Crystal structure of an anti-CRISPR protein, AcrIIA1

Donghyun Ka[1],[†], So Young An[1],[†], Jeong-Yong Suh[1],[2],[3],[*] and Euiyoung Bae[1],[2],[*]

[1]Department of Agricultural Biotechnology, Seoul National University, Seoul 08826, Korea, [2]Research Institute of Agriculture and Life Sciences, Seoul National University, Seoul 08826, Korea and [3]Institute for Biomedical Sciences, Shinshu University, Minamiminowa, Nagano 399 4598, Japan

ABSTRACT

Clustered regularly interspaced short palindromic repeats (CRISPRs) and CRISPR-associated (Cas) proteins provide bacteria with RNA-based adaptive immunity against phage infection. To counteract this defense mechanism, phages evolved anti-CRISPR (Acr) proteins that inactivate the CRISPR-Cas systems. AcrIIA1, encoded by *Listeria monocytogenes* prophages, is the most prevalent among the Acr proteins targeting type II-A CRISPR-Cas systems and has been used as a marker to identify other Acr proteins. Here, we report the crystal structure of AcrIIA1 and its RNA-binding affinity. AcrIIA1 forms a dimer with a novel two helical-domain architecture. The N-terminal domain of AcrIIA1 exhibits a helix-turn-helix motif similar to transcriptional factors. When overexpressed in *Escherichia coli*, AcrIIA1 associates with RNAs, suggesting that AcrIIA1 functions via nucleic acid recognition. Taken together, the unique structural and functional features of AcrIIA1 suggest its distinct mode of Acr activity, expanding the diversity of the inhibitory mechanisms employed by Acr proteins.

INTRODUCTION

The clustered regularly interspaced short palindromic repeats (CRISPR) and CRISPR-associated (Cas) proteins provide bacteria with adaptive immunity against phages (1,2). When phages infect bacteria, bacterial cells activate the CRISPR-Cas system to acquire short segments of the viral DNA into their own genome (3,4). The viral DNA fragments are inserted as variable spacer sequences between invariable pseudo-palindromic repeat sequences in the host genome (4,5). The CRISPR array is transcribed as a long precursor CRISPR RNA (pre-crRNA) that is processed into small mature crRNAs (6,7). The crRNAs guide a single protein effector or a multi-protein effector complex to degrade the complementary sequences of invading viral nucleic acids (7,8).

Anti-CRISPR (Acr) proteins are encoded by phages to neutralize the CRISPR-Cas system of bacteria (9,10). Acr proteins were first discovered in phages infecting the Gram-negative pathogen *Pseudomonas aeruginosa*, which contains the type I-F CRISPR-Cas system (11). Additional Acr proteins have been found against type I-F and type I-E CRISPR-Cas systems by bioinformatics and biochemical studies (12). Structural and functional analyses revealed different inhibitory mechanisms between type I-F Acr proteins (13–16). AcrF1 binds to the effector complex and interferes with the base-pairing between crRNA and its target DNA (15). AcrF2 prevents the binding of double-stranded target DNA to the effector complex (15). AcrF3 interacts with the nuclease Cas3 and inhibits the association between Cas3 and the effector complex as well as substrate DNA binding (16).

More recently, two different groups of Acr proteins, namely, AcrIIA and AcrIIC, were discovered, which target type II-A and type II-C CRISPR-Cas systems of *Listeria monocytogenes* and *Neisseria meningitidis*, respectively (17,18). These Acr proteins inhibit type II CRISPR-Cas systems containing the single effector protein Cas9, and two AcrIIA proteins (AcrIIA2 and AcrIIA4) exhibit cross-reactivity to inhibit *Streptococcus pyogenes* Cas9, which is widely used in the field of genome editing (17). The complex structures of AcrIIA4 and *S. pyogenes* Cas9 have revealed that AcrIIA4 binds to Cas9 with high affinity to block the substrate DNA binding (19–21). Similarly, AcrIIC3 inhibits substrate DNA binding to *N. meningitidis* Cas9, but AcrIIC1 disables the conformational transition of Cas9 to the active state without affecting substrate DNA binding (22).

Accumulating experimental data demonstrate that the sequences, sizes, structures and mechanisms vary among different Acr proteins. Understanding the structural and inhibitory mechanisms of novel Acr proteins would provide a better opportunity to control CRISPR-Cas function for general application in genome editing. In this study, we report the crystal structure of AcrIIA1 and its binding affin-

ity for nucleic acids. AcrIIA1 is the most prevalent among the Acr proteins interfering with the *L. monocytogenes* type II-A CRISPR-Cas system and has been used as a marker to identify other AcrIIA proteins [17]. Our analyses of AcrIIA1 reveal its unique structural and functional features distinct from previously characterized Acr proteins, suggesting a novel mechanism for inactivating CRISPR-mediated immune function.

MATERIALS AND METHODS

Cloning, expression and purification

The synthetic *acrIIA1* gene was cloned into a pBT7-N-His vector with an N-terminal (His)$_6$ tag and a tobacco etch virus (TEV) protease cleavage site. *Escherichia. coli* BL21 (DE3) cells containing these constructs were cultured in lysogeny broth medium at 37°C until the optical density at 600 nm reached 0.8. Protein expression was induced by the addition of 1 mM isopropyl-β-D-thiogalactopyranoside, followed by incubation at 18°C for 20 h. Cells were harvested by centrifugation and resuspended in lysis buffer (200 mM NaCl, 2 mM β-mercaptoethanol (BME), 1 mM ethylenediaminetetraacetic acid (EDTA), 1 mM phenyl-methylsulfonyl fluoride, 20 mM Tris–HCl pH 8.0).

