

# SCIENTIFIC REPORTS



OPEN

## Solution structure and dynamics of anti-CRISPR AcrIIA4, the Cas9 inhibitor

Iktae Kim<sup>1</sup>, Migyeong Jeong<sup>1</sup>, Donghyun Ka<sup>1</sup>, Mookyoung Han<sup>1</sup>, Nak-Kyoon Kim<sup>2</sup>, Euiyoung Bae<sup>1</sup> & Jeong-Yong Suh<sup>1,3</sup>

The bacterial CRISPR-Cas system provides adaptive immunity against invading phages. Cas9, an RNA-guided endonuclease, specifically cleaves target DNA substrates and constitutes a well-established platform for genome editing. Recently, anti-CRISPR (Acr) proteins that inhibit Cas9 have been discovered, promising a useful off-switch for Cas9 to avoid undesirable off-target effects. Here, we report the solution structure and dynamics of *Listeria monocytogenes* AcrIIA4 that inhibits *Streptococcus pyogenes* Cas9 (*SpyCas9*). AcrIIA4 forms a compact monomeric  $\alpha\beta\beta\beta\alpha$  fold comprising three antiparallel  $\beta$  strands flanked by three  $\alpha$ -helices and a short  $3_{10}$ -helix. AcrIIA4 exhibits distinct backbone dynamics in fast and slow timescales at loop regions that form interaction surfaces for *SpyCas9*. In particular, the  $\beta 1$ – $\beta 2$  loop that binds to the RuvC domain of *SpyCas9* is highly mobile, and the  $\beta 1$ – $\beta 2$  and  $\alpha 2$ – $\alpha 3$  loops that bind to the RuvC and C-terminal domains of *SpyCas9*, respectively, undergoes conformational exchanges in microsecond-to-millisecond time scales. AcrIIA4 binds to apo-*SpyCas9* with  $K_D \sim 4.8 \mu\text{M}$ , which compares to  $K_D \sim 0.6 \text{ nM}$  for AcrIIA4 binding to sgRNA-bound *SpyCas9*. Since the binary complex between AcrIIA4 and *SpyCas9* does not compete with the target DNA binding, it can effectively disable the Cas9 nuclease activity by forming a tight ternary complex in the presence of sgRNA.

The clustered regularly interspaced short palindromic repeats (CRISPR) and CRISPR-associated (Cas) proteins constitute a defense system in bacteria and archaea, providing a form of immunological memory against foreign genetic elements<sup>1,2</sup>. Upon infection, the bacterial host cell activates the CRISPR-Cas system to integrate short segments of the invading nucleic acids into the CRISPR loci. The integrated DNA sequences, also known as spacers, are transcribed into CRISPR RNAs that guide a single protein effector or a multi-protein effector complex to degrade the complementary sequences in foreign genes upon subsequent invasions. The CRISPR-Cas system is categorized into two classes, six types and thirty-three subtypes according to different organizations of signature *cas* genes<sup>3</sup>. *Streptococcus pyogenes* Cas9 (*SpyCas9*) belongs to the subtype II-A of the Class 2 CRISPR-Cas system. *SpyCas9* has been extensively investigated and widely applied to genome editing due to its simplicity in generating double-strand DNA breaks at desired sites targeted by single-guide RNA (sgRNA)<sup>4–6</sup>.

Bacteria and phages have long been engaged in an arms race, developing a variety of arsenal to compete for invasion and defense<sup>7</sup>. Recently, anti-CRISPR (Acr) proteins that neutralize Cas9 activity have been discovered in phages targeting pathogenic bacterial strains<sup>8,9</sup>. Among them, AcrIIA2 and AcrIIA4 encoded by *Listeria monocytogenes* prophages have shown cross-strain inhibitory effects against *SpyCas9*, highlighting their potential as an off-switch for Cas9 for temporal and spatial control of genome editing. Recent structures of the AcrIIA4–*SpyCas9*–sgRNA complex obtained by X-ray crystallography and electron microscopy show that *SpyCas9* binds to AcrIIA4 via the protospacer adjacent motif (PAM) interaction site and the RuvC domain<sup>10–12</sup>. Here we report the first solution structure and dynamics of AcrIIA4 using NMR spectroscopy. Backbone dynamics from <sup>15</sup>N NMR relaxation data demonstrates that the binding interfaces for *SpyCas9* exhibit fast internal motions and slow conformational exchanges in variable timescales. Further, we show that AcrIIA4 binds to *SpyCas9* in the absence of sgRNA with an equilibrium dissociation constant ( $K_D$ ) of  $\sim 4.8 \mu\text{M}$ , which can associate with guide RNA to form a tight ternary complex, leading to the loss of Cas9 nuclease activity.

<sup>1</sup>Department of Agricultural Biotechnology and Research Institute of Agriculture and Life Sciences, Seoul National University, Seoul, 08826, South Korea. <sup>2</sup>Advanced Analysis Center, Korea Institute of Science and Technology, Seoul, 02792, Korea. <sup>3</sup>Institute for Biomedical Sciences, Shinshu University, Minamiminowa, Nagano, 399-4598, Japan. Correspondence and requests for materials should be addressed to J.-Y.S. (email: [jysuh@snu.ac.kr](mailto:jysuh@snu.ac.kr))

Experimental restraints	<SA>*
Nonredundant NOEs	1504
Dihedral angles, $\phi/\psi/\chi$	71/71/19
Hydrogen bonds	38
Residual dipolar coupling, $^1D_{NH}$	79
Total number of restraints	1782 (20.5 per residue)
rms deviation from experimental restraints	
Distances (Å) (1504)	0.029 ± 0.001
Torsion angles (°) (161)	0.90 ± 0.10
Residual dipolar coupling R-factor (%) <sup>†</sup>	
$^1D_{NH}$ (%) (79)	1.8 ± 0.2
rms deviation from idealized covalent geometry	
Bonds (Å)	0.002 ± 0
Angles (°)	0.47 ± 0.01
Impropers (°)	0.42 ± 0.01
Coordinate precision (Å)**	
Backbone	0.61 ± 0.11
Heavy atoms	1.38 ± 0.09
Ramachandran statistics (%) <sup>‡§</sup>	
Most favorable regions	91.2 ± 1.1
Allowed regions	8.8 ± 1.1

