ARTICLE

OBC
www.rsc.org/obc

# Expansion of repertoire of modified DNAs prepared by PCR using KOD Dash DNA polymerase†

**Tsutomu Ohbayashi, Masayasu Kuwahara, Masatoshi Hasegawa, Toshiyuki Kasamatsu, Takehiro Tamura and Hiroaki Sawai***

*Department of Applied Chemistry, Gunma University, Kiryu, Gunma 376-8515, Japan. E-mail: sawai@chem.gunma-u.ac.jp*

***Received 29th March 2005, Accepted 12th May 2005***
***First published as an Advance Article on the web 9th June 2005***

Thymidine analogues bearing a variety of functional groups at the C5-position *via* an amino-linker arm were prepared and the substrate activity for PCR using thermophilic KOD Dash DNA polymerase was examined. The enzyme accepted the thymidine analogues bearing pyridine, imidazole, biotin, a cationic-charged guanidinium, a cationic-charged amino, mercaptopyridyl and phenanthrolne groups at the C5-position, forming the corresponding PCR product. However, a thymidine analogue bearing a carboxyl group at the C5-position was a poor substrate and the corresponding PCR products could not be obtained. The thymidine analogue bearing a mercapto group was also a poor substrate for the enzyme, because it dimerized by disulfide linkage under PCR conditions. The enzyme hardly accepts the thymidine analogues with a negatively-charged carboxyl group or a bulky group as a substrate. KOD Dash DNA polymerase, having a broader substrate specificity than any other DNA polymerase, will expand the variety of modified DNAs that can be prepared by PCR.

## Introduction

DNAs that have been chemically modified with a variety of functional groups are important tools for biochemical and biological studies.[1–3] They are also useful for diagnostic and therapeutic purposes. In general, the applications rely on differences in the chemical reactivity, electronic and stereochemical properties of the additional functional groups compared to the naturally occurring DNA. Among the many possible sites of the DNA-modification, the C5-position of the thymidine is an ideal site for the attachment of functional groups, because it is located in the major groove of double-stranded DNA and does not inhibit A–T base pairing. Previously, we reported the synthesis of thymidine analogues, C5-substituted 2′-deoxyuridines bearing an amino-linker arm at the C5-position,[4] and their introduction into oligodeoxyribonucleotides chemically using a DNA synthesizer.[5]

If a DNA polymerase can accept modified nucleotides as substrates and modified DNA can be prepared enzymatically by PCR, the resulting modified DNA could be applied in the production of a DNA aptamer or a catalytic DNA by *in vitro* selection. A variety of DNA aptamers and DNA catalysts have been produced from the natural-type DNAs.[6–10] However, their activity is, in general, not very strong when compared with the corresponding functional proteins, antibodies and enzymes. One reason for the weak activity of the DNA aptamer or catalyst would be a lack of the functional groups which are present within proteins. The modified DNA with a functional group will improve the weak activity of natural-type DNA. Some improved DNA aptamers or catalysts have been prepared from the modified DNAs.[11–14] However, more active DNA aptamers or catalysts are still desired for wide applications. Thus, nucleotide analogues with a variety of functional groups, that are substrates for DNA polymerase, will be desirable for the production of a high-performance functional DNA. Some DNA polymerases can use 5′-triphosphates of modified 2′-deoxyuridines with a C5-side chain carrying an (*E*)-propenyl[15–20] or a propynyl group.[18–24] The differences in reactivity of C5-substituted 2′-deoxyuridines as a substrate may be due to the steric effect of a large side-group and/or to the ionic effect of a substituting group, such as a cationic group or an anionic group. The ability of the modified nucleotides to act as a substrate also depends on the kind of DNA polymerase.[23–29] Furthermore, some DNA polymerases can accept two or four modified nucleotides simultaneously in the primer-extension reaction or polymerase chain reacion (PCR).[23–26] Previously, we have shown that KOD Dash DNA polymerase can readily use some new thymidine analogue nucleotides as a substrate and can read a template DNA containing the modified thymidine, but no other DNA polymerase, including Taq DNA polymerase, can accept the thymidine analogues as a substrate.[27] The new thymidine analogues that have an sp3 hybridized carbon at the C5-position with methyl ester or an amino-linker arm are a good substrate for the enzyme and the corresponding modified DNAs can be obtained by PCR using KOD Dash DNA polymerase.[28]

