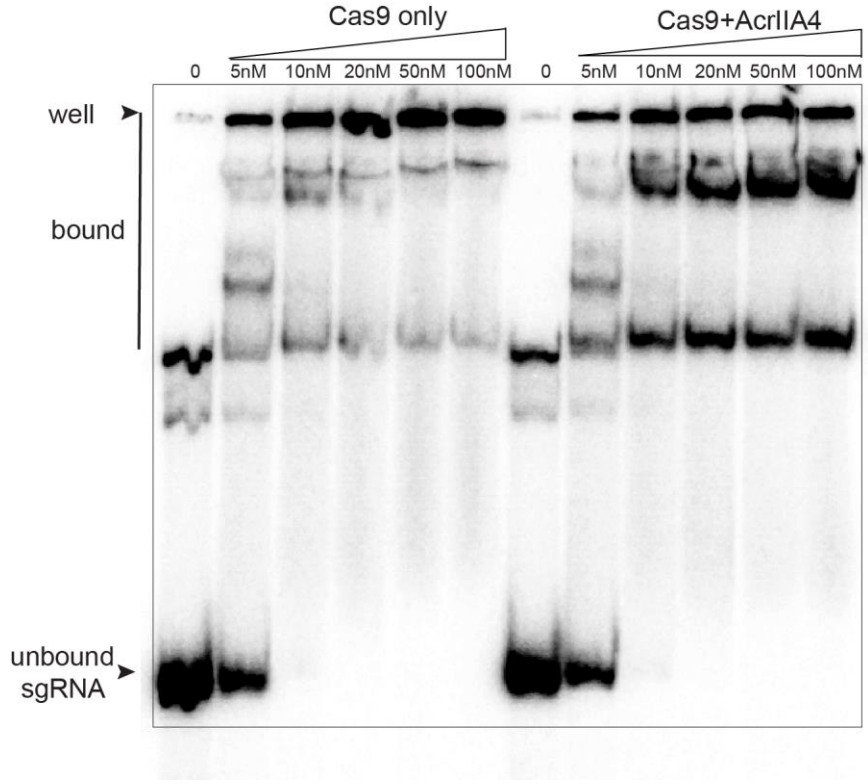

fig. S8. EMSA of increasing concentrations of Cas9 binding to sgRNA in the absence or presence of AcrIIA4. AcrIIA4 does not block sgRNA binding.

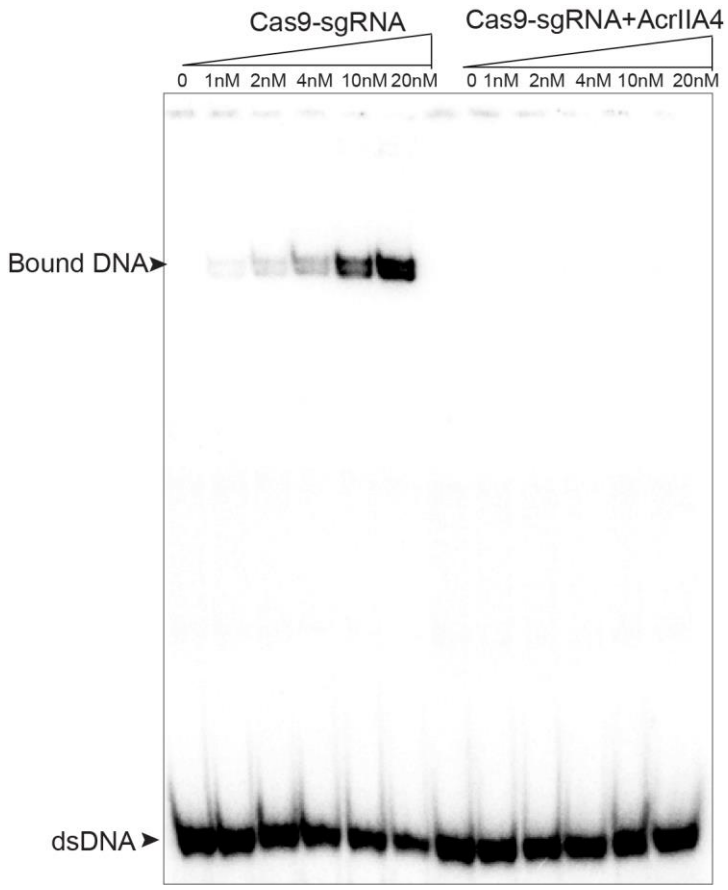

fig. S9. EMSA of Cas9-sgRNA binding to a target DNA with and without AcrIIA4.
The AcrIIA4 inhibitor prevents Cas9-sgRNA binding to target DNA.

