Systematic discovery of natural CRISPR-Cas12a inhibitors

Kyle E. Watters1, Christof Fellmann1-2, Hua B. Bai1, Shawn M. Ren1, Jennifer A. Doudna1,2,3,4,5,6,7*

CRISPR-Cas systems represent the only known adaptive mechanism by which prokaryotes protect themselves from biological attackers (1). Although diverse in composition, all CRISPR-Cas pathways employ RNA-guided enzymes that recognize and destroy foreign nucleic acids, commonly double-stranded DNA (2). The ease of changing the RNA guide molecule, and hence the DNA targeting specificity, has enabled use of CRISPR-Cas9 and Cas12 for programmable genome editing in a wide range of cells and organisms (3, 4). To control Cas9, bacterial inhibitors referred to as anti-CRISPRs (Acrs) have been found to limit or block Cas9 functions (5–9). However, these inhibitors have been found only sporadically, and no such inhibitors have been reported for Cas12a (Cpf1).

The known CRISPR-Cas inhibitors have been identified either through isolation of CRISPR-resistant phages (9–13) or by proximity to anti-CRISPR associated (acra) genes (14, 15). As an alternative inhibitor discovery strategy, the presence of stable self-targeting CRISPR sequences has been proposed as a potential indicator of genomes or mobile genetic elements (MGEs) harboring CRISPR inhibitors (Fig. 1A) (16, 17). Self-targeting CRISPR sequences in CRISPR arrays are expected to be lethal to the host cell by directing cleavage and subsequent degradation of the microbial genome (18). However, such self-targeting CRISPR sequences could exist in cells harboring CRISPR inhibitors (16). To test whether CRISPR inhibitors can be discovered systematically by flagging CRISPR self-targeting genomes, we built a bioinformatic pipeline to search across the NCBI prokaryotic sequence database to locate self-targeting examples within predicted CRISPR arrays (Fig. 1B). The Self-Targeting Spacer Searcher (STSS) first predicts all possible CRISPR arrays using the CRISPR Recognition Tool (CRT) (19) and BLASTs each spacer against the host genome and any associated plasmids. Additionally, STSS collects information to gauge the likelihood that the self-targeting sequence would be lethal to the organism and if the target sequence occurs in a MGE (fig. S1) (20, 21).

Using STSS, we collected self-targeting data for 150,291 genomes, observing 22,125 cases of predicted self-targets, representing 8,917 unique sequences across 9,155 genomes (fig. S2 and data S1). Focusing initially on three species in which multiple Acrs have been previously identified (Pseudomonas aeruginosa, Listeria monocytogenes, and Neisseria meningitidis), we determined the number of genomes that contained at least one lethal self-targeting CRISPR spacer, and the number of those genomes that also contained an Acr using a blastp search (Fig. 1C). In N. meningitidis only 6% of the genomes were observed to contain a potential anti-CRISPR, while in P. aeruginosa and L. monocytogenes the number exceeded 80% or 90%. The self-targeting genomes devoid of known Acrs may also contain inhibitors that have yet to be discovered, especially in N. meningitidis where the number of self-targeting genomes is high but the number containing known Acrs is low.

Based on our observation that self-targeting genomes frequently contain CRISPR inhibitors, we sought to determine whether screening genes in genomes containing self-targeting spacers could uncover new inhibitors. We focused our efforts on the CRISPR-Cas12a system (22–24), which has so far eluded discovery of inhibitory proteins despite the increasing use of Cas12a in gene editing and diagnostic applications (22, 25, 26). From the STSS results, we identified four strains of Moraxella bovoculi that contain self-targeting CRISPR-Cas12a systems as top candidates for containing anti-CRISPRs (see Methods). These strains each contained at least

---

1Department of Molecular and Cell Biology, University of California, Berkeley, Berkeley, CA 94720, USA. 2Gladstone Institutes, San Francisco, CA 94158, USA. 3Department of Chemistry, University of California, Berkeley, Berkeley, CA 94720, USA. 4Innovative Genomics Institute, University of California, Berkeley, CA 94720, USA. 5Howard Hughes Medical Institute, University of California, Berkeley, Berkeley, CA 94720, USA. 6Li Ka Shing Biomedical and Health Sciences Center, University of California, Berkeley, Berkeley, CA 94720, USA. 7Department of Molecular Biology, University of California, Berkeley, Berkeley, CA 94720, USA.

