Nucleic Acids Research, 2026, **54**, gkag849
https://doi.org/10.1093/nar/gkag849
Gene regulation, Chromatin and Epigenetics

OXFORD

Dual-mode application of transcriptional regulation and DNA editing using the ultra-compact TnpB system

Yeounsun Oh[1], Se-Been Jeon[2,3], Lee Wha Gwon[4,5], Bo-Woong Sim[2,6], Seung Hwan Lee[1,*]

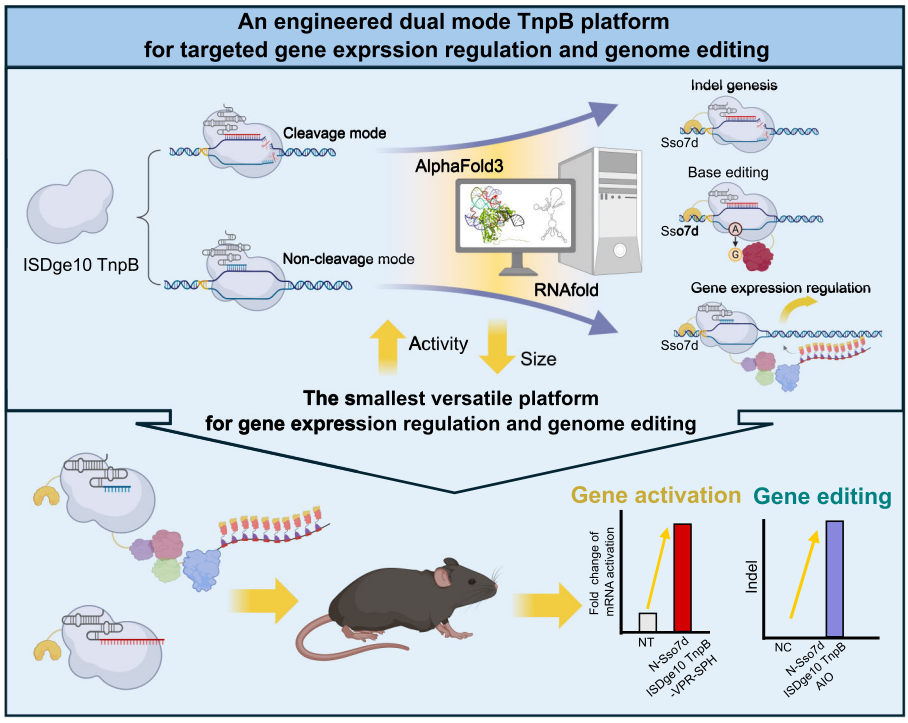
[1]Department of Life Science, Chung-Ang University, Seoul 06974, Republic of Korea
[2]Futuristic Animal Resource & Research Center, Korea Research Institute of Bioscience and Biotechnology, Cheongju 28116, Republic of Korea
[3]Department of Animal Science, College of Natural Resources & Life Science, Pusan National University, Miryang 50463, Republic of Korea
[4]National Primate Research Center, Korea Research Institute of Bioscience and Biotechnology, Cheongju 28116, Republic of Korea
[5]KRIBB School of Bioscience, University of Science and Technology, Daejeon 34113, Republic of Korea
[6]Department of Functional Genomics, University of Science and Technology, Daejeon 34113, Republic of Korea

*To whom correspondence should be addressed. Email: lsh080390@cau.ac.kr

Abstract

RNA-guided obligate mobile element guided activity systems derived from transposable elements have emerged as compact genome-editing tools that may replace clustered regularly interspaced short palindromic repeats platforms. We established a dual-mode genome regulation platform using ISDge10 TnpB effectors and engineered ωRNAs, in which modulation of the ωRNA guide length enables switching between programmable transcriptional activation and genome editing. TnpB programmed with a 10-nt guide region of the ωRNA engages target DNA without inducing double-strand breaks. Fusion of transcriptional activators with Sso7d (DNA-binding protein from *Sulfolobus solfataricus*) enables specific transcriptional upregulation across endogenous loci. Restoring the ωRNA guide length to 20 nt triggers DNA cleavage, thereby supporting homology-directed repair-mediated sequence correction. A catalytically inactivated TnpB-based adenine base editor enabled A-to-G base conversion at genomic targets. TnpB shows strict ωRNA-dependent mismatch sensitivity with low off-target effects, suggesting its potential as a high-fidelity genome regulation platform. Compact ISDge10 TnpB facilitates co-packaging of effector and ωRNA in a single adeno-associated virus vector and co-expression of large functional domains. Thus, this study expands RNA-guided genome-editing capabilities.

Graphical abstract



Received: February 12, 2026. **Revised:** July 25, 2026. **Accepted:** August 11, 2026
© The Author(s) 2026. Published by Oxford University Press.
This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial License (https://creativecommons.org/licenses/by-nc/4.0/), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited. For commercial re-use, please contact reprints@oup.com for reprints and translation rights for reprints. All other permissions can be obtained through our RightsLink service via the Permissions link on the article page on our site—for further information please contact journals.permissions@oup.com.

Downloaded from academic.oup.com/nar/article/54/16/gkag849/8770891 by Chung Ang University, Anseong Campus Library user on 16 October 2026

Introduction

Recently, following the discovery of clustered regularly interspaced short palindromic repeats (CRISPR) systems over the past decade, guide RNA-based reprogramming has enabled targeted editing and gene expression regulation of specific genes [1, 2]. CRISPR-based studies have proliferated across diverse gene-targeted research fields and therapeutic applications in humans [3–8]. Beyond CRISPR-Cas9, the prototypical module of the CRISPR system known for target gene recognition, various alternative DNA recognition modules have been identified, including CRISPR-Cas12 systems that use distinct mechanisms of DNA targeting, as well as obligate mobile element guided activity (OMEGA) effector systems derived from transposon components [9–11]. The discovery of such diverse DNA recognition modules has been used in the development of gene-editing technologies designed to complement and enhance the efficiency and precision of the conventional CRISPR-Cas9 system [12, 13].

Among the recently characterized RNA-guided nucleases, OMEGA-TnpB (hereafter referred to as TnpB) is a transposon-derived effector encoded by insertion sequence (IS) elements that has emerged as a promising platform owing to its potential applicability in genome engineering [14, 15]. TnpB homologs are widely distributed across bacterial and archaeal lineages in association with diverse IS families and function with small non-coding RNAs known as ωRNAs [11, 14]. Comparative phylogenetic and structural analyses indicate that TnpB constitutes an ancestral endonuclease lineage whereby type V CRISPR effectors, including Cas12 enzymes, subsequently diverged [11, 16]. TnpB was initially described as an ultracompact nuclease that facilitates the maintenance and mobility of transposable genetic elements. However, it is now recognized as a broadly distributed and functionally diverse family of reprogrammable endonucleases with distinctive mechanistic and structural features [15, 17–19]. In addition to RNA-guided nucleases, the TnpB protein family also includes catalytically inactive TnpB-like nuclease-dead repressors, which retain RNA-guided DNA binding and function as transcriptional repressors [20], highlighting the functional diversification of TnpB-like proteins. Many TnpB proteins comprise <400 amino acids, placing them among the smallest programmable DNA editors reported to date [21], a feature that is advantageous for human-optimized but delivery-constrained systems, such as adeno-associated virus (AAV) vectors [13, 22–25].

At the mechanistic level, TnpB functions as an RNA-directed DNA endonuclease that generates double-strand breaks proximal to transposon-associated motifs (TAMs) [15]. Although its catalytic fold shares evolutionary similarities with those of Cas12-family nucleases, TnpB exhibits substantial variability in both target recognition stringency and cleavage site determination [17, 26]. This heterogeneity is interpreted as a consequence of selective pressures imposed by its transposon-associated functional role, compared to that of the CRISPR systems [11, 16, 27]. Recent biochemical and structural studies have further demonstrated that the TnpB–ωRNA ribonucleoprotein (RNP) complex adopts a streamlined yet adaptable architecture [13, 15, 17, 19, 28]. This indicates that complex RNA scaffolding is not strictly required for programmable DNA cleavage [13] and suggests that the genome-editing capabilities of this enzyme may be further diversified through rational engineering of the TnpB system [17, 19].

In this study, we developed a technology for finely tuning gene expression regulation and genome editing in the human system using the ISDge10 lineage of miniaturized TnpB, which recognizes the shortest TAM (TTAT) on-target DNA sequences, combined with engineered ωRNAs and functional domains. We demonstrated the conceptual application of miniaturized TnpB for gene targeting and gene expression regulation by simple switching of the guide length of the ωRNA. From an applied perspective, TnpB offers several advantages as a genome-engineering tool, most prominently its markedly smaller size relative to the widely used CRISPR-Cas9 system. This ISDge10 TnpB-based bifunctional platform is optimal for AAV delivery and is expected to enable effective gene therapies for various genetic disorders arising from dysregulated expression or functional defects caused by mutations within human chromosomes. Furthermore, it is anticipated to exert a broad impact on biological research aimed at elucidating gene function, reshaping current trends in advanced genome editing.

Methods

Cell culture

Human-derived cell lines HEK293FT (Invitrogen, R70007), U2OS (ATCC, HTB-96), K562 (ATCC, CCL-243), and AAVpro 293T (Clontech, 632273), as well as the mouse-derived cell line NIH3T3 (ATCC, CRL-1658), were maintained in Dulbecco's modified Eagle's medium (DMEM) (Thermo Fisher Scientific, 11995073) supplemented with 10% fetal bovine serum (Thermo Fisher Scientific, A5669701) at 37°C with 5% CO_2. Cells were passaged every 48 h to maintain ~70% confluency. AAVpro 293T cells were used for AAV packaging.

Bacterial culture/plasmid propagation

Escherichia coli DH5α competent cells (Enzynomics, CP011) were transformed with plasmids for general propagation and grown in LB medium supplemented with appropriate antibiotics at 37°C with shaking. *Escherichia coli* Stbl3 competent cells (Thermo Fisher Scientific, C737303) were transformed with AAV GOI plasmids and cultured under the same conditions for plasmid amplification.

Design and cloning of the ISDge10 TnpB transcriptional activator and gene editor with ωRNA expression vectors

To enable efficient gene editing and transcriptional activation in human cell lines, we used cytomegalovirus (CMV) promoter-driven expression constructs encoding ISDge10 TnpB and ISDge10 TnpB (WT/D187A/D351A)-VPR fusion proteins. To enhance the functional performance of ISDge10 TnpB-VPR, we systematically engineered multiple variants by incorporating linker sequences of different lengths (Supplementary Fig. S4B and Supplementary Table S8). To construct expression vectors incorporating the Sso7d DNA-binding domain into the ISDge10 TnpB system, Sso7d was inserted at the N- or C-termini of ISDge10 TnpB and ISDge10 TnpB-VPR. This resulted in a series of fusion proteins including Sso7d-ISDge10 TnpB, ISDge10 TnpB-

Downloaded from academic.oup.com/nar/article/54/16/gkag849/8770891 by Chung Ang University, Anseong Campus Library user on 16 October 2026

Sso7d, Sso7d-ISDge10 TnpB-VPR, and ISDge10 TnpB-VPR-Sso7d, each with multiple versions for comparative evaluation. All expression plasmids were designed as AIO systems, enabling the simultaneous expression of both the effector protein and corresponding ωRNA (Supplementary Table S1). To develop a minimal ωRNA system compatible with TnpB function, specific regions of the ωRNA were sequentially deleted through *in silico*-based stem-loop prediction. The secondary structure of ωRNA was predicted using the RNAfold webserver (http://rna.tbi.univie.ac.at/cgi-bin/RNAWebSuite/RNAfold.cgi), which predicts RNA secondary structures using thermodynamic nearest-neighbor parameters and dynamic programming to find the minimum free energy conformation, with optional partition function calculations for base-pair probabilities. For gene activation experiments, target sites were selected at various genomic positions relative to the transcription start site (TSS), and the corresponding ωRNAs were designed to match these loci. The activity of these ωRNAs was assessed by quantifying the expression levels of the target genes. The ωRNA expression cassette, driven by a U6 promoter, was designed to allow the straightforward replacement of the guide sequence of the ωRNA using restriction enzyme sites. This modular design enables the rapid customization of ωRNA to target different genomic loci.

Cell transfection

For targeted gene editing and regulation, cells were transfected in 24-well plates, with each well initially seeded with 0.5×10^5 cells. After seeding by 24 h, plasmids expressing guide RNAs (240 pmol), transcriptional activator expression plasmids ISDge10 TnpB or ISDge10 TnpB-VPR (developed in this study; 60 pmol), and reporter plasmids (when required; 40.7 fmol) were delivered to cells at 60%–80% confluency with 1.5 μl Lipofectamine 3000 reagent and 1 μl p3000 reagent according to the manufacturer's protocols (Thermo Fisher Scientific, L3000-015). The cell lines used in this study were authenticated. RNA extraction and fluorescence assays were performed 48–72 h post-transfection; DNA extraction was performed 72 h post-transfection.

Fluorescence reporter assay

Fluorescence reporter assays were performed using a miniCMV-based dTomato reporter system to evaluate ISDge10 TnpB-VPR-mediated transcriptional activation. Various ωRNA guide lengths (8, 10, 12, 14, 16, 18, and 20-nt for wild-type (WT) TnpB, and 20-nt for the D187A mutant) and different ISDge10 TnpB-VPR constructs (ISDge10-VPR, N-Sso7d-TnpB-VPR, and C-Sso7d-TnpB-VPR) were tested. The reporter construct, based on Addgene plasmid #47320, drives dTomato expression under the control of a miniCMV promoter, with the target sequence and TAM positioned immediately upstream of the promoter to enable TnpB-VPR binding and its activation. The sequences used for TnpB-VPR targeting are shown in Supplementary Table S2. Transfected cells were imaged using a fluorescence microscope 48–72 h post-transfection. Gene expression activity was quantified by measuring background-corrected dTomato fluorescence intensity, with the fluorescence signal of the negative control (NC) used as the background and subtracted from each sample. EGFP (green) was co-expressed with each TnpB construct within the same expression vector to facilitate comparison among samples, and the degree of gene expression activation was evaluated by dTomato fluorescence (red) based on the normalized values (extended data files 1 and 2). The mean fluorescence intensity was quantified using ImageJ software by defining regions of interest.

RNA extraction and quantitative PCR to evaluate transcriptional activation

Total RNA was extracted from the cultured cells 48 h post-transfection using the RNeasy Mini Kit (Qiagen). Subsequently, 1 μg of RNA from each sample was used for complementary DNA (cDNA) synthesis with the iScript cDNA synthesis Kit (BioRad), according to the manufacturer's protocol. The resulting cDNA was used in each quantitative polymerase chain reaction (qPCR) reaction. Quantification of specific gene expression was achieved using qPCR with THUNDERBIRD™ Next SYBR® qPCR Mix (Toyobo). qPCR was performed and analyzed using a CFX96 Real-Time PCR Detection System (BioRad). To evaluate transcriptional activation, cDNA was amplified using the qPCR primers listed in Supplementary Table S3. The expression levels of each targeted gene were normalized to the cycle threshold (Ct) of the housekeeping gene GAPDH, yielding ΔCt values. With the Ct method, ΔCt was calculated as ΔCt = Ct (target gene) − Ct (housekeeping gene [GAPDH]), and ΔΔCt was determined as ΔΔCt = ΔCt (experimental sample) − ΔCt (control sample), with the non-targeting experimental sample used as the control. All reactions were performed in triplicate, and the relative expression of each sample was computed as $2^{-\Delta Ct}$ or $2^{-\Delta\Delta Ct}$.

Purification of gDNA and construction of target site amplicons

Genomic DNA (gDNA) was extracted from the cultured cells 72 h after genome editing and was isolated using the DNeasy Blood and Tissue Kit (Qiagen, 69506). Target amplicons were obtained through PCR denaturation at 98°C for 10 s, primer annealing at 58°C for 10 s, elongation at 68°C for 5 s/kb, 35 cycles from gDNA extracted from cells treated with each ISDge10 TnpB editor. Targeted amplicon next-generation sequencing was performed using nested PCR (denaturation: 98°C for 10 s, primer annealing: 58°C for 5 s, and elongation: 68°C for 5 s/kb, 35 cycles) to analyze the gene editing efficiency.

Targeted amplicon sequencing and data analysis

To evaluate mutations at the target genomic locus, gDNA was extracted from cells and amplified using the DNA primers (Supplementary Table S4). Nested PCR (denaturation at 98°C for 10 s, primer annealing at 62°C for 5 s, and elongation at 68°C for 5 s/kb, 35 cycles) was performed to conjugate adapter and index sequences to the amplicons. The resulting barcoded libraries were sequenced on an Illumina MiniSeq system (Illumina, SY-420-1001), and sequencing data were analyzed using Cas-Analyzer (http://www.rgenome.net/cas-analyzer/#!).

Off-target analysis

Potential off-target site candidates were selected using Cas-OFFinder to identify genome-wide off-target sites

Downloaded from academic.oup.com/nar/article/54/16/gkag849/8770891 by Chung Ang University, Anseong Campus Library user on 16 October 2026

(Supplementary Tables S5 and S7). For ISDge10 TnpB-VPR-mediated gene activation, high-probability off-target sites with the same sequence located in coding genes were selected and evaluated by qPCR. For ISDge10 TnpB-mediated genome editing, off-target candidates predicted *in silico* with up to four mismatches were analyzed to determine their editing efficiencies.

