Published online 6 January 2023

Nucleic Acids Research, 2023, Vol. 51, No. 4 ***1501–1511***
https://doi.org/10.1093/nar/gkac1147

NAR Breakthrough Article

CRISPR-Cas9 recognition of enzymatically synthesized base-modified nucleic acids

Hui Yang[1,*], Elena Eremeeva[1,2], Mikhail Abramov[1], Maarten Jacquemyn[3], Elisabetta Groaz[1,4], Dirk Daelemans[3] and Piet Herdewijn[1,*]

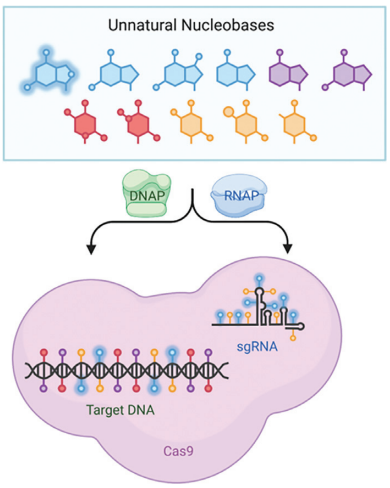
[1]KU Leuven, Department of Pharmaceutical and Pharmacological Sciences, Rega Institute for Medical Research, Medicinal Chemistry, Herestraat 49, Box 1041, 3000 Leuven, Belgium, [2]Queensland University of Technology, Centre for Agriculture and the Bioeconomy, Molecular Engineering Group, George Street 2, 4000 Brisbane, Queensland, Australia, [3]KU Leuven, Department of Microbiology, Immunology and Transplantation, Rega Institute for Medical Research, Laboratory of Virology and Chemotherapy, Herestraat 49, Box 1043, 3000 Leuven, Belgium and [4]University of Padova, Department of Pharmaceutical and Pharmacological Sciences, Via Marzolo 5, 35131 Padova, Italy

Received October 13, 2022; Revised November 09, 2022; Editorial Decision November 10, 2022; Accepted November 17, 2022

ABSTRACT

An enzymatic method has been successfully established enabling the generation of partially base-modified RNA (previously named RZA) constructs, in which all G residues were replaced by isomorphic fluorescent thienoguanosine (thG) analogs, as well as fully modified RZA featuring thG, 5-bromocytosine, 7-deazaadenine and 5-chlorouracil. The transcriptional efficiency of emissive fully modified RZA was found to benefit from the use of various T7 RNA polymerase variants. Moreover, dthG could be incorporated into PCR products by Taq DNA polymerase together with the other three base-modified nucleotides. Notably, the obtained RNA products containing thG as well as thG together with 5-bromocytosine could function as effectively as natural sgRNAs in an *in vitro* CRISPR-Cas9 cleavage assay. N^1-Methylpseudouridine was also demonstrated to be a faithful non-canonical substitute of uridine to direct Cas9 nuclease cleavage when incorporated in sgRNA. The Cas9 inactivation by 7-deazapurines indicated the importance of the 7-nitrogen atom of purines in both sgRNA and PAM site for achieving efficient Cas9 cleavage. Additional aspects of this study are discussed in relation to the significance of sgRNA–protein and PAM—protein interactions that were not highlighted by the Cas9–sgRNA–DNA complex crystal structure. These findings could expand the impact and therapeutic value of CRISPR-Cas9 and other RNA-based technologies.

GRAPHICAL ABSTRACT



INTRODUCTION

RNA plays a central role in all known life forms, as it serves not only as the intermediary of protein synthesis from DNA but can also act as a catalyst, carry viral genetic information and regulate numerous cellular processes (1). Recently, mRNA has attracted considerable attention as a

*To whom correspondence should be addressed. Tel: +32 16 32 26 57; Fax: +32 16 33 73 40; Email: piet.herdewijn@kuleuven.be
Correspondence may also be addressed to Hui Yang. Tel: +32 48 76 98 265; Email: hui.yang@kuleuven.be

versatile therapeutic tool in vaccine development and protein replacement therapy (2,3). In addition, small regulatory RNAs have entered clinical trials, and some have gained FDA approval as RNA-based therapeutics (4). Furthermore, programmable single guide RNAs (sgRNAs) are key components of CRISPR-Cas9 gene editing systems, which have revolutionized the current approach to genome editing and thus biomedical research (5).

However, for all *in vivo* RNA applications, the introduction of chemical modifications is essential to decrease immunogenicity, improve stability, and ultimately enhance efficacy and specificity (6–8). For instance, one prominent RNA modification is N^1-methyl-pseudouridine (m1Ψ, Figure 1), a naturally occurring uridine isomer that is incorporated in COVID-19 mRNA vaccines (9) due to its ability to facilitate translation, increase mRNA half-time as well as reduce immunogenicity *in vivo* (10). The chemical synthesis of modified RNA is challenging and currently limited to 100-nt (11). Moreover, the scope of RNA modifications accessible by solid-phase synthesis is woefully underrepresented. In contrast, *in vitro* transcription reactions using a DNA-dependent RNA polymerase from bacteriophage T7 (T7 RNAP) have been established as a powerful enzymatic approach for generating a diversity of modified RNAs in a cost-effective and efficient way, simply by placing the target sequence immediately downstream of the specific T7 RNA promoter.

Recently, our research group succeeded in generating a fully substituted base-modified RNA transcript using wild-type T7 RNAP, which was composed of 7-deazagunanine, 5-bromocytosine, 7-deazaadenine and 5-chlorouracil (Z: C^7G, BrC, C^7A, and ClU, respectively, Figure 1), denoted as 'RZA' (12). Although an 87-nt full-length RZA product could be synthesized, the introduction of C^7GTP significantly decreased the transcriptional efficiency. Replacing G with chemical analogs that can efficiently initiate transcription and succeed in RNA synthesis is typically not trivial, as wild-type T7 RNAP is known to be sensitive to the sequence composition, especially the presence of three consecutive Gs (-GGG-) as the first three nucleotides of the transcripts (13,14). In addition, the incorporation of chemical modifications is restricted to positions remote from the promoter to avoid interference with the initiation phase (15).

To overcome these shortcomings, we sought for an alternative G surrogate that could more effectively mimic natural GTP in RZA synthesis, while being compatible with the other three modified rNTPs (rZTPs) mentioned above. We examined thienoguanosine triphosphate (^{th}GTP, Figure 1), which is known to bear a high resemblance to natural GTP and act as a substrate for T7 RNAP during both initiation and elongation phases (16,17), while possessing unique photophysical properties including visible emission, high quantum yield, and responsiveness to environmental perturbations (18–21). While short ^{th}G-modified RNA oligonucleotides have been previously produced by solid-phase synthesis for various biophysical applications (22–24), only a few ^{th}G modifications were incorporated at the initial site or specific positions. Therefore, we sought an enzymatic method to generate fully ^{th}G-substituted RZA for nucleic acid technologies.

We anticipated that the synthesis of fully modified RZA containing ^{th}G in combination with BrC + C^7A + ClU could benefit from the use of T7 RNAP variants with improved substrate tolerance towards unnatural nucleotides (25–28). In fact, it was previously demonstrated that Y639F/H784A and P266L/Y639F/H784A variants displayed reduced misincorporation rates of natural ribonucleotides in transcripts in the presence of base-modified substrates (29).

The enzymatic synthesis of RZA is particularly relevant to the development of the CRISPR-Cas9 system into an effective therapeutic modality (30,31), with early examples of sugar and backbone modified guide RNAs establishing the advantages of chemically modified nucleotides (6,32–36). Nonetheless, the influence of modified nucleobase on the function of the different components of the CRISPR system has not yet been explored. Modified nucleobases may result in altered guide-target base-pairing as well as guide–Cas9 interaction. These modifications could vary the free energy barrier between the formation and association of the gRNP : target, as well as the ground-state complex and its transition state, which are important for improving the specificity of Cas9-guided cleavage (37). Additionally, they could affect the dynamics of hybridization of a guide sequence with off-target sites, further impacting specificity.

The protospacer adjacent motif (PAM) site is closely involved in target DNA binding, strand unwinding, and gRNA–DNA hybridization. Previous reports described the recognition of noncanonical PAM sequences (5′-NGNG-3′) using engineered Cas9 variants, demonstrating that the PAM site is not a conserved motive (38,39). Investigating structurally modified PAM sites could also contribute to a better understanding of the mechanism of Cas9 recognition and DNA cleavage, along with a potential enhancement of CRISPR-Cas9 specificity in biological and biotechnology research.

In this study, we evaluate the feasibility and impact of replacing natural counterparts in sgRNAs with different noncanonical nucleobases, including ^{th}G, 2,6-diaminopurine (DAP), N^1-methylpseudouracil (m1Ψ) and 7-deazaadenine (C^7A) (Figure 1). We also examine the introduction of 7-deaza purines in the target DNA in order to elucidate the DNA targeting mechanism, in particular recognition of the PAM region by the Cas9 endonuclease.

