Table of Contents

Name	Sequence (5' to 3')
S1 target RNA	GCU CCC UAC GCA UGC GUC CC
S1 51 nt RNA	GCU CCC UAC GCA UGC GUC CCG UUU UAG AGC UAG AAA UAG CAA GUU AAA AUA
5'P - 30 nt – propylamine RNA	P'-AGG CUAG UCC GUU AUC AAC UUG AAA AAG UG-(N3-3')
IVT 85 RNA	GCU CCC UAC GCA UGC GUC CCG UUU UAG AGC UAG AAA UAG CAA GUU AAA AUA AGG CUA GUC CGU UAU CAA CUU GAA AAA GUG CUC G
S1 sgRNA (81 NH_2)	GCU CCC UAC GCA UGC GUC CCG UUU UAG AGC UAG AAA UAG CAA GUU AAA AUA AGG CUA GUC CGU UAU CAA CUU GAA AAA GUG-(N3-3')
S1 scrambled sgRNA (81)	<u>ACC GUC GCU CCG CAU CCG U</u>CG UUU UAG AGC UAG AAA UAG CAA GUU AAA AUA AGG CUA GUC CGU UAU CAA CUU GAA AAA GUG-(N3-3')
S1 DNA splint	CGA GCA CTT TTT CAA GTT GAT AAC GGA CTA GCC TTA TTT TAA CTT GCT ATT TCT AGC TCT AAA ACG GGA CGC ATG CGT AGG GAG C
S1 DNA splint scrambled target	CAC TTT TTC AAG TTG ATA ACG GAC TAG CCT TAT TTT AAC TTG CTA TTT CTA GCT CTA AAA CGA C ATG CGG AGC GAC GGT
IVT 85mer DNA template	CCA GTG AAT TCT AAT ACG ACT CAC TAT A**G**C TCC CTA CGC ATG CGT CCC GTT TTA GAG CTA GAA ATA GCA AGT TAA AAT AAG GCT AGT CCG TTA TCA ACT TGA AAA AGT GCT CG
IVT F-primer	CCA GTG AAT TCT AAT ACG ACT CAC TAT A
IVT R-primer	CGAGCACTTTTTCAAGTTGATAA
K31C-F	ATTTCCCAGAACCTTGAAACATTTAGACGGAACCTTATATTCATCAGTGATCACC
K31C-R	GGTGATCACTGATGAATATAAGGTTCCGTCTAAATGTTTCAAGGTTCTGGGAAAT
K31* - F	TATTTCCCAGAACCTTGAACTTAGACGGAACCTTATATTCATC
K31* - R	GATGAATATAAGGTTCCGTCTAAATAGTTCAAGGTTCTGGGAAATA
C80S- F	ATTTGAAAAAATCTCCTGTAGATAGCTAATACGATTCTTCCGACGTGTATAC
C80S - R	GTATACACGTCGGAAGAATCGTATTAGCTATCTACAGGAGATTTTTTCAAAT
C574S - F	CCTGAAATTTCAACACTATCAAAACTTTCTATTTTTTTGAAATAATCTTCTTTTAATTGC
C574S - R	GCAATTAAAAGAAGATTATTTCAAAAAAATAGAAAGTTTTGATAGTGTTGAAATTTCAGG
S1 DNA target -F	GGATTGCTCACCAATTCTCAA
S1 DNA target -R	GCTGGGCTTCATGTTTGATT

Table S1. Oligonucleotides used in this study. S1 target scrambled sequence is underlined for clarity. For the *in vitro* transcribed 85mer, the first nucleotide transcribed is bolded for clarity. (N3-3') = 3'-propylamine

RNA Name	Calculated Mass / Observed Mass (Da)
5'P-30 nt-3'-NH_2 (N3-3')	9,838/ 9,830
5'P-30 nt-3'-Me-maleimide	9,977/ 9,980
5'P-30 nt-3'-Et-maleimide	9,991/ 9,991
5'P-30 nt -3'-Biotin	10,065/ 10,074
5'P-30 nt-3'-DBCO	10,126/ 10,118
5'P-30 nt-3'-Cy5	10,304 / 10,311
5'P-30 nt-3'-iodine	10,007 / 10,004

Table S2. Summary of the identity of 5'P-30 nt part, the calculated mass and observed mass for the RNAs used in this study. Identity was confirmed by MALDI-TOF-MS and calibrated to an external DNA standard 6999 Da.

