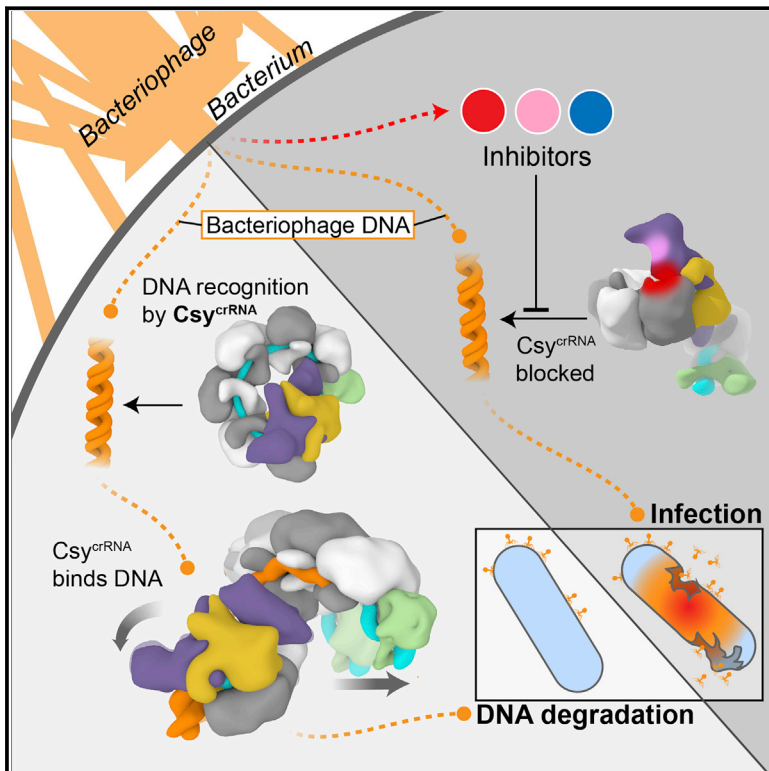


Cryo-EM Structures Reveal Mechanism and Inhibition of DNA Targeting by a CRISPR-Cas Surveillance Complex

Graphical Abstract



Authors

Tai Wei Guo, Alberto Bartesaghi, Hui Yang, ..., Lesley A. Earl, Dinshaw J. Patel, Sriram Subramaniam

Correspondence

subramas@mail.nih.gov

In Brief

Single-particle cryo-EM structures of a type I-F CRISPR surveillance complex before and after target dsDNA binding, as well as after inhibitor binding, demonstrate unique structural features for dsDNA recognition and a large global elongation of the complex, which may facilitate nuclease recruitment and subsequent target degradation.

Highlights

- Cryo-EM structures of Csy surveillance complex with and without bound dsDNA
- dsDNA target binding induces change in pitch of Csy complex
- Visualization of PAM recognition and spacer:protospacer heteroduplex formation
- Binding site definition of anti-CRISPR inhibitors that block dsDNA binding

Data Resources

6B44
6B45
6B46
6B47
6B48



Cryo-EM Structures Reveal Mechanism and Inhibition of DNA Targeting by a CRISPR-Cas Surveillance Complex

Tai Wei Guo,^{1,6} Alberto Bartesaghi,^{1,6} Hui Yang,^{2,6} Veronica Falconieri,¹ Prashant Rao,¹ Alan Merk,¹ Edward T. Eng,³ Ashleigh M. Raczowski,³ Tara Fox,^{4,5} Lesley A. Earl,¹ Dinshaw J. Patel,² and Sriram Subramaniam^{1,4,7,*}

¹Laboratory of Cell Biology, Center for Cancer Research, National Cancer Institute, NIH, Bethesda, MD 20892, USA

²Structural Biology Program, Memorial Sloan Kettering Cancer Center, New York, NY 10065, USA

³Simons Electron Microscopy Center, New York Structural Biology Center, New York, NY 10027, USA

⁴Center for Molecular Microscopy, Center for Cancer Research, National Cancer Institute, NIH, Bethesda, MD 20892, USA

⁵Cancer Research Technology Program, Frederick National Laboratory for Cancer Research, Frederick, MD 21701, USA

⁶These authors contributed equally

⁷Lead Contact

*Correspondence: subramas@mail.nih.gov
<http://dx.doi.org/10.1016/j.cell.2017.09.006>

SUMMARY

Prokaryotic cells possess CRISPR-mediated adaptive immune systems that protect them from foreign genetic elements, such as invading viruses. A central element of this immune system is an RNA-guided surveillance complex capable of targeting non-self DNA or RNA for degradation in a sequence- and site-specific manner analogous to RNA interference. Although the complexes display considerable diversity in their composition and architecture, many basic mechanisms underlying target recognition and cleavage are highly conserved. Using cryoelectron microscopy (cryo-EM), we show that the binding of target double-stranded DNA (dsDNA) to a type I-F CRISPR system *yersinia* (Csy) surveillance complex leads to large quaternary and tertiary structural changes in the complex that are likely necessary in the pathway leading to target dsDNA degradation by a *trans*-acting helicase-nuclease. Comparison of the structure of the surveillance complex before and after dsDNA binding, or in complex with three virally encoded anti-CRISPR suppressors that inhibit dsDNA binding, reveals mechanistic details underlying target recognition and inhibition.

INTRODUCTION

CRISPR and the CRISPR-associated Cas proteins provide bacteria and archaea an adaptable defense against a broad spectrum of invading phages and plasmids (Makarova et al., 2013; Marraffini, 2015). These CRISPR-Cas systems recognize and destroy previously encountered non-self DNA or RNA sequences, functioning as an adaptive “immune system.” Because of the sequence specificity of these complexes,

CRISPR surveillance complexes have been leveraged for use in diverse applications including gene editing to alter genome sequences (Mohanraju et al., 2016; Wright et al., 2016). Although CRISPR-Cas surveillance systems can be found across a broad spectrum of bacteria and archaea, and many of these systems share similar architectural features, there is a wide variation in the sequence and composition of these complexes, which comprise two broad classes, divided into 6 types and 19 subtypes (Makarova et al., 2015; Mohanraju et al., 2016). The type I multi-subunit system is the largest and is further divided into seven subtypes (type I-A through I-F and type I-U).

The structural studies we describe here are focused on the type I-F Csy (CRISPR system *yersinia*) found in *Pseudomonas aeruginosa*, for which phage-derived inhibitors have been newly discovered (Bondy-Denomy et al., 2013, 2015; Pawluk et al., 2016). As with all type I systems, Csy recognizes target DNA sequences via a multi-subunit surveillance complex (Csy complex). In the adaptation phase, a foreign nucleic acid sequence, or “protospacer” (e.g., from a phage or plasmid), is recognized, cleaved, and inserted into the bacterial genome at the CRISPR locus, generating a new “spacer” flanked by repeats. In the second phase (the expression phase), this spacer is transcribed and processed into a CRISPR RNA (crRNA, also called a “guide RNA”) and incorporated into the Csy surveillance complex. In the third phase (the interference phase), the bound crRNA provides a template for the surveillance complex to bind and degrade complementary target nucleic acid sequences. Target recognition first requires detection of a short 2–5 base pair “protospacer adjacent motif” (PAM) that is proximal to the protospacer element in the target sequence (Shah et al., 2013). The PAM itself is not incorporated into the CRISPR locus, allowing the CRISPR-Cas systems to discriminate between self and non-self. PAM recognition is followed by the unwinding of the target dsDNA and hybridization of the target DNA strand with the crRNA spacer, forming an R-loop to displace the non-target DNA strand (Rutkauskas et al., 2015; Szczelkun et al., 2014). Formation of the R-loop in both Csy and the related type I-E Cse

complexes triggers recruitment of the endonuclease Cas3, which degrades the target dsDNA (Rollins et al., 2017; van Duijn et al., 2012).

In a recent publication, Chowdhury et al. (2017) have provided a first glimpse into the molecular architecture of the Csy complex with the determination of a 3.4 Å resolution cryoelectron microscopy (cryo-EM) structure of Csy bound simultaneously to the inhibitors AcrF1 and AcrF2. To understand the mechanistic basis underlying recognition of DNA by Csy, however, it is necessary to determine the structures of the complex before and after DNA binding and to discern exactly how the binding of various inhibitors interferes with DNA binding. In this work, we bridge this gap in understanding by determining cryo-EM structures for the Csy complex in its pre-target-bound state (Csy^{crRNA}), in complex with dsDNA (Csy^{crRNA}·dsDNA), and when it is individually bound to inhibitors AcrF1 (Csy^{crRNA}·AcrF1), AcrF2 (Csy^{crRNA}·AcrF2), and AcrF10 (Csy^{crRNA}·AcrF10). In addition, we also present crystal structures for the AcrF1 and AcrF10 inhibitors. Our work provides structural snapshots of the Csy complex in a variety of functional and inhibited states and leads to a better mechanistic understanding of similarities and differences between Csy and other related type I surveillance complexes (Hayes et al., 2016; Hochstrasser et al., 2016; Xiao et al., 2017).

RESULTS

General Architecture of Csy Complex Bound to dsDNA

We begin with a description of our cryo-EM structure of the type I-F Csy surveillance complex containing a 60-nucleotide (nt) crRNA bound to a DNA target with 17 base pair of PAM-proximal duplex DNA. The double-stranded DNA (dsDNA) includes the G-G/C-C PAM, a 32-nt protospacer complementary to the 32-nt crRNA spacer, and a 10-nt overhang at the 3' end of the non-target DNA strand (Figure 1A). Multiple representations of the 2.9 Å resolution cryo-EM structure of Csy^{crRNA}·dsDNA are shown in Figure 1, including a 2D schematic rendering (Figure 1B), a ribbon diagram (Figure 1C), a surface representation (Figure 1D), and the animation in Movie S1. Technical details relevant to cryo-EM structure determination are included in the STAR Methods and in Figures S1 and S2.

As expected, the general arrangement of the 9 polypeptides in the Csy^{crRNA}·dsDNA is the same as that reported in the cryo-EM structure of Csy^{crRNA} bound simultaneously to inhibitors AcrF1 and AcrF2 by Chowdhury et al. (2017) and by Peng et al. (2017), with the protein subunits arranged around the crRNA in a helically twisted “G” shape (Figures 1C and 1D). Cas6f is located at the 3' stem-loop of the crRNA; Cas8f and Cas5f are located at the 5' handle of crRNA; six interlocking copies of Cas7f are located along the length of the crRNA spacer. Following the convention of earlier publications in the field (Chowdhury et al., 2017; Wiedenheft et al., 2011), the “head” and “tail” of the complex are oriented at the crRNA 3' stem loop and 5' handle, respectively, with a “backbone” region in the middle of the complex. The density map of the Csy^{crRNA}·dsDNA complex displays variations in local resolution (Figure S2B). The density of the backbone region, lined by Cas7f subunits, and the tail region, including Cas5f and Cas8f, is well-defined and could be reliably modeled (Figures 2A and 2B), but no clear density is observed for Cas6f

or the crRNA 3' stem loop at the PAM-distal head, similar to what was reported in the recent cryo-EM studies of the Csy^{crRNA}·AcrF1·AcrF2 complex by Chowdhury et al. (2017) and Peng et al. (2017). For the sake of completeness and consistency, we have included the previously reported atomic model from the crystal structure of Cas6f (PDB ID: 4AL5) (Haurwitz et al., 2012) in the structure shown in Figure 1, with the minor modifications introduced in the cryo-EM structure of the Csy^{crRNA}·AcrF1·AcrF2 complex (PDB ID: 5UZ9) by Chowdhury et al. (2017).

The base pairing between the target DNA and the crRNA threaded through the interlocking Cas7f subunits is clearly visualized in the cryo-EM density map (Figures 2C and 2D). The binding of the target strand to crRNA follows a periodic “5+1” pattern, a recurring motif in class 1 complexes in which 5 consecutive base pairs in pseudo-A conformation are followed by a 1 base pair gap of two unpaired nucleotides that kink out in opposite directions (Hayes et al., 2016; Hochstrasser et al., 2016; Mulepati et al., 2014; Osawa et al., 2015; Xiao et al., 2017). A β-hairpin emanating from the adjacent Cas7f subunit threads through each of these gaps spaced along the backbone, effectively pinning the RNA:DNA heteroduplex to the Csy complex backbone (Figures 2E and S3A). These 5 base pair segments of RNA:DNA hybrid are nearly identical: superimposition of the sugar-phosphate backbone of a central segment (crRNA spacer nucleotides 13–17) on every other segment results in an root-mean-square deviation (RMSD) of 0.3–0.8 Å. The G-G/C-C PAM is situated immediately before bifurcation of the DNA duplex.

PAM Recognition

The recognition of target dsDNA by the Csy complex can be roughly divided into three steps: PAM search, PAM recognition, and target verification (Rollins et al., 2015). In PAM search, the Csy complex randomly samples dsDNA, with rapid association and disassociation between the dsDNA and surveillance complex. In PAM recognition, a weak interaction between the complex and the PAM site temporarily stabilizes the Csy^{crRNA}·dsDNA complex and locally destabilizes the adjacent DNA duplex. The crRNA then hybridizes with the target strand of the DNA duplex in the target verification step, initiating formation of the R-loop.

In the Csy^{crRNA}·dsDNA structure, which includes a partial R-loop (detailed in Figure 2F), the duplex region of the target DNA (including the G-G/C-C PAM) is sandwiched between the Cas8f N-terminal hook domain (residues 1–166), the Cas5f thumb domain (residues 48–109), and Cas7.6f subunit (Figure 3A). Density for ~10 base pair of the PAM-containing DNA duplex is visible, suggesting that the rest of the duplex, which is further away from the PAM site, is likely flexible and either not in direct contact with the Csy complex or at least does not display a single bound conformation. Numerous positively charged residues line this DNA binding pocket, including residues from Cas7.6f (K299), the Cas5f thumb (R90), and especially the Cas8f hook (R24, K28, K31, R59, K71, and R78) (Figure 3B). Although the density map in this region is not resolved at a level of detail to unambiguously discern the atomic interactions between these residues and the DNA, the matched curvature of the region's electrostatic surface to that of DNA strongly suggests that interactions between these residues and the

