bioRχiv

# A census of anti-CRISPR proteins reveals AcrIE9 as an inhibitor of *Escherichia coli* K12 Type IE CRISPR-Cas system

**Dmitry Taranenko,[a] Oksana Kotovskaya,[a] Konstantin Kuznedelov,[b] Daria Yanovskaya,[a] Alina Demkina,[c,d] Sofya Fardeeva,[a] Viktor Mamontov,[a] Kaiya Vierra,[e] Nathaniel Burman,[e] Dan Li,[f] Minggui Wang,[f] Blake Wiedenheft,[e] Konstantin Severinov,[b,c] Ekaterina Semenova,[b,*] and Artem Isaev[a,*]**

[a]Skolkovo Institute of Science and Technology, Moscow, Russia
[b]Waksman Institute for Microbiology, Rutgers, The State University of New Jersey, Piscataway, NJ, USA
[c]Institute of Gene Biology, Moscow, Russia
[d]Shemyakin-Ovchinnikov Institute of Bioorganic Chemistry, Moscow, Russia
[e]Montana State University, Bozeman, Department of Microbiology and Cell Biology, Bozeman, Montana 59717, USA
[f]Institute of Antibiotics, Huashan Hospital, Fudan University
*Correspondence should be addressed to A.I. (Artem.Isaev@skoltech.ru) or E.S. (semenova@waksman.rutgers.edu)

**CRISPR-Cas adaptive immunity systems provide defense against mobile genetic elements and are often countered by diverse anti-CRISPR (Acr) proteins. The Type IE CRISPR-Cas of *Escherichia coli* K12 has been a model for structural and functional studies and is a part of the species' core genome. However, this system is transcriptionally silent, which has fueled questions about its true biological function. To clarify the role of this system in defense, we carried out a census of Acr proteins found in *Enterobacterales* and identified AcrIE9 as a potent inhibitor of the *E. coli* K12 Type IE CRISPR-Cas system. While sharing little sequence identity, AcrIE9 proteins from *Pseudomonas* and *Escherichia* both interact with the Cas7 subunit of the Cascade complex, thus preventing its binding to DNA. We further show that AcrIE9 is genetically linked to AcrIE10, forming the most widespread anti-CRISPR cluster in *Enterobacterales*, and this module often co-occurs with a novel HTH-like protein with unusual architecture.**

## Introduction

CRISPR-Cas (Clustered Regularly Interspaced Short Palindromic Repeats and CRISPR-associated genes) is one of the most widespread prokaryotic immunity systems, playing an important role in the co-evolution of phage and bacteria[1–3]. CRISPR-Cas systems utilize an adaptive immunity principle: at the adaptation stage, small DNA fragments derived from mobile genetic elements (MGE) are introduced into the CRISPR array via the action of conserved Cas1-Cas2 adaptation module[4,5]. Transcription of the CRISPR array and processing of the resulting non-coding RNA generates short guide CRISPR RNA molecules (crRNA)[6,7], which interact with the effector Cas proteins, and mediate degradation of nucleic acid targets complementary to crRNA spacer sequences during the CRISPR interference stage[8,9]. CRISPR-Cas systems are divided into Class 1 with multi-protein effector complexes (e.g., Cascade and Cas3 in Type I systems), and Class 2 with single-protein effectors (e.g., Cas9, Cas12, Cas13)[10]. Based on the composition of the interference module, CRISPR-Cas systems are further classified into seven types and multiple subtypes[10,11].

MGEs can evade CRISPR interference through accumulation of mutations in the targeted DNA regions, called protospacers[12], mutations in the protospacer adjacent motif (PAM)[13], or through activity of anti-CRISPR proteins (Acrs)[14–17] and RNAs (Racrs)[18,19]. The mechanisms of Acr-mediated inhibition are quite diverse: from prevention of crRNA binding[18,20], DNA mimicry[21–23], inhibition of DNA recognition or cleavage by effectors[24–26], to Cas protein modifications[27], inactivation through dimerization[28] or effector complexes disassembly[29]. As a countermeasure, some CRISPR-Cas systems with transcriptional auto-repression circuits could increase Cas protein expression in response to inhibition[30,31]. Acrs have been identified for the majority of CRISPR-Cas types and are found in diverse MGEs (lytic and temperate phages, conjugative plasmids, ICEs)[32–35], where they form so-called anti-defense clusters regulated by Anti-CRISPR-associated (Aca) proteins that function as transcriptional repressors[32,36]. The presence of Acrs in mobile elements could be linked with the spread of antibiotic-resistance genes in clinical isolates of *Klebsiella*[37,38] and other ESKAPE pathogens[39].

Most *Escherichia coli* encode Type IE CRISPR-Cas system (*Ec*-IE) that can be associated with one or two cognate CRISPR arrays, while a few strains contain the Type IF (*Ec*-IF) that can be associated with two other CRISPR arrays[40,41]. The *Ec*-IE system served as a model for determining the principles of adaptation, crRNA maturation, and interference at the dawn of CRISPR-Cas research[42–46]. However, the activity of the *Ec*-IE system was either studied *in vitro* or in *E. coli* overexpression systems, because it is naturally transcriptionally silenced by H-NS[47,48]. This is a common bacterial strategy for the suppression of horizontally acquired xenogenic mobile elemetns[49,50]. Other transcriptional factors, such as LeuO, CRP, StpA, LrhA, and SspA contribute to the regulation of *cas* promoters[48,51–55]. Transcriptional silencing and the lack of spacers matching genomes of lytic phages, have led to a hypothesis that the *Ec*-

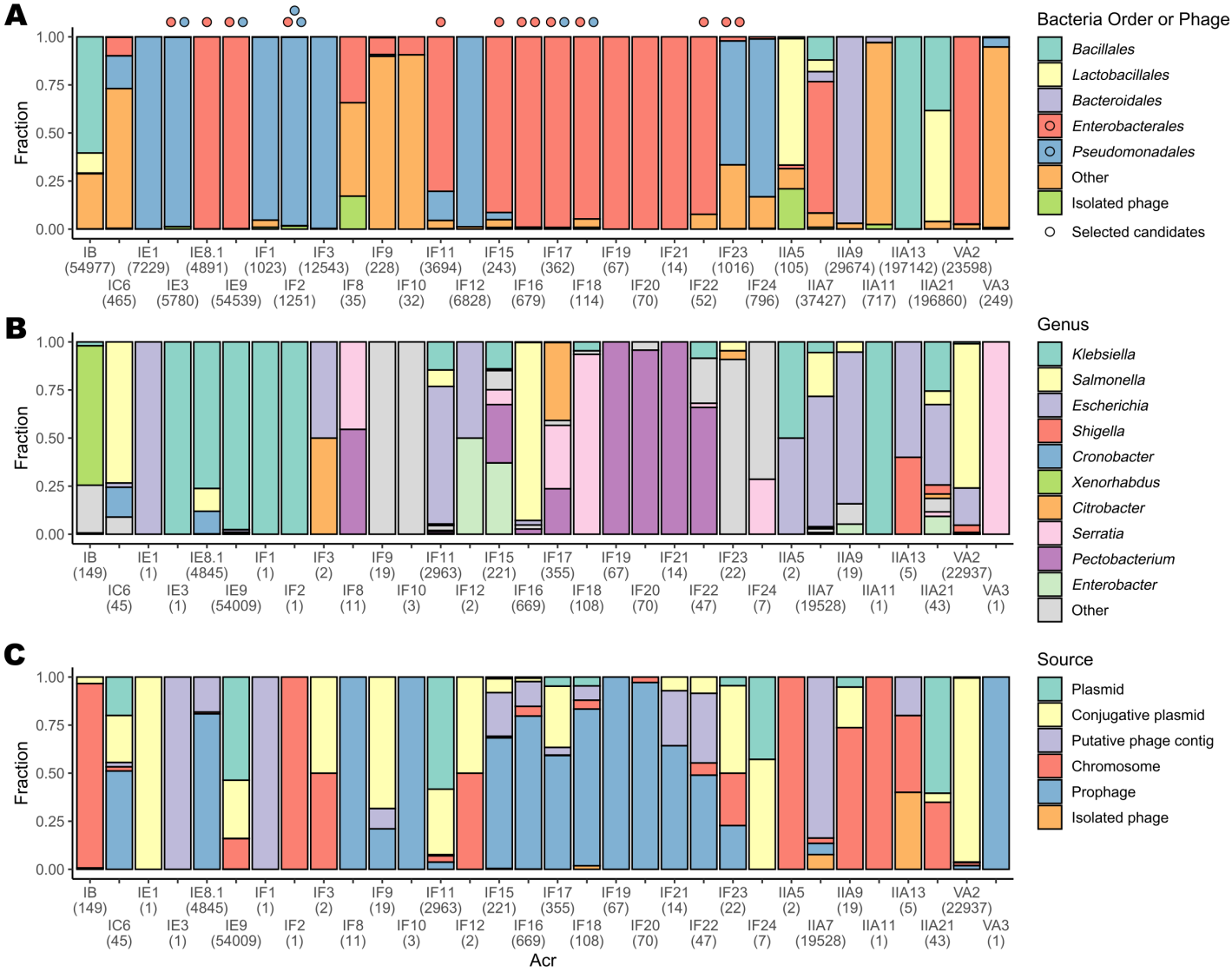


**Figure 1: Acr classes distribution among *Enterobacterales*. (A)** Distribution of known anti-CRISPR proteins among phage genomes and different bacterial groups. Colored circles at the top represent selected candidates for experimental evaluation, red – from *Enterobacterales*, blue – from *Pseudomonadales*. **(B)** Genus level taxonomic distribution of Acrs found in *Enterobacterales*. **(C)** Local genomic context of Acr genes found in *Enterobacterales*, as predicted by geNomad. `Putative phage contig` category includes predicted phage contigs lacking terminal repeats. `Chromosome` category includes all loci, not found to be associated with other categories. Numbers in parenthesis for all panels represent number of loci found.

IE system does not play a major role in anti-viral immunity of the host[56,57]. On the other hand, *Ec*-IE is present in ~70% of sequenced *E. coli* isolates and thus represents a core part of the species pangenome[58,59]. It was speculated that *Ec*-IE can be activated under specific conditions, when transcriptional silencing is relieved[53]. The recent census of *Ec*-IE spacers in the gut microbiome revealed preferential targeting of prophages[60], while independent work proposed the role of *Ec*-IE in control of cryptic prophages[61].

Existence of Acrs active against *Ec*-IE can support a role for this system in genetic conflicts with MGEs in natural populations. Machine learning-based methods have predicted multiple Acr candidates within mobile elements of *E. coli*[62–64]. This could be attributed to the overrepresentation of *E. coli* genomes in the databases or may suggest a hidden diversity of Acrs active against the *Ec*-IE system. Earlier experimental studies have not identified Acrs against the *Ec*-IE system[65], while more recent works demonstrated anti-*Ec*-IE activity of the AcrIE10[38] and AcrIC6*[30] proteins. Nevertheless, a dedicated study of the *Ec*-IE sensitivity to Acrs has not been carried out.

Here, we analyzed the distribution of Acrs within mobile elements of *E. coli* and *Enterobacterales* and experimentally investigated a collection of such Acrs, along with selected Acrs derived from *Pseudomonas aeruginosa*. We demonstrate that *Ec*_AcrIE9 and *Pa*_AcrIE9 robustly inhibit the *Ec*-IE system, while *Ec*_AcrIF15 and *Kp*_AcrIF22 exhibit moderate inhibitory activity against the *Ec*-IF system. We further show that AcrIE9 binds to the Cas7 subunit of Cascade, inhibiting its interaction with target DNA. Analysis of the genomic loci encoding AcrIE9 revealed a strong functional link with the AcrIE10 protein and a novel HTH-like protein with unusual architecture. These results provide the first mechanistic insight into Acr activity against the model Type IE CRISPR-Cas of *E. coli* K12 and have implications for the co-evolution of CRISPR-Cas and MGE in pathogenic bacteria.

## Results

### Multiple Acr classes are found in *Enterobacterales* MGEs

To identify Acr proteins encoded by *Enterobacterales* and their MGEs, we leveraged the dbAPIS database of anti-immune proteins[35] and performed an Acr HMM profile search against prokaryotic and viral protein sequences derived from non-redundant database (nr BLAST) of microbial and MGE genomes (**Extended Data File 1**). **Figure 1A** sums up the phylogenetic distribution of Acr classes, for which at least one hit was found within *Enterobacterales* genomes or within phages infecting *Enterobacterales* hosts. These data reveal that AcrIE8, IE9, IF11, IF15-22, IIA7, and VA2 are predominantly found within *Enterobacterales*. Genus-level analysis further demonstrates a "specialization" of different Acr classes towards specific bacteria, and multiple AcrIE and AcrIF homologs were found within *Escherichia* genomes (**Figure 1B**). In this dataset, AcrIE9 represents the most abundant Acr variant found in more than 54,000 instances. It should be noted that AcrIE10[38] was not present in the dbAPIS HMM profiles database, while our downstream analysis identifies its comparable abundance to AcrIE9. Next, we focused on the genomic environment of Acrs within *Enterobacterales* and found that the majority of Acr genes are encoded by MGEs: phages (including metagenomic viral contigs), prophages, plasmids, and conjugative plasmids (**Figure 1C**). We selected 9 candidates for experimental validation: 2 proteins predicted to be active against IE system, and 7 anti-IF proteins (**Figure 1A**, **Supplementary Table 1**). To balance the collection, we selected homologs of Acr proteins from *P. aeruginosa* as additional controls (2 anti-IE and 4 anti-IF). Since multiple Acrs were shown to possess broad spectrum activity against different types of CRISPR-Cas systems, including cross-inhibitory activity against IE and IF subtypes[23], all candidates were tested against the *Ec*-IE and the *Ec*-IF systems.

### AcrIE9 from *E. coli* and *P. aeruginosa* inhibits Type IE CRISPR-Cas of *E. coli*

To estimate the anti-CRISPR activity against the *Ec*-IE system, we used F' KD263 strain[66], encoding a G8 spacer targeting M13 phage and chromosomal *cas* genes under the control of the arabinose- and IPTG-inducible promoters (**Figure 2A**). Selected Acr candidates were synthesized, cloned, and expressed from a pBAD plasmid under the control of an *araBAD* promoter. While in the presence of an empty pBAD vector (EV), the induced *Ec*-IE system provided robust defence against M13 in phage plaque assay, expression of the *Pa*_AcrIE9 and *Ec*_AcrIE9 restored phage plaquing efficiency (**Figure 2B**). To further validate this effect, we used the SS80 strain, a derivative of the KD263 with a spacer against lytic phage T7[44] (**Figure 2C**). Again, *Pa*_AcrIE9 and *Ec*_AcrIE9 completely suppress the induced *Ec*-IE immune system in both phage plating assay and liquid culture infection (**Figure 2D**, **Supplementary Figure 1A-B**). However, at low Multiplicity of Infection (MOI) *Pa*_AcrIE9 outperforms *Ec*_AcrIE9 (**Supplementary Figure 1C**). Lastly, we investigated Acr activity in conditions of CRISPR-Cas expression from native promoters

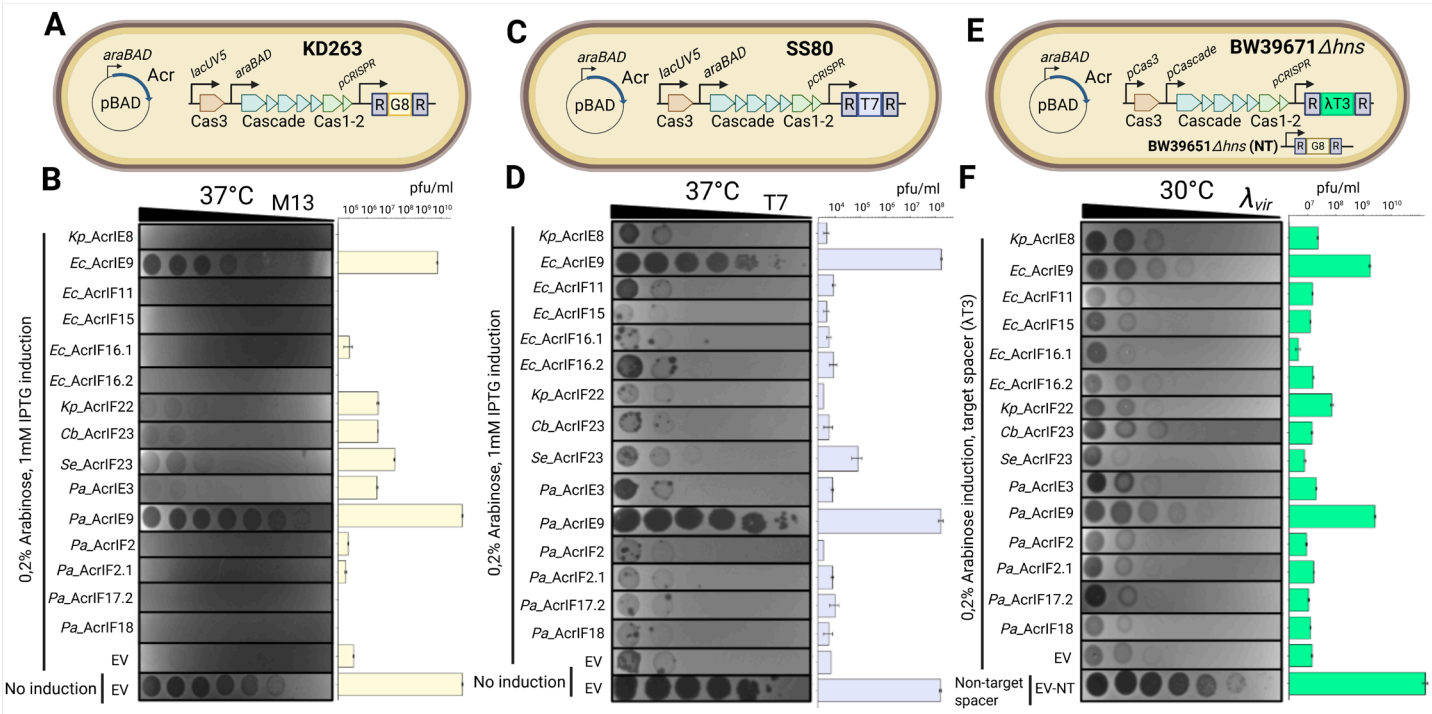


**Figure 2: Screening of Acr candidates reveals anti-CRISPR proteins active against Type I-E CRISPR-Cas of *E.coli*.** **(A)** *E. coli* KD263 strain with a M13 phage-targeting spacer G8 was used for M13 phage infection assay. **(B)** EOP with phage M13 on the lawns of KD263 cells with induced expression of Cas proteins and Acr candidates. Cultures with empty pBAD vector and without induction of CRISPR-Cas expression were used as controls. **(C)** *E. coli* SS80 strain, a KD263 derivative with T7-targeting spacer, was used for T7 infection assay. **(D)** EOP with phage T7 on the lawns of SS80 cells with induced expression of Cas proteins and Acr candidates. Cultures with empty pBAD vector and without induction of CRISPR-Cas expression were used as controls. **(E)** *E. coli* BW39671 strain, a BW25113 *Δhns* derivative with a *λ* phage-targeting spacer T3, was used for $\lambda_{vir}$ phage infection assay. **(F)** EOP with phage $\lambda_{vir}$ on the lawns of BW39671 cells with induced expression of Acr candidates. Cultures without a $\lambda_{vir}$ targeting spacer and with empty pBAD vector were used as controls. EOPs presented on panel **(B)** and **(D)** were performed at 37C, while EOP from the panel **(F)** was performed at 30C. Acr expression was induced with 0.2% L-arabinose, CRISPR-Cas expression at the EOPs presented on panel **(B)** and **(D)** was induced with 0.2% L-arabinose and 1 mM IPTG, while EOP from the panel **(F)** was performed at conditions of native CRISPR-Cas expression. All experiments were performed in biological triplicates and bars on the right represent mean pfu/ml values.

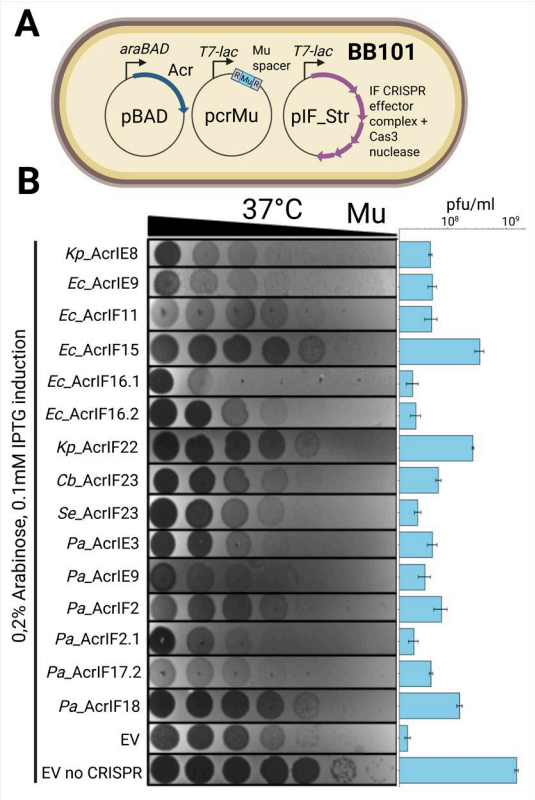


**Figure 3: Screening of Acr candidates reveals potential anti-CRISPR proteins targeting I-F CRISPR of *E.coli* LF82. (A)** *E. coli* BB101 strain with a plasmid-encoded IF LF82 system and phage-targeting spacer was used for the phage Mu phage infection assay. **(B)** EOP with phage Mu on the lawns of BB101 cells with induced expression of Cas proteins and Acr candidates. Cultures with empty pBAD vector and without CRISPR-Cas encoding plasmid were used as controls. Acr expression was induced with 0.2% L-arabinose, CRISPR-Cas and spacer expression was induced with 0.1 mM IPTG. All experiments were performed in biological triplicates and bars on the right represent mean pfu/ml values.

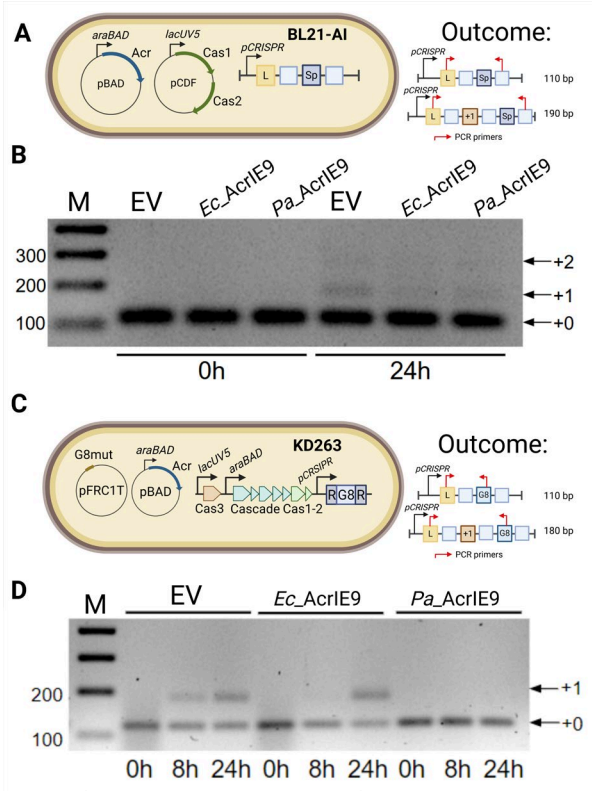


**Figure 4: AcrIE9 inhibits primed but not naïve adaptation. (A)** *E. coli* BL21-AI strain with a plasmid-encoded Cas1-Cas2 was used for the naïve adaptation assay. Acr expression was induced with 0.2% L-arabinose, Cas1-Cas2 expression was induced with 1 mM IPTG. **(B)** Agarose gel representing CRISPR arrays after naïve adaptation. Expanded arrays are labeled with arrows. **(C)** *E. coli* KD263 strain transformed with a priming plasmid was used for the primed adaptation assay. Acr expression was induced with 0.2% L-arabinose, CRISPR-Cas expression was induced with 0.2% L-arabinose and 1 mM IPTG. **(D)** Agarose gel representing CRISPR arrays after primed adaptation. Expanded arrays are labeled with arrows. Representative gels of the experiments carrying in biological triplicates are shown on the panels **(B)** and **(D)**. Expected size of the products after PCR with CRISPR array-specific primers with non-expanded and expanded array templates are shown on the right on the panels **(A)** and **(C)**.

in *Δhns* background, since overexpression of Cas proteins from inducible promoters could mask inhibitory effects of some Acrs. We used BW39671 *Δhns* host[45] encoding λT3 spacer targeting phage λ (**Figure 2F**) and estimated phage $\lambda_{vir}$ plaquing efficiency. Cells with a non-targeting spacer were used as a negative control, and showed a lack of defence. While the inhibitory effect of *Pa*_AcrIE9 and *Ec*_AcrIE9 on the *Ec*-IE system was confirmed, no other Acr protein demonstrated robust inhibition (**Figure 2F**). We additionally monitored the accumulation of the phage particles in the spent liquid media of BW39671 *Δhns* cells infected with $\lambda_{vir}$. In line with the EOP assay, *Pa*_AcrIE9 or *Ec*_AcrIE9 expression enhanced accumulation of phage progeny, although did not completely inhibit the native CRISPR-Cas system (**Supplementary Figure 1D,E**).

