ATPase facilitates PlmCas12e-mediated gene editing through conformational modulation

Jie Zhao[1,2]†, Danyuan Li[1]†‡, Miao Yu[3,4]†, Zongli Zheng[3,4]*, Bing Tian[2]*, Jun-Jie Gogo Liu[1]*

[1]Beijing Frontier Research Center for Biological Structure, State Key Laboratory of Membrane Biology, Center for Synthetic and Systematic Biology, Center for Life Science and Artificial Intelligence, Tsinghua-Peking Center for Life Sciences, School of Life Sciences, Tsinghua University, Beijing 100084, China.

[2]Institute of Biophysics, College of Life Sciences, Zhejiang University, Hangzhou, China.

[3]Department of Biomedical Sciences and Tung Biomedical Sciences Centre, College of Biomedicine, City University of Hong Kong, Kowloon, Hong Kong SAR, China.

[4]Ming Wai Lau Centre for Reparative Medicine, Karolinska Institutet, Shatin, Hong Kong SAR, China.

† These authors contributed equally to this work.

‡ Current address: School of Life Sciences, Ecole Polytechnique Fédérale de Lausanne (EPFL), Lausanne, Switzerland.

* Corresponding author. E-mail: zongli.zheng@cityu.edu.hk; tianbing@zju.edu.cn; junjiegogoliu@tsinghua.edu.cn

Abstract

CRISPR-Cas systems represent powerful and versatile tools for genome engineering, yet the editing efficacy of compact effectors such as Cas12e is often constrained by the intrinsic energy barriers associated with conformational activation. Here, we develop a systematic screening strategy to evaluate the performance of various ATPases fused to SpCas9 and PlmCas12e effectors using diverse linker architectures, aiming to overcome conformational activation barriers through ATPase-driven remodeling. One lead architecture, PlmCas12e fused with DnaA at the N-terminus via a rigid linker, demonstrates significantly enhanced editing efficiencies across exogenous and endogenous targets in HEK293T cells. *In vitro* biochemical assays further suggest enhanced intrinsic catalytic activity of the fusion variant. Molecular dynamics simulations show that DnaA-fused variant exhibits a more confined Free Energy Landscape (FEL) and reduced Root Mean Square Fluctuation (RMSF) in critical domains. These findings support a model in which DnaA facilitates conformational pre-organization, thereby lowering the entropic cost of activation for PlmCas12e to transition into a catalytically active state. Overall, our study not only provides a robustly enhanced PlmCas12e variant for genome editing but also highlights conformational modulation as a strategy for engineering CRISPR effectors.

Keywords: CRISPR-Cas, PlmCas12e, ATPase, genome editing, conformational modulation

Introduction

The rapid development of CRISPR-Cas systems has revolutionized the field of genome engineering, providing powerful tools to manipulate the genetic code. Among the diverse array of CRISPR effectors, Cas12e (also known as CasX), belonging to the Class 2 Type V system, has attracted great interest due to its small molecular size and distinct DNA cleavage mechanism (Liu et al. 2019; Tsuchida et al. 2022). Despite these advantages, at many targeting sites, wild-type Cas12e variants often exhibit lower editing potency compared to the gold-standard *Streptococcus pyogenes* Cas9 (SpCas9) (Li et al. 2024). On the other hand, SpCas9 displays attenuated editing efficiency at some clinical targeting sites as well (Canver et al. 2015; Chari et al. 2015). These limitations motivate further protein engineering of Cas proteins to meet the demands of clinical and industrial applications.

To overcome the performance limitations of wild-type CRISPR effectors, researchers have developed several robust strategies for protein and guide RNA (gRNA) optimization. Rational design based on high-resolution structures has enabled the creation of high-fidelity variants, such as eSpCas9 and SpCas9-HF1, by neutralizing non-specific DNA-binding charges to reduce off-target effects (Chen et al. 2022; Kleinstiver et al. 2016; Slaymaker et al. 2016). Complementary to this, directed evolution, exemplified by methods like arginine scanning mutagenesis or PACE (Phage-assisted continuous evolution), has been successfully employed to broaden the PAM recognition profile and enhance catalytic activity (Hu et al. 2018; Huang et al. 2023; Jin et al. 2025). Beyond the protein, gRNA engineering, including the incorporation of chemical modifications and the optimization of scaffold secondary structures, have further improved the stability and activity of the RNP complex (Bin Moon et al. 2018; Chen et al. 2023; Kim et al. 2022; Park et al. 2018; Su et al. 2024). While these approaches primarily optimize local interactions, a general design principle based on global conformational modulation has yet to be established for broad CRISPR enhancement.

One of the bottlenecks limiting the editing efficiency of structurally dynamic Cas effectors is the energy barriers associated with the transition from DNA binding to catalytic activation (Sternberg et al. 2015; Strohkendl et al. 2018; Xing et al. 2024). We therefore hypothesize that the fusion of the ATPase proteins to the Cas effector could facilitate this transition, leveraging the mechanical work driven by ATP hydrolysis. In fact, many accessory proteins in nature, particularly ATPase-based DNA manipulation proteins, have evolved to facilitate DNA unwinding, protein-DNA assembly, and conformational remodeling (Eustermann et al. 2024; Hauk and Berger 2016). For instance, the DnaA protein, a specialized ATPase, plays a pivotal role in the initiation of DNA replication by binding to the *ori*C region and inducing local DNA strands separation (Messer et al. 1999; Fujikawa et al. 2003; Duderstadt et al. 2011). Thus, we propose that the fusion of such DNA-remodeling related ATPases with CRISPR

effectors could potentially lower the activation energy of the CRISPR complex, either by enhancing DNA accessibility or by stabilizing the effector's catalytically active conformation.

In this study, we systematically engineer Cas-ATPase fusion proteins by varying ATPase variant, fusion orientation, and linker architecture. Our screening identified DnaA ATPase as a potent enhancer of PlmCas12e-mediated genome editing activity. This enhancement was consistent across multiple genomic targets, suggesting a general improvement in enzymatic functionality. *In vitro* assays also indicated enhanced cleavage efficiency of the fusion variant. Molecular dynamics (MD) simulations suggests that DnaA fusion facilitates the conformational pre-organization of the effector into a catalytically competent state, thereby reducing the entropic penalty during the DNA cleavage process. Collectively, our work not only provides a robustly enhanced Cas12e variant but also highlights the feasibility of conformational modulation in the optimization of genome editing tools.

Results

Systematic screening identifies DnaA-fusion as a potent enhancer for PlmCas12e

To investigate whether the integration of ATPase domains could modularly enhance CRISPR effector performance, we firstly curated a library of ATPase proteins with diverse DNA-remodeling capabilities (Figure 1A, Table S1). These ATPase proteins were tethered to either the N- or C-terminus of SpCas9 and PlmCas12e using a flexible (GGSGGGS) or a rigid (PAPAPAP) linker. We then evaluated the designs in HEK293T cells using a GFP reporter assay to quantify gene disruption efficiency (Figure 1B, S1A, B).

In the initial phase, we screened the editing efficacy of SpCas9 protein fused with six ATPases (424 aa–747 aa) at two GFP targets. Most variants exhibited target-dependent improvements, and the constructs with rigid linkers moderately performed better than those with flexible linkers (Figure S2). However, we did not see consistent enhancement when we scaled up to five GFP targets (Figure S3A). Subsequent testing of SpCas9 proteins fused with three additional small ATPases (188 aa–363 aa) across these five GFP targets yielded similar results (Figure S3B).

