


Original Research Article

Semi-rational engineering of terminal deoxynucleotidyl transferase for high-efficiency enzymatic *de novo* DNA synthesisChengjie Zhang^a, Zhicong He^d, Zhengquan Gao^c, Chunxiao Meng^c, Wei Peng^{a,***}, Lei Du^{a,**}, Shengying Li^{a,b,*} ^a State Key Laboratory of Microbial Technology, Shandong University, Qingdao, Shandong, 266237, China^b Laboratory for Marine Biology and Biotechnology, Qingdao Marine Science and Technology Center, Qingdao, Shandong, 266237, China^c School of Pharmacy, Binzhou Medical University, Yantai, 264003, China^d School of Water and Environment, Key Laboratory of Subsurface Hydrology and Ecological Effect in Arid Region of Ministry of Education, Key Laboratory of Ecohydrology and Water Security in Arid and Semi-Arid Regions of Ministry of Water Resources, Chang'an University, Xi'an, 710054, China

ARTICLE INFO

Keywords:

Terminal deoxynucleotidyl transferase
DNA synthesis
3'-ONH₂-dNTPs
Enzyme engineering

ABSTRACT

The 21st century has witnessed rapid advancements in synthetic biology, with DNA synthesis emerging as a foundational technology. Conventional phosphoramidite-based methods face significant limitations, including short DNA elongation lengths (<300 nt), hazardous chemical waste, and low stepwise incorporation efficiency. Enzymatic DNA synthesis using terminal deoxynucleotidyl transferase (TdT) offers a promising alternative, enabling kilobase-scale assembly with greater efficiency and minimal environmental impact. Here, we identified *Bos taurus* TdT (BtTdT) through UniProt database mining as a catalytically active scaffold for natural and 3'-modified dNTPs. Comprehensive characterization of BtTdT's enzymatic properties—including pH, temperature, metal ion dependence, and substrate specificity—revealed its optimal conditions. Truncation of the BRCT domain generated variants with enhanced activity compared to wild-type BtTdT. Guided by AlphaFold3-predicted structural models, we engineered a quintuple mutant (M5: Bt15AA^{R336L/K338G/L397M/E456S/D395G}) optimized for 3'-ONH₂-dNTP incorporation. M5 exhibited 30-fold activity enhancement relative to the triple mutant M3 (Bt15AA^{R336L/K338G/L397M}) and achieved stepwise incorporation efficiency exceeding 98% in *de novo* synthesis of 10-nt ssDNA, demonstrating its potential for scalable enzymatic DNA synthesis. This work establishes a rational framework for TdT engineering through rational domain truncation and computational design, showing potential toward industrial-scale enzymatic DNA manufacturing.

1. Introduction

The next generation of DNA sequencing technologies [1–5], diverse gene editing and assembly tools such as CRISPR/Cas9 [6,7], Red/ET [8–10], Golden Gate [11], and Gibson Assembly [12], as well as other biotechnologies such as DNA origami [13,14] and DNA data storage [15–17], all rely on efficient DNA synthesis, including the synthesis of short primers, long genes, and even entire chromosomes/genomes [18–22]. At present, DNA synthesis is predominantly based on the sophisticated chemical solid-phase synthesis approach, namely, the phosphoramidite method [21] that was first introduced in 1981 [23].

Despite continual refinements, this method remains limited to synthesizing DNA strands up to 200–300 nucleotides in length, thus requiring additional assembly techniques to construct longer sequences [24–26]. The solid-phase DNA synthesis also has other disadvantages, such as the use of toxic reagents and production of hazardous waste, posing significant environmental concerns [20]. Furthermore, depurination during the synthetic process could break the DNA strands, which sometimes severely affect the synthetic efficiency [27–29].

To address these limitations, enzymatic DNA synthesis has been emerging. This kind of DNA synthesis is primarily realized by terminal deoxynucleotidyl transferase (TdT), a template-independent DNA

Peer review under the responsibility of Editorial Board of Synthetic and Systems Biotechnology.

* Corresponding author. State Key Laboratory of Microbial Technology, Shandong University, Qingdao, Shandong, 266237, China.

** Corresponding author.

*** Corresponding author.

E-mail addresses: weipeng@sdu.edu.cn (W. Peng), lei.du@sdu.edu.cn (L. Du), lishengying@sdu.edu.cn (S. Li).<https://doi.org/10.1016/j.synbio.2026.03.009>

Received 5 February 2026; Received in revised form 13 March 2026; Accepted 13 March 2026

2405-805X/© 2026 The Authors. Publishing services by Elsevier B.V. on behalf of KeAi Communications Co. Ltd. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

polymerase discovered in the 1960s that indiscriminately adds dNTP to the 3' end of single-strand DNA (ssDNA), thereby continuously extending the DNA strand [30]. Biologically, TdT was found to participate in the formation of diverse immunoglobulins and T cell receptors via V(D)J recombination (variable (V), diversity (D), and joining (J) gene segments) [31–33]. In recent years, TdT has been adapted to controlled DNA synthesis, having achieved average stepwise coupling efficiency exceeding 97%, making it a promising candidate to synthesize longer DNA sequences [34–38]. As TdT is the core of enzymatic DNA synthesis, great research efforts on optimizing TdT enzymes have been made on improving their catalytic efficiency and accuracy. The main strategies for TdT engineering include constructing combinatorial or saturation mutagenesis libraries for high-throughput screening [39–41], and structure-guided semi-rational design enabled by computational simulations [42]. Although these approaches have significantly improved TdT's catalytic efficiency and thermal stability, major bottlenecks remain that limit the industrial application of the best-in-class engineered TdTs. The key challenges include: (i) incomplete conversion across all substrates while maintaining high synthesis speed and fidelity, (ii) TdT's thermal stability that is still inferior to commercially available DNA polymerases used in PCR—higher operating temperatures would help prevent DNA secondary structure formation and enable enzyme recyclability, thereby reducing synthesis costs, and (iii) markedly reduced catalytic activity on DNA substrates prone to forming secondary structures [43–45].

In this study, we investigated the enzymatic properties of a select number of TdTs and identified *Bos taurus* TdT (BtTdT) as the most active candidate for incorporating 3'-ONH₂ reversible terminator dNTPs under the optimized temperature, pH, and divalent metal ion concentration. Enzyme engineering efforts including *N*-terminal truncation, site-directed mutagenesis, and semi-rational design using AlphaFold 3 (AF3) led to an optimal BtTdT mutant M5, showing the highest reported activity for 3'-ONH₂-dNTP-based DNA synthesis to date. Complete conversions of all 16 initiator DNA (iDNA) sequences with different terminal dinucleotides using all four 3'-ONH₂-dNTP were achieved by M5. While testing iDNAs with variable terminal trinucleotide sequences, M5 is active against almost all combinations except for two combinations (GTC + 3'-ONH₂-dGTP or 3'-ONH₂-dCTP). A stepwise DNA synthesis of ten nucleotides was achieved with a high yield of target product, demonstrating its great potential for precise and scalable enzymatic DNA synthesis.

2. Materials and methods

2.1. Strains, media, and chemicals

Escherichia coli DH5 α (Beijing Tsingke Biotech, China) was used for plasmid construction and cloning. Protein expression was performed in either *E. coli* JM109(DE3) or *E. coli* BL21(DE3) strains (Shanghai Weidi Biotechnology, China). Luria-Bertani (LB) medium (0.5% yeast extract, 1% tryptone, and 1% NaCl) was used to culture *E. coli* strains for plasmids amplification and extraction, seed preparation, and protein expression and purification. The medium was supplemented with 50 mg/L ampicillin for plasmid maintenance, and protein expression was induced with 0.6 mM isopropyl β -D-1-thiogalactopyranoside (IPTG). SYBRTM Gold nucleic acid gel stain was purchased from Thermo Fisher Scientific (USA). 3'-ONH₂-dNTPs were bought from Firebird Biomolecular Sciences (Florida, USA), and 3'-O-CH₂N₃-dNTP and 3'-OCHCHCN-dNTP were purchased from Huana Biomedical Technology Co. Ltd. (Hefei, China). Nitrous acid was acquired from Shanghai Yien Chemical Technology Co. Ltd. (Shanghai, China).

2.2. Construction of plasmids and strains

All TdT genes (see Table S3 in the Supporting Information) were codon-optimized for *E. coli* expression and synthesized by BGI Tech

Solutions (Beijing Liuhe, Beijing, China). The strains, primers, and plasmids used or constructed in this study are shown in the Supporting Information. The TdT genes were cloned by PCR using PrimeSTAR Max DNA Polymerase (Takara Biomedical Technology, Beijing, China). Plasmids were constructed via homologous recombination using the ClonExpress Ultra One Step Cloning Kit purchased from Nanjing Vazyme Biotech (Nanjing, China). TdT genes were inserted into the pETDuet-1 plasmid to express proteins carrying the *N*-terminal His-tag. Vector pETDuet-1 was digested with restriction endonucleases *Bam*HI and *Not*I. The recombinant plasmids were transformed into cloning strains, and single colonies were picked and sent to Sangon Biotech (Qingdao, China) for sequencing. The recombinant plasmids were then transformed into *E. coli* JM109(DE3) for protein expression and purification. Detailed procedures are provided in Tables S3 and S4.

2.3. Protein expression and purification of TdTs

E. coli JM109(DE3) harboring certain recombinant vector was grown overnight in 6 mL LB medium with 100 μ g/mL ampicillin at 37 °C with shaking at 220 rpm. The culture was used as a seed to inoculate 500 mL LB medium containing selective antibiotics and further grown at 37 °C, 220 rpm. When the culture optical density (OD₆₀₀) reached approximately 0.6, the inducing temperature was adjusted to 18 °C, and IPTG was added to a final concentration of 0.6 mM to induce protein expression at 220 rpm for 22 h. Cells were harvested by centrifugation (5000 \times g, 10 min, 4 °C), and cell pellets were re-suspended in 60 mL precooled lysis buffer (50 mM Tris-HCl, 10% glycerol, pH 7.0). All purification steps were performed at 4 °C. First, an Ultrasonic Homogenizer (JY92-IIDN) (Ningbo Scientz Biotechnology, China) was used for ultrasonication-based cell lysis (2 s on, 4 s off, and 30 min for total working time). After centrifugation at 10,000 \times g for 60 min to remove cell debris, the supernatant was transferred to a new 50 mL centrifuge tube. The recombinant His₆-tagged proteins in the supernatant were purified using an Ni-NTA agarose column (Qiagen, Germany). The column was washed eight times with 10 mL lysis buffer with 50 mM imidazole, and the target proteins were eluted using a protein elution buffer (50 mM Tris-HCl, 10% glycerol, 250 mM imidazole, pH 7.0). An Amicon Ultra centrifugal filter with a 30 kDa cutoff (Merck KGaA, Germany) was used to concentrate the protein solution at 5000 \times g for 30 min, and the elution buffer was exchanged with desalting buffer (100 mM Tris-HCl, 50% glycerol, pH 7.0) for three times. The concentration of the concentrated fraction (approximately 200 μ L) was measured by absorbance spectrophotometry using a NanoDrop One^C spectrophotometer (Thermo Fisher Scientific, USA). Finally, protein aliquots (50 μ L per tube) were flash-frozen in liquid nitrogen and stored at -80 °C for later use.

