RESEARCH HIGHLIGHT

Structure-based evolutionary relationship between IscB and Cas9

Ning Jia[1]✉ and Dinshaw J. Patel[2]✉

Cell Research (2022) 32:875–877; https://doi.org/10.1038/s41422-022-00709-8

In a recent study published in *Science*, Ailong Ke and coworkers determined the cryo-EM structure of the dsDNA-bound IscB$^{\omega RNA}$ ribonucleoprotein (RNP) complex, which was supplemented with complementary biochemical experiments, revealing the architectural and mechanistic similarities for RNA-guided DNA recognition and activation-mediated cleavage by IscB and Cas9 RNPs. These findings highlight the evolutionary relationship between IscB and Cas9, and the potential for development of enhanced and programmable genome editing tools by capitalizing on IscB's small size.

CRISPR (clustered regularly interspaced short palindromic repeats)-Cas (CRISPR-associated genes) systems provide prokaryotes with adaptive immunity against invading nucleic acids from viruses and plasmids.[1,2] Based on their genetic composition, CRISPR systems are subdivided into two major classes: class 1 (including types I, III and IV) contains multi-subunit effector complexes, whereas class 2 (including type II, V and VI) is composed of single-subunit protein effectors.[3] Amongst these CRISPR families, the DNA-targeting type II Cas9 effector complex has been widely used for genome editing, with potential impact on translational medicine.[4,5] The type II Cas9 effector complex recognizes and cleaves target double-stranded DNA (dsDNA) using a large single-subunit multidomain protein, termed Cas9, along with two non-coding RNAs, termed CRISPR RNA (crRNA) and *trans*-activating crRNA (tracrRNA), both of which are processed from the CRISPR array in type II CRISPR-Cas9 loci (Fig. 1a). The Cas9 protein consists of two distinct modules termed recognition (REC) and nuclease (NUC) lobes, with the latter containing HNH and split RuvC nuclease domains (Fig. 1b, c).[6] crRNA and tracrRNA form a duplex that guides Cas9 to recognize target dsDNA with a protospacer-adjacent motif (PAM) sequence, resulting in cleavage of the complementary target and non-target strands of the target dsDNA by HNH and RuvC nuclease domains, respectively (Fig. 1c).

The sequence-specific recognition and cleavage of dsDNA target has highlighted Cas9 as an indispensable genome editing tool. However, given the large size of Cas9, it has proved challenging to deliver it into target cells, a feature reflecting the packaging limit associated with delivery of viral vectors such as adeno-associated virus (AAV).[7] Therefore, it has remained a challenge to package genes encoding Cas9, a sgRNA and a donor DNA together with associated promoters into a single AAV vector. Thus, many investigations have focused on identifying smaller alternatives to Cas9, so as to capitalize on its genome editing potential.

Bioinformatic analysis have identified IscB proteins that are only about two fifths of the size of Cas9 but share a similar domain organization.[8] Follow-up biochemical studies have identified the IscB-containing IS200/IS605 transposon family,[9,10] whose loci consist of genes encoding the IscB protein and a non-coding RNA (Fig. 1d), named ωRNA, which contains a highly variable and an evolutionarily conserved region, similar to the spacer and repeat regions in the CRISPR array. The IscB protein has been shown to work as a programmable RNA-guided endonuclease that is functionally similar to Cas9 nuclease, while also requiring a target-adjacent motif (TAM, 5′-NWRRNA-3′) in the target dsDNA, analogous to the indispensable PAM in dsDNA targets of Cas9.

