Dear author,

Please note that changes made in the online proofing system will be added to the article before publication but are not reflected in this PDF.

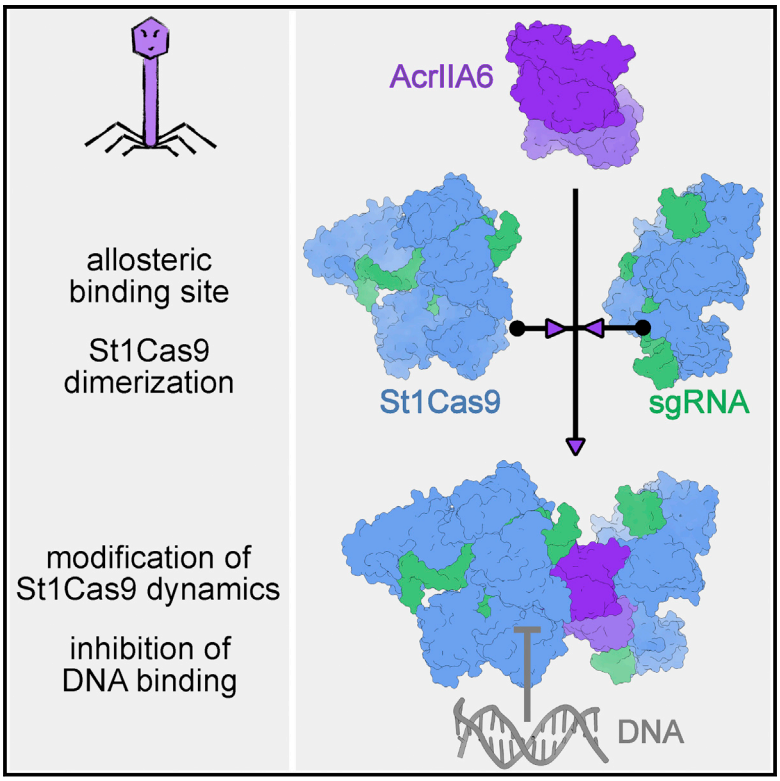
We also ask that this file not be used for submitting corrections.

Article

Molecular Cell

Cas9 Allosteric Inhibition by the Anti-CRISPR Protein AcrIIA6

Graphical Abstract



Authors

Olivier Fuchsbauer, Paolo Swuec, Claire Zimberger, ..., Sylvain Moineau, Yannick Doyon, Adeline Goulet

Correspondence

adeline.goulet@afmb.univ-mrs.fr

In Brief

Fuchsbauer et al. present structural and functional data highlighting the allosteric inhibition mechanism used by AcrIIA6 to inactivate St1Cas9. AcrIIA6 modifies St1Cas9 dynamics and structure, which ultimately prevents its binding to DNA within cells. A naturally resistant St1Cas9 variant illustrates how bacterial cells can escape phages anti-CRISPR systems.

Highlights

- AcrIIA6 allosterically inhibits St1Cas9 and can induce its dimerization
- AcrIIA6 alters St1Cas9 dynamics associated with PAM binding
- AcrIIA6 reduces St1Cas9 DNA binding affinity, thereby blocking DNA binding in cells
- A natural St1Cas9 variant illustrates an anti-CRISPR driven mutational escape

Fuchsbauer et al., 2019, Molecular Cell *76*, 1–16
December 19, 2019
https://doi.org/10.1016/j.molcel.2019.09.012

CellPress

Please cite this article in press as: Fuchsbauer et al., Cas9 Allosteric Inhibition by the Anti-CRISPR Protein AcrIIA6, Molecular Cell (2019), https://doi.org/10.1016/j.molcel.2019.09.012

Molecular Cell
Article

Cas9 Allosteric Inhibition by the Anti-CRISPR Protein AcrIIA6

Olivier Fuchsbauer,[1,2,9] Paolo Swuec,[3,4,9] Claire Zimberger,[1,2] Béatrice Amigues,[1,2] Sébastien Levesque,[5] Daniel Agudelo,[5] Alexis Duringer,[5] Antonio Chaves-Sanjuan,[3] Silvia Spinelli,[1,2] Geneviève M. Rousseau,[6,7] Minja Velimirovic,[5] Martino Bolognesi,[3,4] Alain Roussel,[1,2] Christian Cambillau,[1,2] Sylvain Moineau,[6,7,8] Yannick Doyon,[5] and Adeline Goulet[1,2,10,*]

[1]Architecture et Fonction des Macromolécules Biologiques, Centre National de la Recherche Scientifique (CNRS), Campus de Luminy, Case 932, 13288 Marseille Cedex 09, France

[2]Architecture et Fonction des Macromolécules Biologiques, Aix-Marseille Université, Campus de Luminy, Case 932, 13288 Marseille Cedex 09, France

[3]Dipartimento di Bioscienze, Università degli Studi di Milano, Via Celoria 26, 20133 Milano, Italy

[4]Centro di Ricerca Pediatrica Romeo ed Enrica Invernizzi, Università degli Studi di Milano, Via Celoria 26, 20133 Milano, Italy

[5]Centre Hospitalier Universitaire de Québec–Université Laval Research Center, Québec City, QC G1V 4G2, Canada

[6]Département de biochimie, de microbiologie, et de bio-informatique, Faculté des sciences et de génie, Université Laval, Québec City, QC, G1V 0A6, Canada

[7]Groupe de recherche en écologie buccale, Faculté de médecine dentaire, Université Laval, Québec City, QC, G1V 0A6, Canada

[8]Félix d'Hérelle Reference Center for Bacterial Viruses, Faculté de médecine dentaire, Université Laval, Québec City, QC, G1V 0A6, Canada

[9]These authors contributed equally

[10]Lead Contact

*Correspondence: adeline.goulet@afmb.univ-mrs.fr

https://doi.org/10.1016/j.molcel.2019.09.012

SUMMARY

In the arms race against bacteria, bacteriophages have evolved diverse anti-CRISPR proteins (Acrs) that block CRISPR-Cas immunity. Acrs play key roles in the molecular coevolution of bacteria with their predators, use a variety of mechanisms of action, and provide tools to regulate Cas-based genome manipulation. Here, we present structural and functional analyses of AcrIIA6, an Acr from virulent phages, exploring its unique anti-CRISPR action. Our cryo-EM structures and functional data of AcrIIA6 binding to *Streptococcus thermophilus* Cas9 (St1Cas9) show that AcrIIA6 acts as an allosteric inhibitor and induces St1Cas9 dimerization. AcrIIA6 reduces St1Cas9 binding affinity for DNA and prevents DNA binding within cells. The PAM and AcrIIA6 recognition sites are structurally close and allosterically linked. Mechanistically, AcrIIA6 affects the St1Cas9 conformational dynamics associated with PAM binding. Finally, we identify a natural St1Cas9 variant resistant to AcrIIA6 illustrating Acr-driven mutational escape and molecular diversification of Cas9 proteins.

Q1 Q2 INTRODUCTION

The constant battle between bacteriophages (phages)—the most abundant biological entities on the planet (Cobián Güemes et al., 2016)—and their bacterial preys induces an evolutionary arms race with correlated adaptations of attack and defense mechanisms. Bacteria have evolved an impressive arsenal of antiviral tools, among which the adaptive CRISPR-Cas immune system acts by cleaving invading nucleic acids in a sequence-specific manner (Samson et al., 2013). Upon infection, small pieces of viral genomes are integrated into a CRISPR locus, as spacers flanked by short repeat sequences, to immunize the cell. Spacer-repeat sequences are then transcribed into small RNA guides that associate with Cas nucleases to form ribonucleoproteins (RNPs), named surveillance complexes. In the so-called interference step, these complexes are in charge of recognition and cleavage of invading complementary DNA sequences, known as the protospacers (Garneau et al., 2010).

CRISPR-Cas defense systems are remarkably diverse in the microbial world (Koonin et al., 2017), possibly in response to an even more diversified viral population. Type II-A CRISPR-Cas systems involve a single effector protein for target recognition and cleavage (Cas9) and a single guide RNA (sgRNA) (Jinek et al., 2012). Because the CRISPR-Cas9 tool can easily be programmed to target virtually any sequence, it has given rise to a popular versatile genome-engineering technology, widely used in research and in promising therapeutic applications (Wang et al., 2016a).

From a mechanistic point of view, the Cas9 RNP adopts a DNA recognition-competent conformational state (Jiang et al., 2015; Jinek et al., 2014) to search the genomic landscape for a specific protospacer adjacent motif (PAM), primarily by a three-dimensional diffusion mechanism. PAM binding induces a local separation of neighboring DNA base pairs, which then allows the RNA strand to form an RNA-DNA heteroduplex with the target DNA strand (tDNA), leaving a displaced non-target DNA strand (ntDNA), named the R-loop. Cas9 conformational dynamics assist PAM-binding and R-loop formation, which ultimately

triggers large reorientations of its HNH and RuvC nuclease domains, to achieve the active state for DNA cleavage (reviewed in Wilkinson et al., [2019])(Mekler et al., 2017; Palermo et al., 2016; Szczelkun et al., 2014). Allosteric communications between the PAM-binding site, R-loop, HNH and RuvC domains, and the associated conformational checkpoints between DNA binding and cleavage are crucial for Cas9 activity and cleavage of specific protospacers (Dagdas et al., 2017; Jiang et al., 2016; Josephs et al., 2015; Palermo et al., 2017; Sternberg et al., 2014, 2015; Yang et al., 2018).

It is not surprising that in the battle for persistence, phages have developed means to escape CRISPR-Cas bacterial immunity. While mutations and deletions of the phage targets can generally bypass this antiviral system (Barrangou et al., 2007; Deveau et al., 2008; Garneau et al., 2010; Semenova et al., 2011), some phages can also bring into play anti-CRISPR proteins (Acrs). Acrs are viral proteins produced during phage infection that interact directly with, or modify, components of CRISPR-Cas systems to block the interference step (Bondy-Denomy et al., 2013) (Hwang and Maxwell, 2019) (Dong et al., 2019), therefore allowing completion of the phage lytic cycle.

The discovery of diverse Acrs has rapidly garnered interest, as they form a promising reservoir of biotechnological tools for the control of CRISPR-Cas9-based genome manipulation. In particular, Acrs could be used as genetically encodable off-switches for a precise control of Cas9 activity over time (Shin et al., 2017). To date, nearly 45 diverse Acr families targeting types I, II, and V CRISPR-Cas systems have been identified, mainly in temperate phages, prophages, and mobile genetic elements (Hwang and Maxwell, 2019) (Uribe et al., 2019). Besides being small (<190 amino acids), Acrs share limited sequence similarity with proteins of known function. Structure-function studies have revealed that Acrs can compete with sgRNA-loading (AcrIIC2 [Zhu et al., 2019]), DNA hybridization (AcrIF1, [Chowdhury et al., 2017]), PAM binding (AcrIIA4 [Dong et al., 2017; Shin et al., 2017; Yang and Patel, 2017], AcrIIA2 [Jiang et al., 2019; Liu et al., 2019]), recruitment of the nuclease domain (AcrIF3, [Rollins et al., 2019; Wang et al., 2016b; Wang et al., 2016c]), and interaction with catalytic residues (AcrIIC1 [Harrington et al., 2017]). Also, AcrVA5 covalently modifies Cas12a to prevent PAM binding (Dong et al., 2019).

Recently, we examined virulent phages that are able to bypass the seemingly robust *Streptococcus thermophilus* CRISPR-Cas immunity and identified two Acr families in *S. thermophilus* phages, AcrIIA5 and AcrIIA6 (Hynes et al., 2017, 2018). Interestingly, AcrIIA5 displays a broad type II-A anti-CRISPR activity, while AcrIIA6 specifically inhibits St1Cas9, the model system that kicked off the CRISPR craze (Barrangou et al., 2007). Of interest, both AcrIIA5 and AcrIIA6 block Cas9-mediated genome editing of human cells (Hynes et al., 2018).

The continuous discovery of diverse Acr families raises intriguing questions, particularly relating to the variety of their modes of action, and challenges for their biotechnological use. Besides, their roles in the evolutionary diversification of CRISPR-Cas systems and in the emergence of bacterial anti-Acr strategies remain poorly explored. In this context, we report a structural and functional characterization of AcrIIA6 that uncovers a so-far unknown molecular mechanism for St1Cas9 allosteric inhibition. We demonstrate that AcrIIA6 stably binds to St1Cas9 RNP and can induce its dimerization, and we present four cryo-electron microscopy (cryo-EM) structures of differently assembled complexes in the 3.0–3.3 Å resolution range. We show that AcrIIA6 reduces St1Cas9 RNP binding affinity for the target DNA, preventing DNA binding within cells. Our cryo-EM analyses indicate that the AcrIIA6 inhibitory effects are exerted through modifications of St1Cas9 conformational dynamics needed for PAM binding. Such findings led us to identify a natural variant of St1Cas9 resistant to AcrIIA6, which likely derives from Acr-driven mutational escape.

RESULTS

AcrIIA6 Stably Binds to St1Cas9 Surveillance Complex

We previously reported that AcrIIA6 specifically blocks the interference activity of St1Cas9 in bacteria and genome editing in mammalian cells (Hynes et al., 2018). To determine the inhibitory activity of AcrIIA6, we first investigated whether or not AcrIIA6 binds to the surveillance complexes. We produced recombinant AcrIIA6 (from the virulent phage D1811) (Hynes et al., 2018) and St1Cas9 (from *S. thermophilus* strain DGCC7710) (Karvelis et al., 2015) (Figures S1A and S1B) and performed BioLayer interferometry (BLi) experiments to monitor real-time association and dissociation between AcrIIA6 and apo-St1Cas9 or St1Cas9 RNP. AcrIIA6 strongly interacted with St1Cas9 RNP, as shown by their very slow dissociation rate that indicates the formation of a quasi-irreversible complex (Figure 1A). On the other hand, AcrIIA6 did not bind to apo-St1Cas9, apo-*Streptococcus pyogenes* Cas9 (apo-SpCas9), or SpCas9 RNP. We had previously shown that AcrIIA6 did not block the interference activity of SpCas9 (Hynes et al., 2018). Subsequently, we incubated St1Cas9 with sgRNA molecules that were pre-annealed to a complementary 20-nt-long ssDNA target sequence (tDNA20) and observed comparable association and dissociation rates of AcrIIA6 to the target-bound St1Cas9 RNP (Figure S1C). These results indicate that AcrAll6 recognizes a protein surface or a mixed protein-RNA surface formed upon sgRNA-binding to St1Cas9.

The AcrIIA6 Binding Mode Revealed by Cryo-EM

To further elucidate the inhibition mechanism of AcrIIA6 on St1Cas9, we determined 3D structures of AcrIIA6-bound St1Cas9 RNP using single particle cryo-EM. We obtained a good-quality vitrified sample suitable for high-resolution single-particle analyses (Figure S2A), in terms of particle distribution and minimization of on-grid protein aggregation occurrence, when we performed pre-annealing of sgRNA to tDNA20, subsequently added St1Cas9, and then added AcrIIA6 (Figure 1B). We then inspected and imaged at the electron microscope the vitrified St1Cas9•sgRNA•tDNA20•AcrIIA6 assembly. Because the presence of tDNA20 has no effect on the AcrIIA6 binding properties to St1Cas9 RNP (Figure 1), the St1Cas9•sgRNA-•tDNA20•AcrIIA6 assembly is a good representative model of the St1Cas9 RNP in its AcrIIA6-bound state.

