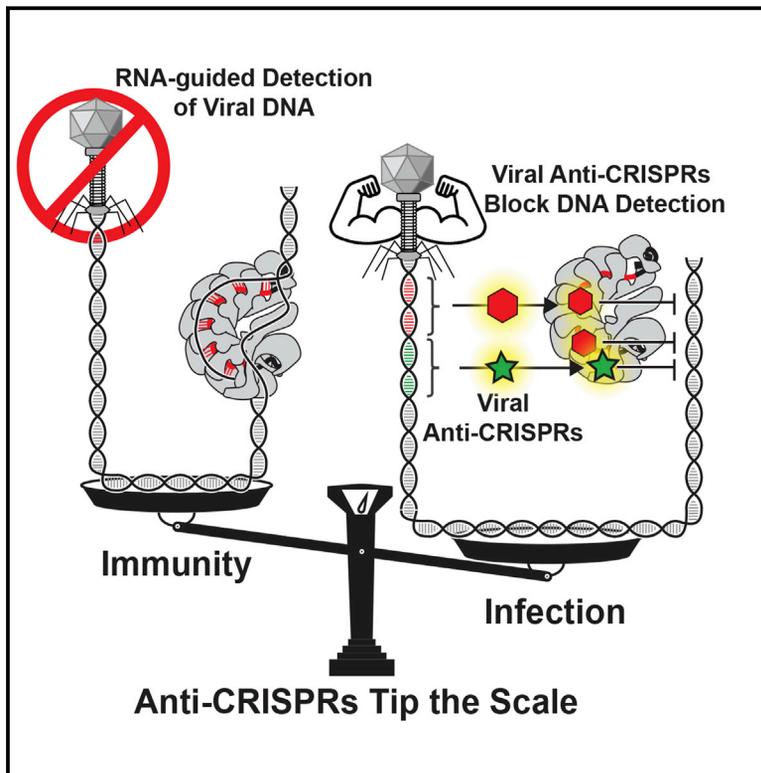


Structure Reveals Mechanisms of Viral Suppressors that Intercept a CRISPR RNA-Guided Surveillance Complex

Graphical Abstract



Authors

Saikat Chowdhury, Joshua Carter, MaryClare F. Rollins, ..., Elizabeth R. Fischer, Gabriel C. Lander, Blake Wiedenheft

Correspondence

glander@scripps.edu (G.C.L.), bwiedenheft@gmail.com (B.W.)

In Brief

The high-resolution structures of a CRISPR surveillance complex with two viral anti-CRISPR proteins reveal different strategies for silencing CRISPR immune function.

Highlights

- Cryo-EM structure of crRNA-guided surveillance complex bound to two anti-CRISPRs
- Anti-CRISPRs bind to residues that are essential for crRNA-guided DNA binding
- Cas7f backbone subunits have unique fold, suggesting a unique evolutionary trajectory
- AcrF2 is a molecular mimic of double-stranded DNA



Structure Reveals Mechanisms of Viral Suppressors that Intercept a CRISPR RNA-Guided Surveillance Complex

Saikat Chowdhury,^{1,7} Joshua Carter,^{2,7} MaryClare F. Rollins,² Sarah M. Golden,² Ryan N. Jackson,^{2,8} Connor Hoffmann,² Lyn'Al Nosaka,¹ Joseph Bondy-Denomy,³ Karen L. Maxwell,⁴ Alan R. Davidson,^{4,5} Elizabeth R. Fischer,⁶ Gabriel C. Lander,^{1,*} and Blake Wiedenheft^{2,9,*}

¹Department of Integrative Structural and Computational Biology, Scripps Research Institute, La Jolla, CA 92037, USA

²Department of Microbiology and Immunology, Montana State University, Bozeman, MT 59717, USA

³Department of Microbiology and Immunology, University of California San Francisco, San Francisco, CA 94158, USA

⁴Department of Biochemistry

⁵Department of Molecular Genetics

University of Toronto, Toronto, ON, Canada

⁶Research Technologies Branch, Rocky Mountain Laboratories, National Institute of Allergy and Infectious Diseases, NIH, Hamilton, MT 59840, USA

⁷Co-first author

⁸Present address: Department of Chemistry and Biochemistry, Utah State University, Logan, UT 84322, USA

⁹Lead Contact

*Correspondence: glander@scripps.edu (G.C.L.), bwiedenheft@gmail.com (B.W.)

<http://dx.doi.org/10.1016/j.cell.2017.03.012>

SUMMARY

Genetic conflict between viruses and their hosts drives evolution and genetic innovation. Prokaryotes evolved CRISPR-mediated adaptive immune systems for protection from viral infection, and viruses have evolved diverse anti-CRISPR (Acr) proteins that subvert these immune systems. The adaptive immune system in *Pseudomonas aeruginosa* (type I-F) relies on a 350 kDa CRISPR RNA (crRNA)-guided surveillance complex (Csy complex) to bind foreign DNA and recruit a *trans*-acting nuclease for target degradation. Here, we report the cryo-electron microscopy (cryo-EM) structure of the Csy complex bound to two different Acr proteins, AcrF1 and AcrF2, at an average resolution of 3.4 Å. The structure explains the molecular mechanism for immune system suppression, and structure-guided mutations show that the Acr proteins bind to residues essential for crRNA-mediated detection of DNA. Collectively, these data provide a snapshot of an ongoing molecular arms race between viral suppressors and the immune system they target.

INTRODUCTION

Viruses that infect bacteria (i.e., phages) are the most diverse and abundant biological agents on the planet, and bacteria have evolved sophisticated adaptive immune systems to protect themselves from phage predation (Rodriguez-Valera et al., 2009; Suttle, 2007; van der Oost et al., 2014). These immune systems rely on CRISPR loci and a diverse cassette of CRISPR-associ-

ated (Cas) genes (Makarova et al., 2011; 2015). Immunity to phages is acquired and passed on to progeny by integrating short fragments of foreign DNA into the host CRISPR locus. These loci are transcribed and processed into short CRISPR-derived RNAs (crRNAs) that are incorporated into multi-subunit crRNA-guided surveillance complexes that recognize, bind, and degrade foreign nucleic acids.

According to the most recent phylogenetic classification, CRISPR systems are divided into two classes, six types, and 19 subtypes (Makarova et al., 2015; Shmakov et al., 2015). Collectively, these immune systems represent a diverse arsenal of sequence-specific defense mechanisms that efficiently block invasion by mobile genetic elements, including phages and plasmids. However, phages and plasmids have evolved a diverse repertoire of anti-CRISPR (Acr) proteins that subvert these immune systems (Bondy-Denomy et al., 2013; 2015; Pawluk et al., 2014; 2016). In general, Acr proteins are difficult to identify using sequence-based bioinformatic techniques due to their low sequence similarity. However, Pawluk et al. (2016) recently identified a conserved transcriptional regulator that serves as a genetic landmark for identifying putative Acr proteins, and phylogenetic analyses indicate that the distribution of these proteins mirrors the distribution of the CRISPR-Cas systems that they inactivate.

To elucidate the mechanistic differences that give rise to phylogenetic distinctions between each of the 19 different CRISPR-mediated immune systems, and to understand the molecular interactions that enable different Acr proteins to specifically subvert the type I-F immune system in *P. aeruginosa*, we determined the cryo-electron microscopy (cryo-EM) structure of the a 350 kDa crRNA-guided surveillance complex (i.e., Csy complex) in association with two distinct anti-CRISPRs (Figures 1, S1, S2, Table S1, and Movie S1). The structure was determined at an average resolution of 3.4 Å, with the most

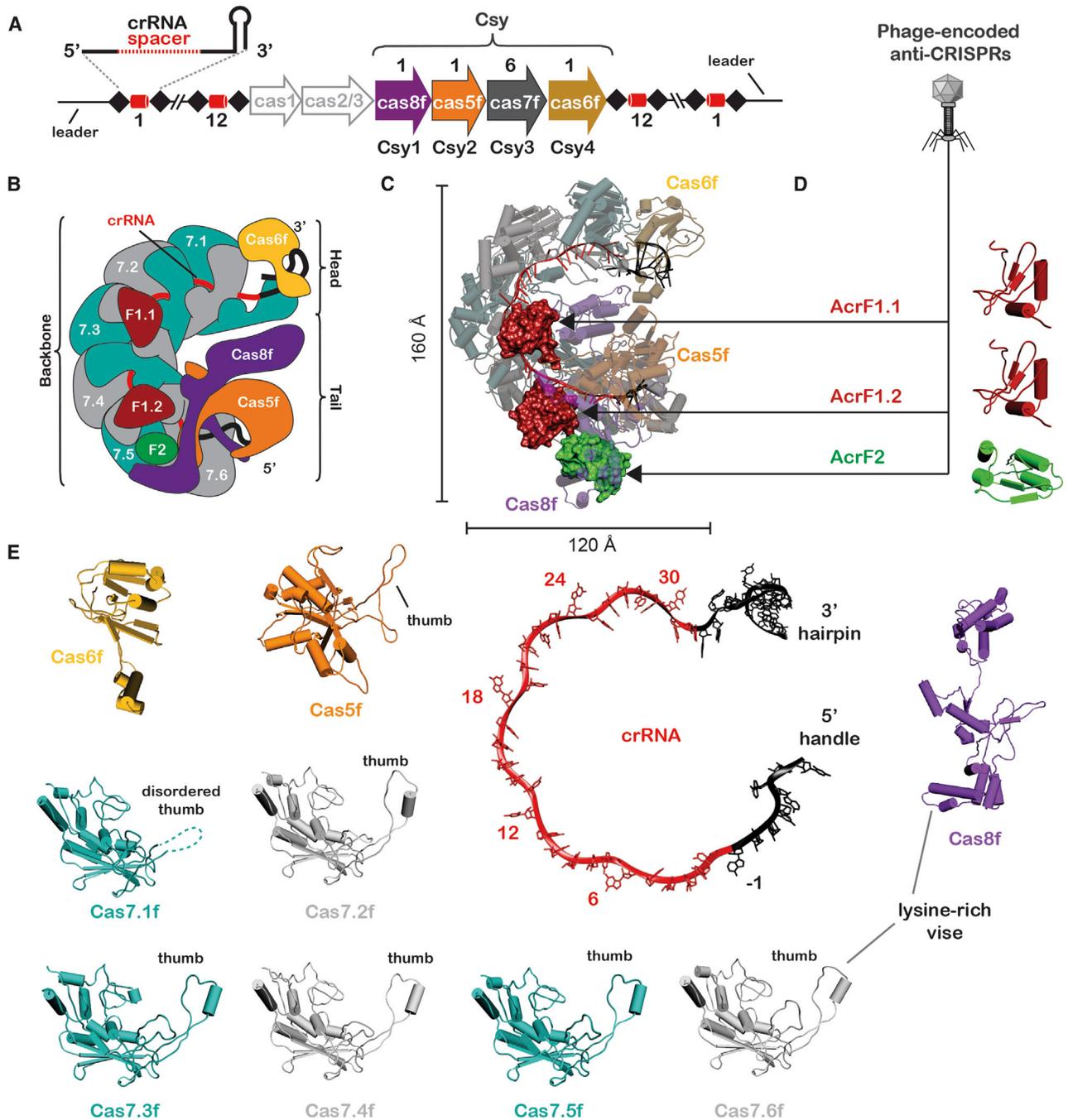


Figure 1. Structure of the Csy Complex Bound to Two Different Virally Encoded Anti-CRISPR Proteins

(A) The type I-F CRISPR-mediated immune system in *P. aeruginosa* (strain PA14) consists of six *cas* genes (legacy names are noted below each arrow) flanked by two CRISPR loci. The CRISPR loci are comprised of 28-nucleotide repeats (black diamonds) separated by 32-nucleotide phage- or plasmid-derived spacer sequences (red cylinders) that lie downstream of an AT-rich leader sequence.

(B) Schematic of Csy complex bound to two molecules of anti-CRISPR protein AcrF1 (red, F1.1 and F1.2) and one molecule of AcrF2 (green, F2).

(C) Atomic model of the Csy complex (transparent) bound to AcrF1.1, AcrF1.2, and AcrF2. AcrF1 and F2 are shown as red and green surfaces, respectively.

(D) Structures of the virally encoded anti-CRISPR proteins (Acr) and their locations in the complex.

(E) Individual subunits of the Csy complex. The “thumb” of each Cas7f subunit and Cas5f folds over the top of the crRNA, creating a kink in the RNA at 6-nucleotide intervals (positions –1, 6, 12, 18, 24, and 30).

See also [Figure S1](#) and [Table S1](#).

