Expanded View Figures

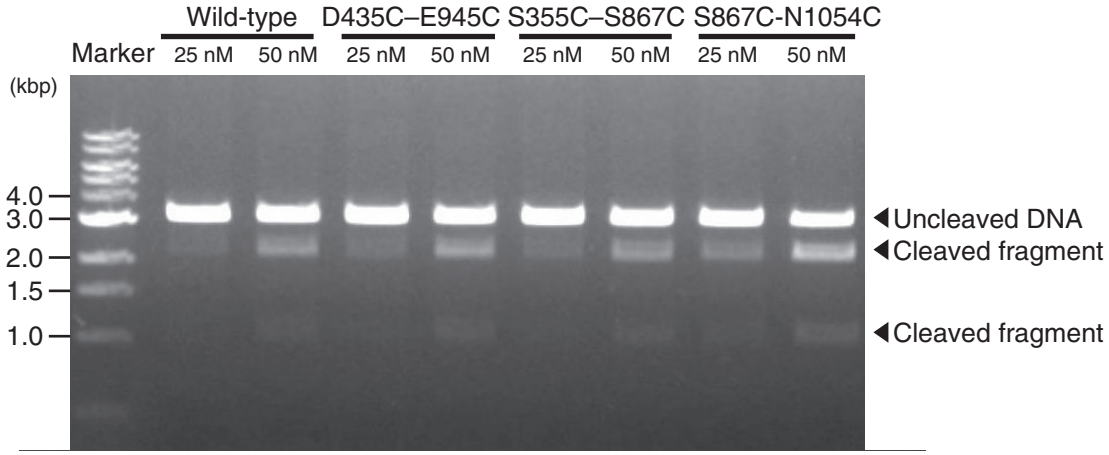

Figure EV1. DNA cleavage activity of fluorescently labeled biotin-Cas9.

All three FRET constructs labeled with Cy3 and Cy5 were tested for nuclease activity. After an incubation of 25 or 50 nM Cas9-sgRNA complex and 5 nM target DNA for 5 min at 37°C, a fraction of the DNA was cleaved into two fragments. Band intensity analysis suggests that the three FRET constructs retain nuclease activity comparable to that of non-labeled wild-type Cas9 (1.1 ± 0.1 for D435C–E945C, 0.9 ± 0.1 for S355C–S867C, and 1.5 ± 0.3 for S867C-N1054C; mean relative activity ± SEM, n = 3).

