PNAS COMMENTARY

Keeping Cas9 out of the (R-)loop

Daan C. Swarts[a,1]

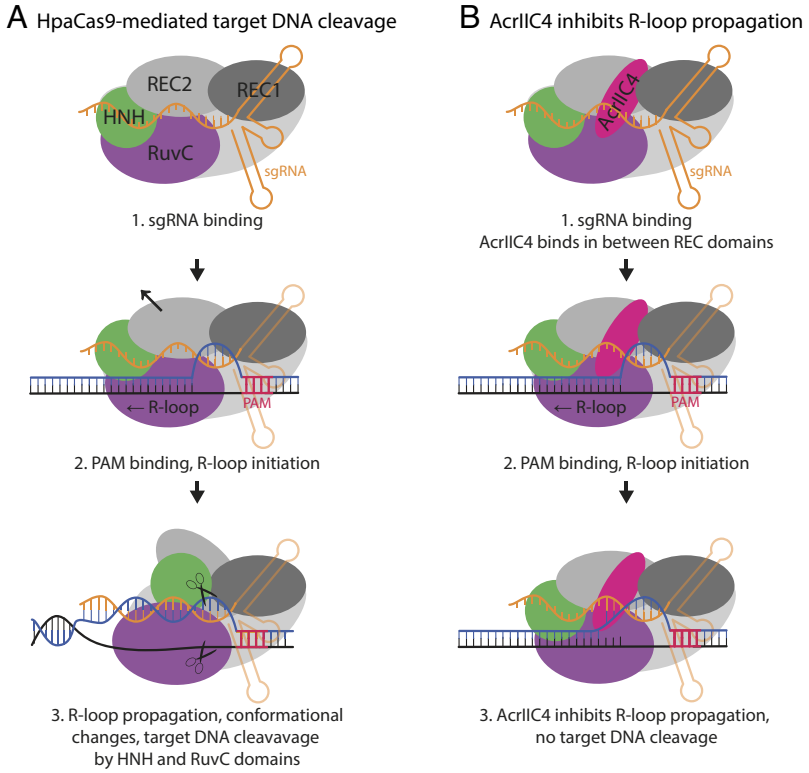

Fig. 1. Simplified model of HpaCas9-mediated target DNA cleavage and inhibition by AcrIIC4. (*A*) 1. sgRNA (orange) binding by HpaCas9. REC1 (dark gray), REC2 (light gray), and nuclease domains HNH (green) and RuvC (purple) are indicated. 2. PAM binding results in the unwinding of the DNA duplex, which allows for base pairing between the sgRNA and the DNA target strand. In this process, the NTS is displaced, forming an R-loop. Conformational changes in the REC domains are required to facilitate DNA unwinding and R-loop propagation. 3. Complete R-loop formation is associated with conformational changes that catalytically activate Cas9, which results in DNA cleavage by the HNH and RuvC domains. (*B*) 1. AcrIIC4 binds to HpaCas9 both in the absence and presence of the sgRNA in a cleft between the REC1 and REC2 domains. It forms interactions with both REC1 and REC2 domains, as well as with the linker between REC1 and REC2 and the sgRNA scaffold (not shown). 2. PAM binding and initial sgRNA-target strand DNA duplex formation is not inhibited by AcrIIC4, but AcrIIC4 restricts conformational changes in the REC2 domain. 3. By restricting REC2 domain movement, AcrIIC4 prevents further DNA duplex unwinding and inhibits complete R-loop formation. Consequentially, the HNH and RuvC domains are not catalytically activated and DNA cleavage is inhibited.