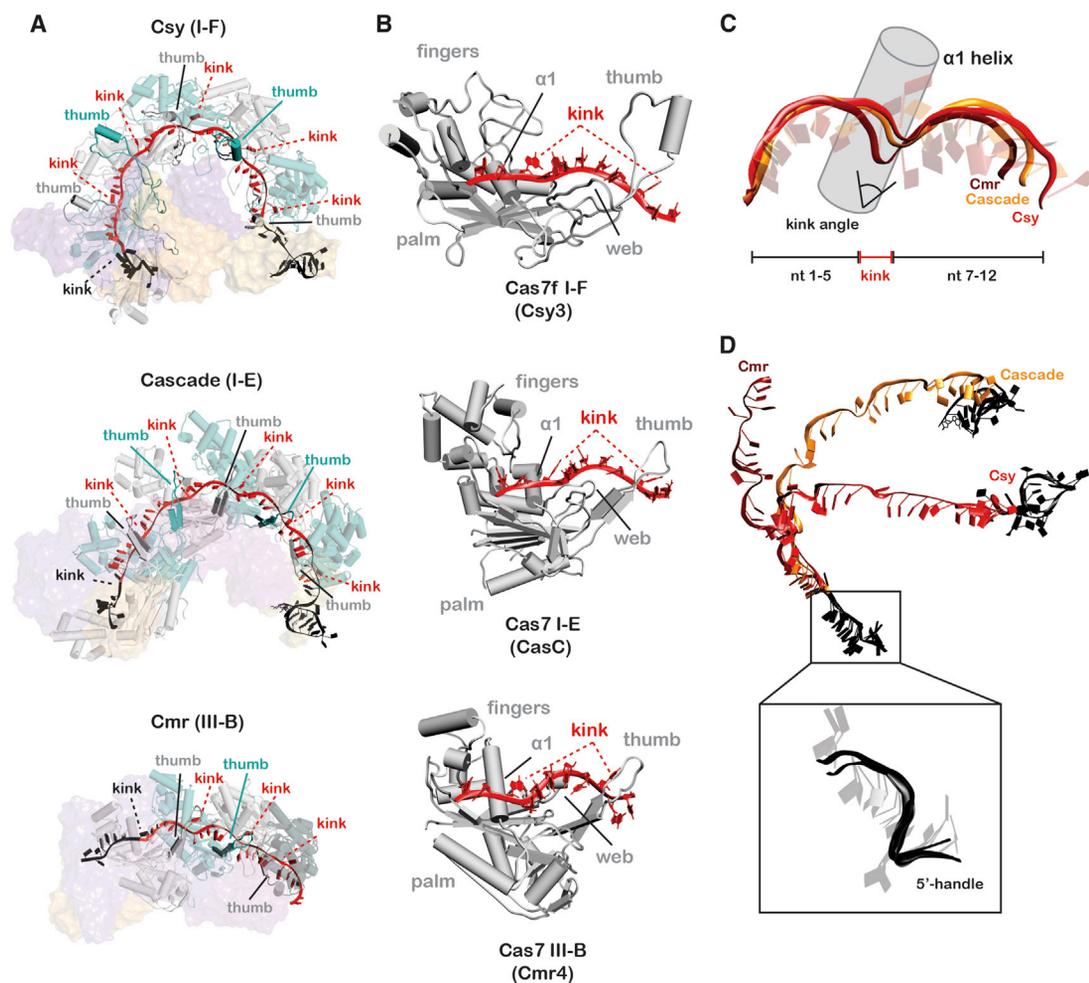


Figure 2. Structural Similarities and Differences between the Cas7-crRNA Backbones of Class 1 CRISPR-Cas Systems

(A) Comparison of type I-F (Csy), I-E (Cascade), and III-B (Cmr). The Cas7 backbone proteins (blue and gray) are shown as pipes and planks with head, tail, and belly proteins shown as transparent surfaces.

(B) A single Cas7 homolog from each system (gray) bound to crRNA (red). Kinked bases formed by the thumbs are highlighted.

(C) Nucleotides 1–5 (nt 1–5) of the crRNA-guide from each system were superimposed. The next six bases of each crRNA diverge after the kink.

(D) Eight repeat-derived nucleotides on the 5' end of the crRNA (black, also called the 5' handle) were aligned using Chimera (Goddard et al., 2005). Differences in kink angles result in crRNAs with very different pitches. Expanded view highlights the conserved structure of the 5' handle.

See also Figure S4.

conformationally stable regions resolved to 3.0 Å, enabling *de novo* atomic modeling and providing a detailed structural understanding of the Acr-Csy interactions (Figures S1 and S2).

RESULTS

Structural Overview of the Csy Complex

The Csy complex is a crRNA-guided surveillance complex composed of nine Cas proteins (one Cas6f, six Cas7fs, one Cas5f, and one Cas8f protein) and a single 60-nt crRNA (Figure 1) (van Duijn et al., 2012; Wiedenheft et al., 2011). Overall, the morphology of the complex is consistent with the seahorse shape that has been described for other class 1 CRISPR-Cas complexes, with subunits referred to as the head (Cas6), backbone (Cas7), and tail (Cas5 and Cas8) (Figures 1 and 2) (Jackson

and Wiedenheft, 2015; Zhang and Sontheimer, 2014). The crRNA is an integral component of the complex, making direct contact with all nine protein subunits (Figure 1). The crRNA performs an important structural role in complex assembly by tethering the protein subunits of the complex together and defining the number of Cas7 subunits incorporated into the backbone (Kuznedelov et al., 2016; Luo et al., 2016; Wiedenheft et al., 2011). The complex is assembled such that a portion of the crRNA (sometimes called the spacer) guides the complex to complementary nucleic acids found in phage and plasmid genomes, and target binding recruits the Cas2/3 nuclease-helicase protein for degradation of the target. Phage-encoded anti-CRISPR proteins AcrF1 and AcrF2 (Figures 1B–1D, red and green, respectively) block crRNA-guided target recognition by binding to either the Cas7f backbone (AcrF1), which blocks

hybridization to a complementary target, or by inserting between the Cas8f and Cas7f subunits in the tail (AcrF2), which prevents interactions with the target DNA duplex (Figure 1). The Acr stoichiometries observed are consistent with our previous work showing AcrF1 bound predominantly to multimeric Cas7f in a ratio of 2.6 ± 0.3 proteins per Csy complex, while AcrF2 bound to Cas8f with a ratio of 0.8 ± 0.1 (Bondy-Denomy et al., 2015).

The CRISPR loci in *P. aeruginosa* are composed of 28-nucleotide repeat sequences separated by 32-nucleotide phage- or plasmid-derived spacer sequences. The repeat sequences are partially palindromic, giving rise to a series of stable stem-loop structures within the long precursor CRISPR transcripts. These stem-loop structures are recognized by a CRISPR-specific endoribonuclease, Cas6f (formerly Csy4), which cleaves the CRISPR RNA at the 3' end of each stem-loop to produce a library of 60-nt crRNAs required for assembly of the Csy surveillance complex (Wiedenheft et al., 2011; Haurwitz et al., 2010). Cas6f and the 3' hairpin of the crRNA are not well resolved in cryo-EM density, suggesting that the linkage tethering Cas6f to the rest of the complex is flexible (Figure S1J). However, a high-resolution crystal structure of the *P. aeruginosa* Cas6f protein bound to the 3' hairpin of the crRNA was previously determined (Haurwitz et al., 2010), and the EM density in this region was sufficiently resolved to enable unambiguous rigid body fitting of the atomic model into position at the head of the complex (Figure 1). In addition, two copies of an NMR structure for AcrF1 were used to facilitate model building (Maxwell et al., 2016). Models for AcrF2 and all other Cas proteins were built de novo using the EM density. The tail of the complex, consisting primarily of Cas5f (formerly Csy2) and Cas8f (formerly Csy1) were not as well resolved as the Cas7f backbone, so a focused classification and refinement strategy using signal subtraction was employed to improve the density in this region for de novo atomic model building (see STAR Methods) (Figure S3).

Ordered Presentation of the crRNA Guide

Phylogenetic and structural studies suggest that multi-subunit crRNA-guided surveillance complexes found in type I and type III immune systems evolved from a common ancestor (Jackson and Wiedenheft, 2015; Makarova et al., 2015; Rouillon et al., 2013; Spilman et al., 2013). A unifying feature of these systems, which are now collectively recognized as class 1 immune systems, is the helical assembly of a Cas7 family protein along the crRNA (Figure 2A). Consistent with previously determined Cas7 structures, the Cas7f protein from *P. aeruginosa* shares a “right-hand” morphology composed of fingers-, palm-, web-, and thumb-shaped domains (Figure 2B). In *P. aeruginosa*, crRNA binding by Cas7f is mediated by non-sequence specific contacts between the sugar-phosphate backbone and residues on the palm (R35, H275, Q277, K278, N281, R284) and web (R169, Q248). The thumb of Cas7.1f is anticipated to be part of the flexible tether that connects the Cas7f backbone to the Cas6 head, but due to the flexibility of the head, the density for the thumb is not well resolved. In contrast, thumbs of the remaining Cas7f subunits (Cas7.2f–7.6f) are well ordered and fold over the top of the crRNA and across the palm of the adjacent subunit (Figures 2A and 2B). Similar to what has been reported for the other

class 1 complexes, Cas7f oligomerization along the crRNA introduces distortions in the sugar-phosphate backbone, resulting in “kinks” at regular 6-nucleotide intervals (Taylor et al., 2015; Hayes et al., 2016; Hochstrasser et al., 2014; Jackson et al., 2014a; van Erp et al., 2015). The thumb of each Cas7f subunit folds over the top of each kink, burying one nucleobase at positions –1, 6, 12, 18, 24, and 30 of the crRNA (Figures 2A–2C). The five nucleotides between each kink are ordered in a pseudo-A-form configuration, and the sugar-phosphate backbone of each helical segment superimposes on the backbone of each of the other segments (e.g., nucleotides 1–5 superimpose on nucleotides 7–11, etc.) with an RMSD of less than 0.37 Å for equivalently positioned phosphates.

Presentation of the crRNA guide in helical segments of 5 nucleotides is a conserved feature of class 1 surveillance complexes. In fact, the sugar-phosphate backbones for helical segments of the crRNA from Csy (type I-F), Cascade (type I-E), and Cmr (type III-B) superimpose with an RMSD of < 1.0 Å (Figure 2C). While the structure of each helical segment is conserved among class 1 surveillance complexes, torsion angles at each kink differ significantly between these systems. Torsion angles are determined by specific interactions between the crRNA phosphate backbone and the palm, web, and thumb domains of the subtype-specific Cas7 proteins. The resulting differences in the torsion angles at each kink correspond with large-scale structural differences in the helical pitch of the Cas7 backbone (Figure 2D). The backbone of the I-F Csy complex adopts a substantially tighter pitch (80 Å) compared to the I-E (127 Å) or III-B complexes (158 Å). The tighter spiral of the Csy backbone results in the head and tail subunits being positioned in close proximity, creating a nearly closed ring-like structure. The functional significance of these structural differences is currently not well understood, but we do know that most anti-CRISPRs are subtype specific (Pawluk et al., 2014; 2016), and these preferences likely reflect structural features that are unique to each of the different subtypes.

While the “right-hand” morphology used to describe Cas7 family proteins is conserved in Cas7f, the Cas7f fold is distinct (Figure S4). Previously determined Cas7 structures contain palm domains with the $\beta\alpha\beta\alpha\beta$ topology of an RNA recognition motif (RRM), where the two antiparallel α helices pack against one face of a four-stranded antiparallel β sheet that has a specific $\beta_2\beta_3\beta_1\beta_4$ arrangement (Topuzlu and Lawrence, 2016). In contrast, the Cas7f palm domain is an antiparallel β sheet with a notably different $\beta_3\beta_2\beta_4\beta_1$ arrangement, and a structural homology search performed using the Cas7f palm domain reveals similarity to the PAD domain of eukaryotic polymerases, rather than other Cas7 family proteins (Holm and Sander, 1993; Trincao et al., 2001). Furthermore, previously determined Cas7 structures contain highly decorated RRM motifs with multiple insertions that form the fingers and thumb domains, while the Cas7f palm contains only one insertion, with the thumb and fingers domains formed by N- and C-terminal extensions, respectively (Figure S4). In addition to the previously described domains, Cas7f contains a 38-residue loop in the C-terminal extension, termed the “extended web,” which packs against the thumb, creating a prominent trough connecting the web and thumb domains. Structurally equivalent loops exist in other type I Cas7 proteins, although they are significantly smaller.

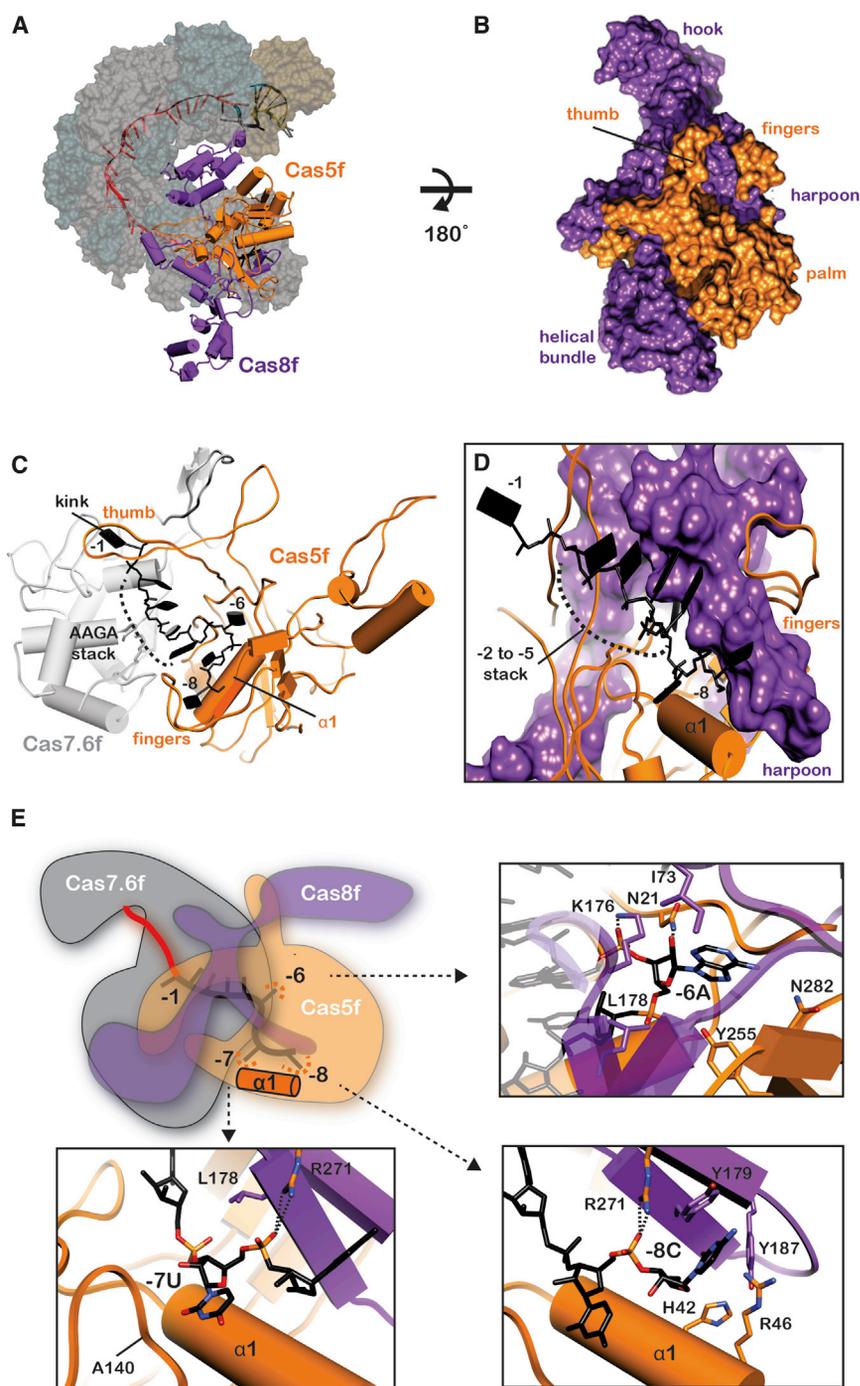


Figure 3. Assembly of the Cas5f-Cas8f Tail through Recognition of the 5' Handle

(A) Surface view of the Csy complex, with Cas5f (orange) and Cas8f (purple) depicted as pipes and planks.