Analysis of single-particles by reference-free 2D classification revealed the presence of two populations: a major particle class (75.2% of identified particles) showing compact particles with

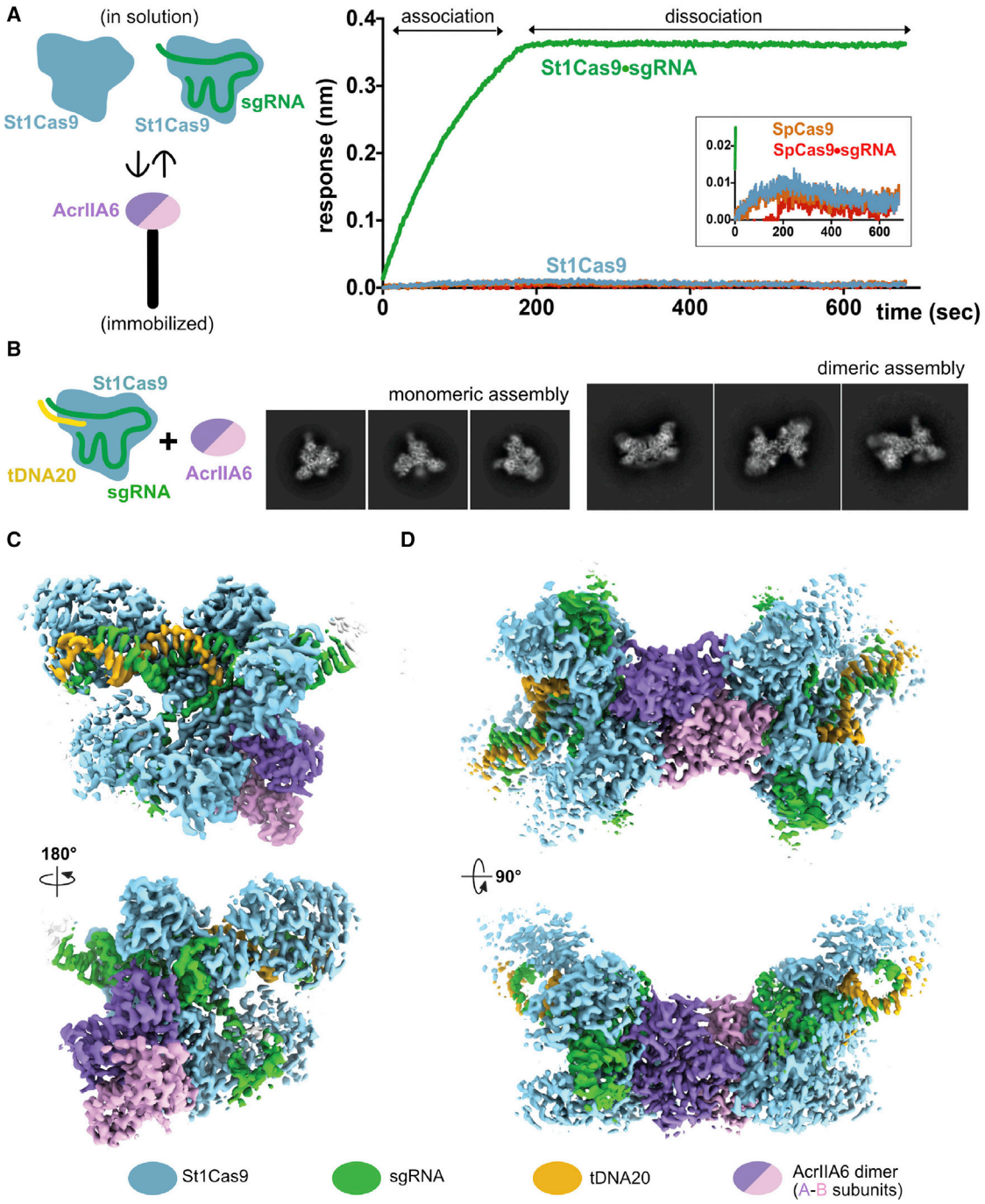


Figure 1. The AcrIIA6 Binding Mode to St1Cas9 Surveillance Complex Revealed by Cryo-EM

(A) Interactions between AcrIIA6 and apo-St1Cas9 (St1Cas9) or St1Cas9 RNP (St1Cas9•sgRNA) were monitored using BLi. AcrIIA6 stably binds to the St1Cas9 RNP, as shown by their slow dissociation rate, but does not bind to apo-St1Cas9. The binding curves of negative controls apo-SpCas9 (SpCas9) and SpCas9 RNP (SpCas9•sgRNA) are shown in the inset. Experiments were performed three times and yielded equivalent results. One experiment is shown.

(B) Representative cryo-EM 2D class averages of monomeric (box size: 263 Å × 263 Å) and dimeric (box size: 334 Å × 334 Å) assemblies of tDNA20-bound St1Cas9 RNP in complex with AcrIIA6.

(C and D) Orthogonal views of cryo-EM 3D reconstructions of the monomeric (C) and dimeric (D) assemblies.

See also Figures S1, S2, and S3.

canonical Cas9 morphology, which we will refer to as the monomeric assembly, together with a minor class (24.8% of identified particles) with a symmetric-elongated shape and dimensions consistent with two Cas9 molecules, which we will refer to as the dimeric assembly (Figures 1B, S2B, and S2C; Table 1). Final 3D reconstructions of monomeric and dimeric assemblies, at overall resolutions of 3.2 and 3.0 Å, respectively, clearly showed the presence of one AcrIIA6 dimer bound to one or two copies of St1Cas9•sgRNA•tDNA20 (Figures 1C, 1D, and S2D–S2I; Table 1).

In both cryo-EM maps, densities of AcrIIA6, sgRNA, and tDNA20 are clearly defined (Figure S2J). The St1Cas9 component is also well defined, except for the HNH and RuvC domains, which are known to be intrinsically flexible (Anders et al., 2014; Huai et al., 2017; Jiang et al., 2015; Jinek et al., 2014; Nishimasu et al., 2014; Palermo et al., 2016; Yang et al., 2018) (Figures 1C, 1D, S2D, and S2E). AcrIIA6 density showed a tightly packed dimer hosting sixteen α helices and two four-stranded β sheets, in agreement with its crystal structure (Hynes et al., 2018). Individual bases of sgRNA and tDNA20 were clearly resolved in both cryo-EM maps (Figure S2J), allowing manual modeling of the nucleic acids. Since the 3D structure of St1Cas9 was not previously known, we first generated a homology model using the crystal structure of *Staphylococcus aureus* Cas9 (SaCas9) (Nishimasu et al., 2015) as a template to fit the cryo-EM maps (St1Cas9 and SaCas9 share 33% amino acid sequence identity) and manually modeled the missing and diverging regions. The generated molecular models of the complexes were refined to optimal stereochemistry and in agreement with experimental data (Table 1).

Finally, we compared the two refined models and showed that the dimeric and monomeric assemblies match closely (St1Cas9: RMSD of 0.7 Å for 812 equivalent Cα atoms; AcrIIA6 subunit A: RMSD of 0.3 Å for 183 equivalent Cα atoms; AcrIIA6 subunit B: RMSD of 0.6 Å for 183 equivalent Cα atoms), thus excluding any major conformational rearrangement induced by St1Cas9 RNP dimerization (Figures S3A and S3B). While mapping the structural bases of AcrIIA6-driven St1Cas9 inhibition, the above results also provide the first experimental 3D structures of St1Cas9 and its cognate sgRNA, which are presented below before analyzing their interaction with AcrIIA6.

Structural Overview of St1Cas9 and Its sgRNA

St1Cas9 (1,121 amino acids [Karvelis et al., 2015]) adopts the canonical bilobed architecture with an α-helical recognition (REC) lobe (residues 74–453) and a nuclease (NUC) lobe (residues 1–39 and 467–1,121), connected by an arginine-rich bridge helix (BH, residues 40–73) and a linker partially resolved in the cryo-EM density maps (residues 454–466) (Figures 2A and 2C). In the NUC lobe, the flexible HNH domain (predicted residues 559–664) is not visible in our cryo-EM density maps, and the RuvC-like domain reconstruction is incomplete (modeled RuvC-I residues 3–39, RuvC-II residues 478–509, RuvC-III residues 690–715 and 736–749) (Figure 2C). The phosphate lock loop (residues 814–826) connects these nuclease domains to the α/β wedge domain (WED, residues 827–963). At the C-terminal end, the topoisomerase-homology (TOPO, residues 964–1,025) and all-β C-terminal domains (CTD, residues 1,026–1,121) together hold the PAM-interacting (PI) domain (Figures 2C and S4A). Overall, as expected, the 3D structures of St1Cas9 and SaCas9 are similar (RMSD of 3.4 Å for 770 equivalent Cα atoms) (Figure S4B).

The sgRNA in complex with St1Cas9 assembles the guide sequence (G1–U20), the repeat-antirepeat duplex (G21–U40, A45–U67) held by a tetraloop (C41–U44), and the consecutive stem loop 1 (A68–A82), linker (A83–C86), and stem loop 2 (A87–U117) (Müller et al., 2016) (Figures 2B and 2C). The tetraloop and the distal region of the repeat-antirepeat duplex (C34–C52) and of stem loop 2 (C93–G109), which are not in close contact with St1Cas9, were poorly resolved in the cryo-EM maps and could not be modeled. St1Cas9's sgRNA exhibits the characteristic T-shape of type II CRISPR-Cas sgRNA molecules (Hirano et al., 2016; Nishimasu et al., 2014, 2015; Yamada et al., 2017), in which the guide and target DNA form a heteroduplex via Watson-Crick base pairs (Figures 2B, 2C, and S4C). The G:U wobble base pair between the guide sequence and the repeat-antirepeat duplex (G21:U67), the bulge distorting the repeat-antirepeat duplex (C30, A56–C58), and the *syn* conformation of the adenosine nucleotide at the junction between the repeat-antirepeat duplex and stem loop1 (A68) (Figures 2B and S3D) are conserved features in type II-A CRISPR-Cas systems that play important roles in folding sgRNA molecules and activating Cas9-catalyzed DNA cleavage (Anders et al., 2014; Hirano et al., 2016; Nishimasu et al., 2014, 2015; Yamada et al., 2017).

The 3D structures of the St1Cas9•sgRNA•tDNA20•AcrIIA6 complex show canonical molecular determinants prompting the formation of Cas9 RNP (Hirano et al., 2016; Nishimasu et al., 2014, 2015; Yamada et al., 2017). The guide RNA-tDNA20 and repeat-antirepeat duplexes are buried between the REC and NUC lobes, while stem loop 1 and stem loop 2 extensively interact with the protein surface on the other side of St1Cas9, referred to as St1Cas9 back face relative to the DNA binding crevasse (Figure 2C). As previously observed for SaCas9 and SpCas9 (Briner et al., 2014; Nishimasu et al., 2015), we demonstrate that truncation of the sgRNA after stem loop 1 abrogates nuclease activity in human cells (Figure S4E). The linker and stem loop 2 make extensive contacts with St1Cas9, connecting Cas9 key structural motifs involved in target binding and target cleavage (the bridge helix, phosphate-lock loop, TOPO, CTD, and RuvC domains) (Figure S4F), which provides the structural basis for their function in St1Cas9 activity.

AcrIIA6 Binds to an Allosteric Site on St1Cas9

Inspection of the 3D structures reveals that AcrIIA6 binds to a large protein-RNA surface area of the St1Cas9•sgRNA complex, which is primarily formed by the WED and PI domains, repeat-antirepeat duplex and stem loop 1 (Figure 3A). In striking contrast with the binding sites previously characterized for different Cas9-Acr pairs (Dong et al., 2017; Harrington et al., 2017; Jiang et al., 2019; Liu et al., 2019; Shin et al., 2017; Yang and Patel, 2017; Zhu et al., 2019), AcrIIA6 does not map to the DNA-binding crevasse or to the catalytic domains. Rather, AcrIIA6 is bound to a St1Cas9 region that is well separated from the above-mentioned functional sites, located on the back face of the enzyme

Table 1. Cryo-EM Data Collection, Image Processing and Model Refinement Statistics

Data Collection and Image Processing				
Microscope	TALOS Arctica	TALOS Arctica	TALOS Arctica	TALOS Arctica
Voltage (kV)	200	200	200	200
Camera	Falcon 3EC	Falcon 3EC	Falcon 3EC	Falcon 3EC
Magnification (kV)	× 120,000	× 120,000	× 120,000	× 120,000
Total electron dose (e^{-}/Å^2)	40.0	40.0	40.0	40.0
Defocus range (μm)	−0.5 to −2.5	−0.5 to −2.5	−0.5 to −2.5	−0.5 to −2.5
Pixel size (Å)	0.889	0.889	0.889	0.889
Micrographs (no.)	2,413	2,413	2,014	2,014
	St1Cas9•sgRNA •tDNA20•AcrIIA6 (monomeric assembly)	St1Cas9•sgRNA •tDNA20•AcrIIA6 (dimeric assembly)	St1Cas9•sgRNA •tDNA59-ntPAM	St1Cas9•sgRNA •AcrIIA6 •tDNA59-ntPAM
Symmetry imposed	C1	C2	C1	C1
Initial particles (no.)	171,585	56,442	190,475	230,397
Final particles (no.)	118,189	43,239	68,361	50,728
Resolution (Å), (FSC threshold = 0.143)	3.2	3.0	3.3	3.2
Sharpening B-factor (Å^2)	−20	−25	−25	−22
EMDB code	EMD-4900	EMD-4901	EMD-4902	EMD-4904
Model refinement				
Non hydrogen atoms	11,874	20,372	9,332	12,220
Protein residues	1,202	1,996	831	1,180
Nucleotides	97	194	123	123
RMSDs, bond lengths (Å)	0.006	0.005	0.006	0.008
RMSD, bond angles (°)	0.775	0.850	0.892	0.948
Ramachandran plot, favored (%)	93.3	94.2	93.1	93.2
Ramachandran plot, allowed (%)	6.7	5.8	6.9	6.8
Ramachandran plot, disallowed (%)	0	0	0	0
Validation, Molprobity score	1.48	1.45	1.52	1.43
Validation, Clashscore	2.59	2.75	2.85	2.15
Validation, Poor rotamers (%)	0.28	0.23	0.27	0.10
Map-model correlation	0.83	0.77	0.80	0.82
PDB code	PDB: 6RJ9	PDB: 6RJA	PDB: 6RJD	PDB: 6RJG

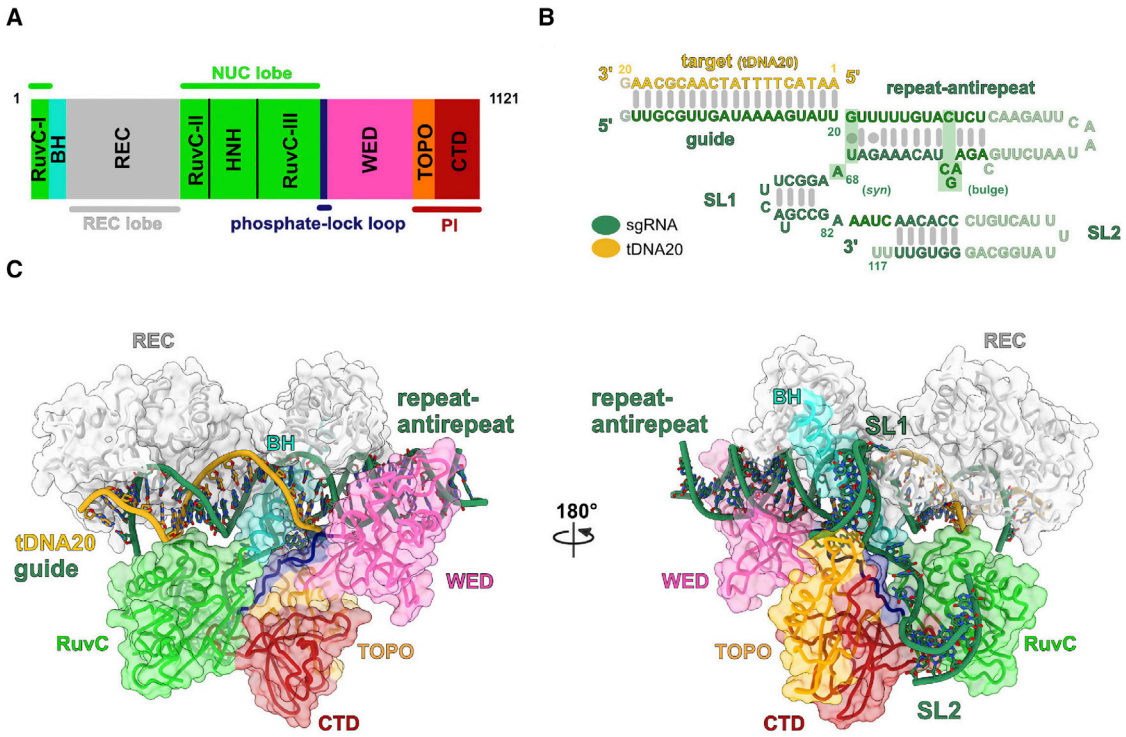


Figure 2. Structural Overview of St1Cas9 and Its sgRNA

(A) Domain organization of St1Cas9. BH: bridge helix, CTD: C-terminal domain, PI: PAM-interacting domain, WED: wedge domain.

(B) Sequences of the sgRNA (SL1: stem loop 1, SL2: stem loop 2) and tDNA20. Grey nucleotides in the target-guide heteroduplex, repeat-antirepeat duplex, and SL2 were not be modeled. The G1 nucleotide in the *in vitro* transcribed sgRNA substitutes the theoretical C1 nucleotide. Watson-Crick and G:U wobble base pairs are shown as lines and filled circles, respectively. Green boxes indicate conserved structural features in CRISPR-Cas9 systems.

(C) Surface and ribbon representations of the St1Cas9•sgRNA•tDNA20 complex (AcrIIA6 is not shown for clarity). The linker between the REC and NUC lobes (457–463) is not modeled.

