

A CRISPR evolutionary arms race: structural insights into viral anti-CRISPR/Cas responses

Cell Research advance online publication 2 September 2016; doi: 10.1038/cr.2016.103

Dear Editor,

CRISPR/Cas systems are widespread RNA-mediated prokaryotic adaptive immune systems providing protection against invading nucleic acids [1]. However, throughout evolution, this host defense system has not resulted in the eradication of phages, suggesting that phages have evolved counter strategies to thrive within bacteria despite these mechanisms [2]. Thus, both bacterial CRISPR system and phage anti-CRISPR system are part of a continuing evolutionary battle between bacterial host and their bacteriophage invaders.

The fight against phages through CRISPR/Cas systems occurs in three stages. For example, for the type I-F CRISPR/Cas system, Cas1 and Cas2 proteins first insert a short segment 'spacer' derived from the foreign DNAs into the CRISPR locus. Subsequently, the CRISPR array is transcribed into a long precursor crRNA, which is further processed into the mature crRNA by Csy4 protein. Finally, this stable Csy4-crRNA complex subsequently assembles with Csy1, Cys2 and Csy3 protein, forming a crRNA-guided surveillance complex (named Csy complex) (Figure 1A). The Csy complex specifically recognizes the complementary target DNA and recruits Cas3 to degrade the invading DNA [3].

To proliferate inside their bacterial hosts, phages must overcome bacterial defense mechanisms, such as the CRISPR/Cas immune system. It was shown earlier that within a population, certain phages display a mutation in the protospacer or in the protospacer adjacent motif (PAM), thus evading CRISPR targeting. Recently, an alternative approach for evading CRISPR/Cas activity, which lacks a requirement for mutation in the protospacer or PAM region [4], was described for *Pseudomonas aeruginosa*. The anti-CRISPR prophages contain a set of atypical genes encoding five small phage proteins involved in inactivating the CRISPR/Cas system. Subsequently, another study characterized three of these anti-CRISPR proteins (AcrF) [5]. AcrF1 and AcrF2 directly bind to the Csy complex, thus blocking target DNA binding. In comparison, AcrF3 interacts directly with the

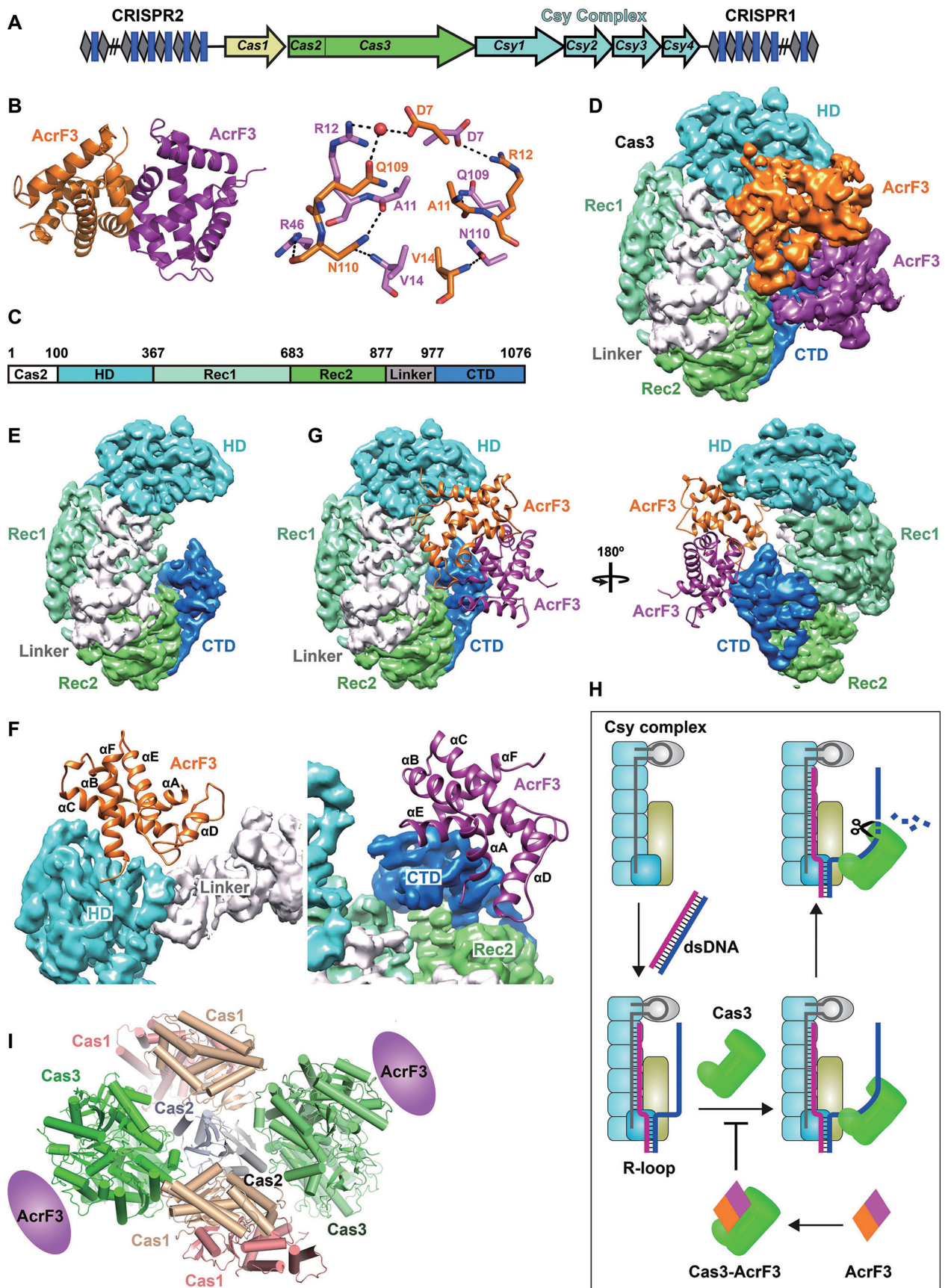
Cas3 protein and blocks its recruitment into Csy complex, thus protecting the phage DNA from being degraded by the CRISPR/Cas system.