2.4. Accurate activity measurement based on fluorescent dye and gray value calculations

The TdT enzymatic reactions for dNTPs were performed as follows: 0.3 μ M purified TdTs, 4 μ M iDNAs (P0; Table S4), 1 mM mixture of all four dNTPs (each at 0.25 mM), 2.5 mM CoCl₂, and 100 mM Tris-HCl buffer (pH 7.0). Samples were mixed and incubated at 37 °C for 30 min to extend the iDNAs and then heated to 75 °C for 6 min to quench the reactions. The DNA products were analyzed on a 20% polyacrylamide gel with 8 M urea electrophoresis (DNA Urea PAGE), using a DNA Urea PAGE Electrophoresis Kit purchased from Real-Times Biotechnology (Beijing, China). The gel was stained with SYBR Gold nucleic acid gel dye and imaged on a Tanon 1600 Gel Image System (Tanon, China), and the products were quantitatively analyzed using ImageJ software based on gray value calculations.

2.5. Determination of the optimal reaction conditions for BtTdT (temperature, pH, and metal ions)

The standard reaction system consisted of 0.3 μM purified BtTdT, 4 μM iDNAs (P0; Table S4), 1 mM dNTPs, 2.5 mM CoCl_2 , and 100 mM Tris-HCl buffer (pH 7.0). The samples were incubated at 37 °C for 30 min and at 75 °C for 6 min. For optimal temperature determination, the reaction system and stopping conditions were the same as mentioned above, and the enzyme reaction mix was aliquoted into different tubes. The tubes were then incubated at temperatures from 20 to 50 °C with an interval of 5 °C. For optimal pH determination, the reaction system and conditions were the same except for the pH level of 100 mM Tris-HCl buffer, which was set from 5.0 to 9.0, with a gap of 0.5. For the effects of metal ions on enzyme activity, the reaction system and conditions were the same apart from the kinds of metal ions used, and eleven kinds of metal ions were compared.

2.6. Preparation of substrate iDNAs for magnetic-bead-based DNA synthesis

First, streptavidin magnetic beads (Dynabeads M-270 Streptavidin, Thermo Fisher Scientific Inc.) were prepared with wash buffer (10 mM Tris-HCl, 0.5 mM EDTA, 1 M NaCl, pH 7.5) to wash the beads twice and magnetic separation rack (Nanjing Vazyme Biotech, China) to separate the beads for 10 s and then remove the supernatant. Beads were suspended with wash buffer (2 \times) and transferred into equivalent volume of biotin-labeled oligonucleotides sample dissolved by ultrapure water. The mixture was blended thoroughly and incubated using a rotary mixer HS-3 (Ningbo Scientz Biotechnology, China) at room temperature for 10–30 min or at 4 °C for 2 h. Then, streptavidin magnetic beads were separated from the mixtures by magnetic separation rack for 1 min, and washed twice with wash buffer. Biotin-labeled oligonucleotides were eluted with DNA/RNA elution buffer (95% formamide, 10 mM EDTA, pH8.2) at 65 °C for 5 min or 90 °C for 2 min. iDNA strands elution condition was incubating samples in DNA/RNA elution buffer at 75–80 °C for 2 min, followed by magnetic separation for 30 s and then transferring the supernatant to a new nuclease-free tube. The products were separated through magnetic rack separation method [46] and washed twice to eliminate other impurities. The clean +1 products were put into next cycle to add another desired 3'-ONH₂-dNTP with 1 min reaction time. Only one nucleotide was incorporated for each cycle, and after 10 cycles, the desired nucleotides were synthesized in a controlled manner.

2.7. De novo DNA synthesis of 10 nt ssDNA by mutant M5

The 5'-biotinylated oligonucleotide P2 (Table S4) was immobilized on streptavidin magnetic beads as iDNA for *de novo* DNA synthesis. All nucleotide additions were performed at 42 °C for 1 min in EDS buffer, which contained 15% (v/v) DMSO, 50 mM *O*-benzylhydroxylamine HCl, 10% (v/v) glycerol, 0.05% (v/v) Tween 20, 2.5 mM tris-HCl, 40 mM NaCl, 0.5 mM Hepes, and 0.5 M cacodylic acid, with a final pH of 7.0 (measured in the absence of DMSO). Blocking groups of products 3'-ONH₂-dNTPs were readily to deblock by treatment with 0.8 M sodium nitrite deblocking buffer at pH 5.2 (adjusted by sodium acetate) for 1 min and generated natural 3'-OH. For the 10 cycles of DNA (5'-ATG-GATCCTC-3') synthesis, the modified nucleotides were added in the following order for each synthetic step: (1) 3'-ONH₂-dATP, (2) 3'-ONH₂-dTTP, (3) 3'-ONH₂-dGTP, (4) 3'-ONH₂-dGTP, (5) 3'-ONH₂-dATP, (6) 3'-ONH₂-dTTP, (7) 3'-ONH₂-dCTP, (8) 3'-ONH₂-dCTP, (9) 3'-ONH₂-dTTP, and (10) 3'-ONH₂-dCTP.

2.8. Enzyme-substrate interaction comparisons

Structural analysis of AlphaFold3-predicted mutant enzymes was performed using PyMOL to visualize substrate-binding interactions. The

workflow included: (i) selecting the analog substrate ATP as the reference molecule, and (ii) identifying polar contacts between ATP and surrounding enzyme atoms. Enzyme–substrate interactions were subsequently visualized as dashed lines.

3. Results and discussions

3.1. Phylogenetic tree analysis and screening of TdTs

Through bioinformatics-based homologous domain searches in the UniProt database, 226 TdTs were selected to build a phylogenetic tree. The sequence similarities between any two TdTs of the analyzed proteins are higher than 40%. These TdTs originate from different vertebrate taxa including *Cyclostomata*, *Pisces* (*Chondrichthyes* and *Osteichthyes*), *Amphibia*, *Reptilia*, *Aves* and *Mammalia*. Since TdTs of different origins show varied polymerization activities [39], it is essential to make parallel comparisons. Thus, we selected 11 representative TdTs having a variety of sequence identities with BtTdT that ranges from 40.9% to 98.8% for *in vitro* enzymatic activity assays (Fig. 1, Tables S1–S2). Their coding sequences were optimized, inserted into pETDuet-1 plasmid, and transformed into *Escherichia coli* JM109(DE3) strains. After heterologous expression and protein purification using Ni-NTA His-tag purification agarose, nine His-tagged recombinant TdTs with high homogeneity were acquired (Fig. S1). Notably, AmTdT from *Ambystoma mexicanum* (O57486) and EbTdT from *Eptatretus burgeri* (A0A8C4QBD0) failed to yield soluble target proteins.

To make accurate activity comparisons, the gray values of each band on a DNA urea PAGE gel were quantified using ImageJ software [47]. Among the tested TdTs, BtTdT exhibited the highest activity towards natural dNTP substrates (Fig. 1b). This result is inconsistent with the previous studies showing that ZaTdT exhibited higher activity than BtTdT [39,42]. We attribute this inconsistency to the combined effects of extended reaction time, elevated assay temperature, and the use of all four native dNTPs, which collectively enabled a more comprehensive assessment of enzyme performance beyond initial velocity measurements. Further experiments showed that BtTdT also demonstrated significant catalytic activities towards the unnatural substrates 3'-ONH₂-dNTP and 3'-OCH₂N₃-dNTP (Fig. S2). The observation that both the P1-CC and P1-TC experimental groups displayed multiple distinct bands at different reaction times might result from TdT's ability to remove a deoxynucleoside triphosphate from 3'-end of one oligonucleotide in the presence of pyrophosphate; that is to catalyze a reverse reaction known as pyrophosphorolysis for DNA polymerases; and the hydrolyzed deoxynucleoside triphosphate can be transferred to the 3'-end of another oligonucleotide, a process known as the pyrophosphorolytic dismutase activity [48,49]. Based on these results, BtTdT was selected as the starting enzyme for subsequent studies on enzyme property characterization, protein engineering, and stepwise DNA synthesis.

3.2. Catalytic properties of BtTdT

To understand the enzyme properties of BtTdT, we evaluated its optimal temperature, pH, and metal ion cofactors. Specifically, the enzymatic activities of BtTdT gradually increased from pH 5.0 to 7.0, and then decreased from pH 7.0 to 9.0 (Fig. 2a). Within the temperatures ranging from 20 to 50 °C, the highest activity was observed at 40 °C (Fig. 2b), which was approximately 58% and 22% higher than that at 20 °C and 30 °C, respectively. To identify the most effective metal ion cofactor and corresponding concentration for BtTdT, 11 different types of monovalent and divalent metal ions were evaluated for their effects on enzyme activity. As a result, Co²⁺/Mg²⁺/Mn²⁺/Zn²⁺ showed positive effects on the polymerization activity (Fig. 2c). Concentrations for these four metal ions were further investigated by testing the concentration range of 0–10 mM with a 2 mM interval. The optimal concentrations for Co²⁺, Mg²⁺, Mn²⁺, and Zn²⁺ were determined to be 2, 4, 4 and 2 mM,

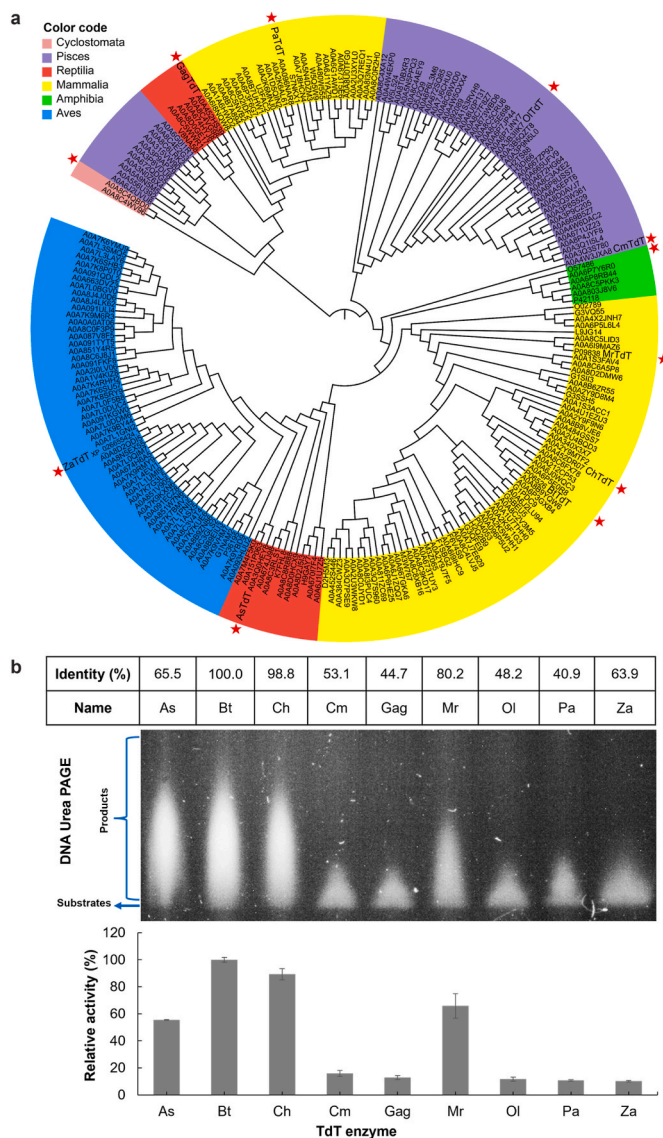


Fig. 1. Evolutionary analysis and functional characterization of TdT enzymes. (a) Phylogenetic analysis of 226 selected TdTs. (b) Protein sequence identity (relative to BtTdT) and activity comparison (4 μ M dT₂₀ and 1 mM mixture of all four dNTPs (each at 0.25 mM) as substrates) of nine selected TdTs (as indicated by red stars in (a) and the following abbreviations). Species name abbreviation: *Ambystoma mexicanum* (Am), *Alligator sinensis* (As), *Bos taurus* (Bt), *Capra hircus* (Ch), *Callorhynchus milii* (Cm), *Eptatretus burgeri* (Eb), *Gopherus agassizii* (Ga), *Mus musculus* (Mm), *Oryzias latipes* (Ol), *Papio anubis* (Pa), and *Zonotrichia albicollis* (Za).

respectively (Fig. 2d). Under combined optimal conditions at pH 7, 43 °C, and 1.5 mM Co²⁺, BtTdT exhibited a maximum activity (Fig. 2e).

