Molecular Cell, Volume *71*

Supplemental Information

Kinetic Basis for DNA Target Specificity of CRISPR-Cas12a

Isabel Strohkendl, Fatema A. Saifuddin, James R. Rybarski, Ilya J. Finkelstein, and Rick Russell

Kinetic basis for DNA target specificity of CRISPR-Cas12a

Isabel Strohkendl, Fatema A. Saifuddin, James R. Rybarski, Ilya J. Finkelstein, and Rick Russell

Supplemental Information

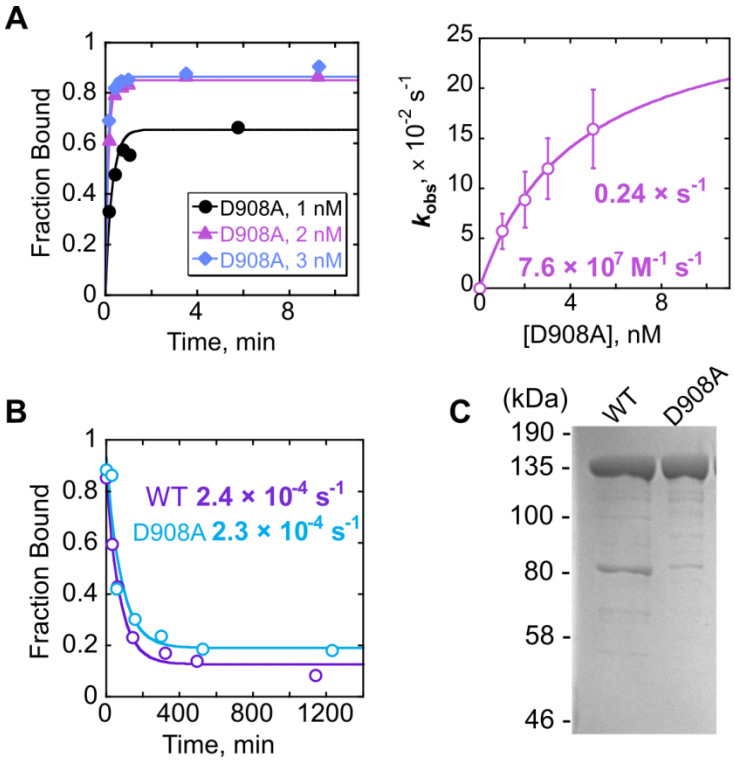

Figure S1. Comparison of Cas12a and Cas12a(D908A) (Related to Figure 1)

(A) Left: representative time courses of Cas12a(D908A) binding to the matched DNA target. Right: hyperbolic fit used to determine the second-order rate constant for binding and the first-order rate constant for R-loop propagation. Error bars represent the SEM of three replicates. The data gave a k_{on} value of 7.6 (± 1.3) × 10^7 M^{-1} s^{-1} and a k_{max} value of 0.24 ± 0.06 s^{-1}.

(B) Comparison of the dissociation time courses of Cas12a (purple) and Cas12a(D908A) (blue) when Mg^{2+} is chelated by EDTA. The data gave k_{off} values of 2.4 (± 0.3) and 2.3 (± 0.3) × 10^{-4} s^{-1} for Cas12a and Cas12a(D908A), respectively. The k_{off} values are given as the mean ± SEM.

(C) Gel image of wild-type Cas12a (WT) and Cas12a(D908A) after purification by affinity and size exclusion chromatography.

