Bioorganic & Medicinal Chemistry 17 (2009) 3782–3788

Contents lists available at ScienceDirect

Bioorganic & Medicinal Chemistry

journal homepage: www.elsevier.com/locate/bmc

Transcription and reverse transcription of artificial nucleic acids involving backbone modification by template-directed DNA polymerase reactions

Masayasu Kuwahara [a,b,*], Hidetoshi Takeshima [a], Jun-ichi Nagashima [a], Satoshi Minezaki [a], Hiroaki Ozaki [a], Hiroaki Sawai [a]

[a] Department of Chemistry and Chemical Biology, Graduate School of Engineering, Gunma University, Gunma 376-8515, Japan
[b] PRESTO, Japan Science and Technology Agency (JST), Chiyodaku, Tokyo 102-0075, Japan

ARTICLE INFO

Article history:
Received 24 February 2009
Revised 21 April 2009
Accepted 22 April 2009
Available online 3 May 2009

Keywords:
Polymerase reaction
Backbone modification
Dinucleotide mimic
Artificial nucleic acids

ABSTRACT

Oligodeoxyribonucleotides (ODN) where the phosphodiester linkage had been replaced with an amide-type linker [–CH_2C═ONH–] or an amine-type linker [–CH_2CH_2NH–] were synthesized to investigate the effect of these backbone modifications on polymerase reactions. In addition, a triphosphate analogue of thymidine dinucleotide with the amide-type linker was synthesized and enzymatic insertion of the amide linkage into ODN was attempted using this analogue for the polymerase reaction. Primer extension reactions using three types of thermostable DNA polymerases, *KOD*(*exo*-), *Vent*(*exo*-) and *Taq* were performed for the assays. Analysis of these data indicate that (i) the polymerase reaction tends to be affected much more by insertion of the cationic flexible amine-type linker than by insertion of the neutral rigid amide-type linker; (ii) the backbone modification has a greater effect on the polymerase reaction when it is adjacent to the 3′-end of a primer as the elongation terminus than when it is on the template, as well as in base or sugar modification; (iii) although the modified linker in the modified DNA template is passed beyond by the polymerase, it still affects the extension reaction several bases downstream from its location; (iv) the modified linker in the template, in some cases, also affects the extension reaction upstream from its location; (v) further improvement of the chemical structure is required for dinucleotide-mimic incorporation.

1. Introduction

Enzymatic transcription and reverse transcription of artificial nucleic acids would be an important technique to allow the application of artificial nucleic acids[1–10] to random screening methods of nucleic acids such as systematic evolution ofl igands by exponential enrichment (SELEX),[11–17] Non-SELEX selection,[18] and one-step selection.[19] These random screening methods would be able to identify various functional nucleic acids such as aptamers and ribozymes with functions similar to antibodies and enzymes, which could be useful not only as research tools for molecular biology but also as diagnostic agents, therapeutic drugs etc. For these applications, the functional nucleic acids have some advantages compared to the corresponding proteins: (i) animals are not required for their production, (ii) large-scale preparation by organic synthesis at low costs is possible, (iii) dry preservation can be used at normal temperatures etc. However, DNA and RNA are rapidly digested by nucleases; nuclease resistance is one of the most important properties for the biological use of the aforementioned functional nucleic acids. Modifications of phosphate or sugar moieties have large effects on the polymerase reaction in many cases, but these effects are also expected to confer nuclease resistance to nucleic acids. Recently, polymerase reactions using artificial nucleic acids with phosphate or sugar modification have been reported; phosphorothioate-oligonucleotide (PS-ODN)[20] and boranophosphate-oligodeoxynucleoside (BH_3^--ODN)[21] are examples of phosphate modification, while locked nucleic acid (LNA),[22] glycol nucleic acid (GNA),[23] cyclohexene nucleic acid (CeNA)[24] and threose nucleic acid (TNA)[25] are examples of sugar modifications.

Previously, we synthesized 2′,4′-bridged nucleoside-5′-triphosphates and templates containing 2′,4′-bridged nucleotides (BNA templates), and using these nucleotides with sugar modification we have demonstrated that modification to the extending strand decreased the catalytic efficiency of DNA polymerase to a far greater extent than modification to the template strand.[26–28] That is, transcription of BNA from DNA was found to be difficult, but reverse transcription of BNA to DNA was possible when certain types of DNA polymerase were used. Applying artificial nucleic acids to SELEX requires both transcription and reverse transcription. However, only conversion of the artificial nucleic acid to natural DNA is needed for non-SELEX selection or one-step selection. In the current study, modification of the phosphate moiety[29–31] was expected to enhance nuclease resistance of nucleic acids much

* Corresponding author. Tel./fax: +81 277 30 1222.
E-mail address: kuwahara@chem-bio.gunma-u.ac.jp (M. Kuwahara).

0968-0896/$ - see front matter © 2009 Elsevier Ltd. All rights reserved.
doi:10.1016/j.bmc.2009.04.045

more than that of the sugar moiety; therefore, we synthesized primers and templates incorporating the amide-type linker '*h*', or the amine-type linker '*j*' (Table 1), and the triphosphate of dinucleotide analogue **3** with the amide-type linker (Fig. 1). We then analyzed how the insertion of phosphate modification to the primer and template affects polymerase reactions and also studied the effects of the chemical structures of modification. We also examined the influence of the type of DNA polymerase on the production of DNA. To enzymatically insert the amide linkage into the elongation strand, the substrate properties of the dinucleotide mimic[32] for the thermostable DNA polymerases were investigated.

2. Results and discussion

2.1. Attempt to incorporate the dinucleotide mimic

To incorporate the dinucleotide mimic, primer extension reactions were performed using the triphosphate analogue **3** with three different enzymes at two different enzyme concentrations. The higher enzyme concentrations were set 10-times higher than the lower enzyme concentrations. Although it is known that 5′-triphosphateterminated polynucleotide is hydrolyzed by DNA polymerase,[33] the mimic might be recognized as a substrate because the amide linker *h* could resist the hydrolysis. Unfortunately, no elongated product could be observed even at the higher concentration for any type of polymerase used (Figs. 2 and S1).