3.3. Truncation of BtTdT to improve catalytic activity

The wild-type BtTdT comprises two structural domains: the BRCT domain and the catalytic domain POLXc. The POLXc domain consists of four subdomains, including the *N*-terminal α -helical subdomain (8 kDa domain), α -helical junction subdomain (finger domain), antiparallel β -sheet central subdomain (palm domain), and *C*-terminal subdomain (thumb domain) (Fig. 3a). Previous studies indicated that removal of the BRCT domain while preserving the complete POLXc domain could enhance enzyme expression, stability, and DNA elongation activity [42, 50]. Additionally, a conserved motif (residues 148–163) preceding the POLXc domain was found to be capable of stabilizing the catalytically

active conformation through intramolecular interactions (e.g., the interaction with Arg462 in the thumb domain) [51]. To investigate the effect of the pre-POLXc motif on the enzymatic performance of BtTdT, we constructed a series of BRCT-deletion variants with incremental retention of the *N*-terminal residues prior to POLXc (Fig. 3b). Enzymatic assays demonstrated that all truncated variants displayed a higher activity than the full-length BtTdT, with Bt15AA (residues 148–509, retaining 15 pre-POLXc amino acids, Fig. 3b) showing the highest catalytic efficiency (Fig. 3c). This activity enhancement was consistently observed for both 3'-ONH₂ and OCH₂N₃ modified dTTP substrates (Fig. S3). Notably, variants with excessive truncations (Bt0AA, Bt2AA, and Bt5AA) showed poor solubility during protein preparation (Fig. S4). Although the GST-tagged versions of Bt2AA and Bt5AA could be partially purified, both recombinant proteins exhibited low expression levels and lost their catalytic activity (Fig. S4), corroborating the essential role of the 16-residue motif in maintaining enzyme functionality [51]. These findings establish an effective truncation strategy for BtTdT engineering and provide mechanistic insights for future structure-guided optimization of TdTs.

3.4. Site-directed mutagenesis of BtTdT

The catalytic performance of TdT could be influenced by three key factors: (i) the terminal dinucleotide composition of iDNA, (ii) natural dNTP substrate specificity, and (iii) compatibility with 3'-modified reversible for terminators (3'-ONH₂, 3'-OCH₂N₃, and 3'-OCHCHCN dNTPs) [34,41,42,52]. To identify the optimal substrate for further engineering, we compared three classes of reversibly terminated dGTP analogues. Both wild-type BtTdT and the Bt15AA truncation mutant exhibited the highest activity toward 3'-ONH₂-dGTP (Fig. S2), thereby establishing this modified nucleotide as the preferred substrate for subsequent mutagenesis efforts.

To enhance catalytic efficiency, we performed structure-guided mutagenesis based on AF3-predicted three-dimensional models. Sequence alignment with MrTdT (*Mus musculus* TdT) revealed a conserved leucine residue (L397 in BtTdT, corresponding to L398 in MrTdT), the substitution of which with methionine was hypothesized to be capable of improving base-stacking interactions between the 3'-end bases of iDNA and the incoming dNTP molecule [39]. Our structural modeling confirmed the spatial proximity of L397 to the terminal nucleotide (Fig. 4a and Fig. S5), and the enzyme activity assays demonstrated that the Bt15AA^{L397M} mutant (M1) achieved a 1.5-fold increase in dNTP incorporation efficiency compared to Bt15AA (Fig. 4b). For the unnatural substrate 3'-ONH₂-dATP, M1 eliminated -1 nt byproducts compared with the wild-type enzyme (Fig. S2), and its +1 nt activity was comparable to that of the wild-type enzyme.

Further structural insights from ZaTdT guided additional mutagenesis. The R336L/K338G double mutant of ZaTdT was previously reported to enhance 3'-ONH₂-dNTP incorporation by enlarging the active site to better accommodate modified nucleotide conformations [39]. Comparative structural analysis revealed conserved spatial positioning of residues R336 and K338 between BtTdT and ZaTdT (Fig. 4c and Fig. S5), suggesting that analogous substitutions in BtTdT might generate similar benefits. To our delight, this was confirmed experimentally that the Bt15AA^{R336L/K338G} mutant (M2) exhibited a 5-fold increase in activity than wild-type BtTdT (Fig. 4d). Incorporation of the L397M mutation yielded the triple mutant Bt15AA^{R336L/K338G/L397M} (M3), which showed an additional 5% increase in catalytic efficiency with 3'-ONH₂-dTTP compared to M2 (Fig. 4d). Given the known influence of iDNA terminal dinucleotides on TdT activity [52], we further evaluated M3 against 16 iDNA substrates with varied terminal dinucleotide sequences and four types of 3'-ONH₂-dNTPs. Remarkably, M3 achieved complete substrate conversion within 10 min for the majority of combinations. Exceptions were observed for TC- and CC-terminated iDNAs extended with 3'-ONH₂-dATP, which only reached 70% and 50% conversions, respectively (Fig. 4e).

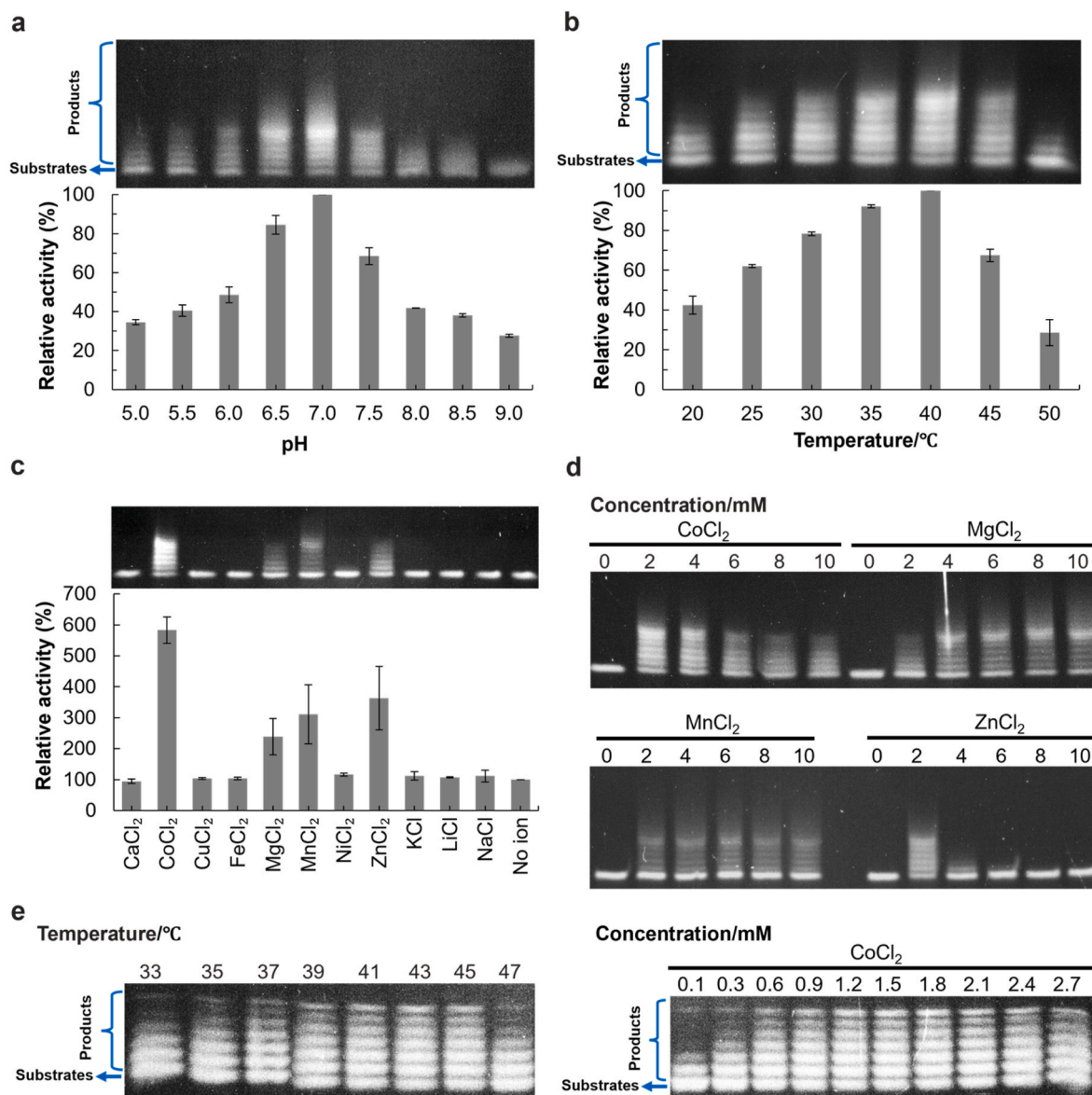


Fig. 2. The activities of BtTdT under different reaction conditions. (a) Optimal reaction pH of BtTdT. (b) Optimal reaction temperature of BtTdT. (c) Effects of different metal ions on BtTdT. (d) Concentration gradients (0–10 mM) of four metal ions ($\text{Co}^{2+}/\text{Mg}^{2+}/\text{Mn}^{2+}/\text{Zn}^{2+}$) for BtTdT. (e) More precise comparisons of the effects of temperature and Co^{2+} concentration on BtTdT activities.

The M3 variant demonstrated partial catalytic limitations, failing to achieve complete conversion of TC- or CC-terminated iDNAs with 3'-ONH₂-dATP within 10 min (Fig. 4e). In contrast, ZaTdT-R335L-K337G achieved full conversion of TC-terminated substrates under identical conditions (Fig. S6), motivating the following structure-guided engineering to address this performance gap. Structurally, there are 21 amino acid residues in total within 5 Å of the catalytic center, which is defined as the bound dATP in the catalytic pocket. By aligning the three-dimensional structures of M3 and ZaTdT-R335L-K337G, we found that two out of these 21 aligned residues exhibit sequence divergence. Specifically, M3 contains Thr396 and Glu456 at the positions occupied by smaller residues (*i.e.*, Ala and Gly) in ZaTdT-R335L-K337G (Fig. 5a). We reasoned that the bulkier side chains in M3 may create steric hindrance impeding 3'-ONH₂-dATP incorporation. To mitigate this constraint, we designed four compensatory mutations including T396G, T396A, E456G, and E456A. As expected, all variants surpassed M3's activity, with the E456G mutant achieving complete conversion of TC-terminated iDNA within 10 min and CC-terminated iDNA within 20 min (Fig. 5b). However, combining Bt15AA^{R336L/K338G/L397M/E456G} with

T396G or T396A (generating two quintuple mutants) conferred no additional benefit beyond that of the quadruple mutant Bt15AA^{R336L/K338G/L397M/E456G} (Fig. 5b). Of note, Bt15AA^{R336L/K338G/L397M/E456G} maintained broad substrate compatibility, achieving complete conversions for 15 out of 16 iDNA types with 3'-ONH₂-dATP (Fig. 5c). And the persistent challenge of CC-terminated iDNA conversion (60% efficiency) still requires additional protein engineering efforts.

