Molecular Cell, Volume *46*

Supplemental Information

Mechanism of Foreign DNA Selection in a Bacterial Adaptive Immune System

Dipali G. Sashital, Blake Wiedenheft, and Jennifer A. Doudna

Supplemental Information Inventory

Figure S1:
a. Sequences of crRNA and DNA targets used in binding experiments (Figure 1b, 2c).
b. Control experiment showing effect of adding excess CasA to Cascade dsDNA binding assays (Figure 1b,c 2c).
c. Electrostatic surface map for CasA structure (Figure 1d).

Figure S2.
a-c. Novel CRISPR constructs used for viral resistance assay shown in Figure 2a.
d. Gel showing data quantified in Figure 2b.
e-g. Controls showing that mutant CasA incorporates into Cascade at same levels as WT under conditions used in binding assays in Figure 2c.
h. Control showing that Cascade binds dsDNA with incorrect PAM sequences (Figure 2c) mainly through non-specific interactions.

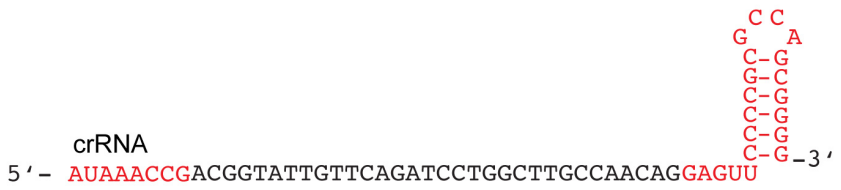
Figure S3.
More detailed look at periodic cleavage pattern observed in hydroxyl radical footprinting of the target strand (Figure 3a,c).

Figure S4.
Alignment of L1 regions of divergent CasA sequences. Figure is too large for main text.

Table S1.
List of plasmids and strains used in this study.

Table S2.
Measure K_d values plotted in Figure 2c,d.

a



O1: WT target
5′-CATGAGGTCCTCGTTTAGTCTGTTGGCAAGCCAGGATCTGAACAATACCGT**CAT**CGGAGGTACGATCAAGG-3′

O2: WT non-target
5′-CCTTGATCGTACCTCCG**ATG**ACGGTATTGTTCAGATCCTGGCTTGCCAACAGACTAAACGAGGACCTCATG-3′

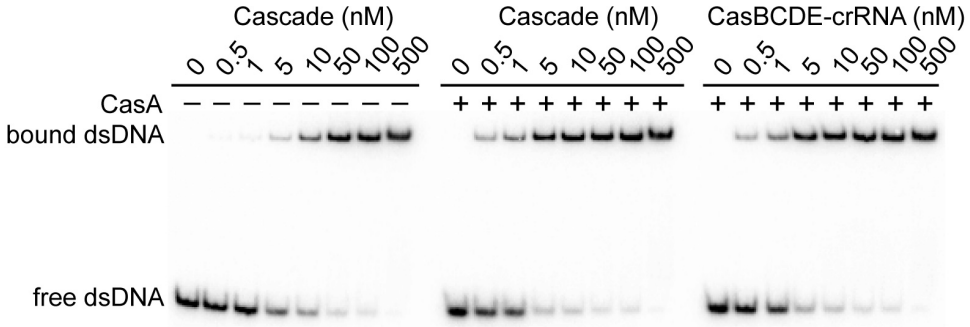
O3: PAM mutant target
5′-CATGAGGTCCTCGTTTAGTCTGTTGGCAAGCCAGGATCTGAACAATACCGT**AGG**CGGAGGTACGATCAAGG-3′