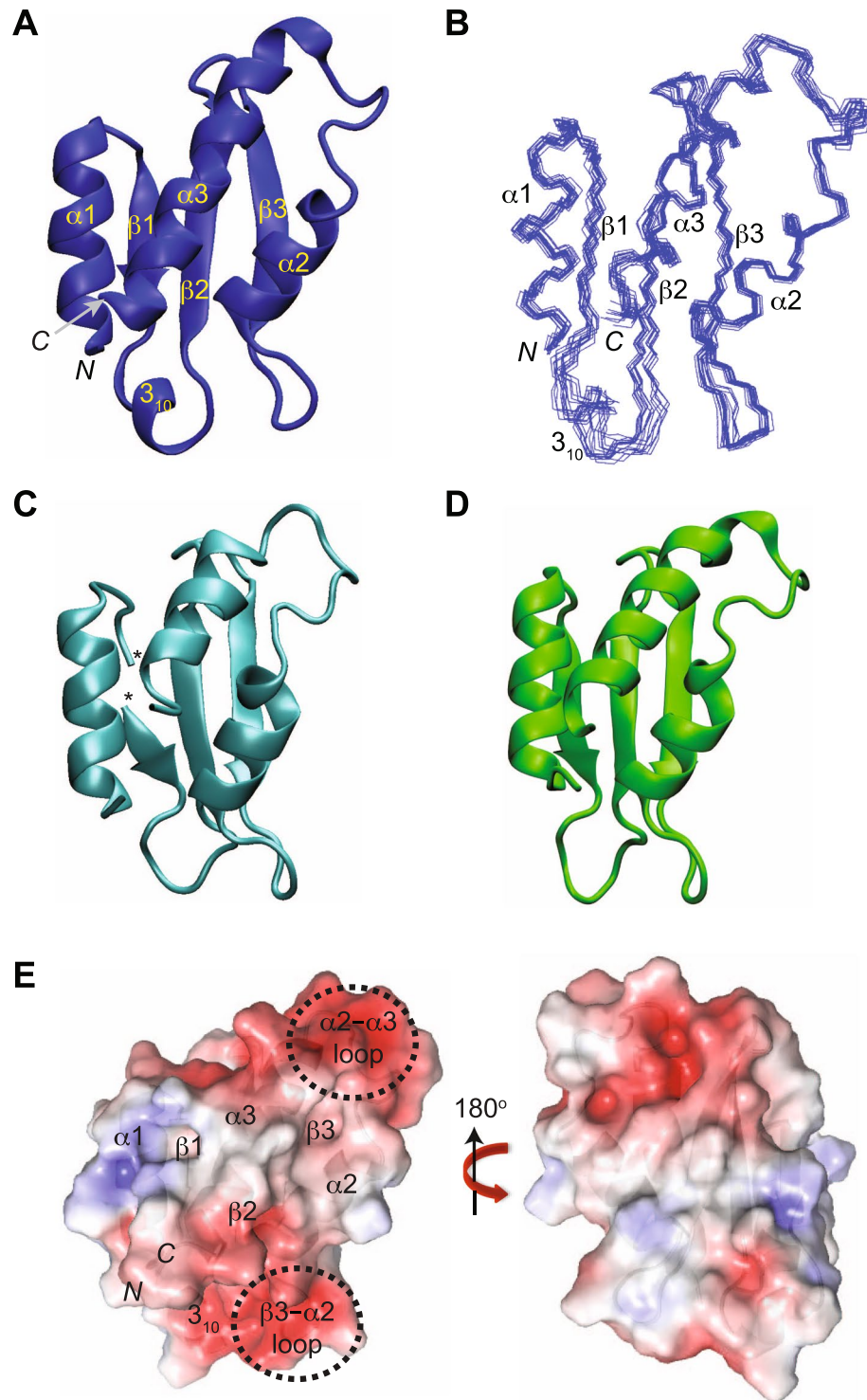
**Table 1.** Restraints and structural statistics for AcrIIA4. \*For the ensemble of the final 20 simulated annealing structures. <sup>†</sup>The magnitudes of the axial and rhombic components of the alignment tensor were 7.3 Hz and 0.35, respectively. <sup>‡</sup>Regions with secondary structures (residues 2–13, 16–19, 29–33, 40–44, 50–59, and 68–85). <sup>§</sup>Calculated using the program PROCHECK<sup>39</sup>.

## Results and Discussion

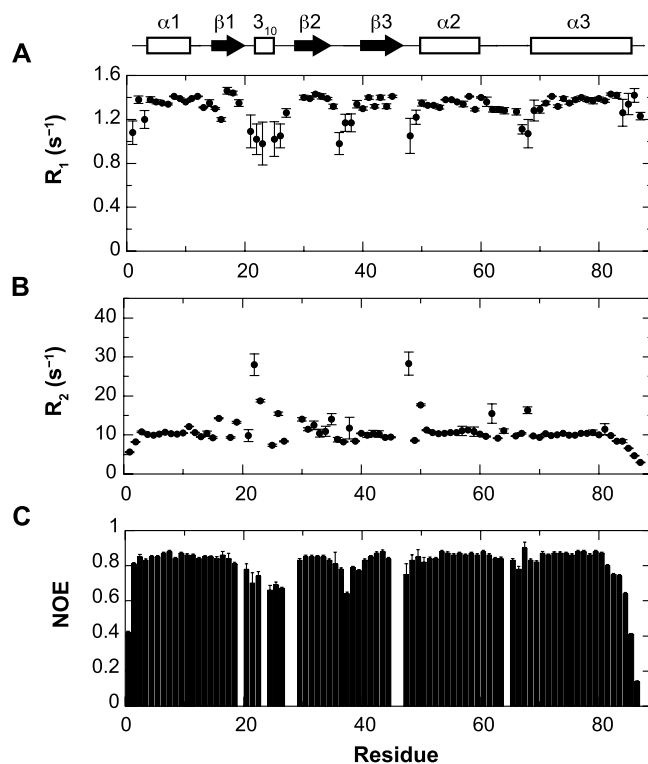
**Structural description of AcrIIA4 determined by NMR spectroscopy.** AcrIIA4 (a.a. 1–87) exists as a monomer in solution, and exhibits a well resolved  $^1H$ – $^{15}N$  HSQC NMR spectrum that is typically observed in compact folded proteins (Figures S1 and S2). We have determined the solution structure of AcrIIA4 using heteronuclear triple-resonance NMR spectroscopy. Backbone and side chain assignments were obtained using a suite of three-dimensional heteronuclear correlation NMR spectroscopy experiments. The backbone chemical shifts were completely assigned except for Thr28 that did not show backbone amide resonance in the 2D  $^1H$ – $^{15}N$  HSQC spectrum. Ser20, Ser24, Gln29, Glu47, and Gln65 did not show backbone amide resonances, but their aliphatic and carbonyl groups could be assigned from their sequential connectivities. The  $^1H$  assignment data covered 602 out of 646  $^1H$  atoms excluding those in exchangeable hydroxyl and amino groups, which amounts to the 93.2% completeness of  $^1H$  assignment. Three-dimensional  $^{13}C$ -separated NOE and  $^{15}N$ -separated NOE restraints were employed for the structure calculation using the Xplor-NIH program<sup>13</sup>. Residual dipolar couplings (RDCs) were measured in 10 mg/ml of *pf1* phage alignment medium. The structure was determined using 1,782 NMR restraints including 1,504 experimental NOE-based distance restraints, 161 dihedral angle restraints, 79 backbone  $^1D_{NH}$  RDC restraints, and 38 hydrogen bonding restraints (Table 1). The hydrogen bond restraints were introduced at the final stage of the structure calculation, based on the secondary structural information from NOE and chemical shift data. The hydrogen bond restraints were in perfect agreement with the short-range and medium range NOEs, and dihedral angle restraints. AcrIIA4 is comprised of three antiparallel  $\beta$  strands ( $\beta$ 1, 16–19;  $\beta$ 2, 29–33;  $\beta$ 3, 40–44) and three flanking  $\alpha$  helices ( $\alpha$ 1, 2–13;  $\alpha$ 2, 50–59;  $\alpha$ 3, 68–85) in an  $\alpha$ 1– $\beta$ 1– $\beta$ 2– $\beta$ 3– $\alpha$ 2– $\alpha$ 3 order (Fig. 1A). In addition, a short  $3_{10}$  helix (22–25) is formed in the loop connecting  $\beta$ 1 and  $\beta$ 2 strands. Superposition of the backbone atoms for the ensemble of the 20 lowest energy structures of AcrIIA4 demonstrates that overall secondary structures are well defined such that three helices are packed together upon a three-stranded  $\beta$  sheet (Fig. 1B). The connecting loop regions are less well-defined, and the  $\beta$ 1– $\beta$ 2 loop among others shows the largest backbone root-mean-square deviations (RMSD  $\sim$ 1.3 Å).