3.5. Semi-rational design based on AlphaFold 3 predicted structures

Mechanistic insights into the critical role of residue E456 in 3'-blocked-dNTP incorporation efficiency prompted our further semi-rational engineering. We hypothesized that smaller amino acid substitution at this site would enlarge the catalytic pocket, thereby enhancing substrate flexibility and optimizing metal ion coordination geometry. To test this hypothesis, we employed AF3 for *in silico* modeling of twenty E456 variants of M3. Computational simulations incorporated ATP (a 3'-ONH₂-dATP analog) and catalytic metal ions to mimic physiological binding states. Enzyme surface topologies were visualized using PyMOL.

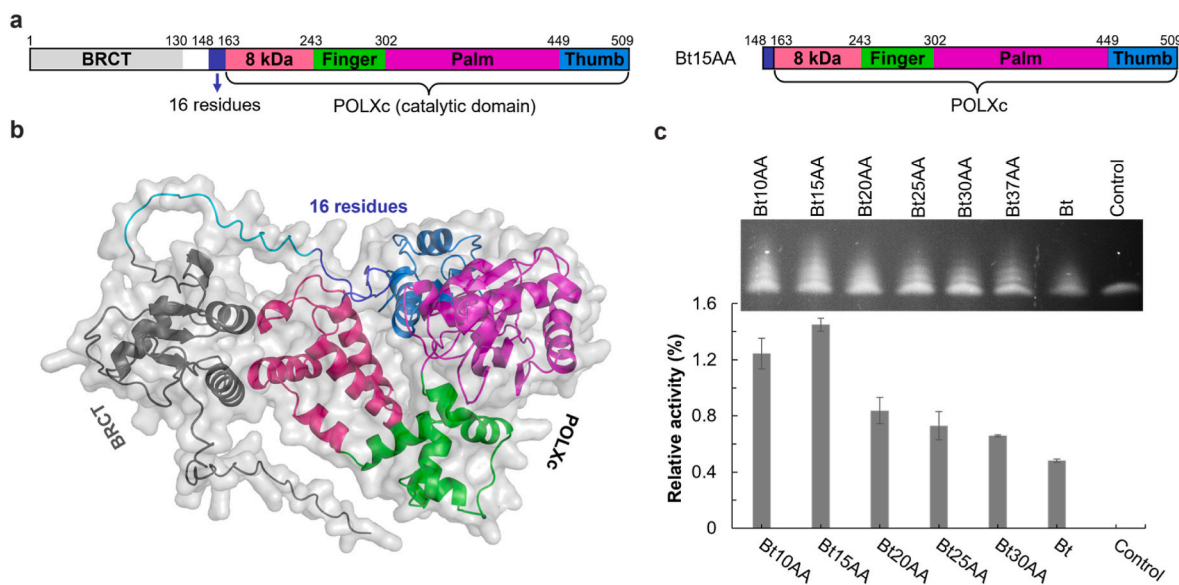


Fig. 3. BtTdT and its truncation variants. (a) Domain organization of intact BtTdT and the optimal truncation mutant Bt15AA. (b) The three-dimensional structure of BtTdT (predicted by AF3) with all structural domains colored using the same color code in (a). (c) Catalytic activities of different BtTdT truncation mutants.

Because direct measurement of the enzyme surface-to-3'-end distance is impractical, changes in this distance in mutants relative to the wild-type—whether increased or decreased—were used to infer corresponding alterations in enzyme surface proximity. AF3 prediction revealed that E456 substitutions increase the spatial separation between the 3'-end of the substrate and enzyme surface, with E456A, E456G, and E456S mutants showing progressive increases of 0.3, 0.6, and 0.7 Å, respectively, when comparing to the wild-type enzyme (Fig. 6a). Among these, E456S exhibits the most elevated position of the substrate sugar ring, suggesting a maximal reduction in steric hindrance and enhancement of conformational flexibility. We reasoned that the spatial reorganization would likely lead to a more stable and catalytically favorable substrate–metal interaction. Further experimental validation supported these computational predictions as the Bt15AA^{R336L/K338G/L397M/E456S} variant (M4) achieved complete conversion of all 16 iDNA types with 3'-ONH₂-dATP within 5 min (Fig. 6b). While testing the elongation time in 1 min with 3'-ONH₂-dATP, the three iDNAs (terminated with TC, CT, or CC) could not be fully transformed. We further tested three other types of 3'-ONH₂-dNTP substrates (16 iDNAs × 3 3'-ONH₂-dNTPs). Results showed that M4 also demonstrated universal high efficiency across these substrates in 5-min reactions (Fig. 6d). To gain mechanistic insight into enzyme specificity, iDNA P1-CC was substituted with P1-AC/TC/GC/AT/TT/CT/GT for structural simulations using AF3. The resulting models were then structurally aligned to identify deviations of the analog substrate ATP induced by iDNA in the catalytic pocket of M4 (Fig. S7). For iDNA P1-NC, the 3'-OH of dATP corresponding to P1-TC/CC was positioned closest to the enzyme surface; and for iDNA P1-NT, the 3'-OH associated with P1-CT protruded nearest to the surface. These structural simulations were basically consistent with the observed activity profiles, suggesting that, in addition to key residue mutations that govern substrate dATP positioning in the catalytic pocket, the 3'-terminal base composition of the iDNA substrate may also influence substrate mononucleotide placement. However, the effect of the 3'-terminal base composition is relatively minor compared with that of key residue mutations. We speculate that this mechanistic insight may also apply to other engineered TdTs, such as ZaTdT or orthologs closely related to BtTdTs.

To elucidate why the introduction of an NH₂ group and the three E456 variants enhanced enzymatic activity, we performed molecular dynamics (MD) simulations using the substrate 3'-ONH₂-dATP to examine its recognition in the wild-type (WT) and mutant enzymes. Our

simulations revealed that, relative to the constructed starting conformation, 3'-ONH₂-dATP readily undergoes conformational flipping, indicating that the NH₂ substitution decreases the stability of substrate binding. In the low-energy conformer overlays, the O atom at the E456A variant exhibits a displacement of 2.2 Å relative to the WT 3'-O atom, whereas the displacements increase to 3.7 Å and 5.5 Å in the E456G and E456S mutants, respectively (Fig. 6e). These positional shifts suggest that the altered low-energy conformations of the substrate reduce steric hindrance within the active site. This trend is consistent with the AF3 simulation results. Accordingly, we quantified the distances between the substrate's O–NH₂ group and the geometric center of the enzyme, the substrate centroid and the geometric center of the enzyme, and the substrate centroid and the geometric center of the active-site residues (G195, G196, H205, D206, D208, D258, T259, M260, D261, G311, W312, T313, G314, and N336), respectively (Fig. 6f and Fig. S8). The results consistently demonstrated a progressive increase in this distance across the WT, E456A, E456G, and E456S enzymes. These findings indicate that the NH₂ substitution at the 3'-O position enhances substrate conformational flexibility and weakens substrate–protein interactions, while mutations at E456 further shift the substrate away from the catalytic center. Collectively, these structural rearrangements enlarge the active-site cavity, thereby facilitating improved enzymatic activity.

To further enhance the catalytic efficiency, we next targeted residue D395 adjacent to T396, based on previous reports implicating it in dNTP binding preferences [53]. AF3 structural modeling reveals significant 3'-end displacement in D395 mutants. Specifically, D395M increases the substrate–enzyme surface distance by 0.7 Å when compared to the wild-type enzyme, while D395G achieves a remarkable 2.7 Å separation (Fig. 6c). This spatial reorganization suggests enhanced substrate freedom, particularly in D395G where the 3'-ONH₂ group adopts the most solvent-exposed conformation. Further functional assays confirmed these structural predictions. M4-derived variants (Bt15AA^{R336L/K338G/L397M/E456S/D395G} and Bt15AA^{R336L/K338G/L397M/E456S/D395M}) displayed significantly improved catalytic activity. Both Bt15AA^{R336L/K338G/L397M/E456S/395G} and Bt15AA^{R336L/K338G/L397M/E456S/D395M} achieved full conversion of 16 iDNAs with 3'-ONH₂-dATP in 3 min, while Bt15AA^{R336L/K338G/L397M/E456S/D395G} (M5) achieved complete conversion within 1 min (Fig. 6g). Furthermore, we synthesized and purified the previously characterized ZaTdT variants [39,40,42] and conducted direct activity comparisons with M5 for P1-CC adding 3'-ONH₂-dATP under identical reaction conditions (Fig. S9).