After lysis by using EmulsiFlex-C3 (AVESTIN) and centrifugation, the supernatant was loaded onto a 5-ml His-Trap HP column (GE Healthcare) that was pre-equilibrated with elution buffer (200 mM NaCl, 2 mM BME, 20 mM Tris–HCl pH 8.0). After washing the column with elution buffer, the bound proteins were eluted by applying a linear gradient of imidazole (up to 500 mM) and dialyzed against TEV proteolysis buffer (5 mM BME, 0.5 mM EDTA, 50 mM Tris–HCl pH 8.0). The (His)$_6$-MBP tag was cleaved by TEV protease and separated on a 5-ml HisTrap HP column (GE Healthcare). The proteins were further purified using a HiLoad 16/60 Superdex 75 column (GE Healthcare) equilibrated with size exclusion chromatography buffer (100 mM NaCl, 2 mM BME, 20 mM Tris–HCl pH 7.4). Finally, AcrIIA1 was loaded onto a Mono-Q anion exchange column (GE Healthcare) equilibrated with buffer (3 mM BME, 20 mM Tris–HCl pH 7.4) and eluted with a linear gradient of NaCl (up to 1 M).

The L52M mutant construct, which was used to introduce an additional anomalous scatterer, was generated using polymerase chain reaction with mismatched primers. The selenomethionyl L52M mutant was expressed in *E. coli* BL21 (DE3) cells grown in M9 medium supplemented with SeMet as previously described [23]. The selenomethionyl mutant protein was purified as described above for the native WT AcrIIA1 protein.

Crystallization, data collection and structure determination

Native WT AcrIIA1 crystals were grown at 20°C by the sitting-drop method from 2.5 mg/ml protein solution in buffer (100 mM NaCl, 5 mM BME, 10 mM Tris–HCl pH 7.4) mixed with an equal amount of reservoir solution (13% [w/v] Polyethylene glycol (PEG) 6000, 0.1 M magnesium acetate, 50 mM sodium cacodylate pH 6.5). Selenomethionyl L52M mutant crystals were grown at 20°C by the sitting-drop method from 6.5 mg/ml protein solution in buffer (100 mM NaCl, 5 mM BME, 10 mM Tris–HCl pH 7.4) mixed with an equal amount of reservoir solution (6% [w/v] PEG 4000, 0.2 M NaCl, 100 mM 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES) pH 7.0). Crystals were cryoprotected in the reservoir solutions supplemented with additional ethylene glycol and flash-frozen in liquid nitrogen. Diffraction data were collected at the beamline 7A of the Pohang Accelerator Laboratory at 100 K. Diffraction images were processed with HKL2000 [24]. The determination of selenium positions, density modification and initial model building for the selenomethionyl L52M mutant structure were performed using PHENIX [25]. The mutant structure was used for molecular replacement phasing of the WT AcrIIA1 structure in PHASER [26]. The final structures were completed using alternate cycles of manual fitting in COOT [27] and refinement in PHENIX [25]. The stereochemical quality of the final models was assessed using MolProbity [28].

Analysis of co-purifying nucleic acids

The histidyl-tagged AcrIIA1 protein was expressed in *E. coli* as described above. *Escherichia coli* cells were harvested by centrifugation and resuspended in lysis buffer (300 mM NaCl, 5 mM BME, 10% [w/v] glycerol, 20 mM Tris–HCl pH 7.4). After sonication and centrifugation, the supernatant was loaded onto a 5 ml HisTrap HP column (GE Healthcare) pre-equilibrated with affinity chromatography buffer (300 mM NaCl, 5 mM BME, 10% [w/v] glycerol, 20 mM Tris–HCl pH 7.4). After washing the column with buffer, the bound samples were eluted by applying a linear gradient of imidazole (up to 450 mM). The eluate was loaded onto a Mono-Q anion exchange column (GE Healthcare) equilibrated with anion exchange chromatography buffer (100 mM NaCl, 5 mM BME, 10% [w/v] glycerol, 20 mM Tris–HCl pH 7.4). Both proteins and nucleic acids were bound to the column but eluted separately by applying a linear gradient of NaCl (up to 1 M). The eluted fractions from the anion exchange column were analyzed by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) and agarose gel electrophoresis. Proteins and nucleic acids were stained with Coomassie Brilliant Blue and ethidium bromide, respectively. For identification of the co-purifying nucleic acids, the eluted fraction containing the nucleic acids was treated with RNase-free DNase I (5 U) and DNase-free RNase A (5 U) at 37°C for 1 h and analyzed on a 2% agarose gel. To estimate the size distribution, the co-purifying nucleic acids were analyzed by 4.5% urea denaturing PAGE and visualized by SYBR Gold (Thermo Fisher Scientific) staining.

RESULTS

AcrIIA1 forms a dimer with a two-domain architecture

To gain structural insight into the inhibitory mechanism of AcrIIA1, we solved its crystal structure to a resolution of 2.0 Å. The structure of the selenomethionine-substituted L52M mutant of AcrIIA1 was determined by single-wavelength anomalous diffraction and was used to find the phase solution for the wild-type (WT) AcrIIA1

Table 1. Data collection, phasing and refinement statistics[a]

	Native WT	SeMet L52M mutant
Space group	C2	P2$_1$2$_1$2$_1$
Unit cell parameters (Å)	$a = 127.4, b = 55.4, c = 46.5, \beta = 96.0°$	$a = 55.1, b = 56.5, c = 90.3$
Wavelength (Å)	0.9793	0.9792
Data collection statistics		
Resolution range (Å)	50.00–2.00 (2.07–2.00)	50.00–1.85 (1.92–1.85)
Number of reflections	151 968 (21 982)	350 325 (24 517)
Completeness (%)	99.9 (100.0)	100.0 (100.0)
R_{merge}[b]	0.112 (0.716)	0.078 (0.765)
Redundancy	6.9 (6.8)	7.5 (7.4)
Mean I/σ	19.9 (3.8)	21.8 (2.5)
Phasing statistics		
f', f'' used in phasing		−8.0, 4.5
Figure of merit		0.504
Refinement statistics		
Resolution range (Å)	28.33–2.00	34.91–1.85
R_{cryst}[c]/R_{free}[d] (%)	18.2/22.8	18.5/23.6
RMSD bonds (Å)	0.007	0.008
RMSD angles (deg)	0.900	0.962
Average B-factor (Å^2)	46.4	32.9
Number of water molecules	134	188
Ramachandran favored (%)	99.0	99.0
Ramachandran allowed (%)	1.0	1.0