MATERIALS AND METHODS

Materials

Deoxyribonucleoside triphosphates (dNTPs), Taq DNA polymerase with 10 × ThermoPol buffer, $MgSO_4$, GC enhancer, Q5 High-Fidelity DNA polymerase with 5 × Q5 reaction buffer, RNase inhibitor, Cas9 Nuclease (*S. pyogenes*) with NEBuffer r3.1, Proteinase K, Monarch DNA Gel Extraction Kit, and the Monarch Plasmid DNA Miniprep Kit were purchased from New England Biolabs. The SYBR Gold Nucleic Acid Gel Stain, Turbo DNase I, guanosine 5′-monophosphate (GMP), T4 RNA ligase with 10 × reaction buffer, and CloneJet PCR cloning kit were obtained from ThermoFisher Scientific. T7 RNA polymerases were purchased either from New England Biolabs (T7 RNA polymerase and HiScribe T7 High Yield RNA Synthesis

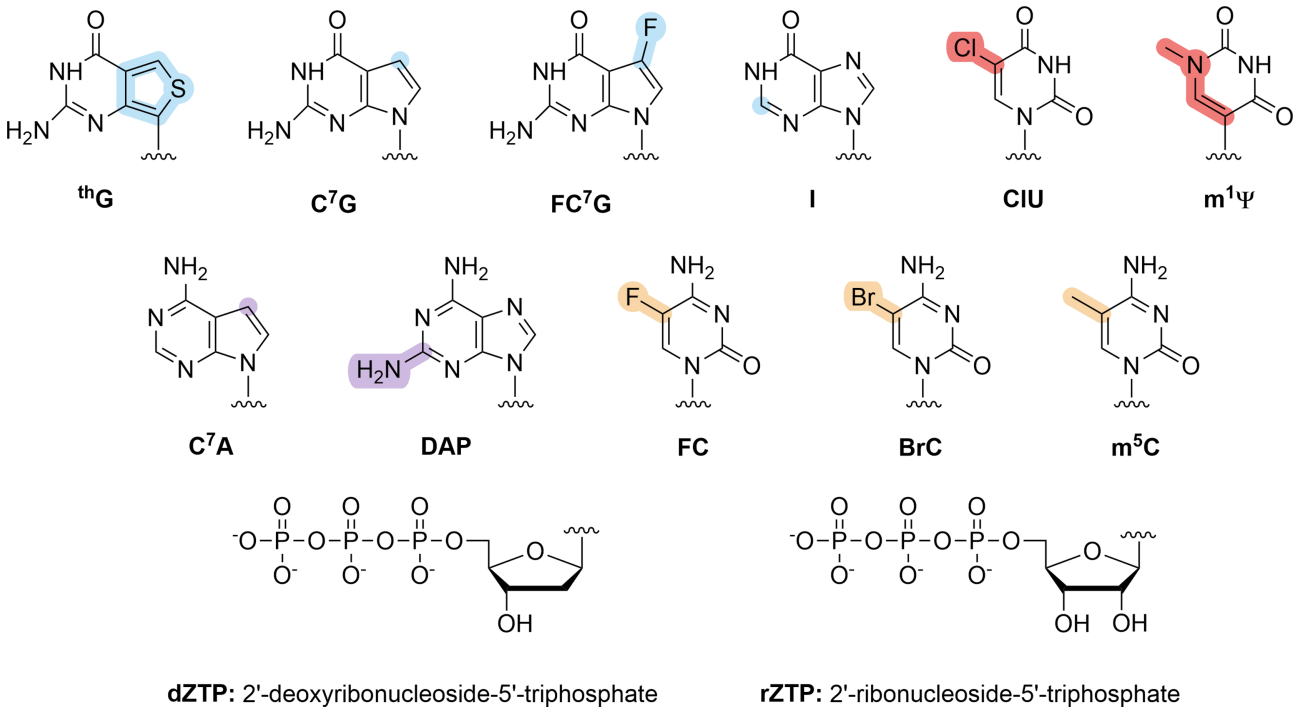


Figure 1. Chemical structures of non-canonical nucleobases and their 2′-deoxyribonucleoside and ribonucleoside 5′-triphosphates (dZTPs and rZTPs, respectively) used in RZA and DZA synthesis: thienoguanine (thG), 7-deazaguanine (C^7G), 7-fluoro-7deazaguanine (FC^7G), inosine (I), 5-chlorouracil (ClU), N^1-methylpseudouracil (m1Ψ), 7-deazaadenine (C^7A), 2,6-diaminopurine (DAP), 5-fluorocytosine (FC), 5-bromocytosine (BrC), 5-methylcytosine (m^5C) 2′-(deoxy) ribo-5′-triphosphates.

Kit) or ThermoFisher Scientific (T7 RNA polymerase with 5× Transcription Buffer). pCp-Cy5 dye was purchased from Jena Bioscience. Gel filtration columns, illustra NAP-25, were acquired from Cytiva. All unmodified oligodeoxyribonucleotides were purchased from Integrated DNA Technologies (IDT, Leuven). All other chemicals were obtained either from VWR or Sigma-Aldrich.

Methods

The chemical synthesis of base-modified ribo- and deoxyribonucleoside triphosphates (rZTP and dZTP, respectively, Supplementary Figure S1), and C^7G-modified oligonucleotides (Supplementary Table S1), as well as preparation of mutant T7 RNAP variants are described in detail in the Supplementary Data section.

PCR amplification of DNA/DZA-1 and 2 in the presence of dZTPs. Either 100 pM of 87-mer ssDNA (Template-1) or 0.2 ng of pET-3a plasmid (Template-2) was PCR amplified in 20 μl of 1 × ThermoPol buffer with 0.2 μM of the corresponding primers (Supplementary Table S2), 200 μM of either natural dNTPs (positive control; PC) or different combinations of modified dZTPs, and 25 U/ml of Taq DNA polymerase. If specified, different additives (0.3 ng of pET-3a plasmid, 2 mM Mg^{2+}, 2% DMSO, and 10% GC enhancer) were used for the PCR amplification of Template-2. The samples were first denatured at 95°C for 60 s, followed by 30 cycles of repetitive denaturation at 95°C for 30 s, annealing at 54°C (50°C for Template-2) for 15 s, extension at 68°C for 15 s, and final extension at 68°C for 10 min. The final PCR products were quenched with a loading buffer (90% formamide, 50 mM EDTA, 0.05% bromophenol blue) and analyzed by 15% denaturing PAGE. The gels of the PCR products obtained using Template-1 were scanned using a Typhoon 9500 imaging system under Cy2 and Cy3 channels and further analyzed by the ImageQuant TL v8.1 Software (both from GE Healthcare Life Science). The relative PCR amplification efficiencies were estimated using the positive control (unmodified PCR product) as 100%. The gels of the PCR products using Template-2 were first stained with 1 × SYBR Gold Nucleic Acid Gel Stain and then analyzed using the above-described method.

Large-scale synthesis of 113-bp DNA-3 used for sgRZA transcription. The synthesis of the 113-bp dsDNA fragment was performed by PCR amplification of 1 nM of 113-nt Template-3 in 50 μl of 1 × Q5 reaction buffer with 0.5 μM of the corresponding primers (Supplementary Table S2), 200 μM of natural dNTPs, and 20 U/ml of Q5 High-Fidelity DNA polymerase. The samples were first denatured at 95°C for 30 s, followed by 30 cycles of repetitive denaturation at 95°C for 10 s, annealing at 70°C for 15 s, extension at 72°C for 15 s and final extension at 72°C for 2 min. The resulting PCR products were pooled and purified by 15% denaturing PAGE according to a standard procedure, followed by elution using 0.3 M sodium acetate (pH 7.2) at 37°C under overnight shaking. The supernatants were collected and desalted on a NAP-25 column according to the manufacturer's recommendations.

In vitro transcription of DNA-1 to generate 87-nt RZA. *In vitro* transcription reactions were carried out in 20 μl of 1 × Transcription Buffer containing 200 nM DNA-1, 5 mM of rNTP (PC) or modified rZTPs, 60 U of T7 RNA polymerase, 0.5 U of RNase inhibitor, 8 mM of Mg^{2+}, and, if specified, different additives (40 U T7 RNA polymerase, 30 mM GMP, 10 mM Mg^{2+}, 10% DMSO). The reaction mixtures were incubated at 37°C for 6 h, followed by digestion using 2 U DNase I at 37°C for 30 min. The final products were quenched by adding EDTA at the final concentration of 15 mM and denaturing at 75°C for 10 min. The resulting transcripts were analyzed by 8% denaturing PAGE and visualized by staining with 1 × SYBR Gold Nucleic Acid Gel Stain. The gels were scanned using Typhoon 9500 imaging system (Cy2 channel) and quantified by the ImageQuant TL v8.1 Software. The relative transcription efficiencies of the dsDZA fragments were calculated by comparison to the transcription efficiency of the unmodified DNA fragment (PC).

In vitro transcription of DZA-1 to generate RNA or RZA. Standard *in vitro* transcription reactions were carried out in 20 μl of 1 × Transcription Buffer containing 200 nM of either natural DNA-1 (PC) or modified 104-bp DZA-1 fragments, 2 mM of natural rNTPs (or 5 mM rZTP for RZA synthesis), 60 U of T7 RNA polymerase, 0.5 U of RNase Inhibitor, and 10 mM of Mg^{2+}. The reaction mixtures were incubated at 37°C for 2 h (or 6 h for RZA synthesis), followed by digestion using 2 U DNase I at 37°C for 30 min. The final products were quenched by adding EDTA at a final concentration of 15 mM and heating at 75°C for 10 min. The resulting transcripts were analyzed as described above.