(B) Surface representation of Cas5f (orange), showcasing the “left-handed fist” morphology and interactions with the Cas8f protein (purple).

(C) Recognition of the S-shaped architecture of the 5' handle by Cas7.6f (gray) and Cas5f (orange). (D) The 5' end of the crRNA (black) is sandwiched between Cas8f harpoon and the first helix ($\alpha 1$) of the Cas5f RRM.

(E) Cartoon depiction of Cas7.6f, Cas5f, Cas8f, and the 5' end of the crRNA. Arrows point to detailed interactions between the Cas5f-Cas8f heterodimer and the crRNA at positions -6 , -7 , and -8 .

that resembles fingers and a loop that forms a thumb-like feature (Figure 3B). Each of these anatomical features plays a role in binding Cas8f, the 5' handle of the crRNA, or both (Figures 3A–3D).

The Cas5f fingers “grip” a long β hairpin on Cas8f (residues 175–191) that extends through the Cas5f subunit. We refer to this β hairpin as the Cas8f harpoon (Figure 3B). The last two nucleotides of the crRNA 5' handle (positions $-8C$ and $-7U$) are sandwiched between the Cas8f harpoon and the first α helix of the Cas5f RRM, while the next nucleotide at position $-6A$ is displaced and the phosphate backbone is kinked by interactions with lysine 176 (K176) on Cas8f and asparagine 21 (N21) on Cas5f (Figures 3C–3E). This network of interactions creates the first curve of the S-shaped 5' handle, which transitions into the next curve consisting of four bases (positions $-5A$, $-4G$, $-3A$, and $-2A$) that are arranged in a pseudo-A-form configuration (Figures 2C and 2D). The phosphate backbone of these four nucleotides is stabilized by non-sequence-specific interactions with Cas7.6f, while bases on either end of the helical stack are stabilized by interactions with Cas5f (i.e., R271 and N86).

Organization of the Tail

The tail of the Csy complex is composed of a stable Cas5f-Cas8f heterodimer that recognizes a conserved S-shaped RNA structure, called the 5' handle, which is formed by the final 8 nucleotides on the 5' end of the crRNA (Figures 1 and 3). The Cas5f protein adopts a “left-handed fist” morphology consisting of a canonical RRM (i.e., $\beta_2\alpha_2\beta_3\beta_1\alpha_1\beta_4$) that forms the palm domain flanked by a four-stranded antiparallel beta sheet ($\beta_2\beta_1\beta_4\beta_3$)

The thumb of Cas5f is formed by a large insertion between β_2 and β_3 of the RRM (residues 48–109). This thumb functions similarly to the Cas7f thumbs by folding over the crRNA at position $-1A$ (Figure 3C). In the type I-E surveillance complex, interactions between the Cas5 thumb and the fingers domain of Cas7.6 coincide with a $\sim 180^\circ$ rotation of the Cas7 fingers. This rotation creates a “lysine-rich vise” between Cas7.5 and 7.6 that is critical for binding dsDNA (Jackson et al., 2014a; van

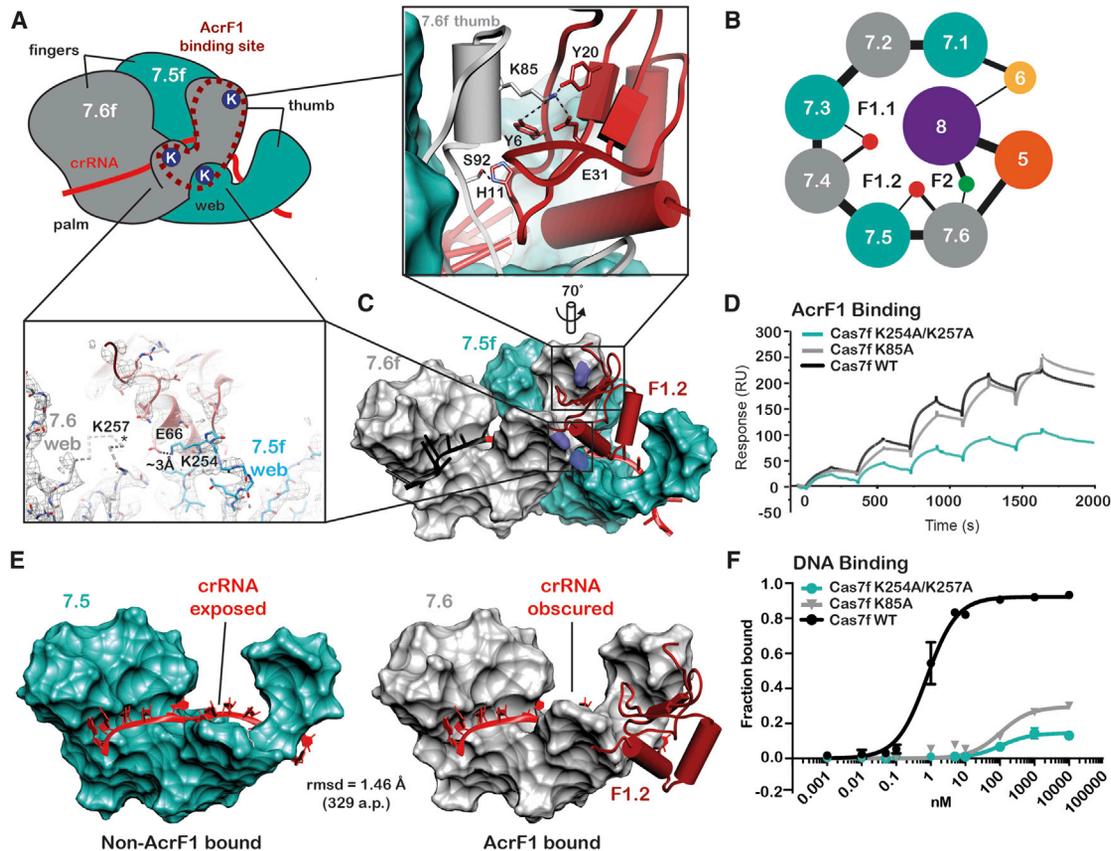


Figure 4. Anti-CRISPR Protein AcrF1 Binds to Residues on Cas7f that Are Essential for crRNA-Guided Recognition of Target DNA

(A) Cartoon showing the location of three lysine residues (blue circles containing the letter K) on two adjacent Cas7f proteins that form a binding site for AcrF1. (B) Schematic of the Csy complex and the Acr proteins illustrating interactions between Acr and Cas proteins. The diameter of lines connecting subunits scales with buried surface.

(C) Structure of two adjacent Cas7f proteins shown as gray (Cas7.6f) and cyan (Cas7.5f) surfaces. The three lysine residues (K85, K254, and K257) highlighted in (A) are shown in blue. AcrF1 is shown in dark red. Above is an expanded view of AcrF1.2 interacting with K85 on the thumb of Cas7.6f. Left is an expanded view of the acidic $\alpha 2$ from AcrF1.2, nestled against the positively charged residues on the extended web domains of Cas7.6f and Cas7.5f, respectively. Side chains involved in specific interactions are labeled. The disordered loop containing K257 is shown as a dotted C- α trace.

(D) Surface plasmon resonance performed with WT and mutant Csy complexes. Mutations in Cas7f (i.e., K85A or K254A and K257A) perturb AcrF1 binding. (E) AcrF1.1 binds to the thumb of Cas7.4f and AcrF1.2 binds to the thumb of Cas7.6f. In these two subunits (gray), but not the other four Cas7f proteins (cyan), the extended web is folded over the crRNA, restricting access to the guide.

(F) Electrophoretic mobility shift assays performed with radiolabeled dsDNA substrates show that Cas7f K85A and K254A and K257A mutations perturb crRNA-guided DNA binding. Error bars, SD; n = 3.

See also Figure S5.

Erp et al., 2015). Notably, no comparable rotation is observed in the sixth Cas7f subunit of the Csy complex (Cas7.6f), suggesting that this complex uses an alternative mechanism for binding dsDNA.

With the exception of a recently described type I-F variant (type I-Fv) (Gleditsch et al., 2016), crRNA-guided surveillance complexes contain a large tail protein called Cas8, but their amino acid sequences are diverse, and Cas8f of *P. aeruginosa* has a distinct fold. Unlike the Cas8 tail in Cascade, which comprises a large globular domain and a C-terminal four-helix bundle, Cas8f has an extended architecture with three discrete domains—an N-terminal “hook” shaped domain (residues

1–166), a central domain (residues 167–264), and a C-terminal helical bundle that occupies the belly of the Csy complex (residues 265–435) (Figure 3B). Although helical bundles have been observed at the C-termini of Cas8 family members, the central location of this helical bundle in *P. aeruginosa* is unique, contacting Cas6, Cas7.2f, and Cas7.3f at the head and upper spiral of the complex, as well as Cas5 at the tail. By contrast, the corresponding helical bundle within the I-E system only contacts Cas5, the crRNA, and the tail-proximal Cse2 “belly” protein. The density for this region of the reconstruction was not as well ordered as the rest of the structure, indicating structural flexibility.

AcrF1 Proteins Prevent DNA Hybridization

AcrF1 binds to the Cas7f backbone and prevents target binding, but the mechanism and location of these interactions has not yet been determined (Bondy-Denomy et al., 2015). The structure presented here explains how two AcrF1 proteins (i.e., AcrF1.1 and 1.2) bind to one Csy complex through interactions with four of the six Cas7f subunits (Figures 1 and 4). AcrF1 is a small protein (78 residues) composed of a four-stranded antiparallel β sheet flanked on one side by two α helices ($\beta_1\beta_2\beta_3\beta_4\alpha_1\alpha_2$) (Figures 1D and 4). The AcrF1 structure was previously determined by NMR, and mutational studies identified three residues (Y6, Y20, and E31) on one face of the β sheet that are critical for AcrF1-mediated repression of type I-F immune response (Maxwell et al., 2016). Here, we show that these three residues interact with a single, conserved lysine (K85) on the Cas7f thumb and that a lysine-to-alanine mutation at this position (K85A) results in a faster dissociation rate from the complex ($k_{d1} = 3.1 \times 10^{-4} \text{ s}^{-1}$) as compared to the dissociation rate of AcrF1 from the wild-type (WT) Csy complex ($k_{d1} = 2.1 \times 10^{-7} \text{ s}^{-1}$) (Figures 4D and S5). In addition to interaction with the thumb, a series of acidic residues on the C-terminal α helix (α_2) of each AcrF1 are wedged between positive charges on the extended web domains of adjacent Cas7f molecules (Figure 4C). We predicted that these interactions might be important for AcrF1 binding. To test this, we mutated lysine 254 and 257 to alanines (i.e., Cas7f K254A/K257A), which resulted in a disruption of AcrF1 binding to the Csy complex (Figures 4D and S5).

Access to the crRNA guide is crucial for target DNA hybridization and subsequent degradation. However, the thumb of each Cas7 protein folds over the crRNA at regular 6-nucleotide intervals, and in Cascade (type I-E), the thumb prevents base pairing at these positions (Jackson et al., 2014a; Mulepati et al., 2014). In the structure of Cascade bound to a single-stranded DNA, the target hybridizes to 5-nucleotide segments of the guide, while the 6th nucleobase is displaced, and the phosphate backbone of the target traverses over the thumb and into the next accessible segment. The two AcrF1 proteins sit on top of the Cas7.6f and Cas7.4f thumbs, respectively, and the binding sites superimpose with an RMSD of 0.1 Å. The AcrF1 proteins block the transition of target DNA from segment 1 to 2, and segment 3 to 4, respectively. Furthermore, the extended web of Cas7.6f and Cas7.4f is ~ 6 Å and ~ 4 Å (respectively) closer to the crRNA-guide than the web of the other subunits, suggesting that AcrF1 binding may induce a conformational change that restricts access to the guide RNA (Figure 4E), though a structure of Csy complex without the anti-CRISPRs will be necessary to know that the AcrF1 protein is directly responsible for this observation.

Since efficient AcrF1 binding relies on interactions with a few key residues on Cas7f (Figure 4D), we expected these residues to be under strong selective pressures. However, using a combination of evolutionary models, we were unable to find evidence for positive selection at these positions (Pond et al., 2005). This suggests that K85, K254, and K257 in Cas7f are important for Csy function, which restricts mutations at these positions. Based on the location and charge of these residues, we hypothesized that they would be involved in DNA binding. To test this hypothesis, we recombinantly expressed and purified the WT Csy complex and complexes containing a mutation in Cas7f at position 85

(K85A) or a double mutation at positions 254 and 257 (i.e., Cas7f K254A and K257A). The mutants express and purify like WT, suggesting that these mutations do not perturb assembly of the complex (Figure S5). Next, we used the purified complexes to perform electrophoretic mobility shift assays (EMSAs) to determine if these mutations perturb DNA binding (Figures 4F and S5). Wild-type Csy complex binds to a 72-base-pair double-stranded DNA target containing a protospacer and PAM with high affinity ($K_D = 1.1 \text{ nM}$), while the mutants result in a severe binding defect ($K_D > 1000 \text{ nM}$), indicating that the lysine residues targeted by AcrF1 are also critical for DNA binding.

