



LETTER TO THE EDITOR

Structural insights into Cas13b-guided CRISPR RNA maturation and recognition

Cell Research (2018) 28:1198–1201; <https://doi.org/10.1038/s41422-018-0109-4>

Dear Editor,

CRISPR-Cas (clustered regularly interspaced short palindromic repeats and CRISPR-associated proteins) systems are RNA-guided adaptive immune systems in prokaryotes.^{1,2} Class 2 CRISPR-Cas systems (including type II, V, and VI) involve large single effector proteins in complex with crRNA for interference.^{3,4} The type II and V effectors, such as Cas9 and Cas12a, have been engineered into powerful tools for genome editing. The type VI system encompasses RNA-guided RNases. Its effectors Cas13a, Cas13b and Cas13d are capable of both precursor CRISPR RNA (pre-crRNA) processing and target RNA cleavage, which protect the host from phage attacks.^{5–7} Once bound to a target RNA, they are activated, switching on a non-specific RNase activity. Moreover, they have been utilized to target and edit RNA as programmable RNA-binding modules.^{6,8–12} Although related to Cas13a and Cas13d, Cas13b possesses many distinctive features. These include the lack of significant sequence similarity with Cas13a and Cas13d, disparate crRNA repeat region, double-sided protospacer flanking sequence (PFS)-dependent target RNA cleavage.^{5–8,13}

To investigate how Cas13b processes pre-crRNA, recognizes crRNA and settles the spacer nucleotides for target recognition, we solved the crystal structure of *Bergeyella zoohelcum* Cas13b (BzCas13b) in complex with its crRNA at 2.79 Å resolution (Supplementary information, Table S1). The binary complex was obtained by the SeMet-derived BzCas13b R1177A mutant co-expressed with CRISPR template in vivo. The architecture of BzCas13b assumes a triangular domain distribution around the central L-shaped crRNA (Fig. 1a–e; Supplementary information, Movie S1). In the binary complex, Helical-1, HEPN-1 and HEPN-2 domains together form one side of the triangular structure. Helical-1 domain comprises six α -helices connected with random loops (Supplementary information, Fig. S1). The second side of the triangle is formed by RRI-1 (the repeat region interacting domain-1), RRI-2 domains and the linker region. RRI-1 domain can be subdivided into two separate motifs (RRI-1 I and II) that stack onto each other. Both motifs contain a short two-stranded, antiparallel β -sheet flanked by five α -helices. RRI-2 domain includes a long central two-stranded, antiparallel β -sheet flanked by two α -helices, and a short central two-stranded, antiparallel β -sheet flanked by three α -helices. The linker region consists of random loops that connect two short α -helices, which shows multiple interactions with RRI-2 domain. Helical-2 domain is composed of nine α -helices and its rather long helix-23 extends in parallel with crRNA, thereby forming the third side of the triangle. Helix-8 of Helical-1 domain and helix-23 of Helical-2 domain protrude out of the complex in a crab claw-like manner to clamp the spacer region of crRNA (Supplementary information, Fig. S1). In addition, HEPN-1 domain bridges Helical-1 and Helical-2 domains.

A mature 52-nt crRNA, originated from a co-expressed CRISPR encoding sequence and being processed by BzCas13b itself in *E. coli* cells, was identified in the binary complex. The mature crRNA contains an intact 36-nt repeat region (G(-1)-C(-36)) and a 5'-

