DNA Polymerase-Mediated Synthesis of Unbiased Threose Nucleic Acid (TNA) Polymers Requires 7-Deazaguanine To Suppress G:G Mispairing during TNA Transcription

ARTICLE *in* JOURNAL OF THE AMERICAN CHEMICAL SOCIETY · MARCH 2015

Impact Factor: 12.11 · DOI: 10.1021/ja511481n · Source: PubMed

7 AUTHORS, INCLUDING:

Matthew Dunn
University of California, Irvine
6 PUBLICATIONS **17** CITATIONS
SEE PROFILE

Andrew Larsen
Arizona State University
9 PUBLICATIONS **26** CITATIONS
SEE PROFILE

Walter J Zahurancik
The Ohio State University
7 PUBLICATIONS **27** CITATIONS
SEE PROFILE

JACS
JOURNAL OF THE AMERICAN CHEMICAL SOCIETY

Communication

pubs.acs.org/JACS

DNA Polymerase-Mediated Synthesis of Unbiased Threose Nucleic Acid (TNA) Polymers Requires 7-Deazaguanine To Suppress G:G Mispairing during TNA Transcription

Matthew R. Dunn,[†,§] Andrew C. Larsen,[§] Walter J. Zahurancik,[∥] Nour Eddine Fahmi,[§] Madeline Meyers,[‡,§] Zucai Suo,[∥] and John C. Chaput*[,‡,§]

[†]School of Life Sciences, [‡]Department of Chemistry and Biochemistry, and [§]The Biodesign Institute, Arizona State University, Tempe, Arizona 85287-5301, United States

[∥]Department of Chemistry and Biochemistry, The Ohio State University, Columbus, Ohio 43210, United States

Supporting Information

ABSTRACT: Threose nucleic acid (TNA) is an unnatural genetic polymer capable of undergoing Darwinian evolution to generate folded molecules with ligand-binding activity. This property, coupled with a nuclease-resistant backbone, makes TNA an attractive candidate for future applications in biotechnology. Previously, we have shown that an engineered form of the *Archaean* replicative DNA polymerase 9°N, known commercially as Therminator DNA polymerase, can copy a three-letter genetic alphabet (A,T,C) from DNA into TNA. However, our ability to transcribe four-nucleotide libraries has been limited by chain termination events that prevent the synthesis of full-length TNA products. Here, we show that chain termination is caused by tG:dG mispairing in the enzyme active site. We demonstrate that the unnatural base analogue 7-deazaguanine (7dG) will suppress tGTP misincorporation by inhibiting the formation of Hoogsteen tG:dG base pairs. DNA templates that contain 7dG in place of natural dG residues replicate with high efficiency and >99% overall fidelity. Pre-steady-state kinetic measurements indicate that the rate of tCTP incorporation is 5-fold higher opposite 7dG than dG and only slightly lower than dCTP incorporation opposite either 7dG or dG. These results provide a chemical solution to the problem of how to synthesize large, unbiased pools of TNA molecules by polymerase-mediated synthesis.

We have developed a polymerase-mediated replication system that makes it possible to copy genetic information back and forth between deoxyribose nucleic acid (DNA) and threose nucleic acid (TNA) (Figure 1).[1] Under appropriate conditions, an engineered form of the *Archaean* replicative DNA polymerase 9°N, known commercially as Therminator DNA polymerase (New England BioLabs, Inc.), can copy sequence-defined DNA templates into TNA through the sequential addition of TNA nucleotide triphosphates.[2,3] TNA templates purified by denaturing polyacrylamide gel electrophoresis (PAGE) can be reverse-transcribed back into cDNA using Superscript II (SSII), which is a highly engineered form of a reverse transcriptase isolated from the Maloney murine leukemia virus.[4] Amplification of the cDNA using the polymerase chain reaction (PCR) followed by strand separation allows the process to be performed iteratively.

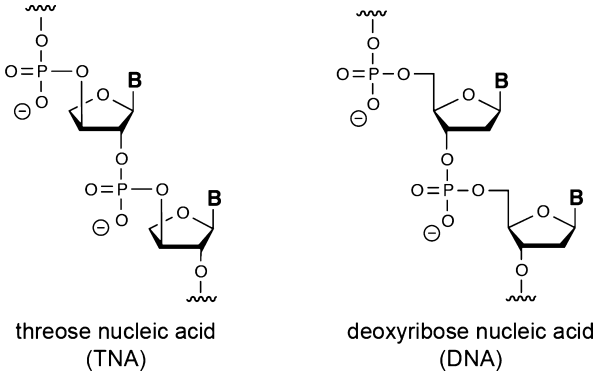


Figure 1. Molecular structure of TNA and DNA. Constitutional structure for the linearized backbone of α-L-threofuranosyl-(3′-2′) nucleic acid, TNA (left), and DNA (right). TNA is an unnatural genetic polymer composed of repeating α-L-threose sugars that are vicinally connected by 2′,3′-phosphodiester linkages.

When transcription and reverse transcription are performed on DNA templates of a known sequence, the overall fidelity of TNA replication can be determined by sequencing the cDNA product. DNA libraries composed of a three-letter genetic alphabet (A,T,C) have been shown to transcribe with high efficiency, and individual sequences replicate with >99% fidelity.[1] This level of fidelity is remarkable, considering the structural differences between DNA and TNA. However, attempts to synthesize TNA libraries that contain a four-letter genetic alphabet (A,C,T,G) have failed due to chain termination events that inhibit TNA synthesis on unbiased DNA templates.[2] Investigation into this problem revealed that DNA templates with low numbers of isolated dG residues could be transcribed into full-length TNA, but the cDNA obtained after reverse transcription contained a high level of G→C mutations.[1] The low fidelity of TNA replication on four-nucleotide templates can be overcome by replacing tCTP in the transcription mixture with dCTP; however, this change leads to a mixed-backbone TNA−DNA heteropolymer that may not be

Received: November 7, 2014

DOI: 10.1021/ja511481n
J. Am. Chem. Soc. XXXX, XXX, XXX−XXX

suitable for some applications.[1] This problem has created a need for new replication strategies that can be used to generate large, unbiased pools of TNA for *in vitro* selection.