In contrast to SpCas9, the fusion of ATPases to PlmCas12e, a compact Type V CRISPR effector, yields robust and consistent enhancement in special designs (Figure 1C, D). Notably, a specific architecture, PlmCas12e fused with an N-terminal DnaA domain via a rigid or a flexible linker (labeled as PlmCas12e-NR-DnaA and PlmCas12e-NF-DnaA), exhibited a significant increase of the GFP-negative HEK293T cells. To further validate this finding, we performed a repeat experiment and further evaluated the editing efficiency of the DnaA fusion variants using both next-generation sequencing (Figure 1E, S4A) and the GFP reporter assay (Figure S4B). The enhancement was consistent across replicates and detection methods. Furthermore, the fusion variant with a rigid linker (PlmCas12e-NR-DnaA) consistently outperformed the flexible-linker construct (DnaA-NF-PlmCas12e), suggesting that precise spatial positioning is critical for the synergy between the DnaA domain and the PlmCas12e scaffold.

PlmCas12e-DnaA exhibits enhanced editing efficiency across multiple genomic targets

The editing superiority of the PlmCas12e-NR-DnaA configuration was further validated across a broader range of genomic targets (Figure 2A, B). ATPase-fused PlmCas12e exhibited significantly higher editing efficiency at the EGFP reporter gene compared to wild-type PlmCas12e (Figure 2C). This enhancement was further corroborated at endogenous genomic loci. In HEK293T cells, PlmCas12e-NR-DnaA exhibited robust indel formation across seven native targets, with an average 1.34-fold increase of editing efficiency compared to the wild-type control (Figure 2D, E).

To determine whether the enhancement of gene editing efficiency depends on the intrinsic catalytic activity of DnaA, we introduced point mutations into the Walker A and Walker B motifs of DnaA (K178A/T179A and D235A/D236A), which are known to abolish ATP binding and hydrolysis (Hou et al. 2022; Katayama et al. 2010). Both DnaA mutant variants exhibited a marked reduction in gene editing activity, reverting to levels lower than the wild-type PlmCas12e (Figure 2F). This result indicates the necessity of a functional ATPase core domain, suggesting that ATP hydrolysis is required for the enhanced editing efficiency.

Having established that ATPase activity underlies the enhanced editing efficiency, we next asked whether this improvement extends across the genome. Using the GenomePAM approach, a GUIDE-seq-based assay targeting high-copy-number repetitive sequences in mammalian cells (Yu et al. 2025), we systematically profiled editing activity at thousands of genomic sites. We selected the Rep-1RC (5′-GGCCAGGCACAGTGGCTCAC-3′) sequence in Alu elements as the spacer, which has 8,471 occurrences across the human genome. The results showed that the total number of effectively edited Alu genome targets (on-target sites) generated by PlmCas12e-NR-DnaA was significantly higher than that of the wild-type control group, which was consistent with the single-locus assays (Figure 2G). Under the same sequencing depth, the number of on-target sites generated by the PlmCas12e-NR-DnaA fusion variant was higher than that of the wild-type PlmCas12e (Figure 2H), and the total number of effective on-target reads captured also showed a notable increase (Figure 2I). Meanwhile, analysis of the edited sequences using GenomePAM revealed that the fusion protein maintained the canonical PAM recognition preference of PlmCas12e (Figure S5A–F). Together, these results demonstrate that the DnaA fusion enhances editing efficiency in a genome-wide manner.

PlmCas12e-NR-DnaA exhibits enhanced DNA cleavage activity *in vitro*

To determine whether the enhanced editing efficiency observed in cells stems from improved chromatin accessibility or intrinsic enzymatic optimization, we next performed in vitro cleavage assays using purified components. Wild-type PlmCas12e (PlmCas12e-WT) and recombinant PlmCas12e-NR-DnaA proteins were reconstituted with sgRNA to form ribonucleoprotein (RNP) complexes, and time-course cleavage assays were carried out using a linearized double-stranded DNA substrate containing the cognate spacer and PAM (Figure 3A).

The PlmCas12e-NR-DnaA variant exhibited superior enzymatic performance compared to the WT enzyme (Figure 3B). The PlmCas12e-WT RNP showed slower cleavage rate and incomplete substrate depletion within the tested timeframe, while the PlmCas12e-NR-DnaA rapidly achieved a maximal cleavage plateau with a higher fraction of converted DNA. This result indicates that the enhancement arises, at least in part, from intrinsic

enzymatic properties. Given that the rate-limiting step for Cas12 effectors typically involves the energy-dependent formation of the R-loop and the subsequent conformational rearrangement required to activate the RuvC catalytic center (Stella et al. 2018; Strohkendl et al. 2024; Swarts and Jinek 2019), our observation suggests that the fused DnaA domain might lower the energetic barrier for these transitions, which facilitates a more rapid progression to the catalytically competent state, even in the absence of cellular factors.

Molecular dynamics reveals DnaA-driven conformational modulation

To elucidate the mechanism underlying the superior activity of the DnaA fusion protein, we performed all-atom molecular dynamics (MD) simulations using the AlphaFold 3-predicted ternary structures of PlmCas12e-WT and PlmCas12e-NR-DnaA in complex with sgRNA and target DNA (Hu et al. 2025). Root Mean Square Deviation (RMSD) analysis indicated that both systems reached equilibrium after 150 ns (Figure 4A, S6A).

We first analyzed the protein-DNA interaction interface. Intriguingly, when considering only the Cas12e effector module, the fusion variant exhibited significantly fewer intermolecular hydrogen bonds and a lower binding free energy (ΔG_{bind}) compared to the WT enzyme (Figure 4B, C). However, the inclusion of the DnaA domain reversed this trend, resulting in a higher total interaction strength for the fusion complex. Structural analysis confirmed that the DnaA DNA-binding domain (DBD) forms stable, canonical interactions with the DNA substrate during the simulation, adopting a conformation similar to the crystal structure of the *E. coli* DnaA-DBD-DNA complex (Fujikawa et al. 2003) (Figure 4D). This is consistent with a model in which DnaA functions as a primary anchor, effectively recruiting the DNA to protein while relieving the Cas12e effector from rigid interfacial constraints.

This “anchoring” effect efficiently enhanced the structural stability of the complex. The RMSD of the DNA substrate within the fusion complex was approximately one third lower than that in the WT complex (Figure 4E), indicating that DnaA acts as a conformational scaffold to minimize DNA fluctuation. Consistent with this, Root Mean Square Fluctuation (RMSF) analysis revealed that the fusion protein exhibits reduced flexibility in critical regions, specifically the Non-Target Strand Binding (NTSB) domain and the RuvC catalytic pocket (including the key catalytic residue D922) (Figure 4F), which are both crucial for Cas12e effector catalytic activity (Li et al. 2024).

Additionally, to characterize the thermodynamic properties of the enzymes, we constructed the Free Energy Landscape (FEL) as a function of RMSD and Radius of Gyration (Rg) (Figure S6B, C). The WT enzyme exhibited a rugged landscape featuring two distinct, deep energy basins separated by energy barriers, indicating that the protein oscillates between multiple metastable states (Figure 4G). In contrast, the Cas12e core of the fusion variant

converged into a single, focused free energy basin (Figure 4H). Collectively, these data suggest that DnaA may facilitate conformational pre-organization, narrowing the conformational search space and stabilizing the catalytic center in an active state (Frederick et al. 2007), thereby lowering the entropy penalty for DNA cleavage and enhancing the gene editing efficiency of PlmCas12e.

Discussion

In this study, we developed an engineered CRISPR-Cas12e system by tethering a DnaA ATPase domain to its N-terminus, significantly enhancing genome editing efficiency both in cells and *in vitro*. Our multi-dimensional analysis suggests that the DnaA domain acts as a conformational scaffold, optimizing the free energy landscape of PlmCas12e and facilitating its transition into a catalytically active state. Together, our results reveal conformational modulation as an underexplored axis for CRISPR engineering.

