Quantification of the affinities of CRISPR-Cas9 nucleases for cognate protospacer adjacent motif (PAM) sequences

Vladimir Mekler, Konstantin Kuznedelov, and Konstantin Severinov

SUPPORTING INFORMATION: SI EXPERIMENTAL PROCEDURES, SI FIGURES AND SI TABLES

SI experimental procedures

Cloning Cas9 variants with altered PAM specificity

Amino acid substitutions introduced into SpCas9 variants Cas9-VQR, xCas9 3.7 and SpCas9-NG are listed in Table S1. The Cas9-VQR and xCas9 3.7 genes were cloned into pMJ806 plasmid by replacing the SpeI-NotI fragment of the Cas9 gene with corresponding Cas9-VQR and xCas9 3.7 fragments which were PCR-amplified from plasmids MSP469 and xCas9(3.7)-BE3, respectively (Table S2). Plasmid encoding SpCas9-NG was prepared from the VQR plasmid by introducing the L1111R substitution and replacement of the distal NheI-NotI fragment (1215-1368 amino acids positions) with a g-block encoding G1218R, E1219F, A1322R, R1335V substitutions. Catalytically inactive D10A/H840A mutants of SpCas9 and its variants were generated using the QuickChange site-directed mutagenesis kit. The constructed plasmids express SpCas9 variants as fusion proteins containing N-terminal His_6-MBP affinity tag.

Protein expression and purification

All Cas9 proteins were expressed in Rosetta 2 (DE3) *E.coli* strain (Novagen) by growing at 37 ºC in 2xYT medium containing 25 mg/L kanamycin until OD600=0.8 followed by 18 hours induction with 0.2 mM IPTG at 16 ºC. The proteins were purified by Ni-NTA and MBP tag affinity chromatographies. Briefly, cells were resuspended in buffer A containing 40 mM Tris-HCl pH 7.5, 300 mM NaCl, 1 mM TCEP, 5% glycerol (supplemented with 0.1 mM PMSF), and lysed by sonication. The clarified lysate was supplemented with 1 mM imidazole and loaded onto a HisPur^{TM} Cobalt Superflow Agarose (Thermo Scientific) column equilibrated with buffer A. The column was extensively washed with buffer A, and the bound protein was eluted in buffer A supplemented with 20 mM imidazole. The eluate was loaded onto amylose resin (NEB) in batch mode, and the resin was extensively washed with buffer A. Affinity bound fusion protein was treated with AcTEV protease (Invitrogen) overnight in cold room, and released Cas9

protein was passed through HisPur™ Cobalt Superflow Agarose (Thermo Scientific) column to remove His-tagged AcTEV protease. Cas9 proteins were concentrated to ~4 mg/ml in a 50,000 MWCO centrifugal filter device (Sartorius) and supplemented with glycerol up to 50% final concentration.

Analysis of Competition Binding Assays

Competition beacon assays were analyzed essentially as described previously (1). In all beacon assay experiments concentration of DNA competitor probe (40-200 nM) was significantly higher than concentration of Cas9/gRNA (5-10 nM). Quantitative analysis of competition binding assays is complicated when beacon binding in the presence of competitors is too slow. In this regard, the K_d values were quantified from experiments in which beacon-binding reactions were at least 80% complete in the presence of competitors (compared with the beacon binding without competitors) after incubation for 3 h. In other words, in these experiments, competitors significantly inhibited the rate of beacon binding to Cas9/gRNA, but not the maximum amount of bound beacon molecules. No significant decreases in fluorescence intensity (less than 5% decrease) were observed upon incubation of preformed Cas9/gRNA-beacon complexes with competitors (200 nM) for 1 hour, indicating that the beacon binding reactions were practically irreversible. The K_d values were calculated assuming that the rate of beacon binding in the absence of competitor (V_0) is proportional to the Cas9/gRNA concentration, whereas the binding rate in the presence of competitor (V_1) is proportional to concentration of Cas9/gRNA molecules that remain unbound to competitor; that is $V_0 = X \times [\text{Cas9/gRNA}]$ and $V_1 = X \times ([\text{Cas9/gRNA}] - [\text{Cas9/gRNA*DNA}])$, where Cas9/gRNA*DNA is the Cas9/gRNA complex with competitor and X is a proportionality coefficient. Consequently, the concentration of Cas9/gRNA*DNA was calculated from Eq. 1:

$$[\text{Cas9/gRNA} * \text{DNA}] = [\text{Cas9/gRNA}] \times (1 - V_1/V_0) \quad (1)$$

