Molecular Cell, Volume *63*

Supplemental Information

DNA Targeting by a Minimal CRISPR RNA-Guided Cascade

Megan L. Hochstrasser, David W. Taylor, Jack E. Kornfeld, Eva Nogales, and Jennifer A. Doudna

Supplemental Figures

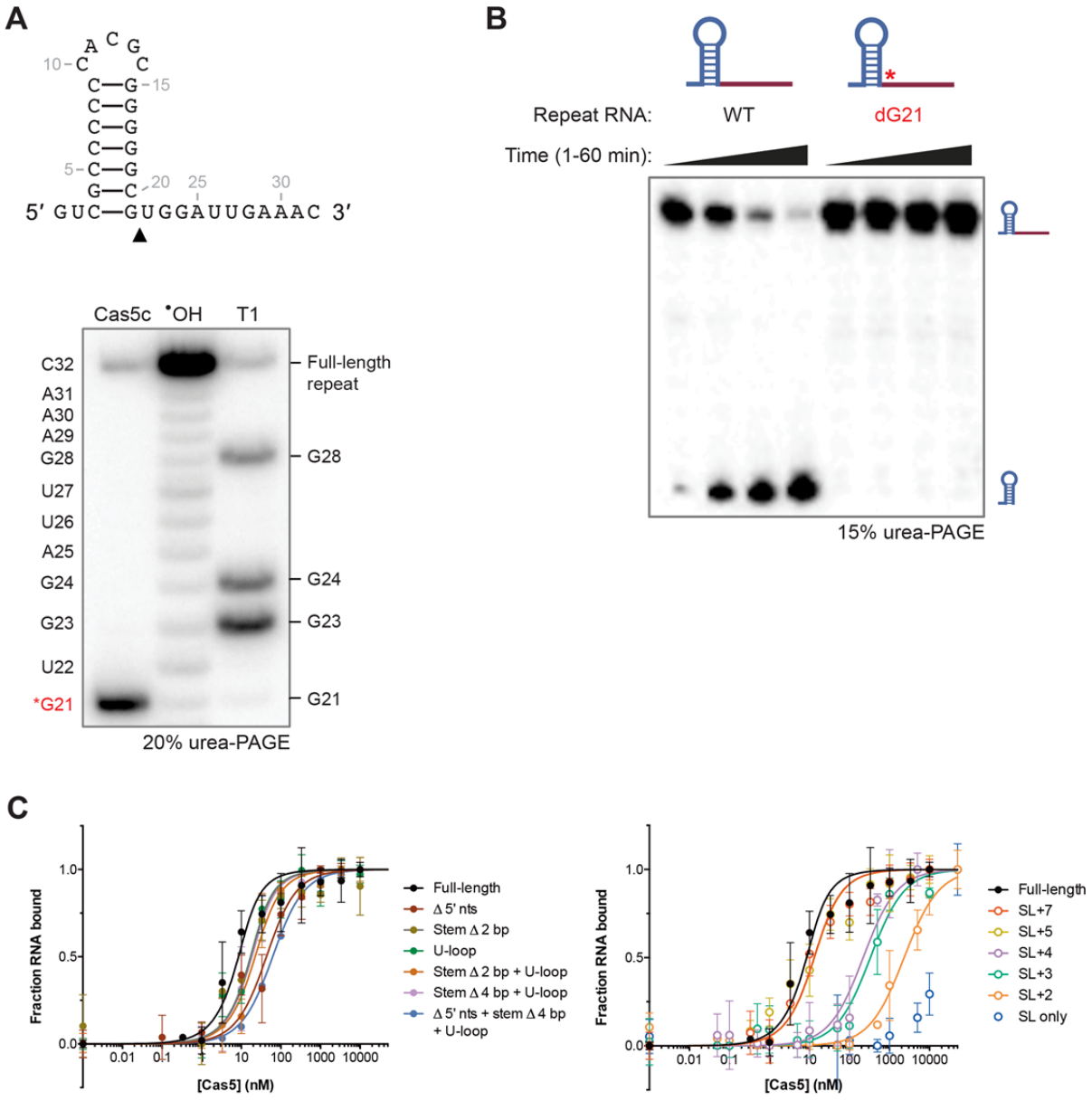

Figure S1. Related to Figure 1. Cas5c cleaves the pre-crRNA repeat at the base of the stem-loop and deoxy substitution at that site abrogates cleavage.
(A) The *D. vulgaris* I-C repeat consensus sequence and its predicted secondary structure (top). Denaturing PAGE comparing the product resulting from Cas5c incubation with the repeat RNA at 37°C to a ladder generated by hydroxyl radical cleavage and RNase T1 digestion products (bottom). The Cas5c cleavage site identified is indicated on the repeat sequence with an arrow.
(B) Denaturing PAGE analysis of radiolabeled pre-crRNA repeat RNAs after incubation with Cas5c at 37°C shows that substitution of a deoxynucleotide 5′ of the cleavage site prevents processing.
(C) Normalized fluorescence polarization binding data for Cas5c binding to variants of the pre-crRNA repeat as shown in Figures 1B and C, but showing all error bars representing ± 1 SD. Data represent at least three independent replicates.

