High fidelity TNA synthesis by Therminator polymerase

Justin K. Ichida\textsuperscript{1,2,3}, Allen Horhota\textsuperscript{4}, Keyong Zou\textsuperscript{1,2,3}, Larry W. McLaughlin\textsuperscript{4} and Jack W. Szostak\textsuperscript{1,2,3,*}

\textsuperscript{1}Howard Hughes Medical Institute and \textsuperscript{2}Department of Molecular Biology, Massachusetts General Hospital, Boston, MA 02114, USA, \textsuperscript{3}Department of Genetics, Harvard Medical School, Boston, MA 02115, USA and \textsuperscript{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 \textit{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 \textit{in vitro} selection with TNA libraries of at least 200 nt in length.

INTRODUCTION

(\textsuperscript{3’}–\textsuperscript{5’}) \textalpha-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 \textit{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 T7 RNA 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 \textit{\textasciitilde}1 \times 10^{-4}, meaning that RNA \textit{in vitro} selection pools of up to 300 nt are \textasciitilde97\% error-free before PCR amplification (5). Typically, \textit{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 \textit{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 functional 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...
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'-GTGGTAGCCAGGGTGTGAAACGAGAGGGCTGCAGACAGAGGNN14CTAC-ACTTGGCCAGCCCTAGCG-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 (GTGGTAGCCAGGGTGTGAAACGAGAGGGCTGCAGACAGAGGNN14CTAC-ACTTGGCCAGCCCTAGCG-3') using either four natural dNTPs or with 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 biotinylated primer 23 (5'-photocleavable biotin-CGCTAGGGCGCTGGCAAG-GGTG-3') using either all four natural dNTPs or with 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′-GTGGTGACGGCCGAGGTTGTGAACCGAGTGCAGCTGC-3′) was annealed to 100 nM ssDNA pool whose sequences contained either adenine or dianaminopurine in 1.1× Thermopol buffer (New England Biolabs) plus 60 μM dTTP, 60 μM dTTP, 18 μM dCTP and 2 μM dGTP. TNA polymerization reactions were started by the addition of 2 mM or no MgCl₂, 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′-GTGGTGACGGCCGAGGTTGTGAACCGAGTGCAGCTGC-3′) labeled at the 5′ end by incubation with [γ-32P]ATP and T4 Polynucleotide Kinase (New England Biolabs) was annealed to 300 nM template ‘T1’ (5′-CCAT TAGTCTCCCTCTCCTATAGTGAGTCTGACCTCCTTTTC TATTGTGAGTTG-3′) or template ‘T2’ (5′-GTGGTGACGGCCGAGGTTGTGAACCGAGTGCAGCTG CTCCCTCTCCTATAGTGAGTCTGACCTCCTTTTC TATTGTGAGTTG-3′) in 1.1× Thermopol buffer. Regions of complementarity between the primers 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 dTTP, 60 μM dCTP, 4 μM dGTP. 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′-AAAAAAACCTTCCCTCCCTTTTCTTCTCTTACCCGCTTTCTCA CACCTCGCCGCTACACC-3′) or template ‘T2B’ (5′-AAAAAAACCTTCCCTCCCTTTTCTTCTCTTACCCGCTTTCTCAA CACCTCGCCGCTACACC-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-CTTCCCTCCCTTTTCTTCTCT-3′) in 1× First Strand buffer (Superscript II, Invitrogen) plus 125 μM dNTPs. Aliquots containing 10 mM DTT, 1.5 mM MnCl₂ 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′-GTGGTGACGGCCGAGGTTGTGTG-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 TOP10 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

\[
P(x; n; \theta) = \binom{n}{x} \theta^x (1 - \theta)^{n-x},
\]

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 dianaminopurine in the template increases the efficiency of TNA polymerization (4), we tested polymerization activity on DNA templates containing either adenine or dianaminopurine. In addition, we tested polymerization in the presence or absence of 1 mM MnCl₂. 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 dianaminopurine-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,
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 Mn²⁺ (Figure 1d). The effect of Mn²⁺ on polymerization was less pronounced, with only a slight difference between 55°C reactions on diaminopurine templates with or without Mn²⁺ (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 0.11 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 $3 \times 10^{-4}$, 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 $2 \times 10^{-2}$ 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).