The K_d for Cas9/gRNA binding to DNA probes was calculated from the chemical equilibrium (Eq. 2):

$$([\text{Cas9/gRNA}] - [\text{Cas9/gRNA} * \text{DNA}]) \times ([\text{DNA}] - [\text{Cas9/gRNA} * \text{DNA}]) = K_d \times [\text{Cas9/gRNA} * \text{DNA}] \quad (2)$$

which can be rewritten as

$$K_d = ([\text{Cas9/gRNA}]/[\text{Cas9/gRNA} * \text{DNA}] - 1) \times ([\text{DNA}] - [\text{Cas9/gRNA} * \text{DNA}]) \quad (3)$$

The V_0 and V_1 rates were determined from the initial stage of the beacon-binding reactions. The rate values were calculated as slopes of initial segments of the kinetic traces measured immediately following beacon addition and mixing of the samples (Fig. S9). The K_d calculation procedure assumes that the equilibrium binding between Cas9/gRNA and the competitor is reached during the reaction incubation time. This assumption is validated by the observation that preincubation of Cas9/gRNA with competitors

before the addition of beacon for either 5 min or 30 min yielded identical beacon binding curves. The equilibrium balance between free and competitor-bound Cas9/gRNA fractions may become disturbed at late stages of beacon binding if the Cas9/gRNA complex with competitor dissociates slowly. Calculation of V_0 and V_1 from the initial stage of the beacon-binding reaction allows one to avoid this potential complication.

SI FIGURES

Guide RNAs

SpCas9 sgRNA

5′GGCUAAAGAGGAAGAGGACAGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCUUUUUUU

SpCas9 crRNA

5′-GGCUAAAGAGGAAGAGGACAGUUUUAGAGCUAUGCUGUUUUG-3′

SpCas9 tracrRNA

5'-AAACAGCAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCU-3′

SaCas9 sgRNA

5′GGGCUAAAGAGGAAGAGGACAGUUUUAGUACUCUGGAAACAGAAUCUACUAAAACAAGGCAAAAUGCCGUGUUUAUCUCGUCAACUUGUUGGCGAGAUUUU-3′

FnCas9 sgRNA

5'GGCUAAAGAGGAAGAGGACAGUUUCAGUUGCUGAAUUAUUUGGUAAACAGUACCAAAUAAUUAAUGCUCUGUAAUCAUUUAAAAGUAUUUUGAACGGACCUCUGUUUGACACGUCUGAAUAACU-3'

Figure S1. The sequences of guide RNAs.

SpCas9 beacon

```
5′Q-ATAGGCTAAAGAGGAAGAGGACGTGGTGAATTCGTAAT-3′
3′F-TATCCGATTTCTCCTTCTCCTGCACCACTTAAGCATTA-5′
```

Cas9-VQR beacon

```
5′Q-GCTAAAGAGGAAGAGGACA-3′
                    5′-TGATGAATTCGTAAT-3′
3′F-CGATTTCTCCTTCTCCTGTACTACTTAAGCATTA-5′
```

xCas9-3.7 and Cas9-NG beacon

```
5′Q-ATAGCTAAAGAGGAAG-3′
                  5′-AGGACATGGTGAATTCGTAAT-3′
3′F-TATCGATTTCTCCTTCTCCTGTACCACTTAAGCATTA-5′
```

SaCas9 beacon

```
5′Q-ATGGCTAAAGAGGAAGAGGATGTCGAGTATACGTACT-3′
3′F-TACCGATTTCTCCTTCTCCTACAGCTCATATGCATGA-5′
```

FnCas9 beacon

```
5′Q-GCTAAAGAGGAAGAGGACG-3′
                    5′-TGGTGAATTCGTAAT-3′
3′F-CGATTTCTCCTTCTCCTGCACCACTTAAGCATTA-5′
```

Figure S2. The structures of Cas9 beacons. The PAM and protospacer sequences are highlighted in yellow and blue, and the mismatched protospacer bases are highlighted in pink.

