



RESEARCH HIGHLIGHT OPEN

Structural insights into Cas9 mismatch: promising for development of high-fidelity Cas9 variants

Honghai Tang^{1,2,3}, Daqi Wang^{1,2,3} and Yilai Shu^{1,2,3}✉

Signal Transduction and Targeted Therapy (2022)7:271

; <https://doi.org/10.1038/s41392-022-01139-z>

The recent research published in *Nature* by Bravo and colleagues has deeply clarified the mechanisms of Cas9 recognizing the base-pair mismatches by kinetics-guided Cryo-EM analysis, which provides a perspective for the development of next-generation Cas9 variants with high-fidelity.¹

CRISPR-Cas9 is a type of endonuclease from bacteria and archaea, and can be programmed to target the genome under the guiding of a specific RNA. Although CRISPR-Cas9 has become one of the most popular tools for genome editing (e.g., the rapid generation of animal models, the steady-state performance optimization of plants, the treatment of hereditary diseases and oncology, etc.), the certain defects of Cas9 hinder its clinical application, especially off-target effects. To date, some strategies (e.g., reducing nonspecific DNA contacts, neutralizing the positive charge) can reduce the off-target effects of Cas9 due to mismatches,^{2,3} but the efficiency of Cas9 variants is also correspondingly decreased in the on-target DNA cleavage. Therefore, it is very urgent to explore the underlying mechanisms that Cas9 recognizes mismatches for developing next-generation Cas9 editing tools with high efficiency and high specificity.

Given that the activated intermediates before Cas9 cleavage are more suitable to help understand the mismatch mechanisms, it is necessary to capture a large amount of the pre-cleavage conformations of Cas9-sgRNA-DNA ternary complexes. To obtain the suitable structures of Cas9-sgRNA-DNA ternary complexes for elucidation of the off-target mechanisms of Cas9, the authors first used kinetic analysis to detect rates of Cas9 cleaving complementary strand (TS), which carries a contiguous 3-nt mismatch with sgRNA. In comparison to rapid on-target cleavage, the authors screened three substrates with 3-nt base mismatches from the PAM distal end to resolve the structures of Cas9-sgRNA-DNA ternary complexes. These substrates included the following DNA segments with triple mismatches 15–17 bp (15–17 mm), 12–14 bp (12–14 mm), and 18–20 bp (18–20 mm), the cleavage rates of which became higher in turn.¹

To understand how Cas9 was activated in mismatched DNA substrates, the authors resolved two structures that Cas9 combined 12–14 mm DNA at different reaction time points. In structure after reaction for 5 min, the TS-sgRNA duplex of the PAM-distal end presented a linear conformation in comparison with the DNA–DNA duplex of the PAM-proximal end, and HNH domain (consists of two beta-sheets and four alpha-helices, which are similar to the structure of HNH endonuclease) was not observed; in the structure after reaction for 60 min, the TS-sgRNA

duplex of the PAM distal end included both linear and kinked duplex conformations.¹ This suggests that the linear duplex conformation resembles a transition state of Cas9 prior to HNH rearranging and docking to TS.¹ Meanwhile, the structure displayed that the REC3 (recognition lobe 3) domain of Cas9 did not directly contact nucleotides from 12 to 14 in the TS-sgRNA, which was also supported by the structure of Cas9 combining 15–17 mm DNA, implying that the REC3 plays an important role in detecting PAM-distal mismatches. These data suggest that the linear conformation preceding the kinked conformation of the TS-sgRNA duplex is necessary for activation, and inhibition of the kinked duplex formation due to mismatches may avoid DNA cleavage by Cas9.¹

To reveal how certain mismatches were recognized by Cas9 to obtain high-effective DNA cleavage in comparison with other mismatches, the authors determined the structure that Cas9 combined 18–20 mm DNA. Particles of the structure contained both linear and kinked duplex conformations, implying that this more tolerant mismatch also experiences a structural transition similar to that of 12–14 mm DNA.¹ Meanwhile, HNH docking at the shearing site of the TS was observed in the structure.¹ These indicate that Cas9 is the fully active conformation in the structure, and the kink of the TS-sgRNA duplex is concerned with HNH docking. In addition, in the structure, RuvC (comprises ten beta-sheets and nine alpha-helices, which share structural similarity with RNase H, for example, *Escherichia coli* RuvC) active site hydrolyzing nucleotide displayed a histidine-mediated catalytic model.¹ The high-quality cryo-EM density of the helices L1 and L2 was observed,¹ the atomic interface interaction of which indicates that TS and NTS cleavage is couple. The 18–20 mm presented a distinctive duplex conformation at the position of the mismatch, which was stabilized by RuvC residues,¹ implying that the residues only participated in the mismatch combining, instead of the on-target activation. Based on these structural works, the authors designed a high-fidelity variant (SuperFi-Cas9) with extreme-low mismatch rates and near wild-type cleavage efficiency, which has an ability to distinguish between on-target DNA and off-target DNA without harming the efficiency of DNA cleavage.¹ But the efficacy of the variant still needs to be verified in cells or animal models in future works.