2.2. Primer extension experiments using modified primers

Next, we investigated the effects of phosphate modification in the primer on DNA polymerization. Primers P3 and P4 contain linkers *h* and *j*, respectively, which replace a phosphodiester linkage closest to the 3′-end (Table 1). The primer is elongated by 5 nt to give the full-length product. In Figure 3, two bands were clearly observed at lane 2 as the positive controls of the reaction with *KOD(exo-)*; here, the lower band corresponds to the full-length product, and the upper band corresponds to the product yielded by a single-nucleotide addition to the full-length product due to nontemplated nucleotide addition at the 3′-end by the action of the polymerases used.[34] The bands of other positive controls seen at lanes 4 and 6 are also one nucleotide longer than the full-length product.

Table 1
Primers and templates used in this study

Primers	
P1	5′-FAM-GGCGTTGAGTGAGTGAATGAGTGAGT-3′
P2	5′-FAM-CGGAGCTGAATGAAGCCATACCTTTT-3′
P3	5′-FAM-CGGAGCTGAATGAAGCCATACCTTT$_h$T-3′
P4	5′-FAM-CGGAGCTGAATGAAGCCATACCTTT$_j$T-3′
Templates	
T1	3′-CCGCAACTCACTCACTTACTCACTCAAAAAAAAA-5′
T2	3′-GCCTCGACTTACTTCGGTATGGAAAACTGTC-5′
T3	3′-CCGCAACTCACTCACTTACTCACTCATTTTTTTTTT-5′
T4	3′-CCGCAACTCACTCACTTACTCACTCAT$_h$TTTTTTTTT-5′
T5	3′-CCGCAACTCACTCACTTACTCACTCAT$_h$TTTT$_h$TTTTT-5′
T6	3′-CCGCAACTCACTCACTTACTCACTCAT$_h$TTTTT$_h$TTTT-5′
T7	3′-CCGCAACTCACTCACTTACTCACTCAT$_h$TTTTTT$_h$TTT-5′
T8	3′-CCGCAACTCACTCACTTACTCACTCAT$_j$TTTTTTTTT-5′
T9	3′-CCGCAACTCACTCACTTACTCACTCAT$_j$TTTT$_j$TTTTT-5′
T10	3′-CCGCAACTCACTCACTTACTCACTCAT$_j$TTTTT$_j$TTTT-5′
T11	3′-CCGCAACTCACTCACTTACTCACTCAT$_j$TTTTTT$_j$TTT-5′
T12	3′-CCGCAACTCACTCACTTACTCACTCATTGGATTGACC-5′
T13	3′-CCGCAACTCACTCACTTACTCACTCAT$_h$TGGATTGACC-5′
T14	3′-CCGCAACTCACTCACTTACTCACTCATTGGAT$_h$TGACC-5′
T15	3′-CCGCAACTCACTCACTTACTCACTCAT$_h$TGGAT$_h$TGACC-5′

In Figure 3A, a fraction of the full-length product was observed only when *KOD*(*exo-*) was used (lane 3). Since the band seems to migrate slightly slower than the lower band of the positive control, it appeared between the upper and lower bands at lane 2. This may be due to linker *h* inserted in the elongation strand. The extension largely stopped after the first nucleotide was incorporated when *KOD*(*exo-*) and *Vent*(*exo-*) were used (lanes 3 and 5), while the extension hardly proceeded when *Taq* was used (lane 7). Conversely in Figure 3B, the reaction using *Taq* proceeded most efficiently among these three reactions (lanes 3, 5 and 7), although the extension largely stopped after the first-nucleotide incorporation and the full-length product could not be observed in any reactions examined. This tendency was also observed if the enzyme concentration was reduced 10 times (*Fig.* S2). It is interesting that the subtle differences in the chemical structures between linkers *h* and *j* adjacent to the elongation terminus caused reversal of the extension efficiency depending on the strain of polymerase used; *KOD(exo-)* and *Vent*(*exo-*) belong to family B, and *Taq* belongs to family A. That is, in case of linker *h*, *KOD(exo-)* and *Vent*(*exo-*) could overcome the first-nucleotide-incorporation step more easily than *Taq*. However, in case of linker *j*, *Taq* could overcome this step more easily than the others. This may reflect the difference in the recognition pattern of the first phosphodiester linkage, numbered from the elongation terminus, by the enzyme. In both linkers, the greatest barrier to the extension was the second-nucleotide-incorporation step for all polymerases used, indicating that the second phosphodiester linkage (numbered from the elongation terminus) could be critically recognized by the enzyme, which would lead to the formation of the proper conformation of the active site.

2.3. Primer extension experiments using modified templates

Finally, we investigated the effects of phosphate-modification in the template on DNA polymerization. Templates T3–T11 include the homosequence with 11 successive Ts from the 5′-ends, while templates T12–T15 include the mixed sequence with CCAGTTAGGTT from the 5′-ends (Table 1). Templates T3 and T12 consisted of all natural nucleotides and were used for positive control reactions. Template T4 contains only a single *h* linker at the position between the two residues opposite the first and second nucleotides incorporated. Templates T5, T6 and T7 contain two *h* linkers at intervals of three, four and five phosphodiester linkages, respectively. Templates T8, T9, T10 and T11 contain linker *j* instead of linker *h* in templates T4, T5, T6 and T7 at the same position, respectively. Templates T13 and T14 contain only a single *h* linker at the position between the two residues opposite the first and second nucleotides, and opposite the sixth and seventh nucleotides incorporated, respectively. Template T15 contains two *h* linkers in templates at the same position of T7.

We examined the effects of insertion intervals and chemical structures of modified linkers in experiments using templates of the homosequence T3–T11 and *KOD(exo-)* DNA polymerase (Fig. 4). The *y*-axis of the graphs in Figure 5 indicates the accumulation of the product yields, and the *x*-axis indicates the number of residues (dAs) incorporated into the extending strand. The product yields were obtained from the band intensities on the gel images (Fig. 4). In these graphs, the graph line rises steeply where the ratio of the extension stop is higher. In this experiment, the primer was elongated by 11 nt to give the full-length product; however, incorporation of more than 11 nt was observed in some cases, for example, the positive controls. This may have occurred because of DNA template slippage[35,36] and/or the aforementioned nontemplated nucleotide addition.