To protect themselves against invading viruses, prokaryotes encode a variety of immune systems (1), including highly diverse CRISPR-Cas systems (2, 3). CRISPR-associated (Cas) effector proteins generally utilize RNA guides to bind DNA targets. In this process, the RNA guide binds the target DNA strand (TS) while the nontarget strand (NTS) is displaced, forming an R-loop structure. Complete R-loop formation generally results in the degradation of targeted DNA or activation of accessory proteins that interfere with virus propagation. The programmable nature of Cas effector proteins (including Cas9 and Cas12a) has facilitated their utilization for applications ranging from nucleic acid detection to genome editing (4, 5). Viruses encode anti-CRISPR (Acr) proteins to evade CRISPR-Cas immunity (6–8). Acr proteins can have different evolutionary origins and allow escape from CRISPR-Cas immunity through distinct mechanisms. Studying Acrs not only provides fundamental insights into the evolutionary arms race between prokaryotes and their viruses but additionally illuminates the key mechanisms of Cas proteins. Furthermore, Acrs can be repurposed for controlling Cas protein activity in genome editing or other biotechnological applications (9). In this issue of PNAS, Sun et al. have studied structures of Cas9 bound to Acr protein AcrIIC4 and thereby uncovered a novel mechanism of Cas9 inhibition: preventing complete R-loop formation (10).

CRISPR-Cas systems are highly diverse and are currently divided into six main types and over 33 subtypes (2). Despite

Author affiliations: [a]Laboratory of Biochemistry, Wageningen University, Wageningen 6708 WE, The Netherlands

Author contributions: D.C.S. wrote the paper.

The author declares no competing interest.

See companion article, "AcrIIC4 inhibits type II-C Cas9 by preventing R-loop formation," 10.1073/pnas.2303675120.

[1]Email: daan.swarts@wur.nl.

Published August 11, 2023.

their high diversity, all CRISPR-Cas systems rely on a similar general mechanism of action (3): CRISPR RNA (crRNA) guides are transcribed from genomic CRISPR loci, which contain sequences from previously encountered viruses. These crRNA sequences guide multiprotein Cas complexes or single multidomain Cas effector proteins to recognize and neutralize viruses and other mobile genetic elements such as plasmids. Type II CRISPSR-Cas systems rely on the multidomain effector protein Cas9 (11). Cas9 utilizes dual-RNA guides composed of a crRNA, which contains the variable segment important for target DNA recognition, and a trans-acting crRNA (tracrRNA) (12). The crRNA and tracrRNA together form several stem-loops that facilitate loading onto Cas9 which enables Cas9-mediated DNA targeting. For genome editing purposes, the dual-RNA structure is often fused to form a single guide RNA (sgRNA) (11). DNA targeting by Cas9 is initiated upon sequence-specific recognition of a protospacer adjacent motif (PAM) downstream of the target DNA sequence. This induces conformational changes that open the DNA duplex, allowing the formation of initial base pairs between the sgRNA and the TS. Further sgRNA-TS DNA hybridization results in the displacement of the NTS and formation of a so-called R-loop structure. In Cas9, R-loop formation is associated with structural rearrangement in the REC domains that together coordinate the crRNA-TS duplex (13). Only upon full R-loop formation, the HNH and RuvC nuclease domains cleave the TS and NTS, respectively.

Studying Acrs not only provides fundamental insights in the evolutionary arms race between prokaryotes and their viruses, but additionally illuminates the key mechanisms of Cas proteins.

Due to its wide application in genome editing, the best characterized Cas9 homolog is the type II-A Cas9 from *Streptococcus pyogenes* (4, 5). However, also type II-C Cas9 homologs have characteristics that make them interesting for genome editing applications: They are generally small (<1,100 amino acids, allowing easier in vivo delivery), they have distinct PAM preferences (expanding the range of targetable sequences), and/or they show low off-target cleavage (reducing undesired off-target editing) (14). As such, type II-C Cas9 homologs provide valuable additions to the genome editing toolbox.