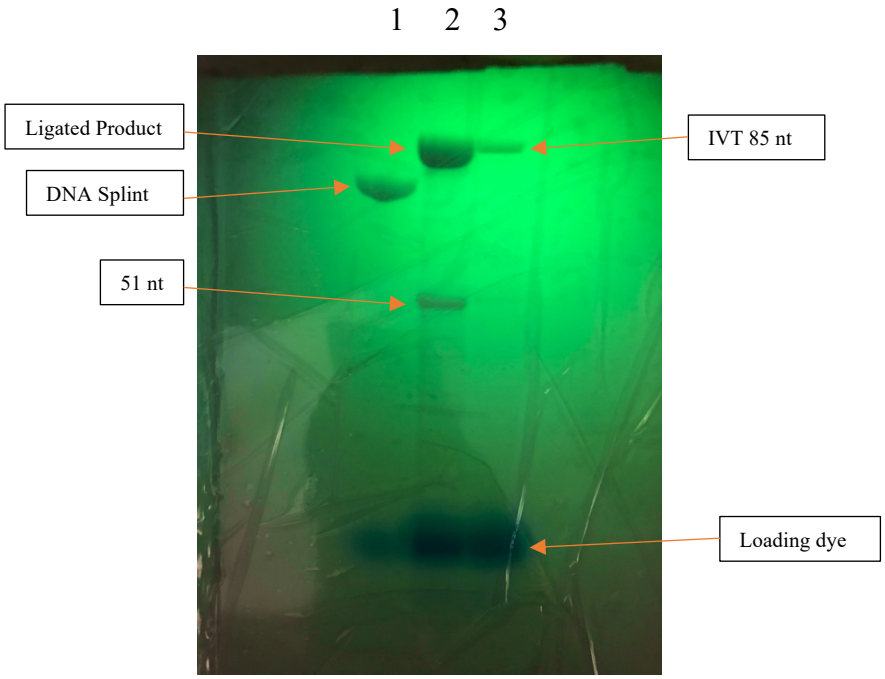

Figure S1. Representative gel image of splint ligation reaction was resolved by 10% denaturing urea PAGE gel. Bands are observed by UV shadow. DNA splint and IVT 85 nt were loaded as running standards. The modified 5'P-30 nt piece was not observed and minimal 51 nt piece was observed. Lane 1 = DNA splint, lane 2 = splint ligation reaction after DNase digestion, lane 3 = IVT 85.

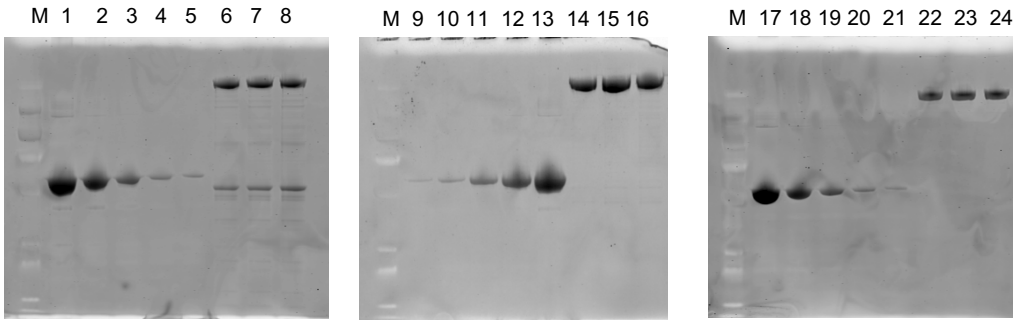

Figure S2. Cas9 protein concentration determination by BSA standards. Cas9 K31pAzF C80S C574S(left) was determined to be 3.35 ug/uL. lane 1 = 10 ug BSA, lane 2 = 5 ug BSA, lane 3 = 2.5 ug BSA, lane 4 = 1 ug BSA, lane 5 = 0.5 ug BSA, lane 6-8 = Cas9 K31pAzF C80S C574S (1 uL purified protein loaded per lane). Cas9 C80S C574S (middle) was determined to be 6.46 ug/uL for a 1:5 dilution. Lane 9 = 0.5 ug BSA, lane 10 = 1 ug BSA, lane 11 = 2.5 ug BSA, lane 12 = 5 ug BSA, lane 13 = 10 ug BSA, lane 14-16 = Cas9 C80S C574S (1 uL of 1:5 dilution loaded per lane). Cas9 K31C C80S C574S (right) was determined to be 3.20 ug/uL for a 1:10 dilution. lane 17 = 10 ug BSA, lane 18 = 5 ug BSA, lane 19 = 2.5 ug, lane 20 = 1 ug BSA, lane 21 = 0.5 ug BSA, lane 22-24 = Cas9 K31C C80S C574S (1 uL of 1:10 dilution loaded per lane).