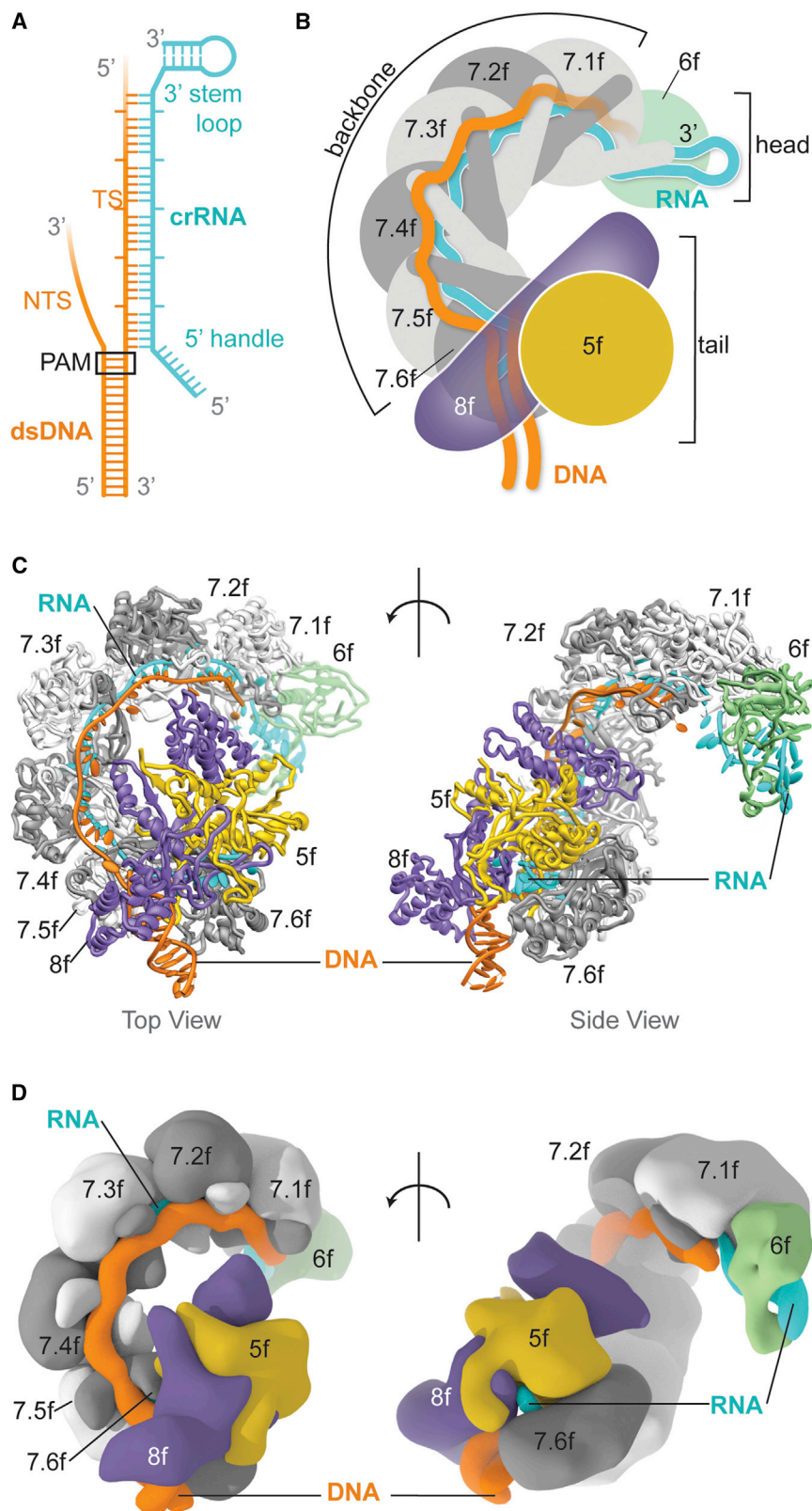


Figure 1. Overall Structure of Csy^{crRNA}.dsDNA

(A) Schematic representation of the partial R-loop in the type I-F Csy surveillance complex from *Pseudomonas aeruginosa* bound to dsDNA (Csy^{crRNA}.dsDNA). The dsDNA is forked, with the target strand (TS) hybridizing to crRNA, and the non-target strand (NTS) displaced.

(B) Schematic rendering of the arrangement of subunits in Csy^{crRNA}.dsDNA. The head of the complex is composed of the crRNA 3' stem loop and Cas6f, while the tail is composed of the crRNA 5' handle, Cas5f, and Cas8f. Six repeating and interlocked Cas7f subunits together with the RNA:DNA heteroduplex make up the helical backbone. Duplex DNA is positioned between Cas8f and Cas7.6f.

(C and D) Cryo-EM-derived atomic model shown as a ribbon diagram (C) and surface rendering (D) of Csy^{crRNA}.dsDNA from two different views (top and side). Shown are Cas6f (green), six identical units of Cas7f (gray), Cas5f (yellow), Cas8f (purple), crRNA (cyan), and the forked dsDNA (orange). Information from the X-ray structure of Cas6f (PDB ID: 4AL5) (Haurwitz et al., 2012), modified as in the previous cryo-EM study (PDB ID: 5ZU9) (Chowdhury et al., 2017) was used to model Cas6f coordinates in the Csy^{crRNA}.dsDNA complex because no clear density was observed in this region in our density map. See also Figures S1, S2, and S3 and Movie S1.

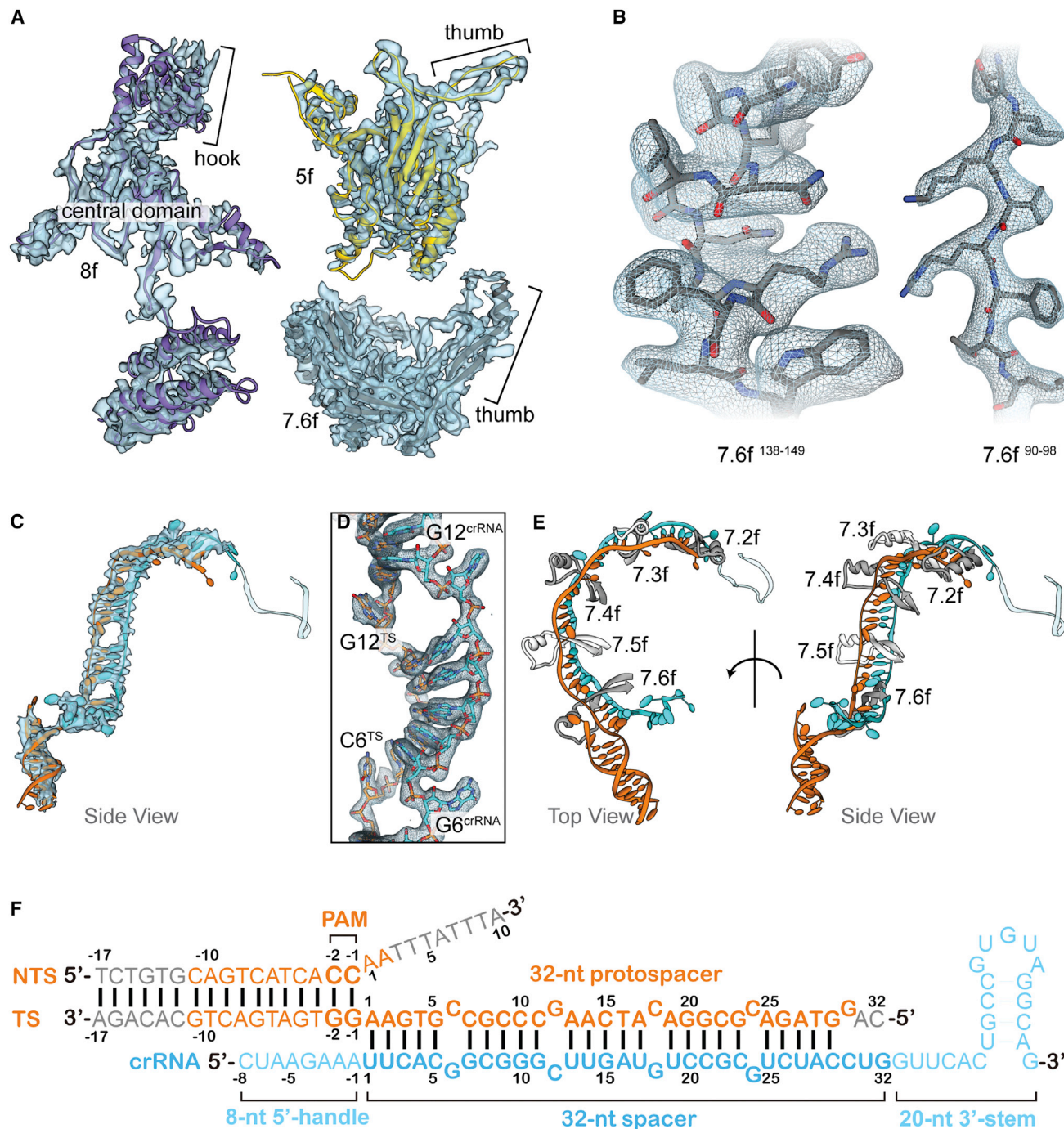


Figure 2. Csy^{crRNA}-dsDNA Components

Visualization of atomic models and corresponding cryo-EM density maps for Csy^{crRNA}-dsDNA components.

(A) Subunits Cas8f, Cas5f, and Cas7.6f, highlighting Cas8f hook (residues 1–166), Cas8f central domain (167–264), Cas5f thumb (48–109), and Cas7f thumb (44–82).

(B and C) Cryo-EM density maps and fits for selected regions of Cas7.6f (B), and the crRNA: dsDNA heteroduplex (C). The 3' stem loop of crRNA is not visualized in our structure and is shown in white to indicate its general location. The RNA:DNA hybrid adopts a pseudo-A conformation resembling a ribbon or ladder with periodic gaps.

(D) A segment of RNA:DNA hybrid (nt 6–12) to show that the heteroduplex follows a “5+1” pattern, in which five base pairs (nt 7–11) are followed by one gap in which the nucleotides are kinked away (nt 6 and nt 12).

(legend continued on next page)

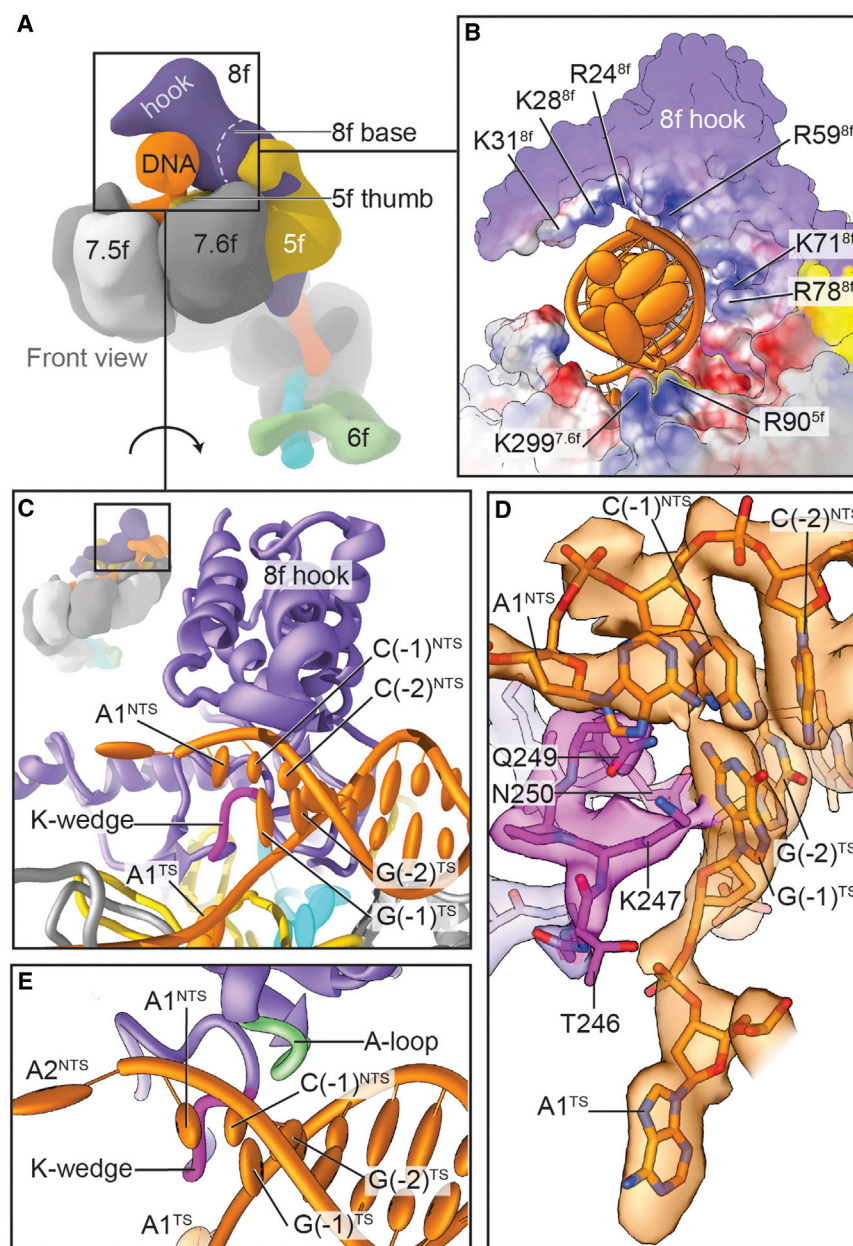


Figure 3. DNA Duplex Binding and PAM Recognition

(A) Surface representation illustrating location of the bound DNA duplex. The DNA duplex binding pocket is located between the Cas8f hook domain, the Cas5f thumb, and Cas7.6f.

(B) Electrostatic surface of the DNA duplex binding pocket, highlighting positively charged residues likely to stabilize duplex-binding via electrostatic interactions with the negatively charged DNA sugar-phosphate backbone.

(C) Detailed view of the region of DNA bifurcation (PAM fork). The Cas8f lysine-containing wedge (residues 246–250, “K-wedge”) is situated at the fork, with two nucleotides (target strand A1^{TS} and non-target strand A1^{NTS}) displaced. Cas8f is shown in purple, K-wedge is shown in dark magenta for emphasis, and DNA is shown in orange.

(D) Cryo-EM density for the K-wedge (magenta) and PAM fork residues (orange). Residue T246 points toward A1^{TS}, while residues Q249 and N250 point toward the PAM site. K247 is situated between A1^{NTS} and G(–1)^{TS}.

(E) Top view of PAM fork, highlighting the K-wedge (magenta) and Ala-rich loop (green) in relation to the dsDNA (orange) and the rest of Cas8f (purple). These two motifs together mediate recognition of both PAM(–1) and PAM(–2) sites.

See also [Figure S3](#) and [Movie S1](#).

Wedge into this fork is the tip of a loop (residues 246–250, sequence TKPQN) emanating from the central domain of Cas8f ([Figure 3C](#)). The placement of this loop, which we refer to as the lysine-containing wedge (or K-wedge), would sterically block a 1^{TS}–1^{NTS} base pair, with K247 stacking with PAM G(–1)^{TS}, thus facilitating the formation of the spacer:protospacer heteroduplex and displacing the non-target DNA strand ([Figure 3D](#)). Accordingly, the first displaced nucleotide of the non-target strand (A1^{NTS}) is splayed away from its theoretical position in an ideal B-form DNA duplex, and the corresponding nucleotide of the target strand (A1^{TS}) is flipped around to base pair with the first nucleotide of the crRNA spacer (U1). Favorable interactions between T246 and the 5' phosphate group of target-strand G(–1)^{TS} probably stabilize the flipping of this nucleotide.

Structural studies of type I-E ([Hayes et al., 2016](#); [Xiao et al., 2017](#)) and other systems ([Anders et al., 2014](#); [Yamano et al.,](#)

negatively charged phosphate groups of the DNA duplex backbone are important for stabilization of the Csy:dsDNA association.

The target and non-target DNA strands separate immediately at the protospacer region next to the PAM site (i.e., G(–1) G(–2)^{TS} and C(–1)C(–2)^{NTS}; see [Figure 2F](#) for nomenclature), with the target strand “flipping” to base pair with crRNA.

(E) crRNA:target DNA heteroduplex and duplex DNA, with Cas7f thumb domains, viewed from the top and side. Cas7f thumbs thread through the RNA:DNA heteroduplex with a β -hairpin, stabilizing the observed 5+1 pattern.

(F) Schematic drawing of the partial R-loop construct with dsDNA and crRNA. Nucleotides are numbered such that the spacer:protospacer heteroduplex are nt 1–32, and the 2-base pair PAM site is at positions (–1) and (–2). Nucleotides with no clear density and no pre-existing model are shown in gray.

See also [Figures S2](#) and [S3](#) and [Movie S1](#).