RNA sequencing and data analysis

Total RNA was isolated using the RNeasy Plus Mini Kit (Qiagen, 74106). Libraries were prepared from total RNA using the CORALL RNA-Seq V2 Library Prep Kit (Lexogen, Inc., Austria). Messenger RNA (mRNA) was isolated using the Poly(A) RNA Selection Kit (Lexogen, Inc.). The isolated mRNAs were used for cDNA synthesis and shearing, following the manufacturer's instructions. Indexing was performed using the Illumina indexes 1–12. The enrichment step was performed using PCR. Subsequently, the libraries were checked using a TapeStation HS D1000 Screen Tape (Agilent Technologies, Amstelveen, The Netherlands) to evaluate the mean fragment size. Quantification was performed using a Library Quantification Kit and a StepOne Real-Time PCR System (Life Technologies, Inc., USA). High-throughput sequencing was performed as paired-end 150 sequencing using NovaSeq XP (Illumina, Inc., USA). Quality control of raw sequencing data was performed using FastQC (Simon, 2010). Adapters and low-quality reads were removed using Fastp (Chen *et al.*, 2018). The trimmed reads were then mapped to the reference genome using STAR (Dobin *et al.*, 2013) and quantified using Salmon (Patro *et al.*, 2017). The read counts were processed based on the TPM normalization method using the Python "conorm" package (Meshcheryakov, 2021). Data mining and graphic visualization were performed using ExDEGA (Ebiogen Inc., Seoul, Korea).

AAV production and cell transduction

AAVpro 293T cells (Clontech) were maintained in DMEM (Thermo Fisher Scientific, 11995073) supplemented with 10% fetal bovine serum (Thermo Fisher Scientific, A5669701) at 37°C in the presence of 5% CO_2. The coding sequences of ISDge10 TnpB-VPR, ωRNA, all-in-one (AIO) ISDge10 TnpB–ωRNA, and split SpCas9 constructs (SpCas9 N-terminal fragment fused to a C-intein and SpCas9 C-terminal fragment fused to the Npu N-terminal intein fragment) were cloned into AAV vector plasmids containing inverted terminal repeats. The single-guide RNA (sgRNA) expression cassette, driven by the U6 promoter, was incorporated into a SpCas9 C-terminal vector. Expression of ISDge10 TnpB-VPR and SpCas9 fragments was driven by CMV-derived promoters (CMV or CMVd) or the CMV enhancer/chicken β-actin (CAG) promoter, whereas ωRNA was expressed under the control of the U6 promoter. Recombinant adeno-associated virus serotype DJ (rAAV-DJ) particles were produced by transient transfection of AAVpro 293T cells using TransIT-VirusGEN (Mirus Bio, MIR 6700). Cells were co-transfected with pAAV-ISDge10 TnpB-VPR, pAAV-ωRNA, pAAV-ISDge10 TnpB AIO, pAAV-SpCas9 N-term, or pAAV-SpCas9 C-term plasmids together with the packaging plasmid pAAV-Rep Cap-DJ and a helper plasmid. At 72 h post-transfection, cells were harvested and lysed, and recombinant AAV particles were purified using the AAVpro® Purification Kit (Takara Bio, 6666) according to the manufacturer's instructions. Viral vector genome titers were determined by real-time qPCR using the AAVpro® Titration Kit for Real-Time PCR (Takara Bio, 6233). For rAAV-mediated gene activation and editing, HEK293FT cells were co-infected with rAAV-DJ particles at a multiplicity of infection of 1×10^4 genome copies/cell. At 4 days post-transduction, total RNA was extracted to measure target gene expression by real-time qPCR, and gDNA was isolated to assess the gene-editing efficiency at 4 and 14 days post-transduction.

Purification of TnpB–ωRNA ribonucleoprotein complex and *in vitro* DNA cleavage assay

The pET28a expression vector encoding the ISDge10 TnpB–ωRNA complex under the control of the T7 promoter was transformed into *E. coli* BL21(DE3) cells. A single colony was cultured at 37°C until the optical density at 600 nm reached 0.6–0.8, after which protein expression was induced with isopropyl β-D-1-thiogalactopyranoside (IPTG) at 16°C overnight. Cells were harvested by centrifugation and resuspended in lysis buffer containing 20 mM Tris–HCl (pH 8.0), 300 mM NaCl, 10 mM β-mercaptoethanol, 1% Triton X-100, and 1 mM phenylmethylsulfonyl fluoride. The cell suspension was lysed by sonication on ice for 3 min, and the lysate was clarified by centrifugation at 20 000 × *g* for 15 min at 4°C. The clarified lysate was incubated with Ni-NTA resin pre-equilibrated with buffer B (20 mM Tris–HCl, pH 8.0, 300 mM NaCl) for 1.5 h at 4°C with gentle agitation. After extensive washing with buffer B to remove non-specifically bound proteins, the TnpB–ωRNA complex was eluted using buffer C (20 mM Tris–HCl, pH 8.0, 300 mM NaCl, and 200 mM imidazole). The eluted RNP complex was exchanged into storage buffer (200 mM NaCl, 50 mM HEPES, pH 7.5, 1 mM DTT, and 40% glycerol) using Amicon Ultra centrifugal filters (Merck, UFC901024). Purity of the purified TnpB–ωRNA RNP complex was confirmed by 10% SDS-PAGE, and protein concentration was determined using the Bradford assay. For the *in vitro* DNA cleavage assay, 151.8 nM purified TnpB–ωRNA RNP complex was incubated with 45 nM PCR-amplified target DNA in reaction buffer containing 10 mM Tris–HCl (pH 7.5), 10 mM $MgCl_2$, 1 mM DTT, 1 mM ethylenediaminetetraacetic acid, and 100 mM NaCl at 37°C for 60 min. Following incubation, DNA was purified using a QIAquick PCR Purification Kit (Qiagen, 28106) and analyzed by agarose gel electrophoresis followed by ethidium bromide staining. Band intensities corresponding to cleaved and uncleaved DNA fragments were quantified using ImageJ (NIH). DNA cleavage efficiency was calculated as the percentage of the combined intensity of the two cleaved fragments relative to the total intensity of the cleaved and uncleaved DNA bands.

$$\text{Cleavage}(\%) = \frac{\text{Intensity1(cleaved fragment1)} + \text{Intensity2(cleaved fragment2)}}{\text{Intensity1(cleaved fragment1)} + \text{Intensity2(cleaved fragment2)} + \text{intensity3(uncleaved band)}} \times 100. \quad (1)$$

Whole genome sequencing

gDNA was used to construct whole-genome sequencing (WGS) libraries using the Illumina TruSeq DNA PCR-Free Library Prep Kit (Illumina, San Diego, CA, USA) according to the manufacturer's instructions. WGS was performed on an Illumina NovaSeq X Plus platform to generate genome-wide sequencing data. Sequencing reads were aligned to the human reference genome, and genomic variants were identified using a standard variant-calling pipeline.

Downloaded from academic.oup.com/nar/article/54/16/gkag849/8770891 by Chung Ang University, Anseong Campus Library user on 16 October 2026

All detected variants were retained without variant allele frequency filtering and were included in the subsequent analyses.

Animal studies

All animal experiments were approved by the Institutional Animal Care and Use Committee (IACUC) (CE26177, CE26376). C57BL/6 mice were maintained under controlled environmental conditions in an animal facility at 23 ± 3°C with a relative humidity of 50 ± 20%, 10–15 air changes per hour, a 12 h light/dark cycle, and an illumination intensity of 150–200 lux. Mice were housed two per individually ventilated polysulfone cage (IVC) containing PG3 beta chip bedding.

In vivo AAV delivery

Six-week-old male C57BL/6 mice were intravenously injected with AAV vectors encoding ISDge10 TnpB. A single dose of 8.12×10^{10} vector genomes (vg) per mouse was administered via tail vein injection. Prior to injection, the tail was warmed using a heat lamp for ∼1 min to dilate the tail vein. Injections were performed using a 31G insulin syringe. At 2 and 6 weeks post-injection, mice were euthanized, and liver tissues were collected. Mice were perfused with phosphate buffered saline (PBS) to remove blood prior to liver harvesting. gDNA was extracted from liver tissues and used to assess genome-editing efficiency.

Hydrodynamic tail vein injection

Plasmid delivery was performed using hydrodynamic tail vein injection (HTVI). Seven-week-old male C57BL/6 mice were used. A total of 10 μg plasmid DNA was rapidly injected within 5–8 s via the tail vein using a 1 ml syringe fitted with a 26G needle. One week after a single injection, mice were euthanized, perfused with PBS, and livers were harvested. Liver tissues were used for gDNA extraction. For genome activation experiments, 10 μg each of the reporter plasmid and N-Sso7d-ISDge10-VPR-SPH plasmid were co-delivered by HTVI. Three days after injection, liver tissues were harvested for flow cytometric analysis. The excised liver was washed with cold PBS and minced into fragments <1–2 mm. Tissue fragments were digested at 37°C in digestion buffer containing Collagenase IV and DNase I. The resulting cell suspension was gently dissociated by pipetting, filtered through a 70-μm cell strainer, and, when necessary, treated with ACK lysis buffer to remove red blood cells. Cells were subsequently washed with fluorescence-activated cell sorting (FACS) buffer before analysis.

Flow cytometry analysis

Single-cell suspensions were analyzed for dTomato expression using flow cytometry. Cells were washed with FACS buffer and resuspended at an appropriate concentration for analysis. DAPI was added immediately prior to analysis at the manufacturer's recommended concentration and incubated for 5–10 min at room temperature in the dark. DAPI-positive cells were detected using a 405-nm laser and the DAPI channel and were considered non-viable and excluded from analysis. Cells were analyzed without fixation. dTomato-positive cells were detected using the PE-Texas Red channel or an equivalent fluorescence channel suitable for dTomato detection on the instrument used.

Quantification and statistical analysis

All statistical analyses were performed using GraphPad Prism version 11.0.0 (GraphPad software, Boston, MA, USA). Data are presented as mean ± standard error (SEM) from at least three independent experiments. The number of independent biological replicates and technical replicates for each experiment are described in the figure legends or the methods section.

Additional resources

Prism https://www.graphpad.com/
SnapGene https://www.snapgene.com/
The PyMOL Molecular Graphics System, Version 3.0 https://www.pymol.org/
RNAfold WebServer http://rna.tbi.univie.ac.at/cgi-bin/RNAWebSuite/RNAfold.cgi
ImageJ https://imagej.nih.gov/ij/
Cas-Analyzer http://www.rgenome.net/cas-analyzer/

Results

Validation of dual-mode TnpB function for gene regulation and genome editing through ωRNA guide length modulation

To optimize the recently identified OMEGA system as a genome-editing tool for human-derived cell lines, we investigated the DNA-targeting properties of the ISDge10 lineage TnpB, which is characterized by the TTAT recognition sequence [26, 29], favoring precise gene targeting (Fig. 1). First, we sought to distinguish between the DNA targeting and subsequent cleavage activities of ISDge10 TnpB according to variations in guide part of the ωRNA length (Fig. 1A). To generate a compact gene regulator, we fused ISDge10 TnpB to the tripartite transcriptional activation domain VPR (VP64-p65-Rta). Therefore, we adopted the same gene expression context previously used in CWCas12f-VPR-based studies [30], substituting the effector with WT ISDge10 TnpB (hereafter referred to as TnpB-VPR). To establish a catalytically inactive TnpB variant suitable for transcriptional activation, we first sought to identify the conserved catalytic residues of ISDge10 TnpB. We performed a broader phylogenetic and sequence conservation analysis across diverse TnpB family members. Multiple sequence alignment revealed that the catalytic motifs are highly conserved among TnpB homologs despite overall sequence divergence (Supplementary Fig. S1A, right). Although ISDge10 TnpB is not the closest homolog of ISDra2 TnpB, both proteins belong to the same evolutionary lineage and share conserved catalytic motifs (Supplementary Fig. S1A,left). Based on the conserved catalytic residues identified in the experimentally characterized ISDra2 TnpB, we identified the corresponding putative catalytic residues in ISDge10 TnpB as Asp187 (D187) and Asp351 (D351). To further investigate structural conservation, the three-dimensional structure of ISDge10 TnpB was predicted using AlphaFold3 [31] and aligned with the experimentally determined ISDra2 TnpB structure (PDB: 8BF8) [15]. Although ISDge10 and ISDra2 TnpB share only 41.5% sequence identity, structural alignment yielded a Template Modeling score (TM-score) of 0.827, indicating strong conser-

Downloaded from academic.oup.com/nar/article/54/16/gkag849/8770891 by Chung Ang University, Anseong Campus Library user on 16 October 2026

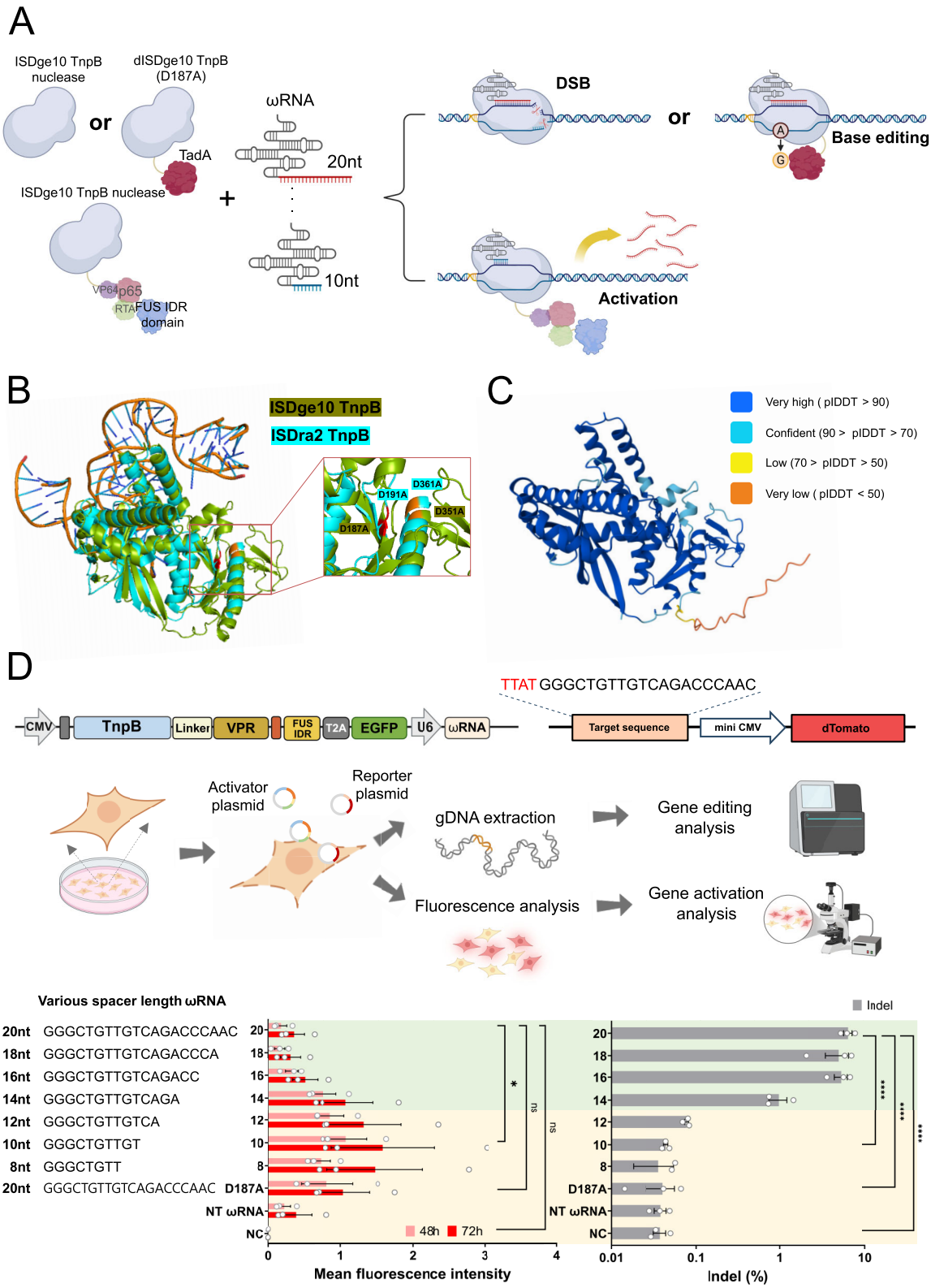


Figure 1. Validation of the bifunctional mode of TnpB for ωRNA-guided target gene editing and transcriptional regulation. (**A**) Schematic illustration of TnpB activity regulated by the guide length of the ωRNA. An ωRNA containing a 20-nt guide induces double-strand DNA breaks (DSBs) or enables base editing when fused to a deaminase, whereas an ωRNA containing a 10-nt guide enables programmable gene regulation without inducing DNA cleavage. Therefore, the TnpB module can be applied to both targeted genome editing and transcriptional activation through guide length switching. A catalytically inactive ISDge10 dTnpB(D187A) variant can be used to induce base editing within target genes when guided by a 20-nt guide length of the ωRNA. Created in BioRender. Lee, S.H. (2026) https://BioRender.com/1hqkh8u. (**B**) Comparative analysis of Alphafold3-predicted structure of ISDge10 TnpB (olive green) and ISDra2 TnpB (cyan) (PDB: 8H1J) with positions aligned to illustrate structural conservation. Catalytic residues of ISDge10 TnpB (red) and ISDra2 TnpB (orange) are indicated. (**C**) AlphaFold3-predicted TnpB structure shown in panel (B), colored according to the pIDDT confidence scores. (**D**) Top: The activator plasmid expresses the TnpB-VPR module linked to an FUS-IDR domain, together with EGFP via a T2A element. The reporter

