



LETTER TO THE EDITOR

Cryo-EM structure of Type III-A CRISPR effector complex

Cell Research (2018) 28:1195–1197; <https://doi.org/10.1038/s41422-018-0115-6>

Dear Editor,

In bacteria and archaea, Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR) and CRISPR-associated (Cas) proteins constitute a crRNA-guided surveillance complex that defends against invading nucleic elements.¹ The CRISPR-Cas system can be divided into six types (Types I–VI), which differ in terms of their nucleic acid target, number and arrangement of *cas* genes, and the composition of a silencing complex. The Type III CRISPR-Cas system can be further classified into two main subtypes: Type III-A and Type III-B. The CRISPR-Cas effector complex of the Type III CRISPR-Cas system consists of five subunits (Csm1–5) in Type III-A (Fig. 1a) and six subunits (Cmr1–6) in Type III-B, both binding a crRNA composed of a typical repeat-derived 8-nucleotide (nt) 5'-handle followed by a long variable RNA-guide region. In contrast to Type I and II effectors that target dsDNA, Type III effectors are sequence-specific RNases that cleave target RNA and also target RNA-stimulated nonspecific DNases.² To date, although atomic resolution structures of Type III-B effector complexes have been obtained,^{3,4} only low resolution electron microscopy (EM) structures of Csm effector complexes from different species have been reported, including *S. solfataricus* Csm complex (SsCsm),⁵ *T. thermophilus* Csm complex (TtCsm),⁶ and *T. onnurineus* complex (ToCsm).⁷ These structures reveal the overall architectural features of the Csm complex, however, the precise mechanisms of the Type III-A complex assembly and action remain elusive because of the lack of the atomic resolution structure. Recently, Type III effector complexes were demonstrated to synthesize a novel second messenger, cyclic oligonucleotide (tetra-adenylate/hexa-adenylate, cOA₄/cOA₆), for activating the nonspecific RNA degradation activity of Csm6/Csx1.^{8,9} The molecule synthesized by Csm1 requires the binding of noncomplementary target RNA to the Csm effector complex, but the mechanism involved in this process is still unclear. Moreover, for the Type III effector complex, it remains unclear how the 3'-flanking sequence of target RNA controls the ssDNA cleavage.¹⁰ To obtain a better understanding of the assembly and molecular mechanisms of the Type III-A complex, we purified the Csm effector complex of *T. onnurineus* (ToCsm effector complex) expressed in *E. coli*, performed single particle Cryo-EM studies and solved its structure at a resolution of 3.35 Å (Supplementary information, Figs. S1–S3 and Table S1).

The overall structure of the ToCsm effector complex is arranged in a 'boot' shape with the stoichiometry of Csm_{1,2,3,2,4,5}:crRNA, significantly smaller than previously reported Type III-A effectors (Fig. 1b, c; Supplementary information, Figs. S4, S5).^{5,6} Csm1, the largest subunit is located at the body of the boot and interacts with Csm4 and Csm2. The C-terminal fragment of Csm1 forms a six-helix bundle, interacting with Csm2, which mainly consists of an eight-helix bundle. Above Csm4, two Csm3 subunits (termed Csm3.1 and Csm3.2) and Csm5 are arranged along the tube of the boot, forming nearly a double-helical backbone. The thumb-like β -hairpin of Csm4 stretches out and inserts into Csm3.1, and a similar thumb-like β -hairpin of Csm3.1/Csm3.2 stretches out and inserts into the next adjacent subunit. According to the Cryo-EM density map, a 24-nt crRNA single strand can be constructed running