In previous studies, many structural forms of CRISPR-Cas9 were determined, including Cas9 monomer (PDB ID: 4CMP), Cas9-sgRNA binary complex (PDB ID: 4ZT0), and Cas9-sgRNA-DNA ternary complexes (PDB IDs: 4O08, 4UN3, 5F9R, 5Y36, and 6O0Y).

¹ENT Institute and Department of Otorhinolaryngology, Eye & ENT Hospital, State Key Laboratory of Medical Neurobiology and MOE Frontiers Center for Brain Science, Fudan University, Shanghai, China; ²Institute of Biomedical Science, Fudan University, Shanghai, China and ³NHC Key Laboratory of Hearing Medicine (Fudan University), Shanghai, China

Correspondence: Yilai Shu (yilai_shu@fudan.edu.cn)

Received: 3 June 2022 Revised: 8 July 2022 Accepted: 19 July 2022

Published online: 04 August 2022

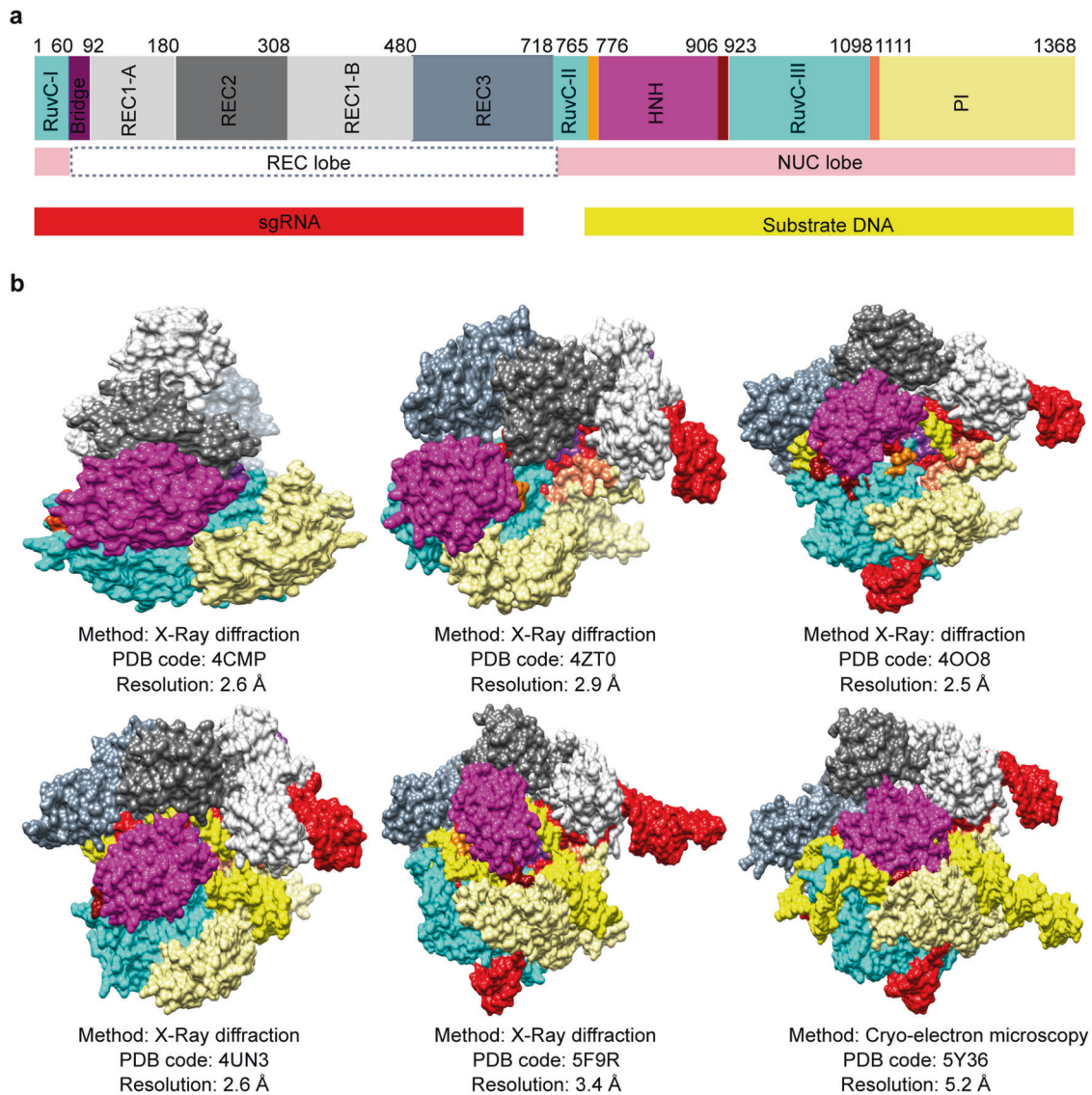


Fig. 1 Structures of CRISPR-Cas9 at different states. **a** Domains of the Cas9-sgRNA-DNA ternary complex are marked by different colors. Cas9 contains 1368 amino acids. **b** The structures are determined by X-ray diffraction or Cryo-electron microscopy, including Cas9 monomer (PDB ID: 4CMP), Cas9-sgRNA binary complex (PDB ID: 4ZT0), and Cas9-sgRNA-DNA ternary complexes (PDB IDs: 4008, 4UN3, 5F9R, 5Y36)

In these structures, only the ternary complexes may be suitable for exploring the off-target effects of Cas9. However, their DNA substrates are only on-target DNA double strands without mismatch. And except 6O0Y,⁴ the residues Asp10 and His840 of the catalytic centers in the other structures are mutated into Ala,⁵ the atomic models of which are not the actually catalytic-active conformation (Fig. 1). For this, it is very difficult to clearly elucidate the mechanisms of Cas9 off-target effects in the active state using these structures. Therefore, the structures of Cas9-sgRNA-DNA ternary complexes in different mismatch states were resolved by Bravo et al. to compensate for the deficiencies of the earlier Cas9 structures, and have important implications for understanding and elucidation of the off-target mechanisms of Cas9. However, now it is not clear whether the rules revealed by Bravo et al. are appropriate for more complex forms of the off-target sites in the entire genome.