A recent study published in *Science* by Schuler et al. reported a 2.78 Å cryo-EM structure of target dsDNA-bound IscB$^{\omega RNA}$ ternary complex, which contains a 496-aa IscB protein and 222-nt ωRNA (Fig. 1e, f).[11] The IscB$^{\omega RNA}$ ribonucleoprotein (RNP) complex adopts a compact two-lobed architecture similar to Cas9 (Fig. 1c), but lacks the REC lobe within the IscB fold. Remarkably, a structured portion of the ωRNA lobe (Fig. 1f, in grey) is located in a similar position to the REC lobe of Cas9 (Fig. 1c, in grey), indicating an evolutionary trend involving replacement of the structured ωRNA from ancestral IscB with protein domains for Cas9. In addition, ωRNA possesses structural elements functionally equivalent to repeat/anti-repeat duplex in crRNA–tracrRNA–Cas9 effector complex, which can be artificially covalently linked into a single sgRNA that has been widely used for Cas9-mediated gene editing applications (Fig. 1c, f), raising the possibility that ωRNA could be a naturally occurring sgRNA. The variable region of ωRNA serves as spacer for recognition of the foreign dsDNA target (Fig. 1f, in red), guiding the IscB protein to recognize and cleave target dsDNA by its HNH and RuvC nuclease domains (Fig. 1f). Furthermore, the IscB protein (Fig. 1f) contains a conserved Arg-rich bridge helix structural element (in marine) in common with Cas9, a P1D domain (in purple) that is essentially the functional equivalent of the Cas9 WED domain, a TAM interaction domain (TID, in yellow) responsible for TAM recognition, and an additional PLMP motif-containing domain (in salmon) which is missing from Cas9 during evolution. Combining structural biology and biochemistry, Schuler et al. present a target dsDNA cleavage mechanism for IscB similar to that proposed for Cas9, in which the HNH domain cleaves the target strand followed by non-target strand cleavage by the RuvC domain (Fig. 1f). Despite the weak genome editing activity reported for IscB–ωRNA RNP observed in human cells, the

[1]Department of Biochemistry, School of Medicine, Shenzhen Key Laboratory of Cell Microenvironment, Guangdong Provincial Key Laboratory of Cell Microenvironment and Disease Research, Southern University of Science and Technology, Shenzhen, Guangdong, China. [2]Structural Biology Program, Memorial Sloan-Kettering Cancer Center, New York, NY, USA. ✉email: jian@sustech.edu.cn; pateld@mskcc.org