Our approach adds a unique dimension to the rapidly evolving field of CRISPR optimization. Traditional strategies have predominantly relied on structure-guided rational design and directed evolution, which focus on local amino acid substitutions to improve DNA binding or PAM recognition (Liu et al. 2022). More recently, AI-driven approaches have transformed the field by enabling the identification of distal mutations with long-range effects on enzymatic activity (Fei et al. 2025; Thomson et al. 2025). While AI performs well at navigating the vast sequence space, they primarily operate at the level of sequence-function relationships. In contrast, our work highlights a complementary biophysical perspective, demonstrating that modular fusion of an accessory domain can reshape the thermodynamic landscape of a CRISPR effector. This framework suggests that incorporating conformational dynamics into computational design may further improve the prediction and optimization of complex domain-fused architectures by AI models (Cavanagh et al. 2026).

Despite the promising results, we acknowledge several limitations in the current study that warrant further investigation. First, our screening library was relatively constrained, focusing on a selected set of well-characterized ATPases. Expanding this library through high-throughput synthesis or metagenomic mining may reveal better fusion partners. Second, while we validated PlmCas12e-NR-DnaA across several endogenous loci, broader genomic characterization, including large-scale off-target assessments and testing in diverse cell types, is necessary to establish its clinical utility. Third, it remains to be seen whether this DnaA-mediated enhancement is universally applicable to other compact Cas12e orthologs or even other CRISPR families like Cas12a. Furthermore, our mechanistic insights into conformational optimization were primarily derived from in silico molecular dynamics

simulations. Although the molecular simulation results provide a convincing theoretical basis, they lack direct experimental structural evidence. Future studies employing Cryo-EM (Cryogenic Electron Microscopy) or single-molecule fluorescence resonance energy transfer (FRET) assays will be crucial to visualize the precise structural transitions and confirm the improved conformational transition of the fusion complex during DNA engagement (Dagdas et al. 2017; Ha et al. 2024).

In summary, we have successfully engineered a PlmCas12e-NR-DnaA variant with enhanced editing capabilities, driven by optimized conformational dynamics. This study not only provides an optimized tool for gene editing but also highlights the potential of using accessory domains to modulate the energy landscapes of enzymes. Looking forward, combining the recruitment of natural functional domains with precise structural and conformational optimization could facilitate the rational design of gene editors with improved efficiency and specificity, thereby broadening their potential in therapeutic and biotechnological applications.

Materials and methods

Strains and cultural conditions

The cloning strain used in this study was *E. coli* DH5α chemically competent cells from Sangon Biotech. Liquid cultures were grown in LB medium (containing 10 g/L NaCl, 5 g/L yeast extract, and 10 g/L tryptone) or in TB medium (containing 24 g/L yeast extract, 12 g/L tryptone, 12.54 g/L K_2HPO_4, 2.31g/L KH_2PO_4). In contrast, solid cultures were maintained on LB agar plates (supplemented with 10 g/L NaCl, 5 g/L yeast extract, and 10 g/L tryptone, and 15 g/L agar). For protein purification, *E. coli* Rosetta (DE3) competent cells from Sangon Biotech were used. Stock solution of ampicillin (100 mg/L) and kanamycin (50 mg/L) were prepared, and unless otherwise indicated, this stock solution was added to the culture medium at a 1:1000 dilution.

Plasmids construction

For mammalian genome editing experiments, the plasmids pBLO62.5 (Addgene #123124) (Liu et al. 2019) and pBFC0398 (Sun et al. 2023) served as the expression backbones for PlmCas12e and SpCas9, respectively. Both vectors were codon-optimized for human cell expression and flanked by SV40 nuclear localization sequences (NLS) at both termini. To construct sgRNA expression vectors, oligonucleotide pairs containing the target spacer sequences were annealed, phosphorylated, and cloned into BbsI-digested backbones via Golden Gate assembly. For the construction of fusion proteins, coding sequences of ATPase domains were PCR-amplified and inserted into the N- or C-terminus of the Cas effectors using Gibson assembly (One Step Seamless Cloning Mix, CWBIO). For bacterial

protein expression and purification, *E. coli* codon-optimized genes encoding PlmCas12e variants were subcloned into a modified pET28b vector containing an N-terminal hexahistidine tag followed by a SUMO tag. Point mutations in the PlmCas12e-NR-DnaA variants were introduced via PCR-based site-directed mutagenesis using mutagenic primers. To generate templates for in vitro gRNA transcription, a DNA fragment containing the T7 promoter and the sgRNA scaffold was inserted into the pUC19 vector. A linear dsDNA template (approximately 0.7 kb) was subsequently PCR-amplified from this plasmid to serve as the substrate for in vitro transcription.

Cell culture and genome editing assay

HEK293T cells and HEK293T-GFP reporter cells were cultured in DMEM (Gibco) supplemented with 10% FBS (VWR). Cells were seeded into 96-well plates and transfected at 60-70% confluency using YALEPIC DNA/RNA Transfection Reagent-LFDR (Yali Biotech) with 200 ng of plasmid DNA. To select successfully transfected populations, puromycin (1.5 μg/mL) was added to cell culture medium 24 hours post-transfection. After 48 hours of selection, the cell culture medium was replaced with fresh medium without puromycin for an additional 24 hours. For GFP reporter assay, cells were harvested and washed twice with 1× PBS containing 0.5% FBS. GFP fluorescence intensity was quantified using a BD LSRFortessa. For subsequent NGS evaluation, genomic DNA was extracted using 50 μL of Lysis Buffer for Direct PCR (Genesand). The lysates were incubated at room temperature for 5 min, followed by incubation at 50°C for 30 min and 95°C for 3 min. All spacer sequences used in this study are listed in Table S2 and Table S3.

For GenomePAM assay, plasmids were transfected into HEK293T cells via Lipofectamine 3000 transfection reagent (Thermo Fisher Scientific) in strict accordance with the manufacturer's operating instructions. Prior to transfection, cells were seeded in 24-well plates at a density of 1×10^5 cells per well. For each well, the transfection mixture was prepared as follows: 500 ng of Cas expression plasmid, and 5 pmol of annealed dsODN were first combined with 1.5 μL of P3000 reagent in 25 μL of Opti-MEM reduced-serum medium. This mixture was then mixed thoroughly with a second solution containing 1.5 μL of Lipofectamine 3000 reagent diluted in 25 μL of Opti-MEM reduced-serum medium, yielding a 50 μL final volume of DNA–lipid transfection complex. The complex was incubated at room temperature for 10 min to allow for lipid–nucleic acid complex formation. Subsequently, the resultant transfection complex was added dropwise to individual wells, and the plates were incubated in a standard cell culture incubator for 72 h.

Next-generation sequencing and data analysis

Target genomic regions were amplified from genomic DNA samples using 2× Phanta Max Master Mix (Vazyme) in a 25 μL reaction volume. Thermal cycling was performed as follows: 98°C for 30 s; 20 cycles of 98°C for 15 s,

59°C for 15 s, and 72°C for 15 s; followed by a final extension at 72°C for 5 min. Illumina adapters and barcodes were added via a second round of PCR using 2× Q5 High-Fidelity Hot-Start Master Mix (New England Biolabs). The resulting amplicons were purified, pooled, and sequenced on an Illumina NextSeq platform with 160-bp paired-end reads. Raw sequencing data were processed using Trim Galore (v0.6.10) to remove adapters and low-quality reads, and paired-end reads were subsequently merged using FLASH (v2.2.00). Indel quantification was performed using CRISPResso2 (v2.3.2) (Clement et al. 2019) with a quantification window centered at the cleavage site. The editing frequency was calculated based on the percentage of reads containing indels. The sequences of PCR primers used for NGS library preparation are listed in Table S3.