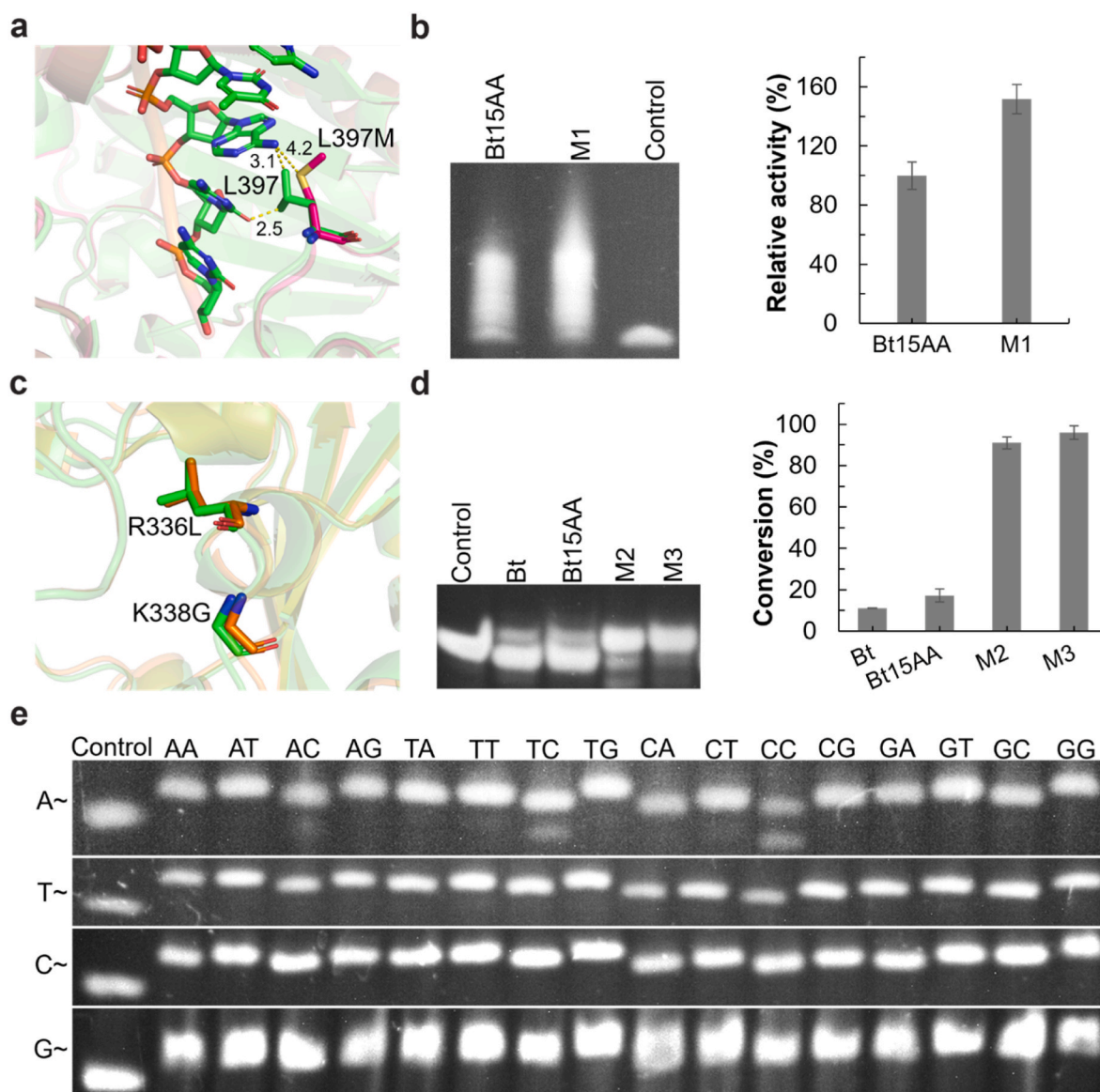


Fig. 4. Mutagenesis analysis of BtTdT variants. (a) L397 (green) or L397M (magenta) in BtTdT. (b) Activity test for M1. (c) R336/K338 in BtTdT (green) or R335/K337 in ZaTdT (orange). (d) Activity comparison of three BtTdT mutants. (e) Elongation by M3 with different iDNAs terminated with 16 dinucleotides and 4 types of 3'-OH₂-dNTPs. Reaction conditions: 1 mg/mL TdT, 1 μM iDNA, 42 °C, 10 min. "A~, T~, C~, and G~" stand for 3'-OH₂-dATP, 3'-OH₂-dTTP, 3'-OH₂-dCTP, and 3'-OH₂-dGTP, respectively. *Note:* Product bands at the same concentration may display different intensities due to sequence-specific effects.

The catalytic activities of the tested mutants at reaction time of 10 min increased as the following order: ZaTdT-R335L-K337G (conversion rate 20%) < ZaTdT-R335L-A193T-G337H-H478G (conversion rate 52%) < M5 ≈ M7-8/LG (conversion rate ~100%). Meanwhile, we compared the catalytic activities of several initiators with lower activities relative to M7-8/LG, including GCC, CGC, CTC, CCT, and CGA. M5 exhibited comparable or higher activities compared to M7-8/LG, such as CGA + A~ (with extension efficiencies of 95% and 68% for M5 and M7-8/LG, respectively), and both showed higher activities than the previously reported mutant ZaTdT-R335L-A193T-G337H-H478G (Supplementary Fig. S9). The M7-8/LG mutant also showed optimal soluble expression in *E. coli*, yielding approximately 12 mg of purified enzyme per liter of LB culture, whereas M5 and ZaTdT-R335L-A193T-G337H-H478G exhibited lower yields of 3.5 and 2.5 mg/L, respectively. To evaluate its full substrate scope, M5 was examined across all 64 combinations of 16 iDNAs and four types of 3'-OH₂-dNTPs. The M5 variant demonstrated universal substrate compatibility, achieving almost complete extension in 30-s reactions (Fig. 6h).

Substrate bias at the last three nucleotides of DNA initiators has been reported for different TdTs [41,42,52]. We tested 64 (4 × 4 × 4) distinct iDNAs with all possible nucleotide combinations at the last three positions for incorporation of 3'-OH₂-dNTPs (Fig. S10). As a result, M5 achieved near-complete conversion for most substrates but showed reduced activity toward CTC/CCC/CGA/CGT/CGC + 3'-OH₂-dATP combinations, and exhibited no detectable activity for GTC + 3'-OH₂-dC/GTP combinations (Fig. S9). Enzymatic assays further revealed divergent activities for GTC + 3'-OH₂-dC/GTP: the wild-type and M1 maintained catalytic turnover, whereas M3 and M4 completely lost activity (Figs. S9 and S10). Previously reported variants ZaTdT/LG and M7-8/LG also displayed substantially attenuated catalytic activity toward GTC + 3'-OH₂-dGTP or 3'-OH₂-dCTP, likely due to the collective influence of multiple residues, including R336L and K338G (Fig. S9).

When catalyzing the P1-CT + 3'-OH₂-dC and dGTP reactions, M5 exhibited a higher coupling rate than with P1-CC + 3'-OH₂-dATP, achieving complete conversion within 2 and 3 min under identical

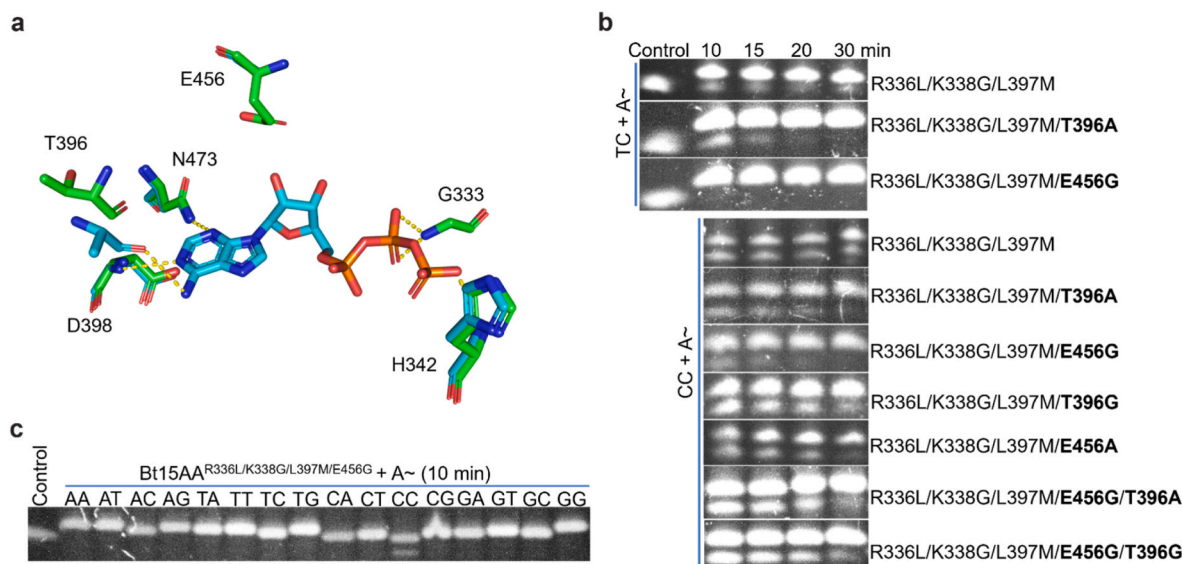


Fig. 5. Further mutational work based on M3. (a) Key amino acids at the catalytic center in BtTdT (green) and ZaTdT (cyan). (b) Activity evaluation for M3-based mutants. (c) Extension of 3'-ONH₂-dATP by Bt15AA^{R336L/K338G/L397M/E456G}. Reaction conditions: 1 mg/mL TdT, 1 μM iDNA, 42 °C, 10 min. *Note:* Product bands at the same concentration may display different intensities due to sequence-specific effects.

conditions, respectively (Fig. S11). Considering that the broad substrate promiscuity of an enzyme could enhance its application potential, we assessed M5's activity toward other reversible nucleotide terminators, 3'-OCH₂N₃ and 3'-OCHCHCN dGTP. Activity profiling showed that M5 lost significant catalytic activity toward these substrates (Fig. S12) relative to wild-type (Fig. S2). The inability of M5 to broaden its substrate scope likely stems from the sterically extended side chains of these additional substrates (when compared with 3'-ONH₂), rendering them incompatible with the NTP analog used in computational saturation mutagenesis.

Thermostability is often a significant challenge in TdT engineering, and trade-offs between enhanced activity and thermostability frequently occur during enzyme directed evolution. Thus, we measured the optimal reaction temperature of M5 (Fig. S12). The temperature optimum of M5 (~40 °C) coincided with that of wild-type BtTdT. Enzyme activity of M5 decreased sharply above 45 °C, indicating its limited thermostability. Further engineering will be required to enhance M5's thermal tolerance to withstand process-favored temperatures, such as 60 °C, which can help minimize potential DNA secondary structures.

We also analyzed the universality of our mutations to other TdT orthologs and the comparative structural features of BtTdT in a wider phylogenetic context (Fig. S13). R336 and K338 are strictly conserved across TdT orthologs, whereas the other three catalytic residue positions exhibit moderate sequence variation. Notably, the residues corresponding to R336 and K338 in previously reported hyperactive variants (ZaTdT-R335L-K337G, ZaTdT-R335L-A193T-G337H-H478G, and M7-8/LG) [39,40,42] were also mutated, highlighting their functional indispensability for enhancing the catalytic activity toward 3'-NH₂-dNTP substrates.

3.6. Mechanism analysis of improved mutants

The remarkable activity enhancement of M5 relative to M3 (extension time reduced from >30 min to 30 s) prompted mechanistic investigation through comparative structural analysis of the serial mutants including Bt15AA, M3, M4, and M5. To identify the structural determinants of the catalytic enhancement, we focused on the P1-CC/ATP substrate pair, a challenging combination that exhibited incomplete conversion by Bt15AA and M3 (>30 min) and only moderate efficiency by M4 (Figs. 5b and 6b). Comparative structural analysis revealed significant alterations in the interactions between the substrate and surrounding residues. Specifically, the number of substrate-interacting

residues decreases from nine in Bt15AA (Fig. 7a) to six in M5 (Fig. 7d), while the number of polar contacts such as hydrogen bonds and ionic interactions is reduced from nine to five. Similarly, the number of 5'-phosphate coordination bonds dramatically declines from seven to two. Importantly, the core phosphate-binding motifs (G341/H342 and G332/G333) remain conserved, thus preserving essential catalytic functions. This simplification of molecular interactions likely contributes to faster catalytic turnover through two synergistic mechanisms: (1) destabilizing pyrophosphate (PPi) retention post-cleavage to expedite product release [54], and (2) enabling 3'-ONH₂-dNTP reorientation for optimal nucleophilic positioning and metal ion coordination (Mg²⁺/Mn²⁺ bridging), which is critical for transition-state stabilization. The inverse correlation between interaction complexity and enzymatic activity perhaps reflects an evolutionary trajectory of optimization through controlled flexibility-minimizing nonessential contacts to improve substrate dynamics without compromising the integrity of the catalytic core.

