

A viral ring nuclease anti-CRISPR subverts type III CRISPR immunity

Januka S Athukoralage¹, Stephen McMahon¹, Changyi Zhang^{3,4}, Sabine Grüschorw¹, Shirley Graham¹, Mart Krupovic², Rachel J Whitaker^{3,4}, Tracey Gloster^{1*} and Malcolm F White^{1*}

¹ Biomedical Sciences Research Complex, School of Biology, University of St Andrews, North Haugh, St Andrews, Fife KY16 9ST, UK. ² Institut Pasteur, Department of Microbiology, 25 rue du Dr Roux, 75015 Paris, FRANCE. ³ Department of Microbiology, University of Illinois at Urbana-Champaign, Urbana, IL 61801. ⁴ Carl R. Woese Institute for Genomic Biology, University of Illinois at Urbana-Champaign, Urbana, IL 61801.

*To whom correspondence should be addressed: email mfw2@st-andrews.ac.uk or tmg@st-andrews.ac.uk

ABSTRACT

The CRISPR system provides adaptive immunity against mobile genetic elements in bacteria and archaea. On detection of viral RNA, type III CRISPR systems generate a cyclic oligoadenylate (cOA) second messenger¹⁻³, activating defence enzymes and sculpting a powerful antiviral response that can drive viruses to extinction^{4,5}. Cyclic nucleotides are increasingly implicated as playing an important role in host-pathogen interactions^{6,7}. Here, we identify a widespread new family of viral anti-CRISPR (Acr) enzymes that rapidly degrade cyclic tetra-adenylate (cA₄). The viral ring nuclease (AcrIII-1) is the first Acr described for type III CRISPR systems and is widely distributed in archaeal and bacterial viruses, and proviruses. The enzyme uses a novel fold to bind cA₄ specifically and utilizes a conserved active site to rapidly cleave the signalling molecule, allowing viruses to neutralise the type III CRISPR defence system. The AcrIII-1 family has a broad host range as it targets cA₄ signalling molecules rather than specific CRISPR effector proteins. This study highlights the crucial role of cyclic nucleotide signalling in the conflict between viruses and their hosts.

Type III CRISPR-Cas systems synthesise the signalling molecule cyclic oligoadenylate (cOA) from ATP^{1,2} when they detect viral RNA. cOA molecules are synthesised with a range of ring sizes with 3-6 AMP subunits (denoted cA₃, cA₄ etc.) by the cyclase domain of the Cas10 protein^{1-3,9,10}. cOA binds to a specific protein domain, known as a CARF (CRISPR Associated Rossman Fold) domain. CARF domains are found fused to a variety of effector domains that are known or predicted to cleave RNA, DNA, or function as transcription factors¹¹. The best characterised CARF protein family is the Csx1/Csm6 family of HEPN (Higher Eukaryotes and Prokaryotes, Nucleotide binding) ribonucleases, which are activated by cOA binding and cleave RNA with minimal sequence dependence¹⁻³. A number of studies have demonstrated that the cOA signalling component of type III systems is crucial for effective immunity against viruses^{4,12-15}, highlighting the importance of this facet of CRISPR immunity.

Recently, we identified a cellular enzyme in *Sulfolobus solfataricus*, hereafter referred to as the Crn1 family (for “CRISPR associated ring nuclease 1”), that degrades cA₄ molecules and thus deactivates the Csx1 ribonuclease *in vitro*¹⁶. These enzymes exhibit very slow kinetics, and are thought to act by mopping up cA₄ molecules in the cell without compromising the immunity provided by the type III CRISPR system. Unsurprisingly, viruses have responded to the threat of the CRISPR system by evolving a range of anti-CRISPR (Acr) proteins, which are used to inhibit and overcome the cell’s CRISPR defences (reviewed in¹⁷). Acr’s have been identified for the type I-D¹⁸, I-F, II-A and V-A effector complexes (reviewed in^{17,19,20}), numbering over 40 families²¹, but importantly not for type III systems. We focussed on one of the protein families, DUF1874, conserved and widespread in a variety of archaeal viruses and plasmids but also bacteriophages and prophages (Extended data figure 1) and has no known function. Structures are available for several family members, including gp29 of *Sulfolobus islandicus* rod-shaped virus 1 (SIRV1)²² and B116 of *Sulfolobus* turreted icosahedral virus (STIV)²³, which is expressed early in the STIV infection cycle²⁴ and is important for normal virus replication kinetics²⁵.

Given the widespread distribution and important role in viral infection, we explored the possibility of an Acr function for this viral protein family. To investigate this, we deleted the genes for the type I-A system in *Sulfolobus islandicus* M.16.4 carrying *pyrEF* and *argD* deletions²⁶ so that it had only a type III-B CRISPR system for defence²⁷. We challenged this strain with the archaeal virus SSeV (Figure. 1a), a lytic virus isolated from Kamchatka Russia with an exact CRISPR-spacer match of 100% in M.16.4 and several other potentially active CRISPR-spacers (Rowland et al. *in prep*). SSeV lacks a *duf1874* gene and fails to form plaques on a lawn of *S. islandicus* M.16.4 with type III-B CRISPR defence (Figure 1a). However, the same cells expressing the SIRV1 gp29 gene from a plasmid are readily infected, giving rise to plaque formation. These data suggest that SIRV1 gp29 is functioning as an anti-CRISPR specific for type III CRISPR defence.

To confirm the specificity of the anti-CRISPR function, we utilised a recently developed recombinant type III CRISPR system from *Mycobacterium tuberculosis*, which we have established in *E. coli*. The system allows the cOA effector protein to be swapped to provide effective immunity against mobile genetic elements based on either cA₆ or cA₄ signalling¹⁰. We set up an assay where cells capable of cA₄ or cA₆-based immunity were transformed with a plasmid that could be targeted for interference due to a match in the tetracycline resistance gene in the plasmid to a spacer in the CRISPR array (Figure 1b). Efficient interference was observed for either strain in the absence of the *duf1874* gene from bacteriophage THSA-485A (Figure 1c). However, the presence of the *duf1874* gene on the plasmid abrogated immunity for cA₄, but not for cA₆-mediated CRISPR defence. Thus, DUF1874 acts as an anti-CRISPR against cA₄-mediated type III CRISPR defence, explaining the previous observations with archaeal viruses^{24,25}.

Hereafter we propose the collective name AcrIII-1 for this family, as it is the first Acr described for type III systems and is not specific for an individual subtype of the type III effectors²⁸.

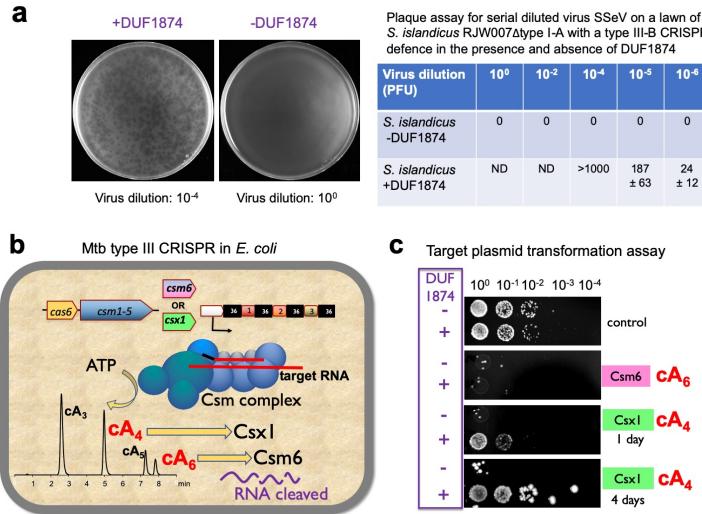


Figure 1. DUF1874 is an anti-CRISPR specific for cA₄ signalling

(a) SSeV infection assay showing the DUF1874 gene SIRV1 gp29 can neutralise the type III-B system in *S. islandicus*. Replicative plasmids with (pOE-gp29) or without (pOE) gp29 were transformed into the *S. islandicus* RJW007Δtype I-A mutant, and the resulting strains were then challenged with SSeV. Plaques are only observed on the lawn of the strain expressing gp29. Data are representative of three biological replicates. **(b)** Schematic showing the recombinant *M. tuberculosis* type III-A CRISPR interference system established in *E. coli*. By swapping the native Csm6 ancillary nuclease for a Csxl protein, the system can be converted from cA₆ to cA₄-mediated immunity. **(c)** Plasmid transformation assay using a plasmid with a match to a spacer in the CRISPR array. If the plasmid is successfully targeted by the CRISPR system, transformants will not be seen. Plasmids with or without the *duf1874* gene are targeted successfully when cA₆ (Csm6) mediated antiviral signalling is active. In contrast, cells using a cA₄-based (Csxl) system only prevent transformation when the DUF1874 protein is not present, showing that the protein is effective in neutralising cA₄-based CRISPR interference. The controls lack cOA-dependent ribonucleases.

To explore the mechanism of the AcrIII-1 family, we cloned and expressed two family members in *E. coli*: the SIRV1 gp29 protein and the YddF protein encoded by an integrative and conjugative element ICEBs1 from *B. subtilis*²⁹. Both proteins possess a potent ring nuclease activity, rapidly degrading cA₄ to generate linear di-adenylate (ApA>P) with a cyclic 2',3' phosphate (Figure 2 and extended data figure 3). With a catalytic rate exceeding 5 min⁻¹, the anti-CRISPR enzyme is at least 60-fold more active than the cellular ring nuclease Crn1 from *S. solfataricus*. We showed previously that the type III-D CRISPR effector of *S. solfataricus* generates cA₄ in proportion to the amount of cognate target RNA present¹⁶. By varying target RNA input and following cA₄ levels and Csxl activity, we compared the abilities of Crn1 and AcrIII-1 to destroy the signalling molecule and deactivate the ancillary defence nuclease Csxl. In keeping with its low turnover number, the Crn1 enzyme was effective at degrading cA₄ and thus deactivating Csxl only at the lowest levels of target RNA (Figure 3a). In contrast, AcrIII-1 degraded cA₄ completely at the highest target RNA concentration examined, preventing Csxl activation. To provide a more rigorous test, we examined the ability of both enzymes to prevent Csxl activation over a range of cA₄ concentrations spanning four orders of magnitude (Figure 3b). Crn1 (2 μM) provided protection only up to 5 μM cA₄, but in contrast 2 μM AcrIII-1 provided complete protection at the highest level of cA₄ tested (500 μM). Thus, AcrIII-1 has the potential to destroy large concentrations of the cA₄ second messenger rapidly, negating the immune response.