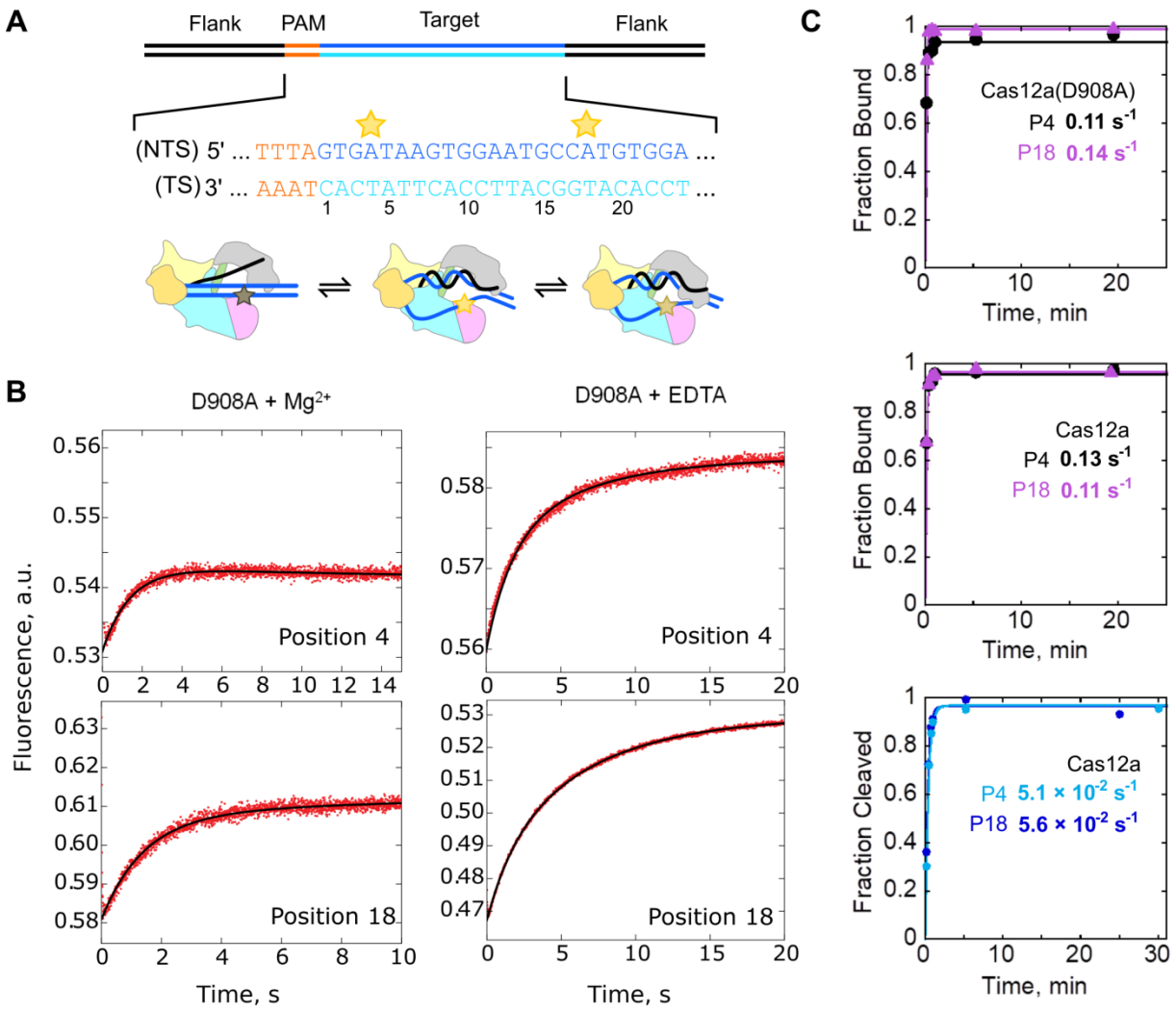

Figure S2. Time courses of Cas12a(D908A) R-loop propagation monitored by 2-aminopurine (2-AP) fluorescence (Related to Figure 1)

(A) Schematic of constructs and 2-AP fluorescence changes. Upon formation of the R-loop, the NTS, carrying the 2-AP base (at the PAM-distal position in the reaction cartoon), is displaced by the crRNA, giving an increase in 2-AP fluorescence. A rapid subsequent step, likely a conformational change within the complex, results in a small decrease in 2-AP fluorescence.

