

UC San Diego

UC San Diego Previously Published Works

Title

Site-Specific Covalent Labeling of DNA Substrates by an RNA Transglycosylase

Permalink

<https://escholarship.org/uc/item/8wq8r7fz>

Journal

Journal of the American Chemical Society, 145(14)

ISSN

0002-7863

Authors

Tota, Ember M
Devaraj, Neal K

Publication Date

2023-04-12

DOI

10.1021/jacs.3c00861

Peer reviewed

Site-Specific Covalent Labeling of DNA Substrates by an RNA Transglycosylase

Ember M. Tota and Neal K. Devaraj*



Cite This: *J. Am. Chem. Soc.* 2023, 145, 8099–8106



Read Online

ACCESS |



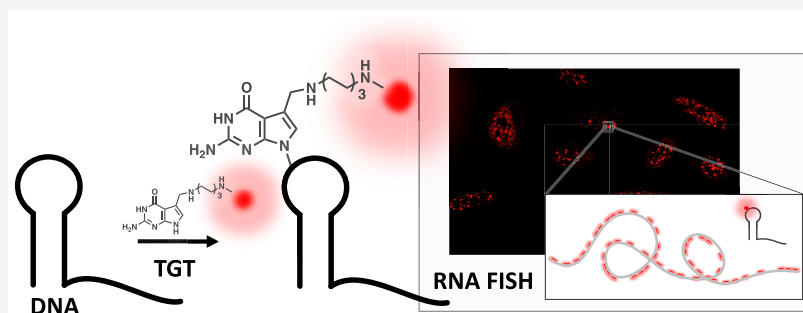
Metrics & More



Article Recommendations



Supporting Information



ABSTRACT: Bacterial tRNA guanine transglycosylases (TGTs) catalyze the exchange of guanine for the 7-deazaguanine queuine precursor, prequeuosine1 (preQ1). While the native nucleic acid substrate for bacterial TGTs is the anticodon loop of queuine-cognate tRNAs, the minimum recognition sequence for the enzyme is a structured hairpin containing the target G nucleobase in a “UGU” loop motif. Previous work has established an RNA modification system, RNA-TAG, in which *Escherichia coli* TGT exchanges the target G on an RNA of interest for chemically modified preQ1 substrates linked to a small-molecule reporter such as biotin or a fluorophore. While extending the substrate scope of RNA transglycosylases to include DNA would enable numerous applications, it has been previously reported that TGT is incapable of modifying native DNA. Here, we demonstrate that TGT can in fact recognize and label specific DNA substrates. Through iterative testing of rationally mutated DNA hairpin sequences, we determined the minimal sequence requirements for transglycosylation of unmodified DNA by *E. coli* TGT. Controlling steric constraint in the DNA hairpin dramatically affects labeling efficiency, and, when optimized, can lead to near-quantitative site-specific modification. We demonstrate the utility of our newly developed DNA-TAG system by rapidly synthesizing probes for fluorescent Northern blotting of spliceosomal U6 RNA and RNA FISH visualization of the long noncoding RNA, metastasis-associated lung adenocarcinoma transcript 1 (MALAT1). The ease and convenience of the DNA-TAG system will provide researchers with a tool for accessing a wide variety of versatile and affordable modified DNA substrates.

INTRODUCTION

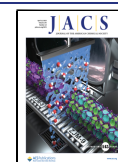
Probe-modified oligonucleotides have widespread applications in imaging, diagnostics, nanotechnology, and medicine.^{1–3} Modified nucleic acids are most commonly produced by incorporating functionalized nucleotides during solid-phase synthesis using phosphoramidite chemistry. In addition, a number of alternative techniques exist that rely on chemical or enzymatic modification of nucleobases or the sugar backbone.^{4–11} Techniques that employ enzymes offer several advantages over those relying on chemical modification as they occur in mild and aqueous conditions, are precisely targeted, and are typically very efficient. The deoxyribose backbone of DNA affords aqueous stability that is lacking in its RNA counterpart, and as such, chemically functionalized single-stranded DNAs (ssDNA) are widely used in biotechnology. Current methods for enzymatic modification of DNA include the 3′ insertion of modified nucleobases using terminal deoxynucleotidyltransferase (TdT), the 5′ insertion of

modified phosphate groups, and the transfer of click handles to a target adenosine (mTAG).^{12–14} Unfortunately, the small-molecule substrate tolerance and precision of insertion for enzyme-mediated ssDNA modification strategies has been limited. The improved ability to site-specifically label ssDNA with enzymes in mild aqueous conditions could have significant applications given the broad use of ssDNA in diagnostics, therapeutics, materials research, imaging, detection, and chemical barcoding.^{11,15–17}

Previously, our group developed RNA transglycosylation at guanosine (RNA-TAG) to site-specifically label RNA using

Received: January 22, 2023

Published: March 29, 2023



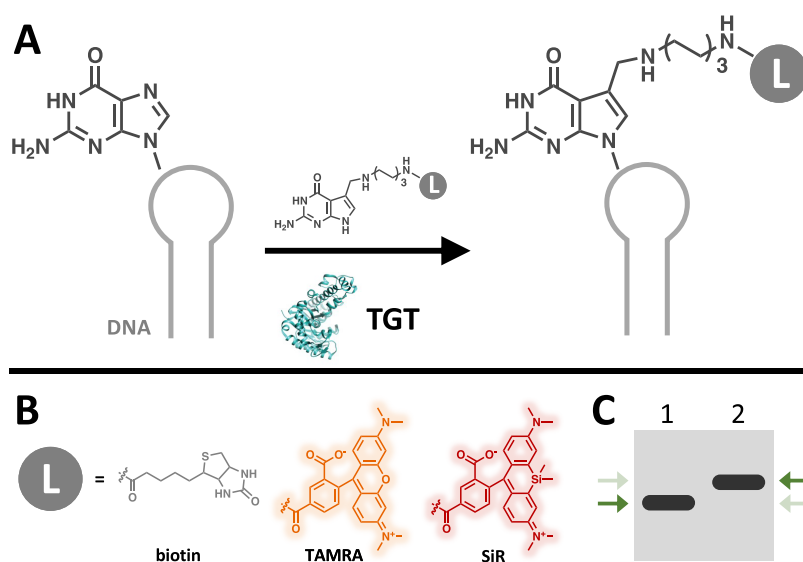


Figure 1. Proposed scheme of *E. coli* TGT mediated modification of DNA with functional small molecules. (A) General scheme of DNA-TAG. *E. coli* TGT irreversibly exchanges a specific guanine with a preQ1-ligand (L) on a DNA hairpin. (B) PreQ1-ligands used in this work include preQ1-biotin, preQ1-tetramethylrhodamine (TAMRA), and preQ1-Silicon Rhodamine (SiR). (C) Cartoon depiction of expected upward gel shift due to increased mass from the base exchange reaction as indicated by the green arrows. Lane 1: DNA; Lane 2: preQ1-L modified DNA.