Recognizing the limitations of G-rich templates, we decided to evaluate the ability of Therminator DNA polymerase to read through G-repeats in a DNA template. We used a primer extension assay in which the DNA template was engineered to contain a central region of 0−3 consecutive dG residues that are flanked by an arbitrary sequence composed of A, C, and T residues (Figure 2A). Therminator was challenged to extend a

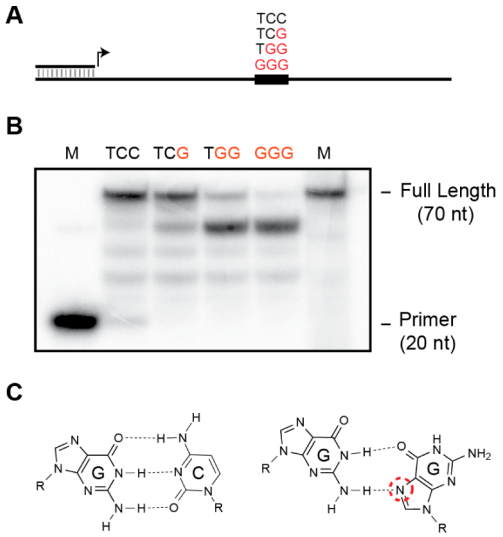


Figure 2. Sequence specificity of Therminator DNA polymerase. (A) Primer extension assay used to evaluate TNA synthesis on dG-containing DNA templates. (B) Primer extension results for DNA templates that contain consecutive G-residues. (C) Molecular structure of a normal G:C Watson−Crick base pair and a G:G Hoogsteen base pair. The analogue of 7-deazaguanine replaces the N-7 nitrogen with a methylene group (red circle).

DNA primer annealed to a DNA template using TNA triphosphates (tNTPs) that were obtained by chemical synthesis.[5] If the polymerase is able to extend the primer to the end of the DNA template, it will produce an elongated product that is extended by 50 TNA residues. However, if the polymerase is unable to read through the G residues, it will produce a truncated product that is easily detected by denaturing PAGE.

Analysis of the resulting primer extension assay products indicates that Therminator is able to generate full-length TNA product when 0 or 1 dG residue is present in the DNA template (Figure 2B). However, the polymerase stalls on templates that contain 2 or more consecutive dG residues, indicating that G-repeats inhibit the transcription of TNA polymers on DNA templates. This result, in conjunction with our previous observation that TNA replication on dG-containing templates leads to an abundance of G→C mutations in the cDNA product,[1] indicates that Therminator misincorporates tGTP monomers opposite dG. While this could be caused by deformation of the duplex in the active site or partial stacking of G in the absence of hydrogen bonds, we postulate that tGTP forms a tG:dG Hoogsteen base pair in the enzyme active site. Hoogsteen base pairing between guanine residues occurs when a guanine base in one strand adopts a *syn* conformation relative to the sugar moiety, which allows the second guanine residue to form two hydrogen bonds on the major groove face of the first guanine base (Figure 2C).[6] Crystallographic analysis of a Watson−Crick DNA duplex containing an internal G:G base pair reveals that a G:G Hoogsteen base pair is structurally similar to a G:C base pair.[7] This structural similarity is one reason why many polymerases (natural and engineered) struggle to read through templates that are rich in G nucleotides.[8]

To test the hypothesis that G:G Hoogsteen base pairing was responsible for the chain termination events that occur when TNA polymers are synthesized on unbiased DNA libraries (L3), we constructed two pools of single-stranded DNA molecules. The first pool contained an unbiased mixture of all four natural nucleobases (A,C,T,G). The second pool was identical in length and composition to the first pool but contained the unnatural base analogue 7dG in place of natural dG. The pools were constructed by PCR using appropriate sets of nucleotide triphosphates and a PEG-modified PCR primer to enable separation of the strands by denaturing PAGE.

Preliminary analysis of the DNA products produced by PCR suggested that the 7dG reaction generated very little product when compared to the all-natural dNTPs reaction (Figure S1). Although we initially interpreted this result to mean that 7dG was a poor substrate for deep vent (exo-) DNA polymerase, further analysis revealed that 7dG quenches the fluorescence of UV-excitable dyes. In fact, only when the DNA was labeled with a fluorescent primer or stained with SYTO 60, a dye that excites in the IR range, were the DNA products from both reactions observed in equal abundance. This unanticipated phenomenon, though not entirely unprecedented, is a reminder of how the chemical and physical properties of modified DNA can differ from those of natural genetic polymers.[9]

Using the dG- and 7dG-containing templates, we performed a primer extension assay for 6 h at 55 °C to compare the efficiency of TNA synthesis on the two pools of DNA molecules. As illustrated in Figure 3, Therminator DNA

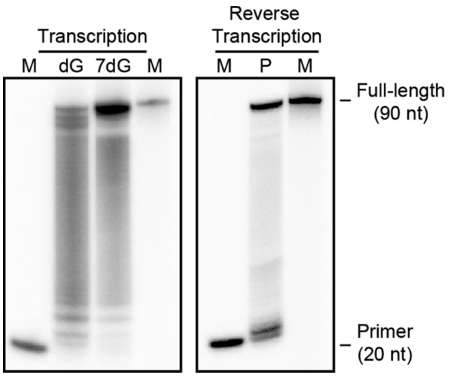


Figure 3. Transcription and reverse transcription of a four-nucleotide TNA library. PAGE analysis reveals that TNA transcription by Therminator DNA polymerase (left panel) and reverse transcription by Superscript II (right panel) proceed with high efficiency on DNA templates that contain 7-deazaguanine in place of the natural guanine base.

polymerase was unable to copy natural DNA templates into TNA, as evidenced by the absence of any significant full-length product and the high degree of banding on the gel. However, efficient synthesis was observed for the primer extension reaction performed on the pool of 7dG-containing DNA templates, which were designed to promote TNA synthesis by inhibiting the formation of G:G mispairs. In this case, the

DOI: 10.1021/ja511481n
J. Am. Chem. Soc. XXXX, XXX, XXX−XXX

polymerase was able to extend the DNA primer by 70 sequential residues to produce a significant amount of full-length product. This result provides the first example of an unbiased TNA library constructed by polymerase-mediated primer extension. It also confirms the hypothesis that G:G mispairing in the enzyme active site was responsible for polymerase stalling on natural DNA templates.