Protein expression and purification

Plasmids encoding His-SUMO-tagged proteins were transformed into *E. coli* Rosetta (DE3) competent cells by heat shock at 42°C for 90 s, followed by incubation on ice for 20 min. Cells were recovered in 900 µL of LB medium at 37°C for 1 h and plated onto LB agar plates containing 50 µg/mL kanamycin. After overnight incubation at 37°C, single colonies were picked and inoculated into 100 mL of LB medium for future culture. After approximately 6 h, 20 mL of the culture was transferred into 750 mL of TB medium supplemented with 50 µg/mL kanamycin and 6 mL of 50% glycerol. When the OD_{600} reached 1.0-1.2, protein expression was induced with 0.4 mM IPTG, and cells were grown overnight at 16°C. Cells were harvested by centrifugation and resuspended in lysis buffer (20 mM HEPES pH 7.5, 500 mM NaCl, 10% glycerol, 50 mM imidazole, 0.5mM ATP, 1 mM TCEP). The suspension was lysed by sonication, and the lysate was clarified by centrifugation at 15,000 r.p.m. for 80 min at 4°C. The supernatant was incubated with pre-equilibrated Ni-NTA agarose beads (QIAGEN) by passing through the resin three times to maximize binding. After washing, the His-tagged protein was eluted using lysis buffer supplemented with 300 mM imidazole. To remove the His-SUMO tag, the eluate was incubated with Ulp1 protease at 4°C for 30 min. The mixture was then loaded onto a 5 mL HiTrap Heparin HP column (GE Healthcare) and eluted with a linear NaCl gradient from 200 mM to 1.5 M. Peak fractions were pooled and further purified by size-exclusion chromatography using a Superdex Increase 200 10/300 GL column (GE Healthcare) equilibrated with SEC buffer (20 mM HEPES pH 7.5, 150 mM NaCl, 10% glycerol, 0.5mM ATP, 1 mM TCEP). Fractions containing purified protein, as confirmed by SDS-PAGE, were aliquoted, flash-frozen in liquid nitrogen, and stored at -80°C. Protein sequences are listed in Table S4.

RNA preparation

The sgRNA was produced using *in vitro* transcription (IVT) as described previously (Sun et al. 2023). Briefly, DNA template was incubated with 0.5 mg/mL T7 polymerase and 4 mM NTP mix in IVT buffer (30 mM Tris-HCl pH

8.1, 25 mM $MgCl_2$, 0.01% Triton, 2 mM spermidine) for 3 h at 37°C. The reaction mixture was resolved on a 10% denaturing urea-polyacrylamide gel (Urea-PAGE). The target RNA band was excised, crushed, and eluted in soaking buffer (0.38 M NaAc pH 5.2, 0.8 mM EDTA, 0.8% SDS) overnight at 4°C. The eluted RNA was buffer-exchanged and concentrated using a 3 kDa MWCO concentrator (Merck Millipore). For precipitation, 0.3 volumes of 3 M NaOAc and 3 volumes of chilled 100% ethanol were added, and the sample was incubated at -80°C for 2 h. After centrifugation at 12,000 × g for 30 min, the pellet was washed twice with 70% ethanol and resuspended in nuclease-free water. To ensure proper folding, the RNA was diluted in refolding buffer (50 mM KCl, 5 mM $MgCl_2$), incubated at 72°C for 5 min, and slowly cooled to room temperature before being frozen and stored at -80°C. SgRNA sequences are listed in Table S5.

In vitro cleavage assay

Ribonucleoprotein (RNP) complexes were assembled by incubating 500 nM of purified Cas protein with 750 nM sgRNA at 25°C for 20 min in cleavage buffer (20 mM Tris-HCl pH 7.5, 150 mM NaCl, 10 mM $MgCl_2$, 1 mM DTT, 1 mM ATP). Subsequently, 30 nM dsDNA substrates were added and mixed with RNP on ice. The reaction mixtures were incubated at 37°C, and aliquots were withdrawn at the time points indicated in the figure legends. Reactions were quenched by the addition of 25 mM EDTA and 200 μg/mL RNase A (Beyotime), followed by incubation at 37°C for 1 h. Then, 1 mg/mL Proteinase K (Solarbio) was added, and samples were incubated for an additional 30 min at 50°C. Reaction products were separated by electrophoresis on a 2% agarose gel and subsequently quantified using ImageJ software (version 1.53k).

Double-stranded oligodeoxynucleotide preparation

Double-stranded oligodeoxynucleotide (dsODN) oligos used for GenomePAM assay were purchased from GENEWIZ and subjected to high-performance liquid chromatography (HPLC) purification. Individual oligos (dsODN-Top and dsODN-Bot) were resuspended in 1× Tris-EDTA (TE) buffer (Thermo Fisher Scientific) to achieve a final concentration of 100 μM. Subsequently, the oligos were annealed at a working concentration of 10 μM in 1× annealing buffer (10 mM Tris-HCl, 50 mM NaCl, 1 mM EDTA, pH 7.4) using a thermocycler. The thermal cycling protocol was set as follows: 95°C for 3 min; followed by 70 cycles of 95°C for 1 min with a ramp rate of -1°C per cycle; and a final hold at 4°C. The sequences of dsODN-Top and dsODN-Bot are listed in Table S6.

GenomePAM assay

Genomic DNA (gDNA) was extracted 72 hours post-transfection, and 500 ng of the purified gDNA was utilized for GenomePAM sequencing library construction in accordance with our previously published protocol (Yu et al. 2025). Briefly, the library preparation workflow comprises enzymatic fragmentation, end repair, A-tailing, adaptor ligation,

and two sequential rounds of anchored nested polymerase chain reaction (PCR). Library quantification was performed using KAPA Library Quantification kits, followed by sequencing on an Illumina NextSeq 1000 System with a 300-cycle kit (2× 150-bp paired-end reads). Sequencing was conducted following the standard Illumina workflow, with no modifications to indexing cycles or the requirement for customized sequencing primers. Raw sequencing data (FASTQ files) were processed and analyzed using the GenomePAM bioinformatics pipeline. To assess PlmCas12e cleavage dynamics as a function of increasing sequencing depth, datasets were subsampled at read counts ranging from 100,000 to 1,000,000 reads for subsequent GenomePAM analysis.

Molecular dynamics simulation and analysis

The initial ternary complex structures of PlmCas12e-WT and PlmCas12e-NR-DnaA were predicted using AlphaFold 3 (Abramson et al. 2024), with the top-ranked model (model_0) selected as the starting conformation for simulations (Table S7). MD simulations were performed using GROMACS (version 2025.1, CUDA-accelerated) employing the Amber99BSC1 force field and the TIP3P water model. Each complex was solvated in a cubic box, maintaining a minimum distance of 1.5 nm between the solute surface and the box edges. The systems were neutralized and adjusted to an ionic concentration of 0.15 M by the addition of Na^+ and Cl^- ions. All simulations applied periodic boundary conditions in three dimensions. Long-range electrostatic interactions were calculated using the Particle Mesh Ewald (PME) method, while short-range electrostatic and van der Waals interactions were cut off at 1.4 nm. The LINCS algorithm was used to constrain bonds involving hydrogen atoms. Energy minimization was performed using the steepest descent algorithm until the maximum force converged to <1,000 kJ mol^{-1} nm^{-1}. The systems were then equilibrated in two phases: a 100 ps NVT ensemble equilibration followed by a 100 ps NPT ensemble equilibration. During equilibration, temperature was maintained at 310 K using the V-rescale thermostat, and pressure was maintained at 1.0 bar using the Parrinell–Rahman barostat. Following equilibration, 300 ns production runs were conducted. To ensure statistical significance and independent sampling, three independent replicates were performed for each system. Each replicate was prepared individually from the solvation step, followed by independent energy minimization and equilibration (NVT and NPT) phases. In the NVT equilibration, initial velocities were generated using random seeds based on a Maxwell-Boltzmann distribution at 310 K.