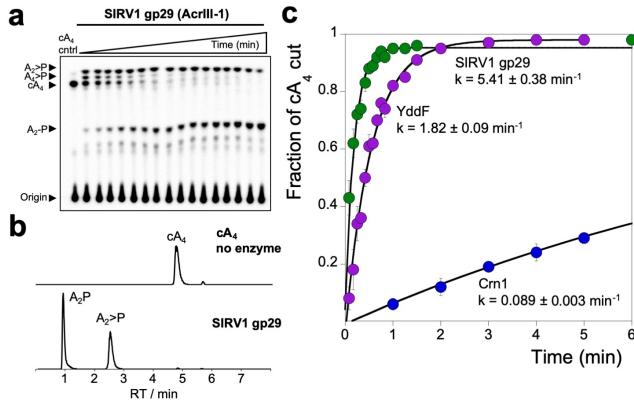


Figure 2. AcrIII-1 rapidly degrades cA₄ to linear products

(a) Single-turnover kinetic analysis of cyclic tetra-adenylate (cA₄, 200 nM) degradation by AcrIII-1 SIRV1 gp29 (4 μM dimer, 50 °C), showing rapid generation of A₂>P (di-adenylate with a 2',3'-cyclic phosphate) via an A₄>P intermediate visualised by phosphorimaging following thin-layer chromatography. A₂>P is further hydrolysed to A₂-P (di-adenylate with a 3' phosphate). Each lane corresponds to reaction quenched at 5 second intervals up to 1 min, and then at 1.25, 1.5, 3, 6 and 12 min. Control reaction (cA₄ cntrl) is cA₄ incubated in the absence of protein for 12 min at 50 °C. **(b)** Liquid chromatography-high resolution mass spectrometry analysis confirms that AcrIII-1 SIRV1 gp29 converts cA₄ (top panel) to A₂>P within 2 min which is then converted to A₂-P (bottom panel). **(c)** Kinetic comparison of cA₄ degradation by the AcrIII-1 enzymes SIRV1 gp29 (4 μM dimer, 50 °C) and YddF (8 μM dimer, 37 °C) with the Crn1 enzyme Sso2081 (4 μM dimer, 50°C). Data were quantified and plotted to show the fraction of cA₄ cut over time and fitted to an exponential equation, as previously described⁸. For YddF and Crn1 the data shown is the average of three independent experiments and for SIRV1 gp29 the data is the average of two biological replicates comprising three independent replicates for each. The error indicated is the standard deviation of the mean.

Several apo-structures for AcrIII-1 family members are available, revealing a dimeric protein of novel fold^{22,23,30}. Importantly, this structure is completely unrelated to the CRISPR associated Rossman fold (CARF) domain, which is the only protein family thus far known to bind cOA¹¹. To elucidate the mechanism of cA₄ binding and cleavage by AcrIII-1, we co-crystallised an inactive variant (H47A) of SIRV1 gp29 with cA₄ and solved the structure to 1.55 Å resolution (Figure 4). The complex reveals a molecule of cA₄ bound at the dimer interface. Comparison of the cA₄-bound and apo structures reveals a significant movement of a loop, comprising residues 82-94, and subsequent α-helix, to bury cA₄ within the dimer. These loops adopt variable or unstructured conformations in the various apo protein structures^{22,23,30}. Once bound, the ligand is completely enclosed by the protein – a considerable accomplishment when one considers the large size of the ligand and the small size of the protein (Figure 4b). By

superimposing the cA₄ ligand on the apo-protein structure, it becomes apparent that the binding site is largely pre-formed, with the exception of the mobile loops that form the lid (Figure 4c). The overall change is like two cupped hands catching a ball, with the loops (fingers) subsequently closing around it.

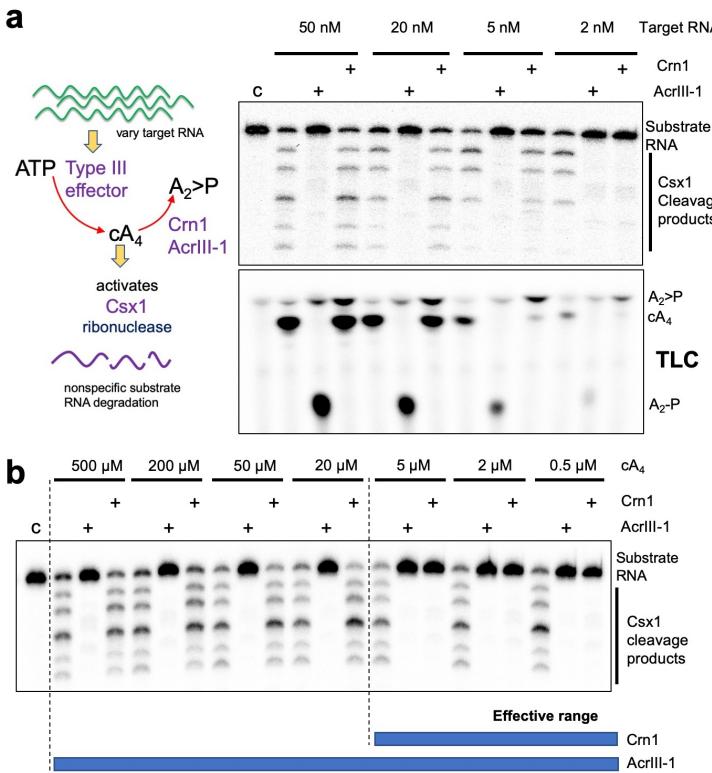


Figure 3. AcrIII-1 neutralises cA₄-activated CRISPR defence enzyme Csx1

(a) Top panel is a phosphor image of denaturing PAGE visualising activation of Csx1 (0.5 μM dimer) in a coupled assay containing type III Csm complex when activated with indicated amounts of (unlabelled) target RNA to initiate cA₄ synthesis. Each set of three lanes after the control (c) reaction with Csx1 and substrate RNA alone, is first in the absence and then with AcrIII-1 SIRV1 gp29 (2 μM dimer) or Crn1 Sso2081 (2 μM dimer), respectively. Whereas AcrIII-1 degraded all cA₄ generated with up to 50 nM RNA target, the Crn1 enzyme deactivated Csx1 only when less than 5 nM RNA target was used to initiate cA₄ synthesis. The lower panel is a phosphorimage of TLC with reactions as above but visualises cA₄ production, by α -ATP incorporation, and degradation, in the presence of indicated amounts of RNA target and absence or presence of either AcrIII-1 or Crn1, and is representative of the results of three independent experiments. Csx1 deactivation correlates with complete cA₄ degradation. (b) Denaturing PAGE showing activation of Csx1 (0.5 μM dimer) by indicated amounts (500-0.5 μM) of HPLC-purified cA₄ and its subsequent deactivation when either AcrIII-1 or Crn1 (2 μM dimer) was present to degrade cA₄. The AcrIII-1 enzyme degraded 100-fold more cA₄ than Crn1. Control reaction (c) shows RNA incubated with Csx1 in the absence of cA₄. The image is representative of the results of three independent experiments.

The cA₄ molecule makes symmetrical interactions with each monomer of AcrIII-1 (Extended data figure 4). Upon binding of cA₄, arginine R85 on the loop from one monomer interacts with the distant half of the cA₄ molecule and appears to 'lock' the closed dimer. Other important interactions are made with main chain L92, I69, and N8, and side chains R66, N8, Q81, S11, T50, S49, and N13, most of which are semi or fully conserved (Extended data figure 1 & 5), suggesting they have important roles in cA₄ binding and/or catalysis in this whole family of enzymes.

At two positions, on opposite sides of the ring, the 2'-hydroxyl of the ribose is positioned correctly for in-line attack on the phosphodiester bond, consistent with the observed bilateral cleavage (Figure 4d). The catalytic power of the AcrIII-1 family likely derives from active site residues that position the 2'-hydroxyl group for in-line nucleophilic attack, stabilise the transition state and protonate the oxyanion leaving group ³¹. For the AcrIII-1 family, the absolutely conserved residue His-47 is suitably positioned to act as a general acid and fulfil the latter role (Figure 4d). To test this hypothesis, we assayed variant H47A of AcrIII-1. The variant enzyme suffered a >2500-fold decrease in catalytic power, which could be partially reversed by chemical rescue with 500 mM imidazole in the reaction buffer (Extended data figure 6).

By targeting a key signalling molecule, a single AcrIII-1 enzyme should have broad utility in the inhibition of endogenous type III CRISPR systems in any species. The only constraint is the specificity for the cA₄ signalling molecule. Of the CRISPR ancillary nucleases studied to date, most are activated by cA₄; activation by cA₆ appears to be limited to certain bacterial phyla including the Firmicutes and Actinobacteria ¹⁰. By targeting the cA₄ activator, AcrIII-1 will neutralize a slew of defence enzymes simultaneously. For example, *Thermus thermophilus* has two Csm6 enzymes ^{32,33} and a recently described DNA nuclease, Can1 ³⁴. All are activated by cA₄ and would thus be deactivated by AcrIII-1. Recently, two other Acr proteins with enzymatic functions have been described: AcrVA1 which catalyses crRNA cleavage of Cas12a ³⁵ and AcrVA5, which acetylates the PAM-sensing site of Cas12a ³⁶. Both, however, target a protein (or protein:nucleic acid complex), implying a requirement for specific interactions that could be evaded by sequence variation.

The gene encoding AcrIII-1 is most prominent in the archaeal viruses, where it is found in representatives of at least five distinct viral families, making it one of the most widely conserved of all archaeal virus proteins ³⁰ (Extended data figure 1). Within archaeal genomes, homologues are typically adjacent to ORFs from mobile genetic elements rather than CRISPR loci, a good example being the STIV virus integrated into *S. acidocaldarius* genomes ³⁷. The distribution of AcrIII-1 in archaea is sporadic but covers most of the lineages, including crenarchaea, euryarchaea, thaumarchaea, Asgard archaea and others. AcrIII-1 is also present in several bacteriophages of the order *Caudovirales*, and there are many instances of *acrIII-1* genes in sequenced bacterial genomes, with homologues found in the Firmicutes, cyanobacteria, proteobacteria, actinobacteria and many more. Maximum likelihood phylogenetic analysis of the AcrIII-1 proteins suggests multiple horizontal gene transfers between unrelated viruses as well as between bacteria and archaea (Extended data figure 7). Sometimes the *acrIII-1* gene is clearly part of an integrated mobile genetic element as is the case for the *yddF* gene in *B. subtilis* ³⁰. However, in other species (n=49) the gene is associated with cellular type

III CRISPR systems (Extended data figures 7-9). In *Marinitoga piezophilia*, AcrIII-1 is fused to a cOA-activated HEPN ribonuclease of the Csx1 family. Since both active sites are conserved, this fusion protein may have cA₄ activated ribonuclease activity coupled with a cA₄ degradative ring nuclease, thus providing an explicit linkage between the AcrIII-1 family and the type III CRISPR system. In this context the enzyme is likely acting as a host-encoded ring nuclease, like Crn1, rather than an Acr. We therefore propose the family name of Crn2 (CRISPR associated ring nuclease 2) to cover DUF1874 family members that are associated with type III CRISPR systems.