The back face of St1Cas RNP, relative to the DNA binding crevasse, is shown on the left-hand side.

See also Figure S4.

molecule. Thus, from a structural viewpoint, AcrIIA6 presents a binding mode with the characteristics of an allosteric inhibitory protein.

Moreover, the AcrIIA6 dimer contains two identical St1Cas9-binding sites, thus promoting the formation of a symmetric dimeric assembly that hosts two St1Cas9 molecules (Figure 3B). Each St1Cas9 binding site is built by regions of both AcrIIA6 subunits A and B. Notably, the recognition interface is identical in the monomeric and dimeric assemblies and is detailed below in the context of the monomeric assembly.

AcrIIA6 subunit A engages ~17% of its surface area to cover a mixed protein and sgRNA region, burying ~1,245 Å^2 of St1Cas9 and ~515 Å^2 of sgRNA surface areas. In contrast, subunit B does not interact with the sgRNA but engages ~5% of its surface area to cover ~530 Å^2 of St1Cas9 only. Altogether, binding of the AcrIIA6 dimer buries an area of ~2,270 Å^2 on St1Cas9 RNP upon complex formation, which contributes to the strong interaction of the St1Cas9 RNP with the inhibitor (Figure 1A).

The intermolecular recognition between St1Cas9 RNP and AcrIIA6 relies on complementary shapes and extensive van der Waals contacts. Polar interactions are also established where AcrIIA6 inserts a protruding β hairpin and loops from subunit A (β2-β3 hairpin: 119–133, L8 loop: 142–148, L9 loop: 156–171), and α helices from subunit B (α5-α6: 74–86) within clefts at the surface of St1Cas9•sgRNA. A global structural comparison between unbound AcrIIA6 (Hynes et al., 2018) and AcrIIA6 bound to St1Cas9•sgRNA shows that the inhibitory protein does not undergo major conformational changes upon association, apart from a ~5 Å displacement of the L9 loop toward the sgRNA (Figure S3C).

The A subunit β2-β3 hairpin is inserted at the junction between the WED and the PI domains, where it forms an intermolecular hydrogen bonds network with St1Cas9 (A subunit β2-β3 hairpin: R120, Y122, A124, N127, Y128, St1Cas9: G943, Q987, K1010) (Figure 3C). Additionally, this β hairpin binds to the sugar-phosphate backbone of stem loop 1 (A subunit β2-β3 hairpin: T118, R120, Y128, R171; sgRNA: A69, G71, C72) (Figure 3C). L8 loop is involved in protein-protein interactions with St1Cas9 residues in the PI domain (A subunit L8 loop: G146-S147, St1Cas9: S1000, E1002) (Figure 3D). In contrast, L9 loop primarily binds to the repeat-antirepeat duplex and stem loop 1 of the sgRNA (A subunit β2-β3 hairpin: G167, R168, Q161, R171; sgRNA: A64, C72, U73) (Figure 3E). Lastly, AcrIIA6 subunit B provides several polar contacts at the interface with the PI domain (B subunit: N78, N81, D82, N83; St1Cas9: Q1001, K1008, D1115, K1116) (Figure 3F).

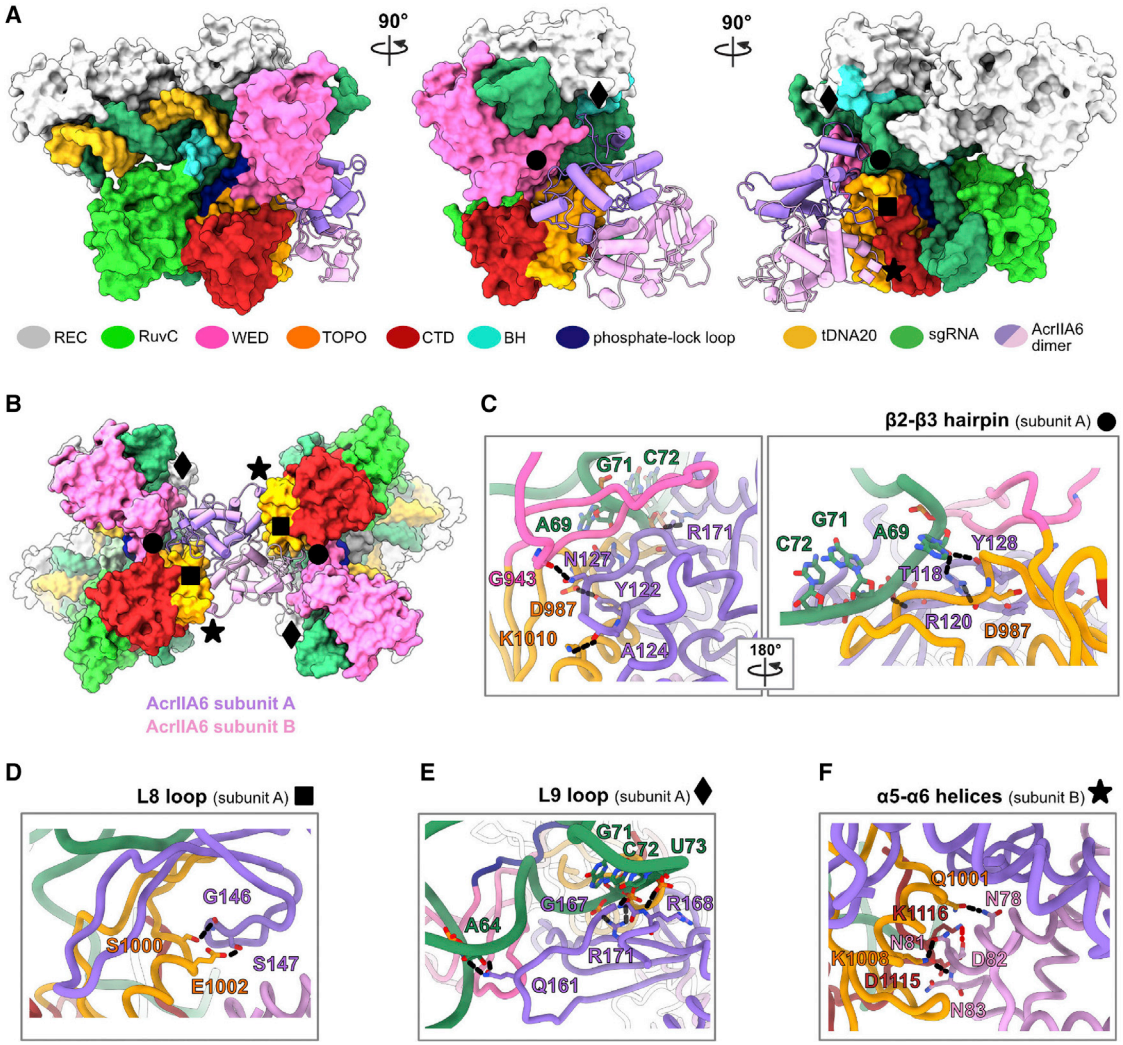


Figure 3. AcrIIA6 Binds to an Allosteric Site on St1Cas9

(A) Surface representations of the monomeric inhibition complex showing the AcrIIA6 dimer bound to an allosteric site at the back face of St1Cas9, relative to the target-binding crevasse and catalytic domains. Orthogonal views of the front (left), side (middle) and back (right) of the complex. The intermolecular interactions between St1Cas9•sgRNA and the β2-β3 hairpin (•), L8 loop (■) and L9 loop (◆) of AcrIIA6 subunit A, and α5-α6 helices (★) of AcrIIA6 subunit B are indicated.

(B) Surface representation of the symmetric dimeric inhibition complex. AcrIIA6 dimer contains two identical St1Cas9-binding sites, each one built by regions of both AcrIIA6 subunits A and B.

(C) Close-up views of the intermolecular interactions between St1Cas9•sgRNA and AcrIIA6 subunit A β2-β3 hairpin (•), L8 loop (■), L9 loop (◆), and subunit B α5-α6 helices (★). Hydrogen bonds are shown as black dotted lines, and the salt bridge is shown as a red dotted line.

See also Figure S4.

Collectively, the above data show that the structural determinants of AcrIIA6 association to St1Cas9 RNP are complex and distributed over a wide recognition interface composed of distinct St1Cas9 (WED and PI domains) and sgRNA (repeat-anti-repeat duplex and stem loop 1) regions and located at the back of the PAM binding site.

AcrIIA6 Reduces St1Cas9 Surveillance Complex DNA Binding Affinity and Blocks DNA Binding in Cells

In principle, AcrIIA6-mediated inhibition of St1Cas9 could result either from inhibition of DNA binding or from inhibition of the nuclease activity. The binding of AcrIIA6 to neither of such St1Cas9 sites and its location in a separate region does not provide direct structural information on its mode of action. To further characterize the molecular mechanism of AcrIIA6, we set up an *in vitro* DNA cleavage assay to follow the inhibitory effect of AcrIIA6 on St1Cas9 enzymatic activity. We concomitantly incubated St1Cas9 RNP with AcrIIA6 and 1500-bp-long dsDNA substrates, containing target and PAM sequences. The dsDNA cleavage products were resolved by electrophoresis at different time points. As expected, St1Cas9 RNP showed specific cleavage of target DNA in the absence of AcrIIA6, but its nuclease activity was abolished in the presence of AcrIIA6 (Figure 4A). DNA cleavage was not affected by the control

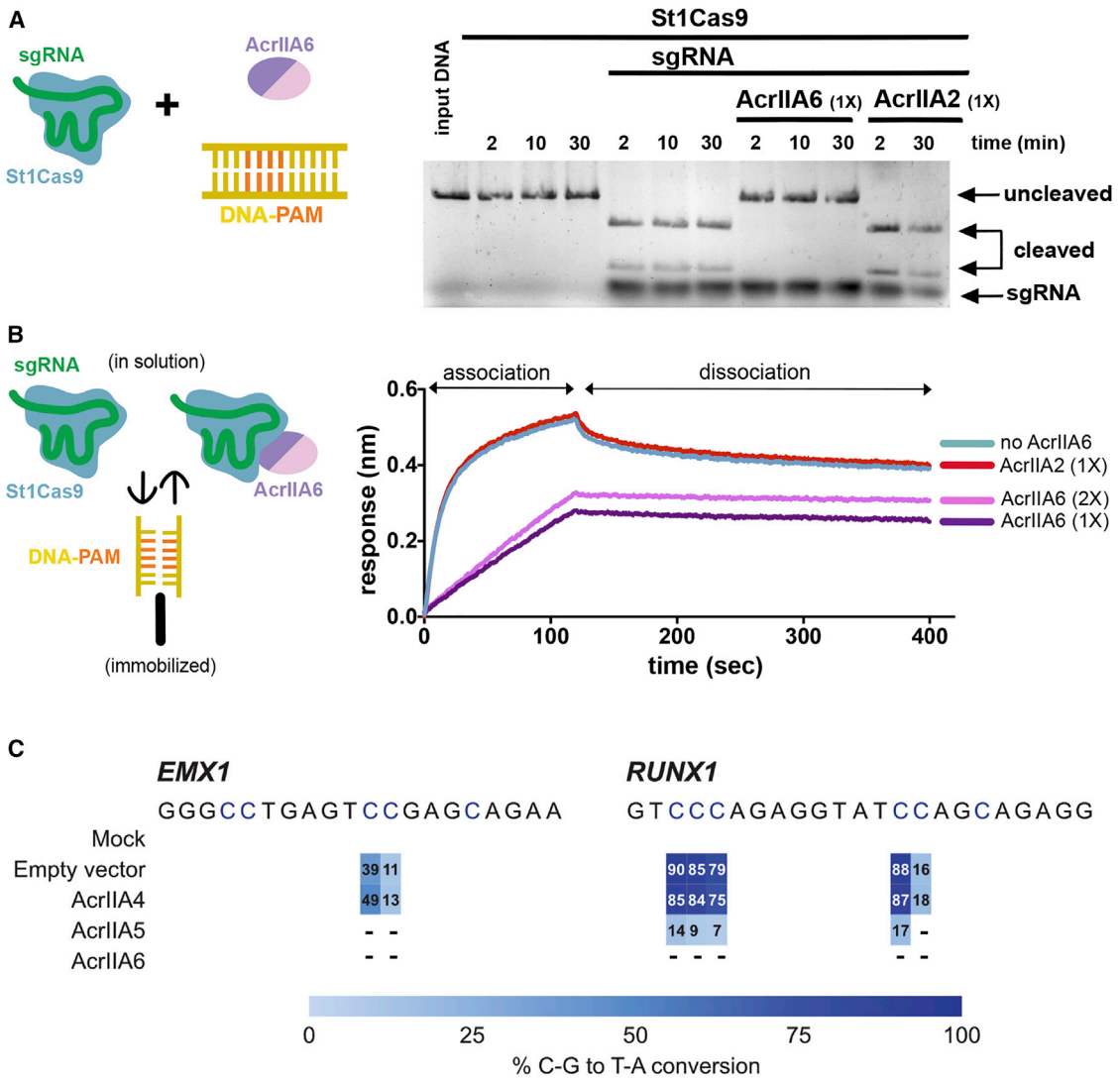


Figure 4. AcrIIA6 Reduces St1Cas9 Surveillance Complex DNA Binding Affinity and Blocks DNA Binding in Cells

(A) Time-course cleavage assay using dsDNA fragments containing a target sequence and a PAM motif. St1Cas9 and AcrIIAs were mixed with a 1:1 molar ratio (1×). Mobilities of input DNA (uncleaved) and cleavage products (cleaved) are indicated with arrows. AcrIIA6 inhibits St1Cas9•sgRNA-mediated DNA cleavage. The negative control AcrIIA2 does not affect St1Cas9 activity. Experiments were performed three times and yielded equivalent results. One experiment is shown.

(B) Interactions between dsDNA fragments (59-bp molecules containing the target and PAM sequences) and pre-formed St1Cas9 RNP or AcrIIAs-bound St1Cas9 RNP were monitored using BLi. St1Cas9 and AcrIIAs were mixed with 1:1 (1×) or 1:2 (2×) molar ratios. AcrIIA6 markedly reduces the St1Cas9 RNP DNA binding affinity. Experiments were performed three times and yielded equivalent results. One experiment is shown.

(C) Quantification of base editing mediated by St1BE4max LMD-9 when co-expressed in presence of the indicated AcrIIAs in K562 cells. Empty pVAX backbone was used as a negative control. Guide sequences are shown with target cytosines highlighted in blue. Dashes indicate that base editing was not detected. Experiments were performed twice for each guide and AcrIIA combinations and yielded equivalent results. One experiment is shown.

See also Figure S5.

AcrIIA2, which had previously been shown to be inactive against St1Cas9 *in vivo* (Hynes et al., 2018).

Next, we used BLi to monitor real-time association and dissociation between immobilized dsDNA fragments, containing target and PAM sequences, and St1Cas9 RNP, alone or bound to AcrIIA6. We ran these experiments in the absence of divalent cations and in the presence of EDTA, to prevent target cleavage. Interestingly, St1Cas9 RNP could still bind to DNA in the presence of AcrIIA6, but with markedly reduced affinity as compared to the experiment in the absence of AcrIIA6 (Figure 4B). Comparable binding curves were observed for St1Cas9 RNP mixed with either an equimolar amount or a two-fold molar excess of AcrIIA6, which rules out the possibility that decreased DNA binding be caused by less inhibitor-free St1Cas9 RNP moieties available. Such results confirm that the observed reduction of St1Cas9 RNP binding affinity to its target DNA is due to the inhibitory action of AcrIIA6. As a negative control, AcrIIA2 did not affect St1Cas9 RNP DNA binding properties. We then

conclude that AcrIIA6 decreases the St1Cas9 RNP affinity for DNA, an observation in agreement with the negative allosteric action of AcrIIA6 on St1Cas9 activity proposed above. In the context of infected cells, in which the concentration of DNA substrates is necessarily lower than in biochemical assays, the decreased St1Cas9 RNP DNA binding affinity may prevent target binding.

To further explore the functional impact of such decreased DNA-binding affinity *in vivo*, we turned to a base editing system that provides a sensitive readout of DNA binding. We adapted the architecture of a previously described DNA base editor (Koblan et al., 2018) and created a fusion between catalytically impaired St1Cas9 (St1Cas9-D9A) and a cytidine deaminase in order to convert C·G base pairs to T·A in human cells, referred to as St1BE4max (Agudelo et al., 2019). While St1BE4max efficiently converted target cytosines into thymines at two distinct loci, co-expression of AcrIIA6 resulted in complete inhibition of base editing (Figures 4C and S5A). To corroborate these findings, we used a transcription reporter system that consists of the catalytically inactive dead St1Cas9 (dSt1Cas9) fused to an artificial activation domain (VPR) and targeted to a minimal promoter driving the expression of a fluorescent reporter (Chavez et al., 2015). While dSt1Cas9-VPR strongly activated transcription, co-expression of AcrIIA6 resulted in complete inhibition of the reporter activity (Figure S5B). AcrIIA6 had no impact on dSpCas9-VPR in this system (Figure S5B). Taken together, the above data show that AcrIIA6 prevents the binding of St1Cas9 surveillance complex to their targets within cells through reduction of DNA binding affinity.