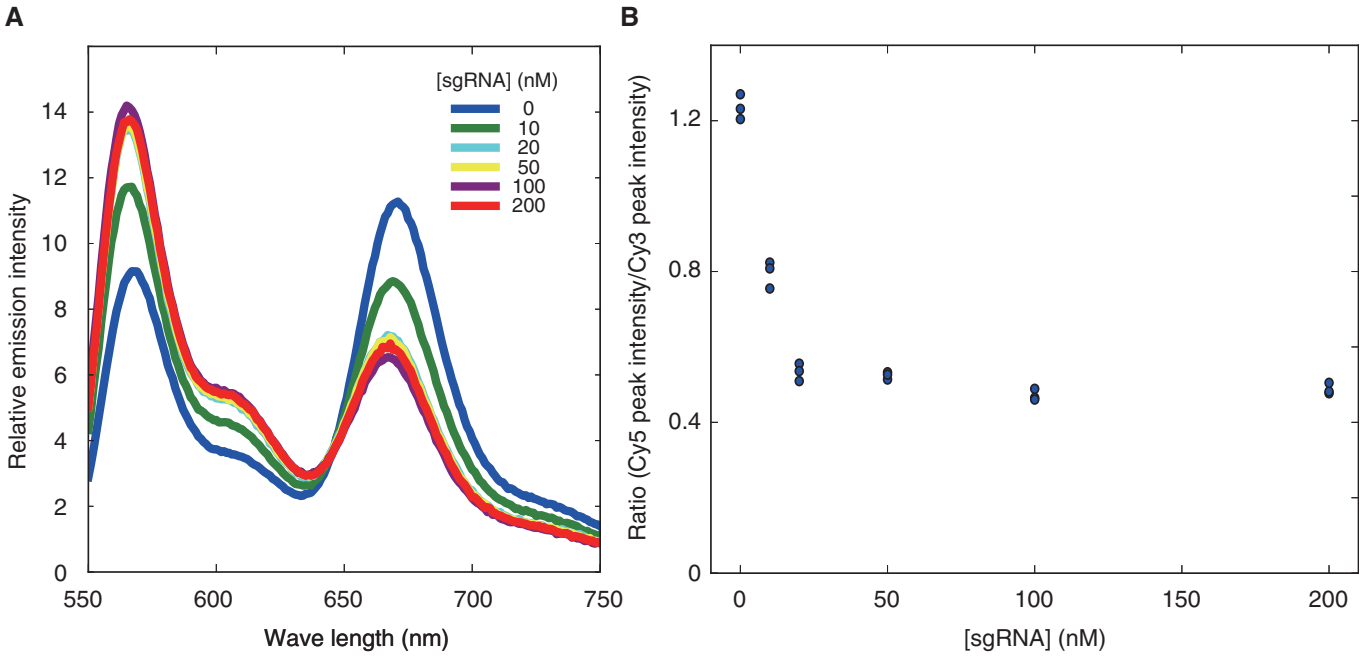

Figure EV2. Stoichiometry of sgRNA binding to Cas9.

A Fluorescence emission spectra of 20 nM fluorescent Cas9 (D435C–E945C) excited at 532 nm. The Cy3 and Cy5 fluorescence intensity changes were coupled with the FRET efficiency change, according to the sgRNA concentration.

B Quantification of the ratio between Cy3 and Cy5 fluorescence intensities. The ratios of Cy5 fluorescence peak intensity over Cy3 fluorescence peak intensity were plotted against the sgRNA concentration (n = 3 for each sgRNA concentration). The FRET efficiency change coupled with the sgRNA binding was almost saturated at the Cas9 to sgRNA ratio of 1:1.

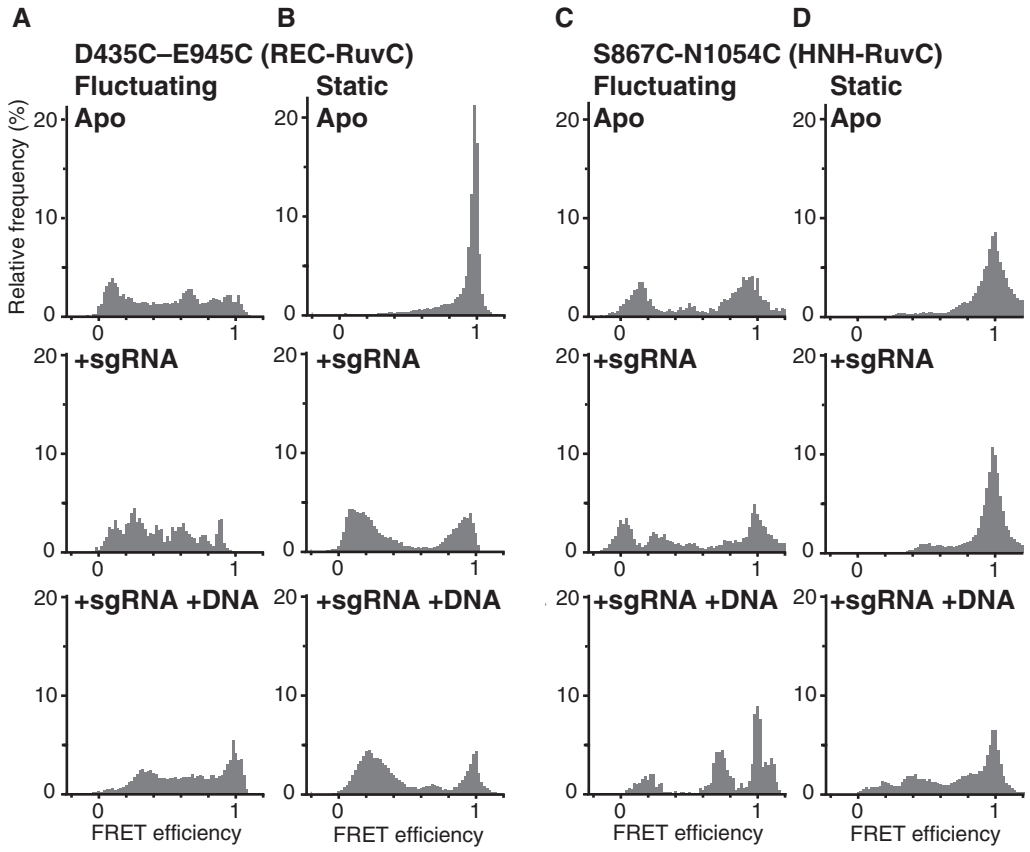

Figure EV3. FRET efficiency histograms of fluctuating and static Cas9 molecules.