*Corresponding author. Email: doudna@berkeley.edu

Cite as: K. E. Watters et al., Science 10.1126/science.aau5138 (2018).
one perfectly matched self-targeting sequence in or near a predicted MGE with a correct TTV PAM sequence and intact Cas12 open reading frame, which should render the self-targeting spacers lethal in the absence of anti-CRISPRs (Fig. 1D and fig. S3) (27).

To test whether the *Moraxella* genomes encode type V-A anti-CRISPRs (AcrVA), we employed a cell-free transcription-translation (TXTL) system (28, 29) to express gene products from *Moraxella* genomic fragments. As an initial test of *M. bovoculi* Cas12a (MbCas12a) protein activity, we PCR-amplified a genomic fragment containing the promoter region and all of the Cas proteins (Cas12a, Cas1, Cas2, and Cas4) from *M. bovoculi* strain 22581. This amplicon was added to TXTL reactions with two reporter plasmids encoding green or red fluorescent protein (GFP or RFP) (Fig. 2A). When supplied with CRISPR RNAs (crRNAs) with base pairing complementary to the GFP and RFP genes, the presence of the MbCas12a-containing genomic fragment greatly reduced expression of both reporters (Fig. 2B and fig. S4). This result suggests that MbCas12 is active in *M. bovoculi*, suggesting the existence of CRISPR inhibitor(s) to prevent the self-targeting spacers from killing the cell.

To identify potential AcrVA-encoding genes, we used a directed screening approach to search the predicted MGEs within three of the *M. bovoculi* strains (strain 283689 was unavailable) containing self-targeting sequences from a type V-A CRISPR array. Interestingly, we also observed 13 self-targeting CRISPR type I-C spacers in strain 58069 that strongly suggest the presence of I-C anti-CRISPRs in that strain (fig. S5). For each of the *M. bovoculi* genomes, pairs of PCR primers were used to make overlapping ~2-10 kb amplicons spanning all of the predicted MGEs in the three strains (generally excluding highly similar sequences) (table S1). These genomic fragments (GFs) were then added to the TXTL cleavage reactions described above.

From a total of 67 GFs that we tested for type V-A CRISPR inhibition activity, four correlated with increased levels of gene expression for both reporters (Fig. 2, C and D, and figs. S6 to S8). We then cloned the individual open reading frames within these fragments (Fig. 2D and table S3) downstream of the Ptet promoter to separately induce transcription and translation of each gene and assessed them for CRISPR inhibition activity using the TXTL Cas12a cleavage assay. From the pool of candidates, three proteins supported high levels of dual reporter gene expression (Fig. 2, E and F, and fig. S9): GF36 candidate 1, GF59 candidate 2, and GF59 candidate 3, hereafter referred to as AcrVA1, AcrVA4, and AcrVA5, respectively, to complement the other AcrVA genes discovered concurrently with this work (30).

To confirm the CRISPR inhibition activity of AcrVA1, AcrVA4, and AcrVA5 in vitro, we first overexpressed and purified each putative Cas12a inhibitor, MbCas12a, and two additional Cas12a enzymes (AsCas12a and LbCas12a) commonly used in genome editing or DNA detection applications (fig. S10) (25, 26). We then generated crRNA–protein (RNP) complexes for each of the purified Cas12a enzymes and added the RNPs to a linearized target plasmid that was preincubated with increasing concentrations of each candidate AcrVA protein. We observed that AcrVA1 inhibited DNA cleavage by all three Cas12a enzymes, with the strongest inhibition observed for MbCas12a and weakest observed for AsCas12a (Fig. 3). AcrVA4 and AcrVA5 inhibited dsDNA cleavage for both MbCas12a and LbCas12a, but did not inhibit AsCas12a. Interestingly, we also observed that AcrVA4 more strongly inhibited the MbCas12a from strain 58069 (fig. S11) than the MbCas12a from strain 22581 (Fig. 3A), and that AcrVA5 was unable to inhibit the MbCas12a from strain 58069 (fig. S11). None of the AcrVA proteins inhibited *S. pyogenes* Cas9 (SpyCas9) cleavage (fig. S12).