AcrIIA6 Alters St1Cas9 Conformational Dynamics Associated with PAM Binding

Having established that AcrIIA6 affects St1Cas9 target DNA binding, we then sought to characterize the associated molecular inhibition mechanisms. Since we showed that AcrIIA6 binds to an allosteric site on St1Cas9 RNP, proximal to the PAM binding site, we investigated the structural events linked to St1Cas9 RNP•AcrIIA6 PAM binding using cryo-EM. We vitrified and analyzed at the electron microscope a sample consisting of the pre-formed St1Cas9•sgRNA•AcrIIA6 complex that was mixed with a partial duplex composed of a 59-nt target DNA strand (tDNA59), and a 23-nt PAM-containing non-target strand (ntPAM) (Figures 5A and S6A; Table 1). The use of this partial double-stranded DNA molecule enabled us to focus our analysis on the effect of AcrIIA6 on PAM recognition, the primary event of St1Cas9 RNP DNA binding. Furthermore, through this complex, which is not functionally relevant *in vivo* as we showed that AcrIIA6 prevents DNA binding within cells (Figures 4C and S5), we specifically aimed to gain insight into the structural details mediating allosteric communication between the AcrIIA6 and PAM binding sites. Reference-free 2D classification of single particles revealed the presence of two particle populations: a major monomeric (97.5% of selected particles) and a minor dimeric (2.5% of selected particles) species, both being morphologically similar to the previously described AcrIIA6-bound St1Cas9 RNP monomeric and dimeric assemblies, respectively. Interestingly, 2D and 3D classification of the monomeric population allowed us to identify and distinguish a class of particles in which AcrIIA6 was bound (53.4% of selected particles) from a class (44.1% of selected particles) in which AcrIIA6 density was absent (Figures 5A, S6B, S6C, and S7A). Refinement of these two monomeric classes led to final 3D reconstructions of St1Cas9•sgRNA•AcrIIA6•tDNA59-ntPAM and St1Cas9•sgRNA•tDNA59-ntPAM to overall resolutions of 3.2 and 3.3 Å, respectively (Figures 5B, 5C, and S6D–S6I). As previously observed, densities of St1Cas9 (except for HNH and RuvC domains), AcrIIA6, sgRNA, and tDNA59-ntPAM were clearly defined in the cryo-EM maps, allowing building and refinement of the two models (Figures 5B, 5C, and S6J; Table 1).

A detailed structural comparison between St1Cas9•sgRNA•tDNA20•AcrIIA6 and St1Cas9•sgRNA•AcrIIA6•tDNA59-ntPAM shows modest perturbation of the overall St1Cas9 structure (RMSD of 1.3 Å for 803 equivalent Cα atoms; notably, the RMSD was 0.7 Å for the AcrIIA6-bound St1Cas9 structures reported above, in the absence of the PAM duplex) (Figure 5D). More specifically, however, St1Cas9 backbone shifts of 1.9–2.7 Å are observed in the PI and WED domains, related to accommodation of the 5′-GCAGAAA-3′-containing PAM duplex that is sandwiched between the WED and PI domains and covers ~1,040 $Å^2$ of St1Cas9 surface area (Figure 5D). The phosphate backbone of the target and non-target strands hydrogen bond with residues in the WED and PI domains. Additionally, the nucleobase dG4 hydrogen bonds with K1086 in the PI domain, providing a structural basis for requirement of the fourth dG nucleotide in the St1Cas9 PAM consensus sequence 5′-NNAGAAW-3′ (Figure S7B) (Deveau et al., 2008). Moreover, in the AcrIIA6 and PAM-bound complex (St1Cas9•sgRNA•AcrIIA6•tDNA59-ntPAM), the AcrIIA6 dimer is shifted away from St1Cas9, more evidently at the St1Cas9-distal end (protein backbone shifts of 1.9–3.5 Å) (Figure 5D).

These structural data show that PAM binding is mirrored by structural rearrangements of the WED and PI domains and of AcrIIA6, in agreement with our proposed allosteric linkage between AcrIIA6 binding and St1Cas9 functional sites. The observed structural effects, which may underline local changes in protein dynamics, also provide a rationale for the coexistence of AcrIIA6-free and AcrIIA6-bound St1Cas9•sgRNA•tDNA59-ntPAM assemblies. In fact, the two species would result from PAM-induced St1Cas9 conformational and dynamics changes, which are transmitted to the AcrIIA6 binding site and vary the affinity for the allosteric inhibitor. In line with such considerations, we demonstrate that AcrIIA6 binds to pre-formed St1Cas9•sgRNA•tDNA59-ntPAM assemblies with a slower association rate, a faster dissociation rate, and overall decreased binding affinity, as compared with the binding properties of AcrIIA6 to St1Cas9 RNP (Figure S7C).

Furthermore, a direct comparison of the PAM-bound 3D structures (St1Cas9•sgRNA•tDNA59-ntPAM and St1Cas9•sgRNA•AcrIIA6•tDNA59-ntPAM) (St1Cas9: RMSD of 0.8 Å for 813 equivalent Cα atoms) reveals localized conformational changes in the WED domain (Figure 5E). In the presence of AcrIIA6, this domain is pushed ~1.8 Å away from the PAM duplex that is shifted by 1.5–2.0 Å, considering that the RNA-DNA heteroduplexes perfectly superimpose onto each other (Figure 5E). The St1Cas9•sgRNA•AcrIIA6•tDNA59-ntPAM

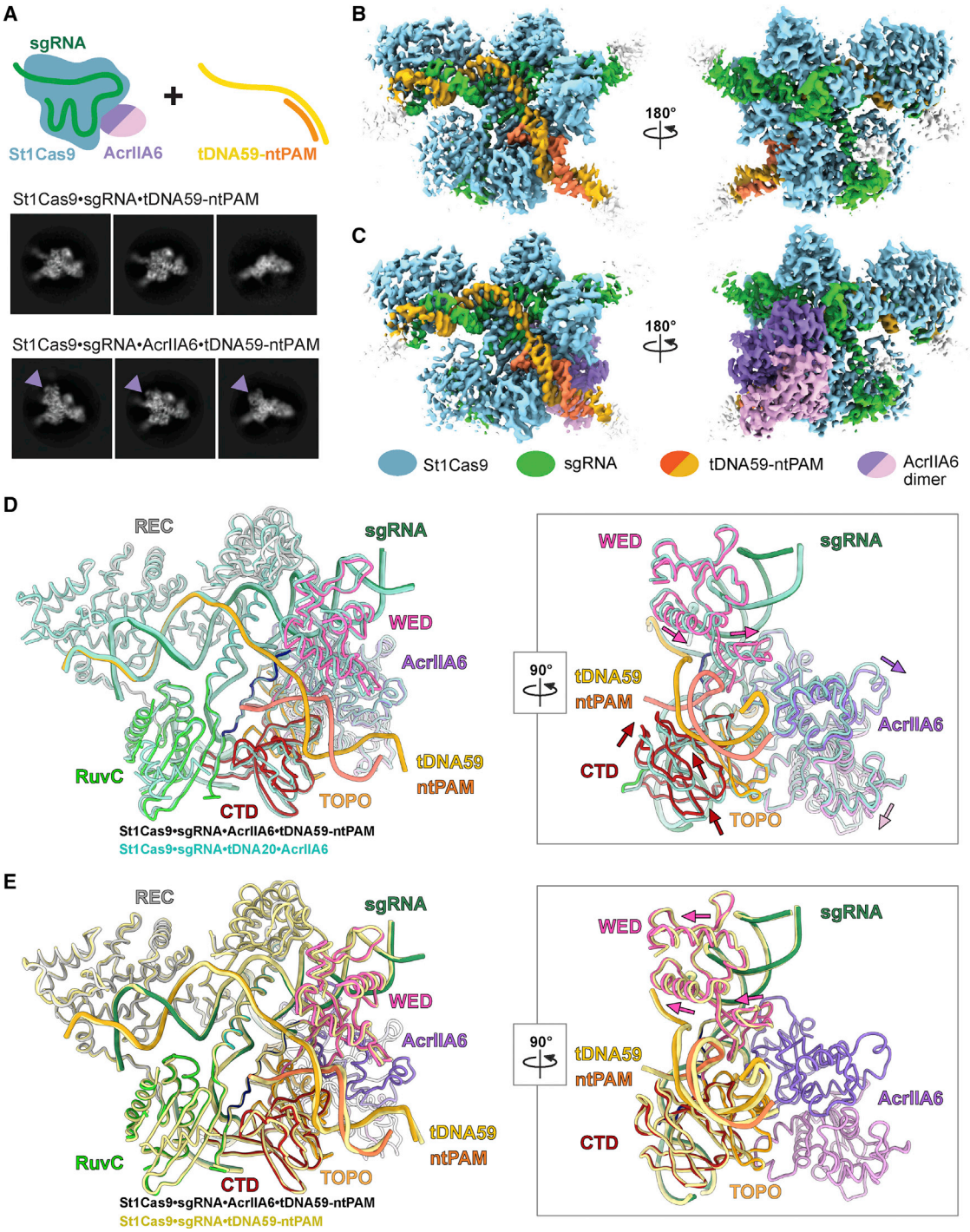


Figure 5. AcrIIA6 Alters St1Cas9 Conformational Dynamics Associated with PAM Binding

(A) Representative cryo-EM 2D class averages showing the absence or presence of AcrIIA6 (marked by purple arrowhead) (box size: 263 Å × 263 Å).

(B and C) Orthogonal views of cryo-EM 3D reconstructions of St1Cas9•sgRNA•tDNA59-ntPAM complex (B) and its AcrIIA6-bound form (C).

(D) (Left) Ribbon representation of the 3D structures of St1Cas9•sgRNA•AcrIIA6•tDNA59-ntPAM (color-coded) superimposed onto the 3D structure of St1Cas9•sgRNA•tDNA20•AcrIIA6 (light blue). (Right) Close-up view of the PAM binding site. The arrows indicate the PAM binding-induced shifts in the St1Cas9 WED and PI domains and in the AcrIIA6 dimer.

(E) (Left) Ribbon representation of the 3D structures of St1Cas9•sgRNA•AcrIIA6•tDNA59-ntPAM (color-coded) superimposed onto the 3D structure of of St1Cas9•sgRNA•tDNA59-ntPAM (khaki). (Right) Close-up view of the PAM binding site. The arrows indicate the movement of the St1Cas WED domain in the AcrIIA6-free St1Cas9•sgRNA•tDNA59-ntPAM complex.

See also Figures S6 and S7.

complex, thus, would represent an energetically unfavorable state in which the AcrIIA6 inhibitor constrains a lower affinity conformation of the PAM binding site.

As a whole, all the above data are coherent with a picture whereby AcrIIA6 modifies the structure, and likely also the dynamics of St1Cas9 PAM binding elements, while binding at a separate St1Cas9 region, as typical of an allosteric inhibitor.

A Natural Variant of St1Cas9 Is Resistant to AcrIIA6 Inhibition

Having previously shown that AcrIIA6 blocks the activity of St1Cas9s from *S. thermophilus* strains DGCC7710 and LMD-9, which differ for two amino acids only in the REC lobe, and recognize the same NNAGAAW PAMs (Hynes et al., 2018), we set up to test whether AcrIIA6 would inhibit other St1Cas9 variants from different *S. thermophilus* strains.

Based on sequences analyses (Bolotin et al., 2004), and at the light of the structural results reported here, St1Cas9s from strains LMG 18311 and CNRZ 1066 showed residue differences in the regions shown to mediate AcrIIA6 recognition. To explore this issue, we engineered and tested hybrid proteins containing the N-terminal domain of St1Cas9 LMD-9 (REC lobe, HNH and RuvC nuclease domains, and phosphate lock loop) and the C-terminal domains of St1Cas9 LMG 18311 or CNRZ 1066 (WED and PI domains), as this is the main interaction interface with AcrIIA6 (Figure 6A). While AcrIIA6 potently inhibited St1Cas9s derived from the LMD-9 and CNRZ 1066 strains, the LMG 18311 fusion construct was largely protected from inhibition in human cells (Figure 6B). Based on our molecular models, three of the four amino acid substitutions detected in the TOPO domain of St1Cas9 LMG 18311 (K993, G1006, M1008 and E1010) directly involve contacts with AcrIIA6. In particular, residues E1010 and M1008 are held to disrupt hydrogen bonds with the AcrIIA6 β2-β3 hairpin and α5-α6 helices, respectively, while residue K993 would provide steric hindrance near the AcrIIA6 L8 loop (Figure 6C). Such capacity of the *S. thermophilus* strain LMG 18311 natural St1Cas9 variant to escape AcrIIA6 further validates the molecular determinants we described as crucial for the St1Cas9-AcrIIA6 interaction.

DISCUSSION

Our work presents cryo-EM structures of four St1Cas9 complexes (dimeric and monomeric St1Cas9•sgRNA•tDNA20•AcrIIA6 assemblies, at 3.0 Å and 3.2 Å resolution; St1Cas9•sgRNA•AcrIIA6•tDNA59-ntPAM, at 3.2 Å resolution; St1Cas9•sgRNA•tDNA59-ntPAM, at 3.3 Å resolution). We provide first insight into the structural and functional landscapes of an Acr from virulent streptococcal phages, highlighting a unique molecular mode of action for inactivation of a CRISPR-Cas system. Notably, we show that AcrIIA6 has the properties of an allosteric inhibitor that can simultaneously bind two St1Cas9 surveillance complexes while reducing their DNA binding affinity, preventing overall binding of the nuclease to the target DNA within cells. Furthermore, our work addresses the next step in the evolutionary arms race between bacteria and phages, regarding mechanisms evolved by bacteria to protect themselves against phages carrying Acrs. The natural St1Cas9 variant from strain LMG 18311, which contains amino acids substitution in the AcrIIA6 binding site and is resistant to AcrIIA6 inhibition, illustrates the Acr-driven molecular diversification of Cas9 enzymes, and qualifies as one of the primary candidates of an "anti-anti-CRISPR" rescue mechanism.

A Potent Viral Immunosuppressant

We showed that viral AcrIIA6 stably binds to bacterial St1Cas9 surveillance complexes—the predominant intracellular form in immune bacteria during a phage infection. The combination of (1) a large association interface, (2) shape complementarity of the interacting partners, and (3) an extensive network of polar interactions, provides the structural rationale for their highly stable association. Such strong complex makes AcrIIA6 a potent inhibitor that could block all St1Cas9 interference activity in infected bacterial cells, even when present at low concentrations. Recent studies described an altruistic cooperation between phages to neutralize CRISPR-Cas immunity whereby, at the level of phages and bacterial populations, Acrs do not totally protect their associated phage genome, but gradually immunosuppress the bacterial community under attack (Borges et al., 2018; Landsberger et al., 2018). Essentially, Acr molecules are accumulated within bacteria through multiple and failed phage infections to ultimately reach the concentration required for an effective inhibition of CRISPR-Cas, thereby enabling phage replication. As a result, the number of failed infections required for bacterial immunosuppression decreases with the Acr inhibitory strength.

AcrIIA6 forms a tight dimer that can sequester and inactivate two Cas9 molecules at a time, thus strategically lowering the critical-threshold number of infection events required for effective phage replication. This is the second example of an Acr, after AcrIIC3 from a *Neisseria meningitidis* prophage, which can induce Cas9 dimerization. Nevertheless, AcrIIC3 displays an entirely different 3D structure, binds to the Cas9 HNH domain, and induces a different dimerization mechanism (Harrington et al., 2017; Zhu et al., 2019). Such "killing two birds with one stone" strategy adopted by AcrIIA6 is particularly suitable for virulent *S. thermophilus* phages that replicate very rapidly and efficiently during milk fermentation (Tremblay and Moineau, 1999).

Cas9 Allosteric Inhibition

We demonstrated that AcrIIA6 blocks the interference activity of St1Cas9 by inhibiting target DNA binding in cells, an approach that, in general terms, appears to be the prevailing mode of action of the so-far characterized Acrs. Yet, AcrIIA6 uses a previously unobserved molecular strategy to block Cas9 activity by tightly binding to a St1Cas9•sgRNA site well separated from the DNA-binding crevasse and the nuclease centers. The inhibitory activity therefore cannot be ascribed to occupation of the DNA binding region or to competition with substrates for the nuclease active centers. Moreover, AcrIIA6 binding does not prevent PAM recognition, in contrast with what has been reported for AcrIIA4 and AcrIIA2 that compete with PAM binding (Jiang et al., 2019; Liu et al., 2019).