While previous studies have provided crucial first insights into the molecular mechanisms of anti-CRISPR activities, it remains to be elucidated as to how the AcrF3 inhibits crRNA interference. To find answers to this question, we set out to solve the crystal structure of AcrF3 and the cryo-electron microscopy (cryo-EM) structure of AcrF3-Cas3 complex. Our structural analysis provides important insights into how the AcrF3 dimer binds Cas3 protein at the entrance of non-complementary strand into Cas3. As a result, recruitment of Cas3 by the Csy complex is inhibited, thus effectively preventing the degradation of phage DNA by the CRISPR/Cas system.

To reveal the molecular details of how AcrF3 inhibits CRISPR/Cas activity, we performed crystal structure analysis of the cloned and purified AcrF3 protein encoded by gene 35 from phage JBD5. The structure of the seleno-methionine substituted AcrF3 was solved to 1.5 Å resolution (Supplementary information, Table S1). As shown in Figure 1B, the AcrF3 is a dimer, with each monomer comprising six α -helices. Dimer formation is facilitated by the interaction between the first α -helices of two AcrF3 subunits (Figure 1B).

To understand how the AcrF3 protein interacts with type I-F *P. aeruginosa* Cas3 (Figure 1C), we prepared the *Pae*Cas3-AcrF3 complex (Supplementary information, Figure S1A and S1B), and solved the cryo-EM structure of *Pae*Cas3-AcrF3 complex at resolution of 4.2 Å using cryo-EM single particle method (Figure 1D). In the 4.2 Å density map, the secondary structures such as α -helix and β -strand were resolved (Supplementary information, Figure S1C-S1F). A model of *Pae*Cas3 was built based on the electron density map and the crystal structure of AcrF3 dimer fits well to the density of the rest of the map. However, due to the flexibility, certain loops were not resolved and remained unidentified in the map.

Similar to *Thermobifida fusca* Cas3 (*Tfu*Cas3) [6] and *Thermobaculum terrenum* Cas3 (*Tte*Cas3) [7], *Pae*Cas3 also contains five conserved domains: HD, Rec1, Rec2,



Linker and CTD, and the helicase domain forms a cleft to accommodate single-stranded DNA (Figure 1E). In order to analyze the conformational change of Cas3 induced by AcrF3 binding, we performed a structural comparison of *PaeCas3*, *TfuCas3* and *TteCas3*. As shown in Supplementary information, Figure S1G, the structures of the Rec1 and Rec2 domains are well conserved, despite their low sequence similarity. The topology of Rec1 and Rec2 are nearly identical for *PaeCas3*, *TfuCas3* and *TteCas3*, except for the presence of two additional α helices in *TfuCas3* and *TteCas3*. In contrast, the HD and CTD domains exhibit significant differences among *PaeCas3*, *TfuCas3* and *TteCas3*.