2016) have shown that PAM recognition, which is crRNA-independent, is mediated by a diversity of peptide motifs. In the structure of Csy^{crRNA}·dsDNA, we observe that the K-wedge also contains two residues, Q249 and N250, that point toward the minor groove of PAM nucleotides G(−1)^{TS} and C(−1)^{NTS} (Figures 3D and S3B). Hydrogen-bond interactions between these polar side chains and the nitrogenous bases are likely. K247 may also contribute to PAM recognition, but the location of its side chain between G(−1)^{TS} and A1^{NTS} makes determination of binding partners difficult. A separate loop, formed by residues N111, A112, A113, and A114 from the “stem” of the Cas8f hook, slots into the PAM minor groove, near G(−2)^{TS} and C(−2)^{NTS} (Figure 3E). Although only main chain density is visible, the positioning of this “Ala-rich loop” (or A-loop) suggests minor groove interaction with the second G(−2)^{TS}·C(−2)^{NTS} PAM pair. Together, the K-wedge and A-loop likely confer the previously observed specificity of the Csy complex for a GG^{TS}·CC^{NTS} PAM sequence (Rollins et al., 2015).

As for the displaced non-target strand, the direction is defined for the first few bases where we observe assignable density, but the rest appears to be disordered. One positively charged residue, K119 from the Cas8f hook, is positioned near the backbone of the first displaced nucleotide (A1), suggesting that interactions with positively charged residues serve to guide the displaced non-target strand, as has been reported for the Cse complex (Hayes et al., 2016). The distribution of positive charges on the surface of Cas8f suggests there may well be multiple plausible paths that provide favorable electrostatic interactions for positioning of the non-target strand (Figure S3F).

Conformational Changes Induced by DNA Binding

To explore the extent and nature of the conformational changes that occur upon dsDNA binding, we determined the structure of the unliganded Csy^{crRNA} complex (Figure 4A). As with the dsDNA complex, all regions except for the PAM-distal head are well-defined in the cryo-EM density map, with a resolution of ~3.5 Å (Figure S4). Comparison of the structures of the Csy^{crRNA} and Csy^{crRNA}·dsDNA complexes reveals several important differences including major changes to the quaternary structure of the complex (Figures 4B–4D). Beginning at the DNA duplex binding site, there is a rearrangement of the Cas8f hook domain, in which the hook “swings” forward upon DNA-binding relative to the proximal Cas7.6f subunit, giving the appearance of clamping onto the bound DNA duplex (Figure 4B). The hook pivots from a stable “base” of the hook (residues 67–107) near the central domain of Cas8f; at the furthest tip of the hook, the movement is as much as 10–16 Å (as measured from Cas8f K31 and D35). Along the backbone, binding of DNA induces a dramatic elongation of the complex, visible in both the overall structure (Figure 4C) and in crRNA (Figure 4D). Compared to Csy^{crRNA}, the dsDNA complex has progressive displacement of each Cas7f subunit along the backbone (Figure S5), for a maximum displacement of ~20 Å when the structures are aligned at one end of the backbone (Figures 4D and S5). This elongation is accomplished by a change in the helical pitch (i.e., rise per turn) of the backbone, increasing from the ~80 Å pitch of Csy^{crRNA} to the ~110 Å pitch of Csy^{crRNA}·dsDNA (Figure 4D). Because the backbone does not comprise a full helical turn,

the net displacement (~20 Å) is less than the change in helical pitch (~30 Å). This change in pitch occurs without a significant change in diameter (compare Figure 4A with Figure 1C). Inspection of individual Cas7f subunits shows remarkably little change before and after DNA binding (Figure S5), with the exception of two domains, the β-hairpin thumb (residues 44–82) and the “extended web” loop (residues 228–253), which also have weaker density in comparison to other regions of Cas7.6f (Figure S2B). Exclusion of these potentially flexible domains results in RMSD of <1.0 Å between Cas7f main chains before and after DNA-binding, suggesting that the quaternary change in helical pitch is effected mainly by a series of rigid body movements.

Inhibition Mechanisms

AcrF10 is a phage-derived polypeptide that inhibits the type I-F CRISPR-Cas system (Pawluk et al., 2016) whose mechanism is unknown. We found that AcrF10 binds stably to Cas5f and Cas8f in solution (Figure S6A) and therefore undertook a structural study of Csy^{crRNA}·AcrF10. Determination of the crystal structure of AcrF10 alone shows that it is composed of a four-stranded β sheet with three α helices positioned along one face of the β sheet (Figure S6B). The 3.6 Å resolution cryo-EM structure of Csy^{crRNA}·AcrF10 (Figures 5A and S4) shows that it occupies a region on the tail that closely overlaps the DNA duplex binding site (see superposition shown in Figure 5B). Based on proximity, AcrF10 likely binds to some of the same sites identified for duplex DNA, including Cas8f K71 and R78, Cas5f R90, and Cas7f K299 (Figure 5C). The relative positioning of AcrF10 and Cas8f at another intermolecular interface, which includes the K-wedge and A-loop, are shown in Figure S7A. Consistent with its role as a DNA mimic, AcrF10 binding causes a swing in the Cas8f hook domain toward Cas7.6f (Figure 5D) that is qualitatively similar to DNA binding (Figure 4C), although the movement is not as large (K31 and D35 at the tip of the hook are displaced 6–9 Å). However, unlike dsDNA binding, AcrF10 binding does not result in helical elongation or other quaternary changes along the backbone. AcrF10 and dsDNA have very different surface charge distribution profiles, and it is interesting that despite this difference, they both occupy closely overlapping positions on Csy.

In recent studies, Chowdhury et al. (2017), as well as Peng et al. (2017), have reported the cryo-EM structures of a Csy complex bound simultaneously to the phage-derived inhibitors AcrF1 and AcrF2, both of which have been shown biochemically to block DNA binding (Bondy-Denomy et al., 2015). The structures show that AcrF1 binds to the Csy backbone, while AcrF2 binds at the junction between Cas7.6f and Cas8f. These results suggest that AcrF1 prevents hybridization of target DNA strand and AcrF2 competes with DNA for a critical binding site (Chowdhury et al., 2017).

To better understand the effects of binding of these inhibitors on Csy conformation and to compare the binding modes of AcrF1, AcrF2, and AcrF10, we determined cryo-EM structures of the complex bound to AcrF1 and AcrF2 individually (Csy^{crRNA}·AcrF1 and Csy^{crRNA}·AcrF2) at nominal resolutions of 3.1 Å and 3.2 Å, respectively (see Figure S4 for FSC plots). Inspection of the structure of Csy^{crRNA}·AcrF2 (Figure 6A) shows that AcrF2 is positioned further away from Cas7.6f and does not occupy the same region as that occupied by either the DNA or

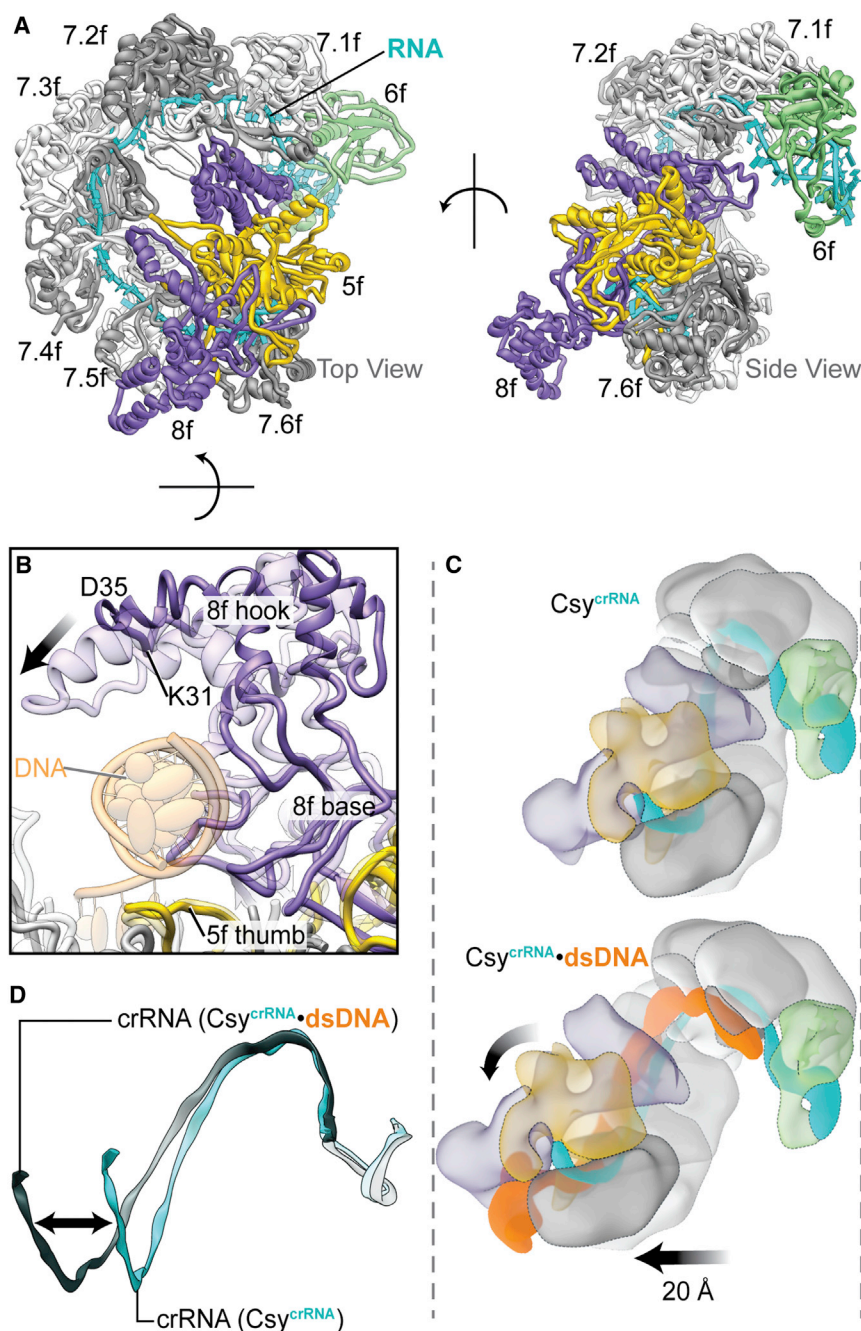


Figure 4. Conformational Changes upon DNA Binding

(A) Top and side views of cryo-EM-derived atomic model for the Csy^{crRNA} complex, shown as a ribbon diagram.

(B) Swing of the Cas8f hook (residues 1–166) upon DNA binding illustrated by superposition of the two structures, with Csy^{crRNA} and Csy^{crRNA}·dsDNA shown in solid and transparent colors, respectively. The hook pivots at a “base” (residues 66–106), with maximum displacement occurring at the tip of the hook, indicated by residues K31 and D35.

(C) Comparison of the structures of Csy^{crRNA} (top) and Csy^{crRNA}·dsDNA (bottom), with the vertical dashed lines provided as a guide to highlight the “unwinding” of the complex. Alignment at the Cas7.1f subunit here results in a net axial displacement of ~20 Å at the distal Cas7.6f subunit as a result of the overall elongation of the complex.

(D) Comparison of the crRNA strands in Csy^{crRNA} (cyan) and Csy^{crRNA}·dsDNA (gray) complexes, illustrating the ~20 Å elongation, which taken in context of the spiral arrangement translates to a change in helical pitch from ~80 Å to ~110 Å with DNA binding.

See also Figures S4 and S5 and Movie S1.

AcrF10 (Figures 6B and 6C). Strikingly, while both DNA binding (Figure 4C) and AcrF10 binding (Figure 5D) result in the hook swinging toward Cas7.6f, AcrF2 has the opposite effect. AcrF2 binding pulls the hook away from Cas7.6f (Figure 6D), with a displacement of ~8 Å at the tip of the hook. This effective difference between AcrF2 and AcrF10 likely results from the differences in their binding sites (Figures S7A and S7B).

Because high-resolution structural information available on AcrF1 at present is from NMR studies (Maxwell et al., 2016), we also determined the crystal structure of AcrF1 to test for

any differences. As shown in Figure S6C, AcrF1 is composed of a three-stranded β sheet with two α helices positioned along one face of the β sheet, and there are minimal differences between the NMR and crystal structures. The cryo-EM structure of Csy^{crRNA}·AcrF1 (Figure 7A) shows that AcrF1 binds with the same stoichiometry and at the same binding location as reported by Chowdhury et al. (2017) for the complex of Csy with both AcrF1 and AcrF2. Densities for tail subunits Cas8f and Cas5f could not be clearly resolved in our cryo-EM structure of Csy with only AcrF1 bound, but the binding sites of AcrF1 to the Cas7f backbone are well-resolved. The relative positioning of AcrF1 and side chains of adjacent Cas7f subunits at the intermolecular interface are shown in Figure S7C. Comparison of Csy^{crRNA}·AcrF1 with Csy^{crRNA}·dsDNA

shows that the bulky AcrF1 inhibitors sit above the nucleic acid trench along the Cas7f subunits and obstruct DNA entrance, with the tip of two AcrF1 loops (residues 8–15 and 33–35) reaching far enough into the trench to sterically clash with the sugar-phosphate backbone of the target DNA (Figures 7B and 7C).

DISCUSSION

While dsDNA target recognition has been extensively analyzed in the Cse (type I-E) system via structural studies (Hayes et al.,

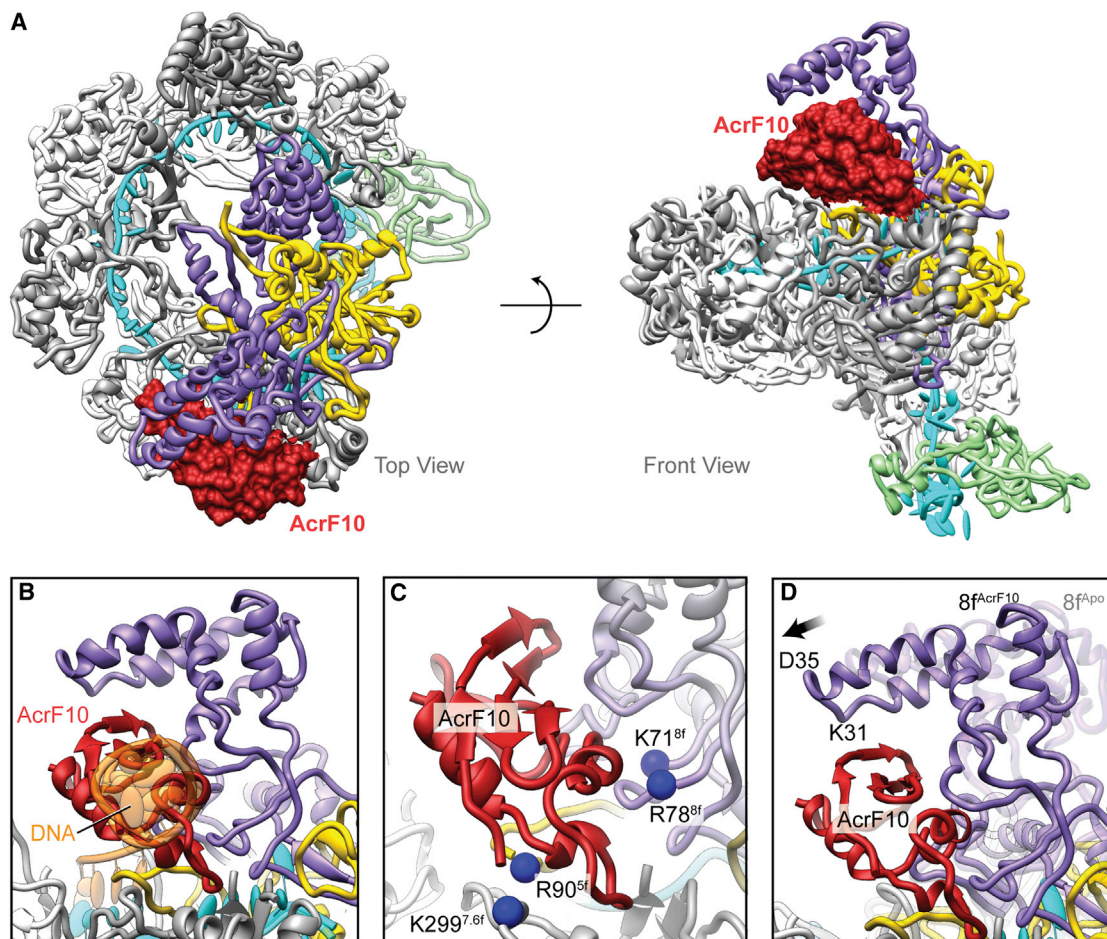


Figure 5. Structure and Mechanism of AcrF10 Binding to Csy^{crRNA}

(A) Top and side views of cryo-EM-derived atomic model for Csy^{crRNA}-AcrF10, shown in ribbon representation. AcrF10 (red) binds between Cas7.6f (gray) and the Cas8f hook (purple).