O4: PAM mutant non-target
5′-CCTTGATCGTACCTCCG**CCT**ACGGTATTGTTCAGATCCTGGCTTGCCAACAGACTAAACGAGGACCTCATG-3′

O5: Non-specific 1
5′-CATGAGGTCCTCGTTTAGTCTGCCATAACAAGTCTAGGACCGAACGGTTGT**CAT**CGGAGGTACGATCAAGG-3′

O6: Non-specific 2
5′-CCTTGATCGTACCTCCG**ATG**ACAACCGTTCGGTCCTAGACTTGTTATGGCAGACTAAACGAGGACCTCATG-3′

b



c

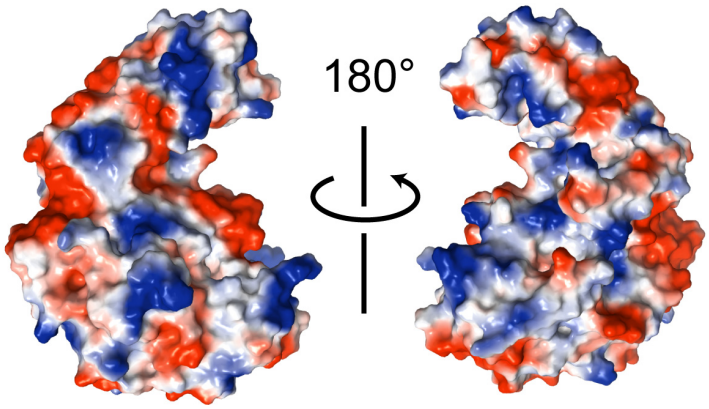


Figure S1. (a) RNA and DNA sequences used for binding assays. The crRNA generated from the CRISPR that was co-expressed with Cascade or CasBCDE protein subunits is shown at the top. For DNA sequences, PAM is in bold and proto-spacer is underlined. Proto-spacers on the target strands are complementary to the crRNA spacer, except for the non-specific DNAs, which contain the reverse of positions 1-31 of the proto-spacer. DNA oligonucleotides are referenced as O1 through O6 according to the nomenclature in this figure throughout the text. (b) Cascade binding O1—O2 dsDNA with and without supplemental CasA. In (+) lanes, excess CasA (2μM) was pre-incubated with either full Cascade or the CasBCDE-crRNA subcomplex prior to addition of dsDNA target. Cascade without supplemental CasA exhibits lower binding affinity for dsDNA, indicating that CasA dissociates from the complex at low concentrations. Full Cascade and CasBCDE-crRNA subcomplex in the presence of excess CasA bind dsDNA identically. (c) Electrostatic surface for CasA (blue = positive, red = negative). No prominent positive patches exist on the protein that could account for non-specific DNA binding. Figure S1 is related to main figure 1.