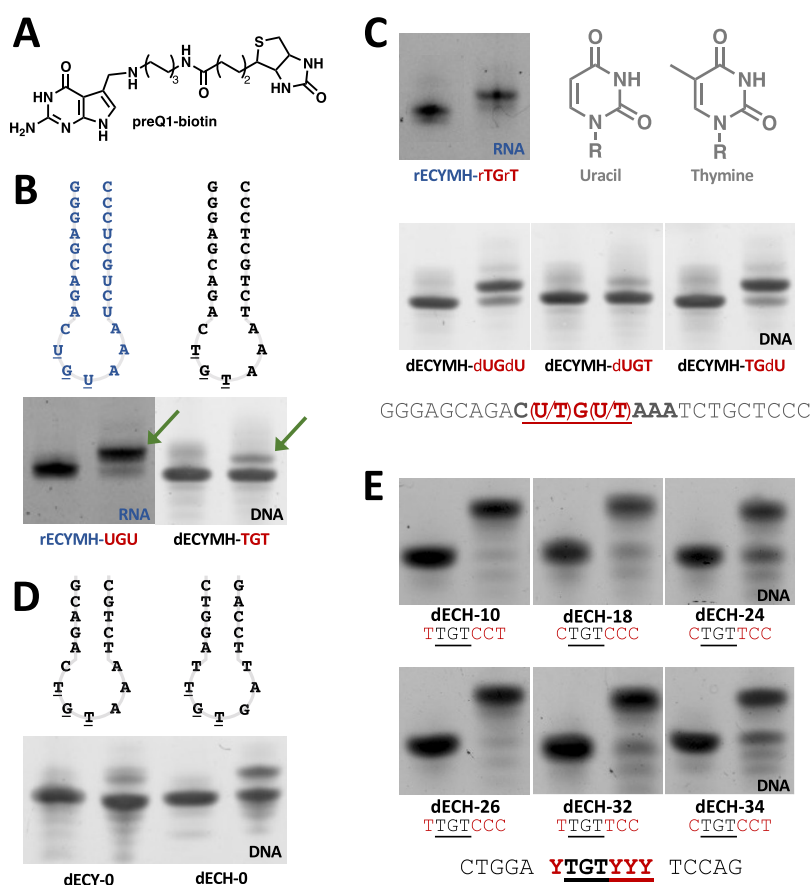


Figure 2. Determining the minimal sequence requirements for DNA modification by *E. coli* TGT. Unmodified oligos are on the left, and reaction products are on the right. Modification is determined by an upward gel shift of the oligo after insertion of preQ1-biotin. (A) PreQ1-biotin probe. (B) TGT modification of extended-stem RNA and DNA hairpins derived from the anticodon loop of tRNA^{Phe}; arrows indicate modified product as is evidenced by the upward gel shift. (C) TGT labeling of T → dU mutants. A single mutation is sufficient for TGT modification of dECYMH. (D) TGT modifies dECY-0 and dECH-0 at ~5 and ~40% efficiency, respectively. (E) TGT modifies DNA-based hairpins with >90% efficiency when the loop sequence is YTGIYCC or YTGITCCY where Y is C or T, except for CTGTTCC, which is labeled with ~85% efficiency.

bacterial tRNA guanine transglycosylase (TGT).^{18,19} TGTs catalyze the exchange of guanine for 7-deazaguanine derivatives on the anticodon loop of queuine-cognate tRNAs.²⁰ These enzymatic modifications are present across all three kingdoms of life, differing slightly in their small-molecule and tRNA substrate counterparts. It had previously been shown that the entire tRNA structure is not necessary for the enzymatic activity of *Escherichia coli* TGT toward RNA substrates, but rather, the minimum recognition sequence is a short 17-nucleotide hairpin (ECY-A1; the anticodon loop of tRNA^{Tyr}), which contains a uridine-flanked target guanine in its loop.²¹ In RNA-TAG, *E. coli* TGT recognizes this minimal hairpin incorporated into an RNA of interest and catalyzes the exchange of the target guanine for synthetic derivatives of the natural preQ1 substrate that are conjugated to functional groups such as fluorophores or affinity ligands.¹⁸ We have demonstrated the utility of RNA-TAG in a range of applications including RNA biotinylation for proteomics studies, controlling mRNA translation, and light-activated CRISPR gene editing in cells.^{18,19,22–27}

Extending the RNA-TAG system to work on DNA would enable one-step, site-specific labeling of ssDNA substrates, increasing the accessibility of small-molecule-modified DNA oligonucleotides and further enabling applications of ssDNA in nanotechnology, therapeutics, and basic research. The exchange of guanine for preQ1 by TGT is electronically dependent on the guanine nucleobase and the endocyclic oxygen of the ribose backbone.²⁰ This suggests that the lack of the exocyclic 2' hydroxyl groups on the backbone of DNA substrates would not impair *E. coli* TGT recognition or activity. However, previous work had deemed *E. coli* TGT incapable of acting on unmodified DNA substrates and showed that unnatural nucleobase substitutions (T → dU) of the thermally stabilized 4 base-pair extended-stem mini helix (dECYMH) were required for efficient recognition.^{20,28}

We speculated that optimizing the hairpin structure could enable recognition of native DNA by *E. coli* TGT. Indeed, through iterative testing of rationally designed DNA hairpin sequences, we were able to reveal the minimal sequence requirements for DNA substrate recognition by the enzyme (Figures 1A and S3), enabling development of a technique we term DNA transglycosylation at deoxyguanosine, or DNA-TAG. DNA-TAG enables the one-step, site-specific labeling of ssDNA substrates bearing a minimal 17-nucleotide hairpin motif with synthetic preQ1 substrates conjugated to functional groups such as biotin and fluorophores (Figure 1B). Next, we characterize the ability of DNA-TAG to label DNAs that include one or more recognition sequences and explore the general substrate scope for DNA-TAG. Finally, we show how DNA-TAG can be used to quickly and inexpensively generate probes for fluorescent Northern blotting and RNA FISH.