(B) Comparison of Csy^{crRNA}-AcrF10 (opaque) and Csy^{crRNA}-DNA (transparent) at the DNA duplex binding pocket. AcrF10 (red) and duplex DNA (orange) occupy similar locations as illustrated by the superposition of the respective coordinates.

(C) The close proximity of potential DNA duplex binding sites (with Cα atoms of relevant residues shown as blue spheres) and AcrF10 (red) is consistent with shared binding sites between AcrF10 and duplex DNA.

(D) Comparison of the structures of Csy^{crRNA} (transparent) and Csy^{crRNA}-AcrF10 (opaque) shows that AcrF10 binding results in movement of the Cas8f hook (tip residues K31 and D35) toward the binding pocket in the direction indicated by the arrow.

See also Figures S4, S6, and S7 and Movie S1.

2016; Hochstrasser et al., 2014; Jackson et al., 2014; Mulepati et al., 2014; van Erp et al., 2015; Wiedenheft et al., 2011; Zhao et al., 2014), there is enormous variation among CRISPR systems in both sequence and structure and the degree to which these mechanisms and structural motifs are conserved or plastic are only beginning to be explored (Cass et al., 2015; Jackson and Wiedenheft, 2015; Nishimasu and Nureki, 2017; van der Oost et al., 2014). Here, we report cryo-EM structures for the *P. aeruginosa* Csy complex before and after binding to either dsDNA or the phage-derived inhibitors AcrF1, AcrF2, or AcrF10. Our findings lead to a structural model for target recognition by a type I-F Csy CRISPR system, beginning with the capture of exogenous dsDNA by electrostatic interactions via a vise-like DNA binding pocket and culminating in a large-scale change in pitch of the Cas7f backbone. We propose that

following contact with dsDNA, the K-wedge pries the dsDNA open by steric displacement of two base-paired nucleotides. Presence of the G-G/C-C PAM produces stabilizing interactions between the K-wedge and A-loop with the PAM bases (Figures 3C–3E) that stabilize the DNA-bound structure long enough for the initial nucleation of a spacer:protospacer duplex. The progressive hybridization of this heteroduplex (i.e., directional target verification) is accompanied by a change in helical pitch along the Cas7f backbone (Figures 4C and 4D) that may, in addition, be essential in the pathway for Cas3 nuclease recruitment to degrade the bound DNA.

DNA Duplex Binding at the Flexible Hook Domain

An interesting feature present in the Csy complex is the flexible Cas8f hook domain, under which the DNA duplex is nestled.

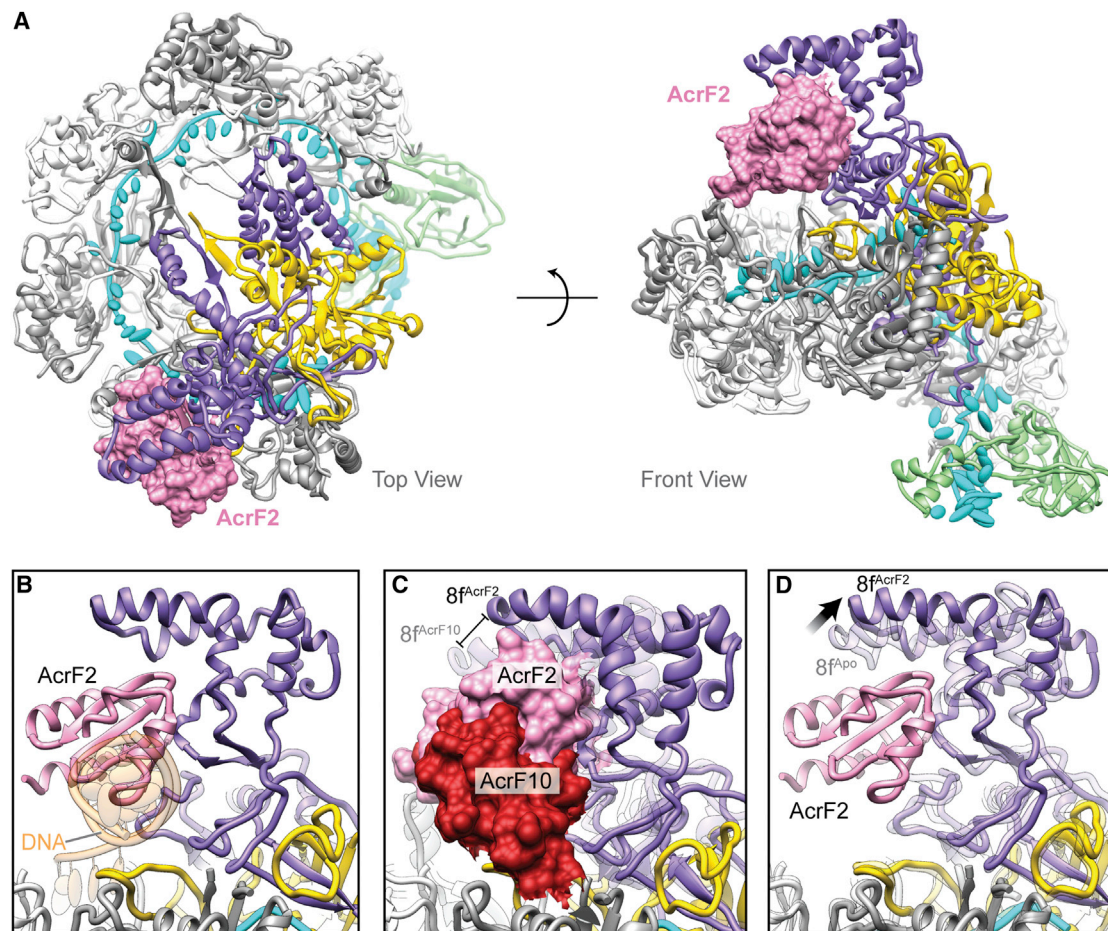


Figure 6. Structure and Mechanism of AcrF2 Binding to Csy^{crRNA}

(A) Top and side views of cryo-EM-derived atomic model of Csy^{crRNA}-AcrF2 shown in ribbon representation. AcrF2 (pink) binds to the Cas8f hook domain (purple). (B) Comparison of the structures of Csy^{crRNA}-AcrF2 (opaque) and Csy^{crRNA}-DNA (transparent) at the DNA duplex binding pocket shows that the site of bound AcrF2 (pink) only partially overlaps with the site of bound DNA duplex (orange). (C) Comparison of the structures of Csy^{crRNA}-AcrF2 and Csy^{crRNA}-AcrF10 illustrates the non-overlapping binding sites, which result in different conformations of the Cas8f hook, shown in transparent color for the AcrF10 bound complex and in solid color for the AcrF2 bound complex. (D) Comparison of the structures of Csy^{crRNA} (transparent) and Csy^{crRNA}-AcrF2 (opaque) show that in contrast to the binding of AcrF10, the binding of AcrF2 results in movement of the Cas8f hook away from the DNA duplex binding pocket. See also [Figures S4 and S7](#) and [Movie S1](#).

Our studies reveal a spectrum of distinct conformations in this hook domain. Csy^{crRNA} initially adopts an “open” hook state, and binding of dsDNA induces transition to a “closed” state ([Figure 4B](#)). Inhibitors binding at this crucial region can also trigger conformational changes in the hook domain. AcrF10 binding induces a partially closed hook state ([Figure 5D](#)), while AcrF2 binding wrenches the hook outward away from the closed state ([Figure 6D](#)), demonstrating the large conformational range of this flexible domain.

We cannot conclude yet whether the closed state is triggered by the mere presence of duplex DNA at its binding site, or if successful PAM recognition is also required. The latter scenario is more appealing, since it provides a plausible mechanism by which the Csy complex temporarily stabilizes association with dsDNA and primes it for strand invasion by crRNA. Notably, although AcrF10 occupies a similar binding location to that of

the DNA duplex ([Figure 5B](#)), and may share common interaction partners ([Figures 5C and S7A](#)), its binding only induces a partial inward movement of the hook domain ([Figure 5D](#)), suggesting that the full closure of the hook domain requires additional specific interactions.

PAM Recognition by the K-Wedge and A-Loop

The Csy PAM site is shorter and less permissive than that of type I-E Cse. While the Cse PAM site is three nucleotides long and recognizes at least four different PAM sequences for interference ([Westra et al., 2013](#)), Csy only recognizes the G-G/C-C PAM ([Mojica et al., 2009](#)). Despite the relative simplicity of the Csy PAM, the higher-level logic of PAM recognition appears to be well conserved between type I-E Cse and type I-F Csy, with differences in execution. As described earlier, Csy employs two structural motifs, the K-wedge and the A-loop for PAM

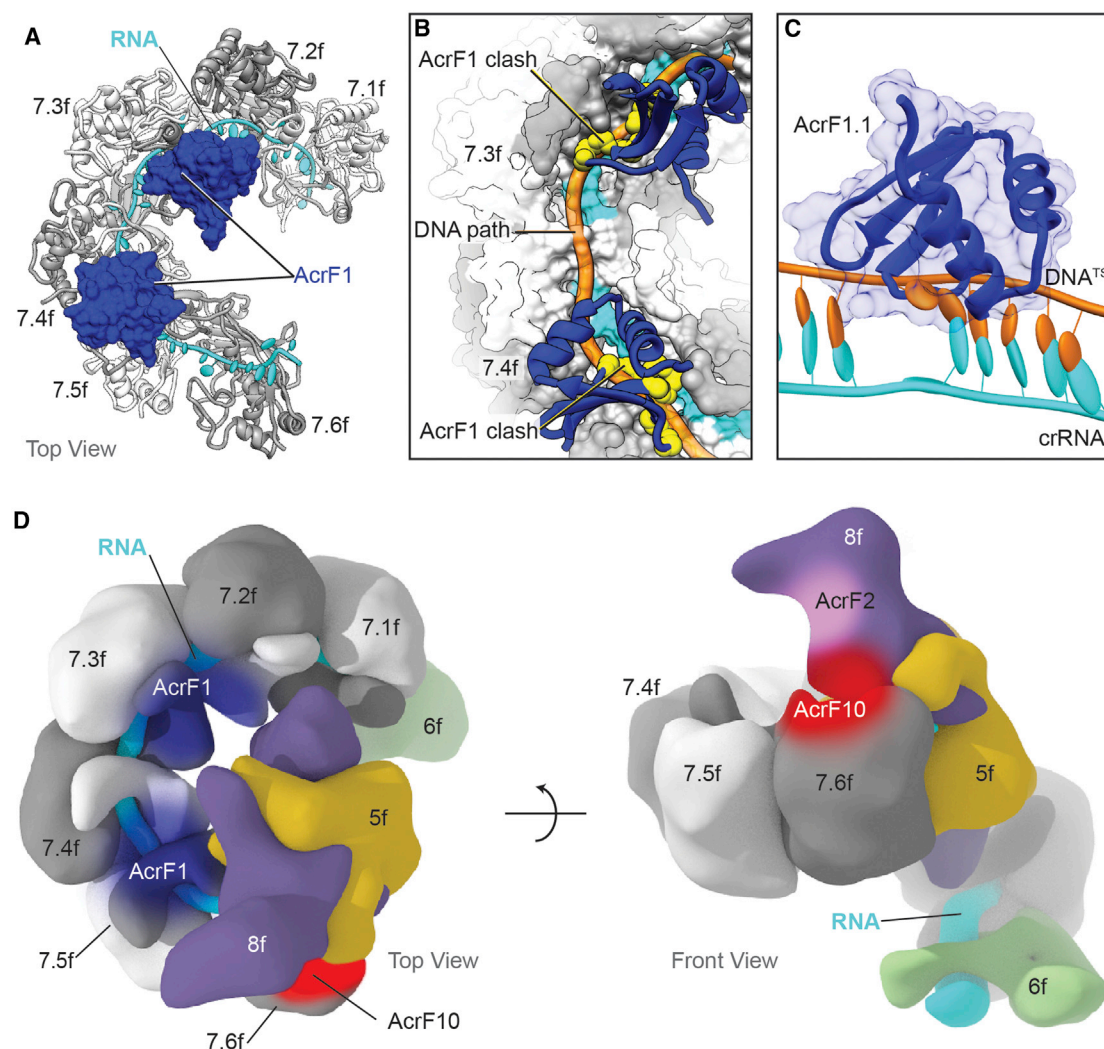


Figure 7. Structure and Mechanism of AcrF1 Binding to Csy^{crRNA}

(A) Top view of cryo-EM-derived atomic model of Csy^{crRNA}-AcrF1 shown in ribbon representation. Two copies of AcrF1 (blue) bind on top of crRNA (cyan) along the backbone Cas7f subunits (gray). The tail and head regions are not shown both for the sake of clarity and due to poor resolution in the density map.

(B and C) Comparison of the structures of Csy^{crRNA}-AcrF1 and Csy^{crRNA}-dsDNA from top (B) and side (C) views. Superposition of the bound DNA and AcrF1 illustrates how steric clashes between AcrF1 and protospacer DNA would prevent target hybridization. AcrF1 residues 8–15 and 33–35, which directly clash with the path of DNA, are highlighted with their C α s as yellow spheres in (B).

(D) A composite surface representation of the Csy complex depicting the binding footprints of AcrF1, AcrF2, and AcrF10.

See also [Figures S4, S6, and S7](#) and [Movie S1](#).

recognition ([Figure S3C](#)). The type I-E Cse complex in *Escherichia coli* (EcoCse) has a Q-wedge, G-loop, and K-finger ([Figure S3D](#)) ([Hayes et al., 2016](#)), while in *Thermobifida fusca*, TfuCse displays a Q-wedge, G-loop, R-finger, and an L1-loop ([Figure S3E](#)) ([Xiao et al., 2017](#)). Despite major differences in the sequence of Cas8, the positions of K- and Q-wedges, as well as the A- and G-loops, are well conserved structurally in these three complexes.

In all three surveillance complexes, the wedge motif is responsible for strand separation. However, the differences in PAM sequences and in promiscuity between CRISPR-Cas systems result in different motifs for sequence discrimination. EcoCse uses four residues across three different structural motifs

(a Q-wedge alanine, two G-loop glycines, and a K-finger lysine) for sequence discrimination ([Hayes et al., 2016](#)); meanwhile, TfuCse relies only on an SGM motif in its G-loop and perhaps the R-finger arginine ([Xiao et al., 2017](#)). As expected, the sequence discrimination motifs in type I-F Csy are different from both type I-E systems. Unlike the previously described Cse Q-wedge, the Csy K-wedge plays a much larger role in sequence discrimination, with Q249, N250, and possibly K247, specifying a G(−1)^{TS}-C(−1)^{NTS} pair ([Figures 3C and 3D](#)). Recognition of the next PAM base pair is independently performed by an Ala-rich loop (A-loop, [Figure 3E](#)), which bears resemblance to the Cse Gly-rich loop (G-loop). The A-loop may recognize the PAM(−2) base pair through main-chain interactions with the

Ala residues (A112, A113, A114), similar to the use of Ala and Gly residues in PAM recognition *Eco*Cse (Hayes et al., 2016) and/or by hydrogen bond interactions with N111. In all three cases (*Eco*Cse, *Tfu*Cse, and Csy), sequence discrimination occurs from the minor groove (Hayes et al., 2016; Xiao et al., 2017).