Downloaded from academic.oup.com/nar/article/54/16/gkag849/8770891 by Chung Ang University, Anseong Campus Library user on 16 October 2026

vation of the overall protein fold (Fig. 1B). Consistent with this result, the AlphaFold3 model of ISDge10 exhibited high predicted local distance difference test (plDDT) confidence scores, supporting the reliability of the predicted structure (Fig. 1C). Structural alignment further showed that the predicted catalytic residues D187 and D351 occupy positions corresponding to the conserved active site of ISDra2 TnpB, demonstrating that the catalytic architecture is well conserved despite sequence divergence. To assess the reliability of the AlphaFold3 predictions, we compared the AlphaFold3-predicted ISDra2 TnpB–ωRNA and ISDge10 TnpB–ωRNA complexes with the experimentally determined cryo-EM structure of ISDra2 TnpB–ωRNA (PDB: 8BF8) (Supplementary Fig. S1B). Structural alignment of the overall complex revealed a high degree of similarity, with a TM-score of 0.96 and root-mean-square deviation (RMSD) of 1.53 Å over 302 aligned residues (Supplementary Fig. S1C). When the protein and ωRNA components were analyzed separately, the TnpB protein showed high structural similarity between the predicted and experimental structures, whereas the ωRNA exhibited substantially lower similarity (TM-score = 0.25, RMSD = 5.93 Å). Structural alignment between the experimentally determined ISDra2 TnpB–ωRNA complex and the AlphaFold3-predicted ISDge10 TnpB–ωRNA complex revealed that the TnpB protein structures are highly conserved, with a TM-score of 0.88 and an RMSD of 2.55 Å (Supplementary Fig. S1D). In contrast, comparison of the predicted ISDge10 ωRNA with the experimentally resolved regions of the ISDra2 ωRNA showed lower structural similarity (TM-score = 0.33, RMSD = 3.51 Å). The experimentally determined ISDra2 ωRNA model contains only the resolved regions, whereas flexible regions are absent from the cryo-EM structure. Despite the lower structural similarity observed for the ωRNA, the TnpB protein scaffold exhibited high structural similarity between the experimental and predicted structures. Based on structural alignment, Asp187 (D187) and Asp351 (D351) of ISDge10 TnpB were identified as putative catalytic residues corresponding to the conserved active site residues in ISDra2 TnpB.

For assays in human-derived HEK293FT cells, plasmids expressing TnpB-VPR and reporter constructs carrying the target DNA sequence were co-delivered (Fig. 1D, extended data file1: Raw data table1, extended data file2). We then quantified both the enhancement of reporter fluorescence and indel frequency at endogenous genomic loci to compare transcriptional activation mediated by binding and cleavage-dependent editing outcomes. Gradual reduction of guide length of the ωRNA from 20 to 8 nt revealed that reporter fluorescence intensity progressively increased from 14 nt and peaked at 10 nt (Fig. 1D, left). Conversely, indel frequencies at endogenous loci decreased from 14 nt onward, with no significant indel induction observed at 10 nt (Fig. 1D,right). Comparative experiments with the evolutionarily related Cas12f system, which shares structural similarity with TnpB, yielded similar reciprocal patterns of reporter activation and indel formation when guide part of the ωRNA was systematically varied (Supplementary Fig. S2A). Quantitative analysis revealed guide length-dependent differences in both transcriptional activation and genome editing activities. Notably, the CWCas12f system programmed with short-spacer guide RNAs retained robust transcriptional activation while exhibiting little or no genome editing activity, a phenotype similar to that of the catalytically inactive D354A variant (Supplementary Fig. S2B). Consistent with these findings, T7 endonuclease I assays performed in HEK293FT cells at 48 and 72 h post-transfection detected indel formation only in cells expressing WT CWCas12f, whereas no detectable cleavage products were observed in cells expressing the D354A mutant (Supplementary Fig. S2C). Accordingly, throughout this study, reporter fluorescence intensity was used as a measure of transcriptional activation mediated by TnpB targeting, whereas indel frequency was used as an indicator of subsequent DNA cleavage activity.

A D187A mutant of ISDge10 TnpB, wherein amino acid substitutions disrupt cleavage domains (Fig. 1B), induced no indels even with 20-nt guide length of the ωRNA, in contrast to that of the WT TnpB (Supplementary Fig. S2D, E). A previous study on ISDra2 TnpB demonstrated that the alanine substitution of the corresponding residue, Asp191 (D191A), eliminated DNA cleavage activity while preserving target DNA binding, enabling its use as a catalytically dead variant [15, 28]. Sequence alignment revealed that Asp187 in ISDge10 TnpB corresponds directly to Asp191 in ISDra2 TnpB and is conserved within the RuvC-I catalytic motif. Considering this strong functional and evolutionary conservation, the D187A substitution in ISDge10 TnpB was designed following the same validated mutational strategy previously applied to ISDra2 TnpB, in which the alanine substitution of the RuvC-I aspartate abolishes catalytic activity while retaining DNA binding. The D187A mutant was selected as a more appropriate inactive TnpB variant than D351A and was used in subsequent experiments. A comparison between catalytically inactive TnpB-VPR (D187A, hereafter referred to as dTnpB-VPR) and TnpB-VPR confirmed that a 10-nt guide length of the ωRNA maximized reporter activation without inducing target DNA cleavage (Fig. 1D). Furthermore, to evaluate endogenous gene activation, we targeted multiple TSS-proximal sites of the human *HBG* gene and compared the fold increase in expression between TnpB-VPR and dTnpB-VPR (Supplementary Fig. S2F). The optimized 10-nt guide length of the ωRNA induced significantly greater gene activation with TnpB-VPR than with dTnpB-VPR. Collectively, these results demonstrate that an optimized 10-nt guide length, when coupled with the cleavage-competent TnpB-VPR module, enables

←

plasmid contained a dTomato expression cassette in which the TnpB target site was inserted immediately upstream of the promoter. Using this system, target DNA-binding efficiency was evaluated based on changes in fluorescence following TnpB-VPR binding, whereas cleavage activity was quantified as editing frequency (%) at the corresponding endogenous loci after extracting gDNA from cells. Bottom: Comparison of target DNA binding and cleavage activity of TnpB according to guide length of the ωRNA. ωRNAs with different guide lengths (8–20 nt) were delivered together with the reporter plasmid to induce transcriptional activation, and the mean dTomato fluorescence intensity was compared as a fold increase relative to a non-target ωRNA control (left histogram). For indel analysis, gene-editing frequencies were measured by NGS at the same binding sequence embedded in the gDNA (right histogram). NC: negative control; NT ωRNA: non-target ωRNA; D187A: inactive ISDge10 TnpB module. Each histogram was plotted by applying the standard error of the mean to the repeated experimental values ($n = 3$). *P*-values were calculated using two-way analysis of variance (ANOVA) and Tukey's test for fluorescence analysis, and one-way ANOVA and Dunnett's test for gene-editing analysis (ns: not significant, $^{*}P = .0332$, $^{**}P = .0021$, $^{***}P = .0002$, $^{****}P < .0001$). Created in BioRender. Lee, S.H. (2026) https://BioRender.com/5vuz6re.

Downloaded from academic.oup.com/nar/article/54/16/gkag849/8770891 by Chung Ang University, Anseong Campus Library user on 16 October 2026

robust activation of target gene expression in human-derived cell lines, highlighting its potential as an effective gene regulation platform.

Minimization and functional optimization of TnpB-VPR through ωRNA engineering

To achieve more effective gene activation using the previously optimized TnpB-VPR, we sought to identify the essential ωRNA regions required for DNA recognition by ISDge10 TnpB (Fig. 2). We engineered the ωRNA scaffold region while retaining the optimized 10-nt guide part of the ωRNA necessary for WT TnpB function (Fig. 2A). To define the 5′-terminal deletion boundaries, the ωRNA secondary structure was predicted based on minimum free energy, and ωRNA variants were designed according to stable hairpin-forming structural elements rather than arbitrary nucleotide positions. This strategy was intended to preserve the integrity of individual structural motifs while systematically evaluating the functional contribution of each 5′-terminal domain. Through *in silico* simulations, we identified stable base-pairing regions within the scaffold that formed stem-loop structures, which we designated as domains D1–D8 (Fig. 2B and Supplementary Fig. S3A). To assess the functionality of each region, we generated an AIO plasmid under the CMV promoter that co-expressed TnpB-VPR with ωRNA and created constructs in which individual or combinatorial domains (ΔD1–ΔD1238) were deleted (Supplementary Table S1). These vectors were delivered alongside fluorescent reporter constructs carrying the target DNA sequence, and both reporter activation and indel frequencies at the endogenous genomic sites were analyzed (Fig. 2C, extended data file1: Raw data table2, and extended data file2). As expected, compared with the 20-nt guide length, the 10-nt guide length ωRNA yielded markedly stronger reporter activation (Fig. 2C, top). Deletion of D4, D5, D6, or D7 reduced reporter expression, whereas removal of D1, D2, D3, or D8 individually or in combination had little effect. Unlike 20-nt guide length ωRNA, none of the truncated ωRNAs with a fixed 10-nt guide length ωRNA (ΔD1–ΔD1238) induced indel formation, indicating the absence of cleavage activity (Fig. 2C,bottom). To further test the functionality of the minimal ωRNA scaffold, we evaluated the ΔD1238 version, which retains only the essential domains, at multiple target sites of *the HBG* endogenous gene (Fig. 2D and Supplementary Tables S2–S4, extended data file1: Raw data table4). When directed to the TSS region of the *HBG* gene, TnpB-VPR with ΔD1238 ωRNA exhibited a gene activation efficiency comparable to that observed with full-length ωRNA. Structural analysis of the AlphaFold3-predicted ISDge10 TnpB–RNA complex demonstrated that the D4–D7 region of the RNA directly contacts the TnpB protein, suggesting a potential role in protein–RNA recognition or stabilization. In contrast, the D1, D2, D3, and D8 regions were structurally exposed to the solution and did not exhibit direct contact with the TnpB protein (Supplementary Fig. S3B and C). These results indicate that domains D4–D7 of the ωRNA scaffold are critical for TnpB-VPR activity; therefore, we selected the ΔD1238 version as the minimal functional scaffold for subsequent experiments. Finally, with the fixed 10-nt guide length ωRNA, we investigated the influence of GC content on targeting efficiency by measuring mRNA expression levels across multiple endogenous sites (Supplementary Tables S2–S4). Consistently, higher GC content within the target DNA correlated with stronger gene activation (Fig. 2E).

Enhanced activation of TnpB-VPR in human-derived cell lines through Sso7d domain fusion

Next, to increase TnpB-VPR activity within the optimized ωRNA design framework, we generated two expression constructs in which the DNA-binding auxiliary domain Sso7d, DNA-binding protein from *Sulfolobus solfataricus*, was fused to either the N- or C-terminus of TnpB-VPR (Fig. 3A). Previous reports have shown that the Sso7d domain assists in the non-specific recognition of double-stranded DNA, facilitating target sequence binding of CRISPR-based modules [32]. To evaluate this function, we co-transfected human-derived cells with either N-Sso7d-TnpB-VPR or C-Sso7d-TnpB-VPR expression constructs together with a fluorescent reporter plasmid carrying the target DNA sequence, and compared the relative fold increases in reporter expression (Fig. 3B, extended data file1: Raw data table3, extended data file2). Among the two versions, N-Sso7d-TnpB-VPR exhibited the highest fold increase in reporter activation. Similar effects were observed in endogenous targets, such as *HBG*, with N-terminal fusion yielding stronger fold increases in target gene activation than in C-terminal fusion (Fig. 3C, extended data file1: Raw data table4). These differences suggest that the positional context of the Sso7d domain influences DNA-binding activity between transfected plasmid and chromosomal DNA.

To further optimize the N-Sso7d-TnpB-VPR module, we tested three flexible linkers inserted between TnpB and the VPR module (Supplementary Fig. S4A and B; Supplementary Table S8). Among these, a 33-amino acid linker yielded the strongest induction of target gene expression, and this construct (designated V1) was selected as the final optimized version (Supplementary Fig. S4C). When combined with the previously optimized minimal ωRNA (ΔD1238) carrying a fixed 10-nt guide part, this optimized module induced robust activation of diverse endogenous target genes without detectable indel formation (Supplementary Fig. S4D). To verify that the absence of indel formation resulted from loss of DNA cleavage rather than reduced editing efficiency, we purified N-Sso7d-TnpB–ΔD1238 ωRNA RNP complexes and performed an *in vitro* DNA cleavage assay (Supplementary Fig. S5A–C). Consistent with the cellular genome-editing results, the RNP containing the 20-nt guide part efficiently cleaved the target DNA, whereas the RNP containing the 10-nt guide part exhibited no detectable DNA cleavage activity (Supplementary Fig. S5D and E). Together, these results demonstrate that shortening the guide part of the ωRNA to 10-nt abolishes the intrinsic DNA cleavage activity of TnpB while preserving its DNA-targeting capability, thereby supporting its application as a programmable transcriptional regulator.

Using the optimized N-Sso7d-TnpB-VPR construct, we tested multiple human-derived target genes (*HBG*, *IFNG*, *CD2*, *HBB*, and *CXCR4*; Fig. 3D; Supplementary Fig. S6, extended data file1: Raw data table5). Targeting upstream of the TSS resulted in varying degrees of fold increase across loci. For targets showing relatively low induction, tiling assays across the upstream and downstream regions relative to the TSS revealed position-dependent variations in gene activation

Downloaded from academic.oup.com/nar/article/54/16/gkag849/8770891 by Chung Ang University, Anseong Campus Library user on 16 October 2026

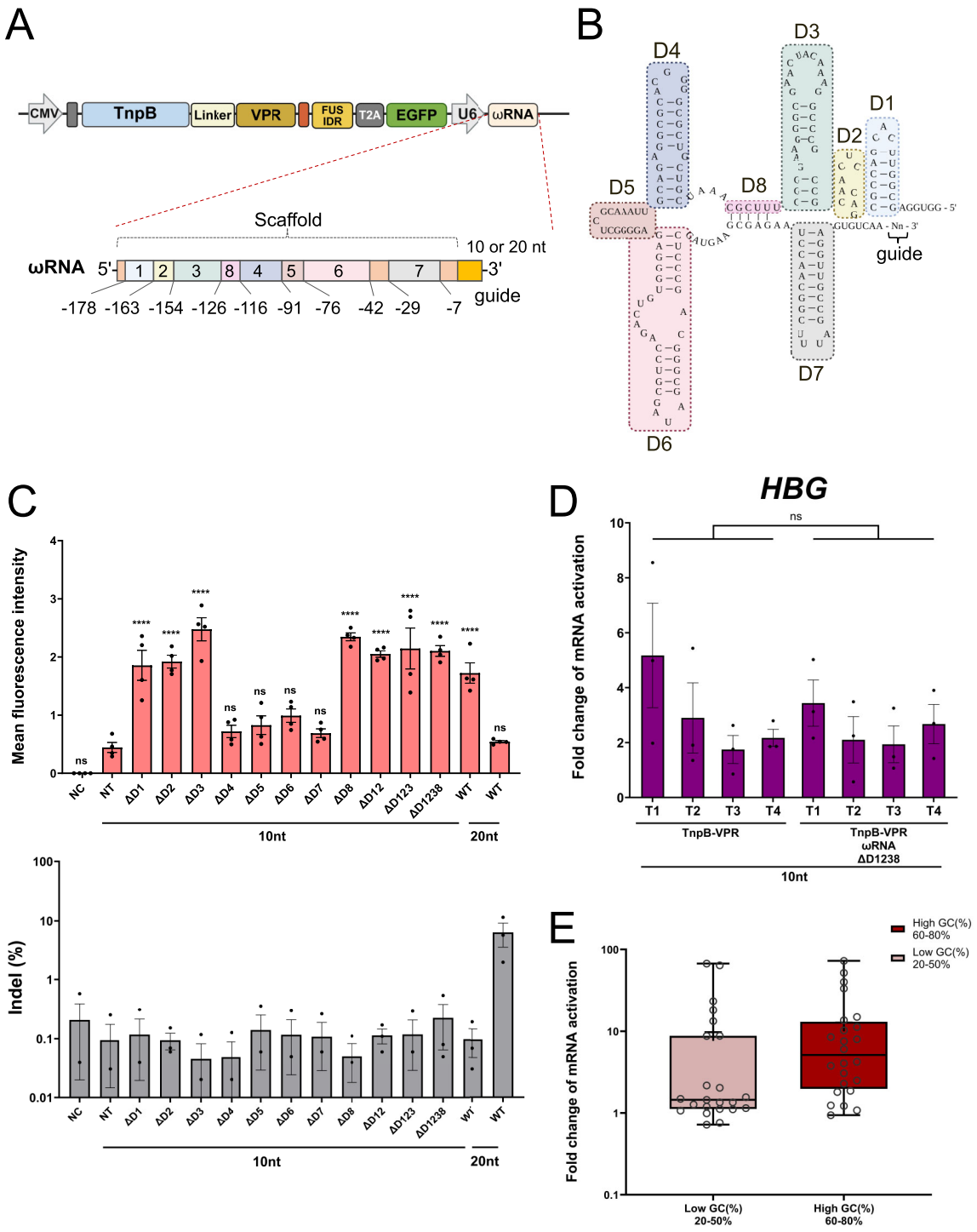


Figure 2. ωRNA engineering and functional assessment to optimize an ultracompact TnpB system. (**A**) Schematic representation of the optimized DNA vector expressing TnpB-VPR under the CMV promoter and ωRNA under the U6 promoter. The secondary structure of U6-driven ωRNA was predicted *in silico* and schematically visualized. (**B**) Predicted secondary structure of ωRNA generated using RNAfold. Distinct hairpin structural domains are indicated by different colors. Created in BioRender. Lee, S.H. (2026) https://BioRender.com/afbd9g9. (**C**) Dual analysis of reporter activation (top) and endogenous gene-editing efficiency (bottom) following cellular delivery of TnpB-VPR with ωRNAs of various structural contexts. NC, negative control; NT, non-target; WT, wild-type full-length ωRNA; ΔD1–Δ1238, deletion mutants of the corresponding ωRNA domains. Each histogram was plotted by applying the standard error of the mean to the repeated experimental values ($n \geq 3$). *P*-values were calculated using one-way ANOVA and Dunnett's test (ns: not significant, $^{*}P = .0332$, $^{**}P = .0021$, $^{***}P = .0002$, $^{****}P < .0001$). (**D**) Comparison of transcriptional activation potency between WT ωRNA and Δ1238 mutant ωRNA targeting the endogenous *HBG* locus in human-derived cells. Each histogram was plotted by applying the standard error of the mean to the repeated experimental values ($n = 3$). *P*-values were calculated using a paired *t*-test, two-tailed (ns: not significant, $^{*}P = .0332$, $^{**}P = .0021$, $^{***}P = .0002$, $^{****}P < .0001$). (**E**) Comparative assessment of the transcriptional activation efficiency of TnpB-VPR according to the GC content of the guide part of the ωRNA.