Published online: 5 August 2022

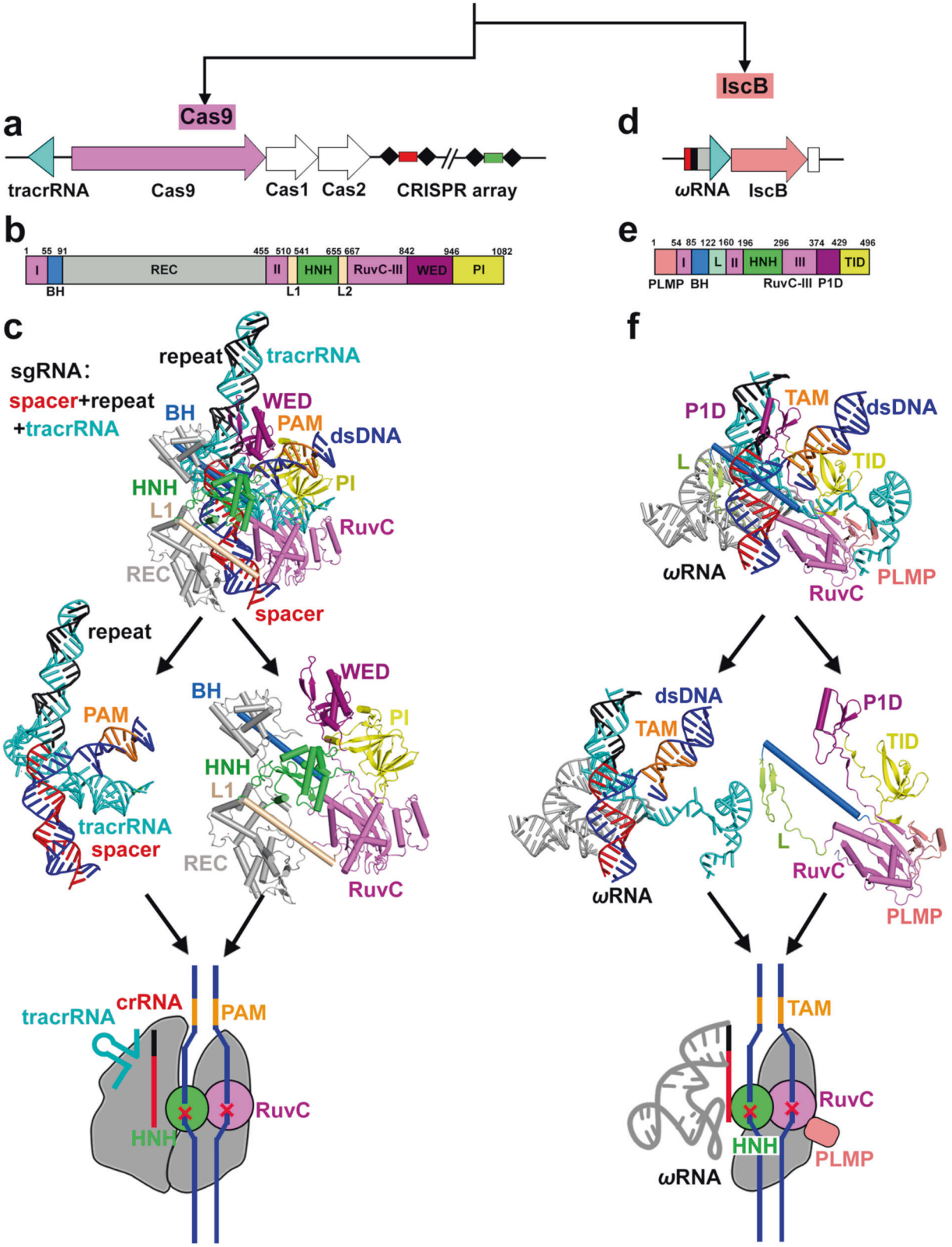

Fig. 1 Structural and mechanistic comparison between Cas9 and IscB. a, **d** The outline of the CRISPR-Cas9 loci (**a**) and IscB-ωRNA loci (**d**). The CRISPR array is composed of the host nucleotide repeats (black diamonds) separated by invading spacer sequences (colored cylinders). The ωRNA is colored according to the functionally equivalent elements related to the crRNA and tracrRNA in CRISPR-Cas9 loci. **b**, **e** Domain organization of *Neisseria meningitidis* Cas9 (**b**; PDB: 6JDV) and *Ogeu*IscB proteins (**e**; PDB: 7UTN). **c**, **f** Structure and mechanistic model of *Neisseria meningitidis* Cas9 (**c**) and *Ogeu*IscB (**f**).

structure of the IscB$^{\omega RNA}$ complex should open opportunities for developing more active versions with potential as genome editing tools.

In summary, IscB proteins share a similar domain organization with Cas9, with the contribution by Schuler et al. providing the first comprehensive view of the high-resolution structure of dsDNA-bound IscB$^{\omega RNA}$ complex, thereby revealing the dsDNA targeting mechanism underlying TAM recognition, R-loop formation and DNA cleavage by IscB$^{\omega RNA}$ RNP complex. Structural comparison between IscB and Cas9 indicates the evolutionary trend involving replacement of structured RNA in IscB with protein domains in Cas9. These findings reveal the architectural and mechanistic similarities between IscB and Cas9, validating the evolutionary relationship between IscB and Cas9. Furthermore, the small size of IscB, which is about two fifths of the size of Cas9, makes it a promising genome editing tool, especially for applications that are limited by packaging gene sizes, thereby opening opportunities for structure-based design aimed towards generation of much smaller, more robust and site-specfic, as well as increasingly active and programmable versions, of IscB-based genome editing tools.

REFERENCES

1. Barrangou, R. et al. *Science* **315**, 1709–1712 (2007).
2. Marraffini, L. A. & Sontheimer, E. J. *Science* **322**, 1843–1845 (2008).
3. Makarova, K. S. et al. *Nat. Rev. Microbiol.* **18**, 67–83 (2020).
4. Jinek, M. et al. *Science* **337**, 816–821 (2012).
5. Cong, L. et al. *Science* **339**, 819–823 (2013).
6. Sun, W. et al. *Mol. Cell* **76**, 938–952.e5 (2019).
7. Wu, Z., Yang, H. & Colosi, P. *Mol. Ther.* **18**, 80–86 (2010).
8. Kapitonov, V. V., Makarova, K. S. & Koonin, E. V. *J. Bacteriol.* **198**, 797–807 (2015).
9. Altae-Tran, H. et al. *Science* **374**, 57–65 (2021).
10. Karvelis, T. et al. *Nature* **599**, 692–696 (2021).
11. Schuler, G., Hu, C. & Ke, A. *Science* **376**, 1476–1481 (2022).

ADDITIONAL INFORMATION

Correspondence and requests for materials should be addressed to Ning Jia or Dinshaw J. Patel.