throughout the whole complex, with its 5'-end bound to Csm4 and 3'-end bound to Csm5, respectively. Notably, some weak densities outside Csm5 and adjacent to the 3'-end of the constructed crRNA can be observed. To investigate the nature of the crRNA, we extracted the crRNA from the effector complex used for Cryo-EM analysis and analyzed its length through denaturing Urea-PAGE. As shown in Supplementary information, Fig. S1c, two adjacent bands of crRNA in length of approximate 30 nt are observed, both of which are far shorter than the predicted 45 nt crRNA from the CRISPR locus cloned. Therefore, the weak densities outside Csm5 could be explained as several flexible nucleotides right after the 24th nt of the crRNA. It is possible that the ~30 nt crRNA is a result of aberrant crRNA processing induced by heterologous expression,¹¹ however, the purified ToCsm effector complex used for Cryo-EM study showed clear target RNA cleavage activity and the ability to synthesize cyclic oligoadenylates in vitro (Supplementary information, Figs. S6, S9), indicating that our structure represents an active complex. Csm4 binds to the 5'-handle of the crRNA with its thumb region interacting with the 8th nt, and the thumbs of two Csm3 subunits induce nucleotide protrusions from the backbone at the 14th and the 20th nts, resulting in 6-nt intervals consistent with the canonical mode of crRNA interaction of the Type III-B Cmr complex.⁴ This indicates that Csm3 acts as an endoribonuclease to degrade target RNA in a regular 6-nt cleavage pattern² (Fig. 1b, c; Supplementary information, Fig. S4). In different Type III CRISPR effectors, multiple copies of Csm3 form a backbone that spans the length of the crRNA.^{3–5} To the best of our knowledge, our structure is the smallest Csm effector complex reported to date (Supplementary information, Fig. S5). This may be beneficial for modifying ToCsm effector complex into an effective gene-editing tool.

Csm1 is composed of N-terminal HD, B, zinc finger, two Palm and C-terminal D domains (Supplementary information, Fig. S7a). The two Palm domains of Csm1, are positioned facing each other and both contain a typical ferredoxin fold, showing the classical $\beta\alpha\beta\alpha\beta$ topology. The Palm domains were identified as the active pocket for producing cyclic oligoadenylates in several species.^{8,9,12} Interestingly, we observed an ATP molecule bound to the GGDD (residues 586–GGDD–589) motif in the Palm2 domain despite no ATP being added during the purification of the ToCsm effector complex (Supplementary information, Fig. S7b). Furthermore, we solved the crystal structure of Csm1 in complex with ATP at a resolution of 1.69 Å. In the crystal structure, two ATP molecules are observed bound to the pocket that is contributed by two Palms (ToCsm1_{crys}–2ATP) (Supplementary information, Fig. S7a and Table S2). One interacts with the GGDD loop, while the other interacts with a similar short loop (residues 292–AGGH–295) in the Palm1 domain (Fig. 1d). Comparison of our Cryo-EM structure with the ToCsm1_{crys}–2ATP structure and the previously reported apo-Csm1 crystal structure shows that the binding of two ATP molecules would result in slight conformational changes of the two loops (Fig. 1d).¹³ The orientation of ATP molecules present in the structure is compatible with the attack of 3'-OH of the ribose of the Palm1-bound ATP towards the α -P atom of the

Received: 29 September 2018 Revised: 29 October 2018 Accepted: 2 November 2018
Published online: 21 November 2018