The overall backbone fold of free AcrIIA4 in solution was similar to that of the crystal structures determined for the AcrIIA4–*Spy*Cas9–sgRNA complexes, indicating that AcrIIA4 does not change the backbone conformation upon binding to *Spy*Cas9–sgRNA (Fig. 1C and D). The  $\beta$ 1 strand was well-defined and the C-terminal  $\alpha$ 3 helix extended between Glu68 and Glu85 in the solution structure, which was consistent with the crystal structure with higher resolution (Fig. 1D). On the other hand, the  $\beta$ 1– $\beta$ 2 loop formed a  $3_{10}$  helix between Thr22 and Asn25 in the solution structure, but did not exhibit a regular secondary structure in complex with *Spy*Cas9–sgRNA.

AcrIIA4 features a cluster of aromatic residues that form a compact hydrophobic core via stacking and edge-to-face packing between the  $\beta$ 3 strand, and  $\alpha$ 2 and  $\alpha$ 3 helices. The tightly packed aromatic side chains result in large ring current effects and unusual chemical shifts, such as  $H\beta$  of Glu71 (–0.1 ppm),  $H\alpha$  of Gln65 (0.4 ppm), and  $H\beta$  and  $H\gamma$  of Met77 (0.17 ppm and –0.05 ppm). AcrIIA4 is an acidic protein (pI  $\sim$ 4.2), and the surface electrostatic potential indicates that negative charges are densely populated in the  $\beta$ 3– $\alpha$ 2 loop, the  $\alpha$ 2– $\alpha$ 3 loop and the beginning of the  $\alpha$ 3 helix (Fig. 1E). It has been suggested that the negatively charged patches mimic



**Figure 1.** Structures of AcrIIA4 as ribbon and surface representations. (A) Solution structure of AcrIIA4 in the free state. Secondary structures, as well as N- and C-termini are annotated. (B) Superposition of the backbone atoms of the final 20 simulated annealing structures of AcrIIA4. The structures are best-fit superposed against well-ordered secondary structures between residues 2–13, 16–19, 29–33, 40–44, 50–59, and 68–85. (C) AcrIIA4 from the crystal structure of the AcrIIA4–SpyCas9–sgRNA complex (PDB code 5XBL). (D) AcrIIA4 from the crystal structure of the AcrIIA4–SpyCas9–sgRNA complex (PDB code 5VW1). (E) Front and backside view of molecular surface representations of AcrIIA4 color-coded by electrostatic potential,  $\pm 5$  kT. AcrIIA4 on the left panel is drawn in the same orientation as in (A).



**Figure 2.** Plots of relaxation parameters of backbone amide groups of AcrIIA4. (A)  $^{15}\text{N}$   $R_1$  relaxation, (B)  $^{15}\text{N}$   $R_2$  relaxation, and (C)  $^1\text{H}$ - $^{15}\text{N}$  heteronuclear NOE data as a function of residue number. The secondary structures of AcrIIA4 are shown on top as a schematic representation. The relaxation parameters were obtained on an 800 MHz Bruker NMR spectrometer.

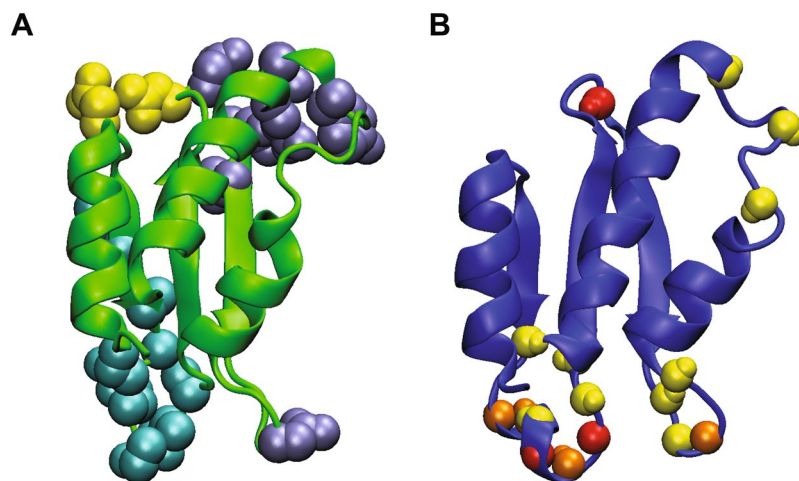
Description	$K_D$ (nM)	$\Delta G$ (kcal/mol)	$\Delta H$ (kcal/mol)	$-T\Delta S$ (kcal/mol)
AcrIIA4: <i>Spy</i> Cas9	$4800 \pm 150$	$-7.3 \pm 0.1$	$-11.5 \pm 0.7$	$4.2 \pm 0.8$
AcrIIA4: <i>Spy</i> Cas9-sgRNA	$0.6 \pm 0.1$	$-12.6 \pm 0.4$	$-10.3 \pm 0.1$	$-2.3 \pm 0.4$

**Table 2.** Thermodynamic parameters for the interaction between AcrIIA4 and *Spy*Cas9 obtained by isothermal titration calorimetry at 25 °C.