AcrF2 Is a DNA Mimic

AcrF2 is a small acidic protein wedged between positively charged residues in the N-terminal hook of Cas8 and the thumb of Cas7.6f (Figures 5A and 5B). Structure homology searches performed using DALI and VAST indicated that AcrF2 adopts a unique fold (Holm and Sander, 1993; Gibrat et al., 1996). The AcrF2 fold is composed of a pair of antiparallel α helices packed on either side of an antiparallel β sheet ($\alpha_1\alpha_2\beta_1\beta_2\beta_3\beta_4\alpha_3\alpha_4$). Interestingly, this is similar to the fold of AcrF1 ($\beta_1\beta_2\beta_3\beta_4\alpha_3\alpha_4$), with the addition of two amino-terminal antiparallel helices, raising the possibility that these two Acr proteins may have evolved from a common ancestor. Acidic residues on AcrF2 (i.e., D30, E77, D76, E94, and E91) are positioned in close proximity to numerous lysines on either Cas7.6f (i.e., K79 and K77) or Cas8f (i.e., K247, K28, and K31). While the density in this region is not sufficient to confidently model the side-chain positions of these residues, we noticed a pseudo-helical display of acidic residues on the surface of AcrF2 mimics the negative charge distribution on the helical backbone of a DNA duplex (Figure 5C) and that distribution of positively charged residues on the N-terminal hook of Cas8f and the thumb of Cas7f.6 form a “lysine-rich, vise-like” structure that has been shown to be crucial for DNA binding by Cascade (van Erp et al., 2015). While the lysine-rich vise in Cascade is composed of positively charged secondary structures on the finger domains of Cas7.5e and Cas7.6e, our structure suggests that the Csy complex may have evolved a functionally analogous DNA vise composed of structurally distinct features. To test the functional role of the positively charged residues on the putative lysine-rich vise, we made charge-swapped mutations on Cas7f (Cas7f K77E and K79E) and Cas8f (Cas8f K247E, K28E, and K31E). These mutants express and purify like WT Csy complex, suggesting that they do not perturb assembly of the Csy complex (Figures S5). However, mutations in either Cas7f or Cas8f result in severe binding defects for AcrF2 and DNA targets (Figures 5D, 5E, and S5). Taken together, these results suggest that AcrF2 is a double-stranded DNA mimic that blocks target recognition by competing for a critical DNA binding site (i.e., lysine-rich vise) and that the lysine-rich vise is structurally plastic but functionally conserved between the I-E and I-F crRNA-guided surveillance systems.

DISCUSSION

Antagonistic interactions between predators (e.g., viruses) and their prey (e.g., bacteria) create dynamic selective pressures that drive diversification and genetic innovation (Van Valen,

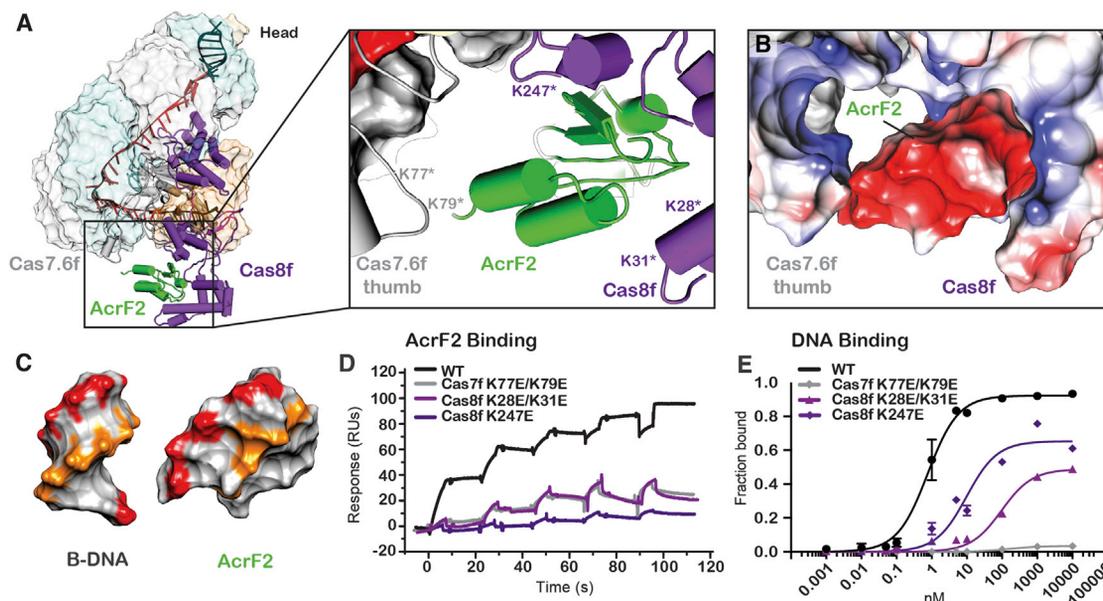


Figure 5. Anti-CRISPR Protein AcrF2 Binds to a Lysine-Rich Vise in the Tail of Csy that Is Essential for DNA Recognition

(A) Location of AcrF2 (green) in the complex with an expanded view of the AcrF2 binding site. Lysines on either Cas7.6f (i.e., K79 and K77) or Cas8f (i.e., K247, K28, and K31) that project into the AcrF2 binding site were mutated to glutamic acids.

(B) Electrostatic representation of AcrF2 in the lysine-rich vise formed by Cas7.6f and Cas8f.

(C) Surface representations of B-form dsDNA and AcrF2. Similarly positioned negative charges shown in red and orange.

(D) Surface plasmon resonance performed with WT and mutant Csy complexes. Mutations in Cas7f (i.e., Cas7f K77E and K79E) or Cas8f (Cas8f K247E and Cas8f K28E and K31E) perturb AcrF2 binding.

(E) Electrophoretic mobility shift assays were performed with radiolabeled dsDNA substrates. Mutations in the lysine-rich vise (Cas7 K79E and K77E, Cas8 K247E, or Cas8 K28E and K31E) perturb binding to dsDNA targets. Error bars, SD; n = 3.

1973). The remarkable diversity of CRISPR-Cas systems (two classes, six types, and 19 subtypes) is consistent with their role in defense against rapidly evolving predators, and the discovery of diverse anti-CRISPRs with distributions that mirror these immune systems is beginning to explain the molecular basis of genetic conflict at the CRISPR-anti-CRISPR interface (Pawluk et al., 2016; Carter et al., 2017).

To determine the mechanistic differences that distinguish type I-F systems from other CRISPR-mediated immune systems, and to understand the molecular mechanisms of AcrF1- and AcrF2-mediated suppression, we determined the cryo-EM structure of the Csy surveillance complex bound to both AcrF1 and AcrF2. The seahorse-shaped morphology of the complex is similar to what has been observed in recent structures of other class 1 surveillance complexes (i.e., type I and III); however, significant structural differences in individual subunits manifest in higher-order morphological distinctions (Figure 2). Like other class 1 surveillance complexes, the backbone of the Csy complex is composed of an assembly of Cas7-family proteins interwoven around the crRNA. While the theme of Cas7 oligomerization along the crRNA is conserved, the Cas7f fold is distinct, and Cas7f-induced kinks in the crRNA result in a tight helical pitch that creates a nearly closed ring-like structure. The appearance of a closed ring is accentuated by an elongated Cas8-family tail protein that extends toward the Cas6f head, and the C-terminal helical domain of Cas8f is wedged in the middle of the Cas7f ring (Figure 1).

While the fold of Cas7f is distinct from previously determined Cas7-family proteins, the familiar right-hand morphology is largely maintained, with one prominent distinction: the web domain in Cas7f extends along the thumb, creating a conspicuous trough between the web and the thumb domains (Figure 4A, 4C, and 4E). We anticipate that this positively charged trough functions as an electrostatic cradle for the target strand as it traverses over each thumb from one complementary segment of the crRNA-guide to the next. The structure explains how two molecules of AcrF1 interact with conserved lysine residues that line this trough and block target binding by obstructing contiguous access of the target to adjacent segments of the crRNA-guide sequences (Figure 4). In addition, the web domains of Cas7.4f and Cas7.6f are closer to the crRNA, suggesting that AcrF1 binding may induce local conformational changes that further restrict target access. However, detection of lysines K85, K254, and K257 by AcrF1 suggests that natural variation at these positions would prevent AcrF1 binding and that these mutants would escape AcrF1-mediated suppression of the immune system. To test this hypothesis, we mutated these lysines and show that they inhibit AcrF1 binding, but these mutations also result in severe DNA-binding defects. Together, these results suggest that AcrF1 has evolved to target specific features of the Csy complex that are essential for DNA binding.

In addition to binding the crRNA, structures of Cascade (type I-E) have shown that a subset of the Cas7e proteins also play a critical role in binding DNA (Hayes et al., 2016; Hochstrasser

et al., 2014; Jackson et al., 2014a; van Erp et al., 2015; Wiedenheft et al., 2011). The fingers domain of Cas7.6e rotates 180°, creating a lysine-rich vise that makes a non-sequence specific interaction with double-stranded DNA, and mutations in Cas7e proteins that eliminate the positive charge no longer bind DNA (van Erp et al., 2015). We anticipated that this would be a conserved feature of the Cas7 backbone in type I surveillance systems; however, no such rotation is observed in the Cas7.6f subunit of the Csy complex. Instead, the structure reveals an analogous lysine-rich, vise-like feature created by lysines displayed on the thumb of Cas7.6f and the N-terminal hook-domain on Cas8f. The importance of these lysines may have escaped our attention if it weren't for the negative charge and positioning of AcrF2 directly in the vise. Interestingly, acidic residues on the surface of AcrF2 mimic the negative charge distribution on the helical backbone of a DNA duplex (Figure 5C), suggesting that AcrF2 is a double-stranded DNA mimic that competes with DNA for binding to the lysine-rich vise. Indeed, mutations in the lysine-rich vise that inhibit AcrF2 binding also prevent DNA binding by the Csy complex, which may explain why type I systems rely on distinct structural innovations to achieve similar functions.

Phylogenetic and structural studies have shown that Cas8 proteins are extremely diverse, which, in light of the results presented here, may be in part due to selective pressures from anti-CRISPR proteins. However, all Cas8 proteins appear to be involved in tail assembly through interactions with both Cas5 and the 5' end of the crRNA (Hayes et al., 2016; Hochstrasser et al., 2016; Jackson et al., 2014a; Jackson and Wiedenheft, 2015; Mulepati et al., 2014; Osawa et al., 2015; Taylor et al., 2015; Zhang and Sontheimer, 2014; Zhao et al., 2014). In *E. coli*, the thumb on Cas5e creates a cylindrical pore that serves as a docking module for a short helix on Cas8e (also called Cse1). This helix is within a loop called loop 1 (L1) that inserts into the Cas5e helix-binding pore and makes base-specific interactions with nucleotides in the 5' handle (Hayes et al., 2016; Jackson et al., 2014a; Mulepati et al., 2014; Zhao et al., 2014). While the structural details are different, similar themes are observed in the Csy complex, wherein the thumb and finger domains of Cas5f protein grip a β hairpin (residues 175–191, called the harpoon) of Cas8f, which makes both specific (e.g., Q177 interacts with adenosine at position –5) and non-sequence-specific interaction with nucleotides in the 5' handle. In addition to the role of Cas8 in tail assembly, a recent structure of Cascade (type I-E) bound to a partial duplexed DNA target explains how the N-terminal globular domain of Cas8e recognizes the PAM sequence and how PAM recognition is coupled to strand separation via insertion of a β hairpin called the glutamine wedge (Hayes et al., 2016). Initial duplex destabilization by the wedge may be coupled to the C-terminal domain of Cas8e, which contains a four-helix bundle that appears to function like a molecular pry bar to maintain separation of the two strands. The displaced strand in this structure traverses over the four-helix bundle, whereupon it is presumably presented to the *trans*-acting Cas3 nuclease-helicase, though the remainder of the displaced strand was not included in the DNA target used in this structural study. Although Cas8f does not share significant sequence similarity with Cas8e, Cas8f also contains a C-terminal four-helix bundle, and the central position and orientation of this structure suggests

that it may also be involved in strand separation during double-stranded DNA binding.

Another conserved feature of all type I systems is that after DNA binding by the crRNA-guided surveillance complex, they all recruit Cas3, a *trans*-acting nuclease-helicase (Jackson et al., 2014b). In type I-F systems, the Cas3 nuclease-helicase is fused to an N-terminal Cas2 protein that is involved in new sequence integration, and an anti-CRISPR protein (i.e., AcrF3) that binds to Cas2/3 prevents recruitment to the Csy complex (Bondy-Denomy et al., 2015; Wang et al., 2016a; 2016b). Using Dali (Holm and Sander, 1993), a structure homology search program, we identified structural similarity between AcrF3 and the C-terminal four-helix bundle of Cas8f (Figure S4). Cas8 proteins in other systems have been shown to be involved in Cas3 recruitment (Hochstrasser et al., 2014; Huo et al., 2014; Xue et al., 2016), and we speculate that AcrF3 may be a molecular mimic of the domain on Cas8f that is involved in Cas2/3 recruitment.

Presumably, CRISPRs evolved in response to antagonistic interactions with molecular parasites such as phages, and phages evolved anti-CRISPRs to subvert these systems. Here, we provide structural and biochemical evidence showing that anti-CRISPR proteins AcrF1 and AcrF2 evolved to target specific structural features that are essential for immune system function, which suggests these antagonists play a central role in driving the diversification of CRISPR-Cas immune systems.

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
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
 - Microbes
- METHOD DETAILS
 - Expression and purification
- ELECTRON MICROSCOPY
 - Electron microscopy sample preparation
 - Electron microscopy data acquisition
 - Electron microscopy data processing
 - Atomic model building and refinement
 - Mutagenesis
 - Electrophoretic mobility shift assay
 - Surface plasmon resonance
- QUANTIFICATION AND STATISTICAL ANALYSIS
- DATA AND SOFTWARE AVAILABILITY

SUPPLEMENTAL INFORMATION

Supplemental Information includes five figures, one table, and one movie and can be found with this article online at <http://dx.doi.org/10.1016/j.cell.2017.03.012>.

AUTHOR CONTRIBUTIONS

M.F.R. purified the Csy complex, Acr proteins, and the Csy-Acr complex. E.R.F. performed initial electron microscopy screening of the complex. L.N. performed the negative-stain electron microscopy and associated image analysis and reconstruction. S.C. performed the cryo electron microscopy, image

analysis, and reconstructions. S.C., R.N.J., and J.C. built the atomic model. M.F.R. and S.M.G. performed the biochemical experiments. All authors contributed to analysis and interpretation of the experimental results and data. J.C., S.C., G.C.L., and B.W. prepared the figures and S.C., J.C., G.C.L., and B.W. wrote the manuscript.