Downloaded from academic.oup.com/nar/article/54/16/gkag849/8770891 by Chung Ang University, Anseong Campus Library user on 16 October 2026

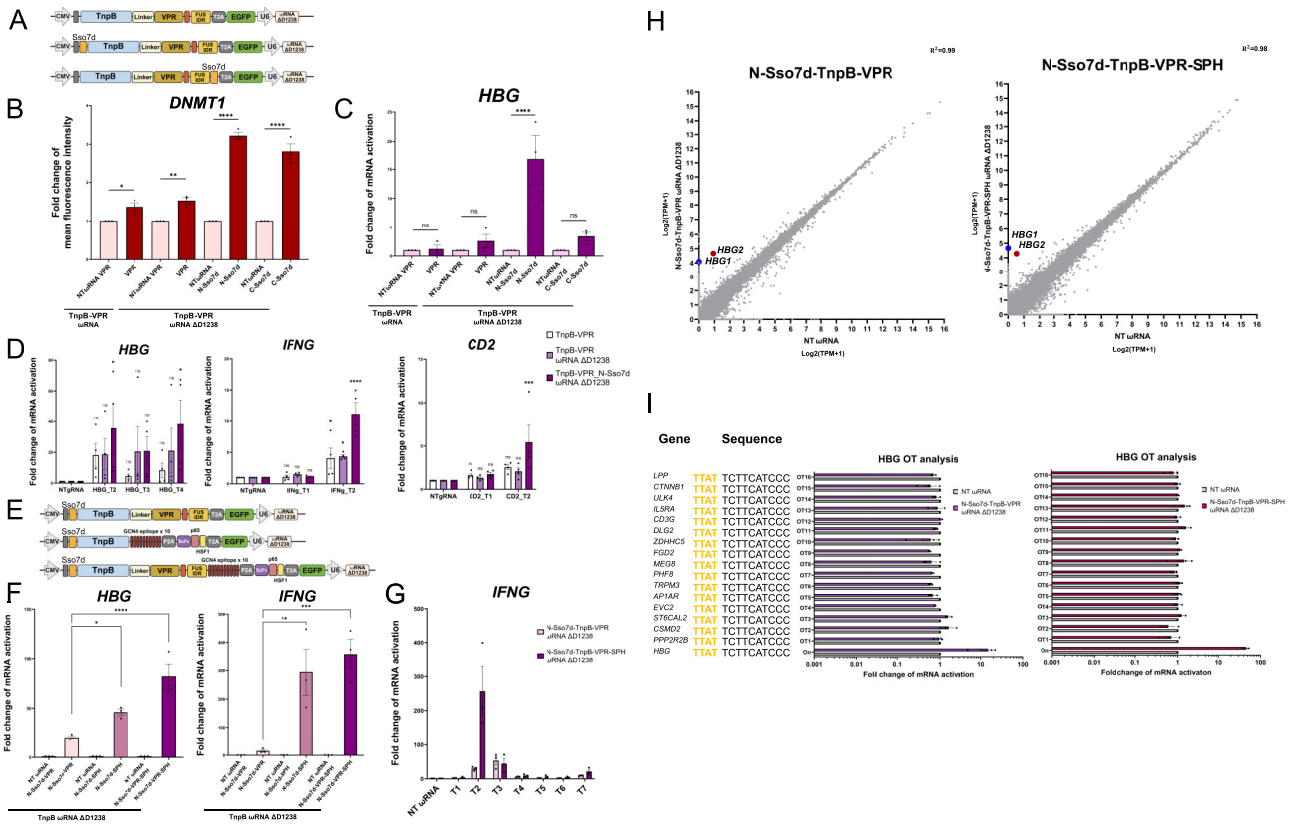


Figure 3. Optimization of the TnpB-VPR system to enhance short guide length of the ωRNA-mediated transcriptional activation. (**A**) Schematic representation of TnpB-VPR variants incorporating the Sso7d domain. From top to bottom: canonical TnpB-VPR activator, N-terminal Sso7d-fused TnpB-VPR, and C-terminal Sso7d-fused TnpB-VPR. Created in BioRender. Lee, S.H. (2026) https://BioRender.com/6ezwz1j. (**B**) Functional comparison of different TnpB-VPR variants using a fluorescent protein reporter plasmid. VPR: canonical TnpB-VPR, VPR-N-Sso7d: N-terminal Sso7d fusion, VPR-C-Sso7d: C-terminal Sso7d fusion. Each histogram was plotted by applying the standard error of the mean to the repeated experimental values ($n = 3$). *P*-values were calculated using a two-way ANOVA and Šídák's test (ns: not significant, $^{*}P = .0332$, $^{**}P = .0021$, $^{***}P = .0002$, $^{****}P < .0001$). (**C**) Performance comparison of TnpB-VPR variants targeting the endogenous HBG locus in HEK293FT cells. Each histogram was plotted by applying the standard error of the mean to the repeated experimental values ($n = 3$). *P*-values were calculated using a two-way ANOVA and Šídák's test (ns: not significant, $^{*}P = .0332$, $^{**}P = .0021$, $^{***}P = .0002$, $^{****}P < .0001$). (**D**) Comparative analysis of transcriptional activation across multiple endogenous loci (*HBG*, *IFNG*, *CD2*) using different TnpB-VPR variants. Each histogram was plotted by applying the standard error of the mean to the repeated experimental values ($n \geq 3$). *P*-values were calculated using a two-way ANOVA and Dunnett's test (ns: not significant, $^{*}P = .0332$, $^{**}P = .0021$, $^{***}P = .0002$, $^{****}P < .0001$). (**E**) Schematic representation of TnpB-VPR variants incorporating the SPH domain. From top to bottom: TnpB-VPR, TnpB-SPH, and TnpB-VPR-SPH. Created in BioRender. Lee, S.H. (2026) https://BioRender.com/ot91qmd. (**F**) Comparative activation efficiencies of TnpB-VPR, TnpB-SPH, and TnpB-VPR-SPH modules targeting *HBG* and *IFNG* loci in HEK293FT cells. Each histogram was plotted by applying the standard error of the mean to the repeated experimental values ($n = 3$). *P*-values were calculated using one-way ANOVA and Dunnett's test (ns: not significant, $^{*}P = .0332$, $^{**}P = .0021$, $^{***}P = .0002$, $^{****}P < .0001$). (**G**) Comparison of transcriptional activation efficiency between TnpB-VPR and TnpB-VPR-SPH at different distances from the TSS of the *IFNG* locus in HEK293FT cells. (**H**) Transcriptome-wide off-target activation analysis of TnpB-VPR and TnpB-VPR-SPH using ωRNAs targeting the *HBG*. RNA-seq data are shown as fold changes relative to the non-target ωRNA control. *HBG1* and *HBG2* are indicated in red and blue, respectively. (**I**) Assessment of gene expression induced by the TnpB-VPR and TnpB-VPR-SPH systems *in silico*-predicted off-target sequence candidates. The nucleotide sequences at non-target loci shown on the left side of the histogram represent off-target candidates, which were computationally predicted by identifying identical sequence matches to the on-target *HBG* locus within the TSS-proximal regions of non-target genes. The left and right histograms depict RT-qPCR-based quantification of on-target and off-target gene activation levels induced by ISDge10 TnpB-VPR and ISDge10 TnpB-VPR-SPH, respectively.

efficiency (Supplementary Fig. S6, extended data file1: Raw data table6). In addition, when the N-Sso7d-TnpB-VPR construct was applied across multiple cell lines (U2OS, HeLa, and K562), distinct levels of target gene induction were observed, reflecting cell type-specific differences in transcriptional regulation (Supplementary Fig. S7, extended data file1: Raw data table7). To further potentiate target-specific activation, we replaced the VPR domain in the N-Sso7d-TnpB-VPR module with an SPH (SunTag-p65-HSF1) domain [33] or combined the VPR and SPH domains into a composite activator (Fig. 3E). When applied to the *HBG* locus, both N-Sso7d-TnpB-SPH and N-Sso7d-TnpB-VPR-SPH modules produced markedly higher fold increases (2.30 ± 0.33 and 3.95 ± 0.29-fold, respectively) in gene expression than the N-Sso7d-TnpB-VPR module (Fig. 3F, extended data file1: Raw data table8). A similar trend was observed at additional genomic loci, including the *IFNG* gene, where the N-Sso7d-TnpB-VPR-SPH module produced the highest fold increase in gene expression (358.42 ± 52.93-fold). However, when transcriptional activation was systematically evaluated across multiple sites within the *IFNG* locus, robust upregulation was restricted to specific target sites, indicating that effective gene activation depends highly on the site (Fig. 3G, extended data file1: Raw data table8).

Downloaded from academic.oup.com/nar/article/54/16/gkag849/8770891 by Chung Ang University, Anseong Campus Library user on 16 October 2026

Following the demonstration of robust target gene activation across multiple loci, we evaluated the specificity of the N-Sso7d-TnpB-VPR and N-Sso7d-TnpB-VPR-SPH modules at the *HBG* locus (Fig. 3H and I; Supplementary Table S5). Whole-transcriptome-wide off-target analysis confirmed significant induction only of the target genes *HBG1* and *HBG2*, whereas no meaningful expression changes were observed for off-target candidate genes with partial sequence similarity near the annotated TSS regions (Fig. 3H). In addition, targeted amplicon sequencing of *in silico*-predicted off-target sites with completely identical guide sequence of the ωRNA contexts revealed no detectable off-target activity above the background signal observed with NT-ωRNA controls (Fig. 3I, extended data file1: Raw data table9). Upon targeting the *IFNG* locus, the version incorporating the VPR-SPH domain exhibited enhanced induction of target gene expression compared with that of the VPR-only conjugated module and was associated with concomitant upregulation of other genes, indicating co-expression effects (Supplementary Fig. S8 and Supplementary Table S6).

We next sought to determine whether the optimized TnpB platform could be adapted for transcriptional repression. The optimized 10-nt guide part of the ωRNA identified for the TnpB activator was retained, whereas the transcriptional activation domain VPR was replaced with the transcriptional repression domains KRAB-MeCP2 and DNMT3L. To further enhance target DNA binding, the Sso7d DNA-binding domain was fused to the N-terminus of TnpB (Supplementary Fig. S9A). To assess transcriptional repression, multiple target sites proximal to the TSS of the endogenous human *HBG* gene were examined. The repression efficiency varied depending on the target position, indicating that the genomic binding site substantially influences TnpB mediated transcriptional repression. N-Sso7d-DNMT3L-TnpB-KRAB-MeCP2 exhibited the strongest repression activity, reducing *HBG* expression to up to 0.58-fold relative to the non-targeting control (Supplementary Fig. S9B). These results demonstrate that the optimized 10-nt ωRNA platform can be readily converted from a transcriptional activator into a transcriptional repressor through replacement of the effector domains while maintaining efficient target DNA binding.

Induction of genome editing in human-derived cell lines using gene-targeting optimized TnpB

To achieve precise and robust sequence editing of target genes in human-derived cells, we optimized ωRNA engineering and target DNA binding using the WT-TnpB module with a previously enhanced transcriptional activation capacity (Fig. 4A). First, we performed comparative analyses of indel induction efficiency by sequential deletion of the ωRNA scaffold domains in the human *DNMT1* locus (Fig. 4B). To induce DNA double-strand breaks in the target gene, we used ωRNAs carrying a full-length 20-nt guide part. As anticipated, the indel induction efficiency exhibited trends that were largely consistent with those of prior TnpB-VPR activation experiments (Fig. 2C). Among the ωRNA domains (D1–D8), regions D1–D3 were relatively dispensable, whereas deletion of D4–D7 markedly reduced the gene-editing efficiency (Fig. 4C). Simultaneous domain removal using the optimized ΔD1238 ωRNA resulted in a 3.47% reduction in indel efficiency compared to that of WT ωRNA, in contrast to the transcriptional activation assays, where ΔD1238 retained activity. Then, we optimized the DNA-binding capacity by fusing the Sso7d domain to either the N- or C-terminus of TnpB, generating N-Sso7d-TnpB and C-Sso7d-TnpB modules (Fig. 4D). When delivered into human HEK293FT cells, the N-terminal fusion construct consistently produced stronger indel induction than in non-conjugated TnpB or the C-terminal fusion (Fig. 4E). N-Sso7d-TnpB exhibited a similar domain dependency across the ωRNA deletion variants, with D1–D3 remaining largely dispensable, whereas deletion of D4–D7 substantially impaired gene-editing activity. Despite the reduced indel efficiency observed with the ΔD1238 ωRNA in the native TnpB system, N-terminal fusion of Sso7d fully compensated for the reduction in gene-editing efficiency and further enhanced editing activity beyond that achieved by WT TnpB programmed with the WT ωRNA (Fig. 4C). These findings suggest that Sso7d enhances productive target DNA engagement, thereby functionally compensating for the loss of ωRNA structural elements required for efficient genome editing. In direct comparison with the previous transcriptional activation assays using TnpB-VPR, the optimized ΔD1238 ωRNAs carrying a 10-nt guide failed to induce detectable indels, whereas extending the guide to 20-nt enabled efficient target DNA cleavage and indel formation across multiple endogenous loci with N-Sso7d-TnpB (Fig. 4F and Supplementary Fig. S10). Notably, the N-terminal fusion construct with ΔD1238 ωRNAs carrying 20-nt guide part significantly improved indel induction (1.57 ± 0.08-fold increase) across multiple endogenous loci in human-derived cells, compared with that of WT-TnpB (Fig. 4G). The N-Sso7d-TnpB module also demonstrated robust editing activity in a mouse cell line (NIH3T3), inducing efficient insertion/deletion (indel) formation (Supplementary Fig. S11). Finally, to establish proof-of-concept for disease modeling applications, we applied the N-Sso7d-TnpB module to engineer a ΔF508 (CTT deletion) mutation in the *CFTR* gene in human-derived cells via homology-directed repair (HDR) [34] (Fig. 4H–L). To identify efficient target sites, we designed a tiling panel of candidate sequences around the F508 locus and compared the indel efficiencies (Fig. 4H and I). Candidate sites exhibited variable levels of indel induction, and T1 and T10 were selected as the most efficient sites (Fig. 4I). Based on these sites, we constructed donor vectors carrying homology arms designed to generate ΔF508 (CTT deletion) and quantified HDR-mediated editing efficiency across multiple conditions using targeted amplicon sequencing (Fig. 4J). Therefore, ΔF508 induction efficiencies of 0.61 ± 0.35 and 1.37 ± 0.31% were observed at T1 and T10, respectively (Fig. 4K and L). To evaluate the genome-editing performance of the optimized TnpB system, we compared its editing efficiency with that of SpCas9 across multiple endogenous target loci. The optimized N-Sso7d-TnpB system programmed with a 20-nt guide part of the ΔD1238 ωRNA was evaluated in parallel with SpCas9. Genome editing efficiencies varied among the target loci examined (Supplementary Fig. S12A). At several endogenous targets, the optimized TnpB system achieved editing efficiencies comparable to those of SpCas9 (Supplementary Fig. S12B). Collectively, these results suggest that the optimized TnpB platform can complement the existing SpCas9 genome-editing system by providing efficient target-dependent genome editing across endogenous loci.