Cas9s are highly dynamic proteins whose DNA cleavage activity is governed by concerted motions of individual domains, allosteric communication, and conformational checkpoints

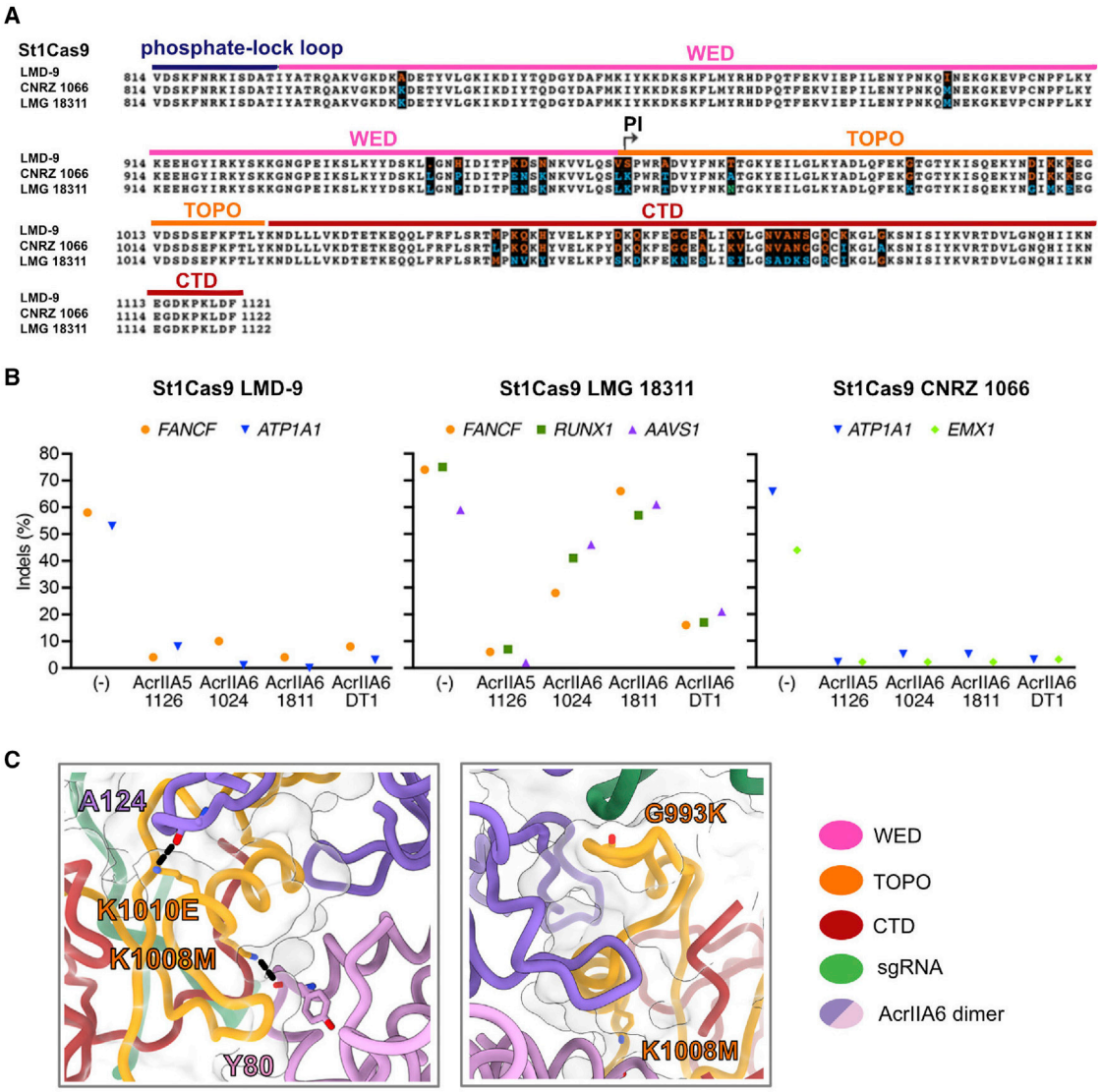


Figure 6. A Natural Variant of St1Cas9 Is Resistant to AcrIIA6 Inhibition

(A) Alignment of the C terminus of St1Cas9 proteins isolated from *S. thermophilus* strains LMD-9, CNRZ 1066, and LMG 18311.

(B) Indel frequencies mediated by the different St1Cas9s when co-expressed in presence of the indicated AcrIIAs in K562 cells. Indels were analyzed by TIDE assay. Empty pVAX backbone was used as a negative control. Experiments were performed twice for each guide, AcrIIA, and St1Cas9 combinations and yielded equivalent results. One experiment is shown.

(C) Surface and ribbon representations of the intermolecular contacts between St1Cas9 and AcrIIA6 that are impaired by the amino acid substitutions identified in the TOPO domain of St1Cas9 LMG 18311 (K1001E, K1008M, and G993K). E1010 and M1008 would disrupt hydrogen bonds (dotted lines) and introduce repulsive forces, while K993 would cause steric hindrance.

(Dagdas et al., 2017; Palermo et al., 2016, 2017; Sternberg et al., 2015; Yang et al., 2018). In particular, PAM binding is known to initiate R-loop formation and trigger Cas9-mediated target DNA cleavage (Palermo et al., 2017; Sternberg et al., 2014), being therefore a key step for Cas9 activation. At the light of such evidences and of our biochemical and structural results, we propose that AcrIIA6 exerts its inhibitory effect on PAM binding through allostery by constraining the dynamics and/or movements of the WED and PI domains, which form the main part of its binding interface. Such mechanism is in agreement with (1) our *in vitro* results showing that AcrIIA6 abolishes DNA cleavage and reduces St1Cas9 DNA binding affinity (PAM binding is indeed the first event in the sequential series leading to target DNA binding and catalytic activation), (2) our 3D structures revealing a relaxation of the PAM binding interface upon release of AcrIIA6, and (3) our interactions BLi experiment showing that PAM binding hampers AcrIIA6 recognition. The AcrIIA6 β2-β3 hairpin structural motif that inserts in between the WED and PI domains (Figure 3C), likely plays an important role in allostery by transducing the inhibitory action

to the PAM binding site. Of note, one of the amino acid substitutions in the AcrIIA6-resistant St1Cas9 natural variant, K1010E, occurs at the AcrIIA6 β2-β3 hairpin interface.

Insights into Molecular Coevolution between Acrs and CRISPR-Cas Systems

The constant battle between phages and their hosts steered the so-called arms race. While bacteria have evolved numerous defense mechanisms, including the CRISPR-Cas systems (Labrie et al., 2010), phages have fought back with countermeasures (Samson et al., 2013), such as Acrs. In this seemingly endless clash, co-evolution of such phage-host interactions is bound to happen. It was shown that overexpression of Cas proteins can inactivate Acrs through a saturation effect (Bondy-Denomy et al., 2015). The accumulation of multiple CRISPR-Cas types and subtypes within bacterial strains, as observed for *Pseudomonas aeruginosa* (van Belkum et al., 2015), forces phages to encode several Acrs (Pawluk et al., 2014) (Bondy-Denomy et al., 2018). Conversely, the diversity within the subtypes of CRISPR-Cas systems (Makarova et al., 2015) likely resulted from Acr-driven selective pressure to escape these widespread viral inhibitors. For example, a type I-F variant lacking the entire Cas8 homolog, and some domains of the Cas7 homolog that form AcrIF1 and AcrIF2 binding sites, has likely evolved to resist these Acrs (Pausch et al., 2017).

Here, we identified a natural variant of St1Cas9 from strain LMG 18311, which is not inhibited by AcrIIA6, and differs in the AcrIIA6-binding site just by four amino acids. Three of these residue substitutions abolish polar contacts and van der Waals contacts with the AcrIIA6 backbone, establishing an efficient molecular evolutionary strategy to bypass AcrIIA6 inhibition. These results allowed us to uncover and rationalize how some St1Cas9s have thus evolved to escape inhibition by specific AcrIIA6.

STAR★METHODS

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

SUPPLEMENTAL INFORMATION

Supplemental Information can be found online at https://doi.org/10.1016/j.molcel.2019.09.012.

ACKNOWLEDGMENTS

We thank Virginijus Siksnys for the plasmid for the recombinant expression of St1Cas9. A.G. acknowledges Christophe Quétard (Field Application Specialist at Molecular Devices) for his help with BLi experiments. A.G. acknowledges funding from the French National Research Agency (ANR-18-CE11-0016-01). This work was supported by the French Infrastructure for Integrated Structural Biology (FRISBI) (ANR-10-INSB-05-01). S.M. and Y.D. acknowledge funding from NSERC of Canada (Discovery program, RGPIN-2014-05132, RGPIN-2014-05698). S.M. holds a Tier 1 Canada Research Chair in Bacteriophages. Salary support to Y.D. was provided by the Fonds de la recherche du Québec-Santé (FRQS 254294). S.L. and D.A. hold Frederick Banting and Charles Best Canada Graduate and Vanier Canada graduate scholarship awards, respectively. A.D. holds a graduate training award from the Fonds de la recherche du Québec-Santé (FRQS). UCSF ChimeraX that was used for molecular graphics and analyses is developed by the Resource for Biocomputing, Visualization, and Informatics at the University of California, San Francisco, and receives support from NIH R01-GM129325 and P41-GM103311. Q5

AUTHOR CONTRIBUTIONS

Conceptualization, C.C., A.G.; Methodology, P.S., A.G., Y.D.; Validation, P.S., A.G., Y.D.; Formal Analysis, A.G., P.S., C.C.; Investigation, O.F., P.S., C.Z., B.A., S.S., A.G., A.C.-S., S.L., D.A., A.D., G.M.R., M.V.; Resources, P.S., A.G., A.R., M.B., Y.D.; Writing – Original Draft, A.G.; Writing – Review and Editing, A.G., C.C., Y.D., P.S., S.M., M.B.; Visualization, A.G., P.S.; Supervision, A.G., C.C., Y.D., S.M.; Project Administration, A.G.; Funding Acquisition, A.G., Y.D., S.M. All authors read and approved the final manuscript.

DECLARATION OF INTERESTS

G.M.R. and S.M. are co-authors on patent applications on AcrIIA6. The other authors declare no competing interests.

Received: May 28, 2019
Revised: August 2, 2019
Accepted: September 6, 2019
Published: October 8, 2019

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.

Afonine, P.V., Poon, B.K., Read, R.J., Sobolev, O.V., Terwilliger, T.C., Urzhumtsev, A., and Adams, P.D. (2018). Real-space refinement in PHENIX for cryo-EM and crystallography. Acta Crystallogr. D Struct. Biol. *74*, 531–544.

Agudelo, D., Duringer, A., Bozoyan, L., Huard, C.C., Carter, S., Loehr, J., Synodinou, D., Drouin, M., Salsman, J., Dellaire, G., et al. (2017). Marker-free coselection for CRISPR-driven genome editing in human cells. Nat. Methods *14*, 615–620.

Agudelo, D., Carter, S., Velimirovic, M., Duringer, A., Levesque, S., Rivest, J.F., Loehr, J., Mouchiroud, M., Cyr, D., Waters, P.J., et al. (2019). Versatile and robust genome editing with *Streptococcus thermophilus* CRISPR1-Cas9. bioRxiv. https://doi.org/10.1101/321208.

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.

Barrangou, R., Fremaux, C., Deveau, H., Richards, M., Boyaval, P., Moineau, S., Romero, D.A., and Horvath, P. (2007). CRISPR provides acquired resistance against viruses in prokaryotes. Science *315*, 1709–1712.

Bolotin, A., Quinquis, B., Renault, P., Sorokin, A., Ehrlich, S.D., Kulakauskas, S., Lapidus, A., Goltsman, E., Mazur, M., Pusch, G.D., et al. (2004). Complete sequence and comparative genome analysis of the dairy bacterium *Streptococcus thermophilus*. Nat. Biotechnol. *22*, 1554–1558.

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.

Bondy-Denomy, J., Davidson, A.R., Doudna, J.A., Fineran, P.C., Maxwell, K.L., Moineau, S., Peng, X., Sontheimer, E.J., and Wiedenheft, B. (2018). A unified resource for tracking anti-CRISPR names. CRISPR J *1*, 304–305.

Borges, A.L., Zhang, J.Y., Rollins, M.F., Osuna, B.A., Wiedenheft, B., and Bondy-Denomy, J. (2018). Bacteriophage cooperation suppresses CRISPR-Cas3 and Cas9 immunity. Cell *174*, 917–925.e10.

Briner, A.E., Donohoue, P.D., Gomaa, A.A., Selle, K., Slorach, E.M., Nye, C.H., Haurwitz, R.E., Beisel, C.L., May, A.P., and Barrangou, R. (2014). Guide RNA functional modules direct Cas9 activity and orthogonality. Mol. Cell *56*, 333–339.

Brinkman, E.K., Chen, T., Amendola, M., and van Steensel, B. (2014). Easy quantitative assessment of genome editing by sequence trace decomposition. Nucleic Acids Res. *42*, e168.

Chavez, A., Scheiman, J., Vora, S., Pruitt, B.W., Tuttle, M., P R Iyer, E., Lin, S., Kiani, S., Guzman, C.D., Wiegand, D.J., et al. (2015). Highly efficient Cas9-mediated transcriptional programming. Nat. Methods *12*, 326–328.

Chen, V.B., Arendall, W.B., 3rd, Headd, J.J., Keedy, D.A., Immormino, R.M., Kapral, G.J., Murray, L.W., Richardson, J.S., and Richardson, D.C. (2010). MolProbity: all-atom structure validation for macromolecular crystallography. Acta Crystallogr. D Biol. Crystallogr. *66*, 12–21.

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.

Cobián Güemes, A.G., Youle, M., Cantú, V.A., Felts, B., Nulton, J., and Rohwer, F. (2016). Viruses as winners in the game of life. Annu. Rev. Virol. *3*, 197–214.

Cong, L., Ran, F.A., Cox, D., Lin, S., Barretto, R., Habib, N., Hsu, P.D., Wu, X., Jiang, W., Marraffini, L.A., and Zhang, F. (2013). Multiplex genome engineering using CRISPR/Cas systems. Science *339*, 819–823.

Dagdas, Y.S., Chen, J.S., Sternberg, S.H., Doudna, J.A., and Yildiz, A. (2017). A conformational checkpoint between DNA binding and cleavage by CRISPR-Cas9. Sci. Adv. *3*, o0027.

Dalvai, M., Loehr, J., Jacquet, K., Huard, C.C., Roques, C., Herst, P., Côté, J., and Doyon, Y. (2015). A scalable genome-editing-based approach for mapping multiprotein complexes in human cells. Cell Rep. *13*, 621–633.

Deveau, H., Barrangou, R., Garneau, J.E., Labonté, J., Fremaux, C., Boyaval, P., Romero, D.A., Horvath, P., and Moineau, S. (2008). Phage response to CRISPR-encoded resistance in Streptococcus thermophilus. J. Bacteriol. *190*, 1390–1400.

Dong, D., Guo, M., Wang, S., Zhu, Y., Wang, S., Xiong, Z., Yang, J., Xu, Z., and Huang, Z. (2017). Structural basis of CRISPR-SpyCas9 inhibition by an anti-CRISPR protein. Nature *546*, 436–439.

Dong, L., Guan, X., Li, N., Zhang, F., Zhu, Y., Ren, K., Yu, L., Zhou, F., Han, Z., Gao, N., and Huang, Z. (2019). An anti-CRISPR protein disables type V Cas12a by acetylation. Nat. Struct. Mol. Biol. *26*, 308–314.

Emsley, P., Lohkamp, B., Scott, W.G., and Cowtan, K. (2010). Features and development of Coot. Acta Crystallogr. D Biol. Crystallogr. *66*, 486–501.

Esvelt, K.M., Mali, P., Braff, J.L., Moosburner, M., Yaung, S.J., and Church, G.M. (2013). Orthogonal Cas9 proteins for RNA-guided gene regulation and editing. Nat. Methods *10*, 1116–1121.

Garneau, J.E., Dupuis, M.E., Villion, M., Romero, D.A., Barrangou, R., Boyaval, P., Fremaux, C., Horvath, P., Magadán, A.H., and Moineau, S. (2010). The CRISPR/Cas bacterial immune system cleaves bacteriophage and plasmid DNA. Nature *468*, 67–71.