(B) Time-dependent changes in 2-AP fluorescence upon DNA binding. Data from all 2-AP fluorescence experiments were best fit by a double exponential equation, indicating the presence of two steps. The model that gave the best global fit to the data includes a first slow step, resulting in a substantial increase in 2-AP fluorescence, followed by a faster step that results in a modest decrease in fluorescence. We infer that the slower step is R-loop formation because of the correspondence with the maximal rate constant determined in binding assays and because unwinding of the DNA helix and exclusion of the NTS during R-loop formation would be expected to increase 2-AP fluorescence. Although the second step decreases the fluorescence value, the time dependences do not display a phase in which the total

fluorescence decreases, instead giving a biphasic curve that increases in value over time. This is because the intermediate with high fluorescence does not accumulate substantially due to its rapid conformational change to the final state. Thus, at any given time, the total fluorescence increase caused by R-loop formation is greater than the decrease caused by the faster conformational change. The rate constants for these transitions were not strongly affected by the presence of 5 mM $MgCl_2$ and were the same within error regardless of the position of the 2-AP. For Cas12a(D908A) with Mg^{2+}, $k_{R\text{-}loop}$ = 0.15 ± 0.01 s^{-1} (left). For Cas12a(D908A) without free Mg^{2+} (+EDTA), $k_{R\text{-}loop}$ = 0.151 ± 0.003 s^{-1} (right). The values of the rate constant for the faster transition were 0.71 ± 0.04 s^{-1} and 0.49 ± 0.03 s^{-1} in the presence and absence of Mg^{2+}, respectively. Values reflect the fit-derived rate ± SE.

(C) Gel-shift assays and cleavage assays of 2-AP-containing target DNAs. Shown are representative time courses of binding by Cas12a(D908A) (top panel) and Cas12a (middle panel) and cleavage by Cas12a (bottom panel). The rates of binding and cleavage with 10 nM enzyme and 5 mM $MgCl_2$ were unaffected by the presence of the 2-AP base at either position 4 (P4) or 18 (P18). For Cas12a(D908A), observed rate constants were 0.11 ± 0.01 s^{-1} and 0.14 ± 0.02 s^{-1} with 2-AP at position 4 and 18, respectively. For wild-type Cas12a, the corresponding rate constants were 0.13 ± 0.01 s^{-1} and 0.11 ± 0.01 s^{-1}, respectively. Observed rate constants for cleavage of the non-target strand (k_{obs}^{NTS}) were 5.1 (± 0.7) × 10^{-2} s^{-1} and 5.6 (± 0.3) × 10^{-2} s^{-1} with 2-AP at position 4 and 18, respectively.