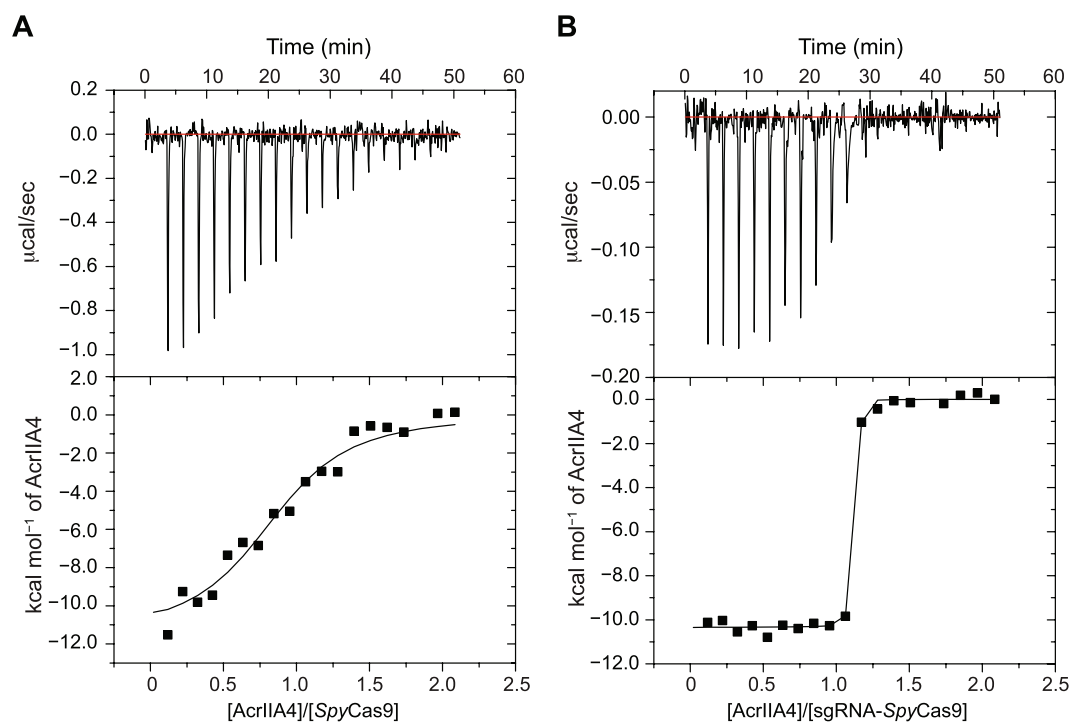
the phosphate groups of nucleic acids that associate with the PAM interaction site of *Spy*Cas9<sup>10–12</sup>, which was also found in AcrF2 proteins inhibiting the subtype I-F effector complex<sup>14</sup>.

**Backbone dynamics reveals characteristic motions at the binding interface.**  $^{15}\text{N}$   $R_1$  and  $R_2$  relaxation rates and  $^1\text{H}$ - $^{15}\text{N}$  heteronuclear NOE were measured to characterize the backbone dynamics of AcrIIA4. The rotational correlation time ( $\tau_c$ ) was calculated as 6.2 ns from the ratio of  $^{15}\text{N}$   $R_2$  and  $R_1$  relaxation rates in the ordered region, which is typical for the size of AcrIIA4 (~10.2 kDa). Secondary structural regions exhibited large  $^1\text{H}$ - $^{15}\text{N}$  heteronuclear NOE (>0.8), and uniform  $^{15}\text{N}$   $R_2$  and  $R_1$  relaxation parameters indicating that the secondary structures were well structured and rigid without internal motions (Fig. 2). On the other hand, loop regions connecting the secondary structures manifested variable dynamic behavior in a wide range of timescales (Fig. 2). Reduced  $^1\text{H}$ - $^{15}\text{N}$  heteronuclear NOE (<0.8) and  $^{15}\text{N}$   $R_1$  relaxation rates indicated that fast backbone dynamics in ps–ns timescales was prevalent in the  $\beta 1$ – $\beta 2$  loop (Gly21–Ile27) region. In addition, increased  $^{15}\text{N}$   $R_2$  relaxation rates indicated the presence of slower conformational exchanges in  $\mu\text{s}$ – $\text{ms}$  timescales in the  $\beta 1$ – $\beta 2$  loop (Thr22, Asn23, and Ser26), the  $\beta 3$ – $\alpha 2$  loop (Asn48 and Ser50) and the  $\alpha 2$ – $\alpha 3$  loop (Gly62 and Glu68) regions. The conformational exchanges also resulted in complete line broadening of backbone amide resonances in the  $\beta 1$ – $\beta 2$  loop (Ser20, Ser24, Thr28, and Gln29), the  $\beta 3$ – $\alpha 2$  loop (Glu47), and the  $\alpha 2$ – $\alpha 3$  loop (Gln65) regions.

It is remarkable that the dynamic loop regions constitute the binding interfaces between AcrIIA4 and *Spy*Cas9 in the crystal structure<sup>10–12</sup>. AcrIIA4 mainly interacts with the TOPO domain, the RuvC domain, and the C-terminal domain (CTD) of *Spy*Cas9. Specifically, AcrIIA4 employs the  $\alpha 1$ – $\beta 1$  and  $\beta 2$ – $\beta 3$  loops to interact with the TOPO domain, the  $\beta 1$  strand and the  $\beta 1$ – $\beta 2$  loop to interact with the RuvC domain, and the  $\beta 2$ – $\beta 3$ ,  $\beta 3$ – $\alpha 2$ , and  $\alpha 2$ – $\alpha 3$  loops to interact with the CTD (Fig. 3A). The interfacial loops exhibited significant backbone dynamics in a wide range of time scales. The  $\beta 1$ – $\beta 2$  loop that binds to the RuvC domain of *Spy*Cas9 was highly mobile with fast internal motions in ps–ns timescales as well as slower conformational exchanges in  $\mu\text{s}$ – $\text{ms}$  timescales (Fig. 3B). On the other hand, the  $\beta 3$ – $\alpha 2$  loop and the  $\alpha 2$ – $\alpha 3$  loop that bind to the CTD of *Spy*Cas9 exhibited



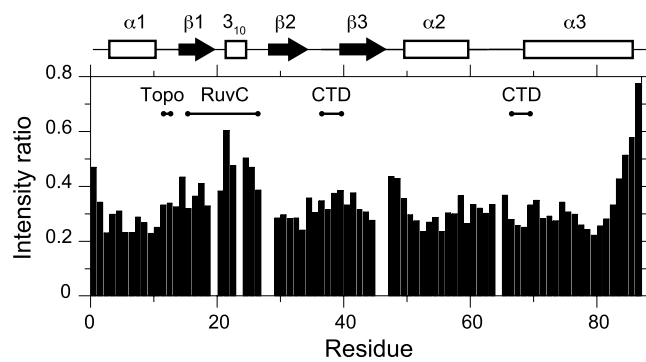
**Figure 3.** Binding interfaces of AcrIIA4 for *SpyCas9*–sgRNA and mobile regions with backbone dynamics. **(A)** Binding interfaces of AcrIIA4 from the crystal structure in complex with *SpyCas9*–sgRNA (PDB code 5VW1). AcrIIA4 is shown as a ribbon diagram, and interfacial residues in the complex are shown as a space-filling model. Interfaces for the TOPO domain, the RuvC domain, and the C-terminal domain of *SpyCas9* are colored in *yellow*, *cyan*, and *ice-blue*, respectively. **(B)** Mobile regions with backbone dynamics of AcrIIA4 from the solution structure. AcrIIA4 is shown as a ribbon diagram, and backbone amide groups are shown as a space-filling model. Residues with fast internal motions (ps–ns time scale) and slower conformational dynamics ( $\mu$ s–ms time scale) are colored in *red* and *yellow*, respectively. Residues with both fast and slower dynamics are colored in *orange*.