Trajectory data were analyzed using standard GROMACS modules. RMSD and RMSF data were calculated using the gmx rms and gmx rmsf commands, respectively. Hydrogen bonds were analyzed using the gmx hbond command. Binding free energies were calculated using the gmxMMPBSA (v16.0) tool (Valdés-Tresanco et al. 2021) at 310 K. Free Energy Landscapes (FEL) were constructed using the gmx sham module with the temperature

parameter set to 310 K, projected onto the coordinates of RMSD and Rg. Data visualization was performed using Python 3.9 with the matplotlib (v3.5.2) package.

Statistics analysis

GraphPad Prism 10.6 software was used to analyze the data. Unless otherwise specified in figure legends, numerical values are presented as means ± SD. Significant differences between controls and treatments were tested using the two-tailed Student's *t* test. $P < 0.05$ was considered statistically significant.

Data availability

All data supporting the findings of this study are available in the main text and the supplementary materials.

Compliance and ethics

Tsinghua University has filed a patent that includes work described in this paper.

Acknowledgements

This work was supported by the STI2030-Major Projects (2022ZD0213900), the National Natural Science Foundation of China (32425030), the Agriculture Science and Technology Major Project and the New Cornerstone Science Foundation through the XPLORER PRIZE (J.J.G.L.).

Figures

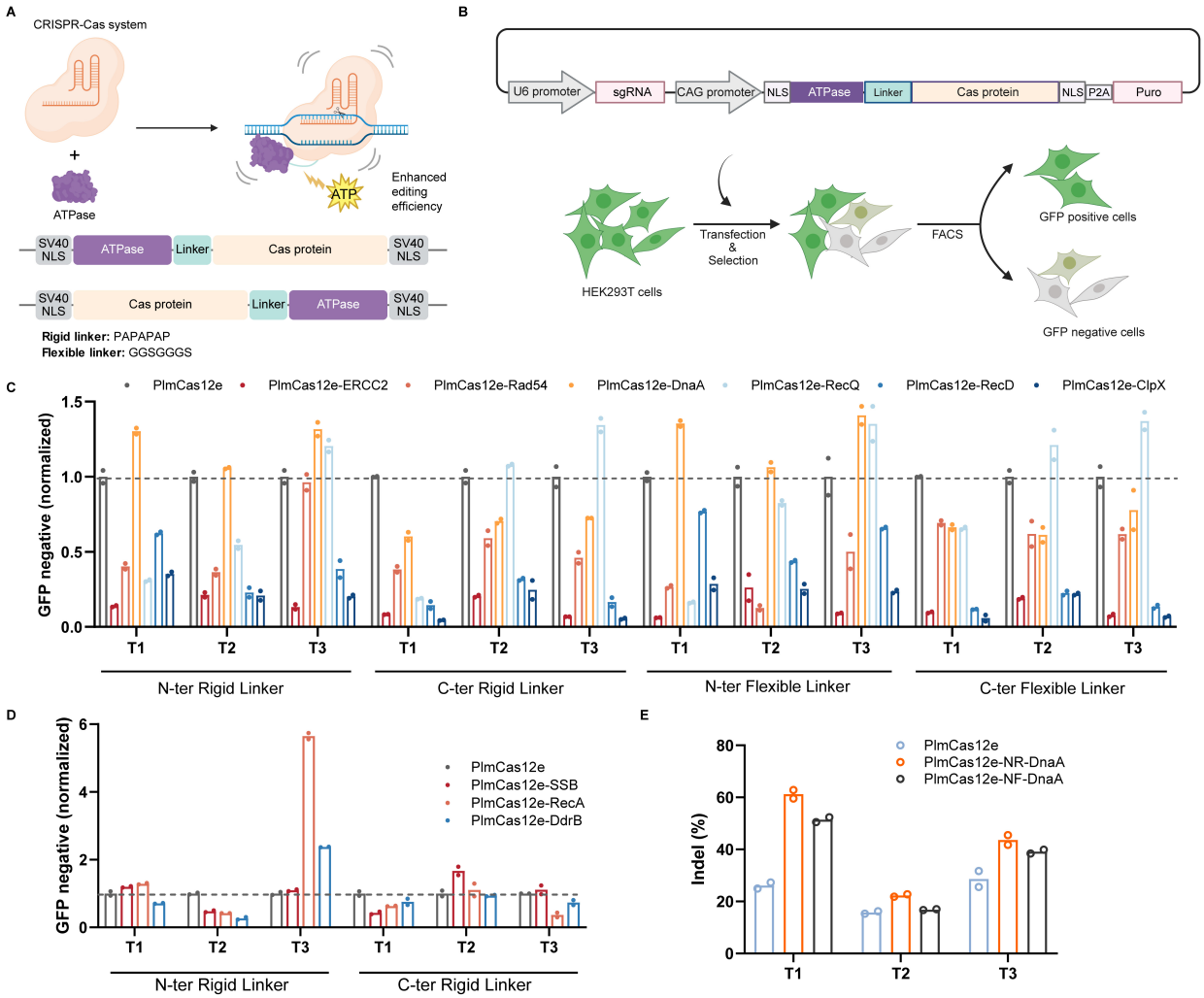

Figure 1 Initial editing efficiency screening of PlmCas12e-ATPase constructs. A, Schematic of Cas-ATPase fusion architectures designed for enhanced editing efficiency. B, Illustration of GFP reporter system for gene editing efficiency evaluation. C, Normalized frequency of GFP-negative cells for six ATPase variants fused to the N- or C-terminus of PlmCas12e via rigid or flexible linkers, targeting three EGFP loci. D, Normalized frequency of GFP-negative cells for three small ATPase variants fused to the N- or C-terminus of PlmCas12e via rigid linkers, targeting three EGFP loci. E, High throughput sequencing indicated indels generated by PlmCas12e-DnaA variants. Data are presented as the mean from two biological replicates.