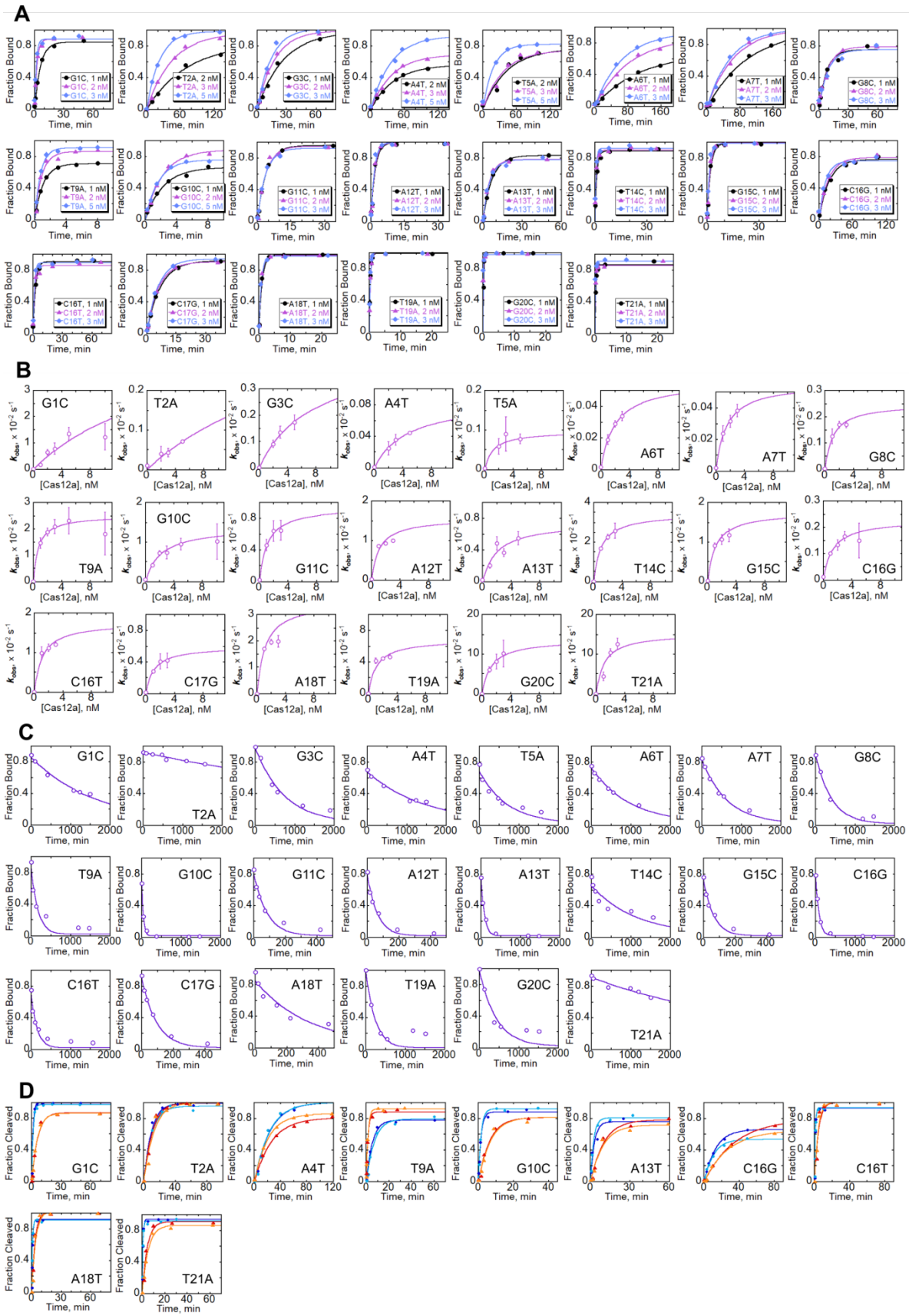
A
Fraction Bound
Time, min
G1C, 1 nM
G1C, 2 nM
G1C, 3 nM
T2A, 2 nM
T2A, 3 nM
T2A, 5 nM
G3C, 1 nM
G3C, 2 nM
G3C, 3 nM
A4T, 2 nM
A4T, 3 nM
A4T, 5 nM
T5A, 2 nM
T5A, 3 nM
T5A, 5 nM
A6T, 1 nM
A6T, 2 nM
A6T, 3 nM
A7T, 1 nM
A7T, 2 nM
A7T, 3 nM
G8C, 1 nM
G8C, 2 nM
G8C, 3 nM
T9A, 1 nM
T9A, 2 nM
T9A, 5 nM
G10C, 1 nM
G10C, 2 nM
G10C, 5 nM
G11C, 1 nM
G11C, 2 nM
G11C, 3 nM
A12T, 1 nM
A12T, 2 nM
A12T, 3 nM
A13T, 1 nM
A13T, 2 nM
A13T, 3 nM
T14C, 1 nM
T14C, 2 nM
T14C, 3 nM
G15C, 1 nM
G15C, 2 nM
G15C, 3 nM
C16G, 1 nM
C16G, 2 nM
C16G, 3 nM
C16T, 1 nM
C16T, 2 nM
C16T, 3 nM
C17G, 1 nM
C17G, 2 nM
C17G, 3 nM
A18T, 1 nM
A18T, 2 nM
A18T, 3 nM
T19A, 1 nM
T19A, 2 nM
T19A, 3 nM
G20C, 1 nM
G20C, 2 nM
G20C, 3 nM
T21A, 1 nM
T21A, 2 nM
T21A, 3 nM
B
k_{obs}, x 10^{-2} s^{-1}
[Cas12a], nM
G1C
T2A
G3C
A4T
T5A
A6T
A7T
G8C
T9A
G10C
G11C
A12T
A13T
T14C
G15C
C16G
C16T
C17G
A18T
T19A
G20C
T21A
C
Fraction Bound
Time, min
G1C
T2A
G3C
A4T
T5A
A6T
A7T
G8C
T9A
G10C
G11C
A12T
A13T
T14C
G15C
C16G
C16T
C17G
A18T
T19A
G20C
T21A
D
Fraction Cleaved
Time, min
G1C
T2A
A4T
T9A
G10C
A13T
C16G
C16T
A18T
T21A