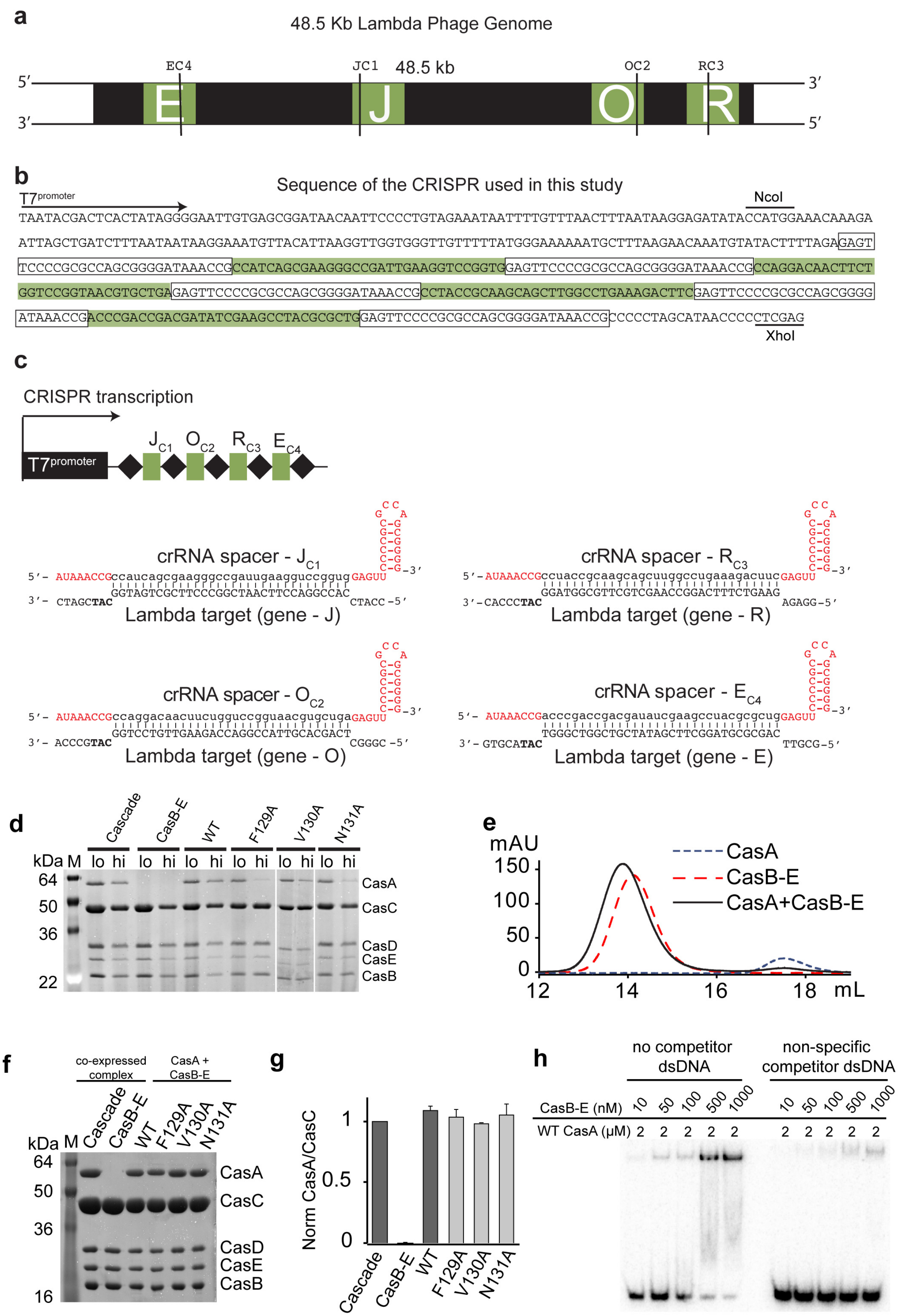

a
48.5 Kb Lambda Phage Genome
EC4
JC1
48.5 kb
OC2
RC3
5′
3′
E
J
O
R
3′
5′
b
Sequence of the CRISPR used in this study
T7promoter
NcoI
TAATACGACTCACTATAGGGGAATTGTGAGCGGATAACAATTCCCCTGTAGAAATAATTTTGTTTAACTTTAATAAGGAGATATACCATGGAAACAAAGA
ATTAGCTGATCTTTAATAATAAGGAAATGTTACATTAAGGTTGGTGGGTTGTTTTTATGGGAAAAAATGCTTTAAGAACAAATGTATACTTTTAGAGAGT
TCCCCGCGCCAGCGGGGATAAACCGCCATCAGCGAAGGGCCGATTGAAGGTCCGGTGGAGTTCCCCGCGCCAGCGGGGATAAACCGCCAGGACAACTTCT
GGTCCGGTAACGTGCTGAGAGTTCCCCGCGCCAGCGGGGATAAACCGCCTACCGCAAGCAGCTTGGCCTGAAAGACTTCGAGTTCCCCGCGCCAGCGGGG
ATAAACCGACCCGACCGACGATATCGAAGCCTACGCGCTGGAGTTCCCCGCGCCAGCGGGGATAAACCGCCCCCTAGCATAACCCCCTCGAG
XhoI
c
CRISPR transcription
T7promoter
JC1
OC2
RC3
EC4
crRNA spacer - JC1
5′- AUAAACCGccaucagcgaagggccgauugaagguccggugGAGUU
3′- CTAGCTAC GGTAGTCGCTTCCCGGCTAACTTCCAGGCCAC CTACC-5′
Lambda target (gene - J)
crRNA spacer - RC3
5′- AUAAACCGccuaccgcaagcagcuuggccugaaagacuucGAGUU
3′-CACCCTAC GGATGGCGTTCGTCGAACCGGACTTTCTGAAG AGAGG-5′
Lambda target (gene - R)
crRNA spacer - OC2
5′- AUAAACCGccaggacaacuucugguccgguaacgugcugaGAGUU
3′- ACCCGTAC GGTCCTGTTGAAGACCAGGCCATTGCACGACT CGGGC -5′
Lambda target (gene - O)
crRNA spacer - EC4
5′- AUAAACCGacccgaccgacgauaucgaagccuacgcgcugGAGUU
3′- GTGCATAC TGGGCTGGCTGCTATAGCTTCGGATGCGCGAC TTGCG-5′
Lambda target (gene - E)
d
Cascade
CasB-E
WT
F129A
V130A
N131A
kDa M lo hi lo hi lo hi lo hi lo hi lo hi
64
50
36
22
CasA
CasC
CasD
CasE
CasB
e
mAU
150
100
50
0
12
14
16
18
mL
CasA
CasB-E
CasA+CasB-E
f
co-expressed complex
CasA + CasB-E
kDa M Cascade CasB-E WT F129A V130A N131A
64
50
36
16
CasA
CasC
CasD
CasE
CasB
g
Norm CasA/CasC
1
0.5
0
Cascade
CasB-E
WT
F129A
V130A
N131A
h
no competitor dsDNA