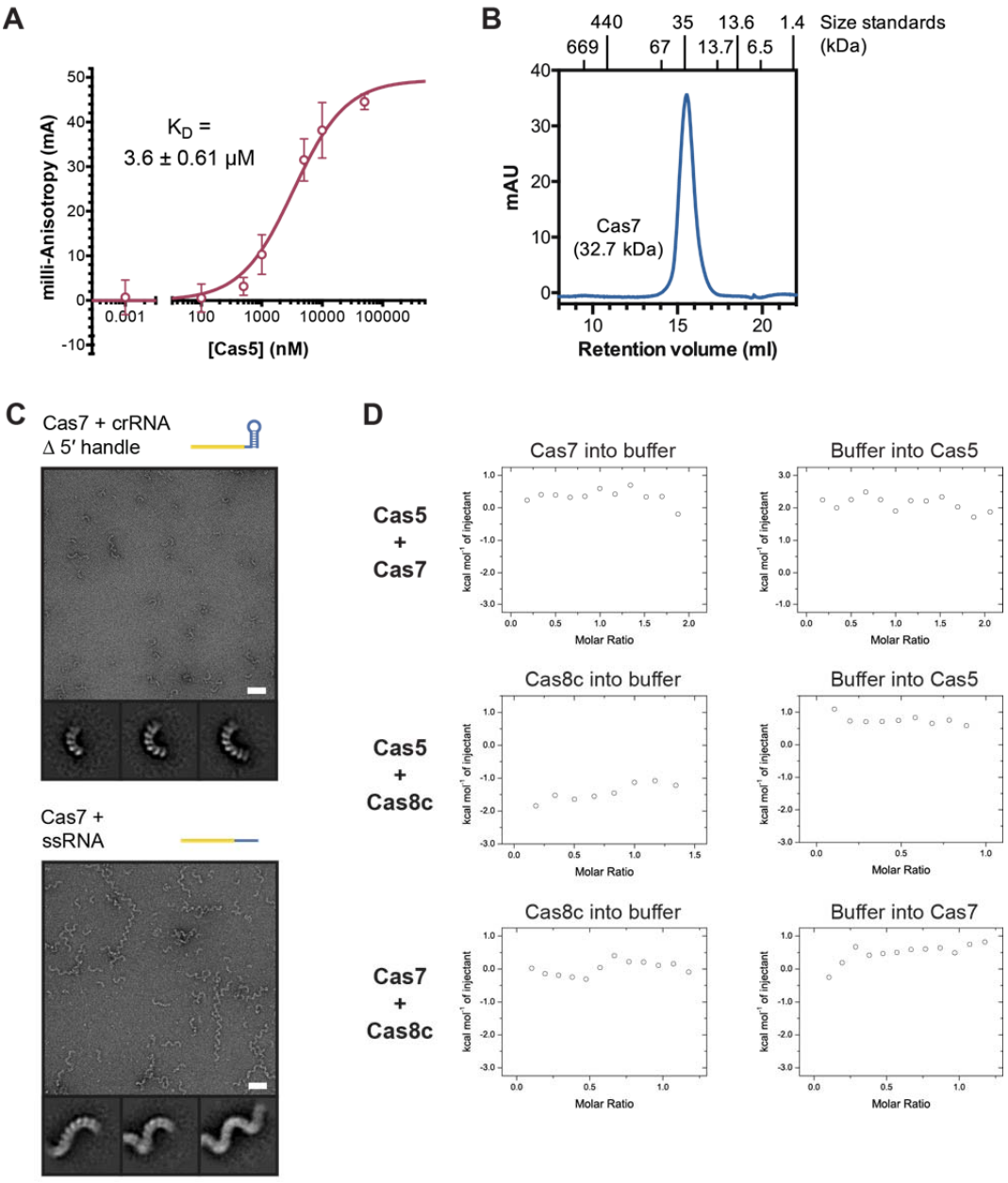

Figure S2. Related to Figures 2 and 3. Cas5c fluorescence polarization control experiment with 3′ fluorescein-labeled 5′ handle RNA, Cas7 elution profile, class averages from RNA-induced oligomerization EM studies, and isothermal titration calorimetry control experiments.

(A) Fluorescence anisotropy data showing binding of Cas5c to a 5′ handle RNA labeled at the 3′ end with fluorescein and the calculated dissociation constant (K_D value). Each point is the average of at least three independent replicates; error bars represent ± 1 SD.

(B) Chromatogram depicting Cas7 elution from a Superdex 200 size exclusion column with typical retention volumes of size standards denoted for comparison.

(C) Sample negative stain electron microscopy micrographs and class averages of Cas7 mixed with different RNA ligands. The width of the boxes in the class averages is 432 Å. Micrograph scale bar, 500 Å.

(D) For each pair of binding partners analyzed in Figure 3, control experiments were performed in which one or the other component was swapped out with buffer. These data indicate that observed experimental heats are caused by binding between two Cascade/I-C subunits rather than mismatched buffers, heats of dilution, oligomer dissociation, or other misleading background signals.

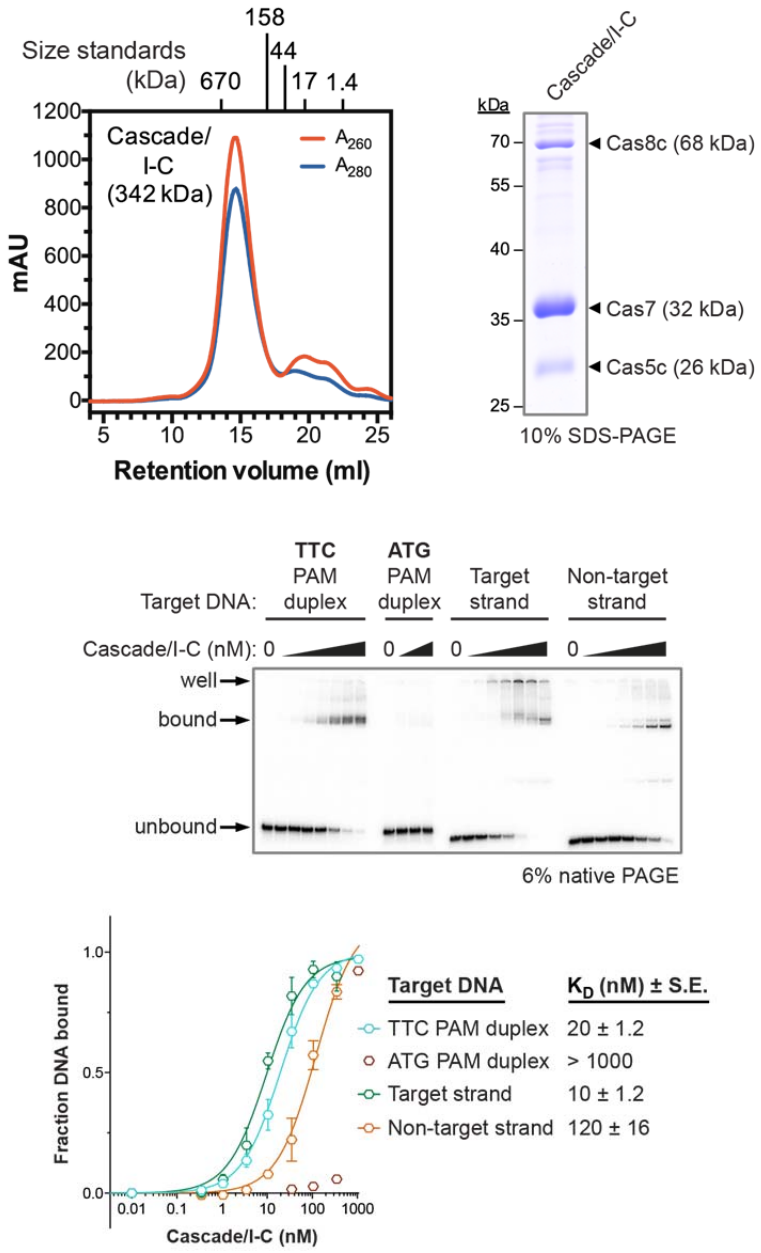

Figure S3. Related to Figure 4. Purification and binding behavior of Cascade/I-C.
Top left, chromatogram depicting Cascade/I-C elution from a Superose 6 10/300 GL size exclusion column with retention volumes of size standards denoted for comparison. Top right, 10% SDS-PAGE analysis of purified Cascade/I-C. Middle, electrophoretic mobility shift assay (EMSA) measuring binding of 0, 0.35, 1.1, 3.5, 11, 35, 110, and 350 nM Cascade/I-C to DNA duplexes containing a matching protospacer and a WT (TTC) PAM (5′ to 3′ on the non-target strand), the target strand, and the non-target strand after incubation at 37°C for 30 min. Only the 0, 35, 110, and 350 nM Cascade/I-C concentrations are shown for a DNA duplex with a mutated (ATG) PAM. Bottom, quantification of gel shift data. Each point is the average of at least three independent replicates; error bars represent ± 1 SD. A binding curve for the mutated PAM dsDNA data is omitted due to its poor fit.