#	Identified Proteins/Cluster	Accession Number	Alternate ID	Molecular Weight	Band e	Band d	Band c	Band b	Band a
1	Cationic trypsin OS=Bos taurus PE=1 SV=3	TRY1_BOVIN		26 kDa	152	103	114	136	174
2	Cluster of Keratin, type II cytoskeletal 1	K2C1_HUMAN [37]	KTR1	66 kDa	46	18	100	355	85
3	Cluster of Keratin, type I cytoskeletal 10	K1C10_HUMAN [20]	KTR10	59 kDa	19	40	58	245	67
4	Cluster of Nuclease-sensitive element-binding protein 1	H0Y449_HUMAN [5]	YB1	42 kDa	21	202	11	1	0
5	Cluster of Heterogeneous nuclear ribonucleoproteins C1/C2	G3V2Q1_HUMAN [10]	hnRNP C1/C2	34 kDa	228	9	5	3	2
6	Cluster of Keratin, type I cytoskeletal 9	K1C9_HUMAN [2]	KTR9	62 kDa	20	5	42	124	12
7	Cluster of Serum albumin	ALBU_BOVIN [4]	ALB	69 kDa	18	19	24	82	16
8	Cluster of Trypsin-1	E7EQ64_HUMAN [3]	PRSS1	28 kDa	5	1	5	5	9
9	Cluster of Heterogeneous nuclear ribonucleoprotein K	HNRPK_HUMAN [2]	hnRNP K	51 kDa	16	15	126	5	2
10	Cluster of POTE ankyrin domain family member F	POTEF_HUMAN [16]	POTEF	121 kDa	11	7	1	6	0
11	Cluster of Heterogeneous nuclear ribonucleoprotein U	HNRPU_HUMAN [4]	hnRNP U	91 kDa	5	0	1	6	79
12	Cluster of Matrin-3	D6RB45_HUMAN [7]	MATR3	5 kDa	5	0	0	7	90
13	Cluster of Eukaryotic initiation factor 4A-III	IF4A3_HUMAN [9]	EIF4A3	47 kDa	0	50	0	0	0
14	Insulin-like growth factor 2 mRNA-binding protein 1	IF2B1_HUMAN	IGF2BP1	63 kDa	1	5	0	117	0
15	Cluster of RNA-binding motif protein, X chromosome	RBMX_HUMAN [3]	RBMX	42 kDa	1	37	1	0	0
16	Cluster of Polyadenylate-binding protein 1-like	PAP1L_HUMAN [7]	PABPC1L	68 kDa	0	8	9	0	0
17	60S ribosomal protein L3	RL3_HUMAN	RPL3	46 kDa	0	31	0	0	0
18	Heterogeneous nuclear ribonucleoprotein A3	ROA3_HUMAN	hnRNP A	40 kDa	61	0	0	0	0
19	Cluster of Heterogeneous nuclear ribonucleoprotein M	A0A087X0X3_HUMAN [4]	hnRNP M	78 kDa	8	1	7	7	0
20	Cluster of Ribonuclease inhibitor	RINI_HUMAN [4]	RNH1	50 kDa	2	12	1	0	1
21	Cluster of Plastin-3	A0A0A0MSQ0_HUMAN [3	PLS3	69 kDa	0	0	0	31	0
22	Probable ATP-dependent RNA helicase DDX5	DDX5_HUMAN	DDX5	69 kDa	0	0	25	27	0
23	Cluster of Non-POU domain-containing octamer-binding protein	NONO_HUMAN [2]	NONO	54 kDa	0	3	24	0	0
24	Desmoplakin	DESP_HUMAN	DSP	332 kDa	0	0	0	27	0
25	Cluster of Elongation factor 1-alpha	A0A2U3TZH3_HUMAN [5]	EEF1A2	54 kDa	2	10	0	2	0
26	Cluster of T-complex protein 1 subunit theta	TCPQ_HUMAN [2]	CCT8	60 kDa	0	0	45	0	0
27	Cluster of 60 kDa heat shock protein, mitochondrial	CH60_HUMAN [3]	HSPD1	61 kDa	6	5	19	1	0
28	Collagen alpha-1(I) chain	CO1A1_HUMAN	COL1A1	139 kDa	0	0	22	1	0
29	Nucleolin	NUCL_HUMAN	NCL	77 kDa	0	0	11	26	5
30	Elongation factor 1-gamma	EF1G_HUMAN	EEF1G	50 kDa	0	11	0	0	0

Table S3. Top thirty protein candidates from the mass spectroscopy proteomics analysis based on total spectrum counts. Highlighted in yellow were candidates chosen for verification by Western blot analysis. For clarity, boxes are highlighted in orange to indicate sample identity (e.g. YB1 was found in band d)