Downloaded from academic.oup.com/nar/article/54/16/gkag849/8770891 by Chung Ang University, Anseong Campus Library user on 16 October 2026

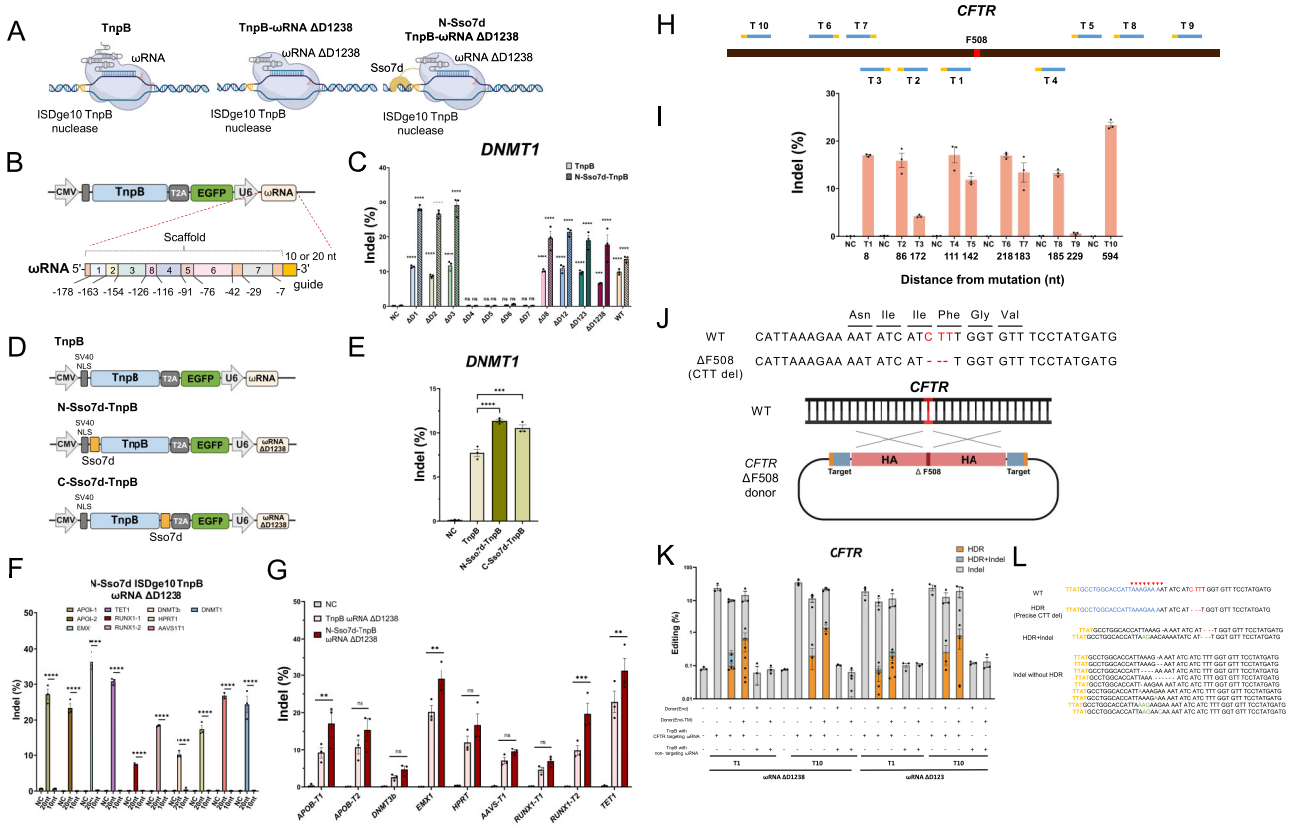


Figure 4. Efficient genome editing using engineered TnpB guided by extended guide length of the ωRNA. (**A**) Schematic representation of TnpB-mediated genome editing using 20-nt guide part of the ωRNAs. (**B**) Schematic representation of CMV-driven TnpB expression constructs with or without an N-terminal Sso7d fusion and the U6-driven ωRNA expression construct. The ωRNA hairpin domains predicted by RNAfold are highlighted in different colors. (**C**) Comparative analysis of genome-editing efficiencies using ωRNA deletion mutants lacking specific structural domains. Editing of endogenous target loci was quantified in HEK293FT cells. TnpB: WT TnpB, N-Sso7d-TnpB: N-terminus Sso7d conjugated TnpB. (**D**) Schematic representation of TnpB variants engineered by fusion with the Sso7d domain. From top to bottom: WT TnpB, N-terminal Sso7d-TnpB, and C-terminal Sso7d-TnpB. All constructs were co-expressed with the optimized ΔD1238 mutant ωRNAs. (**E**) Comparison of indel frequencies at the endogenous DNMT1 locus among WT TnpB, N-terminal Sso7d-TnpB, and C-terminal Sso7d-TnpB. (**F**) Analysis of editing efficiency using ΔD1238 ωRNA containing either a 10- or 20-nt guide length of the ωRNA at diverse endogenous genes with an N-terminal Sso7d fusion. (**G**) Comparison of indel induction efficiencies between WT-TnpB and N-Sso7d-TnpB across multiple endogenous loci using ΔD1238 mutant ωRNAs. (**H**) Schematic of the TnpB-targeting strategy used to generate a ΔF508 cystic fibrosis (CF) disease-model cell line with high precision and efficiency. Multiple *CFTR* target sites were selected for N-Sso7d-TnpB-mediated mutagenesis. The ΔF508 mutation is highlighted in red, TAM sequence in yellow, and target sites in blue. (**I**) Screening of ωRNAs to identify high-efficiency guides for N-Sso7d-TnpB-mediated indel induction at the *CFTR*. (**J**) Schematic of HDR donor design and strategy for ΔF508 knock-in. The pathogenic ΔF508 mutation is indicated in red. Homology arms (pink) flank the editing site, and target cleavage sites are embedded near the arm termini to generate linearized donor DNA in cells and enhance HDR efficiency. (**K**) Comparison of HDR efficiencies for ΔF508 *CFTR* editing depending on whether cleavage sites were inserted at the ends of donor homology arms. (**L**) Representative sequencing patterns from HDR-mediated ΔF508 editing using N-Sso7d-TnpB, showing correct HDR, indel-only, and mixed HDR + indel alleles. The target sequence is highlighted in blue, and TAM sequence is shown in yellow. Deletions are indicated by dashes, insertions by red, and substitutions by green. Red triangles indicate the cleavage sites. All data are presented as mean ± standard error from three independent experiments ($n = 3$). All *P*-values shown in the histograms were calculated using two-way ANOVA followed by Dunnett's multiple-comparison test (ns: not significant, $^{*}P = .0332$, $^{**}P = .0021$, $^{***}P = .0002$, $^{****}P < .0001$). The figures were created in BioRender. Lee, S.H. (2026) https://BioRender.com/o0fa9kv.

ωRNA-guided TnpB-derived adenine base editor (dTnpB-ABE) enables programmable adenine base conversion

To further assess the feasibility of a bifunctional module based on switching the guide length of the ωRNA, we fused the TadA deaminase domain to the N-Sso7d-TnpB module. We then evaluated whether this system could induce the intended single-nucleotide substitutions at the target genomic loci (Fig. 5). Single-base editing was achieved using an inactivated ISDge10 dTnpB-ABE targeting endogenous loci within the *HEK1* gene in human cells. In this study, two base editor architectures were designed by fusing the TadA deaminase [35] to either the N-terminus or C-terminus of catalytically inactivated ISDge10 dTnpB (D187A, hereafter referred to as dTnpB-ABE; N-term) and dTnpB-ABE (C-term) with 20-nt guide length of ωRNA, respectively, and their base-editing activities were systematically compared (Fig. 5A). Across all four target sites within the *HEK1* locus, A-to-G base conversion was detected exclusively under conditions in which dTnpB-ABE was co-delivered with the targeting ωRNA, whereas no significant editing signals were observed in the NC group. At the *HEK1*-target1 site, base editing was observed across a relatively broad region spanning positions a2–a18, with the highest A-to-G conversion efficiencies detected at positions a3–a6 (Fig. 5B). The N-terminal TadA fusion (dTnpB-ABE; N-term) exhibited higher editing efficiencies than that of the C-terminal fusion (dTnpB-ABE, C-term) at most adenine positions, with a gradual decline in editing efficiency observed at more distal sites. Similarly, at *HEK1*-Target2, robust editing activity was detected within the a4–a10 region,

Downloaded from academic.oup.com/nar/article/54/16/gkag849/8770891 by Chung Ang University, Anseong Campus Library user on 16 October 2026

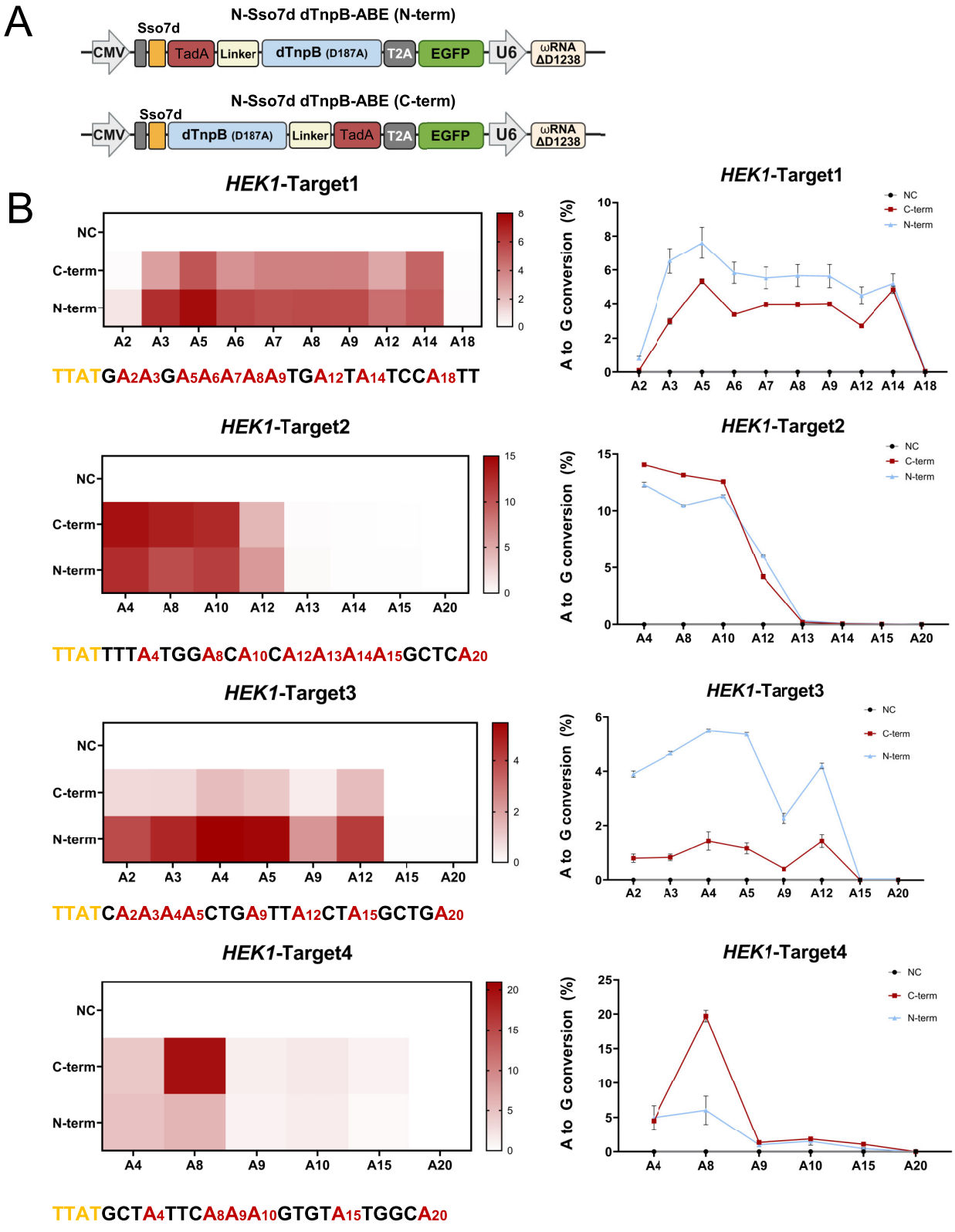


Figure 5. Efficient adenine base editing mediated by dTnpB-ABE. (**A**) Schematic representation of the TnpB-based adenine base editor. Adenine deaminase was fused to either the N- or C-terminus of catalytically inactive dTnpB, generating dTnpB-ABE (N-terminus) and dTnpB-ABE (C-terminus). Each construct was expressed together with the optimized ΔD1238 ωRNA. Created in BioRender. Lee, S.H. (2026) https://BioRender.com/w9y0yn8. (**B**) Editing efficiencies of dTnpB-ABE (N-terminus) and dTnpB-ABE (C-terminus) at multiple *HEK1* target sites containing clustered adenine. Left: Heat maps displaying normalized A-to-G conversion frequencies across target windows (0–1 scale). Right: Line graphs showing the measured A-to-G substitution efficiencies for each construct. All data points were plotted by applying the standard error of the mean to repeated experimental values ($n \geq 3$). NC: negative control, C-term: N-Sso7d dTnpB-ABE (C-term), N-term: N-Sso7d dTnpB-ABE (N-term).

Downloaded from academic.oup.com/nar/article/54/16/gkag849/8770891 by Chung Ang University, Anseong Campus Library user on 16 October 2026

whereas editing efficiencies decreased sharply beyond position a12, indicating a relatively constrained base-editing window. At *HEK1*-Target3, the dTnpB-ABE (N-term) module demonstrated pronounced base-editing activity across multiple adenine residues, whereas the dTnpB-ABE (C-term) module induced only low-level or sporadic editing. In contrast, *HEK1*-Target4 displayed a localized editing pattern with a prominent editing peak at position a8, which was more pronounced in the dTnpB-ABE (C-term) configuration. Thus, both the efficiency and spatial distribution of dTnpB-ABE-mediated base editing are influenced by dTnpB-ABE architecture and by the local sequence context of the target DNA. N-terminal fusion of the TadA domain conferred overall superior editing performance across multiple targets, and the effective base-editing window for both N- and C-terminal configurations was consistently observed ∼3–12 bp downstream of the TAM (Fig. 5B). Collectively, the dTnpB-ABE system can reliably induce position-specific adenine base editing at endogenous loci.

Stringent mismatch sensitivity confers high target specificity to TnpB

Subsequently, we evaluated the target specificity of the ISDge10 TnpB module in direct comparison with SpCas9 using the *DNMT1* locus as a shared genomic target (Fig. 6). For this comparison, genome-editing efficiencies were quantified as normalized indel ratios across a series of gRNA-target DNA pairs harboring defined mismatches, indicated by red lowercase nucleotides within the target sequence (Fig. 6A). When perfectly matched guides were used, both ISDge10 TnpB and SpCas9 induced robust indel formation at the *DNMT1* site. However, the introduction of single mismatches within the guide-target duplex sequentially revealed marked differences in mismatch tolerance between the two systems. While SpCas9 retained substantial editing activity across several mismatch positions, ISDge10 TnpB exhibited a pronounced and position-dependent reduction in indel efficiency, with editing activity sharply diminished upon the introduction of mismatches at multiple sites. These results indicate that TnpB has a stricter requirement for guide-target complementarity than did Cas9.

To evaluate off-target propensity, *in silico* off-target prediction analyses were performed for each guide sequence (*DNMT1*, *EMX1*, *RUNX1-2*, and *TET1*; Fig. 6B, Supplementary Table S7). For ISDge10 TnpB, predicted off-target sites containing mismatches between the ωRNA and gDNA hardly exhibited indel frequencies. Compared with untreated NCs, WT-TnpB and N-Sso7d-TnpB displayed no significant off-target activity, except for a weak indel signal observed at OT8 of *RUNX1-2*. Collectively, ISDge10 TnpB-mediated genome editing is characterized by heightened sensitivity to guide-target mismatches and a minor, *in silico* off-target profile. To further assess potential genome-wide off-target effects of the TnpB system in an unbiased manner, we performed WGS of *EMX1*-targeted, TnpB-treated samples and corresponding NCs. After excluding variants detected in the control, the remaining TnpB treatment-specific variants were analyzed for similarity to the target sequence. Comprehensive analyses of candidate loci containing 0–20 mismatches relative to the *EMX1* target sequence did not reveal evidence of severe off-target editing events (Fig. 6C, extended data file3). Inspection of sequencing read alignments revealed a general pattern consistent with on-target *EMX1* editing, without consistent evidence supporting bona fide off-target editing at the candidate loci (Supplementary Fig. S13). Collectively, these findings indicate that only a limited number of candidate sequence variants were detected under the experimental conditions tested.