Figure S3. Time courses and concentration dependences of Cas12a binding and cleavage of DNA targets with single mismatches (Related to Figure 3 and Figure 4)

(A) Representative time courses of binding reactions at multiple Cas12a concentrations.

(B) Observed rate constant for binding as a function of Cas12a concentration. These plots were used to derive the second-order rate constant for binding (k_{on}) and maximal rate constant for R-loop propagation ($k_{R\text{-}loop}$) from the fit of a hyperbolic function. Error bars represent the SEM of at least three replicates.

(C) Representative time courses of Cas12a(D908A) complex dissociation. Dissociation time courses were fit to a single exponential to determine the rate constant. We note that the time dependences of PAM-distal mismatches appear to include a second phase, perhaps reflecting a sub-population in which the R-loop is incompletely formed. Further work will be required to dissect this behavior, but its presence does not affect the conclusions from the current work.

(D) Representative time courses of cleavage for both DNA strands of a subset of mismatch targets with saturating concentrations of Cas12a: NTS at 10 nM Cas12a (blue), NTS at 50 nM Cas12a (cyan), TS at 10 nM Cas12a (red), and TS at 50 nM Cas12a (orange).

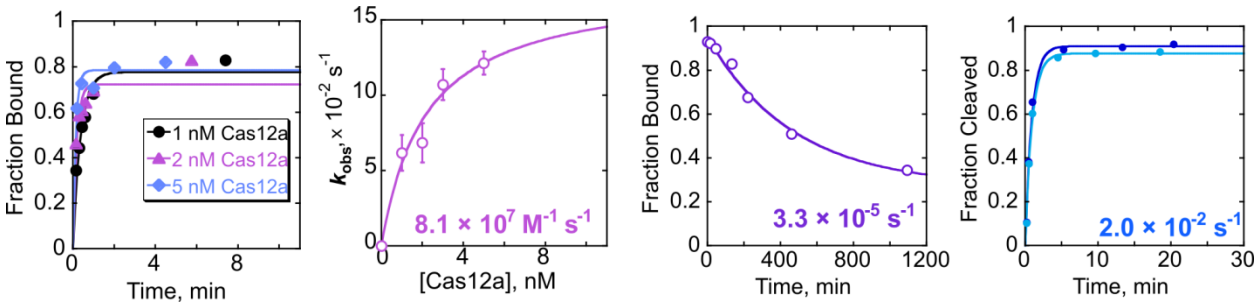

Figure S4. Cas12a interactions and cleavage of a matched target with a physiological Mg^{2+} concentration (Related to Figure 1)

Graphs from left to right depict the same type of experiments as shown in Figure S3 A-D, respectively, but represent experiments performed using 1 mM $MgCl_2$. Error bars in the plot of observed binding rate constant versus Cas12a concentration represent the SEM of at least three replicates. Wild-type Cas12a bound matched target DNA with a k_{on} value of 8.1 (± 1.2) × 10^7 M^{-1} s^{-1} and cleaved the NTS with a k_c^{NTS} value of 2.0 (± 0.1) × 10^{-2} s^{-1}. Cas12a(D908A) dissociated from matched target DNA with a k_{off} value of 3.3 (± 0.6) × 10^{-5} s^{-1}. In the plot at the right, NTS cleavage is measured with 10 nM (blue) or 50 nM (cyan) Cas12a. The value of k_{off} represents the mean ± SEM of three independent determinations. The values of k_{on} and k_c represent the best fit ± SE.