At the lower enzyme concentration, the polymerase passed beyond the first barrier, linker *h*, in all cases when templates T4–T7 were used (graph A1). However, in the case of T8–T11, the

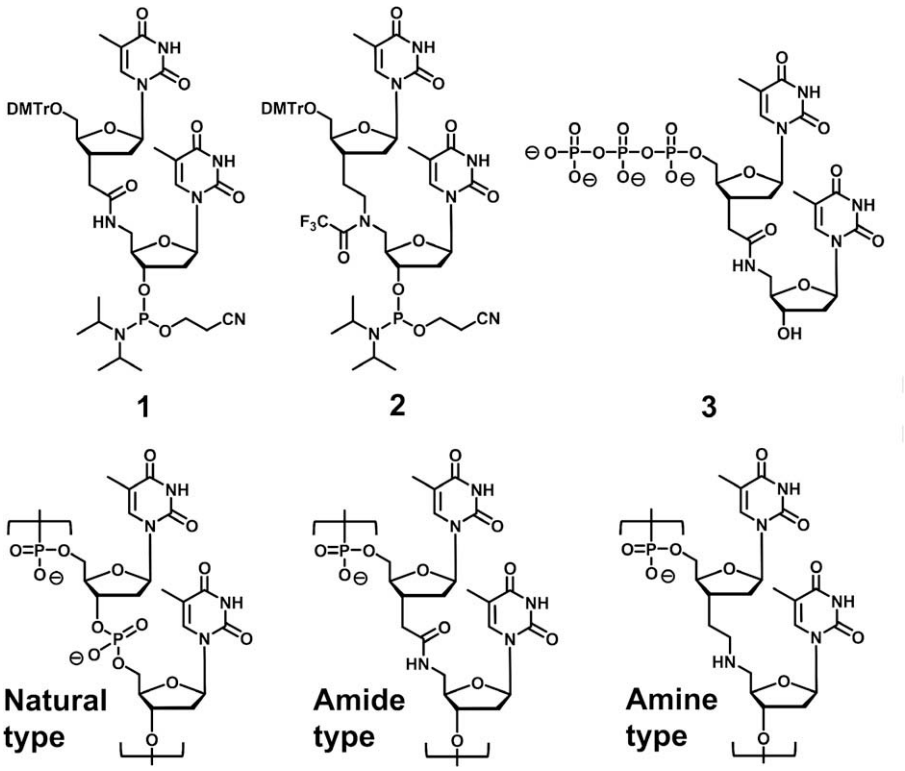

Figure 1. Chemical structures of the amidite derivatives (**1**, **2**) and the 5′-triphosphate analogue **3**. Nucleic acids involving natural phosphodiester linker, the amide-type linker *h* and amine-type linker *j* are also shown.

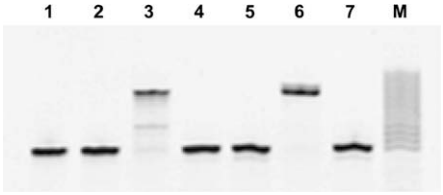

Figure 2. Representative gel images of the reactions using analogue **3** with *KOD*(*exo*-) DNA polymerase. The reaction mixtures for the negative control (lanes 2 and 5), for the positive control (lanes 3 and 6) and those containing **3** (lanes 4 and 7). Primer P1 only migrated in lane 1. Extension was performed at lower (lanes 2–4) and higher concentrations (lanes 5–7).

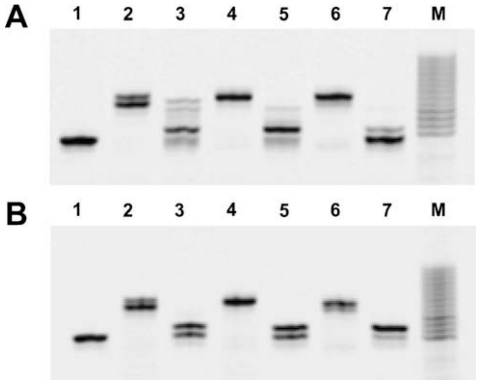

Figure 3. Representative gel images of the reactions using primer containing linker *h* or *j* with *KOD*(*exo*-), *Vent*(*exo*-) and *Taq* DNA polymerases. The reaction mixtures contained primer P2 (lanes 2, 4 and 6 in images A and B), P3 (lanes 3, 5 and 7 in image A), P4 (lanes 3, 5 and 7 in image B). Extension was performed at higher enzyme concentrations. Primer P2 migrated only in lane 1. The thermostable DNA polymerases used were *KOD*(*exo*-) (lanes 2 and 3), *Vent*(*exo*-) (lanes 4 and 5) and *Taq* (lanes 6 and 7).

polymerase passed beyond the first barrier, linker *j*, but the extension stopped substantially after the third dA was incorporated (graph A2). These results indicate that the extension is affected much more by insertion of linker *j* than by that of linker *h*. Furthermore, there are two more notable results. First, in the case of T4 and T8, the full-length product was not obtained, that is, no strand elongated after the ninth dA was incorporated, although the polymerase could pass beyond linker *h* or *j* incorporated as a single barrier. These results indicate that the barrier still affected the extension reaction several bases downstream from its location. Second, the graph line of T6 and T7 rose up earlier than that of T4 in the range of the first barrier to the two residues and before the second barrier. Also, the graph line of T9, T10 and T11 rose up earlier than that of T8 in the same range. In addition, as the

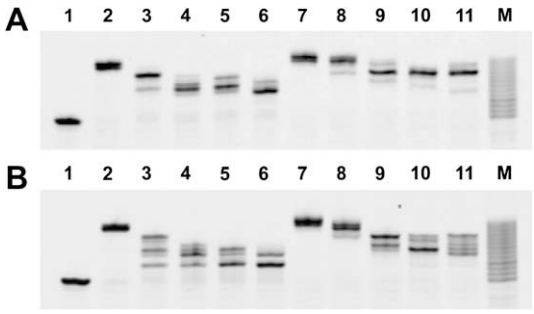

Figure 4. Representative gel images of the reactions using the template containing linker *h* or *j* with *KOD*(*exo*-) DNA polymerase. The reaction mixtures contained template T3 (lanes 2 and 7 in images A and B), T4 (lanes 3 and 8 in image A), T5 (lanes 4 and 9 in image A), T6 (lanes 5 and 10 in image A), T7 (lanes 6 and 11 in image A), T8 (lanes 3 and 8 in image B), T9 (lanes 4 and 9 in image B), T10 (lanes 5 and 10 in image B) or T11 (lanes 6 and 11 in image B). Extension was performed at lower (lanes 2–6) and higher enzyme concentrations (lanes 7–11). Primer P1 only migrated in lane 1.