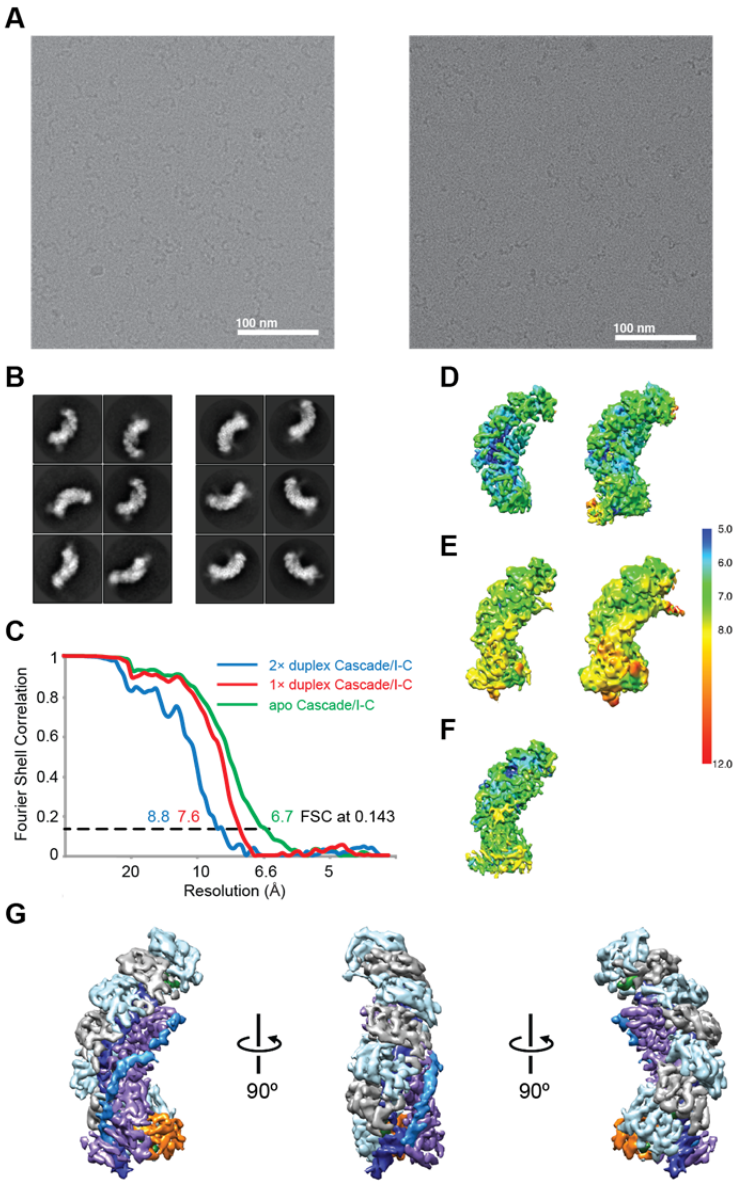

Figure S4. Related to Figure 4. Cryo-EM of Cascade/I-C ribonucleoprotein particles.
(A) A drift-corrected cryo-EM micrograph of apo-Cascade/I-C complexes (left) and dsDNA target-bound Cascade/I-C complexes (right) recorded on a Gatan K2 Summit direct electron detector. The scale bar indicates 100 nm.
(B) Reference-free 2D class averages of apo-Cascade/I-C complexes (left) and target-bound Cascade/I-C complexes (right) showing the caterpillar-like particles with some α-helices and DNA extensions clearly visible. The width of the boxes is ~288 Å.
(C) Fourier shell correlation (FSC) curves for the final reconstructions, showing the resolution to be ~6.7, ~7.6, and 8.8 Å for the apo-, PAM-proximal duplex (1× duplex) target-bound, and full R-loop (2× duplex) target-bound Cascade/I-C structures.
(D–F) Local resolution analysis of apo- (D) 2× duplex DNA target-bound (E) and 1× duplex target-bound (F) Cascade/I-C structures. For D and E two different thresholds are shown.
(G) Cryo-EM reconstruction of 1× duplex target-bound Cascade/I-C at 7.6-Å resolution (using the 0.143 gold standard Fourier Shell Correlation criterion). Subunits are segmented and colored as indicated in Figure 4.

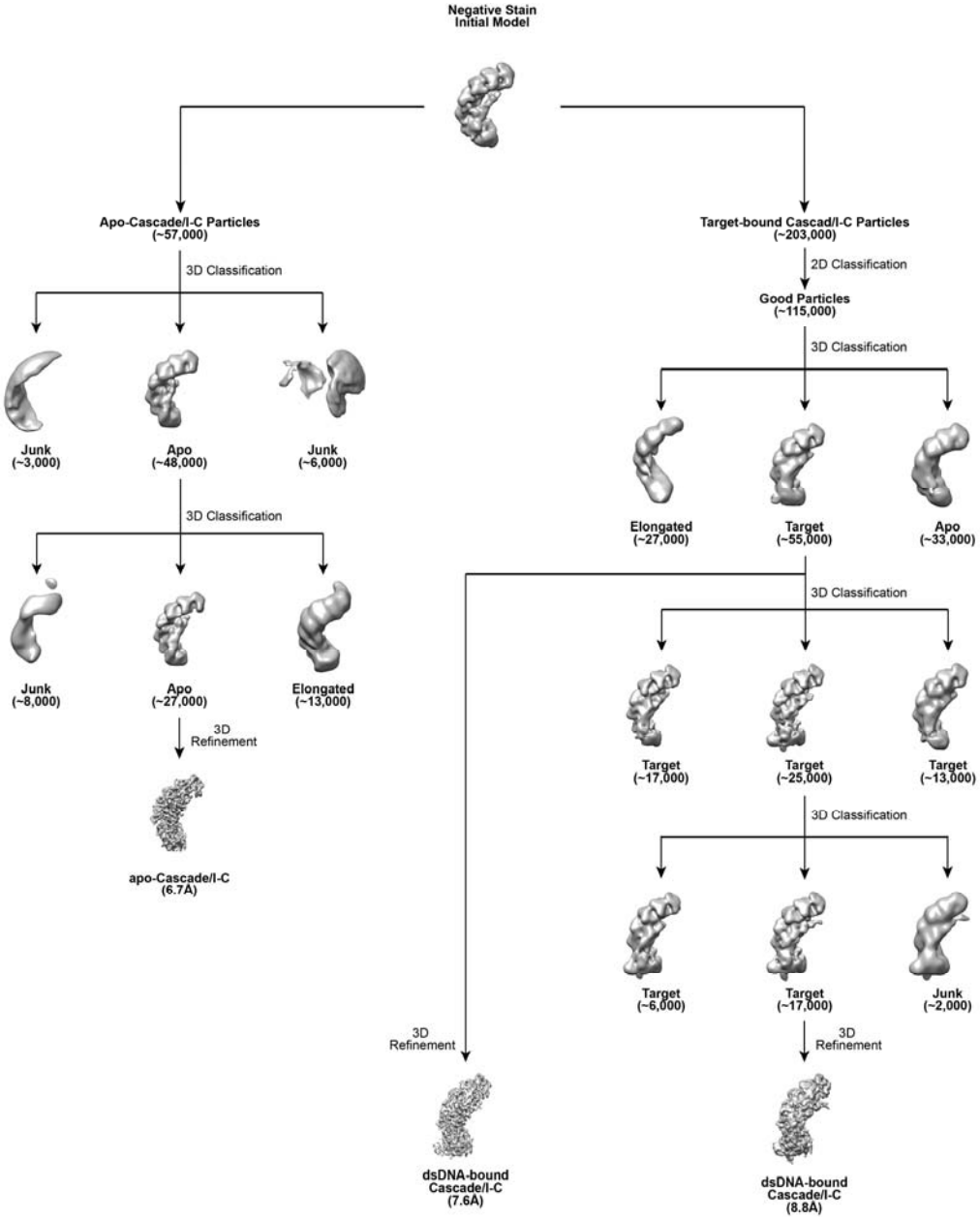

Figure S5. Related to Figure 4. Classification and refinement workflow for Cascade/I-C structures. A total starting stack of ~57,000 apo-Cascade/I-C particles (left) were subjected to 3D classification using 3 classes and a negative stain target-bound Cascade/I-C structure low-pass filtered to 60-Å resolution as a reference. The class with the most particles (~48,000) was subjected to an additional round of 3D classification using its model as a reference. This yielded a final apo-Cascade/I-C dataset of ~27,000 particles, which was subsequently refined to 6.7-Å resolution. A total starting stack of ~203,000 target-bound Cascade/I-C particles (right) were subjected to 2D reference-free alignment and classification. A subset of ~115,000 particles from good 2D classes were selected for 3D classification using 3 classes and a negative stain target-bound Cascade/I-C structure low-pass filtered to 60-Å resolution as a reference. The class with the most particles (~55,000) was subsequently refined to 7.6-Å resolution and showed clear density for the PAM-proximal duplex. To obtain better density for the non-target strand and PAM-distal dsDNA, the same class was subjected to two additional rounds of 3D classification. This yielded a final target-bound Cascade/I-C dataset of ~17,000 particles, which was subsequently refined to 8.8-Å resolution. All steps were performed in RELION (Scheres, 2012).