ACKNOWLEDGMENTS

We thank Jeff Speir and Anchi Cheng, who aided with early EM data collection and analyses at the National Resource for Automated Molecular Microscopy, which is supported by National Institutes of Health grant (GM103310). We thank Mark A. Herzik, Jr. for advice in model building and C. Martin Lawrence for pointing out the distinct fold of Cas7f. We are grateful to Bill Anderson, TSRI Electron Microscopy facility manager, and Jean-Christophe Ducom at TSRI High Performance Computing for support during EM data collection and processing. Research in the Wiedenheft lab is supported by the National Institutes of Health (P20GM103500, P30GM110732-03, R01GM110270, and R01GM108888), the National Science Foundation EPSCoR (EPS-110134), the M. J. Murdock Charitable Trust, a young investigator award from Amgen, and the Montana State University Agricultural Experimental Station. J.C. is supported by the National Institutes of Health (P20GM103474) through Montana INBRE. R.N.J. is a recipient of the Ruth L. Kirschstein National Research Service Award from the National Institutes of Health (F32 GM108436). G.C.L. is supported as a Searle Scholar, as a Pew Scholar, and by the National Institutes of Health (DP2EB020402). J.B.-D. is supported by the University of California San Francisco Program for Breakthrough in Biomedical Research, funded in part by the Sandler Foundation, and an NIH Office of the Director Early Independence Award (DP5-OD021344). A.R.D. and K.L.M. are supported by the Canadian Institutes of Health Research (MOP-130482 and MOP-136845, respectively). E.R.F. is supported by the NIH intramural research program.

Received: January 12, 2017

Revised: February 23, 2017

Accepted: March 6, 2017

Published: March 23, 2017

REFERENCES

- Adams, P.D., Afonine, P.V., Bunkoczi, 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.
- Bai, X.C., Rajendra, E., Yang, G., Shi, Y., and Scheres, S.H. (2015). Sampling the conformational space of the catalytic subunit of human gamma-secretase. *Elife* **4**, e11182.
- 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.
- Carter, J., Hoffman, C., and Wiedenheft, B. (2017). The interfaces of genetic conflict are hot spots for innovation. *Cell* **168**, 9–11.
- DiMaio, F., Song, Y., Li, X., Brunner, M.J., Xu, C., Conticello, V., Egelman, E., Marlovits, T.C., Cheng, Y., and Baker, D. (2015). Atomic-accuracy models from 4.5-Å cryo-electron microscopy data with density-guided iterative local refinement. *Nat Methods* **12**, 361–365.
- Emsley, P., and Cowtan, K. (2004). Coot: model-building tools for molecular graphics. *Acta Crystallogr. D. Biol. Crystallogr* **60**, 2126–2132.
- Gibrat, J.F., Madej, T., and Bryant, S.H. (1996). Surprising similarities in structure comparison. *Curr. Opin. Struct. Biol.* **6**, 377–385.
- Gleditsch, D., Müller-Esparza, H., Pausch, P., Sharma, K., Dwarakanath, S., Urlaub, H., Bange, G., and Randau, L. (2016). Modulating the Cascade architecture of a minimal Type I-F CRISPR-Cas system. *Nucleic Acids Res.* **44**, 5872–5882.
- Goddard, T.D., Huang, C.C., and Ferrin, T.E. (2005). Software extensions to UCSF chimera for interactive visualization of large molecular assemblies. *Structure* **13**, 473–482.
- Goddard, T.D., Huang, C.C., and Ferrin, T.E. (2007). Visualizing density maps with UCSF Chimera. *J. Struct. Biol.* **157**, 281–287.
- Haurwitz, R.E., Jinek, M., Wiedenheft, B., Zhou, K., and Doudna, J.A. (2010). Sequence- and structure-specific RNA processing by a CRISPR endonuclease. *Science* **329**, 1355–1358.
- 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.
- Heymann, J.B., and Belnap, D.M. (2007). Bsoft: image processing and molecular modeling for electron microscopy. *J. Struct. Biol.* **157**, 3–18.
- 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.
- Holm, L., and Sander, C. (1993). Protein structure comparison by alignment of distance matrices. *J. Mol. Biol.* **233**, 123–138.
- Huo, Y., Nam, K.H., Ding, F., Lee, H., Wu, L., Xiao, Y., Farchione, M.D., Jr., Zhou, S., Rajashankar, K., Kurinov, I., et al. (2014). Structures of CRISPR Cas3 offer mechanistic insights into Cascade-activated DNA unwinding and degradation. *Nat. Struct. Mol. Biol.* **21**, 771–777.
- 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. (2014a). Structural biology. Crystal structure of the CRISPR RNA-guided surveillance complex from *Escherichia coli*. *Science* **345**, 1473–1479.
- Jackson, R.N., Lavin, M., Carter, J., and Wiedenheft, B. (2014b). Fitting CRISPR-associated Cas3 into the helicase family tree. *Curr. Opin. Struct. Biol.* **24**, 106–114.
- Kuznedelov, K., Mekler, V., Lemak, S., Tokmina-Lukaszewska, M., Datsenko, K.A., Jain, I., Savitskaya, E., Mallon, J., Shmakov, S., Bothner, B., et al. (2016). Altered stoichiometry *Escherichia coli* Cascade complexes with shortened CRISPR RNA spacers are capable of interference and primed adaptation. *Nucleic Acids Res.* **44**, 10849–10861.
- 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. *J. Struct. Biol.* **166**, 95–102.
- Li, X., Mooney, P., Zheng, S., Booth, C.R., Braunfeld, 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.
- Luo, M.L., Jackson, R.N., Denny, S.R., Tokmina-Lukaszewska, M., Maksimchuk, K.R., Lin, W., Bothner, B., Wiedenheft, B., and Beisel, C.L. (2016). The CRISPR RNA-guided surveillance complex in *Escherichia coli* accommodates extended RNA spacers. *Nucleic Acids Res.* **44**, 7385–7394.
- Makarova, K.S., Haft, D.H., Barrangou, R., Brouns, S.J., Charpentier, E., Horvath, P., Moineau, S., Mojica, F.J., Wolf, Y.I., Yakunin, A.F., et al. (2011). Evolution and classification of the CRISPR-Cas systems. *Nat. Rev. Microbiol.* **9**, 467–477.

- 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.
- 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.
- Mindell, J.A., and Grigorieff, N. (2003). Accurate determination of local defocus and specimen tilt in electron microscopy. *J. Struct. Biol.* **142**, 334–347.
- 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.
- Ogura, T., Iwasaki, K., and Sato, C. (2003). Topology representing network enables highly accurate classification of protein images taken by cryo electron-microscope without masking. *J. Struct. Biol.* **143**, 185–200.
- 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.
- Pawluk, A., Bondy-Denomy, J., Cheung, V.H.W., Maxwell, K.L., and Davidson, A.R. (2014). A new group of phage anti-CRISPR genes inhibits the type I-E CRISPR-Cas system of *Pseudomonas aeruginosa*. *MBio* **5**, e00896.
- 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.
- Pond, S.L.K., Frost, S.D.W., and Muse, S.V. (2005). HyPhy: hypothesis testing using phylogenies. *Bioinformatics* **21**, 676–679.
- Rodriguez-Valera, F., Martin-Cuadrado, A.B., Rodriguez-Brito, B., Pasić, L., Thingstad, T.F., Rohwer, F., and Mira, A. (2009). Explaining microbial population genomics through phage predation. *Nat. Rev. Microbiol.* **7**, 828–836.
- 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.
- Roseman, A.M. (2004). FindEM—a fast, efficient program for automatic selection of particles from electron micrographs. *J. Struct. Biol.* **145**, 91–99.
- Rouillon, C., Zhou, M., Zhang, J., Politis, A., Beilsten-Edmands, V., Cannone, G., Graham, S., Robinson, C.V., Spagnolo, L., and White, M.F. (2013). Structure of the CRISPR interference complex CSM reveals key similarities with cascade. *Mol. Cell* **52**, 124–134.
- Scheres, S.H. (2012). RELION: implementation of a Bayesian approach to cryo-EM structure determination. *J. Struct. Biol.* **180**, 519–530.
- Scheres, S.H. (2014). Beam-induced motion correction for sub-megadalton cryo-EM particles. *Elife* **3**, e03665.
- Shmakov, S., Abudayyeh, O.O., Makarova, K.S., Wolf, Y.I., Gootenberg, J.S., Semenova, E., Minakhin, L., Joung, J., Konermann, S., Severinov, K., et al. (2015). Discovery and Functional Characterization of Diverse Class 2 CRISPR-Cas Systems. *Mol. Cell* **60**, 385–397.
- Sorzano, C.O., Marabini, R., Velazquez-Muriel, J., Bilbao-Castro, J.R., Scheres, S.H., Carazo, J.M., and Pascual-Montano, A. (2004). XMIPP: a new generation of an open-source image processing package for electron microscopy. *J. Struct. Biol.* **148**, 194–204.
- Spilman, M., Ccozaki, A., Hale, C., Shao, Y., Ramia, N., Terns, R., Terns, M., Li, H., and Stagg, S. (2013). Structure of an RNA silencing complex of the CRISPR-Cas immune system. *Mol. Cell* **52**, 146–152.
- 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 Legimon system. *J. Struct. Biol.* **151**, 41–60.
- Suttle, C.A. (2007). Marine viruses—major players in the global ecosystem. *Nat. Rev. Microbiol.* **5**, 801–812.
- Taylor, D.W., Zhu, Y., Staals, R.H., Kornfeld, J.E., Shinkai, A., van der Oost, J., Nogales, E., and Doudna, J.A. (2015). Structural biology. Structures of the CRISPR-Cmr complex reveal mode of RNA target positioning. *Science* **348**, 581–585.
- Topuzlu, E., and Lawrence, C.M. (2016). Recognition of a pseudo-symmetric RNA tetranucleotide by Csx3, a new member of the CRISPR associated Ross-mann fold superfamily. *RNA Biol.* **13**, 254–257.
- Trincao, J., Johnson, R.E., Escalante, C.R., Prakash, S., Prakash, L., and Agarwal, A.K. (2001). Structure of the catalytic core of *S. cerevisiae* DNA polymerase eta: implications for translesion DNA synthesis. *Mol. Cell* **8**, 417–426.
- 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.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.
- Van Valen, L. (1973). A new evolutionary law. *Evol. Theory* **1**, 1–30.
- 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.
- Wang, J., Ma, J., Cheng, Z., Meng, X., You, L., Wang, M., Zhang, X., and Wang, Y. (2016a). A CRISPR evolutionary arms race: structural insights into viral anti-CRISPR/Cas responses. *Cell Res.* **26**, 1165–1168.
- Wang, X., Yao, D., Xu, J.G., Li, A.R., Xu, J., Fu, P., Zhou, Y., and Zhu, Y. (2016b). Structural basis of Cas3 inhibition by the bacteriophage protein AcrF3. *Nat. Struct. Mol. Biol.* **23**, 868–870.
- Wiedenheft, B., van Duijn, E., Bultema, J.B., Waghmare, S.P., Zhou, K., Barendregt, A., Westphal, W., Heck, A.J., Boekema, E.J., Dickman, M.J., and Doudna, J.A. (2011). RNA-guided complex from a bacterial immune system enhances target recognition through seed sequence interactions. *Proc. Natl. Acad. Sci. USA* **108**, 10092–10097.
- Wiedenheft, B., Lander, G.C., Zhou, K., Jore, M.M., Brouns, S.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.
- Xue, C., Whittis, N.R., and Sashital, D.G. (2016). Conformational Control of Cascade Interference and Priming Activities in CRISPR Immunity. *Mol. Cell* **64**, 826–834.
- Zhang, Y., and Sontheimer, E.J. (2014). Structural biology. Cascading into focus. *Science* **345**, 1452–1453.
- 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.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT OR RESOURCE	SOURCE	IDENTIFIER
Bacterial and virus strains		
<i>E. coli</i> : BI21 DE3 competent cells	NEB	Cat# C25271
<i>E. coli</i> : BI21 DH5 α competent cells	ThermoFisher	Cat# 18265017
Chemicals, Peptides, and Recombinant Proteins		
TCEP	Soltec	Cat# M115
Protease inhibitor cocktail	Thermo Scientific	Cat# 1861278
T4 DNA ligase	NEB	Cat# M0202S
PNK	NEB	Cat# M0201s
DPN I	NEB	Cat# R0176s
³² P-ATP	Perkin Elmer	Cat# 100539131
poly-L-lysine hydrobromide	Polysciences	Cat# 09730
Deposited Data		
Csy complex bound to AcrF1 and AcrF2	This paper	PDB: 5ZU9
Csy complex bound to AcrF1 and AcrF2	This paper	EMDB: 8624
pCsy_complex expression vector	Wiedenheft lab	Addgene plasmid # 89232
pAcrF1 expression vector	Bondy-Denomy et al., 2015	Addgene plasmid # 89233
pAcrF2 expression vector	Bondy-Denomy et al., 2015	Addgene plasmid # 89234
pCRISPR_DMS3g24 expression vector	Wiedenheft lab	Addgene plasmid # 89244
Oligonucleotides		
Primer: Cas7_K85A.F: gaccgtgaccggccGCGctggacgcctcgatc	Eurofins	n/a
Primer: Cas7_K85A.R: gatcgaggcgctccagCGCggccgggtcacggtc	Eurofins	n/a
Primer: Cas7_K77E,K79E.F: cGAGgaccgtgaccggccaag	Eurofins	n/a
Primer: Cas7_K77E,K79E.R: gtCTCgaggcggttgagatggtg	Eurofins	n/a
Primer: Cas7_K254A,K257A.F: gacGCGaaaggccagaagagcaag	Eurofins	n/a
Primer: Cas7_K254A,K257A.R: gccTGCgtcgaggatcagttcctg	Eurofins	n/a
Primer: Cas8_K28E,K31E.F: gcctccagggcGAGctggacGAGctccaaccgac	Eurofins	n/a
Primer: Cas8_K28E,K31E.R: gtcgggttgagCTCtccagCTCgccctggaggcg	Eurofins	n/a
Primer: Cas8_K247E.F: agttcggcggtaccGAGccgcagaacatc	Eurofins	n/a
Primer: Cas8_K247E.R: gatgtctgctgCTCggtaccgccgaact	Eurofins	n/a
Oligonucleotides for gel shift, Target (blue is protospacer, red is PAM) GCTGTACGTCACTATCGAAGCAATACAGGTAGACGCGGACATCAAGC CCGCCGTGAAGGTGCAGCTTCTTACAGAGTGC	Eurofins	n/a
Oligonucleotides for gel shift, Non-Target CGACATGCAGTGATAGCTTC GTTATGTCCATCTGCGCCTGTAGTTCGGGCGGCACTTCCACGTCAA GAGATGTCTCACG	Eurofins	n/a
Recombinant DNA		
Plasmid: pCsy_complex cas8f K28E/K31E	This study	n/a
Plasmid: pCsy_complex cas8f K247E	This study	n/a
Plasmid: pCsy_complex cas7f K77E/K79E	This study	n/a
Plasmid: pCsy_complex cas7f K85A	This study	n/a
Plasmid: pCsy_complex cas7f K254A/K257E	This study	n/a
Plasmid: pCsy_complex	Wiedenheft lab	Addgene plasmid # 89232
Plasmid: pCRISPR_DMS3g24	Wiedenheft lab	Addgene plasmid # 89244
Plasmid: pAcrF1	Bondy-Denomy et al., 2015	Addgene plasmid # 89233