Goddard, T.D., Huang, C.C., Meng, E.C., Pettersen, E.F., Couch, G.S., Morris, J.H., and Ferrin, T.E. (2018). UCSF ChimeraX: Meeting modern challenges in visualization and analysis. Protein Sci. *27*, 14–25.

Guschin, D.Y., Waite, A.J., Katibah, G.E., Miller, J.C., Holmes, M.C., and Rebar, E.J. (2010). A rapid and general assay for monitoring endogenous gene modification. Methods Mol. Biol. *649*, 247–256.

Harrington, L.B., Doxzen, K.W., Ma, E., Liu, J.J., Knott, G.J., Edraki, A., Garcia, B., Amrani, N., Chen, J.S., Cofsky, J.C., et al. (2017). A broad-spectrum inhibitor of CRISPR-Cas9. Cell *170*, 1224–1233.e15.

Hirano, H., Gootenberg, J.S., Horii, T., Abudayyeh, O.O., Kimura, M., Hsu, P.D., Nakane, T., Ishitani, R., Hatada, I., Zhang, F., et al. (2016). Structure and engineering of *Francisella novicida* Cas9. Cell *164*, 950–961.

Holm, L., and Laakso, L.M. (2016). Dali server update. Nucleic Acids Res. *44* (W1), W351–W355.

Huai, C., Li, G., Yao, R., Zhang, Y., Cao, M., Kong, L., Jia, C., Yuan, H., Chen, H., Lu, D., and Huang, Q. (2017). Structural insights into DNA cleavage activation of CRISPR-Cas9 system. Nat. Commun. *8*, 1375.

Hwang, S., and Maxwell, K.L. (2019). Meet the anti-CRISPRs: widespread protein inhibitors of CRISPR-Cas systems. CRISPR J *2*, 23–30.

Hynes, A.P., Rousseau, G.M., Lemay, M.L., Horvath, P., Romero, D.A., Fremaux, C., and Moineau, S. (2017). An anti-CRISPR from a virulent streptococcal phage inhibits Streptococcus pyogenes Cas9. Nat. Microbiol. *2*, 1374–1380.

Hynes, A.P., Rousseau, G.M., Agudelo, D., Goulet, A., Amigues, B., Loehr, J., Romero, D.A., Fremaux, C., Horvath, P., Doyon, Y., et al. (2018). Widespread anti-CRISPR proteins in virulent bacteriophages inhibit a range of Cas9 proteins. Nat. Commun. *9*, 2919.

Jiang, F., Zhou, K., Ma, L., Gressel, S., and Doudna, J.A. (2015). STRUCTURAL BIOLOGY. A Cas9-guide RNA complex preorganized for target DNA recognition. Science *348*, 1477–1481.

Jiang, F., Taylor, D.W., Chen, J.S., Kornfeld, J.E., Zhou, K., Thompson, A.J., Nogales, E., and Doudna, J.A. (2016). Structures of a CRISPR-Cas9 R-loop complex primed for DNA cleavage. Science *351*, 867–871.

Jiang, F., Liu, J.J., Osuna, B.A., Xu, M., Berry, J.D., Rauch, B.J., Nogales, E., Bondy-Denomy, J., and Doudna, J.A. (2019). Temperature-responsive competitive inhibition of CRISPR-Cas9. Mol. Cell *73*, 601–610.e5.

Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J.A., and Charpentier, E. (2012). A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. Science *337*, 816–821.

Jinek, M., Jiang, F., Taylor, D.W., Sternberg, S.H., Kaya, E., Ma, E., Anders, C., Hauer, M., Zhou, K., Lin, S., et al. (2014). Structures of Cas9 endonucleases reveal RNA-mediated conformational activation. Science *343*, 1247997.

Josephs, E.A., Kocak, D.D., Fitzgibbon, C.J., McMenemy, J., Gersbach, C.A., and Marszalek, P.E. (2015). Structure and specificity of the RNA-guided endonuclease Cas9 during DNA interrogation, target binding and cleavage. Nucleic Acids Res. *43*, 8924–8941.

Karvelis, T., Gasiunas, G., Young, J., Bigelyte, G., Silanskas, A., Cigan, M., and Siksnys, V. (2015). Rapid characterization of CRISPR-Cas9 protospacer adjacent motif sequence elements. Genome Biol. *16*, 253.

Kelley, L.A., Mezulis, S., Yates, C.M., Wass, M.N., and Sternberg, M.J. (2015). The Phyre2 web portal for protein modeling, prediction and analysis. Nat. Protoc. *10*, 845–858.

Kluesner, M.G., Nedveck, D.A., Lahr, W.S., Garbe, J.R., Abrahante, J.E., Webber, B.R., and Moriarity, B.S. (2018). EditR: A Method to Quantify Base Editing from Sanger Sequencing. CRISPR J *1*, 239–250.

Koblan, L.W., Doman, J.L., Wilson, C., Levy, J.M., Tay, T., Newby, G.A., Maianti, J.P., Raguram, A., and Liu, D.R. (2018). Improving cytidine and adenine base editors by expression optimization and ancestral reconstruction. Nat. Biotechnol. *36*, 843–846.

Koonin, E.V., Makarova, K.S., and Zhang, F. (2017). Diversity, classification and evolution of CRISPR-Cas systems. Curr. Opin. Microbiol. *37*, 67–78.

Krissinel, E., and Henrick, K. (2007). Inference of macromolecular assemblies from crystalline state. J. Mol. Biol. *372*, 774–797.

Labrie, S.J., Samson, J.E., and Moineau, S. (2010). Bacteriophage resistance mechanisms. Nat. Rev. Microbiol. *8*, 317–327.

Landsberger, M., Gandon, S., Meaden, S., Rollie, C., Chevallereau, A., Chabas, H., Buckling, A., Westra, E.R., and van Houte, S. (2018). Anti-CRISPR phages cooperate to overcome CRISPR-Cas immunity. Cell *174*, 908–916.e12.

Liu, L., Yin, M., Wang, M., and Wang, Y. (2019). Phage AcrIIA2 DNA mimicry: Structural basis of the CRISPR and anti-CRISPR arms race. Mol. Cell *73*, 611–620.e3.

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.

Mekler, V., Minakhin, L., and Severinov, K. (2017). Mechanism of duplex DNA destabilization by RNA-guided Cas9 nuclease during target interrogation. Proc. Natl. Acad. Sci. USA *114*, 5443–5448.

Müller, M., Lee, C.M., Gasiunas, G., Davis, T.H., Cradick, T.J., Siksnys, V., Bao, G., Cathomen, T., and Mussolino, C. (2016). *Streptococcus thermophilus* CRISPR-Cas9 systems enable specific editing of the human genome. Mol. Ther. *24*, 636–644.

Nishimasu, H., Ran, F.A., Hsu, P.D., Konermann, S., Shehata, S.I., Dohmae, N., Ishitani, R., Zhang, F., and Nureki, O. (2014). Crystal structure of Cas9 in complex with guide RNA and target DNA. Cell *156*, 935–949.

Nishimasu, H., Cong, L., Yan, W.X., Ran, F.A., Zetsche, B., Li, Y., Kurabayashi, A., Ishitani, R., Zhang, F., and Nureki, O. (2015). Crystal structure of *Staphylococcus aureus* Cas9. Cell *162*, 1113–1126.

Palermo, G., Miao, Y., Walker, R.C., Jinek, M., and McCammon, J.A. (2016). Striking plasticity of CRISPR-Cas9 and key role of non-target DNA, as revealed by molecular simulations. ACS Cent Sci *2*, 756–763.

Palermo, G., Ricci, C.G., Fernando, A., Basak, R., Jinek, M., Rivalta, I., Batista, V.S., and McCammon, J.A. (2017). Protospacer Adjacent Motif-Induced Allostery Activates CRISPR-Cas9. J. Am. Chem. Soc. *139*, 16028–16031.

Pausch, P., Müller-Esparza, H., Gleditzsch, 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., Bondy-Denomy, J., Cheung, V.H., 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.

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

Ran, F.A., Cong, L., Yan, W.X., Scott, D.A., Gootenberg, J.S., Kriz, A.J., Zetsche, B., Shalem, O., Wu, X., Makarova, K.S., et al. (2015). In vivo genome editing using Staphylococcus aureus Cas9. Nature *520*, 186–191.

Rohou, A., and Grigorieff, N. (2015). CTFFIND4: Fast and accurate defocus estimation from electron micrographs. J. Struct. Biol. *192*, 216–221.

Rollins, M.F., Chowdhury, S., Carter, J., Golden, S.M., Miettinen, H.M., Santiago-Frangos, A., Faith, D., Lawrence, C.M., Lander, G.C., and Wiedenheft, B. (2019). Structure reveals a mechanism of CRISPR-RNA-guided nuclease recruitment and anti-CRISPR viral mimicry. Mol. Cell *74*, 132–142.e5.

Samson, J.E., Magadán, A.H., Sabri, M., and Moineau, S. (2013). Revenge of the phages: defeating bacterial defences. Nat. Rev. Microbiol. *11*, 675–687.

Semenova, E., Jore, M.M., Datsenko, K.A., Semenova, A., Westra, E.R., Wanner, B., van der Oost, J., Brouns, S.J., and Severinov, K. (2011). Interference by clustered regularly interspaced short palindromic repeat (CRISPR) RNA is governed by a seed sequence. Proc. Natl. Acad. Sci. USA *108*, 10098–10103.

Shin, J., Jiang, F., Liu, J.J., Bray, N.L., Rauch, B.J., Baik, S.H., Nogales, E., Bondy-Denomy, J., Corn, J.E., and Doudna, J.A. (2017). Disabling Cas9 by an anti-CRISPR DNA mimic. Sci. Adv. *3*, e1701620.

Sternberg, S.H., Redding, S., Jinek, M., Greene, E.C., and Doudna, J.A. (2014). DNA interrogation by the CRISPR RNA-guided endonuclease Cas9. Nature *507*, 62–67.

Sternberg, S.H., LaFrance, B., Kaplan, M., and Doudna, J.A. (2015). Conformational control of DNA target cleavage by CRISPR-Cas9. Nature *527*, 110–113.

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.

Tremblay, D.M., and Moineau, S. (1999). Complete genomic sequence of the lytic bacteriophage DT1 of Streptococcus thermophilus. Virology *255*, 63–76.

Uribe, R.V., van der Helm, E., Misiakou, M.A., Lee, S.W., Kol, S., and Sommer, M.O.A. (2019). Discovery and characterization of Cas9 inhibitors disseminated across seven bacterial phyla. Cell Host Microbe *25*, 233–241.e5.

van Belkum, A., Soriaga, L.B., LaFave, M.C., Akella, S., Veyrieras, J.B., Barbu, E.M., Shortridge, D., Blanc, B., Hannum, G., Zambardi, G., et al. (2015). Phylogenetic distribution of CRISPR-Cas systems in antibiotic-resistant *Pseudomonas aeruginosa*. MBio *6*, e01796–e15.

Wang, H., La Russa, M., and Qi, L.S. (2016a). CRISPR/Cas9 in genome editing and beyond. Annu. Rev. Biochem. *85*, 227–264.

Wang, J., Ma, J., Cheng, Z., Meng, X., You, L., Wang, M., Zhang, X., and Wang, Y. (2016b). 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. (2016c). Structural basis of Cas3 inhibition by the bacteriophage protein AcrF3. Nat. Struct. Mol. Biol. *23*, 868–870.

Wilkinson, R.A., Martin, C., Nemudryi, A.A., and Wiedenheft, B. (2019). CRISPR RNA-guided autonomous delivery of Cas9. Nat. Struct. Mol. Biol. *26*, 14–24.

Yamada, M., Watanabe, Y., Gootenberg, J.S., Hirano, H., Ran, F.A., Nakane, T., Ishitani, R., Zhang, F., Nishimasu, H., and Nureki, O. (2017). Crystal structure of the minimal Cas9 from *Campylobacter jejuni* reveals the molecular diversity in the CRISPR-Cas9 systems. Mol. Cell *65*, 1109–1121.e3.

Yang, H., and Patel, D.J. (2017). Inhibition mechanism of an anti-CRISPR suppressor AcrIIA4 targeting SpyCas9. Mol. Cell *67*, 117–127.e5.

Yang, M., Peng, S., Sun, R., Lin, J., Wang, N., and Chen, C. (2018). The conformational dynamics of Cas9 governing DNA cleavage are revealed by single-molecule FRET. Cell Rep. *22*, 372–382.

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.

Zhu, Y., Gao, A., Zhan, Q., Wang, Y., Feng, H., Liu, S., Gao, G., Serganov, A., and Gao, P. (2019). Diverse mechanisms of CRISPR-Cas9 inhibition by type IIC anti-CRISPR proteins. Mol. Cell *74*, 296–309.e7.

Zivanov, J., Nakane, T., Forsberg, B.O., Kimanius, D., Hagen, W.J., Lindahl, E., and Scheres, S.H. (2018). New tools for automated high-resolution cryo-EM structure determination in RELION-3. eLife *7*, e42166.

Zuris, J.A., Thompson, D.B., Shu, Y., Guilinger, J.P., Bessen, J.L., Hu, J.H., Maeder, M.L., Joung, J.K., Chen, Z.Y., and Liu, D.R. (2015). Cationic lipid-mediated delivery of proteins enables efficient protein-based genome editing in vitro and in vivo. Nat. Biotechnol. *33*, 73–80.

Please cite this article in press as: Fuchsbauer et al., Cas9 Allosteric Inhibition by the Anti-CRISPR Protein AcrIIA6, Molecular Cell (2019), https://doi.org/10.1016/j.molcel.2019.09.012

Q3 Q4 STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Bacterial and Virus Strains		
Escherichia coli BL21 Star DE3	Novagen	Cat#69450
Escherichia coli Rosetta 2 DE3 pLysS	Novagen	Cat#71401
Escherichia coli T7 Express	New England Biolabs	Cat#C2566H
Chemicals, Peptides, and Recombinant Proteins		
TEV protease	Homemade	N/A
T7 polymerase	Homemade	N/A
EZ-link NHS-PEG4-biotin	ThermoFischer Scientific	Cat#21362
proteinase K	Fischer Scientific	Cat#10259184
Taq polymerase	New England Biolabs	Cat#M0267L
RNasin	Promega	Cat#N2511
DNase I	Roche	Cat#11284932001
Biocytin	Sigma-Aldrich	Cat#B4261
Critical Commercial Assays		
QuickExtract DNA extraction solution	Lucigen	Cat#QE09050
Surveyor mutation detection kit	Integrated DNA technologies	Cat#706020
Deposited Data		
cryo-EM map of the St1Cas9●sgRNA●tDNA20●AcrIIA6 complex (monomeric assembly)	This paper	EMDB-4900
coordinates of the St1Cas9●sgRNA●tDNA20●AcrIIA6 complex (monomeric assembly)	This paper	PDB: 6RJ9
cryo-EM map of the St1Cas9●sgRNA●tDNA20●AcrIIA6 complex (dimeric assembly)	This paper	EMDB-4901
coordinates of the St1Cas9●sgRNA●tDNA20●AcrIIA6 complex (dimeric assembly)	This paper	PDB: 6RJA
cryo-EM map of the St1Cas9●sgRNA●AcrIIA6●tDNA59-ntPAM complex	This paper	EMDB-4904
coordinates of the St1Cas9●sgRNA●AcrIIA6●tDNA59-ntPAM complex	This paper	PDB: 6RJG
cryo-EM map of the St1Cas9●sgRNA● tDNA59-ntPAM complex	This paper	EMDB-4902
coordinates of the St1Cas9●sgRNA● tDNA59-ntPAM complex	This paper	PDB: 6RJD
coordinates of the AcrIIA6 crystal structure	Hynes et al., 2018	PDB: 6EYX
coordinates of the SaCas9 crystal structure	Nishimasu et al., 2015	PDB: 5AXW
Experimental Models: Cell Lines		
K-562	ATCC	Cat#CCL-243; RRID:CVCL_0004