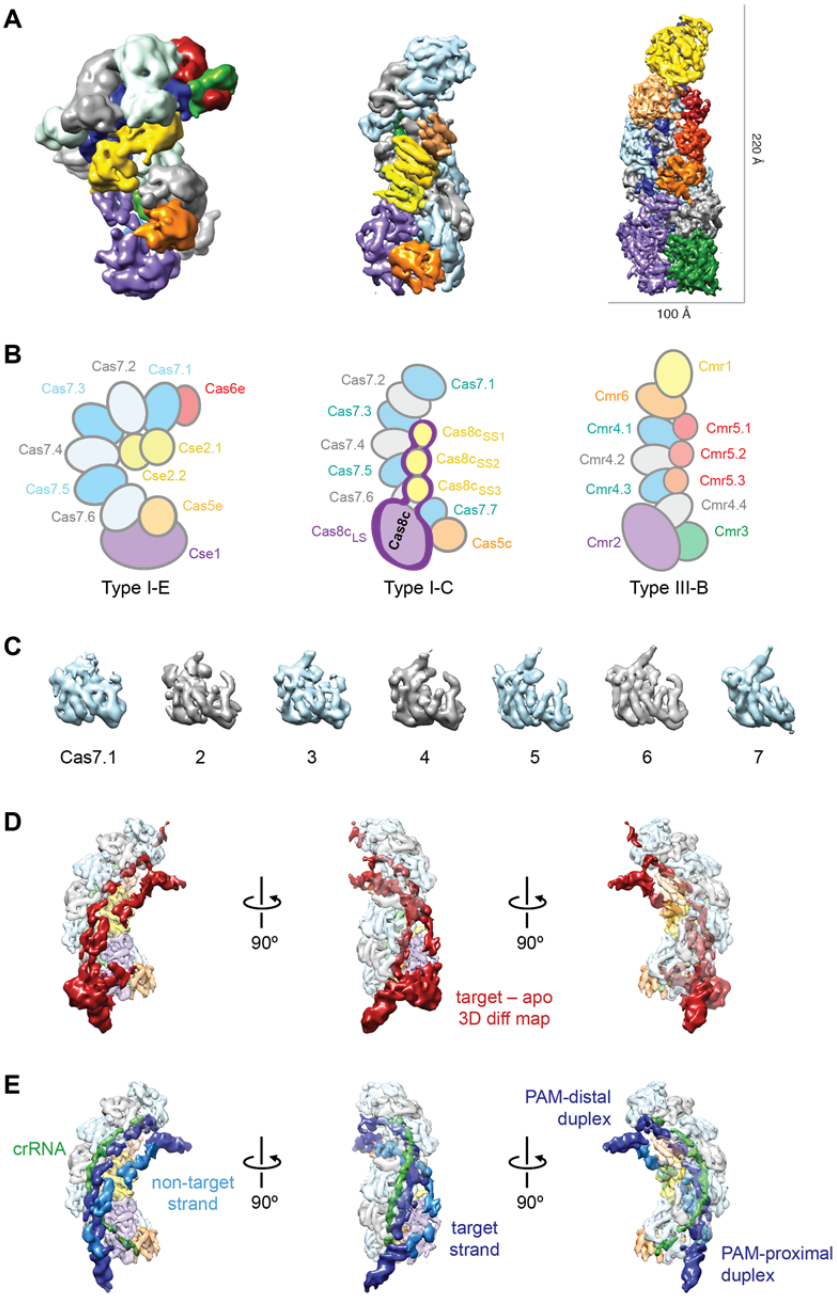

Figure S6. Related to Figure 4 and Figure 5. Analysis of Cascade/I-C structures and comparison to Cascade/I-E and Type III-B Cmr architectures.

(A) Comparison of Cascade/I-E (left), Cascade/I-C (middle), and Type III-B Cmr surveillance complexes. Subunits are segmented and colored as indicated in cartoon below.

(B) Cartoon representation of Cascade/I-E (left), Cascade/I-C (middle), and Type III-B Cmr (right) complexes.

(C) Comparison of Cas7 subunits within apo-Cascade/I-C. The bottom Cas7 subunit (Cas7.7) is strikingly different from the other backbone subunits.

(D) Difference map between intact apo-Cascade/I-C and target-bound Cascade/I-C at 3.5-σ (solid, red) superimposed onto the apo-Cascade/I-C structure (transparent) showing the additional nucleic acid density and the stabilization of a part of Cas8c$_{LS}$.

(E) Segmented difference map showing the full architecture of the R-loop (solid) superimposed onto the apo-Cascade/I-C structure (transparent). Subunits and nucleic acid are colored as in Figure 5.