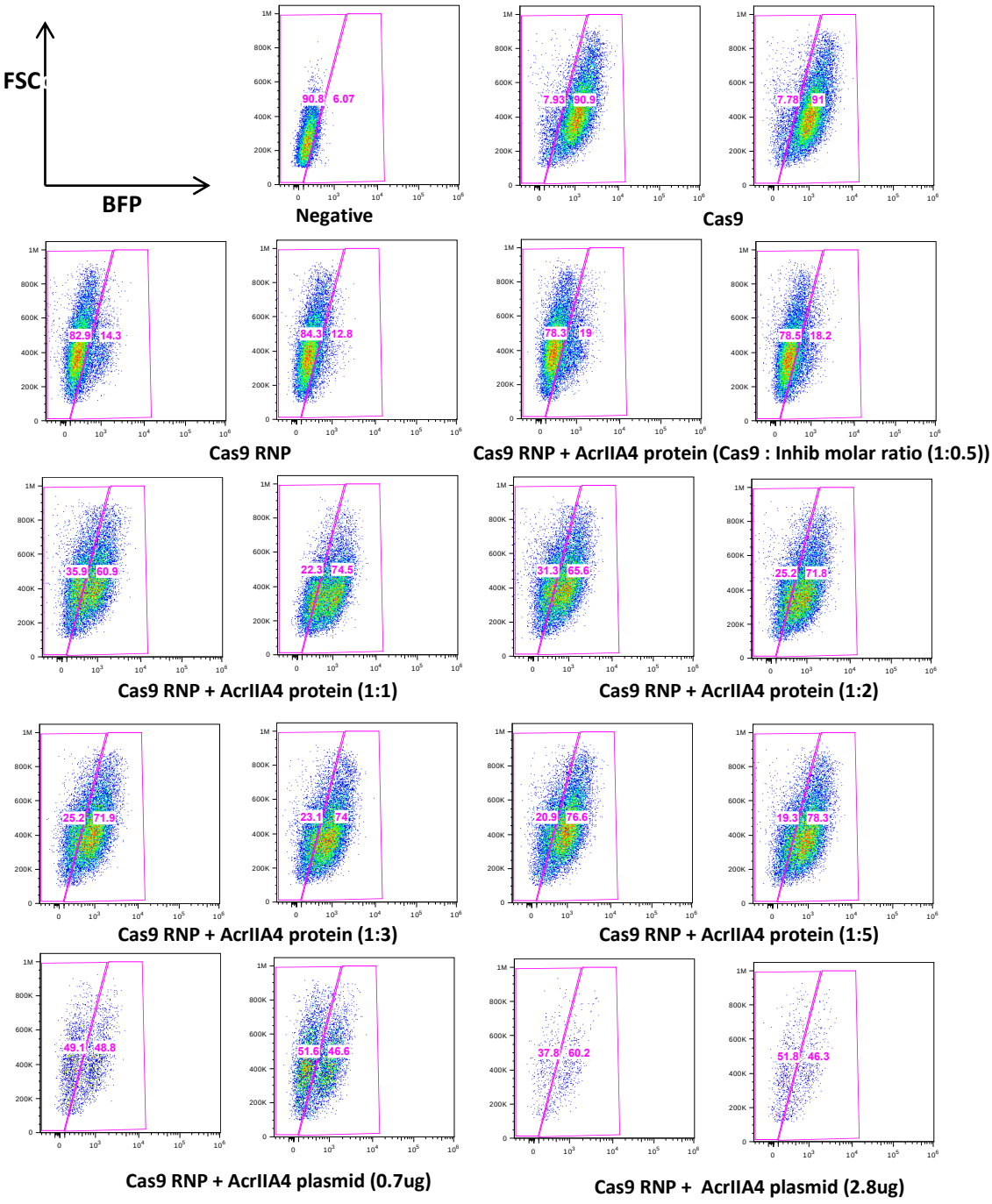

fig. S10. Representative flow cytometry data used to create Fig. 4A.