RESULTS AND DISCUSSION

Previous studies reported that TGT does not modify the DNA equivalent substrate (dECYMH) of the extended-stem tRNA^{Tyr}-based RNA mini helix (ECYMH). However, it was also reported that when all “T” nucleobases in the hairpin were replaced with “U”, the enzyme was able to recognize and form the covalent intermediate with a deoxyribose-based substrate (dUdECYMH).²⁸ While we confirmed only trace modification (less than 5% by densitometry) of dECYMH with our preQ1-biotin probe (Figure 2A,B), the published mechanism of TGT catalysis alludes to the possibility that the enzyme could

recognize unmodified DNA substrates, as the lack of a 2' hydroxyl group on the substrate nucleotide does not prevent catalysis.^{20,28} With this information, we designed a series of experiments to assess the limitations of TGT toward accepting DNA, restricting the number of T → dU mutations to only those that are part of the minimal “UGU” recognition element in the extended-stem DNA analog dECYMH: GGGAGCA-GACXGXAAATCTGC **TCCC** (dECY-A1 is italicized, the 4 base-pair extension and loop sequences are underlined, and the minimum recognition element is bolded). For all DNA substrate modification experiments, we relied on an established oligonucleotide gel shift assay.^{19,24,27} Observation of an upward gel shift of the product in relation to the starting material reports an increase in mass from the enzymatic exchange of guanine for a preQ1-biotin small-molecule reporter conjugate (Figures 1B,C and 2A). We surveyed transglycosylase activity toward dECYMH hairpin variants containing the following minimal recognition element identities: TGT, dUGdU, dUGT, and TGdU. While the TGT- and dUGT-containing hairpins showed no appreciable modification by the enzyme, those containing dUGdU and TGdU were modified to a similar extent as the wild-type RNA substrate, as evidenced by the upward gel shift in the product lane (Figure 2B,C). It appears that a single nucleobase mutation of the minimal recognition element from TGT to TGdU is sufficient for TGT to recognize and modify the dECYMH substrate. This finding expands on the previous observation that turnover of a DNA substrate only occurred when all T nucleotides in the hairpin were replaced with dU.²⁸ Encouraged by this result, we next set out to ask which aspects of dECYMH were limiting TGT enzyme activity. We synthesized an RNA hairpin (rECYMH-rTGrT) in which all U residues were replaced with 5-meU (rT). Surprisingly, we found that this substrate had a nearly quantitative turnover, similar to the wild-type RNA hairpin (Figure 2B,C), indicating that TGT can efficiently modify nucleic acid substrates containing thymine bases. Combined with the fact that a minimally modified DNA sequence with a single T → dU mutation is also recognized, our initial findings suggested that the inability of *E. coli* TGT to modify dECYMH likely arises from the combined steric constraints of the deoxyribose backbone and the methyl groups on the T bases of the putative dTdGdT recognition element.

Based on these observations, we hypothesized that, rather than a fundamental incompatibility of TGT with DNA oligonucleotides, the seeming inability of TGT to modify DNA was steric and could be overcome by using an alternative DNA substrate. Previous studies have focused on derivatives of the 17-nucleotide ECY-A1 and 25-nucleotide ECYMH hairpins which are derived from the anticodon loop of the tyrosine tRNA containing the GUA anticodon. However, all four tRNAs with GUN anticodons are substrates for TGT, including tRNA^{Tyr}, tRNA^{His}, tRNA^{Asp}, and tRNA^{Asn}. While each of these four tRNAs have the same “UGU” minimal recognition element, the rest of the anticodon loop sequence varies. Therefore, we decided that hairpin substrates based on these alternative tRNAs would be a good starting point for engineering a novel DNA-TAG substrate. Adapting the already established nomenclature, and to account for mutations, we refer to the 17-nucleotide wild-type hairpins as dECY-0 (originally dECY-A1), dECH-0, dECD-0, and dECN-0, corresponding to the hairpins derived from tRNA^{Tyr}, tRNA^{His}, tRNA^{Asp}, and tRNA^{Asn}, respectively. Mutant hairpins are referred to using a sequential numbering scheme based on

their parent hairpin (e.g., dECY-1, dECY-2, etc.). All hairpins were treated with *E. coli* TGT and preQ1-biotin and the resulting products analyzed via urea polyacrylamide gel electrophoresis (PAGE) gel shift. Interestingly, both the dECH-0 and dECD-0, hairpins showed an appreciable increase in TGT labeling activity compared to their dECY-0 counterpart, with labeled product increasing from ~5 to ~40% by densitometry (Figures 2D and S1). We note that dECN-0 shows a shift indicative of degradation rather than modification, a phenomenon that is consistent for most oligos bearing a 3'T (Figures S1, S2, and S4). While the data was included here for completeness, it is outside the scope of this work and will not be discussed further. The approximate increase in labeling of the dECH-0 and dECD-0 substrates to 40% was promising, but not comparable to the near-quantitative labeling observed with RNA-TAG.¹⁸ To further improve the DNA-TAG substrate, we next sought to determine the importance of the stems and/or loops for each of the tRNA-derived hairpins. We designed composite substrates, swapping the stems and loops of all four cognate DNA analogs. All hairpins with the aspartate-based loop (dECD) demonstrated labeling of over 40%, with the histidine stem-aspartate loop substrate reaching over 50% by densitometry (dECH-dECD; Figure S1). The results of this experiment suggested that while both the stem and the loop sequences contribute to the ability of TGT to turnover a given substrate, the loop sequence plays the more critical role in substrate compatibility, in line with previous observations during biochemical and structural studies of the enzyme and tRNA substrate.^{21,29}