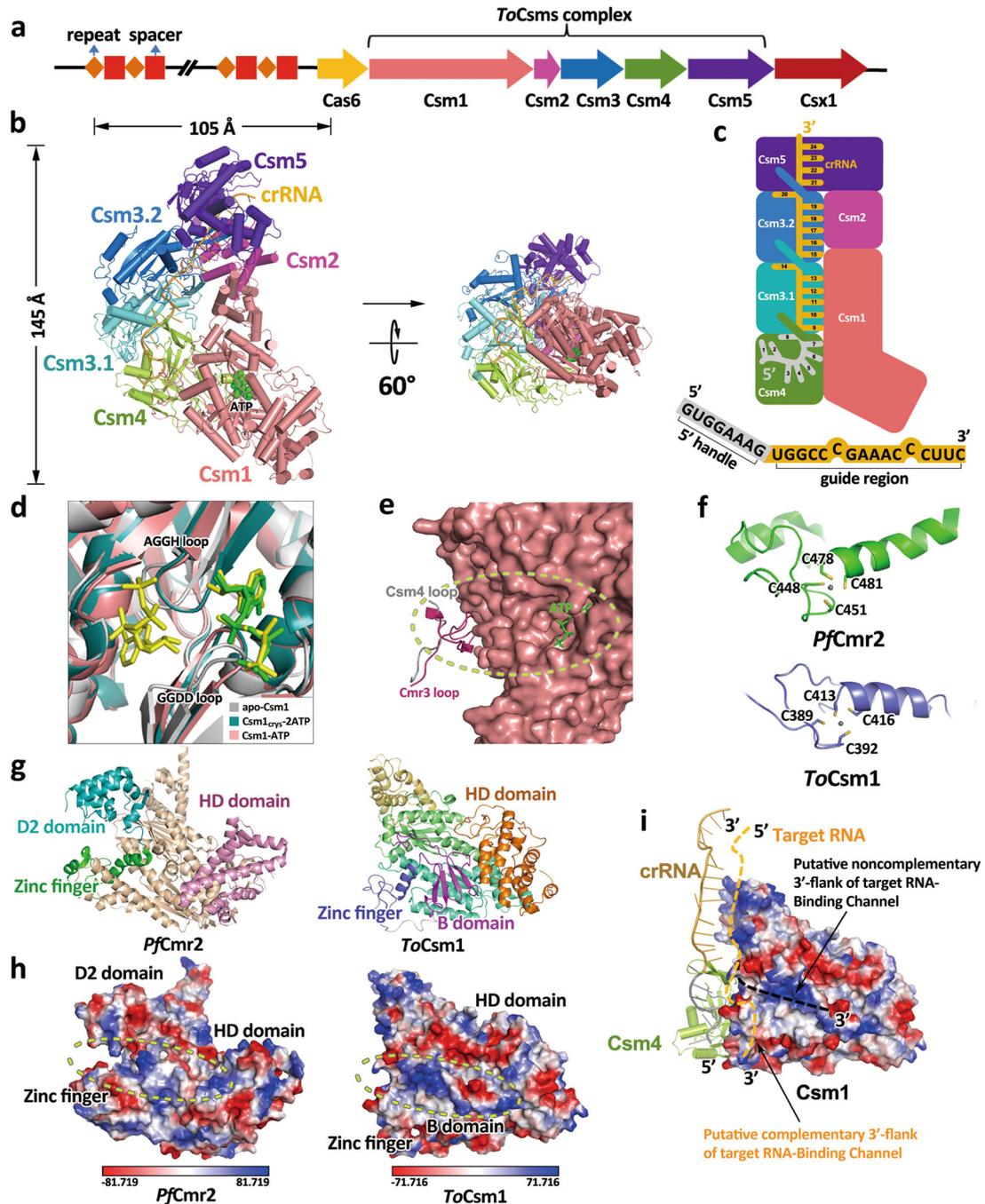


Fig. 1 Cryo-EM structure study of Type III-A CRISPR effector complex. **a** Graphic representation of the *T. onnurineus* CRISPR/Cas locus. **b** Overall Cryo-EM structure of *ToCsm* effector complex. Each subunit is colored according to **(a)**, ATP is shown in green. **c** Schematic representation of the *ToCsm* effector complex. 5'-handle of the crRNA (1–8 nt) is shown in grey, and the guide region (9–24 nt) is shown in orange. **d** Close-up view of superimposed ATP-binding pocket of apo-Csm1 (grey, PDB: 4UW2), Csm1_{crys}-2ATP (teal, with ATP in yellow), and Csm1-ATP (salmon, with ATP in green). **e** The loop (G13-D17) of *ToCsm4* is close to the ATP-binding pocket. **f** Zinc finger domains from *PfCmr2* (up, PDB: 4W8Y) and *ToCsm1* (down). **g** Comparison of *PfCmr2* (left, PDB: 4W8Y) with *ToCsm1* (right) shows significant structural differences between D2 and B domains. **h** Electrostatic surface potential comparison of *PfCmr2* (left) with *ToCsm1* (right) shows differences in their RNA-binding channels for noncomplementary 3'-flank of target RNA. **i** Hypothetical model of *ToCsm* effector complex binding with target RNA. The location of the two channels for binding noncomplementary and complementary 3'-flank of target RNA are indicated. The zinc finger partially occupies the putative channel for complementary 3'-flank of target RNA

Palm2-bound ATP molecule (Supplementary information, Fig. S7c). Our structure seems to present a pre-reaction state of the previously proposed cyclic oligonucleotide formation pocket,^{8,9} which should facilitate future investigation.