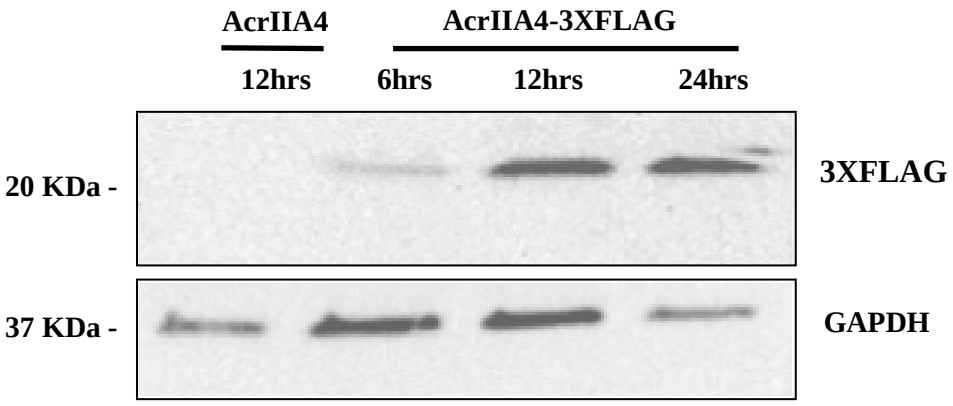

fig. S11. Western blot of AcrIIA4-3XFLAG expression. The expression level of AcrIIA4-3XFLAG was determined by Western blot 6,12,and 24 hours after plasmid nucleofection. Glyceraldehyde 3-phosphate dehydrogenase (GAPDH) was used as a loading control.

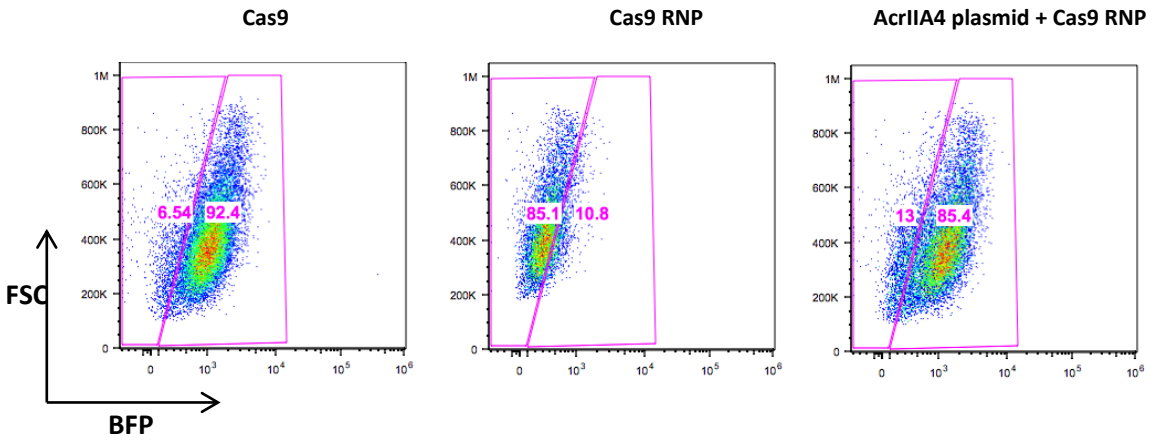

fig. S12. Representative flow cytometry data used to create the graph shown in Fig. 4B.