[a]Values in parentheses are for the highest-resolution shell.
[b]$R_{merge} = \Sigma_h \Sigma |I_i(h) - <I(h)>|/ \Sigma_h \Sigma_i I_i(h)$, where $I_i(h)$ is the intensity of an individual measurement of the reflection and $<I(h)>$ is the mean intensity of the reflection.
[c]$R_{cryst} = \Sigma_h ||F_{obs}|-|F_{calc}||/\Sigma_h |F_{obs}|$, where F_{obs} and F_{calc} are the observed and calculated structure factor amplitudes, respectively.
[d]R_{free} was calculated as R_{cryst} using ∼5% of the randomly selected unique reflections that were omitted from structure refinement.

structure. Data collection, phasing and refinement statistics are summarized in Table 1. The asymmetric units of both WT and mutant structures, which were crystallized at disparate conditions in different space groups, contained two AcrIIA1 protomers forming a single dimer with pseudo 2-fold symmetry. The dimeric state of AcrIIA1 was also supported by size exclusion chromatography analysis (Supplementary Figure S1). The root-mean-square deviation (RMSD) values of the Cα atomic positions between the WT and mutant structures were only 0.92 Å, indicating that the L52M mutation, which was used to introduce the additional anomalous scatterer, did not disrupt the structural integrity of the protein and thus the two structures are very similar (Supplementary Figure S2). We hereafter describe the WT structure since several residues in the mutant structure were missing in the final model due to insufficient electron density despite its slightly higher resolution (1.85 Å) than that of the WT structure.

The crystal structure of AcrIIA1 revealed an all-helical two-domain architecture (Figure 1A and B). The AcrIIA1 protomer consists of a globular N-terminal domain (residues 4–72) and an extended C-terminal domain (residues 73–149). The N-terminal domain comprises five α-helices (α1–α5) and a single 3$_{10}$ helix (η1). The first three α-helices (α1–α3) form an approximately triangular arrangement. The two additional α-helices (α4 and α5) of the N-terminal domain are almost perpendicular to one another, and pack against the α1 helix to make a compact helical assembly. As predicted from its amino acid sequence (17), helices α2 and α3 (residues 17–36) constitute a helix-turn-helix (HTH) motif to form the preceding and recognition helices, respectively, joined by the characteristic sharp turn (29). The C-terminal domain of AcrIIA1 consists of three

α-helices (α6–α8) and two 3$_{10}$ helices (η2 and η3). Except for the η2 helix, the other four helices (α6–α8 and η3) of the C-terminal domain are approximately parallel or antiparallel to one another and are stabilized by a closely packed hydrophobic core reminiscent of the coiled-coil motif. The α7 and η3 helices are connected by a long loop (residues 106–117) spanning ∼30 Å, in which the electron density is relatively poorly defined. Residues in or adjacent to the η3 helix show a notable structural deviation when the two AcrIIA1 protomers are structurally aligned. In fact, the η3 helix is not recognizable as a secondary structural element in chain B of the crystal structure. These observations suggest the flexible nature of this region in AcrIIA1.

The overall shape of the AcrIIA1 homodimer resembles a bird with open wings in which the N- and C-terminal domains represent the body and the wings, respectively (Figure 1C). The two N-terminal domains interact closely to form a central core of the dimer structure with a cleft between the two HTH motifs. The distance between the N-termini of the two α3 helices is ∼35 Å. The C-terminal domains comprise the wings of the bird-like structure with a wingspan of ∼83 Å. The two extended C-terminal domains protrude from the central core in opposite directions nearly parallel to the dimer interface formed by the N-terminal domains. The dimerization of AcrIIA1 buries 2464 Å^2 of the accessible surface area, and forms 11 hydrogen bonds and two salt bridges between the two protomers. The dimer interface can be divided mainly into three regions (Figure 2). In the N-terminal domain, three helices (η1, α4 and α5) following the α3 recognition helix of the HTH motif make extensive intermolecular contacts with those of the other protomer. Residues at the helical interface form a network of polar and hydrophobic interactions of a nearly 2-fold symmetry, in-
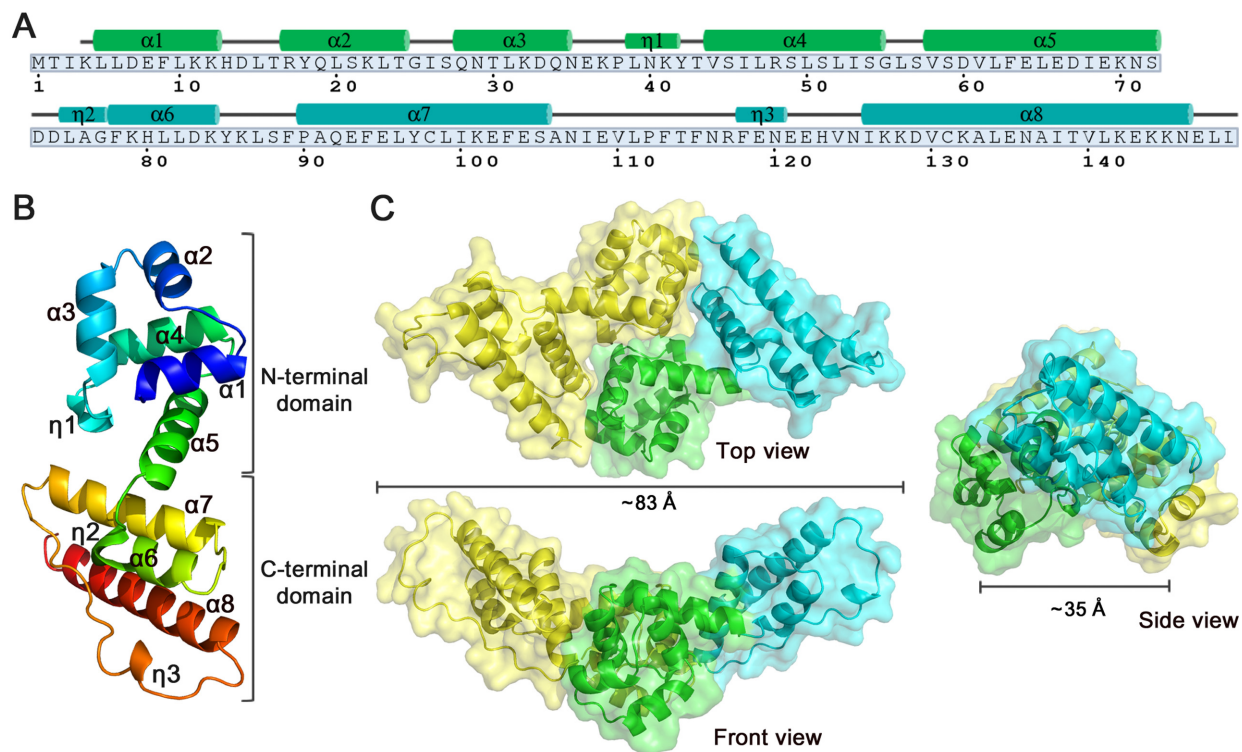
Figure 1. Crystal structure of AcrIIA1. (The reader is referred to the online version for color indication.) (**A**) Schematic representation of the secondary structure of AcrIIA1. The amino acid sequence of AcrIIA1 is shown and numbered below. (**B**) Protomer structure of AcrIIA1. Protomer A is colored in rainbow format from N-terminus (blue) to C-terminus (red). Secondary structure elements are also indicated. (**C**) Dimeric structure of AcrIIA1. N- and C-terminal domains of protomer A are shown in green and cyan, respectively. Protomer B is shown in yellow.
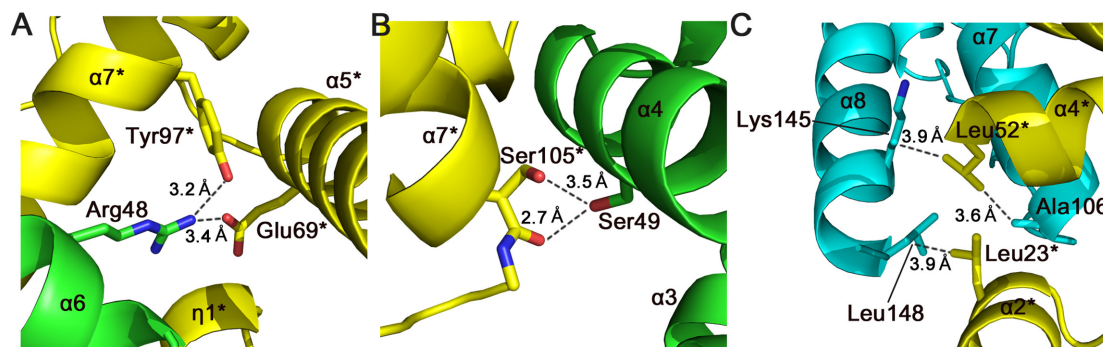