Cas9-mediated DNA cleavage can be inhibited by Acr proteins by preventing crRNA loading, blocking DNA binding, or inhibiting DNA cleavage (8). To date, six different type II-C-specific inhibitors have been identified (AcrIIC1-AcrIIC6): AcrIIC1 and AcrIIC3 inhibit the nuclease activity of HNH, AcrIIC2 prevents crRNA loading, and AcrIIC5 and AcrIIC6 inhibit DNA binding (8). The mode of action and structural mechanism by which AcrIIC4 inhibits Cas9 were not completely clear, as some ambiguity was observed in structures and inhibitory effects of AcrIIC4 (15–17).

AcrIIC4 is a 88 amino acid Acr encoded by a putative prophage, a virus integrated into the genome of *Haemophilus parainfluenzae* (15). *H. parainfluenzae* also encodes a type II-C Cas9 (HpaCas9) which recognizes a 5′-N_4GNTT-3′ PAM and can be used for genome editing in mammalian cells (15). In its initial characterization study, it was demonstrated that ActIIC4 inhibits HpaCas9-mediated antiviral defense and that it interferes with DNA binding and cleavage but not crRNA binding (15). In pursuit of uncovering the mechanism underlying AcrIIC4-mediated Cas9 inhibition, two additional studies focusing on the structure and biochemistry of AcrIIC4 were published (16, 17). However, these studies showed ambiguous results in terms of structural conformation and mode of action. Kim et al. revealed a monomeric and compact AcrIIC4 conformation in which it forms four short α-helices (16). Based on the observation that it interferes with DNA binding, it was hypothesized that AcrIIC4 blocks DNA-binding sites in Cas9 (16). In contrast, Hwang et al. revealed a homodimeric elongated conformation AcrIIC4, in which each AcrIIC4 monomer forms one long and two short α-helices (17). In contrast to the two other studies, Hwang et al. revealed that AcrIIC4 does not strictly inhibit DNA binding. Based on the observation that AcrIIC4 interacts with the REC1 and REC2 domains, it was hypothesized that AcrIIC4 prevents HNH movement, thereby inhibiting target cleavage (17). Yet, the exact molecular mechanism by which AcrIIC4 inhibits Cas9 remained unclear.

Sun et al. have now resolved structures of free AcrIIC4, the HpaCas9–sgRNA-DNA complex in the absence of AcrIIC4, and two structures of AcrIIC4-bound HpaCas9–sgRNA complexes in the presence and absence of target DNA (10). Based on their studies, Sun et al. not only explain previously made observations but reveal the molecular mechanism of AcrIIC4-based Cas9 inhibition (Fig. 1*B*). Similar to the structure by Hwang et al., Sun et al. report an elongated homodimeric AcrIIC4 conformation in the absence of HpaCas9. However, it is shown that soluble AcrIIC4 is monomeric and when bound to HpaCas9 monomeric AcrIIC4 assumes the more compact conformation similar to the conformation observed by Kim et al. Explaining the ambiguity in the previously reported structural conformations, Sun et al. reveal that the long central alpha helix in the elongated structure of AcrIIC4 bends to facilitate AcrIIC4 interactions with HpaCas9 (10).

The structures of AcrIIC4 bound to HpaCas9–sgRNA in the presence and absence of target DNA confirm that AcrIIC4 indeed does not prevent sgRNA or target DNA binding (10). Instead, it forms long-range interactions within the REC lobe and the gRNA scaffold, thereby restricting REC2 movement. In the absence of AcrIIC4, crRNA-TS duplex formation is associated with conformational changes in the REC2 and HNH domains. These conformational changes facilitate the formation of the full R-loop in which the sgRNA/TS duplex spans 24 base pairs. In contrast, in the presence of AcrIIC4, REC2 movement is impaired. This stalls sgRNA/TS duplex formation after seven base pairs, preventing formation of a full R-loop. By providing mismatched dsDNA targets (resembling pre-unwound DNA substrates), the authors show that the DNA can be cleaved by HpaCas9 even in the presence of AcrIIC4. This underscores the importance of REC2 flexibility for complete R-loop formation and consequential target cleavage.