3.7. De novo DNA synthesis by mutant M5

To evaluate the potential of M5 for *de novo* synthesis of long DNA sequences with 3'-ONH₂-dNTPs as building blocks, a two-step oligonucleotide extension protocol was established to elongate an 18-nucleotide iDNA. We employed a streptavidin-biotin affinity system utilizing magnetic beads to capture the iDNA products, taking advantage of the high-affinity interactions between streptavidin and biotin-labeled iDNAs. The optimal BtTdT mutant M5 catalyzed DNA synthesis using 3'-ONH₂-dNTPs as substrates in 1-min extension cycles. Each extension step was followed by a 1-min deprotection using 0.8 M sodium nitrite buffer (pH 5.2, adjusted using sodium acetate), which efficiently regenerated the natural 3'-OH terminus. By iteratively alternating between nucleotide incorporation and deprotection, full-length single-stranded DNA was synthesized *de novo*. After synthesis, DNA products were purified via magnetic separation and eluted for yield quantification. In a proof-of-concept demonstration, a 10-nucleotide ssDNA sequence (5'-ATGGATCCTC-3') was synthesized achieving an overall cumulative yield of >80% over 10 iterative cycles, with about 98% incorporation efficiency per step (0.8^{1/10}) (Fig. 6i). The synthesized 10-nucleotide ssDNA sample was verified by mass spectrometry analysis (Sangon Biotech, China), as the molecular weight of the enzymatic sample was consistent with that of the commercially synthesized standard, confirming the successful synthesis of the DNA with the intended length

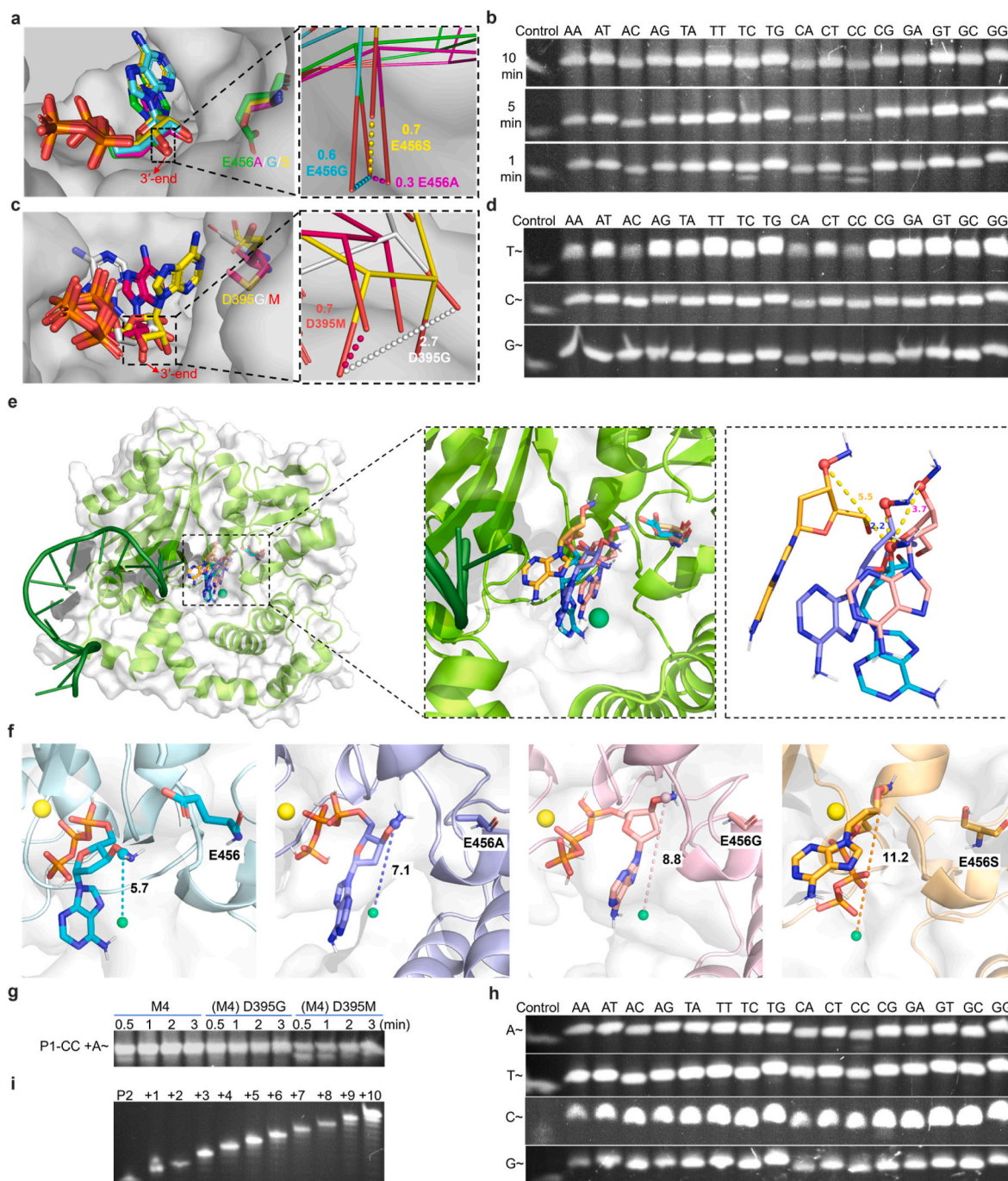


Fig. 6. Mutagenesis based on AF3-based structural analysis. (a) Position of surrogate substrate ATP in M3 variants with an amino acid substitution at E456. Enzyme-ligand complexes were predicted using AF3. E456S (yellow), E456G (aquamarine), E456 (green), and E456A (magenta). (b) Activity test of M4 mutant after adding 3'-ONH₂-dATP for different time. (c) Position of surrogate substrate ATP in M4 variants with an amino acid substitution at D395. The distances in angstrom from the substrate's 3'-end to the reference residues E456 or D395 (wild-type) are shown in different color for each mutant. Enzyme-ligand complexes were predicted using AF3. D395 (M4, yellow), D395G (white), and D395M (red). (d) Extension of 3'-ONH₂-dNTP by M4 in 5 min. Reaction conditions: 1 mg/mL TdT and 1 μM iDNA. (e) Low-energy binding conformations of 3'-ONH₂-dNTP in the WT and three E456 variants, along with their structural overlays. The distances between the 3'-O atom in the WT conformation and those in the mutant conformations are indicated. (f) The average distance between the substrate's O-NH₂ moiety and the geometric center of the WT and three E456 mutant backbones. (g) Time course assays of M4 mutations at D395. (h) Extension of 3'-ONH₂-dNTP by M5 in 30 s. Reaction conditions: 1 mg/mL BtTdT and 1 μM iDNA. (i) PAGE analysis of the 10 steps of extension by M5. Extension conditions: 1 mg/mL M5, 1 μM iDNA P2, 1 min. "A~", "T~", "C~", and "G~" stand for 3'-ONH₂-dATP, 3'-ONH₂-dTTP, 3'-ONH₂-dCTP, and 3'-ONH₂-dGTP, respectively. In panels (e) and (f), blue, purple, pink, and orange denote wild type, E456A, E456G, and E456S, respectively. *Note:* Product bands at the same concentration may display different intensities due to sequence-specific effects.

(Fig. S14).

4. Conclusions

Although enzyme-mediated *de novo* DNA synthesis represents a transformative technology attracting global research interest, its

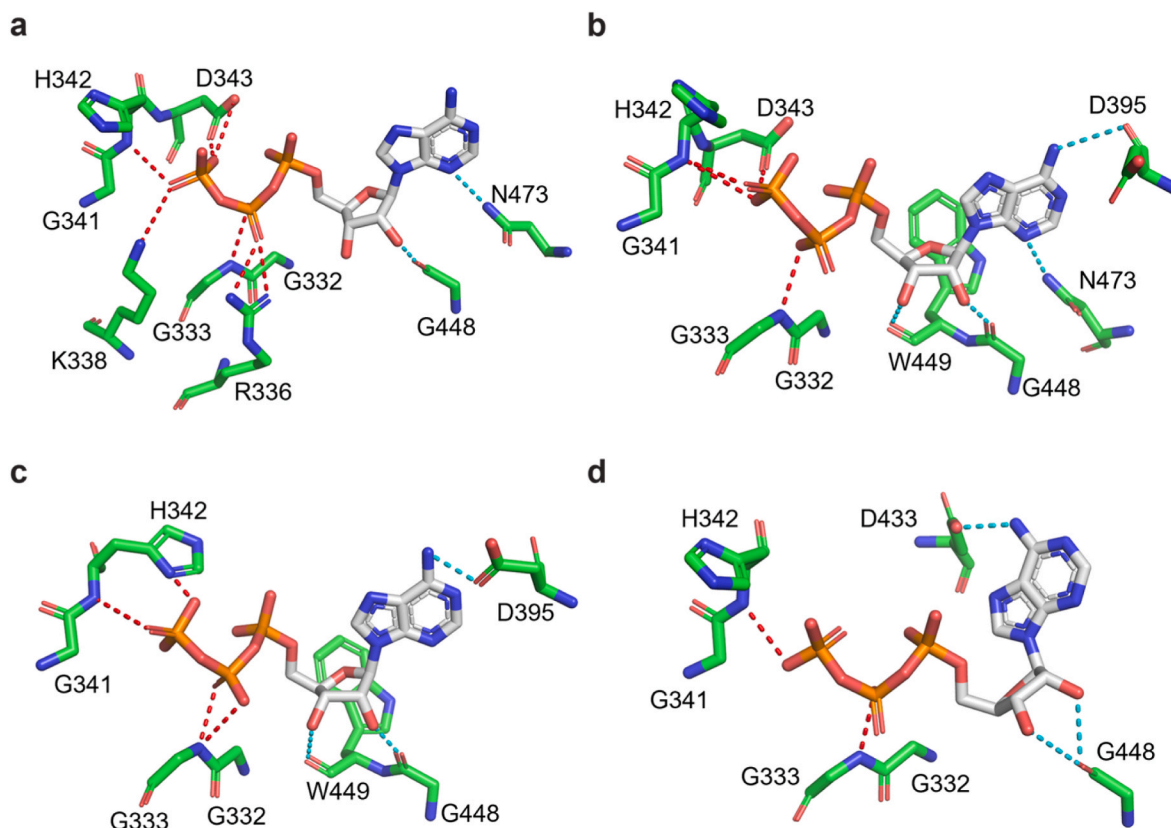


Fig. 7. Mechanism analysis of improved mutants through enzyme-substrate interactions comparisons. (a) The truncation mutant Bt15AA. (b) M3 (Bt15AA^{R336L/K338G/L397M}). (c) M4 (Bt15AA^{R336L/K338G/L397M/E456S}). (d) M5 (Bt15AA^{R336L/K338G/L397M/E456S/D395G}). The polar contacts between 5'-phosphate group of surrogate substrate ATP and surrounding amino acids are highlighted in red dashed lines. Amino acid residues are shown as green sticks, and substrate as white sticks.

practical application has been constrained by limited performance of TdT systems and inefficient synthetic workflows. In this work, through comparative analysis of natural TdTs, we identified BtTdT as a catalytically competent scaffold with significant engineering potential. Comprehensive characterization of BtTdT revealed critical limitations in substrate acceptance and reaction efficiency, motivating our structure-guided engineering efforts using 3'-ONH₂-blocked nucleotides as substrates. Sequential introduction of mutations (Fig. 8) effectively alleviated steric constraints on 3'-ONH₂-dNTP binding, driving a progressive increase in substrate conversions and shifting reaction equilibria toward near-complete incorporation of 3'-ONH₂-dNTPs. These results underscore that enhancing activity toward noncanonical substrates necessitates strategic disruption and reconfiguration of native substrate binding modes—rather than passive expansion of the binding pocket—to balance substrate mobility with precision alignment.