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Oligonucleotides		
sgRNA (St1Cas9): CUUGCGUUGAUAAAAGUA UUGUUUUUGUACUCUCAAGAUUCAAUAAUC UUGCAGAAGCUACAAAGAUAAGGCUUCA UGCCGAAAUCAACACCCUGUCAUUUUAUG GCAGGGUGUUUU	This paper	N/A
sgRNA (SpCas9): UUAGAACAAAAUUCUUG UUUGUUUUAGAGCUAGAAAUAGCAAG UUAAAAUAAGGCUAGUCCGUUAUCAAC UUGAAAAAGUGGCACCGAGUCGGUGCUU	This paper	N/A
tDNA20: AATACTTTTATCAACGCAAG	This paper	N/A
tDNA59: ATGGTTCTGGTTTCATTTTCTGCAA TACTTTTATCAACGCAAGAGGTGCTTCT GTTATG	This paper	N/A
ntPAM: GCAGAAAATGAAACCAGAACCAT	This paper	N/A
tDNA20-PAM sequence inserted into the pUC19 plasmid: TTTCTGCAATACTTTTA TCAACGCAAG	This paper	N/A
biotin-tDNA59: ATGGTTCTGGTTTCATTTTC TGCAATACTTTTATCAACGCAAGAGGTGC TTCTGTTATG	This paper	N/A
ntDNA59: CATAACAGAAGCACCTCTTGCG TTGATAAAAGTATTGCAGAAAA TGAAACCAGAACCAT	This paper	N/A
Recombinant DNA		
pBAD24-CHis-St1Cas9	Provided by Virginijus Siksnys, Karvelis et al., 2015	N/A
pET-Cas9-6xHis	Zuris et al., 2015	RRID:Addgene #62374
pETG-20A-AcrIIA6	Hynes et al., 2018	N/A
pET-28a(+)-AcrIIA2	This paper	N/A
MSP1594_2x_NLS	Agudelo et al., 2019	RRID:Addgene #110625
AAVS1_Puro_PGK1_3xFLAG_Twin_Strep	Dalvai et al., 2015	RRID:Addgene #68375
St1Cas9_LMD-9_sgRNA_pUC19	Agudelo et al., 2019	RRID:Addgene #110627
U6_sgRNA_CAG_hSt1Cas9_LMD-9	Agudelo et al., 2019	RRID:Addgene #110626
M_ST1n_VPR	Chavez et al., 2015	RRID:Addgene #63799
M-tdTom-ST1	Esvelt et al., 2013	RRID:Addgene #48678
M-ST1-sgRNA	Esvelt et al., 2013	RRID:Addgene #48672
SP-dCas9-VPR	Chavez et al., 2015	RRID:Addgene #63798
M-tdTom-SP	Esvelt et al., 2013	RRID:Addgene #48677
M-SP-sgRNA	Esvelt et al., 2013	RRID:Addgene #48671
pCMV_BE4max_3xHA	Koblan et al., 2018	RRID:Addgene #112096
Software and Algorithms		
FortéBio Data Analysis software 8.2	Molecular Devices	N/A
Prism8	GraphPad	N/A
Relion-3.0	Zivanov et al., 2018	https://www3.mrc-lmb.cam.ac.uk/relion/index.php/Main_Page
MotionCor2 (version 1.2.1)	Zheng et al., 2017	https://hpc.nih.gov/apps/MotionCor2.html
CTFFIND4 (version 4.1.10)	Rohou and Grigorieff, 2015	http://grigoriefflab.janelia.org/ctffind4
Gautomatch (version 0.56)	written by Dr. Kai Zhang	https://www.mrc-lmb.cam.ac.uk/kzhang/
PHYRE2	Kelley et al., 2015	http://www.sbg.bio.ic.ac.uk/phyre2/html/page.cgi?id=index

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
UCSF Chimera	Pettersen et al., 2004	https://www.cgl.ucsf.edu/chimera/
PHENIX	Adams et al., 2010	https://www.phenix-online.org/
Coot	Emsley et al., 2010	https://www2.mrc-lmb.cam.ac.uk/personal/pemsley/coot/
Molprobity	Chen et al., 2010	http://molprobity.biochem.duke.edu/
DALI	Holm and Laakso, 2016	http://ekhidna2.biocenter.helsinki.fi/dali/
PISA	Krissinel and Henrick, 2007	http://www.ebi.ac.uk/msd-srv/prot_int/
UCSF Chimera X	Goddard et al., 2018	https://www.rbvi.ucsf.edu/chimerax/
EditR	Kluesner et al., 2018	http://baseeditr.com/
Other		
Biosensor streptavidin (SA)	Molecular Devices	Cat#18-5021

LEAD CONTACT AND MATERIAL AVAILABILITY

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Adeline Goulet (adeline.goulet@afmb.univ-mrs.fr).

EXPERIMENTAL MODELS AND SUBJECT DETAILS

Escherichia coli strains

E. coli BL21 Star (DE3) cells (Novagen) were used to produce St1Cas9 and SpCas9 for *in vitro* analysis and cryo-EM. *E. coli* Rosetta 2 (DE3) pLysS (Novagen) were used to produce AcrIIA6 for *in vitro* analysis and cryo-EM. *E. coli* T7 Express cells (New England Biolabs) were used to produce AcrIIA2 for *in vitro* analysis. Cells were grown in Terrific Broth (TB) or NZY auto-induction (NZYtech) medium supplemented with appropriate antibiotics, at 37°C with shaking.

Cell culture and transfection

K562 cells were obtained from the ATCC (CCL-243) and maintained at 37°C under 5% CO_2 in RPMI medium supplemented with 10% FBS, penicillin-streptomycin and GlutaMAX, and were periodically tested for absence of mycoplasma contamination. Cells (2E5 per transfection) were transfected using the Amaxa 4D-Nucleofector (Lonza), per manufacturer's recommendations. The K562 cell line expressing St1Cas9 from the *AAVS1* safe harbor locus was generated as described previously (Agudelo et al., 2017; Dalvai et al., 2015). Briefly, simultaneous selection and cloning was performed for 10 days in methylcellulose-based semi-solid RPMI medium, supplemented with 0.5 μg/mL puromycin, starting 3 days post-transfection. Clones were picked and expanded in 96-wells plates for 3 days, and transferred to 12-well plates for another 3 days before cells were harvested for western blot analysis.

METHOD DETAILS

Plasmids preparation for protein expression in *E. coli*

Plasmids expressing St1Cas9 of *S. thermophilus* DGCC7710 (pBAD24-CHis-St1Cas9, a gift from Virginijus Siksnys), SpCas9 (pET-Cas9-6xHis, Addgene #62374, a gift from David Liu), and AcrIIA6 of phage D1811 (pETG20A-AcrIIA6) were previously described (Hynes et al., 2018; Karvelis et al., 2015; Zuris et al., 2015). The DNA sequence encoding for AcrIIA2 (GenBank: AEO04363.1) was synthesized (GenScript) and cloned into a pET-28a (+) vector. The final construct encoded for AcrIIA2 with a N-terminal hexahistidine tag (6xHis) and a TEV protease recognition site.

Production and purification of anti-CRISPR proteins

AcrIIA6 was produced and purified as described previously (Hynes et al., 2018). AcrIIA2 was produced in *E. coli* T7 Express cells by the addition of 1 mM isopropyl β-D-1-thiogalactopyranoside (IPTG) and incubation for 18 hours at 18°C. Cells were harvested by centrifugation (6,000 x g, 15 min, 4°C), resuspended in buffer A (50 mM Tris pH 7.5, 300 mM NaCl, 10 mM imidazole) supplemented with 1 mM phenylmethylsulfonyl fluoride (PMSF) and 0.5 mg/mL lysozyme, and disrupted by French Press. The lysate was cleared by centrifugation (13,000 x g, 30 min, 4°C), and the supernatant was applied onto a 5-mL HisTrap FF column (GE Healthcare) for Ni-NTA affinity chromatography. The immobilized proteins were eluted in buffer B (50 mM Tris pH 7.5, 300 mM NaCl, 250 mM imidazole). Cleavage of the 6xHis tag was performed with 2 mg of homemade TEV protease for 12 hours at 4°C. The untagged AcrII2 was recovered in the flow-through of a HisTrap column, while the 6xHis-tagged TEV protease and uncleaved proteins remained bound to the column. AcrIIA2 was further purified by size exclusion chromatography using a HiLoad 16/600 Superdex 75 column (GE Healthcare)

equilibrated in buffer C (20 mM HEPES pH 7.5, NaCl 150 mM). The AcrIIA2 peak was concentrated to 6 mg/mL using Amicon Ultra Centrifugal Filters, flash-frozen in liquid nitrogen, and stored at −80°C.

Production and purification of St1Cas9

St1Cas9 was produced in *E. coli* BL21 Star (DE3) by the addition of 0.2% (w/v) arabinose and incubation for 18 hours at 25°C. Cells were harvested by centrifugation (6,000 x g, 15 min, 4°C), resuspended in buffer A supplemented with 5 mM 2-mercaptoethanol, 1 mM PMSF and 0.5 mg/mL lysozyme, and disrupted by sonication. The lysate was cleared by centrifugation (13, 000 x g, 30 min, 4°C), and the supernatant was applied onto a 5-mL HisTrap FF column. After a wash step with 4 column volumes (CV) of a buffer containing 50 mM Tris pH 7.5, 2M NaCl, the immobilized proteins were eluted in buffer B supplemented with 5mM 2-mercaptoethanol and 1 mM PMSF. Next, the St1Cas9-containing fractions were buffer exchanged in buffer D (20 mM HEPES pH 7, 100 mM NaCl, 5 mM 2-mercaptoethanol) and applied onto a 1mL Resource S column (GE Healthcare) for cation exchange chromatography. St1Cas9 was eluted by a linear salt gradient in 20 mM HEPES pH 7, 320 mM NaCl. St1Cas9 containing fractions were pooled and concentrated to 1.6 mg/mL using Amicon Ultra Centrifugal Filters, flash-frozen in liquid nitrogen, and stored at −80°C.

Production and purification of SpCas9

SpCas9 was produced in *E. coli* BL21 Star (DE3) by the addition of 0.5 mM IPTG and incubation for 18 hours at 25°C. Cells were harvested by centrifugation (6,000 x g, 15 min, 4°C), resuspended in buffer A containing 5 mM 2-mercaptoethanol, 1 mM PMSF and 0.5 mg/mL lysozyme, and disrupted by sonication. The lysate was cleared by centrifugation (17, 000 x g, 30 min, 4°C), and the supernatant was applied onto a 5-mL HisTrap FF column. The immobilized proteins were eluted in buffer B containing 5mM 2-mercaptoethanol and 1 mM PMSF. SpCas9 was further purified by size exclusion chromatography using a HiLoad 26/600 Superdex 200 column (GE Healthcare) equilibrated in buffer C containing 1 mM Tris(2-carboxyethyl)phosphine (TCEP). The SpCas9 peak was concentrated to 5 mg/mL using Amicon Ultra Centrifugal Filters, flash-frozen in liquid nitrogen, and stored at −80°C.

In vitro transcription and purification of sgRNAs

St1Cas9 and SpCas9-specific sgRNAs were transcribed *in vitro* using linear dsDNA templates that were generated by PCR from pUC19 plasmids containing the sgRNA-encoding sequences downstream a T7 promoter. Transcription reactions were performed by mixing 40 μg of DNA template with 0.7 mg of homemade T7 polymerase at 37°C for 16 hours, in a buffer containing 50 mM Tris pH 8, 6 mM $MgCl_2$, 10 mM DTT, 0.01% Triton X-100, 2 mM spermidine, 10 mM NTPs, and RNasin (Promega) After digestion with DNase I, sgRNAs were purified by denaturing urea polyacrylamide gel electrophoresis (acrylamide 12%, urea 8M, TBE 1X). sgRNAs were gel-extracted in a buffer containing 500 mM sodium acetate pH5.6, 0.1 mM EDTA, 0.1% SDS for 18 hours at 4°C, ethanol precipitated, and resuspended in RNase-free water for storage at −80°C. The first nucleotide of the sgRNAs theoretical sequence was substituted into a G nucleotide in the *in vitro* transcribed sgRNAs to optimize their production (theoretical sequence of the St1Cas9-specific sgRNA: 5′-CUUGCGUUGAUAAAAGUAUUGUUUUUGUACUCUCAAGAUUCAAUAAUCUUGCAGAAG CUACAAAGAUAAGGCUUCAUGCCGAAAUCAACACCCUGUCAUUUUAUG GCAGGGUGUUUU-3′; theoretical sequence of the SpCas9-specific sgRNA: 5′-UUAGAACAAAAUUCUUGUUUGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUA UCAACUUGAAAAAGUGGCACCGAGUCGGUGCUU-3′).

Analysis of the interactions between AcrIIA6 and Cas9 using BioLayer Interferometry

Streptavidin-coated biosensors (Molecular Devices) were used to immobilize biotinylated AcrIIA6 and monitor its association and dissociation with different Cas9 molecular species. AcrIIA6 was biotinylated using the EZ-link NHS-PEG4-biotin (ThermoFischer Scientific). AcrIIA6 concentrated at 1 mg/mL in buffer C was incubated with 1 mM NHS-PEG4-biotin for 30 min at room temperature. The reaction was stopped by removing the excess of biotin using a Zeba spin desalting column (ThermoFisher Scientific). St1Cas9 and SpCas9 RNP were prepared by incubating Cas9 molecules with a 1.2 molar excess of their cognate sgRNA for 20 min at room temperature. St1Cas9 in complex with the sgRNA•tDNA20 duplex (target sequence, tDNA20): 5′-AATACTTTTAT CAACGCAAG-3′) was assembled following the same procedure, except that the sgRNA was pre-mixed with a 1.2 molar excess of the tDNA20, heated to 90°C for 1 min, and slowly cooled down to room temperature. The PAM-bound St1Cas9 RNP was assembled by incubating, for 20 min at room temperature, St1Cas9 RNP with a 1.2 molar excess of partially duplexed tDNA59-ntPAM molecules, which were pre-formed by mixing equal molar amounts of 59-nt-long ssDNA containing the target (underlined) and PAM (bold) sequences (tDNA59: 5′- ATGGTTCTGGTTTCAT**TTTCTGC**AATACTTTTATCAACGCAAGAGGTGCTTCTGTTATG-3′) and 23-nt-long ssDNA containing the PAM sequence (ntPAM: 5′-**GCAGAAAA**TGAAACCAGAACCAT-3′) in the annealing buffer, heating to 94°C for 5 min, and slowly cooled down to room temperature. All experiments were performed using an Octet® Red96 system (FortéBio, Molecular Devices) at 27°C with an agitation speed of 1000 rpm in the interaction buffer (50 mM HEPES pH 7.5, 150 mM NaCl, 5 mM DTT, 0.02% Tween 20, 0.1% (w/v) BSA). The consecutive steps of the BLi measurements were as follows: 60 s incubation in interaction buffer, 300 s in 150 nM of biotinylated AcrIIA6, 60 s incubation in biocytin (quenching of biosensors surface), 60 s wash in buffer, 120 s baseline measurement, 180 s association in 375 nM of analytes (apo-St1Cas9, St1Cas9 RNP, St1Cas9 in complex with the sgRNA•tDNA20 duplex, St1Cas9 RNP in complex with the tDNA59-ntPAM partial duplex, apo-SpCas9, and SpCas9 RNP) and 500 s dissociation in buffer. Association and dissociation curves were double referenced against the buffer reference signal (biosensors coated with ligands and incubated in buffer only) and the reference

sensors signal (biosensors without AcrIIA6 and incubated in analytes), using the FortéBio Data Analysis software 8.2 (Molecular Devices). The figures were prepared with Prism8 (GraphPad).

Preparation of protein-nucleic acids complexes for cryo-EM

The St1Cas9•sgRNA•tDNA20•AcrIIA6 complex was prepared by first incubating St1Cas9 with a 1.2 molar excess of sgRNA•tDNA20 duplexes (formed as described in the previous section) for 20 min at room temperature, in a buffer containing 50 mM HEPES pH 7.5, 150 mM NaCl. Then, a two-fold molar excess of AcrIIA6 was added and the mixture was incubated for 20 min at room temperature. The complex was flash-frozen in liquid nitrogen and stored at −80°C.

The St1Cas9•sgRNA•AcrIIA6•tDNA59-ntPAM complex was prepared by first incubating St1Cas9 RNP (formed as described in the previous section) with a two-fold molar excess of AcrIIA6 for 20 min at room temperature. Then, the pre-formed St1Cas9•sgRNA•AcrIIA6 assembly was mixed with a 1.2 molar excess of the tDNA59-ntPAM partial duplex (formed as described in the previous section), and incubated for 20 min at room temperature. The complex was flash-frozen in liquid nitrogen and stored at −80°C.

Cryo-EM sample preparation

Freshly prepared samples were first analyzed by negative staining. Briefly, a 5 μL droplet of complexes at a final concentration of 15 μg/mL was applied onto a 300-mesh copper carbon-coated grids (Agar Scientific), glow discharged for 1 min at 15 mA using a PELCO easiGlow cleaning system (Ted Pella, Inc). After incubation of 45 s, the excess of sample on grid was removed by gentle side-blotting. Then, the grid was washed with a 10 μL droplet of water and stained with 1% (w/v) uranyl acetate solution for 30 s. Grids were blotted dry and imaged on a T12 Spirit transmission electron microscope operating at 120 kV. For cryo-EM, a 3.5 μL droplet of samples at final concentration of 0.5-0.8 mg/mL was applied onto a R1.2/1.3 300-mesh copper holey carbon grid (Quantifoil), previously glow discharged for 30 s at 30mA using a GloQube system (Quorum Technologies). Following incubation for 30-60 s at 4°C and 100% relative humidity, the grid was blotted for 2-3 s and plunge-frozen in liquid ethane using a Vitrobot Mk IV (ThermoFischer Scientific).