The detailed structural basis for the molecular mechanism by which the cyclic oligonucleotide is synthesized by Csm1 remains

unclear. To better understand the underlying molecular determinants, we first superimposed the crystal structures of *PfCmr2*dHD-Cmr3 (Type III-B)¹⁴ and the Cmr/crRNA/target-DNA-analog ternary complex.⁴ Compared to Cmr2dHD-Cmr3, an N-terminal loop (residues D10–S26) of Cmr3 in the ternary Cmr complex swings a far distance to coordinate with crRNA (Supplementary

information, Fig. S8a). Intriguingly, we found that the corresponding loop in *ToCsm4* is much shorter (residues P13-D17) (Supplementary information, Fig. S8b). Our in vitro assay results showed that the *ToCsm* effector complex produces cOA_3 and cOA_4 apparently, but not cOA_5 or cOA_6 (Supplementary information, Fig. S9). As this loop is close to the ATP binding pocket (Fig. 1e), we speculate that crRNA binding may lead to varied loop conformation and in turn affect cOA_4 or cOA_6 production by impacting the reaction pocket.

Intriguingly, sequence alignment of *Csm4* between *T. onnurineus* and several other *Csm* species reported to be able to produce cOA_4 or cOA_6 showed that all the cOA_4 -producing *Csm4*s have a short corresponding loop (4–5 amino acids), whereas *StCsm4* which produces cOA_6 bears a loop with a longer length (13–14 amino acids) (Supplementary information, Fig. S8c). Therefore, it is worth investigating in future work whether the loop length of *Csm4* is correlated to the product species.

The HD domain of *Csm1* performs nonspecific ssDNA degradation in a target RNA-stimulated manner. As Type III effectors discriminate self and non-self DNA/RNA by checking the complementarity of the 5'-handle region, the activity of the HD domain is most likely modulated by the complementarity of the 5'-handle,¹⁰ although the mechanism involved in this is unclear. It was previously predicted that a C4 zinc finger in *Csm1* (residues A386-P427) may be related to the activation or repression of the HD DNase activity (Supplementary information, Fig. S10a).² We found that this region exhibits poor density probably due to its disordered conformation and thus the residues cannot be built in *ToCsm1_{crys}*-2ATP structure. Intriguingly, the corresponding region is well behaved in the *ToCsm* effector complex as a C4 zinc finger motif (Supplementary information, Fig. S10b), with C389 and C392 located in a loop and C413 and C416 located on the head of the following α -helix (Fig. 1f), showing similarity to the C4 zinc finger of *Cmr2* in the *Cmr* complex⁴ and the full-length *PfCmr2* alone (Fig. 1f).¹⁵ Previous studies showed that the mechanisms underlying HD domain activation of Types III-A and III-B may not be identical. For instance, in the absence of the noncomplementary 3'-flanking region of the target RNA, DNase is activated in *PfCmr* (Type III-B), *TtCmr* (Type III-B) but repressed in *StCsm1* (Type III-A).² Moreover, it was predicted that an RNA-binding channel for noncomplementary 3'-flank of target RNA is located between the D2 domain and the zinc finger motif of *Cmr2*. We found that the B domain of *Csm1*, corresponding to the D2 domain in *Cmr2*, forms a four-stranded β -sheet, which is markedly different from the D2 domain containing a four-helix bundle (Fig. 1g). Whereas the above-mentioned RNA-binding channel is well formed in *PfCmr*, the end of the channel that is close to the 5'-handle is narrower in *ToCsm* effector complex than that in the *PfCmr* complex, which may be due to the different folding type of the B domain compared with the D2 domain of *PfCmr2* (Fig. 1h). These observations may support the previous hypothesis that the zinc finger motif as well as the structure of the D2/B domain may be associated with the regulation of DNase activation.² Furthermore, the DNase activity of both Type III-A and III-B effectors is completely repressed when the complementary 3'-flank of target RNA is present (the self-protective state), and we found the zinc finger of *ToCsm* effector complex partially occupies the putative RNA-binding channel for complementary 3'-flank of target RNA. Therefore, we speculate that the complementary 3'-flank of target RNA may push the zinc finger aside from its present position to block the putative channel for noncomplementary 3'-flank of target RNA and in turn impairs the DNase activity of the HD

domain (Fig. 1i). However, further investigations are required to examine this hypothesis.