Because a complete replication cycle requires copying the pool of TNA strands back into DNA, we tested the ability for SSII to reverse-transcribe the TNA back into DNA. Full-length TNA strands isolated from the transcription of 7dG-containing templates were purified by denaturing PAGE and incubated with SSII for 2 h at 55 °C. Consistent with our previous analysis of a three-nucleotide library,[1] primer extension analysis revealed that the four-nucleotide library was an efficient template for reverse transcription (Figure 3).

To ensure that Therminator and SSII produced an accurate copy of each template, we measured the fidelity of TNA replication by sequencing the cDNA product. This assay measures the overall fidelity of replication, which includes the transcription of DNA into TNA, followed by the reverse transcription of TNA back into DNA (Figure 4). Several controls were used to eliminate possible contamination by the starting DNA template. First, we showed that PAGE-purified TNA could not be amplified by PCR unless it was reverse-transcribed into cDNA (Figure S2). Second, the primer–template complex used for TNA synthesis was designed to contain a T:T mismatch that would produce a T:A watermark in the cDNA product. Only those sequences that were transcribed into TNA and reverse-transcribed back into cDNA would carry the watermark. With these controls in place, we sequenced 850 nucleotide positions from cDNA isolated after TNA transcription and reverse transcription. Analysis of the DNA sequences yielded an overall fidelity of replication of >99% (Figure S3). This level of fidelity is similar to what we observed for the replication of three-nucleotide templates and supports the robust transcription of TNA polymers on 7dG-containing DNA templates.[1]

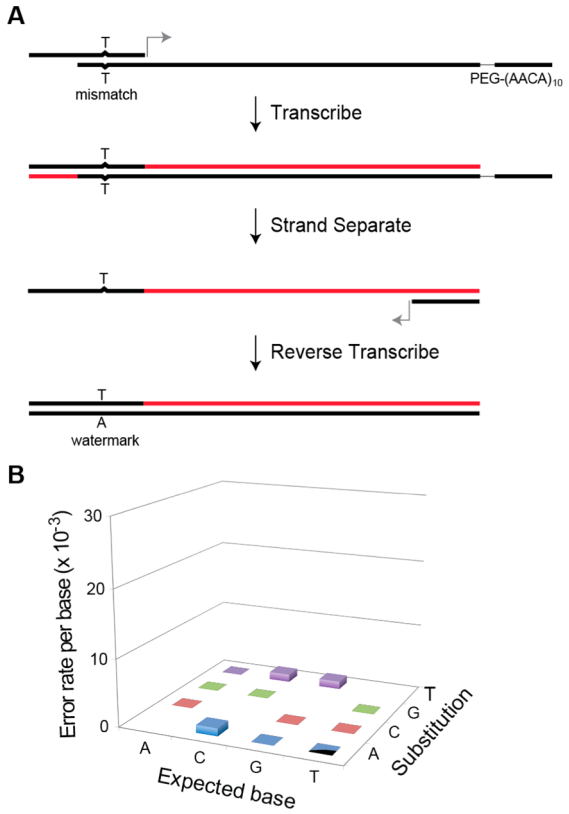


Figure 4. Fidelity of TNA replication using 7-deazaguanine. (A) Schematic representation of the transcription and reverse transcription process used to evaluate the fidelity of TNA replication. DNA is shown in black, and TNA is shown in red. The primer–template complex contains a T:T mismatch, which produces a T-to-A transversion in the cDNA strand. The transversion represents a watermark to ensure that the sequenced DNA was produced by TNA transcription and reverse transcription. (B) The mutation profile reveals that TNA transcription and reverse transcription proceed with an error rate of 3.8×10^{-3} (>99% fidelity).

To determine if 7dG overcomes polymerase stalling by increasing the rate of tCTP incorporation, we used pre-steady-state kinetics to measure the rate constants for incorporation of dCTP or tCTP opposite a templating dG or 7dG residue. Briefly, a pre-incubated solution of Therminator DNA polymerase (100 nM) and a 5′-radiolabeled DNA primer–template substrate (20 nM) in a reaction buffer at 37 °C was rapidly mixed with dCTP or tCTP (100 μM) for various times before being quenched with EDTA. The data were fit to a single-exponential equation, $[\text{product}] = A[1 - \exp(-k_{obs}t)]$, to yield k_{obs}, the observed rate constant of nucleotide incorporation (Table 1). Notably, the rate constant of correct dCTP or incorrect dGTP incorporation was not significantly affected by substituting the templating dG with 7dG. This indicates that the unnatural templating nucleotide 7dG did not significantly alter the interactions between the polymerase–DNA complex and incoming nucleotide at the enzyme active site. In contrast, the rate constant of tCTP incorporation increased by 4.2-fold opposite the templating 7dG relative to dG and is comparable to the rate constant for normal dCTP incorporation opposite dG (Table 1). Thus, tCTP is able to pair with 7dG nearly as well as a natural dCTP pairs with dG. Furthermore, opposite the templating 7dG, the rate constant for tGTP misincorporation is 6.1-fold lower than that for correct tCTP incorporation, which is comparable the 10.8-fold difference between dCTP:dG and dGTP:dG incorporations. Thus, the switching of dG to 7dG in the DNA template not only solves the polymerase stalling problem but also maintains relatively high DNA→TNA transcription fidelity.

Table 1. Pre-Steady-State Kinetic Parameters for Therminator-Mediated Single-Nucleotide Extension

5′-d-CGCAGCCGTCCAACCAACTCA
3′-d-GCGTCGGCAGGTTGGTTGAGT-**X**-GCAGCTAGGTTACGGCAGG

NTP	X	k_{obs} (s^{-1})
tCTP	dG	0.9 ± 0.2
tCTP	7dG	3.8 ± 0.5
tGTP	7dG	0.62 ± 0.09
tGTP	dG	0.68 ± 0.06
dCTP	dG	4.2 ± 0.9
dCTP	7dG	5.0 ± 0.7
dGTP	dG	0.39 ± 0.01
dGTP	7dG	0.40 ± 0.06

The ability to code and decode sequence-defined genetic polymers, like TNA, provides access to many non-biological applications that will benefit materials science, nanotechnology, and molecular medicine.[10] For example, TNA is highly resistant

DOI: 10.1021/ja511481n
J. Am. Chem. Soc. XXXX, XXX, XXX–XXX

to nuclease degradation, making it a stable scaffold for future diagnostic and therapeutic applications. In addition, because TNA has the ability to evolve in response to imposed selection constraints, it could also be used to enhance our understanding of why nature chose RNA as the molecular basis of life's genetic material.[11] Previous work in this area has been limited by the absence of polymerases that could be used to study alternative chemistries of life. As this paradigm is now changing, we may soon discover that many different types of genetic polymers exhibit the characteristic signatures of heredity and evolution—two important hallmarks of life.[12,2a]

In summary, we have developed a method for replicating unbiased pools of TNA polymers that uses 7-deazaguanine in the DNA template to suppress G:G mispairing during TNA transcription with Therminator DNA polymerase. Characterization of the replication cycle reveals that the TNA synthesis proceeds with high efficiency and high overall sequence fidelity. We suggest that this approach could be used to explore the functional properties of TNA by *in vitro* selection.