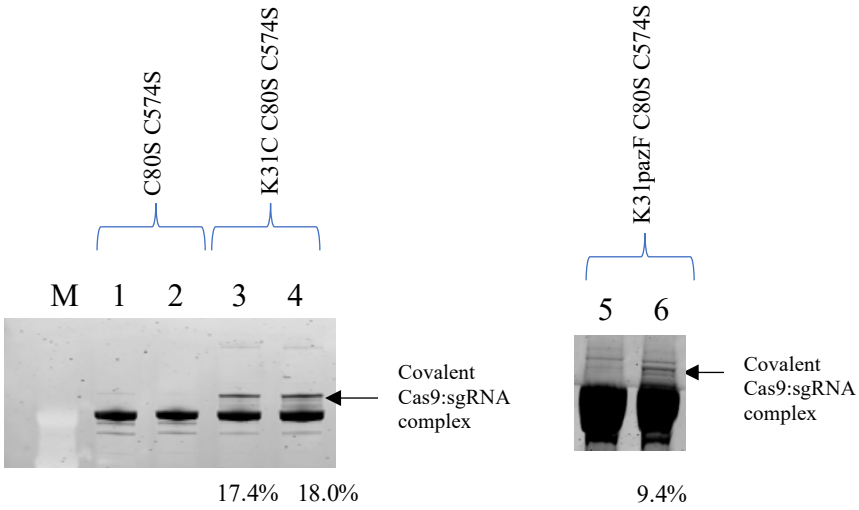

Figure S3. Representative crosslinking experiment. Lane 1 = Cas9 C80S C574S + 81 BMPS, lane 2 = Cas9 C80S C574S + 81 SIA, lane 3 = Cas9 K31C C80S C574S + 81 BMPS, lane 4 = Cas9 K31C C80S C574S + 81 SIA, lane 5 = K31pAzF C80S C574S + IVT 85, lane 6 = K31pAzF C80S C574S + 81 DBCO. Numbers below represent crosslinking percentage from this experiment.

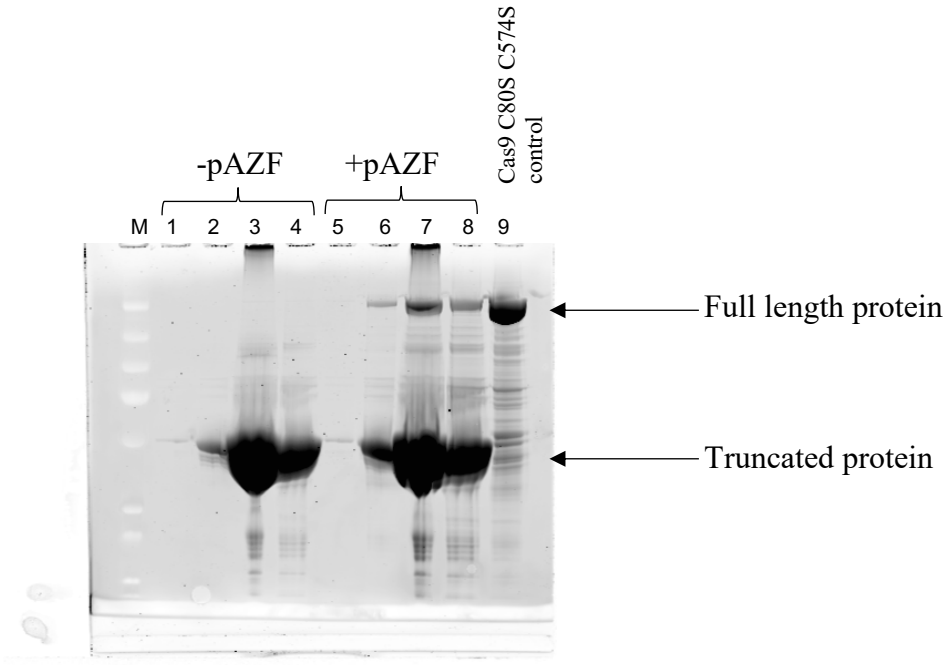

Figure S4. Eluted protein from Ni-NTA column, fractions 1-4 were resolved down SDS-PAGE gel. Full length protein is produced only when the cell culture was supplied with the unnatural amino acid, p-azidophenylalanine. Cas9 C80S C574S sample was used as a positive control. Cas9 K31pAzF C80S C574S was further purified by Hi-Trap SPHP cation exchange column. Lane 1-4 = no pAzF added, fractions 1-4. Lane 5-8 = pAzF added, fractions 1-4. Lane 9 = Cas9 C80S C574S used control