**Figure 4.** ITC for the interaction **(A)** between AcrIIA4 and *SpyCas9*, and **(B)** between AcrIIA4 and sgRNA-bound *SpyCas9*. Raw ITC data (Top panel) and integrated heats of injection (bottom panel) are presented for the titration between AcrIIA4 and *SpyCas9*. In the bottom panel, the experimental data are shown as solid squares and the least squares best-fit curves derived from a simple one-site binding model are shown as a black line. The thermodynamic parameters for the complex formation are listed in Table 2.

extensive conformation exchanges in  $\mu$ s–ms timescales. The binding interface for the TOPO domain appeared less dynamic compared to the other interfaces.

It has been reported that the fast and slow motions characterized by NMR relaxation play an important role in the target search and recognition process of biomolecular interactions<sup>15–18</sup>. It is notable that the dynamic  $\beta$ 1– $\beta$ 2



**Figure 5.** Intensity ratio of backbone amide resonances between AcrIIA4 in complex with *SpyCas9* (10% bound) and free AcrIIA4. The intensity ratio of backbone amide groups of AcrIIA4 measured from  $^1\text{H}$ - $^{15}\text{N}$  HSQC spectra at 25 °C is shown as a function of residue number. The intensity of 0.2 mM AcrIIA4 in free state ( $I_{\text{free}}$ ), and in complex with 0.02 mM *SpyCas9* ( $I_{\text{bound}}$ ) are measured and the ratio is presented as  $I_{\text{bound}}/I_{\text{free}}$ . The secondary structures of AcrIIA4 are shown on top as a schematic representation, and the binding interfaces of AcrIIA4 in the AcrIIA4-*SpyCas9*-sgRNA complex (PDB code 5VW1) are shown above the ratio as a visual guidance.

loop of AcrIIA4 directly interacts with the active site of the *SpyCas9* nuclease, preventing the substrate DNA binding. *SpyCas9* cleaves double-stranded DNA substrates using two nuclease domains, the HNH nuclease domain that processes the target DNA strand complementary to the guide RNA, and the RuvC nuclease domain that processes the non-target DNA strand. The  $\beta 1$ - $\beta 2$  loop of AcrIIA4 extends to the active site of RuvC, and participates in a hydrogen bonding network with catalytic residues, blocking the access of the non-target DNA strand. Meanwhile, the  $\alpha 2$ - $\alpha 3$  loop that recognizes the PAM interaction site for the target DNA strand manifests slower conformational exchanges. The  $\alpha 2$ - $\alpha 3$  loop does not directly block the active site of the nuclease, but occludes the binding site of the PAM sequence that is located immediately after the cleavage site of the target DNA strand. The loop dynamics of AcrIIA4 are thus linked to the target recognition and binding, with different time scales between the interaction surfaces.

**AcrIIA4 binds to *SpyCas9* in the absence of sgRNA.** We quantitatively measured the binding thermodynamics between AcrIIA4 and apo-*SpyCas9* using isothermal titration calorimetry (ITC). Our measurement revealed a large exothermic reaction between AcrIIA4 and *SpyCas9* with  $K_D = 4.8 \pm 1.5 \mu\text{M}$  and a 1:1 stoichiometry (Fig. 4A). The interaction between AcrIIA4 and sgRNA bound-*SpyCas9* showed a significantly higher affinity, with the  $K_D$  value of  $0.6 \pm 0.1 \text{ nM}$  (Fig. 4B). The binding affinity measured by ITC was similar to the previous report of  $K_D = 4.1 \pm 1.1 \text{ nM}$  from the microscale thermophoresis data<sup>10</sup>. Thus, the presence of sgRNA promoted the interaction between AcrIIA4 and *SpyCas9* by a factor of 8000 (Table 2). The weaker binding between AcrIIA4 and apo-*SpyCas9* explains previous observations that AcrIIA4 did not co-elute with *SpyCas9* in the absence of sgRNA from the size exclusion chromatography, though MBP pull-down assays showed a possible weak interaction between AcrIIA4 and apo-*SpyCas9*<sup>10-12</sup>. We also note that AcrIIA4 did not exhibit measurable affinity to sgRNA alone (Figure S3). The binding of AcrIIA4 and *SpyCas9* was largely an enthalpy-driven reaction, which can be explained by their binding surfaces comprising extensive electrostatic interactions and hydrogen bonding networks<sup>11</sup>. The entropic contribution to the binding was unfavorable, which reflects a decrease in conformational flexibility at the interfaces upon complex formation. Indeed, the increased affinity of AcrIIA4 to sgRNA-bound *SpyCas9* largely originated from favorable entropic contribution, demonstrating that sgRNA binding to *SpyCas9* relieved the entropic cost of *SpyCas9* to accommodate AcrIIA4.

$^1\text{H}$ - $^{15}\text{N}$  HSQC spectra of  $^{15}\text{N}$ -AcrIIA4 titrated with unlabeled *SpyCas9* exhibited gradual line broadening without chemical shift changes. Due to the large molecular weight of *SpyCas9* (m.w.  $\sim 158.4 \text{ kDa}$ ), backbone amide resonances of free AcrIIA4 gradually disappeared without any sign of new signals from the complex, such that 0.2 mM  $^{15}\text{N}$ -AcrIIA4 complexed with *SpyCas9* at a 1:1 ratio resulted in complete line broadening, except for the C-terminal Asn87. Most amide resonances of free AcrIIA4 exhibited intensity reduction of  $>60\%$ , when the complex formation reached 10% from the stoichiometric titration (Fig. 5). This indicates that the line broadening is caused not only by the large complex formation, but also by intermediate exchange from the interaction. It has been reported that the interaction surfaces in slow exchanging large macromolecular complexes can be identified by differential line broadening upon binding<sup>19</sup>. We examined the line broadening of individual residues of AcrIIA4 upon apo-*SpyCas9* binding. The line broadening was less severe at the terminal and internal loop regions, and interfacial residues such as Tyr67 and Asp69 exhibited larger line broadening than others. The differential line broadening, however, did not unambiguously identify the binding interface of AcrIIA4. Comparison of amide resonance intensities of AcrIIA4 between free and 10% bound states indicated that the intensity reduction spanned most backbone amide groups. We examined the line broadening upon titration at 10 °C and 45 °C, but could not distinguish the binding interfaces based on the differential line broadening. We infer that the dynamics of AcrIIA4 and *SpyCas9* in complex may have contributed to the intermediate exchange, obscuring the precise mapping of the binding interface.