■ ASSOCIATED CONTENT

Ⓢ Supporting Information

Additional information and experimental methods. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

*john.chaput@asu.edu

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

We thank other members of the Chaput laboratory for helpful comments and suggestions. This work was supported by National Science Foundation grants 1304583 to J.C.C. and MCB-0960961 to Z.S. W.J.Z. is supported by graduate training grant T32 GM008512 from National Institutes of Health.

■ REFERENCES

(1) Yu, H.; Zhang, S.; Dunn, M. R.; Chaput, J. C. *J. Am. Chem. Soc.* **2013**, *135*, 3583.

(2) For examples of polymerase-mediated TNA synthesis, see: (a) Yu, H.; Zhang, S.; Chaput, J. C. *Nat. Chem.* **2012**, *4*, 183. (b) Horhota, A.; Zou, K.; Ichida, J. K.; Yu, B.; McLaughlin, L. W.; Szostak, J. W.; Chaput, J. C. *J. Am. Chem. Soc.* **2005**, *127*, 7427. (c) Chaput, J. C.; Szostak, J. W. *J. Am. Chem. Soc.* **2003**, *125*, 9274. (d) Kempeneers, V.; Vastmans, K.; Rozenski, J.; Herdewijn, P. *Nucleic Acids Res.* **2003**, *31*, 6221. (e) Ichida, J. K.; Zou, K.; Horhota, A.; Yu, B.; McLaughlin, L. W.; Szostak, J. W. *J. Am. Chem. Soc.* **2005**, *127*, 2802. (f) Ichida, J. K.; Horhota, A.; Zou, K.; McLaughlin, L. W.; Szostak, J. W. *Nucleic Acids Res.* **2005**, *33*, 5219. (g) Zhang, S.; Chaput, J. C. *Bioorg. Med. Chem. Lett.* **2013**, *23*, 1447.

(3) For papers on Therminator DNA polymerase, see: (a) Gardner, A. F.; Jack, W. E. *Nucleic Acids Res.* **2002**, *30*, 605. (b) Gardner, A. F.; Jack, W. E. *Nucleic Acids Res.* **1999**, *27*, 2545.

(4) Arezi, B.; Hogrefe, H. *Nucleic Acids Res.* **2009**, *37*, 473.

(5) (a) Zhang, S.; Chaput, J. C. *Curr. Protein Nucleic Acid Chem.* **2012**, *4*, No. 4.51.1. (b) Zhang, S.; Yu, H.; Chaput, J. C. *Curr. Protein Nucleic Acid Chem.* **2013**, *4*, No. 4.54.1.

(6) Cantor, C. R. *Principles of Nucleic Acid Chemistry*; Springer-Verlag: New York, 1984; p 120.

(7) Skelly, J. V.; Edwards, K. J.; Jenkins, T. C.; Neidle, S. *Proc. Natl. Acad. Sci. U.S.A.* **1993**, *90*, 804.

(8) Staiger, N.; Marx, A. *ChemBioChem* **2010**, *11*, 1963.

(9) (a) Kelley, S. O.; Barton, J. K. *Chem. Biol.* **1998**, *5*, 413. (b) Innis, M. A.; Myambo, K. B.; Gelfand, D. H.; Brow, M. A. D. *Proc. Natl. Acad. Sci. U.S.A.* **1988**, *85*, 9436.

(10) For reviews, see: (a) Lutz, J.-F.; Ouchi, M.; Liu, D. R.; Sawamoto, M. *Science* **2013**, *341*, 628. (b) Chaput, J. C.; Yu, H.; Zhang, S. *Chem. Biol.* **2012**, *19*, 1360.

(11) (a) Eschenmoser, A. E. *Science* **1999**, *284*, 2118. (b) Joyce, G. F. *Nature* **2002**, *418*, 214.

(12) Pinheiro, V. B.; Taylor, A. I.; Cozens, C.; Abramov, M.; Renders, M.; Zhang, S.; Chaput, J. C.; Wengel, J.; Peak-Chew, S.-Y.; McLaughlin, S. H.; Herewihjn, P.; Holliger, P.; et al. *Science* **2012**, *336*, 341.

DOI: 10.1021/ja511481n
J. Am. Chem. Soc. XXXX, XXX, XXX−XXX

Supplemental Information

Therminator-Mediated Synthesis of Unbiased TNA Polymers Requires 7-Deazaguanine to Suppress G-G Mispairing during TNA Transcription

Matthew R. Dunn[1,3], Andrew C. Larsen[3], Walter J. Zahurancik[4], Nour Eddine Fahmi[3], Madeline Meyers[2,3], Zucai Suo[4], John C. Chaput[2,3*]

[1]School of Life Sciences, Arizona State University, Tempe, AZ 85287. [2]Department of Chemistry and Biochemistry, Arizona State University, Tempe, AZ 85287. [3]The Biodesign Institute at Arizona State University, Tempe, AZ 85287. [4]Department of Chemistry and Biochemistry, The Ohio State University, Columbus, OH 43210, USA

* To whom correspondence should be addressed. E-mail: john.chaput@asu.edu

Materials and Methods

General information.