We decided to move forward with the dECH-0 substrate for further development and designed nine hairpin mutants (dECH-1–9; Figure S2). These experiments confirmed the importance of the loop and led us to design and test another 31 mutants (dECH-10–40; Figure S3). This round of mutants produced optimal DNA-TAG substrates and revealed the minimal substrate loop requirement for efficient recognition and modification of DNA by TGT to be either YTGTCY or YTGTYCC, where Y represents smaller pyrimidine bases (C or T), further corroborating our hypothesis that steric effects play a major role in substrate recognition (Figure 2E). We chose one of the hairpins with the highest labeling efficiency, dECH-10, as our model substrate to probe how stem sequence affects DNA modification by TGT. Stems derived from the anticodon arm sequences of all *E. coli* tRNAs were appended with the dECH-10 "TTGTCCT" loop (Figure S4). Nearly all of the anticodon-based hairpins were appreciably labeled regardless of the stem sequence, apart from some apparent multilabeling and degradation products, similar to that seen for dECN-0. Ultimately, our data suggest that, with the right loop, the stem is far less important, as 16 DNA hairpins with different stem sequences are nearly quantitatively labeled by *E. coli* TGT (all greater than 95% by densitometry; Figure S5).

While any of the 16 substrates are valid candidates for use with DNA-TAG, we performed subsequent experiments with dECH-10 (hairpin 176, CTGGATTGTCCTTCCAG). We analyzed the kinetics of preQ1-biotin insertion by TGT and found that modification was nearly complete after 4 h (Figure S6). To confirm the result of our gel shift assay, we verified labeling with liquid chromatography–mass spectrometry analysis (LCMS). The LC trace showed a clear product peak with the expected mass, matching the conclusion derived from urea PAGE gel shift analysis (Figure S7). To confirm that the base being exchanged by the enzyme was the expected G

residue, we tested whether dECH-10ΔC (CTGGATTCTCCTTCCAG) could be modified with TGT and found that the mutant hairpin remained unlabeled (Figure S8).

Having optimized DNA hairpin labeling by TGT, we sought to explore the utility of our newly developed DNA-TAG system. A key benefit of RNA-TAG is the wide substrate tolerance of the enzyme for preQ1 derivatives. We compared the substrate tolerance for DNA modification and found that, similar to RNA, TGT can modify the dECH-10 substrate with a variety of functional preQ1 probes in high yield. We used the established gel shift assay to assess TGT's ability to modify dECH-10 with various probes: preQ1-biotin, TAMRA, Cy5, Alexa Fluor 647, Alexa Fluor 488, and Silicon Rhodamine (Figure S9) and confirmed with high-resolution mass spectrometry (HRMS; see the Supporting Information).

DNA-TAG modification of DNA requires that a TGT recognition hairpin be appended to the DNA of interest. We purchased 59 mer DNA constructs containing 5', 3', and internal hairpins and tested the extent of their labeling. All three hairpin adducts were efficiently labeled using the preQ1-biotin probe (Figure S10). Additionally, we tested the ability of DNA-TAG to insert multiple labels into a single DNA oligonucleotide. Indeed, TGT modification of transcripts bearing two hairpins either in tandem at the same end (5', 5' or 3', 3') or separated at opposite ends (5', 3') was observed via gel shift and confirmed with liquid chromatography–electrospray ionization-time of flight-mass spectrometry (LC-ESI-TOF-MS) (Figure S11). Our experiments demonstrated that DNA-TAG has the potential to be used in a range of applications as it is compatible with several small-molecule probes varying in size and overall charge. Furthermore, DNA-TAG can be used for labeling DNA oligonucleotides at a location of choice, both singly or in tandem.

DNA-TAG could offer a simple, efficient, and cost-effective means for generating fluorescent probes to reliably detect nucleic acid species of interest. We initially explored the suitability of DNA-TAG for generating fluorescent probes for Northern blotting. Northern blotting is a technique used to visualize or quantify an RNA species from a given sample. In short, the sample is separated by gel electrophoresis, transferred to a membrane, and detected via hybridization of reporter-labeled antisense oligonucleotide probes. Conventional Northern blotting utilizes radiolabeled probes because they are highly sensitive and the methods are well established. However, radioactive probes require significant safety measures due to their toxicity and only produce a single output signal, restricting multiplexing capabilities. Recently developed fluorescent blotting strategies, on the other hand, have many advantages over radiolabeling including mitigating safety concerns and allowing for the detection of different targets via the use of orthogonal fluorophores.³⁰ We designed a single fluorescently labeled probe to detect U6 RNA, a small nuclear RNA that plays a catalytic role in the spliceosome.³¹ A 50-mer DNA oligonucleotide U6 antisense probe was appended with a 5' DNA-TAG hairpin (Figure 3A) and labeled with preQ1-Silicon Rhodamine (preQ1-SiR; Figure 1B), a near-infrared fluorophore used in cell imaging.³² Samples containing *in vitro* transcribed (IVT) U6 RNA or crude total RNA from U2OS cellular extracts were separated via urea PAGE in dilution series. The gels were transferred to a nitrocellulose membrane, incubated with the SiR-labeled antisense probe, and the signal was detected via fluorescence scanning. The DNA-TAG