To expand the structural and functional repertoire of modified DNAs, we have attached a variety of functional groups at the C5-position of the thymidine analogue *via* the amino-linker arm and examined whether they work as a substrate for KOD Dash DNA polymerase in forming the corresponding modified DNA by PCR. The functional groups include pyridine, imidazole, biotinyl, cationic-charged guanidinium, cationic-charged amino, anionic-charged carboxyl, mercapto, mercaptopyridyl and phenanthroline groups. Among the thymidine analogues examined, all analogues except the ones with a carboxyl or a sterically bulky group could be accepted as a substrate for the enzyme, yielding the modified DNA by PCR. The finding that the thymidine analogues with an anionic-charged carboxyl and bulky groups are a poor substrate for the enzyme may be useful for study of the recognition process of the enzyme and substrate. KOD Dash DNA polymerase has broader substrate specificity which will expand the variety of modified DNA that can be prepared by PCR.

## Results and discussion

### Synthesis of modified substrate

The modified thymidine analogue, 5-*N*-(6-aminohexyl)carbamoylmethyl-2′-deoxyuridine 5′-triphosphate (**1**) was prepared according to the method described previously.[27] This analogue is a good substrate for the thermostable KOD Dash DNA

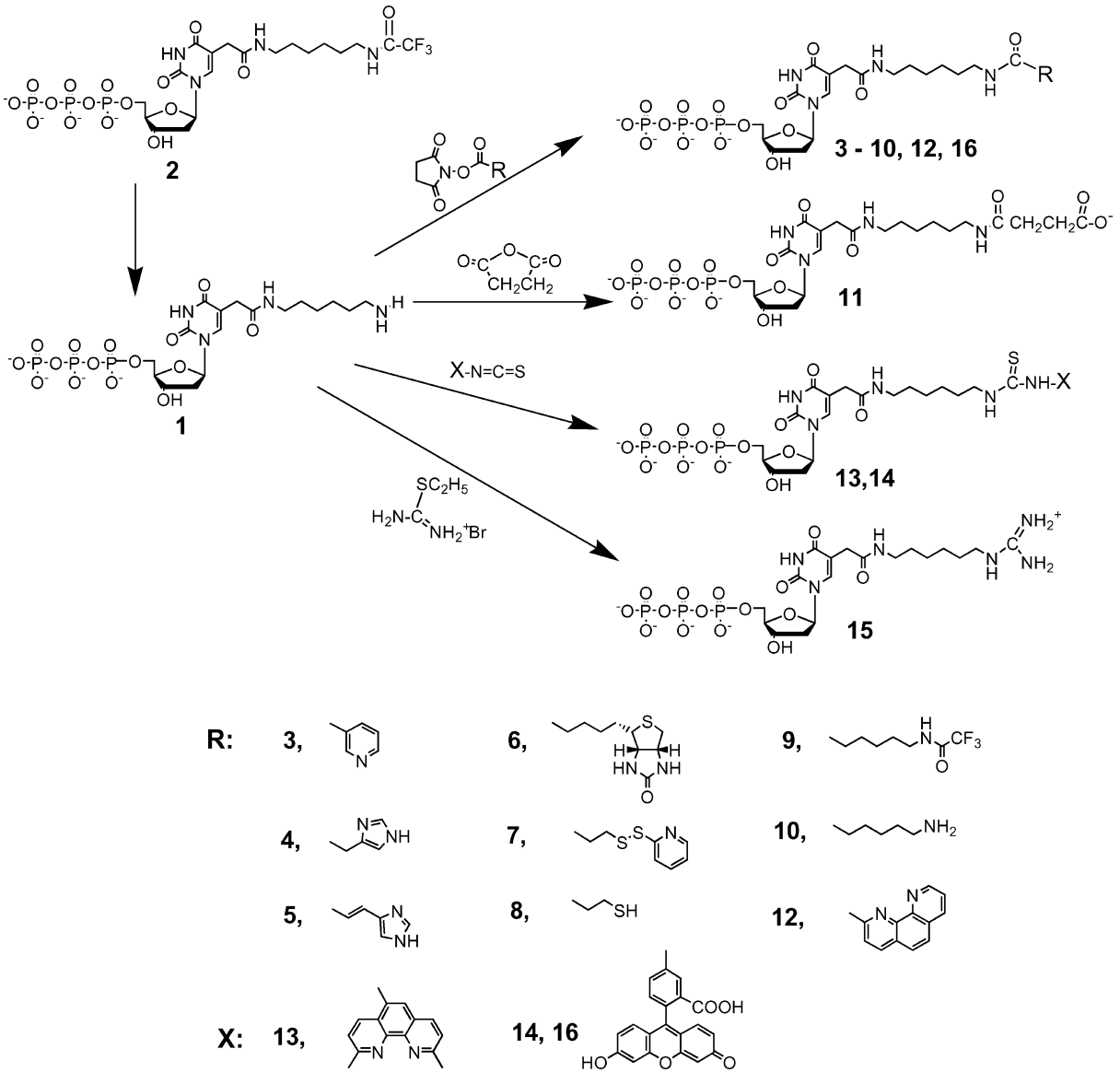
† Electronic supplementary information (ESI) available: synthetic procedure for the 5′-triphosphates of thymidine analogues, **4**–**14** and **16**. See http://www.rsc.org/suppdata/ob/b5/b504330a/