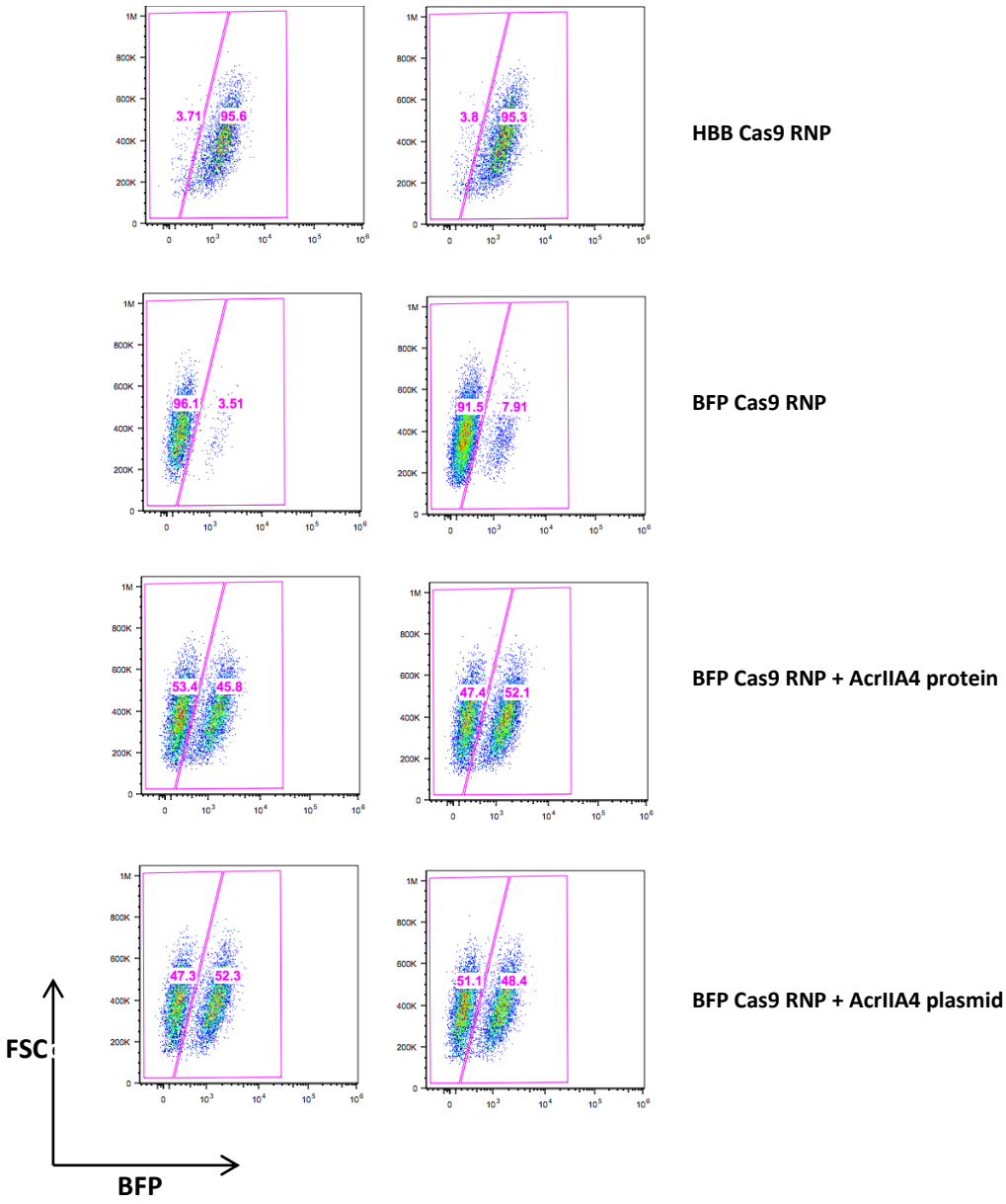

fig. S13. Representative flow cytometry data used to create the graph shown in Fig. 4C.

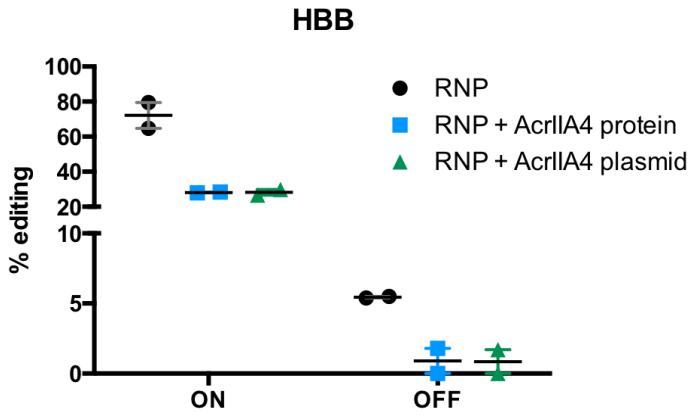

fig. S14. Quantification of on- and off-target editing at HBB, as measured by TIDE analysis.

table S1. Data collection and model refinement statistics.