non-specific competitor dsDNA
CasB-E (nM) 10 50 100 500 1000 10 50 100 500 1000
WT CasA (μM) 2 2 2 2 2 2 2 2 2 2

Figure S2. (a-c) CRISPR construct targeting lambda sequences within genes E, J, O and R containing PAMs. (a) Schematic of lambda genome showing location of proto-spacers. (b) Sequence of CRISPR. Repeat sequences are boxed with a black line, spacers are boxed in green. The CRISPR was synthesized by DNA 2.0, then cloned into pACYCDuet-1 between NcoI and XhoI restriction sites. (c) CRISPR transcription and pre-crRNA processing yields the four crRNAs shown at the bottom. PAMs are shown in bold on the target sequences. (d-g) Reconstitution of Cascade using exogenous CasA added to CasBCDE-crRNA (CasB-E) complex. (d) SDS-PAGE of Cascade pull-downs performed in low (100 mM) and high (500 mM) concentrations of NaCl. Normalized levels of CasA are shown in Figure 2b. (e) Purification of reconstituted Cascade by size exclusion chromatography. Overlay of chromatograms for WT CasA (blue), CasBCDE-crRNA (CasB-E) complex (red) and reconstituted Cascade (2:1 CasA:CasBCDE-crRNA, black) run on a Superose 6 column. (f) SDS-PAGE of co-expressed and purified Cascade and CasBCDE-crRNA (CasB-E) complexes compared with reconstituted WT, F129A and N131A CasA + CasBCDE-crRNA. (g) Bar graph plotting amounts of CasA relative to CasC as detected by SDS-PAGE, normalized to relative amount of CasA for co-expressed Cascade. (h) Gel shift assay showing CasA+CasBCDE-crRNA (CasB-E) binding to O3—O4 dsDNA containing the correct proto-spacer sequence but incorrect PAM in the absence or presence of 1 μM non-specific competitor O5—O6 dsDNA (Fig. S1a). Figure S2 is related to main figure 2.