Sequencing analysis. The 104-bp DNA/DZA-1 fragments after PCR, *in vitro* transcription (with modified rZTPs), and RT-PCR assays, or 97-bp DNA/DZA-2 fragments, directly after PCR amplification with dZTPs, were cloned into a pJET1.2 vector using a CloneJet PCR cloning kit, according to the manufacturer's recommendations. Several colonies were collected from each group, and plasmids from each colony were isolated using the Monarch Plasmid DNA Miniprep Kit, followed by sequencing using pJET-fw and pJET-rv primers (Supplementary Table S2). A fully modified DZA fragment (thG + BrC + C^{7}A + ClU) was amplified using natural dNTPs prior to sequencing, followed by cloning and transformation.

sgRZA production and purification. *In vitro* transcription reactions using 1 μg of 113-bp dsDNA-3 were performed in the presence of different variants of T7 RNA polymerase, including commercial enzymes from New England Biolabs (NEB and HiScribe T7 RNAP) and ThermoFisher Scientific (TF T7 RNAP) as well as home-made mutants (with NEB buffer) in 20 μl of their corresponding buffers. Each reaction contained 3 mM of either rNTPs or modified rZTPs, 100 U of certain T7 RNA polymerase variants, 0.5 U of RNase inhibitor, 15 mM of Mg^{2+}, and 30 mM GMP. The reaction mixtures were incubated at 37°C for 12 h, followed by digestion using 2 U DNase I at 37°C for 1 h. The resulting transcripts were analyzed as described above by 15% denaturing PAGE gel. The resulting sgRZA transcripts were pooled and purified by 15% denaturing PAGE according to a standard procedure, followed by elution using 0.3 M sodium acetate (pH 5.4) at 37°C under overnight shaking. The supernatants were collected and desalted on a NAP-25 column, according to the manufacturer's recommendations. The concentration of each sample was measured using a NanoDrop spectrophotometer (ThermoFisher Scientific).

Post-transcriptional labeling. Post-transcriptional labeling of both RNA and RZA sequences was performed in 20 μl of 1 × T4 RNA Ligase Reaction Buffer containing 12 μl of transcript solutions, 50 μM of Cy5-labeled cytidine-5′-phosphate-3′-(6-aminohexyl) phosphate (pCp-Cy5), 0.1 mg μl^{-1} of BSA and 20 U of T4 RNA ligase. The reactions were incubated at 16°C for 16 h, followed by inactivation at 70°C for 10 min. The samples (15 μl) were analyzed by 8% denaturing PAGE. The gels were scanned using a Typhoon 9500 imaging system (Cy5 channel) and quantified by the ImageQuant TL v8.1 Software.

CRISPR target DZA synthesis. Long and short target DNAs (622- and 90-bp, respectively, Supplementary Table S3) were PCR amplified using 0.5 ng of mCherry plasmid (Supplementary Figure S2A) with 0.2 μM of the corresponding primers (Supplementary Table S2) in 50 μl of 1 × ThermoPol buffer with 200 μM of either natural dNTPs or different combinations of modified dZTPs and 50 U/ml of Taq DNA polymerase. The target DZA samples were first denatured at 95°C for 30 s, followed by 30 cycles of repetitive denaturation at 95°C for 15 s, annealing at 54°C for 15 s (56°C for short target DZA), extension at 68°C for 30 s (15 s for short target DZA) and final extension at 68°C for 5 min. Several PCR reactions of long target DZA samples were collected and purified by 1% agarose gel, followed by extraction using the Monarch DNA Gel Extraction Kit, according to the manufacturer's recommendations. Short DZA were purified as described above using 15% denaturing PAGE gel.

Evaluation of sgRZAs in the CRISPR-Cas9 in vitro cleavage assay. *In vitro* cleavage assays were performed in a 9 μl of 1 × NEB r3.1 buffer containing 80 ng of various modified sgRZAs and 200 nM of Cas9 nuclease. The reaction mixtures were incubated at 25°C for 10 min (longer times were used when specified, Supplementary Figure S3), 1 μl of long target DNA (100 ng, Supplementary Table S3) was added to each reaction followed by incubation at 37°C for 15 min (longer times were used when specified, Supplementary Figure S3), then, 2 μl of proteinase K was added to each sample followed by incubation at RT for 10 min and inactivation at 65°C for 10 min. The resulting mixtures were analyzed by 1.5% agarose gel containing 1 × SYBR Gold Nucleic Acid Gel Stain. The gels were scanned using a Typhoon 9500 imaging system (Cy2 channel) and quantified by the ImageQuant TL v8.1 Software. The efficiency of the cleavage reaction was estimated using the band intensity of the cleaved product to a total band intensity for the sample (i.e., cut/cut + uncut). The relative cleavage efficiencies of the sgRZAs were normalized to those of the unmodified sgRNA.

CRISPR-Cas9 in vitro cleavage assay with target DZA. *In vitro* cleavage assays were performed in a 10 μl reaction. The reaction mixtures containing 1 μl of NEB r3.1 buffer, 80 ng of sgRNA and 200 nM of Cas9 nuclease were first incubated at 25°C for 10 min, then 100 ng of different target DZAs were added and incubated at 37°C. After 15 min, 2 μl of proteinase K was added to each sample and incubated at RT for 10 min, followed by inactivation at 65°C for 10 min. The resulting mixtures of long and short target DZAs were analyzed by 1.5% agarose gel and 15% denaturing PAGE gel as described above.

Emission spectra of thG-RZA and thG-DZA. Fluorescence spectra were recorded at 25°C on a CLARIOstar Plus Multi-mode Plate Reader (BMG LABTECH). The thG-sgRZA/DZA samples in 50 mM Tris pH 7.5 buffer were excited at 350 nm wavelength in a 96-well black flat-bottom plate, and the fluorescence intensity was recorded within the 380–650 nm range, with a slit width for excitation and emission of 10 nm, resolution of 1 nm and integration time of 0.5 s. The emission spectra were blanked against 50 mM Tris pH 7.5 buffer.

RESULTS

Synthesis of RZA with intrinsic fluorescence

To examine the efficiency of synthesis of modified RNA containing thG as a fluorescent surrogate of all natural G residues, a series of *in vitro* transcription experiments were performed (Figure 2A), using a 104-bp dsDNA sequence containing the T7 promoter as template in the *in vitro* transcription with wild-type T7 RNA polymerase (wtT7 RNAP) and thGTP as substrate (Supplementary Figure S4) (12,40). The reaction using thGTP natural A-, C-, and UTP yielded 19% full-length thG-RZA transcript (Figure 2B), comparable to synthesis of a C^{7}G-RZA transcript (16%) previously reported by our group (12). When A-, C-, and UTP were substituted with base-modified nucleotides BrCTP, C^{7}ATP, and ClUTP (Figure 1), the corresponding transcriptional efficiency was not significantly affected (yielding 14% full-length RZA) (Figure 2B), which was three times higher than that previously described for the C^{7}G + BrC + C^{7}A + ClU transcript (5%) (12). In contrast, an alternative combination, thG + FC + C^{7}A + ClU resulted in no RZA product (Figure 2B).

To further improve the yield of fully modified RZA composed of thG + BrC + C^{7}A + ClU, additives that previously proved to be beneficial for improving fully modified RZA synthesis were included in the transcription reactions (Figure 2C) (12,41). The addition of either 10 mM Mg^{2+} or 10% DMSO led to an 8-fold enhanced transcriptional yield in comparison to the reaction without additives, while the addition of rGMP or a higher concentration of wtT7 RNAP led only to a slight (2- to 3-fold) improvement of the reaction efficiency.

To verify whether thG affected the accuracy of the *in vitro* transcription, fully modified thG + BrC + C^{7}A + ClU-RZA transcripts were analyzed by sequencing after performing RT-PCR. It was found that thG + BrC + C^{7}A + ClU-RZA was synthesized by the wtT7 RNAP and converted by reverse transcriptase M-MuLV and Taq DNA polymerase with the same accuracy as the natural RNA counterpart (Supplementary Table S4A and Figure S5A).

Encouraged by that the emissive thGTP substrate could substitute GTP in RNA synthesis by wtT7 RNAP, we subsequently attempted the enzymatic synthesis of functional RZA: a single guide RNA (sgRNA) for further evaluation in a CRISPR-Cas9 gene cleavage assay. To this end, a 113-bp dsDNA was used as template to generate a 96-nt 'sgRZA' containing 27 Gs (cf 17 Gs in the 87-nt RZA) (Supplementary Table S3 and Figure S6). The synthesis of this longer sgRZA proved to be more challenging, with the optimal conditions described above (10 mM Mg^{2+} and 30 mM rGMP additives) yielding no, or substantially reduced, full-length RZA transcript (6, 0, or 8% for C^{7}G + BrC + C^{7}A + ClU, I + BrC + C^{7}A + ClU or thG + BrC + C^{7}A + ClU, respectively) (Supplementary Figure S7A). Among various factors, the reduced yield of 96-nt sgRZA compared with 87-nt RZA transcripts might be associated with a strong wtT7 RNAP preference for the initiating sequence composition, specifically -GGG- as the three starting nucleotides, where -GGA- was used here (13,14).