The crystal structure of the AcrIIA4–*SpyCas9*–sgRNA complex has shown that *SpyCas9* employs the RuvC domain, the TOPO domain, and the CTD to interact with AcrIIA4<sup>10–12</sup>. The interfaces at the TOPO domain and the CTD (red spheres in Figure S4) overlap with the PAM interaction site of *SpyCas9* for the target strand of the DNA substrate. The interface at the RuvC domain (orange spheres in Figure S4) is the active site pocket that cleaves the non-target strand of the double-stranded DNA substrate. When the binding interfaces in the complex are mapped onto apo–*SpyCas9*, interfaces from the RuvC domain and the CTD are relatively well maintained in apo–*SpyCas9* (dashed circle in red; Figure S4), but those from the TOPO domain are completely lost (dashed circle in black; Figure S4). Based on the structural alignment, we speculate that AcrIIA4 binds to apo–*SpyCas9* in a similar manner to sgRNA-bound *SpyCas9*, and that the RuvC domain and the CTD provide the interaction surfaces for AcrIIA4 in the absence of sgRNA. Binding of *SpyCas9* to sgRNA induces a major rearrangement of the helical REC domain, as well as ordering of the PAM interaction site at the TOPO domain and the CTD<sup>20,21</sup>. It is remarkable that the affinity of the AcrIIA4–*SpyCas9* interaction increases by a thousand-fold upon sgRNA binding<sup>10</sup>. Thus, the guide RNA binding and concomitant conformational rearrangement of *SpyCas9* provide key interfaces not only for the DNA substrate but also for the inhibitor AcrIIA4.

It has been shown that AcrIIA4 competes with the double-strand DNA substrate for *SpyCas9* binding with a ~17-fold higher affinity<sup>10</sup>. AcrIIA4 inhibited the substrate DNA binding to *SpyCas9* in a concentration-dependent manner, but did not affect the dissociation of the DNA substrate from *SpyCas9*, indicating a slow dissociation kinetics<sup>12</sup>. In other words, AcrIIA4 effectively inhibits the nuclease activity of *SpyCas9*, but there is competition between the DNA substrate and AcrIIA4 targeting for the same interface on *SpyCas9*. We demonstrated that AcrIIA4 bound sequentially to *SpyCas9* and then to guide RNA with increasing affinity. In the sequential binding mechanism, AcrIIA4 does not need to compete with target DNA for *SpyCas9* binding, since target DNA recognition requires a pre-formed assembly of guide RNA and *SpyCas9*. The binary complex between AcrIIA4 and apo–*SpyCas9* has a moderate affinity, but it can rapidly turn into a tight ternary complex in the presence of mature guide RNA, effectively preventing the target DNA binding. The sequential binding mechanism would be advantageous for the anti-CRISPR function, especially when the anti-CRISPR genes are integrated as a prophage, and constitutively inactivates the host CRISPR-Cas system<sup>22</sup>.

In summary, we have determined the solution structure of the anti-CRISPR protein AcrIIA4, and investigated the backbone dynamics using NMR spectroscopy. AcrIIA4 adopts a small and compact backbone fold, but exhibits significant dynamics in the loop region that serves as the binding interface for *SpyCas9*. Notably, AcrIIA4 can bind to *SpyCas9* in the absence of sgRNA, and the binary complex can associate with sgRNA to form a tight ternary complex, avoiding competition with the DNA substrate for *SpyCas9* binding.

## Materials and Methods

**Cloning, protein expression, and purification.** The synthetic AcrIIA4 (a.a. 1–87) gene was cloned into a modified pBT7-N-His vector (ATCC). The plasmid was introduced into *Escherichia coli* strain BL21star(DE3) cells (Invitrogen), and grown in LB or minimal media with <sup>15</sup>NH<sub>4</sub>Cl and/or <sup>13</sup>C<sub>6</sub>-glucose as sole nitrogen or carbon sources, respectively. Cells were grown at 37 °C to OD<sub>600</sub> ~0.8, induced with 1 mM isopropyl-β-D-thiogalactopyranoside (IPTG) at 16 °C, and harvested by centrifugation after 16 hours of induction. Cell pellets were resuspended in 20 mM Tris-HCl, pH 7.4, 200 mM NaCl, 20 mM imidazole, and 1 mM phenylmethylsulfonyl-fluoride, lysed using Emulsiflex C3 (AVESTIN) and centrifuged at 20,000 g for 30 min. The clear supernatant was loaded onto a HisTrap column (GE Healthcare) equilibrated with 20 mM Tris-HCl, pH 7.4, 200 mM NaCl, and 20 mM imidazole, and eluted with 500 mM imidazole. The His<sub>6</sub>-tag was removed using TEV protease in 20 mM Tris-HCl, pH 8.0, 5 mM β-mercaptoethanol (BME), and 200 mM NaCl, and the digestion reaction was loaded onto the HisTrap column. The flow-through was collected and loaded onto a Superdex 75 column (GE Healthcare) equilibrated with 20 mM Tris-HCl, pH 7.4, 200 mM NaCl. Lastly, AcrIIA4 was loaded onto a MonoQ column (GE Healthcare) and eluted with a gradient of 0.1–1 M NaCl.

The synthetic *S. pyogenes* Cas9 (a.a. 1–1368) gene was cloned into a pET28b vector with a C-terminal His<sub>6</sub>-tag, and transformed into *E. coli* strain BL21(DE3) cells. The cells were cultured in LB medium at 37 °C to OD<sub>600</sub> ~0.7. Protein expression was induced by the addition of 0.5 mM IPTG, followed by incubation at 17 °C for 16 hours. Cells were harvested by centrifugation and resuspended in 20 mM Tris, pH 7.5, 300 mM NaCl, 5 mM BME, and 10% (w/v) glycerol. After sonication and centrifugation, the supernatant was loaded onto a 5 mL HisTrap HP column (GE Healthcare) pre-equilibrated with 20 mM Tris, pH 7.5, 300 mM NaCl, 5 mM BME, 10% (w/v) glycerol, and 30 mM imidazole. *SpyCas9* was eluted with a linear gradient of 450 mM imidazole and then dialyzed against 20 mM Tris-HCl, pH 7.5, 175 mM NaCl, 10% (w/v) glycerol, and 5 mM BME. The *SpyCas9* fraction was loaded onto a 5 mL HiTrap SP column (GE Healthcare) pre-equilibrated with the dialysis buffer. *SpyCas9* was eluted with a linear gradient of 1 M NaCl. The protein was further purified using a HiLoad 16/60 Superdex200 column (GE Healthcare) equilibrated with 20 mM Tris-HCl pH 8.0, 200 mM NaCl, 5% (w/v) glycerol, and 2 mM dithiothreitol.

**Preparation of sgRNA.** The DNA template for RNA transcription was prepared by using the Giga-prep kit (Epigenetics). The sgRNA was prepared *in vitro* by mixing rNTPs (rATP, rGTP, rCTP, and cUTP), MgCl<sub>2</sub>, T7 RNA polymerase (P266L mutant)<sup>23</sup>, inorganic pyrophosphatase, and the DNA template in the transcription buffer. After 6 hour transcription at 37 °C, the synthesized RNA was precipitated by ethanol treatment overnight, purified using 12% denaturing PAGE (19:1 cross-linking ratio), and electro-eluted (Elutrap, Whatman). The purified RNA was desalted and exchanged into water using Amicon (Millipore).