Dual-mode application of TnpB in human-derived cells via AAV delivery

Finally, we assessed whether the optimized TnpB-VPR and TnpB modules could be effectively deployed in human-derived cells through AAV-mediated delivery to enable transcriptional regulation and genome editing (Fig. 7). Compared with other known CRISPR modules, TnpB possesses a relatively small gene size, making it well-suited to the established packaging capacity of AAV vectors (Supplementary Fig. S14A). When large transcriptional activator domains, such as VPR, are fused to CRISPR modules, the resulting constructs often exceed the AAV size limit, rendering recombinant virus packaging and purification difficult (Supplementary Fig. S14B). In the case of CRISPR-Cas9, even when the nuclease is used in a split configuration, fusion with the VPR transcriptional activation domain results in a construct that exceeds the packaging capacity (∼4.7 kb) of AAV. Consequently, AAV-based delivery for gene expression regulation is technically impracticable. In contrast, the optimized TnpB-VPR module had a total length of <4.7 kb, falling within the AAV packaging threshold (Fig. 7A and B). Accordingly, we designed a dual-delivery system in which a recombinant AAV (rAAV) expressing the optimized TnpB-VPR module was co-delivered with a separate rAAV expressing target-specific ωRNA (Fig. 7A and Supplementary Fig. S14C). When applied to human-derived HEK293FT cells, delivery of the N-Sso7d-TnpB-VPR module with target ωRNA rAAVs induced robust activation of the *HBG* gene, yielding an 8.7-fold increase relative to non-target ωRNA controls (Fig. 7B).

Subsequently, we attempted to induce indel formation at the target locus by modifying the guide length of the optimized ωRNA to 20-nt (Fig. 7C and D). We compared the indel-inducing efficiencies of rAAV-delivered SpCas9 and ISDge10 TnpB systems. Owing to gene size constraints, SpCas9 was packaged into a split-configured rAAV vector that enabled the separate expression of its N-terminal and C-terminal domains (Supplementary Fig. S14D). In contrast, VPR-deleted TnpB is sufficiently compact to be fully packaged with ωRNA within a single rAAV vector (Fig. 7C and Supplementary Fig. S14D). Thus, efficient delivery into human-derived cells was achieved using a single rAAV system, leading to robust induction of indels at the target gene (average 19.61 ± 1.01% at Day 4 and 42.16 ± 3.96% at Day 14; Fig. 7D). Editing efficiency further increased over time following rAAV transduction. Compared with the split-SpCas9 rAAV system, the TnpB AIO rAAV system exhibited markedly higher editing efficiencies, showing 75.5-fold and 12.7-fold increases at Day 4 and Day 14, respectively.

To compare the performance of our compact TnpB platform with existing compact gene activation and genome-editing systems under AAV delivery, we first evaluated transcriptional activation using the previously reported Cas12f-based activator, CasMini-VPR [36]. Because both systems are compact and compatible with AAV packaging, each construct was packaged under identical conditions and delivered using the same AAV configuration (Fig. 7E). Under these same conditions, the N-Sso7d-TnpB-VPR system exhibited significantly higher transcriptional activation than CasMini-VPR, result-

Downloaded from academic.oup.com/nar/article/54/16/gkag849/8770891 by Chung Ang University, Anseong Campus Library user on 16 October 2026

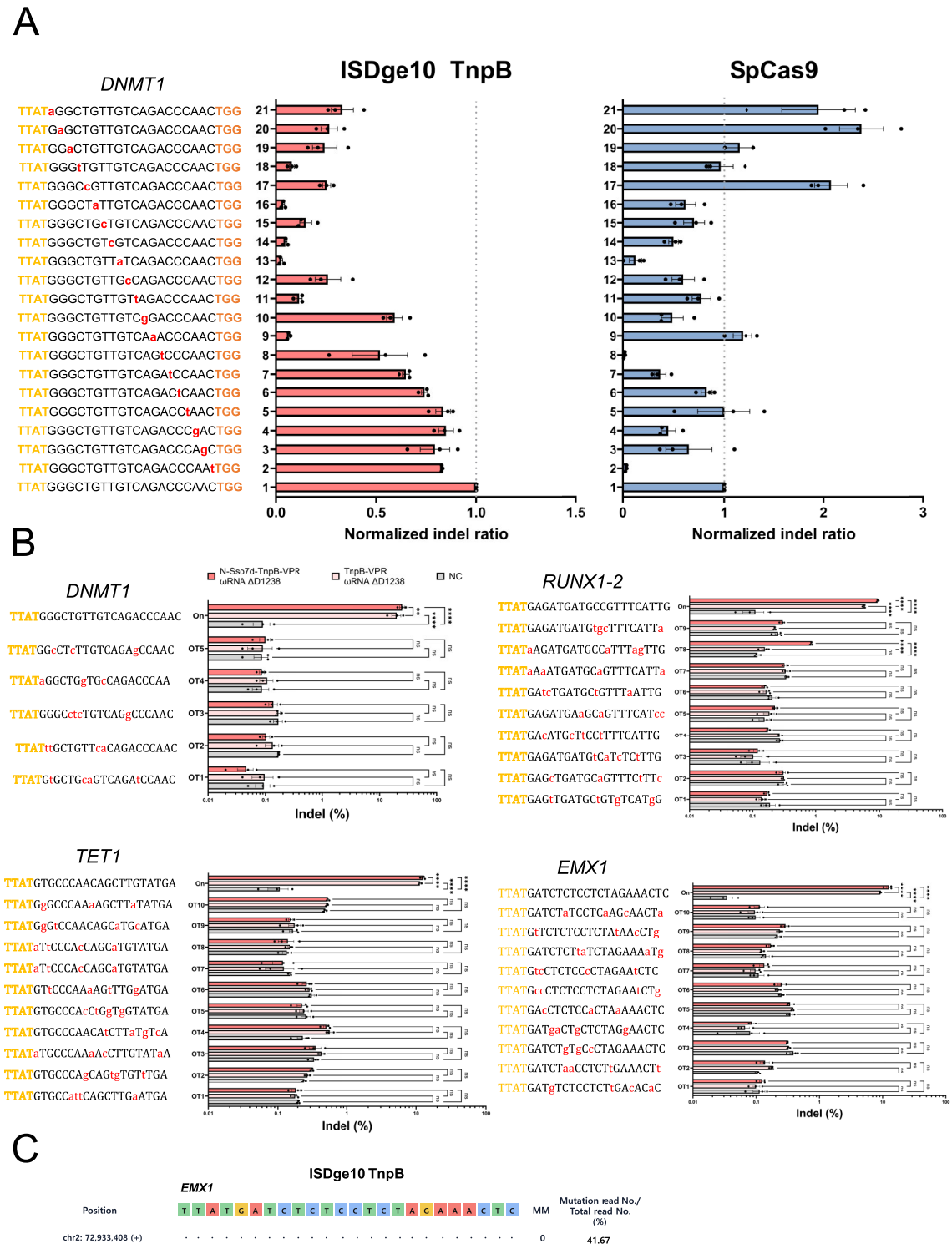


Figure 6. Evaluation of DNA-targeting specificity and off-target activity of TnpB. (**A**) Comparison of mismatch tolerance between ISDge10 TnpB and SpCas9 at identical target sites. TAM (or PAM) sequences are indicated in yellow and orange, respectively. A series of ωRNA (or gRNA) variants carrying a single predefined mismatch at different guide positions relative to the target sequence was generated. The guide RNA-DNA mismatches are highlighted in red. The indel frequencies (%) induced by mismatched guides were normalized to those obtained with perfect-match guides. The data in each histogram are presented as the mean ± SEM of repeated experimental values ($n = 3$). (**B**) Off-target candidates containing up to four mismatches relative to the 20-nt guide length of the ωRNA were predicted *in silico* based on sequence similarity. The mismatched nucleotides are highlighted in red. Each histogram represents the measured indel frequencies (%) at each predicted off-target site. Each histogram was plotted by applying the standard error of the mean to the repeated experimental values ($n = 3$). *P*-values were calculated using a two-way ANOVA and Tukey's test (ns: not significant, $^{*}P = .0332$, $^{**}P = .0021$, $^{***}P = .0002$, $^{****}P < .0001$). (**C**) Unbiased genome-wide off-target analysis of ISDge10 TnpB targeting the *EMX1* locus by WGS. Colored A, C, T, and G indicate mismatched nucleotides relative to the target DNA sequence, whereas dots (·) denote matched nucleotides. The left panel shows the chromosomal coordinates of the on-target and candidate off-target sites in the human genome, whereas the right panel shows the proportion of variant reads among the total sequencing reads detected at each locus. MM, number of mismatches between the guide RNA and the target DNA sequence.

Downloaded from academic.oup.com/nar/article/54/16/gkag849/8770891 by Chung Ang University, Anseong Campus Library user on 16 October 2026

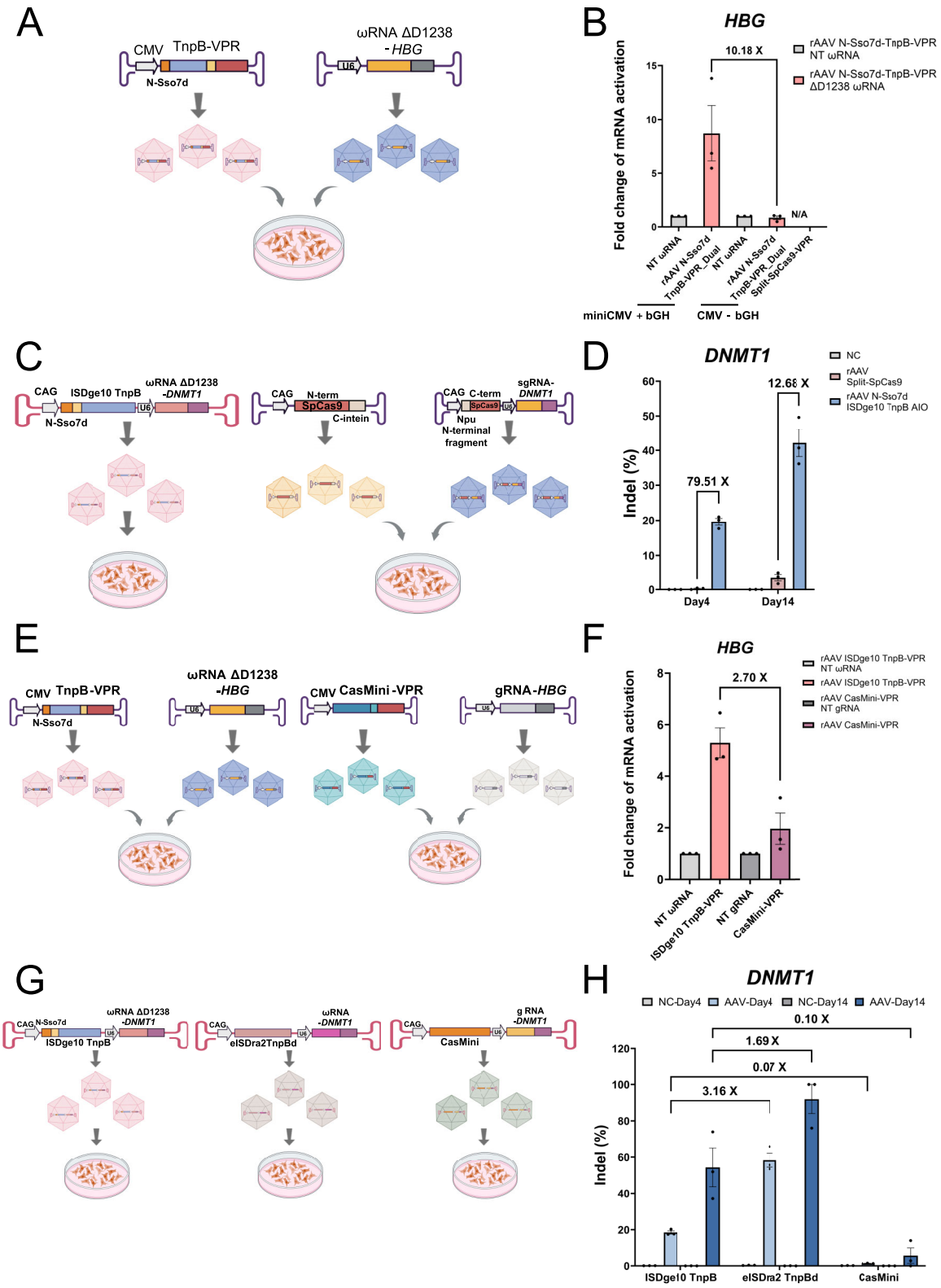


Figure 7. rAAV-mediated delivery of ultracompact TnpB/TnpB-VPR enables transcriptional activation and genome editing in human cells. (**A**) Schematic illustration of the dual-rAAV transcriptional activation system. The CMV promoter-driven N-Sso7d-TnpB-VPR expression cassette and the U6 promoter-driven ΔD1238 ωRNA expression cassette targeting the *HBG* locus were packaged into separate rAAV vectors and co-delivered into HEK293FT cells. (**B**) Comparison of transcriptional activation efficiencies following rAAV-mediated delivery of the dual-rAAV TnpB-VPR system shown in panel (A). *HBG* mRNA expression was quantified by RT-qPCR after co-transduction with the indicated rAAV vectors. The dual-rAAV split-SpCas9 system exceeded the packaging capacity required for VPR-mediated transcriptional activation and therefore was not applicable (N/A). Data are presented as mean ± standard error from three independent experiments ($n = 3$). rAAV N-Sso7d-TnpB-VPR NT ωRNA: rAAV encoding N-Sso7d-TnpB-VPR with a non-targeting ωRNA; rAAV N-Sso7d-TnpB-VPR ΔD1238 ωRNA: rAAV encoding N-Sso7d-TnpB-VPR with a targeting ωRNA; rAAV N-Sso7d-TnpB-VPR_Dual: Dual-rAAV system in which N-Sso7d-TnpB-VPR and ωRNA are delivered in separate rAAV vectors; miniCMV + bGH: miniCMV

Downloaded from academic.oup.com/nar/article/54/16/gkag849/8770891 by Chung Ang University, Anseong Campus Library user on 16 October 2026

ing in a 2.70-fold increase (Fig. 7F). These results demonstrate that the optimized N-Sso7d-TnpB-VPR system can be efficiently packaged into AAV and mediates robust transcriptional activation in human cells. We next compared genome-editing activity with the recently reported engineered ISDra2 TnpB (eISDra2 TnpBd) [37] and CasMini [36] using an AIO AAV system (Fig. 7G). The optimized N-Sso7d-ISDge10 TnpB system exhibited substantially higher editing efficiencies than CasMini at both Day 4 and Day 14 and achieved editing efficiencies comparable to those of eISDra2 TnpBd (Fig. 7H). Collectively, these results demonstrate that the optimized ISDge10 TnpB platform provides an efficient AAV-compatible system for both transcriptional activation and genome editing.

In vivo validation of a compact TnpB platform for transcriptional activation and genome editing

To evaluate the *in vivo* applicability of the TnpB system for gene therapy, we first assessed its transcriptional activation and genome-editing activities in mice using plasmid delivery via HTVI (Fig. 8A–D). Plasmids encoding the AIO N-Sso7d-TnpB-VPR-SPH and 10-nt ΔD1238 ωRNA targeting the mini-CMV promoter, together with the corresponding reporter plasmid used in Fig. 1D, were co-delivered into mice by HTVI. Three days after HTVI, dTomato expression was analyzed by flow cytometry (Fig. 8A). Mice treated with the targeting ωRNA exhibited a 2.13-fold increase in dTomato expression compared with those receiving the non-targeting (NT) ωRNA, demonstrating that the engineered TnpB activator efficiently induces target gene activation *in vivo* (Fig. 8B). We next investigated whether the TnpB system could also mediate genome editing *in vivo*. Plasmids encoding N-Sso7d-TnpB and the corresponding 20-nt ωRNA ΔD1238 were delivered to mouse by HTVI, and genome-editing efficiency was assessed 1 week after injection (Fig. 8C). Targeted indel formation was readily detected, showing a 20.15-fold increase compared with untreated control mice, indicating that the engineered TnpB nuclease efficiently mediates genome editing *in vivo* (Fig. 8D). Finally, to evaluate the therapeutic potential of the compact TnpB system, we packaged the AIO TnpB with 20-nt ωRNA ΔD1238 construct into recombinant adeno-associated virus (rAAV) and targeted the endogenous *Pcsk9* locus in mice. Following injection of purified rAAV, targeted indels at the endogenous *Pcsk9* locus were detected at both 2 and 6 weeks post-injection (Fig. 8E), with editing efficiencies of average 1.73 ± 0.54% and 12.44 ± 0.79%, respectively. Editing efficiency further increased between 2 and 6 weeks, demonstrating sustained genome editing following rAAV delivery (Fig. 8F). These findings establish that the compact TnpB system is compatible with single-vector rAAV delivery and efficiently mediates genome editing at an endogenous therapeutic locus *in vivo*, underscoring its potential for AAV-based gene therapy [9, 22, 38].