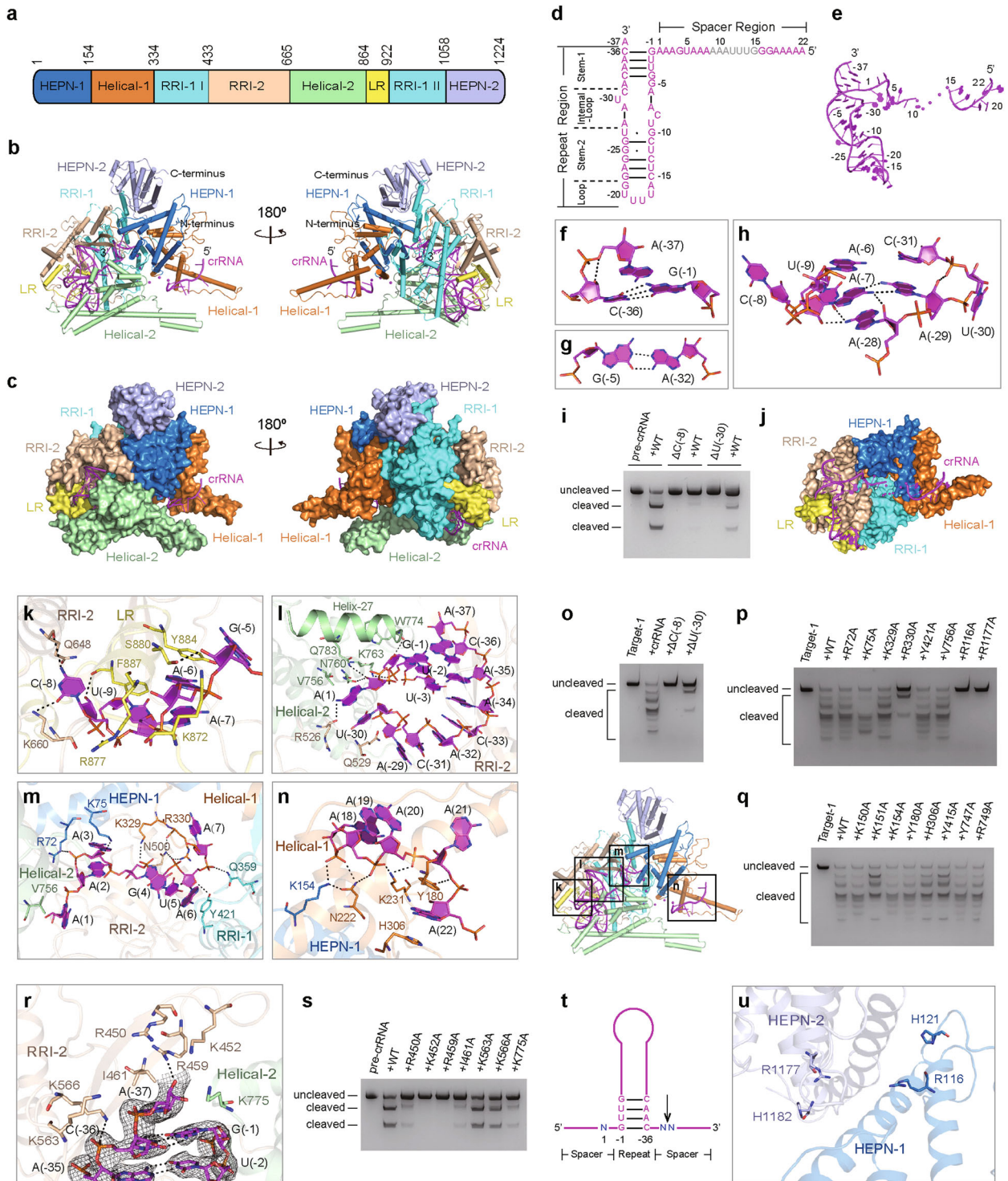
upstream spacer region (A(22)-A(1)) (Fig. 1d, e). In particular, an additional nucleotide A(-37) from the spacer region was also identified in the mature crRNA, which locates 3'-downstream of the repeat region (Fig. 1d–f). On the whole, crRNA adopts the L-shaped architecture (Fig. 1e). The conformation of the repeat region shows a distorted RNA duplex, while the spacer region makes a $\sim 90^\circ$ turn. The repeat region can be further divided into four sub-regions: stem-1, internal-loop, stem-2 and loop regions. In the stem-1 region, the nucleotides G(-1)-G(-4) and C(-33)-C(-36) form four canonical Watson-Crick base pairs, and G(-5) forms two hydrogen bonds with A(-32) (Fig. 1d, g). In the stem-2 region, the nucleotides G(-10)-C(-15) and G(-22)-U(-27) form four Watson-Crick base pairs and two wobble G-U base pairs. In the internal-loop region, two nucleotides U(-30) and C(-8) twist away from the RNA duplex, and A(-7) forms hydrogen bonds with A(-29) (Fig. 1h). Notably, the nucleotide U(-30) stacks with the spacer nucleotide A(1), which stabilizes their conformations (Fig. 1e). While deletion of U(-30) significantly decreased the pre-crRNA processing by the protein, the deletion of C(-8) almost abolished it (Fig. 1i). The nucleotide U(-9) interacts with A(-28) through a hydrogen bond, but does not form a Watson-Crick base pair (Fig. 1h). The nucleotides A(-6), C(-8), U(-30) and C(-31) remain unpaired within this region. In addition, the nucleotides A(-16)-G(-21) constitute the loop region.