The density map of *ToCsm* effector complex is available through EMDB with entry code: EMD-9708. The atomic coordinates of *ToCsm* effector complex are deposited in the Protein Data Bank with entry code: 6IQW.

ACKNOWLEDGEMENTS

We thank Dr. Boling Zhu, Dr. Xiaojun Huang, Dr. Gang Ji and Dr. Zhenxi Guo for Cryo-EM data collection at the Center for Biological Imaging (CBI, <http://cbi.ibp.ac.cn>), Institute of Biophysics. We are grateful to the staffs from BL19U beamline at Shanghai Synchrotron Radiation Facility for their assistance during data collection. We acknowledge Dr. Zhensheng Xie and Dr. Lili Niu for Mass spectrometry analysis at Lab of Proteomics of Institute of Biophysics, Chinese Academy of Sciences. This work was supported by grants from Strategic Priority Research Program of the Chinese Academy of Sciences (XDB08010301 and XDB29010000) and the National Science Foundation of China (31670750).

AUTHOR CONTRIBUTIONS

T.J., Y.H., X.W. and T.L. designed the project; Y.H. and T.L. performed the Cryo-EM sample preparation, characterization, and data collection; X.W. and N.W. processed the Cryo-EM data and reconstructed the Cryo-EM map; T.L., Y.H. and T.J. built and refined the structure model; Q.D. assisted in the model building and artwork preparing; Y.H. and T.L. performed the *Csm1*-2ATP sample preparation, crystallization, diffraction data collection and structure determination. T.J., Y.H., T.L. and X.W. analyzed the structure and wrote the manuscript; all authors discussed and commented on the results and the manuscript.

ADDITIONAL INFORMATION

Supplementary information accompanies this paper at <https://doi.org/10.1038/s41422-018-0115-6>.

Competing interests: The authors declare no competing interests.

Yangao Huo^{1,2}, Tao Li^{1,3}, Nan Wang^{1,3}, Qinghua Dong¹,
Xiangxi Wang¹ and Tao Jiang^{1,3}

¹National Laboratory of Macromolecules, Institute of Biophysics, Chinese Academy of Sciences, Beijing, China; ²School of Life Science and Engineering, Foshan University, Guangdong, China and ³University of Chinese Academy of Sciences, Beijing, China
These authors contributed equally: Yangao Huo, Tao Li, Nan Wang
Correspondence: Yangao Huo (huoyangao@moon.ibp.ac.cn) or Xiangxi Wang (xiangxi@ibp.ac.cn) or Tao Jiang (tjiang@ibp.ac.cn)

REFERENCES

1. Makarova, K. S. et al. *Nat. Rev. Microbiol* **13**, 722–736 (2015).
2. Tamulaitis, G., Venclovas, C. & Siksnys, V. *Trends Microbiol* **25**, 49–61 (2017).
3. Taylor, D. W. et al. *Science* **348**, 581–585 (2015).
4. Osawa, T., Inanaga, H., Sato, C. & Numata, T. *Mol. Cell* **58**, 418–430 (2015).
5. Rouillon, C. et al. *Mol. Cell* **52**, 124–134 (2013).
6. Staals, R. H. et al. *Mol. Cell* **56**, 518–530 (2014).
7. Park, K. H. et al. *EMBO Rep.* **18**, 826–840 (2017).
8. Kazlauskienė, M., Kostiuik, G., Venclovas, C., Tamulaitis, G. & Siksnys, V. *Science* **357**, 605–609 (2017).
9. Niewoehner, O. et al. *Nature* **548**, 543–548 (2017).
10. Kazlauskienė, M., Tamulaitis, G., Kostiuik, G., Venclovas, C. & Siksnys, V. *Mol. Cell* **62**, 295–306 (2016).
11. Sapranuskas, R. et al. *Nucleic Acids Res* **39**, 9275–9282 (2011).
12. Shah, S. A. et al. *RNA Biol* **19**, 1–13 (2018).
13. Jung, T. Y. et al. *Structure* **23**, 782–790 (2015).
14. Osawa, T., Inanaga, H. & Numata, T. *J. Mol. Biol.* **425**, 3811–3823 (2013).
15. Benda, C. et al. *Mol. Cell* **56**, 43–54 (2014).