Discussion

In this study, we established a platform technology that enables the simultaneous and efficient induction of target gene activation and editing by reprogramming ωRNAs within the novel transposon-derived ISDge10 TnpB system, demonstrating its applicability to biological systems. The structural comparison between ISDge10 and ISDra2 TnpB provides several insights into the conservation and evolution of compact TnpB systems (Fig. 1B and Supplementary Fig. S1). Although the TnpB protein structure was highly conserved, the ωRNA showed relatively low structural similarity. This discrepancy likely reflects the incomplete experimentally determined ωRNA structure, in which only the resolved regions are represented, together with the inherent difficulty of accurately predicting RNA tertiary structures compared with protein structures because of their intrinsic conformational flexibility and dynamic nature. Therefore, the observed differences in ωRNA structure may partly reflect methodological limitations rather than solely biological differences between the two systems. In contrast, the high structural similarity among TnpB proteins suggests that an AlphaFold3-based *in silico* approach can accurately predict their overall architecture, thereby supporting the use of AlphaFold3-generated models for the comparative structural analyses performed in this study. Consistent with prior findings, the TnpB system, which harbors only the core motifs of Cas12 [17],

promoter with the bovine growth hormone (bGH) polyadenylation signal; CMV – bGH: CMV promoter without the bGH polyadenylation signal. (**C**) Schematic comparison of rAAV delivery strategies for genome editing.Left, AIO rAAV system in which the N-Sso7d-ISDge10 TnpB expression cassette and ΔD1238 ωRNA are packaged within a single rAAV vector. Right, conventional dual-rAAV split-SpCas9 system in which the N-terminal and C-terminal fragments of SpCas9 are separately packaged and reconstituted by Npu intein-mediated protein trans-splicing, while the sgRNA is co-expressed with the C-terminal fragment. (**D**) Comparison of genome-editing efficiencies between the AIO TnpB–ωRNA rAAV system and the dual split-SpCas9 rAAV system. Targeted indel frequencies at the *DNMT1* locus were quantified in HEK293FT cells at 4 and 14 days following delivery of the AIO TnpB–ωRNA rAAV system and the dual SpCas9 split rAAV system. Each histogram was plotted by applying the standard error of the mean to the repeated experimental values ($n = 3$). NC: negative control; rAAV Split-SpCas9: Npu intein-mediated split-SpCas9 dual-rAAV delivery system; rAAV N-Sso7d ISDge10 TnpB AIO: AIO rAAV system expressing both TnpB and ωRNA. (**E**) Schematic comparison of compact CRISPR-based transcriptional activation systems compatible with AIO rAAV delivery. TheN-Sso7d-TnpB-VPR/ΔD1238 ωRNA activation system was compared with the CasMini-VPR/sgRNA (Cas12f-based) activation system targeting the *HBG* locus. (**F**) Comparison of transcriptional activation efficiencies between the TnpB-VPR and CasMini-VPR systems following rAAV-mediated TnpB-VPR and CasMini-VPR delivery. Relative *HBG* mRNA expression was measured by RT-qPCR after transduction of HEK293FT cells with the indicated AIO rAAV vectors. Data are presented as mean ± standard error from three independent experiments ($n = 3$). rAAV ISDge10 TnpB-VPR NT ωRNA: rAAV encoding N-Sso7d-ISDge10 TnpB-VPR with a non-targeting ωRNA; rAAV ISDge10 TnpB-VPR: rAAV encoding N-Sso7d-ISDge10 TnpB-VPR with a targeting ωRNA; rAAV CasMini-VPR NT gRNA: rAAV encoding CasMini-VPR with a non-targeting gRNA; rAAV CasMini-VPR: rAAV encoding CasMini-VPR with a targeting gRNA. (**G**) Schematic comparison of compact genome-editing systems compatible with AIO rAAV delivery, including N-Sso7d-ISDge10 TnpB, the engineered eISDra2 TnpBd, and CasMini, each packaged together with its corresponding guide RNA expression cassette within a single rAAV vector. (**H**) Comparison of genome-editing efficiencies of compact editors delivered by AIO rAAV vectors. Indel frequencies at the *DNMT1* locus were measured in HEK293FT cells at 4 and 14 days after rAAV transduction. Genome-editing efficiencies of N-Sso7d-ISDge10 TnpB, eISDra2 TnpBd, and CasMini were compared under identical delivery conditions. Data are presented as mean ± standard error from three independent experiments ($n = 3$). All *P*-values shown in the histograms were calculated using two-way ANOVA followed by Dunnett's multiple-comparison test (ns: not significant, $^{*}P = .0332$, $^{**}P = .0021$, $^{***}P = .0002$, $^{****}P < .0001$). NC: negative control; AAV: rAAV-treated group; Day4 and Day14 indicate the time points after rAAV transduction. The figures were created in BioRender. Lee, S.H. (2026) https://BioRender.com/4dv5jjk.

Downloaded from academic.oup.com/nar/article/54/16/gkag849/8770891 by Chung Ang University, Anseong Campus Library user on 16 October 2026

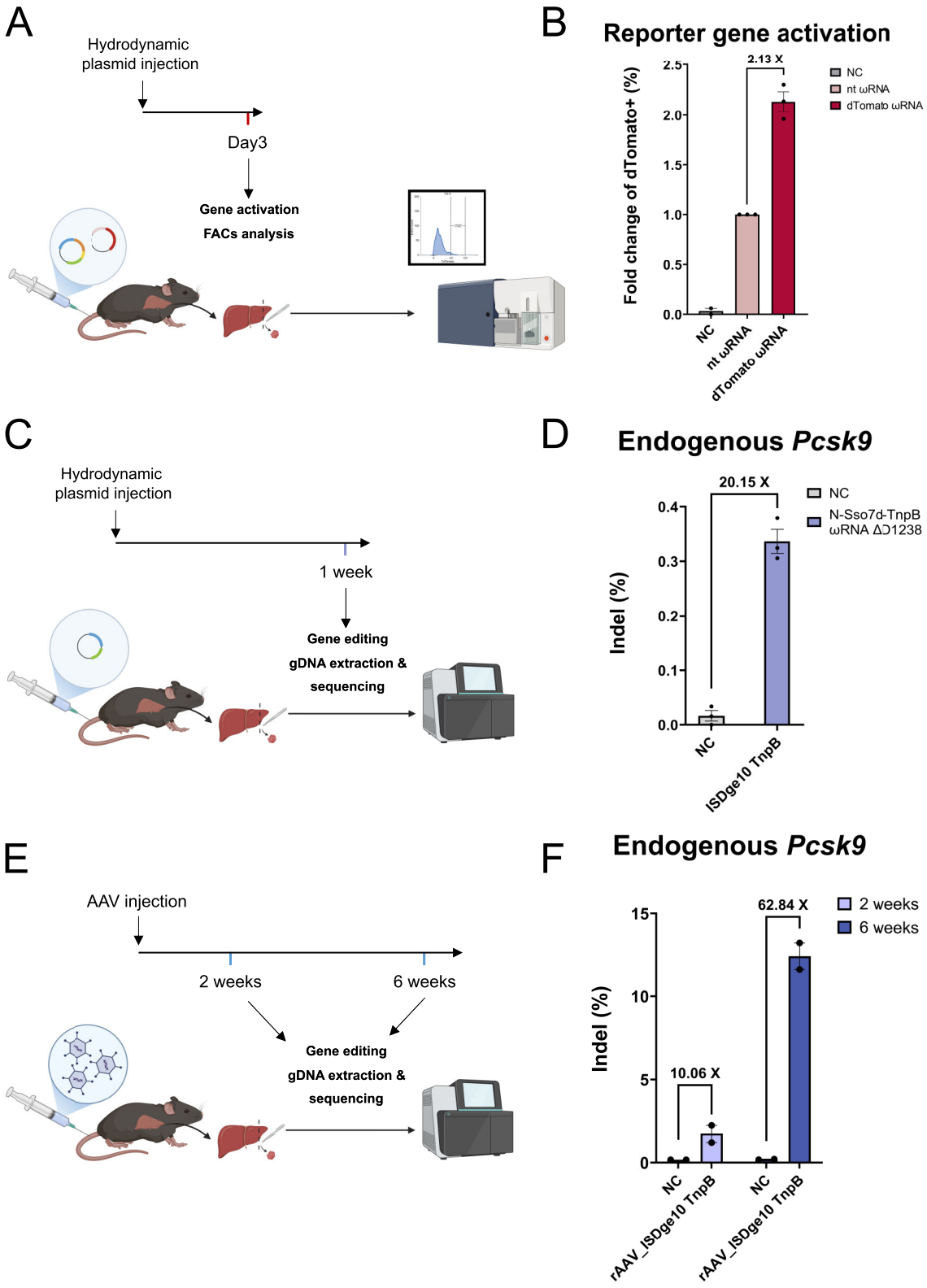


Figure 8. *In vivo* applications of the TnpB system for transcriptional activation and genome editing. (**A**) Schematic illustration of the hydrodynamic plasmid delivery strategy and experimental timeline for *in vivo* reporter gene activation. Plasmids encoding the TnpB-VPR-SPH system were delivered to mouse liver by hydrodynamic tail-vein injection. Reporter gene activation was analyzed by FACS 3 days after plasmid delivery. (**B**) Reporter gene activation following hydrodynamic plasmid delivery of the TnpB-VPR-SPH system. Histograms show the fold change in reporter fluorescence intensity relative to the untreated control measured by FACS analysis 3 days after plasmid delivery. Data are presented as mean ± SEM from three independent experiments ($n = 3$). NC: negative control, nt ωRNA: non-targeting ωRNA, dTomato ωRNA: dTomato reporter gene targeting ωRNA. (**C**) Schematic illustration of the hydrodynamic plasmid delivery strategy and experimental timeline for endogenous *Pcsk9* genome editing. Plasmids encoding the TnpB genome-editing system were delivered by hydrodynamic tail-vein injection, and liver tissues were collected 1 week after injection for gDNA extraction and targeted sequencing. (**D**) Genome-editing efficiency at the endogenous *Pcsk9* locus following hydrodynamic plasmid delivery of the TnpB

Downloaded from academic.oup.com/nar/article/54/16/gkag849/8770891 by Chung Ang University, Anseong Campus Library user on 16 October 2026

can be reprogrammed through ωRNA design and coupled to various functional domains to achieve robust genetic modification. Gene expression activation was induced using short guide length ωRNA (10-nt) that enables reprogrammed DNA targeting without inducing double-strand cleavage, reflecting the bifurcated nature of target recognition and cleavage observed in CRISPR systems [17, 39]. Based on the mechanism previously established for ISDra2 TnpB [17], we reasoned that ISDge10 TnpB operates in a similar manner. When the guide region of the ωRNA forms a sufficiently long and stable RNA–DNA heteroduplex with the target DNA, TnpB is likely able to recognize the heteroduplex and undergo the conformational rearrangements required to position its catalytic domain for DNA cleavage. In contrast, shortening the ωRNA guide region is expected to generate a heteroduplex that is too short to support these conformational changes, thereby preventing the catalytic domain from adopting a cleavage-competent conformation and consequently abolishing DNA cleavage.

These considerations suggest that preserving the native TnpB architecture while regulating cleavage through ωRNA design may be preferable to introducing catalytic mutations into the protein. Consistent with this interpretation, TnpB-VPR consistently exhibited higher transcriptional activation activity than dTnpB-VPR, particularly when programmed with 10-nt wRNAs. Although the precise molecular basis for this difference remains to be elucidated, we speculate that the catalytic residue substitutions introduced into dTnpB may alter the conformational dynamics of the TnpB–ωRNA complex, thereby reducing its ability to adopt a stable DNA-targeting state required for efficient transcriptional activation. In contrast, guide-length modulation preserves the WT protein architecture while selectively suppressing DNA cleavage activity, allowing efficient target recognition without catalytic DNA cleavage. This interpretation is consistent with recent structural studies describing distinct non-cleavage and cleavage-proficient states of TnpB. While the precise mechanism will require further structural and biochemical investigation, our findings suggest that regulating ωRNA guide length while retaining WT TnpB represents a more effective strategy for transcriptional activation than introducing catalytic mutations into the protein.

From an evolutionary perspective, OMEGA-TnpB and CRISPR-Cas systems likely share a guide RNA-dependent DNA recognition and cleavage mechanism, wherein full guide length ωRNA annealing to the target DNA induces a catalytically competent conformational state. Direct side-by-side analysis of TnpB and Cas12f demonstrated highly concordant guide length-dependent profiles of target engagement and cleavage. Both DNA non-cleavage and cleavage property of TnpB were highly sensitive to structural variations in the ωRNA scaffold, particularly within domains predicted *in silico* to form critical stem-loop elements, enabling the identification of minimal ωRNA variants that retained their full activity. These results showed that both the DNA cleavage and non-cleavage states were strongly affected by deletion of the ωRNA scaffold domains (ΔD4–D7), highlighting the importance of protein–RNA interactions for target recognition. This observation is consistent with the predicted structure of the ISDge10 TnpB–ωRNA complex. Furthermore, the pre-organized ωRNA architecture may facilitate target DNA recognition and cleavage in a manner analogous to that proposed for Cas9 [40]. To overcome the relatively weaker gene expression regulation compared to that of indel induction, we introduced the Sso7d domain, known to promote non-specific DNA binding, into the ISDge10-TnpB module. Additional fusions with transcriptional activators (VPR, SPH, or VPR-SPH composites) further strengthened gene activation, with the optimized N-Sso7d-TnpB-VPR construct exhibiting the greatest efficacy. Systematic targeting of TSS regions revealed variable fold induction depending on the chromosomal context, implying that the recruitment of transcriptional activators is influenced by spatial positioning. Comparative experiments across multiple cell types also demonstrated cell line-specific differences in activation, further supporting the role of chromatin architecture. Composite VPR-SPH fusions provided additive effects at responsive loci (e.g. *HBG*) but not at weaker targets, underscoring the need for locus-specific optimization. In contrast, restoring the guide length of the ωRNA to 20-nt enabled the ISDge10-TnpB system to induce efficient double-strand breaks and subsequent indel formation. Direct fusion of the Sso7d domain at the N-terminus of ISDge10 enhanced indel induction, likely because of non-specific DNA binding, which assisted target engagement. While these results demonstrate that fusion-based engineering can substantially improve the performance of the ISDge10-TnpB system, engineering of the TnpB effector itself represents another promising avenue for further enhancing efficiency [13, 28, 37]. Future engineering efforts integrating structure-guided design, directed evolution [41], and high-throughput mutational scanning [37] may enable the identification of variants with enhanced catalytic efficiency, target DNA recognition, ωRNA engagement, and structural stability while maintaining the compact architecture required for AAV-compatible delivery. Recent studies demonstrating engineered TnpB variants with substantially improved editing activity support the feasibility of this strategy and underscore the considerable potential of protein engineering to further expand the efficiency and versatility of compact TnpB-based genome editing platforms [28, 37].

Based on this optimized module, we demonstrated proof-of-concept disease modeling by targeting the *CFTR* locus to generate the common ΔF508 mutation via donor DNA-mediated HDR and conversely, to correct the mutation in ΔF508 disease-model cells. An HDR efficiency of up to 1.4% was obtained, establishing the practical feasibility of disease-model generation via selection-based strategies. The relatively low HDR efficiency may be attributable to multiple factors,

system. Histograms show indel frequencies determined by targeted deep sequencing of gDNA isolated from mouse liver 1 week after injection. Data are presented as mean ± SEM from three independent experiments ($n = 3$). NC: negative control. ISDge10: plasmid encoding N-Sso7d-ISDge10 TnpB and its corresponding ωRNA. (**E**) Schematic illustration of the rAAV-mediated *in vivo* delivery strategy and experimental timeline for genome editing at the mouse *Pcsk9* locus. Mice were intravenously injected with the AIO TnpB rAAV vector, and liver tissues were harvested 2 and 6 weeks post-injection for gDNA extraction and targeted deep sequencing. (**F**) Genome-editing efficiencies at the endogenous *Pcsk9* locus following administration of the AIO TnpB rAAV system. Histograms show indel frequencies determined by targeted deep sequencing of gDNA isolated from mouse liver at 2 and 6 weeks after rAAV administration. Individual biological replicate values ($n = 2$) are shown. NC: negative control. rAAV ISDge10 TnpB: rAAV encoding the AIO N-Sso7d-ISDge10 TnpB genome-editing system. Schematics were created in BioRender. Lee, S.H. (2026) https://BioRender.com/ni4b4rb.

Downloaded from academic.oup.com/nar/article/54/16/gkag849/8770891 by Chung Ang University, Anseong Campus Library user on 16 October 2026

including the inherently limited activity of the cellular HDR machinery, as well as the design and delivery efficiency of the donor template. Future strategies aimed at enhancing the HDR pathway, together with optimization of factors involved in HDR, may further improve HDR-mediated genome editing efficiency. In addition, the relatively broad editing window observed in the N-Sso7d-TnpB-based base editor represents a limitation that will require further refinement to achieve more precise base editing. Future optimization of the linker architecture between TnpB and the deaminase domain, together with engineering of the deaminase itself, may provide effective strategies for narrowing the editing window and improving editing precision.