(Continued on next page)

Continued

REAGENT OR RESOURCE	SOURCE	IDENTIFIER
Plasmid: pAcrF2	Bondy-Denomy et al., 2015	Addgene plasmid # 89234
Software and Algorithms		
ImageQuant software	GE	n/a
Biacore evaluation software	GE	Cat# BR100798
Leginon automated software	FEI	n/a
CTFFind3	Mindell and Grigorieff, 2003	http://grigoriefflab.janelia.org/ctf
Coot	Emsley and Cowtan, 2004	https://www2.mrc-lmb.cam.ac.uk/personal/pemsley/coot/
PHENIX v1.11	Adams et al., 2010	https://www.phenix-online.org/
USCF Chimera	Goddard et al., 2007	https://www.cgl.ucsf.edu/chimera/
EMAN	Ludtke et al., 1999	http://www.msg.ucsf.edu/local/programs/eman/
DoG picker	Voss et al., 2009	http://emg.nysbc.org/redmine/projects/software/wiki/DoGpicker
XMIPP	Sorzano et al., 2004	http://xmipp.cnb.csic.es/twiki/bin/view/Xmipp/WebHome
RELION	Scheres, 2012	https://www2.mrc-lmb.cam.ac.uk/relion/index.php/Main_Page
FindEM	Roseman, 2004	http://www.ccpem.ac.uk/download.php
Other		
Strep trap HP column	GE	Cat# 28-9075-47
Strep-tactin superflow column	QIAGEN	Cat# 30160
Spin concentrators	Corning	Cat# 431491
Superdex 200 column	GE	Cat# 28-9893-95
Ni NTA superflow column	QIAGEN	Cat# 30760
Superdex 75 column	GE	Cat# 28-9893-33
G-25 spin columns	GE	Cat# 27-5325-01
Ni-NTA sensor chip	GE	Cat# BR1000-12
CM5 sensor chip	GE	Cat# BR1000-34
400mesh Cu-Rh maxtaform grids	Electron Microscopy Sciences	Cat# M400-CR
c-flat grids (2 μ m holes and spacing)	Protochips	Cat# CF-2/2-2C

CONTACT FOR REAGENT AND RESOURCE SHARING

Please direct any requests for further information or reagents to the Lead Contact Blake Wiedenheft (bwiedenheft@gmail.com).

EXPERIMENTAL MODEL AND SUBJECT DETAILS**Microbes**

Escherichia coli cells were cultured on LB medium.

METHOD DETAILS**Expression and purification**

The *cas* genes and a synthetic CRISPR were co-expressed on separate vectors in *E. coli* BL21 (DE3) cells as previously described (Rollins et al., 2015; Wiedenheft et al., 2011). The expression vectors have been deposited at Addgene (see [Key Resources Table](#)). In brief, expression was induced with 0.5 mM isopropyl- β -D-1-thiogalactopyranoside (IPTG) when the culture reached an optical density of 0.5 (OD_{600 nm}). Cells were incubated overnight at 16°C, then pelleted by centrifugation (5000 x g for 15 min at 4°C) and re-suspended in lysis buffer (50 mM 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES) pH 7.5, 300 mM potassium chloride, 5%

glycerol, 1 mM Tris(2-carboxyethyl) phosphine hydrochloride (TCEP), 1x protease inhibitor cocktail (Thermo Scientific)). Cell pellets were sonicated on ice for 3 × 2.5 min (1 s on, 3 s off), lysate was clarified by centrifugation at 22,000 × g for 30 min at 4°C. The Cas7f subunit includes an N-terminal strep-II tag and the Csy complex self-assembles in vivo. The intact complex was affinity purified using StrepTrap HP resin (GE), eluted with lysis buffer supplemented with 2.5 mM desthiobiotin, and then concentrated (Corning Spin-X concentrators) at 4°C before further purification over a Superdex 200 size-exclusion column (GE Healthcare) in 20 mM HEPES pH 7.5, 100 mM KCl, 5% glycerol, 1 mM TCEP.

Gene 35 from phage JBD30 (AcrF1) and gene 30 from phage D3112 (AcrF2) were cloned in a p15TV-L vector with N-terminal His₆ tags and overexpressed in *E. coli* BL21 DE3 cells as previously described (Bondy-Denomy et al., 2015). Briefly, *E. coli* BL21 DE3 cells were grown to an OD_{600 nm} of 0.5 and then induced with IPTG for 3 hr at 37°C. Cells were collected by centrifugation at 5,000 g for 10 min and suspended in a lysis buffer containing 50 mM Tris, pH 7.5, 300 mM NaCl, 5% glycerol, and 0.5x protease inhibitor cocktail (Thermo Scientific). The cells were lysed by sonication as described above and the lysate was centrifuged at 15,000 g for 15 min to remove cell debris. The supernatant was injected into a Ni-NTA column, washed in with 20 mL of the lysis buffer supplemented with 20 mM imidazole and eluted from the column using a linear gradient from 20 mM to 300 mM imidazole in the wash buffer. Fractions were collected, and concentrated (Corning Spin-X concentrators) at 4°C before further purification over a Superdex 75 size-exclusion column (GE Healthcare) in 20 mM Tris, pH 7.5, 100 mM NaCl, 1 mM TCEP.

10- and 30-fold molar excess of the purified Acr proteins F1 and F2 (respectively), were added simultaneously to purified Csy complex. The mixture was incubated on ice for 30 min and free Acr proteins were separated from the Csy bound proteins using a Superdex 200 size-exclusion column (GE Healthcare) in 20 mM Tris pH 7.5, 100 mM KCl, 1 mM TCEP.

ELECTRON MICROSCOPY

Electron microscopy sample preparation

For negative stain electron microscopy (EM), 3 μL of purified (0.05 mg/ml) Csy complex bound to AcrF1 and AcrF2 (Csy-Acr complex) was applied to freshly plasma cleaned 400 mesh Cu-Rh maxtaform grids (Electron Microscopy Sciences) that were coated with a thin layer of amorphous carbon. After incubating for 1 min at room temperature, excess protein was wicked away with a filter paper (Whatman No. 1), the grids were then stained with 2% (w/v) uranyl formate solution, and blotted to dryness.

Initial cryo-EM screening of the purified Csy-Acr complex on C-flat grids with 2 μm holes and 2 μm spacing (Protochips) showed an orientation bias in ice. To overcome this issue, a thin amorphous carbon film was floated onto the holey C-flat grids. After plasma-cleaning these grids, 5 μL of 0.1% (w/v) poly-L-lysine hydrobromide (Polysciences) was deposited on the carbon surface for 90 s, blotted to dryness, followed by three successive washes with 10 μL drops of water at a time, then blotted to dryness. 4 μL of 0.1 mg/mL purified Csy-Acr complex was then applied on the poly-lysine-treated grids. Excess sample was manually blotted with filter paper for ~3–5 s, and the sample was immediately vitrified by plunge freezing in liquid-ethane slurry at –179°C. The entire procedure was carried out at 4°C and 98% relative humidity.

Electron microscopy data acquisition

Negative-stain EM micrographs were collected using the Legikon automated data acquisition software on a Tecnai Spirit (FEI) transmission electron microscope operating at 120 keV (Suloway et al., 2005). Images were collected at a nominal magnification of 52,000x on an F416 CMOS 4K × 4K-pixel camera (TVIPS) with a pixel size of 2.05 Å/pixel at specimen level. A total of 97 negative-stain micrographs were collected with an electron dose of 20 e⁻/Å² using a defocus range of 0.5 μm to 1.5 μm.

Cryo grids were loaded into a Titan Krios transmission electron microscope operating at 300 keV. Micrographs were acquired using the Legikon data collection software (Suloway et al., 2005) at a nominal magnification of 29,000X (1.026 Å/pixel at the specimen level) using a total dose of ~46 e⁻/Å² and a nominal defocus range of 1.2 to 2.5 μm. A total of 2,261 micrographs were acquired using a Gatan K2 Summit direct electron detector, operated in electron counting mode. Each micrograph was acquired as a 30-frame dose-fractionated movie during a 6 s exposure.

Electron microscopy data processing

The Appion image processing pipeline was used for processing of micrographs and initial 2D analyses of both negative stain and cryo-EM data (Lander et al., 2009).

Negative stain data

CTFFind3 was used to determine the contrast transfer function (CTF) of each negatively stained micrograph, and particles were selected from micrographs using a Difference of Gaussians (DoG)-based automated particle picker (Mindell and Grigorieff, 2003; Voss et al., 2009). Phases for each micrograph were corrected using the EMAN software package (Ludtke et al., 1999), and 14,542 particles were extracted using a 160 × 160 pixel box. Individual particles were normalized by eliminating pixels with values above or below 4.5 σ of the mean pixel value using the normalization function in the XMIPP package (Sorzano et al., 2004) and the particles were then binned by a factor of four to speed up computation. The extracted particle dataset was subjected to five rounds of iterative multivariate statistical analysis (MSA) (Ogura et al., 2003), and multi-reference alignment (MRA) in Appion to remove any erroneously picked non-particle features and aggregates. This resulted in a final stack of 13,476 particles, which was used for 3D processing using RELION (Scheres, 2012). A cylindrical disc having roughly the same dimensions as the Csy-Acr complex was used as

starting volume for 3D processing. Multiple iterations of 3D classification and auto-refinement yielded a final 3D reconstruction (Figures S1A–S1C) with a resolution of ~ 25 Å (at 0.143 gold-standard Fourier Shell Correlation (FSC)).

Cryo-EM data

The individual frames of each cryo micrograph were translationally aligned to account for beam-induced motion and drift using the GPU frame alignment program MotionCorr (Li et al., 2013), implemented in the Appion workflow. A frame offset of 7 and a B-factor of 1000 pixels was used for frame alignment. CTF parameters were estimated with CTFFind3 (Mindell and Grigorieff, 2003) (Figures S1D–S1F), and images reporting a confidence value of less than 90% accuracy were discarded. Initially DoGPicker was used for automated particle picking from 200 micrographs, and a stack of $\sim 50,000$ particles was extracted using a box size of 320 pixels. These particles were then subjected to reference-free alignment and classification using MSA/MRA after binning to a pixel size of 2.05 Å/pixel. Selected 2D class averages were then used as templates for reference-based automated particle picking from all the micrographs using the program FindEM (Roseman, 2004), and a new stack of 199,348 particles was extracted with a 320-pixel box size. This stack was binned to a pixel size of 2.05 Å/pixel and subjected to MSA/MRA 2D classification, as described before. After discarding particles belonging to 2D classes depicting non-particle features or aggregates, a stack of 145,874 particles was created. Coordinates of these particles were imported into Relion1.4 (Scheres, 2014), and particles were re-extracted with a pixel size of 2.05 Å/pixel and box size of 160 pixels. These particles were then subjected to 25 iterations of reference-free 2D classification within Relion. After discarding particles belonging to 2D class averages lacking high-resolution structural features, a new stack of 68,548 particles was created (Figure S3). This particle stack was then subjected to 3D auto-refinement within RELION, using a 40 Å low-pass filtered negative stain reconstruction of the Csy-Acr complex as a starting volume. After this 3D refinement, the particle stack was subjected to 3D classification, requesting 3 classes and without performing angular or translational searches, in order to sort out structural heterogeneity. After 25 iterations of 3D classification, 57,647 particles belonging to the class that represented the complete Csy-Acr complex and showed the highest levels of structural details, were re-centered and re-extracted from the micrographs with a pixel size of 1.026 Å/pixel and box size of 320 pixels. These particles were then subjected to further 3D refinement runs that resulted in a reconstruction with an overall resolution of 3.9 Å (by gold-standard FSC at 0.143). In order to improve the resolution of the reconstruction, we performed the particle polishing procedure in RELION 1.4 to correct for individual particle movement and radiation damage (Scheres, 2014). A new round of 3D refinement with the polished particle stack resulted in a 3.5 Å (at an FSC of 0.143) reconstruction.

Density corresponding to the Cas6 head and Cas8 tail was poorly resolved in the 3D reconstruction, due to intrinsic flexibility. A soft-edged binary mask was created by applying an 8-pixel extension and a 10-pixel cosine edge fall-off to the 3D reconstruction, with density corresponding to the flexible head and tail portions removed. The polished particle stack after 3D refinement was subjected to 3D classification without angular or translational search, with this binary mask applied. This eliminated particles that lacked high-resolution structural information for the stable core region of the complex, leading to a new stack of 51,212 particles. This particle stack was then subjected to 3D auto-refinement in RELION to generate the final 3D map of the Csy-Acr complex (Figure S3). The overall resolution of this map estimated using a gold-standard FSC at a cutoff of 0.143 is 3.4 Å, and a B-factor of -71 Å² was applied during RELION post-processing to generate a sharpened map that was used for atomic modeling. The local resolution estimation was performed using the two half volumes generated during 3D auto-refinement, with the ‘bloccres’ function in the Bsoft package (Figure S1H) (Heymann and Belnap, 2007).