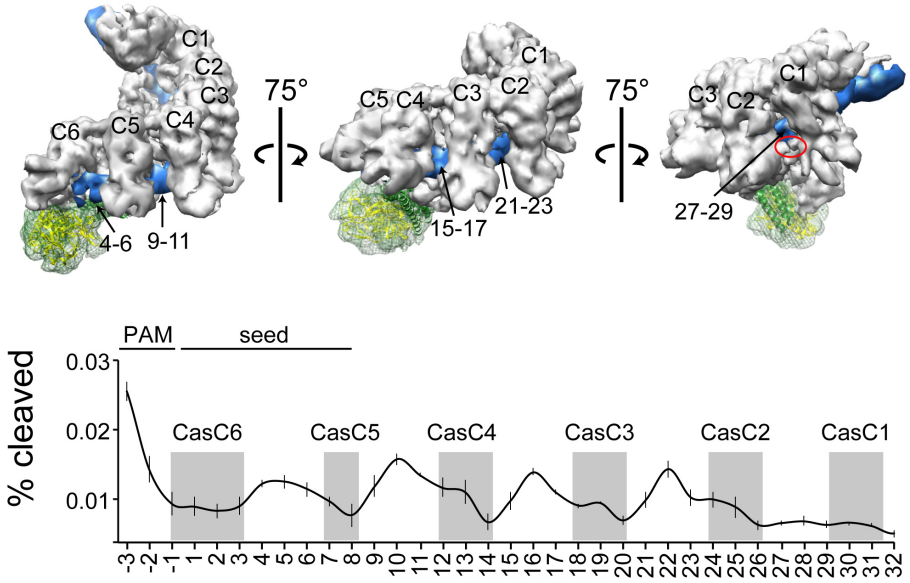


Figure S3. Periodic hydroxyl radical cleavage of the target DNA strand is observed between CasC subunits, indicating that each CasC subunit protects ~3 nt. Cascade is colored as in Figure 1. Extra density (red circle) is observed between CasC1 and C2, which may lead to protection of nucleotides 27-29. Figure S3 is related to main figure 3.