Elongated Helical Pitch in DNA-Bound Csy

The large-scale change in pitch of the complex that we observe between Csy^{crRNA} and Csy^{crRNA}.dsDNA is achieved by each Cas7f subunit rotating as a near-rigid body relative to its predecessor (Figure S5). One explanation for this change is that while single-stranded RNA can adopt this unusual periodically kinked pseudo-A conformation at a tight helical pitch of 80 Å, the RNA:DNA heteroduplex has more stringent torsional constraints and requires a greater pitch. As the target protospacer directionally hybridizes to crRNA, the Cas7f subunits are likely to rotate to accommodate heteroduplex formation. Although this type of large-scale conformational change has not been observed for other CRISPR complexes, it is interesting to note that the Class I CRISPR systems are known to vary in helical pitch based on subtype. Of the complexes with known structures, Cse (type I-E) at ~130 Å and Cmr (type III-B) at ~160 Å (Hayes et al., 2016; Osawa et al., 2015) have a pitch closer to that of dsDNA-bound Csy (type I-F, at ~110 Å). Even across subtypes, the helical radius remains the same and the 5+1 repeating RNA:DNA segments overlap well with each other (Chowdhury et al., 2017).

Biophysical experiments with the type I-E surveillance complex show that the only requirement for the recruitment and activity of the Cas3 nuclease/helicase is the completion of target verification across the entire length of crRNA (and stabilization of the R-loop) (Rutkauskas et al., 2015). If the same is true for Csy, then the structural changes observed upon DNA binding could be necessary and on the pathway for the recruitment and activity of the 121 kDa Cas3/2 nuclease fusion protein (Rollins et al., 2015). Extensive studies of type I-E surveillance complexes provide an interesting example of how Cas3 recruitment to the Cas8 subunit at the tail can be dependent on allosteric target verification along the Cas7 backbone. As with Csy, the incoming DNA duplex in the Cse complex is primed at Cas8e with only a transiently stable conformation (Rutkauskas et al., 2015). Only upon successful verification of the entire target (i.e., template matching with crRNA) do the two “belly” subunits (Cse2.1 and Cse2.2) slide toward Cas8e, triggering local rearrangements in Cas8e that “flag” for Cas3 recruitment in the Cse system (Hayes et al., 2016; van Erp et al., 2015; Wiedenheft et al., 2011; Xiao et al., 2017). The unanticipated and large helical elongation observed upon DNA-binding in Csy may thus serve as a recruitment beacon for Cas3/2, perhaps by opening up the complex to reveal new binding sites.

Structural Comparisons between Type I-E Cse and Type I-F Csy Surveillance Complexes

Comparisons of Csy and Cse show that the structural plasticity of type I CRISPR systems can result in an incredible diversity of structural and mechanistic means to common functional ends.

The structural motifs for PAM recognition in type I-F as we have described resemble type I-E more closely than they do type I-Fv, a minimal variant of type I-F that recognizes the same G-G/C-C PAM

(Pausch et al., 2017). Relative to the DNA duplex, the K-wedge and Q-wedge of Csy and Cse are in analogous positions, as are the A-loop and G-loop, allowing both complexes to read the PAM from the minor groove (Figures S3C–S3E) (Hayes et al., 2016; Xiao et al., 2017). Thus, the PAM recognition principles in Cse and Csy can be described as comprising “variations on a theme” despite the fact they utilize different targeting sequences. However, the Cse and Csy complexes employ very different mechanisms to engage target DNA. DNA interaction with Csy involves the Cas8f “hook” domain, which flexibly clamps onto duplex DNA and holds it in a secure vise. No corresponding domain exists in Cse; instead, duplex DNA association appears to be mediated by Cas7e lysine-rich α helices that “squeeze” the duplex from opposite sides (van Erp et al., 2015). Further, in the Cse complex, two copies of type-specific Cse2 “belly” subunit stabilize the target DNA strand on one side (van Erp et al., 2015; Hayes et al., 2016; Xiao et al., 2017) and lock the displaced non-target strand on the “backside” (Hayes et al., 2016; Xiao et al., 2017); these proteins are not present in Csy. The closest analogy of this type of interaction in the Csy complex is the 25 aa extension of the web domains in each of the six Cas7f subunits: these “extended webs” form wide loops that occlude the target strand bases from solvent and may compensate for Cse2. Perhaps the most remarkable mechanistic difference between Cse and Csy complexes is in the quaternary structural changes that occur in response to DNA binding. In Cse, complete target verification is followed by displacement of the two Cse2 subunits and rotation of Cas8e, with little change in the overall shape of the complex (van Erp et al., 2015; Hayes et al., 2016). In Csy, however, complete target verification results in displacement of the Cas7 backbone subunits and a dramatic elongation of the entire helical complex, as we have demonstrated here. Finally, in all three cryo-EM studies of Csy (this paper; Chowdhury et al., 2017; Peng et al., 2017), the Cas6f “head” is poorly resolved, indicating inherent flexibility of the region; such flexibility has not been reported for the analogous Cas6e in Cse (Xiao et al., 2017). Future studies investigating the assembly of the Csy complex (that is initiated by Cas6f) or examining the PAM-distal DNA duplex may shed light on the functional significance of this flexibility.

Anti-CRISPR Inhibitors Sterically Block DNA Recognition

Anti-CRISPR inhibitors can block different stages of DNA recognition and exploit various parts of the DNA recognition machinery. Phage-derived anti-CRISPRs may delay escape mutations by targeting functionally important residues of the surveillance complex (Chowdhury et al., 2017). In our present work, we have analyzed the binding of AcrF1, AcrF2, and AcrF10, each of which have distinct binding sites on the Csy complex (Figure 7D). AcrF10 is a DNA mimic that triggers a smaller conformational change in the Cas8f hook along the same trajectory that is observed with DNA binding (Figure 5). AcrF2 was previously believed to be a DNA mimic; we show here that its binding site only partially overlaps with that of the DNA duplex, but does cause an alternative conformational change in the hook in a direction opposite to that seen with DNA binding (Figure 6). Finally, we show that the site of AcrF1 binding directly clashes with the path of the target protospacer (Figure 7).

Taken together, our findings provide new and unexpected insights to the structure and dynamics associated with DNA targeting and binding inhibition in the Csy surveillance complex. The large and striking conformational change that is observed with DNA binding has not yet been observed in other CRISPR-Cas systems, and it will be interesting to see if this mechanism is more general. All three anti-CRISPR inhibitors we have analyzed function by steric occlusion, precluding DNA from binding, but it is plausible that other inhibitors that act allosterically without directly competing with the DNA binding site await discovery. Given that the various conformational states of CRISPR-Cas complexes such as Csy will likely be flexible and hence challenging for analysis by X-ray crystallography, the use of cryo-EM methods will undoubtedly represent an increasingly practical approach to deciphering the structural and mechanistic basis of their function.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- **KEY RESOURCES TABLE**
- **CONTACT FOR REAGENT AND RESOURCE SHARING**
- **METHOD DETAILS**
 - Protein Expression and Purification
 - Crystallization, Data Collection, and Structure Determination
 - Electron Microscopy: grid preparation and data acquisition
 - Image processing
 - Atomic model building and refinement
 - Figure generation
- **DATA AND SOFTWARE AVAILABILITY**
 - Data Resources

SUPPLEMENTAL INFORMATION

Supplemental Information includes seven figures, two tables, and one movie and can be found with this article online at <http://dx.doi.org/10.1016/j.cell.2017.09.006>.

AUTHOR CONTRIBUTIONS

Conceptualization, T.W.G., A.B., H.Y., D.J.P., and S.S.; Methodology, T.W.G., A.B., H.Y., P.R., A.M., and S.S.; Software, T.W.G., A.B., H.Y., and E.T.E.; Data Acquisition, P.R., A.M., E.T.E., A.M.R., and T.F.; Writing, T.G., A.B., H.Y., D.J.P., and S.S.; Editing, T.G., A.B., H.Y., L.A.E., D.J.P., and S.S.; Visualization, V.F.; Supervision, D.J.P. and S.S.; Project Administration, D.J.P. and S.S.; Funding Acquisition, D.J.P. and S.S.

ACKNOWLEDGMENTS

We thank K. Podolsky for initial evaluation of Csy^{crRNA} by negative stain electron microscopy. We also thank Drs. Bridget Carragher and Clint Potter for access to the facilities at the New York Structural Biology Center for high-throughput electron microscopy data collection. The structural work with the Csy^{crRNA}-AcrF2 was performed at the Simons Electron Microscopy Center and National Resource for Automated Molecular Microscopy located at the New York Structural Biology Center, supported by grants from the Simons Foundation (349247), NYSTAR, and the NIH National Institute of General Med-

ical Sciences (GM103310) with additional support from the Agouron Institute (F00316) and NIH S10 Od019994-01. This research was supported by funds from the Center for Cancer Research, National Cancer Institute, NIH, Bethesda, MD (to S.S.), by NIH grant GM104962 and a Geoffrey Beene Cancer Research Center grant to D.J.P., by Memorial Sloan Kettering Cancer Center Core Grant (P30CA008748), and by federal funds from the Frederick National Laboratory for Cancer Research, National Institutes of Health, under contract HHSN261200800001E. The computational analyses utilized the high-performance computational capabilities of the Biowulf Linux cluster at the NIH, Bethesda, MD (<https://hpc.nih.gov/>).

Received: July 17, 2017

Revised: August 15, 2017

Accepted: September 6, 2017

Published: October 5, 2017

REFERENCES

- Adams, P.D., Afonine, P.V., Bunkóczi, G., Chen, V.B., Davis, I.W., Echols, N., Headd, J.J., Hung, L.W., Kapral, G.J., Grosse-Kunstleve, R.W., et al. (2010). PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D Biol. Crystallogr.* 66, 213–221.
- Anders, C., Niewoehner, O., Duerst, A., and Jinek, M. (2014). Structural basis of PAM-dependent target DNA recognition by the Cas9 endonuclease. *Nature* 513, 569–573.
- Bartasaghi, A., Matthies, D., Banerjee, S., Merk, A., and Subramaniam, S. (2014). Structure of β -galactosidase at 3.2-Å resolution obtained by cryo-electron microscopy. *Proc. Natl. Acad. Sci. USA* 111, 11709–11714.
- Bondy-Denomy, J., Pawluk, A., Maxwell, K.L., and Davidson, A.R. (2013). Bacteriophage genes that inactivate the CRISPR/Cas bacterial immune system. *Nature* 493, 429–432.
- Bondy-Denomy, J., Garcia, B., Strum, S., Du, M., Rollins, M.F., Hidalgo-Reyes, Y., Wiedenheft, B., Maxwell, K.L., and Davidson, A.R. (2015). Multiple mechanisms for CRISPR-Cas inhibition by anti-CRISPR proteins. *Nature* 526, 136–139.
- Cardone, G., Heymann, J.B., and Steven, A.C. (2013). One number does not fit all: mapping local variations in resolution in cryo-EM reconstructions. *J. Struct. Biol.* 184, 226–236.
- Case, D.A., Cerutti, D.S., Cheatham, T.E., III, Darden, T.A., Duke, R.E., Giese, T.J., Gohlke, H., Goetz, A.W., Greene, D., Homeyer, N., et al. (2017). AMBER 2017 (University of California, San Francisco).
- Cass, S.D., Haas, K.A., Stoll, B., Alkhnbashi, O.S., Sharma, K., Urlaub, H., Backofen, R., Marchfelder, A., and Bolt, E.L. (2015). The role of Cas8 in type I CRISPR interference. *Biosci. Rep.* 35, e00197.
- Chowdhury, S., Carter, J., Rollins, M.F., Golden, S.M., Jackson, R.N., Hoffmann, C., Nosaka, L., Bondy-Denomy, J., Maxwell, K.L., Davidson, A.R., et al. (2017). Structure reveals mechanisms of viral suppressors that intercept a CRISPR RNA-guided surveillance complex. *Cell* 169, 47–57.e11.
- Emsley, P., Lohkamp, B., Scott, W.G., and Cowtan, K. (2010). Features and development of Coot. *Acta Crystallogr. D Biol. Crystallogr.* 66, 486–501.
- Goddard, T.D., Huang, C.C., and Ferrin, T.E. (2007). Visualizing density maps with UCSF Chimera. *J. Struct. Biol.* 157, 281–287.
- Grigorieff, N. (2016). FREALIGN: an exploratory tool for single-particle cryo-EM. *Methods Enzymol.* 579, 191–226.
- Haurwitz, R.E., Sternberg, S.H., and Doudna, J.A. (2012). Csy4 relies on an unusual catalytic dyad to position and cleave CRISPR RNA. *EMBO J.* 31, 2824–2832.
- Hayes, R.P., Xiao, Y., Ding, F., van Erp, P.B., 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. *Proc. Natl. Acad. Sci. USA* 111, 6618–6623.