AcrIII-1 is the first Acr have functional roles in both viruses and cells³⁸. It remains to be determined whether the *acrIII-1* gene arose in viruses and was appropriated by cellular type III systems or *vice versa*. However, the extremely broad distribution of *acrIII-1* and limited distribution of *crn2* suggests the former. Adoption of an anti-CRISPR protein as a component of a cellular CRISPR defence system seems counter-intuitive. However, the enzyme could be harnessed for a role in defence by either “detuning” its activity to make it a slower enzyme, or by putting it under tight transcriptional control so that it is expressed at very low levels. The unprecedented wide occurrence of this Acr across many archaeal and bacterial virus families reflects the fact that this enzyme degrades a key signalling molecule to subvert cellular immunity. This makes it very hard for cells to evolve counter-resistance, other than by switching to a different signalling molecule. The recent discovery of multiple cyclic nucleotide signalling systems in bacteria may thus reflect the ongoing conflict between cells and viruses.

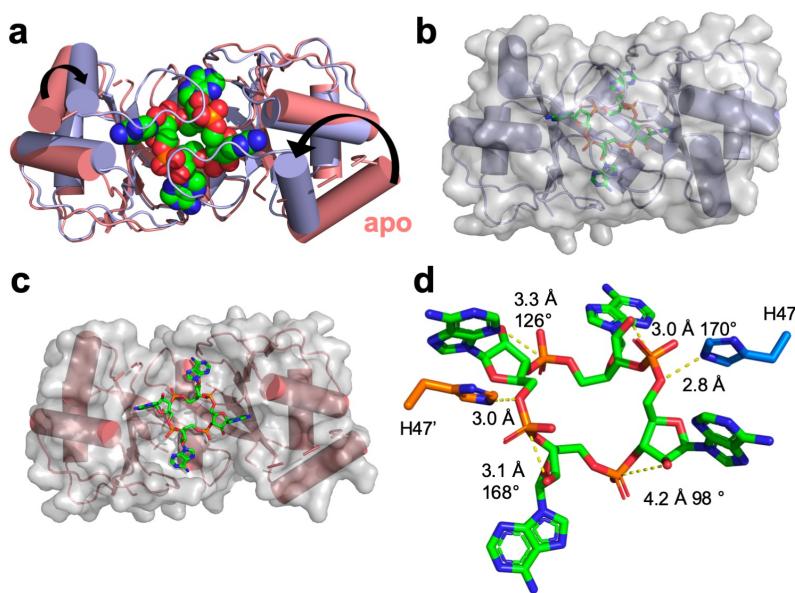


Figure 4. Structure of AcrIII-1 bound to cA₄

(a) Superimposition of the apo SIRV1 gp29 structure (salmon) and in complex with cA₄ (blue), highlighting the movement of the loop and α -helix upon cA₄ binding. cA₄ is shown in green spheres. (b) Surface representation of the structure of SIRV1 gp29 (blue) in complex with cA₄ (green), emphasising the complete burial of the ligand. (c) Surface representation of the apo structure of SIRV1 gp29 (salmon) with cA₄ (green) in the position observed in the complex structure, indicating that the binding site is pre-formed. (d) Structure of cA₄ bound to SIRV1 gp29. The two active site histidine residues (modelled based on the position of the alanine side chain in the H47A variant crystallised with cA₄; coloured to represent residues from different monomers) are in suitable positions to act as the general acid, protonating the oxyanion leaving group. The corresponding ribose sugars have 2'-hydroxyl groups suitably positioned for in-line nucleophilic attack on the phosphodiester bond.

Acknowledgements This work was supported by grants from the Biotechnology and Biological Sciences Research Council (REF: BB/S000313/1 to MFW and REF: BB/R008035/1 to TMG) and by a NASA Exobiology and Evolutionary Biology grant (NNX14AK23G to RJW). We thank Jesse Black and Maria Alejandra-Bautista for isolating and characterizing the SSeV virus, and thank Rebecca Wipfler and Wenlong Zhu for technical assistance.

Author contributions J.S.A. designed experiments and carried out the enzyme assays and analysis; S.M. carried out the structural biology; C.Z. constructed the *S. islandicus* strains and performed the virus infection assays; S.Grü. carried out the plasmid transformation assays and mass spectrometry; S.Gra. generated expression plasmids and purified proteins; M.R. contributed to the conception of the project; T.M.G., R.J.W. and M.F.W. oversaw the work, analysed the data and wrote the manuscript. All authors contributed to data analysis and writing.

Competing interests The University of St Andrews has filed a patent application related to this work for which J.S.A. and M.F.W. are inventors. The other authors declare no competing interests.

Correspondence should be addressed to M.F.W. or T.M.G.

METHODS

Construction of *S. islandicus* strains

The type I-A CRISPR defence module including seven genes i.e. *cas3b*, *csa5*, *cas7*, *cas5*, *cas3'*, *cas3'*, and *casX*³⁹ was in-frame deleted from the genetic host *S. islandicus* RJW007, derived from wild-type strain *S. islandicus* M.16.4 carrying a double *pyrEF* and *argD* deletion²⁶, by employing Plasmid Integration and Segregation knockout strategy⁴⁰. The type I-A deletion mutants were confirmed by PCR using primers that bind outside of the homologous flanking arms.

Synthesised SIRV1 *gp29* gene was purchased from IDT as a g-block and was cloned into a *Sulfolobus-E. coli* shuttle vector pSeSd-SsoargD²⁶ (referred to as pOE hereafter) at the *NdeI* and *NotI* sites, generating the *gp29* expression plasmid pOE-gp29 in which the *gp29* gene was placed under the control of the arabinose promoter. The pOE-gp29 and pOE plasmids were then transformed into competent cells of the Δ type I-A mutant via electroporation as described previously²⁶, generating strains expressing and not expressing SIRV1 *gp29*, respectively.

Viral quantification

To calculate the titer of SSeV, 100 μ L diluted virus (10^{-5} , 10^{-6} , and 10^{-7}) was co-incubated with 500 μ L *S. islandicus* Y08.82.36 host²⁷ (10-fold concentrated) without shaking at 76-78 °C for 30 min. Afterwards, the virus-infected cells were transferred into a glass test tube containing 5 mL of pre-warmed Sucrose-Yeast (SY) and 0.8% gelrite mixture, and plated onto SY plates. The plates were put into a plastic bag, and incubated for two days at 76-78 °C. Plaques were counted in plates with proper virus dilutions, and the titer of SSeV were determined as 4.96×10^8 plaque forming units (PFU) per ml.

Infection by virus SSeV of *S. islandicus* M.16.4 with a type III CRISPR defence

The SSeV infection assay was carried out according to the procedure as described previously⁴¹ with minor modifications. In brief, approximately 6×10^8 cells of RJW007 Δ type I-A/pOE and RJW007 Δ type I-A/pOE-gp29 taken from the exponential stage were spun down at 4000 rpm \times 12 min, and resuspended in 1mL of Arabinose-Tryptone (AT) medium. The resuspensions were then co-incubated with 20 mL of fresh AT medium or SSeV supernatant at different dilutions (10^0 , 10^{-1} , 10^{-2} , 10^{-3} , 10^{-4} , 10^{-5} , and 10^{-6}) in a Falcon tube at 76-78 °C for 1 h without shaking. The SSeV-infected cells were washed twice with 10 mL of AT medium and resuspended into 500 μ L of AT medium. Afterwards, the concentrated SSeV-infected cells were mixed with 5 mL of top layer (2.5 mL of 2 \times Arabinose-Yeast medium+2.5 mL of 0.8% gelrite), and then plated onto the Arabinose-Yeast plates. PFU were counted after 4 days of incubation at 76-78 °C. Three independent experiments were performed.

Cloning and purification

For cloning, synthetic genes (g-blocks) were purchased from Integrated DNA Technologies (IDT), Coralville, USA, and cloned into the vector pEhisV5spacerTev between the *NcoI* and *BamHI* sites⁸. Competent DH5 α (*Escherichia coli*) cells were transformed with the construct and sequence integrity confirmed by sequencing (Eurofins Genomics). The plasmid was transformed into *Escherichia coli* C43 (DE3) cells for protein expression. Cloning of AcrIII-1 SIRV1 *gp29*, Crn1 Sso2081 and SsoCsx1 has been previously described^{16,22}. For expression of SIRV1 *gp29* and *Bacillus subtilis* YddF, 2 L of Luria-Broth culture was grown at 37 °C to an OD₆₀₀ of 0.8 with shaking at 180 rpm. Protein expression was induced with 0.4 mM Isopropyl β -D-1-thiogalactopyranoside and cells were grown at 25 °C overnight before harvesting by centrifugation at 4000 rpm (Beckman Coulter Avanti JXN-26; JLA8.1 rotor) at 4 °C for 15 min.

For protein purification the cell pellet was resuspended in four volumes equivalent of lysis buffer containing 50 mM Tris-HCl 7.5, 0.5 M NaCl, 10 mM imidazole and 10% glycerol supplemented with EDTA-free protease inhibitor tablets (Roche; 1 tablet per 100 ml buffer) and lysozyme (1 mg/ml). Cells were lysed by sonicating six times 1 min with 1 min rest intervals on ice at 4 °C, and the lysate was ultracentrifuged at 40,000 rpm (70 Ti rotor) at 4 °C for 35 min. The lysate was then loaded onto a 5 ml HisTrap FF Crude column (GE Healthcare) equilibrated with wash buffer containing 50 mM Tris-HCl pH 7.5, 0.5 M NaCl, 30 mM imidazole and 10% glycerol. Unbound protein was washed away with 20 column volumes (CV) of wash buffer prior to elution of his-tagged protein using a linear gradient (holding at 10% for 3 CV, and 50% for 3 CV) of elution buffer containing 50 mM Tris-HCl pH 7.5, 0.5 M NaCl, 0.5 M imidazole and 10% glycerol. SDS-PAGE was carried out to identify fractions containing the protein of interest, and relevant fractions were pooled and concentrated using a 10 kDa molecular weight cut-off centrifugal concentrator (Merck). The his-tag was removed by incubating concentrated protein overnight with Tobacco Etch Virus (TEV) protease (1 mg per 10 mg protein) while dialysing in buffer containing 50 mM Tris-HCl pH 7.5, 0.5 M NaCl, 30 mM imidazole and 10% glycerol at room temperature. The protein with his-tag removed was isolated using a 5 ml HisTrapFF column, eluting the protein using 4 CV wash buffer. His-tag removed protein was further purified by size-exclusion chromatography (S200 16/60; GE Healthcare) in buffer containing 20 mM Tris-HCl pH 7.5, 0.125 M NaCl using an isocratic gradient. After SDS-PAGE, fractions containing protein of interest were concentrated and protein was aliquoted and stored at -80 °C. Variant enzymes were generated using the QuickChange

Site-Directed Mutagenesis kit as per manufacturer's instructions (Agilent technologies) and purified as for the wild-type proteins.

