

An in Vitro Selection System for TNA

Justin K. Ichida,[†] Keyong Zou,[†] Allen Horhota,[‡] Biao Yu,[§] Larry W. McLaughlin,[‡] and Jack W. Szostak^{*†}

Howard Hughes Medical Institute, and Department of Molecular Biology, Massachusetts General Hospital, Boston, Massachusetts 02114, Department of Chemistry, Boston College, Chestnut Hill, Massachusetts 02467, and State Key Laboratory of Bioorganic and Natural Products Chemistry, Shanghai Institute of Organic Chemistry, Chinese Academy of Sciences, Shanghai 200032, China

Received August 1, 2004; E-mail: szostak@molbio.mgh.harvard.edu

(3'-2')- α -L-Threose nucleic acid (TNA, Figure 1A) is an unnatural nucleic acid that was identified during an extensive evaluation of alternative sugar-phosphate backbones aimed at explaining the structure of the biological nucleic acids.^{1,2} TNA possesses the ability to specifically base-pair with RNA, DNA, and itself.² This capability, together with the chemical simplicity of threose relative to ribose, suggests that TNA could have acted as an evolutionary competitor of RNA or even have preceded RNA as the genetic molecule of life. We are attempting to investigate the functional potential of TNA by implementing an in vitro selection scheme for TNA.^{3,4} Here, we show that a mutant archaeal family B DNA polymerase is capable of polymerizing more than 50 nucleotides of TNA on a DNA template. We also demonstrate the display of single-stranded TNA covalently linked to its encoding duplex DNA, thus enabling the selection of functional TNA sequences and the amplification or recovery of the attached DNA.

We and others have previously shown that certain family B archaeal DNA polymerases possess the ability to synthesize limited stretches of TNA on a DNA template.^{5,6} Our recent synthesis of all four TNA triphosphates (tNTPs) enabled us to test polymerases for more extensive activity.⁷ The Terminator DNA polymerase is an engineered exonuclease-deficient form of "9°N" DNA polymerase containing an A485L mutation.⁸ It is capable of efficiently incorporating a wide spectrum of modified nucleotides. We tested the ability of this polymerase to accept tNTPs as substrates using a DNA primer/template construct containing a 50-nucleotide single-stranded template region in which all four DNA nucleobases were represented (Figure 1B). Since previous work had shown that pairing diaminopurine opposite thymine increases the efficiencies of both template-directed ligation and polymerization,^{5,9} we used diaminopurine triphosphate (tDTP) instead of tATP. The Terminator polymerase catalyzed the synthesis of >20% full-length 50-nucleotide TNA product within 24 h (Figure 1C).

We reasoned that if transcribed TNA could be covalently linked to its DNA template, we could perform functional selections for TNA molecules and rescue the successful genotypes by PCR amplification of the attached DNA. This approach is analogous to the selection of functional peptides and proteins by mRNA display;¹⁰ the use of DNA display for peptide and PNA selections has also been proposed.^{10,11} By starting with a library of single-stranded DNA hairpins, the 3' end of each hairpin could act as a primer for TNA transcription across the randomized DNA template region (Figure 2A). A primer annealed to the loop region of the hairpin could then initiate strand-displacement synthesis, liberating the TNA strand to allow folding and linearizing the DNA template by making it double-stranded.

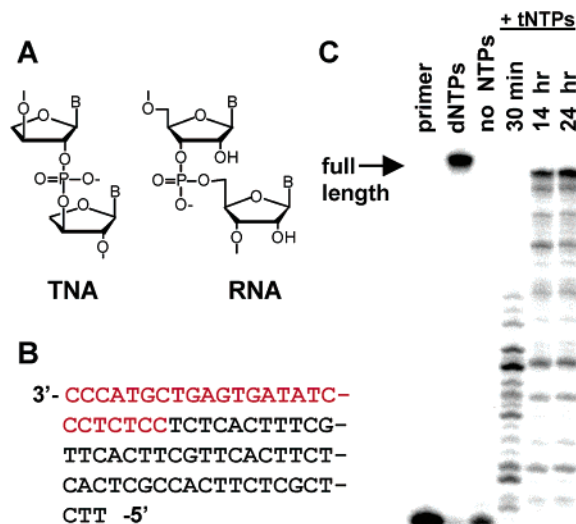


Figure 1. DNA-dependent TNA polymerization with the Terminator DNA polymerase. (A) Structures of TNA and RNA (B) Template sequence. Red = primer binding site; black = ssDNA template. (C) Time course of TNA polymerization. Reaction was performed with 5 nM primer/template, 60 μ M tDTP, 31 μ M tTTP, 18 μ M tCTP, 2 μ M tGTP, 1 unit of *Th* pyrophosphatase, and 0.5 units of Terminator polymerase in 10 μ L of 1 \times thermopol buffer (NEB) at 75 $^{\circ}$ C. Samples were boiled in 7 M urea and 400 mM NaOH and analyzed by 20% denaturing PAGE.