GCAACGCGTTACACGCAGGTGCCGCACTAACACACCCAAGGAGTCCGTCTGCCG
GCTCCAAAGGATTGCTCACCAATTCTCAAAAATAAAGGTATCTGGGTTGTCAAAAAT
CTTAATAAGCCCAAATCTAGAATGTTTTGGTCAAATAGGATGCACTTTCACTACTAG
AATCCACAAGCCCAGCTGACCTTCCTCGCTCCATTGCGTTGCAAATCACTCCTCAG
AACCAAGCGTCTGGCATCTCCGGCTCCCTCTCCTCTCTGCGCTAGTCTTGCCGCA
ATGGCTCAGGTTTGTCGCGCGGCTCCCTACGCATGCGTCCCAGGCAATGGCTGCA
CATGCGCACATTTTTGCCGCAATGCAGGGGTCTCTGTCCCTCTGACCGGAATGTC
CTGCCAAAAGGCAGCTACCAAAAGGAATGCTTGAGCATTCCTACTGCGGAAATTTG
AGCGTGTATTAAAAAAAAAAAAATCACTGTCAATCTGGAATAATGGATATAATTGCC
TATTACTAAATACATTTAATTTTTTAAAAAAATTGCTCCCAAGTGATTGCAAGCATCA
CACAGATTTTTATTATGTGGGAGAGGAAAAAAGGTTTGATCTCCAATATGAGGAAG
ACTGGTTTCTAGGCTATATCATGTTACATAGTGTATCACACACACACACACACACAC
ACACACACACACAATCAAACATGAAGCCCAGCCAAATTGTCAGATACTTAAAATTCT
GTACAGAAAAAAAAGTTATCCTGATTCCTAATTCTACATTCTACATTAATGTAGTATT
ACTTGTATTTTCTTTCTCTGTGTTAGAACATATGA

S1 target sequence. Cyan highlighted sequence represents primers used to amplify the S1 target. Highlighted in green is the S1 target with the PAM site being “AGG”. Cleavage of this target will result in 257 bp and 385 bp products.

Amino acid sequences of proteins used in this publication:

MBP-Cas9 (C80S, C574S)

HHHHHHGKIEEGKLVIWINGDKGYNGLAEVGKKFEKDTGIKVTVEHPDKLEEKFPQVAATGDGPDIIFWA
HDRFGGYAQSGLLAEITPDKAFQDKLYPFTWDAVRYNGKLIAYPIAVEALSLIYNKDLLPNPPKTWEEIPAL
DKELKAKGKSALMFNLQEPYFTWPLIAADGGYAFKYENGKYDIKDVGVDNAGAKAGLTFLVDLIKNKHMN
ADTDYSIAEAAFNKGETAMTINGPWAWSNIDTSKVNYGVTVLPTFKGQPSKPFVGVLSAGINAASPNKEL
AKEFLENYLLTDEGLEAVNKDKPLGAVALKSYEEELAKDPRIAATMENAQKGEIMPNIPQMSAFWYAVRT
AVINAASGRQTVDEALKDAQTNSSSNNNNTSENLYFQGAASMDKKYSIGLDIGTNSVGWAVITDEYKVPS
KKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRISYLQEIFSNEMAKVDDSFFHR
LEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGD
LNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSL
GLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSA
SMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELL
VKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRF
AWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTE
GMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIESFDSVEISGVEDRFNASLGTYHDLLKIIKDK
DFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQS
GKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDEL
VKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQ
NGRDMYVDQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQL
LNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLK
SKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIG
KATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQT
GGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSS
FEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKL
KGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNL
GAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQLGGD*

MBP-Cas9 (K31C, C80S, C574S)

HHHHHHGKIEEGKLVIWINGDKGYNGLAEVGKKFEKDTGIKVTVEHPDKLEEKFPQVAATGDGPDIIFWA
HDRFGGYAQSGLLAEITPDKAFQDKLYPFTWDAVRYNGKLIAYPIAVEALSLIYNKDLLPNPPKTWEEIPAL
DKELKAKGKSALMFNLQEPYFTWPLIAADGGYAFKYENGKYDIKDVGVDNAGAKAGLTFLVDLIKNKHMN
ADTDYSIAEAAFNKGETAMTINGPWAWSNIDTSKVNYGVTVLPTFKGQPSKPFVGVLSAGINAASPNKEL
AKEFLENYLLTDEGLEAVNKDKPLGAVALKSYEEELAKDPRIAATMENAQKGEIMPNIPQMSAFWYAVRT
AVINAASGRQTVDEALKDAQTNSSSNNNNTSENLYFQGAASMDKKYSIGLDIGTNSVGWAVITDEYKVPS
KCFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRISYLQEIFSNEMAKVDDSFFHR
LEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGD
LNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSL
GLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSA
SMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELL
VKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRF
AWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTE
GMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIESFDSVEISGVEDRFNASLGTYHDLLKIIKDK
DFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQS
GKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDEL
VKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQ
NGRDMYVDQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQL
LNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLK
SKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIG
KATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQT
GGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSS
FEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKL
KGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNL
GAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQLGGD*