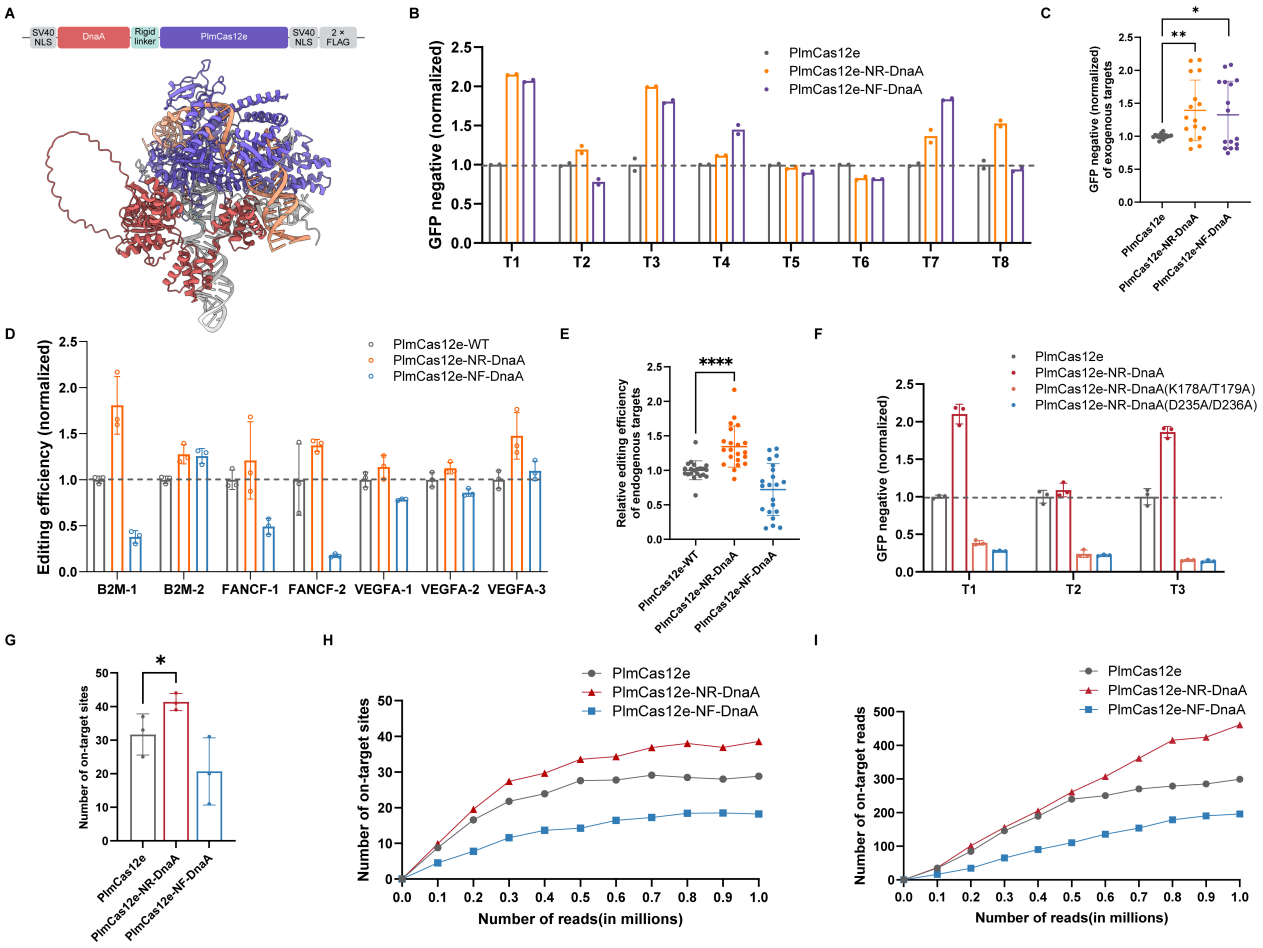

Figure 2 Enhanced editing efficiency of PlmCas12e-NR-DnaA at both exogenous and endogenous targets. A, Schematic representation of the PlmCas12e-NR-DnaA fusion construct. The ternary complex structure was predicted using AlphaFold 3. B, Normalized frequency of GFP-negative cells for PlmCas12e-DnaA variants targeting eight distinct EGFP sites in HEK293T cells. Data are presented as the mean from two biological replicates. C, Aggregate analysis of normalized GFP-negative cells across the eight exogenous targets shown in (B). D, Normalized indel frequencies generated by PlmCas12e variants at seven endogenous genomic loci, determined by high-throughput sequencing. E, Aggregate statistics of normalized indel frequencies across the seven endogenous targets shown in (D). F, Normalized frequency of GFP-negative cells for PlmCas12e and DnaA ATPase-deficient mutants (Walker A/B) targeting three EGFP sites. G, Number of effectively edited sites targeting Alu elements in HEK293T cells. H, I, The number of on-target sites (H) and total on-target reads (I) detected in the GenomePAM assay, plotted against varying raw sequencing depths (from 0.1 to 1 million reads). Data in D–I are presented as mean ± SD from three biological replicates. *P* values were calculated using a two-tailed Student's *t* test. $*P < 0.05$, $**P < 0.01$, $****P < 0.0001$.

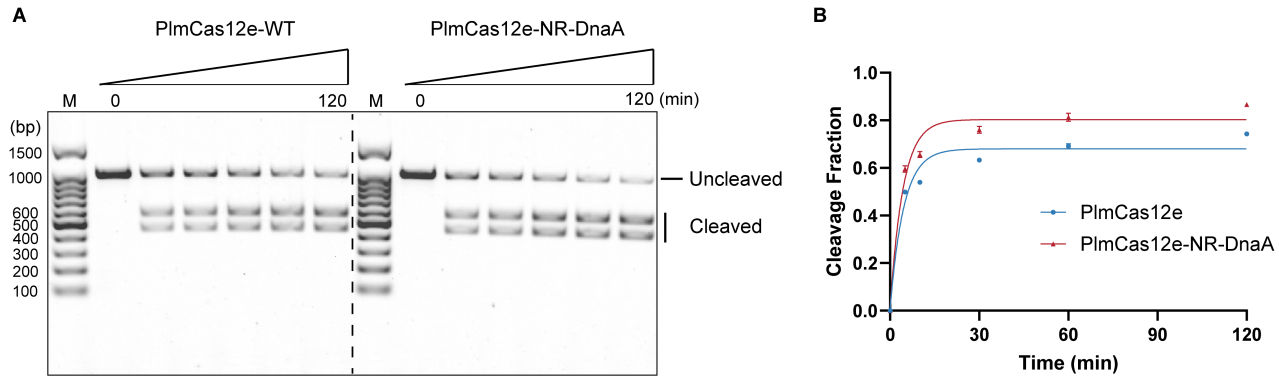

Figure 3 DnaA fusion enhances the DNA cleavage activity of PlmCas12e *in vitro*. A, Representative agarose gel showing the time-course cleavage of the linear dsDNA substrate by PlmCas12e-WT and PlmCas12e-NR-DnaA RNP complexes. Time points: 0, 5, 10, 30, 60, 120 min. M, Marker. B, Quantification of dsDNA cleavage rate. The fraction of cleaved DNA substrate is plotted against time. Data were fitted into a one-phase decay model. Data are presented as mean ± SD from three biological replicates.