ID	Description	Sequence (5′ → 3′)	Length (nt)
1	Cas5c TOPO forward primer	CACCATGACACATGGGGCTGTTAAGACATACGGC	34
2	Cas5c reverse primer	TCATGCCCTCACCTCCGGTGACGTGCGC	28
3	Cas8c TOPO forward primer	CACCATGATCCTGCAGGCATTGCATGGCTATTACC	35
4	Cas8c reverse primer	CTAGTTCTCCTTGTTCTTCTTGGTGAAAAGGGCC	34
5	Cas7 TOPO forward primer	CACCATGACCGCCATTGCCAACAGATACGAGTTCG	35
6	Cas7 reverse primer	CTACAGCATCTCTTTGACCTCTACGCCTTCGG	32
7	EcoNI sticky end-T7-repeat top	ATTAGGTAATACGACTCACTATAGGGAGAGTCGCCCCC CACGCGGGGGCGTGGATTGAAAC	61
8	EcoNI sticky end-T7-repeat bottom	CCGCGTGGGGGGCGACTCTCCCTATAGTGAGTCGTATT ACCTAA	44
9	Spacer-repeat top	GCCATGCTCAGGCTGGCGAGTGCGCCACTCATCAAGTC GCCCCCCACGCGGGGGCGTGGATTGAAAC	67
10	Repeat-spacer-repeat bottom	CCGCGTGGGGGGCGACTTGATGAGTGGCGCACTCGCCA GCCTGAGCATGGCGTTTCAATCCACGCCC	67
11	Repeat-KpnI sticky end top	GCCATGCTCAGGCTGGCGAGTGCGCCACTCATCAAGTC GCCCCCCACGCGGGGGCGTGGATTGAAACGGTAC	72
12	Repeat-KpnI sticky end bottom	CGTTTCAATCCACGCCC	17
13	Full-length repeat RNA	GUCGCCCCCCACGCGGGGGCGUGGAUUGAAAC	32
14	Full-length dG21 repeat RNA	GUCGCCCCCCACGCGGGGGC**dG**UGGAUUGAAAC	32
15	5′ fluorescein-dG21 repeat RNA	Fluor-GUCGCCCCCCACGCGGGGGC**dG**UGGAUUGAAAC	32
16	5′ fluorescein-dG21 Δ 5′ nts	Fluor-CGCCCCCCACGCGGGGGC**dG**UGGAUUGAAAC	30
17	5′ fluorescein-dG21 stem Δ 2 bp	Fluor-GUCGCCCCACGCGGGC**d**GUGGAUUGAAAC	28
18	5′ fluorescein-dG21 U-loop	Fluor-GUCGCCCCCUUUUUGGGGGC**d**GUGGAUUGAAAC	32
19	5′ fluorescein-dG21 SL Δ 2 bp + U-loop	Fluor-GUCGCCCUUUUUGGGC**d**GUGGAUUGAAAC	28
20	5′ fluorescein-dG21 SL Δ 4 bp + U-loop	Fluor-GUCGCUUUUUGC**d**GUGGAUUGAAAC	26
21	5′ fluorescein-dG21 Δ 5′ nts + SL Δ 2 bp + U-loop	Fluor-CGCUUUUUGC**d**GUGGAUUGAAAC	24
22	5′ fluorescein-dG21 SL+7	Fluor-GUCGCCCCCCACGCGGGGGC**dG**UGGAUUG	28
23	5′ fluorescein-dG21 SL+5	Fluor-GUCGCCCCCCACGCGGGGGC**dG**UGGAU	26
24	SL+5	GUCGCCCCCCACGCGGGGGCGUGGAU	26
25	5′ fluorescein-dG21 SL+4	Fluor-GUCGCCCCCCACGCGGGGGC**dG**UGGA	25
26	SL+4	GUCGCCCCCCACGCGGGGGCGUGGA	25
27	5′ fluorescein-dG21 SL+3	Fluor-GUCGCCCCCCACGCGGGGGC**dG**UGG	24
28	SL+3	GUCGCCCCCCACGCGGGGGCGUGG	24
29	5′ fluorescein-dG21 SL+2	Fluor-GUCGCCCCCCACGCGGGGGC**dG**UG	23
30	SL+2	GUCGCCCCCCACGCGGGGGCGUG	23
31	SL+1	GUCGCCCCCCACGCGGGGGCGU	22
32	5′ fluorescein-5′ handle	Fluor-UGGAUUGAAAC	11
33	3′ fluorescein-5′ handle	UGGAUUGAAAC-Fluor	11
34	5′ fluorescein-protospacer	Fluor-GAGUAAAAUGUGAAUGAGGAGAGAAUAUAGGAAA	34
35	5′ fluorescein 3′ stem-loop	Fluor-GUCGCCCCCCACGCGGGGGCG	21
36	crRNA	UGGAUUGAAACGAGUAAAAUGUGAAUGAGGAGAGAAUA UAGGAAAGUCGCCCCCCACGCGGGGGCG	66
37	5′ fluorescein-crRNA	Fluor-UGGAUUGAAACGAGUAAAAUGUGAAUGAGGAGAGAAUA UAGGAAAGUCGCCCCCCACGCGGGGGCG	66
38	crRNA Δ 5′ handle	GAGUAAAAUGUGAAUGAGGAGAGAAUAUAGGAAAGUCG CCCCCCACGCGGGGGCG	55
39	5′ fluorescein-crRNA Δ 5′ handle	Fluor-GAGUAAAAUGUGAAUGAGGAGAGAAUAUAGGAAAGUCG CCCCCCACGCGGGGGCG	55
40	crRNA Δ 5′ handle, linearized (ssRNA)	GAGUAAAAUGUGAAUGAGGAGAGAAUAUAGGAAAGUCG CCCCCC	44
41	5′ fluorescein-crRNA Δ 3′ stem-loop	Fluor-UGGAUUGAAACGAGUAAAAUGUGAAUGAGGAGAGAAUA UAGGAAA	45
42	ssRNA annealed to #41 to make “dsRNA”	UUUCCUAUAUUCUCUCCUCAUUCACAUUUUACUCGUUU CAAUCCA	45
43	5′ fluorescein-ssDNA	Fluor-TGGATTGAAACGAGTAAAATGTGAATGAGGAGAGAATA TAGGAAA	45
44	ssDNA annealed to #43 to make “dsDNA”	TTTCCTATATTCTCTCCTCATTCACATTTTACTCGTTT CAATCCA	45
45	Sequence of crRNA co-purified with Cascade/I-C	UGGAUUGAAACGCCAUGCUCAGGCUGGCGAGUGCGCCA CUCAUCAAGUCGCCCCCCACGCGGGGGCG	67
46	Cascade/I-C sp2 WT PAM non-target strand, annealed to #47 to make target duplex	AGCAGACTGGAGGAGTTTTTCGCCATGCTCAGGCTGGCG AGTGCGCCACTCATCAAGCCATGTGGGCTGTCAAAAT	75

47	Cascade/I-C sp2 WT PAM target strand, annealed to #46 to make WT target duplex	ATTTTGACAGCCCACATGGCTTGATGAGTGGCGCACTC GCCAGCCTGAGCATGGCGAAAACTCCTCCAGTCTGCT	75
48	Cascade/I-C sp2 mutated PAM non-target strand, annealed to #49 to make mut. PAM target duplex	AGCAGACTGGAGGAGTTATGGCCATGCTCAGGCTGGCG AGTGCGCCACTCATCAAGCCATGTGGGCTGTCAAAAT	75
49	Cascade/I-C sp2 mutated PAM target strand, annealed to #48 to make mut. PAM target duplex	ATTTTGACAGCCCACATGGCTTGATGAGTGGCGCACTC GCCAGCCTGAGCATGGCCATAACTCCTCCAGTCTGCT	75

Table S1. Related to Figures 1–5. Oligonucleotides used in this study.
Bold indicates a deoxy-guanosine, red indicates a mutated region, blue indicates a protospacer, and underlining indicates a protospacer adjacent motif (PAM).

Plasmid	Name	Description	Oligos
1	Cas5c/pHMGWA	*D. vulgaris cas5c* expression vector with N-terminal His_6-MBP-TEV site tag	1 & 2
2	Cas8c/pHMGWA	*D. vulgaris cas8c* expression vector with N-terminal His_6-MBP-TEV site tag	3 & 4
3	Cas7/pHMGWA	*D. vulgaris cas7* expression vector with N-terminal His_6-MBP-TEV site tag	5 & 6
4	Cas5c-Cas8c-Cas7/pHMGWA	*D. vulgaris Cascade/I-C* expression vector with N-terminal His_6-MBP-TEV site tag	1 & 6
5	Dvu sp2 CRISPR/pACYCDuet-1	*D. vulgaris* CRISPR, 3x genomic Dvu spacer #2 in pACYCDuet-1 expression vector	7–12

Table S2. Related to Figures 1–5. Plasmids used in this study.

Supplemental Movie Legends

Movie S1. Related to Figure 4. Architecture of apo-Cascade/I-C. This movie shows 360° rotations of the apo-Cascade/I-C reconstruction. First, the crRNA and the seven Cas7 subunits are shown; second, the Cas8c subunit is shown; third, Cas8c is removed and $Cas8c_{SS1}$, $Cas8c_{SS2}$, $Cas8c_{SS3}$, $Cas8c_{LS}$, and Cas5c are added to show the overall architecture of the apo-Cascade/I-C complex.

Movie S2. Related to Figure 4. Architecture of dsDNA target-bound Cascade/I-C. This movie shows 360° rotations of the target-bound Cascade/I-C reconstruction. First, the seven Cas7 subunits are shown; second, Cas8c is shown; third, Cas8c is removed and $Cas8c_{SS1}$, $Cas8c_{SS2}$, $Cas8c_{SS3}$, $Cas8c_{LS}$, and Cas5c are added; last, the crRNA:target DNA strand heteroduplex, the non-target DNA strand, the PAM-proximal duplex, and the PAM-distal DNA duplex are added and rotated to show the architecture of the dsDNA target-bound Cascade/I-C complex.

Movie S3. Related to Figure 5. Conformational rearrangements of apo-Cascade/I-C upon target-dsDNA binding and R-loop formation. This movie shows a morph between apo-Cascade/I-C and target-bound Cascade/I-C. First, apo-Cascade/I-C is shown. Next, to emphasize the conformational rearrangements required for binding target dsDNA and formation of an R-loop, the subunits of Cascade/I-C from the apo structure are morphed from their apo- to target-bound positions. Once in the target-bound state, the crRNA:target DNA strand heteroduplex, PAM-proximal duplex, and PAM-distal DNA duplex density are shown in navy blue and the non-target DNA strand density is shown in light blue. The movie then repeats in reverse (from target-bound to the apo state). The morph maps were generated using the morph map function in UCSF Chimera (Pettersen et al., 2004) for each individual subunit. The two maps were aligned based on the density of the middle five Cas7 subunits.