MBP-Cas9 (K31stop, C80S C574S)

HHHHHHGKIEEGKLVIWINGDKGYNGLAEVGKKFEKDTGIKVTVEHPDKLEEKFPQVAATGDGPDIIFWA
HDRFGGYAQSGLLAEITPDKAFQDKLYPFTWDAVRYNGKLIAYPIAVEALSLIYNKDLLPNPPKTWEEIPAL
DKELKAKGKSALMFNLQEPYFTWPLIAADGGYAFKYENGKYDIKDVGVDNAGAKAGLTFLVDLIKNKHMN
ADTDYSIAEAAFNKGETAMTINGPWAWSNIDTSKVNYGVTVLPTFKGQPSKPFVGVLSAGINAASPNKEL
AKEFLENYLLTDEGLEAVNKDKPLGAVALKSYEEELAKDPRIAATMENAQKGEIMPNIPQMSAFWYAVRT
AVINAASGRQTVDEALKDAQTNSSSNNNNTSENLYFQGAASMDKKYSIGLDIGTNSVGWAVITDEYKVPS
K*FKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRISYLQEIFSNEMAKVDDSFFHRL
EESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDL
NPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSL
GLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSA
SMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELL
VKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRF
AWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTE
GMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIESFDSVEISGVEDRFNASLGTYHDLLKIIKDK
DFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQS
GKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDEL
VKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQ
NGRDMYVDQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQL
LNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLK
SKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIG
KATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQT
GGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSS
FEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKL
KGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNL
GAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQLGGD*

Note: Mutations are shown in red, * = indicates stop codon

Supplemental Experimental Section:

In vitro transcription of 85 nt (IVT 85) sgRNA

For transcription of the 85 nt sgRNA, a T7 RNA polymerase template was constructed by PCR assembly. Briefly, a T7 RNA polymerase substrate template was constructed containing 11 nt prior to the T7 promotor region, followed by the 20 nt *Snurf1* target sequence[1] followed by the RNA scaffold (See Supporting Information, Table 1). To assemble the duplex T7 RNA polymerase template, 250 uL 2x Q5 Hot Start Master Mix (New England Biolabs), 25 uL forward/reverse primers (100 uM), 5 uL (100 uM) single stranded DNA template and 195 uL of H_2O were combined in a single tube then split into 50 uL reactions. PCR products were purified using DNA Clean and Concentrate -50 (Zymo) according to the manufacturer's instructions. To transcribe sgRNAs from the template HiScribe T7 High Yielding RNA Synthesis kit (New England Biolabs) was used according to the manufacturer's instructions for short RNAs. Transcribed sgRNA was purified by 10% denaturing polyacrylamide gel electrophoresis (PAGE) gel, bands were visualized by UV shadowing and excised. To elute the RNA, gel slices were crush and soaked overnight at room temperature in 500 mM NH_4OAc and 100 mM EDTA. Polyacrylamide fragments were removed by 0.2-μM filter (Costar) followed by ethanol precipitation and lyophilized to dryness. RNAs were resuspended in nuclease free water and stored in -20°C. All primers and sgRNA sequences can be found in Supporting Information, Table 1.

Genomic DNA isolation and Amplification of the S1 target

The mouse neuroblastoma cell line Neuro-2a (ATCC#CCL-131) was grown in Dulbecco's modified Eagle's medium (DMEM) supplemented with 10% bovine calf serum (BCS). Neuro-2a cells were grown to 70% confluency and genomic DNA was extracted by Puregene Core kit (Qiagen) according to the manufacturer's protocol. The amplicon for the *Snurf1* target was obtained by amplification of 1 ng of genomic DNA using the S1 DNA -F and S1 DNA -R primers (2 min at 95°C; 10 s at 95°C, 30 s at 64.3°C, 60 s at 72°C, 35 cycles; 10 min at 72°C). Primer sequences can be found in Supporting Information, Table 1.