DNA Target Oligo	**Sequence** (5' to 3')
PT NTS	ACGCTCTTCCGATCTTTTAGTGATAAGTGGAATGCCATGTGGAGATCGGAAGAGCAC
G1C NTS	ACGCTCTTCCGATCTTTTACTGATAAGTGGAATGCCATGTGGAGATCGGAAGAGCAC
T2A NTS	CGCTCTTCCGATCTTTTAGAGATAAGTGGAATGCCATGTGGAGCTACTGTGCTGCT
G3C NTS	CGCTCTTCCGATCTTTTAGTCATAAGTGGAATGCCATGTGGAGTAGCTACTGTGCT
A4T NTS	CGCTCTTCCGATCTTTTAGTGTTAAGTGGAATGCCATGTGGAGCTACTGTGCTGCT
T5A NTS	ACGCTCTTCCGATCTTTTAGTGAAAAGTGGAATGCCATGTGGAGATCGGAAGAGCAC
A6T NTS	CGCTCTTCCGATCTTTTAGTGATTAGTGGAATGCCATGTGGAGTAGCTACTGTGCT
A7T NTS	CGCTCTTCCGATCTTTTAGTGATATGTGGAATGCCATGTGGAGTAGCTACTGTGCT
G8C NTS	CGCTCTTCCGATCTTTTAGTGATAACTGGAATGCCATGTGGAGTAGCTACTGTGCT
T9A NTS	ACGCTCTTCCGATCTTTTAGTGATAAGAGGAATGCCATGTGGAGATCGGAAGAGCAC
G10C NTS	ACGCTCTTCCGATCTTTTAGTGATAAGTCGAATGCCATGTGGAGATCGGAAGAGCAC
G11C NTS	CGCTCTTCCGATCTTTTAGTGATAAGTGCAATGCCATGTGGAGTAGCTACTGTGCT
A12T NTS	CGCTCTTCCGATCTTTTAGTGATAAGTGGTATGCCATGTGGAGTAGCTACTGTGCT
A13T NTS	ACGCTCTTCCGATCTTTTAGTGATAAGTGGATTGCCATGTGGAGATCGGAAGAGCAC
T14C NTS	CGCTCTTCCGATCTTTTAGTGATAAGTGGAACGCCATGTGGAGTAGCTACTGTGCT
G15C NTS	CGCTCTTCCGATCTTTTAGTGATAAGTGGAATCCCATGTGGAGTAGCTACTGTGCT
C16G NTS	ACGCTCTTCCGATCTTTTAGTGATAAGTGGAATGGCATGTGGAGATCGGAAGAGCAC
C16T NTS	ACGCTCTTCCGATCTTTTAGTGATAAGTGGAATGTCATGTGGAGATCGGAAGAGCAC
C17G NTS	CGCTCTTCCGATCTTTTAGTGATAAGTGGAATGCGATGTGGAGTAGCTACTGTGCT
A18T NTS	CGCTCTTCCGATCTTTTAGTGATAAGTGGAATGCCTTGTGGAGTAGCTACTGTGCT
T19A NTS	CGCTCTTCCGATCTTTTAGTGATAAGTGGAATGCCAAGTGGAGTAGCTACTGTGCT
G20C NTS	CGCTCTTCCGATCTTTTAGTGATAAGTGGAATGCCATCTGGAGTAGCTACTGTGCT
T21A NTS	ACGCTCTTCCGATCTTTTAGTGATAAGTGGAATGCCATGAGGAGATCGGAAGAGCAC
PT TS	GTGCTCTTCCGATCTCCACATGGCATTCCACTTATCACTAAAAGATCGGAAGAGCGT