Radiolabelled cA₄ cleavage assays

Cyclic oligoadenylate (cOA) was generated by incubating 120 µg *Sulfolobus solfataricus* (Sso) type III-D (Csm) complex with 5 nM α -³²P-ATP, 1 mM ATP, 120 nM A26 RNA target and 2 mM MgCl₂ in Cs_x1 buffer containing 20 mM 2-(N-morpholino)ethanesulfonic acid (MES) pH 5.5, 100 mM K-glutamate, 1 mM DTT and 3 units SUPERase•In™ Inhibitor for 2 h at 70 °C in a 100 µl reaction volume. cOA was extracted by phenol-chloroform (Ambion) extraction followed by chloroform extraction (Sigma-Aldrich), and stored at -20 °C.

For single turnover kinetics experiments AcrIII-1 SIRV1 gp29 and variants (4 µM protein dimer) were assayed for radiolabelled cA₄ degradation by incubating with 1/400 diluted ³²P-labelled SsoCsm cOA (~200 nM cA₄; when generated in a 100 µl cOA synthesis reaction as described above) in Cs_x1 buffer supplemented with 1 mM EDTA at 50 °C. AcrIII-1 YddF (8 µM dimer) was incubated with cOA in buffer containing 20 mM MES pH 6.0, 100 mM NaCl, 1 mM DTT, 1 mM EDTA and 3 units SUPERase•In™ Inhibitor at 37 °C. For SIRV1 gp29 H47A chemical rescue, reactions were supplemented with 0.5 M imidazole. Two experimenters were involved in kinetic experiments involving 5 s time points. At desired time points, a 10 µl aliquot of the reaction was removed and quenched by adding to phenol-chloroform and vortexing. Subsequently, 5 µl of deproteinised reaction product was extracted into 5 µl 100% formamide xylene-cyanol loading dye if intended for denaturing polyacrylamide gel electrophoresis (PAGE), or products were further isolated by chloroform extraction if intended for thin-layer chromatography (TLC). A reaction incubating cOA in buffer without protein to the endpoint of each experiment was included as a negative control. All experiments were carried out in triplicate. For SIRV1 gp29 two biological samples were assayed in triplicate. cA₄ degradation was visualised by phosphor imaging following denaturing PAGE (7M urea, 20% acrylamide, 1x TBE) or TLC.

For TLC, 1 µl of radiolabelled product was spotted 1 cm from the bottom of a 20 x 20 cm silica gel TLC plate (Supelco Sigma-Aldrich). The TLC plate was placed in a sealed glass chamber pre-warmed and humidified at 37 °C and containing 0.5 cm of a running buffer composed of 30% H₂O, 70% ethanol and 0.2 M ammonium bicarbonate, pH 9.2. The temperature was lowered to 35 °C and the buffer was allowed to rise along the plate through capillary action until the migration front reached 17 cm. The plate was air dried and sample migration was visualised by phosphor imaging.

For kinetic analysis, cA₄ cleavage was quantified using the Bio-Formats plugin⁴² of ImageJ as distributed in the Fiji package⁴³ and fitted to a single exponential curve ($y = m1 + m2 * (1 - \exp(-m3 * x))$; $m1=0.1$; $m2=1$; $m3=1$) using Kaleidagraph (Synergy Software), as described previously⁴⁴. The cA₄ cleavage rate by the H47A variant in the absence of imidazole was obtained by linear fit.

Deactivation of HEPN nucleases by ring nucleases in coupled assays

In the absence and presence of Crn1 Sso2081 (2 µM dimer) or AcrIII-1 SIRV1 gp29 (2 µM dimer) 4 µg *S. solfataricus* Csm complex (~140 nM Csm carrying crRNA targeting A26 RNA target) was incubated with A26 RNA target (50, 20, 5, 2, or 0.5 nM) in buffer containing 20 mM MES pH 6.0, 100 mM NaCl, 1 mM DTT and 3 units SUPERase•In™ Inhibitor supplemented with 2 mM MgCl₂ and 0.5 mM ATP at 70 °C for 60 min. 5'-end ³²P-labelled A1 RNA (AGGGUAUUUUUGUUUCUUCAAACUAUAAGCUAGUUCUGGAGA) and 0.5 µM dimer SsoCs_x1 was added to the reaction at 60 min and the reaction was allowed to proceed for a further 60 min before quenching by adding phenol-chloroform. A1 RNA cleavage was visualised by phosphor imaging after denaturing PAGE. A control reaction incubating SsoCs_x1 with A1 RNA in the absence of cOA was carried out to determine SsoCs_x1 background activity. cA₄ synthesis by Csm in response to A26 RNA target and subsequent cA₄ degradation in the presence of Crn1 Sso2081 or AcrIII-1 SIRV1 gp29, was visualised by adding 5 nM α -³²P-ATP with 0.5 mM ATP at the start of the reaction. Reactions were quenched at 60 min with phenol-chloroform and cA₄ degradation products were visualised by phosphor imaging following TLC. A control reaction incubating Csm with ATP and α -³²P-ATP in the absence of A26 RNA target was also carried out, quenching the reaction after 60 min.

cA₄ degradation capacity of AcrIII-1 SIRV1 gp29 versus the Crn1 enzyme (Sso2081) was determined by incubating 2 µM dimer of each enzyme with 500-0.5 µM unlabelled cA₄ (BIOLOG Life Science Institute, Bremen, Germany) in Cs_x1 buffer at 70 °C for 20 min before introducing SsoCs_x1 (0.5 µM dimer) and ³²P-labelled A1 RNA (50 nM). The reaction was left to proceed for a further 60 min at 70 °C before quenching by adding phenol-chloroform. Deproteinised products were separated by denaturing PAGE to visualise RNA degradation.

Plasmid transformation assays

Tsac 2833 mediated plasmid immunity from a reprogrammed type III system in *E. coli* Plasmids pCsm1-5_ΔCsm6 (containing the type III Csm interference genes *cas10*, *csm3*, *csm4*, *csm5* from *M. tuberculosis* and *csm2* from *M. canettii*), pCRISPR_TetR (containing *M. tuberculosis* *cas6* and tetracycline resistance gene-targeting CRISPR array), pRAT-Target (tetracycline-resistance, target plasmid) and *M. tuberculosis* (MtB)Csm6/ *Thioalkalivibrio sulfidiphilus* (Tsu)Csx1 expression constructs have been described previously¹⁰. pRAT-Duet was constructed by replacing the pUC19 *lacZα* gene of pRAT-Target with the MCSs of pACYCDuet-1 by restriction digest (5'-NcoI, 3'-XhoI). The viral ring nuclease (*duf1784*) gene from *Thermoanaerobacterium* phage THSA_485A, *tsac_2833*, was PCR-amplified from its pEHisTEV expression construct and cloned into the 5'-NdeI, 3'-XhoI sites of MCS-2. The cOA-dependent nucleases (*mtb csm6*, *tsu csx1*) were cloned into the 5'-NcoI, 3'-SalI sites of MCS-1 by restriction digest from their respective expression constructs. Each nuclease was cloned with and without the viral ring nuclease; pRAT-Duet without insert and pRAT-Duet containing only the viral ring nuclease were used as controls. The plasmid transformation assay was carried out essentially as described in¹⁰. *E. coli* C43 containing pCsm1-5_DCsm6 and pCRISPR_TetR were transformed by heat shock with 100 ng of pRAT-Duet target plasmid containing different combinations of cOA-dependent nuclease and viral ring nuclease. After outgrowth at 37 °C for 2 h, cells were collected and resuspended in 200 μl LB. A series of 10-fold dilutions was applied onto LB agar containing 100 μg ml⁻¹ ampicillin and 50 μg ml⁻¹ spectinomycin to determine the cell density of the recipient cells and onto LB agar additionally containing 25 μg ml⁻¹ tetracycline, 0.2% (w/v) D-lactose and 0.2% (w/v) L-arabinose to determine the cell density of viable transformants. Plates were incubated at 37 °C for 16 – 18 h; further incubation was carried out at room temperature.

Liquid chromatography high-resolution mass spectrometry

AcrIII-1 SIRV1 gp29 (40 μM dimer) was incubated with 400 μM cA₄ in Csx1 buffer for 2 min at 70 °C and deproteinised by phenol-chloroform extraction followed by chloroform extraction. Liquid chromatography-high resolution mass spectrometry (LC-HRMS) analysis was performed on a Thermo Scientific™ Velos Pro instrument equipped with HESI source and Dionex UltiMate 3000 chromatography system. Compounds were separated on a Kinetex® EVO C18 column (2.6 μm, 2.1 × 50 mm, Phenomenex) using the following gradient of acetonitrile (B) against 20 mM ammonium bicarbonate (A): 0 – 2 min 2% B, 2 – 10 min 2 – 8% B, 10 – 11 min 8 – 95% B, 11 – 14 min 95% B, 14 – 15 min 95 – 2% B, 15 – 20 min 2% B at a flow rate of 300 μl min⁻¹ and column temperature of 40 °C. UV data were recorded at 254 nm. Mass data were acquired on the FT mass analyzer in negative ion mode with scan range *m/z* 150 – 1500 at a resolution of 30,000. Source voltage was set to 3.5 kV, capillary temperature was 350 °C, and source heater temperature was 250 °C. Data were analysed using Xcalibur™ (Thermo Scientific).

Phylogenetic analysis

AcrIII-1 homologs were collected by using gp29 (NP_666617) of SIRV1 as a query and running two iterations (E=1e-05) of PSI-BLAST⁴⁵ against the non-redundant protein database at NCBI. The sequences were aligned using PROMALS3D⁴⁶. Redundant sequences (95% identity threshold) and sequences with mutated H47 active site residue were removed from the alignment. Poorly aligned (low information content) positions were removed using the gt 0.2 function of Trimal⁴⁷. The final alignment contained 124 positions. The maximum likelihood phylogenetic tree was constructed using the PhyML program⁴⁸ with the automatic selection of the best-fit substitution model for a given alignment. The best model identified by PhyML was LG +G+I. The branch support was assessed using aBayes implemented in PhyML. The tree was visualized using iTOL⁴⁹.