DNA oligonucleotides were purchased from Integrated DNA Technologies (Coralville, IA), purified by denaturing polyacrylamide gel electrophoresis, electroeluted, concentrated by ethanol precipitation, and quantified by UV absorbance. TNA triphosphates (tNTPs) were obtained through a multistep chemical synthesis as previously reported.[1,2] 7-deaza-dGTP was purchased from Roche (Mannheim, Germany). dNTPs were purchased from Sigma (St. Louis, MO). [γ-^{32}P]ATP was purchased from Perkin Elmer (Walktham, MA). 7-deaza-dG-CE Phosphoramidite was purchased from Glen Research (Sterling, VA). The enzymes T4 polynucleotide kinase, Therminator™ DNA polymerase, and exonuclease-deficient Deep Vent DNA polymerase (DV exo-) were purchased from New England Biolabs (Ipswich, MA). Optikinase™ was purchased from Affymetrix (Santa Clara, CA). SuperScript II™ reverse transcriptase, Sybr Gold, and Sybr Safe were purchased from Invitrogen (Grand Island, NY). CloneJET™ PCR cloning kit was purchased from Fermentas (Waltham, MA). Syto60 dye was purchased from LI-COR (Lincoln, NE).

DNA sequences.

The sequences below are written from 5' to 3'. Primer and template sequences were synthesized by IDT, except for L3 and the 7dG-modified 41-mer template, which were synthesized in house.

Primers

PBS2: GACACTCGTATGCAGTAGCC

PBS2 FAM: /56-FAM/GACACTCGTATGCAGTAGCC

PBS2.mismatch: CTTTTAAGAACCGGACGAACGACACTCGTTTGCAGTAGCC

PBS1: TGTCTACACGCAAGCTTACA

PBS1.PEG:
AACAAACAAACAAACAAACAAACAAACAAACAAACAAACA/iSp9/TGTCTACACGCAAGCTTACA

extra.primer: CTTTTAAGAACCGGACGAAC

21-mer primer: CGCAGCCGTCCAACCAACTCA

Templates

4NT.0g:
ACTATTCAACTTACAATCCTATCAACCTTATAATCCACTTGGCTACTGCATACGAGTGTC

4NT.1g:
ACTATTCAACTTACAATCGTATCAACCTTATAATCCACTTGGCTACTGCATACGAGTGTC

4NT.2g:
ACTATTCAACTTACAATGGTATCAACCTTATAATCCACTTGGCTACTGCATACGAGTGTC

4NT.3g:
ACTATTCAACTTACAATGGGATCAACCTTATAATCCACTTGGCTACTGCATACGAGTGTC

FI.4NT.9G: TGTCTACACGCAAGCTTACATTAAGACTCGCCATGTTACGATCTGCCAAGTACA
GCCTTGAATCGTCACTGGCTACTGCATACGAGTGTC

L3 library: TGTCTACACGCAAGCTTACA - N_{50} - GGCTACTGCATACGAGTGTC

41mer template: GGACGGCATTGGATCGACG**X**TGAGTTGGTTGGACGGCTGCG

Imaging DNA strands that contain 7dG nucleotides in place of dG.

DNA strands that contain 7dG are known to quench the fluorescence of UV excitable dyes like ethidium bromide and Sybr Gold.[3-6] During imaging, this can lead to weaker than expected band intensity on a gel that can be misconstrued as a low yielding PCR reaction. Because this phenomenon is not widely known, we demonstrated this problem using various imaging techniques to observe PCR amplified DNA on a denaturing polyacrylamide gel. PCR reactions were performed in 100 μL volumes that contained 100 pmol of DNA primer (FAM PBS2 and PBS1) and 5 pmol of DNA template in 1x Thermopol buffer supplemented with 400 nM dATP, dTTP, dCTP, either dGTP or 7-deaza-dGTP, and 0.04 U/μL DV exo-. The solution was thermocycled as follows: 2 min at 95°C followed by 15 cycles of 15 sec at 95°C, 15 sec at 58°C, and 45 sec for dGTP or 2.25 min for 7dGTP at 72°C. PCR reactions were separately analyzed on four different PAGE gels. The fluorescein signal was imaged using a Typhoon TRIO scanner. The gels were then individually stained with either ethidium bromide, Sybr Gold, Sybr Green, or Syto60. Ethidium bromide, Sybr Gold, and Sybr Green stained gels were imaged on a Biorad GelDoc, and the Syto60 stained gel was imaged on a LI-COR Odyssey CLx imager.

Synthesis of DNA templates containing 7-deaza-deoxyguanosine (7dG).

Single-stranded DNA templates containing 7dG in place of dG were generated by PCR. Each PCR reaction (1,000 μL total volume) contained 1,000 pmol of each DNA primer (PBS2 and PBS1.PEG.long) and 50 pmol of DNA template or single-stranded library. Reactions were performed in 1x Thermopol buffer [20 mM Tris-HCl, 10 mM $(NH4)_2SO_4$, 10 mM KCl, 2 mM $MgSO_4$, 0.1% Triton X-100, pH 8.8] supplemented with 400 nM dATP, dTTP, dCTP, 7-deaza-dGTP, and 0.04 U/μL DV exo-. The solution was divided into 50 μL reactions and thermocycled as follows: 2 min at 95°C followed by 15 cycles of 15 sec at 95°C, 15 sec at 58°C, and 2.25 min at 72°C. The reactions were combined, concentrated by lyophilization, and supplemented with 5 mM EDTA and 50% (w/V) urea. The solution was heat denatured for 5 min at 95°C, and the PEG-modified strand was purified by 10% denaturing urea PAGE. The corresponding band was excised, electroeluted, and concentrated by ethanol precipitation. The precipitated pellet was re-suspended in 20 μL of nuclease free water and quantified by UV absorbance. A reaction schematic and example PAGE of the PCR are shown in figure S2.

Small-scale TNA synthesis on dG- and 7dG-containing DNA templates.

PBS2 primer (50 pmol) was radiolabeled using T4 polynucleotide kinase and [γ-^{32}P]ATP. TNA transcriptions reactions were performed in a 10 μL volume containing 10 pmol of synthetic template or 7dG template generated by PCR and 5 pmol of radiolabeled PBS2 primer. The primer and template were annealed in 1x ThermoPol buffer by heating for 5 min at 95°C and cooling for 10 min at 4°C. 0.1 U/μL Therminator DNA polymerase was premixed with 1.25 mM $MnCl_2$. The premixed Therminator/$MnCl_2$ solution and 100 μM tNTPs were added to the primer template complex and the solution was incubated for 3 hours at 55°C for 4NT.0g, 4NT.1g, 4NT.2g, and 4NT.3g templates (Table S1). A 6-hour incubation was used for the L3 library. The reaction products were analyzed by 20% denaturing PAGE and visualized by phosphor-imaging with a Storm scanner.