Data collection and model refinement statistics of SpyCas9-sgRNA-AcrIIA4	
Data Collection	
EM	Titan Krios 300kV, K2 Gatan Summit
Pixel size (Å)	1.04
Defocus range (μm)	–1.5 to –3.0
Reconstruction (Relion)	Relion2.0
Accuracy of rotations (°)	2.485
Accuracy of transitions (pixel)	1.310
Final resolution (Å)	3.9
Refinement (Phenix)	
Map CC (whole unit cell)	0.895
Map CC (around atoms)	0.829
R.m.s. deviations	
Bond lengths (Å)	0.01
Bond angles (°)	1
Ramachandran plot	
% favoured	91.91
%allowed	8.09
% outliers	0.00
Molprobity	
Clashscore	2.5

table S2. Oligonucleotides used in this study.

Type	Name	Sequence
Oligos used for testing inhibition of SpyCas9 activity with radiolaveled target DNA	Alt-R crRNA	/AltR1/rGrA rCrCrC rCrCrU rCrCrA rCrCrC rCrGrC rCrUrC rGrUrU rUrUrA rGrArG rCrUrA rUrGrC rU/AltR2/
	Alt-R tracrRNA	IDT 1072532
	target DNA (complementary strand)	GGT ATG AGG TTT ATT CTT TCC TGA GGC GGG GTG GGG GGG AAT TGC TTT CTG TTT TGT TAA
	target DNA (non-complementary strand)	TTA ACA AAA CAG AAA GCA ATT CCC CCC CAC CCC GCC TCA GGA AAG AAT AAA CCT CAT ACC
Oligos used in BLI assay	target	/5Biosg/agcagaaatctctgctGACGCATAAAGATGAG ACGCTGGagtacaaacgtcagct
	target_rc	agctgacgtttgtactCCAGCGTCTCATCTTTATGCGT Cagcagagatttctgct
	target-PAM	/5Biosg/agcagaaatctctgctGACGCATAAAGATGAG ACGCTCGagtacaaacgtcagct
	target-PAM_rc	agctgacgtttgtactCCAGCGTCTCATCTTTATGCGT Cagcagagatttctgct
	FwdVar oligo	ggatcctaatacgactcactataGACGCATAAAGATGAGA CGCgttttagagctagaaatagc
Assembly of T7 polymerase substrates for sgRNA synthesis by in vitro transcription (UPPERCASE = guide)	FwdVar-HBB	ggatcctaatacgactcactatagCTTGCCCCA CAGGGCAGTAAgttttagagctagaa
	FwdVar-VEGFAs2	ggatcctaatacgactcactatagGACCCCCTCCACCCCGC CTCgttttagagctagaa
	T7RevLong	AAAAAAGCACCGACTCGGTGCCACTTTTTCA AGTTGATAACGGACTAGCCTTATTTTAACTTG CTA TTTCTAGCTCTAAAAC
	T7FwdAmp	GGATCCTAATACGACTCACTATAG
	T7RevAmp	AAAAAAGCACCGACTCGG
Primers for cloning 3X-FLAG AcrIIA4	RB_AcrIIA4 Fwd 1	AACATTAACGACCTCATACG
	RB_AcrIIA4 RV1	GTTCAGTTCACTTTTCAACG
	RB_AcrIIA	CGAGCTCGGATCCGCCACCATGAAC