A recently published ML-based model predicted multiple novel Acr classes and *E. coli* represented the largest source of identified canidates[62]. To test the activity of these candidates we selected, synthesized, and cloned representatives from the top 13 anti-IEs found in *E. coli* (**Supplementary Table 2**). While the two candidates representing Acr candidate clusters 2320 and 1797, exhibited high toxicity to the host, the expression of the remaining candidates neither resulted in toxic phenotypes nor showed notable *Ec*-IE suppression (**Supplementary Figure 2A,B**).

## Acr proteins provide moderate anti-CRISPR activity against Type IF CRISPR-Cas of E. coli LF82

Although not as widespread as Type IE, Type IF CRISPR-Cas system can be found in many *E. coli* isolates[40,41]. Previous work reported that the chromosomally encoded *Ec*-IF system from LF82 strain interferes with the plasmid stability[67]. Despite multiple attempts, we were unable to demonstrate LF82 *Ec*-IF system interference against M13 phage infection or plasmid transformation under native conditions. Thus, we turned to the plasmid-mediated expression of the *Ec*-IF LF82 system in a CRISPR-deficient BB101 *E. coli* host[68] (**Figure 3A**). In the presence of the targeting spacer, the *Ec*-IF LF82 system reduced the phage Mu plaquing ability by two orders of magnitude (**Figure 3B**). Expression of *Ec*_AcrIF15, *Kp*_AcrIF22, and *Pa*_AcrIF18 provides moderate (i.e., 10x) inhibition of the *Ec*-IF LF82 interference, supporting the existence of active anti-IF Acrs in MGEs of *Entrobacterales* (**Figure 3B**). ML-predicted clusters lacked an activity against the *Ec*-IF system (**Supplementary Figure 2C,D**). Given the relatively low Acr activity against the IF system, we focused on elucidating the mechanisms of *Ec*-IE inhibition by AcrIE9 in subsequent analyses.

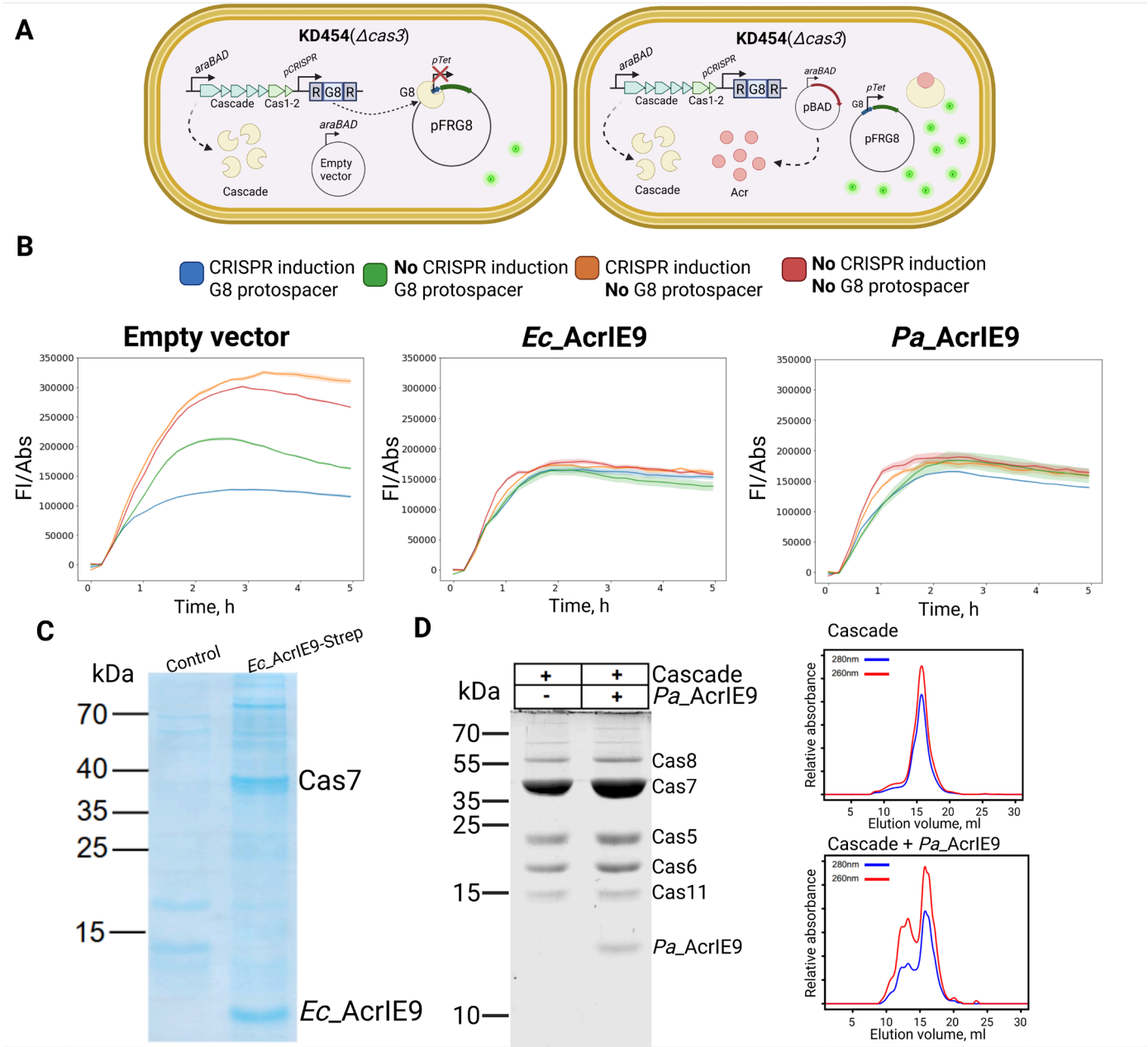


**Figure 5: AcrIE9 targets Cas7 subunit of Cascade *in vivo* and interferes with Cascade DNA binding. (A)** *E. coli* KD454 strain, a KD263 derivative lacking Cas3, was used for CRISPR transcriptional silencing assay. sfGFP was expressed from *pTet* containing or lacking G8 protospacer. Acr expression was induced with 0.2% L-arabinose, CRISPR-Cas expression was induced with 0.2% L-arabinose and 1 mM IPTG, sfGFP expression was induced with 200 ng/mL aTc. A scheme of the assay and conditions color code is shown on the right. **(B)** Dynamics of sfGFP production in KD454 cells carrying an empty pBAD vector or expressing *Ec*_AcrIE9/*Pa*_AcrIE9. Fluorescence intensity was normalized to the culture optical density. **(C)** *In vivo* pull-down of *Ec*_AcrIE9 with a C-terminal Strep tag in *E.coli* KD263 after induction of untagged *Ec*_IE Cascade expression. Control represents a similar experiment carried with AcrIE9_Ec without the Strep tag. Identity of the labeled proteins on was determined via MALDI-TOF mass-spectrometry. **(D)** SDS-PAGE analysis of Cascade purified from cell cultures with or without *Pa*_AcrIE9 co-expression (left panel). SEC elution profiles of Cascade and Cascade:*Pa*_AcrIE9 complexes run on Superose 6 10/300 column (right panel).

## AcrIE9 inhibits primed, but not naïve, adaptation

We started with testing whether AcrIE9 inhibits the incorporation of new spacers into the CRISPR array. For this, we overexpressed Cas1-Cas2 proteins in the *E. coli* BL21-AI host carrying CRISPR array but lacking other Cas proteins. We then monitored CRISPR array expansion via PCR in the presence and absence of AcrIE9 (**Figure 4A**). After 24h of cultivation, the CRISPR array expanded for one or two new spacers, and spacer acquisition remained unaffected by the presence of AcrIE9 (**Figure 4B**).

Some CRISPR-Cas systems, including *Ec*-IE, have evolved a primed adaptation mechanism, which facilitates efficient spacer acquisition from CRISPR-targeted DNA[69]. This process is triggered by a pre-existing spacer that matches a sequence within the genome of MGE. Primed adaptation efficiently restricts the accumulation of escapers. This mode of spacer acquisition relies on the coordinated action of both the interference and adaptation modules and is, therefore, strictly dependent on the presence of Cas interference proteins[69]. To investigate the effect of AcrIE9 on primed adaptation, we transformed *E. coli* KD263 with a targeted plasmid, carrying a single mismatch (C1T) in the protospacer region complementary to the G8 spacer[70] (**Figure 4C**). In the absence

of AcrIE9, expansion of the CRISPR array was already observed 8h after *Ec*-IE induction (**Figure 4D**). In contrast, expression of *Ec*_AcrIE9 resulted in CRISPR array expansion only 24h after *Ec*-IE induction, whereas expression of *Pa*_AcrIE9 completely inhibited primed adaptation. This aligns with the more pronounced anti-CRISPR effect observed in liquid culture infection (**Supplementary Figure 1C**). Together, these results indicate that AcrIE9 does not affect naïve adaptation but specifically inhibits primed adaptation, which depends on the activity of the Cascade and Cas3.

### AcrIE9 binds to the Cas7 subunit and prevents Cascade interaction with DNA *in vivo* and *in vitro*

To identify the mechanism of CRISPR inhibition, we tested the ability of *Ec*-IE Cascade to bind DNA in the presence of AcrIE9. In Type I CRISPR-Cas systems, the Cascade complex is guided by crRNA to a complementary DNA target. Binding of the target DNA leads to the formation of an R-loop, which induces conformational changes required for the recruitment of the processive nuclease-helicase Cas3[71,72]. In the absence of Cas3, Cascade remains bound to the protospacer and can act as a transcriptional silencer[73,74]. This property allows for the discrimination of Acr mechanisms (i.e., blocking DNA degradation versus inhibition of Cascade assembly, or target DNA recognition). We used *Δcas3* KD454 strain with inducible *Ec*-IE system[75] and introduced a G8 protospacer in the promoter of sfGFP gene expressed from a plasmid (**Figure 5A**). As expected, expression of sfGFP was repressed in conditions of CRISPR-Cas induction and the presence of the targeted G8 protospacer (**Figure 5B**). Expression of *Ec*_AcrIE9 and *Pa*_AcrIE9 slightly reduced cell growth, indicating intermediate toxicity, yet restored sfGFP production. This suggests that AcrIE9 prevents DNA recognition by the *Ec*-IE Cascade complex (**Figure 5B**, **Supplementary Figure 3**).

To check if AcrIE9 directly interacts with Cascade, we constructed a C-terminally Strep-tagged version of *Ec*_AcrIE9 and performed an *in vivo* pull-down experiment in KD263 cells producing all *Ec*_IE subunits. Cells expressing untagged AcrIE9 were used as a negative control. A coomassie stained SDS-PAGE of the pulldown reveals Cas7 subunit as the major AcrIE9 binding partner (**Figure 5C**). To clarify whether AcrIE9 binds to the free Cas7 or the assembled Cascade complex, we performed a total proteomic analysis of the StrepTrap column eluates that additionally revealed Cas11 and Cas8e among the most enriched proteins (**Supplementary Figure 4A**). The lack of other Cascade subunits could reflect they lower abundance and/or decreased stability of the protein complexes not directly bound to a bait (AcrIE9). A similar analysis performed with *Pa*_AcrIE9 also confirmed co-purification of the Cas7 and Cas8 subunits, despite a weaker binding of the bait protein to the resin (**Supplementary Figure 4B**). We further performed purification of Strep-tag labeled *Ec*_IE Cascade from cells co-expressing untagged *Pa*_AcrIE9 and via SEC and SDS-PAGE analysis demonstrated the formation of a stable complex, containing all expected Cascade subunits (**Figure 5D**). Together, the results confirm direct interaction of AcrIE9 with complete Cascade complex *in vivo*.

To validate these results, we purified AcrIE9 and studied its interaction with Cascade *in vitro*. Since *Pa*_AcrIE9 was more

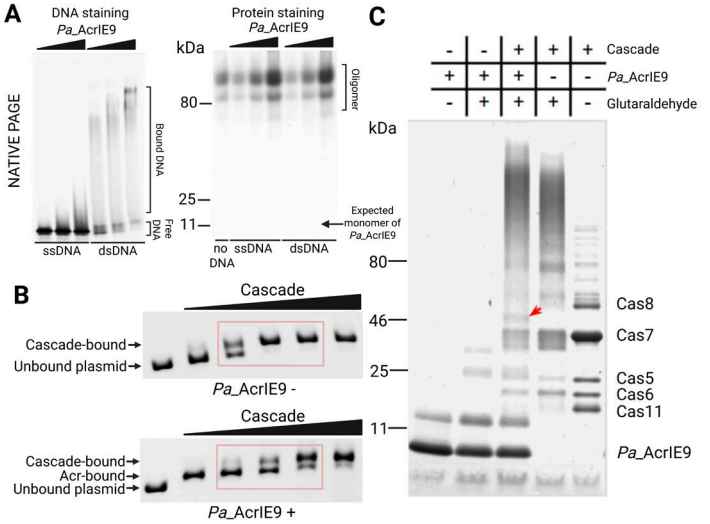


**Figure 6: AcrIE9 binds to the Cas7 subunit of the assembled Cascade complex and interferes with target DNA recognition *in vitro*. (A)** EMSA demonstrates *Pa*_AcrIE9 binding to dsDNA. 0.25 μM of 52-bp single-stranded or double-stranded DNA was incubated with *Pa*_AcrIE9 at increasing concentrations (17, 33 and 66 μM). **(B)** EMSA with increasing concentrations of *Ec*_IE Cascade (20 – 320 nM) and 4.5 nM of target plasmid in the presence or absence of an excess (1.5 μM) of *Pa*_AcrIE9. Without AcrIE9, a complete plasmid shift is achieved (top panel). In the presence of *Pa*_AcrIE9, two gel shifts associated with consequent binding of the *Pa*_AcrIE9 and *Ec*_IE Cascade could be observed, and Cascade binding is incomplete (bottom panel). **(C)** *In vitro* glutaraldehyde (0.5%) cross-linking assay between 23 μM *Pa*_AcrIE9 and 6 μM of *Ec*_IE Cascade, resulting in the appearance of novel band migrating above Cas7 (red arrow), that was confirmed to represent Cas7-AcrIE9 adduct.

stable after purification, most downstream experiments were performed with this variant. First, we noticed that AcrIE9 on it's own has double-stranded but not single-stranded DNA binding activity (**Figure 6A**). Incubation with DNA binding competitor heparin reduces gel shift (**Supplementary Figure 5A)**. In addition, protein staining after native PAGE revealed that irrelevant ot the presence of DNA AcrIE9 migrates as a higher molecular weight complex, while incubation with heparin reduces its size (**Figure 6A, Supplementary Figure 5A**). Since Acr proteins can inhibit CRISPR-Cas DNA cleavage activity through effects on DNA topology[76], we studied whether AcrIE9 can provide a non-sequence specific protection against nucleases. Excluding this hypothesis, plasmid DNA remained sensitive to Cas9 cleavage even in the presence of high concentrations of AcrIE9 (**Supplementary Figure 5B**).

We further studied whether AcrIE9 inhibits *in vitro* interaction of *Ec*_IE Cascade with the target plasmid DNA. First, we established the range of Cascade concentrations that produce gel shift in the absence of AcrIE9 (**Figure 6B**). Next, we performed EMSA in the presence of an excess of AcrIE9. Here, two gel shifts can be observed, first associated with the DNA binding by AcrIE9, and the second shift mediated by Cascade binding (**Figure 6B**). However, the second shift required higher concentrations of Cascade, compared to conditions without AcrIE9. To prove that AcrIE9 inhibits Cascade DNA binding, we performed EMSA with a fixed concentration of Cascade and increasing concentrations of AcrIE9. This assay confirmed disruption of a complex with the target plasmid DNA, and comparable results were achieved for both, *Pa*_AcrIE9 and *Ec*_AcrIE9 (**Supplementary Figure 5C**).

*In vivo* data demonstrated that Cas7 subunit is the major AcrIE9 binding partner (**Figure 5C**). To confirm this observation *in vitro*, we performed cross-linking of the Cascade in the presence or absence of AcrIE9; a control reaction contained AcrIE9 only. When both Cascade and AcrIE9 were present in the cross-linking reaction, we observed a novel band migrating above Cas7 (**Figure 6C**). Mass-spectrometry analysis confirmed the presence of AcrIE9- and Cas7-derived peptides in this band, supporting the formation of a covalent adduct between two proteins. To investigate a possible role of other Cas subunits, we repeated cross-linking experiments with Cascade subcomplexes (Cas8-11-7-5-6, Cas11-7-5 and Cas11-7-5-6) at increasing concentrations of AcrIE9. Intensity of the fusion protein increased with AcrIE9 concentration, and this band was observed for all conditions tested, suggesting that Cas8 and Cas6 subunits are unlikely to contribute to AcrIE9 binding (**Supplementary Figure 5D**).

Together, *in vivo* and *in vitro* data support the model of AcrIE9 binding to the Cas7 subunit of Cascade, leading to compromised target DNA recognition.

## AcrIE9 can be harnessed for the control of Cas3-mediated genome editing

While single-subunit Class 2 CRISPR-Cas effectors become the standard genome editing tools, effectors of the Class 1 complexes, including Type IE Cascade+Cas3, also have promising applications. Cas3 is a processive nuclease/helicase and thus can be harnessed for the introduction of long chromosomal deletions[77–79]. Acr proteins have already proved useful for the temporal control of CRISPR-Cas activity and reduction of the off-target effects[80–82], and AcrIE9 can potentially be used to control the activity of Cascade+Cas3 complexes. To investigate this possibility, we used *E. coli* KD504 strain, encoding a self-targeting spacer (**Figure 7A**). Induction of *Ec*-IE system expression in this strain resulted in the rapid loss of cell viability, while the presence of *Ec*_AcrIE9 or *Pa*_AcrIE9 prevented cell death caused by the Cas3-mediated degradation of the host DNA (**Figure 7B**). We sequenced genomic DNA purified from cells with induced CRISPR self-targeting, and compared genome coverage obtained from cells containing empty pBAD plasmid (EV) to those expressing *Pa*_AcrIE9 (**Figure 7C**). The results reveal extensive degradation of the chromosomal DNA stretching up to 1.8 Mb downstream of the *yihN* protospacer and about 30 kB upstream, 3 hours after *Ec*-IE induction. In contrast, genomic DNA remained intact in the AcrIE9-expressing cells (**Figure 7C**, **Supplementary Figure 6**). The results confirm that AcrIE9 blocks target DNA degradation *in vivo* and demonstrate how this Acr can be used to regulate chromosomal deletions by Type IE effector complexes.

## AcrIE9 and AcrIE10 form a widespread functional module co-occurring with a novel HTH-like protein

*Ec*_AcrIE9 and *Pa*_AcrIE9 are both highly active against the *E. coli* Type IE CRISPR-Cas system. A homolog of AcrIE9, named AcrIE9.2, was also recently described in *K. pneumonia*[37]. This protein demonstrates activity against *K. pneumonia* CRISPR-Cas of the subtype IE*, that is distinct from *Ec*-IE system. Compared to the *Kp*_AcrIE9.2, *Ec*_AcrIE9 contains a single amino-acid change, while it shares only ~35% sequence identity with *Pa*_AcrIE9, which lacks a C-terminal extension (**Supplementary Figure 7A**). The corresponding sequences of the targeted CRISPR-Cas systems are also quite divergent, and Cas7 of *E. coli* shares only 32% and 27% amino-acid identity with Cas7 from *P. aeruginosa* and *K. pneumonia* (**Supplementary Figure 7B**). This suggests that despite high sequence divergence, AcrIE9 retained the ability to target a conserved interface within the Cas7 subunit of multiple Type

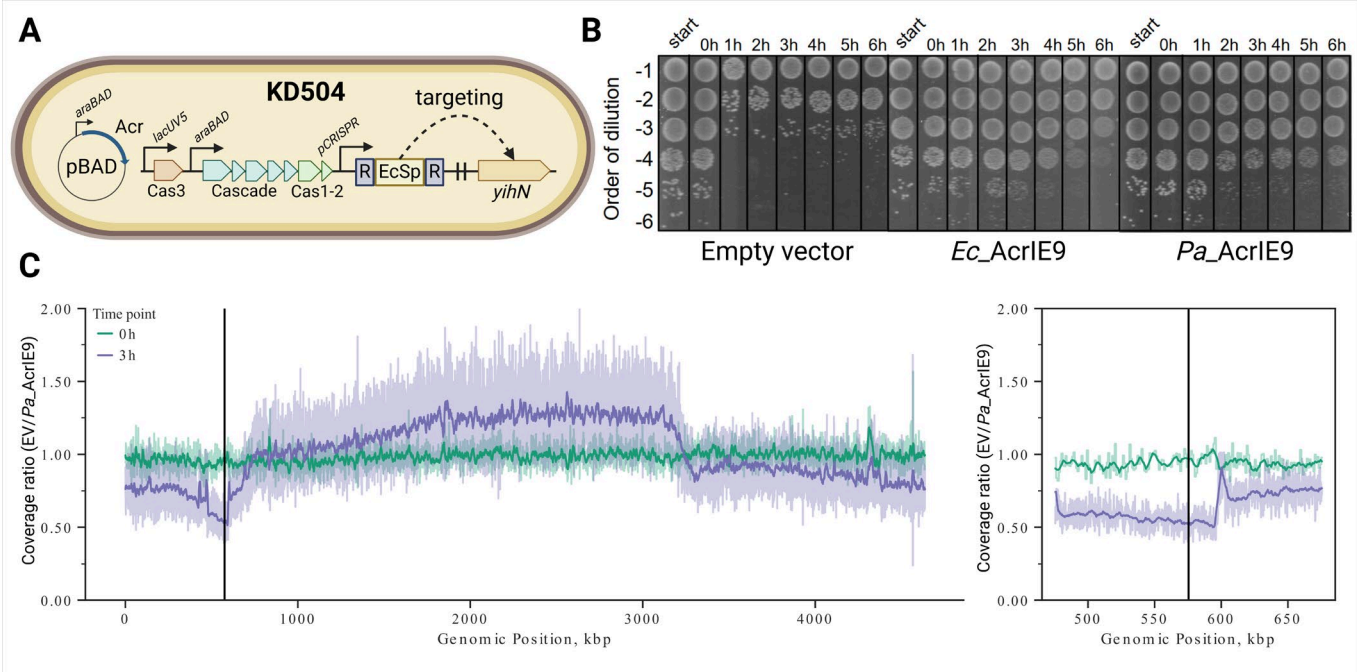


**Figure 7: AcrIE9 prevents genomic DNA degradation in a self-targeting CRISPR-Cas system. (A)** *E. coli* KD504 strain, a KD263 derivative encoding a self-targeting *yihN* spacer, was used for the CRISPR toxicity assay. Acr expression was induced with 0.2% L-arabinose, CRISPR-Cas expression was induced with 0.2% L-arabinose and 1 mM IPTG. **(B)** Cell toxicity after induction of CRISPR-Cas self-targeting. CRISPR-Cas was induced at time zero and at indicated time points cells were plated on M9 media supplemented with 0.2% glucose. Representative plates from an experiment performed in biological triplicates are presented. **(C)** Genomic DNA coverage in self-targeting KD504 cells after CRISPR-Cas induction in conditions of *Pa*_AcrIE9 expression or in the presence of the empty pBAD vector. Bold curves represent moving average for coverage ratios before and 3h after induction of CRISPR-Cas expression. Window size is 10kb for the left panel and 5kbp for the right panel zoomed in at the *yihN* protospacer region (black vertical line). Semi-transparent curves represent 1x-normalized coverage ratios for each nucleotide position.