Crystallisation

AcrIII-1 H47A variant was concentrated to 10 mg ml⁻¹, incubated at 293 K for 1 hour with a 1.2 M excess of cA₄, and centrifuged at 13,000 rpm for 10 minutes prior to crystallization. Sitting drop vapor diffusion experiments were set up at the nanoliter scale using commercially available and in-house crystallization screens and incubated at 293 K. Crystals appeared in various conditions, but those used for data collection grew from 40% 2-methyl-2,4-pentanediol, 5% polyethylene glycol 8000 and 0.1 M sodium cacodylate, pH 6.5. Crystals were harvested and transferred briefly into cryoprotectant containing mother liquor with 20% glycerol immediately before cryo-cooling in liquid nitrogen. The H47A variant was used to avoid cleavage of the cA₄ substrate during co-crystallisation. The position of the active site histidine was inferred from the structure of the apo-protein.

Data collection and processing

X-ray data were collected from two crystals at 100 K, at a wavelength 0.9686 Å, on beamline I24 at Diamond Light Source, to 1.49 and 1.60 Å resolution. Both data sets were automatically processed using Xia2⁵⁰, utilizing XDS and XSCALE⁵¹. The data were merged in Aimless⁵² and the overall resolution truncated to 1.55 Å. The data were phased by molecular

replacement using Phaser⁵³ with a monomer from PDB file 2X4I stripped of water molecules as the search model. Model refinement of AcrIII-1 was achieved by iterative cycles of REFMAC5⁵⁴ in the CCP4 suite⁵⁵ and manual manipulation in COOT⁵⁶. Electron density for cA₄ was clearly visible in the maximum likelihood/ σ_A weighted $F_{\text{obs}} - F_{\text{calc}}$ electron density map at 3 σ . The coordinates for cA₄ were generated in ChemDraw (Perkin Elmer) and the library was generated using acedrg⁵⁷, before fitting of the molecule in COOT. Model quality was monitored throughout using Molprobity⁵⁸. Data and refinement statistics are shown in Table S1.

Data availability statement

The structural coordinates and data have been deposited in the Protein Data Bank with deposition code 6SCF.

Raw data is available for the plasmid immunity analysis presented in figure 1.

Raw data is available for the kinetic analysis presented in figure 2 and extended data figure 6.

REFERENCES:

- 1 Kazlauskiene, M., Kostiuk, G., Venclovas, C., Tamulaitis, G. & Siksny, V. A cyclic oligonucleotide signaling pathway in type III CRISPR-Cas systems. *Science* **357**, 605-609, doi:10.1126/science.aao100 (2017).
- 2 Niewohner, O. *et al.* Type III CRISPR-Cas systems produce cyclic oligoadenylate second messengers. *Nature* **548**, 543-548, doi:10.1038/nature23467 (2017).
- 3 Rouillon, C., Athukoralage, J. S., Graham, S., Grüschor, S. & White, M. F. Control of cyclic oligoadenylate synthesis in a type III CRISPR system. *eLife* **7**, e36734 (2018).
- 4 Rostol, J. T. & Marraffini, L. A. Non-specific degradation of transcripts promotes plasmid clearance during type III-A CRISPR-Cas immunity. *Nat Microbiol* **4**, 656-662, doi:10.1038/s41564-018-0353-x (2019).
- 5 Pyenson, N. C., Gayvert, K., Varble, A., Elemento, O. & Marraffini, L. A. Broad Targeting Specificity during Bacterial Type III CRISPR-Cas Immunity Constrains Viral Escape. *Cell Host Microbe* **22**, 343-353 e343, doi:10.1016/j.chom.2017.07.016 (2017).
- 6 Whiteley, A. T. *et al.* Bacterial cGAS-like enzymes synthesize diverse nucleotide signals. *Nature* **567**, 194-199, doi:10.1038/s41586-019-0953-5 (2019).
- 7 Maelfait, J. & Rehwinkel, J. RECONsidering Sensing of Cyclic Dinucleotides. *Immunity* **46**, 337-339, doi:10.1016/j.immuni.2017.03.005 (2017).
- 8 Rouillon, C., Athukoralage, J. S., Graham, S., Grüschor, S. & White, M. F. Investigation of the cyclic oligoadenylate signalling pathway of type III CRISPR systems. *Methods Enzymol* **616**, 191-218 (2019).
- 9 Nasef, M. *et al.* Regulation of cyclic oligoadenylate synthesis by the S. epidermidis Cas10-Csm complex. *RNA*, pii: rna.070417.070119, doi:10.1261/rna.070417.119 (2019).
- 10 Grüschor, S., Athukoralage, J. S., Graham, S., Hoogendoorn, T. & White, M. F. Cyclic oligoadenylate signalling mediates Mycobacterium tuberculosis CRISPR defence. *BioRxiv* **667758**, doi:<https://doi.org/10.1101/667758> (2019).
- 11 Makarova, K. S., Anantharaman, V., Grishin, N. V., Koonin, E. V. & Aravind, L. CARF and WYL domains: ligand-binding regulators of prokaryotic defense systems. *Frontiers in genetics* **5**, 102, doi:10.3389/fgene.2014.00102 (2014).
- 12 Deng, L., Garrett, R. A., Shah, S. A., Peng, X. & She, Q. A novel interference mechanism by a type IIIB CRISPR-Cmr module in Sulfolobus. *Mol. Microbiol.* **87**, 1088-1099, doi:10.1111/mmi.12152 (2013).
- 13 Jiang, W., Samai, P. & Marraffini, L. A. Degradation of Phage Transcripts by CRISPR-Associated RNases Enables Type III CRISPR-Cas Immunity. *Cell* **164**, 710-721, doi:10.1016/j.cell.2015.12.053 (2016).
- 14 Foster, K., Kalter, J., Woodside, W., Terns, R. M. & Terns, M. P. The ribonuclease activity of Csm6 is required for anti-plasmid immunity by Type III-A CRISPR-Cas systems. *RNA Biol.* **1**, 1-12, doi:10.1080/15476286.2018.1493334 (2018).
- 15 Hatoum-Aslan, A., Maniv, I., Samai, P. & Marraffini, L. A. Genetic Characterization of Antiplasmid Immunity through a Type III-A CRISPR-Cas System. *J. Bacteriol.* **196**, 310-317, doi:10.1128/JB.01130-13 (2014).
- 16 Athukoralage, J. S., Rouillon, C., Graham, S., Grüschor, S. & White, M. F. Ring nucleases deactivate Type III CRISPR ribonucleases by degrading cyclic oligoadenylate. *Nature* **562**, 277-280, doi:<https://doi.org/10.1101/380436> (2018).
- 17 Borges, A. L., Davidson, A. R. & Bondy-Denomy, J. The Discovery, Mechanisms, and Evolutionary Impact of Anti-CRISPRs. *Annu Rev Virol* **4**, 37-59, doi:10.1146/annurev-virology-101416-041616 (2017).
- 18 He, F. *et al.* Publisher Correction: Anti-CRISPR proteins encoded by archaeal lytic viruses inhibit subtype I-D immunity. *Nat Microbiol* **3**, 1076, doi:10.1038/s41564-018-0184-9 (2018).
- 19 Pawluk, A. *et al.* Inactivation of CRISPR-Cas systems by anti-CRISPR proteins in diverse bacterial species. *Nat Microbiol* **1**, 16085, doi:10.1038/nmicrobiol.2016.85 (2016).
- 20 Maxwell, K. L. Phages Fight Back: Inactivation of the CRISPR-Cas Bacterial Immune System by Anti-CRISPR Proteins. *PLoS Pathog* **12**, e1005282, doi:10.1371/journal.ppat.1005282 (2016).
- 21 Hwang, S. & Maxwell, K. L. Meet the Anti-CRISPRs: Widespread Protein Inhibitors of CRISPR-Cas Systems. *CRISPR J* **2**, 23-30, doi:10.1089/crispr.2018.0052 (2019).
- 22 Oke, M. *et al.* The Scottish Structural Proteomics Facility: targets, methods and outputs. *J Struct Funct Genomics* **11**, 167-180, doi:10.1007/s10969-010-9090-y (2010).
- 23 Larson, E. T. *et al.* A new DNA binding protein highly conserved in diverse crenarchaeal viruses. *Virology* **363**, 387-396, doi:10.1016/j.virol.2007.01.027 (2007).
- 24 Ortmann, A. C. *et al.* Transcriptome analysis of infection of the archaeon Sulfolobus solfataricus with Sulfolobus turreted icosahedral virus. *J Virol* **82**, 4874-4883, doi:10.1128/JVI.02583-07 (2008).
- 25 Wirth, J. F. *et al.* Development of a genetic system for the archaeal virus Sulfolobus turreted icosahedral virus (STIV). *Virology* **415**, 6-11, doi:10.1016/j.virol.2011.03.023 (2011).
- 26 Zhang, C. & Whitaker, R. J. Microhomology-Mediated High-Throughput Gene Inactivation Strategy for the Hyperthermophilic Crenarchaeon Sulfolobus islandicus. *Appl Environ Microbiol* **84**, doi:10.1128/AEM.02167-17 (2018).
- 27 Bautista, M. A., Zhang, C. & Whitaker, R. J. Virus-induced dormancy in the archaeon Sulfolobus islandicus. *MBio* **6**, e02565-02514, doi:10.1128/mBio.02565-14 (2015).
- 28 Bondy-Denomy, J. *et al.* A Unified Resource for Tracking Anti-CRISPR Names. *CRISPR J* **1**, 304-305, doi:10.1089/crispr.2018.0043 (2018).
- 29 Auchtung, J. M., Aleksanyan, N., Bulku, A. & Berkmen, M. B. Biology of ICEBs1, an integrative and conjugative element in *Bacillus subtilis*. *Plasmid* **86**, 14-25, doi:10.1016/j.plasmid.2016.07.001 (2016).
- 30 Keller, J. *et al.* Crystal structure of AFV3-109, a highly conserved protein from crenarchaeal viruses. *Virol J* **4**, 12, doi:10.1186/1743-422X-4-12 (2007).
- 31 Yang, W. Nucleases: diversity of structure, function and mechanism. *Q Rev Biophys* **44**, 1-93, doi:10.1017/S0033583510000181 (2011).

32 Athukoralage, J. S., Graham, S., Grüschor, S., Rouillon, C. & White, M. F. A type III CRISPR ancillary ribonuclease degrades its cyclic oligoadenylate activator. *J. Mol. Biol.* **431**, 2894-2899, doi:<https://doi.org/10.1016/j.jmb.2019.04.041> (2019).

33 Niewoehner, O. & Jinek, M. Structural basis for the endoribonuclease activity of the type III-A CRISPR-associated protein Csm6. *RNA* **22**, 318-329, doi:10.1261/rna.054098.115 (2016).