As shown in Figure 1D, the AcrF3 is present as a dimer in the Cas3-AcrF3 complex. One AcrF3 monomer interacts with the HD and Linker of Cas3 protein, and the other monomer interacts with the CTD domain (Figure 1F). Thus, AcrF3 dimer connects the HD and CTD domains of Cas3, and covers the cleft formed inside Cas3 protein (Figure 1G). The Cas3 protein is recruited by the Csy complex following pairing of the crRNA bases with those of the complementary strand, generating an R-loop conformation with a displaced non-complementary strand. Interestingly, the AcrF3 dimer binding site of Cas3 is the entrance of the displaced single-stranded DNA associated with Cas3, suggesting that the AcrF3 dimer blocks the access to Cas3 of the non-target DNA (Figure 1H). Without Cas3 recruitment by the Csy-dsDNA complex, the CRISPR/Cas system is unable to efficiently destroy the invading phage DNA, resulting in the escape of phages from the immune response.

As shown in Figure 1F the loop between α_6 and α_7 , and the loop between α_5 and α_6 of HD domain interact with one AcrF3. In the *PaeCas3* CTD domain, the loop between α_{36} and β_{22} , and the loop between β_{19} and β_{20} interact with the other AcrF3 monomer. In order to investigate whether AcrF3 is able to bind other Cas3, we aligned the *PaeCas3*, *TfuCas3* and *TteCas3* protein sequences. As shown in Supplementary information, Figure S1H, the Rec1 and Rec2 sequences are conserved, whereas the regions of Cas3 interacting with AcrF3

(highlighted by a blue box in Supplementary information, Figure S1H) are not conserved, suggesting that AcrF3 cannot bind *TfuCas3* and *TteCas3* proteins. This is in agreement with earlier result showing that the AcrF3 protein is highly specific for the *Pseudomonas* I-F CRISPR system [4].

Cas3 plays critical roles during the process of primed spacer acquisition. When crRNA interference is unsuccessful due to the phage mutation, some infected bacterial host cells survive by acquiring new spacer sequences directly from the invading phage genome [8]. In addition to Cas1 and Cas2, Csy complex and the Cas3 nuclease are involved during the primed acquisition. Recently, it was found that the Cas1-Cas2 complex associates with the Cas3 protein during primed acquisition [9]. The degradation products of Cas3 are used by Cas1-Cas2 as precursors for new spacers.

In the type I-F CRISPR system, Cas2 and Cas3 are coded by one gene as a fused protein, with Cas2 located at the N-terminal of Cas3. During the preparation of our manuscript, a similar crystal structure of AcrF3 bound with Cas3 complex was published [10] (Supplementary information, Figure S1I). While Cas2 was clearly observed in this reported AcrF3-Cas3 structure, Cas2 is disordered in our cryo-EM complex structure due to the flexibility of this region. To reveal whether the presence of AcrF3 protein affects spacer acquisition and to gain knowledge of the Cas1-Cas2-Cas3 complex, we made the model of Cas1-Cas2-Cas3 complex (Figure 1I) by superposing the Cas2 from structure 5B7I and the Cas2 protein in the Cas1-Cas2 complex [11]. The Cas1-Cas2-Cas3 complex contains two Cas1 dimers, one Cas2 dimer and two Cas3 proteins, with Cas2 dimer sandwiched by two Cas1 dimers and two Cas3 proteins from orthogonal direction. Cas3 associates with Cas1-Cas2 complex by the interaction between its Rec1 domain and Cas2. The AcrF3 dimer binding surface is located on the opposite site of Cas2, suggesting that AcrF3 has little effect on the assembly of the Cas1-Cas2-Cas3 complex. Thus is unlikely to interfere with primed acquisition directly. However, AcrF3 inhibits Cas3 recruitment by Csy-dsD-

Figure 1 Viral AcrF3 dimer binds to *PaeCas3* protein, inactivating the bacterial CRISPR system. **(A)** Graphic representation of the *P. aeruginosa* CRISPR/Cas locus. The CRISPR locus consists of series of repeats (gray diamonds) separated by spacer sequences (blue rectangles). Four *Csy* genes are shown in cyan, *Cas1* in yellow and the fused *Cas2-Cas3* in green. **(B)** Overall crystal structure of AcrF3 (left panel) and the interactions between two monomers (right panel). **(C)** Domain organization of *PaeCas3*. **(D)** Architecture of the *PaeCas3*-AcrF3 complex. Cryo-EM density map of *PaeCas3*-AcrF3 complex at 4.2 Å resolution was segmented and colored according to subunit and domains. **(E)** Segmented *PaeCas3* with domains color-coded. **(F)** Close-up view of interactions between *PaeCas3* and AcrF3 dimer. One AcrF3 monomer interacts with HD domain and Linker of Cas3 (left panel); and the other AcrF3 monomer interacts with CTD domain of Cas3 (right panel). **(G)** AcrF3 dimer blocks the replaced non-complementary DNA access of Cas3. AcrF3 is shown as cartoon. **(H)** Model of AcrF3 inactivating CRISPR/Cas immune system. **(I)** Model of Cas1-Cas2-Cas3 complex.