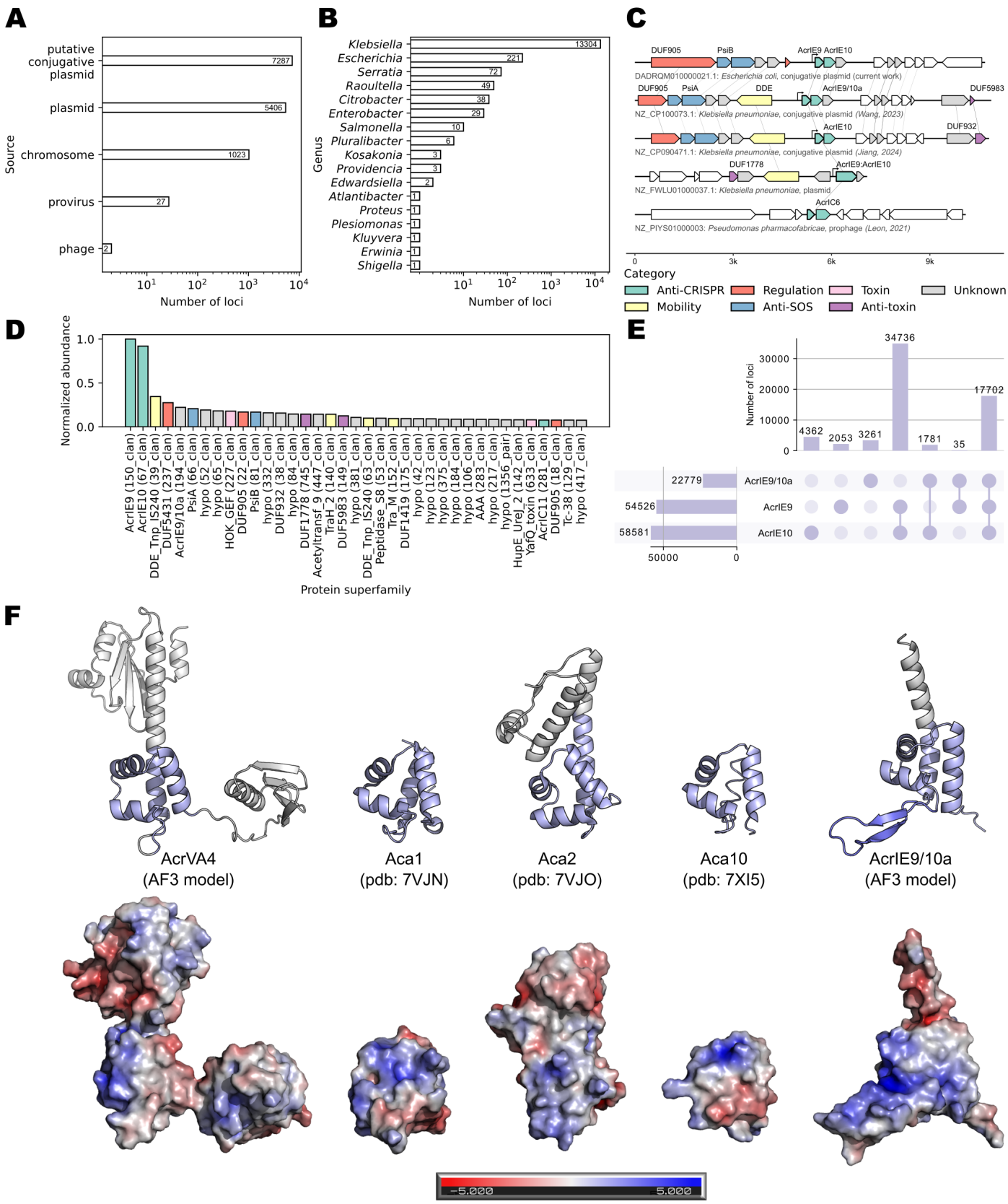


**Figure 8: AcrIE9 and AcrIE10 form a highly abundant genetic module associated with a novel HTH-like protein. (A-B)** Distribution of AcrIE9 across mobile genetic elements **(A)** and genera **(B)** within *Enterobacterales* order. **(C)** Representative genomic loci harboring *acrIE9*. Genes encoding proteins from one superfamily are connected with lines. σ70 promoters predicted with BPROM are shown as arrows. **(D)** Normalized abundance of most common protein superfamilies found in AcrIE9-encoding loci. Bars are colored according to manually assigned functional categories, identical to panel **(C)**. **(E)** Co-localization of *acrIE9*, *acrIE10*, and *acrIE9/10a* in genomes of *Enterobacterales*. Genomic distance lower than 2 kB was accounted as co-localization. **(F)** Structure and surface charge distribution of AcrIE9/10a in comparison with distantly related Aca proteins and AcrVA4. Common HTH protein fold is colored in light blue. Comparison of AcrIE9/10a protein models with other Aca types is provided in **Supplementary Figure 9**. All used Aca protein sequences are listed in **Supplementary Table 3**.

IE systems, a feature that could provide a significant evolutionary advantage.

Indeed, AcrIE9 is the most widespread Acr protein among *Enterobacterales* (**Figure 1**), and we decided to further investigate its distribution and local genomic context. We extracted all *Enterobacterales* loci encoding AcrIE9 and classified them to groups based on the content of encoded proteins (see Methods, **Supplementary Figure 8**). AcrIE9 was most frequently found encoded on plasmids and conjugative plasmids (12,693 groups), and rarely in prophages and phages (29 groups), while chromosomal occurrences (1,023 groups) may also represent unrecognized MGEs (**Figure 8A**). Conjugative plasmids are known to encode hotspots of anti-immune proteins, including Acr clusters, that are expressed during transfer of ssDNA into the recipient cell[32,83]. While we were unable to predict ssDNA-specific promoters upstream of AcrIE9 operons, the location of AcrIE9 genes might support its role in the early suppression of the CRISPR-Cas response during plasmid conjugation. AcrIE9 is most prevalent in *Klebsiella,* followed by *Escherichia* and *Serratia* (**Figure 8B**), the structures of representative loci encoding AcrIE9 are provided in **Figure 8C**.

We further investigated AcrIE9 potential functional partners based on their co-localization frequency (**Figure 8D**). The AcrIE9 was often encoded close to FlmC (DUF5431) and IS*Kpn28* DDE transposase[84], suggesting the role of this element in the Acr mobility. However, the most widespread AcrIE9 partner is AcrIE10, recently described as a novel inhibitor of the *K. pneumonia* Type IE* CRISPR-Cas system that binds to the free Cas7* subunit[38]. These proteins were almost exclusively co-encoded together in our dataset, sometimes even forming a fusion (**Figure 8C,D**). AcrIE9 co-localizes with AcrIE10 in ~96% of loci, and in ~32.5% of instances this pair is flanked by a novel gene, encoding a protein of unknown function that we tentatively named AcrIE9/10-associated or AcrIE9/10a (**Figure 8E**). HHPRED analysis identified distant similarity of this protein with multiple HTH anti-CRISPR associated (Aca) transcriptional repressors: Aca1[85], Aca2[85,86], Aca10[87], and AcrVA4[88,89] protein (**Figure 8F**).

Since Acr operons are frequently controlled by Aca proteins, we questioned whether AcrIE9/10 expression is regulated. Given its genetic location and distant similarity with known Acas, we initially considered AcrIE9/10a as a transcriptional repressor. We modeled its structure with AlphaFold3 (AF3)[90] and aligned it with known or predicted structures of all confirmed Aca proteins and AcrVA4 (**Figure 8F**, **Supplementary Figure 9A, Supplementary Table 3**). AcrIE9/10a is predicted to possess a helix-turn-helix motif with insertion of antiparallel beta-sheet, reminiscent of a winged HTH domains. Amino-acid conservation analysis confirmed the presence of beta-sheet inserts in all AcrIE9/10a homologs, while in a few instances, protein terminated in this region (**Supplementary Figure 9B,C**).

We cloned and produced AcrIE9/10a protein from a clinically-relevant *K. pneumonia* ST15 plasmid RJKP41[37], that is nearly identical to *E.coli* AcrIE9/10a (**Supplementary Figure 7C,D)**, and performed a direct EMSA experiment with 5-FAM-labeled dsDNA substrate containing cognate *acrIE9* promoter. AcrIE9/10a demonstrated weak DNA binding that was completely prevented by an excess of non-labeled competing DNA (**Supplementary Figure 10**). Thus, although HTH-like AcrIE9/10a protein demonstrates non-specific DNA binding, it is unlikely to play a role of the Aca repressor for the AcrIE9/10 operon.

Our analysis establishes AcrIE9/AcrIE10 as the most widespread anti-CRISPR locus in sequenced *Enterobacterales* genomes and reveals an unusual co-occurrence of two Acr proteins targeting the same subunit of IE Cascade complex. In addition, we identified a novel protein with HTH-like architecture as a frequent component of the AcrIE9/AcrIE10 locus. Further research is needed to establish the role for each component of this locus and determine details of its transcriptional regulation.

## Discussion

In this study, we aimed to identify an anti-CRISPR protein active against the model Type IE CRISPR-Cas system of *Escherichia coli* K12. The biological role of this system remains enigmatic due to H-NS-mediated silencing of *cas* promoters and the absence of known conditions leading to activation of interference or adaptation[47,48]. However, the natural diversity of *E. coli* Type IE spacers suggests potential involvement in conflict[60], which would imply that MGEs infecting *E. coli* could encode Acr proteins promoting their coexistence with the resident CRISPR-Cas systems. To test this hypothesis, we conducted a large-scale analysis of known Acr protein distribution within the *Enterobacterales* order, revealing an abundance of potential inhibitors targeting IB, IC, IE, IF, IIA, and VA CRISPR-Cas systems (**Figure 1B**). This finding aligns with the fact that many *Enterobacterales* species (e.g., *Klebsiella*, *Serratia*, *Pectobacterium*) serve as model organisms in studies of naturally active CRISPR-Cas systems[91,92]. Among Type IE inhibitors, AcrIE9 and AcrIE10 emerged as the most prevalent. We experimentally evaluated 11 representative AcrIE and AcrIF homologs against *E. coli* K12 Type IE system and demonstrated that AcrIE9 exhibits robust anti-CRISPR activity.

Originally identified as a prophage-encoded inhibitor of *Pseudomonas aeruginosa* Type IE CRISPR-Cas system[23], AcrIE9 was recently implicated in the dissemination of $bla_{KPC}$ carbapenem resistance plasmids in *Klebsiella pneumoniae* ST15[37]. This homolog (designated AcrIE9.2) shares ~35% amino acid sequence identity with the *P. aeruginosa* variant and inhibits the Type IE* CRISPR-Cas system. *E. coli* AcrIE9 discovered in this work is nearly identical to AcrIE9.2 and is also encoded on conjugative plasmid, that are likely transferred between *Klebsiella* and *Escherichia*. Remarkably, despite low sequence conservation between the Cas proteins of *P. aeruginosa*, *E. coli*, and *K. pneumoniae* (**Figure S7**), *Pa*_AcrIE9 and *Ec*_AcrIE9 effectively inhibit diverse systems. This suggests that Acr proteins often evolve to target the most conserved interfaces of CRISPR-Cas complexes, in line with the evidence that non-homologous Type IE and IF Acrs, repeatedly target the same surfaces of the Cascade complex[15].

Does the discovery of AcrIE9 as a potent inhibitor of *E. coli* K12 Type IE CRISPR-Cas support the biological role of this system in anti-MGE immunity? Although AcrIE9 is found in

726 *E. coli* genomes, its prevalence in *Klebsiella* (>52,000 genomes, **Extended Data File 1**) suggests stronger selection pressure from *Klebsiella* CRISPR-Cas systems, while *E. coli* could serve only as a temporary reservoir of AcrIE9-encoding conjugative plasmids. Otherwise, it can be speculated that, unlike the model K12 strain, other strains of *E. coli* could encode active Type IE CRISPR-Cas systems that drive the accumulation of Acr proteins in mobile elements.

Characterization of the anti-CRISPR mechanism revealed that AcrIE9 binds the Cas7 subunit of the assembled Cascade complex, disrupting target DNA recognition both *in vivo* and *in vitro*. Intriguingly, AcrIE9 itself exhibits DNA-binding activity, though this does not confer nonspecific protection against other nucleases, such as Cas9. While it remains unclear whether DNA binding activity contributes to the anti-CRISPR function, it can explain the moderate toxicity of the protein. Other Acrs, like AcrIF9, can promote Cascade complexes binding to non-specific DNA[93,94]. Our results indicate decreased affinity of the *Ec*Cascade complexes to the target DNA, yet we cannot exclude formation of the triple Cascade:AcrIE9:DNA complexes with non-target DNA.

We have shown that AcrIE9 is sufficient to inactivate Cascade complexes *in vivo*. However, AcrIE9 is almost invariably co-encoded with AcrIE10, another Type IE inhibitor recently reported to target *K. pneumoniae* Type IE and IE* CRISPR-Cas systems[38]. AcrIE10 binds the free Cas7 subunit to prevent Cascade assembly, and it is not clear if this mechanism provides protection from pre-existing surveillance complexes during the entry of conjugative plasmid to the host cell. The co-occurrence of two Acrs targeting the same subunit implies a novel strategy for CRISPR-Cas suppression, warranting further study of this widespread anti-CRISPR locus.

We hypothesize that two Acrs may function at distinct stages of plasmid conjugation: AcrIE9 could neutralize preassembled complexes at the onset of DNA transfer, while AcrIE10 could prevent the assembly of novel Cascade complexes at later stages. The clustering of multiple Acrs and Racrs, targeting the same CRISPR system, in other mobile elements suggests that coordinated suppression with multiple inhibitors may represent a widespread anti-CRISPR strategy[18,19]. Alternatively, AcrIE9 and AcrIE10 can act synergistically to achieve complete inhibition of CRISPR interference at the early stage of conjugation.

Stable coexistence of MGEs and CRISPR-Cas systems necessitates tight regulation to balance the negative costs of toxic Acr production against the risk of CRISPR-mediated cleavage[36]. Typically, Acr transcription is controlled by cognate Acr-associated (Aca) helix-turn-helix (HTH) transcription factors, and 10 Aca classes have been experimentally validated[95]. We identified an HTH-like protein, AcrIE9/10a, in one third of AcrIE9/10 loci, though it lacks sequence-specific binding affinity for *acrIE9* promoter and its function remains to be shown.

In summary, we analyzed anti-CRISPR proteins distribution in *Enterobacterales*, identified AcrIE9 as a potent inhibitor of *E. coli* K12 Type IE CRISPR-Cas system, and elucidated its mechanism of action via Cas7 binding and inhibition of the surveillance complex interaction with target DNA. Our study highlights that AcrIE9 is genetically linked with AcrIE10, thus forming the most prevalent anti-CRISPR locus in *Enterobacterales* MGEs. While the individual components of this operon have been studied[37,38], future work should elucidate their collective roles under native expression conditions in conjugative plasmids. This task is particularly urgent given the role of AcrIE9/10 loci in the dissemination of antimicrobial resistance genes in pathogenic bacteria[37,38].

## Materials and Methods

### Bacterial strains and phages

All strains of bacteria, plasmids, and phages used in this study are listed in **Supplementary Table 4**. For the initial test of anti-CRISPR activity, the KD263 strain[66] and its derivatives were used. KD263 contains a g8 spacer targeting a protospacer in the M13 phage genome. The *cas3* gene expression is driven by the *lac* promoter, the remaining *cas* genes are transcribed from the *araBAD* promoter. SS80 is a derivative of KD263, with a spacer targeting T7 phage[44]. KD454 is a derivative of KD263 carrying a deletion of the *cas3* gene. Self-targeting experiments were performed in KD504, a derivative of KD263 with spacer targeting *yihN* gene. For testing Acrs in a natively active *Ec*-IE system, a pair of BW25113 derivative strains was used. BW39651 is a BW25113-derivative encoding T3 spacer targeting phage lambda (*λ*), while BW39671 is the same strain with the deletion of *hns* repressor gene[45]. Acr activity against the Ec_IF system was initially investigated in the LF82 *E. coli* strain[67]. Plasmid-encoded Ec_IF system was investigated in a CRISPR-deficient *E. coli* BB101 strain, susceptible to Mu phage infection[68].

Bacterial cells were grown in LB (1% Bacto tryptone; 0.5% Bacto Yeast Extract; 1% NaCl) medium at 37°C overnight. Long-term storage stocks were prepared in LB supplemented with 10% glycerol and stored at -70 °C. To obtain a phage stock, an overnight bacterial culture was diluted 100-fold in LB medium and incubated at 37°C until $OD_{600}$ reached 0.6. The culture was infected with phage at low MOI and further incubated overnight at 37°C. For phage *λ* production, the culture was supplemented with 0.2% maltose and 10 mM MgCl2. Phage M13 was produced on F' KD263 host. Phage Mu was produced from a lysogenic culture, upon switching from 30°C to 45°C for 45 minutes at $OD_{600}$=0.6. After induction, cells were grown at 37°C until lysis. The resulting lysates were purified from cell debris by centrifugation (6000g, 10 min, 4°C) and treated with chloroform (10 µl per 1 ml of lysate). Phage stocks were stored at +4°C.

### Plasmid construction

All primers used in this study are listed in **Supplementary Table 5**. Candidate anti-CRISPR genes were synthesized and cloned in a pBAD vector by Proteogenix company (France). Strep-tag II coding sequence with GSG linker (GSGSHPQFEK) was fused to the C-terminus of *Ec*_AcrIE9 and Pa_AcrIE9, encoded on pBAD via Gibson assembly (Gibson Assembly® Master Mix, NEB). Plasmid pFRG8 for Cascade-mediated sfGFP transcription silencing was constructed by KLD mutagenesis with the use of KLD reaction mix (NEB). Similarly, plasmid pFRC1T, promoting primed adaptation, was constructed from pFRG8 using KLD reaction. To enhance Acr expression, a variant of the pBAD_Ec_AcrIE9/Pa vector with modified RBS site was constructed via KLD reaction. These vectors were used in the pull-down experiments. For Ec_IF system expression, plasmid pET21d-LF82IFCR, kindly provided by A. Davidson, was modified to make it compatible with the pBAD vector. The ampicillin resistance gene was substituted with the streptomycin resistance gene via Gibson assembly (Gibson Assembly® Master Mix, NEB).

For heterologous expression and preparative purification of *Pa*_AcrIE9 protein, corresponding *acr* gene was re-cloned in pET28 vector under control of T7 promoter and in frame with C-terminal 6His-Tag sequence. *Ec*_AcrIE9 gene was re-cloned from plasmid pBAD-AcrIE9_Ec using primers 5'-GGAATTAATCATGAACTTCACCATTAAGAGTCGC-3' and 5'-GCAGATCTCGAGTTACTGATTATCTTCGC-3' into the expression vector pET30 restricted by NcoI and XhoI enzymes. AcrIE9/10a from RJKP41 plasmid was cloned with primers 5'-CATGCCATGGCCATGAGTGACGTACTTTCCTG-3' and 5'-CCGCTCGAGTTACTTATCATTGTTTTCACTCACATTATTCGTTTCAAACGAGACAAG-3' into pET28a-SUMO2 cloning vector restricted by NcoI and XhoI enzymes. The target plasmids pJ3 and pλ1-on-target were generated for gel shift binding assay with J3-Cascade and cleavage activity assay for SpCas9 nuclease, respectively. The pT7Blue-based plasmid pG8_dir carrying a 209-bp M13 fragment[69] was used to replace g8 protospacer sequence 5'-GCTGTCTTTCGCTGCTGAGGGTGACGATCCCGC-3' with J3 protospacer 5'-

GCCAGTGATAAGTGGAATGCCATGTGGGCTGTC-3' resulting in pJ3 plasmid[48]. The plasmid pλ1-on-target was obtained by blunt-end cloning synthetic DNA containing Cas9 protospacer[96] in pT7Blue vector. For preparation of J3-Cascade (Cascade loaded with crRNA containing J3 spacer) we constructed pCDF-casBCDE(J3) plasmid by replacing g8 spacer in pCDF-casBCDE(g8)[97] with J3 spacer[48].

All constructs were transformed into XL1 Blue (Evrogen, Russia) or NEB® 10-beta (NEB) commercial chemically competent *E. coli* cells. Sequences of assembled plasmids were confirmed by Sanger sequencing (Genotech, Russia).

### Efficiency of plating (EOP) assay

To estimate the level of CRISPR-Cas defense and Acr activity, the titer of phage particles was measured by the double agar plating method. Plates were prepared with two layers: the bottom layer contained 1,5% LB agar media, while the top layer was mixed with overnight bacterial cell culture and included corresponding antibiotics and inductors, if necessary (1 mM IPTG, 0,2% Arabinose). For phage λ infection, the top layer of agar also contained 0.2% maltose and 10 mM $MgCl_2$; for phage Mu infection, the top agar was supplemented with 10 mM $MgCl_2$. Phage lysate was spotted on agar top in serial ten-fold dilutions. Plates were incubated overnight at 37°C. EOP with BW39651 and BW39671 strains were carried at 30°C. EOP with the phage T7 were carried at room temperature.

### Monitoring dynamics of phage infection in liquid culture

The overnight liquid cultures were diluted 100-fold and grew in liquid LB media at 37°C until $OD_{600}$ = 0.6 with required antibiotics and in the presence/absence of inductors. The cultures were equilibrated to $OD_{600}$ = 0.3 and 200 µl of equilibrated culture was loaded to 96 well plates and infected with the phage at indicated MOI (Multiplicity of Infection). The plate was loaded to the EnSpire Multimode Plate Reader (PerkinElmer, USA). Optical density was monitored for 20 hours at 37°C with intense orbital shaking (200 rpm).

### Phage release

To measure phage burst and accumulation of phage particles in liquid media, an extended one-step growth curve assay was performed. The overnight cultures of *E. coli* BW39671 encoding λT3 spacer (and BW39651 as control) containing pBAD plasmids encoding Acrs were diluted 100-fold in 5 ml in liquid LB media supplemented with corresponding antibiotics before $OD_{600}$ reaches 0.6. The cultures were mixed with phage $\lambda_{vir}$ to achieve MOI=$10^{-3}$. Aliquots were collected immediately after phage addition, and 1, 3, 5 hours post-infection. 600 µl of liquid media was collected, centrifuged at 13.000g for 5 minutes to remove cell debris and supernatant was plated in serial ten-fold dilutions on a lawn sensitive BW25113 *E. coli* strain to enumerate phage titer. The experiment was performed in three replicates.

### Self-toxicity inhibition assay

The overnight culture of *E. coli* KD504 cells was diluted 100-fold and grew in liquid LB media supplemented with corresponding antibiotics before $OD_{600}$ reaches 0,2-0,3. At this point, 0,2% L-Arabinose and 1 mM IPTG were added to induce CRISPR and Acr expression. At every chosen time point 100 µl of cultures were collected from each sample and plated in a serial ten-fold dilutions on plates with 1% agar M9 media. M9 was prepared by diluting autoclaved 5x M9 salts (250mM $Na_2HPO_4$, 100mM $KH_2PO_4$, 50mM NaCl, 10mM $NH_4Cl$) in mQ and mixing with $MgSO_4$ (2 mM final concentration) and $CaCl_2$ (0.1 mM final concentration). This media was additionally supplemented with 1% agar and 0,2% D-Glucose as carbon source. Plates were incubated at 37°C overnight. For NGS sequencing, 5 mL of cell culture aliquots were collected to extract genomic DNA. Genomic DNA purification was performed with Genomic DNA Purification Kit (Evrogen) under manufacturer protocol.