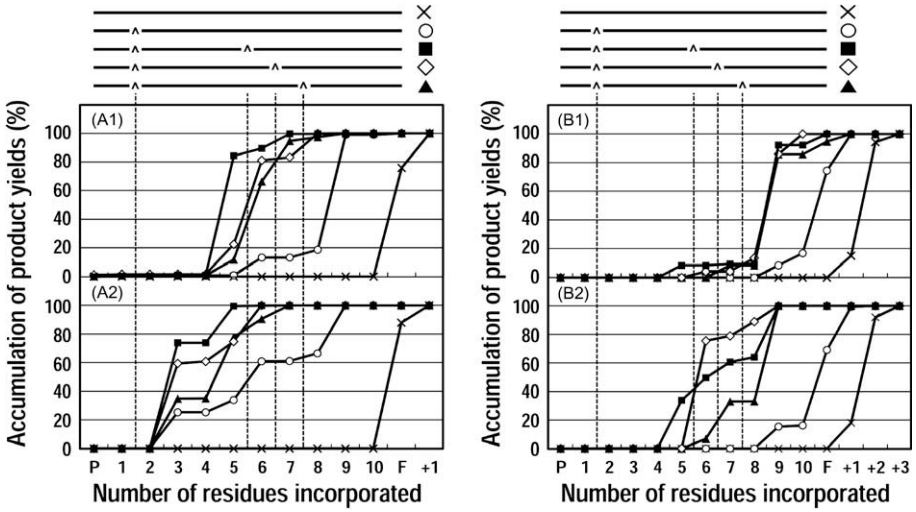

Figure 5. Accumulation of relative yield of the natural DNA generated by primer extension reactions involving natural DNA templates or various modified templates together with *KOD*(*exo*-) DNA polymerase. The reaction mixture contained templates T3–T7 at the lower (A1) and higher (B1) enzyme concentration; T3 and T8–T11 at the lower (A2) and higher (B2) enzyme concentration; reactions using T3 (crosses), T4 and T8 (open circles), T5 and T9 (closed squares), T6 and T10 (open diamonds), T7 and T11 (closed triangles). The *x*-axis represents the number of residues incorporated, and the *y*-axis indicates the accumulation of the yield of the products. The caret (^) shows the insertion site of linker *h* or *j*. P and F on the *x*-axis represent the primer and the full-length product, respectively. The relative standard deviations were less than ±5% for all reactions.

intervals became shorter ($\infty \sim 4$ residues), the ratio of the extension stop at this position increased to 25%, 35%, 59% and 74% in the case of T8, T9, T10 and T11, respectively. These results indicate that the second barrier would also affect the extension reaction upstream from its location.

At the higher enzyme concentrations, the polymerase could pass beyond both the first and the second barriers (graphs B1 and B2). However, extension almost stopped before the 11th dA incorporated and did not give the full-length product when templates containing two insertion sites of the modified linkers T5–T7 and T9–T11 were used. The polymerase might pass through without incorporating two or more residues because the homosequence would be liable to cause template slippage. When other polymerases *Vent*(*exo*-) and *Taq* were used, increasing enzyme concentrations also improved primer extensions. Use of *Taq* provided full-length and longer products much more efficiently than *KOD*(*exo*-) and *Vent*(*exo*-) in many cases (graphs B1 and B2 in Figs. 5, *S4* and S5). This is interesting because *Taq* was liable to stop extension earlier than the other polymerases at the lower enzyme concentration; about 50–70% of the extension was stopped by just the first dA incorporated, while 0% was stopped for *KOD*(*exo*-) and about 0–10% was stopped for *Vent*(*exo*-) in the case of T4–T7 involving the amide linker *h*. Furthermore, 50–80% of the extension was stopped by the third dA incorporated, about 25–75% for *KOD*(*exo*-) and about 40–70% for *Vent*(*exo*-), in the case of T8–T11 involving the amide linker *j* (graphs A1 and A2 in Figs. 5, *S4* and S5). These data may suggest that the use of *Taq* at high enzyme concentrations may be liable to cause more template slippage compared with the other polymerases.

We therefore next examined the primer extension using templates with mixed sequences T12–T15. In this experiment, incorporation of 12 nt was observed in some cases such as the positive controls. This may have occurred because of the aforementioned nontemplated single-nucleotide addition (Fig. 6). Figure 7 shows the putative results obtained from a simple model, assuming that (i) there is no long-range effect of linker *h* on the extension both upstream and downstream of its location, that is, the polymerase would stop extension only when it passes over the modified linkage *h*; (ii) there is no effect of instability of the duplex near the 5′-end of the template; (iii) the first and the second barriers could cause provisionally 45% and 35% of the extension stop, respectively. The duplex formed when the polymerase passes over the second barrier would gain more stability (compared to that from the first barrier) corresponding to the six additional base pairs. Therefore, the ratio of the extension stop at the second barrier was assumed to be set lower than that of the first one. Among the graphs in Figure 6, graphs A3 and particularly B3, where *Taq* was used, showed a profile most similar to Figure 7. However, a difference between the graphs of T13 and T15 was observed for the first barrier to the two residues and before the second. This was also observed in graphs A1 and A2, for which *KOD*(*exo*-) and *Vent*(*exo*-) were used at the lower enzyme concentration. Thus, the modified linker on the template would affect the extension reaction upstream from its location. It is difficult to imagine from the X-ray crystal structure[37–41] that the polymerase recognizes the upcoming modified linker on the single-strand part of the template in the primer/template complex, that is, that the polymerase can look forward. However, it is possible that the insertion of the linker influences the dynamic flexibility of the single-strand part of the template, which could disrupt the formation of the stable enzyme/primer/template ternary complex, resulting in a slowing of the extension speed. The graphs of T13 in A1 and A2 rose up gradually after the second nucleotide was incorporated, indicating that even though the polymerase could pass beyond the modified linker, the linker would keep affecting the extension reaction downstream from its location. This can be also seen in the aforementioned experiments using the templates with the homosequence. Furthermore, the results of the experiments using the templates with the mixed sequence suggest that the instability of the duplex near the 5′-end of the template strand also affects the extension reaction. The graphs of T13 and T14 in A1–A3 showed that the polymerase left the last nucleotide incorporation and could not complete the extension, although it passed beyond the modified linker. Polymerase, in fact, may dissociate from the primer/template complex around the 5′-end of the template strand, presumably because the ends of the duplex would be liable to unwind. Therefore, the structure around the 5′-end of the template strand became more sensitive to subtle distortion or destabilization of the local duplex structure by a modified linker insertion, even if it is located less than nine residues upstream from the 5′-terminus. Comparing the graphs of T14 and T15 in B1 and B2, extension was largely inhibited at the second barrier in T15, while