DOI: 10.1039/b504330a

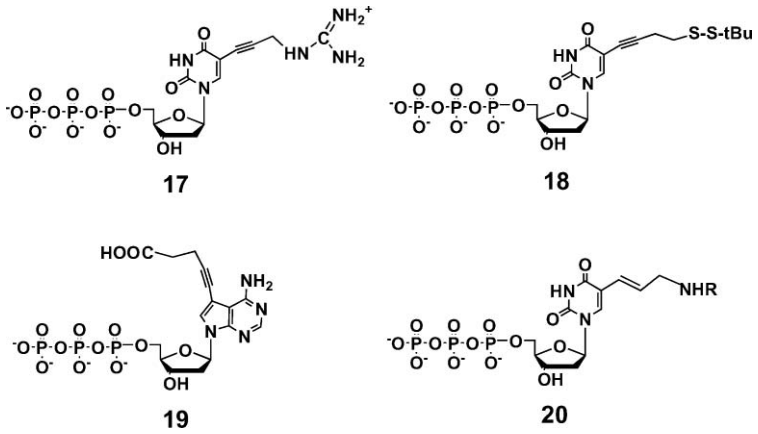
polymerase, which belongs to the B-family DNA polymerase derived from thermophilic archae.[30] The nucleotide **1** bears a primary amino-group at the C5-position of thymidine *via* a linker arm. Many kinds of functional groups can be introduced to the terminal amino-group by amide bond formation or by other reactions with the amino group, as shown in Scheme 1. The hydroxysuccinimide activated esters of urocanic acid, imidazole acetic acid, nicotinic acid, biotin, 3-(2-pyridyldithio)propionic acid, trifluoroacetylaminohexanoic acid and 2-phenanthroline carboxylic acid were reacted easily with **1** in aqueous DMF solution. The corresponding thymidine analogues bearing functional groups, **3–10**, **12** and **16**, were obtained in moderate to high yields. The reaction of succinic anhydride with the amino group of **1** gave a thymidine analogue with an anionic-charged carboxyl group at the C5-position of **11**. Isothiocyanate of 2,9-dimethylphenanthroline or fluorescence reacts with the amino group of **1** to give the 2,9-dimethylphenanthroline- or fluorescence-substituted thymidine analogue, **13** or **14**, respectively. Similarly, guanidinium-substituted thymidine analogue **15** was obtained by reaction of *S*-ethylthiopseudourea hydrobromide with **1**. Thus, thymidine analogues with catalytic function, an aromatic compound hydrophobic moiety, a hydrophilic moiety, an anionic-charged group, a cationic charged group, a chelating group and fluorescent groups were obtained from the nucleotide **1**.