### Naive adaptation assay

*E. coli* BL21AI culture was co-transformed with plasmids pCDFCas1-2 and pBAD empty vector or pBAD_Pa_AcrIE9/Ec by electroporation. The next day, a single colony from each plate was inoculated into liquid LB media with corresponding antibiotics and grew until $OD_{600}$ = 0,2-0,3. Cells were pelleted by centrifugation (3.200 g at room temperature) and put in 5 ml of fresh LB media without antibiotics. Cas1-2 and Acr expression was induced with 1 mM IPTG and 0,2% L-arabinose and following induction 100µl cell aliquots were collected each hour by centrifugation at (3.200 g at 4°C) and stored at -80°C in 10% glycerol. The naive adaptation was detected by PCR amplification from 100-fold diluted bacterial culture, using leader - repeat pair of primers Moji3 and Moji4 (**Figure 4A**, **Supplementary Table 5**). PCR products were loaded on 2% TAE Agarose gel and ran for 1 hour at 70W.

### Primed adaptation assay

*E. coli* KD263 culture was co-transformed with pFR66_G8(C1T) and pBAD empty vector or pBAD_Pa_AcrIE9/Ec by electroporation. The next day, a single colony from each plate was inoculated into liquid LB media with corresponding antibiotics and grew overnight at 37°C. Overnight culture was diluted 100-fold and grown in a fresh LB media without antibiotics until $OD_{600}$ reaches 0.2-0.3. Then, 0,2% L-arabinose and 1 mM IPTG were added to induce CRISPR-Cas and Acr expression. 3h and 24h after induction, cell aliquots were collected by centrifugation at (3.200 g at 4°C) and frozen at -80°C in 10% glycerol. The primed adaptation was detected from 100-fold diluted bacterial culture, by PCR amplification using leader - spacer (G8) pair of primers EcLdr_F and M13G8_R (**Figure 4B**, **Supplementary Table 5**). PCR products were loaded into 2% TAE Agarose gel and ran for 1 hour at 70W.

### sfGFP transcriptional silencing

To investigate *Ec*-IE Cascade DNA binding activity *in vivo*, we measured sfGFP expression from a promoter overlapping with *Ec*-IE protospacer. *E.coli* KD263 strain with G8 spacer carried pBAD empty vector or pBAD encoding Pa_AcrIE9/Ec, and sfGFP-encoding plasmid with (pFR66) or without (pFR66) a G8 protospacer. The overnight liquid cultures were diluted 100-fold and grew in LB supplemented with required antibiotics/inducers at 37°C until $OD_{600}$ reaches 0.6. The cultures were equilibrated to $OD_{600}$ = 0.3 and 200 µl was loaded to 96 well plate. Cell growth continued in EnSpire Multimode Plate Reader (PerkinElmer, USA) at 37°C with monitoring sfGFP fluorescence (Ex λ = 480, Em λ = 510) and absorbance ($OD_{600}$). The resulting fluorescence data was normalized to an optical density. The experiment was performed in three replicates.

### Protein pull-down

Overnight culture of *E. coli* BW25113 carrying pBAD_Pa_AcrIE9/Ec C-Strep was diluted 100-fold in 3 L of LB media with corresponding antibiotics. When the culture reached $OD_{600}$ = 0.2-0.3 CRISPR-Cas and C-terminally Strep-tagged AcrIE9 were induced with 1 mM IPTG and 0,2% L-Arabinose were added and cells grew overnight at 37°C. Thereafter, cells were spun down by centrifugation at 3500 g for 20 min at 4°C and cell pellets were resuspended in StrepA buffer (20 mM Tris-HCl pH 8.0, 50 mM NaCl, 5 mM β-mercaptoethanol, 1 mM EDTA) supplemented with 5 mg/ml of lysozyme and protease inhibitor cocktail (Roche). Cells were lysed by sonication on ice for 30 min with using 10s impulse - 20s rest circle at power 20% (Qsonica sonicator with 6 mm sonotrode). Cell debris was removed by centrifugation at 25.000 g for 45 min at 4°C. Purification of Strep-tagged protein together with suggestive partners was carried out with a BioRad NGC Chromatography System on a 5 mL StrepTrap HP column (GE Healthcare). Before applying the bacterial lysate, the column was equilibrated with a StrepA buffer. Proteins bound to the column were eluted with a StrepB buffer (StrepA supplemented with 2.5 mM desthiobiotin). The protein composition of the eluted fractions was analyzed by Laemli SDS-PAGE after concentration with 3-kDa ultra centrifugal Amicon filters. The identity of protein bands was determined by matrix-assisted laser desorption/ionization time-of-flight (MALDI-TOF) mass spectrometry. Samples were treated with Trypsin Gold (Promega) in accordance with manufacturer's instructions. Mass spectra were obtained using the rapifleX system (Bruker).

To purify Ec_IE Cascade from cell culture expressing untagged *Pa*_AcrIE9, *E.coli* BL-21 AI cells were transformed with plasmids pWUR515[98] (Strep-CasC,CasD,CasE), pWUR407, and EcCRISPR P7-7x, and pET30-AcrIE9_Pa, and plated on LB agar supplemented with appropriate antibiotics. Expression of *Pa*_AcrIE9, Cascade and CRISPR RNA was induced using 0.2% Arabinose and 0.5mM IPTG when cells reached an $OD_{600}$ of 0.4. After expressing at 37 °C for 16 hours, cells were harvested by centrifugation at 5,000G for 20 minutes. Cell pellets were resuspended in lysis buffer containing 20 mM Tris (pH 8.0), 100 mM NaCl, and 1 mM DTT before lysing by sonication. Cellular debris was removed by centrifugation at 26,892g for 30 min and the cleared lysate was loaded onto a 5mL StrepTrap HP column. After washing with 15 column volumes of lysis buffer, the Cascade complex was eluted using lysis buffer supplemented with 50mM desthiobiotin. Elutions were concentrated before being resolved on a Superose 6 10/300 size exclusion column equilibrated in SEC buffer (lysis buffer + 5% glycerol). Fractions from the main peak were concentrated to ≈6uM and proteins were visualized on a 12% acrylamide gel via SDS-PAGE.

### Protein purification

Cascade (Cas8,11,7,5,6) or Cascade subcomplexes lacking Cas8 (Cas11,7,5,6), or Cas8 and Cas6 (Cas11,7,5) were prepared from E. coli KD418

cells co-expressing appropriate *cas* genes with a source of crRNA (**Table S4**) and were one-step affinity-purified on Strep-Tactin® column (IBA) using the N-terminal Strep-tag attached to the Cas11 subunit[99] according to manufacturer's protocol.

N-terminal 6His-tagged Cas8 protein was IMAC purified from *E.coli* BL21 Star (DE3) pLysS strain (Invitrogen) transformed with pET30-*casA* expression plasmid (**Supplementary Table 4**) followed by gel-filtration on Superose-6 column (Pharmacia) equilibrated with 40 mM Tris-HCl buffer (pH8) containing 200 mM NaCl, 5% Glycerol, 1mM TCEP.

C-terminal 6His-tagged *Pa*_AcrIE9 protein was IMAC purified from *E.coli* Rosetta 2 (DE3) strain (Novagen) transformed with pET28-*Pa*_AcrIE9 expression plasmid (**Supplementary Table 4**) followed by ultrafiltration through VivaSpin 5 kDa MWCO concentrator (EMD Millipore) for concentrating and buffer exchange in 40mM Tris-HCl, pH 7.5, 300 mM NaCl, 1 mM TCEP, 5% Glycerol.

N-terminal 6His-tagged AcrIE9_Ec protein was IMAC purified from *E.coli* Rosetta 2 (DE3) strain (Novagen) transformed with pET30-AcrIE9_Ec expression plasmid (**Supplementary Table 4**) followed by ultrafiltration through VivaSpin 5 kDa MWCO concentrator (EMD Millipore) for concentrating and buffer exchange in 40mM Tris-HCl, pH 7.5, 300 mM NaCl, 1 mM TCEP, 5% Glycerol. Before use in the binding reactions with target plasmid DNA, the N-terminal 43 aa tag has been removed by Enterokinase, light chain (NEB) according to the manufacture protocol.

N-terminal 6His-SUMO-tagged AcrIE9/10a protein was IMAC purified from *E.coli* BL21 (DE3) strain (Novagen) transformed with pET28a-AcrIE9/10a expression plasmid (**Supplementary Table 4**) followed by ultrafiltration through VivaSpin 5 kDa MWCO concentrator (EMD Millipore) for concentrating and buffer exchange in 40mM Tris-HCl, pH 7.5, 300 mM NaCl, 1 mM TCEP, 5% Glycerol. Before use in the binding reactions with target DNA, His6-SUMO tag was removed by SENP1 protease digestion overnight followed by Ni-NTA purification to obtain tag-free AcrIE9/10a.

SpCas9 nuclease from *Streptococcus pyogenes* was expressed in *E.coli* Rosetta 2 (DE3) strain transformed with pMJ806 (Addgene) and purified by protocol published earlier[100]. To obtain catalytically active Cas9 nuclease we used synthetic guide RNAs: tracrRNA and crRNA ordered in IDT (**Supplementary Table 5**). Reconstituted complex SpCas9/tracrRNA/crRNA was validated for binding and cleavage activity towards a target plasmid DNA.

### Glutaraldehyde crosslinking between AcrIE9 and E.coli IE Cascade

Crosslinking between *Pa*_AcrIE9 and Cascade or its sub-complexes was initialized by addition of glutaraldehyde (GA) to a final concentration of 0.5% into the mix of 20 µM *Pa*_AcrIE9 and 1 µM Cascade in the buffer containing 40 mM MOPS, pH 8.0, 50 mM NaCl at room temperature for 2-3 min. Then reaction was quenched by addition of TRIS solution to a final concentration of 100 mM. Crosslinking products were resolved by denaturing SDS-PAGE and analyzed by mass spectrometry.

### Electrophoretic Mobility Shift Assay (EMSA)

J3-Cascade was mixed with *Pa*_AcrIE9 or *Ec*_AcrIE9 in buffer containing 40 mM Tris-HCl (pH7.5), 50 mM NaCl, 0.5 mM TCEP, and 0.1 mg/ml BSA for 10 min at 37°C, then target plasmid DNA pJ3 was added and incubated for 20 min at 37°C. Samples containing Cascade:DNA complexes engaged by AcrIE9 were run in 0.7% TBE-agarose gel for 3 hours at 60V, 30 mA, 2W. Then the gel was stained with EtBr and visualized by GelDoc XR+(Bio-Rad).

Binding reaction between *Pa*_AcrIE9 and 52-bp short single-stranded and double-stranded oligos labeled by fluorescein (**Supplementary Table 5**) were carried out in the same binding buffer and samples were loaded on native precast 10 to 20% gradient polyacrylamide gel (Novex, Invitrogen), and run in tris-glycine buffer at 20 mA for 2 hours. The bound complexes were visualized with GelDoc XR+(Bio-Rad).

Binding reaction between AcrIE9/10a and *acrIE9* promoter reaction was performed using 300 bp PCR product amplified with 5-FAM labeled forward primer (**Supplementary Table 5**). EMSAs were performed in a reaction volume of 20 µl containing binding buffer (5% glycerol,pH 7.8 20 mM Tris HCl ,150mM KCl, 2 mM $MgCl_2$, 0.1µg/µl BSA,0.5 mM DTT), 20 nM of *acrIE9* promoter substrate and a gradient concentration of Aca14 (range 0–800 nM). 0.5µl of herring sperm DNA (5mg/ml, Sangon Biotech (Shanghai) Co., Ltd.), was used to exclude non-specific binding of the protein to DNA. Binding reactions were mixed on ice and incubated for 30 min. Then 16 µL of each reaction was loaded on a 5% Tris-Glycine native PAGE gel and run at 130 V for 55 min on ice. An Amersham TyphoonTM 5 Biomolecular Imager was used for gel analysis.

### SpCas9 nuclease cleavage activity assay

For cleavage experiments, 75 nM of SpCa9 and gRNA (tracrRNA+crRNA) were preincubated at room temperature for 10 min in binding buffer containing 40mM Tris-HCl(pH7.5), 100 mM NaCl, 1mM DTT, 5mM MgCl2, before initiating reaction by addition of 1.2 nM of target plasmid DNA pl1-on-target with or without 14 µM of AcrIE9 for indicated time at room temperature. Cleavage products were analyzed by 0.8% agarose gel electrophoresis and visualized by GelDoc XR+(Bio-Rad) after EtBr staining.

### Proteomic analysis

Protein-containing fractions obtained from pull-down with Strep-tagged AcrIE9 variants (or non-tagged proteins as controls) were loaded on 5% SDS-PAGE after boiling for 10 minutes in standard SDS-loading dye. Proteins were allowed to completely enter the resolving gel then the current was stopped. The first 3-4 mm of resolving gel containing the samples were sliced and digested with Trypsin Gold (Promega). To elute peptides, gel was incubated in a microtube shaker with 50 µl of 50% acetonitrile/5% formic acid for 45 min at room temperature. Liquid was removed and extraction was repeated. Final extraction step was carried with 50 µl of 90% acetonitrile/5% formic acid for 5 min. Collected peptides were vacuum dried and solubilized in 5 µl of 0.1% formic acid. Peptides were subjected to LC-MS/MS analysis with Q Exactive HF-X mass spectrometer (Thermo Scientific) as described before[101]. Raw data was processed with MaxQuant software and peptides were searched against *E.coli* BW25113 proteome extended with AcrIE9 sequences.

### NGS library preparation and sequencing

Sequencing libraries were constructed from 1 µg of total genomic DNA using MGI Easy PCR-Free Library Prep Set in accordance with User Manual v1.1 (MGI Tech Co., Shenzen, China), size-selected and cleaned up with the provided DNA Clean Beads. Quality control was performed with Agilent High Sensitivity D1000 ScreenTape Assay on 4200 TapeStation System (Agilent Technologies Inc., Santa Clara, CA, USA) and QuDye ssDNA kit (Lumiprobe, Moscow, Russia). Barcoded libraries were pooled and sequenced using DNBSEQ-G400 in 2x150bp PE mode.

### Self-targeting NGS data analysis

The obtained reads were trimmed and filtered with fastp v. 0.23.2[102]. Then reads were aligned on a reference genome (*Escherichia coli* KD403[103] and cloning vector pBAD) with BWA mem v. 0.7.18[104]. Optical duplicates were marked with Picard v. 2.18.7. The alignments were sorted and filtered with samtools v.1.21[105] to retain only primary aligned paired reads (`-f 2 -q 20 -F 1024,256,2048`). The coverage was extracted for each position in bacterial genome and normalized with 1x normalization with bamCoverage utility from deepTools v.3.5.5[106]. Effective genome size was calculated with `unique-kmers.py` from khmer v.2.1.1[107]. Ratios of normalized coverage for samples with/without *Pa*_AcrIE9 expression in each time point were calculated with bigwigCompare from deepTools.

### Identification of known anti-CRISPRs within *Enterobacterales* order

***Search for known anti-CRISPR proteins in non-redundant protein database.*** An exhaustive search of known Acr proteins in non-redundant protein database was performed. hmmsearch v.3.3.2 (http://eddylab.org/software/hmmer/) was used to scan non-redundant protein sequences of bacteria, viruses, and plasmids (taxid: 2, 10239, and 36549 from nr BLAST, v.5) against HMM profiles of Acr proteins from a database of anti-prokaryotic immune systems (dbAPIS) v.1[35]. The protein hits with E-value less than 1e-10 were considered as Acrs and were included in further analysis.

***Taxonomic and genetic localization assignment for acr genes.*** For each obtained hit we explored contigs, in which they are encoded, and defined their taxonomy based on sequence metadata. Taxon ID for genome assemblies were obtained with NCBI datasets toolkit v.16.21.0[108], and the corresponding lineages were fetched with taxonkit v. 0.14.2[109]. Contigs encoding potential Acrs of *Enterobacterales* were selected, and mobile genetic elements were predicted with geNomad v. 1.8 using end-to-end pipeline[110].

***Loci extraction.*** Genome assemblies of *Enterobacterales* contigs harboring Acrs were fetched with `datasets download` tool from NCBI datasets toolkit v.16.21.0. Contigs encoding Acrs were clustered with MMseqs2 v. 15-6f452 cluster[111] to remove redundancy with following parameters: minimal coverage 0.9, minimal sequence identity 0.99, clustering mode 0, coverage mode 0. Then the nucleotide sequences of Acrs genomic contexts (5 kb upstream and

downstream) were extracted and used for further analysis of Acrs genomic context.

### Analysis of AcrIE9 genomic context

***Search for proteins sequence homology groups.*** Coding regions in extracted genomic loci containing Acr hits were predicted with prodigal v. 2.6.3[112]. Predicted protein sequences were clustered with MMSeqs2 v. 15-6f452[111] with the following parameters: minimal sequence identity 0.3, minimal coverage 0.8, clustering mode 0, coverage mode 0. Set of experimentally validated Acr sequences were added to the clustering procedure to remove possible false-positive hmmsearch hits. Next, multiple sequence alignments (set of MSAs) for each protein cluster were extracted, and MSAs for clusters with more than 3 representatives were aggregated into a HH-suite3 database v. 3.3.0[113]. Set of MSAs were searched against formed database of HMM-profiles resulting in a list of pairs of homologous clusters (E-value threshold 5e-5). Communities, or locally dense subgraphs, were obtained from this similarity network with ClusterONE v.1.2[114], and were considered as protein clans (or protein superfamilies).

***Loci diversity analysis.*** The diversity of local genomic contexts for AcrIE9 was described as follows. To leverage the effect of disproportionate abundance of different bacterial species sequences, groups of loci similarity were defined. First, the target loci were represented as sets of protein clans, encoded in these loci. Second, Jaccard similarity between each pair of loci was calculated. Third, for estimation of similarity threshold, Jaccard similarity for 200 permuted sets were calculated (protein clans were randomly permuted among loci). The value of 0.01 of permuted 99th quantiles observed in permutations was selected as a threshold (0.4), and all pairs of loci with Jaccard similarity more than the threshold were considered as edges within the loci content similarity network. Same as for protein cluster similarity network, communities of the network were found. For each of 67 found communities, the within-community abundance of protein clans was defined as the proportion of loci within the community, in which the genes corresponding for protein clans are present. Then the normalized abundance was calculated for each protein clan as a sum of within-community abundance divided by the number of communities. Given that the abundance of protein groups is calculated only for loci encoding AcrIE9 homologs, the higher normalized abundance reflects higher AcrIE9 co-localization rate with respect to a more variable 10 kbp genomic context. For the analysis of genomic context, only proteins with normalized abundance higher than 0.075 were selected.

***Construction of HMM profiles.*** For the analysis of AcrIE9, AcrIE10 and AcrIE9/10a co-occurrence, we performed an HMM profile search against the nr BLAST database. To build HMM profiles, target protein sequences were extracted for each protein superfamily with seqkit v2.8.2[115], and clustered with MMSeqs2 to remove redundancy with following parameters: minimal coverage 0.9, minimal sequence identity 0.95, clustering mode 0, coverage mode 0. Resulting non-redundant sequence sets were aligned with MAFFT v.7.490[116], and then converted into HMM profiles with hmmbuild from HMMER package.

***Protein sequences annotation.*** Alignments of protein clusters obtained with MMSeqs2 were searched against Pfam-A database[117] with hhsearch tool from HH-Suite3 v.3.3.0[113] (E-value threshold 1e-5). Then the hit with lowest E-value was assigned as a true hit for each protein cluster. Of note, such an approach excludes assignment of more than one found domain to the protein cluster.

***AcrIE9 promoters' prediction.*** Upstream regions of AcrIE9 gene for four representative loci were selected with bedtools getfasta v2.30.0[118] (-200nt and +50nt with respect to gene start). The promoter regions within obtained nucleotide sequences were predicted with BRPROM[119].

***AcrIE9/10a identification and structure comparison analysis.*** Amino acid sequence of AcrIE9/10a protein was queried to the HHPRED WebServer[120]. After the homology was established, the protein structure was predicted with AlphaFold3[90] (WP_013023810.1 was used as a representative AcrIE9/10a), and the obtained molecule model was searched with DALI[121] against the PDB database. To evaluate whether AcrIE9/10a belongs to one of the previously described Aca classes, the protein structure comparison was performed. Protein structures of all known Acas were modeled with AlphaFold3 (Protein sequences/structures listed in **Supplementary Table 3**). To build **Figure 8F**, alignment of 3D structures was performed with jFATCAT[122] in a rigid mode (all-vs-AcrIE9/10a), while for the analysis presented at the **Figure S9A**, FoldMason WebServer[123] in all-against-all mode was used. Resulted structural alignments were visualized with PyMOL v.3.0.0. The electrostatic potentials obtained with APBS[124]. For the coloring of protein structures according to LDDT for MSA, the custom python3 script was utilized, and available on GitHub (https://github.com/ombystoma-young/annotationtoolkit/foldmason2pdb). The global alignment of homologs of AcrIE9/10a was performed with MAFFT on a subset of deduplicated sequences of AcrIE9/10a, belonging to protein clusters with a number of representatives more than 2 (resulted in alignment of 89 unique sequences).

### Workflow

We processed data and visualized results using Python3 v.3.11.2 (https://docs.python.org/3.11/) and R v.4.2.3 (https://www.R-project.org/). To work with networks, we used the NetworkX v.2.7.1[125]. The multiple sequence alignments were visualised with pyMSAviz v.0.4.2 (https://github.com/moshi4/pyMSAviz). We aggregated the resulting figures and drew a model schema with the help of Inkscape v.1.3.2 (https://inkscape.org/credits/) or Biorender. The proposed workflow is formalized in a Snakemake-based pipeline (Snakemake v.7.25.2[126]). The developed pipeline and all supplementary scripts are available at GitHub: https://github.com/ombystoma-young/loci-search-pipe.

## Acknowledgements

Mass-spectrometry was performed at Skoltech Advanced Mass-Spectrometry Core Facility with support of the internal Skoltech grant, we thank Dr. Maria Zavialova for the samples handling. We thank Dr. Alan Davidson, who kindly shared *E. coli* Type IF CRISPR-Cas system plasmids. We express deep gratitude to Dr. David Bikard and members of his laboratory for the support of this research project.

## Author contributions

Conceptualization, A.I.; Molecular cloning, microbiology and phage methods, D.T., S.F., D.Y.; Bioinformatic analysis, O.K., and V.M.; *In vitro* assays, K.K., K.V., D.L., and N.B.; BGI sequencing, A.D.; Resources, A.I., B.W., M.W., E.S., K.S.; Writing—original draft, A.I., D.T., O.K.; Writing—review and editing, A.I., B.W., E.S., K.S; visualization, D.T., O.K. All authors have read and agreed to the published version of the manuscript.

## Funding

This research was funded by the Russian Scientific Foundation grant (24-74-10089) to A.I., and by the National Institutes of Health, grant GM104071 (K.S. and E.S.).

## Competing interest statement

Authors declare no competing interests.

## Data availability statement

Plasmids and bacterial strains used in this work are available upon request from the lead contact, Dr. Artem Isaev.

## References

1. Koonin, E. V, Makarova, K. S. & Zhang, F. Diversity, classification and evolution of CRISPR-Cas systems. *Curr Opin Microbiol* **37**, 67–78 (2017).
2. Koonin, E. V & Makarova, K. S. Origins and evolution of CRISPR-Cas systems. *Philosophical Transactions of the Royal Society B* **374**, 20180087 (2019).
3. Koonin, E. V & Makarova, K. S. Mobile genetic elements and evolution of CRISPR-Cas systems: all the way there and back. *Genome Biol Evol* **9**, 2812–2825 (2017).
4. Sternberg, S. H., Richter, H., Charpentier, E. & Qimron, U. Adaptation in CRISPR-Cas systems. *Mol Cell* **61**, 797–808 (2016).