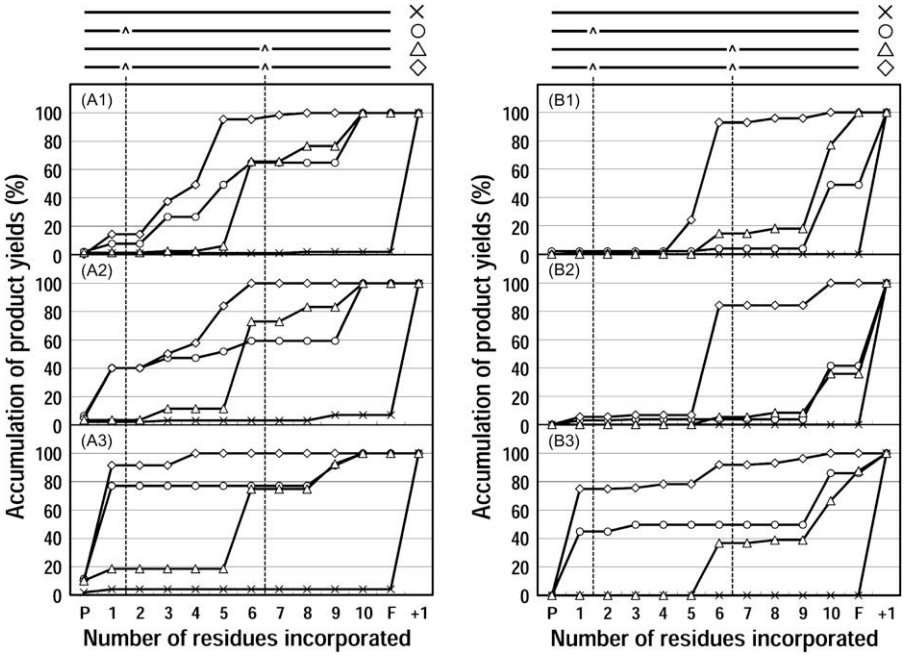

Figure 6. Accumulation of relative yield of the natural DNA generated by primer extension reactions involving natural DNA templates or modified templates. The reaction mixture contained templates T12–T15 with *KOD*(*exo*-), *Vent*(*exo*-) and *Taq* DNA polymerase at lower enzyme concentration (A1, A2 and A3) and at higher enzyme concentration (B1, B2 and B3), respectively; reactions using T12 (crosses), T13 (open circles), T14 (open squares), and T15 (open diamonds). The *x*-axis indicates the number of residues incorporated, and the *y*-axis represents the accumulation of yield of the products. The caret (^) shows the insertion site of linker *h*. P and F on the *x*-axis represents the primer and the full-length product, respectively. The relative standard deviations were less than ±3% for all reactions.

it was hardly interrupted at the first barrier in T14, although both of the barriers were located at the same position in the same sequence of the template. Thus, the upstream barrier cooperatively comes down through the backbone string and affects the downstream extension reaction; therefore, even minor effects such as the instability of the duplex ends could stop the polymerase from working.

3. Conclusions

In dynamic molecular environments such as hot water thermostable DNA polymerases would 'bronco ride' rather than 'tightrope walk'. It seems to be a miracle that polymerases can quickly and accurately produce a complimentary strand based on sequence information of the template strand. The slight structural differences between the amide and amine linkages were found to clearly reflect the efficiency of strand extensions. Insertions of linker *h* into the backbone, in general, would help stabilize the duplex, while insertions of linker *j* would decrease the stability of the duplex.[42] This would explain why linker *j* was a greater barrier for extensions than linker *h* in many cases. However, there were some cases that could not sufficiently be explained solely by thermodynamical stabilization or destabilization of the primer/template duplex due to the insertions. For example, *Taq* DNA polymerase tended to prefer linker *j* to *h* (Figs. 3 and *S2*, graphs A1 and A2 in Fig. S5). Also, the modified linker located on the primer strand adjacent to the 3′-end affected the extension much more than that located on the template strand opposite the elongation terminus (Figs. 3 and 4); the thermostability of the duplex should be almost same regardless of the strand on which the linker is located if the same type of linker is inserted. Thus, our results reflect an enhancement of slight and local effects such as electrostatic interaction between DNA polymerase and the DNA backbone, dynamic flexibility of the single-strand part of the template, and instability of the duplex near the 5′-end of the template.

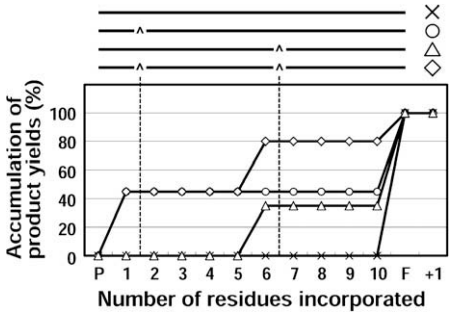

Figure 7. Putative results of accumulation of relative yield of the natural DNA generated by primer extension reactions involving natural DNA templates or modified templates. Reactions using T12 (crosses), T13 (open circles), T14 (open squares), and T15 (open diamonds). The *x*-axis indicates the number of residues incorporated, and the *y*-axis indicates the accumulation of yield of the products. The caret (^) shows the insertion site of the linker *h*. P and F on the *x*-axis represent the primer and the full-length product, respectively.

Enzymatic incorporation of the dinucleotide mimic with an amide bond could not, only confer nuclease resistance on DNA, but could also be used to construct a *n*-mer artificial-nucleic-acid library with a diversity equal to the *n*th power of 16 (=4 × 4) if 16 types of functional group were allocated to each dinucleotide mimic as building blocks. Unfortunately this challenge, the transcription of artificial nucleic acids, has failed so far. Even if transcription was successful, the further extension could not proceed efficiently as shown in Figure 3. However, the reverse transcription of artificial nucleic acids would be possible if the reaction conditions or insertion intervals and the chemical structures of the modified linker were well optimized. We are currently improving the genetic engineering of thermostable DNA polymerases to apply artificial nucleic acids with backbone modifications to systems of random screening for functional nucleic acids.