G1C TS	GTGCTCTTCCGATCTCCACATGGCATTCCACTTATCAGTAAAAGATCGGAAGAGCGT
T2A TS	AGCACAGTAGCTACTCCACATGGCATTCCACTTATCTCTAAAAGATCGGAAGAGCG
G3C TS	AGCACAGTAGCTACTCCACATGGCATTCCACTTATGACTAAAAGATCGGAAGAGCG
A4T TS	AGCAGCACAGTAGCTCCACATGGCATTCCACTTAACACTAAAAGATCGGAAGAGCG
T5A TS	GTGCTCTTCCGATCTCCACATGGCATTCCACTTTTCACTAAAAGATCGGAAGAGCGT
A6T TS	AGCACAGTAGCTACTCCACATGGCATTCCACTAATCACTAAAAGATCGGAAGAGCG
A7T TS	AGCACAGTAGCTACTCCACATGGCATTCCACATATCACTAAAAGATCGGAAGAGCG
G8C TS	AGCACAGTAGCTACTCCACATGGCATTCCAGTTATCACTAAAAGATCGGAAGAGCG
T9A TS	GTGCTCTTCCGATCTCCACATGGCATTCCTCTTATCACTAAAAGATCGGAAGAGCGT
G10C TS	GTGCTCTTCCGATCTCCACATGGCATTCGACTTATCACTAAAAGATCGGAAGAGCGT
G11C TS	AGCACAGTAGCTACTCCACATGGCATTGCACTTATCACTAAAAGATCGGAAGAGCG
A12T TS	AGCACAGTAGCTACTCCACATGGCATACCACTTATCACTAAAAGATCGGAAGAGCG
A13T TS	GTGCTCTTCCGATCTCCACATGGCAATCCACTTATCACTAAAAGATCGGAAGAGCGT
T14C TS	AGCACAGTAGCTACTCCACATGGCGTTCCACTTATCACTAAAAGATCGGAAGAGCG
G15C TS	AGCACAGTAGCTACTCCACATGGGATTCCACTTATCACTAAAAGATCGGAAGAGCG
C16G TS	GTGCTCTTCCGATCTCCACATGCCATTCCACTTATCACTAAAAGATCGGAAGAGCGT
C16T TS	GTGCTCTTCCGATCTCCACATGACATTCCACTTATCACTAAAAGATCGGAAGAGCGT
C17G TS	AGCACAGTAGCTACTCCACATCGCATTCCACTTATCACTAAAAGATCGGAAGAGCG
A18T TS	AGCACAGTAGCTACTCCACAAGGCATTCCACTTATCACTAAAAGATCGGAAGAGCG
T19A TS	AGCACAGTAGCTACTCCACTTGGCATTCCACTTATCACTAAAAGATCGGAAGAGCG
G20C TS	AGCACAGTAGCTACTCCAGATGGCATTCCACTTATCACTAAAAGATCGGAAGAGCG
T21A TS	GTGCTCTTCCGATCTCCTCATGGCATTCCACTTATCACTAAAAGATCGGAAGAGCGT
2-AP PT P4	CGCTCTTCCGATCTTTTAGTG-2AP-TAAGTGGAATGCCATGTGGAGATCGGAAGAGCAC
2-AP PT P18	CGCTCTTCCGATCTTTTAGTGATAAGTGGAATGCC-2AP-TGTGGAGATCGGAAGAGCAC