34 McMahon, S. A. *et al.* Structure and mechanism of a Type III CRISPR defence DNA nuclease activated by cyclic oligoadenylate. *submitted* (2019).

35 Knott, G. J. *et al.* Broad-spectrum enzymatic inhibition of CRISPR-Cas12a. *Nat Struct Mol Biol* **26**, 315-321, doi:10.1038/s41594-019-0208-z (2019).

36 Dong, L. *et al.* An anti-CRISPR protein disables type V Cas12a by acetylation. *Nat Struct Mol Biol* **26**, 308-314, doi:10.1038/s41594-019-0206-1 (2019).

37 Anderson, R. E., Kouris, A., Seward, C. H., Campbell, K. M. & Whitaker, R. J. Structured Populations of *Sulfolobus acidocaldarius* with Susceptibility to Mobile Genetic Elements. *Genome biology and evolution* **9**, 1699-1710, doi:10.1093/gbe/evx104 (2017).

38 Koonin, E. V. & Makarova, K. S. Anti-CRISPRs on the march. *Science* **362**, 156-157, doi:10.1126/science.aav2440 (2018).

39 Held, N. L., Herrera, A. & Whitaker, R. J. Reassortment of CRISPR repeat-spacer loci in *Sulfolobus islandicus*. *Environmental microbiology* **15**, 3065-3076, doi:10.1111/1462-2920.12146 (2013).

40 Zhang, C., Cooper, T. E., Krause, D. J. & Whitaker, R. J. Augmenting the genetic toolbox for *Sulfolobus islandicus* with a stringent positive selectable marker for agmatine prototrophy. *Appl Environ Microbiol* **79**, 5539-5549, doi:10.1128/AEM.01608-13 (2013).

41 Bautista, M. A., Zhang, C. & Whitaker, R. J. Virus-induced dormancy in the archaeon *Sulfolobus islandicus*. *MBio* **6**, doi:10.1128/mBio.02565-14 (2015).

42 Linkert, M. *et al.* Metadata matters: access to image data in the real world. *J Cell Biol* **189**, 777-782, doi:10.1083/jcb.201004104 (2010).

43 Schindelin, J. *et al.* Fiji: an open-source platform for biological-image analysis. *Nat Methods* **9**, 676-682, doi:10.1038/nmeth.2019 (2012).

44 Sternberg, S. H., Haurwitz, R. E. & Doudna, J. A. Mechanism of substrate selection by a highly specific CRISPR endoribonuclease. *RNA* **18**, 661-672, doi:10.1261/rna.030882.111 (2012).

45 Altschul, S. F. *et al.* Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucl. Acids Res.* **25**, 3389-3402, doi:10.1093/nar/25.17.3389 (1997).

46 Pei, J. & Grishin, N. V. PROMALS3D: multiple protein sequence alignment enhanced with evolutionary and three-dimensional structural information. *Methods Mol Biol* **1079**, 263-271, doi:10.1007/978-1-62703-646-7_17 (2014).

47 Capella-Gutierrez, S., Silla-Martinez, J. M. & Gabaldon, T. trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics* **25**, 1972-1973, doi:10.1093/bioinformatics/btp348 (2009).

48 Guindon, S. *et al.* New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Syst Biol* **59**, 307-321, doi:10.1093/sysbio/syq010 (2010).

49 Letunic, I. & Bork, P. Interactive Tree Of Life (iTOL) v4: recent updates and new developments. *Nucl. Acids Res.* **47**, W256-W259, doi:10.1093/nar/gkz239 (2019).

50 Winter, G. xia2: an expert system for macromolecular crystallography data reduction. *Journal of Applied Crystallography* **43**, 186-190, doi:10.1107/S0021889809045701 (2010).

51 Kabsch, W. Xds. *Acta crystallographica. Section D, Biological crystallography* **66**, 125-132, doi:10.1107/S0907444909047337 (2010).

52 Evans, P. R. An introduction to data reduction: space-group determination, scaling and intensity statistics. *Acta Crystallogr D Biol Crystallogr* **67**, 282-292, doi:10.1107/S090744491003982X (2011).

53 McCoy, A. J. *et al.* Phaser crystallographic software. *Journal of Applied Crystallography* **40**, 658-674, doi:10.1107/S0021889807021206 (2007).

54 Murshudov, G. N., Vagin, A. A. & Dodson, E. J. Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr D Biol Crystallogr* **53**, 240-255, doi:10.1107/S0907444996012255 (1997).

55 Winn, M. D. *et al.* Overview of the CCP4 suite and current developments. *Acta Crystallogr D Biol Crystallogr* **67**, 235-242, doi:10.1107/S0907444910045749 (2011).

56 Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta Crystallogr D Biol Crystallogr* **66**, 486-501, doi:10.1107/S0907444910007493 (2010).

57 Long, F. *et al.* AceDRG: a stereochemical description generator for ligands. *Acta Crystallogr D Struct Biol* **73**, 112-122, doi:10.1107/S2059798317000067 (2017).

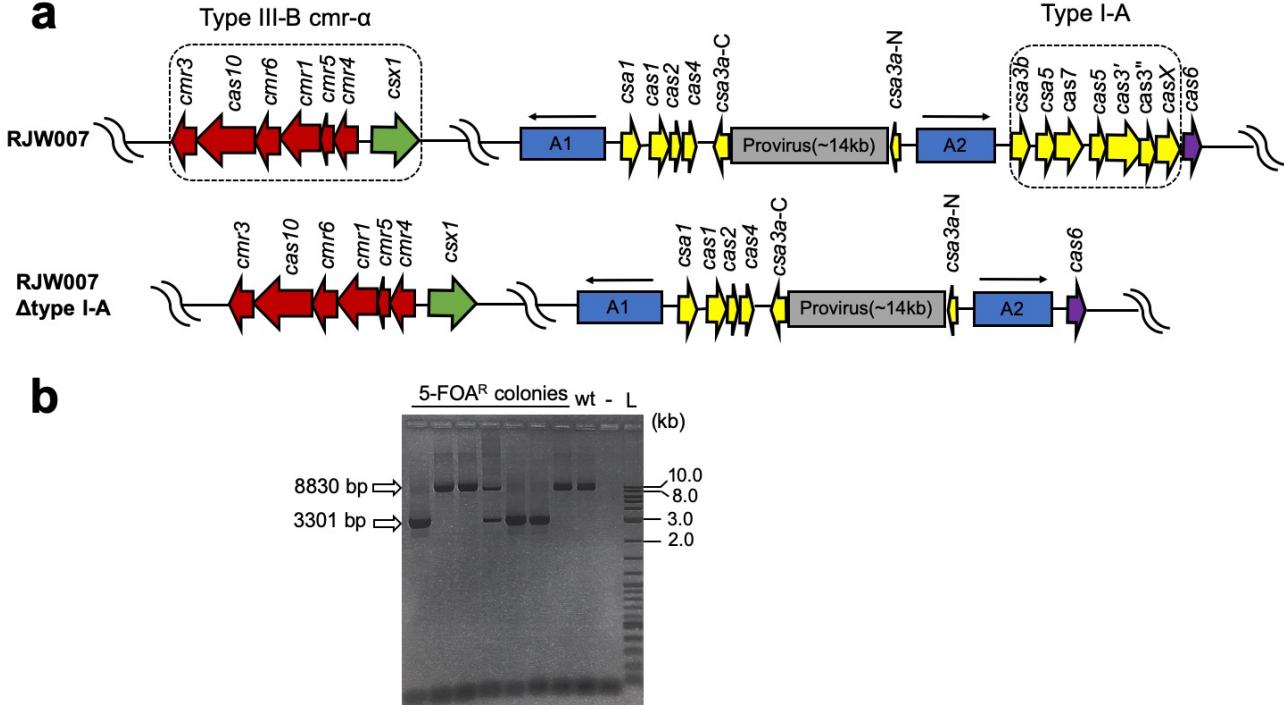
58 Chen, V. B. *et al.* MolProbity: all-atom structure validation for macromolecular crystallography. *Acta crystallographica. Section D, Biological crystallography* **66**, 12-21, doi:10.1107/S0907444909042073 (2010).

59 Gerlt, J. A. Genomic Enzymology: Web Tools for Leveraging Protein Family Sequence-Function Space and Genome Context to Discover Novel Functions. *Biochemistry* **56**, 4293-4308, doi:10.1021/acs.biochem.7b00614 (2017).