NA complex, thus stopping the generation of the precursor protospacer DNA. As a result, phage AcrF3 protein blocks both crRNA interference and spacer acquisition.

In conclusion, our structural analysis of the Cas3-AcrF3 complex revealed that the AcrF3 dimer binds to HD, Linker and CTD domains of the Cas3 protein, therefore blocking the replaced non-complementary DNA access of Cas3. Therefore, AcrF3 inhibits crRNA-mediated DNA degradation via inhibiting recruitment of Cas3 by Csy complex. In addition, inhibition of Cas3-guided DNA degradation also blocks primed spacer acquisition. In summary, the AcrF3 dimer overcomes the CRISPR system at both the crRNA interference and the primed acquisition stages. Our studies provide new insights into the ongoing molecular arms race between viral parasites and the immune systems of their hosts.

Acknowledgments

We thank the staffs from BL19U1 beamline at Shanghai Synchrotron Radiation Facility. We thank X J Huang, G Ji, W Ding, F Sun and other staff members at the Center for Biological Imaging (IBP, CAS) for the assistance of data collection. This work was supported by grants from the National Natural Science Foundation of China (91440201, 31571335, 31400640 and 31570874), the Chinese Ministry of Science and Technology (2014CB910102) and the Strategic Priority Research program of the Chinese Academy of Sciences (XDB08010203). XZ received scholarships from the 'National Thousand (Young) Talents Program' from the Office of Global Experts Recruitment in China. The atomic coordinates of the AcrF3 and *Pae*Cas3-AcrF3 have been deposited in the Protein Data Bank with accession numbers 5GNF and 5GQH. The cryo-EM map has been deposited in the Electron Microscopy Data Bank under accession code EMD-9535.

Jiuyu Wang^{1,2,*}, Jun Ma^{3,4,*}, Zhi Cheng^{1,2}, Xu Meng^{1,5}, Lilan You^{1,2,6}, Min Wang^{1,2}, Xinzheng Zhang^{3,4}, Yanli Wang^{1,2,6}

¹Key Laboratory of RNA Biology, Beijing 100101, China; ²Beijing Key Laboratory of Noncoding RNA, Beijing 100101, China; ³National Laboratory of Biomacromolecules, Beijing 100101, China; ⁴Center for Biological Imaging, CAS Center

for Excellence in Biomacromolecules, Institute of Biophysics, Chinese Academy of Sciences, Beijing 100101, China; ⁵Institute of Life Sciences, Jiangsu University, Zhenjiang, Jiangsu 212013, China; ⁶University of Chinese Academy of Sciences, Beijing 100049, China

*These two authors contributed equally to this work.

Correspondence: Yanli Wang^a, Xinzheng Zhang^b

^aE-mail: ylwang@ibp.ac.cn

^bE-mail: xzzhang@ibp.ac.cn

References

- 1 Barrangou R, Fremaux C, Deveau H, *et al.* *Science* 2007; **315**:1709-1712.
- 2 Samson JE, Magadan AH, Sabri M, *et al.* *Nat Rev Microbiol* 2013; **11**:675-687.
- 3 van der Oost J, Westra ER, Jackson RN, *et al.* *Nat Rev Microbiol* 2014; **12**:479-492.
- 4 Bondy-Denomy J, Pawluk A, Maxwell KL, *et al.* *Nature* 2013; **493**:429-432.
- 5 Bondy-Denomy J, Garcia B, Strum S, *et al.* *Nature* 2015; **526**:136-139.
- 6 Huo Y, Nam KH, Ding F, *et al.* *Nat Struct Mol Biol* 2014; **21**:771-777.
- 7 Gong B, Shin M, Sun J, *et al.* *Proc Natl Acad Sci USA* 2014; **111**:16359-16364.
- 8 Richter C, Chang JT, Fineran PC. *Viruses* 2012; **4**:2291-2311.
- 9 Richter C, Dy RL, McKenzie RE, *et al.* *Nucleic Acids Res* 2014; **42**:8516-8526.
- 10 Wang X, Yao D, Xu J, *et al.* *Nat Struct Mol Biol* 2016 Jul 25. doi:10.1038/nsmb.3269
- 11 Wang J, Li J, Zhao H, *et al.* *Cell* 2015; **163**:840-853.

(Supplementary information is linked to the online version of the paper on the *Cell Research* website.)



This license allows readers to copy, distribute and transmit the Contribution as long as it attributed back to the author. Readers are permitted to alter, transform or build upon the Contribution as long as the resulting work is then distributed under this is a similar license. Readers are not permitted to use the Contribution for commercial purposes. Please read the full license for further details at - <http://creativecommons.org/licenses/by-nc-sa/4.0/>

© The Author(s) 2016