To overcome this limitation in the sgRZA synthesis, we examined several T7 RNAP variants that were demonstrated to exert increased accuracy and incorporation efficiency with modified rNTP substrates (25,28). To this end, four T7 RNAP mutants, i.e., P266L, P266L/H784A, P266L/Y639F, and P266L/Y639F/H784A (L, LA, FL, and FAL, respectively; Supplementary Figure S2B and Table S5), were engineered by site-directed mutagenesis and expressed in *Escherichia coli*. These T7 RNAP variants were compared with commercially available enzymes acquired from different sources, using *in vitro* transcription assays with four modified rZTP substrates (thG + BrC + C^{7}A + ClU). All four mutant T7 RNAPs (L, LA, FL, and FAL) showed increased efficiency of fully modified sgRZA synthesis compared to commercial enzymes (Figure 2D), to different extents; the T7 RNAP mutants L and FL enhanced the transcription by 3- and 2-fold compared to TF, yielding 31 and 22% full-length transcripts, respectively. In contrast, variants containing the H784A mutation, such as LA and FAL, exhibited negligible improvement compared to the TF polymerase (16 and 11%, respectively, versus 10%). The most profound effect on sgRZA transcription was observed with the HiScribe T7 High Yield RNA Synthesis Kit (with proprietary composition), which produced 40% full-length yield (Figure 2D).

To verify whether the resulting thG-RZA transcripts retained the fluorescence properties of the individual thG nucleotides, the fluorescence spectra of different sgRZAs were recorded at 350 nm excitation (Supplementary Figure S8A). The 96-nt sgRZA containing thG + C + A + U exhibited a high quantum fluorescence performance, suggesting potential utility for monitoring and probing enzymatic processes. Although the introduction of additional base-modifications (BrC or BrC + C^{7}A + ClU) somewhat quenched the fluorescence intensity, the fully modified sgRZA nevertheless exhibited strong emission, as further corroborated under 365 nm UV irradiation (Supplementary Figure S8B).

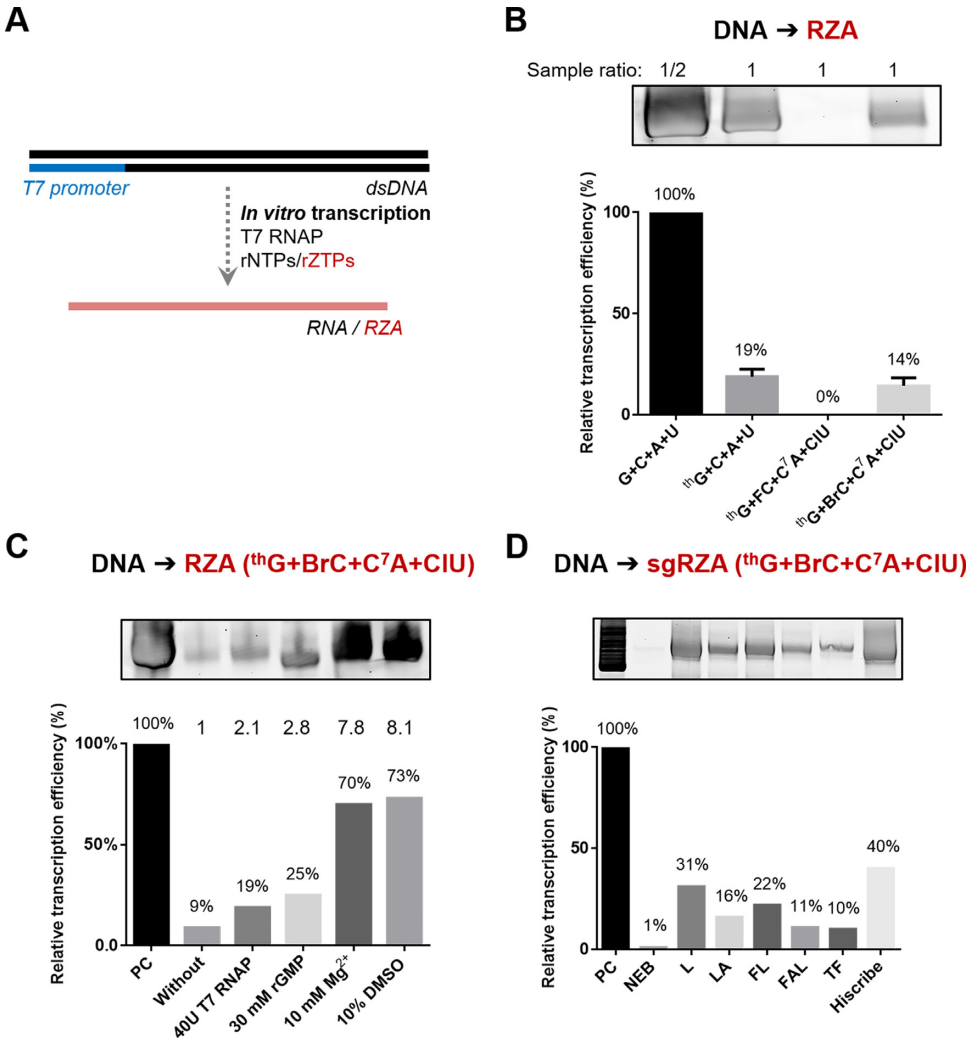


Figure 2. Enzymatic synthesis of thG-RZA fragments. (**A**) General experimental scheme of thG-RZA *in vitro* transcription. (**B**) Relative efficiencies of *in vitro* transcription using thGTP with or without the three other modified nucleotides. (**C**) Optimization of thG + BrC + C^{7}A + ClU RZA *in vitro* transcription. Transformational efficiency [i.e., yield of full-length (%)] relative to positive control reaction with all four natural substrates (PC) compared to reactions with thG + BrC + C^{7}A + ClU RZA without any additives or with an additional 40U of T7 RNAP, 30 mM rGTP, 10 mM Mg^{2+} and 10% DMSO. (**D**) Relative efficiency of thG + BrC + C^{7}A + ClU sgRZA synthesis using T7 RNAP variants. PC indicates the positive control with all natural rNTPs and T7 RNAP from New England Biolabs. The sgRZA enzymatic synthesis with T7 RNAP from New England Biolabs, T7 RNAP mutants (L, LA, FL, and FAL), ThermoFisher Scientific, or HiScribe T7 High Yield RNA Synthesis Kit are indicated as NEB, L, LA, FL, FAL, TF, and Hiscribe, respectively.

thG-containing DZA to RZA information exchange

We have previously demonstrated that genetic information can be transferred within a simulated circuit that uses non-canonical nucleotides (12), but the capacity of thG-containing DZA↔RZA to function as a synthetic genetic cassette was unknown. To explore this, we synthesized thG-substituted DZA genetic templates by PCR amplification, using dthGTP as a substrate for Taq DNA polymerase (Figure 3). A full-length PCR product was obtained using dthGTP and the other three natural dNTPs with 6% yield relative to a natural control DNA product (Figure 3B). The amplification efficiency was slightly improved when dthGTP was paired with 2′-deoxyribonucleotides bearing either FC or BrC, resulting in 9 and 11% relative yield, respectively. In contrast, when dCTP was substituted with the 2′-deoxyribonucleotide of m^{5}C, no PCR product was formed under the same reaction conditions. This was somewhat surprising as m^{5}C is known to be a naturally occurring epigenetic modification, although consistent with previous observations that some modified base pair combinations can synergistically either improve or inhibit PCR amplification efficiency compared to single modifications (42,43). This finding also highlighted that, among all tested combinations, BrC was the most favourable pairing with thG in both *in vitro* transcription and PCR amplification experiments. Interestingly, a fully modified dsDZA containing thG + BrC + C^{7}A + ClU was obtained without loss in amplification efficiency, producing the same yield as the thG + BrC + A + T-DZA fragment (Figure 3B). Although thG-DZA PCR amplification was clearly feasible, the overall amplification efficiency remained rela-