To test this idea, we synthesized a single-stranded DNA hairpin and transcribed the 60-nucleotide single-stranded region using the Terminator DNA polymerase and tNTPs (Figure 2B). Native gel analysis of the reaction indicated efficient TNA transcription (Figure 2C). Full-length transcripts were purified from the gel and subjected to a strand-displacement reaction with dNTPs, Terminator polymerase, and an end-labeled primer. With the hairpin construct in excess, most of the labeled primer was extended to full length, indicating very efficient strand displacement (Figure 2D). To verify that the resulting DNA-DNA duplexes were stable against strand-invasion by TNA, the strand-displaced constructs were treated with the restriction enzyme BsrBI, which specifically cleaves 15 base pairs from the end of the DNA template. The strand-displaced products were completely digested by BsrBI. BbsI, a control enzyme without a site in the DNA sequence, showed no activity (Figure 2D). BsrBI did not cleave ssDNA (data not shown).

In vitro selection for TNA function will only work if the TNA transcript is a reasonably faithful copy of its DNA template. The number of errors in a full-length TNA transcript is a function of both the misincorporation rate and the rate of primer extension from mismatched sites and is therefore time-dependent (since at longer times, full-length material accumulates by extension of strands that had paused at sites of misincorporation). To study the time-

[†] Howard Hughes Medical Institute and Massachusetts General Hospital.

[‡] Boston College.

[§] Chinese Academy of Sciences.

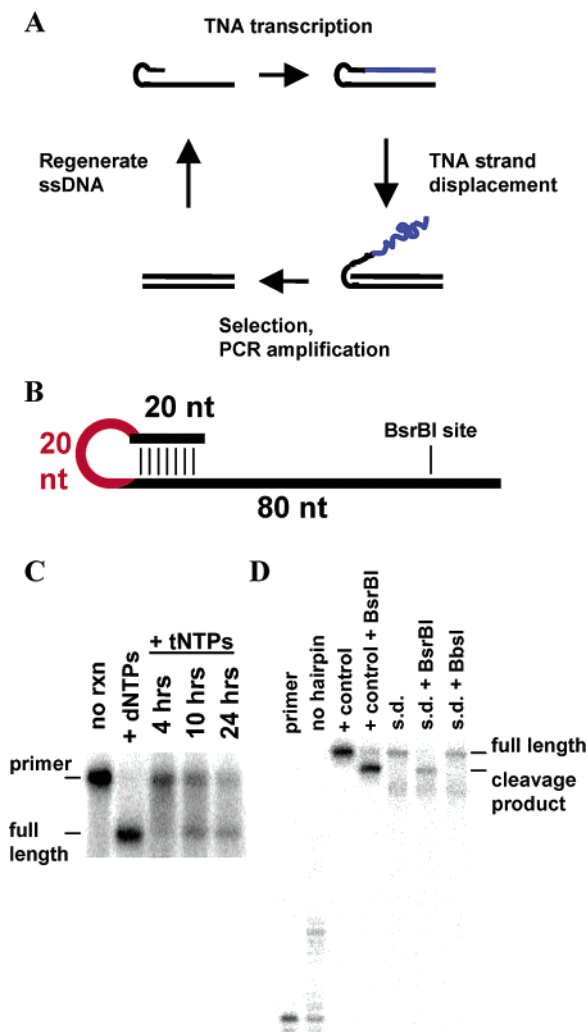


Figure 2. DNA-displayed TNA molecules. (A) In vitro selection scheme for TNA. (B) Synthetic DNA construct. Red represents loop region/strand displacement primer binding site. (C) TNA transcription of hairpin construct. 50 nM hairpin was transcribed with tNTPs for 12 h (~13% fully extended) as in Figure 1C. Control reaction (+dNTPs) was performed for 10 min. Reactions were analyzed by 12% native PAGE. (D) Strand displacement (s.d.) of TNA strand. Full-length TNA-transcribed hairpins were gel-purified by native PAGE. End-labeled DNA primer (0.1 nM) was annealed to 1 nM transcribed hairpin by heating to 55 °C for 1 min in 1.1× thermopol buffer and 250 μM of each dNTP. One unit of Therminator was added, and the reactions were incubated for 30 min at 55 °C followed by 90 min at 75 °C. Restriction digests: strand displacement reactions were diluted 4-fold in 1× thermopol buffer, and 10 units of restriction enzyme were added for 2 h at 37 °C. Reactions were analyzed by 20% denaturing PAGE. “No hairpin” lane is a primer extension without the hairpin template. “+ control” lane is a primer extension on a single-stranded version of the hairpin construct.