A–D FRET efficiency histograms of fluctuating D435C–E945C (A), static D435C–E945C (B), fluctuating S867C-N1054C (C), and static S867C-N1054C (D). The numbers of measured molecules are summarized in Table EV1. The top, middle, and bottom panels show data acquired in the absence of the nucleic acids, in the presence of 200 nM sgRNA, and in the presence of 200 nM sgRNA and 200 nM target DNA, respectively. All of the data in this figure were obtained in assay buffer containing 2 mM $MgCl_2$.

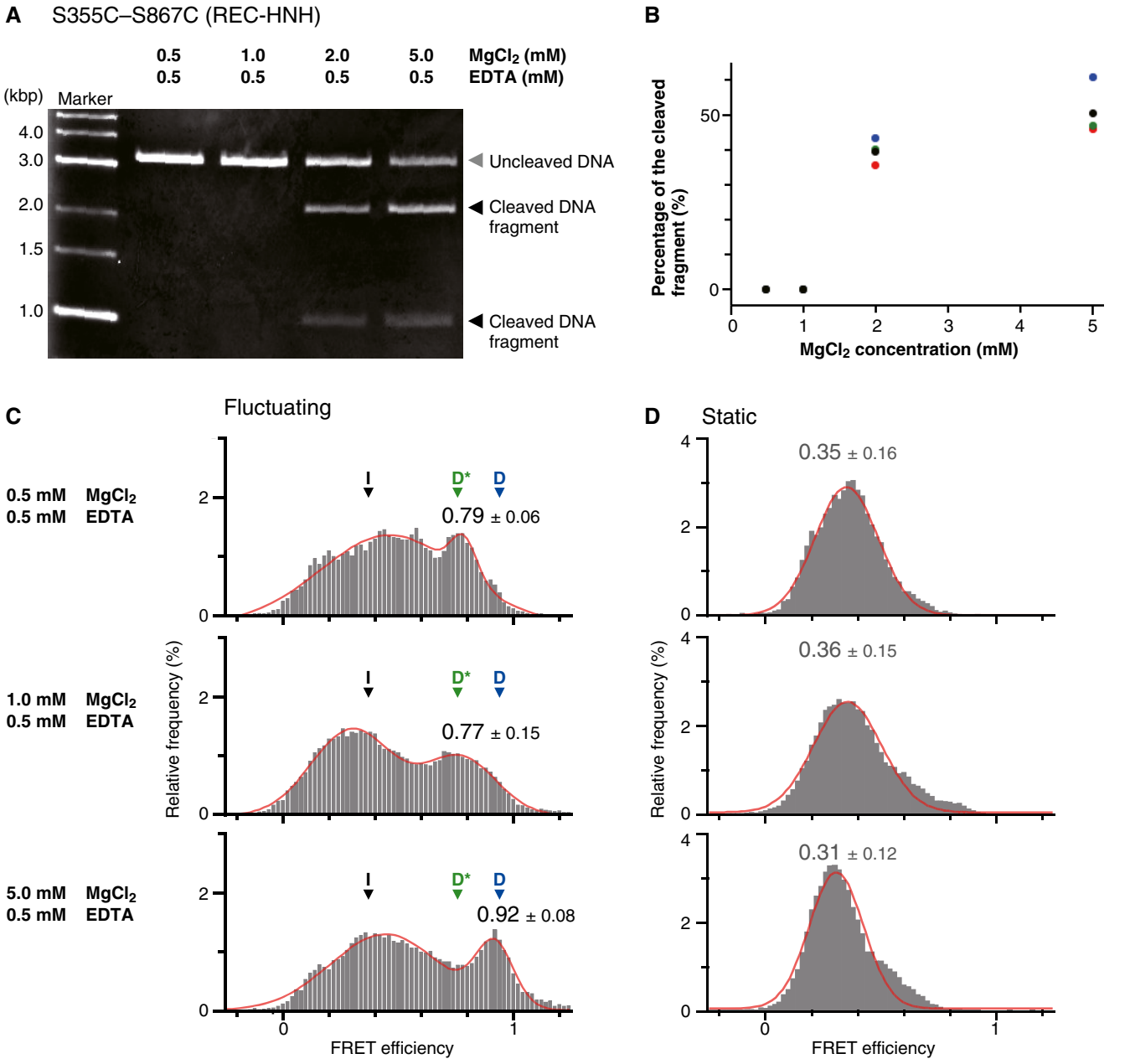

Figure EV4. Effects of Mg^{2+} concentration on the DNA cleavage activity and the HNH location in the sgRNA/DNA-Cas9 ternary complex.

