Published online September 12, 2005

Nucleic Acids Research, 2005, Vol. 33, No. 16 ***5219–5225***
doi:10.1093/nar/gki840

High fidelity TNA synthesis by Therminator polymerase

Justin K. Ichida[1,2,3], Allen Horhota[4], Keyong Zou[1,2,3], Larry W. McLaughlin[4] and Jack W. Szostak[1,2,3,*]

[1]Howard Hughes Medical Institute and [2]Department of Molecular Biology, Massachusetts General Hospital, Boston, MA 02114, USA, [3]Department of Genetics, Harvard Medical School, Boston, MA 02115, USA and [4]Department of Chemistry, Boston College, Chestnut Hill, MA 02467, USA

Received July 25, 2005; Revised and Accepted August 30, 2005

ABSTRACT

Therminator DNA polymerase is an efficient DNA-dependent TNA polymerase capable of polymerizing TNA oligomers of at least 80 nt in length. In order for Therminator to be useful for the *in vitro* selection of functional TNA sequences, its TNA synthesis fidelity must be high enough to preserve successful sequences. We used sequencing to examine the fidelity of Therminator-catalyzed TNA synthesis at different temperatures, incubation times, tNTP ratios and primer/template combinations. TNA synthesis by Therminator exhibits high fidelity under optimal conditions; the observed fidelity is sufficient to allow *in vitro* selection with TNA libraries of at least 200 nt in length.

INTRODUCTION

(3′–2′) α-L-threose nucleic acid (TNA) (Figure 1a) is an unnatural nucleic acid discovered during an extensive examination of the base-pairing properties of nucleic acids containing alternative sugar-phosphate backbones (1,2). Despite having a repeat unit one atom shorter than natural nucleic acids, TNA is able to 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 been an evolutionary progenitor or competitor of RNA during the origin of life on Earth. We are attempting to investigate the plausibility of this hypothesis by examining the functional potential of TNA oligomers. This will require a series of *in vitro* selection experiments, which in turn will require the accurate synthesis of long TNA sequences (3).

Therminator DNA polymerase, an A485L point mutant of the 9°N DNA polymerase, efficiently synthesizes long TNA oligonucleotides on a DNA template, using threonucleoside triphosphates (tNTPs, Figure 1b) as substrates (3,4). The fidelity of TNA polymerization is a critical factor, because an excessive mutation rate would make the selection of functional TNA sequences impossible. For example, if only 10% of pool molecules are correctly transcribed, there could be a reduction of up to 10-fold in the enrichment factor per round due to background from inactive mutant sequences. T7 RNA polymerase and AMV reverse transcriptase have a combined error rate of $\sim 1 \times 10^{-4}$, meaning that RNA *in vitro* selection pools of up to 300 nt are >97% error-free before PCR amplification (5). Typically, *in vitro* selection experiments based on differential binding effects have an enrichment factor of 100- to 1000-fold per round.

We previously estimated Therminator TNA polymerization fidelity by an assay in which rates of polymerization were compared between a reaction with all four tNTPs and others lacking one of the tNTPs (3). This 'dropout assay' suggested that the TNA polymerization error rate exceeds that of natural DNA polymerases with dNTPs, but would still allow *in vitro* selection experiments using short TNA sequences (3). However, the dropout assay is likely to underestimate the fidelity of TNA synthesis, because the extended pausing at sites where the required nucleotide is absent provides more opportunity for misincorporation to occur. Direct sequencing of TNA is not trivial, because there are no highly efficient TNA-directed DNA polymerases. We therefore developed a scheme to sequence TNA polymerization products of 10 nt formed in the presence of all four tNTPs. The fidelity determined by this sequencing assay is the actual accuracy with which full-length TNA is synthesized, and therefore reflects the combined effects of misincorporation and the rate of extension from mismatched primer-termini. In this paper, we use the term fidelity in this operational sense of the accuracy with which full-length TNA is synthesized, as opposed to the more restricted sense of the accuracy of single-nucleotide incorporation. We examined the fidelity of Therminator-mediated TNA

*To whom correspondence should be addressed. Tel: +1 617 726 5981; Fax: +1 617 726 6893; Email: szostak@molbio.mgh.harvard.edu