**Synthesis of modified DNA by incorporation of the thymidine analogues by PCR using KOD Dash DNA polymerase**

We studied the ability of the thymidine analogues **1–16** to act as a substrate in place of deoxythymidine triphosphate (TTP) for KOD Dash DNA polymerase under typical PCR conditions, using pUC18 plasmid DNA as a template, with oligonucleotide **A**, 5′-GGAAACAGCTATGACCATGATTAC-3′, and oligonucleotide **B**, 5′-CGACGTTGTAAAACGACGGCCAGT-3′, as primers. Fig. 1 shows the results of PCR with the substrates. We used ten times the concentration of the enzyme for the PCR with the modified thymidine analogues compared to that with natural substrate, TTP, to increase the yield of the modified DNA. Non-specific amplification of DNA took place when the ten times concentration of the enzyme was used for PCR with natural TTP (lane 4 of Fig. 1B). We confirmed the previous report[25] that KOD Dash DNA polymerase can accept **1** or **2** as a substrate, giving a 108 base-pair DNA product (**2**) (lane 4 and 5 of Fig. 1A). The formation of the corresponding natural-type DNA from the natural substrate TTP was 1.4 times larger compared to that from **1** under the optimum conditions for the formation of natural type DNA, where the enzyme quantity was reduced to suppress the non-specific amplification. The 108 base-pair DNA containing the primers and the template region has 40 modified thymidines with a single-stretch of four successive thymidine residues. The




**Scheme 1** Synthesis of 5′-triphosphates of thymidine analogues bearing various functional groups.




**Scheme 2** Structures of DNA polymerase-tolerated modified nucleotides reported by other workers.

enzyme could accept thymidine analogues **3**–**7**, **9**–**12** and **15**, forming the corresponding 108 bp PCR product. The mobility shift of the modified DNA on the gel was observed depending on the mass increase and the charge associated with modification. However, the enzyme could not tolerate thymidine analogues, **8**, **11**, **13**, **14** and **16** as a substrate for the PCR. Table 1 shows the relative incorporation rate of the modified substrate, estimated from the band intensity of the resulting modified DNA on the gel after staining with ethidium bromide and quantification by a molecular imager. The incorporation rate is expressed as the ratio of the modified DNA from each substrate analogue to the modified DNA from **1**. Thus, modified DNAs bearing imidazole acetic acid, urocanic acid nicotinic acid, phenanthroline 9-carboxylic acid or biotin *via* an amide linkage were obtained by PCR. The thymidine analogue with imidazole acetic acid or urocanic acid *via* hexamethylenediamine linker, **4** or **5**, works as a substrate for KOD Dash DNA polymerase, giving the PCR product (lane 7 or 8 of Fig. 1A). However, the enzyme could not tolerate a similar analogue bearing an imidazole group without a hexamenthylenediamine linker.[28] Aminohexanoic acid was added to the hexamethylenediamine linker of the thymidine analogue **1** to increase the length of the linker, forming **10**. The yield of the modified DNA from **10** decreased slightly compared to that from **1** (lane 13). The thymidine analogue containing a cationic-charged guanidinium group at the C5-position was also a good substrate for the enzyme. The modified DNA with a guanidinium group was formed efficiently by PCR from the modified substrate **15** (lane 6 of Fig. 1B). This is the first example that a guanidinium-substituted thymidine analogue works as a substrate during PCR, although a thymidine analogue bearing a guanidinium group *via* a propynyl linker **17** has been reported to be accepted by Tth or Pwo DNA polymerase in the template-directed primer extension reactions using a natural or modified DNA template (Scheme 2).[25,26] The oligonucleotides with a guanidinium group have been prepared by a DNA synthesizer using phosphoramidite chemistry and are shown to have high hybridization ability.[31,32] The PCR with the modified substrate **15** using KOD Dash DNA polymerase will provide a simple synthetic method for the modified DNA with a guanidinium ion, which is expected to have high hybridization and cell permeation ability and thus may be useful as an antisense agent. The analogue with pydridyldithio group **7** was slightly incorporated into the DNA by PCR. On the other hand, thymidine analogues bearing the anionic-charged carboxyl group **11** (lane 14), mercapto group **8** (lane 11) and 2,9-dimethylphenanthroline *via* thiourea linkage **13** (lane 16) were a poor substrate for the enzyme. The analogue with phenanthroline 9-carboxylic acid *via* amide linkage **12** works as a substrate and could be incorporated into the modified DNA (lane 15), whereas the analogue with 2,9-dimethylphenanthroline *via* 5-thiourea linkage was a very poor substrate for the enzyme. The analogue bearing fluorecsein *via* thiourea linkage **14** was also a poor substrate and no fluorescein-labelled DNA could be obtained when **14** was used in place of TTP (lane 17). However, combination use of TTP and **14** in a molar ratio of 5 : 5 gave the fluorescein-labelled DNA by PCR.[33] The result indicates that the enzyme could tolerate a single incorporation of **14**, but not successive multiple incorporation. The corresponding analogue with fluorescein carboxylic acid *via* amide linkage **16** also could not work as a substrate for the PCR, although the mixture of **16** and **1** in ratio of 9 : 1 yielded the resulting fluorescent-modified DNA. The enzyme could not accept the substrate with a mercapto group at the C5-position of **8**, because the mercapto group of **8** was converted to the disulfide-bridged dimer under the PCR conditions and the resulting dimer could not work as a substrate. We confirmed that the disulfide bridged dimer of **8** was formed easily by air oxidation when **8** was kept under the PCR conditions, as the buffer for the PCR did not contain a reducing agent such as DTT. Held and Benner reported that the thymidine analogue with a C5-position side chain carrying a thiol group protected as the disulfide with *t*-butyl group **18** can be incorporated as a substrate for PCR with Pwo DNA polymerase, although the corresponding analogue with a free thiol group was not a substrate for the enzyme.[22] These results suggest that steric