Figure 2. Dimerization interface of AcrIIA1. (**A**) Arg48 forms a salt bridge and a hydrogen bond with Glu69* and Tyr97*, respectively. (**B**) Hydrogen bonds are formed between Ser49 and Ser105*. (**C**) Ala106, Lys145 and Leu148 make hydrophobic contacts with Leu23* and Leu52*. The AcrIIA1 structure is colored as in Figure 1C. The asterisk denotes residues from the other protomer.

cluding the two symmetric Arg48–Glu69* salt bridges (the asterisk denotes residues from the other protomer; Figure 2A). Another dimer interface is formed between the α7 helix in the C-terminal domain of one protomer and the α4 helix in the N-terminal domain of the other protomer. This second interface region is stabilized by hydrogen bonds between Arg48 and Tyr97* (Figure 2A), and between Ser49 and Ser105* (Figure 2B), and a hydrophobic interaction between Leu52 and Ala106* (Figure 2C). In addition to these interface regions, Lys145 and Leu148 at the C-terminus make hydrophobic contacts with Leu23* and Leu52*, re-

spectively, in the N-terminal domain of the other protomer (Figure 2C).

AcrIIA1 exhibits structural similarity to HTH transcription factors

A search for structural neighbors using the Dali server (30) did not identify any significant structural matches to the entire AcrIIA1 protomer in the Protein Data Bank (PDB) (31). However, its N-terminal domain (residues 4–72) reveals a compelling structural resemblance to the HTH domains of many transcription factors, including the well-
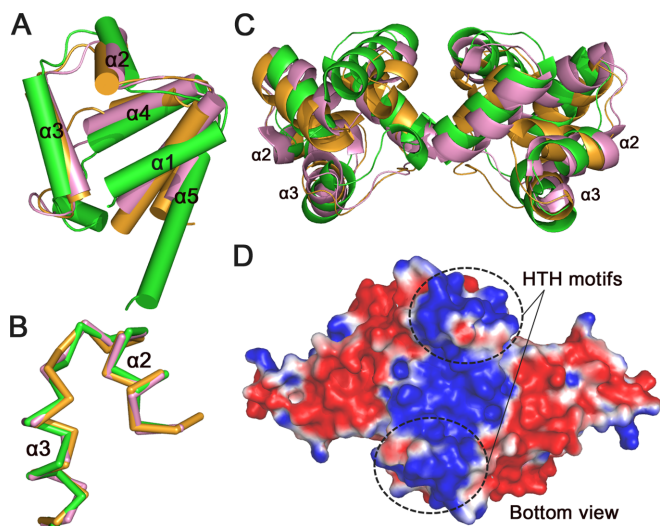
Figure 3. Comparison of AcrIIA1 and HTH transcription factors. (The reader is referred to the online version for color indication.) (**A**) Structure superposition of the N-terminal domain of AcrIIA1 (green) and DNA-binding domains of the *Bacillu subtilis* SinR protein (pink) and the phage 434 cI repressor (orange). Helices are shown as solid cylinders. (**B**) Structural alignment of the HTH motifs of AcrIIA1 and the two transcription factors. The Cα traces of the proteins are colored as in (A). (**C**) Overlay of the AcrIIA1 N-terminal dimer with HTH domains of the SinR and 434 repressor dimers (PDB IDs: 3ZKC and 2OR1, respectively). AcrIIA1 and the transcription factors are shown in cartoon representation and colored as in (A). (**D**) Electrostatic potential surface of AcrIIA1 dimer with positive charges around HTH motifs, typical of HTH transcription factors. Pymol (www.pymol.org) was used to generate the surface (red = −1.0 kT, blue = +1.0 kT). The HTH motifs are indicated by dashed lines.