	DNA probes	Used in
1.1	GTTGAGTATACGTACTGCAA CAACTCATATGCATGACGTT	Fig. 2B, S4
1.2	GTTGAATATACGTACTGCAA CAACTTATATGCATGACGTT	Fig. 2B
1.3	GTTGATTATACGTACTGCAA CAACTAATATGCATGACGTT	Fig. 2B, S4
1.4	GTTGACTATACGTACTGCAA CAACTGATATGCATGACGTT	Fig. 2B
1.5	CATGAGTATACGTACTGCAA TGACTCATATGCATGACGTT	Fig. 2C
1.6	CATGAATATACGTACTGCAA TGACTTATATGCATGACGTT	Fig. 2C
1.7	CATGATTATACGTACTGCAA TGACTAATATGCATGACGTT	Fig. 2C
1.8	CATGACTATACGTACTGCAA TGACTGATATGCATGACGTT	Fig. 2C
2.1	CATGGAGTATACGTACTGCAA GTACCTCATATGCATGACGTT	Fig. 3
2.2	CATGAAGTATACGTACTGCAA GTACTTCATATGCATGACGTT	Fig. 3
2.3	CATAGAGTATACGTACTGCAA GTATCTCATATGCATGACGTT	Fig. 3
2.4	CATGTAGTATACGTACTGCAA GTACATCATATGCATGACGTT	Fig. 3
2.5	CATAAAGTATACGTACTGCAA GTATTTCATATGCATGACGTT	Fig. 3
2.6	CATGCAGTATACGTACTGCAA GTACGTCATATGCATGACGTT	Fig. 3
2.7	CATTGAGTATACGTACTGCAA GTAACTCATATGCATGACGTT	Fig. 3
2.8	CATCGAGTATACGTACTGCAA GTAGCTCATATGCATGACGTT	Fig. 3
3.1	CAAGGAGATACGTACTGGAA GTTCCTCTATGCATGACCTT	Fig. 4 and S5
3.2	CAAGAAGATACGTACTGGAA GTTCTTCTATGCATGACCTT	Fig. 4 and S5
3.3	CAAGTAGATACGTACTGGAA GTTCATCTATGCATGACCTT	Fig. 4 and S5
3.4	CAAGCAGATACGTACTGGAA GTTCGTCTATGCATGACCTT	Fig. 4 and S5

```
3.5      CAATCAGATACGTACTGGAA                                  Fig. 4
         GTTAGTCTATGCATGACCTT


           +1
4.1          TGGAGTATACGTACTGCAA                    Figs. 5, 7, S7, S9
             ACCTCATATGCATGACGTT
          -1
4.2        TTGGAGTATACGTACTGCAA                       Figs. 5 and 7
           AACCTCATATGCATGACGTT
         -2
4.3       GTTGGAGTATACGTACTGCAA                       Figs. 5, 7, S6
          CAACCTCATATGCATGACGTT
        -3
4.4      TGTTGGAGTATACGTACTGCAA                       Figs. 5 and 7
         ACAACCTCATATGCATGACGTT
     -5
4.5   ACTGTTGGAGTATACGTACTGCAA                        Figs. 5 and 7
      TGACAACCTCATATGCATGACGTT
   -7
4.6 CTACTGTTGGAGTATACGTACTGCAA                       Figs. 5 and 7
    GATGACAACCTCATATGCATGACGTT
         -2
4.7       GTTGTAGTATACGTACTGCAA                      Fig. S6
          CAACATCATATGCATGACGTT
            +2
4.8           GGAGTATACGTACTGCAA                     Fig. S7
              CCTCATATGCATGACGTT
```

Figure S3. Sequences of competitor DNA probes. The probe numbers used in the text are shown on the left of the sequences. The PAM sequences are highlighted in yellow, 2-bp upstream segments matching the spacer sequence are highlighted in blue, mutated PAM bases are highlighted in pink.