Large-scale TNA synthesis of 7dG-modified DNA template for the fidelity assay.

Large-scale TNA synthesis reactions were performed in a 250 μL volume reaction containing 250 pmol of extra.PBS2.mismatch DNA primer and an equal amount of FI.4NT.9G DNA template generated by PCR with 7-deaza-dGTP. The primer and template were annealed in 1x ThermoPol buffer by heating for 5 min at 95°C and cooling for 10 min at 4°C. 0.1 U/μL Therminator DNA polymerase was premixed with 1.25 mM $MnCl_2$. The premixed Therminator/$MnCl_2$ solution and 100 μM tNTPs were added to the primer template complex and the solution was incubated for 6 hours at 55°C. TNA was purified from the PEG-modified DNA template by 10% denaturing urea PAGE, electroeluted, and concentrated by ethanol precipitation. The

precipitated pellet was re-suspended in 20 μL of nuclease free water and quantified by Nanodrop. This strategy produces a TNA strand with a single T-T mismatch that can be identified as a watermark (point mutation) only after reverse transcription, PCR amplification, and DNA sequencing. A reaction schematic and example PAGE of the transcription are shown in figure S3.

TNA reverse transcription.

PBS1 (50 pmol) primer was radiolabeled using T4 polynucleotide kinase and [γ-^{32}P]ATP. TNA was reverse transcribed in a final volume of 10 μL. 5 pmol radiolabeled DNA primer PBS1 was annealed to 10 pmol TNA template in 1× First Strand Buffer [50 mM Tris-HCl, 75 mM KCl, 3 mM $MgCl_2$, pH 8.3] by heating for 5 min at 95°C and cooling for 10 min at 4°C. 500 μM dNTPs and 10 mM DTT was added to the primer template complex and incubated for 2 min at 42°C. Finally, 3 mM $MgCl_2$, 1.5 mM $MnCl_2$, and 10 U/μL SuperScript II reverse transcriptase was added to the reaction and incubated for 1 h at 42°C. The reaction products were analyzed by 20% denaturing PAGE and visualized by a storage phosphor screen with a Storm scanner.

PCR analysis of the TNA sample before and after reverse transcription into cDNA.

PCR was used to assay the TNA sample for DNA contaminants that could interfere with the fidelity study. Two 100 μL PCR reactions were performed on 1% volume of the TNA product and TNA RT solutions in 1x ThermoPol buffer, supplemented with 400 nM dNTPs, 100 pmol of DNA primer (extra.primer and PBS1), and 0.04 U/μL DV exo-. The reactions were cycled as follows: 2 min at 95°C followed by 27 cycles of 15 sec at 95°C, 15 sec at 58°C, and 45 sec at 72°C. To monitor DNA amplification, aliquots (8 μL) were removed every 3 cycles and stored on ice. Once the PCR reaction was complete, the samples were analyzed on a 3% agarose gel stained with ethidium bromide.

Fidelity analysis by DNA sequencing.

PCR amplified DNA (1 pmol) was ligated into a pJET vector following manufacturer's protocol. The ligated product was transformed into XL1-blue *E. coli*, cloned, and sequenced at the ASU sequencing facility. Sequencing results were analyzed using CLC Main Workbench.

Pre-steady-state kinetic analysis.

The 21-mer primer strand was radiolabeled using Optikinase and [γ-^{32}P]ATP for 3 hours at 37°C following manufacturer's protocol. The 5□-radiolabeled primer was then annealed to each 41-mer template by incubating the primer and template in a 1:1.15 ratio for 5 min at 95°C before cooling slowly to room temperature over several hours.

All pre-steady-state assays were performed at 37°C in reaction buffer R (20 mM Tris-HCl, 10 mM $(NH_4)_2SO_4$, 10 mM KCl, 0.1 mg/mL BSA, 0.1 mM EDTA, 5 mM DTT, 10% glycerol, pH 8.8), and were carried out using a rapid chemical quench-flow apparatus. A pre-incubated solution of 100 nM Therminator and 20 nM 5□-radiolabeled DNA substrate in buffer R containing 1.25 mM $MnCl_2$ was rapidly mixed with 100 μM dNTP or tNTP and 2 mM Mg^{2+} in reaction buffer R for various times before quenching with 0.37 M EDTA. All reported concentrations are final after mixing.

The reaction products were separated by 17% denaturing urea PAGE. The gel image was exposed to a storage phosphor screen, imaged using a Typhoon TRIO, and analyzed by ImageQuant. The plots of product concentration versus time were prepared using Kaleidagraph. Data were fit to a single-exponential equation ($[\text{product}] = A[1 - \exp(-k_{obs}t)]$), where k_{obs} is the observed reaction rate and A is the product amplitude. All reactions were performed in duplicate.

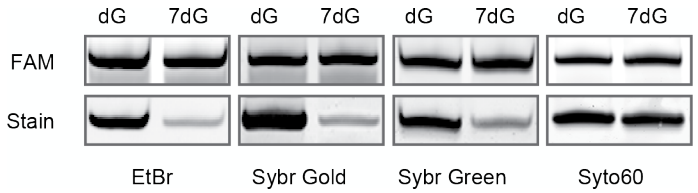
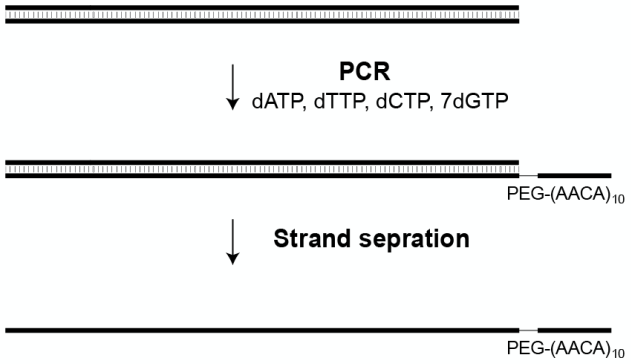
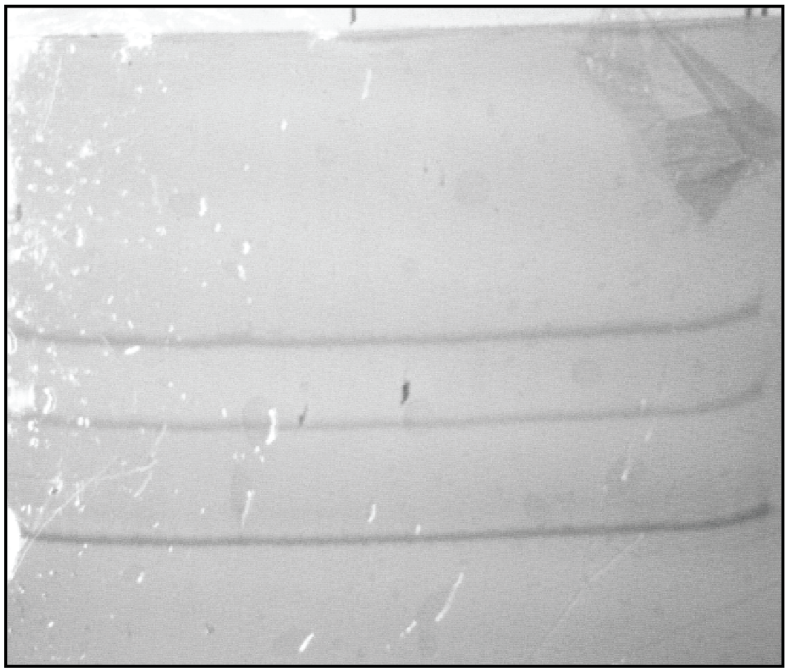