studied *Bacillus subtilis* sporulation inhibitor SinR (PDB ID: 1B0N; *Z*-score: 6.9; 2.4 Å RMSD for 61 Cα atoms) (32) and the phage 434 cI repressor (PDB ID: 1R69; *Z*-score: 5.8; 2.6 Å RMSD for 58 Cα atoms) (33). The topology and fold of the AcrIIA1 N-terminal domain is similar to those of the DNA-binding domains of the SinR and 434 cI transcriptional regulators (Figure 3A); all of them contain five α-helices that fold in a nearly identical manner into a compact globular conformation. The HTH motifs, composed of the second and third α-helices (α2 and α3), display higher structural similarity than the rest of the domains (Figure 3B). The structural similarity extends to the level of the quaternary structure (Figure 3C). When the dimers of the HTH domains are structurally aligned, the protomers are arranged in a very similar orientation to each other with pseudo 2-fold symmetry. The dimer interfaces are essentially identical. The RMSD values between the AcrIIA1 dimer and the two structural neighbors are 2.5 Å (for 102 Cα atoms) and 3.2 Å (for 101 Cα atoms) for the SinR and 434 cI repressor dimers, respectively.

Analysis of the electrostatic potential surface also revealed the similarity between AcrIIA1 and the HTH transcription factors. The theoretical pI of AcrIIA1 is calculated to be ∼5.5, suggesting that the Acr protein possesses a net negative charge at neutral pH. However, the electrostatic potential of the AcrIIA1 structure revealed that positive charges are densely populated at the HTH motifs and the cleft formed between them (Figure 3D). The positively charged surface around the HTH motifs has been identified

in many transcription factors and is thought to play crucial roles in DNA recognition through electrostatic interactions with the phosphate backbone (34–36).

The possibility of DNA recognition by AcrIIA1 led us to test its binding affinity for DNAs (Supplementary Figure S3). It was previously shown that AcrIIA1 is not involved in the transcriptional control of Cas9 (17). Thus, if AcrIIA1 is a transcriptional repressor like the SinR protein and the phage 434 cI repressor, it is likely that the protein binds to the promoter regions of the tracrRNA or pre-crRNA for the CRISPR inhibitory function. However, we were unable to demonstrate the binding of AcrIIA1 to specific double-stranded (ds) and single-stranded (ss) DNA sequences in the putative promoter regions in our electrophoretic mobility shift assay (Supplementary Figure S3A, B and D). Moreover, AcrIIA1 did not show any affinity for the linearized pUC19 plasmids upon testing its non-specific DNA binding (Supplementary Figure S3C). We also tested the AcrIIA1 interaction with the CRISPR repeat RNA, its complementary tracrRNA fragment and the RNA duplex formed by them, but no binding was detected (Supplementary Figure S3A and E). These results suggest that sequences and/or types of biologically relevant binding partner(s) for AcrIIA1 are different from those of the tested nucleic acids.

AcrIIA1 associates with RNAs when overexpressed in *E. coli*

When overexpressed in *E. coli*, AcrIIA1 associated with a significant amount of nucleic acids. During purification of the His$_6$-tagged version of AcrIIA1, abundant co-purifying nucleic acids were observed following Ni-affinity chromatography. Gel electrophoretic analyses revealed the association between nucleic acids and AcrIIA1 (Figure 4B, lane 2). We were able to separate the nucleic acids from AcrIIA1 using additional anion exchange chromatography (Figure 4A and B) and tested their sensitivity to DNA and RNA nucleases (DNase and RNase, respectively; Figure 4C). Treatment with RNase A resulted in almost complete degradation of the nucleic acids, whereas such degradation was not observed when DNase I was added, indicating that the co-purifying nucleic acids were mostly RNAs. The sizes of the co-purifying RNAs were heterogeneous, ranging from ∼100 to ∼2000 nt (Figure 4D). These results are consistent with the observation that the AcrIIA1 structure contains the nucleic acid binding motifs with the positively charged surface (Figure 3), supporting that AcrIIA1 functions via nucleic acid recognition.

DISCUSSION

Despite its structural similarity to HTH transcription factors, AcrIIA1 did not display DNA- or RNA-binding affinity in the electrophoretic mobility shift assays (Supplementary Figure S3). We cannot exclude the possibility that our experimental conditions are not optimal for the binding of AcrIIA1. However, it is also conceivable that AcrIIA1 possesses intrinsic binding affinity for distinct, untested nucleic acids. We found abundant co-purifying RNAs when AcrIIA1 was overexpressed in *E. coli*, suggesting that AcrIIA1 can interact with nucleic acids. It is premature to conclude that AcrIIA1 interacts with RNAs more
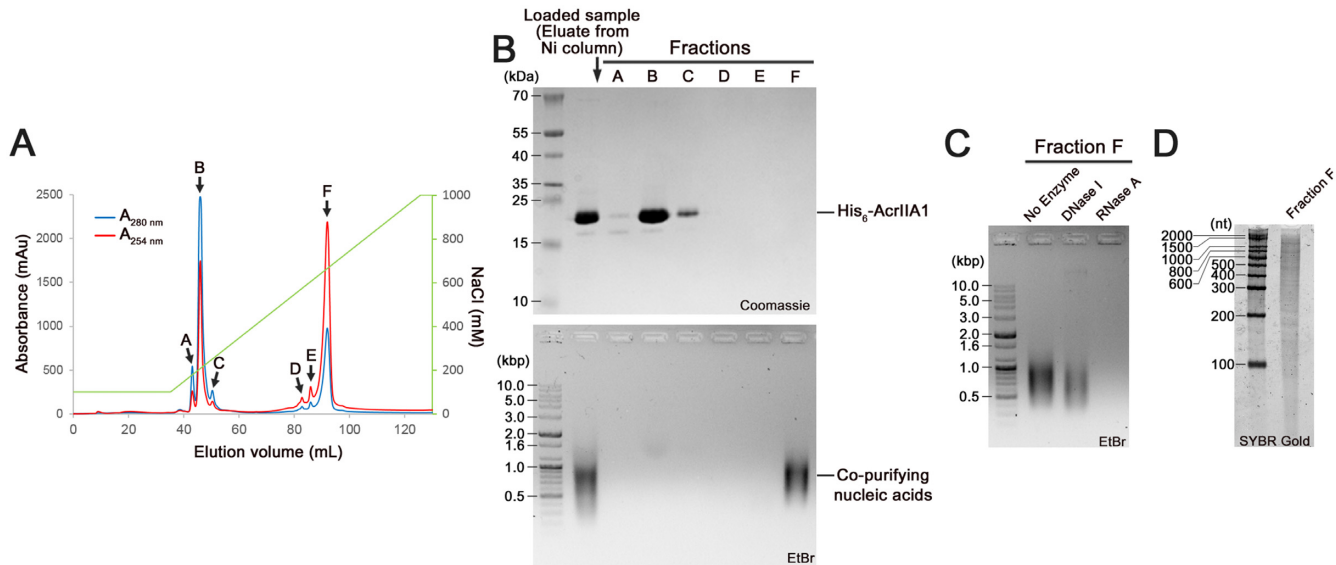
Figure 4. Analysis of nucleic acids co-purifying with AcrIIA1. (**A**) Separation of AcrIIA1 and co-purifying nucleic acids using anion exchange chromatography. The eluate from Ni-affinity chromatography was loaded onto an anion exchange column. AcrIIA1 and the co-purifying nucleic acids were eluted separately by applying a linear gradient of NaCl. (**B**) Gel electrophoretic analyses of AcrIIA1 and co-purifying nucleic acids. The loaded sample (lane 2) and the fractions corresponding to the elution peaks from the anion exchange chromatography (Lanes 3–8) were analyzed by SDS-PAGE (top) and on an agarose gel (bottom). (**C**) Identification of the co-purifying nucleic acids as RNAs. The fraction F containing the co-purifying nucleic acids was treated with DNase 1 and RNase A separately and analyzed on an agarose gel. (**D**) Size distribution of the co-purifying RNAs. The fraction F was analyzed on a urea denaturing polyacrylamide gel with RNA ladder.