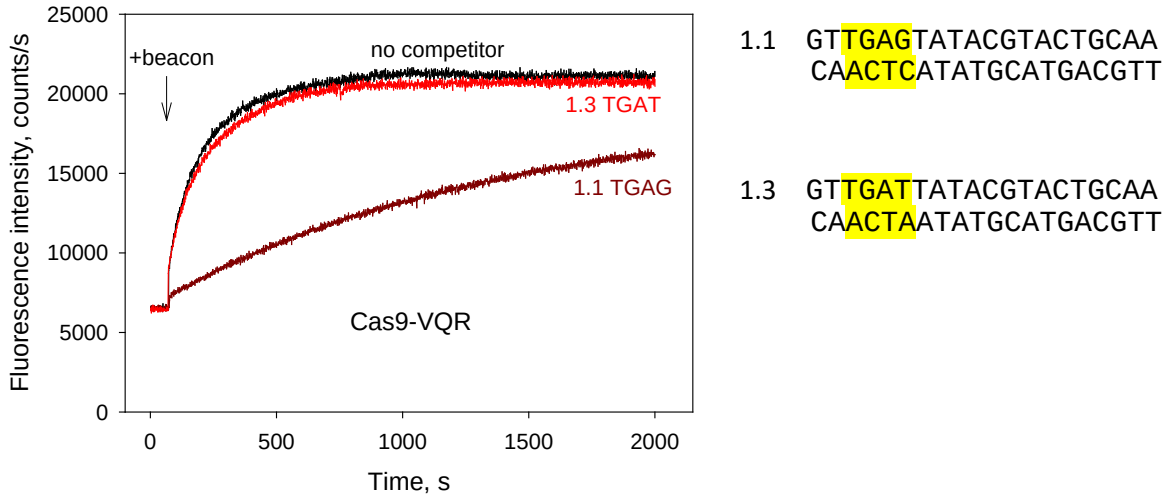


Figure S4. Effect of competitor DNA probes 1.1 and 1.3 on the kinetics of dCas9-VQR/gRNA binding to beacon (related to Fig. 2B). The probe concentrations were 50 nM. The probes bore non-complementary to gRNA 2 bp segments upstream of indicated in the panel PAM sequences.

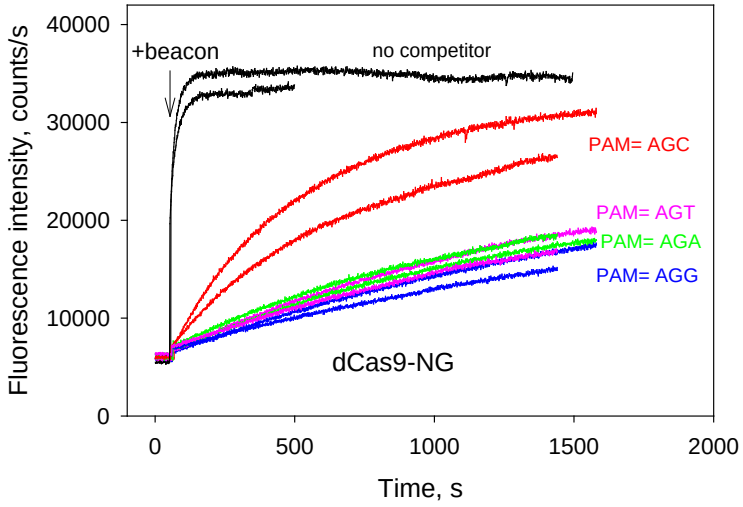


Figure S5. Effect of competitor probes with different PAM sequences on the kinetics of dCas9-NG/gRNA binding to beacon measured in two replicate experiments (related to Figs. 4A and C). The identity of PAM sequences in competitor probes is indicated at the right and curves are colored correspondingly. The figure provides a comparison between the kinetic curves shown in Fig. 4C (longer traces) and kinetic curves obtained from independent replicate measurements (shorter traces). Structures of competitor DNA probes are shown in Fig. 4A, the concentrations of competitor DNA probes were 40 nM.

