RESEARCH HIGHLIGHT

IS607 TnpB is a hypercompact RNA-guided DNA endonuclease

Hui Yang [1]✉ and Dinshaw J. Patel [2]✉

Cell Research (2024) 0:1–2; https://doi.org/10.1038/s41422-024-00973-w

The Zhiwei Huang lab report on a hypercompact IS607 TnpB protein that naturally co-occurs with the conserved RAGATH-18 RNA, thereby identifying a new family of compact RNA-guided DNA endonucleases. Such an IS607 TnpB-based approach for programmable genome editing in both *Escherichia coli* and human cells, combined with its structural and target-adjacent motif divergence from earlier reported IS200/IS605 TnpB systems, highlights the diversity of the non-Cas genome editor toolbox.

Transposable elements (TEs) are mobile DNA sequences capable of moving around or generating copies of themselves in a host organism's genome. The IS200/IS605 and IS607 transposons constitute the simplest mobile elements and are widespread in prokaryotes. These transposons typically contain subterminal left end (LE) and right end (RE) palindromic elements, *tnpA* and *tnpB* genes, and encode a non-coding RNA transcribed from the transposon right-end element, called right-end RNA (reRNA) or ωRNA[1,2] (Fig. 1a, i). For IS200/IS605 transposon, TnpA transposase binds to both the LE and RE elements and catalyzes the cleavage and rejoining of DNA substrates, mediating 'Peel-and-Paste' transposition of single-stranded DNA (ssDNA)[1,2] (Fig. 1a). By contrast, IS607 transposons encode a serine recombinase TnpA for the insertion of double-stranded DNA (dsDNA).[3] Although IS200/IS605 and IS607 transposons encode distinct TnpA proteins, they share *tnpB* genes. TnpB forms a ribonucleoprotein (RNP) complex with ωRNA and induces double-strand breaks. Recent studies have revealed that TnpB is likely the ancestor of Cas12 in prokaryotic CRISPR-Cas immune systems and Fanzor in eukaryotic transposons[1,2,4–6] (Fig. 1e). Similar to Cas12,[7] TnpB and Fanzor recognize the target-adjacent motif (TAM) and guided by ωRNA, cleave target dsDNA using their single RuvC domains[8,9] (Fig. 1b–d, f–h). Due to their RNA-guided DNA-targeting activities, IS605 TnpB and Fanzor enable genome editing in human cells.[1,2,4] The functional and mechanistic details of IS607 TnpB systems are largely unknown.

Ren et al.[10] identified the co-occurrence of a conserved 50–60-nt non-coding RNA, called Twister and Hammerhead 18 (RAGATH-18), with IS607 TnpB through the analysis of terabase-scale genomic and metagenomic datasets from the microbiome. These RAGATH-18 RNAs from different species share highly conserved sequences and secondary structures, forming an E-loop and kink-turn motifs (Fig. 1l). Additionally, the 5′- and 3′-flanking regions of RAGATH-18 also share conserved sequences. Further small RNA-sequencing revealed that *Firmicutes bacterium* AM43-11BH IS607 TnpB (ISFba1 TnpB) forms a stable RNP complex with the reRNA encoded within the transposon (Fig. 1j). Similar to IS200/IS605 TnpB, IS607 TnpB acts as RNA-guided and TAM-dependent endonuclease, with the ~20-nt 3′ end variable region serving as a guide for sequence-specific DNA targeting. Sequencing data established the requirement of either 5′-AGGAG or 5′-GAGGG TAM sequences in 22 tested IS607 TnpB systems from different bacterial strains. Further biochemical characterization demonstrated that ISFba1 TnpB generates 5′-overhang double-strand breaks at the TAM-distal region using its single RuvC domain. Ren et al. then elucidated the optimal conditions, including temperature, salt concentration, guide length, and the dependence on divalent ions. Additionally, ISFba1 TnpB exhibited target DNA-activated collateral *trans*-cleavage on non-specific ssDNA,[10] which is also observed for IS605 TnpB[9] and Cas12.[7]

As a reRNA-guided endonuclease, IS607 TnpB exhibited programmable and robust genomic DNA interference activity in *E. coli* cells and genome editing ability for endogenous genes in human cells.[10] Among the 37 tested IS607 TnpBs from different species, ISFba1 TnpB showed comparable gene editing efficiency in human cells to those originally reported for SpCas9 and LbCas12a. Single or double mismatches within the 1–13-nt TAM-adjacent base pair region between the reRNA guide and target DNA impaired or abolished the editing efficiency, indicating high specificity in gene editing.