dependent fidelity of TNA synthesis by the Therminator polymerase, we examined the effect of omitting one tNTP at a time from short primer-extension reactions (Supporting Information Figure 1). For each reaction, the DNA template contained only one template site complementary to the omitted tNTP. This site was positioned 11 nucleotides downstream of the DNA primer terminus to ensure that fidelity was measured when the polymerase was fully contacting the TNA strand. A control reaction containing all four tNTPs was performed in parallel for each template. We compared the amount of full-length product in each –tNTP reaction versus control reactions at multiple time points to estimate the fidelity for each base. This assay accounts for both the slow rate of misincor-

poration as well as the slow rate of extension from mismatched primer termini, but underestimates the actual fidelity because of the lack of competition between correct and incorrect nucleotides. These experiments suggest that the Therminator polymerase incorporates tGTP and tCTP with ~99 and ~98% accuracy, respectively, in full-length TNA strands, even at relatively long incubation times (Supporting Information Table 1). The fidelities for incorporation of tDTP and tTTP were much lower (~94% at 1 h, 10–20% of the primer fully extended), presumably reflecting misincorporation of tGTP and tCTP because of wobble base-pairing. Although the overall fidelity of TNA synthesis is lower than previously measured DNA synthesis fidelities,¹² more than 10% of TNA pool molecules would be error-free in material purified when TNA synthesis had reached 10–20% as we would do during a selection (Supporting Information Table 2). The actual error rate in material synthesized in the presence of all four tNTPs, and with high D/G and T/C ratios, may be much lower. Since typical in vitro selection experiments involve enrichment factors of more than 100–1000-fold per round, even a 10-fold reduction due to incorrectly transcribed sequences would yield an acceptable enrichment factor of at least 10–100-fold per round.

We have identified the Therminator DNA polymerase as an enzyme capable of reasonably efficient and faithful DNA-templated TNA polymerization. The DNA-display technique developed here for the selection of functional TNA molecules is uniquely suited for the directed evolution of unnatural nucleic acids because it does not require a polymerase that can copy the unnatural nucleic acid back into DNA. Thus, any polymer that can be synthesized by a DNA-templated reaction, either enzymatically or nonenzymatically,¹¹ can be used for selection. We hope to investigate the potential of TNA for forming ligand-binding sites or catalytically active structures in order to evaluate its possible role as a progenitor of RNA.

Acknowledgment. We thank R. Krishnamurthy and A. Eschenmoser for the generous gift of advanced intermediates for the synthesis of D, G, and T tNTPs. We are grateful to W. Jack and A. Gardner for very helpful discussions. This work was supported by grants from the NIH (R01 GM053936) and the NASA Astrobiology Institute (NNA04CC12A) to J.W.S. and a grant from the NSF (MCB0077667) to LWM. J.W.S. is an Investigator of the Howard Hughes Medical Institute. J.K.I. was supported in part by a Ford Foundation predoctoral fellowship.

Supporting Information Available: List of oligonucleotide sequences, denaturing PAGE analysis of fidelity assays (Supporting Figure 1), and the fidelity values (Supporting Table 1). This material is available free of charge via the Internet at <http://pubs.acs.org>.

References

- (1) Eschenmoser, A. *Science* **1999**, *284*, 2118–2124.
- (2) Schoning, K.; Scholz, P.; Guntha, S.; Wu, X.; Krishnamurthy, R.; Eschenmoser, A. *Science* **2000**, *290*, 1347–1351.
- (3) Carothers, J. M.; Oestreich, S. C.; Davis, J. H.; Szostak, J. W. *J. Am. Chem. Soc.* **2004**, *126*, 5130–5137.
- (4) Chaput, J. C.; Ichida, J. K.; Szostak, J. W. *J. Am. Chem. Soc.* **2003**, *125*, 856–857.
- (5) Chaput, J. C.; Szostak, J. W. *J. Am. Chem. Soc.* **2003**, *125*, 9274–9275.
- (6) Kempeneers, V.; Vastmans, K.; Rozenski, J.; Herdewijn, P. *Nucleic Acids Res.* **2003**, *31*, 6221–6226.
- (7) Zou, K.; Horhota, A.; Chaput, J. C.; Szostak, J. W.; McLaughlin, L. W. *Org. Lett.*, to be submitted for publication.
- (8) Gardner, A. F.; Jack, W. E. *Nucleic Acids Res.* **2002**, *30*, 605–613.
- (9) Wu, X.; Delgado, G.; Krishnamurthy, R.; Eschenmoser, A. *Org. Lett.* **2002**, *4*, 1283–1286.
- (10) Kurz, M.; Gu, K.; Al-Gawari, A.; Lohse, P. A. *ChemBioChem* **2001**, *2*, 666–672.
- (11) Rosenbaum, D. M.; Liu, D. R. *J. Am. Chem. Soc.* **2003**, *125*, 13924–13925.
- (12) Kunkel, T. A.; Bebenek, K. *Annu. Rev. Biochem.* **2000**, *69*, 497–529.

JA045364W