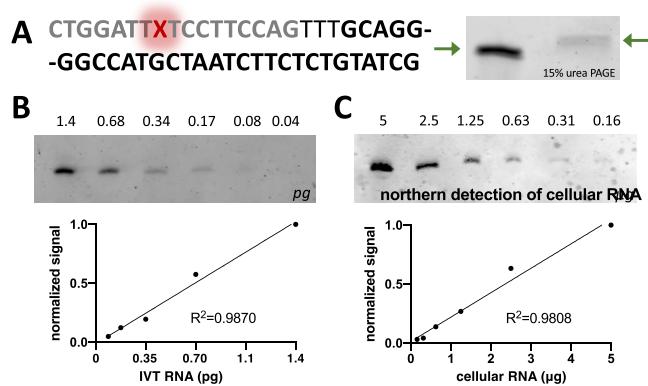


Figure 3. Near-infrared fluorescent Northern blot of U6 RNA using a DNA-TAG generated SiR-antisense probe. (A) SiR-labeled antisense probe. The U6 antisense sequence is black; the DNA-TAG recognition sequence is gray; the inserted preQ1-SiR (X) probe is red. Modification of the antisense probe with preQ1-SiR is evidenced by an upward gel shift; unmodified antisense probe is on the left, and SiR-labeled antisense probe is on the right; note that the gel shift signal is quenched by the SiR dye. (B) Dose-dependent fluorescence detection of picogram (pg) quantities of in vitro transcribed U6 RNA. (C) Dose-dependent fluorescence detection of U6 RNA from microgram (μg) quantities of total cellular RNA extract from U2OS cells.

generated U6 Northern blot probe was successful in detecting the U6 RNA in a dose-dependent manner with a reliable R^2 value when analyzed via linear fit for both IVT and cellular U6 RNA (Figure 3B,C, respectively). The SiR fluorescent Northern blot detection appears to be comparable to previously reported detection limits (Figure 3B).^{30,33} In the future, the sensitivity and specificity of the technique might be improved by using reporter fluorophores with better photo-

physical properties, including multiple fluorescent tags on a single probe, or generating multiple probes and using a tiled detection approach.

Next, we used DNA-TAG to generate probe sets for single-molecule FISH (smFISH), a powerful technique that has been widely used to visualize specific DNA and RNA species in cells using tiled antisense probes labeled with fluorescent reporters.^{34–37} There are a variety of smFISH techniques, all of which rely on the use of fluorescently labeled nucleic acid probes. FISH probes have been generated using naturally occurring nucleic acids such as DNA and RNA and synthetic peptide nucleic acids (PNA).³⁸ Imaging signals can be increased with a variety of strategies, including increasing the number of probes or reporter molecules³⁹ and using turn-on intercalating dyes, such as thiazole orange, to minimize background.¹⁷ A standard smFISH probe set is made up of 30–48 distinct probes which tile the target transcript.⁴⁰ Given the large number of probes required, probe set synthesis can be very costly and needs to be high-yielding. While labeled probe sets can be purchased commercially or generated using TdT along with a limited set of accepted ddUTP-fluorophores, DNA-TAG offers a unique strategy for quickly generating affordable probe sets with a wide small-molecule substrate scope.⁴¹ We tested the ability of DNA-TAG to generate a fluorescent-tiled antisense oligo probe set against metastasis-associated lung adenocarcinoma transcript 1 (MALAT1).

MALAT1 is an abundant long non-coding RNA (lncRNA) found in the nuclei of cells that is implicated in cancers including those affecting the lungs, breasts, prostate, pancreas, glia, and blood.⁴² Using the Stellaris probe design tool, a set of 48 unique antisense oligonucleotide probe sequences was generated to target MALAT1 and appended with DNA-TAG hairpins. We designed two distinct probe sets containing either one or two TGT modification sites. To maintain consistency between the sets, both constructs had two hairpins, with the

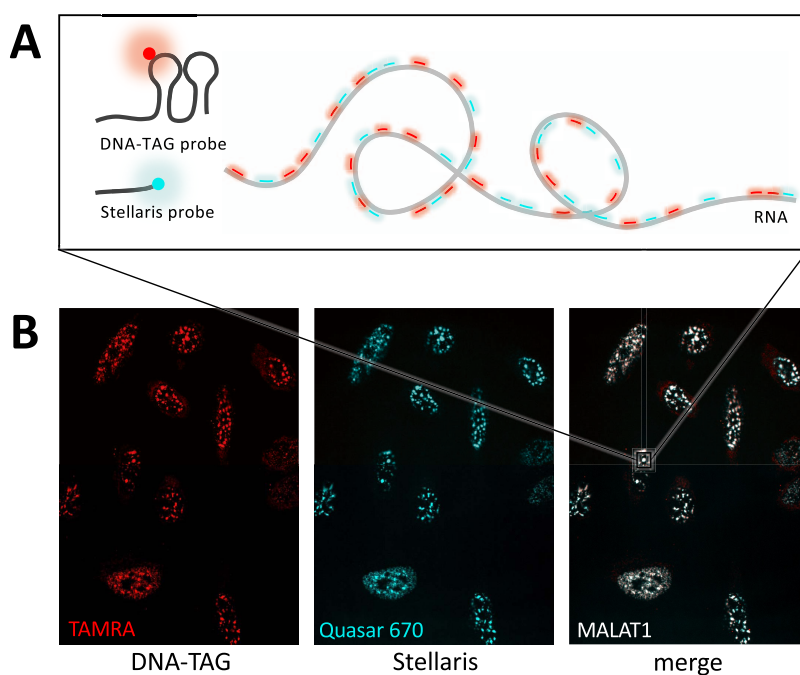


Figure 4. RNA-FISH detection of MALAT1 RNA in U2OS cells using both a DNA-TAG-generated probe set and a commercial probe set from Stellaris. (A) Cartoon depiction of RNA-FISH using the tiled probe sets. The DNA-TAG probe set is modified with preQ1-TAMRA (red) and the Stellaris probe set is labeled with Quasar 670 (teal). (B) Both probe sets hybridize with the MALAT1 puncta in the nuclei of U2OS cells.

single modification set containing one hairpin bearing the “TGT” minimal recognition element and the other containing the inactive “TCT” minimal recognition element (Tables S7 and S8). The probes were individually purchased from a commercial vendor (Integrated DNA Technologies), pooled, and fluorescently labeled using TGT and preQ1-TAMRA. The combined oligonucleotide probe set was then purified by spin filtration. U2OS cells were treated with either a DNA-TAG-labeled TAMRA probe set, an orthogonal Quasar 570-labeled commercial probe set (Stellaris), or a combination of the two. As expected for MALAT1, treatment with all probe sets resulted in punctate stains localized in the nuclei of the cells, with both DNA-TAG labeled probe sets colocalizing with the Stellaris probe set (Figure 4). The signal from the single-label probe set was sufficient for reliably detecting MALAT1 and we did not observe significant improvement in signal from doubly labeled probes (Figures S12 and S13). The latter observation may be due to the well-known self-quenching of TAMRA and further optimization of fluorophore spacing may result in improved signals.⁴³ A single experimental design was shown here; however, our strategy should be amenable with a variety of amplification strategies, for instance, by including additional hairpins on each probe to increase the fluorescence signal. We note that the additional bulk of the TAG hairpin and spacing between probes should be considered when using this technology for FISH.