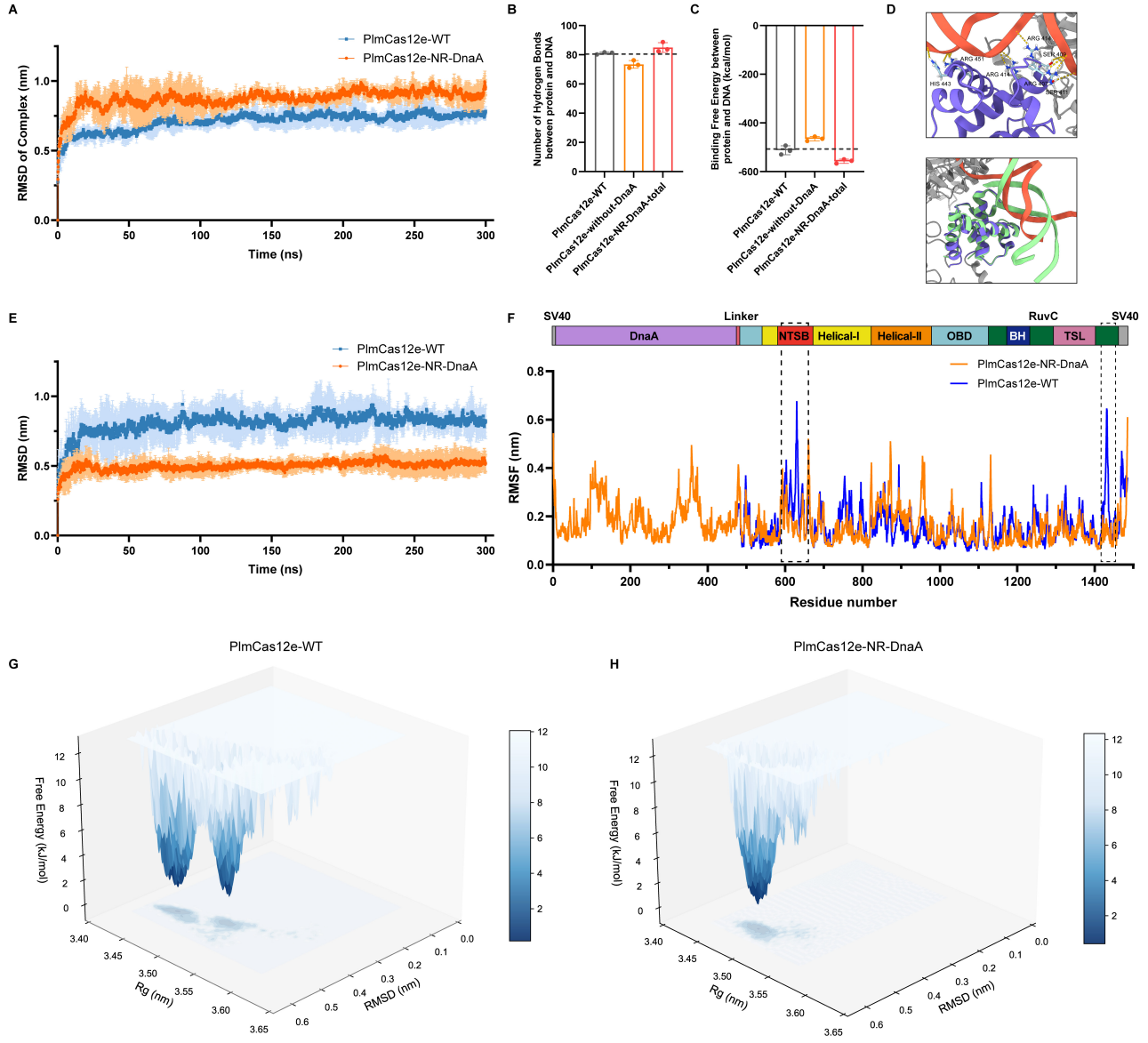

Figure 4 Molecular dynamics simulations reveal conformational optimization mediated by DnaA. A, RMSD trajectories for PlmCas12e-WT and PlmCas12e-NR-DnaA ternary complex in the MD simulation, the systems reached equilibrium after 150 ns. B, Average number of hydrogen bonds formed between the protein components and DNA, calculated over the final 150 ns of the simulation. C, Binding free energy between protein and DNA calculated over the final 150 ns of the simulation. D, Structural illustrations of the DnaA-DNA interface. Top: Snapshot from the 300 ns MD simulation showing hydrogen bond networks (yellow dashed lines) between the DnaA-DBD (purple) and DNA (orange). Bottom: Structural alignment of the simulated DnaA-DBD with the *E. coli* DnaA-DBD-DNA crystal structure (PDB: 1J1V, green). E, RMSD trajectories for DNA substrates within the PlmCas12e-WT and PlmCas12e-NR-DnaA complexes during the MD simulation. F, Per-residue RMSF plot for PlmCas12e-WT and PlmCas12e-NR-DnaA in the MD simulation over the last 150 ns. Key domains (NTSB, RuvC) showing reduced fluctuation are highlighted in dashed box. G, H, The free energy landscape against RMSD and Rg

for Cas effector protein of PlmCas12e-WT (G) and PlmCas12e-NR-DnaA (H) during a 300 ns MD simulation. Data in B, C are presented as mean ± SD from three independent replicates.

References

Abramson, J., Adler, J., Dunger, J., Evans, R., Green, T., Pritzel, A., Ronneberger, O., Willmore, L., Ballard, A.J., Bambrick, J., et al. (2024). Accurate structure prediction of biomolecular interactions with AlphaFold 3. Nature 630, 493–500.

Bin Moon, S., Lee, J.M., Kang, J.G., Lee, N.-E., Ha, D.-I., Kim, D.Y., Kim, S.H., Yoo, K., Kim, D., Ko, J.-H., et al. (2018). Highly efficient genome editing by CRISPR-Cpf1 using CRISPR RNA with a uridinylate-rich 3′-overhang. Nat Commun 9, 3651.

Canver, M.C., Smith, E.C., Sher, F., Pinello, L., Sanjana, N.E., Shalem, O., Chen, D.D., Schupp, P.G., Vinjamur, D.S., Garcia, S.P., et al. (2015). BCL11A enhancer dissection by Cas9-mediated in situ saturating mutagenesis. Nature 527, 192–197.

Cavanagh, P.E., Xue, A.G., Dai, S.A., Qiang, A., Matsui, T., Ting, A.Y. (2026). Computational design of conformation-biasing mutations to alter protein functions. Science 0, eadv7953.

Chari, R., Mali, P., Moosburner, M., Church, G.M. (2015). Unraveling CRISPR-Cas9 genome engineering parameters via a library-on-library approach. Nat Methods 12, 823–826.

Chen, W., Ma, J., Wu, Z., Wang, Z., Zhang, H., Fu, W., Pan, D., Shi, J., Ji, Q. (2023). Cas12n nucleases, early evolutionary intermediates of type V CRISPR, comprise a distinct family of miniature genome editors. Molecular Cell 83, 2768-2780.e6.

Chen, Y., Hu, Y., Wang, X., Luo, S., Yang, N., Chen, Y., Li, Z., Zhou, Q., Li, W. (2022). Synergistic engineering of CRISPR-Cas nucleases enables robust mammalian genome editing. Innovation 3.

Clement, K., Rees, H., Canver, M.C., Gehrke, J.M., Farouni, R., Hsu, J.Y., Cole, M.A., Liu, D.R., Joung, J.K., Bauer, D.E., et al. (2019). CRISPResso2 provides accurate and rapid genome editing sequence analysis. Nat Biotechnol 37, 224–226.

Dagdas, Y.S., Chen, J.S., Sternberg, S.H., Doudna, J.A., Yildiz, A. (2017). A conformational checkpoint between DNA binding and cleavage by CRISPR-Cas9. Science Advances 3, eaao0027.

Duderstadt, K.E., Chuang, K., Berger, J.M. (2011). DNA stretching by bacterial initiators promotes replication origin opening. Nature 478, 209–213.

Eustermann, S., Patel, A.B., Hopfner, K.-P., He, Y., Korber, P. (2024). Energy-driven genome regulation by ATP-dependent chromatin remodellers. Nat Rev Mol Cell Biol 25, 309–332.

Fei, H., Li, Y., Liu, Y., Wei, J., Chen, A., Gao, C. (2025). Advancing protein evolution with inverse folding models integrating structural and evolutionary constraints. Cell 188, 4674-4692.e19.

Frederick, K.K., Marlow, M.S., Valentine, K.G., Wand, A.J. (2007). Conformational entropy in molecular recognition by proteins. Nature 448, 325–329.

Fujikawa, N., Kurumizaka, H., Nureki, O., Terada, T., Shirouzu, M., Katayama, T., Yokoyama, S. (2003). Structural basis of replication origin recognition by the DnaA protein. Nucleic Acids Res 31, 2077–2086.