5. Amitai, G. & Sorek, R. CRISPR–Cas adaptation: insights into the mechanism of action. *Nat Rev Microbiol* **14**, 67–76 (2016).
6. Haurwitz, R. E., Jinek, M., Wiedenheft, B., Zhou, K. & Doudna, J. A. Sequence-and structure-specific RNA processing by a CRISPR endonuclease. *Science (1979)* **329**, 1355–1358 (2010).
7. Li, H. Structural principles of CRISPR RNA processing. *Structure* **23**, 13–20 (2015).
8. Barrangou, R. CRISPR-Cas systems and RNA-guided interference. *Wiley Interdiscip Rev RNA* **4**, 267–278 (2013).
9. Marraffini, L. A. & Sontheimer, E. J. CRISPR interference: RNA-directed adaptive immunity in bacteria and archaea. *Nat Rev Genet* **11**, 181–190 (2010).
10. Makarova, K. S. *et al.* Evolutionary classification of CRISPR–Cas systems: a burst of class 2 and derived variants. *Nat Rev Microbiol* **18**, 67–83 (2020).
11. Altae-Tran, H. *et al.* Uncovering the functional diversity of rare CRISPR-Cas systems with deep terascale clustering. *Science (1979)* **382**, eadi1910 (2023).
12. Schelling, M. A., Nguyen, G. T. & Sashital, D. G. CRISPR-Cas effector specificity and cleavage site determine phage escape outcomes. *PLoS Biol* **21**, e3002065 (2023).
13. Westra, E. R. *et al.* Type IE CRISPR-cas systems discriminate target from non-target DNA through base pairing-independent PAM recognition. *PLoS Genet* **9**, e1003742 (2013).
14. Pawluk, A., Davidson, A. R. & Maxwell, K. L. Anti-CRISPR: discovery, mechanism and function. *Nat Rev Microbiol* **16**, 12–17 (2018).
15. Davidson, A. R. *et al.* Anti-CRISPRs: protein inhibitors of CRISPR-Cas systems. *Annu Rev Biochem* **89**, 309–332 (2020).
16. Bondy-Denomy, J., Pawluk, A., Maxwell, K. L. & Davidson, A. R. Bacteriophage genes that inactivate the CRISPR/Cas bacterial immune system. *Nature* **493**, 429–432 (2013).
17. Bondy-Denomy, J. *et al.* Multiple mechanisms for CRISPR–Cas inhibition by anti-CRISPR proteins. *Nature* **526**, 136–139 (2015).
18. Camara-Wilpert, S. *et al.* Bacteriophages suppress CRISPR–Cas immunity using RNA-based anti-CRISPRs. *Nature* **623**, 601–607 (2023).
19. Hayes, V. M. *et al.* RNA-mediated CRISPR-Cas13 inhibition through crRNA structural mimicry. *Science (1979)* **388**, 387–391 (2025).
20. Thavalingam, A. *et al.* Inhibition of CRISPR-Cas9 ribonucleoprotein complex assembly by anti-CRISPR AcrIIC2. *Nat Commun* **10**, 2806 (2019).
21. Guo, T. W. *et al.* Cryo-EM structures reveal mechanism and inhibition of DNA targeting by a CRISPR-Cas surveillance complex. *Cell* **171**, 414–426 (2017).
22. Gabel, C., Li, Z., Zhang, H. & Chang, L. Structural basis for inhibition of the type IF CRISPR–Cas surveillance complex by AcrIF4, AcrIF7 and AcrIF14. *Nucleic Acids Res* **49**, 584–594 (2021).
23. León, L. M., Park, A. E., Borges, A. L., Zhang, J. Y. & Bondy-Denomy, J. Mobile element warfare via CRISPR and anti-CRISPR in Pseudomonas aeruginosa. *Nucleic Acids Res* **49**, 2114–2125 (2021).
24. Lu, W.-T., Trost, C. N., Müller-Esparza, H., Randau, L. & Davidson, A. R. Anti-CRISPR AcrIF9 functions by inducing the CRISPR–Cas complex to bind DNA non-specifically. *Nucleic Acids Res* **49**, 3381–3393 (2021).
25. Song, G. *et al.* AcrIIA5 inhibits a broad range of Cas9 orthologs by preventing DNA target cleavage. *Cell Rep* **29**, 2579–2589 (2019).
26. Wiegand, T., Karambelkar, S., Bondy-Denomy, J. & Wiedenheft, B. Structures and strategies of anti-CRISPR-mediated immune suppression. *Annu Rev Microbiol* **74**, 21–37 (2020).
27. Niu, Y. *et al.* A type IF anti-CRISPR protein inhibits the CRISPR-Cas surveillance complex by ADP-ribosylation. *Mol Cell* **80**, 512–524 (2020).
28. Zhu, Y. *et al.* Diverse mechanisms of CRISPR-Cas9 inhibition by type IIC anti-CRISPR proteins. *Mol Cell* **74**, 296–309 (2019).
29. Trost, C. N. *et al.* An anti-CRISPR that pulls apart a CRISPR–Cas complex. *Nature* **632**, 375–382 (2024).
30. Liu, C. *et al.* Widespread RNA-based cas regulation monitors crRNA abundance and anti-CRISPR proteins. *Cell Host Microbe* **31**, 1481–1493 (2023).
31. Workman, R. E. *et al.* Anti-CRISPR proteins trigger a burst of CRISPR-Cas9 expression that enhances phage defense. *Cell Rep* **43**, (2024).
32. Pinilla-Redondo, R. *et al.* Discovery of multiple anti-CRISPRs highlights anti-defense gene clustering in mobile genetic elements. *Nat Commun* **11**, 5652 (2020).
33. Pons, B. J., van Houte, S., Westra, E. R. & Chevallereau, A. Ecology and evolution of phages encoding anti-CRISPR proteins. *J Mol Biol* **435**, 167974 (2023).
34. Dong, C. *et al.* Anti-CRISPRdb v2. 2: an online repository of anti-CRISPR proteins including information on inhibitory mechanisms, activities and neighbors of curated anti-CRISPR proteins. *Database* **2022**, baac010 (2022).
35. Yan, Y., Zheng, J., Zhang, X. & Yin, Y. dbAPIS: a database of a nti-p rokaryotic i mmune s ystem genes. *Nucleic Acids Res* **52**, D419–D425 (2024).
36. Stanley, S. Y. *et al.* Anti-CRISPR-associated proteins are crucial repressors of anti-CRISPR transcription. *Cell* **178**, 1452–1464 (2019).
37. Wang, C. *et al.* A Novel Anti-CRISPR AcrIE9. 2 Is Associated with Dissemination of bla KPC Plasmids in Klebsiella pneumoniae Sequence Type 15. *Antimicrob Agents Chemother* **67**, e01547-22 (2023).
38. Jiang, C. *et al.* A new anti-CRISPR gene promotes the spread of drug-resistance plasmids in Klebsiella pneumoniae. *Nucleic Acids Res* gkae516 (2024).
39. Botelho, J., Cazares, A. & Schulenburg, H. The ESKAPE mobilome contributes to the spread of antimicrobial resistance and CRISPR-mediated conflict between mobile genetic elements. *Nucleic Acids Res* **51**, 236–252 (2023).

40. Díez-Villaseñor, C., Almendros, C., García-Martínez, J. & Mojica, F. J. M. Diversity of CRISPR loci in Escherichia coli. *Microbiology (N Y)* **156**, 1351–1361 (2010).
41. Xue, C. & Sashital, D. G. Mechanisms of type IE and IF CRISPR-Cas systems in Enterobacteriaceae. *EcoSal Plus* **8**, 10–1128 (2019).
42. Savitskaya, E., Semenova, E., Dedkov, V., Metlitskaya, A. & Severinov, K. High-throughput analysis of type IE CRISPR/Cas spacer acquisition in E. coli. *RNA Biol* **10**, 716–725 (2013).
43. Redding, S. *et al.* Surveillance and processing of foreign DNA by the Escherichia coli CRISPR-Cas system. *Cell* **163**, 854–865 (2015).
44. Strotskaya, A. *et al.* The action of Escherichia coli CRISPR–Cas system on lytic bacteriophages with different lifestyles and development strategies. *Nucleic Acids Res* **45**, 1946–1957 (2017).
45. Pougach, K. *et al.* Transcription, processing and function of CRISPR cassettes in Escherichia coli. *Mol Microbiol* **77**, 1367–1379 (2010).
46. Jore, M. M. *et al.* Structural basis for CRISPR RNA-guided DNA recognition by Cascade. *Nat Struct Mol Biol* **18**, 529–536 (2011).
47. Pul, Ü. *et al.* Identification and characterization of E. coli CRISPR-cas promoters and their silencing by H-NS. *Mol Microbiol* **75**, 1495–1512 (2010).
48. Westra, E. R. *et al.* H-NS-mediated repression of CRISPR-based immunity in Escherichia coli K12 can be relieved by the transcription activator LeuO. *Mol Microbiol* **77**, 1380–1393 (2010).
49. Singh, K., Milstein, J. N. & Navarre, W. W. Xenogeneic silencing and its impact on bacterial genomes. *Annu Rev Microbiol* **70**, 199–213 (2016).
50. Cooper, C. *et al.* H-NS is a bacterial transposon capture protein. *Nat Commun* **15**, 7137 (2024).
51. Fang, M. *et al.* LrhA promotes CRISPR-Cas immunity by promoting reciprocal interplay between interference and primed adaptation in Escherichia coli. *bioRxiv* 2023–2028 (2023).
52. Lopez, S. C., Lee, Y., Zhang, K. & Shipman, S. SspA is a transcriptional regulator of CRISPR adaptation in E. coli. *bioRxiv* 2024–2025 (2024).
53. Yang, C., Chen, Y., Huang, H., Huang, H. & Tseng, C. CRP represses the CRISPR/Cas system in E scherichia coli: evidence that endogenous CRISPR spacers impede phage P 1 replication. *Mol Microbiol* **92**, 1072–1091 (2014).
54. Mitić, D., Radovčić, M., Markulin, D. & Ivančić-Baće, I. StpA represses CRISPR-Cas immunity in H-NS deficient Escherichia coli. *Biochimie* **174**, 136–143 (2020).
55. Patterson, A. G., Yevstigneyeva, M. S. & Fineran, P. C. Regulation of CRISPR–Cas adaptive immune systems. *Curr Opin Microbiol* **37**, 1–7 (2017).
56. Touchon, M. *et al.* CRISPR distribution within the Escherichia coli species is not suggestive of immunity-associated diversifying selection. *J Bacteriol* **193**, 2460–2467 (2011).
57. Savitskaya, E. *et al.* Dynamics of Escherichia coli type I-E CRISPR spacers over 42 000 years. *Mol Ecol* **26**, 2019–2026 (2017).
58. Almendros, C., Mojica, F. J. M., Díez-Villaseñor, C., Guzmán, N. M. & García-Martínez, J. CRISPR-Cas functional module exchange in Escherichia coli. *mBio* **5**, 10–1128 (2014).
59. Hochhauser, D., Millman, A. & Sorek, R. The defense island repertoire of the Escherichia coli pan-genome. *PLoS Genet* **19**, e1010694 (2023).
60. Dion, M. B. *et al.* Escherichia coli CRISPR arrays from early life fecal samples preferentially target prophages. *ISME J* **18**, wrae005 (2024).
61. Song, S. *et al.* CRISPR-Cas controls cryptic prophages. *Int J Mol Sci* **23**, 16195 (2022).
62. Gussow, A. B. *et al.* Machine-learning approach expands the repertoire of anti-CRISPR protein families. *Nat Commun* **11**, 1–12 (2020).
63. Eitzinger, S. *et al.* Machine learning predicts new anti-CRISPR proteins. *Nucleic Acids Res* **48**, 4698–4708 (2020).
64. Dao, F.-Y. *et al.* AcrPred: A hybrid optimization with enumerated machine learning algorithm to predict Anti-CRISPR proteins. *Int J Biol Macromol* **228**, 706–714 (2023).
65. Pawluk, A., Bondy-Denomy, J., Cheung, V. H. W., Maxwell, K. L. & Davidson, A. R. A new group of phage anti-CRISPR genes inhibits the type IE CRISPR-Cas system of Pseudomonas aeruginosa. *mBio* **5**, 10–1128 (2014).
66. Shmakov, S. *et al.* Pervasive generation of oppositely oriented spacers during CRISPR adaptation. *Nucleic Acids Res* **42**, 5907–5916 (2014).
67. Almendros, C., Guzman, N. M., Diez-Villasenor, C., García-Martínez, J. & Mojica, F. J. M. Target motifs affecting natural immunity by a constitutive CRISPR-Cas system in Escherichia coli. *PLoS One* **7**, e50797 (2012).
68. Lu, W.-T. Investigation of non-specific DNA binding by the type IF CRISPR-Cas complex when bound to anti-CRISPR AcrIF9. (2022).
69. Datsenko, K. A. *et al.* Molecular memory of prior infections activates the CRISPR/Cas adaptive bacterial immunity system. *Nat Commun* **3**, 1–7 (2012).
70. Semenova, E., … E. S.-P. of the & 2016, undefined. Highly efficient primed spacer acquisition from targets destroyed by the Escherichia coli type IE CRISPR-Cas interfering complex. *National Acad SciencesE Semenova, E Savitskaya, O Musharova, A Strotskaya, D Vorontsova, KA DatsenkoProceedings of the National Academy of Sciences, 2016•National Acad Sciences*.
71. Rutkauskas, M. *et al.* Directional R-loop formation by the CRISPR-Cas surveillance complex cascade provides efficient off-target site rejection. *Cell Rep* **10**, 1534–1543 (2015).
72. Lemak, S., Brown, G., Makarova, K. S., Koonin, E. V. & Yakunin, A. F. Biochemical plasticity of the Escherichia coli CRISPR Cascade revealed by in vitro reconstitution of Cascade activities from purified Cas proteins. *FEBS J* (2024) doi:10.1111/FEBS.17295.

73. Luo, M. L., Mullis, A. S., Leenay, R. T. & Beisel, C. L. Repurposing endogenous type I CRISPR-Cas systems for programmable gene repression. *Nucleic Acids Res* **43**, 674–681 (2015).
74. Rath, D., Amlinger, L., Hoekzema, M., Devulapally, P. R. & Lundgren, M. Efficient programmable gene silencing by Cascade. *Nucleic Acids Res* **43**, 237–246 (2015).
75. Musharova, O. *et al.* Systematic analysis of Type I-E Escherichia coli CRISPR-Cas PAM sequences ability to promote interference and primed adaptation. *Mol Microbiol* **111**, 1558–1570 (2019).
76. Forsberg, K. J. *et al.* The novel anti-CRISPR AcrIIA22 relieves DNA torsion in target plasmids and impairs SpyCas9 activity. *PLoS Biol* **19**, e3001428 (2021).
77. Dolan, A. E. *et al.* Introducing a spectrum of long-range genomic deletions in human embryonic stem cells using type I CRISPR-Cas. *Mol Cell* **74**, 936–950 (2019).
78. Hou, Z., Hu, C., Ke, A. & Zhang, Y. Introducing Large Genomic Deletions in Human Pluripotent Stem Cells Using CRISPR-Cas3. *Curr Protoc* **2**, e361 (2022).
79. Hu, C. *et al.* Allosteric control of type IA CRISPR-Cas3 complexes and establishment as effective nucleic acid detection and human genome editing tools. *Mol Cell* **82**, 2754–2768 (2022).
80. Hu, C. *et al.* Exploiting activation and inactivation mechanisms in type IC CRISPR-Cas3 for genome-editing applications. *Mol Cell* **84**, 463–475 (2024).
81. Asmamaw Mengstie, M. *et al.* Recent Advancements in Reducing the Off-Target Effect of CRISPR-Cas9 Genome Editing. *Biologics* 21–28 (2024).
82. Marino, N. D., Pinilla-Redondo, R., Csörgő, B. & Bondy-Denomy, J. Anti-CRISPR protein applications: natural brakes for CRISPR-Cas technologies. *Nat Methods* **17**, 471–479 (2020).
83. Samuel, B., Mittelman, K., Croitoru, S. Y., Ben Haim, M. & Burstein, D. Diverse anti-defence systems are encoded in the leading region of plasmids. *Nature* 1–7 (2024).
84. Zaman, T. U., Albladi, M., Siddique, M. I., Aljohani, S. M. & Balkhy, H. H. Insertion element mediated mgrB disruption and presence of ISKpn28 in colistin-resistant Klebsiella pneumoniae isolates from Saudi Arabia. *Infect Drug Resist* 1183–1187 (2018).
85. Liu, Y. *et al.* Structural basis for anti-CRISPR repression mediated by bacterial operon proteins Aca1 and Aca2. *Journal of Biological Chemistry* **297**, (2021).
86. Birkholz, N. *et al.* Phage anti-CRISPR control by an RNA- and DNA-binding helix–turn–helix protein. *Nature* **631**, 670–677 (2024).
87. Lee, S. Y., Birkholz, N., Fineran, P. C. & Park, H. H. Molecular basis of anti-CRISPR operon repression by Aca10. *Nucleic Acids Res* **50**, 8919–8928 (2022).
88. Knott, G. J. *et al.* Structural basis for AcrVA4 inhibition of specific CRISPR-Cas12a. *Elife* **8**, e49110 (2019).
89. Peng, R. *et al.* Structural insight into multistage inhibition of CRISPR-Cas12a by AcrVA4. *Proceedings of the National Academy of Sciences* **116**, 18928–18936 (2019).
90. Abramson, J. *et al.* Accurate structure prediction of biomolecular interactions with AlphaFold 3. *Nature* 1–3 (2024).
91. Medina-Aparicio, L., Davila, S., Rebollar-Flores, J. E., Calva, E. & Hernandez-Lucas, I. The CRISPR-Cas system in enterobacteriaceae. *Pathog Dis* **76**, fty002 (2018).
92. Xue, C. & Sashital, D. G. Mechanisms of type IE and IF CRISPR-Cas systems in Enterobacteriaceae. *EcoSal Plus* **8**, 10–1128 (2019).
93. Lu, W.-T., Trost, C. N., Müller-Esparza, H., Randau, L. & Davidson, A. R. Anti-CRISPR AcrIF9 functions by inducing the CRISPR–Cas complex to bind DNA non-specifically. *Nucleic Acids Res* **49**, 3381–3393 (2021).
94. Hirschi, M. *et al.* AcrIF9 tethers non-sequence specific dsDNA to the CRISPR RNA-guided surveillance complex. *Nat Commun* **11**, 2730 (2020).
95. Shehreen, S., Birkholz, N., Fineran, P. C. & Brown, C. M. Widespread repression of anti-CRISPR production by anti-CRISPR-associated proteins. *Nucleic Acids Res* **50**, 8615 (2022).
96. Sternberg, S. H., LaFrance, B., Kaplan, M. & Doudna, J. A. Conformational control of DNA target cleavage by CRISPR–Cas9. *Nature* **527**, 110–113 (2015).
97. Kuznedelov, K. *et al.* Altered stoichiometry Escherichia coli Cascade complexes with shortened CRISPR RNA spacers are capable of interference and primed adaptation. *Nucleic Acids Res* gkw914 (2016).
98. Jackson, R. N. *et al.* Crystal structure of the CRISPR RNA–guided surveillance complex from Escherichia coli. *Science (1979)* **345**, 1473–1479 (2014).
99. Beloglazova, N. *et al.* CRISPR RNA binding and DNA target recognition by purified Cascade complexes from Escherichia coli. *Nucleic Acids Res* **43**, 530–543 (2015).
100. Mekler, V., Kuznedelov, K. & Severinov, K. Quantification of the affinities of CRISPR–Cas9 nucleases for cognate protospacer adjacent motif (PAM) sequences. *Journal of Biological Chemistry* **295**, 6509–6517 (2020).
101. Laptev, I. *et al.* METTL15 interacts with the assembly intermediate of murine mitochondrial small ribosomal subunit to form m4C840 12S rRNA residue. *Nucleic Acids Res* **48**, 8022–8034 (2020).
102. Chen, S. Ultrafast one-pass FASTQ data preprocessing, quality control, and deduplication using fastp. *Imeta* **2**, e107 (2023).
103. Shiriaeva, A. A. *et al.* Detection of spacer precursors formed in vivo during primed CRISPR adaptation. *Nat Commun* **10**, 4603 (2019).
104. Li, H. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. *arXiv preprint arXiv:1303.3997* (2013).
105. Danecek, P. *et al.* Twelve years of SAMtools and BCFtools. *Gigascience* **10**, giab008 (2021).
106. Ramírez, F. *et al.* deepTools2: a next generation web server for deep-sequencing data analysis. *Nucleic Acids Res* **44**, W160 (2016).
107. Crusoe, M. R. *et al.* The khmer software package: enabling efficient nucleotide sequence analysis. *F1000Res* **4**, (2015).

108. O'Leary, N. A. *et al.* Exploring and retrieving sequence and metadata for species across the tree of life with NCBI Datasets. *Sci Data* **11**, 732 (2024).
109. Shen, W. & Ren, H. TaxonKit: A practical and efficient NCBI taxonomy toolkit. *Journal of genetics and genomics* **48**, 844–850 (2021).
110. Camargo, A. P. *et al.* Identification of mobile genetic elements with geNomad. *Nat Biotechnol* **42**, 1303–1312 (2024).
111. Steinegger, M. & Söding, J. Clustering huge protein sequence sets in linear time. *Nat Commun* **9**, 2542 (2018).
112. Hyatt, D. *et al.* Prodigal: prokaryotic gene recognition and translation initiation site identification. *BMC Bioinformatics* **11**, 1–11 (2010).
113. Steinegger, M. *et al.* HH-suite3 for fast remote homology detection and deep protein annotation. *BMC Bioinformatics* **20**, 1–15 (2019).
114. Nepusz, T., Yu, H. & Paccanaro, A. Detecting overlapping protein complexes in protein-protein interaction networks. *Nat Methods* **9**, 471–472 (2012).
115. Shen, W., Le, S., Li, Y. & Hu, F. SeqKit: a cross-platform and ultrafast toolkit for FASTA/Q file manipulation. *PLoS One* **11**, e0163962 (2016).
116. Katoh, K. & Standley, D. M. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol Biol Evol* **30**, 772–780 (2013).
117. Mistry, J. *et al.* Pfam: The protein families database in 2021. *Nucleic Acids Res* **49**, D412–D419 (2021).
118. Quinlan, A. R. & Hall, I. M. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* **26**, 841–842 (2010).
119. Salamov, V. S. A. & Solovyevand, A. Automatic annotation of microbial genomes and metagenomic sequences. *Metagenomics and its applications in agriculture, biomedicine and environmental studies* 61–78 (2011).
120. Söding, J., Biegert, A. & Lupas, A. N. The HHpred interactive server for protein homology detection and structure prediction. *Nucleic Acids Res* **33**, W244–W248 (2005).
121. Holm, L., Laiho, A., Törönen, P. & Salgado, M. DALI shines a light on remote homologs: One hundred discoveries. *Protein Science* **32**, e4519 (2023).
122. Li, Z., Jaroszewski, L., Iyer, M., Sedova, M. & Godzik, A. FATCAT 2.0: towards a better understanding of the structural diversity of proteins. *Nucleic Acids Res* **48**, W60–W64 (2020).
123. Gilchrist, C. L. M., Mirdita, M. & Steinegger, M. Multiple Protein Structure Alignment at Scale with FoldMason. *bioRxiv* 2024–2028 (2024).
124. Jurrus, E. *et al.* Improvements to the APBS biomolecular solvation software suite. *Protein Science* **27**, 112–128 (2018).
125. Hagberg, A., Swart, P. J. & Schult, D. A. *Exploring Network Structure, Dynamics, and Function Using NetworkX*. (2008).
126. Mölder, F. *et al.* Sustainable data analysis with Snakemake. *F1000Res* **10**, (2021).