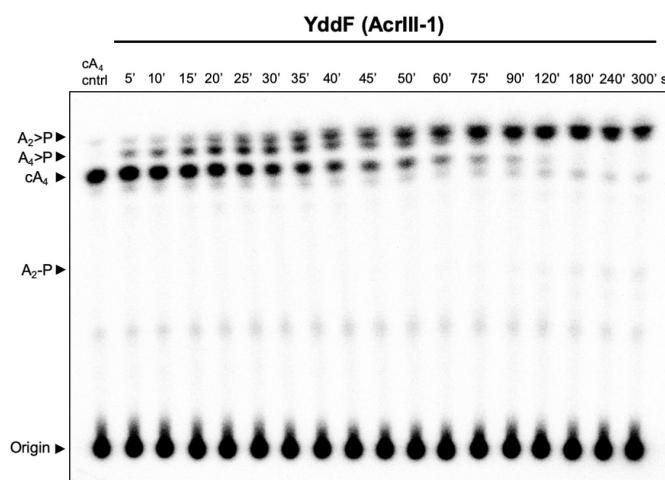
Extended dataset

| | | | |
|---------------------------------------|-------|--|------|
| SRV1_gp29 | 1 | MNKVYLANAFSINM--LTKFPTKVVIDKIDRLEFCEN--IDNED | 40 |
| STIV_B116 | 1 | MGKVFLNTAFAINSNM--LKEFPTTITIDKLDEDFCLKLELRLEDGT | 44 |
| AFV3-109 | 1 | MLYILNSAIALPL--KPGEEYTVKAKEITIQEAKEL--VTKEQ | 38 |
| ARV1_gp13 | 1 | MLYILNAQIITP--FEAGAQATFVERRIDVNEAKKI--VNSQ | 37 |
| SIVF_118 | 1 | MLYILSATLPL--KPGKEYVIHAKELTIEEAKELE--LENER | 38 |
| SMV4_113 | 1 | MTVYLANAFSPSM--LNKLPSAVEFQRVDQKEFCEA--IHGG- | 38 |
| ATV_gp06 | 1 | MGVWSVVLYLLNTLIVP--FRDERAKFEIERSAEEAKKI IQ--MHNSQ | 45 |
| Thermoanaerobacterium phage THSA-485A | 1 | MFIANAFSLQMQ--LSQFPAHIDIEEVATSAVAKL--D | 33 |
| Synechococcus phage S-CBWM1 | 1 | MACCVVP--KGAPGLWSVVEISLEEVIQD--LEEGE | 32 |
| Fusobacterium phage Fnu1 | 1 | MTIGILNTPTILT--GEGTYKLSNITLEQAQKL--VNENE | 35 |
| Hydrogenobaculum phage 1 | 1 | MLYVLNNSLIIVPVDFQNQKQGYIVSLWKIDLETARKI--VREMP | 40 |
| ICEBs1 Yddf | 1 | MEIAFLNNSLVVT--SPGFYKAEKITDELKHY--DGR | 33 |
| Crenothrix polyspora | 1 | MTLFIIINAPILT--SYGDWRFEG-PLSIDKTRKL--LR-EG | 35 |
| | - - - | N. - - | - - |
| SRV1_gp29 | 41 | IINSIGHDSTIQLINSLCGTTFQKNRVEIKLEKEDKLYVVQI-SQRLEEG | 89 |
| STIV_B116 | 45 | LINAIGHDSTINLVNTLCGTQLQKRNVEVKMNEGDEALIIMI-SQRLEEG | 93 |
| AFV3-109 | 39 | FTSAIGHQATAELLSILGVNVPMNRQVICKVTHGDRILAFML-KQRLEPEG | 87 |
| ARV1_gp13 | 38 | FVSAVGHAATAQLLSKLDASIPTNRTQVFLKPGDMALAIVL-KSRIPEG | 86 |
| SIVF_118 | 39 | FISAVGHEATAKMLTNIFDVEIPMNRIQIFLDDGDKLLSIIIL-KTRLEEG | 87 |
| SMV4_113 | 39 | VNSAIGHKGKTIEFVNTLCNTLNQTLNVEIKAGINDVIVIYL-GFRLLEEG | 87 |
| ATV_gp06 | 46 | FVSAIGHASANALSLLLGVAVPVNTEREVFFNVGDEIAAMAL-KKRLAEG | 94 |
| Thermoanaerobacterium phage THSA-485A | 34 | LQSAIGHADTAVVLNLGKDIESNRVNVQLQPGDSLIVAQLMGGRLPEG | 83 |
| Synechococcus phage S-CBWM1 | 33 | FISTIGHPSSAHILETLTGFPFEACRREADPRPGDEFYCFIL-NSRAPEG | 81 |
| Fusobacterium phage Fnu1 | 36 | FISYIGHQATAEIISIILGTEVPMNRQGFKQEVGQKAIIFKL-KSRLLEG | 84 |
| Hydrogenobaculum phage 1 | 47 | FTSAVGHEATAKVLSELLGVEISFNRITVKMKGEDAGLHFVL-RTRLPEG | 89 |
| ICEBs1 Yddf | 34 | YKSFIGHKSTAQFLQKLLGIRIEQNRKTFRHMKYQKAICFSL-YERYPEN | 82 |
| Crenothrix polyspora | 36 | FTSAIGHAASAEMLARLLAMDIPIVNRIAITMEAGDRALILRL-LQRLEPEG | 84 |
| | GH | NR | R EG |
| SRV1_gp29 | 90 | KILTL-EEILKLYESGKVQFFEIIIVD | 114 |
| STIV_B116 | 94 | KVLSKD-KEIKDMYRQGKISFYEVW | 116 |
| AFV3-109 | 88 | VVVKTTEELEKI----GYELWLFEIQ | 109 |
| ARV1_gp13 | 87 | VVLDE-QAIRNI----GFEIVVIERSV | 108 |
| SIVF_118 | 88 | KVIKTVEELEQI----GYNIWLFEVVTYEHNVKYE | 118 |
| SMV4_113 | 88 | KVLSA-GEVQKAYDEGKVLLKAIIGK | 113 |
| ATV_gp06 | 95 | QVLRVTQELEAV----GFDLYYIKRQV | 117 |
| Thermoanaerobacterium phage THSA-485A | 84 | STTLP-AGF----SFKFFKVTVQA | 102 |
| Synechococcus phage S-CBWM1 | 82 | KILDE-HEIYKI----GFSFRKMTYVLGKIPPTAPD | 111 |
| Fusobacterium phage Fnu1 | 85 | QILTIQEIEI----GYEFQLLERKN | 107 |
| Hydrogenobaculum phage 1 | 90 | KVLS-EELRL----DFDVLVLSRVS | 110 |
| ICEBs1 Yddf | 83 | VLLTQ-RDLEKA----RYQFYVLLTRLD | 104 |
| Crenothrix polyspora | 85 | KVLNH-HEMMAT----PFELALLTKLK | 106 |

Extended data figure 1. Multiple sequence alignment of DUF1874 family members. Includes the AcrIII-1 proteins from the archaeal viruses SIRV1, STIV, AFV3, ARV1, SIFV, SMV4 and ATV, the integrated conjugative element ICEBs1 protein YddF from *B. subtilis*, the bacteriophage proteins from *Thermoanaerobacterium* phageTHSA-485A, *Synechococcus* phage S-CBWM1, *Fusobacterium* phage Fnu1, *Hydrogenobaculum* phage 1 and the Crn2 protein from *Crenothrix polyspora*.

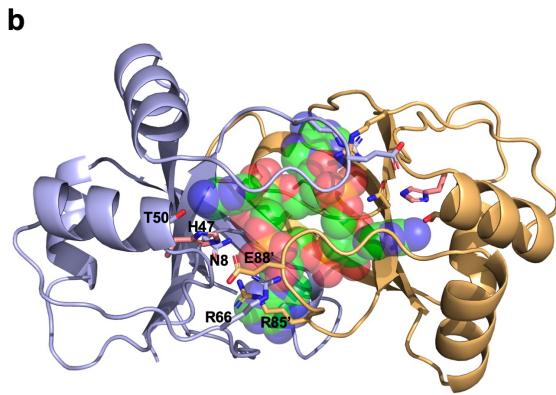
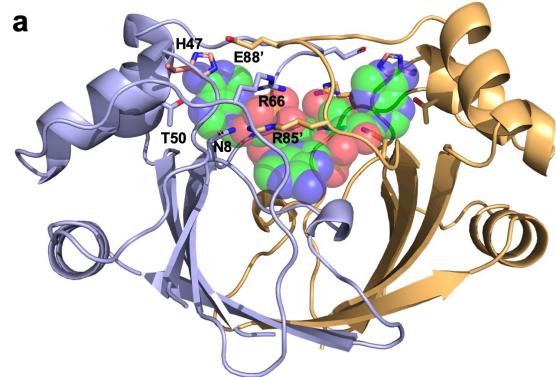


Extended data figure 2 – Construction of type I-A deletion mutant in the genetic host *S. islandicus* RJW007 (*S. islandicus* M.16.4 Δ pyrEF Δ argD). (a) Genomic context of CRISPR system in the genetic host and the mutant strain. A1 and A2 denote two different CRISPR arrays, the orientation of which are indicated with arrows. (b) PCR verification of Δ type I-A mutants. A representative *Sulfolobus* transformant with type I-A knockout plasmid integrated was grown in Dextrin-Tryptone liquid medium, and the cell cultures were subsequently plated on Dextrin-Tryptone plates containing 5-FOA (5-fluoroorotic acid, 50 μ g/mg), uracil (20 μ g/ml), and agamatin (1 mg/ml). Seven randomly selected 5-FOA resistant (5-FOA^R) colonies were screened using the primer set type IA-flankP-For/Rev that binds outside of the flanking homologous regions to confirm the type I-A deletion. A representative Δ type I-A mutant was further colony-purified for following experiment. The expected size of PCR product amplified from the genomic DNA of parental strain (referred to wt) and Δ type I-A mutant is 8830 bp and 3301 bp, respectively. -, negative control (water as DNA template for PCR). L, 2-log DNA ladder (NEB, USA).



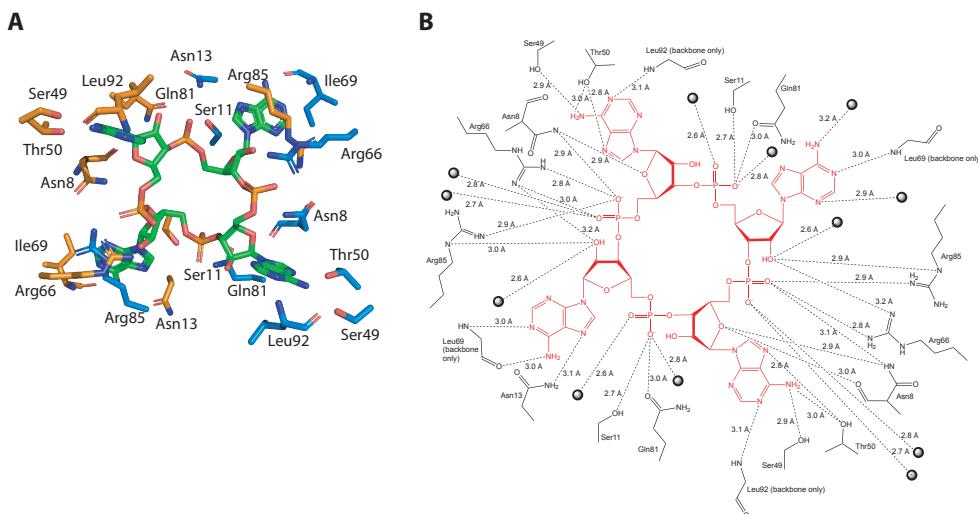
Extended data figure 3. Single turnover kinetics of cA₄ cleavage by *B. subtilis* YddF.

Phosphorimage of thin-layer chromatography (TLC) visualising cA₄ cleavage by YddF (8 μ M dimer, 37 °C) over a period of 5 min. The rate of cA₄ cleavage to generate A₄>P (intermediate product tetra-adenylate containing 2',3' phosphate and a 5' hydroxyl moiety) and A₂>P over time (as shown in figure 2c) was calculated by quantifying densiometric signals from the phosphorimage, which is representative of three independent experiments. Control reaction (cA₄ cntrl) is cA₄ incubated in the absence of protein for 5 min at 37 °C.