Of note, the engineering strategy in this study presents the following advantages: 1) It is a rational and targeted strategy; 2) The AlphaFold3-guided saturation mutagenesis at key sites covers all 20 amino acids, reducing the risk of missing beneficial mutants; and 3) The saturation mutagenesis of several rationally selected sites is more time-saving and cost-effective. By selecting only the top two or three mutants for protein purification and activity analysis, this approach significantly lowers cost, reduces workload, and improves efficiency, while still enabling the identification of mutants with substantially enhanced activity (e.g., the M5 variant).

The optimal mutant, Bt15AA^{R336L/K338G/L397M/E456S/D395G} (M5), demonstrated robust catalytic enhancement, enabling the *de novo* synthesis of medium-purity ssDNA of defined lengths (10 nt) with per-step incorporation efficiency exceeding 98%. Compared to its predecessor M3 (M3 and M5 exhibited conversion of ~2.7% and 87% at 5 min,

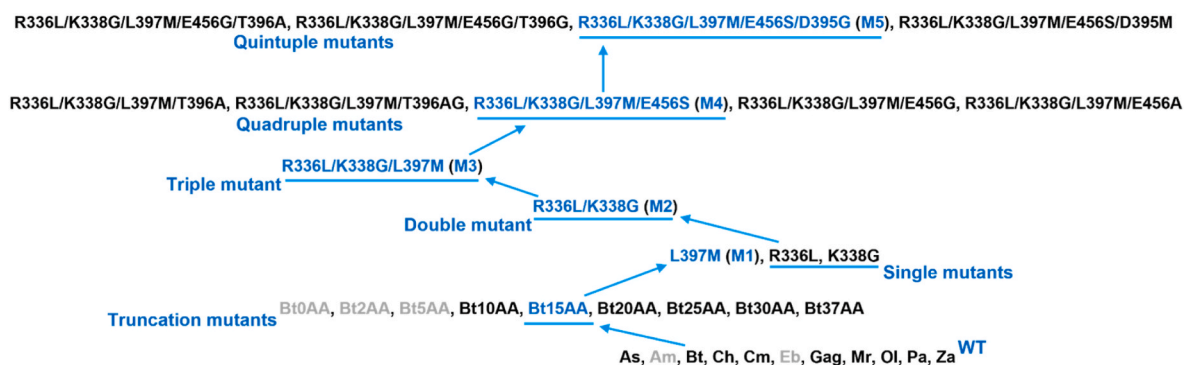


Fig. 8. The summary of the TdT mutants generated in this study. The enzymes in gray stand for insoluble proteins.

respectively), M5 exhibited a >30-fold increase in catalytic efficiency with 3'-ONH₂-dATP (Fig. S11). This improvement was achieved through semi-rational design using AF3, which provided accurate structural predictions to guide substrate-binding site remodeling. The five key mutation sites identified in this work provide a rational blueprint for further engineering TdT activity toward both canonical and noncanonical dNTPs. This framework may also be applicable for other structurally homologous PolX family enzymes. For dNTPs with bulky 3'-blocking groups, the situation is more complex. On one hand, their blocking groups are larger than that of 3'-ONH₂, thus requiring a larger binding pocket to reduce clashes with the surface of enzyme activity pocket and facilitate the search for more suitable catalytic coordination. On the other hand, their interactions with key amino acids in the enzyme's binding pocket differ more significantly from those of the natural 3'-OH compared to 3'-ONH₂, requiring the introduction of more diverse mutations to enhance catalytic activity. We suggest that combining saturation mutagenesis at key residue sites with MD simulations (or reliable molecular docking) could accelerate the development of high-activity variants. This approach may particularly enable targeted enhancement of catalytic efficiency for specific and challenging iDNA/mononucleotide substrate combinations.

Findings from our study and previous reports [34,35,39,42] lay the foundation for systematic screening and engineering TdT enzymes across broader orthologous families, with the aim of achieving high catalytic efficiency, broad substrate tolerance, and favorable expression profiles, thereby facilitating their potential application in industrial-scale DNA synthesis. Since thermostability of TdT enzyme is identified as a limitation for its catalytic reactions and industrial application in *de novo* DNA synthesis, it is necessary to engineer TdTs to obtain variants with enhanced thermal stability. Currently, many strategies are available to enhance enzyme thermostability, such as consensus design, stabilizing mutations, directed evolution, ancestral enzyme reconstruction, and utilizing deep learning tools (such as ProteinMPNN) for enzyme backbone design. As artificial intelligence (AI) technology becomes increasingly mature, leveraging AI to continuously learn, summarize and feedback until uncovering more fundamental principles for improving enzyme thermostability and facilitating efficient engineering, will become more viable. Furthermore, exploring new natural thermophilic enzymes can also contribute to engineering highly thermostable enzymes, such as other non-animal-derived PolX enzymes capable of *de novo* DNA synthesis. In summary, this work highlights two key advances: a mechanistic model for TdT engineering through the controlled destabilization of nonproductive substrate interactions, and a predictive design framework that integrates AF3 to streamline the enzyme optimization process. These advances represent a significant step toward scalable, high-fidelity enzymatic DNA synthesis for industrial application.

CRedit authorship contribution statement

Chengjie Zhang: Writing – original draft, Methodology, Formal analysis. **Zhicong He:** Writing – review & editing, Formal analysis. **Zhengquan Gao:** Writing – review & editing. **Chunxiao Meng:** Writing – review & editing. **Wei Peng:** Writing – review & editing, Formal analysis. **Lei Du:** Writing – review & editing, Supervision, Funding acquisition, Formal analysis. **Shengying Li:** Writing – review & editing, Supervision.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This work was supported by National Key Research and Development Program of China (2021YFA0911500), Shandong Provincial Natural Science Foundation (ZR2020ZD23), National Natural Science Foundation of China (32370032, 32170088), Taishan Young Scholars (tsqn202312032), and Fundamental Research Funds of Shandong University (2023QNTD001).

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.synbio.2026.03.009>.

References

- [1] Shendure J, Ji H. Next-generation DNA sequencing. *Nat Biotechnol* 2008;26(10):1135–45. <https://doi.org/10.1038/nbt1486>.
- [2] Metzker ML. Emerging technologies in DNA sequencing. *Genome Res* 2005;15(12):1767–76. <https://doi.org/10.1101/gr.3770505>.
- [3] Hutter D, Kim MJ, Karalkar N, Leal NA, Chen F, Guggenheim E, Visalakshi V, Olejnik J, Gordon S, Benner SA. Labeled nucleoside triphosphates with reversibly terminating aminoalkoxyl groups. *Nucleosides Nucleotides Nucleic Acids* 2010;29(11):879–95. <https://doi.org/10.1080/15257770.2010.536191>.
- [4] Guo J, Xu N, Li Z, Zhang S, Wu J, Kim DH, Sano Marmar M, Meng Q, Cao H, Li X, Shi S, Yu L, Kalachikov S, Russo JJ, Turro NJ, Ju J. Four-color DNA sequencing with 3'-O-modified nucleotide reversibly terminating terminators and chemically cleavable fluorescent dideoxynucleotides. *Proc Natl Acad Sci USA* 2008;105(27):9145–50. <https://doi.org/10.1073/pnas.0804023105>.
- [5] He S, Liu Y, Fang S, Li Y, Weng T, Tian R, Yin Y, Zhou D, Yin B, Wang Y, Liang L, Xie W, Wang D. Solid-state nanopore DNA sequencing: advances, challenges and prospects. *Coord Chem Rev* 2024;510:215816. <https://doi.org/10.1016/j.ccr.2024.215816>.
- [6] Sun X, Zhang H, Jia Y, Li J, Jia M. CRISPR-Cas9-based genome-editing technologies in engineering bacteria for the production of plant-derived terpenoids. *Eng Microbiol* 2024;4(3):100154. <https://doi.org/10.1016/j.engmic.2024.100154>.
- [7] Cong L, Ran FA, Cox D, Lin S, Barretto R, Habib N, Hsu PD, Wu X, Jiang W, Marraffini LA, Zhang F. Multiplex genome engineering using CRISPR/Cas systems. *Science* 2013;339(6121):819–23. <https://doi.org/10.1126/science.1231143>.
- [8] Li R, Li A, Zhang Y, Fu J. The emerging role of recombining in microbiology. *Eng Microbiol* 2023;3(3):100097. <https://doi.org/10.1016/j.engmic.2023.100097>.
- [9] Myrers JP, Zhang Y, Stewart AF. Techniques: recombinogenic engineering—new options for cloning and manipulating DNA. *Trends Biochem Sci* 2001;26(5):325–31. [https://doi.org/10.1016/s0968-0004\(00\)01757-6](https://doi.org/10.1016/s0968-0004(00)01757-6).
- [10] Wang H, Li Z, Jia R, Hou Y, Yin J, Bian X, Li A, Müller R, Stewart AF, Fu J, Zhang Y. RecET direct cloning and Red α β recombining of biosynthetic gene clusters, large operons or single genes for heterologous expression. *Nat Protoc* 2016;11(7):1175–90. <https://doi.org/10.1038/nprot.2016.054>.
- [11] Engler C, Kandzia R, Marillonnet S. A one pot, one step, precision cloning method with high throughput capability. *PLoS One* 2008;3(11):e3647. <https://doi.org/10.1371/journal.pone.0003647>.
- [12] Gibson DG, Young L, Chuang RY, Venter JC, Hutchison 3rd CA, Smith HO. Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat Methods* 2009;6(5):343–5. <https://doi.org/10.1038/nmeth.1318>.
- [13] Dey S, Fan C, Gotherl KV, Li J, Lin C, Liu L, Liu N, Nijenhuis MAD, Sacca B, Simmel FC, Yan H, Zhan P. DNA origami. *Nat Rev Methods Primers* 2021;1(1):13. <https://doi.org/10.1038/s43586-020-00009-8>.
- [14] Praetorius F, Kick B, Behler KL, Honemann MN, Weuster-Botz D, Dietz H. Biotechnological mass production of DNA origami. *Nature* 2017;552(7683):84–7. <https://doi.org/10.1038/nature24650>.
- [15] Church GM, Gao Y, Kosuri S. Next-generation digital information storage in DNA. *Science* 2012;337(6102):1628. <https://doi.org/10.1126/science.1226355>.
- [16] Yu M, Tang X, Li Z, Wang W, Wang S, Li M, Yu Q, Xie S, Zuo X, Chen C. High-throughput DNA synthesis for data storage. *Chem Soc Rev* 2024;53(9):4463–89. <https://doi.org/10.1039/D3CS00469D>.
- [17] Weng Z, Li J, Wu Y, Xiu X, Wang F, Zuo X, Song P, Fan C. Massively parallel homogeneous amplification of chip-scale DNA for DNA information storage (MPHAC-DIS). *Nat Commun* 2025;16(1):667. <https://doi.org/10.1038/s41467-025-55986-9>.
- [18] Wu Y, Li BZ, Zhao M, Mitchell LA, Xie ZX, Lin QH, Wang X, Xiao WH, Wang Y, Zhou X, Liu H, Li X, Ding MZ, Liu D, Zhang L, Liu BL, Wu XL, Li FF, Dong XT, Jia B, Zhang WZ, Jiang GZ, Liu Y, Bai X, Song TQ, Chen Y, Zhou SJ, Zhu RY, Gao F, Kuang Z, Wang X, Shen M, Yang K, Stracquadanio G, Richardson SM, Lin Y, Wang L, Walker R, Luo Y, Ma PS, Yang H, Cai Y, Dai J, Bader JS, Boeke JD, Yuan YJ. Bug mapping and fitness testing of chemically synthesized chromosome X. *Science* 2017;355(6329):eaaf4706. <https://doi.org/10.1126/science.aaf4706>.
- [19] Richardson SM, Mitchell LA, Stracquadanio G, Yang K, Dymond JS, DiCarlo JE, Lee D, Huang CLV, Chandrasegaran S, Cai Y, Boeke JD, Bader JS. Design of a synthetic yeast genome. *Science* 2017;355(6329):1040–4. <https://doi.org/10.1126/science.aaf4557>.