Movie S4. Related to Figure 5. Architecture of the R-loop of Cascade/I-C. This movie shows a 360° rotation of the R-loop difference map in the context of the apo-Cascade/I-C reconstruction. First, the entire apo-Cascade/I-C structure is shown; second, the R-loop from the difference map is added, the apo-Cascade/I-C structure is changed to transparent surface, and the structure is rotated to show the overall architecture of the R-loop.

Supplemental Experimental Procedures

Cloning

Genomic DNA from *Desulfovibrio vulgaris str. Hildenborough* was obtained from the Joint Bioenergy Institute (JBEI). The genes encoding Cas5c (YP_009170.1), Cas8c (YP_009171.1), and Cas7 (YP_009172.1) were PCR amplified as a single operon for Cascade/I-C expression as well as individually using primers (Table S1) that added 'CACC' 5′ of the start codon to allow TOPO cloning into the pENTR/D-TOPO entry vector (Invitrogen). The constructs were then cloned into the pHMGWA destination vector by LR recombination (Invitrogen) and sequence-verified. The pHMGWA expression vector encodes an N-terminal His_6-MBP (maltose binding protein) tag followed by a tobacco etch virus (TEV) protease recognition site.

To generate a plasmid for expressing pre-crRNAs, pACYCDuet-1 (Novagen) was digested with EcoNI and KpnI and a series of partially overlapping oligos (Table S2) were mixed, annealed, and ligated into this vector. The oligos contained appropriate sticky ends, a T7 promoter, and overlapping pieces of the *D. vulgaris* repeat and spacer #2 sequences. Depending on how many copies of the repeat-spacer oligos annealed and were ligated into the vector, different numbers of spacers were cloned into the final plasmid. Several clones were sequenced and we chose a final construct containing a T7 promoter sequence followed by three copies of the spacer sequence, each flanked by repeats.

Protein Purification

Cas5c, Cas8c, and Cas7 were overexpressed in BL21(DE3) Rosetta pLysS cells. For purification of Cascade/I-C, the protein subunits (Cas5c-Cas8c-Cas7) and a synthetic CRISPR array containing three copies of *D. vulgaris* spacer #2 were co-expressed in BL21(DE3) cells. Cells were grown at 37°C to an OD_{600} of approximately 0.5–0.8 prior to induction of protein expression by the addition of 0.5 mM isopropyl-β-D-thiogalactopyranoside (IPTG). After overnight growth at 16°C, cells were harvested and lysed by sonication in buffer containing 50 mM HEPES–NaOH (pH 7.5), 500 mM KCl, 10 mM imidazole, 5% (v/v) glycerol, 1 mM tris(2-carboxyethyl)phosphine (TCEP), 1 mM EDTA, 0.01% Triton X-100, 0.5 mM PMSF, and complete mini protease inhibitor tablets (Roche). Cas8c lysis buffer was identical except it contained 10% (v/v) glycerol and Cascade/I-C lysis buffer contained 100 mM KCl. Lysate was clarified by centrifugation at 27,000 × g and purified using Ni-NTA affinity resin (QIAGEN). The resin was washed with buffer containing 50 mM HEPES–NaOH (pH 7.5), 500 mM KCl, 20 mM imidazole, 5% (v/v) glycerol, and 1 mM TCEP. Cas8c wash buffer was identical except it contained 10% (v/v) glycerol and 1 mM PMSF, and Cascade/I-C wash buffer contained 100 mM KCl. Protein was eluted with 50 mM HEPES–NaOH (pH 7.5), 500 mM KCl, 300 mM imidazole, 5% (v/v) glycerol, and 1 mM TCEP. Cas8c elution buffer was identical except it contained 10% (v/v) glycerol and 1 mM PMSF, and Cascade/I-C elution buffer instead contained 100 mM KCl. Approximately 1 mg of TEV protease was added per 25 mg of protein and the protein-TEV mixture was dialyzed at 4°C overnight against wash buffer. Cas8c was concentrated to <10 ml and incubated with TEV overnight at room temperature and buffer exchanged into wash buffer. Cleaved His_6-MBP tag, uncut protein, and His-TEV were removed by an additional Ni purification.

Cas5c and Cas7 were further purified by size exclusion chromatography using a Superdex 75 16/60 or Superdex 200 10/300 GL column (GE Healthcare), respectively, pre-equilibrated in buffer containing 50 mM HEPES–NaOH (pH 7.5), 500 mM KCl, 5% (v/v) glycerol, and 1 mM TCEP. Cascade/I-C was purified by size exclusion chromatography using a Superose 6 prep grade or 10/300 GL column pre-equilibrated in buffer containing 50 mM HEPES–NaOH (pH 7.5), 500 mM KCl, 5% (v/v) glycerol, and 1 mM TCEP. Cas5c, Cas7, and Cascade/I-C were either buffer exchanged or dialyzed into storage buffer containing 20 mM HEPES–NaOH (pH 7.5), 100 mM KCl, 5% (v/v) glycerol, and 1 mM TCEP. Cas8c was concentrated after the ortho-Ni step and dialyzed into storage buffer containing 20 mM HEPES–NaOH (pH 7.5), 300 mM KCl, 10% (v/v) glycerol, and 1 mM TCEP. All proteins and the Cascade/I-C complex were analyzed for purity by SDS-PAGE. Final stocks were concentrated, flash frozen in liquid nitrogen, and stored at −80°C.

Oligonucleotide Preparation
RNA and DNA oligonucleotides used in cleavage assays, gel shifts, and electron microscopy were purchased from Integrated DNA Technologies (IDT) (Table S1). Oligonucleotides were purified on denaturing gels containing 15% (v/v) 29:1 polyacrylamide, 7M urea, and 1× Tris/Borate/EDTA (TBE). RNA or DNA bands were visualized by UV light, excised, and eluted by soaking gel pieces in deionized H_2O. Gel pieces were filtered out, and DNA was ethanol-precipitated and resuspended in deionized H_2O. DNA duplexes were formed by mixing equimolar amounts of each DNA strand in 40 mM Tris (pH 8.0), 38 mM $MgCl_2$, and 1 mM spermidine, heating at 95°C for 2 min, and slow cooling at room temperature for at least 10 min. Duplexes were resolved on a native gel containing 6% (v/v) 29:1 polyacrylamide and 1× TBE and purified as described above. RNA and DNA samples were 5′ end labeled with [γ-^{32}P] ATP using polynucleotide kinase (PNK, New England Biolabs) for 30 min at 37°C. PNK was heat-denatured at 65°C for 20 min, and excess ATP was removed using a G-25 spin column (GE Healthcare).

RNA Cleavage Assays
All cleavage assays were performed in 1× cleavage buffer [20 mM HEPES–NaOH (pH 7.5), 100 mM KCl, 5% glycerol, and 1 mM TCEP]. 1 μM Cas5c was incubated with 0.2–2.5 nM 5′ radiolabeled RNA at 37°C for 30 min and resolved on denaturing gels containing 15–20% (v/v) 29:1 polyacrylamide, 7 M urea, and 1× TBE. Gels were dried, RNA was visualized by phosphorimaging, and images were quantified using ImageQuant software (GE Healthcare). The percentage of RNA cleaved was determined by dividing cleaved RNA by the sum of uncleaved and cleaved RNA. The background percentage cleaved in control lanes containing no protein was subtracted to generate the corrected percent cleaved. Reported values are the average of at least three independent replicates, and error bars represent the SD.