CONCLUSIONS

Beyond being central to life, nucleic acids have proven to be invaluable tools in both basic and biomedical research. Harnessing the unparalleled programmability afforded by oligonucleotides requires tools to install functional handles or facilitate detection. Furthermore, increasing the availability, versatility, and diversity of oligonucleotide tools is critical to the continued advancement of their use in multiple fields. We have shown that bacterial TGT enzymes, which normally modify tRNA, can be repurposed to site-specifically label DNA. Enzymatic modification using the DNA-TAG methodology facilitates the one-step, site-specific insertion of a variety of functional small molecules into ssDNA substrates of interest for downstream applications. This system is compatible with both internal and terminal insertions of the hairpin sequence into DNAs of interest, is tolerant of tandem modifications, and works with a variety of small-molecule substrates. We believe that future studies will lead to further optimization of the labeling conditions, reducing the time and costs of the reaction. While the possible biological relevance of our findings is unclear, it is worth noting that recent studies have discovered 7-deazaguanine modifications in phage DNA.^{44,45}

DNA-TAG allows an inexpensive and straightforward method by which researchers can quickly label several DNA oligos in a single step, either simultaneously or in parallel, followed by short spin column purification. We have demonstrated that, with a one-time upfront cost of materials needed for small-molecule probe synthesis and protein expression, DNA-TAG offers researchers a versatile, affordable means to generate fluorescent DNA probes of their choice from commercially available DNA oligonucleotides for applications including fluorescent Northern blotting and smFISH.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at <https://pubs.acs.org/doi/10.1021/jacs.3c00861>.

Experimental procedures; general information; synthesis of preQ1 probes with commercially purchased preQ1-C₆NHBoc (Scheme S1); hairpin sequences, *in vitro* transcription template sequences for the rT hairpin, sequences for U6 RNA northern blot, and MALAT1 FISH probe set sequences (Tables S1–S8); TGT modification of histidine hairpin loop mutants and anticodon stem/dECH-10 loop chimeras, time course analysis of TGT modification of dECH-10, confirmation of specific guanine exchange, small molecule substrate scope, single or double labeling of DNA substrates bearing orthogonal hairpins, and RNA FISH detection of MALAT1 (Figures S1–S13); and MS of preQ1 probes and oligo modifications (PDF)

AUTHOR INFORMATION

Corresponding Author

Neal K. Devaraj – Department of Chemistry and Biochemistry, University of California, San Diego, La Jolla, California 92093, United States; orcid.org/0000-0002-8033-9973; Email: ndevaraj@ucsd.edu

Author

Ember M. Tota – Department of Chemistry and Biochemistry, University of California, San Diego, La Jolla, California 92093, United States

Complete contact information is available at:

<https://pubs.acs.org/10.1021/jacs.3c00861>

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

The authors acknowledge the UCSD Molecular Mass Spectrometry Facility for small-molecule high-resolution and oligonucleotide mass spectrometry support. This work was supported by the National Institutes of Health [R35 GM141939]. The content of the information does not necessarily reflect the position or the policy of the government, and no official endorsement should be inferred.

REFERENCES

- (1) Wang, D.-X.; Wang, J.; Wang, Y.-X.; Du, Y.-C.; Huang, Y.; Tang, A.-N.; Cui, Y.-X.; Kong, D.-M. DNA Nanostructure-Based Nucleic Acid Probes: Construction and Biological Applications. *Chem. Sci.* **2021**, *12*, 7602–7622.
- (2) Li, L.; Xu, S.; Yan, H.; Li, X.; Yazd, H. S.; Li, X.; Huang, T.; Cui, C. H.; Jiang, J.; Tan, W. Nucleic Acid Aptamers for Molecular Diagnostics and Therapeutics: Advances and Perspectives. *Angew. Chem., Int. Ed.* **2021**, *60*, 2221–2231.
- (3) Le, P.; Ahmed, N.; Yeo, G. W. Illuminating RNA Biology through Imaging. *Nat. Cell Biol.* **2022**, *24*, 815–824.
- (4) Holstein, J. M.; Schulz, D.; Rentmeister, A. Bioorthogonal Site-Specific Labeling of the 5'-Cap Structure in Eukaryotic MRNAs. *Chem. Commun.* **2014**, *50*, 4478–4481.
- (5) Holstein, J. M.; Stummer, D.; Rentmeister, A. Enzymatic Modification of 5'-Capped RNA with a 4-Vinylbenzyl Group Provides a Platform for Photoclick and Inverse Electron-Demand Diels–Alder Reaction. *Chem. Sci.* **2015**, *6*, 1362–1369.