Overexpression and purification of MBP-Cas9 mutants: The pMJ806 expression vector for the MBP-Cas9 fusion protein was a gift from Prof. Jennifer Doudna, UC Berkeley (AddGene plasmid #39312). The pEVOL-pAzF plasmid for incorporation of azido-phenylalanine was a gift from Prof. Peter Schultz, Scripps (Addgene #31186). Each mutation was introduced using QuikChange Lightning Kit (Agilent) and used to transform XL-10 Ultra Competent Cells (Agilent) according to the manufacturer's protocol. DNA sequences were confirmed by Sanger sequencing. Primer for mutagenesis and sequencing are provided in the Supplemental Information. *E. coli* BL21 STAR (DE3)-competent cells were transformed by electroporation with the plasmid encoding the *S. pyogenes* Cas9 (K31C C80S C574S or C80S C574S) fused to an N-terminal 10×His-tag/maltose binding protein. The resulting expression strain was inoculated in 5 mL Luria-Broth (LB) overnight at 37 °C in 50 ug/mL kanamycin. The cells were diluted in 1 L of LB containing 50 ug/mL kanamycin and grown at 37 °C to an OD = ~0.6-0.8. The culture was induced with IPTG (final concentration 0.5 mM) and incubated 16-20 h at 25 °C. Cells were collected by centrifugation (7500 rpm x 25 min) and resuspended in 30 mL lysis buffer (20 mM Tris-HCl pH 8.0, 500 mM NaCl,

1mM TCEP) supplemented with protease cocktail inhibitor (Roche). Cells were lysed by sonication (6W, 1 s pulse-on/ 1 s pulse-off) for 4 min, cooled on ice for 5 min and then sonicated (6W,1 s pulse-on/ 1 s pulse-off) for 3 min. Clarified cell lysate was obtained by centrifugation (13,000 rpm x 25 min). Cell lysate was batch bound with Ni-NTA resin (3 mL, Qiagen) rocking at 4 °C for 45 min to capture His-tagged MBP-Cas9. The mixture was then transferred to a 25 mL column (Bio-rad) and washed with 50 mL of lysis buffer. Cas9 was eluted in 1.5 mL fractions in 50 mM Tris-HCl pH 8, 0.1 M NaCl, 10% glycerol, 1 mM BME and 300 mM imidazole. Fractions containing MBP-Cas9 were pooled and filtered through a 0.22-micron sterile filter (Millex-HV). MBP-Cas9 was further purified by injecting Cas9 onto a HiTrap SPHP (GE Healthcare) chromatography column. MBP-Cas9 was eluted with a linear gradient 0.1 mM KCl to 1 M KCl containing 20 mM HEPES pH 7.5 and 10% glycerol. Fractions containing MBP-Cas9 were pooled and concentrated with 100 kDa MWCO Amicon Ultra centrifugation units (Millipore) to a volume of ~150 uL. Concentrated Cas9 was supplemented with TCEP (final concentration 1 mM), centrifuged briefly, and frozen in 2 uL aliquots at -70°C. Protein concentration was determined with BSA standards and visualized with SYPRO Orange (Thermo Fischer Scientific) staining on a 4-12% SDS-polyacrylamide gel (Supporting Information, Figure 3). For expression of the K31pAzF mutant, *E. coli* BL21 STAR (DE3)-competent cells were transformed by electroporation expression of with the plasmid encoding Cas9 with the amber stop codon at position K31 (K31TAG(stop) C80S C574S) and the pEVOL-pAzF plasmid and plated on an LB agar plate supplemented with 25 ug/mL kanamycin and 40 ug/mL chloramphenicol. LB media (5 mL) supplemented with 50 ug/mL kanamycin and 40 ug/mL chloramphenicol, was inoculated with the resulting expression strain and grown overnight at 37 °C. The cells were then diluted in 1 L of the same LB media and grown

at 37 °C to an OD = 0.7-0.9. The culture was induced with IPTG (1 mL, 0.5 M), L-arabinose (1 mL, 20% w/v) and p-azidophenylalanine (pAzF) (1 mL, 0.5 M pAzF dissolved in 0.5 M NaOH). pAzF is sparingly soluble in water alone therefore a stock concentration of 0.5 M pAzF in 0.5 M NaOH was used. As a control, a culture was also induced in the absence of pAzF and full length MBP-Cas9 was not observed (Figure S4). Cells were incubated at 25 °C for 16-18 h and collected by centrifugation (7,500 rpm x 25 min). The protein was obtained as described above except that none of the buffers contained reducing agent (i.e BME or DTT or TCEP) as they are known to reduce the azide to an amine[2].