	4 Fwd3	
	RB_AcrIIA 4 RV3 3XFLAG	ATATCTGCAGAATTCTTATCACTTGTCGTCGT CGTCCTTGTAGTCGATGTCGTGGTCCTTGTA GTCACCGTCGTGGTCCTTGTAGTCGTTCAG TTCACTTTT
Primers for T7E1 and TIDE analysis	HBB on-target F1	TCACTTAGACCTCACCCTGTG
	HBB on-target R1	TATGGGACGCTTGATGTTTTCT
	HBB off-target F1	TGGATAGGAAAGGTGAAGT
	HBB off-target R1	ATATTTGAGAGCCACCGC
Primers for Next Generation Amplicon Sequencing (First PCR)	HBB stubbed on-target F	GCTCTTCCGATCTactgtgttcactagcaacctcaa
	HBB stubbed on-target R	GCTCTTCCGATCTtgggaaaatagaccaataggcagag
	HBB stubbed off-target F	GCTCTTCCGATCTtaccctttcccgttctccac
	HBB off-target with stub R	GCTCTTCCGATCTggtacggcctaagaaattatagttta
	VEGFAs2 stubbed on-target F1	GCTCTTCCGATCTcgacaggggcaaagtgagtgacc
	VEGFAs2 stubbed on-target R1	GCTCTTCCGATCTcctccgaagcgagaacagccc
	VEGFAs2 stubbed off-target F2	GCTCTTCCGATCTgtccaggaacccctagcccaaac
	VEGFAs2 stubbed off-target R2	GCTCTTCCGATCTgctctttcattcgttccatctctcg
	VEGFAs2 stubbed off-target F15	GCTCTTCCGATCTcaagatgtgcacttgggctagc
	VEGFAs2 stubbed off-target R15	GCTCTTCCGATCTcgcagcctattgtctcctgg
	VEGFAs2 stubbed off-target F22	GCTCTTCCGATCTgggctctggggtgactccaag

	VEGFAs2 stubbed off-target R22	GCTCTTCCGATCTgactttcccaagcccatcctccc
Primers for Next Generation Amplicon Sequencing (Second PCR)	Forward primers	AATGATACGGCGACCACCGAGATCTACAC[i5] ACACTCTTTCCCTACACGACGCTCTTCCGATC * T
	Reverse primers	CAAGCAGAAGACGGCATACGAGAT[i7]GTGAC TGGAGTTCAGACGTGTGCTCTTCCGA
	Indexes used	

i7 Index	i7_rc index	i5 Index	
ATAGGTCC	GGACCTAT	ACAGGCAT	
CAGTGCTT	AAGCACTG	ACAGGCAT	
ACCATAGG	CCTATGGT	ACAGGCAT	
CAACTTGG	CCAAGTTG	ACAGGCAT	
GACGAACT	AGTTCGTC	ACAGGCAT	
TCGATGAC	GTCATCGA	ACAGGCAT	
CAAGTCGT	ACGACTTG	ACAGGCAT	
AGTTCGCA	TGCGAACT	ACAGGCAT	
CTGTATGC	GCATACAG	ACAGGCAT	
TCTAGGAG	CTCCTAGA	ACAGGCAT	
TGCTTGCT	AGCAAGCA	ACAGGCAT	
ACGAACGA	TCGTTCGT	ACAGGCAT	
TTCGCCAT	ATGGCGAA	AATGGTCG	
GAGCAATC	GATTGCTC	AATGGTCG	
AACACTGG	CCAGTGTT	AATGGTCG	
CCATGAAC	GTTCATGG	AATGGTCG	
ATAGTCGG	CCGACTAT	AATGGTCG	
GATTGTCC	GGACAATC	AATGGTCG	
TCCACGTT	AACGTGGA	AATGGTCG	
CAACTCCA	TGGAGTTG	AATGGTCG	
TGGTGAAG	CTTCACCA	AATGGTCG	
TGAGCTGT	ACAGCTCA	AATGGTCG	
AACCAGAG	CTCTGGTT	AATGGTCG	
AAGTCCTC	GAGGACTT	AATGGTCG	
GTTCTTCG	CGAAGAAC	TTGCTTGG	
ACAGTGAC	GTCACTGT	TTGCTTGG	
CGCAATGT	ACATTGCG	TTGCTTGG	
AACCGTGT	ACACGGTT	TTGCTTGG	
CCGTTATG	CATAACGG	TTGCTTGG	
CCACAACA	TGTTGTGG	TTGCTTGG	
GGTACGAA	TTCGTACC	TTGCTTGG	
TACTGCTC	GAGCAGTA	TTGCTTGG	
ACCTCTTC	GAAGAGGT	TTGCTTGG	
TGGATGGT	ACCATCCA	TTGCTTGG	
CTTCACTG	CAGTGAAG	TTGCTTGG	