- (6) Gillingham, D.; Shahid, R. Catalysts for RNA and DNA Modification. *Curr. Opin. Chem. Biol.* **2015**, *25*, 110–114.
- (7) Büttner, L.; Javadi-Zarnaghi, F.; Höbartner, C. Site-Specific Labeling of RNA at Internal Ribose Hydroxyl Groups: Terbium-Assisted Deoxyribozymes at Work. *J. Am. Chem. Soc.* **2014**, *136*, 8131–8137.
- (8) Holstein, J. M.; Rentmeister, A. Current Covalent Modification Methods for Detecting RNA in Fixed and Living Cells. *Methods* **2016**, *98*, 18–25.
- (9) Holstein, J. M.; Stummer, D.; Rentmeister, A. Engineering Giardia Lamblia Trimethylguanosine Synthase (GlaTgs2) to Transfer Non-Natural Modifications to the RNA 5'-Cap. *Protein Eng., Des. Sel.* **2014**, *28*, 179–186.
- (10) Muttach, F.; Rentmeister, A. A Biocatalytic Cascade for Versatile One-Pot Modification of mRNA Starting from Methionine Analogues. *Angew. Chem.* **2016**, *128*, 1951–1954.
- (11) Flood, D. T.; Knouse, K. W.; Vantourout, J. C.; Kitamura, S.; Sanchez, B. B.; Sturgell, E. J.; Chen, J. S.; Wolan, D. W.; Baran, P. S.; Dawson, P. E. Synthetic Elaboration of Native DNA by RASS (SENDR). *ACS Cent. Sci.* **2020**, *6*, 1789–1799.
- (12) Ono, K. Inhibitory Effects of Various 2', 3'-Dideoxynucleoside 5'-Triphosphates on the Utilization of 2'-Deoxynucleoside 5'-Triphosphates by Terminal Deoxynucleotidyltransferase from Calf Thymus. *Biochim. Biophys. Acta, Gene Struct. Expression* **1990**, *1049*, 15–20.
- (13) Zearfoss, N. R.; Ryder, S. P. End-Labeling Oligonucleotides with Chemical Tags After Synthesis. In *Recombinant and In Vitro RNA Synthesis*; Springer, 2012; Vol. 941, pp 181–193.
- (14) Lukinavičius, G.; Lapienė, V.; Staševskij, Z.; Dalhoff, C.; Weinhold, E.; Klimašauskas, S. Targeted Labeling of DNA by Methyltransferase-Directed Transfer of Activated Groups (MTAG). *J. Am. Chem. Soc.* **2007**, *129*, 2758–2759.
- (15) Flood, D. T.; Asai, S.; Zhang, X.; Wang, J.; Yoon, L.; Adams, Z. C.; Dillingham, B. C.; Sanchez, B. B.; Vantourout, J. C.; Flanagan, M. E.; Piotrowski, D. W.; Richardson, P.; Green, S. A.; Shenvi, R. A.; Chen, J. S.; Baran, P. S.; Dawson, P. E. Expanding Reactivity in DNA-Encoded Library Synthesis via Reversible Binding of DNA to an Inert Quaternary Ammonium Support. *J. Am. Chem. Soc.* **2019**, *141*, 9998–10006.
- (16) Flood, D. T.; Kingston, C.; Vantourout, J. C.; Dawson, P. E.; Baran, P. S. DNA Encoded Libraries: A Visitor's Guide. *Isr. J. Chem.* **2020**, *60*, 268–280.
- (17) Hövelmann, F.; Gaspar, I.; Loibl, S.; Ermilov, E. A.; Röder, B.; Wengel, J.; Ephrussi, A.; Seitz, O. Brightness through Local Constrain-LNA-Enhanced FIT Hybridization Probes for In Vivo Ribonucleotide Particle Tracking. *Angew. Chem., Int. Ed.* **2014**, *53*, 11370–11375.
- (18) Alexander, S. C.; Busby, K. N.; Cole, C. M.; Zhou, C. Y.; Devaraj, N. K. Site-Specific Covalent Labeling of RNA by Enzymatic Transglycosylation. *J. Am. Chem. Soc.* **2015**, *137*, 12756–12759.
- (19) Busby, K. N.; Devaraj, N. K. Enzymatic Covalent Labeling of RNA with RNA Transglycosylation at Guanosine (RNA-TAG). In *Methods in Enzymology* 2020; Vol. 641, pp 373–399.
- (20) Stengl, B.; Reuter, K.; Klebe, G. Mechanism and Substrate Specificity of TRNA-Guanine Transglycosylases (TGTs): TRNA-Modifying Enzymes from the Three Different Kingdoms of Life Share a Common Catalytic Mechanism. *ChemBioChem* **2005**, *6*, 1926–1939.
- (21) Nakanishi, S.; Ueda, T.; Hori, H.; Yamazaki, N.; Watanabe, N.; Okada, K. A UGU Sequence in the Anticodon Loop Is a Minimum Requirement for Recognition by *Escherichia coli* TRNA-Guanine Transglycosylase. *J. Biol. Chem.* **1994**, *269*, 32221–32225.
- (22) Ehret, F.; Zhou, C. Y.; Alexander, S. C.; Zhang, D.; Devaraj, N. K. Site-Specific Covalent Conjugation of Modified mRNA by TRNA Guanine Transglycosylase. *Mol. Pharm* **2018**, *15*, 737–742.
- (23) Busby, K. N.; Fulzele, A.; Zhang, D.; Bennett, E. J.; Devaraj, N. K. Enzymatic RNA Biotinylation for Affinity Purification and Identification of RNA – Protein Interactions. *ACS Chem. Biol.* **2020**, *15*, 2247–2258.
- (24) Zhang, D.; Jin, S.; Piao, X.; Devaraj, N. K. Multiplexed Photoactivation of mRNA with Single-Cell Resolution. *ACS Chem. Biol.* **2020**, *15*, 1773–1779.
- (25) Zhang, D.; Zhou, C. Y.; Busby, K. N.; Alexander, S. C.; Devaraj, N. K. Light-Activated Control of Translation by Enzymatic Covalent mRNA Labeling. *Angew. Chem., Int. Ed.* **2018**, *57*, 2822–2826.