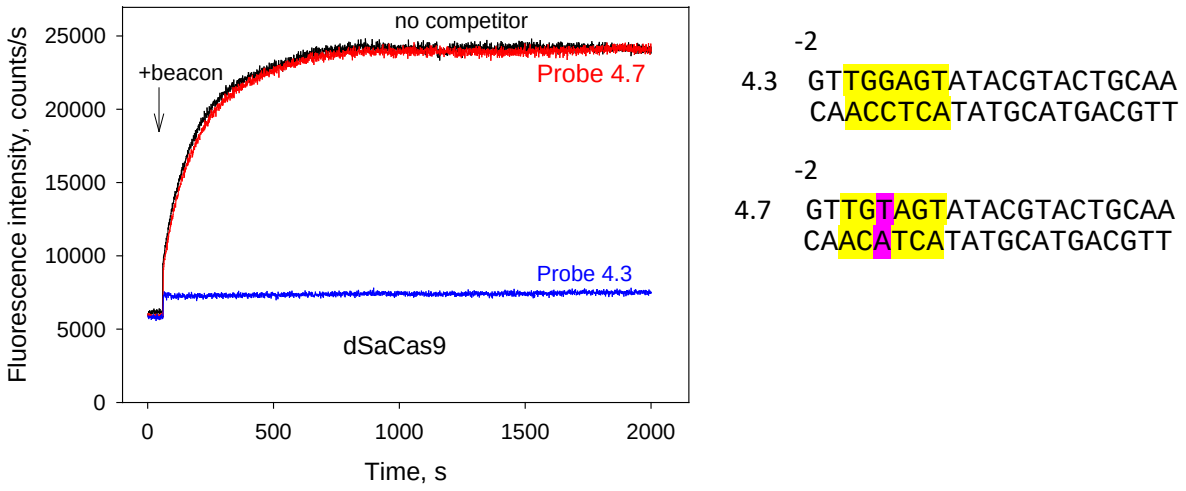
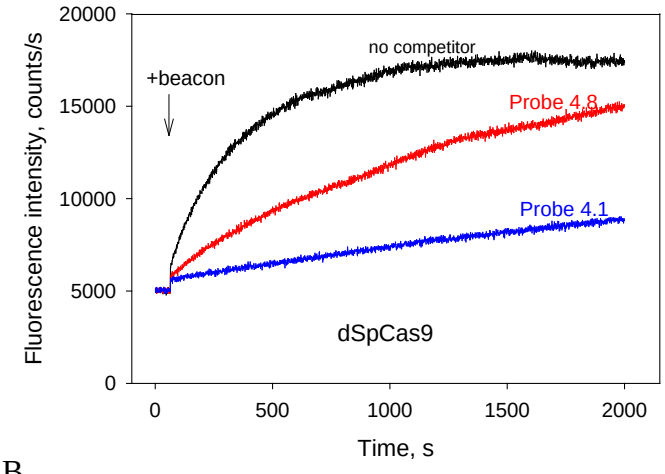
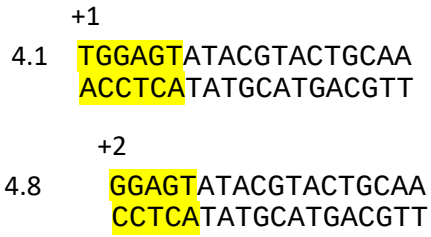


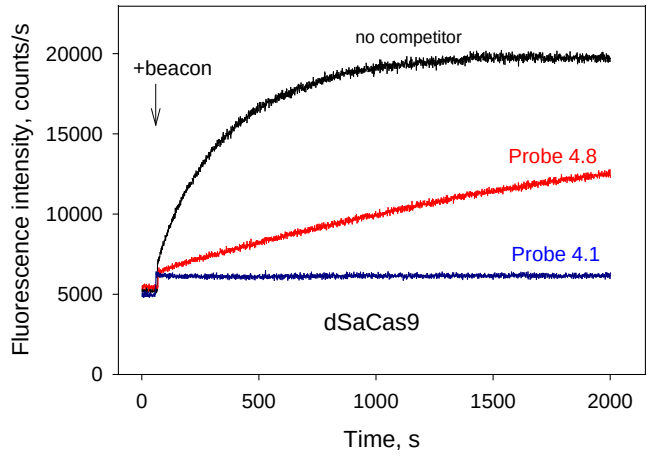
Figure S6. Effect of competitor DNA probes 4.3 and 4.7 on the kinetics of dSaCas9/gRNA binding to beacon. The probe concentrations were 40 nM.