Cas5c Cleavage Site Mapping
The Cas5c cleavage reactions were performed in 1× cleavage buffer [20 mM HEPES–NaOH (pH 7.5), 100 mM KCl, 5% glycerol, and 1 mM TCEP]. 1 μM Cas5c was incubated with 10 nM 5′ radiolabeled repeat RNA at 37°C for 30 min. The hydrolysis ladder was generated by incubating 10 nM 5′ radiolabeled repeat RNA in 25 mM sodium bicarbonate and 0.25 mM EDTA at 95°C for 10 min prior to quenching on ice. The RNase T1 digest was performed by incubating 1 unit of RNase T1 (ThermoFisher) with 10 nM 5′ radiolabeled repeat RNA in buffer containing 20 mM sodium citrate (pH 5), 400 mM EDTA, and 6.4 M urea at 37°C for 30 min. Samples were resolved on a denaturing gel containing 20% (v/v) 29:1 polyacrylamide, 7 M urea, and 1× TBE. Gels were dried and RNA was visualized by phosphorimaging.

Fluorescence Polarization
Fluorescence polarization assays were performed in 1× binding buffer [20 mM HEPES–NaOH (pH 7.5), 100 mM KCl, 5% (v/v) glycerol, 1 mM TCEP, 10 ug/ml heparin, and 0.01% Igepal CA-630]. Cas8c assays were performed in 1× binding buffer containing 150 mM KCl. 5′ fluorescein-labeled RNAs (IDT) were added to protein solution at a final concentration of 10–20 nM and the RNA-protein mixture was incubated at 37°C for 30 min. Measurements were made by excitation at 485 nm and monitoring emission at 535 nm. All data analysis was performed using Prism software (GraphPad). Data reported as fraction RNA bound have been normalized by setting the lowest anisotropy value to 0 and the highest value to 1. The data were fit using a solution of the quadratic binding equation (all R-square values ≥ 0.87), shown below, to avoid any effects of protein depletion when [RNA] > K_D (Pollard, 2010; Sternberg et al., 2012).

$$Y = \frac{([RNA] + [protein] + K_D) - \sqrt{([RNA] + [protein] + K_D)^2 - (4 \times [RNA] \times [protein])}}{2 \times [RNA]}$$

Data reported in milli-Anisotropy units have been normalized to zero by subtracting average [0 nM] protein value from all points. The data were fit using a standard one-site binding model (all R-square values ≥ 0.87), shown below.

$$Y = \frac{B_{max} \times X}{(K_D + X)}$$

Cas7 data were fit using a one-site binding model with Hill slope to account for the apparent cooperativity observed in raw data (all R-square values ≥ 0.98), shown below.

$$Y = \frac{B_{max} \times X^h}{({K_D}^h + X^h)}$$

Each experiment was performed in at least triplicate and error bars, when included, represent the standard deviation.

Isothermal Titration Calorimetry (ITC)

All ITC experiments were conducted on a MicroCal Auto-iTC$_{200}$ system (GE Healthcare). Purified Cas5c and Cas7 were dialyzed at 4°C against 20 mM HEPES–NaOH (pH 7.5), 100 mM KCl, 5% (v/v) glycerol, and 1 mM TCEP. 100-150 μM Cas7 was titrated into the cell containing 10-30 μM Cas5c using one 0.5 μl injection followed by twelve 3.1 μl injections at 37°C. For Cas8c ITC experiments, all proteins were dialyzed at 4°C against 20 mM HEPES–NaOH (pH 7.5), 300 mM KCl, 10% (v/v) glycerol, and 1 mM TCEP. 175 μM Cas8c was titrated into the cell containing 17.5-30 μM Cas5c and 200 μM Cas8c was titrated into 20-35 μM Cas7 using one 0.5 μl injection followed by twelve 3.1 μl injections at 37°C. Reported values for each interaction are the average of three independent experiments. Modified Origin software (GE Healthcare) was used for integration, baseline correction, and fitting to a one-site binding curve. Control experiments with each protein injected into buffer or buffer injected into each protein were performed to account for heats associated with buffer mismatch, oligomer dissociation, or other factors but only minor background heats were observed.

Electrophoretic Mobility Shift Assays

Gel shift assays were performed in 1× binding buffer [20 mM HEPES–NaOH (pH 7.5), 100 mM KCl, 5% (v/v) glycerol, 1 mM TCEP, 10 μg/ml heparin, and 0.01% Igepal CA-630]. All reaction components were diluted into 1× binding buffer and radiolabeled DNA was added to a final concentration of 0.15–0.5 nM. Samples were incubated at 37°C for 30 min and resolved at 4°C on 6% (v/v) 29:1 polyacrylamide gels containing 1× TBE. Gels were dried, DNA was visualized by phosphorimaging, and images were quantified using ImageQuant software (GE Healthcare). The fraction of DNA bound (amount of bound DNA divided by the sum of free and bound DNA) was plotted versus the concentration of Cascade/I-C and fit to standard one-site binding isotherm (all R-square values ≥ 0.98) using Prism (GraphPad), as described in the “Fluorescence polarization” analysis section. Reported K_D values are the average of at least three independent experiments, and error bars represent the standard deviation. A binding curve is not depicted for a duplex with a mutated PAM because at concentrations of Cascade/I-C 1 μM and higher, the mutated PAM duplex jumps to the fully bound state, as observed in *E. coli* Cascade/I-E (Hochstrasser et al., 2014), which is not fit well by a standard binding curve.

Negative Stain Electron Microscopy

Purified Cas7 was added to RNA oligonucleotides ordered from Integrated DNA Technologies (IDT) at a 10:1 molar ratio of protein to RNA. The mixture was incubated at 37°C for 30 min in 1× storage buffer [20 mM HEPES–NaOH pH 7.5, 100 mM KCl, 5% (v/v) glycerol, and 1 mM TCEP] and then kept on ice until grids were made.

The Cas7 samples were applied to glow-discharged 400-mesh continuous carbon grids. After adsorption for 1 min, we stained the samples consecutively with five droplets of 2% (w/v) uranyl acetate solution, blotted off the residual stain, and air-dried the sample in a fume hood. Data were acquired using a Tecnai F20 Twin transmission electron microscope operated at 120 keV at a nominal magnification of ×80,000 (1.45 Å at the specimen level) using low-dose exposures (~20 $e^{-}Å^{-2}$) with a random defocus ranging from −0.5 to −1.3 µm.

Cryo-Electron Microscopy

Apo-Cascade/I-C was diluted in 1× cryo buffer [20 mM HEPES–NaOH (pH 7.5), 100 mM KCl, and 1 mM TCEP]. dsDNA-bound Cascade/I-C was generated by mixing an approximately equimolar ratio of Cascade/I-C with gel-purified target dsDNA and incubating at 37°C for 15 min in 1× cryo buffer. The sample was centrifuged at 16 × g for 5 min at 4°C and transferred to a new tube to remove any precipitated material. Both apo and dsDNA-bound samples were kept on ice until grids were prepared.