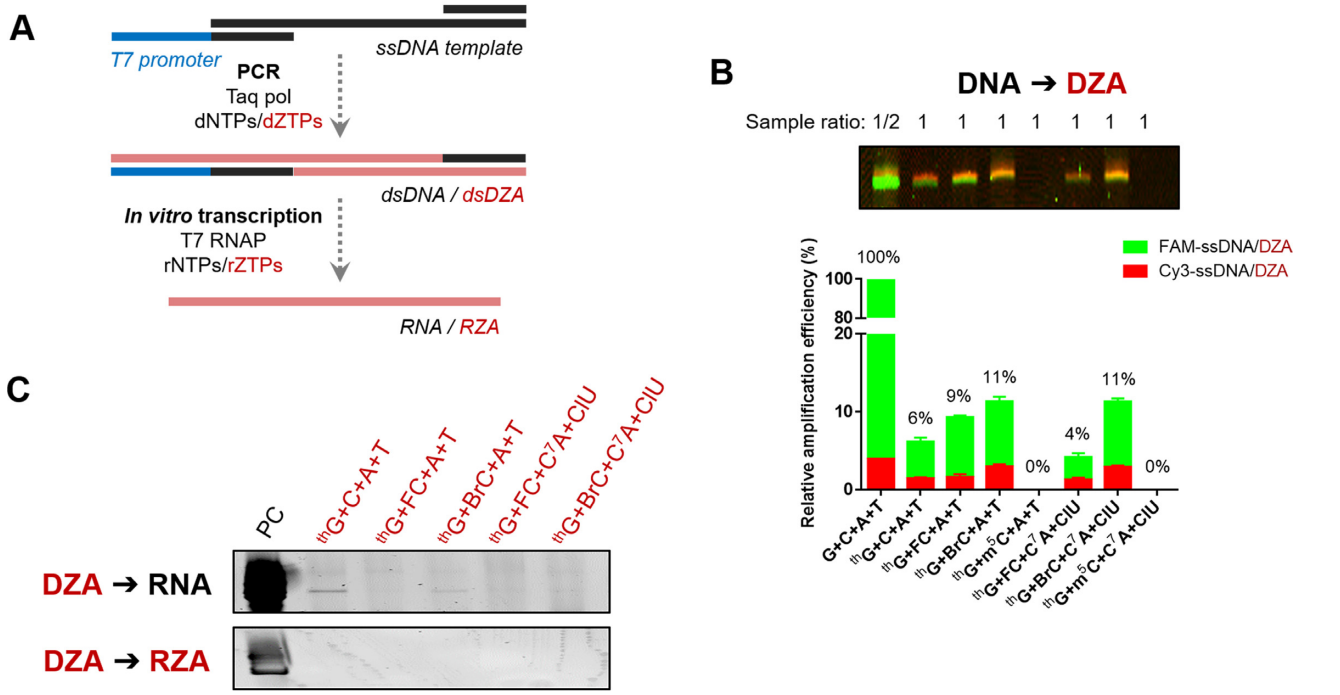


Figure 3. *In vitro* information transfer of thG-containing fragments. (**A**) General experimental scheme of *in vitro* PCR amplification and transcription studies. (**B**) Relative PCR amplification efficiencies with various compositions of dZTPs compared to natural product formation. The relative yield of the dsDZA products are shown as the total yield from Cy3- and FAM-labeled ssDZA fragments. (**C**) *In vitro* transcription of either natural RNA or thG-containing RZAs using the corresponding DNA or DZA templates. PC refers to the positive control, reactions with natural DNA templates and rNTPs.

tively low compared to C^{7}G + FC + C^{7}A + ClU and C^{7}G + BrC + C^{7}A + ClU (which yielded 96 and 77% full-length product relative to control, respectively) (12). To determine whether the low PCR efficiency observed in the case of dthGTP was sequence-dependent, PCR amplification was carried out using different DNA templates. However, yields were found to be substantially reduced, even when various additives were included (Supplementary Figure S7C).

Next, the amplification accuracy of the thG-DZA fragments was analyzed by sequencing (Supplementary Table S4B and Figure S5B). The presence of thG in the DNA sequence increased the mutation frequency compared with natural DNA (0.34% versus 0%). Interestingly, the inclusion of dBrC instead of dC in the PCR abolished the infidelity of the Taq DNA polymerase observed with a single thG modification. The presence of additional substitutions in the DZA sequence, such as thG + BrC + C^{7}A + ClU, restored base pairing errors, although with a 2-fold lower frequency than thG-modified DZA (0.17% versus 0.34%). This observation might explain the differences in amplification yields of thG, thG + BrC, and fully modified thG + BrC + C^{7}A + ClU DZAs, as the error rate of PCR with a single dthGTP substrate could be improved by the 'stabilizing' effect of the dBrC modification, leading to increased PCR yields (44).

In order to complete a circuit of genetic information transfer from thG-DZA to thG-RZA, thG-DZA templates were used in the *in vitro* transcription assays with either natural rNTPs or modified rZTPs (Figure 3C). thG-containing DZAs were poorly recognized as templates by wtT7 RNAP, with only trace amounts of RNA transcribed using the thG and thG + BrC DZAs, and no detectable RNA products with thG + FC or the fully modified DZA templates, using rNTP (Figure 3C). Similarly, no RZA transcripts were produced from the corresponding DZA templates using rZTP substrates. Despite the ability of thG to perfectly mimic G within the nucleic acid structure (24), our results demonstrated that thG-DZA templates inhibited *in vitro* transcription catalyzed by wtT7 RNAP. Although the efficiency of thG-DZA synthesis was poor, the resulting fragments nevertheless displayed high fluorescence emission, demonstrating the potential utility of thG-DNA templates for biotechnology research (Supplementary Figure S8C). Further optimization, such as the use of mutant polymerases and additives, is still needed to render thG suitable for genetic information storage and transfer.

To the best of our knowledge, the successful enzymatic synthesis of thG-modified DNA fragments in which the emissive thG fully substitutes all G residues has not yet been reported. The enzymatic production of such fluorescent fragments would open new avenues for developing diverse RNA and DNA probes to study nucleic acid–nucleic acid and nucleic acid–protein interactions of interest *in vitro* and *in vivo* (24,45). The presence of multiple thG residues along with other modifications within an oligonucleotide sequence can potentially improve their sensitivity when used as probes in living cell assays while providing resistance to restriction enzyme hydrolysis and suppression of immune responses (42,43,46).

Evaluation of modified nucleotides in a CRISPR-Cas9 *in vitro* cleavage assay

Previously, fluorescent thG-containing RNA oligonucleotides have been shown to function as probes in real-time assays for monitoring translation-related events (47–49). Here, we sought to exploit the unique photophysical properties of thG, such as visible light emission, high quantum yield and long excited state lifetime, in the context of sgRNA in the CRISPR-Cas9 gene-editing system. sgRNA guides the Cas9 complex towards a specific region on a target dsDNA, fluorescence could therefore enable imaging of the gene-editing process.

To evaluate the impact of the thG substitution on the sgRNA functionality, different thG-containing sgRZAs, with or without other modifications, were assayed for their ability to direct *in vitro* CRISPR-Cas9 cleavage (Figure 4A) of a fragment of the mCherry gene (Supplementary Table S3). sgRZAs composed of thG + C + A + U and thG + BrC + A + U were indeed able to guide the Cas9 protein to cleave the target DNA fragment with similar efficiencies to that of natural sgRNA (Figure 4B). Although sgRNA G nucleobases have been reported to interact with Cas9 residues at different sites (G18, G43, G62, G81, and G89) (50), the thieno[3,4-d] pyrimidine core showed no obvious impairment of sgRNA activity, demonstrating that emissive thG alone, or in combination with BrC, can potentially replace natural counterparts in sgRNA while retaining full functionality *in vitro*. On the other hand, the introduction of two additional modifications, i.e., C^7A and ClU, into the sgRZA sequence abolished its activity in the cleavage assay (Figure 4B).

To explore this observation further, cleavage activity was examined using sgRZAs bearing single modifications (Supplementary Figure S9A). Pyrimidine-substituted (FC, BrC, ClU) sgRZAs achieved equivalent or even higher DNA cleavage efficiency than natural sgRNA. A previously solved crystal structure of a Cas9–sgRNA–DNA complex revealed that sgRNA nucleobases C67, C91, U23, U44, U50, U59, and U90 (Supplementary Figure S6) form interactions with Cas9 residues (51). It has also been widely reported that the 10–20 bp PAM-proximal seed region is critical for the Cas9-catalyzed DNA cleavage (52). However, from our observations, halogens at the C5 position of pyrimidines do not impair sgRNA recognition, and the presence of such modifications at seed region were well-tolerated *in vitro*. In contrast, no cleavage was observed when sgRZAs contained C^7A, regardless natural or modified pyrimidines were included in the sequence (Supplementary Figure S9A). This suggests that the N7 of adenine in sgRNAs might be essential for DNA cleavage processes. Consistent with this, the crystal structure of a Cas9–sgRNA–DNA complex shows that sgRNA nucleobases A42, A49, A51, and A88 directly interact with Cas9 residues (51), which may be abolished by deletion of the N7 of adenine in C^7A-modified sgRZAs.

To further substantiate this, we tested another A analog, 2,6-diaminopurine (DAP, Figure 1), which has an intact N7 position and forms three hydrogen bonds with its T/U base counterpart, stabilizing the nucleic acid duplex structure (53) (Figure 4C). Although fully modified thG + BrC + DAP + ClU-sgRZA failed to promote DNA hydrolysis even with prolonged reaction time (Supplementary Figure S3), sgRNA modified only with DAP resulted in 24% DNA cleavage compared to natural sgRNA (Figure 4C). Next, we examined the uridine analog m1Ψ and found that this could replace U in sgRNA, enhancing hydrolysis of target DNA (103%), similar to ClU-sgRZA (Figure 4C and Supplementary Figure S9A). This result indicates the possibility of utilizing pyrimidine analogs such as ClU, m1Ψ, BrC, and FC in sgRZA production to reduce immunogenicity, thereby increasing the efficiency of *in vivo* CRISPR-Cas9 gene editing. However, the A-modification of sgRZA adversely affected cleavage efficiency and should therefore be avoided.