preferentially than with DNAs *in vivo* and functions via RNA recognition for the CRISPR-Cas inhibition. In our gel shift assays (Supplementary Figure S3E), AcrIIA1 did not bind to the tested RNAs. The co-purification of RNAs can be an artifact caused by the overexpression in *E. coli* in the absence of biologically relevant binding partner(s). Indeed, previous studies have shown that certain HTH structures can binds to both RNAs and dsDNAs (37,38).

Acr proteins (AcrIIA1–4) that suppress type II-A CRISPR-Cas systems were first identified in *L. monocytogenes* prophages (17). Inhibition against *S. pyogenes* Cas9 by direct interaction was demonstrated for two (AcrIIA2 and AcrIIA4) of the four Acr proteins (17,19–21), whereas AcrIIA1 did not inactivate *S. pyogenes* Cas9 function (17). The lack of cross-strain inhibition of AcrIIA1 may be attributed to the species barrier, indicating that the binding interface for AcrIIA1 may not be conserved between *L. monocytogenes* and *S. pyogenes* Cas9 homologs. Difficulty in purification of *L. monocytogenes* Cas9, which was also reported in a previous study (19), impeded a direct *in vitro* comparison of AcrIIA1 binding to the recombinant Cas9 homologs. Our results, however, raises another possible explanation that AcrIIA1 functions via a different inhibitory mechanism from those of the two Cas9-binding Acr inhibitors. AcrIIA1 reveals structural and functional features that are unique among the Acr proteins against the type II-A CRISPR-Cas systems, supporting its distinct mode of action. According to our crystal structure, AcrIIA1 forms a dimer with a two-domain arrangement, while AcrIIA4 forms a single-domain structure and AcrIIA2 is monomeric in solution (20). AcrIIA1 manifests nucleic acid-binding affinity, which has not been reported for AcrIIA2 and AcrIIA4. The genomic distribution of

acrIIA genes also suggests the potential mechanistic distinction between AcrIIA1 and the Cas9-interacting Acr proteins. It was previously noted that the *acrIIA2* and *acrIIA4* genes do not co-occur at the same *acr* loci, but frequently with the *acrIIA1* gene, although each gene alone is sufficient to inhibit the type II-A CRISPR-Cas system (17). It is conceivable that the redundant inhibitory mechanism may have relieved the need for the co-occurrence of the *acrIIA2* and *acrIIA4* genes, while the distinct mode of action of AcrIIA1 has evolved into the co-existence of the *acrIIA1* gene with those encoding the Cas9-binding Acr proteins.

The role of the C-terminal domain in AcrIIA1 is more elusive. A Dali search (30) found structural neighbors for the C-terminal domain (residues 73–149). The top three PDB entries with the highest *Z*-scores are the potato disease resistance protein Rx (PDB ID: 4M70; *Z*-score: 5.4; 3.1 Å RMSD for 74 Cα atoms) (39), the *Lachancea thermotolerans* autophagy-related protein Atg29 (PDB ID: 4P1W; *Z*-score: 4.5; 4.1 Å RMSD for 61 Cα atoms) (40), and the *Streptococcus mutans* adhesion P1 (PDB ID: 4TSH; *Z*-score: 4.4; 2.8 Å RMSD for 61 Cα atoms) (41). The structural similarities were less significant than those observed between the AcrIIA1 N-terminal domain and its structural homologs. It is difficult to deduce common functionality relevant to the structural features of the C-terminal structural neighbors, which can provide information on the biological role of the C-terminal domain. Nevertheless, all three structural neighbors form multi-protein complexes in the crystal structures, suggesting that the C-terminal domain of AcrIIA1 functions via interactions with other proteins. In many HTH transcription factors, DNA-binding HTH domains are connected with regulatory domains that are controlled by protein- or ligand-binding (42–44). It is

also notable that AcrIIA5, another Acr protein recently identified in a phage of *Streptococcus thermophiles*, is predicted to possess a coiled-coil motif (45), in which helices are arranged in an approximately parallel or antiparallel manner, as seen in the C-terminal domain of AcrIIA1.

Together with previous studies, our results expand the diversity of the inhibitory mechanisms employed by Acr proteins. The unique structural and functional features of AcrIIA1 suggest that AcrIIA1 function through a distinct mechanism that is yet unknown in Acr activities. AcrIIA1 may not interact directly with Cas9, but rather with other components of type II-A CRISPR-Cas systems and/or host factors. Details of the inhibitory mechanism of AcrIIA1, including the identity of its interacting partners and the role of the C-terminal domain, remain to be determined.