		TACTAGCG	CGCTAGTA	TTGCTTGG	
		CTAGCAGT	ACTGCTAG	AGCCTATC	
		GAGAGTAC	GTACTCTC	AGCCTATC	
		AGGTGTTG	CAACACCT	AGCCTATC	
		TTACCGAC	GTCGGTAA	AGCCTATC	
		ACCGACAA	TTGTCGGT	AGCCTATC	
		CAACGAGT	ACTCGTTG	AGCCTATC	
		AGTATGCC	GGCATACT	AGCCTATC	
		TCACCTAG	CTAGGTGA	AGCCTATC	
		ATCGGAGA	TCTCCGAT	AGCCTATC	
		CTGTACCA	TGGTACAG	AGCCTATC	
		ACAACAGC	GCTGTTGT	AGCCTATC	
		GACTTGTG	CACAAGTC	AGCCTATC	
		ATAGGTCC	GGACCTAT	GGTCGTAT	
		CAGTGCTT	AAGCACTG	GGTCGTAT	
		ACCATAGG	CCTATGGT	GGTCGTAT	
		CAACTTGG	CCAAGTTG	GGTCGTAT	
		GACGAACT	AGTTCGTC	GGTCGTAT	
		TCGATGAC	GTCATCGA	GGTCGTAT	
		CAAGTCGT	ACGACTTG	GGTCGTAT	
		AGTTCGCA	TGCGAACT	GGTCGTAT	
		CTGTATGC	GCATACAG	GGTCGTAT	
		TCTAGGAG	CTCCTAGA	GGTCGTAT	
		TGCTTGCT	AGCAAGCA	GGTCGTAT	
		ACGAACGA	TCGTTCGT	GGTCGTAT	
		TTCGCCAT	ATGGCGAA	CTCCTGAA	
		GAGCAATC	GATTGCTC	CTCCTGAA	
		AACACTGG	CCAGTGTT	CTCCTGAA	
		CCATGAAC	GTTCATGG	CTCCTGAA	
		ATAGTCGG	CCGACTAT	CTCCTGAA	
		GATTGTCC	GGACAATC	CTCCTGAA	
		TCCACGTT	AACGTGGA	CTCCTGAA	
		CAACTCCA	TGGAGTTG	CTCCTGAA	
		TGGTGAAG	CTTCACCA	CTCCTGAA	
		TGAGCTGT	ACAGCTCA	CTCCTGAA	
		AACCAGAG	CTCTGGTT	CTCCTGAA	
		AAGTCCTC	GAGGACTT	CTCCTGAA	
		GTTCTTCG	CGAAGAAC	TGTTCCGT	
		ACAGTGAC	GTCACTGT	TGTTCCGT	
		CGCAATGT	ACATTGCG	TGTTCCGT	
		AACCGTGT	ACACGGTT	TGTTCCGT	
		CCGTTATG	CATAACGG	TGTTCCGT	
		CCACAACA	TGTTGTGG	TGTTCCGT	
		GGTACGAA	TTCGTACC	TGTTCCGT	
		TACTGCTC	GAGCAGTA	TGTTCCGT	

		ACCTCTTC	GAAGAGGT	TGTTCCGT	
		TGGATGGT	ACCATCCA	TGTTCCGT	
		CTTCACTG	CAGTGAAG	TGTTCCGT	
		TACTAGCG	CGCTAGTA	TGTTCCGT	
		CTAGCAGT	ACTGCTAG	CCACATTG	
		GAGAGTAC	GTACTCTC	CCACATTG	
		AGGTGTTG	CAACACCT	CCACATTG	
		TTACCGAC	GTCGGTAA	CCACATTG	
		ACCGACAA	TTGTCGGT	CCACATTG	
		CAACGAGT	ACTCGTTG	CCACATTG	
		AGTATGCC	GGCATACT	CCACATTG	
		TCACCTAG	CTAGGTGA	CCACATTG	
		ATCGGAGA	TCTCCGAT	CCACATTG	
		CTGTACCA	TGGTACAG	CCACATTG	
		ACAACAGC	GCTGTTGT	CCACATTG	
		GACTTGTG	CACAAGTC	CCACATTG	