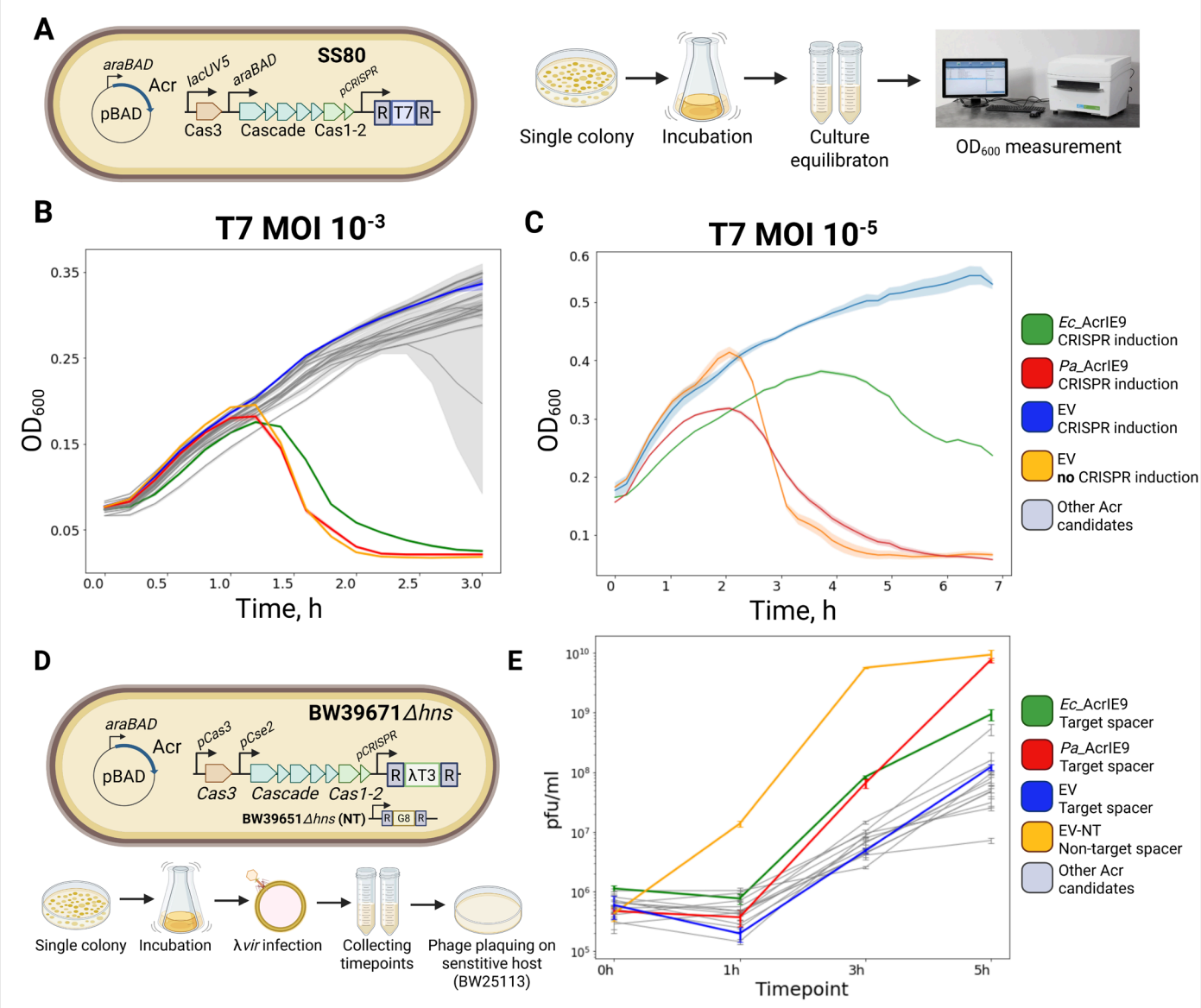


**Figure S1. Experimental validation of *Pa*_AcrIE9 and *Ec*_AcrIE9 *Ec*-IE-inhibiting activity. (A)** *E. coli* SS80 strain, a KD263 derivative with T7-targeting spacer, was used for T7 infection assay. **(B-C)** Growth curves of SS80 culture infected with T7 in conditions of Acr expression. Cells were grown with induction until $OD_{600}$ ~0.6, equilibrated to $OD_{600}$ = 0,3 and infected with the T7 phage at MOI=$10^{-3}$ **(B)**, MOI=$10^{-5}$ **(C)**. **(D)** *E. coli* BW39671 *Δhns* strain, a BW25113 *Δhns* derivative with a *λ* phage-targeting spacer T3, was used for $\lambda_{vir}$ phage infection assay. **(E)** Dynamics of phage $\lambda_{vir}$ release from infected *E. coli* BW39671 *Δhns* cells in conditions of Acr expression. Cells were grown at 30°C with induction until $OD_{600}$=0.6 and then infected with $\lambda_{vir}$ at MOI=$10^{-2}$. Phage titer in the spent media was measured at indicated time points on a lawn of sensitive *E.coli* BW25113 host. Acr expression was induced with 0.2% L-arabinose, CRISPR-Cas system was expressed from native promoters. All experiments were performed in three biological replicates.

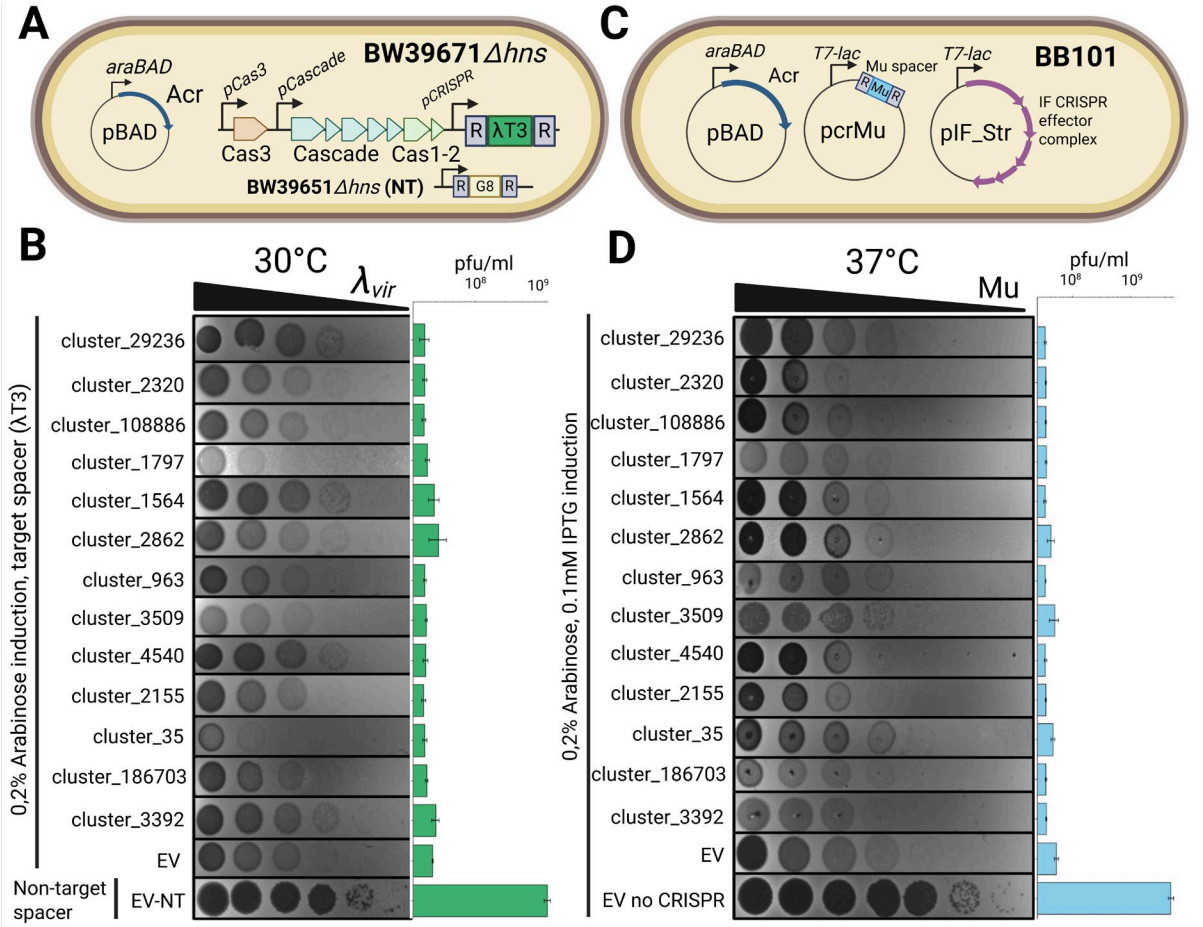


**Figure S2. Screening of ML-predicted collection of anti-CRISPR candidates. (A)** *E. coli* BW39671 *Δhns* strain, a BW25113 *Δhns* derivative with a *λ* phage-targeting spacer T3, was used for $\lambda_{vir}$ phage infection assay. **(B)** EOP with phage $\lambda_{vir}$ on the lawns of BW39671 *Δhns* cells with induced expression of Acr ML-predicted candidates. Cultures without a $\lambda_{vir}$ targeting spacer (EV-NT) and with empty pBAD vector (EV +ind) were used as controls. **(C)** *E. coli* BB101 strain with a plasmid-encoded IF LF82 system and phage-targeting spacer was used for the phage Mu phage infection assay. Acr expression was induced with 0.2% L-arabinose, CRISPR-Cas expression was induced with 0.2% L-arabinose and 1 mM IPTG. **(D)** EOP with phage Mu on the lawns of BB101 cells with induced expression of Cas proteins and Acr ML-predicted candidates. Cultures with empty pBAD vector (EV-NT) and without CRISPR-Cas encoding plasmid (no CRISPR) were used as controls. Acr expression was induced with 0.2% L-arabinose, CRISPR-Cas and spacer expression were induced with 0.1 mM IPTG. All experiments were performed in biological triplicates and bars on the right represent mean pfu/ml values.

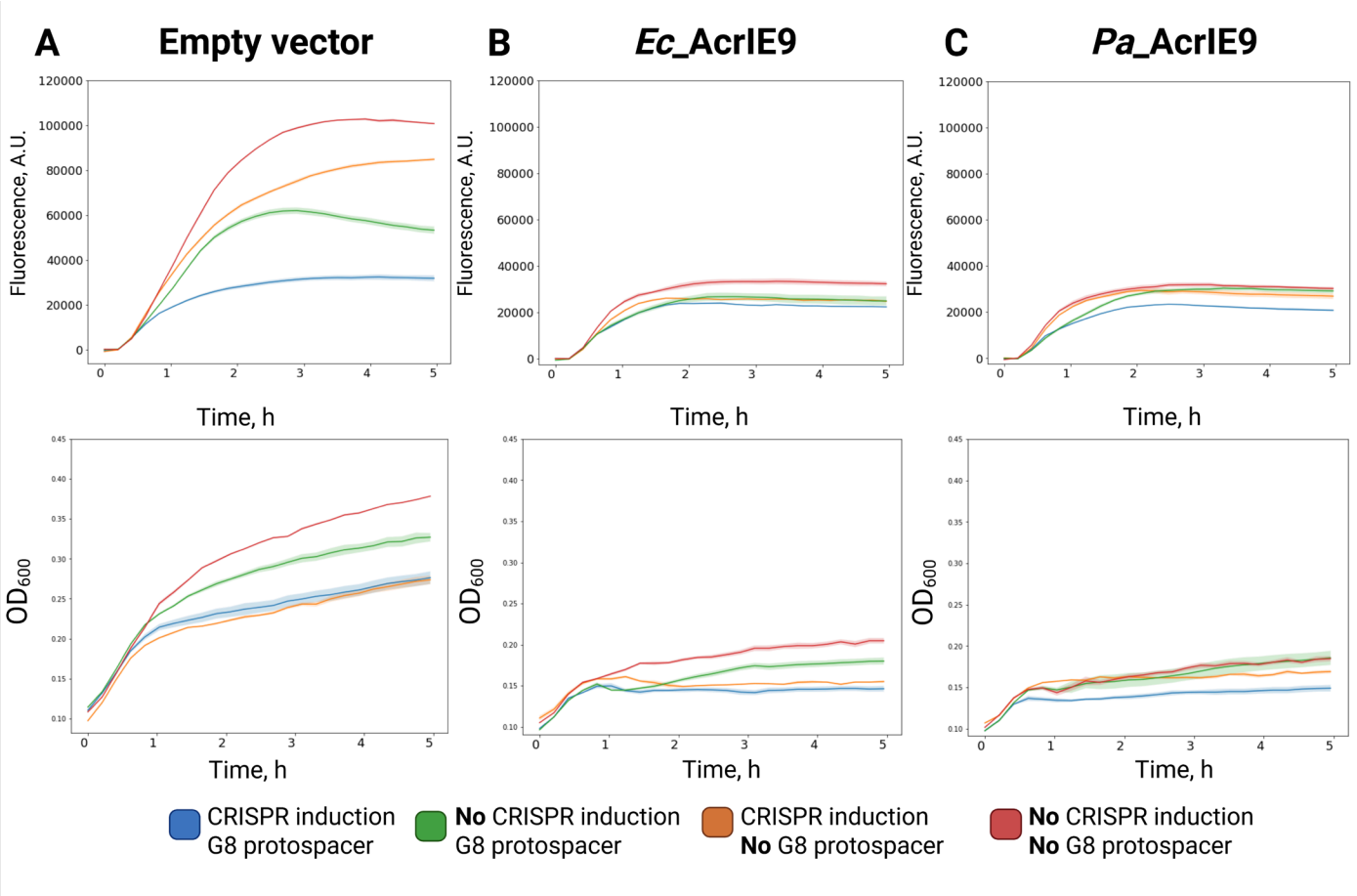


**Figure S3. Additional data for Figure 5B.** Dynamics of sfGFP production in KD454 cells carrying an empty pBAD vector **(A)** or expressing *Ec*_AcrIE9 **(B)** or *Pa*_AcrIE9 **(C)**. Top - Dynamics of sfGFP production measured via fluorescence accumulation. Bottom – optical density for the respective cultures. Acr expression inhibits *E.coli* growth, yet suppresses Cascade-mediated transcriptional silencing.

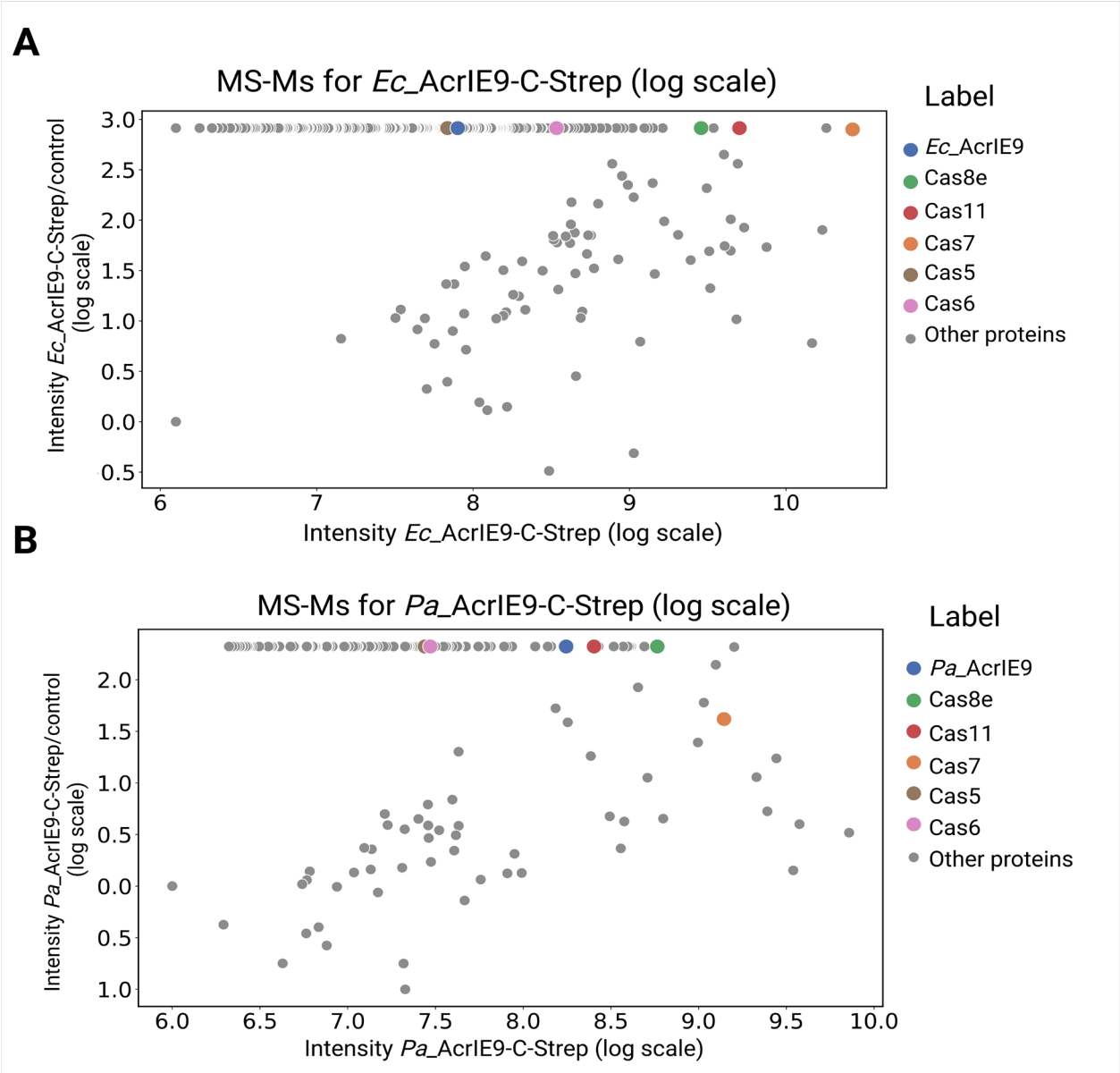


**Figure S4. AcrIE9 interacts with Cas7 *in vivo*.** Tandem mass-spectrometry analysis of the *in vivo* pull-downs performed with C-Strep *Ec*_AcrIE9 **(A)** or *Pa*_AcrIE9 **(B)** in *E.coli* KD263 after induction of CRISPR-Cas expression. Control pull-downs performed with AcrIE9 without Strep tag account for non-specific binding of proteins from the lysate to the column and were used to normalize protein intensity.

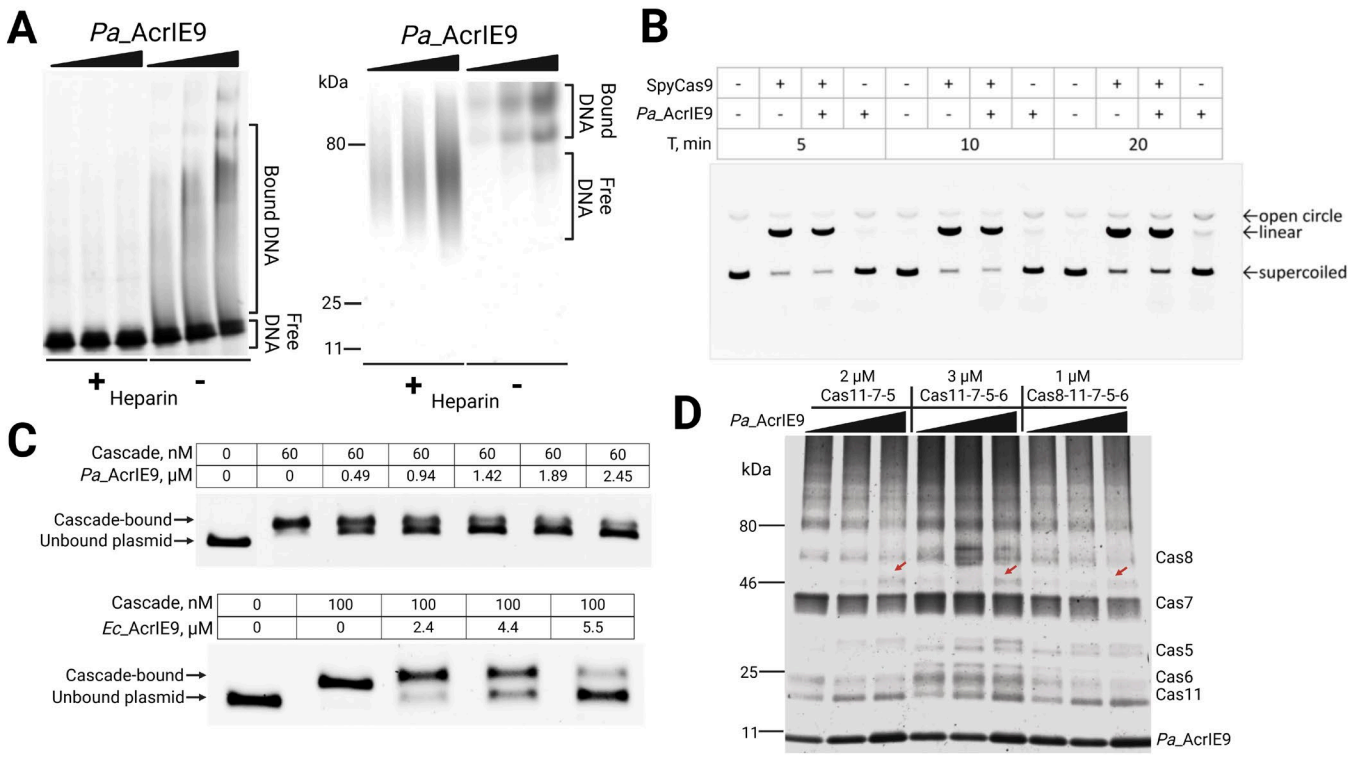


**Figure S5. AcrIE9 binds to dsDNA and inhibits Cascade interaction with DNA target. (A)** EMSA of *Pa*_AcrIE9 with dsDNA performed in conditions of competition with heparin. DNA (left) and protein (right) staining indicates formation of high-molecular-weight complexes and reduction of their size in the presence of heparin. Positions of control proteins with known molecular masses are labeled. 0.125 µM of 52-bp dsDNA was incubated with increasing concentrations of *Pa*_AcrIE9 (26, 46 and 92 µM). **(B)** Target plasmid DNA (1.2 nM) cleavage with 75 nM of SpyCas9 in the presence of 14 µM AcrIE9. **(C)** EMSA with a fixed concentration of *Ec*_IE Cascade and increasing concentration of AcrIE9 demonstrates disruption of *Ec*_IE Cascade:DNA complexes. Top panel: 60 nM of *Ec*_IE Cascade has been incubated with 4.5nM of target plasmid and *Pa*_AcrIE9. Bottom panel: 100 nM of *Ec*_IE Cascade has been incubated with 1 nM of target plasmid and *Ec*_AcrIE9. **(D)** *In vitro* glutaraldehyde (0.5%) crosslinking assay between *Pa*_AcrIE9 and subassemblies of Cascade complexes with different subunits composition. Each complex was incubated with increasing concentrations of *Pa*_AcrIE9 (5,14,19 µM), resulting in the appearance of novel band migrating above Cas7 (red arrows), that was confirmed to represent Cas7-AcrIE9 adduct.

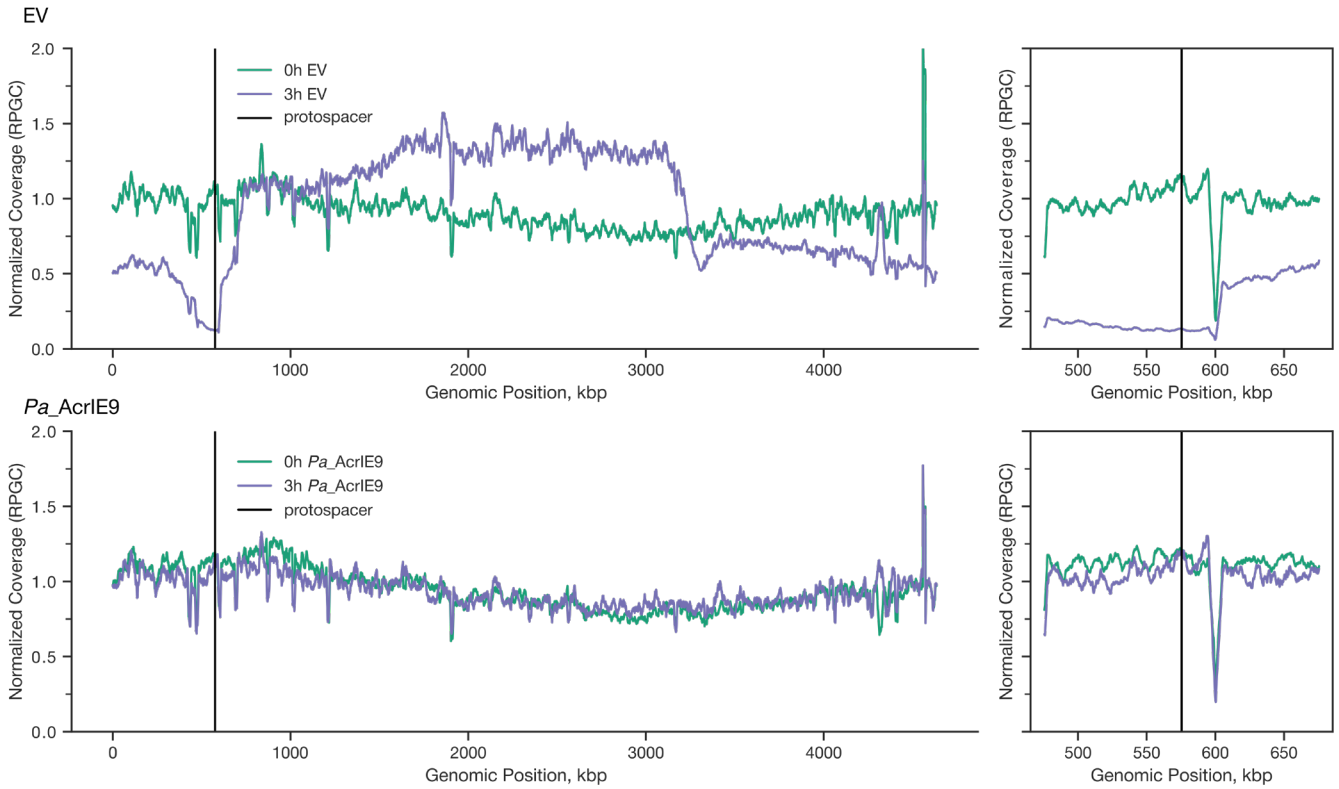


**Figure S6. Additional data for Figure 7C.** Genomic DNA coverage in self-targeting KD504 cells after CRISPR-Cas induction in conditions of *Pa*_AcrIE9 expression or in the presence of the empty pBAD vector. Bold curves represent moving average for coverage ratios before and 3h after induction of CRISPR-Cas expression. Window size is 10 kbp for the left panel and 5 kbp for the right panel zoomed in at the *yihN* protospacer region (black vertical line).