To better understand the mechanism of target DNA recognition, variously modified target DZAs were tested in the *in vitro* cleavage assay using natural sgRNA (Supplementary Figure S9B). DZAs comprising single modifications FC, ClU, C^7A or DAP, and a G + FC + C^7A + ClU-DZA were hydrolyzed as efficiently as a natural DNA target (99–101% DNA cleavage). In contrast, all the C^7G-containing DZAs were resistant to DNA cleavage. It has been reported that the guanosine dinucleotide at the PAM site (5′-NGG-3′) forms essential hydrogen-bonding interactions with the Arg 1333 and Arg 1335 residues of Cas9 endonuclease through the N7 and O6-guanine positions located in the major groove (Supplementary Figure S10) (51,54). Therefore, it follows that the deletion of N7 on the GG dinucleotide reduces or abolishes PAM recognition. The introduction of a fluorine atom, which could act as an H-bond acceptor, bonded to a carbon at a position analogous to N7 of guanine (i.e., FC^7G modification, Figure 1) did not restore Cas9 cleavage (Supplementary Figure S9B). Other than the G analog, all other tested modifications in target DNA, including C^7A at the PAM site 5′-AGG-3′, were well tolerated by the Cas9 enzyme.

To verify this observation, we chemically synthesized primers containing one or two C^7G modifications at the 3′-end (Supplementary Table S2 and Figure S11) and, together with a Cy5-labeled natural reverse primer, generated short target DZAs by PCR with a chemically altered PAM site, i.e., AG*G, AGG* and AG*G* (G* indicates C^7G incorporation). The resulting 90-bp DZAs were evaluated as substrates of the Cas9 protein in the *in vitro* cleavage assay (Figure 4D). DNA cleavage occurred in the absence of a 7-nitrogen atom of the first G in the PAM site (65%), while C^7G at the second position presumably inhibited PAM recognition by Cas9, resulting in 8% DNA cleavage compared to a natural target DNA. These observations indicate that the interactions formed by the N7 of the second G at the PAM site play a more critical role in Cas9 recognition compared with those of the first G, but that deletion of the N7 of both completely inhibits DNA cleavage (Figure 4D). Furthermore, when all Gs in the sequence except for the PAM site were substituted by C^7G, efficient DNA cleavage similar to natural DNA could still be observed (98%; Supplementary Figure S9C), underscoring the importance of the N7 atom at the second G position of the PAM site over other positions in the target DNA sequence.

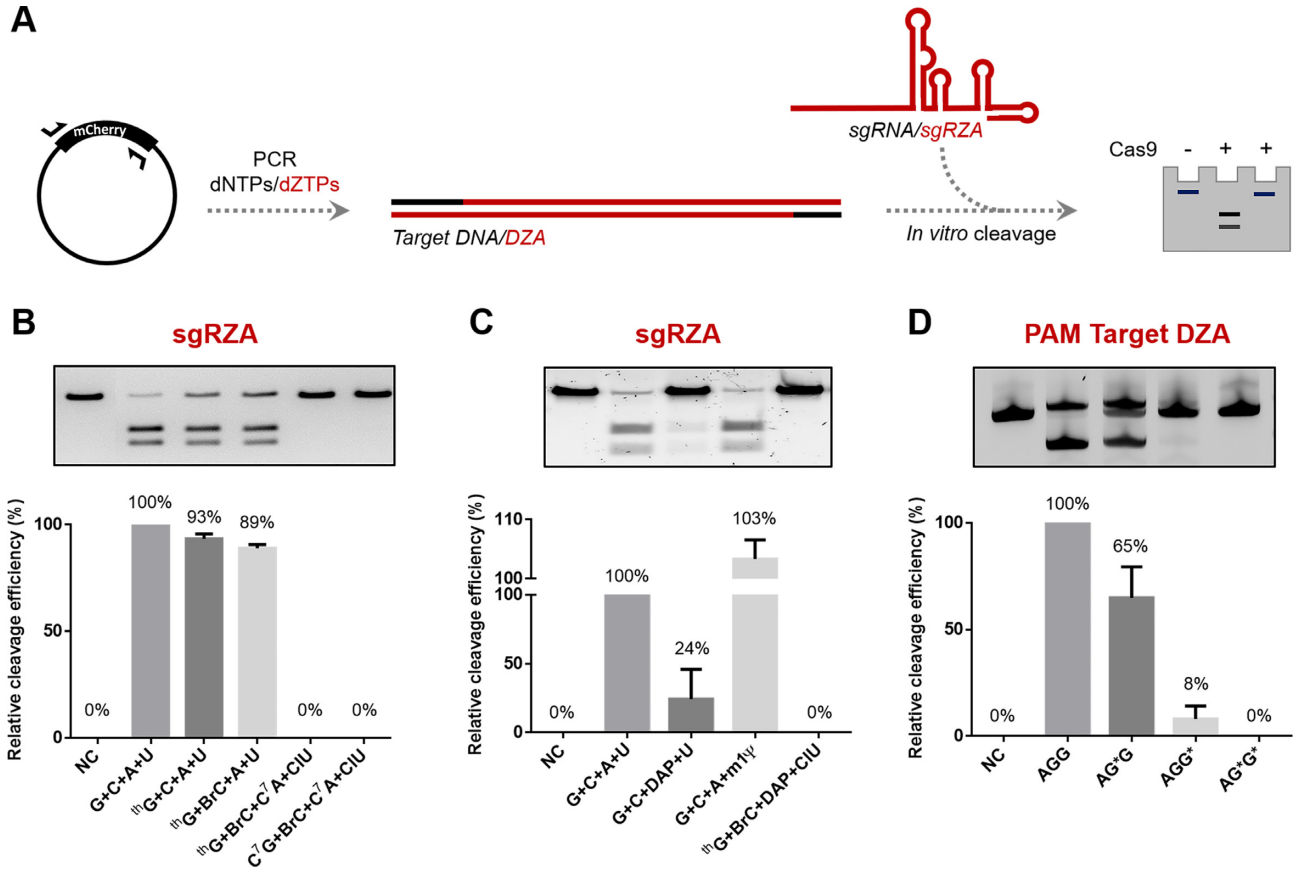


Figure 4. The CRISPR-Cas9 *in vitro* cleavage assay with base-modified sgRZA and target DZA. (**A**) General experimental scheme of the CRISPR-Cas9 *in vitro* cleavage assay with sgRZAs and target DZAs. The sgRZA and DZA sequences are shown in Supplementary Table S3. (**B**, **C**) Relative cleavage efficiency of various sgRZAs in the CRISPR-Cas9 DNA cleavage assay. (**D**) Relative cleavage efficiency of short target DZAs with different modified PAM sites, where G* refers to the C^7G modification. NC indicates the negative control, i.e., DNA cleavage assay without sgRNA. DNA cleavage efficiency with natural sgRNA (G + C + A + U; **B**, **C**) or natural PAM site (AGG; **D**) is taken as 100%. Data from three independent experiments are shown as mean ± SEM.

DISCUSSION

In a previous study, we described how the enzymatic synthesis of RZA was hampered by the poor incorporation efficiency of C^7GTP and ITP as surrogates of GTP by wtT7 RNAP (12). Here, we demonstrate that the replacement of GTP with an isomorphic thGTP enables *in vitro* transcription of 87- and 96-nt single-modified RZA transcripts. Transcriptional efficiency was maintained even with further substitution of the remaining three natural monomers with other base-modified nucleotides (BrC + C^7A + ClU). Important for the conversion of genetic information, the four nucleotide analogs thG + BrC + C^7A + ClU were incorporated into RZA with an accuracy comparable to a natural counterpart. Two T7 RNAP variants bearing the P266L and P266L/Y639F mutations were found to substantially improve the yield of full-length fully modified RZA products, producing approximately 2–3 times higher yield compared with the commercial TF polymerase.

Subsequently, we explored the ability of dthGTP to serve as a substrate of DNA polymerase in a PCR assay. Although we found that Taq DNA polymerase possessed a limited capability to incorporate dthGTP, the PCR yield could be doubled by replacing dC with dBrC as a pair for dthG (6% versus 11%). Incorporation of additional base modifications, such as C^7A and ClU, did not affect the PCR efficiency leading to the formation of a fully-modified ds-DZA fragment. However, the PCR yield of thG-containing products was generally lower compared to other base-modifications, suggesting that testing different polymerases may enable further gains to be made in efficient PCR amplification of nucleotide analogs containing a thieno[3,4-d] pyrimidine core. Overall, the successful enzymatic production of fluorescent thG-DZA/RZA fragments provides a promising route to access modified emissive nucleic acids for a wide variety of applications.