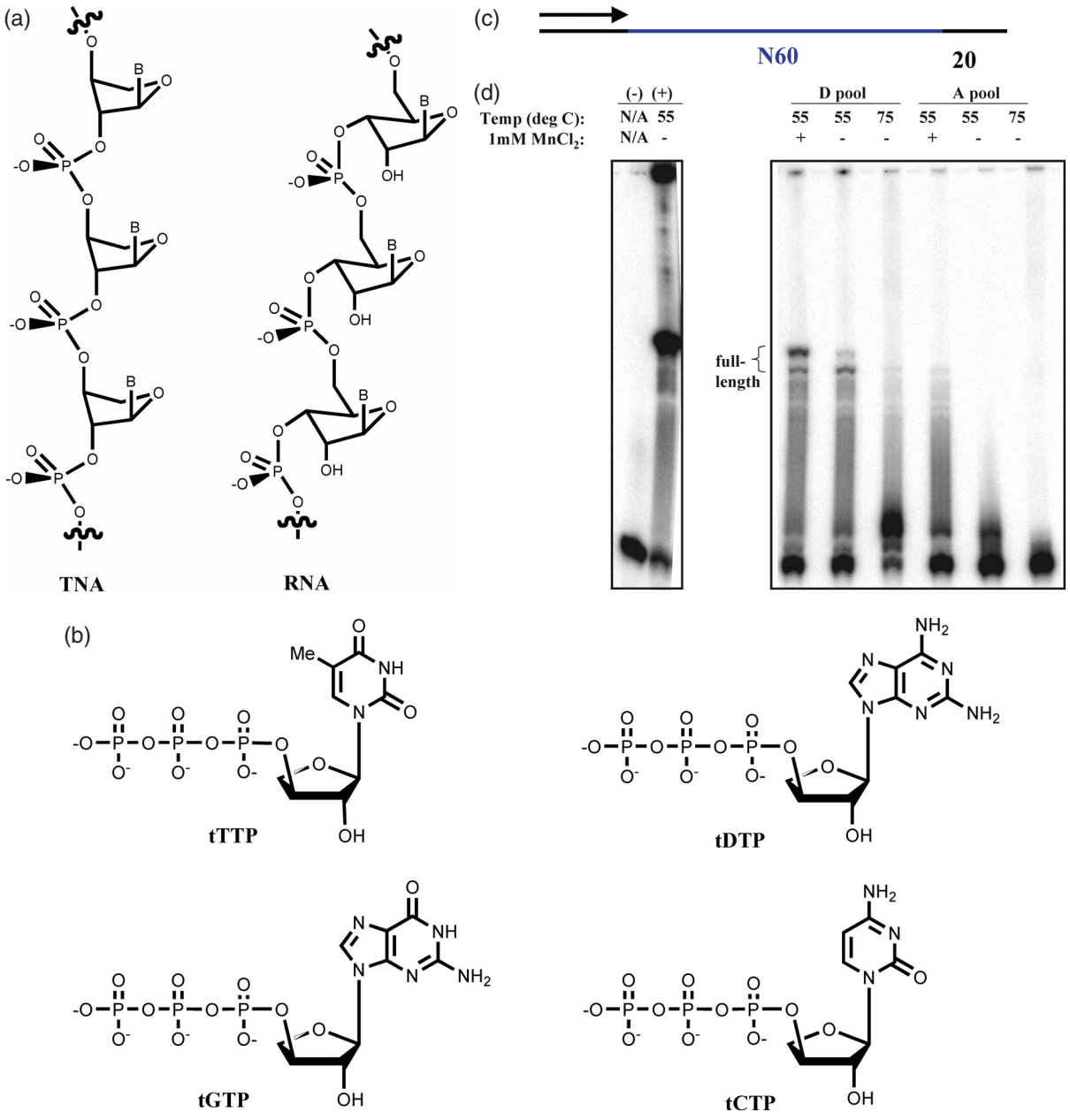

Figure 1. TNA polymerization. (**a**) Structure of TNA and RNA. (**b**) Structures of the TNA triphosphates. tDTP was used instead of tATP to increase polymerization efficiency (4,23). (**c**) Construct used to test TNA polymerization conditions. Blue region is randomized, black region is constant. (**d**) Denaturing PAGE analysis of 24 h TNA polymerization reactions. Left panel: control reactions (−) without polymerase or (+) with polymerase and dNTPs. Right panel: tNTP reactions. The lower band in 'full-length' products corresponds to the end of the random region.

synthesis with different templates, reaction temperatures, incubation times and tNTP concentrations. Under optimal reaction conditions with purification of full-length material, Therminator catalyzes TNA polymerization with high fidelity sufficient for *in vitro* selection with TNA libraries of at least 200 nt in length.

MATERIALS AND METHODS

TNA polymerization on random sequence templates

The 123 nt pool 123 (5′-GTGGTGACGGCCGAGGTTG-TAACACGCGAGTGCACGCTGCN$_{40}$GAGCGGN$_{14}$CTAC-ACTTGCCAGCGCCCTAGCG-3′) was synthesized at the 0.2 μmol scale (Massachusetts General Hospital DNA Core) and was gel purified. An aliquot of 100 ng of pool was amplified by PCR for 10 cycles with primers 20PCR (GTGGTGACGGCCGAGGTTGT) and biotinylated primer 23 (5′-photocleavable biotin-CGCTAGGGCGCTGGCAAG-TGTAG-3′) using either all four natural dNTPs or with diaminopurine-2′-deoxyribose triphosphate (dDTP, Trilink Biotech) substituted for adenosine-2′-deoxyribose triphosphate (dATP) in a total reaction volume of 200 μl. Each PCR was immobilized on Immunopure streptavidin agarose (Pierce Biotechnology) and washed with 40 column volumes of 75 mM NaOH to remove the unbiotinylated strands. The remaining single-stranded DNA (ssDNA) was released from