In conclusion, the study by Sun et al. provides the basis for the different AcrIIC4 conformations observed in previous

studies, reveals how AcrIIC4 reduces the affinity of Cas9 towards target DNA without completely inhibiting DNA binding, and provides a molecular basis for AcrIIC4-mediated inhibition of Cas9. The study highlights the importance of REC domain movement in Cas9-mediated target cleavage and presents a novel mechanism by which Acrs can inhibit CRISPR-Cas-mediated immunity: keeping Cas9 out of the (R-)loop.

ACKNOWLEDGMENTS. I apologize to the colleagues whose relevant work cannot be highlighted here due to space limitations. I thank Martin Pacesa for assisting me in the preparation of the manuscript.

1. H. Georjon, A. Bernheim, The highly diverse antiphage defence systems of bacteria. *Nat. Rev. Microbiol.*, 10.1038/s41579-023-00934-x (2023).
2. K. S. Makarova *et al.*, Evolutionary classification of CRISPR–Cas systems: A burst of class 2 and derived variants. *Nat. Rev. Microbiol.* **18**, 67–83 (2020).
3. P. Mohanraju *et al.*, Diverse evolutionary roots and mechanistic variations of the CRISPR-Cas systems. *Science* **353**, aad5147 (2016).
4. J. Y. Wang, P. Pausch, J. A. Doudna, Structural biology of CRISPR–Cas immunity and genome editing enzymes. *Nat. Rev. Microbiol.* **20**, 641–656 (2022).
5. J. P. K. Bravo, G. N. Hibshman, D. W. Taylor, Constructing next-generation CRISPR–Cas tools from structural blueprints. *Curr. Opin. Biotechnol.* **78**, 102839 (2022).
6. A. R. Davidson *et al.*, Anti-CRISPRs: Protein inhibitors of CRISPR-Cas systems. *Annu. Rev. Biochem.* **89**, 309–332 (2020).
7. Y. Li, J. Bondy-Denomy, Anti-CRISPRs go viral: The infection biology of CRISPR-Cas inhibitors. *Cell Host Microbe* **29**, 704–714 (2021).
8. S. Hwang, K. L. Maxwell, Diverse mechanisms of CRISPR-Cas9 inhibition by type II anti-CRISPR proteins. *J. Mol. Biol.* **435**, 168041 (2023).
9. C. Kraus, E. J. Sontheimer, Applications of anti-CRISPR proteins in genome editing and biotechnology. *J. Mol. Biol.* **435**, 168120 (2023).
10. W. Sun *et al.*, AcrIIC4 inhibits type II-C Cas9 by preventing R-loop formation. *Proc. Natl. Acad. Sci. U.S.A.* **120**, e2303675120 (2023).
11. M. Jinek *et al.*, A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* **337**, 816–821 (2012).
12. E. Deltcheva *et al.*, CRISPR RNA maturation by trans-encoded small RNA and host factor RNase III. *Nature* **471**, 602–607 (2011).
13. M. Pacesa *et al.*, R-loop formation and conformational activation mechanisms of Cas9. *Nature* **609**, 191–196 (2022).
14. A. Mir, A. Edraki, J. Lee, E. J. Sontheimer, Type II-C CRISPR-Cas9 biology, mechanism, and application. *ACS Chem. Biol.* **13**, 357–365 (2018).
15. J. Lee *et al.*, Potent cas9 inhibition in bacterial and human cells by AcrIIC4 and AcrIIC5 anti-CRISPR proteins. *mBio* **9**, e02321-18 (2018).
16. G. E. Kim, S. Y. Lee, H. H. Park, Crystal structure of the anti-CRISPR, AcrIIC4. *Protein Sci.* **30**, 2474–2481 (2021).
17. S. Hwang *et al.*, Structural and mechanistic insight into CRISPR-Cas9 inhibition by anti-cRISPR protein AcrIIC4Hpa. *J. Mol. Biol.* **434**, 167420 (2022).