In order to improve the quality of the reconstructed map at the flexible tail region of the Csy-Acr complex, we used the signal-subtracted focus classification and refinement technique in RELION (Figure S3) (Bai et al., 2015). A soft-edged binary mask for the tail region was applied to the polished particle stack from the first 3D refinement to generate a signal-subtracted stack of particles for focused analysis. Initiating a completely de novo 3D classification and auto-refinement run (starting with larger angular and translational search) using the signal-subtracted particle stack resulted in a 3D density of very poor quality. Instead, we replaced the original stack file with the signal-subtracted stack in the final optimizer file of the 3D refinement and “continued” the refinement. This was followed by a 3D classification run requesting 2 classes, without performing angular or translational search, and using a regularization parameter value of 15. 33,553 signal-subtracted particles from the 3D class, containing the highest resolution structural information, were then subjected to 3D auto-refinement by continuing with the last angular and translational search parameters from the 3D refinement step using polished particles. This procedure produced a better quality final map at ~ 4 Å resolution (gold-standard FSC at 0.143), and was used for building atomic model for the tail region. Similar attempts for improving the quality of the EM density map for the head region did not result in an improved map.

Atomic model building and refinement

De novo atomic models were built for six Cas7f, one Cas8f, one Cas5f, two AcrF2, and 40 nucleotides of the crRNA (without the 3'-hairpin), by using a combination of the stable core map and the focused map for the tail region. COOT was used for model building and visualization (Emsley and Cowtan, 2004). Two copies of the NMR structure of AcrF1 (PDB: 2LW5) were initially rigid-body fit into the EM density map using UCSF Chimera (Goddard et al., 2007) and manually manipulated to fit the EM density using refinement tools in COOT. These coarse models were then refined using real-space refinement in PHENIX v1.11 (Adams et al., 2010). After relaxing this initial, refined model using the Rosetta suite (DiMaio et al., 2015), 200 models were generated using Rosetta refinement program. The resulting models were sorted based on Rosetta score, clash score, Ramachandran outliers, and MolProbity score. The top scoring model was then re-refined using PHENIX real-space refinement, and regions with bad geometries were manually

inspected using COOT. Restraints for secondary structure, Ramachandran, and non-crystallographic symmetry were used during PHENIX real-space refinement runs. Several iterative rounds of refinement and inspection of the model were performed until a consensus model of the complex was obtained. The crystal structure of Cas6 and crRNA 3'-hairpin (PDB: 4AL5) was rigid-body fit into a lower resolution, unsharpened map of the Csy-Acr complex, using UCSF Chimera (Figures S1 and S2). To limit over-interpretation of the atomic model in regions of weak electron density, side chains lacking well-resolved features in the reconstruction were truncated to the C-beta carbon. The docked Cas6 crystal structure was reduced to a C-alpha backbone for deposition. The top 5 atomic models resulting from PHENIX real-space refinement were deposited together to the PDB databank (PDB: 5UZ9).

Mutagenesis

Mutants were made by site directed mutagenesis using quick change (Aglient) or Q5 method (NEB). Primers for mutagenesis are listed in the [Key Resources Table](#).

Electrophoretic mobility shift assay

Oligonucleotides (Eurofins) listed in the [Key Resources Table](#) were 5' end labeled with ^{32}P -ATP (PerkinElmer) using T4 polynucleotide kinase (NEB). Labeled oligonucleotides were purified by phenol/chloroform extraction followed by MicroSpin G-25 columns (GE Healthcare). dsDNA was prepared by mixing labeled oligonucleotides with more than five-fold molar excess of the complementary oligonucleotide in hybridization buffer (20 mM HEPES pH 7.5, 75 mM NaCl, 2mM EDTA, 10% glycerol and 0.01% bromophenol blue). The mixture was incubated at 95°C for 5 min, and gradually cooled to 25°C in a thermocycler. Oligonucleotide duplexes were gel purified, ethanol precipitated, and recovered in hybridization buffer. Increasing concentrations of WT or Csy mutants were incubated with oligonucleotides in hybridization buffer plus 1 mM TCEP. Samples were incubated for 15 min at 37°C, loaded onto a 6% native polyacrylamide gel and run for 3 hr at 150 V at 4°C. Gels were dried, exposed to phosphor storage screens, and scanned with a Typhoon (GE Healthcare) phosphorimager. Bound and unbound DNA fractions were quantified using ImageQuant software (GE Healthcare). After background subtraction, the fractions of bound oligonucleotides were plotted against total Csy concentration. The data were fit by nonlinear regression analysis using the equation:

$$\text{Fraction bound DNA} = (M1 * [\text{Csy}]_{\text{total}}) / (K_D + [\text{Csy}]_{\text{total}})$$

Where M1 is the amplitude of the binding curve. Equilibrium dissociation constants (K_D) are the average of three independent experiments and error bars represent standard deviations.

Surface plasmon resonance

Experiments were conducted with a Biacore X100 SPR instrument (GE Healthcare). Purified Csy complex with 6his-tagged Cas7 subunits was immobilized on a Ni-NTA coated sensor chip. Purified AcrF1 was injected into the buffer flow in increasing concentrations (30.9 nM, 92.6 nM, 278 nM, 833 nM, 2.5 μM). Experiments were conducted at 37°C, in 20 mM Tris, 100 mM NaCl, 1mM TCEP, 0.005% Tween. Data were fit with a model describing a single ligand protein (Csy complex) bound by two analyte proteins (AcrF1). Kinetic rate constants were extracted from this curve fit using Biacore evaluation software (GE). A similar method was used to measure interactions between Csy and AcrF2; however, immobilization of Csy complex on the Ni-NTA chip was perturbed by the addition of AcrF2, so we covalently immobilized Csy complex by amine coupling to the surface of a carboxymethyl-dextran-modified (CM5) sensor chip. Purified AcrF2 was injected into the buffer flow in increasing concentrations (12.1 nM, 40.5 nM, 135 nM, 450 nM, 1.5 μM), and Csy complex-AcrF2 binding events were recorded in real time. Experiments were conducted at 37°C, in 20 mM Tris, 100 mM NaCl, 1mM TCEP, 0.005% Tween. Data were fit with a model describing a bivalent analyte. Kinetic rate constants were extracted from this curve fit using Biacore evaluation software (GE).

QUANTIFICATION AND STATISTICAL ANALYSIS

All experiments were conducted with at least three biological replicates ($n = 3$). Statistical parameters are reported in the Figures and the Figure Legends. Cryo-EM data collection, refinement, and validation statistics are reported in [Table S1](#).

DATA AND SOFTWARE AVAILABILITY

Electron microscopy density maps, including a focused map of the tail, and an atomic model of the complex including a C α trace of Cas6f (based on rigid body fitting of 4AL5) have been deposited at the Electron Microscopy Data Bank and Protein Data Bank under accession numbers EMDB: 8624 and PDB: 5UZ9, respectively. Plasmids used for overexpression and purification of the Csy complex, AcrF1, and AcrF2 have been deposited at Addgene (see [Key Resources Table](#)).

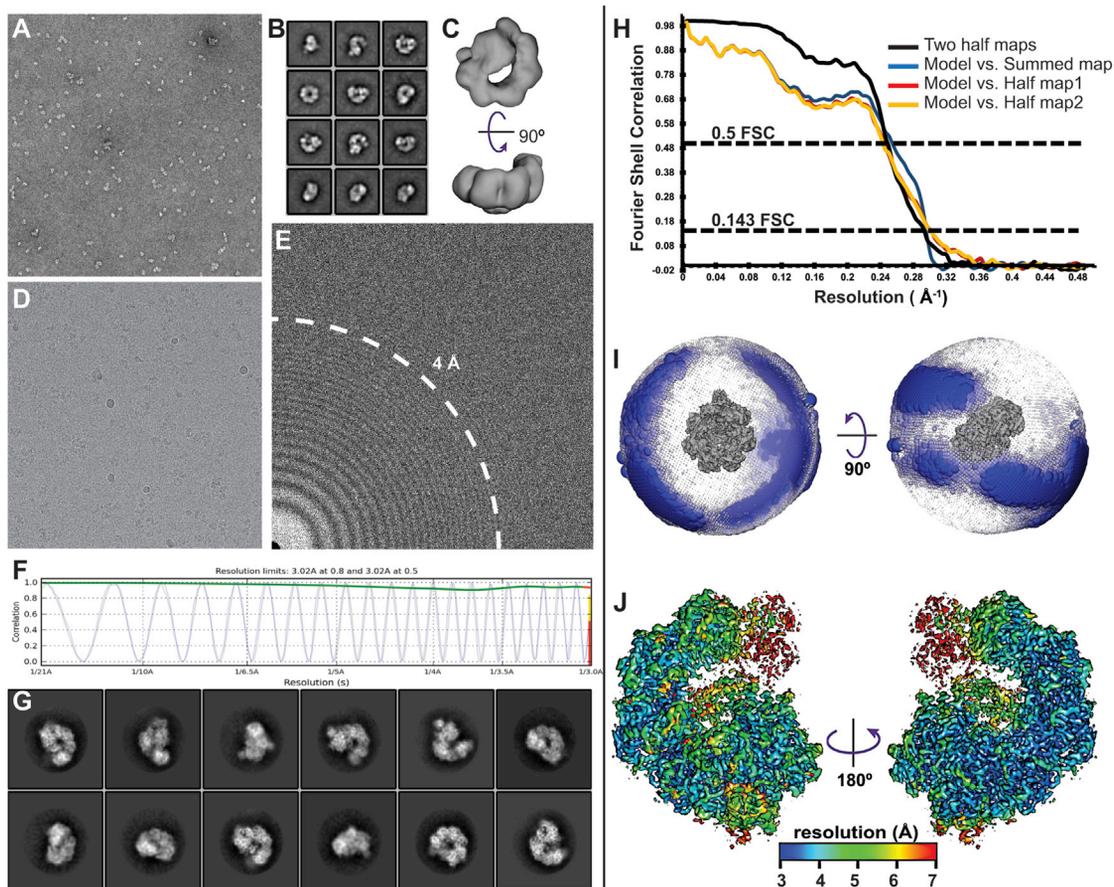


Figure S1. Electron Microscopy of Csy Complex Bound to Two Different Virally Encoded Anti-CRISPR Proteins, Related to Figure 1

(A) A negatively stained micrograph of the Csy-Acr complex showing monodisperse particles.

(B) Representative reference-free 2D class averages of negative stained particles.

(C) Negative stain reconstruction of the Csy-Acr complex at $\sim 25\text{\AA}$ resolution. This was used as initial reference volume for cryo-EM reconstruction.

(D) A representative cryo-EM micrograph of the Csy-Acr complexes in vitreous ice over thin carbon coated with polylysine.

(E) Fourier transform of the micrograph shown in (D), with Thon rings extending beyond 4\AA (white dotted line).

(F) 1D plot showing the correlation (green line) between the experimental and estimated contrast transfer function of the micrograph (blue and black lines, respectively).

(G) Representative reference-free 2D class averages of particles extracted from cryo-EM micrographs, shows different orientations of the Csy-Acr complex in ice.

(H) Fourier Shell Correlation (FSC) curves between the two independently reconstructed half maps (black curve, gold standard FSC curve), atomic model versus summed map (blue curve), and model versus half map1 (red curve) and half map2 (yellow curve).

(I) Euler angle distribution of all particles that contributed to the final 3D reconstructed map of the Csy-Acr complex. The position of each blue sphere relative to the 3D density map (gray in the center) corresponds to its angular assignment, and the radius of the sphere is proportional to the number of particles in that orientation.

(J) The final reconstructed map colored by local resolution estimation using Bsoft's "blores" function (Heymann and Belnap, 2007). The majority of the map is close to $\sim 3\text{\AA}$ resolution, with peripheral regions in the flexible tail and head are not as well resolved.

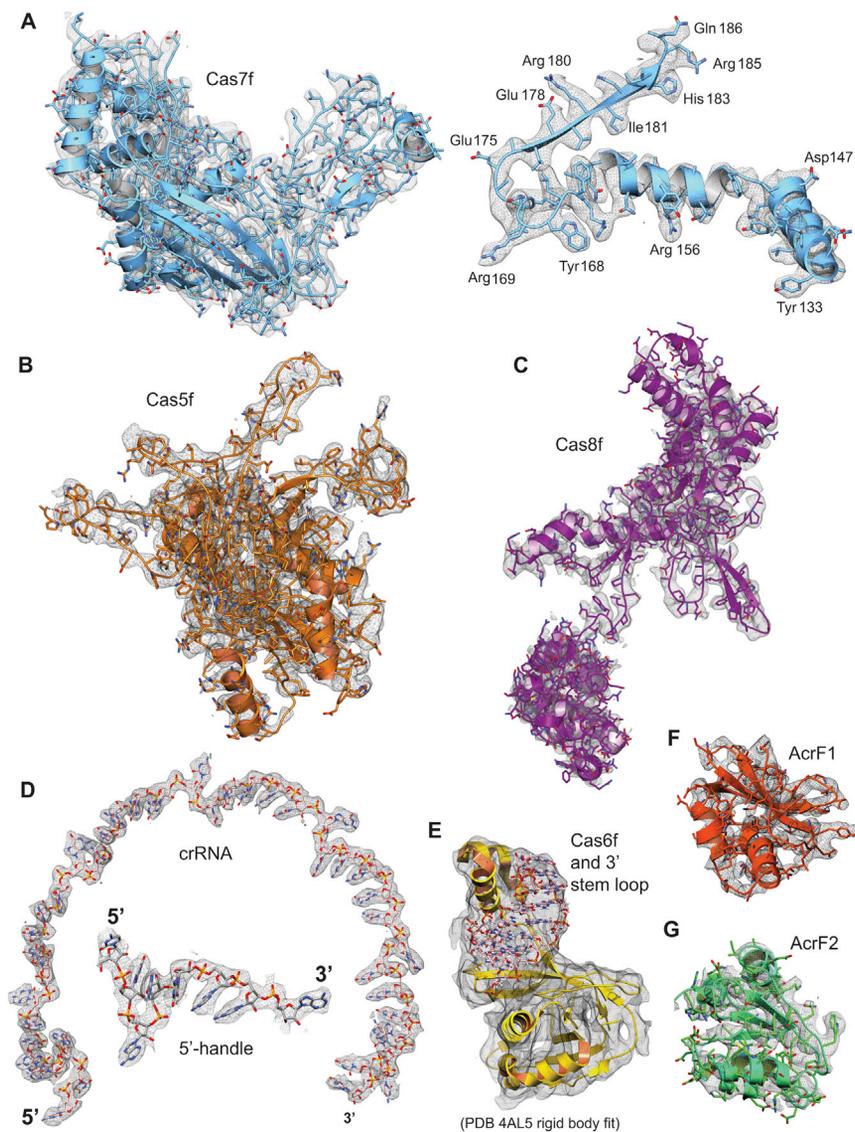


Figure S2. Atomic Models of Different Subunits of Csy-Acr Complex and Corresponding Cryo-EM Density, Related to Figure 1

EM density for each subunit is shown as a gray mesh.