the column by irradiation at 365 nm for 30 min and gel purified by denaturing PAGE, electroeluted in 1× TAE buffer, ethanol precipitated and resuspended in water. An aliquot of 50 nM end-labeled primer 40 (5′-GTGGTGACGGCCGAGGTTG-TAACACGCGAGTGCACGCTGC-3′) was annealed to 100 nM ssDNA pool whose sequences contained either adenine or diaminopurine in 1.1× Thermopol buffer (New England Biolabs) plus 60 μM tDTP, 60 μM tTTP, 18 μM tCTP and 2 μM tGTP. TNA polymerization reactions were started by the addition of 1 mM or no $MnCl_2$, 1 U of *Tth* pyrophosphatase (Roche) and 0.5 U of Therminator DNA polymerase (New England Biolabs). Reactions were incubated at 55 or 75°C for 24 h and stopped by the addition of 7 M urea/40 mM EDTA/400 mM NaOH. Samples were analyzed by denaturing PAGE and phosphorimager (Molecular Dynamics) and quantified by ImageQuant software.

TNA sequencing

An aliquot of 200 nM primer 'P1/2' (5′-GTGGTGACGGCC-GAGGTTGTATAGAAAGAGGGTtCGACTCACTATAGG-GAGAGG-3′), labeled at the 5′ end by incubation with [γ-^{32}P]ATP and T4 Polynucleotide Kinase (New England Biolabs), was annealed to 300 nM template 'T1' (5′-CCAT-TAGTCTCCTCTCCCTATAGTGAGTCGTACCCTCTTTC-TATTGTTGG-3′) or template 'T2' (5′-GTDTCGTDC-TCCTCTCCCTATAGTGAGTCGTACCCTCTTTCTATTG-TTGG-3′) in 1.1× Thermopol buffer. Regions of complementarity between the primer and templates are underlined. The lower case 't' in the primer indicates a marker nucleotide that was intentionally designed to be non-complementary to the template in order to distinguish clones that resulted from amplification of an actual primer extended with TNA from clones that resulted from amplification of leftover template strands. tNTP concentrations were either 120 μM tDTP, 60 μM tTTP, 18 μM tCTP, 2 μM tGTP or 30 μM tDTP, 30 μM tTTP, 30 μM tCTP, 4 μM tGTP. Control reactions contained 250 μM of each dNTP instead of tNTPs. The 5 μl reactions were initiated by the addition of 1 U of *Tth* pyrophosphatase and 0.5 U of Therminator DNA polymerase. Reactions were incubated at 55 or 75°C for the times specified in Table 2 and stopped by the addition of 7 M urea/40 mM EDTA/400 mM NaOH. Full-length material (~10–25% of reaction) was gel purified by denaturing PAGE, electroeluted, ethanol precipitated and resuspended in water. Purified full-length material was annealed to 100 nM template 'T1B' (5′-AAAAAAAAAA-AAAAAAAAAACTTCCCTCCTCTTTCTTCTCTCCATT-AGTCTCCTCTCCCTATAGTGAGTCGTACCCTCTTTC-TATACAACCTCGGCCGTCACCAC-3′) or template 'T2B' (5′-AAAAAAAAAAAAAAAAAAAAACTTCCCTCCTCT-TTCTTCTCTGTATCGTACTCCTCTCCCTATAGTGAGT-CGTACCCTCTTTCTATACAACCTCGGCCGTCACCA-C-3′) in 1.1× Thermopol buffer plus 250 μM dATP and dGTP. Therminator DNA polymerase (0.5 U) was added and the 5 μl reactions were incubated at 55°C for 1 h. Fully extended primer was purified by denaturing PAGE, electroeluted, ethanol precipitated and resuspended in water. One half of each TNA/DNA sample was annealed to 1 nM biotinylated primer 'RT' (5′-biotin-CTTCCCTCCTCTTTCTTCTCT-3′) in 1.5× First Strand buffer (Superscript II, Invitrogen) plus 125 μM dNTPs. Aliquots containing 10 mM DTT, 1.5 mM $MnCl_2$ and 1 μl Superscript II reverse transcriptase were added to a total volume of 20 μl and reactions were incubated at 37°C for 5 min followed by 1 h at 42°C and 10 min at 80°C. Reverse-transcribed material was gel purified by denaturing PAGE, electroeluted and immobilized on Immunopure streptavidin agarose overnight. Samples were washed with 100 column volumes of 100 mM NaOH to remove non-biotinylated DNA and TNA, 40 column volumes of water, 60 column volumes of 1× Thermopol buffer, and resuspended in 80 μl 1× Thermopol buffer. An aliquot of 20 μl from each sample was PCR amplified for 25 rounds with dNTPs, a non-biotinylated version of primer 'RT', primer 'PCR' (5′-GTGGTGACGGCCGAGGTTGT-3′) and *Taq* polymerase. PCR products were gel purified by agarose gel electrophoresis and Gel Extraction Kit (Qiagen) and re-amplified by 10 rounds of PCR with the same primers. Aliquots containing 4 μl of each PCR were TOPO TA cloned (Invitrogen) and 24–72 colonies were sequenced. The presence of the 'marker' nucleotide indicated that the product was not a result of DNA template contamination and was confirmed for each sequence counted. Statistical significance probabilities were calculated using the binomial probability mass function $P(x,p,n) = \binom{n}{x}(p)^x(1-p)^{(n-x)}$, where $\binom{n}{x} = n!/(x!(n-x!))$. P was the probability of observing x successes in n number of trials where p was the probability of success on each trial. For example, to determine whether the difference between two fidelity measurements, A and B, was significant, the raw fidelity of A was used as p, the number of correct bases in B was x, and the total number of bases sequenced in B was n. In each case, P was also calculated using the fidelity of B as p, the number of correct bases in A as x, and the number of total bases sequenced in A as n. In all cases, switching A and B did not change the significance of the difference between fidelities. In using the binomial probability mass function, we assume that each nucleotide incorporation event is independent of all other incorporation events, that only two outcomes were possible (correct insertion or incorrect insertion) and that the probability of success was constant throughout the experiment.