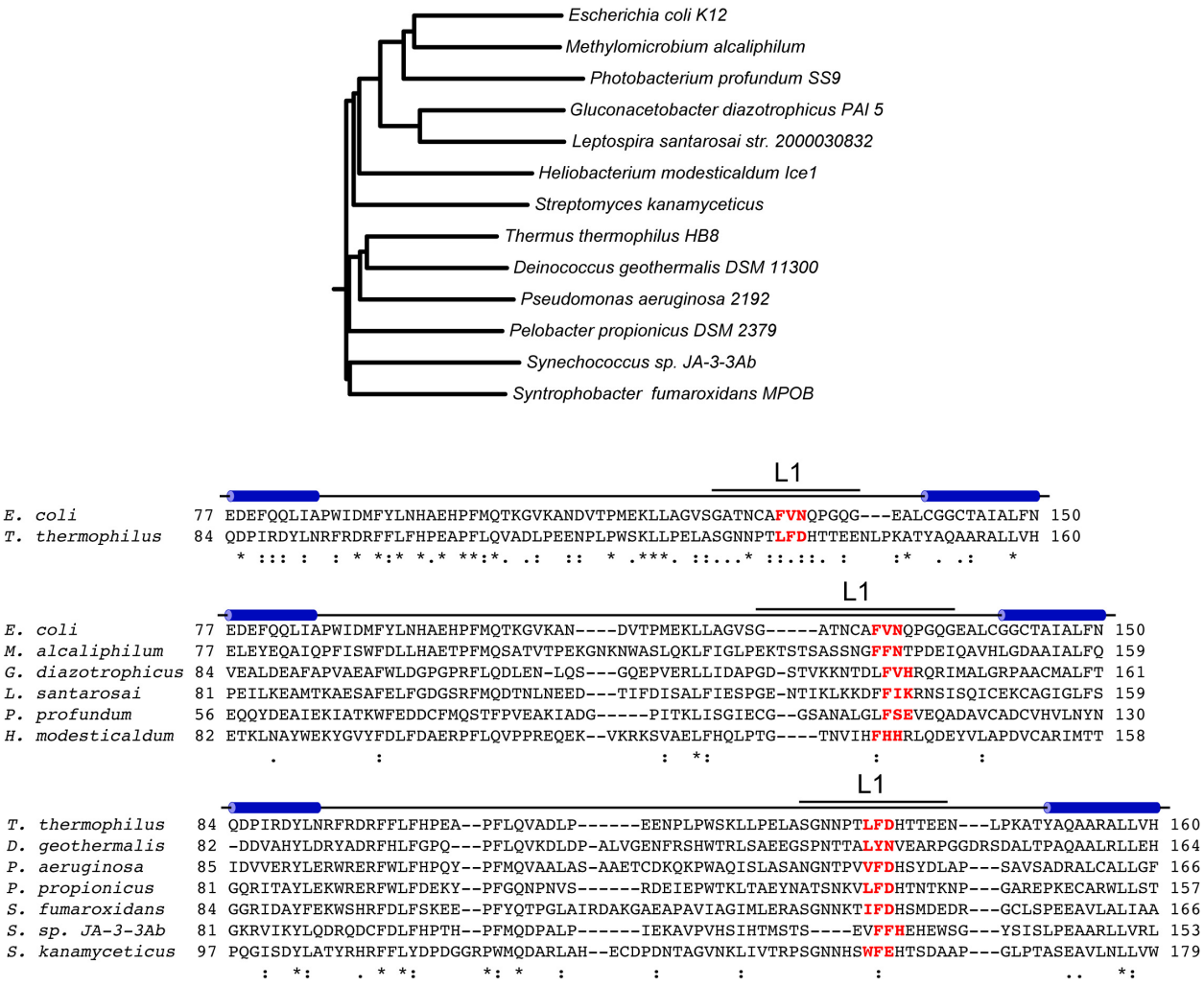


Figure S4. Alignment of L1 region of CasA sequences. The *E. coli* and *T. thermophilus* sequences were used as queries in separate protein-protein BLAST (Altschul et al., 1990) searches, and representatives from the most divergent branches of the resulting phylograms were used in multiple sequence alignments using ClustalW (Larkin et al., 2007). The L1 region and surrounding secondary structure (blue cylinders: α-helices) present in the *T. thermophilus* structure and predicted for all sequences using PSIPRED (Jones, 1999) are shown. The phylogram at the top shows the relationship of the two

groups of sequences used for the alignments. A pairwise alignment of the *E. coli* and *T. thermophilus* sequences was used to identify similar residues in L1 for mutational analysis. Residues between 129 and131 vary in divergent CasA sequences, but an aromatic residue (usually Phe) is present at nearly the same location. This three amino acid motif generally contains another hydrophobic residue and a polar residue at the third position. Figure S4 is related to main figure 4.