Ren et al. further determined the cryo-EM structure of the ISFba1 TnpB–reRNA–target DNA ternary complex, elucidating the molecular mechanisms defining RNP assembly and target DNA recognition[10] (Fig. 1j–l). Similar to IS605 *Deinococcus radiodurans* ISDra2 TnpB[8,9] and *Spizellomyces punctatus* Fanzor1 (SpuFz1),[4] ISFba1 TnpB adopts a bilobed architecture including REC and NUC lobes. The TAM-interacting domain (TID), REC1, REC2 and RuvC domains of ISFba1 TnpB form the central channel to accommodate the guide:target duplex, with the REC1 and TID domains responsible for sequence-specific recognition of TAM-containing duplex. The reRNA scaffold, composed of 5′ flanking, RAGATH and 3′ flanking motifs, is stabilized by extensive hydrophobic and hydrophilic interactions with ISFba1 TnpB (Fig. 1k, l). The residues involved in nucleic acid recognition are conserved across IS607 TnpBs from different species, suggesting a conserved reRNA scaffold and binding mode in IS607 TnpB systems.[10]

Compared with SpuFz1,[4] ISFba1 TnpB–reRNA exhibits differences in domain composition and architecture of both protein and reRNA, including smaller REC1 and REC2 domains and additional SL1 and SL3 motifs within the reRNA scaffold[10] (Fig. 1f–h, j–l). Although IS607 ISFba1 TnpB[10] and IS605 ISDra2 TnpB[8,9] both belong to transposon family, they exhibited distinct features on both protein

[1]Key Laboratory of RNA Innovation, Science and Engineering, Shanghai Institute of Biochemistry and Cell Biology, Center for Excellence in Molecular Cell Science, Chinese Academy of Sciences, University of Chinese Academy of Sciences, Shanghai, China. [2]Structural Biology Program, Memorial Sloan-Kettering Cancer Center, New York, NY, USA. ✉email: yanghui@sibcb.ac.cn; pateld@mskcc.org