The crRNA makes extensive intermolecular interactions with multiple domains of BzCas13b except HEPN-2 domain (Fig. 1b). The majority of the interactions play a key role in stabilizing the crRNA phosphodiester backbone (Supplementary information, Fig. S2). The repeat region of crRNA is anchored inside the pocket formed by Helical-2, two RRI domains and the linker region (Fig. 1b, j). Helix-34 of RRI-1 II motif interacts with the backbone phosphate groups of the crRNA stem-2 and loop regions (Supplementary information, Fig. S3a, b). RRI-1 I motif has no direct contact with the crRNA repeat region. Helical-2 domain maintains electrostatic interactions with the phosphate groups of U(-27) and A(-28) via K680 and K677, as well as with those of U(-2) and U(-3) via K763 and K771 (Supplementary information, Fig. S2). In addition, N772 and W774 form hydrogen bonds with 2'-hydroxyl groups of U(-2) and G(-1), respectively.

RRI-2 domain is most intensively involved in interactions with the crRNA repeat region. It makes direct contacts with the crRNA stem-1, stem-2 and internal-loop regions. More than 10 residues in RRI-2 domain interact with the sugar-phosphate backbone of the nucleotides G(-25)-A(-37) and G(-10)-C(-11) through their side chains or backbone atoms (Supplementary information, Figs. S2 and S3c–e).

The linker region of BzCas13b entangles with the crRNA duplex region and contributes a number of interactions with four sub-regions of the repeat region. About 11 residues of the linker region interact with the sugar-phosphate backbone of the nucleotides G(-5)-A(-29) (Supplementary information, Fig. S3f–h). Particularly, the side chain of F887 stacks with U(-9), which significantly displaces the orientation of C(-8). As a result, C(-8)

Received: 3 September 2018 Revised: 15 October 2018 Accepted: 24 October 2018
Published online: 13 November 2018



flips outward from the crRNA duplex. The N4, O2 atoms and the 2'-hydroxyl group of C(-8) form hydrogen bonds with Q648, K660 and F887, respectively, which further stabilize the conformation of C(-8) (Fig. 1k). Furthermore, the neighboring A(-7) stacks between A(-28) and A(-6), with the 2'-hydroxyl group hydrogen-bonding with the N6 atom of A(-28) (Fig. 1h). Additionally, the side chain of Y884 wedges between G(-5) and A(-6), stacking with them. Helix-27 of Helical-2 domain disturbs the orientation of the nucleotide A(1), and there is a swinging of A(1) relative to G(-1)

(Fig. 1l). Interestingly, the nucleotide A(1) stacks with U(-30) and their spatial orientations show a certain degree of continuity. This suggests that the nucleotide U(-30) may play a specific role in 5'-PFS recognition through base pairing. To verify this notion, the nucleotide U(-30) was replaced by G(-30) and the crRNA mutant was tested for the target RNA cleavage. Unexpectedly, the mutation did not detectably alter pre-crRNA processing or the non-cytosine 5'-PFS targeting preference of BzCas13b (Supplementary information, Fig. S4a, b). One possible explanation is that the cytosine of 5'-PFS may

Fig. 1 Structural and biochemical studies of the BzCas13b-crRNA binary complex. **a** Domain organization of BzCas13b. **b** Overall structure of the BzCas13b-crRNA binary complex, color-coded as defined in (a). **c** Surface representations of the BzCas13b-crRNA binary complex. **d** Schematic representation of the crRNA secondary structure. Bars between nucleotide pairs represent Watson-Crick base pairs and dots represent wobble G-U base pairs. The disordered nucleotides are colored in grey. **e** Structure of the crRNA in the BzCas13b-crRNA binary complex. Dots represent possible positions of the disordered nucleotides within the spacer region. **f-h** Detailed base pair interactions within crRNA. Hydrogen bonds are shown as dashed lines. **i** Denaturing gel demonstrating the cleavage of pre-crRNA, the Δ C(-8) and Δ U(-30) pre-crRNA mutants by wild-type BzCas13b. **j** Surface representation of BzCas13b HEPN-1, RRI-1, RRI-2, Helical-1 domains and the linker region (LR) in complex with crRNA. **k-n** Detailed interactions between crRNA and different domains of BzCas13b. **o** Denaturing gel demonstrating the cleavage of target-1 RNA by wild-type BzCas13b in complex with crRNA, the Δ C(-8) and Δ U(-30) crRNA mutants, respectively. **p, q** Denaturing gel demonstrating the cleavage of target-1 RNA by wild-type BzCas13b and the mutants in complex with crRNA. **r** Close-up view of BzCas13b active site for the pre-crRNA processing. The *2Fo-Fc* omit map is contoured at 1.0 σ level. **s** Denaturing gel demonstrating the pre-crRNA cleavage by wild-type BzCas13b and the mutants speculatively involved in the pre-crRNA processing. **t** Schematic representation of BzCas13b processing pre-crRNA at the phosphodiester bond connecting two spacer nucleotides located directly 3'-downstream of the repeat region. **u** Close-up view of BzCas13b active site for the target RNA cleavage. Residue A1177 of the BzCas13b R1177A mutant binary complex was virtually mutated back to R1177

compete with C(-36) to pair with G(-1), which may hinder the conformational activation for the target cleavage. Additionally, mutations of residues involved in the interactions with the repeat region did not obviously influence the pre-crRNA processing and target RNA cleavage (Supplementary information, Fig. S4c, d).