Having confirmed robust DNA cleavage inhibition by AcrVA1, AcrVA4, and AcrVA5 using purified protein samples, we next tested whether these Cas12a inhibitors could block or reduce Cas12a-mediated genome editing in human cells. We cloned each AcrVA candidate, AcrIIA4 (a SpyCas9 inhibitor), or negative controls, into a lentiviral expression vector and stably transduced HEK293T-derived genome editing reporter cells containing a doxycycline-inducible GFP marker. Purified AsCas12a, LbCas12a, MbCas12a, or SpyCas9 protein was assembled with a GFP-targeting guide RNA and transfected into the AcrVA-expressing reporter cell lines (Fig. 4A). At 24 hours post RNP delivery, cells were induced by doxycycline for another 24 hours before quantifying editing efficiency by flow cytometry (Fig. 4B and fig. S13A) and a T7 endonuclease 1 assay (Fig. 4C and fig. S13B). We observed high levels of genome editing induced by SpyCas9 with no inhibition by any of the AcrVA proteins or negative controls, but virtually complete inhibition by AcrIIA4. We also observed strong Cas12a RNP editing inhibition that generally matched the in vitro cleavage results. Mirroring their biochemical behavior, AcrVA1 provided the broadest inhibition of Cas12a and fully blocked AsCas12a with efficiencies comparable to AcrIIA4’s inhibition of SpyCas9. AcrVA4 and AcrVA5 only inhibited LbCas12a. RNP-based delivery of MbCas12a did not edit efficiently enough to determine the effectiveness of the AcrVA genes on its activity (fig. S13), consistent with previous findings (22).

Together, these results establish a new approach for systematic discovery and validation of CRISPR-Cas inhibitors hidden within self-targeting genomes. Importantly, the Cas12a inhibitors revealed by this approach are only found within a few genomes within the NCBI database, with AcrVA4 and AcrVA5 being particularly rare genes, only occurring with each other and thus intractable to an acas-based search approach (fig. S14). While we expect the
extensive set of yet-to-be-analyzed CRISPR self-targeting genomes (data SI) will lead to the discovery of many more Acrs across all CRISPR subtypes, the AcrVs discovered and validated here provide a toolkit for selective Cas12a regulation both in vitro and in mammalian systems, with the potential to advance synthetic biology, CRISPR diagnostics, and therapeutic genome editing.