4. Experimental

4.1. General experimental procedures

A TC-312 thermal cycler (Techne, Stone, Staffordshire, UK) was used for primer extension experiments. Reaction products were resolved by denaturing PAGE using a vertical electrophoresis unit (Nihon Eido, Tokyo, Japan) at 48 °C in an M-260F incubator (Taitec, Saitama, Japan). Bands were imaged using a Molecular Imager FX (Bio-Rad, Hercules, CA, USA) equipped with an external laser module and quantified with the software Quantity One (Bio-Rad). Mass spectral analysis for nucleoside analogues was performed on an ABI MDS-Sciex API-100 spectrometer under atmospheric pressure ionization conditions. Ultraviolet (UV) analyses were performed on a Shimadzu UV-1200 spectrometer. ^{1}H and ^{31}P nuclear magnetic resonance (NMR) spectra were recorded on a JEOL JNM-AL300 or JNM-LA500 Fourier-transform-NMR spectrometer. Tetramethylsilane and 85% phosphoric acid were used as the internal standards for ^{1}H and ^{31}P NMR, respectively. The sodium salt of the triphosphate analogue **3**, generated from the corresponding triethylammonium salt using Dowex 50WX8 (Na^+ form), was used for NMR measurements. Reversed-phase high-performance liquid chromatography (HPLC) was performed using a JASCO Gulliver system with UV detection at 260 nm and a packed Wakosil 5C18 (φ4.6 × 250 mm; Wako) or TSKgel ODS-80Ts (φ20 × 250 mm; Tosoh) column. Reversed-phase medium-pressure liquid chromatography (MPLC) was performed using an YFLC-Wprep system (Yamazen) with a glass column (φ33 × 250 mm) filled with Wakosil 40C18 (Wako). Ion exchange column chromatography was performed using an ECONO system (Bio-Rad) with a glass column (φ25 × 500 mm) filled with diethylaminoethyl (DEAE) A-25-Sephadex (Amersham Biosciences). Molecular absorption coefficients were 18,600 M^{-1} cm^{-1} at 260 nm for **3**.

4.2. Materials

The following commercially available thermostable DNA polymerases were purchased: *Taq* (Takara Bio, Sigma, Japan) and *Vent*(*exo*-) (New England Biolabs, Hitchin, Herts, UK). *KOD(exo-)* DNA polymerase was supplied by Toyobo. *KOD(exo-)* is an enzyme genetically engineered to eliminate 3′,5′ exonuclease activity from *KOD*. Natural 2′-deoxynucleoside-5′-triphosphates (dATP, dGTP, dCTP and TTP) were obtained from Roche Diagnostics. Two amidite derivatives (**1** and **2**) of thymidine dimer with the amide-type linker *h*, or with the amine-type linker *j*, were synthesized according to previously published procedures[43] (Fig. 1). The corresponding oligodeoxynucleotides (P3–P4, T4–T11, and T13–T15) were prepared and characterized by MALDI-TOF mass spectrometry. Characterization of these primers and templates are provided in Supplementary data. Primers P1 and P2 and templates T1–T3 and T12 were purchased from JBioS. To detect and quantify extension products, the 5′-ends of the primers were labelled with 6-carboxyfluorescein (6-FAM). Spectroscopic data of **1** and **2** and synthesis of **3** are given below.

Amidite derivative ***1***: ^{1}H NMR (300 MHz, $CDCl_3$) δ 7.59, 7.42 (s, 2H), 7.32–7.21 (m, 9H), 6.82 (d, 4H), 6.20 (t, 1H), 5.80-5.70 (t, 1H), 4.45 (m, 1H), 4.13 (m, 1H), 3.90–3.58 (m, 8H), 3.48–3.25 (d, 5H), 2.73–2.58 (m, 4H), 2.34–2.16 (m, 7H), 1.89, 1.45 (s, 6H), 1.18, 1.16 (d, 12H); ^{31}P NMR (500 MHz, $CDCl_3$) δ 149.80, 149.40; ESI-MS (positive ion mode) *m/z*, found = 1032.6, calcd for $[(M+Na)^+]$ = 1032.4.

Amidite derivative ***2***: ^{1}H NMR (300 MHz, $CDCl_3$) δ 7.72, 7.40 (s, 2H), 7.39–7.25 (m, 9H), 6.84-6.81 (m, 4H), 6.09, 5.91 (m, 2H), 4.43, 4.21 (m, 2H), 3.90–3.85 (m, 2H), 3.78 (s, 6H), 3.64–3.38 (m, 5H), 3.28–3.14 (m, 2H), 2.72–2.56 (m, 4H), 2.65–2.02 (m, 5H), 1.87 (s, 3H), 1.78–1.55 (m, 2H), 1.46 (s, 3H), 1.18, 1.12 (d, 12H); ^{31}P NMR (500 MHz, $CDCl_3$) δ 149.13, 148.91; ESI-MS (positive ion mode) *m/z*, found = 1114.5, calculated for $[(M+Na)^+]$ = 1114.4.

Triphosphate analog ***3***: Thymidinylamidoethylthymidine[43] (100 mg) and *N*,*N*,*N*′,*N*′-tetramethyl-1,8-naphthalenediamine (Proton Sponge®; 63 mg, 1.5 equiv) was dried in a flask under vacuum overnight. Trimethylphosphate (1.5 mL) was added to the flask under argon, and the solution was cooled to 0 °C. Distilled phosphorus oxychloride (27 μL, 1.5 equiv) was then added dropwise using a micro syringe, and the reaction mixture was stirred at 0 °C. After 45 min, *n*-tributylamine (188 μL, 4.0 equiv) and *n*-tributylamine pyrophosphate (1.9 mL of a 0.5 M solution in DMF; 5.0 equiv) were added at 0 °C, and the reaction mixture was warmed to room temperature and stirred for an additional 1 h. The reaction was quenched with triethylammonium bicarbonate (1.0 M aqueous solution). The solvents were removed in vacuo, and the remaining crude mixture was dissolved in water. The product was purified using a Sephadex DEAE A-25 column with a linear gradient of 0–1.0 M triethylammonium bicarbonate buffer (pH 8). The corresponding fractions were combined and evaporated under reduced pressure. To remove the excess of pyrophosphate, the residue was purified by reversed-phase MPLC with a linear gradient of 0–100% acetonitrile in 10 mM triethylammonium acetate buffer (pH 7). Further purification was performed on reversed-phase HPLC (φ20 × 250 mm) with a linear gradient of 0–5% acetonitrile in 50 mM triethylammonium acetate buffer to give **3** (38.4 μmol) in 19% yield: ESI-MS (nagative ion mode) *m/z*, found = 746.1, calculated for $[(M-H)^-]$ = 746.1.