Generation of human cell lysate

HEK293T cells were cultured in Dulbecco's modified Eagle's medium (DMEM), 10 % fetal bovine serum (Gibco), and 1 % anti-anti at 37°C, 5 % CO_2. Cells were cultured in 25 cm^2 flask until they reached a 70-90% confluency. Media was removed and the cells were washed three times with 10 mL of 1x Dulbecco's Phosphate Buffered Saline (DPBS, Gibco). DPBS (4 mL) of was added and cells were collected using a cell scraper. Collected cells were pelleted by centrifugation (500 g x 5 min). DPBS was removed and lysis buffer (4 mL) (20 mM Tris-HCl pH 8, 100 mM KCl, 5 mM $MgCl_2$, 1% NP-40) supplemented with fresh HALT protease inhibitor and RiboLock RNase inhibitor (Thermo Fischer Scientific). The cell pellet was resuspended thoroughly by pipetting up and down 10-20 times and then incubated on ice for 10 min. Cellular debris was removed by centrifugation (15,000 g x 10 min) at 4 °C. The supernatant was removed, and protein concentration was determined using Bradford Assay (Bio-rad) using BSA standards. Protein aliquots were stored at -70°C until needed.

Mass Spectroscopy Analysis of Present Proteins from Excised Bands

Briefly, proteins were reduced and alkylated according to previously described procedures[3], and digested with sequencing grade tryspin per manufacturer's recommendations (Promega). Peptides were dried in a vacuum concentrator after digestion, then resolubilized in 2% acetonitrile/ 0.1 % trifluoroacetic acid. Digested peptides were analyzed by LC-MS/MS on a Thermo Scientific Q Exactive Orbitrap Mass spectrometer in conjunction Proxeon Easy-nLC II HPLC (Thermo Scientific) and Proxeon nanospray source. The digested peptides were loaded a 100 micron x 25 mm Magic C18 100Å 5U reverse phase trap where they were desalted inline before being separated using a 75 micron x 150 mm Magic C18 200Å 3U reverse phase column. Peptides were eluted using a 90 min gradient with a flow rate of 300 nl/min. A survey scan was obtained for the m/z range 300-1600, MS/MS spectra were acquired using a top 15 method, where the top 15 ions in the MS spectra were subjected to HCD (High Energy Collisional Dissociation). An isolation mass window of 2.0 m/z was for the precursor ion selection, and normalized collision energy of 27% was used for fragmentation. A five second duration was used for the dynamic exclusion. Tandem mass spectra were extracted and charge state deconvoluted by Proteome Discoverer (Thermo Scientific). All MS/MS samples were analyzed using X! Tandem (thegpm.org; version TORNADO (2013.02.01.1)). X! Tandem was set up to search (20190304-human-FRDB-wcrap) database (147956 entries), the cRAP database of common laboratory contaminants (www.thegpm.org/crap; 114 entries) plus an equal number of reverse protein sequences assuming the digestion enzyme trypsin. X! Tandem was searched with a fragment ion mass tolerance of 20 PPM and a parent ion tolerance of 20 PPM. Iodoacetamide derivative of cysteine was specified in X! Tandem as a fixed modification. Deamidation of asparagine and glutamine, oxidation of methionine and tryptophan, sulphone of methionine, tryptophan

oxidation to formylkynurenin of tryptophan and acetylation of the n-terminus were specified in X! Tandem as variable modifications.

Criteria for Protein Identification

Scaffold (version Scaffold_4.4.1, Proteome Software Inc.) was used to validate MS/MS based peptide and protein identifications. Peptide identifications were accepted if they could be established at greater than 85.0% probability by the Scaffold Local False Discovery Rate (FDR) algorithm. Protein identifications were accepted if they could be established at greater than 80.0% probability to achieve a FDR less than 5.0% and contained at least 1 identified peptide. Protein probabilities were assigned by the Protein Prophet algorithm[4]. Proteins that contained similar peptides and could not be differentiated based on MS/MS analysis alone were grouped to satisfy the principles of parsimony. Proteins sharing significant peptide evidence were grouped into clusters. From this data, sgRNA binding protein candidates were established by having greater than 90 total spectrum counts present primarily from only one sample band from the same SDS-PAGE. From this analysis, IGF2BP1, hnRNP K, hnRNP C1/C2, MATR3, and YB1, were determined to be candidates (Supporting Information, Table 3).

References

[1] H. O'Geen, I. M. Henry, M. S. Bhakta, J. F. Meckler, D. J. Segal, *Nucleic Acids Res.* **2015**, *43*, 3389–3404.

[2] C. J. López, M. R. Fleissner, E. K. Brooks, W. L. Hubbell, *Biochemistry* **2014**, *53*, 7067–7075.

[3] A. Shevchenko, M. Wilm, O. Vorm, M. Mann, *Anal. Chem.* **1996**, *68*, 850–858.

[4] A. I. Nesvizhskii, A. Keller, E. Kolker, R. Aebersold, *Anal. Chem.* **2003**, *75*, 4646–4658.