**NMR Spectroscopy.** The NMR sample contained 2 mM <sup>13</sup>C,<sup>15</sup>N-AcrIIA4 in 20 mM sodium phosphate, pH 7.4, 100 mM NaCl, 0.01% NaN<sub>3</sub>. NMR spectra were recorded at 25 °C on Bruker 600 MHz and 800 MHz spectrometers equipped with a z-shielded gradient triple resonance probe. Sequential and side chain assignments of <sup>1</sup>H, <sup>15</sup>N and <sup>13</sup>C resonances were achieved by three-dimensional triple resonance through-bond scalar

correlation experiments CBCA(CO)NH, HNCACB, HNCO, HN(CA)CO, HBHA(CO)NH, HCCH-TOCSY, and  $^{15}\text{N}$ -TOCSY-HSQC. 3D  $^{13}\text{C}$ -separated NOESY and  $^{15}\text{N}$ -separated NOESY experiments were obtained using a mixing time of 120 ms. The acquisition parameters for the 3D experiments are described in detail as Supplementary Information. The  $\chi^1$  angle restraints were obtained based on the quantification of intraresidue NOEs between backbone amide proton,  $\alpha$  proton, and  $\beta$  methylene protons<sup>24</sup>. Residual  $^1\text{D}_{\text{NH}}$  dipolar couplings were obtained by taking the difference in the J splitting values measured in oriented (10 mg/ml of *pf1* phage alignment media) and isotropic (water) AcrIIA4 using 2D in-phase/antiphase  $^1\text{H}$ - $^{15}\text{N}$  HSQC spectra<sup>25</sup>.  $^{15}\text{N}$ - $R_1$  and  $^{15}\text{N}$ - $R_{1\rho}$  relaxation rates, and heteronuclear  $^1\text{H}$ - $^{15}\text{N}$  NOE were measured using pulse schemes described previously<sup>26,27</sup> on the Bruker 800 MHz spectrometer. Delays of 10, 50, 100, 400, 800, 1200, 1500 ms were used for the  $R_1$  relaxation measurement, and 0, 4, 12, 24, 40, 60, 80, 96 ms used for  $R_{1\rho}$  experiment.  $^{15}\text{N}$ - $R_2$  relaxation rates were obtained from the  $^{15}\text{N}$ - $R_1$  and  $^{15}\text{N}$ - $R_{1\rho}$  relaxation rates using the equation,  $R_2 = (R_{1\rho} - R_1 \sin^2 \theta) / \cos^2 \theta$ , where  $\theta = \arctan(\Omega_N / \gamma_{\text{N}} B_1)$ ,  $\Omega_N$  is the resonance offset, and  $\gamma_{\text{N}} B_1$  is the spin-lock field strength<sup>28</sup>. For the heteronuclear  $^1\text{H}$ - $^{15}\text{N}$  NOE measurement,  $^1\text{H}$  saturation was achieved using 3 s of  $120^\circ$   $^1\text{H}$  pulses separated by 5 ms intervals. 2 s of relaxation delays were used for  $^{15}\text{N}$ - $R_1$  and  $^{15}\text{N}$ - $R_{1\rho}$  relaxation measurements, and 4 s of the relaxation delay was used for the heteronuclear  $^1\text{H}$ - $^{15}\text{N}$  NOE measurement. NMR spectra were processed using the NMRPipe<sup>29</sup>, and analyzed using PIPP<sup>30</sup>, and NMRView<sup>31</sup> programs.

**Structure calculation.** Interproton distance restraints were derived from the NOE spectra and classified into distance ranges according to the peak intensity.  $\phi/\psi$  torsion angle restraints were derived from backbone chemical shifts using the program TALOS+<sup>32</sup>. The solution structure of AcrIIA4 was calculated by simulated annealing in torsion angle space using the Xplor-NIH program<sup>13</sup>. The target function of simulated annealing included a covalent geometry, a quadratic van der Waals repulsion potential<sup>33</sup>, square-well potentials for interproton distance and torsion angle restraints, hydrogen bonding, RDC restraints<sup>34</sup>, harmonic potentials for  $^{13}\text{C}\alpha^{13}\text{C}\beta$  chemical shift restraints<sup>35</sup>, a multidimensional torsion angle database potential of mean force<sup>36</sup>, and a radius of gyration term<sup>37</sup>. The radius of gyration represents a weak overall packing potential, and residues 1–86 of AcrIIA4 were selected for the term. Structures were displayed using VMD-XPLOR software<sup>38</sup>. The violations from the structure calculation are described in detail as Supplementary Information.

**Isothermal Titration Calorimetry.** The ITC experiment was performed at  $25^\circ\text{C}$  using an iTC200 calorimeter (Malvern). To characterize the affinity of AcrIIA4 to SpyCas9,  $50\ \mu\text{M}$  SpyCas9 in the cell was titrated with  $500\ \mu\text{M}$  of the AcrIIA4 in  $20\ \text{mM}$  sodium phosphate,  $\text{pH } 7.4$ ,  $100\ \text{mM}$  NaCl,  $5\ \text{mM}$  BME, and  $0.01\%$   $\text{NaN}_3$ . Twenty consecutive  $2\ \mu\text{L}$  aliquots of AcrIIA4 were titrated into the cell and the duration of each injection was 4 sec, with injections occurring at intervals of 150 sec. The heats associated with dilution of the substrates were subtracted from the measured heats of binding. ITC titration data were analyzed using the Origin version 7.0 program provided with the instrument.

**Data availability.** Atomic coordinates and NMR restraints of the reported solution structure have been deposited in the Protein Data Bank under accession code 5XN4 and in the Biological Magnetic Resonance Bank under accession code 36085, respectively.