The spacer region is wrapped by Helical-1, HEPN-1, two RRI domains on one side, and by Helical-2 domain on the other side. It adopts a U-shaped turn in the region of A(8)-A(1). Due to the flexibility and accessibility to the environmental solvent, the electron density of the nucleotides G(15)-A(9) within the spacer region is poor and these nucleotides are not determined in the present model (Supplementary information, Fig. S5a). However, the discontinuous electron density of G(15)-A(9) located between Helical-1 and HEPN-1 domains still provides a hint about how the nucleotides pass through this region (Supplementary information, Fig. S5b). Residues K150, K151, K157, Y415, Y747 and R749 are speculated to interact with the nucleotides U(14)-A(10). In contrast, the nucleotides A(22)-G(16) are clearly determined and mainly interact with Helical-1 domain. A number of residues contribute to the conformational stabilization of the spacer region (Fig. 1m, n). The side chain of residue V756 generates an obstruction to the motion of the nucleotide A(1). In line with the pre-crRNA processing experiment, the deletion of the nucleotide U(-30) greatly decreased target RNA degradation, and the deletion of C(-8) resulted in no observable cleavage activity (Fig. 1o). However, the mutation of C(-8) or U(-30) did not notably affect the pre-crRNA processing and the cleavage of target RNA (Supplementary information, Fig. S4e, f). In addition, the R330A mutation significantly decreased the cleavage activity, and the K151A, H306A and Y415A mutations only slightly reduced the cleavage activity (Fig. 1p, q).

The mature crRNA contains an intact 36-nt repeat region and an additional 3'-downstream spacer nucleotide A(-37) (Fig. 1r; Supplementary information, Fig. S5a). Residues surrounding the nucleotide A(-37) have been speculated to play a critical role during the pre-crRNA processing. To determine residues that comprise the catalytic site, mutational analysis has been carried out and a pre-crRNA consisting of two repeats separated by a spacer has been used as the processing target (Fig. 1s). BzCas13b is active in pre-crRNA processing, whereas alanine substitution of K452 or R459 abolished this activity. Considering that these two residues from RRI-2 domain are conserved among Cas13b homologs (Supplementary information, Fig. S6), they play an essential role for the catalytic activity. The pre-crRNA cleavage may be through a base-catalyzed process. Moreover, mutations of R450A, I461A and K775A lead to significantly reduced activity, whereas the mutation of K563A or K566A results in a limited decrease of the catalytic activity. In addition, in order to verify whether A(-37) possesses the base specificity for the pre-crRNA cleavage, A(-37) was mutated to G, C and U, respectively. Results indicate that base type change at this position does not influence the cleavage activity, whereas the mutation of K452A or R459A still abolishes the catalytic activity (Supplementary information, Fig. S4g). Thus, BzCas13b processes pre-crRNA one nucleotide 3'-downstream of the repeat RNA

duplex structure, which resides at the spacer region regardless of the base type (Fig. 1t). This feature has been further verified by RNA sequencing (Supplementary information, Fig. S7).

The previous study has revealed that residues R116, H121, R1177 and H1182 from two HEPN domains of BzCas13b are involved in the cleavage of the target and collateral RNAs.⁶ Here, we demonstrated that a single R116A or R1177A mutation was sufficient to abolish the target RNA cleavage activity (Fig. 1p). Despite the fact that two R-X₄-H motifs face each other, the distance between the backbone Ca atoms of residues R116 and R1177 in the present model is ~18.6 Å, which remains too distant for the catalysis (Fig. 1u). Therefore, further conformational adjustment is required for BzCas13b to exhibit its RNase activity.

Although Cas13a and Cas13b exhibit distinct domain architectures, they may employ a similar strategy to fulfill the crRNA-mediated RNA targeting and cleavage (Supplementary information, Figs. S8 and S9).^{5,6,14,15} In summary, the present study determines the molecular architecture of the BzCas13b binary complex, which reveals the catalytic site for pre-crRNA cleavage and elucidates the crRNA recognition pattern. Our study can be further used to leverage Cas13b in a wide range of potential applications.