Furthermore, a thG-modified sgRZA was enzymatically generated using the above optimized conditions and demonstrated an ability to direct the Cas9 nuclease to its target DNA fragment in a CRISPR-Cas9 *in vitro* assay with cleavage efficiency similar to natural sgRNA. These results confirmed that thG-substitution did not fundamentally alter the properties of the Cas9-RNA complex, allowing successful cleavage of DNA. The same result was observed upon the introduction of modified pyrimidines, such as FC, BrC, ClU, and m1Ψ, whose presence was tolerated throughout the guide RNA. To our knowledge, this pro-

vides the first evidence supporting the use of base-modified monomers, in particular m1Ψ, in sgRZA to enable enhanced CRISPR-based therapy. The N7 of adenine was found to be an essential element for maintaining sgRNA activity and although an additional 2-amino group (as occurs in DAP) disrupted residue interactions, DNA cleavage was nevertheless possible with low efficiency. The introduction of base-modified analogs into the sgRNA structure can aid our understanding of structural and biological interactions within the guide RNA-Cas9 complex–target DNA involved in CRISPR gene-editing processes.

Lastly, the modification of N7 in guanine residues at the PAM site was demonstrated to be a major reason for the loss of target DNA recognition. Our data suggest that interactions between the PAM site and Arg 1333 are of greater significance than those with the Arg 1335 residue in terms of enzymatic recognition, which may inform the engineering of Cas9 variants with altered PAM site recognition. Alternatively, it may be beneficial in some instances to introduce chemical modification at the PAM sites, namely C^7G, in order to protect the upstream gene from bacterial Cas9 cleavage, thus reducing off-site targeting and improving the specificity of CRISPR-Cas9 gene editing.

In summary, this study demonstrates the successful generation of fully base-modified thG-dsDZA and thG-RZA with intrinsic fluorescence, the first to do so using an enzymatic approach. Although dthGTP was found to be a poor substrate for dsDZA synthesis, and the resulting dsDZA templates inhibited *in vitro* transcription, thGTP was found to be a good G substitute for RZA synthesis. The efficiency of thG-RZA transcription could be further improved using T7 RNAP variants and additives. In principle, thGTP could be used as an attractive guanosine analog to produce different functional RNAs, including noncoding RNA, aptamers, ribozymes, or mRNAs, by enzymatic synthesis with T7 RNAP. Modified RZA with strong visible emission due to thG might enable fluorescent detection of discrete biomolecular events without altering the RNA structure and function and thus can be utilized for the generation of fluorescent probes in cell assays, RNA catalysis, and RNA structural analysis. In addition, we have demonstrated the importance of the 7-nitrogen atom of adenines in sgRNA as well as the tolerance of sgRNA to the m1Ψ base, and further elucidated the contribution of the 7-nitrogen atom of guanines, especially at the second position of the PAM recognition site of the DNA target, to the mechanism of cleavage by CRISPR-Cas9 systems.

DATA AVAILABILITY

The data underlying this article are available in the article and in its online supplementary material.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

We thank Alexander I. Taylor for proofreading this manuscript and Chantal Biernaux for editorial help. Authors are indebted to Prof. J. Rozenski for mass spectrometry analysis and G. Schepers for the chemical synthesis of oligonucleotides.

FUNDING

Fonds Wetenschappelijk Onderzoek (FWO), Flanders Research Foundation [12Q8619N and 1509920N to E.E]; European Union's Horizon 2020 research and innovation program (FetOpen) [965135]; China Scholarship Council (CSC) [201707040069 to H.Y.]. Funding for open access charge: FWO [G085321N].
Conflict of interest statement. None declared.

REFERENCES

1. Alberts,B., Johnson,A., Lewis,J., Morgan,D., Raff,M., Roberts,K. and Walter,P. (2014) *Molecular Biology of the Cell Garland Science*. New York and Abingdon, UK.
2. Pardi,N., Hogan,M.J., Porter,F.W. and Weissman,D. (2018) mRNA vaccines-a new era in vaccinology. *Nat. Rev. Drug Discov.*, **17**, 261–279.
3. Qin,S., Tang,X., Chen,Y., Chen,K., Fan,N., Xiao,W., Zheng,Q., Li,G., Teng,Y., Wu,M. *et al.* (2022) mRNA-based therapeutics: powerful and versatile tools to combat diseases. *Signal Transduct. Target. Ther.*, **7**, 166.
4. Zhu,Y., Zhu,L., Wang,X. and Jin,H. (2022) RNA-based therapeutics: an overview and prospectus. *Cell Death. Dis.*, **13**, 1–15.
5. Martin,J., Krzysztof,C., Ines,F., Michael,H., A.,D.J. and Emmanuelle,C. (2012) A programmable dual-rna–guided DNA endonuclease in adaptive bacterial immunity. *Science (80-).*, **337**, 816–821.
6. Ryan,D.E., Taussig,D., Steinfeld,I., Phadnis,S.M., Lunstad,B.D., Singh,M., Vuong,X., Okochi,K.D., McCaffrey,R., Olesiak,M. *et al.* (2018) Improving CRISPR–Cas specificity with chemical modifications in single-guide RNAs. *Nucleic Acids Res.*, **46**, 792–803.
7. Winkle,M., El-Daly,S.M., Fabbri,M. and Calin,G.A. (2021) Noncoding RNA therapeutics — challenges and potential solutions. *Nat. Rev. Drug. Discov.*, **20**, 629–651.
8. Kulkarni,J.A., Witzigmann,D., Thomson,S.B., Chen,S., Leavitt,B.R., Cullis,P.R. and van der Meel,R. (2021) The current landscape of nucleic acid therapeutics. *Nat. Nanotechnol.*, **16**, 630–643.
9. Nance,K.D. and Meier,J.L. (2021) Modifications in an emergency: the role of N1-Methylpseudouridine in COVID-19 vaccines. *ACS Cent. Sci.*, **7**, 748–756.
10. Parr,C.J.C., Wada,S., Kotake,K., Kameda,S., Matsuura,S., Sakashita,S., Park,S., Sugiyama,H., Kuang,Y. and Saito,H. (2020) N1-Methylpseudouridine substitution enhances the performance of synthetic mRNA switches in cells. *Nucleic Acids Res.*, **48**, e35.
11. Sekine,M. (2018) *Recent Development of Chemical Synthesis of RNA BT - Synthesis of Therapeutic Oligonucleotides*. In: Obika,S. and Sekine,M. (eds.) Springer Singapore, Singapore, pp. 41–65.
12. Yang,H., Eremeeva,E., Abramov,M. and Herdewijn,P. (2021) The network of replication, transcription, and reverse transcription of a synthetic genetic cassette. *Angew. Chemie - Int. Ed.*, **60**, 4175–4182.
13. Milligan,J.F., Groebe,D.R., Witherell,G.W. and Uhlenbeck,O.C. (1987) Oligoribonucleotide synthesis using T7 RNA polymerase and synthetic DNA templates. *Nucleic Acids Res.*, **15**, 8783–8798.
14. Milligan,J.F. and Uhlenbeck, O.C.B.T.-M. in E.Uhlenbeck, O.C.B.T.-M. in E. (1989) [5]Synthesis of small RNAs using T7 RNA polymerase. In: *RNA Processing Part A: General Methods*. Academic Press, Vol. **180**, pp. 51–62.
15. Milisavljevič,N., Perlíková,P., Pohl,R. and Hocek,M. (2018) Enzymatic synthesis of base-modified RNA by T7 RNA polymerase. A systematic study and comparison of 5-substituted pyrimidine and 7-substituted 7-deazapurine nucleoside triphosphates as substrates. *Org. Biomol. Chem.*, **16**, 5800–5807.
16. McCoy,L.S., Shin,D. and Tor,Y. (2014) Isomorphic emissive GTP surrogate facilitates initiation and elongation of in vitro transcription reactions. *J. Am. Chem. Soc.*, **136**, 15176–15184.