Ha, T., Fei, J., Schmid, S., Lee, N.K., Gonzalez, R.L., Paul, S., Yeou, S. (2024). Fluorescence resonance energy transfer at the single-molecule level. Nat Rev Methods Primers 4, 21.

Hauk, G., Berger, J.M. (2016). The role of ATP-dependent machines in regulating genome topology. Current Opinion in Structural Biology 36, 85–96.

Hou, Y., Kumar, P., Aggarwal, M., Sarkari, F., Wolcott, K.M., Chattoraj, D.K., Crooke, E., Saxena, R. (2022). The linker domain of the initiator DnaA contributes to its ATP binding and membrane association in E. coli chromosomal replication. Science Advances 8, eabq6657.

Hu, J., Guo, M., Gao, Q., Jia, H., He, M., Wang, Z., Guo, L., Liu, G., Gao, Q., Zhao, K.T. (2025). QBEmax is a sequence-permuted and internally protected base editor. Nat Biotechnol , 1–7.

Hu, J.H., Miller, S.M., Geurts, M.H., Tang, W., Chen, L., Sun, N., Zeina, C.M., Gao, X., Rees, H.A., Lin, Z., et al. (2018). Evolved Cas9 variants with broad PAM compatibility and high DNA specificity. Nature 556, 57–63.

Huang, T.P., Heins, Z.J., Miller, S.M., Wong, B.G., Balivada, P.A., Wang, T., Khalil, A.S., Liu, D.R. (2023). High-throughput continuous evolution of compact Cas9 variants targeting single-nucleotide-pyrimidine PAMs. Nat Biotechnol 41, 96–107.

Jin, S., Zhu, Z., Li, Y., Zhang, S., Liu, Y., Li, D., Li, Y., Luo, Y., Cheng, Z., Zhao, K.T., et al. (2025). Functional RNA splitting drove the evolutionary emergence of type V CRISPR-Cas systems from transposons. Cell 188, 6283-6300.e22.

Katayama, T., Ozaki, S., Keyamura, K., Fujimitsu, K. (2010). Regulation of the replication cycle: conserved and diverse regulatory systems for DnaA and oriC. Nat Rev Microbiol 8, 163–170.

Kim, D.Y., Lee, J.M., Moon, S.B., Chin, H.J., Park, S., Lim, Y., Kim, D., Koo, T., Ko, J.-H., Kim, Y.-S. (2022). Efficient CRISPR editing with a hypercompact Cas12f1 and engineered guide RNAs delivered by adeno-associated virus. Nat Biotechnol 40, 94–102.

Kleinstiver, B.P., Pattanayak, V., Prew, M.S., Tsai, S.Q., Nguyen, N.T., Zheng, Z., Joung, J.K. (2016). High-fidelity CRISPR–Cas9 nucleases with no detectable genome-wide off-target effects. Nature 529, 490–495.

Li, D., Zhang, S., Lin, S., Xing, W., Yang, Y., Zhu, F., Su, D., Chen, C., Liu, J.-J.G. (2024). Cas12e orthologs evolve variable structural elements to facilitate dsDNA cleavage. Nat Commun 15, 10727.

Liu, G., Lin, Q., Jin, S., Gao, C. (2022). The CRISPR-Cas toolbox and gene editing technologies. Molecular Cell 82, 333–347.

Liu, J.-J., Orlova, N., Oakes, B.L., Ma, E., Spinner, H.B., Baney, K.L.M., Chuck, J., Tan, D., Knott, G.J., Harrington, L.B., et al. (2019). CasX enzymes comprise a distinct family of RNA-guided genome editors. Nature 566, 218–223.

Messer, W., Blaesing, F., Majka, J., Nardmann, J., Schaper, S., Schmidt, A., Seitz, H., Speck, C., Tüngler, D., Wegrzyn, G., et al. (1999). Functional domains of DnaA proteins. Biochimie 81, 819–825.

Park, H.M., Liu, H., Wu, J., Chong, A., Mackley, V., Fellmann, C., Rao, A., Jiang, F., Chu, H., Murthy, N., et al. (2018). Extension of the crRNA enhances Cpf1 gene editing in vitro and in vivo. Nat Commun 9, 3313.

Slaymaker, I.M., Gao, L., Zetsche, B., Scott, D.A., Yan, W.X., Zhang, F. (2016). Rationally engineered Cas9 nucleases with improved specificity. Science 351, 84–88.

Stella, S., Mesa, P., Thomsen, J., Paul, B., Alcón, P., Jensen, S.B., Saligram, B., Moses, M.E., Hatzakis, N.S., Montoya, G. (2018). Conformational Activation Promotes CRISPR-Cas12a Catalysis and Resetting of the Endonuclease Activity. Cell 175, 1856-1871.e21.

Sternberg, S.H., LaFrance, B., Kaplan, M., Doudna, J.A. (2015). Conformational control of DNA target cleavage by CRISPR–Cas9. Nature 527, 110–113.

Strohkendl, I., Saha, A., Moy, C., Nguyen, A.-H., Ahsan, M., Russell, R., Palermo, G., Taylor, D.W. (2024). Cas12a domain flexibility guides R-loop formation and forces RuvC resetting. Molecular Cell 84, 2717-2731.e6.

Strohkendl, I., Saifuddin, F.A., Rybarski, J.R., Finkelstein, I.J., Russell, R. (2018). Kinetic Basis for DNA Target Specificity of CRISPR-Cas12a. Molecular Cell 71, 816-824.e3.

Su, M., Li, F., Wang, Y., Gao, Y., Lan, W., Shao, Z., Zhu, C., Tang, N., Gan, J., Wu, Z., et al. (2024). Molecular basis and engineering of miniature Cas12f with C-rich PAM specificity. Nat Chem Biol 20, 180–189.

Sun, A., Li, C.-P., Chen, Z., Zhang, S., Li, D.-Y., Yang, Y., Li, L.-Q., Zhao, Y., Wang, K., Li, Z., et al. (2023). The compact Casπ (Cas12l) 'bracelet' provides a unique structural platform for DNA manipulation. Cell Res 33, 229–244.

Swarts, D.C., Jinek, M. (2019). Mechanistic Insights into the cis- and trans-Acting DNase Activities of Cas12a. Molecular Cell 73, 589-600.e4.

Thomson, T., Li, G., Strilchuk, A., Cui, H., Wang, B., Li, B. (2025). Harnessing artificial intelligence to advance CRISPR-based genome editing technologies. Nat Rev Genet , 1–19.

Tsuchida, C.A., Zhang, S., Doost, M.S., Zhao, Y., Wang, J., O'Brien, E., Fang, H., Li, C.-P., Li, D., Hai, Z.-Y., et al. (2022). Chimeric CRISPR-CasX enzymes and guide RNAs for improved genome editing activity. Molecular Cell 82, 1199-1209.e6.

Valdés-Tresanco, M.S., Valdés-Tresanco, M.E., Valiente, P.A., Moreno, E. (2021). gmx_MMPBSA: A New Tool to Perform End-State Free Energy Calculations with GROMACS. J. Chem. Theory Comput. 17, 6281–6291.

Xing, W., Li, D., Wang, W., Liu, J.-J.G., Chen, C. (2024). Conformational dynamics of CasX (Cas12e) in mediating DNA cleavage revealed by single-molecule FRET. Nucleic Acids Res 52, 9014–9027.

Yu, M., Ai, L., Wang, B., Lian, S., Ip, L., Liu, J., Li, L., Tsai, S.Q., Kleinstiver, B.P., Zheng, Z. (2025). GenomePAM directs PAM characterization and engineering of CRISPR-Cas nucleases using mammalian genome repeats. Nat. Biomed. Eng , 1–14.