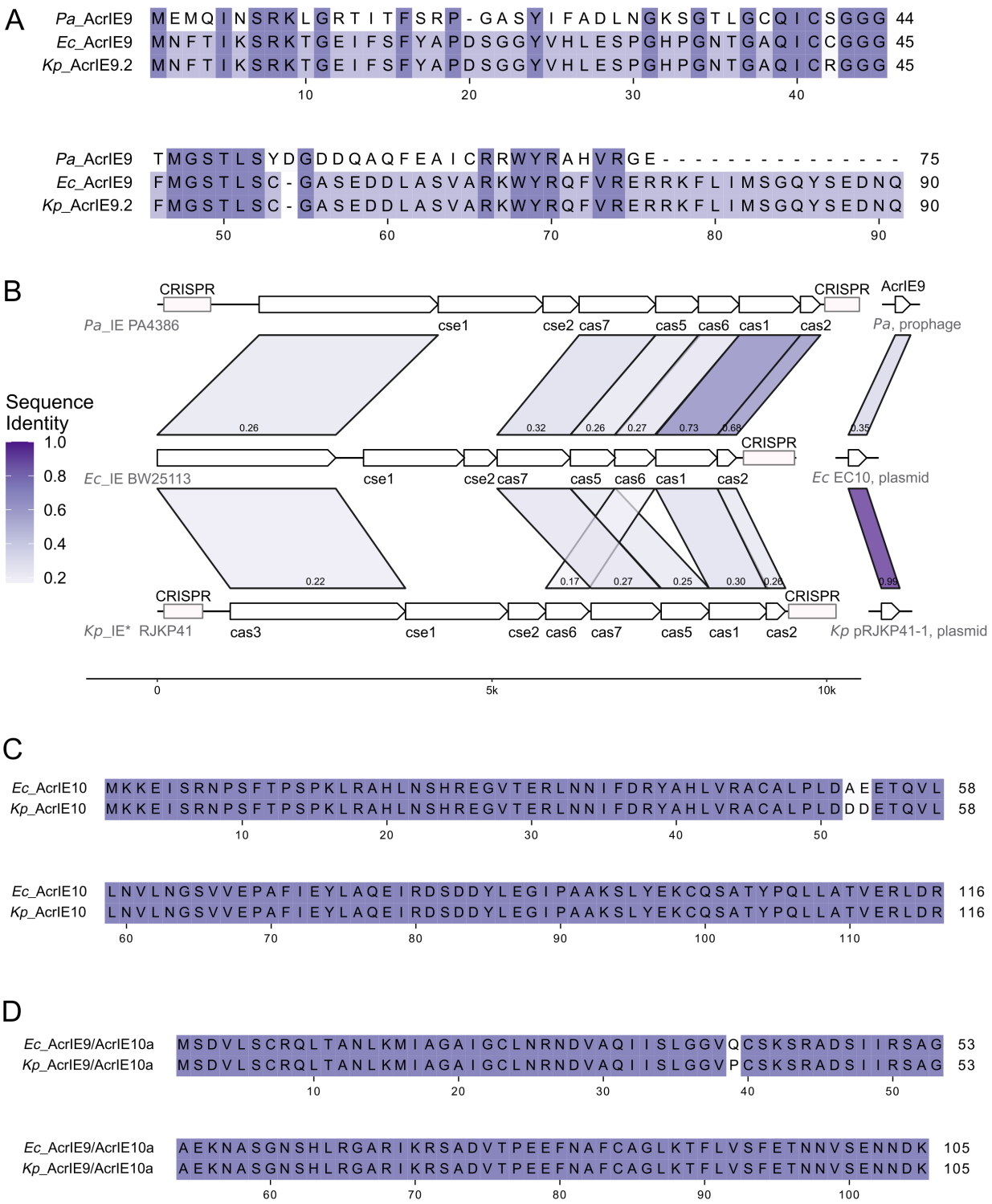


**Figure S7. Different Type I-E CRISPR-Cas systems are inhibited by AcrIE9 proteins. (A)** Multiple sequence alignment of experimentally studied AcrIE9 proteins. **(B)** Sequence identity between CRISPR-Cas systems and corresponding AcrIE9 proteins. Color represents the identity level (shown at the bottom on each colored segment). Sequence identity less than 0.2 is not shown. **(C)** Sequence alignment between AcrIE10 from *E. coli* and *K. pneumoniae*. **(D)** Sequence alignment between AcrIE9/10a from *E. coli* and *K. pneumoniae*. Pa – *Pseudomonas aeruginosa*, Ec – *Escherichia coli*, Kp – *Klebsiella pneumoniae*.

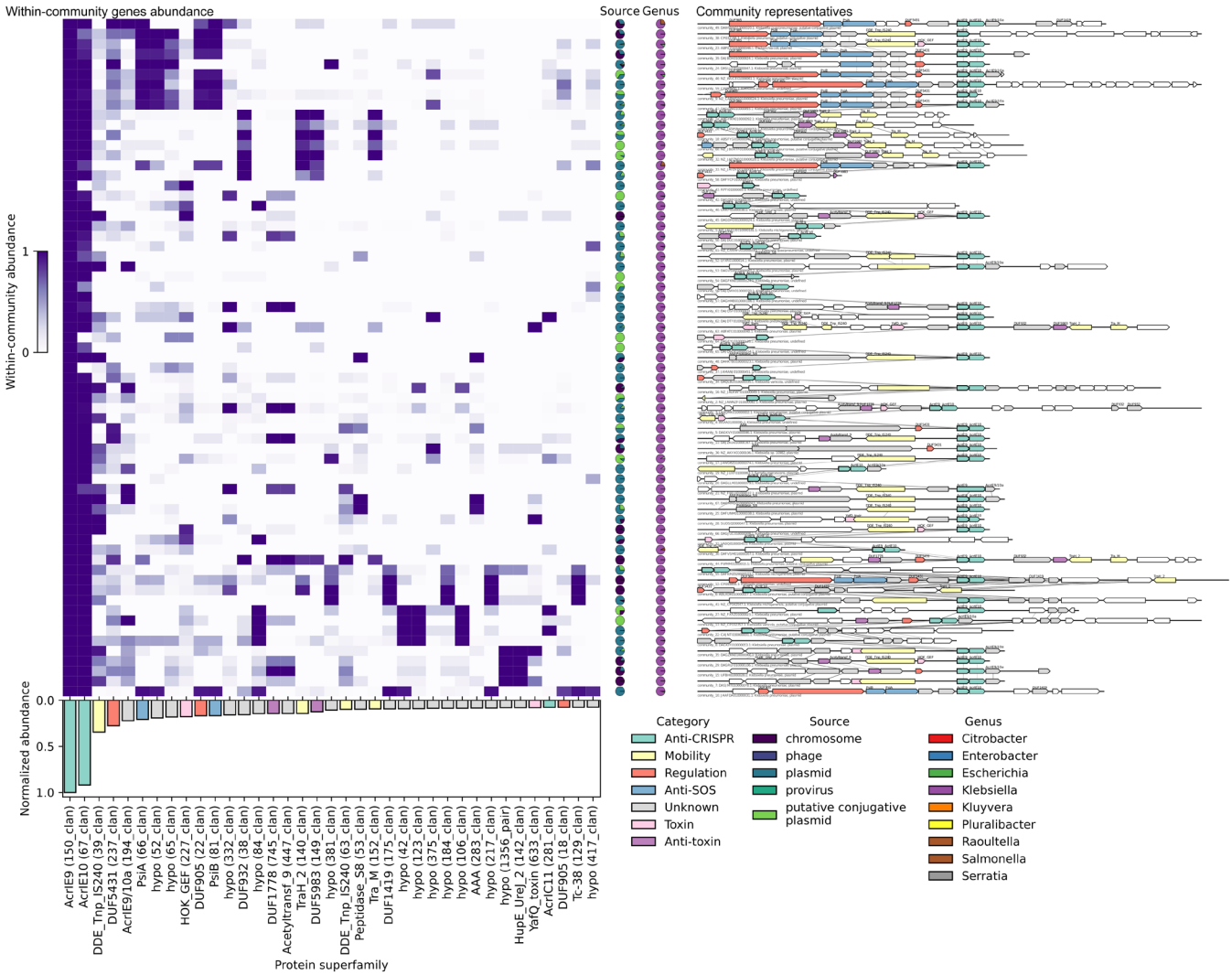


**Figure S8. The diversity of loci, encoding AcrIE9. (Left upper),** within-community abundance of the most common protein superfamilies, encoded in different communities of loci, encoding AcrIE9. The color of each cell indicates the fraction of loci in the community (rows), encoding the corresponding protein superfamily (columns). Pie-charts indicate the localization (source) and taxonomy (genus) content of a community. **(Left bottom)**, community-normalized abundance of proteins for the most common proteins, encoded among all loci (normalized abundance more than 0.075, See `Loci diversity analysis` in `Methods` for details). The bars are colored according to a functional category that has been assigned manually. **(Right)** Representative genomic loci of the corresponding communities. Shown is the locus with the highest degree of centrality within the given community. Non-compressed variant of the Figure S8 can be found as an **Extended Data File 2**.

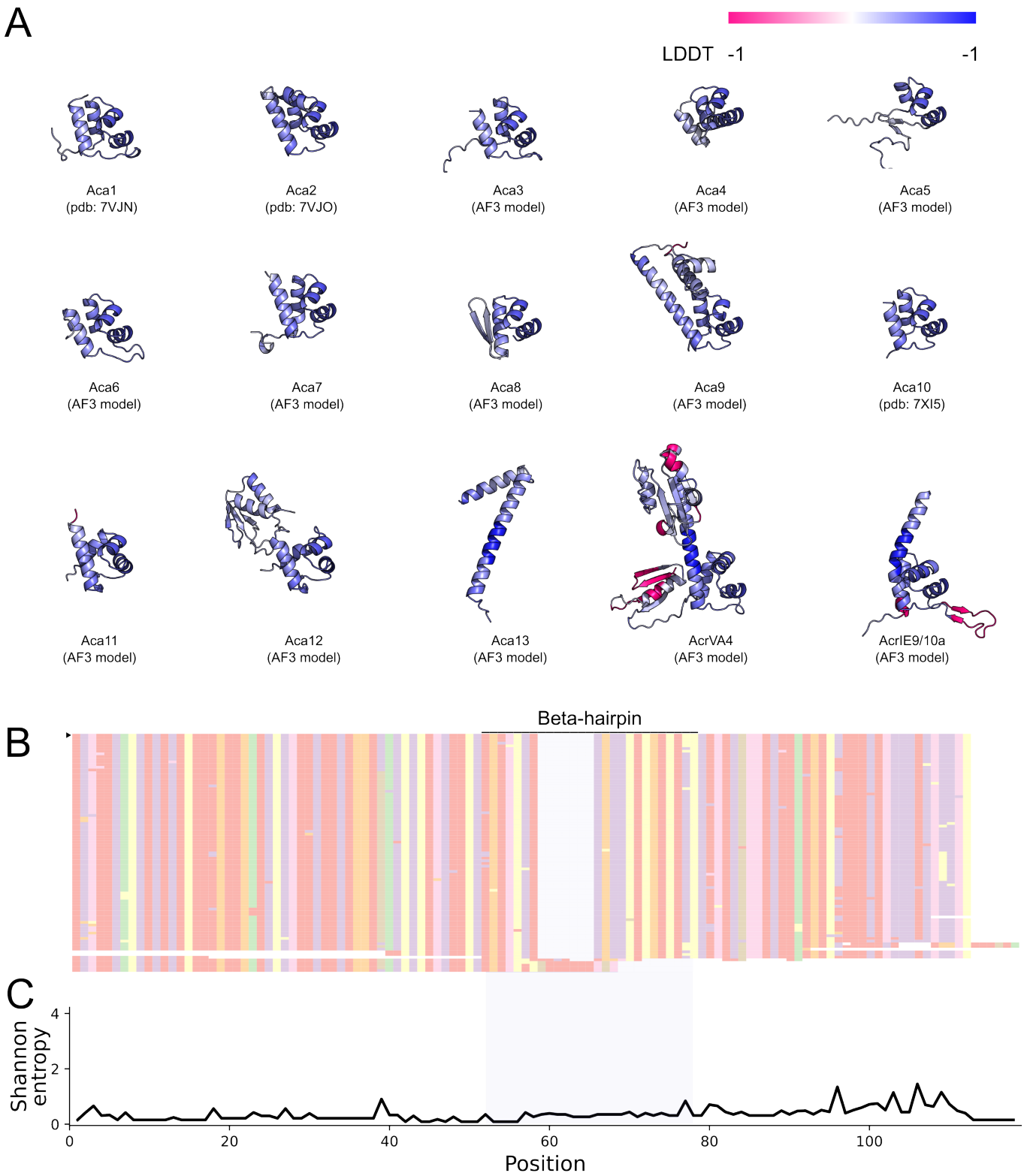


**Figure S9. AcrIE9/10a is predicted to be structurally similar to Aca proteins and is characterized by insertion of β-hairpin. (A)** Comparison of protein structures for previously discovered Aca proteins. Multiple sequence alignment of protein structures was performed with FoldMason WebServer[123]. Protein chains are colored according to the obtained structure-based score (LDDT) of a MSA, where -1 represents unaligned regions. All used protein sequences are listed in **Supplementary Table 3**. **(B)** Multiple sequence alignment of unique protein sequences belonging to AcrIE9/10a protein superfamily. AcrIE9/10a from AcrIE9 locus studied in this paper is marked with a triangle. The amino acids color scheme is adapted from Clustal color scheme. Residues are hidden for simplicity. **(C)** Shannon entropy calculated to 21-letters alphabet (gaps counted) for MSA on **(B)**.

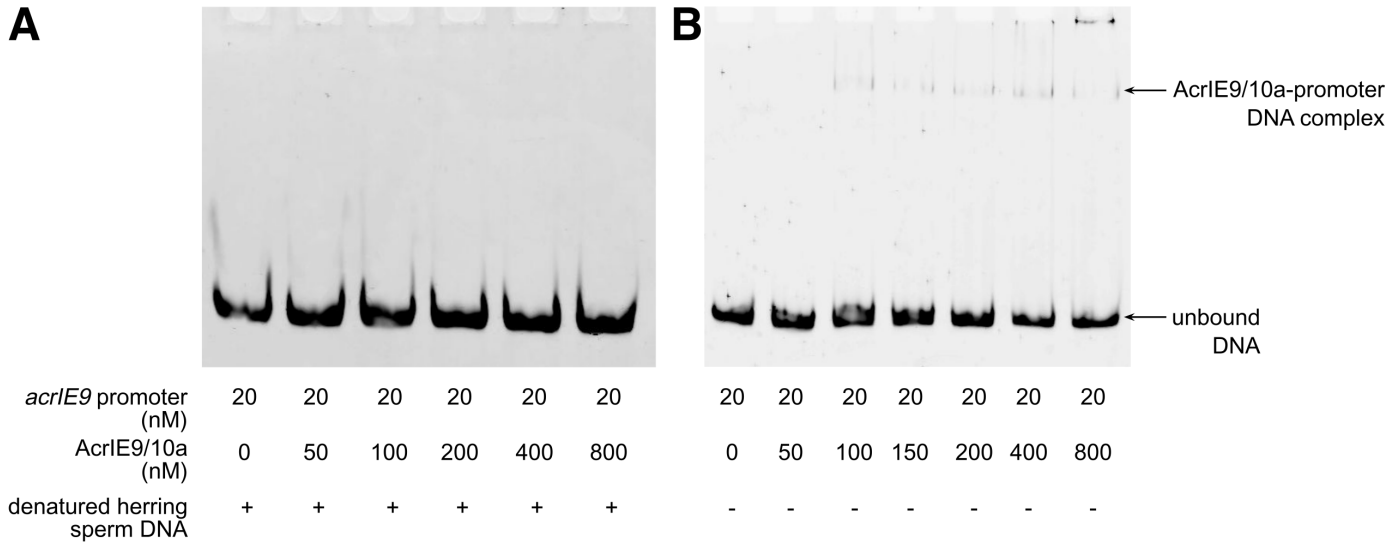


**Figure S10. AcrIE9/10a has non-specific DNA binding activity.** EMSA of *Kp*_AcrIE9/10a with 300 bp 5-FAM labeled dsDNA containing *acrIE9* promoter performed in conditions of competition with an excess of non-specific DNA **(A)** or without competition **(B)**.

**Supplementary Table 1. Dataset of Acr homologs studied in this work**

| Name | Protein id | Sequence | Verified Acr homolog | Percent Identity | Accession | Reference |
|---|---|---|---|---|---|---|
| *Pa*_AcrIE3 | HBQ8682788.1 | MKITNDTTTYEVAELMGSEADELD-GRIMMGLLSRECVVDTDDLSEDQW-LALIDESQKVRREQFESDEA | *Pa*_AcrIE3 | 100% | YP_950454 | (1) |
| *Kp*_AcrIE8 | APV19442.1 | MTTITINTYDPEARFNMDKDEARS-FFEFVEKKATDAG-FNVQYDSCNYVDEESERFVEKCFEDY | *Kp*_AcrIE8.1 | 100% | WP_009483893.1 | (2) |
| *Ec*_AcrIE9 | HBA3502305.1 | MNFTIKSRKTGEIFSFYAPDSG-GYVHLESPGHPGNTGAQIC-CGGGFMGSTLSCGASEDD-LASVARKWYRQFVRERRKFLIMSGQYSEDNQ | *Pa*_AcrIE9 | 41,89% | WP_101192668.1 | (3) |
| *Ec*_AcrIF11 | EEY5704741.1 | MKLFHGTYEKTAPIIKLGEFAMS-GNNVFDGLFACLDRDIAASHGNSIFI-YHVDDD-KIAESRDLNARFQEVYAFLRNELDTADVEEIADRVIWDNDSDIENFADIL-SPRLDSDISGAYSWEL-QRLRGRVAAYL-GFDAIEMNDEHGTSYLIVNPQIKGE | *Pa*_AcrIF11 | 32,90% | 6KYF_A | (4) |
| *Ec*_AcrIF15 | EFD0645621.1 | MTTIVIAHEVQDDKIHMI-EELLHDVALRDVNWSGAKFEIERG-DFTCIPDDESADAVKLLNKIQEIISGY | *Pa*_AcrIF15 | 55,70% | WP_117085604.1 | (2) |
| *Ec*_AcrIF16.1 | EFH8163204.1 | MPKNKTLDGPWVRHKYLLKGEMTSL-TIAAWTLGFANCRSLYRYLCSHKIKP-GADISHIRN-PIAGRTNVKKFIVNGELVSMNRAAALLGYSAASGLHHRIRRGSIEPGSDISH-LVKMSKVKK | *Pa*_AcrIF16 | 44% | WP_121296237.1 | (2) |
| *Ec*_AcrIF16.2 | EIS4806069.1 | MAIDTSSDALFEKIVKYG-MKPNYVAPLLALSQRDFFTLTAGI-ERAYQRQITRKELAEVAG-FSSAAIYSYTASPSARDYRTISEDMRIAAIWRVSCALPRL-PETKRGPREKYVIDGKPTTLSAAAKLLGYSKGSNLLRAIKGEGIPRGGDISH-VKIKPTKY-YIVDGQQLELAPAAFALGYRSSTAIRK-KNQS | *Pa*_AcrIF16 | 40,70% | WP_121296237.1 | (2) |
| *Pa*_AcrIF17 | ASE84371.1 | MTELDLRKLSKKALLQLISE-AANELHRRESSSSFVEPAAAEETAL-SPGQADLVFINNCLNASDYVHADAKDRYKQLASKYSAW-FAYKGYPRDLRGSDLKKWKQYFRP-NTREQ | *Pa*_AcrIF17 | 100% | WP_039476511.1 | (2) |
| *Pa*_AcrIF18 | HBQ7840699.1 | MTAIKATYISKDQNLTDGTTVY-WFDVNGETYGVVHGGESYNAK-VVDCDGAPSDQYTIDQFNISEEMIAE | *Pa*_AcrIF18.1 | 85,50% | WP_049300010.1 | (5) |

| | | | | | | |
|---|---|---|---|---|---|---|
| *Pa*_AcrIF2.1 | HBQ8682756.1 | MIAQQHKDTVAACEAAEAIAI-AKDQVWDGEGYTKYTFDDNSVLIQS GTTQYAMDAD-DADSIKGYADWLDDEART-AEDSEIKRLLEAVEDEA | *Pa*_AcrIF2 | 94.44% | WP_015994 820.1 | (2) |
| *Kp*_AcrIF22 | HBR0862247.1 | MSKMKFIENVGSSSTTAHLGTKKLGH-DVNAAAYECDGKIVIKLE-SNGVPVKGVSHI-EMTREQYDEFCLPQARKLFVRGIELFG ASTIFGE | *Pa*_AcrIF22 | 55,70% | WP_109463 511.1 | (6) |
| *Cb*_AcrIF23 | NCL83697.1 | MDENYIIARSIKEANKFIQTWEEAD-IEKLTDDQTRAAIG-FASKINSELREWIRMHLDGEG-TAHEEGYLKEQQAPWKKANTGDLFT DFGWWHRIANLMLHTAYINHAM-LGGDQYHSRLMKIFRDRFSYPEE | *Pa*_AcrIF3 | 44,90% | WP_181853 181.1 | (2) |
| *Se*_AcrIF23 | HAV1239868.1 | MNDNFRYWYNEI-ADMPEYDETGKYWYDLETGLRYDP-SCRYLPQNSRKPVSDKA-LAARSVAKKFGGRALKGTSKQVRWAE EI-RASVITQIPENAARICCNPDGLMKSAK FWIENRHKPATAFADFAINHRRLLA-DYQAAHNAGDAPLVAEIAGKYN-ALTTLWGFTG | *Pa*_AcrIF23 | 47,50% | WP_052155 777.1 | (7) |

Pa – *Pseudomonas aeruginosa*, Kp – *Klebsiella pneumonia*, Ec – *Escherichia coli*, Cb - *Citrobacter braakii*, Se - *Salmonella enterica*.