- Hochstrasser, M.L., Taylor, D.W., Kornfeld, J.E., Nogales, E., and Doudna, J.A. (2016). DNA targeting by a minimal CRISPR RNA-guided cascade. *Mol. Cell* 63, 840–851.
- Jackson, R.N., and Wiedenheft, B. (2015). A conserved structural chassis for mounting versatile CRISPR RNA-guided immune responses. *Mol. Cell* 58, 722–728.
- Jackson, R.N., Golden, S.M., van Erp, P.B., Carter, J., Westra, E.R., Brouns, S.J., van der Oost, J., Terwilliger, T.C., Read, R.J., and Wiedenheft, B. (2014). Structural biology. Crystal structure of the CRISPR RNA-guided surveillance complex from *Escherichia coli*. *Science* 345, 1473–1479.
- Johnson, G.T., Autin, L., Goodsell, D.S., Sanner, M.F., and Olson, A.J. (2011). ePMV embeds molecular modeling into professional animation software environments. *Structure* 19, 293–303.
- Li, X., Mooney, P., Zheng, S., Booth, C.R., Braunkopf, M.B., Gubbens, S., Agard, D.A., and Cheng, Y. (2013). Electron counting and beam-induced motion correction enable near-atomic-resolution single-particle cryo-EM. *Nat. Methods* 10, 584–590.
- Ludtke, S.J., Baldwin, P.R., and Chiu, W. (1999). EMAN: semiautomated software for high-resolution single-particle reconstructions. *J. Struct. Biol.* 128, 82–97.
- Makarova, K.S., Wolf, Y.I., and Koonin, E.V. (2013). Comparative genomics of defense systems in archaea and bacteria. *Nucleic Acids Res.* 41, 4360–4377.
- Makarova, K.S., Wolf, Y.I., Alkhnbashi, O.S., Costa, F., Shah, S.A., Saunders, S.J., Barrangou, R., Brouns, S.J., Charpentier, E., Haft, D.H., et al. (2015). An updated evolutionary classification of CRISPR-Cas systems. *Nat. Rev. Microbiol.* 13, 722–736.
- Maraffini, L.A. (2015). CRISPR-Cas immunity in prokaryotes. *Nature* 526, 55–61.
- Maxwell, K.L., Garcia, B., Bondy-Denomy, J., Bona, D., Hidalgo-Reyes, Y., and Davidson, A.R. (2016). The solution structure of an anti-CRISPR protein. *Nat. Commun.* 7, 13134.
- Mohanraju, P., Makarova, K.S., Zetsche, B., Zhang, F., Koonin, E.V., and van der Oost, J. (2016). Diverse evolutionary roots and mechanistic variations of the CRISPR-Cas systems. *Science* 353, aad5147.
- Mojica, F.J., Díez-Villaseñor, C., García-Martínez, J., and Almendros, C. (2009). Short motif sequences determine the targets of the prokaryotic CRISPR defence system. *Microbiology* 155, 733–740.
- Mulepati, S., Héroux, A., and Bailey, S. (2014). Structural biology. Crystal structure of a CRISPR RNA-guided surveillance complex bound to a ssDNA target. *Science* 345, 1479–1484.
- Nishimasu, H., and Nureki, O. (2017). Structures and mechanisms of CRISPR RNA-guided effector nucleases. *Curr. Opin. Struct. Biol.* 43, 68–78.
- Osawa, T., Inanaga, H., Sato, C., and Numata, T. (2015). Crystal structure of the CRISPR-Cas RNA silencing Cmr complex bound to a target analog. *Mol. Cell* 58, 418–430.
- Otwinowski, Z., and Minor, W. (1997). Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.* 276, 307–326.
- Pausch, P., Müller-Esparza, H., Gleditsch, D., Altegoer, F., Randau, L., and Bange, G. (2017). Structural variation of type I-F CRISPR RNA guided DNA surveillance. *Mol. Cell* 67, 622–632.e4.
- Pawluk, A., Staals, R.H., Taylor, C., Watson, B.N., Saha, S., Fineran, P.C., Maxwell, K.L., and Davidson, A.R. (2016). Inactivation of CRISPR-Cas systems by anti-CRISPR proteins in diverse bacterial species. *Nat. Microbiol.* 1, 16085.
- Peng, R., Xu, Y., Zhu, T., Li, N., Qi, J., Chai, Y., Wu, M., Zhang, X., Shi, Y., Wang, P., et al. (2017). Alternate binding modes of anti-CRISPR viral suppressors AcrF1/2 to Csy surveillance complex revealed by cryo-EM structures. *Cell Res.* 27, 853–864.
- Rohou, A., and Grigorieff, N. (2015). CTFFIND4: fast and accurate defocus estimation from electron micrographs. *J. Struct. Biol.* 192, 216–221.
- Rollins, M.F., Schuman, J.T., Paulus, K., Bukhari, H.S., and Wiedenheft, B. (2015). Mechanism of foreign DNA recognition by a CRISPR RNA-guided surveillance complex from *Pseudomonas aeruginosa*. *Nucleic Acids Res.* 43, 2216–2222.
- Rollins, M.F., Chowdhury, S., Carter, J., Golden, S.M., Wilkinson, R.A., Bondy-Denomy, J., Lander, G.C., and Wiedenheft, B. (2017). Cas1 and the Csy complex are opposing regulators of Cas2/3 nuclease activity. *Proc. Natl. Acad. Sci. USA* 114, E5113–E5121.
- Rutkauskas, M., Sinkunas, T., Songailiene, I., Tikhomirova, M.S., Siksnys, V., and Seidel, R. (2015). Directional R-loop formation by the CRISPR-Cas surveillance complex cascade provides efficient off-target site rejection. *Cell Rep.* 10, 1534–1543.
- Scheres, S.H. (2012). RELION: implementation of a Bayesian approach to cryo-EM structure determination. *J. Struct. Biol.* 180, 519–530.
- Schrodinger. (2015). The PyMOL Molecular Graphics System, Version 1.8. (New York, NY: Schrödinger, LLC).
- Shah, S.A., Erdmann, S., Mojica, F.J., and Garrett, R.A. (2013). Protospacer recognition motifs: mixed identities and functional diversity. *RNA Biol.* 10, 891–899.
- 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. *J. Struct. Biol.* 151, 41–60.
- Szczelkun, M.D., Tikhomirova, M.S., Sinkunas, T., Gasiunas, G., Karvelis, T., Pschera, P., Siksnys, V., and Seidel, R. (2014). Direct observation of R-loop formation by single RNA-guided Cas9 and Cascade effector complexes. *Proc. Natl. Acad. Sci. USA* 111, 9798–9803.
- Tang, G., Peng, L., Baldwin, P.R., Mann, D.S., Jiang, W., Rees, I., and Ludtke, S.J. (2007). EMAN2: an extensible image processing suite for electron microscopy. *J. Struct. Biol.* 157, 38–46.
- van der Oost, J., Westra, E.R., Jackson, R.N., and Wiedenheft, B. (2014). Unravelling the structural and mechanistic basis of CRISPR-Cas systems. *Nat. Rev. Microbiol.* 12, 479–492.
- van Duijn, E., Barbu, I.M., Barendregt, A., Jore, M.M., Wiedenheft, B., Lundgren, M., Westra, E.R., Brouns, S.J., Doudna, J.A., van der Oost, J., and Heck, A.J. (2012). Native tandem and ion mobility mass spectrometry highlight structural and modular similarities in clustered-regularly-interspaced short-palindromic-repeats (CRISPR)-associated protein complexes from *Escherichia coli* and *Pseudomonas aeruginosa*. *Mol. Cell. Proteomics* 11, 1430–1441.
- van Erp, P.B., Jackson, R.N., Carter, J., Golden, S.M., Bailey, S., and Wiedenheft, B. (2015). Mechanism of CRISPR-RNA guided recognition of DNA targets in *Escherichia coli*. *Nucleic Acids Res.* 43, 8381–8391.
- Voss, N.R., Yoshioka, C.K., Radermacher, M., Potter, C.S., and Carragher, B. (2009). DoG Picker and TiltPicker: software tools to facilitate particle selection in single particle electron microscopy. *J. Struct. Biol.* 166, 205–213.
- Westra, E.R., Semenova, E., Datsenko, K.A., Jackson, R.N., Wiedenheft, B., Severinov, K., and Brouns, S.J.J. (2013). Type I-E CRISPR-cas systems discriminate target from non-target DNA through base pairing-independent PAM recognition. *PLoS Genet.* 9, e1003742.
- Wiedenheft, B., Lander, G.C., Zhou, K., Jore, M.M., Brouns, S.J.J., van der Oost, J., Doudna, J.A., and Nogales, E. (2011). Structures of the RNA-guided surveillance complex from a bacterial immune system. *Nature* 477, 486–489.
- Wright, A.V., Nuñez, J.K., and Doudna, J.A. (2016). Biology and Applications of CRISPR Systems: Harnessing Nature's Toolbox for Genome Engineering. *Cell* 164, 29–44.
- Xiao, Y., Luo, M., Hayes, R.P., Kim, J., Ng, S., Ding, F., Liao, M., and Ke, A. (2017). Structure basis for directional R-loop formation and substrate hand-over mechanisms in type I CRISPR-Cas system. *Cell* 170, 48–60.e11.
- Yamano, T., Nishimasu, H., Zetsche, B., Hirano, H., Slaymaker, I.M., Li, Y., Fedorova, I., Nakane, T., Makarova, K.S., Koonin, E.V., et al. (2016). Crystal structure of Cpf1 in complex with guide RNA and target DNA. *Cell* 165, 949–962.
- Zhao, H., Sheng, G., Wang, J., Wang, M., Bunkoczi, G., Gong, W., Wei, Z., and Wang, Y. (2014). Crystal structure of the RNA-guided immune surveillance Cascade complex in *Escherichia coli*. *Nature* 515, 147–150.
- Zheng, S.Q., Palovcak, E., Armache, J.P., Verba, K.A., Cheng, Y., and Agard, D.A. (2017). MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy. *Nat. Methods* 14, 331–332.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Chemicals, Peptides, and Recombinant Proteins		
ULP1	Home-made	N/A
β -mercaptoethanol	Sigma-Aldrich	CAS: 60-24-2 Cat#M6250-100ML
1 M MES, pH 6.2	Hampton	Cat#HR2-243
1 M HEPES sodium salt, pH 7.5	QIAGEN	Cat#133050
1 M tri-sodium citrate	QIAGEN	Cat#132629
50% w/v PEG 3000	QIAGEN	Cat#133082
100% v/v PEG 400	QIAGEN	Cat#133086
100% v/v glycerol anhydrous	QIAGEN	Cat#133047
M9 Broth	Sigma-Aldrich	Cat#63011-500G-F
Critical Commercial Assays		
NeXtal Tubes Classics Suite	QIAGEN	Cat#130701
NeXtal Tubes JCSG Core Suite III	QIAGEN	Cat#130726
Deposited Data		
AcrF1-MBP	This paper	PDB: 6ANV
AcrF10	This paper	PDB: 6ANW
Model of Csy complex bound to AcrF1 and AcrF2	Chowdhury et al., 2017	PDB: 5UZ9
Model of Type I-E Cse complex bound to dsDNA	Hayes et al., 2016	PDB: 5H9F
Map of Csy complex bound to dsDNA	This paper	EMDB: 7048
Map of Csy complex	This paper	EMDB: 7049
Map of Csy complex bound to AcrF1	This paper	EMDB: 7050
Map of Csy complex bound to AcrF2	This paper	EMDB: 7051
Map of Csy complex bound to AcrF10	This paper	EMDB: 7052
Model of Csy complex bound to dsDNA	This paper	PDB: 6B44
Model of Csy complex	This paper	PDB: 6B45
Model of Csy complex bound to AcrF1	This paper	PDB: 6B46
Model of Csy complex bound to AcrF2	This paper	PDB: 6B47
Model of Csy complex bound to AcrF10	This paper	PDB: 6B48
Experimental Models: Organisms/Strains		
E.coli: BL21(DE3) Competent cells	Novagen	Cat#69450-3
Recombinant DNA		
pRSFDuet-1-His ₆ -SUMO-Cas8f-Cas5f	This paper	N/A
pCDFDuet-1-Cas7f	This paper	N/A
pACYDuet-1-Cas6f-CRISPR	This paper	N/A
pRSFDuet-1-His ₆ -SUMO-AcrF1	This paper	N/A
pRSFDuet-1-His ₆ -SUMO-AcrF2	This paper	N/A
pRSFDuet-1-His ₆ -SUMO-AcrF10	This paper	N/A
pRSFDuet-1-His ₆ -SUMO-AcrF1-MBP	This paper	N/A
Sequence-Based Reagents		
49-nt target DNA strand CAGGTAGACGCGGACATCAAGCCCGCC GTGAAGGTGATGACTGCACAGA	IDT	N/A
27-nt non-target DNA strand TCTGTGCAGTCATACCAATTTATTTA	IDT	N/A
Software and Algorithms		
COOT	Emsley et al., 2010	http://www2.mrc-lmb.cam.ac.uk/personal/pemsley/coot

(Continued on next page)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
PHENIX	Adams et al., 2010	https://www.phenix-online.org
HKL2000	Otwinowski and Minor, 1997	http://www.hkl-xray.com/
PyMOL	The PyMOL Molecular Graphics System, Version 1.8.4.0, Schrodinger, LLC	http://www.pymol.org
MotionCorr	Li et al., 2013	http://cryoem.ucsf.edu/software/driftcorr.html
MotionCor2	Zheng et al., 2017	http://msg.ucsf.edu/em/software/motioncor2.html
CTFFIND4	Rhou and Grigorieff, 2015	http://grigoriefflab.janelia.org/ctf
FREALIGN	Grigorieff, 2016	http://grigoriefflab.janelia.org/frealign
RELION	Scheres, 2012	http://www2.mrc-lmb.cam.ac.uk/relion/index.php/Main_Page
EMAN2	Ludtke et al., 1999	http://blake.bcm.tmc.edu/EMAN2/
UCSF Chimera	Goddard et al., 2007	https://www.cgl.ucsf.edu/chimera/
Nucleic Acid Builder (AmberTools)	Case et al., 2017	http://casegroup.rutgers.edu/
Other		
Amicon concentrators (30K)	Millipore	Cat#UFC903024
Amicon concentrators (3K)	Millipore	Cat#UFC900324
HisTrap FF (5 ml)	GE Healthcare	Cat#17-5255-01
HiTrap Q Sepharose FF (5 ml)	GE Healthcare	Cat#17-5156-01
HiLoad 16/600 Superdex 200 pg	GE Healthcare	Cat#28989335
Superose 6 10/300 GL	GE Healthcare	Cat#17517201

CONTACT FOR REAGENT AND RESOURCE SHARING

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Sriram Subramaniam (ss1@nih.gov).

METHOD DETAILS**Protein Expression and Purification**

Recombinant proteins were overexpressed in *Escherichia coli* BL21 (DE3) strain in Lysogeny broth (LB) medium or M9 for selenomethionine-derivatized proteins. The cells were grown at 37°C until OD600 reached 0.8 and then induced with 0.25 mM isopropyl β-D-1-thiogalactopyranoside (GoldBio) at 18°C for 20 hr. The full-length *Pseudomonas aeruginosa* (PA14) *cas* genes, *cas5f*, *cas6f*, *cas7f*, and *cas8f* were synthesized and inserted into different expression vectors. *Cas8f* and *cas5f* genes were subcloned into a modified pRSF-Duet-1 vector (Novagen), in which *cas8f* was attached with N-terminal His6-SUMO tag following an ubiquitin-like protease (ULP1) cleavage site. *Cas7f* and *cas6f* genes were subcloned into the first multiple clone sites (MCS) of pCDFDuet-1 and pACYDuet-1 vectors (Novagen), respectively. The synthetic CRISPR was inserted into the second MCS of pACYDuet-1 vector. All the three vectors were co-transformed into *E. coli* BL21 (DE3) strain and co-expressed as described above. Cells were harvested by centrifugation and frozen at –80°C until purification. Cells were resuspended in lysis buffer (20 mM Tris-HCl, pH 8.0, 300 mM NaCl, 5% glycerol, 20 mM imidazole, 1 mM phenylmethylsulfonyl fluoride, 5 mM β-mercaptoethanol), lysed by the EmulsiFlex-C3 homogenizer (Avestin), and centrifuged at 20,000 rpm for 1 hr in a JA-20 fixed angle rotor (Avanti J-E series centrifuge, Beckman Coulter). The supernatant was loaded to 5 mL HisTrap Fast flow column (GE Healthcare) pre-equilibrated in lysis buffer. The column was washed with 20 column volumes of lysis buffer, and the intact complex was eluted with lysis buffer supplemented with 480 mM imidazole. The His6-SUMO tag was removed by ULP1 during dialysis against lysis buffer overnight and then separated by re-loading to Ni-NTA column. The flow-through fraction containing the intact Csy complex was further dialyzed against buffer A (20 mM Tris-HCl, pH 7.5, 150 mM NaCl, 5% glycerol, 5 mM β-mercaptoethanol) and loaded on 5 mL HiTrap Q Fast flow column (GE Healthcare) pre-equilibrated in buffer A. Elution of the Csy complex was achieved by a linear gradient from 150 mM to 1 M NaCl in 20 column volumes.

For Csy^{crRNA}·dsDNA and Csy^{crRNA}·AcrF1, the camera was also on the end of the GIF Quantum energy filter operated in zero-energy-loss mode with a slit width of 20 eV. Movies were recorded in super-resolution mode at a dose rate of 2.6 e⁻/Å²/s with a total exposure time of 15.2 s, for an accumulated dose of 40 e⁻/Å². Intermediate frames were recorded every 0.4 s for a total number of 38 frames per micrograph. Defocus values range from approximately 0.5 – 3 μm. The physical pixel size was 0.825 Å for the Csy^{crRNA}·dsDNA and Csy^{crRNA}·AcrF1 micrographs, and 0.84 Å for the Csy^{crRNA} and Csy^{crRNA}·AcrF10 micrographs.