RESULTS

TNA polymerization of a 60 nt random pool

To determine the optimal conditions for efficient TNA polymerization on long templates, we tested the ability of Therminator to synthesize TNA on a library of random-sequence DNA templates (Figure 1c). Since our kinetic studies have shown that replacement of adenine by diaminopurine in the template increases the efficiency of TNA polymerization (4), we tested polymerization activity on DNA templates containing either adenine or diaminopurine. In addition, we tested polymerization in the presence or absence of 1 mM $MnCl_2$. Finally, since the dropout fidelity assay had indicated that raising the reaction temperature from 55 to 75°C increased fidelity, we tested activity at both temperatures.

Polymerization on diaminopurine-containing templates was dramatically better than on adenine-containing templates, as indicated by the synthesis of at least 10 times more full-length product at early time points (Figure 1d and Table 1). Similarly,

Table 1. Percentage of full-length pool TNA polymerization

	Template bases	Temperature (°C)	$MnCl_2$ (mM)	% Full length
1	DTGC	55	1	10
2	DTGC	55	0	5
3	DTGC	75	0	1
4	ATGC	55	1	1
5	ATGC	55	0	<1
6	ATGC	75	0	<1

TNA polymerization was much better at 55°C than at 75°C, with barely detectable amounts of full-length product being formed at 75°C in the absence of Mn^{2+} (Figure 1d). The effect of Mn^{2+} on polymerization was less pronounced, with only a slight difference between 55°C reactions on diaminopurine templates with or without Mn^{2+} (Figure 1d). Based on these results, we compared the fidelity of reactions carried out at 55 or 75°C and on diaminopurine or adenine templates.

Design of sequencing scheme

We sequenced TNA polymerization products by copying the TNA into DNA and sequencing those DNA copies. First, we performed the TNA polymerization reactions using either primer/template pair 1 (P/T1) or primer/template pair 2 (P/T2) (Figure 2a). At different time points, we stopped the reactions and gel purified the full-length material, as we would do during *in vitro* selection in order to enrich for correctly transcribed products. Next, we annealed the products to an extended form of the original template and supplied dATP, dGTP and Therminator polymerase to add a DNA tail to the end of the TNA transcript (Figure 2b). This tail served as a primer binding site for reverse transcription of the TNA/DNA chimera. After gel purification of the TNA transcript with the DNA tail, we reverse transcribed the transcript using Superscript II reverse transcriptase because we had previously shown that this enzyme is capable of limited TNA-dependent DNA polymerization (6). We purified the cDNA first by polyacrylamide gel electrophoresis and then by immobilization on streptavidin beads followed by washing with 100 mM NaOH. Finally, we PCR amplified, cloned and sequenced the resulting products. As a control, we measured the fidelity of *Taq* DNA polymerase and Therminator DNA polymerase using dNTPs by our sequencing assay. We observed an error rate of .011 for *Taq* polymerase with dNTPs (data not shown). Since the *Taq*-polymerized fragments were PCR amplified for 35 rounds prior to cloning and sequencing, it is necessary to divide the observed error rate by 35 to compensate for errors introduced during PCR. This gives a corrected error measurement of 3×10^{-4}, consistent with previously published values of $\sim 2 \times 10^{-4}$ obtained under similar conditions (7,8). Therminator DNA polymerization fidelity has not been rigorously measured previously. We observed a corrected error rate of 2×10^{-2} after compensating for 35 rounds of PCR (data not shown). This is significantly higher than the *Taq* error rate in our assay ($p < 0.03$), and may be due to a less discriminating active site, paralleling Therminator's ability to incorporate modified nucleotides (9–11).