Cryo-EM data collection and image processing

All data were acquired on a 200 kV Talos Arctica (ThermoFischer Scientific) transmission electron microscope aligned to operate in parallel illumination. Cryo-EM data were acquired using EPU automated data collection software (ThermoFischer Scientific). Images were collected at nominal magnification of 120,000 × , corresponding to a pixel size of 0.889 Å/pixel at the specimen level, with an applied defocus values between −0.5 and −2.5 μm. Movies were acquired using Falcon 3EC direct electron detector (ThermoFischer Scientific) operating in electron counting mode (0.5 e^-/pix/sec) with a total exposure time of ~61 s and a total accumulated dose of 40 e^-/A^2, equally distributed over 40 movie fractions (1 e^-/A^2 per fraction). Movie frames were aligned before summing in individual fractions within EPU software. A total of 2,413 movies were collected for the St1Cas9•sgRNA•tDNA20•AcrIIA6 sample, and 2,014 movies were collected for the St1Cas9•sgRNA•AcrIIA6•tDNA59-ntPAM sample. Micrograph movies were imported in RELION-3.0 for all subsequent image processing tasks (Zivanov et al., 2018). Prior to particle picking, all images were subjected to motion-correction and dose-weighting using MotionCor2 (version 1.2.1) (Zheng et al., 2017). Estimation of the contrast transfer function (CTF) was performed on aligned, not weighted sum images using CTFFIND4 (version 4.1.10) (Rohou and Grigorieff, 2015). Particles were automatically picked using Gautomatch (version 0.56) (https://www.mrc-lmb.cam.ac.uk/kzhang/). Coordinates were then imported in RELION for visual inspection and manual removal of erroneously picked contaminants and aggregates.

For the St1Cas9•sgRNA•tDNA20•AcrIIA6 sample, a total of 599,758 particles were initially normalized and extracted in RELION-3.0. Dataset was subjected to several rounds of reference-free 2D classification to separate monomeric and dimeric forms of the complex (171,585 and 56,442 particles, respectively) and to also remove junk particles. For the monomeric form, an initial 3D auto-refinement was performed using a reference map of the St1Cas9•sgRNA complex, low-pass filtered at 30 Å, reconstructed from negative stain EM data. For the dimeric form, an initial 3D reference model was created by applying C2 symmetry to the monomeric form low-pass filtered at 40 Å. Subsequent rounds of 3D classification were performed to isolate particles best contributing to isotropic, high-resolution reconstructions. A final dataset of 118,189 and 43,239 particles were selected for the monomeric and dimeric forms, respectively. Final particle stacks were subjected to two rounds of CTF refinement, beam tilt correction and 3D auto-refinement followed by Bayesian polishing as implemented in RELION-3.0. A final 3D auto-refinement was performed on polished particles by applying a soft-edged and extended mask to a 15 Å low-pass filtered reference map and enabling solvent-flattened FSC calculation. Refined 3D reconstructions were sharpened using RELION-3.0 standard post-processing procedure applying a soft-edge and extended solvent mask. Overall resolutions estimates were calculated from Fourier shell correlations at 0.143 (applied B-factors and estimated resolutions listed in Table 1). Estimates of local resolutions were performed using RELION-3.0 local resolution tool.

For the St1Cas9•sgRNA•AcrIIA6•tDNA59-ntPAM sample, 592,839 particles in total were normalized and extracted in RELION-3.0. Dataset was subjected to several rounds of reference-free 2D classification to separate monomeric and dimeric forms of the complex and to remove junk particles. A total of 10,848 particles corresponding to the dimeric form of the complex were selected and re-centered upon extraction using RELION-3.0. For the particles stack corresponding to monomeric forms, an initial 3D auto-refinement was performed using the previously obtained map of the monomeric St1Cas9•sgRNA•tDNA20•AcrIIA6 assembly as reference, low-pass filtered at 30 Å. Particles were then subjected to 3D classification to separate particles contributing to 3D classes showing clear AcrIIA6 density. A total of 190,475 particles for St1Cas9•sgRNA•tDNA59-ntPAM and 230,397 particles for

St1Cas9•sgRNA•AcrIIA6•tDNA59-ntPAM were re-extracted and re-centered. Individual stacks were subjected to subsequent rounds of 3D classification to isolate particles best contributing to isotropic, high-resolution reconstructions. A final dataset of 68,361 particles were selected for the St1Cas9•sgRNA•tDNA59-ntPAM complex and 50,728 particles for the St1Cas9•sgRNA•AcrIIA6•tDNA59-ntPAM complex. Final particle stacks were subjected to two rounds of CTF refinement, beam tilt correction and 3D auto-refinement followed by Bayesian polishing as implemented in RELION-3.0. A final 3D auto-refinement was performed on polished particles by applying a soft-edged and extended mask to a 15 Å low-pass filtered reference map and enabling solvent-flattened FSC calculation. Refined 3D reconstructions were sharpened using RELION-3.0 standard post-processing procedure applying a soft-edge and extended solvent mask. Overall resolutions estimates were calculated from Fourier shell correlations at 0.143 (applied B-factors and estimated resolutions listed in Table 1). Estimates of local resolutions were performed using RELION-3.0 local resolution tool.

Models building, refinement and validation

The initial model of AcrIIA6 was based on its crystal structure (PDB: 6EYX (Hynes et al., 2018)). The initial model of St1Cas9 was generated by threading with PHYRE2 (Kelley et al., 2015) using the crystal structure of the target-bound form of the SaCas9 RNP (PDB: 5AXW) (Nishimasu et al., 2015). These initial models were aligned with cryo-EM reconstructions using UCSF Chimera (Pettersen et al., 2004) and the resulting aligned models were subjected to rigid body fitting and real-space refinement using PHENIX (Adams et al., 2010) before manual modifications, rebuilding and local refinement in *Coot* (Emsley et al., 2010). RNA and DNA molecules were manually built in *Coot* (Emsley et al., 2010). The full atomic models of St1Cas9•sgRNA•tDNA20•AcrIIA6 and St1Cas9•sgRNA•AcrIIA6•tDNA59-ntPAM complexes were subjected to multiple rounds of real-space refinement in PHENIX including global minimization and refinement of atomic displacement parameters, and applying secondary structures, Ramachandran and nucleic acids restraints (Afonine et al., 2018). NCS constraints were also applied in the case of the dimeric St1Cas9•sgRNA•tDNA20•AcrIIA6 assembly. The geometry and all-atoms contacts of the final models were validated by MolProbity (Chen et al., 2010) and the RCSB PDB validation server. *Coot*, UCSF Chimera, DALI (Holm and Laakso, 2016) and PISA (Krissinel and Henrick, 2007) were used to analyze the 3D structures, and UCSF ChimeraX was used to make figures (Goddard et al., 2018).

In vitro DNA cleavage assay

1540-bp long dsDNA substrates containing the target (underlined) and PAM (bold) sequences were generated by PCR (target and PAM sequences inserted into the pUC19 plasmid: **TTTCTGC**AATACTTTTATCAACGCAAG). DNA cleavage reactions were performed in cleavage buffer (20 mM HEPES pH 7.5, 100 mM KCl, 1 mM DTT, 10 mM $MgCl_2$). St1Cas9 (150 nM) was incubated with a 1.2 molar excess of sgRNA for 20 min at room temperature. Then, the dsDNA substrates (3nM) and AcrIIA6 or AcrIIA2 (150 nM) were simultaneously added, and the reaction was incubated at 37°C and allowed to react for 2, 10, and 30 min. Samples were flash-frozen in liquid nitrogen, thawed at 95°C for 5 min, and treated with proteinase K (Fisher Scientific). Products were analyzed by electrophoresis in a 1% agarose, 0.5X TBE gel stained with SYBR Gold (ThermoFisher Scientific).

Analysis of the interactions between dsDNA and St1Cas9 RNP using BLi

Streptavidin-coated biosensors (Molecular Devices) were used to immobilize biotinylated dsDNA molecules and monitor their association and dissociation with different St1Cas9 molecular species. Biotinylated dsDNA, containing target (underlined sequence) and PAM (bold sequences) sequences, were prepared by mixing equal molar amounts of biotinylated 59-mer ssDNA (biotin-tDNA59: 5′-ATGGTTCTGGTTTCAT**TTTCTGC**AATACTTTTATCAACGCAAGAGGTGCTTCTGTTATG-3′) with its complement (ntDNA59: 5′-CATAACAGAAGCACCTCTTGCGTTGATAAAAGTATT**GCAGAAA**ATGAAACCAGAACCAT-3′) in annealing buffer. The mixture was heated to 94°C for 5 min and slowly cooled down to room temperature. St1Cas9 RNP was prepared as describe above. AcrIIA6-bound St1Cas9 RNP was prepared by incubating St1Cas9 RNP with AcrIIA6 (1:1 and 1:2 molar ratios) for 20 min at room temperature. The negative control was prepared by incubating St1Cas RNP with AcrIIA2 (1:1molar ratio) for 20 min at room temperature. All experiments were performed using an Octet® Red96 system (FortéBio, Molecular Devices) at 27°C with an agitation speed of 1000 rpm in the interaction buffer. The consecutive steps of the BLi measurements were as follows: 60 s incubation in interaction buffer, 180 s in 150 nM of biotinylated dsDNA, 60 s incubation in biocytin (quenching of biosensors surface), 60 s wash in buffer, 120 s baseline measurement, 120 s association in 150 nM of analytes (St1Cas9 RNP, AcrIIA6-bound St1Cas9 RNP, St1Cas9 RNP mixed with AcrIIA2), and 280 s dissociation in buffer. Association and dissociation curves were double referenced against the buffer reference signal (biosensors coated with ligands and incubated in buffer only) and the reference sensors signal (biosensors without dsDNA and incubated in analytes), using the Fortébio Data Analysis software 8.2 (Molecular Devices). The figure was prepared with Prism8 (GraphPad).

Genome editing vectors

To establish the clonal K562 cell line constitutively expressing C-terminally tagged St1Cas9 under the control of the h*PGK1* promoter, the Cas9 ORFs from MSP1594_2x_NLS (Addgene #110625) was subcloned into AAVS1_Puro_PGK1_3xFLAG_Twin_Strep (Dalvai et al., 2015) (Addgene #68375). The U6-driven sgRNA expression plasmid for St1Cas9 (St1Cas9_LMD-9_sgRNA_pUC19; Addgene #110627) was synthesized as a gBlock gene fragment (Integrated DNA Technologies) and cloned into pUC19. The same approach was used to create the truncated sgRNA expression plasmids for St1Cas9, SpCas9, and SaCas9 (Cong et al., 2013; Ran et al., 2015).

The single vector mammalian expression system containing a CAG promoter-driven St1Cas9 LMD-9 and its U6-driven sgRNA (U6_sgRNA_CAG_hSt1Cas9_LMD-9; Addgene #110626) was built from the above-described plasmids. Hybrid St1Cas9 proteins containing the N terminus of LMD-9 and the C terminus of LMG 18311 or CNRZ 1066 were generated using gBlock gene fragments (Integrated DNA Technologies) subcloned into (U6_sgRNA_CAG_hSt1Cas9_LMD-9; Addgene #110626). The following guides targeting St1Cas9 LMD-9 to *EMX1* (GAAGGGCCTGAGTCCGAGCA), *FANCF* (GTAGGTAGTGCTTGAGACCG), *RUNX1* (GAGG TATCCAGCAGAGGGGA), and *ATP1A1* (GCAGCTTGGATGCTATAAGC) have been used. The St1Cas9 LMG 18311 chimera was targeted using the *FANCF* (GGCGGCTGCACAACCAGTGG), *RUNX1* (GCCATTTCATTACAGGCAAAGCT) and *AAVS1* (GGACACAG GATCCCTGGAGG) guides. For the St1Cas9 CNRZ 1066 hybrid, the *EMX1* (GTTCCAGAACCGGAGGACAAA) and *ATP1A1* (GGTAATTGAGAAGAAGTGGG) guides were used. The AcrIIAs expression vectors have been described (Hynes et al., 2018). Each ORFs were codon-optimized for expression in human cells, synthesized as gBlock gene fragments (Integrated DNA Technologies), and cloned into a modified pVAX1 vector (ThermoFisher Scientific) containing a beta-globin intron.

Surveyor nuclease and TIDE assays

Genomic DNA from 2.5E5 cells was extracted with 250 μL of QuickExtract DNA extraction solution (Lucigen) per manufacturer's protocol, 3 days after transfection. The various loci were amplified by PCR and assays were performed with the Surveyor mutation detection kit ((Integrated DNA Technologies) as described (Agudelo et al., 2017; Guschin et al., 2010). Samples were separated on 10% PAGE gels in TBE buffer. Gels were imaged using a ChemiDoc MP (Bio-Rad) system and quantifications were performed using the Image lab software (Bio-Rad).

Base editing vectors

Base editors were constructed into U6_sgRNA_CAG_hSt1Cas9_LMD9 (Addgene #110626) using fragments derived from pCMV_BE4max_3xHA (Koblan et al., 2018) (Addgene #112096, a gift from David Liu). Protein and DNA sequences for the St1Cas9 LMD-9 base editors are available at https://doi.org/10.1101/321208.

Base editing assays

Genomic DNA from 2.5E5 cells was extracted with 250 μL of QuickExtract DNA extraction solution (Lucigen) per manufacturer's recommendations, 3 days after transfection. The various loci were amplified by 30 cycles of PCR using *EMX1*-F 5′-CCATCCCCTTCTGTGAATGT-3′, *EMX1*-R 5′-GGAGATTGGAGACACGGAGA-3′, *RUNX1*-F 5′-CCAGCACAACTTACTCGCACTT GAC-3′, *RUNX1*-R 5′-CATCACCAACCCACAGCCAAGG-3′. Base editing was quantified from Sanger sequencing reads using EditR (Kluesner et al., 2018).

Transcriptional activation system

K562 cells were transfected with M_ST1n_VPR (Addgene #63799, a gift from George Church) (Chavez et al., 2015), M-tdTom-ST1 (Addgene #48678, a gift from George Church) (Esvelt et al., 2013), and M-ST1-sgRNA (Addgene #48672, a gift from George Church) (Esvelt et al., 2013) in combination with the indicated anti-CRISPR vectors. For transfection involving transcriptional activation with dSpCas9-VPR, K562 cells were transfected with SP-dCas9-VPR (Addgene #63798, a gift from George Church) (Chavez et al., 2015) #123), (Addgene #63798, a gift from George Church), M-tdTom-SP (Addgene #48677, a gift from George Church, a gift from George Church) (Esvelt et al., 2013), M-SP-sgRNA (Esvelt et al., 2013) (Addgene #48671, a gift from George Church, a gift from George Church) (Esvelt et al., 2013). Empty pVAX1 and pUC19 vectors were used to normalize DNA concentration in all transfections. Fluorescence microscopy images were taken with an EVOS FL Cell Imaging System three days post-transfection. The intensity and the frequency of cells expressing tdTomato were assessed with a BD LSR II flow cytometer three days post-transfection.

QUANTIFICATION AND STATISTICAL ANALYSIS

The figure legends describe the number of biological replicates for the different assays (BLi, DNA cleavage, base editing, transcriptional activation, TIDE). EditR (Kluesner et al., 2018) was used to quantify base editing from Sanger sequencing reads with a p value cut-off < 0.01. TIDE analysis was performed using a significance cut-off value for decomposition of $p < 0.001$ (Brinkman et al., 2014).

DATA AND CODE AVAILABILITY

The cryo-EM maps and the corresponding refined molecular models have been deposited in the Electron Microscopy Data Bank and RCSB Protein Data Bank under the accession codes EMD-4900 and PDB: 6RJ9 for St1Cas9•sgRNA•tDNA20•AcrIIA6 (monomeric assembly), EMD-4901 and PDB: 6RJA for St1Cas9•sgRNA•tDNA20•AcrIIA6 (dimeric assembly), EMD-4904 and PDB: 6RJG for St1Cas9•sgRNA•AcrIIA6•tDNA59-ntPAM, and EMD-4902 and PDB: 6RJD for St1Cas9•sgRNA•tDNA59-ntPAM. Unprocessed gel images of Figures 4A, S1B, and S4E are available at Mendeley data (https://doi.org/10.17632/hrd973rwsf.1). Other data are available from the corresponding author upon request.