DATA AVAILABILITY

Atomic coordinates and structure factors for the WT and the L52M mutant AcrIIA1 structures have been deposited with the PDB under accession numbers 5Y6A and 5Y69, respectively.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

FUNDING

REFERENCES

1. Barrangou,R., Fremaux,C., Deveau,H., Richards,M., Boyaval,P., Moineau,S., Romero,D.A. and Horvath,P. (2007) CRISPR provides acquired resistance against viruses in prokaryotes. *Science*, **315**, 1709–1712.
2. Brouns,S.J., Jore,M.M., Lundgren,M., Westra,E.R., Slijkhuis,R.J., Snijders,A.P., Dickman,M.J., Makarova,K.S., Koonin,E.V. and van der Oost,J. (2008) Small CRISPR RNAs guide antiviral defense in prokaryotes. *Science*, **321**, 960–964.
3. Sternberg,S.H., Richter,H., Charpentier,E. and Qimron,U. (2016) Adaptation in CRISPR-Cas systems. *Mol. Cell*, **61**, 797–808.
4. Jackson,S.A., McKenzie,R.E., Fagerlund,R.D., Kieper,S.N., Fineran,P.C. and Brouns,S.J. (2017) CRISPR-Cas: adapting to change. *Science*, **356**, eaal5056.
5. Amitai,G. and Sorek,R. (2016) CRISPR-Cas adaptation: insights into the mechanism of action. *Nat. Rev. Microbiol.*, **14**, 67–76.
6. Mohanraju,P., Makarova,K.S., Zetsche,B., Zhang,F., Koonin,E.V. and van der Oost,J. (2016) Diverse evolutionary roots and mechanistic variations of the CRISPR-Cas systems. *Science*, **353**, aad5147.
7. van der Oost,J., Westra,E.R., Jackson,R.N. and Wiedenheft,B. (2014) Unravelling the structural and mechanistic basis of CRISPR-Cas systems. *Nat. Rev. Microbiol.*, **12**, 479–492.
8. Nishimasu,H. and Nureki,O. (2017) Structures and mechanisms of CRISPR RNA-guided effector nucleases. *Curr. Opin. Struct. Biol.*, **43**, 68–78.
9. Sontheimer,E.J. and Davidson,A.R. (2017) Inhibition of CRISPR-Cas systems by mobile genetic elements. *Curr. Opin. Microbiol.*, **37**, 120–127.
10. Borges,A.L., Davidson,A.R. and Bondy-Denomy,J. (2017) The discovery, mechanisms, and evolutionary impact of anti-CRISPRs. *Ann. Rev. Virol.*, **4**, 37–59.
11. Bondy-Denomy,J., Pawluk,A., Maxwell,K.L. and Davidson,A.R. (2013) Bacteriophage genes that inactivate the CRISPR/Cas bacterial immune system. *Nature*, **493**, 429–432.
12. Pawluk,A., Staals,R.H., Taylor,C., Watson,B.N., Saha,S., Fineran,P.C., Maxwell,K.L. and Davidson,A.R. (2016) Inactivation of CRISPR-Cas systems by anti-CRISPR proteins in diverse bacterial species. *Nat. Microbiol.*, **1**, 16085.
13. Bondy-Denomy,J., Garcia,B., Strum,S., Du,M., Rollins,M.F., Hidalgo-Reyes,Y., Wiedenheft,B., Maxwell,K.L. and Davidson,A.R. (2015) Multiple mechanisms for CRISPR-Cas inhibition by anti-CRISPR proteins. *Nature*, **526**, 136–139.
14. Maxwell,K.L., Garcia,B., Bondy-Denomy,J., Bona,D., Hidalgo-Reyes,Y. and Davidson,A.R. (2016) The solution structure of an anti-CRISPR protein. *Nat. Commun.*, **7**, 13134.
15. Chowdhury,S., Carter,J., Rollins,M.F., Golden,S.M., Jackson,R.N., Hoffmann,C., Nosaka,L., Bondy-Denomy,J., Maxwell,K.L., Davidson,A.R. *et al.* (2017) Structure reveals mechanisms of viral suppressors that intercept a CRISPR RNA-guided surveillance complex. *Cell*, **169**, 47–57.
16. Wang,X., Yao,D., Xu,J.G., Li,A.R., Xu,J., Fu,P., Zhou,Y. and Zhu,Y. (2016) Structural basis of Cas3 inhibition by the bacteriophage protein AcrF3. *Nat. Struct. Mol. Biol.*, **23**, 868–870.
17. Rauch,B.J., Silvis,M.R., Hultquist,J.F., Waters,C.S., McGregor,M.J., Krogan,N.J. and Bondy-Denomy,J. (2017) Inhibition of CRISPR-Cas9 with bacteriophage proteins. *Cell*, **168**, 150–158.
18. Pawluk,A., Amrani,N., Zhang,Y., Garcia,B., Hidalgo-Reyes,Y., Lee,J., Edraki,A., Shah,M., Sontheimer,E.J., Maxwell,K.L. *et al.* (2016) Naturally occurring off-switches for CRISPR-Cas9. *Cell*, **167**, 1829–1838.
19. Dong, Guo,M., Wang,S., Zhu,Y., Wang,S., Xiong,Z., Yang,J., Xu,Z. and Huang,Z. (2017) Structural basis of CRISPR-SpyCas9 inhibition by an anti-CRISPR protein. *Nature*, **546**, 436–439.
20. Yang,H. and Patel,D.J. (2017) Inhibition mechanism of an anti-CRISPR suppressor AcrIIA4 targeting SpyCas9. *Mol. Cell*, **67**, 117–127.
21. Shin,J., Jiang,F., Liu,J.J., Bray,N.L., Rauch,B.J., Baik,S.H., Nogales,E., Bondy-Denomy,J., Corn,J.E. and Doudna,J.A. (2017) Disabling Cas9 by an anti-CRISPR DNA mimic. *Sci. Adv.*, **3**, e1701620.
22. Harrington,L.B., Doxzen,K.W., Ma,E., Liu,J.J., Knott,G.J., Edraki,A., Garcia,B., Amrani,N., Chen,J.S., Cofsky,J.C. *et al.* (2017) A broad-spectrum inhibitor of CRISPR-Cas9. *Cell*, **170**, 1224–1233.
23. Mark,B.L., Vocadlo,D.J., Knapp,S., Triggs-Raine,B.L., Withers,S.G. and James,M.N. (2001) Crystallographic evidence for substrate-assisted catalysis in a bacterial beta-hexosaminidase. *J. Biol. Chem.*, **276**, 10330–10337.
24. Otwinowski,Z. and Minor,W. (1997) Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol*, **276**, 307–326.
25. Adams,P.D., Afonine,P.V., Bunkoczi,G., Chen,V.B., Davis,I.W., Echols,N., Headd,J.J., Hung,L.W., Kapral,G.J., Grosse-Kunstleve,R.W. *et al.* (2010) PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D Biol. Crystallogr.*, **66**, 213–221.
26. McCoy,A.J., Grosse-Kunstleve,R.W., Adams,P.D., Winn,M.D., Storoni,L.C. and Read,R.J. (2007) Phaser crystallographic software. *J. Appl. Crystallogr.*, **40**, 658–674.
27. Emsley,P. and Cowtan,K. (2004) Coot: model-building tools for molecular graphics. *Acta Crystallogr. D Biol. Crystallogr.*, **60**, 2126–2132.
28. Chen,V.B., Arendall,W.B. 3rd, Headd,J.J., Keedy,D.A., Immormino,R.M., Kapral,G.J., Murray,L.W., Richardson,J.S. and Richardson,D.C. (2010) MolProbity: all-atom structure validation for