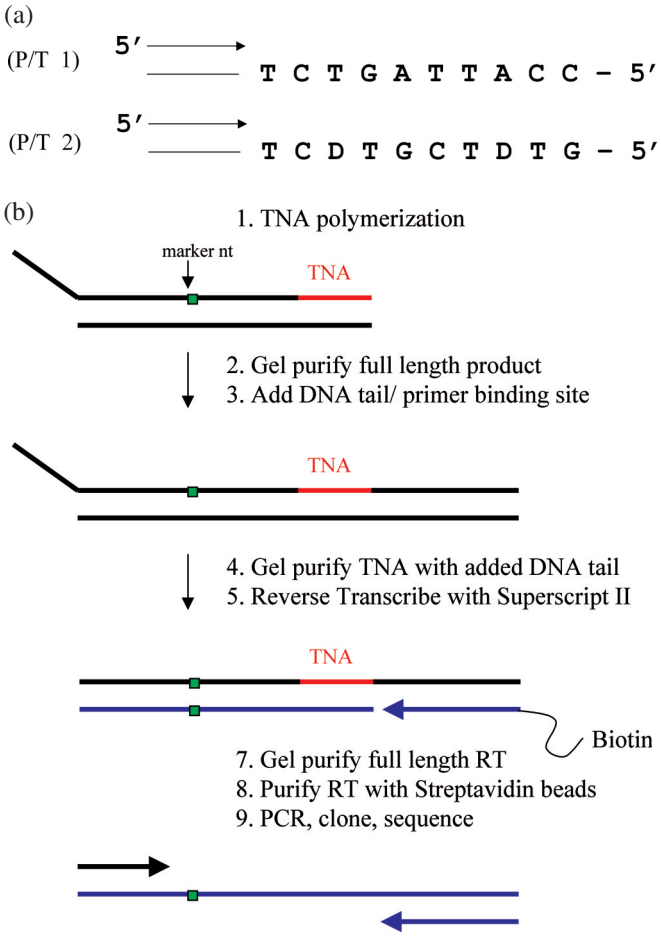

Figure 2. TNA sequencing design. (**a**) Primer/template combinations used for TNA polymerization sequencing. (**b**) Sequencing scheme.

High fidelity of TNA synthesis

We initially measured Therminator TNA polymerization fidelity under conditions we had previously estimated to be most favorable for fidelity (3), namely short reaction times, a reaction temperature of 75°C and highly skewed tNTP concentrations. Under these conditions, Therminator-catalyzed TNA polymerization had a corrected error rate of $\sim 1 \times 10^{-3}$, below the level that can be accurately measured by our assay (Table 2, column 1). The raw error rate for this reaction was not significantly different from the raw error rate of *Taq* polymerase with dNTPs ($p > 0.1$). However, our assay measures *Taq* fidelity after most of the primer has been fully extended. Thus, most of the mismatches that occurred were incorporated into the full-length product that was eventually sequenced. This is not the case for the Therminator-catalyzed TNA reactions where many of the mutated transcripts were eliminated by stopping the reaction with a low fraction of fully extended primer and gel purifying only the full-length product. Thus, the TNA reactions are enriched for faithful transcripts whereas the DNA reactions are not.

For comparison, we also sequenced TNA polymerization products from a different primer/template pair (Figure 2a, P/T 2) that contained diaminopurine in place of adenine. Again, Therminator fidelity was above 99% (Table 2, columns

Table 2. TNA polymerization fidelity

	1 Therm/tNTPs P/T1 75°C/1 h	2 Therm/tNTPs P/T2 75°C/20 min	3 Therm/tNTPs P/T2 75°C/40min	4 Therm/tNTPs P/T2 55°C/10min	5 Therm/*tNTPs P/T2 55°C/40min	6 Therm/*tNTPs P/T2 55°C/3 h
Transitions	1	4	2	3	4	8
Transversions	1	0	0	1	0	22
Insertions	0	0	0	0	0	0
Deletions	0	0	0	1	0	1
Errors	2	4	2	5	4	31
Total bases	240	272	275	309	356	644
Raw error-rate	8×10^{-3}	15×10^{-3}	7×10^{-3}	16×10^{-3}	11×10^{-3}	48×10^{-3}
Corrected error-rate	1×10^{-3}	8×10^{-3}	1×10^{-3}	9×10^{-3}	4×10^{-3}	41×10^{-3}

Raw error rate was determined by dividing the number of substitutions and frame-shift errors by the total number bases sequenced. Corrected error rate = raw error rate minus the error rate of *Taq* over 35 rounds of PCR. Values of 1×10^{-3} (columns 1 and 3) indicate error rates below the level of detection in this assay. 'tNTPs' indicates 120 μM tDTP, 60 μM tTTP, 18 μM tCTP and 2 μM tGTP. '*tNTPs' indicates 30 μM tDTP, 30 μM tTTP, 30 μM tCTP and 4 μM tGTP.