- [20] Hoose A, Vellacott R, Storch M, Freemont PS, Ryadnov MG. DNA synthesis technologies to close the gene writing gap. *Nat Rev Chem* 2023;7(3):144–61. <https://doi.org/10.1038/s41570-022-00456-9>.
- [21] Ma Y, Zhang Z, Jia B, Yuan Y. Automated high-throughput DNA synthesis and assembly. *Heliyon* 2024;10(6):e26967. <https://doi.org/10.1016/j.heliyon.2024.e26967>.
- [22] Yu W, Dai J, Ma Y. Recent development on DNA & genome synthesis. *Curr Opin Syst Biol* 2024;37:100490. <https://doi.org/10.1016/j.coisb.2023.100490>.
- [23] Beaucage SL, Caruthers MH. Deoxynucleoside phosphoramidites—A new class of key intermediates for deoxypolynucleotide synthesis. *Tetrahedron Lett* 1981;22(20):1859–62. [https://doi.org/10.1016/S0040-4039\(01\)90461-7](https://doi.org/10.1016/S0040-4039(01)90461-7).
- [24] Caruthers MH. The chemical synthesis of DNA/RNA: our gift to science. *J Biol Chem* 2013;288(2):1420–7. <https://doi.org/10.1074/jbc.X112.442855>.
- [25] Li H, Huang Y, Wei Z, Wang W, Yang Z, Liang Z, Li Z. An oligonucleotide synthesizer based on a microreactor chip and an inkjet printer. *Sci Rep* 2019;9(1):5058. <https://doi.org/10.1038/s41598-019-41519-0>.
- [26] Liu C, Zhuang W, Liu L, Chen Y, Yang C, Chen H, Yao Y, Sun X, Hu WW. Research progress in high-throughput DNA synthesis and its applications. *J Mater Chem B* 2025;13(27):7973–8004. <https://doi.org/10.1039/d5tb00869g>.
- [27] An R, Jia Y, Wan B, Zhang Y, Dong P, Li J, Liang X. Non-enzymatic depurination of nucleic acids: factors and mechanisms. *PLoS One* 2014;9(12):e115950. <https://doi.org/10.1371/journal.pone.0115950>.
- [28] LeProust EM, Peck BJ, Spirin K, McCuen HB, Moore B, Namsaraev E, Caruthers MH. Synthesis of high-quality libraries of long (150 mer) oligonucleotides by a novel depurination controlled process. *Nucleic Acids Res* 2010;38(8):2522–40. <https://doi.org/10.1093/nar/gkq163>.
- [29] Deng X, Li H, Song Y. Inkjet printing-based high-throughput DNA synthesis. *Giant* 2024;17:100222. <https://doi.org/10.1016/j.giant.2023.100222>.
- [30] Bollum FJ. Calf thymus polymerase. *J Biol Chem* 1960;235(8):2399–403. [https://doi.org/10.1016/S0021-9258\(18\)64634-4](https://doi.org/10.1016/S0021-9258(18)64634-4).
- [31] Desiderio SV, Yancopoulos GD, Paskind M, Thomas E, Boss MA, Landau N, Alt FW, Baltimore D. Insertion of N regions into heavy-chain genes is correlated with expression of terminal deoxytransferase in B cells. *Nature* 1984;311(5988):752–5. <https://doi.org/10.1038/311752a0>.
- [32] Jung D, Alt FW. Unraveling V(D)J recombination; insights into gene regulation. *Cell* 2004;116(2):299–311. [https://doi.org/10.1016/s0092-8674\(04\)00039-x](https://doi.org/10.1016/s0092-8674(04)00039-x).
- [33] Boubakour-Azzouz I, Bertrand P, Claes A, Lopez BS, Rougeon F. Terminal deoxynucleotidyl transferase requires KU80 and XRCC4 to promote N-addition at non-V(D)J chromosomal breaks in non-lymphoid cells. *Nucleic Acids Res* 2012;40(17):8381–91. <https://doi.org/10.1093/nar/gks585>.
- [34] Palluk S, Arlow DH, de Rond T, Barthel S, Kang JS, Bector R, Baghdassarian HM, Truong AN, Kim PW, Singh AK, Hillson NJ, Keasling JD. *De novo* DNA synthesis using polymerase-nucleotide conjugates. *Nat Biotechnol* 2018;36(7):645–50. <https://doi.org/10.1038/nbt.4173>.
- [35] Verardo D, Adelizzi B, Rodriguez-Pinzon DA, Moghaddam N, Thomée E, Loman T, Godron X, Horgan A. Multiplex enzymatic synthesis of DNA with single-base resolution. *Sci Adv* 2023;9(27):ead0263. <https://doi.org/10.1126/sciadv.adi0263>.
- [36] Zhang C, Subthain H, Guo F, Fang P, Zheng S, Shen M, Jiang X, Gao Z, Meng C, Li S, Du L. Terminal deoxynucleotidyl transferase: properties and applications. *Eng Microbiol* 2025;5(1):100179. <https://doi.org/10.1016/j.engmic.2024.100179>.
- [37] Ashley J, Potts IM, Olorunniji FJ. Applications of terminal deoxynucleotidyl transferase enzyme in biotechnology. *Chembiochem* 2023;24(5):e202200510. <https://doi.org/10.1002/cbic.202200510>.
- [38] Pichon M, Hollenstein M. Controlled enzymatic synthesis of oligonucleotides. *Commun Chem* 2024;7(1):138. <https://doi.org/10.1038/s42004-024-01216-0>.
- [39] Lu X, Li J, Li C, Lou Q, Peng K, Cai B, Liu Y, Yao Y, Lu L, Tian Z, Ma H, Wang W, Cheng J, Guo X, Jiang H, Ma Y. Enzymatic DNA synthesis by engineering terminal deoxynucleotidyl transferase. *ACS Catal* 2022;12(5):2988–97. <https://doi.org/10.1021/acscatal.1c04879>.
- [40] Hu L, Zhang Z, Li C, Han M, Hao M, Zhang X, Ahmed N, Luo J, Lu X, Sun J, Jiang H. High-throughput screening for oligonucleotide detection by ADE-OPI-MS. *Anal Chem* 2024;96(29):12040–8. <https://doi.org/10.1021/acs.analchem.4c02110>.
- [41] Forget SM, Krawczyk MJ, Knight AM, Ching C, Copeland RA, Mahmoodi N, Mayo MA, Nguyen J, Tan A, Miller M, Vroom J, Lutz S. Evolving a terminal deoxynucleotidyl transferase for commercial enzymatic DNA synthesis. *Nucleic Acids Res* 2025;53(4):gkaf115. <https://doi.org/10.1093/nar/gkaf115>.
- [42] Niu Y, Chen B, Zhang H, Zheng W, Wu J, Yang L, Yang M, Yu H. Computational design of a thermostable and highly active terminal deoxynucleotidyl transferase for synthesis of long *de novo* DNA molecules. *ACS Catal* 2025;15(9):7201–16. <https://doi.org/10.1021/acscatal.5c01571>.
- [43] Fessner ND. Enzymatic strategies for next-generation DNA synthesis: boosting efficiency and overcoming secondary structures. *ACS Catal* 2025;15:16106–14. <https://doi.org/10.1021/acscatal.5c05189>.
- [44] Gao N, Yu A, Yang W, Zhang X, Shen Y, Fu X. Enzymatic *de novo* oligonucleotide synthesis: emerging techniques and advancements. *Biotechnol Adv* 2025;82:108604. <https://doi.org/10.1016/j.biotechadv.2025.108604>.
- [45] Zhou H, Li H, Jia Z, Niu S, Song Y. Reaction pathways and technologies of *in vitro* DNA synthesis. *Cell Rep Phys Sci* 2025;6(8):102777. <https://doi.org/10.1016/j.xcrp.2025.102777>.
- [46] Haukanes B-I, Kvam C. Application of magnetic beads in bioassays. *Nat Biotechnol* 1993;11(1):60–3. <https://doi.org/10.1038/nbt0193-60>.
- [47] Schneider CA, Rasband WS, Eliceiri KW. NIH image to ImageJ: 25 years of image analysis. *Nat Methods* 2012;9(7):671–5. <https://doi.org/10.1038/nmeth.2089>.
- [48] Anderson RS, Bollum FJ, Beattie KL. Pyrophosphorylytic dismutation of oligodeoxy-nucleotides by terminal deoxynucleotidyltransferase. *Nucleic Acids Res* 1999;27(15):3190–6. <https://doi.org/10.1093/nar/27.15.3190>.
- [49] Shock DD, Freudenthal BD, Beard WA, Wilson SH. Modulating the DNA polymerase β reaction equilibrium to dissect the reverse reaction. *Nat Chem Biol* 2017;13(10):1074–80. <https://doi.org/10.1038/nchembio.2450>.
- [50] Li A-N, Shi K, Zeng B-B, Xu J-H, Yu H-L. Enhancing the expression of terminal deoxynucleotidyl transferases by N-terminal truncation. *Biotechnol J* 2024;19(9):2400226. <https://doi.org/10.1002/biot.202400226>.
- [51] Delarue M, Boulé JB, Lescar J, Expert-Bezançon N, Jourdan N, Sukumar N, Rougeon F, Papanicolaou C. Crystal structures of a template-independent DNA polymerase: murine terminal deoxynucleotidyltransferase. *EMBO J* 2002;21(3):427–39. <https://doi.org/10.1093/emboj/21.3.427>.
- [52] Schaudy E, Lietard J, Somoza MM. Sequence preference and initiator promiscuity for *de novo* DNA synthesis by terminal deoxynucleotidyl transferase. *ACS Synth Biol* 2021;10(7):1750–60. <https://doi.org/10.1021/acssynbio.1c00142>.
- [53] Ukladov EO, Tyugashev TE, Kuznetsov NA. Computational modeling study of the molecular basis of dNTP selectivity in human terminal deoxynucleotidyltransferase. *Biomolecules* 2024;14(8):961. <https://doi.org/10.3390/biom14080961>.
- [54] Gouge J, Rosario S, Romain F, Beguin P, Delarue M. Structures of intermediates along the catalytic cycle of terminal deoxynucleotidyltransferase: dynamical aspects of the two-metal ion mechanism. *J Mol Biol* 2013;425(22):4334–52. <https://doi.org/10.1016/j.jmb.2013.07.009>.