Table S1: Plasmids and strains used in this study

Plasmids	Name	Description	Source
1	TtCasA-pSV272	*T. thermophilus casA* expression vector with N-terminal His_6-MBP-TEV site tag	This study
2	EcCasA-pSV272	*E. coli casA* expression vector with N-terminal His_6-MBP-TEV site tag	This study
3	EcCasA F129A-pSV272	*E. coli* F129A *casA* expression vector with N-terminal His_6-MBP-TEV site tag	This study
4	EcCasA V130A-pSV272	*E. coli* V130A *casA* expression vector with N-terminal His_6-MBP-TEV site tag	This study
5	EcCasA N131A-pSV272	*E. coli* N131A *casA* expression vector with N-terminal His_6-MBP-TEV site tag	This study
6	pWUR404	*casE* in pCDF-1b expression vector, no tags	(Brouns, 2007)
7	pWUR408	*casA* in pRSF-1b expression vector, no tags	(Brouns, 2007)
8	pWUR480	*casB* with N-terminal Strep-tag II, *casC*, *casD* in pET52b expression vector	(Brouns, 2007)
9	pWUR547	*E. coli* R44 CRISPR, 7x spacer number 2 in pACYCDuet-1 expression vector	(Brouns, 2007)
10	JORE-CRISPR-DNA	CRISPR targetting PAM-adjacent regions in lambda genes J, O, R and E in pACYCDuet-1 expression vector, Figure S2a-c	This study
11	pWUR477	Non-targeting CRISPR in pACYCDuet-1	(Brouns, 2007)
12	pWUR478	CRISPR targeting regions in lambda genes J, O, R and E with incorrect PAM sequences in pACYCDuet-1 expression vector	(Brouns, 2007)
13	pWUR397	*cas3* in pRSF-1b expression vector, no tags	(Brouns, 2007)
14	pWUR400	*casA-casB-casC-casE* in pCDF-1b expression vector, no tags	(Brouns, 2007)
15	pWUR400 N126A	*casA* with N126A mutation, *casB-casC-casE* in pCDF-1b expression vector, no tags	This study
16	pWUR400 F129A	*casA* with F129A mutation, *casB-casC-casE* in pCDF-1b expression vector, no tags	This study
17	pWUR400 V130A	*casA* with V130A mutation, *casB-casC-casE* in pCDF-1b expression vector, no tags	This study
18	pWUR400 N131A	*casA* with N131A mutation, *casB-casC-casE* in pCDF-1b expression vector, no tags	This study
19	pWUR400 Q132A	*casA* with Q132A mutation, *casB-casC-casE* in pCDF-1b expression vector, no tags	This study
20	pWUR400 Q135A	*casA* with Q135A mutation, *casB-casC-casE* in pCDF-1b expression vector, no tags	This study

Strains	*E. coli* host	Purpose	Plasmid 1	Plasmid 2	Plasmid 3	Plasmid 4
1	BL21 AI	Viral resistance assay WT	10	13	14	--
2	BL21 AI	Viral resistance assay non-targeting CRISPR	11	13	14	--
3	BL21 AI	Viral resistance assay incorrect PAM CRISPR	12	13	14	--
4	BL21 AI	Viral resistance assay N126A CasA	10	13	15	--
5	BL21 AI	Viral resistance assay F129A CasA	10	13	16	--
6	BL21 AI	Viral resistance assay V130A CasA	10	13	17	--
7	BL21 AI	Viral resistance assay N131A CasA	10	13	18	--
8	BL21 AI	Viral resistance assay Q132A CasA	10	13	19	--
9	BL21 AI	Viral resistance assay Q135A CasA	10	13	20	--
10	BL21(DE3)	Cascade expression	6	7	8	9
11	BL21(DE3)	CasBCDE-crRNA expression	6	8	9	--

Table S2: Measured K_d (nM) values

	dsDNA	ssDNA
WT	1.4 ± 0.2	0.25 ± 0.01
CasA mutants, correct PAM		
F129A	720 ± 120	1.1 ± 0.2
V130A	12 ± 1	0.44 ± 0.11
N131A	93 ± 5	1.0 ± 0.1
WT CasA, incorrect PAM		
Fully incorrect	310 ± 70	0.55 ± 0.03
Target strand incorrect	320 ± 80	NA
Non-target strand incorrect	20 ± 3	NA

Supplemental References

Altschul, S.F., Gish, W., Miller, W., Myers, E.W., and Lipman, D.J. (1990). Basic local alignment search tool. J Mol Biol *215*, 403-410.

Brouns, S.J., Jore, M.M., Lundgren, M., Westra, E.R., Slijkhuis, R.J., Snijders, A.P., Dickman, M.J., Makarova, K.S., Koonin, E.V., and van der Oost, J. (2008). Small CRISPR RNAs guide antiviral defense in prokaryotes. Science *321*, 960-964.

Jones, D.T. (1999). Protein secondary structure prediction based on position-specific scoring matrices. J Mol Biol *292*, 195-202.

Larkin, M.A., Blackshields, G., Brown, N.P., Chenna, R., McGettigan, P.A., McWilliam, H., Valentin, F., Wallace, I.M., Wilm, A., Lopez, R.*, et al.* (2007). Clustal W and Clustal X version 2.0. Bioinformatics *23*, 2947-2948.