Table 3. TNA polymerization fidelity by position and sequence

Therminator + tNTPs, 75°C, 1 h										
Total	D (24)	G (24)	D (24)	C (24)	T (24)	D (24)	D (24)	T (24)	G (24)	G (24)
Errors		C							D	
Therminator + tNTPs, 75°C, 20 min										
Total	D (28)	G (28)	T (28)	D (28)	C (28)	G (28)	D (28)	T (28)	D (24)	C (24)
Errors					TT	D			G	
Therminator + tNTPs, 75°C, 40 min										
total:	D (30)	G (29)	T (30)	D (30)	C (30)	G (29)	D (29)	T (29)	D (20)	C (19)
errors:		D			T					
Therminator + tNTPs, 55°C, 10 min										
total:	D (32)	G (33)	T (33)	D (33)	C (33)	G (32)	D (32)	T (32)	D (25)	C (24)
errors:	T	D			T				G_	
Therminator + *tNTPs, 55°C, 40 min										
Total	D (37)	G (37)	T (37)	D (37)	C (37)	G (37)	D (37)	T (37)	D (30)	C (30)
Errors		D			TT					
Therminator + *tNTPs, 55°C, 3 h										
Total	D (70)	G (70)	T (66)	D (68)	C (68)	G (66)	D (66)	T (66)	D (52)	C (51)
Errors	_	DDCD	DDDD		TTTG					
			DDDD		GGGG					
			DDD		GGGG					

Numbers in parentheses indicate the total number of bases sequenced at each position. '*tNTPs' indicates 30 μM tDTP, 30 μM tTTP, 30 μM tCTP and 4 μM tGTP.

2 and 3). No difference was seen in tTTP insertion fidelity between P/T 1 and P/T 2, indicating that adenine and diaminopurine are equally faithful template residues (Table 3), within the limits of this assay. Furthermore, fidelity was general with respect to sequence context. For example, tGTP incorporation immediately downstream of all four different nucleobases and directly preceding three of the four nucleobases occurred with similar fidelity, using data from both primer/template pairs 1 and 2 (Table 3).

Fidelity under conditions optimized for efficient polymerization

Given that TNA polymerization is most efficient at 55°C, we compared fidelity at 55°C versus 75°C. Previously, we had determined by the dropout assay that fidelity was higher at 75°C (3), possibly because the higher temperature decreased polymerase extension from mismatches more than from correctly matched termini. However, by the TNA sequencing assay, the fidelity was essentially equivalent at 55 and 75°C (Table 2, columns 2 and 4) ($p > 0.05$). The fact that fidelity at 55°C as determined by sequencing is much higher than that measured by the dropout assay indicates that Therminator discriminates strongly against incorporating an incorrect base when the correct nucleotide is present.

Next, we tested the effect of using more equimolar tNTP ratios, which would allow us to conserve tNTPs during TNA synthesis. The dropout fidelity assay had indicated that tDTP incorporation was very error prone, presumably because of tGTP wobble-pairing with thymine. Our original concentrations therefore included 60-fold more tDTP than tGTP. With these concentrations, a large amount of tDTP would be needed for TNA synthesis. We changed the tNTP concentrations from 120 μM tDTP, 31 μM tTTP, 18 μM tCTP and 2 μM tGTP to 30 μM tDTP, tTTP and tCTP, and 4 μM tGTP. The TNA polymerization fidelity remained above 99% with the new nucleotide concentrations (Table 2, column 5) ($p > 0.1$), indicating that heavily biased tNTP concentrations are unnecessary to maintain accurate TNA synthesis.

The yield of full-length TNA products increases with time, but we expect transcripts completed at later times to have a higher error rate, since molecules that were paused due to a misincorporation event would eventually be extended to full-length product. Indeed, when we incubated a reaction at 55°C for 3 h (~20% full-length) instead of only 40 min (~10% full-length), the error rate increased 10-fold from 4×10^{-3} to 4×10^{-2} (Table 2, columns 5 and 6) ($p < 0.001$). Surprisingly, the increased error rate was due to a drastic increase in transversion mutations (Table 3). It is possible that most of the mismatches incorporated by Therminator are actually

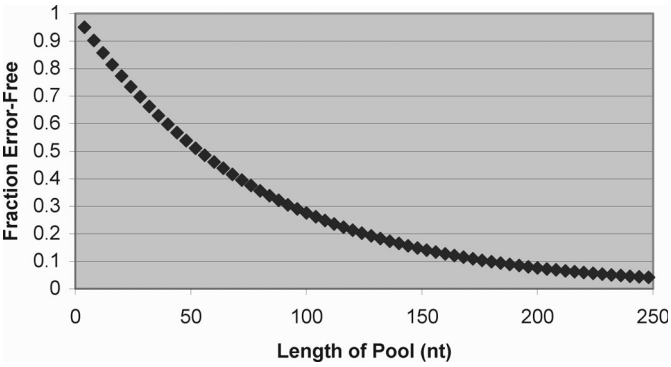

Figure 3. Fraction of error-free TNA transcripts as a function of pool length. Data points were generated using the fidelity for individual tNTPs, assuming equal amounts of all 4 nt in the final product and a short reaction time. [tNTP] = 30 μM tDTP, tTTP, tCTP and 4 μM tGTP, temperature = 55°C.

transversions, but since they are extended so slowly compared with transition mutations, they only accumulate in full-length product after long incubations. Assuming equimolar representation of all four TNA bases, a short reaction time with 30 μM tDTP, tTTP, and tCTP and 4 μM tGTP at 55°C would yield a 60mer TNA pool with about half of the molecules free of mutations (Figure 3). If the TNA polymerization reaction is allowed to incubate longer, the majority of the later transcripts would contain mutations since the fidelity would drop to 96% or less (Table 2). Thus, the additional transcripts generated by a long incubation time would merely dilute the TNA pool with erroneous material, and their production should be minimized by keeping reaction times as short as possible.