4 µl droplets of the sample (~0.25 $mg•ml^{-1}$) were placed onto C-flat grids immediately after glow-discharging for 8 seconds using a Solaris plasma cleaner. We used C-flat grids with 4 µm holes for apo-Cascade/I-C and 2 µm holes for target-bound Cascade/I-C, with 2 µm spacing between holes (Protochips Inc.) and a thin layer of carbon over the holes. The grids were rapidly plunged into liquid ethane using an FEI Vitrobot MarkIV maintained at 4°C and 100% humidity after being blotted for 4 seconds with a blot force of 15 or 20. For the target bound complex, data were acquired using a FEI Titan Krios transmission electron microscope (at Janelia Research Campus Cryo-EM Shared Resource), operated at 300 keV at a nominal magnification of ×29,000 (1.02 Å at the specimen level), using a defocus ranging from −2.0 to −4.5 µm. For the apo complex, data were acquired using a FEI Titan electron microscope (at UC Berkeley/LBNL) operating at 300 keV at a nominal magnification of ×17,000 (1.05 Å at the specimen level) with defocus ranging from −2.0 to −4.5 µm. A total of ~170 and ~4,200 micrographs of apo-Cascade/I-C and dsDNA target-bound Cascade/I-C were automatically recorded on a Gatan K2 Summit direct electron detector operated in super-resolution mode and counting mode, respectively, using the MSI-Template application within the automated macromolecular microscopy software LEGINON (Suloway et al., 2005). While there is a large difference in the total number of micrographs for each sample, the final structures were obtained from a similar number of particle images (see below). For the apo-Cascade/I-C complexes, we collected total exposures of 18 s, fractionated into 30, 600 ms frames with a dose of 2.25 $e^{-}Å^{-2}s^{-1}$ (for a total dose of 40.5 $e^{-}Å^{-2}s^{-1}$). For the target-bound complexes, we collected total exposures of 6 s, fractionated into 20, 300 ms frames with a dose of 8 $e^{-}Å^{-2}s^{-1}$ (for a total dose of 48 $e^{-}Å^{-2}s^{-1}$).

3D Reconstruction and Analysis

Individual movie frames were aligned and averaged using 'dosef_gpu_driftcorr' drift-correction software from UCSF. These drift-corrected micrographs were binned by 8, and bad micrographs and bad regions of micrographs were removed using the 'manual masking' command within Appion (Lander et al., 2009). All image pre-processing was performed in Appion (Lander et al., 2009). The contrast transfer function (CTF) of each micrograph was estimated using CTFFind3 (Mindell and Grigorieff, 2003). Particles were selected concurrently with data collection with a template-based particle picker using reference-free 2D class averages of negatively stained Cmr complexes as templates (Staals et al., 2013). We selected ~57,000 and ~203,000 apo- and target-bound Cascade/I-C complexes, respectively, using this approach.

We used two different data processing pipelines for the apo- and target-bound Cascade/I-C datasets. For the apo dataset, we performed 3D classification within RELION (Scheres, 2012) into 3 classes, using a negative stain apo-Cascade/I-C structure low-pass filtered to 60 Å as a starting model and selected the best class, which corresponded to the model with the largest number of particles (~48,000), for further processing. The other two classes likely represented ice contaminants and partially assembled complexes. We performed an additional round of 3D classification on this class of ~48,000 particles to obtain a final dataset of ~28,000 'good' particles. Finally, we performed 3D autorefine in RELION using these particles and their corresponding class reconstruction low-pass filtered to 60 Å as a starting model. The final apo-Cascade/I-C structure had a resolution of 6.7 Å after post-processing.

For the target-bound dataset, reference-free 2D alignment and classification was performed in RELION using 100 classes to remove partially assembled particles and ice contaminants, resulting in a starting stack of ~115,000 target-bound Cascade/I-C particles. Using the same negative stain structure of apo-Cascade/I-C low-pass filtered to 60 Å as an initial model, we performed 3D classification using 3 classes within RELION and selected the class with the largest number of particles (~55,000) for further processing. The other two classes likely correspond to the apo-Cascade/I-C complex and elongated chains of Cascade/I-C particles. Next, we performed 3D autorefine in RELION on these particles using their corresponding class structure low-pass filtered to 60 Å as a starting model. The resulting structure had a resolution of 7.6 Å after post-processing and showed only the PAM-proximal duplex end (1x duplex). To further improve the density of the non-target strand and the dsDNA at the PAM-distal end, we performed two additional rounds of 3D classification. A homogenous class of ~17,000 particles emerged from the classification that showed unambiguous density for the non-target strand and PAM-distal DNA duplex. We then carried out 3D autorefine for those particles using RELION to obtain a final structure of target-bound Cascade/I-C at 8.8 Å resolution after post-processing (2x duplex). We used RELION to automatically calculate B-factors and apply them to all post-processed maps. The applied B-factors ranged between –400 and –750 $Å^2$. All reported resolutions are based on the gold standard 0.143 FSC criterion.

Docking and Analysis

The final reconstructions were segmented using Segger in Chimera (Pettersen et al., 2004) based on inspection of the maps and available crystal structures (Hayes et al., 2016; Nam et al., 2012) that we rigidly docked into the maps using Fit in Map. To elucidate the path of the target and non-target DNA strands and to visualize subunit level conformational changes, we performed 3D difference mapping between the apo- and target-bound Cascade/I-C structures. To calculate the difference density corresponding to the nucleic acid, we created a 'Franken'-apo map by fitting each of the apo subunits into the corresponding target-bound subunits and adding the volumes together using the Chimera command 'vop add' to create a single map. The difference map was calculated by subtracting the 'Franken'-apo structure from the target-bound structure after normalization of the two maps using SPIDER (Frank et al., 1996). All figures and movies were generated using UCSF Chimera (Pettersen et al., 2004).

Supplemental References

Frank, J., Radermacher, M., Penczek, P., Zhu, J., Li, Y., Ladjadj, M., and Leith, A. (1996). SPIDER and WEB: processing and visualization of images in 3D electron microscopy and related fields. Journal of Structural Biology *116*, 190–199.

Hayes, R.P., Xiao, Y., Ding, F., van Erp, P.B.G., Rajashankar, K., Bailey, S., Wiedenheft, B., and Ke, A. (2016). Structural basis for promiscuous PAM recognition in type I-E Cascade from E. coli. Nature *530*, 499–503.

Hochstrasser, M.L., Taylor, D.W., Bhat, P., Guegler, C.K., Sternberg, S.H., Nogales, E., and Doudna, J.A. (2014). CasA mediates Cas3-catalyzed target degradation during CRISPR RNA-guided interference. PNAS *111*, 6618–6623.

Lander, G.C., Stagg, S.M., Voss, N.R., Cheng, A., Fellmann, D., Pulokas, J., Yoshioka, C., Irving, C., Mulder, A., Lau, P.-W., et al. (2009). Appion: An integrated, database-driven pipeline to facilitate EM image processing. Journal of Structural Biology *166*, 95–102.

Nam, K.H., Haitjema, C., Liu, X., Ding, F., Wang, H., DeLisa, M.P., and Ke, A. (2012). Cas5d protein processes pre-crRNA and assembles into a cascade-like interference complex in subtype I-C/Dvulg CRISPR-Cas system. Structure *20*, 1574–1584.

Pettersen, E.F., Goddard, T.D., Huang, C.C., Couch, G.S., Greenblatt, D.M., Meng, E.C., and Ferrin, T.E. (2004). UCSF Chimera--A visualization system for exploratory research and analysis. J. Comput. Chem. *25*, 1605–1612.

Pollard, T.D. (2010). A Guide to Simple and Informative Binding Assays. Molecular Biology of the Cell *21*, 4061–4067.

Scheres, S.H.W. (2012). RELION: implementation of a Bayesian approach to cryo-EM structure determination. Journal of Structural Biology *180*, 519–530.

Staals, R.H.J., Agari, Y., Maki-Yonekura, S., Zhu, Y., Taylor, D.W., van Duijn, E., Barendregt, A., Vlot, M., Koehorst, J.J., Sakamoto, K., et al. (2013). Structure and Activity of the RNA-Targeting Type III-B CRISPR-Cas Complex of Thermus thermophilus. Molecular Cell *52*, 135–145.

Sternberg, S.H., Haurwitz, R.E., and Doudna, J.A. (2012). Mechanism of substrate selection by a highly specific CRISPR endoribonuclease. RNA *18*, 661–672.

Suloway, C., Pulokas, J., Fellmann, D., Cheng, A., Guerra, F., Quispe, J., Stagg, S., Potter, C.S., and Carragher, B. (2005). Automated molecular microscopy: The new Leginon system. Journal of Structural Biology *151*, 41–60.