A





B



C

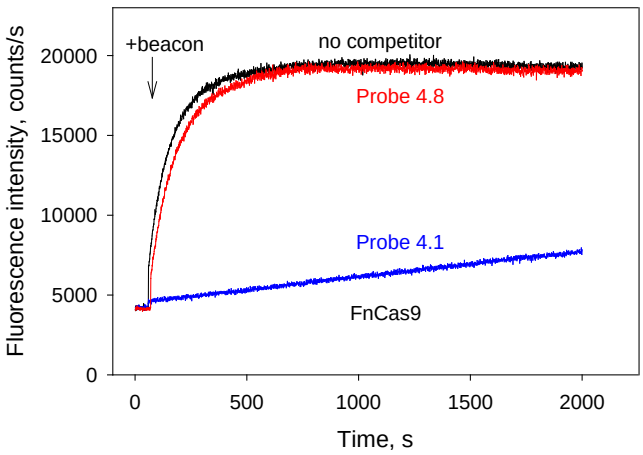


Figure S7. Effect of competitor DNA probes 4.1 and 4.8 on the kinetics of dSpCas9/gRNA (A), dSaCas9/gRNA (B) and FnCas9/gRNA (C) binding to beacons. The probe concentrations in experiments shown in panels A, B and C were 200, 40 and 200 nM, respectively.

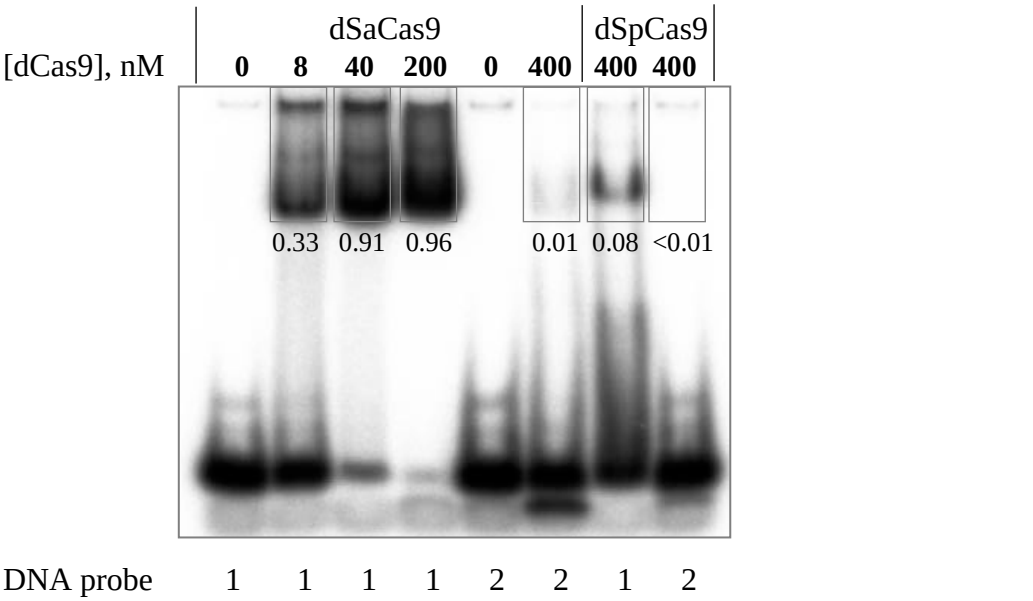
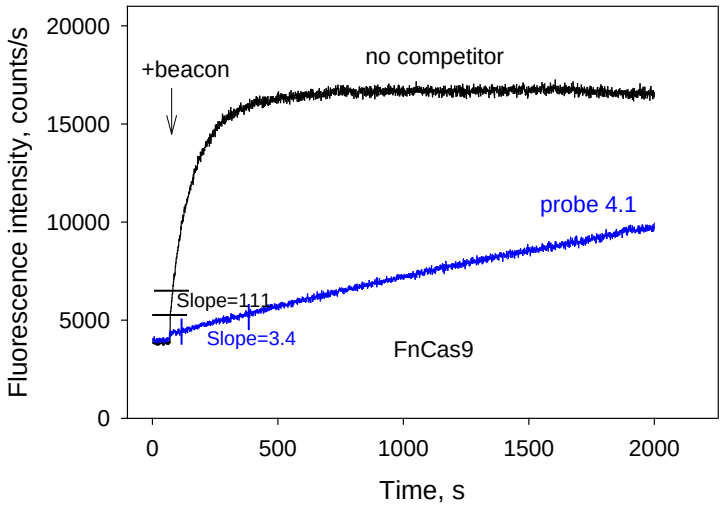