The 10 bp sequenced in our study constitute nearly a complete turn of a DNA double helix (12). X-ray crystallography has shown that DNA polymerases make extensive contacts with the primer/template up to 5 nt upstream of the primer terminus (13,14). Therefore, positional analysis of the sequencing results should give information regarding fidelity when the polymerase is partly contacting the DNA section of the primer versus when it is completely contacting the TNA portion. In all of the sequencing results, we observed no decrease in fidelity as the polymerase reached the last few template nucleotides (Table 3), indicating that longer TNA polymerization should remain just as faithful as polymerization near the DNA primer terminus. This is consistent with our previous kinetic studies that indicated that tNTP incorporation remained highly efficient whether a DNA or TNA primer was used (4).

DISCUSSION

We have determined that TNA synthesized by the Therminator DNA polymerase can have an error rate of <1%, as long as full-length material is purified after a short reaction time. The high fidelity obtained under these conditions means that it should be practical to use TNA pools as large as 200 nt for *in vitro* selection (Figure 3), provided that Therminator could polymerize enough 200 nt-long TNA. Such long pool lengths increase the probability of finding more complex active molecules within a given pool diversity (15,16). Our previous efforts to estimate TNA polymerization fidelity utilizing a dropout assay overestimated the error rate significantly and predicted that only pools smaller than 70 nt could be used for TNA selections (3). It is likely that misincorporation was artificially forced due to the absence of the correct tNTP, accounting for much of the apparent infidelity in the dropout assay.

The reaction parameter most essential for obtaining accurately synthesized TNA is the reaction time. Long incubations allow mismatches to be extended and incorporated into full-length product, whereas short incubations limit mismatch extension and allow the truncated, error-prone molecules to be purified away. Since shorter incubation times translate into lower yields of full-length product, a balance must be struck between yield and fidelity. This can be done by first performing a time course of the reaction and then stopping the actual reaction when the polymerization starts to slow down significantly. For the short templates used in these experiments, this occurred when 10–25% of the primer was converted to full-length product (data not shown), but the percentage will be lower for longer template sequences due to a higher probability of misincorporation per template.

Remarkably, TNA polymerization and polymerization fidelity seem to be relatively unaffected by the length of the transcript even though the distance between phosphorus atoms across a TNA nucleotide within a DNA duplex is 0.5–0.8 Å shorter than the corresponding distance for a DNA nucleotide (17). An elevated level of frame-shift mutations due to slippage might be expected if helical strain builds up along the TNA–DNA heteroduplex. Although we cannot rule out the possibility that a higher frequency of frame-shift mutations would be seen on longer templates, we did not observe an increase in frame-shift mutations as polymerization proceeded downstream from the DNA primer terminus. Moreover, previous kinetic measurements showed that TNA polymerization from a primer ending with 5 TNA residues is more efficient than TNA polymerization from an all DNA primer (4), suggesting that the active site of the Therminator polymerase is able to adapt to the distorted TNA/DNA structure, or that the polymerase alters the TNA conformation when it is at or near the active site.

It would be useful to improve the ability of the Therminator polymerase to discriminate against incorrect nucleotides at the incorporation step, because this would increase product yield by reducing the amount of truncated transcripts. Furthermore, the resulting higher fidelity would increase the fraction of error-free TNA transcripts. In addition to the conditions tested here, we tried lower $MgCl_2$ concentrations and lower pH, both of which improve the fidelity of *Taq* polymerase (18). Unfortunately, these conditions drastically reduce TNA polymerization yields (data not shown). It may be possible to evolve the Therminator polymerase into a more faithful and efficient TNA polymerase by *in vitro* selection (19–21) or rational design (22). Since Therminator itself was identified by testing a tiny fraction of all possible mutants (9,10), it is likely that other beneficial mutations exist in sequence space.

ACKNOWLEDGEMENTS

This work was supported in part by grants from the NIH to J.W.S. (GM053936) and to L.W.M. NIH (GM067986). 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. Funding to pay the Open Access publication charges for this article was provided by the Department of Molecular Biology, Massachusetts General Hospital.

Conflict of interest statement. None declared.