For Csy^{crRNA}·AcrF2, 3 μL of freshly thawed protein at a concentration of 0.1 mg/ml was applied to plasma-cleaned C-flat R1.2/1.3 400 mesh Cu holey carbon grids (EMS), blotted for 2.5 s after 30 s wait time, and then plunge frozen in liquid ethane, cooled by liquid nitrogen, using the Cryoplunge 3 (Gatan) at 75% relative humidity. Movies collected in counting mode using Legikon (Suloway et al., 2005) at a dose rate of 7.0 e⁻/Å²/s with a total exposure time of 10 – 14 s, for an accumulated dose of 70 – 98 e⁻/Å². Intermediate frames were recorded every 0.2 – 0.25 s for a total of 40 – 50 frames per micrograph. Defocus values range from approximately 1.0 – 3.0 μm, and the physical pixel size was 1.07 Å.

Image processing

Movies recorded on the K2 were aligned by cross-correlation using the cumulative average of previously aligned frames as a reference (Bartesaghi et al., 2014) and CTF estimation performed with CTFFIND4 using a frequency range of 30–3.5 Å for the defocus fit. Particles were picked automatically from the aligned frame averages using a Gaussian disk of 80 Å in radius, extracted using a binning factor of 8 and subjected to 3D refinement in FREALIGN (Grigorieff, 2016). A bimodal distribution of FREALIGN scores was observed and only the particles assigned to the lobe with the highest scores were subjected to 3D classification. Particles corresponding to the best class showing the highest-resolution features were then re-extracted from the original micrographs using a binning factor of 2 and a box size of 512x512 pixels and subjected to an additional 8 rounds of local refinement in FREALIGN (MODE = 1), followed by dose-weighting. The highest resolution information used during all stages of refinement carried out in FREALIGN was set to 4.5 Å for Csy^{crRNA} and Csy^{crRNA}·AcrF10 and to 4.0 Å for Csy^{crRNA}·AcrF1 and Csy^{crRNA}·dsDNA.

In the case of Csy^{crRNA}·AcrF2, movies were aligned using MotionCorr (Li et al., 2013) and MotionCor2 with dose weighting (Zheng et al., 2017) and CTF estimation was performed with CTFFIND4 (Rohou and Grigorieff, 2015). Particles were picked automatically using the Appion DoG Picker (Voss et al., 2009), extracted and subjected to 2D classification in RELION (Scheres, 2012). The best classes were selected for further 3D classification into 10 classes, using an initial model generated *ab initio* by EMAN2's e2initialmodel.py program (Ludtke et al., 1999). For the final reconstruction, particles corresponding to the best two 3D classes were selected and subjected to 3D refinement in RELION using a box size of 256x256 pixels. Micrographs aligned using MotionCorr also underwent particle polishing in RELION.

For all datasets, masks for the FSC calculations were obtained by thresholding low pass filtered versions of the final maps to 20 Å resolution followed by mask apodization using the EMAN2 command e2proc3d.py (–process = mask.auto3d:nshellsgauss = 6) (Tang et al., 2007). Negative B-factors were applied to the final maps for visualization. Local resolution measurements were carried out with the program blocres (Cardone et al., 2013) using unfiltered half maps. Additional processing details specific to each of the datasets are presented in Table S2.

Atomic model building and refinement

Initial models for all five complexes were generated by per-chain rigid-body fitting into their respective electron density maps. The initial models for AcrF1 and AcrF10 are from the crystal structures described in this work. Initial model building into the cryo-EM density maps resulted in coordinates for Csy^{crRNA} subunits that were closely comparable to those in the previously reported atomic model for Csy complexed to AcrF1 and AcrF2 (PDB: 5UZ9) (Chowdhury et al., 2017); we therefore proceeded with the reported model as a starting point for real space refinement. The initial model of the RNA:DNA heteroduplex was generated through rigid-body fitting short segments of a model of Cse-dsDNA (PDB: 5H9F) (Hayes et al., 2016). To model the DNA duplex, we used an ideal B-form DNA structure generated by Nucleic Acid Builder (Case et al., 2017). Due to the observed flexibility of Cas8f, we considered the flexible portion of the hook domain (residues 1–65, 107–168) and the four-helix bundle domain (residues 265–435) as separate rigid bodies for fitting.

In the next stage, models were refined into cryo-EM density maps using real-space refinement in PHENIX with rotamer, Ramachandran plot and Cβ deviation restraints (Adams et al., 2010). After automated refinement, models were manually inspected in COOT (Emsley et al., 2010) and UCSF Chimera (Goddard et al., 2007) and were further refined or modified. Several such iterative rounds of automated refinement in PHENIX and manual inspection in COOT and Chimera were performed. In the first round, we used simulated annealing, morphing (local rigid body fit), and gradient-driven minimization against an unsharpened map to capture tertiary structural changes. In intermediate rounds, we used morphing, minimization, and local rotamer fitting to a sharpened map. In the final round, we use only minimization and local rotamer fitting against a sharpened map. Secondary structure restraints were applied for Cas8f and Cas5f, and base-pair restraints were applied to the duplex DNA to prevent these structures from overfitting to noise. Secondary structure and non-crystallographic symmetry (NCS) restraints were also applied to Cas7f as necessary.

Figure generation

Figures of map density and coordinates were created in UCSF Chimera (Goddard et al., 2007), with the exception of Figures S6 and S7, which were created in PyMol (Schrodinger, 2015). For the 3D schematic models (Figures 1D, 3A, 4C, and 7D), the coordinates were imported to Maxon Cinema4D using the plugin Embedded Python Molecular Viewer (ePMV) (Johnson et al., 2011) and a Coarse Molecular Surface was generated. The figures were then rendered from Cinema4D.

DATA AND SOFTWARE AVAILABILITY**Data Resources**

The accession numbers for the maps and coordinates for the five cryo-EM structures reported in this paper are EMDB: 7048, 7049, 7050, 7051, and 7052 and PDB: 6B44, 6B45, 6B46, 6B47, and 6B48.

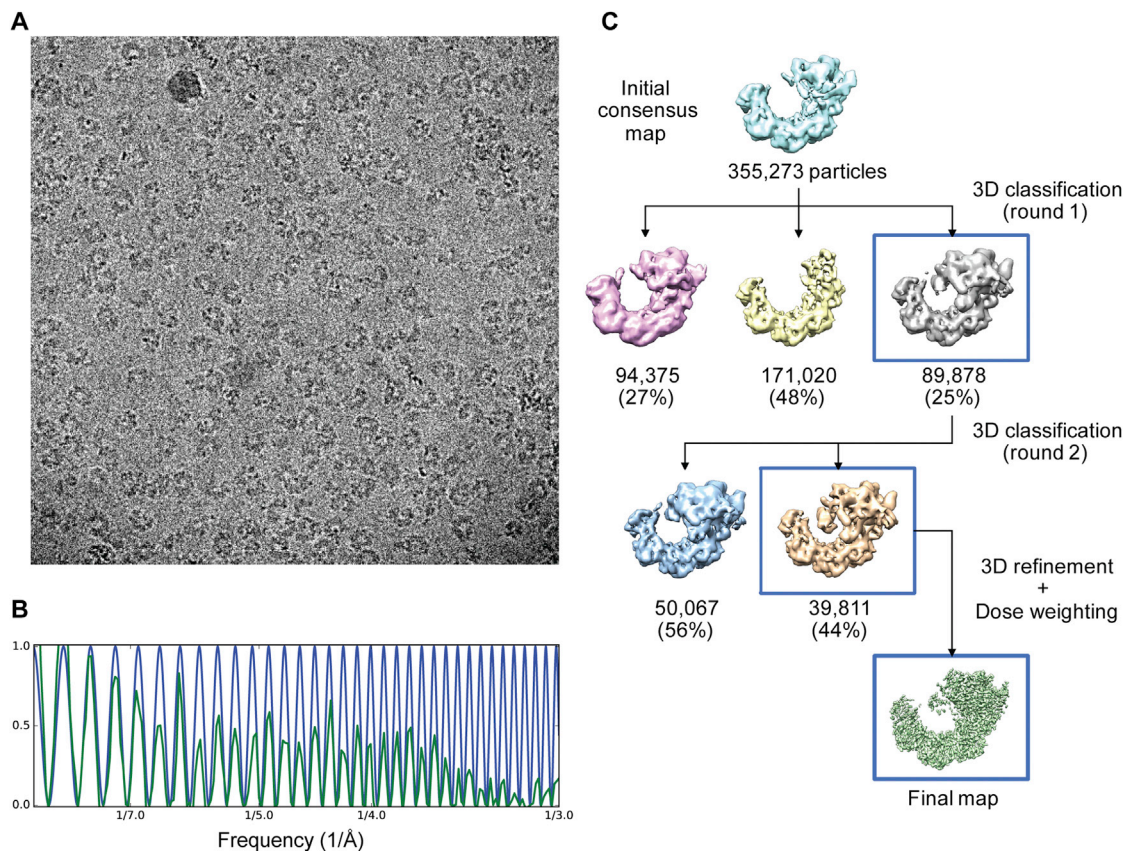


Figure S1. Cryo-EM of Csy Complex Bound to dsDNA, Related to Figure 1

(A) Representative cryo-EM micrograph, averaged over 38 aligned movie frames, of the Csy complex bound to dsDNA.

(B) 1D profile obtained by rotationally averaging the 2D power spectrum of image shown in (panel A) (green curve) and corresponding theoretical CTF fit at 1.8 μm defocus (blue curve) showing signal in the raw data extending to 3.0 \AA resolution.

(C) Overview of image refinement protocol used to obtain the Csy^{crRNA}-dsDNA cryo-EM density map. An initial consensus map was obtained from the starting set of 355,273 particles using binned data (8x), followed by two successive rounds of 3D classification into three and two classes, respectively. The highest resolution class was selected at the end of each round yielding a final homogeneous set of 39,811 particles which were then re-extracted from the original micrographs using a lower binning factor (2x) and subjected to an additional round of 3D refinement combined with dose weighting, resulting in a final density map at 2.9 \AA resolution. See also Table S2.

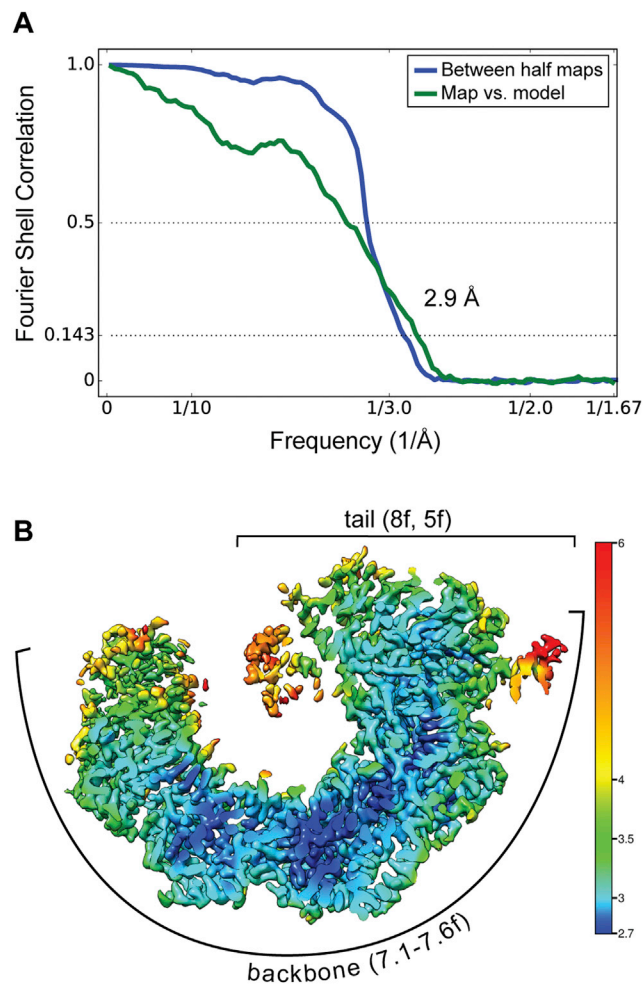


Figure S2. Assessment of Map Resolution for Csy Complex Bound to dsDNA, Related to Figure 1

(A) Fourier Shell Correlation (FSC) curve between two semi-independently refined halves of the data (blue curve), and between the cryo-EM map and corresponding refined model (green curve).

(B) Cut-away view of 3D reconstruction of Csy^{crRNA}.dsDNA complex colored according to local resolution as reported by blocres. See also Table S2.

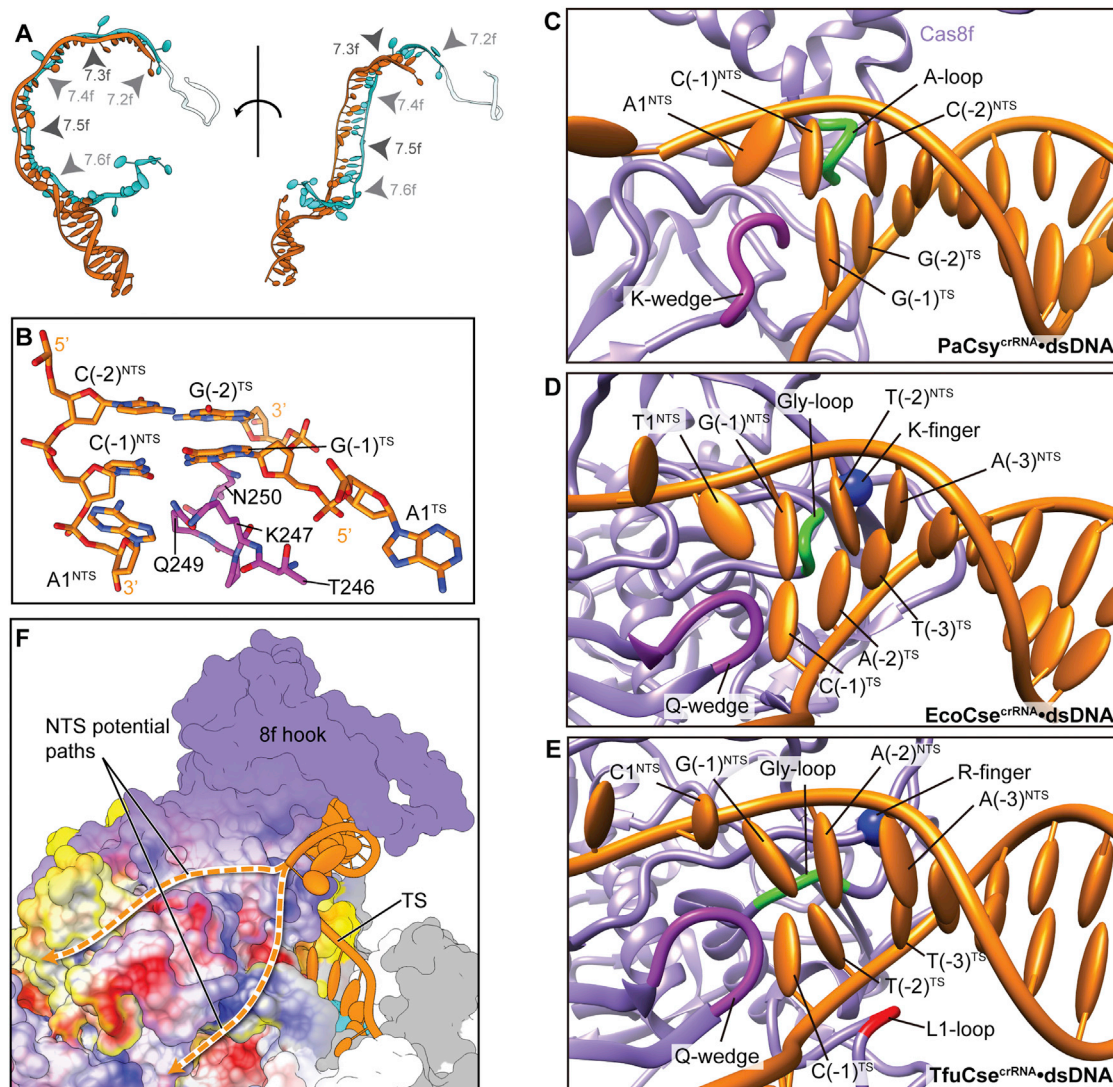


Figure S3. RNA/DNA Conformation, Recognition, and Potential Paths, Related to Figures 2 and 3

(A) Atomic model of crRNA (cyan) and dsDNA (orange) in two orthogonal views. The 3' stem loop of crRNA is not visualized in the cryo-EM density map and is shown in white. The gaps of RNA:DNA hybrid and binding site for Cas7f thumbs are highlighted. Compare to Figure 2E.