**Table 1** Relative yield of the modified DNA from thymidine analogues by PCR with KOD Dash DNA polymerase

| Nucleotide | Relative yield[a] | Nucleotide | Relative yield[a] | Nucleotide | Relative yield[a] |
|---|---|---|---|---|---|
| **1** | 1 | **7** | 0.04 | **13** | 0 |
| **2** | 0.90 | **8** | 0 | **14** | 0. |
| **3** | 0.67 | **9** | 0.43 | **15** | 1.09 |
| **4** | 0.80 | **10** | 0.76 | **16** | 0 |
| **5** | 0.60 | **11** | 0 | | |
| **6** | 0.41 | **12** | 0.32 | | |

[a] Relative yield is expressed as the ratio of the resulting modified DNA from each thymidine analogue to that from **1** under the same PCR conditions, estimated from the quantification of the modified DNA on the gel by a molecular imager.

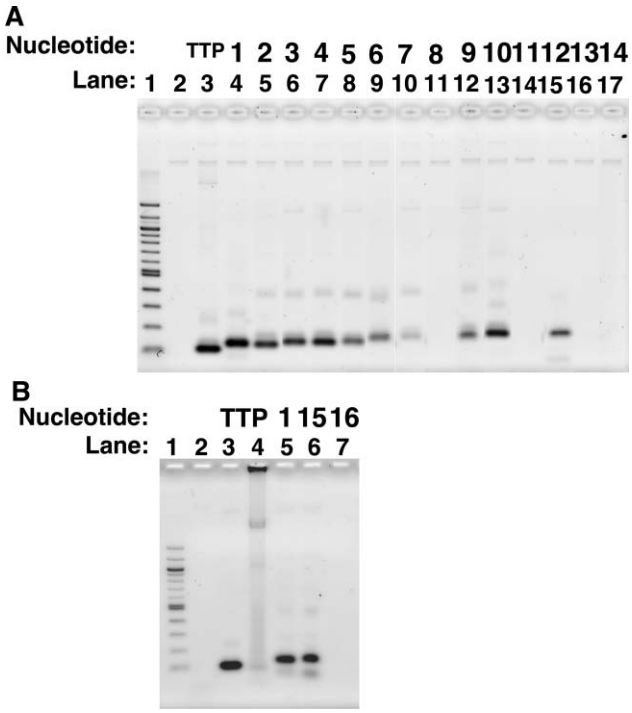



**Fig. 1** Modified DNA synthesis by PCR from dTTP and analogues bearing various functional groups using KOD Dash DNA polymerase. Total reaction mixture (20 μl) contained 0.5 ng $\mu l^{-1}$ of DNA template (pUC18 2686 bp), 0.2 μM of each primer, natural dNTPs or modified dNTP mix (0.2 mM of each nucleotide) and 0.5 unit per 10 μl of DNA polymerase (except lane 3, where natural TTP mix and 0.05 unit per 10 μl of DNA polymerase was used) in the buffer supplied by the maker for the DNA polymerase reaction. PCR assays were carried out at 94 °C for 1 min, 30 cycles of 94 °C for 30 sec, 52 °C for 30 sec, 74 °C for 1 min and 74 °C for 5 min. The PCR reaction mixture was quenched by the addition of a formamide-dye solution and aliquots of PCR products were analyzed by 2% agarose gel electrophoresis and visualized by staining with ethidium bromide. A) lane 1: marker DNA (100–1200 bp); lane 2: negative control without TTP (dATP + dCTP + dGTP); lane 3: positive control: TTP + dATP + dCTP + dGTP; lane 4: **1** + dATP + dCTP + dGTP; lane 5: **2** + dATP + dCTP + dGTP: lane 6: **3** + dATP+ dCTP + dGTP; lane 7: **4** + dATP + dCTP + dGTP; lane 8: **5** + dATP + dCTP + dGTP; lane 9: **6** + dATP + dCTP + dGTP; lane 10: **7** + dATP + dCTP + dGTP; lane 11: **8** + dATP + dCTP +dGTP; lane 12: **9** + dATP + dCTP + dGTP: lane 13: **10** + dATP + dCTP + dGTP; lane 14: **11** + dATP + dCTP + dGTP; lane 15: **12** + dATP + dCTP + dGTP; lane 16: **13** + dATP + dCTP + dGTP; lane 17: **14** + dATP + dCTP + dGTP. B) lane 1: marker DNA (100–1200 bp); lane 2: negative control without TTP (dATP + dCTP + dGTP); lane 3: positive control: TTP + dATP + dCTP + dGTP; lane 4: TTP + dATP + dCTP + dGTP (0.5 unit per 10 μl of DNA polymerase was used); lane 5: **1** + dATP + dCTP + dGTP: lane 6: **15** + dATP + dCTP + dGTP: lane 7: **16** + dATP + dCTP + dGTP.

effects of the substituent group and the length of a linker arm at C5 of thymidine may be important for the acceptance of the substrate by the enzyme. The enzyme also hardly accepted the thymidine analogue with anionic-charged carboxyl group **11** (lane 11). The analogue bearing a carboxyl group without a linker arm was also a poor substrate for the KOD DNA polumerase.[28] On the other hand, Pwo DNA polymerase accept the 7-deazaadenosine nucleotide bearing a carboxyl group with a propynyl linker **19** in place of dATP, forming the corresponding modified DNA.[25,26] KOD DNA polymerase has been reported to have two aspartic acid residues in the active site and the carboxyl group of the active site of the enzyme is crucial for the DNA polymerization ability.[34] The electrostatic repulsion between the carboxylic acids at the active site of the enzyme and the carboxylic acid of the substrate might be responsible for the poor substrate activity of the substrate with carboxylic acid. A cationic group at the C5-position of the thymidine analogues does not impede the interaction of the substrate with the active site of the enzyme, hence maintaining their substrate activity.

Accurate incorporation of 5′-triphosphate of the modified thymidine in place of TTP during PCR was checked by sequencing of the resulting modified DNA. The modified DNA was enzymatically converted into natural-type DNA, inserted into plasmid DNA and then sequenced with an ABI genetic analyzer. The sequencing of the modified DNA from 5′-triphosphate of the thymidine analogue bearing the guanidinium group is shown in Fig. 2. The result demonstrates that the modified DNA obtained by PCR has an identical sequence to that of the original one. Several other modified DNAs from other thymidine analogues have also the same sequence.[28,29] Thus, the modified thymidine analogue was correctly incorporated into the modified DNA and the resulting modified DNA can serve as a template for sequence-specific DNA amplification.

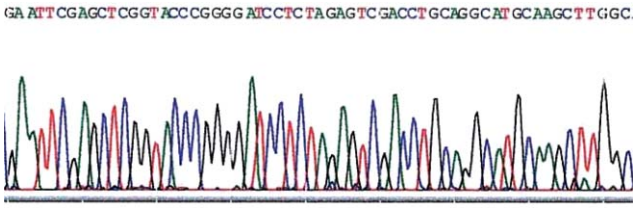



**Fig. 2** Sequencing of the PCR product from the 5′-triphosphate of modified thymidine **15** in place of TTP. The figure shows DNA sequence excluding the primer region.