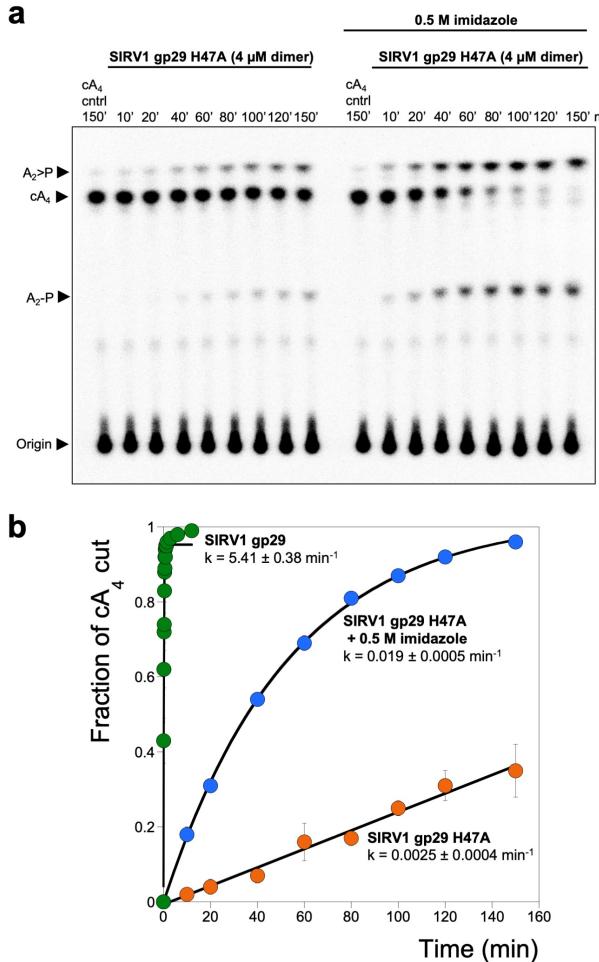


Extended data figure 4. Structure of SIRV1 gp29 bound to cA₄.

Orthogonal views of SIRV1 gp29 dimer in complex with cA₄. The monomers are coloured blue and orange, with catalytic residue H47 from the apo structure shown in salmon. cA₄ is shown in green spheres. Conserved residues (Extended data figure 1) in the AcrIII-1 family are indicated and discussed in the text.

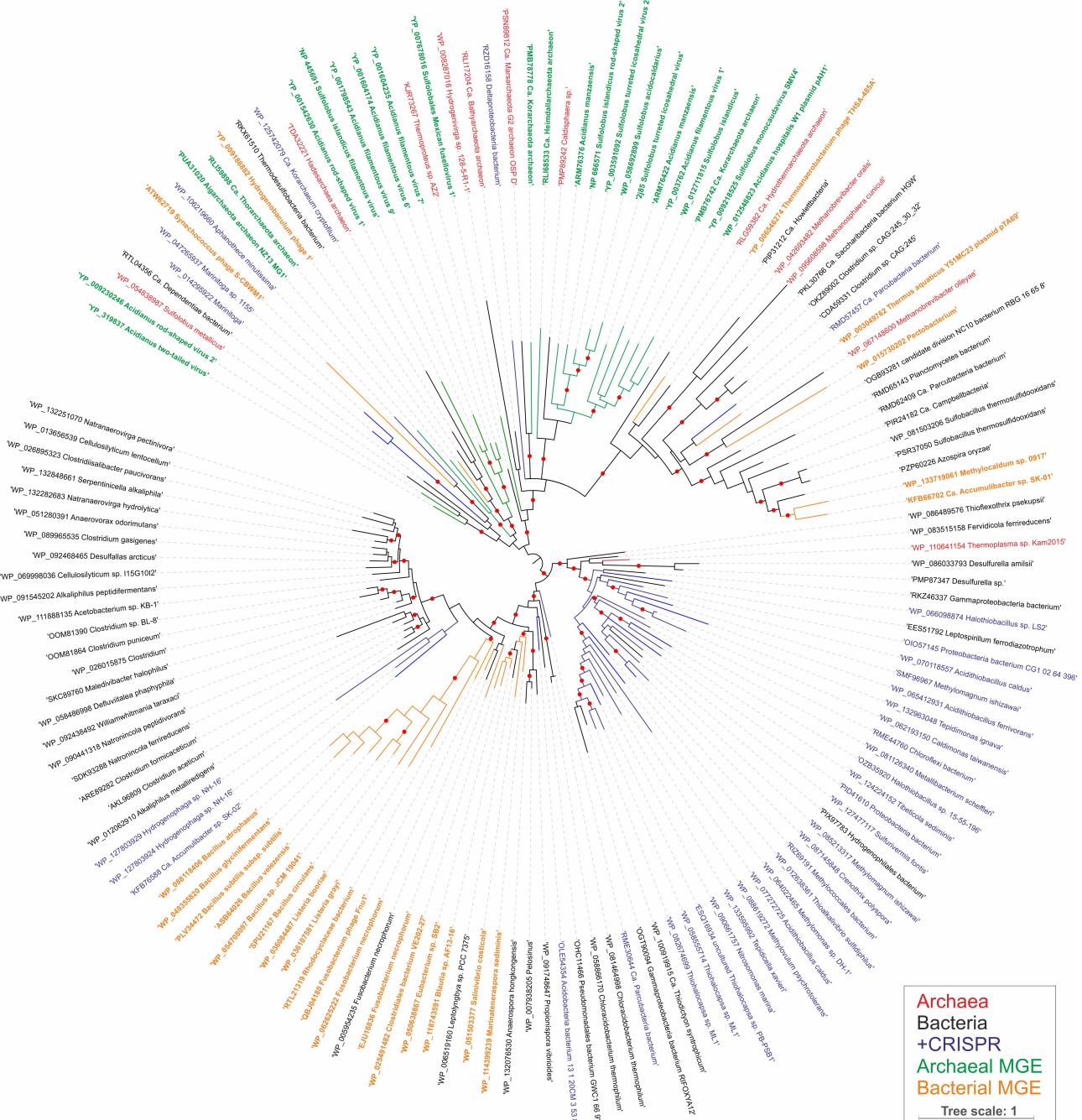


Extended data figure 5. Interactions made in the active site of SIRV1 gp29 in complex with cA₄. (a) Interactions between each monomer of the SIRV1 dimer (orange and blue), with cA₄ shown in green. (b) Schematic showing the interaction. Dotted lines represent hydrogen bonds, with the distance annotated. Spheres represent water molecules.



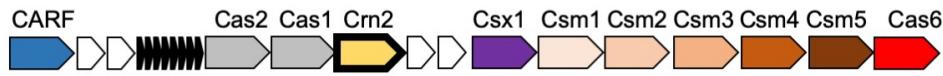
Extended data figure 6. Single turnover cA₄ cleavage by SIRV1 gp29 H47A and chemical rescue with imidazole

(a) Phosphorimage of TLC visualising cA₄ cleavage by SIRV1 gp29 H47A (4 μ M dimer, 50 °C) in the presence or absence of 500 mM imidazole, over time. The rate of cA₄ cleavage to generate A₂>P and A₂-P was calculated by quantifying densiometric signals from the phosphorimage, which is representative of three independent experiments. (b) Plot comparing the single turnover rates of cA₄ by SIRV1 gp29 and its H47A variant in the presence or absence of imidazole. cA₄ cleavage by the H47A variant can be partially restored when the reaction is supplemented with 500 mM imidazole. Errors bars indicate the standard deviation of the mean.



Extended data figure 7. Maximum likelihood phylogeny of AcrIII-1 homologs. The maximum likelihood phylogenetic tree was constructed with the automatic selection of the best-fit substitution model for a given alignment (LG +G+I). Red circles indicate 95-100% branch support, as assessed using aBayes implemented in PhyML. The scale bar represents the number of substitutions per site. Branches and labels are colour coded: archaea, red; bacteria, black; bacteria and archaea in which AcrIII-1 homologs are associated with CRISPR loci, blue; archaeal viruses and plasmids, green; bacteriophages, orange.

Crenothrix polyspora



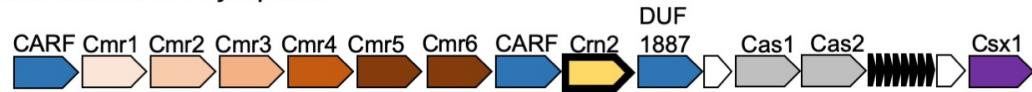
Methylovulum psychrotolerans



Methylovulum psychrotolerans



Thioalkalivibrio sulfidiphilus



Marinitoga piezophilia



Extended data figure 8. Genomic context of *crn2* genes in selected bacteria

Type III CRISPR loci in the bacterial species *Crenothrix polyspora*, *Methylovulum psychrotolerans*, *Methylovulum psychrotolerans*, *Methylovulum psychrotolerans*, *Thioalkalivibrio sulfidiphilus* and *Marinitoga piezophilia* are shown in cartoon form, with genes labelled and colour-coded. The *crn2* gene is shown in pale yellow with a bold outline; CRISPRs are indicated by small black arrows and unrelated/hypothetical genes shown as small white arrows. The size and orientation of genes is not reflected in the cartoon. Key to gene labels: RT, Reverse Transcriptase; Ago, Argonaute; CARF, CRISPR associated Rossman fold; DUF1887, predicted CARF nuclease; CARF-RelE, CARF domain fused to the RelE toxin.



Extended data figure 9. CRISPR-associated ACRII-1 homologs. Genomic neighbourhoods were analysed using the Enzyme Function Initiative-Genome Neighbourhood Tool (EFI-GNT) against the Pfam profile database⁵⁹. Gene annotations are colour coded and the key is provided on the right of the figure.

Table S1: Data collection and refinement statistics for AcrIII-1 in complex with cA₄

| AcrIII-1 – cA ₄ | |
|---------------------------------------|---------------------------|
| Data Collection | |
| Space group | P1 |
| Cell dimensions | |
| a, b, c (Å) | 49.8, 51.7, 85.6 |
| α, β, γ (°) | 80.2, 89.7, 83.4 |
| Resolution (Å) | 50.63 – 1.55 (1.58-1.55)* |
| R _{merge} | 0.12 (0.36)* |
| I/σ(I) | 12.3 (1.7)* |
| Completeness (%) | 98.6 (92.4)* |
| Multiplicity | 2.9 (1.8)* |
| CC _{1/2} | 0.99 (0.79) |
| V _m (Å ³ /Da) | 2.04 |
| Solvent (%) | 39.8 |
| Refinement | |
| Unique reflections | 119765 |
| R _{work} / R _{free} | 19.8 / 24.3 |
| RMSD bonds (Å) | 0.012 |
| RMSD angles (°) | 1.61 |
| Number of atoms (non hydrogen): | |
| Protein | 7383 |
| Water | 598 |
| cA ₄ | 352 |
| B factors (Å ²): | |
| Protein | 19.8 |
| Water | 30.2 |
| cA ₄ | 12.9 |
| Ramachandran plot: | |
| favoured / disallowed (%) | 98.5 / 0 |
| Molprobity score / centile | 1.13 / 99 |

* Values in parentheses refer to the high resolution shell
RMSD, root mean square deviation