macromolecular crystallography. *Acta Crystallogr. D Biol. Crystallogr.*, **66**, 12–21.

29. Wintjens,R. and Rooman,M. (1996) Structural classification of HTH DNA-binding domains and protein-DNA interaction modes. *J. Mol. Biol.*, **262**, 294–313.

30. Holm,L. and Rosenstrom,P. (2010) Dali server: conservation mapping in 3D. *Nucleic Acids Res.*, **38**, W545–W549.

31. Berman,H.M., Westbrook,J., Feng,Z., Gilliland,G., Bhat,T.N., Weissig,H., Shindyalov,I.N. and Bourne,P.E. (2000) The Protein Data Bank. *Nucleic Acids Res.*, **28**, 235–242.

32. Lewis,R.J., Brannigan,J.A., Offen,W.A., Smith,I. and Wilkinson,A.J. (1998) An evolutionary link between sporulation and prophage induction in the structure of a repressor:anti-repressor complex. *J. Mol. Biol.*, **283**, 907–912.

33. Mondragon,A., Subbiah,S., Almo,S.C., Drottar,M. and Harrison,S.C. (1989) Structure of the amino-terminal domain of phage 434 repressor at 2.0 A resolution. *J. Mol. Biol.*, **205**, 189–200.

34. Newman,J.A., Rodrigues,C. and Lewis,R.J. (2013) Molecular basis of the activity of SinR protein, the master regulator of biofilm formation in Bacillus subtilis. *J. Biol. Chem.*, **288**, 10766–10778.

35. Aggarwal,A.K., Rodgers,D.W., Drottar,M., Ptashne,M. and Harrison,S.C. (1988) Recognition of a DNA operator by the repressor of phage 434: a view at high resolution. *Science*, **242**, 899–907.

36. Rumpel,S., Razeto,A., Pillar,C.M., Vijayan,V., Taylor,A., Giller,K., Gilmore,M.S., Becker,S. and Zweckstetter,M. (2004) Structure and DNA-binding properties of the cytolysin regulator CylR2 from Enterococcus faecalis. *EMBO J.*, **23**, 3632–3642.

37. Liu,W., Seto,J., Sibille,E. and Toth,M. (2003) The RNA binding domain of Jerky consists of tandemly arranged helix-turn-helix/homeodomain-like motifs and binds specific sets of mRNAs. *Mol. Cell. Biol.*, **23**, 4083–4093.

38. Niessing,D., Driever,W., Sprenger,F., Taubert,H., Jackle,H. and Rivera-Pomar,R. (2000) Homeodomain position 54 specifies transcriptional versus translational control by Bicoid. *Mol. Cell*, **5**, 395–401.

39. Hao,W., Collier,S.M., Moffett,P. and Chai,J. (2013) Structural basis for the interaction between the potato virus X resistance protein (Rx) and its cofactor Ran GTPase-activating protein 2 (RanGAP2). *J. Biol. Chem.*, **288**, 35868–35876.

40. Fujioka,Y., Suzuki,S.W., Yamamoto,H., Kondo-Kakuta,C., Kimura,Y., Hirano,H., Akada,R., Inagaki,F., Ohsumi,Y. and Noda,N.N. (2014) Structural basis of starvation-induced assembly of the autophagy initiation complex. *Nat. Struct. Mol. Biol.*, **21**, 513–521.

41. Heim,K.P., Crowley,P.J., Long,J.R., Kailasan,S., McKenna,R. and Brady,L.J. (2014) An intramolecular lock facilitates folding and stabilizes the tertiary structure of Streptococcus mutans adhesin P1. *Proc. Natl. Acad. Sci. U.S.A.*, **111**, 15746–15751.

42. Aravind,L., Anantharaman,V., Balaji,S., Babu,M.M. and Iyer,L.M. (2005) The many faces of the helix-turn-helix domain: transcription regulation and beyond. *FEMS Microbiol. Rev.*, **29**, 231–262.

43. Colledge,V.L., Fogg,M.J., Levdikov,V.M., Leech,A., Dodson,E.J. and Wilkinson,A.J. (2011) Structure and organisation of SinR, the master regulator of biofilm formation in Bacillus subtilis. *J. Mol. Biol.*, **411**, 597–613.

44. Shi,K., Brown,C.K., Gu,Z.Y., Kozlowicz,B.K., Dunny,G.M., Ohlendorf,D.H. and Earhart,C.A. (2005) Structure of peptide sex pheromone receptor PrgX and PrgX/pheromone complexes and regulation of conjugation in Enterococcus faecalis. *Proc. Natl. Acad. Sci. U.S.A*, **102**, 18596–18601.

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