A Representative gel image of the DNA cleavage assay using the fluorescently labeled S355C–S867C construct. The sgRNA/DNA-Cas9 ternary complex was incubated at room temperature (25°C) for 30 min. This condition is equivalent to that of the smFRET measurement, as we observed smFRET for approximately 30–40 min at room temperature.

B Percentages of cleaved DNA against $MgCl_2$ concentration. The plot shows the results of four individual assays (black, blue, green, and red balls) for each $MgCl_2$ concentration. While the ternary complex with 0.5 or 1.0 mM $MgCl_2$ did not cleave the DNA, the complex with 2.0 and 5.0 mM $MgCl_2$ cleaved 39 ± 3% and 51 ± 6% (mean ± SEM, n = 4) of the DNA, respectively.

C, D FRET efficiency histograms of fluctuating (C) and static (D) S355C–S867C molecules. The panels from top to bottom show data in the presence of 0.5, 1, and 5 mM $MgCl_2$. All of the assays were performed in the presence of 0.5 mM EDTA. The low, middle, and high FRET efficiencies corresponding the I, D*, and D positions of the HNH domain are indicated by black, green, and blue arrowheads, respectively (C). The numbers of observed molecules were 57, 72, and 40 for 0.5, 1, and 5 mM $MgCl_2$ conditions in (C), and 57, 55, and 70 for those in (D), respectively. The number of peaks in the histograms was estimated to be two for the histograms in (C) and one for the histograms in (D), using Silverman's test (threshold: P = 0.01). Thus, we fitted the histograms with two-peak (C) or one-peak (D) Gaussian distribution function (red curves). The peak values and widths of the higher FRET peaks were indicated on the histograms (median ± HWHM). The DNA cleavage activity (A, B) correlated well with the appearance of the highest FRET efficiency peak (C), indicating that the HNH domain in the D position adopts the cleavage-competent state.

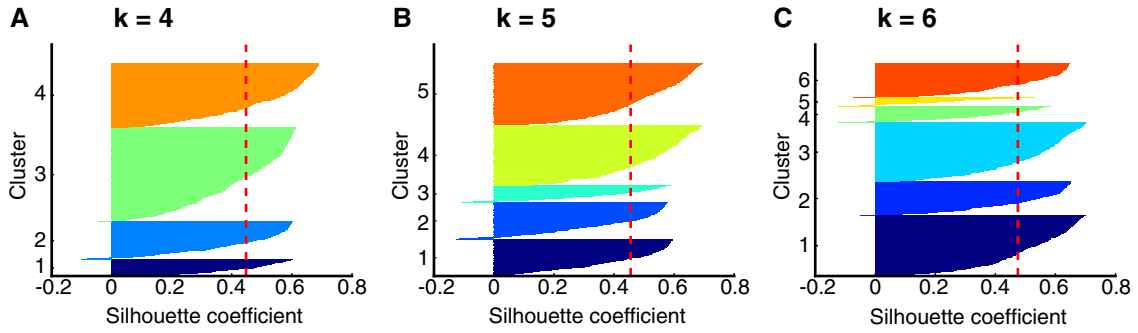

Figure EV5. Silhouette analysis on *k*-means clustering of the FRET efficiency shift to determine the number of clusters.

The Silhouette coefficients, which were calculated using the machine learning Python Package Scikit learn, were plotted for each cluster in the cases of $k = 4$, 5, and 6, respectively (A-C). The vertical red dashed lines indicate the mean value of the Silhouette coefficients. In the cases of $k = 4$ (A) and 5 (B), all clusters showed higher Silhouette coefficients than the mean values. This was not true for $k = 6$, meaning that $k = 5$ is the most probable number of clusters for the transition density plot shown in Fig 5B.