Figure S8. EMSA assay of dSpCas9/gRNA and dSaCas9/gRNA binding to probe 1 containing both SpCas9 and SaCas9 PAM sequences and to control probe 2 bearing mutated PAM sequences. The probe segments containing PAM sequences are highlighted in yellow, mutated PAM bases in probe 2 are highlighted in pink. The probe 1 is similar to probe 4.3 (shown in Figs. 5A and S3) except that the sequence of probe 1 was modified in order to prevent possible formation of hairpin structures under the low temperature EMSA conditions used, which may be more favorable for hairpin formation than the conditions of the fluorescence experiments. The ratios of the bound DNA probe band intensity to the total intensity of both the bound and unbound DNA bands are shown in the panel.



4.1 TGGAGTATACGTACTGCAA
ACCTCATATGCATGACGTT

Figure S9. Effect of competitor DNA probe 4.1 on the kinetics of FnCas9/gRNA binding to beacon (related to Fig. 7). The probe concentration was 100 nM. Slopes of the initial portions of the beacon binding kinetic traces recorded immediately following beacon addition and mixing of the sample are indicated; the mixing dead-time was 15 s.

SI tables

Table S1. *S. pyogenes* Cas9 variants with altered PAM specificity used in this work (2-4). The substituted amino acids are shown in bold.

Cas9 amino acid variable positions													
Domain	Recognition						PAM-interacting						
# / **ID**	262	324	409	480	543	694	1111	1135	1218	1219	1322	1335	1337
SpCas9	A	R	S	E	E	M	L	D	G	E	A	R	T
Cas9-VQR	A	R	S	E	E	M	L	**V**	G	E	A	**Q**	**R**
SpCas9-NG	A	R	S	E	E	M	**R**	**V**	**R**	**F**	**R**	**V**	**R**
xCas9-3.7	**T**	**L**	**I**	**K**	**D**	**I**	L	D	G	**V**	A	R	T

Table S2. Plasmids used in this study.

Plasmids	**Relevant Description**	**Source/Reference**
pMJ806	pET-based His_6-MBP expression vector with SpCas9 (1-1368)	Addgene / (2)
pMJ841	pET-based His_6-MBP expression vector with SpCas9 (1-1368), D10A/H840A double mutant, catalytically inactive	Addgene / (2)
MSP469	Human expression vector for SpCas9 VQR variant: D1135V/ R1335Q/T1337R	Addgene / (3)
pVQR	pET-based His_6-MBP expression vector with SpCas9 VQR variant: D1135V/R1335Q/T1337R	This study/ (3)
pdVQR	pVQR-derived by D10A/H840A mutations in VQR, catalytically inactive	This study
pSpCas9-NG	pVQR-derived by mutations L1111R, G1218R, E1219F, A1322R, R1335V in VQR (SpCas9-NG variant)	This study/ (5)
pdSpCas9-NG	pSpCas9-NG-derived by D10A/H840A mutations in SpCas9-NG, catalytically inactive	This study
xCas9(3.7)-BE3	Mammalian expression vector for xCas9(3.7)-BE3: A262T, R324L, S409I, E480K, E543D, M694I, E1219V (xCas9 3.7 variant)	Addgene / (4)
pxCas9-3.7	pET-based His_6-MBP expression vector with xCas9 3.7 variant: A262T, R324L, S409I, E480K, E543D, M694I, E1219V	This study/ (4)
pdxCas9-3.7	pxCas9-3.7-derived by D10A/H840A mutations in xCas9 3.7 variant, catalytically inactive	This study

Table S3. Oligonucleotides for in vitro transcription.