Published online: 14 May 2024

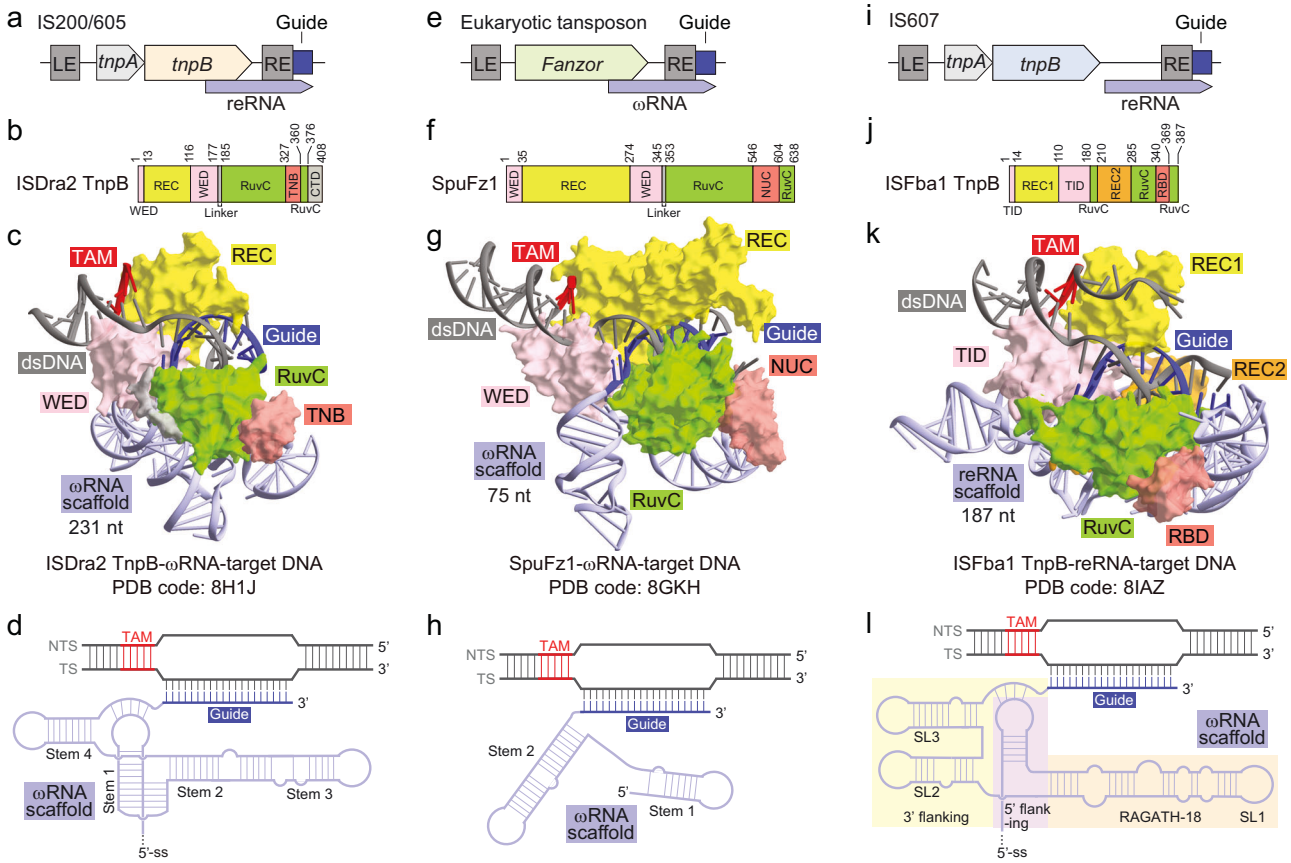

Fig. 1 Naturally occurring RNA-guided DNA-targeting TnpB and Fanzor systems. **a**, **e**, **i** Depiction of locus architecture of IS200/605 TnpB (**a**), eukaryotic Fanzor (**e**) and IS607 TnpB transposons (**i**). LE and RE correspond to left-end and right-end transposon elements. The guide elements and transcribed ωRNA/reRNA are indicated, respectively. **b**, **f**, **j** Domain organization of ISDra2 TnpB (**b**), SpuFz1 (**f**) and ISFba1 TnpB (**j**). **c**, **g**, **k** Overall structures of ISDra2 TnpB (**c**), SpuFz1 (**g**) and ISFba1 TnpB (**k**) in complex with the corresponding reRNA/ωRNA and target dsDNA. The protein components are shown as surface. **d**, **h**, **l** Schematic diagram of the ωRNA/reRNA and target DNA in ISDra2 TnpB (**d**), SpuFz1 (**h**) and ISFba1 TnpB (**l**) systems. The structural elements of reRNA in ISFba1 TnpB system are indicated in colored backgrounds (**l**). TS target DNA strand, NTS non-target DNA strand.

architectures and reRNA scaffolds. The SL1 and SL3 motifs of reRNA, as well as the SL3-interacting helix in the TID domain and SL1-interacting reRNA-binding domain (RBD), are missing in ISDra2 TnpB–reRNA complex (Fig. 1b–d, j–l). Moreover, the protein–nucleic acid recognition modes involving TAM region are different, resulting in G-rich TAM preference for ISFba1 TnpB[10] and AT-rich TAM preference for ISDra2 TnpB.[8,9] The substantial structural differences in protein counterpart and reRNA scaffold suggest that IS607 and IS605 TnpBs belong to two different clades.

Collectively, Ren et al. identified and biochemically characterized IS607 TnpB, which naturally co-occurs with RAGATH-18, as reRNA-guided DNA endonuclease.[10] IS607 TnpB forms a stable RNP complex with reRNA and cleaves dsDNA using its single RuvC domain. Importantly, IS607 TnpB exhibited genome editing activities in bacterial and human cells and displayed high sensitivity to base mismatches within the guide:target duplex. Additionally, IS607 TnpB exhibits structural divergence from IS605 TnpB and displays a distinct requirement for the G-rich TAM, which would expand the targeting range of TnpB family proteins. The robust editing activity of IS607 TnpB, combined with its small size (~390 aa), suggests its great potential as a hypercompact genome editor. Nevertheless, future trimming of non-essential elements within its reRNA component should further facilitate entry into cells. The widespread distribution and diversity of TnpB systems from prokaryotes to eukaryotes suggest functional diversity across life cycles. Elucidating the biology of these RNA-guided and TAM-dependent endonucleases would increase the diversity of the non-Cas genome editor toolbox.

REFERENCES

1. Altae-Tran, H. et al. *Science* **374**, 57–65 (2021).
2. Karvelis, T. et al. *Nature* **599**, 692–696 (2021).
3. Chen, W. et al. *Elife* **7**, e399611 (2018).
4. Saito, M. et al. *Nature* **620**, 660–668 (2023).
5. Yoon, P. H. et al. *Nucleic Acids Res.* **51**, 12414–12427 (2023).
6. Jiang, K. et al. *Sci. Adv.* **9**, eadk0171 (2023).
7. Wang, J. Y. et al. *Nat. Rev. Microbiol.* **20**, 641–656 (2022).
8. Nakagawa, R. et al. *Nature* **616**, 390–397 (2023).
9. Sasnauskas, G. et al. *Nature* **616**, 384–389 (2023).
10. Ren, K. et al. *Cell Res.* https://doi.org/10.1038/s41422-024-00952-1 (2024).

COMPETING INTERESTS

The authors declare no competing interests.

ADDITIONAL INFORMATION

Correspondence and requests for materials should be addressed to Hui Yang or Dinshaw J. Patel.