## Conclusions

We have shown that triphosphates of thymidine analogues bearing various functional groups except an anionic, mercapto or bulky group can be incorporated into DNA by PCR using KOD Dash DNA polymerase, yielding the corresponding modified DNA. The modified DNAs with imidazole, pyridine and phenanthroline derivatives obtained by PCR in this method will be useful for the production of catalytic DNA by applying *in vitro* selection. The *in vitro* selection using the modified DNA with a cationic-charged guanidinium group or amino group may help in the production of highly active DNA aptamers against the anionic-charged target molecule, that cannot be obtained from the natural-type DNA.

## Experimental

### Analytical methods and gel electrophoresis

ESI-Mass spectra were recorded on a PE-Sciex API-100 mass spectrometer. UV spectra were obtained with a Shimazu 1200 spectrometer. $^{1}$H- and $^{31}$P-NMR spectra were obtained with a JEOL α-500 or a JEOL AL-300 spectrometer. Triphosphates of 2′-deoxyuridine derivatives for NMR measurement were passed through a Dowex 50-WX-8 ($Na^+$ form) column to convert the triethylammonium salt to the sodium salt. Tetramethylsilane (TMS) and 85% phosphoric acid were used as the internal standards for $^{1}$H- and $^{31}$P-NMR, respectively. High-pressure liquid chromatography (HPLC) on an ODS-silica gel column (4 mm × 250 mm) was carried out with a linear gradient elution of acetonitrile in 50 mM triethylammonium acetate at a flow rate of 1.0 ml $min^{-1}$. Gel electrophoresis was carried out using 2% agarose gel electrophoresis at 100 V for 45 min and was visualized by staining with ethidium bromide. The band intensities of the DNA on the gel were quantified with a Bio-Rad Molecular Imager FX Pro. The sequencing of the PCR product was carried out using a BigDye Terminator Cycle Sequencing Kit (Applied Biosystems) and an ABI Prism 310NT Genetic Analyzer.

### Materials

2′-Deoxynucleoside 5′-triphosphates, dATP, dGTP, dCTP and TTP, were purchased from Roche Diagnostics Co. KOD Dash

DNA polymerases were obtained from Toyobo Co. 5′-Triphosphates of 5-*N*-(6-aminohexyl)carbamoylmethyl-2′-deoxyuridine 5′-triphosphate (**1**) and 5-*N*-(6-trifluoroacetylaminohexyl)carbamoylmethyl-2′-deoxyuridine (**2**) were prepared as described previously.[27] Other C5-substituted deoxyuridine derivatives were prepared from **1**. The pUC18 plasmid DNA was from Toyobo Co. Oligonucleotides used as primers, DNA **A**, 5′-GGAAACAGCTATGACCATGATTAC-3′ and DNA **B**, 5′-CGACGTTGTAAAACGACGGCCAGT-3′, were from Sawaday Technology Co.

DNA ligation kit was purchased from Takara Co. GenElute plasmid purification kit and X-Gal were from Sigma. All other chemicals were reagent grade and were used without further purification.

### Synthesis of 5′-triphosphates of 2′-deoxyuridine derivatives

**5-*N*-(6-[Nicotinyl]aminohexyl)carbamoylmethyl-2′-deoxyuridine 5′-triphospahte (3).** A solution of dicyclohexylcarbodiimide (309 mg, 1.50 mmol) in dry DMF (5 ml) was added gradually to a solution of nicotinic acid (123 mg, 1.00 mmol) and *N*-hydroxysuccinimide (138 mg, 1.20 mmol) in dry DMF (5 ml), under nitrogen, in an ice bath, with stirring. The mixture was stirred for 1 h at 0 °C, then overnight at rt. The solution was evaporated to dryness and the residue was dissolved in ethyl acetate. The solution was filtered to remove the precipitate and concentrated by evaporation to give the *N*-hydroxysuccinimide ester of biotin as a white powder, 307 mg. The activated ester was used without further purification and was generated immediately prior to use.