Sequence	description	source	ref.
5'taatacgactcactatagGGCTAAAGAGGAAGAGGACAGTTTTAGTACTCTGGAAACAGAATCTACTAAAACAAGGCAAAATGCCGTGTTTATCTCGTCAACTTGTTGGCGAGATTTT-3'	Non-template DNA strand for SaCas9 sgRNA transcription	IDT	6
5'AAAATCTCGCCAACAAGTTGACGAGATAAACACGGCATTTTGCCTTGTTTTAGTAGATTCTGTTTCCAGAGTACTAAAACTGTCCTCTTCCTCTTTAGCCctatagtgagtcgtatta-3'	Template DNA strand for SaCas9 sgRNA transcription	IDT	6

Synthetic gBlock sequence used for pSpCas9-NG plasmid construction (*Nhe*I and *Not*I cloning sites are shown in italic):

5'cggatgttg*gctagc*gccagattccttcaaaaggggaacgaactcgcactaccgtctaaatacgtgaatttcctgtatttagcgtcccattacgagaagttgaaaggttcacctgaagataacgaacagaagcaacttttgttgagcagcacaaacattatctcgacgaaatcatagagcaaatttcggaattcagtaagagagtcatcctagctgatgccaatctggacaaagtattaagcgcatacaacaagcacagggataaacccatacgtgagcaggcggaaaatattatccatttgtttactcttaccaacctcggcgctccacgcgcattcaagtattttgacacaacgatagatcgcaaagtgtacagatctaccaaggaggtgctagacgcgacactgattcaccaatccatcacgggattatatgaaactcggatagatttgtcacagcttgggggtgactaa*gcggccgc*actcgagcacc-3'

References

1. Mekler, V., Minakhin, L., and Severinov, K. (2017) Mechanism of duplex DNA destabilization by RNA-guided Cas9 nuclease during target interrogation. *Proc. Natl. Acad. Sci. U.S.A.* **114,** 5443–5448

2. Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J. A., and Charpentier, E. (2012) A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science*. **337**, 816–821

3. Kleinstiver, B. P., Prew, M. S., Tsai, S. Q., Topkar, V. V., Nguyen, N. T., Zheng, Z. L., Gonzales, A. P. W., Li, Z. Y., Peterson, R. T., Yeh, J. R. J., Aryee, M.J., and Joung, J.K. (2015) Engineered CRISPR-Cas9 nucleases with altered PAM specificities. *Nature.* **523,** 481–485

4. Hu, J. H., Miller, S. M., Geurts, M. H., Tang, W., Chen, L., Sun, N., Zeina, C. M., Gao, X., Rees, H. A., Lin, Z., and Liu, D.R. (2018) Evolved Cas9 variants with broad PAM compatibility and high DNA specificity. *Nature.* **556,** 57–63

5. Nishimasu, H., Shi, X., Ishiguro, S., Gao, L., Hirano, S., Okazaki, S., Noda, T., Abudayyeh, O. O., Gootenberg, J. S., Mori, H., Oura, S., Holmes, B., Tanaka, M., Seki, M., Hirano, H., Aburatani, H., Ishitani, R., Ikawa, M., Yachie, N., Zhang, F., and Nureki, O. (2018) Engineered CRISPR-Cas9 nuclease with expanded targeting space. *Science.* **361,** 1259–1262

6. Ran, F. A., Cong, L., Yan, W. X., Scott, D. A., Gootenberg, J. S., Kriz, A. J., Zetsche, B., Shalem, O., Wu, X., Makarova, K. S., Koonin, E.V., Sharp, P.A., and Zhang, F. (2015) In vivo genome editing using *Staphylococcus aureus* Cas9. *Nature.* **520,** 186–191