## References

- Barrangou, R. *et al.* CRISPR provides acquired resistance against viruses in prokaryotes. *Science* **315**, 1709–1712 (2007).
- Brouns, S. J. *et al.* Small CRISPR RNAs guide antiviral defense in prokaryotes. *Science* **321**, 960–964 (2008).
- Koonin, E. V., Makarava, K. S. & Zhang, F. Diversity, classification and evolution of CRISPR-Cas systems. *Curr. Opin. Microbiol.* **37**, 67–78 (2017).
- Jiang, W. & Marraffini, L. A. CRISPR-Cas: new tools for genetic manipulations from bacterial immunity systems. *Annu. Rev. Microbiol.* **69**, 209–228 (2015).
- Sternberg, S. H. & Doudna, J. A. Expanding the biologist's toolkit with CRISPR-Cas9. *Mol. Cell* **568**, 568–574 (2015).
- Wright, A. V., Nunez, J. K. & Doudna, J. A. Biology and applications of CRISPR systems: harnessing nature's toolbox for genome engineering. *Cell* **164**, 29–44 (2016).
- Samson, J. E., Magadán, A. H., Sabri, M. & Moineau, S. Revenge of the phages: defeating bacterial defences. *Nat. Rev. Microbiol.* **11**, 675–687 (2013).
- Pawluk, A. *et al.* Naturally occurring off-switches for CRISPR-Cas9. *Cell* **167**, 1829–1838 (2016).
- Rauch, B. J. *et al.* Inhibition of CRISPR-Cas9 with bacteriophage proteins. *Cell* **168**, 150–158 (2017).
- Dong, D. *et al.* Structural basis of CRISPR-SpyCas9 inhibition by an anti-CRISPR protein. *Nature* **546**, 436–439 (2017).
- Yang, H. & Patel, D. J. Inhibition mechanism of an anti-CRISPR suppressor AcrIIA4 targeting SpyCas9. *Mol. Cell* **67**, 117–127 (2017).
- Shin, J. *et al.* Disabling Cas9 by an anti-CRISPR DNA mimic. *Sci. Adv.* **3**, e1701620 (2017).
- Schwieters, C. D., Kuszewski, J. & Clore, G. M. Using Xplor-NIH for NMR molecular structure determination. *Prog. Nucl. Magn. Reson. Spectrosc.* **48**, 47–62 (2006).
- Chowdhury, S. *et al.* Structure reveals mechanisms of viral suppressors that intercept a CRISPR RNA-guided surveillance complex. *Cell* **169**, 47–57 (2017).
- Kalodimos, C. G. *et al.* Structure and flexibility adaptation in nonspecific and specific protein-DNA complexes. *Science* **305**, 386–389 (2004).
- Lee, G. M. *et al.* The structural and dynamic basis of Ets-1 DNA binding autoinhibition. *J. Biol. Chem.* **280**, 7088–7099 (2005).
- Bhabha, G. *et al.* A dynamic knockout reveals that conformational fluctuations influence the chemical step of enzyme catalysis. *Science* **332**, 234–238 (2011).
- Kovermann, M., Rogne, P. & Wolf-Watz, M. Protein dynamics and function from solution state NMR spectroscopy. *Q. Rev. Biophys.* **49**, e6 (2016).
- Matsuo, H. *et al.* Identification by NMR spectroscopy of residues at contact surfaces in large, slowly exchanging macromolecular complexes. *J. Am. Chem. Soc.* **121**, 9903–9904 (1999).
- Jiang, F., Zhou, K., Ma, L., Gressel, S. & Doudna, J. A. A Cas9-guide RNA complex preorganized for target DNA recognition. *Science* **348**, 1477–1482 (2015).
- Jinek, M. *et al.* Structures of Cas9 endonucleases reveal RNA-mediated conformational activation. *Science* **343**, 1247997 (2014).



22. 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).
23. Guillerez, J., Lopez, P. J., Proux, F., Launay, H. & Dreyfus, M. A mutation in T7 RNA polymerase that facilitates promoter clearance. *Proc. Natl. Acad. Sci. USA* **102**, 5958–5963 (2005).
24. Wagner, G. *et al.* Protein structures in solution by nuclear magnetic resonance and distance geometry. The polypeptide fold of the basic pancreatic trypsin inhibitor determined using two different algorithms, DISGEO and DISMAN. *J. Mol. Biol.* **196**, 611–639 (1987).
25. Bax, A., Kontaxis, G. & Tjandra, N. Dipolar couplings in macromolecular structure determination. *Methods. Enzymol.* **339**, 127–174 (2001).
26. Farrow, N. A., Zhang, O. W., Forman-Kay, J. D. & Kay, L. E. A heteronuclear correlation experiments for simultaneous determination of  $^{15}\text{N}$  longitudinal decay and chemical exchange rates of systems in slow equilibrium. *J. Biomol. NMR* **4**, 727–734 (1994).
27. Tjandra, N., Wingfield, P., Stahl, S. & Bax, A. Anisotropic rotational diffusion of perdeuterated HIV protease from  $^{15}\text{N}$  NMR relaxation measurements at two magnetic fields. *J. Biomol. NMR* **8**, 273–284 (1996).
28. Davis, D. G., Perlman, M. E. & London, R. E. Direct measurements of the dissociation-rate constant for inhibitor-enzyme complexes via the T1 rho and T2 (CPMG) methods. *J. Magn. Reson. Ser. B.* **104**, 266–275 (1994).
29. Delaglio, F. *et al.* NMRPipe: A multidimensional spectral processing system based on UNIX pipes. *J. Biomol. NMR* **6**, 277–293 (1995).
30. Garrett, D. S., Powers, R., Gronenborn, A. M. & Clore, G. M. A common sense approach to peak picking in two-, three-, and four-dimensional spectra using automatic computer analysis of contour diagrams. *J. Magn. Reson.* **95**, 214–220 (1991).
31. Johnson, B. A. & Blevins, R. A. NMRView: A computer program for the visualization and analysis of NMR data. *J. Biomol. NMR* **4**, 603–614 (1994).
32. Shen, Y., Delaglio, F., Cornilescu, G. & Bax, A. TALOS+: A hybrid method for predicting protein backbone torsion angles from NMR chemical shifts. *J. Biomol. NMR* **44**, 213–223 (2009).
33. Nilges, M., Gronenborn, A. M., Brünger, A. T. & Clore, G. M. Determination of three-dimensional structures of proteins by simulated annealing with interproton distance restraints. Application to crambin, potato carboxypeptidase inhibitor and barley serine proteinase inhibitor 2. *Protein Eng.* **2**, 27–38 (1988).
34. Clore, G. M., Gronenborn, A. M. & Tjandra, N. Direct refinement against residual dipolar couplings in the presence of rhombicity of unknown magnitude. *J. Magn. Reson.* **131**, 159–162 (1998).
35. Kuszewski, J., Gronenborn, A. M. & Clore, G. M. The impact of direct refinement against  $^{13}\text{C}\alpha$  and  $^{13}\text{C}\beta$  chemical shifts on protein structure determination by NMR. *J. Magn. Reson. Ser. B.* **106**, 92–96 (1995).
36. Clore, G. M. *et al.* The three-dimensional structure of  $\alpha$ 1-purothionin in solution: combined use of nuclear magnetic resonance, distance geometry and restrained molecular dynamics. *EMBO J.* **5**, 2729–2735 (1986).
37. Kuszewski, J., Gronenborn, A. M. & Clore, G. M. Improving the packing and accuracy of NMR structures with a pseudopotential for the radius of gyration. *J. Am. Chem. Soc.* **121**, 2337–2338 (1999).
38. Schwieters, C. D. & Clore, G. M. The VMD-XPLOR visualization package for NMR structure refinement. *J. Magn. Reson.* **149**, 239–244 (2001).
39. Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Cryst.* **26**, 283–291 (1993).

## Acknowledgements

This work was supported by the Cooperative Research Program for Agriculture Science & Technology Development, Rural Development Administration (PJ011112), and the Creative-Pioneering Researchers Program through Seoul National University (500-20170161). We thank the Research Institute of Agriculture and Life Sciences, the high-field NMR facility at the Korea Institute of Science and Technology, the Korea Basic Science Institute, and the National Center for Inter-University Research Facilities. We thank Ms. So Young An and Ms. Ji Yeon Shin for the help with sgRNA preparation.

## Author Contributions

I.K. and J.Y.S. designed the study; I.K., D.K., M.H., N.K.K., and J.Y.S. performed the experiments; I.K., M.J., E.B., and J.Y.S. analyzed the data; I.K., M.J., E.B., and J.Y.S. wrote the manuscript.

## Additional Information

**Supplementary information** accompanies this paper at <https://doi.org/10.1038/s41598-018-22177-0>.

**Competing Interests:** The authors declare no competing interests.

**Publisher's note:** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2018