Taken together, Bravo and colleagues' work provides a systematic study for Cas9 recognizing mismatches at the atomic level, and offers underlying structural insights into the off-target

mechanisms of Cas9. Meanwhile, this work also provides structural guidance for the development of next-generation Cas9 variants with high efficiency and high specificity.

ACKNOWLEDGEMENTS

This work was supported by the National Key Research and Development Program of China grant (2020YFA0908201), National Natural Science Foundation of China grant (82171148), Clinical Research Plan of SHDC (SHDC2020CR4083), the Science and Technology Commission of Shanghai Municipality (21511905100), Special Project for Clinical Research in Health Industry of Shanghai Municipal Health Commission (20224Z0003), Excellent Doctors-Excellent Clinical Researchers Program of Eye & ENT Hospital of Fudan University (SYB202005), and "Shuguang Program" supported by Shanghai Education Development Foundation and Shanghai Municipal Education Commission (205G08).

ADDITIONAL INFORMATION

Competing interests: The authors declare no competing interests.

REFERENCES

1. Bravo, J. P. K. et al. Structural basis for mismatch surveillance by CRISPR-Cas9. *Nature* **603**, 343–347 (2022).
2. Slaymaker, I. M. et al. Rationally engineered Cas9 nucleases with improved specificity. *Science* **351**, 84–88 (2016).
3. Kleinstiver, B. P. et al. High-fidelity CRISPR-Cas9 nucleases with no detectable genome-wide off-target effects. *Nature* **529**, 490–495 (2016).
4. Zhu, X. et al. Cryo-EM structures reveal coordinated domain motions that govern DNA cleavage by Cas9. *Nat. Struct. Mol. Biol.* **26**, 679–685 (2019).
5. Jiang, F. & Doudna, J. A. CRISPR-Cas9 structures and mechanisms. *Annu Rev. Biophys.* **46**, 505–529 (2017).



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2022