- (26) Zhou, C. Y.; Alexander, S. C.; Devaraj, N. K. Fluorescent Turn-on Probes for Wash-Free mRNA Imaging via Covalent Site-Specific Enzymatic Labeling. *Chem. Sci.* **2017**, *8*, 7169–7173.
- (27) Zhang, D.; Liu, L.; Jin, S.; Tota, E.; Li, Z.; Piao, X.; Zhang, X.; Fu, X.-D.; K Devaraj, N. Site-Specific and Enzymatic Cross-Linking of SgRNA Enables Wavelength-Selectable Photoactivated Control of CRISPR Gene Editing. *J. Am. Chem. Soc.* **2022**, *144*, 4487–4495.
- (28) Nonekowsky, S. T.; Kung, F. L.; Garcia, G. A. The *Escherichia coli* TRNA-Guanine Transglycosylase Can Recognize and Modify DNA*. *J. Biol. Chem.* **2002**, *277*, 7178–7182.
- (29) Curnow, A. W.; Garcia, G. A. TRNA-Guanine Transglycosylase from *Escherichia coli*: Minimal TRNA Structure and Sequence Requirements for Recognition. *J. Biol. Chem.* **1995**, *270*, 17264–17267.
- (30) Miller, B. R.; Wei, T.; Fields, C. J.; Sheng, P.; Xie, M. Near-Infrared Fluorescent Northern Blot. *RNA* **2018**, *24*, 1871–1877.
- (31) Didychuk, A. L.; Butcher, S. E.; Brow, D. A. The Life of U6 Small Nuclear RNA, from Cradle to Grave. *RNA* **2018**, *24*, 437–460.
- (32) Lukinavičius, G.; Umezawa, K.; Olivier, N.; Honigmann, A.; Yang, G.; Plass, T.; Mueller, V.; Reymond, L.; Corrêa, I. R., Jr.; Luo, Z.-G.; Schultz, C.; Lemke, E. A.; Heppenstall, P.; Eggeling, C.; Manley, S.; Johnsson, K. A Near-Infrared Fluorophore for Live-Cell Super-Resolution Microscopy of Cellular Proteins. *Nat. Chem.* **2013**, *5*, 132–139.
- (33) Iwatate, R. J.; Kamiya, M.; Umezawa, K.; Kashima, H.; Nakadate, M.; Kojima, R.; Urano, Y. Silicon Rhodamine-Based Near-Infrared Fluorescent Probe for γ -Glutamyltransferase. *Bioconjugate Chem.* **2018**, *29*, 241–244.
- (34) Femino, A. M.; Fay, F. S.; Fogarty, K.; Singer, R. H. Visualization of Single RNA Transcripts in Situ. *Science* **1998**, *280*, 585–590.
- (35) Raj, A.; van den Bogaard, P.; Rifkin, S. A.; van Oudenaarden, A.; Tyagi, S. Imaging Individual mRNA Molecules Using Multiple Singly Labeled Probes. *Nat. Methods* **2008**, *5*, 877–879.
- (36) Taniguchi, Y.; Choi, P. J.; Li, G. W.; Chen, H.; Babu, M.; Hearn, J.; Emili, A.; Sunney Xie, X. Quantifying *E. coli* Proteome and Transcriptome with Single-Molecule Sensitivity in Single Cells. *Science* **2010**, *329*, 533–538.
- (37) Paré, A.; Lemons, D.; Kosman, D.; Beaver, W.; Freund, Y.; McGinnis, W. Visualization of individual Scr mRNAs during *Drosophila* embryogenesis yields evidence for transcriptional bursting. *Curr. Biol.* **2009**, *19*, 2037–2042.
- (38) Patel, N.; Miller, D.; Relhan, N.; Flynn, H. W. Peptide Nucleic Acid-Fluorescence In Situ Hybridization for Detection of Staphylococci From Endophthalmitis Isolates: A Proof-of-Concept Study. *Invest. Ophthalmol. Vis. Sci.* **2017**, *58*, 4027–4029.
- (39) Kishi, J. Y.; Lapan, S. W.; Beliveau, B. J.; West, E. R.; Zhu, A.; Sasaki, H. M.; Saka, S. K.; Wang, Y.; Cepko, C. L.; Yin, P. SABER Amplifies FISH: Enhanced Multiplexed Imaging of RNA and DNA in Cells and Tissues. *Nat. Methods* **2019**, *16*, 533–544.
- (40) Orjalo, A.; Johansson, H. E.; Ruth, J. L. Stellaris Fluorescence in Situ Hybridization (FISH) Probes: A Powerful Tool for mRNA Detection. *Nat. Methods* **2011**, *8*, 1–2.
- (41) Gáspár, I.; Wippich, F.; Ephrussi, A. Terminal Deoxynucleotidyl Transferase Mediated Production of Labeled Probes for Single-Molecule FISH or RNA Capture. *Bio-Protocol* **2018**, *8*, No. e2750.
- (42) Zhao, M.; Wang, S.; Li, Q.; Ji, Q.; Guo, P.; Liu, X. Malat1: A Long Non-Coding RNA Highly Associated with Human Cancers (Review). *Oncol. Lett.* **2018**, *16*, 19–26.
- (43) Shiba, A.; Kinoshita-Kikuta, E.; Kinoshita, E.; Koike, T. TAMRA/TAMRA Fluorescence Quenching Systems for the Activity Assay of Alkaline Phosphatase. *Sensors* **2017**, *17*, No. 1877.

(44) Hutinet, G.; Kot, W.; Cui, L.; Hillebrand, R.; Balamkundu, S.; Gnanakalai, S.; Neelakandan, R.; Carstens, A. B.; Lui, C. F.; Tremblay, D.; Jacobs-Sera, D.; Sassanfar, M.; Lee, Y.-J.; Weigele, P.; Moineau, S.; Hatfull, G. F.; Dedon, P. C.; Hansen, L. H.; De Crécy-Lagard, V. 7-Deazaguanine Modifications Protect Phage DNA from Host Restriction Systems. *Nat. Commun.* **2019**, *10*, No. 5442.

(45) Thiaville, J. J.; Kellner, S. M.; Yuan, Y.; Hutinet, G.; Thiaville, P. C.; Jumpathong, W.; Mohapatra, S.; Brochier-Armanet, C.; Letarov, A. V.; Hillebrand, R.; Malik, C. K.; Rizzo, C. J.; Dedon, P. C.; Crécy-Lagard, V. Novel Genomic Island Modifies DNA with 7-Deazaguanine Derivatives. *Proc. Natl. Acad. Sci. U.S.A.* **2016**, *113*, E1452–E1459.