The N-Sso7d-TnpB and N-Sso7d-TnpB-VPR systems developed in this study exhibited distinct advantages over canonical TnpB and CRISPR platforms. First, by fusing the Sso7d domain to TnpB or TnpB-VPR, these systems achieve enhanced DNA targeting, enabling either gene editing or transcriptional activation, depending on the programmed ωRNA. Thus, gene expression or editing outcomes can be flexibly controlled by adjusting the guide length of the ωRNA. Second, the ISDge10 TnpB-based modules demonstrated a relatively small gene size (1173 bp), which provides a major advantage for *in vivo* delivery using AAV vectors. N-Sso7d-TnpB and N-Sso7d-TnpB-VPR can be packaged intact within a single AAV, allowing efficient delivery depending on the serotype. In contrast, large effectors, such as Cas9 or Cas12a, often exceed the AAV packaging limit in their native or modified forms, requiring split-protein strategies that result in reduced efficiency. Third, beyond editing efficiency, ISDge10 TnpB provides an additional advantage in target accessibility. Although ISDra2 TnpB is one of the most extensively characterized and engineered TnpB genome editors, ISDge10 TnpB offers an important advantage in target accessibility owing to its shorter TAM requirement. Whereas ISDra2 TnpB requires a 5-bp TTGAT TAM, ISDge10 TnpB recognizes a 4-bp TTAT TAM, substantially expanding the number of targetable genomic sites. To further evaluate this advantage in therapeutically relevant genes, we analyzed the distribution of TTAT and TTGAT TAM sequences within the genomic regions of *CFTR*, *PCSK9*, and *TP53*. We identified 3 427 TTAT sites and 571 TTGAT sites within the *CFTR* locus, indicating that ISDge10 TnpB provides approximately a 6.00-fold broader targeting range than eISDra2 TnpBd. Similarly, *PCSK9* contained 172 TTAT and 52 TTGAT sites, corresponding to an ∼3.3-fold increase in targeSupplementary Table Sites, whereas TP53 contained 121 TTAT and 18 TTGAT sites, representing an ∼6.7-fold increase. These analyses demonstrate that the shorter TTAT TAM substantially expands the number of targetable genomic sites compared with the TTGAT TAM required by eISDra2 TnpBd (Supplementary Table S9). These analyses demonstrate that the shorter TTAT TAM substantially expands the number of targetable genomic sites compared with the TTGAT TAM required by eISDra2 TnpBd, highlighting the broader targeting scope and potential translational utility of the ISDge10 TnpB system. Finally, the short guide length ωRNA-based Sso7d-WT ISDge10 module offers versatile functionality through facile fusion with effector-domain. Beyond VPR-mediated transcriptional activation, linkage to methylases or demethylases, similar to those used in CRISPR systems, could enable epigenetic regulation *in vivo* [42].

This study is the first to demonstrate the functioning of an ISDge10 lineage endonuclease in human-derived cells to mediate both gene activation and editing. Considering that TnpB effectors remain relatively unexplored in mammalian contexts, further mechanistic studies are required to elucidate their DNA recognition principles and refine their precision while minimizing unintended off-target events. The dual-mode capabilities developed here, transcriptional activation and sequence editing, will require additional optimization of gene activation potency, editing efficiency, and *in vivo* delivery strategies to advance their therapeutic applications. Taken together, the N-Sso7d-TnpB/N-Sso7d-TnpB-VPR systems developed in this study exhibited both reprogrammable targeting ability and an ultracompact genetic footprint, rendering them highly compatible with AAV-mediated delivery. Our optimized ISDge10 TnpB platform enables efficient AAV delivery while supporting both transcriptional activation and genome editing. Compared with representative genome editors and activators, including compact CRISPR and OMEGA systems and the widely used SpCas9 platform, our system exhibited robust transcriptional activation and efficient genome editing activity. Importantly, the compact size of ISDge10 TnpB enabled robust transcriptional activation and genome editing following *in vivo* delivery. Furthermore, preclinical evaluation demonstrated efficient editing of the endogenous *Pcsk9* locus following systemic AAV administration, supporting the feasibility of the ISDge10 TnpB platform for gene therapy. Moreover, its modularity enables expansion into a multifunctional platform for gene activation, epigenetic regulation, and base editing through fusion with diverse functional domains. Despite these promising therapeutic features, comprehensive evaluation of genome-wide specificity remains essential for the further development and clinical translation of the ISDge10 TnpB platform. As targeted deep sequencing demonstrated that TAM-distal mismatches are relatively well tolerated by ISDge10 TnpB, similar off-target editing could potentially occur *in vivo* and lead to unintended consequences. Therefore, further studies employing orthogonal off-target detection methods for individual target sites will be necessary to comprehensively characterize the off-target landscape of the ISDge10 TnpB system. In addition, future studies should include rigorous preclinical validation in disease models, with particular emphasis on genome-editing and transcriptional activation efficiencies, long-term stability, and immunological safety. Successful validation in these areas would position this system as a promising next-generation platform for precision medicine.

Limitations of the study

In this study, the accuracy and efficacy of TnpB-based gene expression regulation and genome editing were rigorously validated in both cultured cells and mouse models, including AAV-mediated *in vivo* delivery. Future work should extend these investigations to disease-relevant animal models, where long-term safety, therapeutic efficacy, and tissue-specific delivery efficiency can be further established. Such studies will be essential to support the clinical translation of this platform.

Resource availability

Lead contact

For inquiries regarding resources, reagents, or materials generated in this study, please contact the lead author, Seung Hwan Lee (lsh080390@cau.ac.kr).

Downloaded from academic.oup.com/nar/article/54/16/gkag849/8770891 by Chung Ang University, Anseong Campus Library user on 16 October 2026

Materials availability

Plasmids and other unique, stable materials will be made available upon reasonable request and execution of an appropriate Materials Transfer Agreement.

Acknowledgements

Graphical abstract created in BioRender. Lee, S.H. (2026) https://BioRender.com/gjsbl0k.

Author contributions: Conceptualization, Y.O. and S.H.L.; Software, Y.O. and S.H.L.; Writing—original draft, Y.O. and S.H.L.; Writing—review & editing, S.H.L.; Materials, S.B.J., L.W.G., and B.W.S.; Visualization, Y.O. and S.H.L.; Supervision, S.H.L.; Project administration, S.H.L.; Funding acquisition, S.H.L.

Supplementary data

Supplementary data is available at NAR online.

Conflict of interest

The authors have filed a patent related to the content of this manuscript.

Declaration of generative AI and AI-assisted technologies

During the preparation of this manuscript, we used AI-assisted tools solely to refine grammar and improve linguistic clarity. All AI-generated suggestions were critically reviewed and, where appropriate, incorporated by the authors, who assume full responsibility for the final content of the publication.

Funding

This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT) (RS-2025-00554011, RS-2025-02218918, RS-2026-25526540, RS-2026-25505249). This research was also supported by the Korea Research Institute of Bioscience and Biotechnology (KRIBB) Research Initiative Program (KGM5382632), Republic of Korea, and grants from the Korean Fund for Regenerative Medicine funded by the Korean government (the Ministry of Science and ICT, and the Ministry of Health & Welfare; RS-2022-00090257). Funding to pay the Open Access publication charges for this article was provided by Republic of Korea, and grants from the Korean Fund for Regenerative Medicine funded by the Korean government (the Ministry of Science and ICT, and the Ministry of Health & Welfare; RS-2022-00090257).

Data availability

High-throughput sequencing, RNA sequencing, and whole-genome sequencing data have been deposited in the NCBI Sequence Read Archive (SRA) under BioProject accession numbers PRJNA1415478, PRJNA1415496, and PRJNA1499350, respectively.

References

1. Jinek M, Chylinski K, Fonfara I *et al.* A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* 2012;337:816–21. https://doi.org/10.1126/science.1225829
2. Qi LS, Larson MH, Gilbert LA *et al.* Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression. *Cell* 2021;184:844. https://doi.org/10.1016/j.cell.2021.01.019
3. Hsu PD, Lander ES, Zhang F. Development and applications of CRISPR-Cas9 for genome engineering. *Cell* 2014;157:1262–78. https://doi.org/10.1016/j.cell.2014.05.010
4. Doudna JA, Charpentier E. Genome editing. The new frontier of genome engineering with CRISPR-Cas9. *Science* 2014;346:1258096. https://doi.org/10.1126/science.1258096
5. Park CY, Kim DH, Son JS *et al.* Functional correction of large factor VIII gene chromosomal inversions in hemophilia a patient-derived iPSCs Using CRISPR-Cas9. *Cell Stem Cell* 2015;17:213–20. https://doi.org/10.1016/j.stem.2015.07.001
6. Nguyen TH, Anegon I. Successful correction of hemophilia by CRISPR/Cas9 genome editing *in vivo*: delivery vector and immune responses are the key to success. *EMBO Mol Med* 2016;8:439–41. https://doi.org/10.15252/emmm.201606325
7. Han JP, Kim M, Choi BS *et al. In vivo* delivery of CRISPR-Cas9 using lipid nanoparticles enables antithrombin gene editing for sustainable hemophilia A and B therapy. *Sci Adv* 2022;8:eabj6901. https://doi.org/10.1126/sciadv.abj6901
8. Zaib S, Saleem MA, Khan I. CRISPR-Cas9 genome engineering: trends in medicine and health. *Mini Rev Med Chem* 2022;22:410–21. https://doi.org/10.2174/1389557521666210913112030
9. Badon IW, Oh Y, Kim HJ *et al.* Recent application of CRISPR-Cas12 and OMEGA system for genome editing. *Mol Ther* 2024;32:32–43. https://doi.org/10.1016/j.ymthe.2023.11.013
10. Wu WY, Adiego-Perez B, van der Oost J. Biology and applications of CRISPR-Cas12 and transposon-associated homologs. *Nat Biotechnol* 2024;42:1807–21. https://doi.org/10.1038/s41587-024-02485-9
11. Altae-Tran H, Shmakov SA, Makarova KS *et al.* Diversity, evolution, and classification of the RNA-guided nucleases TnpB and Cas12. *Proc Natl Acad Sci USA* 2023;120:e2308224120. https://doi.org/10.1073/pnas.2308224120
12. Kim DY, Lee JM, Moon SB *et al.* Efficient CRISPR editing with a hypercompact Cas12f1 and engineered guide RNAs delivered by adeno-associated virus. *Nat Biotechnol* 2022;40:94–102. https://doi.org/10.1038/s41587-021-01009-z
13. Li Z, Guo R, Sun X *et al.* Engineering a transposon-associated TnpB-omegaRNA system for efficient gene editing and phenotypic correction of a tyrosinaemia mouse model. *Nat Commun* 2024;15:831. https://doi.org/10.1038/s41467-024-45197-z
14. Altae-Tran H, Kannan S, Demircioglu FE *et al.* The widespread IS200/IS605 transposon family encodes diverse programmable RNA-guided endonucleases. *Science* 2021;374:57–65. https://doi.org/10.1126/science.abj6856
15. Karvelis T, Druteika G, Bigelyte G *et al.* Transposon-associated TnpB is a programmable RNA-guided DNA endonuclease. *Nature* 2021;599:692–6. https://doi.org/10.1038/s41586-021-04058-1
16. Shmakov S, Smargon A, Scott D *et al.* Diversity and evolution of class 2 CRISPR-Cas systems. *Nat Rev Micro* 2017;15:169–82. https://doi.org/10.1038/nrmicro.2016.184
17. Sasnauskas G, Tamulaitiene G, Druteika G *et al.* TnpB structure reveals minimal functional core of Cas12 nuclease family. *Nature* 2023;616:384–9. https://doi.org/10.1038/s41586-023-05826-x
18. Schargel RD, Qayyum MZ, Tanwar AS *et al.* Structure of Fanzor2 reveals insights into the evolution of the TnpB superfamily. *Nat*

Downloaded from academic.oup.com/nar/article/54/16/gkag849/8770891 by Chung Ang University, Anseong Campus Library user on 16 October 2026

Struct Mol Biol 2025;32:243–6. https://doi.org/10.1038/s41594-024-01394-4
19. Nakagawa R, Hirano H, Omura SN *et al.* Cryo-EM structure of the transposon-associated TnpB enzyme. *Nature* 2023;616:390–7. https://doi.org/10.1038/s41586-023-05933-9
20. Mohanraju P, Wu WY. TldR: tnpB's evolutionary shift from transposon nucleases to RNA-guided transcriptional regulators. *Trends Microbiol* 2024;32:1039–41. https://doi.org/10.1016/j.tim.2024.09.004
21. High-throughput screening of TnpB proteins identifies hypercompact genome editors. *Nat Biotechnol* 2024;42:704–5. https://doi.org/10.1038/s41587-023-01858-w
22. Wang D, Tai PWL, Gao G. Adeno-associated virus vector as a platform for gene therapy delivery. *Nat Rev Drug Discov* 2019;18:358–78. https://doi.org/10.1038/s41573-019-0012-9
23. Wang JH, Zhan W, Gallagher TL *et al.* Recombinant adeno-associated virus as a delivery platform for ocular gene therapy: a comprehensive review. *Mol Ther* 2024;32:4185–207. https://doi.org/10.1016/j.ymthe.2024.10.017
24. Wan H, Kong D, Yan T *et al.* A compact and inducible dCas12f-based CRISPRa platform for programmable *in vivo* gene activation. *Nat Commun* 2026.17:1447. https://doi.org/10.1038/s41467-025-68183-5
25. Guo R, Li Z, Li G *et al.* *In vivo* treatment of tyrosinaemia with hypercompact Cas12f1. *Cell Discov* 2023;9:73. https://doi.org/10.1038/s41421-023-00554-y
26. Xiang G, Li Y, Sun J *et al.* Evolutionary mining and functional characterization of TnpB nucleases identify efficient miniature genome editors. *Nat Biotechnol* 2024;42:745–57. https://doi.org/10.1038/s41587-023-01857-x
27. Koonin EV, Makarova KS. Origins and evolution of CRISPR-Cas systems. *Phil Trans R Soc B* 2019;374:20180087. https://doi.org/10.1098/rstb.2018.0087
28. Marquart KF, Mathis N, Mollaysa A *et al.* Effective genome editing with an enhanced ISDra2 TnpB system and deep learning-predicted omegaRNAs. *Nat Methods* 2024;21:2084–93. https://doi.org/10.1038/s41592-024-02418-z
29. Zhang R, Tang X, He Y *et al.* IsDge10 is a hypercompact TnpB nuclease that confers efficient genome editing in rice. *Plant Commun* 2024;5:101068. https://doi.org/10.1016/j.xplc.2024.101068
30. Oh Y, Gwon LW, Lee HK *et al.* Highly efficient and specific regulation of gene expression using enhanced CRISPR-Cas12f system. *Gene Ther* 2024;31:358–65. https://doi.org/10.1038/s41434-024-00458-w
31. Abramson J, Adler J, Dunger J *et al.* Accurate structure prediction of biomolecular interactions with AlphaFold 3. *Nature* 2024;630:493–500. https://doi.org/10.1038/s41586-024-07487-w
32. Wang X, Li L, Guo L *et al.* Robust miniature Cas-based transcriptional modulation by engineering Un1Cas12f1 and tethering Sso7d. *Mol Ther* 2024;32:910–9. https://doi.org/10.1016/j.ymthe.2024.02.013
33. Zhou H, Liu J, Zhou C *et al.* *In vivo* simultaneous transcriptional activation of multiple genes in the brain using CRISPR-dCas9-activator transgenic mice. *Nat Neurosci* 2018;21:440–6. https://doi.org/10.1038/s41593-017-0060-6
34. Schwank G, Koo BK, Sasselli V *et al.* Functional repair of CFTR by CRISPR/Cas9 in intestinal stem cell organoids of cystic fibrosis patients. *Cell Stem Cell* 2013;13:653–8. https://doi.org/10.1016/j.stem.2013.11.002
35. Lapinaite A, Knott GJ, Palumbo CM *et al.* DNA capture by a CRISPR-Cas9-guided adenine base editor. *Science* 2020;369:566–71. https://doi.org/10.1126/science.abb1390
36. Xu X, Chemparathy A, Zeng L *et al.* Engineered miniature CRISPR-Cas system for mammalian genome regulation and editing. *Mol Cell* 2021;81:4333–45.e4. https://doi.org/10.1016/j.molcel.2021.08.008
37. Thornton BW, Weissman RF, Rodriguez JE *et al.* Engineered TnpB genome editors for plants and human cells identified by ribonucleoprotein mutational scanning. *Nat Biotechnol* 2026. https://doi.org/10.1038/s41587-026-03059-7
38. Wang M, Sun Z, Liu Y *et al.* Hypercompact TnpB and truncated TnpB systems enable efficient genome editing *in vitro* and *in vivo*. *Cell Discov* 2024;10:31. https://doi.org/10.1038/s41421-023-00645-w
39. Dahlman JE, Abudayyeh OO, Joung J *et al.* Orthogonal gene knockout and activation with a catalytically active Cas9 nuclease. *Nat Biotechnol* 2015;33:1159–61. https://doi.org/10.1038/nbt.3390
40. Jinek M, Jiang F, Taylor DW *et al.* Structures of Cas9 endonucleases reveal RNA-mediated conformational activation. *Science* 2014;343:1247997. https://doi.org/10.1126/science.1247997
41. Jiang K, Yan Z, Di Bernardo M *et al.* Rapid *in silico* directed evolution by a protein language model with EVOLVEpro. *Science* 2025;387:eadr6006. https://doi.org/10.1126/science.adr6006
42. Nunez JK, Chen J, Pommier GC *et al.* Genome-wide programmable transcriptional memory by CRISPR-based epigenome editing. *Cell* 2021;184:2503–19.e17. https://doi.org/10.1016/j.cell.2021.03.025

Downloaded from academic.oup.com/nar/article/54/16/gkag849/8770891 by Chung Ang University, Anseong Campus Library user on 16 October 2026

Received: February 12, 2026. **Revised:** July 25, 2026. **Accepted:** August 11, 2026
© The Author(s) 2026. Published by Oxford University Press.
This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial License (https://creativecommons.org/licenses/by-nc/4.0/), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited. For commercial re-use, please contact reprints@oup.com for reprints and translation rights for reprints. All other permissions can be obtained through our RightsLink service via the Permissions link on the article page on our site—for further information please contact journals.permissions@oup.com.