REFERENCES AND NOTES


SUPPLEMENTARY MATERIALS
www.sciencemag.org/cgi/content/full/science.aau5138/DC1

Materials and Methods
Figs. S1 to S14
Tables S1 to S4
References (31–44)
Data S1

18 June 2018; accepted 23 August 2018
Published online 6 September 2018
10.1126/science.aau5138

ACKNOWLEDGMENTS
We thank Rachel Lew for technical support. David Burstein for Cas protein HMMs, Mary West and the CIRM/QB3 Shared Stem Cell Facility/High-Throughput Screening Facility for flow cytometry equipment support, Adair Borges for discussions on self-targeting in Pseudomonas, and Doudna lab members for thoughtfull discussion. Funding: The authors acknowledge financial support from the Defense Advanced Research Projects Agency (DARPA) (award HR0011-17-2-0043 to J.A.D.), the Paul G. Allen Frontiers Group and the National Science Foundation (MCB-1244557 to J.A.D.). C.F. is supported by a US National Institutes of Health K99/R00 Pathway to Independence Award (K99GM118909, 1R01GM118909) from the National Institute of General Medical Sciences (NIGMS). J.A.D. is an investigator of the Howard Hughes Medical Institute (HHMI), and this study was supported in part by HHMI; J.A.D; Supervision, K.E.W. and J.A.D.; Writing – Review & Editing, K.E.W., C.F., J.A.D.; Funding Acquisition, K.E.W., C.F., and J.A.D.; Competing interests: J.A.D. is a co-founder of Caribou Biosciences, Editas Medicine, Intellia Therapeutics, Scribe Therapeutics, and Mammoth Biosciences, a scientific adviser to Caribou, Intellia, Scribe, Synthego, Metagenomi, Inari, and eFFECTOR Therapeutics, and a director of Driver and Johnson & Johnson. The Regents of the University of California have patents pending for CRISPR related technologies on which the authors are inventors. Data and materials availability: All data are available in the main text or the supplementary materials. Plasmids for expression of AcrVA1, AcrVA4, AcrVA5, and Mbcas12 are available on Addgene (IDs 115653-115664, 115669-115670, and 115795-115797) under a Uniform Biological Material Transfer Agreement. Code for STSS is available on GitHub. M. bovoculi strains are available from the University of Nebraska (strain 58069) or the US Meat Animal Research Center (strains 22581 and 33362) under a Material Transfer Agreement.
Fig. 1. Bioinformatic approach for discovering Acr genes. (A) Anti-CRISPRs allow survival of cells containing self-targeting CRISPR arrays. (B) STSS finds self-targeting CRISPR spacers in genomic DNA, predicts the type of CRISPR system involved, and obtains information about the targeted sequence. (C) A large percentage of genomes containing self-targeting (ST) spacers predicted to be lethal contain previously-identified Acr genes. (D) *Moraxella bovoculi* strain 22581 contains three self-targeting spacers in two different prophages in the genome. All of the protospacers contain a TTV protospacer-adjacent motif (PAM) (22).
Fig. 2. TXTL screening for Acr gene candidates. (A) Overview of the transcription-translation (TXTL) reaction. DNA expressing Cas12a, two fluorescent reporters, and two gRNAs are mixed with or without DNA potentially containing Acr genes. (B) Cleavage of the reporter plasmid results in a reduced fluorescent output that is rescued by Acr genes (Acr-absent data in triplicate). (C) Amount of relative inhibition observed for 67 genomic fragments (GFs) across three self-targeting M. bovoculi strains. Four GFs (bold) exhibited inhibition in both fluorescence channels. (D) Genomic fragments GF29, GF35, GF36, and GF59 (99% nucleotide identity to GF29) exhibited high levels of expression for both reporters. (E) Testing the individual genes from the fragments in (D) (table S2) resulted in the identification AcrVA1 (GF36 candidate 1), AcrVA4 (GF59 candidate 2), and AcrVA5 (GF59 candidate 3). (F) Kinetic TXTL data for the AcrVA genes measured over the course of 10 hours of gene expression.
Fig. 3. Biochemical validation of AcrVA inhibitors. (A) Moraxella bovoculi Cas12a (MbCas12a) in vitro dsDNA cleavage is inhibited by increasing concentrations of AcrVA1, AcrVA4, and AcrVA5 (0 - 1.25 μM; see Methods). (B) LbCas12a, a Cas12a commonly used for gene editing and diagnostics (4, 22, 26), is also inhibited by all three AcrVA proteins. (C) High concentrations of AcrVA1 inhibit AsCas12a-mediated dsDNA cleavage, but AcrVA4 and AcrVA5 have no effect. Triangles indicate uncleaved (black) or cleaved (gray) DNA.

Fig. 4. AcrVAs robustly inhibit genome editing by specific CRISPR-Cas12a nucleases in mammalian cells. (A) Overview of the editing reporter assay in human cells. (B) Quantification of genome editing in reporter cell lines stably expressing the indicated CRISPR-Cas12a inhibitors (AcrVAs) or a control (mTagBFP2, mCherry). The scale of each plot is adjusted to compensate for differences in editing efficiency. Error bars indicate standard deviations of triplicates. (C) Biochemical analysis of AcrVA-mediated inhibition in representative samples shown in (B). Editing was assessed by the T7 endonuclease 1 (T7E1) assay.
Systematic discovery of natural CRISPR-Cas12a inhibitors
Kyle E. Watters, Christof Fellmann, Hua B. Bai, Shawn M. Ren and Jennifer A. Doudna

published online September 6, 2018

http://science.sciencemag.org/content/early/2018/09/05/science.aau5138

http://science.sciencemag.org/content/suppl/2018/09/05/science.aau5138.DC1

This article cites 43 articles, 8 of which you can access for free
http://science.sciencemag.org/content/early/2018/09/05/science.aau5138#BIBL

http://www.sciencemag.org/help/reprints-and-permissions

Use of this article is subject to the Terms of Service