Table S1. Table of oligonucleotide sequences. (Related to Figure 3 and Figure 4)

Target DNA substrates were constructed from two fully-complementary oligonucleotides representing the non-target strand (NTS) and target strand (TS). Mismatched sequences were created by inverting the nucleotides of the NTS and TS, except at position 14 where this substitution would have created an undesired PAM.

DNA Substrate	k_{on} ($\times 10^7$ M^{-1} s^{-1})	k_{max} ($\times 10^{-2}$ s^{-1})	k_{off} ($\times 10^{-4}$ s^{-1})	k_{obs}^{NTS} ($\times 10^{-2}$ s^{-1})	k_{obs}^{TS} ($\times 10^{-2}$ s^{-1})
PT	11 ± 1	13 ± 1	0.059 ± 0.004	5.2 ± 0.6	0.51 ± 0.02
G1C	0.32 ± 0.03	2.6 ± 0.3	0.101 ± 0.002	1.7 ± 0.2	0.33 ± 0.02
T2A	0.028 ± 0.002	0.14 ± 0.01	0.018 ± 0.000	0.16 ± 0.02	0.23 ± 0.02
G3C	0.059 ± 0.006	0.45 ± 0.37	0.22 ± 0.01	N.D.	N.D.
A4T	0.019 ± 0.003	0.08 ± 0.04	0.14 ± 0.02	0.069 ± 0.004	0.090 ± 0.009
T5A	0.082 ± 0.024	0.097 ± 0.046	0.28 ± 0.02	N.D.	N.D.
A6T	0.030 ± 0.002	0.056 ± 0.018	0.16 ± 0.01	N.D.	N.D.
A7T	0.040 ± 0.005	0.056 ± 0.019	0.22 ± 0.03	N.D.	N.D.
G8C	0.22 ± 0.02	0.25 ± 0.01	0.36 ± 0.06	N.D.	N.D.
T9A	2.6 ± 0.2	2.9 ± 0.2	1.1 ± 0.2	1.8 ± 0.2	0.42 ± 0.05
G10C	0.64 ± 0.05	1.4 ± 0.3	4.6 ± 0.6	1.1 ± 0.1	0.22 ± 0.01
G11C	0.78 ± 0.10	0.95 ± 0.10	2.2 ± 0.3	N.D.	N.D.
A12T	1.43 ± 0.04	1.57 ± 0.03	2.3 ± 0.4	N.D.	N.D.
A13T	0.28 ± 0.04	0.82 ± 0.26	2.4 ± 0.2	0.81 ± 0.22	0.23 ± 0.02
T14C	3.18 ± 0.03	3.53 ± 0.04	0.16 ± 0.02	N.D.	N.D.
G15C	1.6 ± 0.1	1.3 ± 0.3	2.9 ± 0.2	N.D.	N.D.
C16G	0.18 ± 0.01	0.23 ± 0.05	1.5 ± 0.3	0.26 ± 0.05	0.072 ± 0.006
C16T	1.5 ± 0.1	1.7 ± 0.1	0.89 ± 0.26	2.7 ± 0.4	0.31 ± 0.05
C17G	0.55 ± 0.03	0.60 ± 0.03	1.86 ± 0.04	N.D.	N.D.
A18T	3.2 ± 0.1	3.5 ± 0.1	1.1 ± 0.3	2.3 ± 0.6	0.57 ± 0.07
T19A	6.3 ± 0.1	6.9 ± 0.1	0.70 ± 0.04	N.D.	N.D.
G20C	11 ± 1	14 ± 8	0.47 ± 0.03	N.D.	N.D.
T21A	11 ± 1	15 ± 1	0.032 ± 0.001	3.0 ± 0.1	0.43 ± 0.01

Table S2. Table of measured rates for matched and all mismatched targets. (Related to Figure 3 and Figure 4)

All rates were measured in the presence of 5 mM $MgCl_2$. Values of k_{off} represent the mean ± SEM. Values of k_{on}, k_{max} (representing the R-loop formation rate), and observed rate constants for strand cleavage (k_{obs}^{NTS} and k_{obs}^{TS}) represent the fit parameter ± SE. At least three reaction replicates were performed for each Cas12a concentration and used to perform a weighted hyperbolic fit. PT, matched (perfect) target. N.D., not determined.