4.3. Primer extension experiments using triphosphate analogue

To investigate the enzymatic incorporation of the dinucleotide mimic into a DNA strand, primer extension reactions were performed in a 20 μl reaction volume containing 0.4 μM of a primer (P1), 0.4 μM of a template (T1), an appropriate concentration of a thermostable DNA polymerase, a reaction buffer supplied with the enzyme (at 1 × concentration) and a nucleoside triphosphate at 200 μM. A reaction mixture with natural TTP (thymidine-5′-triphosphate) was used as a positive control. The assays were performed with the triphosphate analogue **3** in place of TTP; a reaction using water in place of TTP was used as a negative control. The final concentrations of the thermostable DNA polymerase in each reaction mixture were 0.00315 U/μl for *KOD(exo-)* and 0.050 U/μl for *Vent(exo-)* and *Taq*, as the 'lower enzyme concentrations'. The concentrations were 0.0315 U/μl for *KOD*(*exo*-) and 0.50 U/μl for *Vent(exo-)* and *Taq* for the 'higher enzyme concentrations'. All reactions were performed by denaturation for 1.5 min at 94 °C, annealing for 0.5 min at 52 °C and extension for 5 min at 74 °C, successively. The reaction mixtures were mixed with 4 μL of 40 mM EDTA containing 0.1% bromophenol blue and 24 μL of 7 M urea containing 3 mM EDTA. The sample solutions were resolved by denaturing PAGE (300 V, 3 h, 48 °C) and gel images were recorded with excitation of the 5′-labelled fluorophore at 488 nm (Figs. 2 and S1).

4.4. Primer extension experiments using modified primers

Primer extension reactions were performed in a 20 μl reaction volume containing one of the primers (P2–P4) at 0.4 μM, 0.4 μM of a template (T2), an appropriate concentration of thermostable DNA polymerase, reaction buffer supplied with an enzyme (at 1 × concentration) and 200 μM each of the four natural 2′-deoxynucleoside-5′-triphosphates. A reaction with a natural primer (P2) was used as a positive control. The assays were performed with one of the modified primers (P3–P4) containing linker *h* or *j*

in place of P2. The final concentrations of the thermostable DNA polymerase in each reaction mixture were 0.0025 U/μl for *KOD(exo-)*, 0.020 U/μl for *Vent*(*exo-*) and 0.025 U/μl for *Taq* as the 'lower enzyme concentrations'. The concentrations were 0.025 U/μl for *KOD(exo-)*, 0.20 U/μl for *Vent*(*exo-*) and 0.25 U/μl for *Taq* as the 'higher enzyme concentrations'. The lower enzyme concentrations are almost the same as the standard conditions recommended by manufacturers except for *KOD(exo-)*; the recommended concentrations are 0.025–0.050 U/μl for *KOD(exo-)*, 0.010–0.020 U/μl for *Vent*(*exo-*) and 0.025 U/μl for *Taq*. The higher concentrations were set 10 times higher than the lower concentrations. All reactions were performed by denaturation for 1.5 min at 94 °C, annealing for 0.5 min at 52 °C and extension for 5 min at 74 °C, successively. The reaction products were resolved by denaturing PAGE, and gel images were obtained as described above (Figs. 3 and S2).

4.5. Primer extension experiments using modified templates

Primer extension reactions were performed in a 20 μl reaction volume containing 0.4 μM of primer (P1), one of the templates (T3–T15) at 0.4 μM, an appropriate concentration of thermostable DNA polymerase, reaction buffer supplied with an enzyme (at 1 × concentration) and 200 μM of dATP (2′-deoxyadenosine-5′-triphosphate) when templates (T3–T11) were used, or 200 μM each of the four natural 2′-deoxynucleoside-5′-triphosphates when templates (T12–T15) were used. A reaction with a natural template (T3 or T12) was used as a positive control. The assays were performed with one of the modified templates (T4–T11) containing linker *h* or *j* in place of T3, or with one of the modified templates (T13–T15) instead of T12. The final concentrations of the thermostable DNA polymerase in each reaction mixture were 0.0025 U/μl for *KOD*(*exo-*), 0.020 U/μl for *Vent*(*exo-*) and 0.025 U/μl for *Taq*, as the 'lower enzyme concentrations'. The concentrations were 0.025 U/μl for *KOD(exo-)*, 0.20 U/μl for *Vent*(*exo-*) and 0.25 U/μl for *Taq*, as the 'higher enzyme concentrations'. The higher concentrations were set 10 times higher than the lower concentrations. All reactions were performed by denaturation for 1.5 min at 94 °C, annealing for 0.5 min at 52 °C and extension for 5 min at 74 °C, successively. The reaction products were resolved by denaturing PAGE, and gel images were obtained as described above (Figs. 4 and Supplementary data S3). The yields of elongated products were calculated from the intensity of each band on gel images visualized by the detection of the 5′-labelled fluorophore. The total amount of the products was set at 100% in each reaction mixture, and the calculated yields were the averages of three independent experiments (Figs. 5 and 6, S4 and S5).

Acknowledgements

We are grateful to Dr. Steven A. Benner and Dr. Nicole A. Leal at the Foundation for Applied Molecular Evolution (FAME) for useful discussions and helpful comments. This study was partly supported by a Grant-in-Aid for Scientific Research on Priority Areas from the Ministry of Education, Culture, Sports, Science and Technology of Japan, and a Grant for Industrial Technology Research from the New Energy and Industrial Technology Development Organization (NEDO) of Japan. We would like to thank Toyobo Ltd for generously providing *KOD(exo-)* DNA polymerase.

Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.bmc.2009.04.045.