### Table 1. Percentage of full-length pool TNA polymerization

<table>
<thead>
<tr>
<th>Template bases</th>
<th>Temperature (°C)</th>
<th>MnCl₂ (mM)</th>
<th>% Full length</th>
</tr>
</thead>
<tbody>
<tr>
<td>DTGC</td>
<td>55</td>
<td>1</td>
<td>10</td>
</tr>
<tr>
<td>DTGC</td>
<td>55</td>
<td>0</td>
<td>5</td>
</tr>
<tr>
<td>DTGC</td>
<td>75</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>ATGC</td>
<td>55</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>ATGC</td>
<td>55</td>
<td>0</td>
<td>&lt;1</td>
</tr>
<tr>
<td>ATGC</td>
<td>75</td>
<td>0</td>
<td>&lt;1</td>
</tr>
</tbody>
</table>

### 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 TNP 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

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

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 Tag over 35 rounds of PCR. Values of 1 x 10^{-3} (columns 1 and 3) indicate error rates below the level of detection in this assay. tNTPs indicates 30 μM tDTP, 30 μM tTTP, and 30 μM tCTP. *tNTPs indicates 30 μM tDTP, 30 μM tTTP, 30 μM tCTP, and 4 μM tGTP.

Table 3. TNA polymerization fidelity by position and sequence

<table>
<thead>
<tr>
<th></th>
<th>Therminator + tNTPs, 75°C, 1 h</th>
<th>Therminator + tNTPs, 75°C, 20 min</th>
<th>Therminator + tNTPs, 75°C, 40 min</th>
<th>Therminator + tNTPs, 55°C, 10 min</th>
<th>Therminator + tNTPs, 55°C, 40 min</th>
<th>Therminator + tNTPs, 33°C, 3 h</th>
</tr>
</thead>
<tbody>
<tr>
<td>Transitions</td>
<td>T (28)</td>
<td>C (28)</td>
<td>T (30)</td>
<td>C (33)</td>
<td>T (37)</td>
<td>C (37)</td>
</tr>
<tr>
<td>Transversions</td>
<td>T (28)</td>
<td>C (28)</td>
<td>T (30)</td>
<td>C (33)</td>
<td>T (37)</td>
<td>C (37)</td>
</tr>
<tr>
<td>Insertions</td>
<td>T (28)</td>
<td>C (28)</td>
<td>T (30)</td>
<td>C (33)</td>
<td>T (37)</td>
<td>C (37)</td>
</tr>
<tr>
<td>Deletions</td>
<td>T (28)</td>
<td>C (28)</td>
<td>T (30)</td>
<td>C (33)</td>
<td>T (37)</td>
<td>C (37)</td>
</tr>
<tr>
<td>Errors</td>
<td>T (28)</td>
<td>C (28)</td>
<td>T (30)</td>
<td>C (33)</td>
<td>T (37)</td>
<td>C (37)</td>
</tr>
<tr>
<td>Total bases</td>
<td>240</td>
<td>272</td>
<td>275</td>
<td>309</td>
<td>356</td>
<td>644</td>
</tr>
<tr>
<td>Raw error-rate</td>
<td>8 x 10^{-3}</td>
<td>15 x 10^{-3}</td>
<td>7 x 10^{-3}</td>
<td>16 x 10^{-3}</td>
<td>11 x 10^{-3}</td>
<td>48 x 10^{-3}</td>
</tr>
<tr>
<td>Corrected error-rate</td>
<td>1 x 10^{-3}</td>
<td>8 x 10^{-3}</td>
<td>1 x 10^{-3}</td>
<td>9 x 10^{-3}</td>
<td>4 x 10^{-3}</td>
<td>41 x 10^{-3}</td>
</tr>
</tbody>
</table>

Table 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 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 55°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 these concentrations, a large amount of tDTP would be needed.
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).

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).

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 MgCl₂ 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