17. Li,Y., Fin,A., McCoy,L. and Tor,Y. (2017) Polymerase-Mediated site-specific incorporation of a synthetic fluorescent isomorphic g surrogate into RNA. *Angew. Chemie - Int. Ed.*, **56**, 1303–1307.
18. Shin,D., Sinkeldam,R.W. and Tor,Y. (2011) Emissive RNA alphabet. *J. Am. Chem. Soc.*, **133**, 14912–14915.
19. Samanta,P.K., Manna,A.K. and Pati,S.K. (2012) Thieno analogues of RNA nucleosides: a detailed theoretical study. *J. Phys. Chem. B*, **116**, 7618–7626.
20. Sinkeldam,R.W., McCoy,L.S., Shin,D. and Tor,Y. (2013) Enzymatic interconversion of isomorphic fluorescent nucleosides: adenosine deaminase transforms an adenosine analogue into an inosine analogue. *Angew. Chemie Int. Ed.*, **52**, 14026–14030.
21. Didier,P., Kuchlyan,J., Martinez-Fernandez,L., Gosset,P., Leónard,J., Tor,Y., Improta,R. and Mély,Y. (2020) Deciphering the pH-dependence of ground- A nd excited-state equilibria of thienoguanine. *Phys. Chem. Chem. Phys.*, **22**, 7381–7391.
22. Sinkeldam,R.W., Greco,N.J. and Tor,Y. (2010) Fluorescent analogs of biomolecular building blocks: design, properties, and applications. *Chem. Rev.*, **110**, 2579–2619.
23. Sholokh,M., Sharma,R., Shin,D., Das,R., Zaporozhets,O.A., Tor,Y. and Mély,Y. (2015) Conquering 2-aminopurines deficiencies: highly emissive isomorphic guanosine surrogate faithfully monitors guanosine conformation and dynamics in DNA. *J. Am. Chem. Soc.*, **137**, 3185–3188.
24. Kuchlyan,J., Martinez-Fernandez,L., Mori,M., Gavvala,K., Ciaco,S., Boudier,C., Richert,L., Didier,P., Tor,Y., Improta,R. *et al.* (2020) What makes thienoguanosine an outstanding fluorescent DNA probe? *J. Am. Chem. Soc.*, **142**, 16999–17014.
25. Meyer,A.J., Garry,D.J., Hall,B., Byrom,M.M., McDonald,H.G., Yang,X., Yin,Y.W. and Ellington,A.D. (2015) Transcription yield of fully 2′-modified RNA can be increased by the addition of thermostabilizing mutations to T7 RNA polymerase mutants. *Nucleic Acids Res.*, **43**, 7480–7488.
26. Boulain,J.C., Dassa,J., Mesta,L., Savatier,A., Costa,N., H.Muller,B., L'hostis,G., A.Stura,E., Troesch,A. and Ducancel,F. (2013) Mutants with higher stability and specific activity from a single thermosensitive variant of T7 RNA polymerase. *Protein Eng. Des. Sel.*, **26**, 725–734.
27. Chelliserrykattil,J. and Ellington,A.D. (2004) Evolution of a T7 RNA polymerase variant that transcribes 2′-O-methyl RNA. *Nat. Biotechnol.*, **22**, 1155–1160.
28. Padilla,R. and Sousa,R. (2002) A Y639F/H784A T7 RNA polymerase double mutant displays superior properties for synthesizing RNAs with non-canonical NTPs. *Nucleic Acids Res.*, **30**, 138e.
29. Hoshika,S., Leal,N.A., Kim,M.J., Kim,M.S., Karalkar,N.B., Kim,H.J., Bates,A.M., Watkins,N.E., SantaLucia,H.A., Meyer,A.J. *et al.* (2019) Hachimoji DNA and RNA: a genetic system with eight building blocks. *Science (80-.).*, **363**, 884–887.
30. Sander,J.D. and Joung,J.K. (2014) CRISPR-Cas systems for editing, regulating and targeting genomes. *Nat. Biotechnol.*, **32**, 347–350.
31. Cong,L., Ran,F.A., Cox,D., Lin,S., Barretto,R., Habib,N., Hsu,P.D., Wu,X., Jiang,W., Marraffini,L.A. *et al.* (2013) Multiplex genome engineering using CRISPR/Cas systems. *Science*, **339**, 819–823.
32. Hendel,A., Bak,R.O., Clark,J.T., Kennedy,A.B., Ryan,D.E., Roy,S., Steinfeld,I., Lunstad,B.D., Kaiser,R.J., Wilkens,A.B. *et al.* (2015) Chemically modified guide RNAs enhance CRISPR-Cas genome editing in human primary cells. *Nat. Biotechnol.*, **33**, 985–989.
33. Rahdar,M., McMahon,M.A., Prakash,T.P., Swayze,E.E., Bennett,C.F. and Cleveland,D.W. (2015) Synthetic CRISPR RNA-Cas9–guided genome editing in human cells. *Proc. Natl. Acad. Sci. USA*, **112**, E7110–E7117.
34. Shapiro,J., Tovin,A., Iancu,O., Allen,D. and Hendel,A. (2021) *Chemical Modification of Guide RNAs for Improved CRISPR Activity in CD34+ Human Hematopoietic Stem and Progenitor Cells BT - CRISPR Guide RNA Design: Methods and Protocols*. In: Fulga,T.A., Knapp,D.J.H.F. and Ferry,Q.R.V. (eds.), Springer US, New York, NY, pp. 37–48.
35. Moon,S.B, Kim,D.Y., Ko,J.-H., Kim,J.-S. and Kim,Y.-S. (2019) Improving CRISPR genome editing by engineering guide RNAs. *Trends Biotechnol.*, **37**, 870–881.
36. Yin,H., Song,C.-Q., Suresh,S., Wu,Q., Walsh,S., Rhym,L.H., Mintzer,E., Bolukbasi,M.F., Zhu,L.J., Kauffman,K. *et al.* (2017) Structure-guided chemical modification of guide RNA enables potent non-viral in vivo genome editing. *Nat. Biotechnol.*, **35**, 1179–1187.
37. Bisaria,N., Jarmoskaite,I. and Herschlag,D. (2017) Lessons from enzyme kinetics reveal specificity principles for RNA-Guided nucleases in RNA interference and CRISPR-Based genome editing. *Cell Syst.*, **4**, 21–29.
38. Wu,X., Scott,D.A., Kriz,A.J., Chiu,A.C., Hsu,P.D., Dadon,D.B., Cheng,A.W., Trevino,A.E., Konermann,S., Chen,S. *et al.* (2014) Genome-wide binding of the CRISPR endonuclease cas9 in mammalian cells. *Nat. Biotechnol.*, **32**, 670–676.
39. Fu,Y., Foden,J.A., Khayter,C., Maeder,M.L., Reyon,D., Joung,J.K. and Sander,J.D. (2013) High-frequency off-target mutagenesis induced by CRISPR-Cas nucleases in human cells. *Nat. Biotechnol.*, **31**, 822–826.
40. Kojima,T., Furukawa,K., Maruyama,H., Inoue,N., Tarashima,N., Matsuda,A. and Minakawa,N. (2013) PCR amplification of 4′-ThioDNA using 2′-Deoxy-4′-thionucleoside 5′-Triphosphates. *ACS Synth. Biol.*, **2**, 529–536.
41. Chen,Z. and Zhang,Y. (2005) Dimethyl sulfoxide targets phage RNA polymerases to promote transcription. *Biochem. Biophys. Res. Commun.*, **333**, 664–670.
42. Eremeeva,E., Abramov,M., Margamuljana,L. and Herdewijn,P. (2017) Base-modified nucleic acids as a powerful tool for synthetic biology and biotechnology. *Chem. - A Eur. J.*, **23**, 9560–9576.
43. Eremeeva,E., Abramov,M., Marlière,P. and Herdewijn,P. (2017) The 5-chlorouracil:7-deazaadenine base pair as an alternative to the dT:dA base pair. *Org. Biomol. Chem.*, **15**, 168–176.
44. Guza,R., Kotandeniya,D., Murphy,K., Dissanayake,T., Lin,C., Giambasu,G.M., Lad,R.R., Wojciechowski,F., Amin,S., Sturla,S.J. *et al.* (2011) Influence of C-5 substituted cytosine and related nucleoside analogs on the formation of benzo[a]pyrene diol epoxide-dG adducts at CG base pairs of DNA. *Nucleic Acids Res.*, **39**, 3988–4006.
45. Park,S., Otomo,H., Zheng,L. and Sugiyama,H. (2014) Highly emissive deoxyguanosine analogue capable of direct visualization of B–Z transition. *Chem.*, **50**, 1573–1575.
46. Eremeeva,E., Abramov,M., Margamuljana,L., Rozenski,J., Pezo,V., Marlière,P. and Herdewijn,P. (2016) Chemical morphing of DNA containing four noncanonical bases. *Angew. Chemie Int. Ed.*, **55**, 7515–7519.
47. Liu,W., Shin,D., Tor,Y. and Cooperman,B.S. (2013) Monitoring translation with modified mRNAs strategically labeled with isomorphic fluorescent guanosine mimetics. *ACS Chem. Biol.*, **8**, 2017–2023.
48. Cong,D., Li,Y., Ludford,P.T. and Tor,Y. (2022) Isomorphic fluorescent nucleosides facilitate a real-time monitoring of RNA depurination by ribosome inactivating proteins. *Chem. – A Eur. J.*, **28**, e202200994.
49. Li,Y., Fin,A., McCoy,L. and Tor,Y. (2017) Polymerase-Mediated site-specific incorporation of a synthetic fluorescent isomorphic g surrogate into RNA. *Angew. Chemie Int. Ed.*, **56**, 1303–1307.
50. Nishimasu,H., Ran,F.A., Hsu,P.D., Konermann,S., Shehata,S.I., Dohmae,N., Ishitani,R., Zhang,F. and Nureki,O. (2014) Crystal structure of cas9 in complex with guide RNA and target DNA. *Cell*, **156**, 935–949.
51. Anders,C., Niewoehner,O., Duerst,A. and Jinek,M. (2014) Structural basis of PAM-dependent target DNA recognition by the cas9 endonuclease. *Nature*, **513**, 569–573.
52. Kartje,Z.J., Barkau,C.L., Rohilla,K.J., Ageely,E.A. and Gagnon,K.T. (2018) Chimeric guides probe and enhance cas9 biochemical activity. *Biochemistry*, **57**, 3027–3031.
53. Bailly,C. (1998) The use of diaminopurine to investigate structural properties of nucleic acids and molecular recognition between ligands and DNA. *Nucleic Acids Res.*, **26**, 4309–4314.
54. Anders,C., Bargsten,K. and Jinek,M. (2016) Structural plasticity of PAM recognition by engineered variants of the RNA-Guided endonuclease cas9. *Mol. Cell*, **61**, 895–902.