ACCESSION NUMBER

The atomic coordinates and structure factors of the BzCas13b-crRNA complex has been deposited in the Protein Data Bank under the accession code 6AAY.

ACKNOWLEDGEMENTS

This work was supported by the Ministry of Science and Technology of China grants 2014CB910400 and the National Nature Science Foundation of China grants 31770948, 31570875 and 81590761. We thank Fujian Normal University for financial support. The diffraction data were collected at the beamline BL-17U1 of Shanghai Synchrotron Radiation Facility (SSRF). We thank Gonghong Wei and Zhaoqing Luo for discussions. We thank Wei Yang and Zhijian J. Chen for critical review of the manuscript.

AUTHOR CONTRIBUTIONS

B.Z. and W.Y. purified the BzCas13b-crRNA complex and the BzCas13b mutants. B.Z. and Y.Y. purified the in vitro transcribed RNAs and performed the cleavage assays. B.Z., W.Y. and Y.Y. performed the crystallization screening. B.Z. optimized the crystallization condition. H.Z. and B.Z. collected X-ray diffraction data. B.Z. solved the crystal structure. J.C., J.L. and V.P. contributed to some experiments. B.Z., S.O., Q.C., C.C., M.C. and M.I.C. designed the experiments. S.O. and B.Z. supervised the study. B.Z., A.S., S.O. wrote the manuscript.

ADDITIONAL INFORMATION

Supplementary information accompanies this paper at <https://doi.org/10.1038/s41422-018-0109-4>.

Competing interests: The authors declare no competing interests.

Bo Zhang^{1,2}, Weiwei Ye¹, Yangmiao Ye¹, Huan Zhou³,
Abdullah F. U. H. Saeed¹, Jing Chen¹, Jinying Lin¹, Vanja Perčulija¹,
Qi Chen¹, Chun-Jung Chen⁴, Ming-Xian Chang⁵,
Muhammad Iqbal Choudhary⁶ and Songying Ouyang^{1,2,7}

¹The Key Laboratory of Innate Immune Biology of Fujian Province, Biomedical Research Center of South China, College of Life Sciences, Fujian Normal University, Fuzhou 350117, China; ²Provincial University Key Laboratory of Cellular Stress Response and Metabolic Regulation, College of Life Sciences, Fujian Normal University, Fuzhou 350117, China; ³Shanghai Institute of Applied Physics, Chinese Academy of Sciences, Shanghai 201204, China; ⁴Institute of Biotechnology, National Cheng Kung University, Tainan, 701 Taiwan, China; ⁵State Key Laboratory of Freshwater Ecology and Biotechnology, Institute of Hydrobiology, Chinese Academy of Sciences, Wuhan 430072, China; ⁶H.E.J. Research Institute of Chemistry, International Center for Chemical and Biological Sciences, University of Karachi, Karachi 75270, Pakistan and ⁷National Laboratory of Biomacromolecules, Institute of Biophysics, Chinese Academy of Sciences, Beijing 100101, China

These authors contributed equally: Bo Zhang, Weiwei Ye
Correspondence: Songying Ouyang (ouyangsy@fjnu.edu.cn)

REFERENCES

1. Koonin, E. V., Makarova, K. S. & Wolf, Y. I. *Annu. Rev. Microbiol.* **71**, 233–261 (2017).
2. Hille, F. et al. *Cell* **172**, 1239–1259 (2018).
3. Shmakov, S. et al. *Nat. Rev. Microbiol.* **15**, 169–182 (2017).
4. Jiang, F. & Doudna, J. A. *Annu. Rev. Biophys.* **46**, 505–529 (2017).
5. Abudayyeh, O. O. et al. *Science* **353**, aaf5573 (2016).
6. Smargon, A. A. et al. *Mol. Cell* **65**, 618–630 e617 (2017).
7. Yan, W. X. et al. *Mol. Cell* **70**, 327–339 e325 (2018).
8. Konermann, S. et al. *Cell* **173**, 665–676 e614 (2018).
9. Abudayyeh, O. O. et al. *Nature* **550**, 280–284 (2017).
10. Cox, D. B. T. et al. *Science* **358**, 1019–1027 (2017).
11. Gootenberg, J. S. et al. *Science* **356**, 438–442 (2017).
12. Gootenberg, J. S. et al. *Science* **360**, 439–444 (2018).
13. East-Seletsky, A. et al. *Nature* **538**, 270–273 (2016).
14. Liu, L. et al. *Cell* **170**, 714–726 e710 (2017).
15. Liu, L. et al. *Cell* **168**, 121–134 e112 (2017).