## Supplementary Table 2. Dataset of ML-predicted Acr candidates studied in this work

| Name | Protein Sequence | Accession |
|---|---|---|
| cluster_29236 | MPTNWKFIRLKIEDELIQPVNNLAGIRDELKADI-ASEIVEWYLNQRKTGHVHIRY-FSSPKDSDYQGIWFHPDTIKRAKQQAADDQQPLNRLIYTALARYFEQSI | WP_044863886.1 |
| cluster_2320 | MMLFIETEIFTRDVQELLSDDEYSEFQFFLADNPDAGDVIPGTG-GLRKVRWARGGKGKRGGVRVIYYYQVEHHQIWLLLVYRKGEKDDLPPEEKKMLRELMERW | WP_000897305.1 |
| cluster_108886 | MLTMLIYVKFYVATLKPHFNHHDKYKINIYLCRKYQYT | AIF63969.1 |
| cluster_1797 | MDTIITWMGDRLLGWITNKSDLRQRAITGLTPAIYSTI-LYTEKLKRGEPNNPNEEEKLYRL-WYEASSQVVDFDRELAKRCLDKSEYWLHSELYSPEKVGELNISLVGMKATLEGIKHN | WP_000379698.1 |
| cluster_1564 | MNKNAKIIKIHERILIILKHLHDDELDEALLVAIELT-KNLVDGTCPEHKPGLFMVLLTA-KALLEELSSTINSLLTEIDELHDHIEGLEIQLSRVRRQKQSGERNGN | WP_032235563.1 |
| cluster_2862 | MKIEQIG-MAAFRQQAESGGVDEFVVQRFGGVYHLFAVNRRAGVSYFLQERRGDYKTWLSLDRAAAFLSGIGVSRFTVRFEDNKHAEKDDRGH | ENC42600.1 |
| cluster_963 | MRVNKGDRQYYLNKEGDNFHLVKRVKTFSKSATLGKTKATVKT-VADLVFNENEFDTIDFTCDGLRENDKEIVSMMIQEMSEGKNAR | WP_000702249.1 |
| cluster_3509 | MLRTNQKILKDAYFDRGRDLAKLCAALKEISQYFTDIEDDIANKGASITLAN-RAHFMGFNVFTAIRTLKEQKEEIEKLKYRHKAKRV | WP_000849549.1 |
| cluster_4540 | MTTHTKNHYQQQSNIGLANAGLFYVRSKTRVAAHNIPES-FRKLSLRTAVYGGDHLGAAFLCEQAQLSKR | AKH22938.1 |
| cluster_2155 | MTLKMESSRTSKPEGPFVLIT-FEGDEFLFDDRHNVVKINGRPKKMDVKRNYFESDLGEGKAKYWTLNIEEAHRFNSIDEATEQLCKLSRPHLIKIRKVEGVSE | WP_000172888.1 |
| cluster_35 | MAIKHFPVVRFTSRGREYEVDERLITTIDKHRSEKDAHHIYLTDGTYFCAT-NVARVNLIRQVQEPRR | CAU97960.1 |
| cluster_186703 | MVSRFQSAGQRSINVIFGLFCCAGITCGCRHILRRLLCVSGCGP-GALLRLRHHFLKAFDDIRFQVALIRSNQKVIRLAP | EDU63394.1 |
| cluster_3392 | MGPQIAGLADYVNVPRTIAT-VISSGKCSLTELDTVLGVQDLWWWLEIITVDNYNQMVIDRAQEAW | WP_000228829.1 |

## Supplementary Table 3. Aca proteins analyzed in this work

| Protein | Accession | Source | Reference |
|---|---|---|---|
| Aca1 | 7VJN | PDB | (8) |
| Aca2 | 7VJO | PDB | (8) |
| Aca3 | WP_049360086.1 | NCBI Protein | (9) |
| Aca4 | WP_079381596.1 | NCBI Protein | (9) |
| Aca5 | WP_074032234.1 | NCBI Protein | (9) |
| Aca6 | WP_035450933.1 | NCBI Protein | (9) |
| Aca7 | WP_064700810.1 | NCBI Protein | (9) |
| Aca8 | YP_009272953.1 | NCBI Protein | (9) |
| Aca9 | WP_131138857.1 | NCBI Protein | (9) |
| Aca10 | 7XI5 | PDB | (10) |
| Aca11 | WP_009730540.1 | NCBI Protein | (11) |
| Aca12 | WP_033585136.1 | NCBI Protein | (11) |
| Aca13 | WP_094140901.1 | NCBI Protein | (11) |
| AcrVA4 | WP_046699156.1 | NCBI Protein | (12) |
| AcrIE9/10a | WP_013023810.1 | NCBI Protein | |

**Supplementary Table 4. List of strains, phages and plasmids used in this work**

| Bacterial strains | | |
|---|---|---|
| **Name** | **Comments** | **Source** |
| KD263 | *E. coli K-12 F+ , lacUV5-cas3 araBp8-cse1,* M13-targeting g8 spacer *(5'CTGTCTTTCGCTGCTGAGGGTGAC-GATCCCGC-3')* | Lab Stock |
| KD454 | KD263 derivative, ΔCas3 | Lab Stock |
| KD504 | KD263 derivative Encodes self-targeting *yihN* spacer *(5'-TCAAACAACCGACCTTGTTGTTCGCTATTGCC-3')* | Lab Stock |
| SS80 | KD263 derivative, T7-targeting first spacer, CmR | Lab Stock |
| BW25113 | *E. coli K12 F – Δ(araD-araB)567 ΔlacZ4787(::rrnB3) λ– rph-1 Δ(rhaD-rhaB)568 hsdR514* | Lab Stock |
| BW39651 | BW25113 derivative, λ-targeting T3 spacer in CRISPRI *(5'-TGGGATGCCTACCGCAAGCAGCTTGGCCTGAA-3')* | Lab Stock |
| BW39671 | BW39651 derivative*, Δhns::kan* | Lab Stock |
| BB101 | *ara Δ(lac pro) halA argE(Am) rif thi-1 ΔslyDF' lacIq lacZ::Tn5 pro+ (λDE3)* | kindly provided by prof. Alan Davidson |
| XL1-blue | *E. coli K12 recA1 endA1 gyrA96 thi-1 hsdR17 supE44 relA1 lac [F'proAB lacI q ZΔM15 Tn10],* TetR | Evrogen |
| BL21-AI | F- ompT hsdSB (rB- mB-) gal dcm araB::T7RNAP-tetA | Invitrogen |
| NEB 10-beta | *Δ(ara-leu) 7697 araD139 fhuA ΔlacX74 galK16 galE15 e14-ϕ80dlacZΔM15 recA1 relA1 endA1 nupG rpsL* (StrR) *rph spoT1 Δ(mrr-hsdRMS-mcrBC)* | NEB |
| KD418 | BL21 Star (DE3) with CRISPR1 deleted | (13) |
| LF82 | Adherent-invasive *Escherichia coli* strain (clinical isolate) | kindly provided by Alan Davidson |
| Rosetta 2 (DE3) | F- *ompT hsdSB(rB- mB-) gal dcm (DE3) pRARE2, CmR* | Novagen |
| BL21 Star (DE3) pLysS | F- *ompT hsdSB (rB-, mB-) gal dcm rne131 (DE3) pLysS, CmR* | Invitrogen |
| **Phages** | | |
| M13 | | Lab Stock |
| $\lambda_{vir}$ | Obligatory lytic phage λ derivative | Lab Stock |
| T7 | | Lab Stock |
| Mu WT | | Kindly provided by Alan Davidson |

| Plasmids | | |
|---|---|---|
| pBAD/HisA | Commertial plasmid, contains MCS under araBAD promoter, AmpR | Lab Stock |
| pBAD256 | Like pBAD/HisA except different SD sequence, AmpR | Lab Stock |
| pBAD_X | *araBAD* promoter, X denotes Acr genes cloned in this vector, listed in the Supplementary Tables 1 and 2, AmpR | This work |
| pBAD_Pa_AcrIE9/Ec C-strep | pBAD encoding C-terminally Strep-tagged versions of *Ec*_AcrIE9 or Pa_AcrIE9, contains enhanced RBS, AmpR | This work |
| pET30-AcrIE9_Ec | AcrIE9 from *Escherichia coli* cloned in pET30 with N-terminal 6His-Tag from pBAD-AcrIE9_Ec, KnR | This work |
| pET30-AcrIE9_Pa | AcrIE9 from *Pseudomonas aeruginosa* cloned in pET30 with N-terminal 6His-Tag from pBAD-AcrIE9_Ec, KnR | This work |
| pET28a-AcrIE9/10a | AcrIE9/10a from *Klebsiella pneumoniae* cloned in pET28a with N-terminal 6His-SUMO Tag from, KnR | This work |
| pAcr_crMu | Derivative of pACYCduet, with MuSp (5'-GCACTAGAGCAGGTGTCAAAGGTGGTAAAGCA-3') under *lac-T7* promoter targeting transposase A of Mu phage, CmR | Kindly provided by Alan Davidson |
| pET21d-LF82CR | Derivative of pET21d, contains IF CRISPR-Cas system from *E.coli* strain LF82 under lac-T7 promoter, AmpR | Kindly provided by Alan Davidson |
| pIF_Str | pET21d-LF82CR derivative, StrR | This work |
| pFR66 | sfGFP cloned under aTc-inducible promoter, KnR | Kindly provided by David Bikard |
| pFRG8 | pFR66 derivative, contains G8 protospacer of the phage M13 in sfGFP promoter region, KnR | This work |
| pFRC1T | pFR66 derivative, contains G8 protospacer with C1T mutation in sfGFP promoter region, KnR | This work |
| pG8_dir | g8 protospacer cloned in pT7Blue, AmpR | (14) |
| pJ3 | A derivative of pG8_dir where g8 protospacer replaced with J3 protospacer*, AmpR | This study |
| pCDF-casBCD | *cas11-cas7-cas5* in pCDF-1, *cas11* with N-terminal StrepII-tag, StrR | (15) |
| pCDF-casBCDE(g8) | *cas11-cas7-cas5-cas6* in pCDF-1, *cas11* with N-terminal StrepIItag and g8 spacer CRISPR array, StrR | (16) |
| pCDF-casBCDE(J3) | A derivative of pCDF-casBCDE(g8), where g8 spacer is replaced with J3 spacer* in CRISPR array, StrR | This study |
| pET30-casA | *cas8* cloned in pET30 with N-terminal 6His-Tag, KnR | This study |
| pMJ806 | pET-based His6-MBP expression vector with SpCas9 (1-1368), KnR | (17) |
| pl1-on-target | l1-on-target protospacer** cloned in pT7Blue, AmpR | This study |

* GCCAGTGATAAGTGGAATGCCATGTGGGCTGTC (18)

**AGCAGAAATCTCTGCTGACGCATAAAGATGAGACGCTGGAGTACAAACGTGAGCT (19)

**Supplementary Table 5. List of primers used in this work**

| Primer | Sequence (5'->3') | Description |
|---|---|---|
| pBAD_for | ATGCCATAGCATTTTTATCC | pBAD sequencing primer |
| pBAD_rev | GATTTAATCTG-TATCAGGCTG | pBAD sequencing primer |
| G8_sfGFP_Prom_F | GCAGCGAAAGACAG-CATTTTGATATGCCTCCG-GATATC | Cloning of G8 protospacer into sfGFP promoter region of pFR66 plasmid (KLD) |
| G8_sfGFP_Prom_R | TGAGGGTGAC-GATCCCGCTGACCTAG-TTAGGAGGC | Cloning of G8 protospacer into sfGFP promoter region of pFR66 plasmid (KLD) |
| pFR66-pTet-R | CCAGGCATCAAATAAAAC-GAAAGG | pFR66 linearization |
| G8_check_F | CACCCTCAGCAGCGAAA-GAC | G8 insert verification in pFR66 |
| pFR-g8-check-R | GCAGCTCTAATGCGCTGTT AAT | G8 insert verification in pFR66 |
| pFR_g8_C1T_R | GACAACATTTTGATAGCC | Cloning of C1T mutation in G8 protospacer (KLD), pFR66 plasmid |
| Moji3 | GTGGTTTGAGCGATGATAT | Naïve adaptation detection via PCR, anneals to the leader sequence of BL21AI *E.coli* CRISPR array |
| Moji4 | AGTTGGTAGATTGTGACTG | Naïve adaptation detection via PCR, anneals to the repeat sequence of BL21AI *E.coli* CRISPR array |
| EcLDR_F | AAGGTTGGTGGGTT-GTTTTTATGG | Primed adaptation detection via PCR, anneals to the leader sequence of KD263 *E.coli* CRISPR array |
| M13_G8_rev | GGATCGTCACCCTCAG-CAGCG | Primed adaptation detection via PCR, anneals to the G8 spacer sequence of KD263 *E.coli* CRISPR array |
| GA_Gene_LF82IFCR_F | CAATAA-TATTGAAAAAGGAAGAG-TATGAGGGAA-GCGGTGATCGCCGAAG | Cloning of StrR gene into pET21d-LF82CR (Gibson assembly) |
| GA_Gene_LF82IFCR_R | GAGTAAACTT-GGTCTGACAGTTATTT-GCCGACTACCTT-GGTGATCTCGC | Cloning of StrR gene into pET21d-LF82CR (Gibson assembly) |

| | | |
|---|---|---|
| GA_Back-bone_LF82IFCR_F | CTGTCAGAC-CAAGTTTACTCATATATAC | Amplification of the pET21d-LF82CR backbone (Gibson assembly) |
| GA_Back-bone_LF82IFCR_R | ACTCTTCCTTTTTCAA-TATTATTGAAG | Amplification of the pET21d-LF82CR backbone (Gibson assembly) |
| Ec_AcrIE9_SDpBAD256_F | CTCTAAGGAGGTTATAAAAAATGAACTTCACCATTAA-GAGTCG | Cloning of enhanced RBS sequence to the pBAD_Ec_AcrIE9 |
| Ec_AcrIE9_SDpBAD256_R | CTCGAATTCGCTAGCCCAAAAAACGGGTATG | Cloning of enhanced RBS sequence to the pBAD_Ec_AcrIE9 |
| Pa_AcrIE9_SDpBAD256_F | CTCTAAGGAGGTTATAAAAAATGCAGATTAACAGTCG-CAAAC | Cloning of enhanced RBS sequence to the pBAD_Pa_AcrIE9 |
| Pa_AcrIE9_SDpBAD256_R | CTCGAATTCGCTAGCCCAAAAAACGGGTATG | Cloning of enhanced RBS sequence to the pBAD_Pa_AcrIE9 |
| GA_Back-bone_AcrIE9strepC_Ec_F | TGA-GATCCGGCTGCTAACAAAG | Linearization of pBAD plasmid (Gibson assembly) |
| GA_Back-bone_AcrIE9strepC_Ec_R | GGTATATCTCCTTCTTAAAGTTAAACAAAATT | Linearization of pBAD plasmid (Gibson assembly) |
| GA_Gene_AcrIE9strepC_Ec_F | GTTTAACTTTAAGAAGGA-GATATACCATGAACTTCAC-CATTAAGAGTCGCA | Cloning of Strep-tag II at the C-terminus of pBAD_Ec_AcrIE9 (Gibson assembly) |
| GA_Gene_AcrIE9strepC_Ec_R | CGGGCTTTGTTAGCAGCCG-GATCTCATTATTTTTCGAACTGCGGGTGGC | Cloning of Strep-tag II at the C-terminus of pBAD_Ec_AcrIE9 (Gibson assembly) |
| GA_Gene_AcrIE9strepC_Ps_F | GTTTAACTTTAAGAAGGA-GATATACCATGCAGAT-TAACAGTCGCAAACTGG | Cloning of Strep-tag II at the C-terminus of pBAD_Pa_AcrIE9 (Gibson assembly) |
| GA_Gene_AcrIE9strepC_Ps_R | CGGGCTTTGTTAGCAGCCG-GATCTCATTATTTTTCGAACTGCGGGTGGCTC | Cloning of Strep-tag II at the C-terminus of pBAD_Pa_AcrIE9 (Gibson assembly) |

| | | |
|---|---|---|
| SpCas9 crRNA | GACGCAUAAAGAUGAGAC-GCGUUUUAGAGCUAUGCU | Part of synthetic SpyCas9 sgRNA |
| SpCas9 tracrRNA | AAACAGCAUAGCAAGUU-AAAAU-AAGGCUAGUCCGUUAUCA ACUUGAAAAAGUGGCAC-CGAGUCGGUGCU | Part of synthetic SpyCas9 sgRNA |
| ssDNA | TCGTTCCGATGCCAGTGA-TAAGTG-GAATGCCATGTGGGCTGTC AAAAGCGGC/36-FAM/ | Single-stranded substrate for EMSA |
| dsDNA_top | TCGTTCCGATGCCAGTGA-TAAGTG-GAATGCCATGTGGGCTGTC AAAAGCGGC/36-FAM/ | Double-stranded substrate for EMSA |
| dsDNA_bot | AGCAAGGCTACGGTCAC-TATTCACCTTACGG-TACACCCGACAG-TTTTCGCCG | Double-stranded substrate for EMSA |
| AcrIE9_promoter_top | CGCTACATCAGGTT-GCATAG | 5'FAM labeled primer for amplification of *acrIE9* promoter for EMSA |
| AcrIE9_promoter_bot | GCTGGCTCTCCTTTGGCTCA | Primer for amplification of *acrIE9* promoter for EMSA |

## Supplementary references


1. Pawluk A, Bondy-Denomy J, Cheung VH, Maxwell KL, Davidson AR. A new group of phage anti-CRISPR genes inhibits the type I-E CRISPR-Cas system of Pseudomonas aeruginosa. mBio. 2014 Apr 15;5(2):e00896. doi: 10.1128/mBio.00896-14. PMID: 24736222; PMCID: PMC3993853.
2. Pinilla-Redondo R, Shehreen S, Marino ND, Fagerlund RD, Brown CM, Sørensen SJ, Fineran PC, Bondy-Denomy J. Discovery of multiple anti-CRISPRs highlights anti-defense gene clustering in mobile genetic elements. Nat Commun. 2020 Nov 6;11(1):5652. doi: 10.1038/s41467-020-19415-3. PMID: 33159058; PMCID: PMC7648647.
3. Wang C, Sun Z, Hu Y, Li D, Guo Q, Wang M. A Novel Anti-CRISPR AcrIE9.2 Is Associated with Dissemination of blaKPC Plasmids in Klebsiella pneumoniae Sequence Type 15. Antimicrob Agents Chemother. 2023 Apr 18;67(4):e0154722. doi: 10.1128/aac.01547-22. Epub 2023 Mar 28. PMID: 36976005; PMCID: PMC10112057.

4. Niu Y, Yang L, Gao T, Dong C, Zhang B, Yin P, Hopp AK, Li D, Gan R, Wang H, Liu X, Cao X, Xie Y, Meng X, Deng H, Zhang X, Ren J, Hottiger MO, Chen Z, Zhang Y, Liu X, Feng Y. A Type I-F Anti-CRISPR Protein Inhibits the CRISPR-Cas Surveillance Complex by ADP-Ribosylation. Mol Cell. 2020 Nov 5;80(3):512-524.e5. doi: 10.1016/j.molcel.2020.09.015. Epub 2020 Oct 12. PMID: 33049228.
5. Peng R, Xu Y, Zhu T, Li N, Qi J, Chai Y, Wu M, Zhang X, Shi Y, Wang P, Wang J, Gao N, Gao GF. Alternate binding modes of anti-CRISPR viral suppressors AcrF1/2 to Csy surveillance complex revealed by cryo-EM structures. Cell Res. 2017 Jul;27(7):853-864. doi: 10.1038/cr.2017.79. Epub 2017 Jun 2. PMID: 28574055; PMCID: PMC5518991.
6. Bondy-Denomy J, Pawluk A, Maxwell KL, Davidson AR. Bacteriophage genes that inactivate the CRISPR/Cas bacterial immune system. Nature. 2013 Jan 17;493(7432):429-32. doi: 10.1038/nature11723. Epub 2012 Dec 16. PMID: 23242138; PMCID: PMC4931913.
7. Ren J, Wang H, Yang L, Li F, Wu Y, Luo Z, Chen Z, Zhang Y, Feng Y. Structural and mechanistic insights into the inhibition of type I-F CRISPR-Cas system by anti-CRISPR protein AcrIF23. J Biol Chem. 2022 Jul;298(7):102124. doi: 10.1016/j.jbc.2022.102124. Epub 2022 Jun 10. PMID: 35697070; PMCID: PMC9270243.
8. Liu Y, Zhang L, Guo M, Chen L, Wu B, Huang H. Structural basis for anti-CRISPR repression mediated by bacterial operon proteins Aca1 and Aca2. J Biol Chem. 2021 Dec;297(6):101357. doi: 10.1016/j.jbc.2021.101357. Epub 2021 Oct 28. PMID: 34756887; PMCID: PMC8633003.
9. Shehreen S, Birkholz N, Fineran PC, Brown CM. Widespread repression of anti-CRISPR production by anti-CRISPR-associated proteins. Nucleic Acids Res. 2022 Aug 26;50(15):8615-8625. doi: 10.1093/nar/gkac674. PMID: 35947749; PMCID: PMC9410906.
10. Lee SY, Birkholz N, Fineran PC, Park HH. Molecular basis of anti-CRISPR operon repression by Aca10. Nucleic Acids Res. 2022 Aug 26;50(15):8919-8928. doi: 10.1093/nar/gkac656. PMID: 35920325; PMCID: PMC9410881.
11. Song G, Zhang F, Tian C, Gao X, Zhu X, Fan D, Tian Y. Discovery of potent and versatile CRISPR-Cas9 inhibitors engineered for chemically controllable genome editing. Nucleic Acids Res. 2022 Mar 21;50(5):2836-2853. doi: 10.1093/nar/gkac099. PMID: 35188577; PMCID: PMC8934645.
12. Knott GJ, Cress BF, Liu JJ, Thornton BW, Lew RJ, Al-Shayeb B, Rosenberg DJ, Hammel M, Adler BA, Lobba MJ, Xu M, Arkin AP, Fellmann C, Doudna JA. Structural basis for AcrVA4 inhibition of specific CRISPR-Cas12a. Elife. 2019 Aug 9;8:e49110. doi: 10.7554/eLife.49110. PMID: 31397669; PMCID: PMC6711708.
13. Beloglazova N, Kuznedelov K, Flick R, Datsenko KA, Brown G, Popovic A, Lemak S, Semenova E, Severinov K, Yakunin AF. CRISPR RNA binding and DNA target recognition by purified Cascade complexes from Escherichia coli. Nucleic Acids Res. 2015 Jan;43(1):530-43. doi: 10.1093/nar/gku1285. Epub 2014 Dec 8. PMID: 25488810; PMCID: PMC4288178.
14. Datsenko, K.A., Pougach, K., Tikhonov, A., Wanner, B.L., Severinov, K., and Semenova, E. (2012). Molecular memory of prior infections activates the CRISPR/Cas adaptive bacterial immunity system. Nat Commun 3, 945.
15. Semenova E, Kuznedelov K, Datsenko KA, Boudry PM, Savitskaya EE, Medvedeva S, Beloglazova N, Logacheva M, Yakunin AF, Severinov K. The Cas6e ribonuclease is not required for interference and adaptation by the E. coli type I-E CRISPR-Cas system. Nucleic Acids Res. 2015 Jul 13;43(12):6049-61. doi: 10.1093/nar/gkv546. Epub 2015 May 26. PMID: 26013814; PMCID: PMC4499155.
16. Kuznedelov K, Mekler V, Lemak S, Tokmina-Lukaszewska M, Datsenko KA, Jain I, Savitskaya E,

Mallon J, Shmakov S, Bothner B, Bailey S, Yakunin AF, Severinov K, Semenova E. Altered stoichiometry Escherichia coli Cascade complexes with shortened CRISPR RNA spacers are capable of interference and primed adaptation. Nucleic Acids Res. 2016 Dec 15; 44(22):10849-10861. doi: 10.1093/nar/gkw914. Epub 2016 Oct 13. PMID: 27738137;

17. 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
18. Westra, E.R., Pul, U., Heidrich, N., Jore, M.M., Lundgren, M., Stratmann, T.,Wurm, R., Raine, A., Mescher, M., Van Heereveld, L., et al. (2010). H-NS-mediated repression of CRISPR-based immunity in Escherichia coli K12 can be relieved by the transcription activator LeuO. Mol. Microbiol. 77, 1380–1393.
19. Sternberg SH, LaFrance B, Kaplan M, and Doudna JA (2015) Conformational control of DNA target cleavage by CRISPR-Cas9. Nature 527, 110–113.