References and notes

1. Latham, J. A.; Johnson, R.; Toole, J. J. *Nucleic Acids Res.* **1994**, *22*, 2817.
2. Battersby, T. R.; Ang, D. N.; Burgstaller, P.; Jurczyk, S. C.; Bowser, M. T.; Buchanan, D. D.; Kennedy, R. T.; Benner, S. A. *J. Am. Chem. Soc.* **1999**, *121*, 9781.
3. Santoro, S. W.; Joyce, G. F.; Sakthivel, K.; Gramatikova, S.; Barbas, C. F., III *J. Am. Chem. Soc.* **2000**, *122*, 2433.
4. Kusser, W. *Rev. Mol. Biotechnol.* **2000**, *74*, 27.
5. Perrin, D. M.; Garestier, T.; Hélène, C. *J. Am. Chem. Soc.* **2001**, *123*, 1556.
6. May, J. P.; Ting, R.; Lermer, L.; Thomas, J. M.; Roupioz, Y.; Perrin, D. M. *J. Am. Chem. Soc.* **2004**, *126*, 4145.
7. Sidorov, A. V.; Grasby, J. A.; Williams, D. M. *Nucleic Acids Res.* **2004**, *32*, 1591.
8. Masud, M. M.; Kuwahara, M.; Ozaki, H.; Sawai, H. *Bioorg. Med. Chem.* **2004**, *12*, 1111.
9. Shoji, A.; Kuwahara, M.; Ozaki, H.; Sawai, H. *J. Am. Chem. Soc.* **2007**, *129*, 1456.
10. Ohsawa, K.; Kasamatsu, T.; Nagashima, J.; Hanawa, K.; Kuwahara, M.; Ozaki, H.; Sawai, H. *Anal Sci.* **2008**, *24*, 167.
11. Gold, L.; Polisky, B.; Uhlenbeck, O.; Yarus, M. *Annu. Rev. Biochem.* **1995**, *64*, 763.
12. Santoro, S. W.; Joyce, G. F. *Proc. Natl. Acad. Sci. U.S.A.* **1997**, *94*, 4262.
13. Breaker, R. R. *Chem. Rev.* **1997**, *97*, 371.
14. Osborne, S. E.; Ellington, A. D. *Chem. Rev.* **1997**, *97*, 349.
15. Li, Y.; Breaker, R. R. *Curr. Opin. Struct. Biol.* **1999**, *9*, 315.
16. Wilson, D. S.; Szostak, J. W. *Annu. Rev. Biochem.* **1999**, *68*, 611.
17. Famulok, M.; Blind, M.; Mayer, G. *Acc. Chem. Res.* **2000**, *33*, 591.
18. Berezovski, M.; Musheev, M.; Drabovich, A.; Krylov, S. N. *J. Am. Chem. Soc.* **2006**, *128*, 1410.
19. Nitsche, A.; Kurth, A.; Dunkhorst, A.; Pänke, O.; Sielaff, H.; Junge, W.; Muth, D.; Scheller, F.; Stöcklein, W.; Dahmen, C.; Pauli, G.; Kage, A. *BMC Biotechnol.* **2007**, *7*, 48.
20. Andreola, M. L.; Calmels, C.; Michel, J.; Toulmé, J. J.; Litvak, S. *Eur. J. Biochem.* **2000**, *267*, 5032.
21. He, K.; Porter, K. W.; Hasan, A.; Briley, J. D.; Shaw, B. R. *Nucleic Acids Res.* **1999**, *27*, 1788.
22. Veedu, R. N.; Vester, B.; Wengel, J. *Chembiochem.* **2007**, *8*, 490.
23. Tsai, C. H.; Chen, J.; Szostak, J. W. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 14598.
24. Kempeneers, V.; Renders, M.; Froeyen, M.; Herdewijn, P. *Nucleic Acids Res.* **2005**, *33*, 3828.
25. 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.
26. Held, H. A.; Benner, S. A. *Nucleic Acids Res.* **2002**, *30*, 3857.
27. Kuwahara, M.; Nagashima, J.; Hasegawa, M.; Tamura, T.; Kitagata, R.; Hanawa, K.; Hososhima, S.; Kasamatsu, T.; Ozaki, H.; Sawai, H. *Nucleic Acids Res.* **2006**, *34*, 5383.
28. Kuwahara, M.; Obika, S.; Nagashima, J.; Ohta, Y.; Suto, Y.; Ozaki, H.; Sawai, H.; Imanishi, T. *Nucleic Acids Res.* **2008**, *36*, 4257.
29. Gat, Y.; Lynn, D. G. *Biopolymers* **1998**, *48*, 19.
30. Leal, N. A.; Sukeda, M.; Benner, S. A. *Nucleic Acids Res.* **2006**, *34*, 4702.
31. Kuwahara, M.; Minezaki, S.; Nagashima, J.; Ozaki, H.; Sawai, H. *Nucleic Acids Symp. Ser.* **2008**, *52*, 453.
32. Abramova, T. V.; Vasileva, S. V.; Serpokrylova, I. Y.; Kless, H.; Silnikov, V. N. *Bioorg. Med. Chem.* **2007**, *15*, 6549.
33. Cozzarelli, N. R.; Kelly, R. B.; Kornberg, A. *J. Mol. Biol.* **1969**, *45*, 513.
34. Clark, J. M. *Nucleic Acids Res.* **1988**, *16*, 9677.
35. Karthikeyan, G.; Chary, K. V.; Rao, B. J. *Nucleic Acids Res.* **1999**, *27*, 3851.
36. Viguera, E.; Canceill, D.; Ehrlich, S. D. *J. Mol. Biol.* **2001**, *312*, 323.
37. Hashimoto, H.; Nishioka, M.; Fujiwara, S.; Takagi, M.; Imanaka, T.; Inoue, T.; Kai, Y. *J. Mol. Biol.* **2001**, *306*, 469.
38. Eom, S. H.; Wang, J.; Steitz, T. A. *Nature* **1996**, *382*, 278.
39. Kim, Y.; Eom, S. H.; Wang, J.; Lee, D. S.; Suh, S. W.; Steitz, T. A. *Nature* **1995**, *376*, 612.
40. Kiefer, J. R.; Mao, C.; Braman, J. C.; Beese, L. S. *Nature* **1998**, *391*, 304.
41. Damsma, G. E.; Alt, A.; Brueckner, F.; Carell, T.; Cramer, P. *Nat. Struct. Mol. Biol.* **2007**, *14*, 1127.
42. Freier, S. M.; Altmann, K. H. *Nucleic Acids Res.* **1997**, *25*, 4429.
43. Luo, P.; Leitzel, J. C.; Zhan, Z.-Y. J.; Lynn, D. G. *J. Am. Chem. Soc.* **1998**, *120*, 3019.