Figure S1. Evidence of fluorescence quenching by 7dG-containing DNA. PCR amplified DNA carrying a fluorescein-label was analyzed on four different denaturing PAGE gels. Each gel was imaged by fluorescence and separately stained with ethidium bromide (EtBr), Sybr Gold, Sybr Green, or Syto60. Fluorescence imaging (top) demonstrates that an equivalent amount of dG- and 7dG-containing DNA is present in each lane. When the same gels are stained with ethidium bromide, Sybr Gold, Sybr Green, and Syto60 (bottom) only the Syto60 stain shows equal amounts of PCR product for the dG and 7dG amplicons. Reduced band intensity for the EtBr, Sybyr Gold, and Sybr Green stains is due to quenching of the fluorescent dyes by 7dG residues in the DNA. The Syto60 stain does not lead to fluorescence quenching because it emits in the infrared region.

A



B



- Templating strand (130 nt)

- Non-templating strand (90 nt)

- PBS.PEG (60 nt)

Figure S2. Amplification and gel mobility of the 7dG DNA template. (A) Double-stranded DNA library L3 was amplified by PCR with 7dGTP in place of natural dGTP. One of the PCR primers (PBS1.PEG) is modified with a PEG spacer and a 40-nt overhang for separation by denaturing urea PAGE. (B) Representative image showing the different DNA products separated on a denaturing polyacrylamide gel.

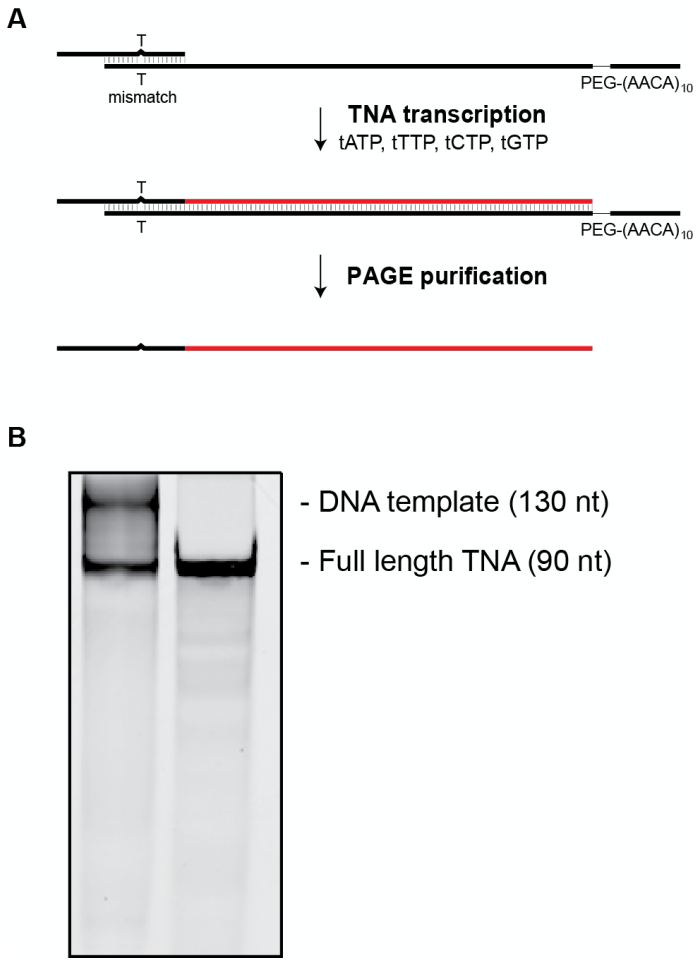


Figure S3. TNA transcription. (A) Cartoon imaging showing single-stranded DNA template containing either dG or 7dG transcribed into TNA. The resulting TNA-DNA duplex is purified by denaturing PAGE. (B) Representative image showing the TNA product obtained from a typical primer-extension reaction observed by denaturing PAGE with ethidium bromide staining (left lane). Size-matched DNA marker (right lane).

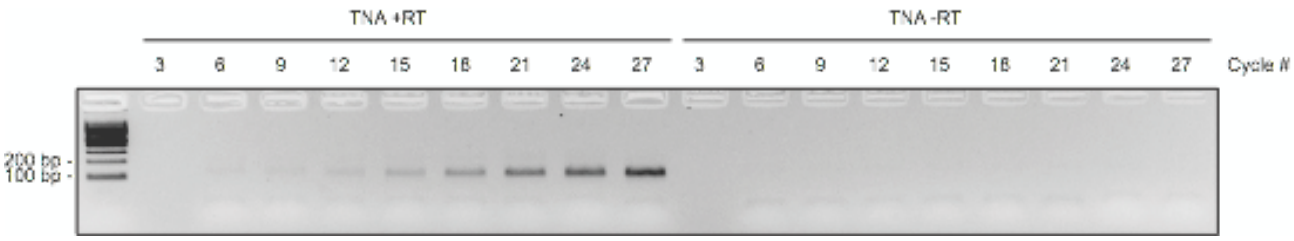


Figure S4. PCR assay for DNA contaminants. PCR analysis confirms that DNA amplicons are only observed when the TNA sample is reverse transcribed into cDNA.