(B) Model for specific interactions between the K-wedge (magenta) and PAM fork residues (orange). The K-wedge appears to be responsible for both DNA strand separation and PAM(-1) sequence discrimination. Compare to Figure 3D.

(C) Interface between PAM fork (orange) and Cas8f (purple) in *PaCsy*^{crRNA}-dsDNA complex, highlighting the K-wedge (magenta) and A-loop (green) which are responsible for the recognition of PAM(-1) and PAM(-2) sites.

(D) Interface between PAM fork (orange) and Cas8e (purple) in *EcoCse*^{crRNA}-dsDNA complex (PDB ID: 5H9F), highlighting the Q-wedge (magenta), Gly-loop (green), and K-finger (blue sphere) which are responsible for the recognition of PAM(-1), PAM(-2), and PAM(-3) sites.

(E) Interface between PAM fork (orange) and Cas8e (purple) in *TfuCse*^{crRNA}-dsDNA complex (PDB ID: 5U0A), highlighting the Q-wedge (magenta), Gly-loop (green), L1-loop (red), and R-finger (blue sphere) which are responsible for the recognition of PAM(-1), PAM(-2), and PAM(-3) sites.

(F) Electrostatic surface of Cas8f and Cas5f, with potential binding paths for the displaced non-target DNA strand highlighted.

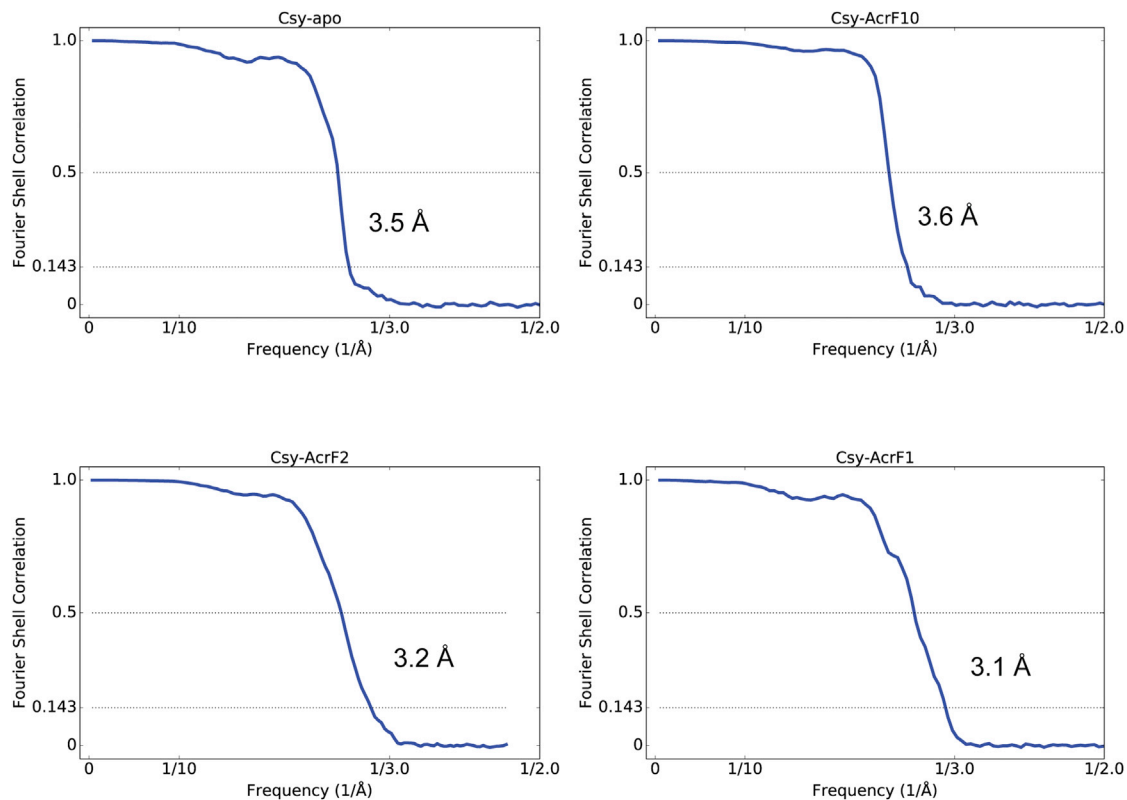


Figure S4. Assessment of Map Resolution for Apo and Inhibitors-Bound Csy Complexes, Related to Figures 4–7

Fourier Shell Correlation curves estimating the resolution of the final maps for unbound Csy, and inhibitor-bound complexes Csy^{crRNA}·AcrF10, Csy^{crRNA}·AcrF2 and Csy^{crRNA}·AcrF1.

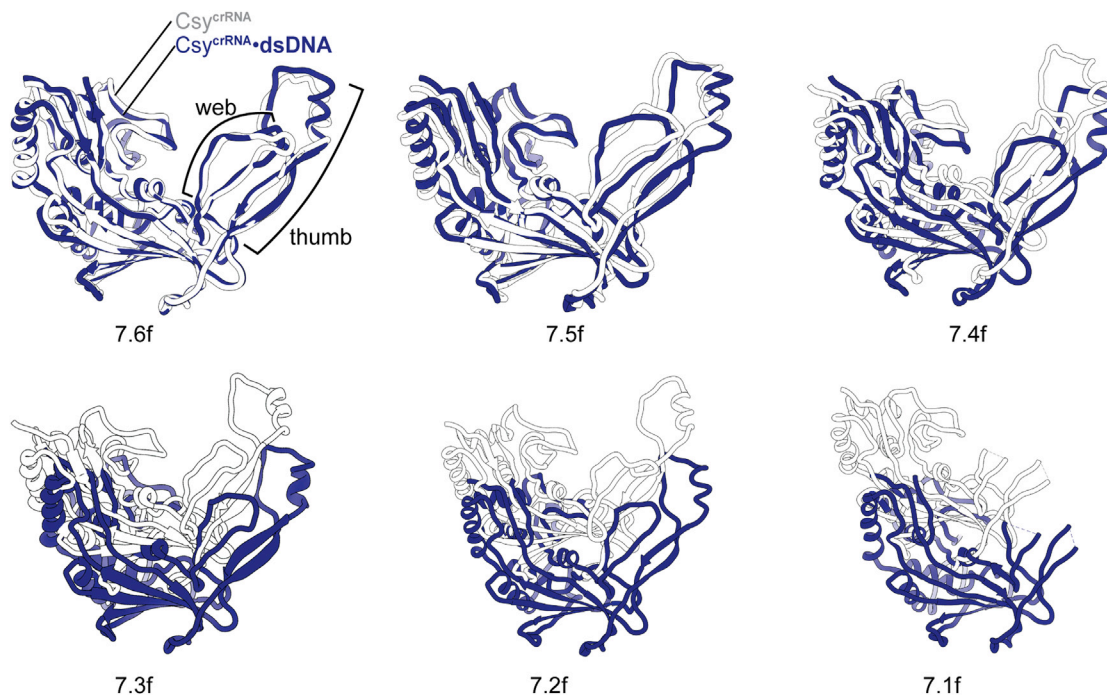


Figure S5. Progressive Displacement of Cas7f Subunits upon DNA Binding, Related to Figure 4

Comparison of each Cas7f subunit in Csy^{crRNA} (white) and $Csy^{crRNA} \cdot dsDNA$ (blue). The subunits move primarily as rigid bodies and become progressively displaced as a result of DNA binding. Structures are aligned at Cas7.6f.

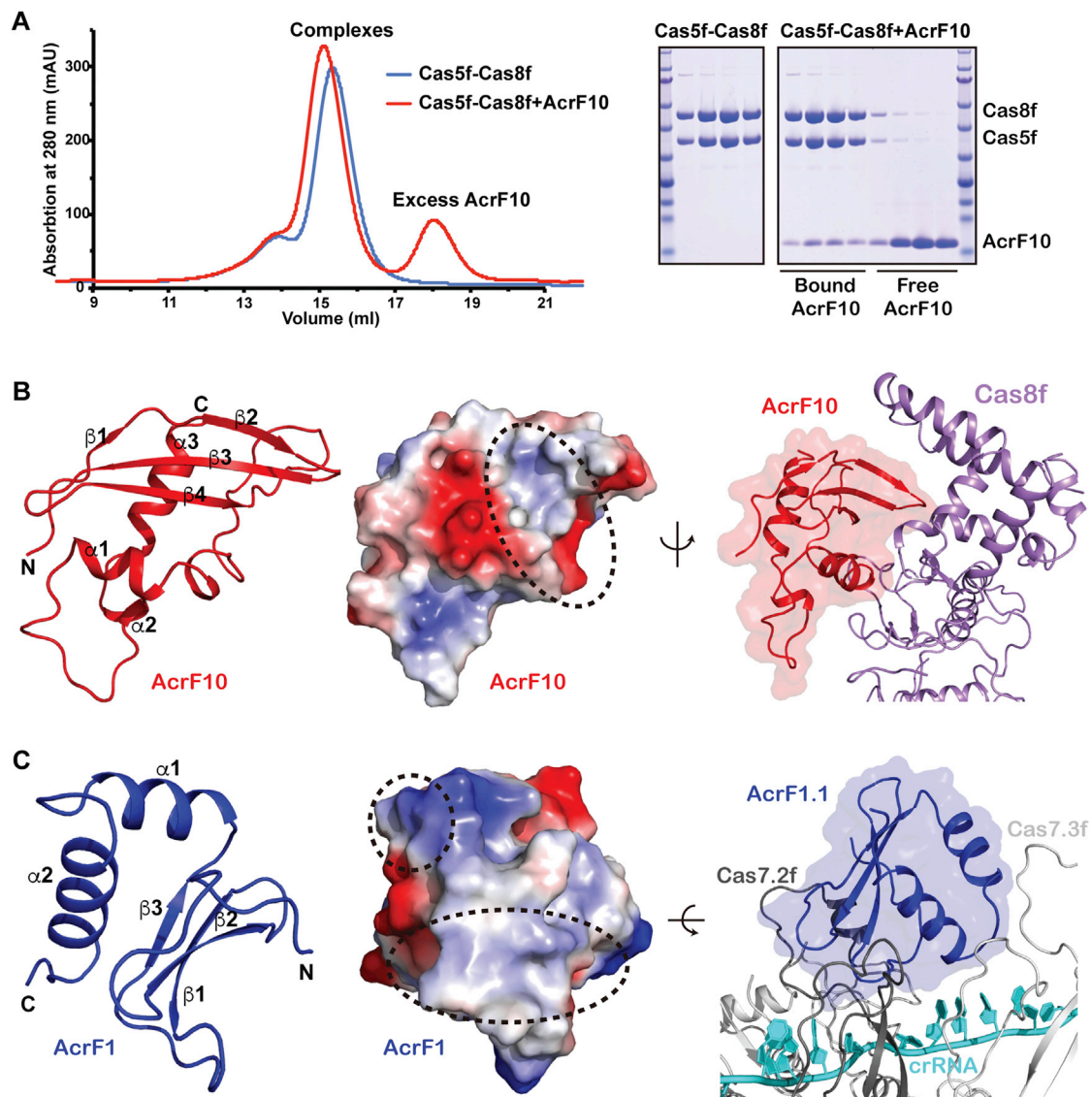


Figure S6. Illustration of Biochemical Differences in Binding of AcrF10 and AcrF1 to Csy, Related to Figures 5 and 7

(A) AcrF10 forms a stable complex with Cas5f-Cas8f in solution. Size-exclusion chromatography (SEC) (left panel) was performed in absence or presence of AcrF10. The fractions were analyzed by SDS-polyacrylamide gel electrophoresis (SDS-PAGE) (right panel).

(B) Crystal structure of AcrF10 (left panel) and electrostatic surface of AcrF10 (middle panel), as well as the interface between AcrF10 and Cas8f (right panel). The interaction surface with the Csy complex is circled with a dotted line.

(C) Crystal structure of AcrF1 (left panel) and electrostatic surface of AcrF1 (middle panel), as well as interface between AcrF1 and Cas7f (right panel). The interaction surface with the Csy complex is circled with a dotted line. See also Table S1.

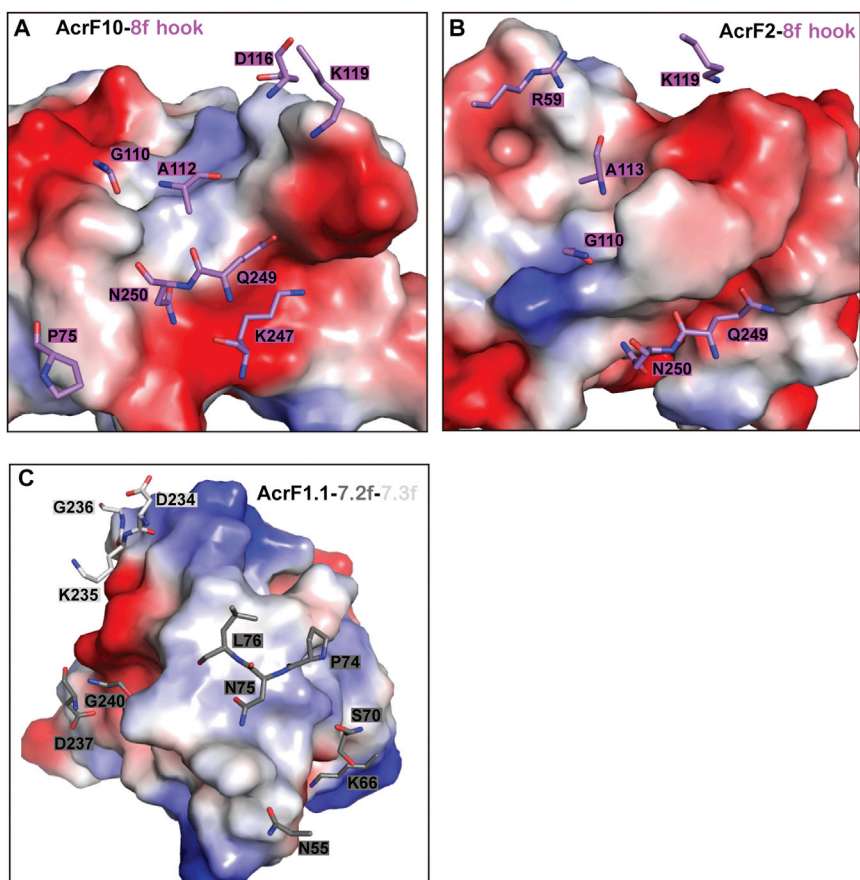


Figure S7. Intermolecular Contacts between AcrF Proteins and Subunits of Csy Complex, Related to Figures 5–7

(A) Interfaces between AcrF10 and Cas8f hook domain, with the binding site on AcrF10 represented as electrostatic surface and the potential interaction partners of Cas8f represented as sticks.

(B) Interfaces between AcrF2 and Cas8f hook domain, with the binding site on AcrF2 represented as electrostatic surface and the potential interaction partners of Cas8f represented as sticks.

(C) Interfaces between AcrF1.1 and Cas7f subunits, with the binding site on AcrF1.1 represented as electrostatic surface and the potential interaction partners of Cas7f subunits represented as sticks.