(A) Density from one of the six Cas7f subunits, with the atomic model is shown in blue. To the right, a segment of this subunit containing helices, turns, and a β strand is shown. Side chains of certain residues are labeled to illustrate that individual residues could be unambiguously built into the map.

(B) Cas5f subunit map and model (orange).

(C) Cas8f subunit map and model (purple).

(D) The crRNA map and model. Density corresponding to the 3'-hairpin of the crRNA was poorly resolved and is not included.

(E) Density for the Cas6f "head" and the 3'-hairpin was not sufficiently for model building, but a co-crystal structure of *P. aeruginosa* Cas6f bound to the 3'-hairpin (PDB: 4AL5, yellow) was rigid-body docked into the low-resolution density.

(F) AcrF1 model (red) and map.

(G) AcrF2 model (green) and map.

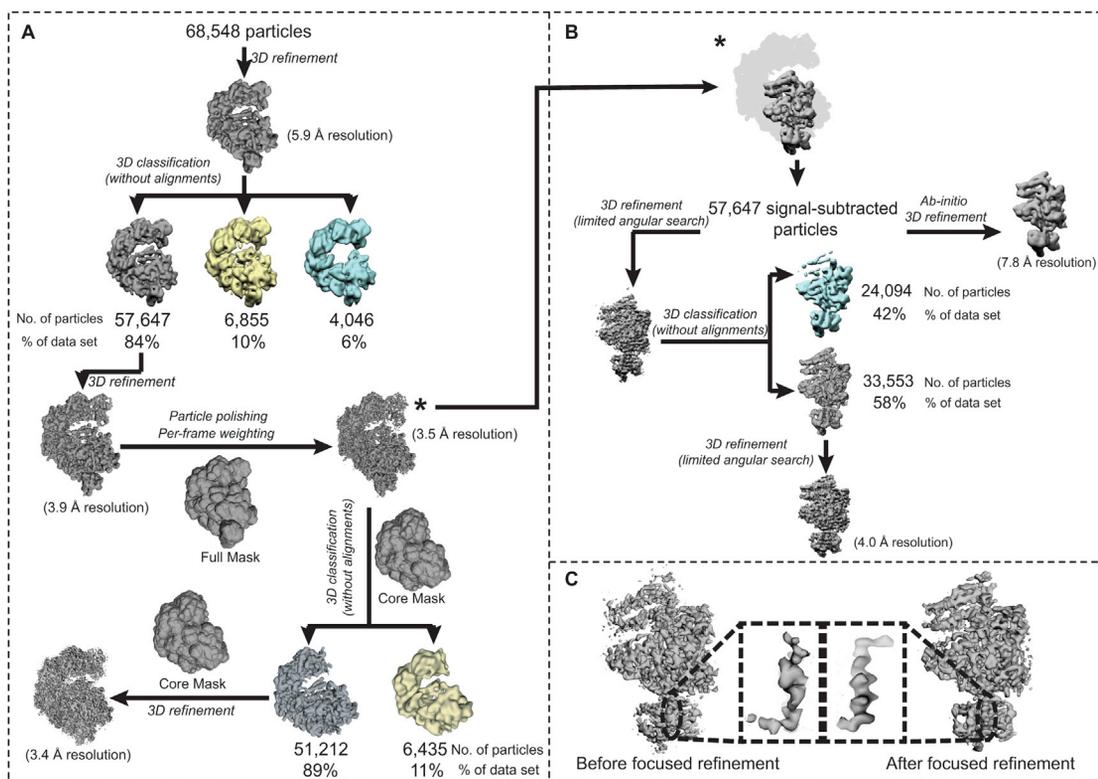


Figure S3. Cryo-EM Data Processing Workflow for 3D Reconstruction and Focused Refinement, Related to Figure 1

(A) For initial refinement and 3D classification, data were binned to 2.05 Å/pixel and particles extracted with a box size of 160 pixels. For all subsequent processing steps, data were extracted at 1.026 Å/pixel with a box size of 320 pixels. All the 3D classification steps were performed without angular or translational searches. The stack of polished/dose-weighted particles (indicated with *) was used for focused 3D analysis to better resolve the “tail-region” of the complex (described in STAR Methods). The 3D masks used for particle polishing and the final refinements are shown.

(B) Focused 3D analysis of the Csy-Acr complex “tail-region.” *Ab-initio* 3D refinement with the signal subtracted particle stack using the segmented tail density as an initial model did not improve the resolution the tail. An improved map was obtained by performing 3D refinement using local angular and translational searches (local 3D refinement).

(C) The final map for the tail after focused 3D refinement (shown on the right) had better resolved structural details than the corresponding region in the original 3D refinement (left). A helix density from the Cas8f subunit in the tail has been magnified to exemplify the improvement in map quality using the focused refinement approach.

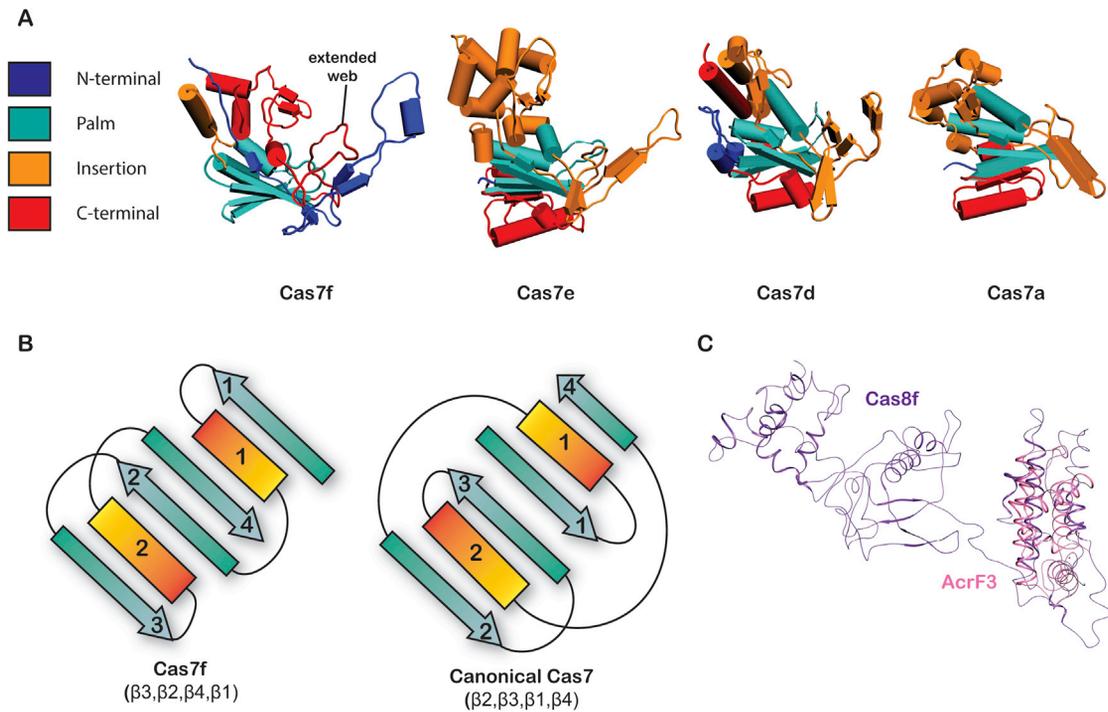


Figure S4. Cas7f Has a Distinct Fold, and the C-terminal Helical Bundle of Cas8f is Structurally Similar to AcrF3, Related to Figures 2, 3, and 5
 (A) Cas7 family proteins presented in the “right-hand” orientation. In contrast to other Cas7 family proteins where the fingers and thumb domains are insertions in an RRM (palm), the palm of Cas7f is not an RRM. The Cas7f palm contains only a one small insertion (orange), and the thumb and fingers domains are composed of N- and C-terminal extensions (blue and red, respectively).

(B) Cartoon showing the topology of the Cas7f palm compared to the canonical Cas7 family RRM.

(C) The Cas8f structure was used to perform structural homology search using Dali (Holm and Sander, 1993). AcrF3 was identified as one of the top hits (z-score of 5.3). Superposition of AcrF3 (pink) structure on the C-terminal helix-bundle domain of Cas8f (violet). The RMSD for 114 equivalently positioned C-alpha atoms is 4.1 Å.

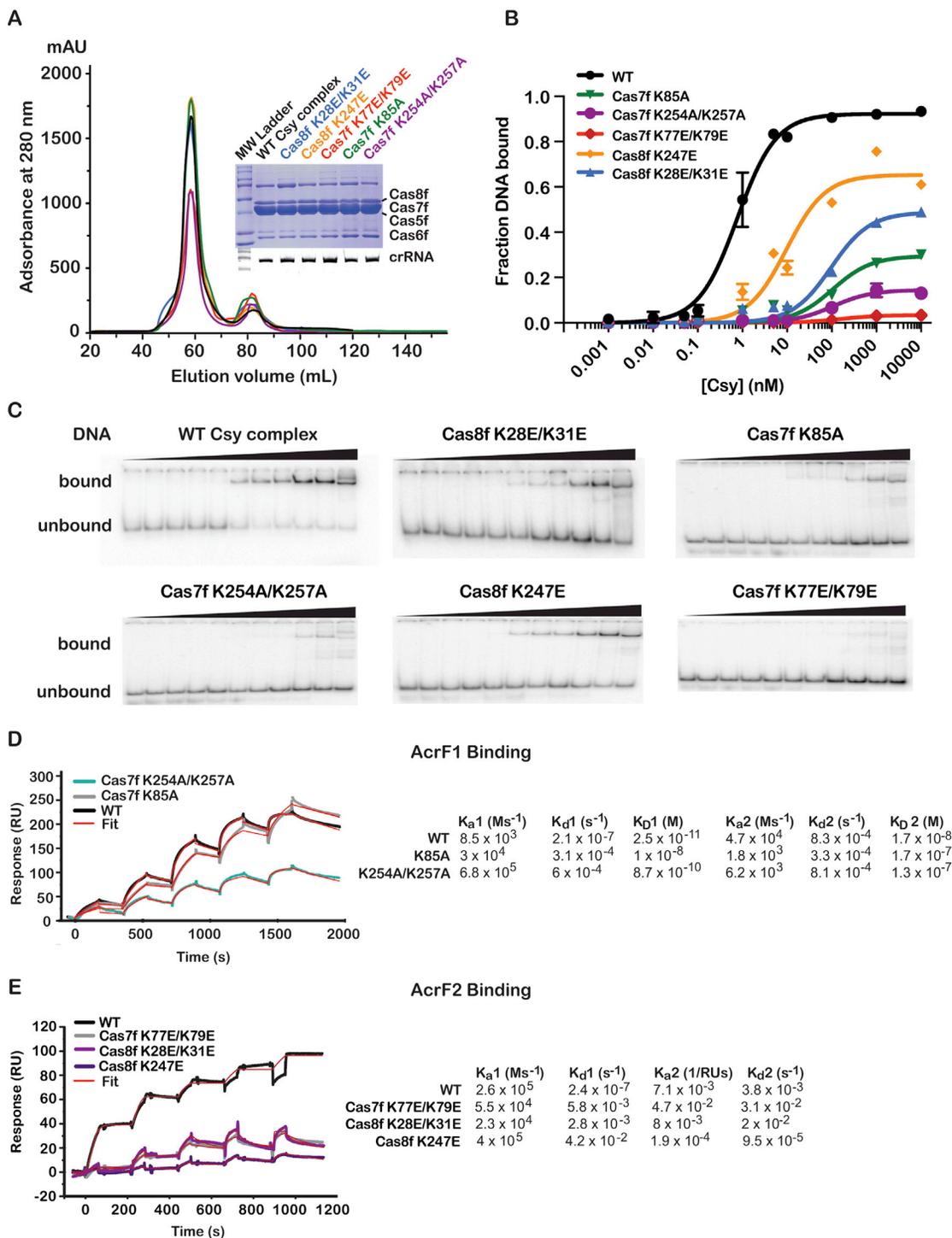


Figure S5. Mutations in the Csy Complex Perturb DNA Binding and Anti-CRISPR Binding Kinetics, Related to Figures 4 and 5

(A) Elution profile of the WT Csy complex and different mutants. The insert shows a Coomassie blue-stained SDS-PAGE gel (top) and a denaturing polyacrylamide gel of phenol:chloroform extracted crRNA isolated for the each of the Csy complexes (bottom).

(B) Binding isotherms based on electrophoretic mobility shift assays performed by adding increasing concentrations (0, 0.001, 0.01, 0.05, 0.1, 1, 5, 10, 100, 1000, 10000 nM) of purified WT or mutant Csy complexes to a 72-nt double-stranded DNA that was ^{32}P -labeled on the complementary strand. The fraction of DNA bound was plotted as a function of increasing Csy concentration. Equilibrium dissociation constants (K_D) were determined from an average of three independent experiments and the error bars represent standard deviations.

(C) Representative gels for each of the six different Csy complexes.

(legend continued on next page)

(D) Overlay of sensorgrams from surface plasmon resonance (SPR) experiments, used to determine kinetics of anti-CRISPR AcrF1 binding to WT Csy complex (black) or Csy complex with mutations in Cas7f subunit (green and gray). Data were fit with a model describing binding for a heterogeneous ligand (i.e., binding of two analyte proteins to a single ligand protein) (red lines). Rate constants are reported in the table on the right.

(E) Overlay of sensorgrams from surface plasmon resonance (SPR) experiments, used to determine kinetics of anti-CRISPR AcrF2 binding to WT Csy complex (black), or Csy complex with mutations in Cas7f (gray) or Cas8f subunits (violet and purple). Data were fit with a model describing binding of a bivalent analyte (i.e., one analyte protein binding spatially discrete areas on the same ligand protein). Rate constants are reported in the table on the right.