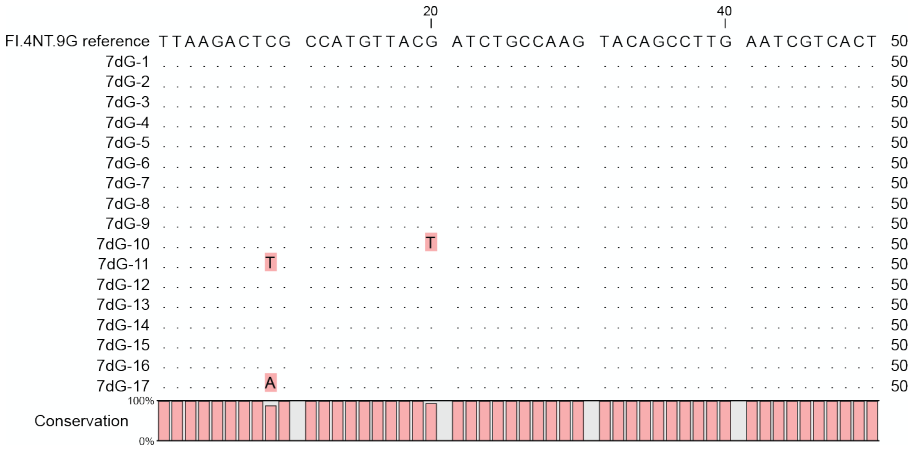


Figure S5. DNA sequence alignment used to assess TNA replication fidelity. An alignment of the sequence reads to the FI.4NT.9G reference sequence indicates that 3 mutations were observed in a total of 850-nt positions. This yields an overall fidelity of 99.6% for TNA replication using 7dG containing templates. Interestingly, no G to C transversions were observed in the data, demonstrating that 7dG inhibits G:G mispairing in the enzyme active site.

Table S1. List of sequencing results from the 7dG sequencing of FI.4NT.9G template.

Name	Sequence
7dG-1	TGTCTACACGCAAGCTTACATTAAGACTCGCCATGTTACGATCTGCCAAGTA CAGCCTTGAATCGTCACTGGCTACTGCAAACGAGTGTC
7dG-2	TGTCTACACGCAAGCTTACATTAAGACTCGCCATGTTACGATCTGCCAAGTA CAGCCTTGAATCGTCACTGGCTACTGCAAACGAGTGTC
7dG-3	TGTCTACACGCAAGCTTACATTAAGACTCGCCATGTTACGATCTGCCAAGTA CAGCCTTGAATCGTCACTAGCTACTGCAAACGAGTGTC
7dG-4	TGTCTACACGCAAGCTTACATTAAGACTCGCCATGTTACGATCTGCCAAGTA CAGCCTTGAATCGTCACTGGCTACTGCAAACGAGTGTC
7dG-5	TGTCTACACGCAAGCTTACATTAAGACTCGCCATGTTACGATCTGCCAAGTA CAGCCTTGAATCGTCACTGGCTACTGCAAACGAGTGTC
7dG-6	TGTCTACACGCAAGCTTACATTAAGACTCGCCATGTTACGATCTGCCAAGTA CAGCCTTGAATCGTCACTGGCTACTGCAACGAGTGTC
7dG-7	TGTCTACACGCAAGCTTACATTAAGACTCGCCATGTTACGATCTGCCAAGTA CAGCCTTGAATCGTCACTGGCTACTGCAAACGAGTGTC
7dG-8	TGTCTACACGCAAGCTTACATTAAGACTCGCCATGTTACGATCTGCCAAGTA CAGCCTTGAATCGTCACTGGCTACTGCAAACGAGTGTC
7dG-9	TGTCTACACGCAAGCTTACATTAAGACTCGCCATGTTACGATCTGCCAAGTA CAGCCTTGAATCGTCACTGGCTACTGCAAACGAGTGTC
7dG-10	TGTCTACACGCAAGCTTACATTAAGACTCGCCATGTTACTATCTGCCAAGTAC AGCCTTGAATCGTCACTGGCTACTGCAAACGAGTGTC
7dG-11	TGTCTACACGCAAGCTTACATTAAGACTTGCCATGTTACGATCTGCCAAGTAC AGCCTTGAATCGTCACTGGCTACTGCAAACGAGTGTC
7dG-12	TGTCTACACGCAAGCTTACATTAAGACTCGCCATGTTACGATCTGCCAAGTA CAGCCTTGAATCGTCACTGGCTACTGCAAACGAGTGTC
7dG-13	TGTCTACACGCAAGCTTACATTAAGACTCGCCATGTTACGATCTGCCAAGTA CAGCCTTGAATCGTCACTGGCTACTGCAAACGAGTGTC
7dG-14	TGTCTACACGCAAGCTTACATTAAGACTCGCCATGTTACGATCTGCCAAGTA CAGCCTTGAATCGTCACTGGCTACTGCAAACGAGTGTC
7dG-15	TGTCTACACGCAAGCTTACATTAAGACTCGCCATGTTACGATCTGCCAAGTA CAGCCTTGAATCGTCACTGGCTACTGCAAACGAGTGTC
7dG-16	TGTCTACACGCAAGCTTACATTAAGACTCGCCATGTTACGATCTGCCAAGTA CAGCCTTGAATCGTCACTGGCTACTGCAAACGAGTGTC
7dG-17	TGTCTACACGCAAGCTTACATTAAGACTAGCCATGTTACGATCTGCCAAGTA CAGCCTTGAATCGTCACTGGCTACTGCAAACGAGTGTC

Table S2. Comparison of the replication fidelity for TNA transcription.

Temp.	XNA	Transcription	Reverse Transcription	Bases Read	Substitution	Insertion + Deletion	Total Error
3NT	TNA	Therminator	SuperScript II	1050	1.9×10^{-3}	1.9×10^{-3}	3.8×10^{-3}
4NT	TNA	Therminator	SuperScript II	1650	34.5×10^{-3}	1.2×10^{-3}	35.8×10^{-3}
4NT-7	TNA	Therminator	SuperScript II	850	3.5×10^{-3}	1.2×10^{-3}	3.5×10^{-3}