REFERENCES

1. Eschenmoser,A. (1999) Chemical etiology of nucleic acid structure. *Science*, **284**, 2118–2124.
2. Schoning,K., Scholz,P., Guntha,S., Wu,X., Krishnamurthy,R. and Eschenmoser,A. (2000) Chemical etiology of nucleic acid structure: the alpha-threofuranosyl-(3′→2′) oligonucleotide system. *Science*, **290**, 1347–1351.
3. Ichida,J.K., Zou,K., Horhota,A., Yu,B., McLaughlin,L.W. and Szostak,J.W. (2005) An *in vitro* selection system for TNA. *J. Am. Chem. Soc*., **127**, 2802–2803.
4. Horhota,A., Zou,K., Ichida,J.K., Yu,B., McLaughlin,L.W., Szostak,J.W. and Chaput,J.C. (2005) Kinetic analysis of an efficient DNA-dependent TNA polymerase. *J. Am. Chem. Soc*., **127**, 7427–7434.
5. Boyer,J.C., Bebenek,K. and Kunkel,T.A. (1996) Analyzing the fidelity of reverse transcription and transcription. *Methods Enzymol*., **275**, 523–537.
6. Chaput,J.C., Ichida,J.K. and Szostak,J.W. (2003) DNA polymerase-mediated DNA synthesis on a TNA template. *J. Am. Chem. Soc*., **125**, 856–857.
7. Keohavong,P. and Thilly,W.G. (1989) Fidelity of DNA polymerases in DNA amplification. *Proc. Natl Acad. Sci. USA*, **86**, 9253–9257.
8. Tindall,K.R. and Kunkel,T.A. (1988) Fidelity of DNA synthesis by the Thermus aquaticus DNA polymerase. *Biochemistry*, **27**, 6008–6013.
9. Gardner,A.F. and Jack,W.E. (1999) Determinants of nucleotide sugar recognition in an archaeon DNA polymerase. *Nucleic Acids Res*., **27**, 2545–2553.
10. Gardner,A.F. and Jack,W.E. (2002) Acyclic and dideoxy terminator preferences denote divergent sugar recognition by archaeon and *Taq* DNA polymerases. *Nucleic Acids Res*., **30**, 605–613.
11. Gardner,A.F., Joyce,C.M. and Jack,W.E. (2004) Comparative kinetics of nucleotide analog incorporation by vent DNA polymerase. *J. Biol. Chem*., **279**, 11834–11842.
12. Wing,R., Drew,H., Takano,T., Broka,C., Tanaka,S., Itakura,K. and Dickerson,R.E. (1980) Crystal structure analysis of a complete turn of B-DNA. *Nature*, **287**, 755–758.
13. Kiefer,J.R., Mao,C., Braman,J.C. and Beese,L.S. (1998) Visualizing DNA replication in a catalytically active Bacillus DNA polymerase crystal. *Nature*, **391**, 304–307.
14. Li,Y., Korolev,S. and Waksman,G. (1998) Crystal structures of open and closed forms of binary and ternary complexes of the large fragment of *Thermus aquaticus* DNA polymerase I: structural basis for nucleotide incorporation. *EMBO J*., **17**, 7514–7525.
15. Sabeti,P.C., Unrau,P.J. and Bartel,D.P. (1997) Accessing rare activities from random RNA sequences: the importance of the length of molecules in the starting pool. *Chem. Biol*., **4**, 767–774.
16. Gevertz,J., Gan,H.H. and Schlick,T. (2005) *In vitro* RNA random pools are not structurally diverse: a computational analysis. *RNA*, **11**, 853–863.
17. Wilds,C.J., Wawrzak,Z., Krishnamurthy,R., Eschenmoser,A. and Egli,M. (2002) Crystal structure of a B-form DNA duplex containing (L)-alpha-threofuranosyl (3′→2′) nucleosides: a four-carbon sugar is easily accommodated into the backbone of DNA. *J. Am. Chem. Soc*., **124**, 13716–13721.
18. Eckert,K.A. and Kunkel,T.A. (1990) High fidelity DNA synthesis by the *Thermus aquaticus* DNA polymerase. *Nucleic Acids Res*., **18**, 3739–3744.
19. Fa,M., Radeghieri,A., Henry,A.A. and Romesberg,F.E. (2004) Expanding the substrate repertoire of a DNA polymerase by directed evolution. *J. Am. Chem. Soc*., **126**, 1748–1754.
20. Ghadessy,F.J., Ramsay,N., Boudsocq,F., Loakes,D., Brown,A., Iwai,S., Vaisman,A., Woodgate,R. and Holliger,P. (2004) Generic expansion of the substrate spectrum of a DNA polymerase by directed evolution. *Nat. Biotechnol*., **22**, 755–759.
21. Holmberg,R.C., Henry,A.A. and Romesberg,F.E. (2005) Directed evolution of novel polymerases. *Biomol. Eng*., **22**, 39–49.
22. Suzuki,M., Yoshida,S., Adman,E.T., Blank,A. and Loeb,L.A. (2000) *Thermus aquaticus* DNA polymerase I mutants with altered fidelity. Interacting mutations in the O-helix. *J. Biol. Chem*., **275**, 32728–32735.
23. Wu,X., Delgado,G., Krishnamurthy,R. and Eschenmoser,A. (2002) 2,6-Diaminopurine in TNA: effect on duplex stabilities and on the efficiency of template-controlled ligations. *Org. Lett*., **4**, 1283–1286.