5-*N*-(6-Aminohexyl)carbamoylmethyl-2′-deoxyuridine 5′-triphosphate (**1**) (60 $OD_{260\ nm}$, 6.5 μmol) was dissolved in a solution containing distilled water (300 μl) and 1.0 M sodium hydrogen carbonate buffer (pH 9.5) (200 μl) at rt. The solution of the active ester (40 mg, 180 μmol) in dry DMF (500 μl) was added and stirred for 20 h at 37 °C. The mixture was concentrated and purified by HPLC on an ODS-silica gel column (10 mm × 250 mm), with a linear gradient elution of acetonitrile (2.1–30.1%) in 50 mM triethylammonium acetate (pH 7.2) for 28 min at a flow rate of 3.0 ml $min^{-1}$. The purified **3** was obtained in 46% yield (42 $OD_{260\ nm}$, 3.0 μmol). The yield was calculated from the estimation that the molar absorption coefficient of the product is the sum of that of 5-*N*-(6-aminohexyl)carbamoylmethyl-2′-deoxyuridine and nicotinic acid. ESI-Mass (negative mode) *m*/*z*: found 728.0; calcd for $[M–H]^-$ 728.1. $^1H$ NMR ($D_2O$) $\delta$: 8.86 (1H, s, Py–H2), 8.69 (1H, d, Py–H6), 8.17 (1H, d, $J$ = 7.9 Hz, Py–H4), 7.91 (1H, s, H6), 7.58 (1H, dd, Py–H5), 6.29 (1H, dd, H1′), 4.69 (1H, m, H3′), 4.24–4.13 (3H, m, H4′ and H5′), 3.40 (2H, t, $J$ = 6.9 Hz, $NHCH_2$), 3.35 (2H, s, C5–$CH_2$), 3.20 (2H, m, $CH_2NH$), 2.40 (2H, m, H2′), 1.63 (2H, m, $NHCH_2CH_2$), 1.53 (2H, m, $CH_2CH_2NH$), 1.37 (4H, m, $CH_2CH_2$). $^{31}P$ NMR $\delta$: −5.7 (d, $J$ = 60 Hz, γ-P), −10.3 (d, $J$ = 60 Hz, α-P), −21.3 (t, $J$ = 60 Hz, β-P).

**5-*N*-(6-Guanidinumhexyl)carbamoylmethyl-2′-deoxyuridine 5′-triphospahte (15).** The guanidylation reagent, *S*-ethylthiopseudourea hydrobromide was prepared according to the published procedure.[35] Triethylamine (659 μl, 4.75 μmol) was added to a solution of **1** (100 $OD_{260\ nm}$, 10.8 μmol) and *S*-ethylthiopseudourea hydrobromide (440 mg, 2.38 μmol) in DMF (2.38 μl) with stirring at rt. The reaction mixture was stirred for 10 h at rt. The mixture was concentrated and purified by HPLC on an ODS-silica gel column (20 mm × 250 mm) with a linear gradient elution of acetonitrile (3.5–14%) in 50 mM triethylammonium acetate (pH 7.2) for 45 min at a flow rate of 8.0 ml $min^{-1}$. The purified nucleotide **15** was obtained in 70% yield (70 $OD_{260\ nm}$, 7.6 μmol). ESI-Mass (negative mode) *m*/*z*: found 665.1; calcd for $[M–H]^-$ 665.1. $^1H$ NMR ($D_2O$) $\delta$ 7.82 (s, 1H), 6.23 (t, 1H), 4.59 (m, 1H), 4.18–3.99 (m, 3H), 3.26 (s, 2H), 3.08 (m, 4H), 2.30 (m, 2H), 1.43 (m, 4H), 1.22 (m, 4H); $^{31}P$ NMR ($D_2O$) $\delta$ −5.6 (d), −10.6 (d), −21.4 (t).

Synthetic procedure of 5′-triphosphates of thymidine analogues, **4–14** and **16** was similar to that described above for the synthesis of **3** and **15** and a detailed procedure is described in the supporting information.† The yield and ESI-mass data of the nucleotides, **4–14** and **16**, are listed in the following.

**5-*N*-(6-[4-Imidazoleacetyl]aminohexyl)carbamoylmethyl-2′-deoxyuridine 5′-triphospahte (4).** The yield of **4** was 43% from **1**. ESI-Mass (negative mode) *m*/*z*: found 731.3; calcd for $[M–H]^-$ 731.1.

**5-*N*-(6-[Urocanyl]aminohexyl)carbamoylmethyl-2′-deoxyuridine 5′-triphospahte (5).** The nucleotide **5** was obtained in 96% yield from **1**. ESI-Mass (negative mode) *m*/*z*: found 723.1; calcd for $[M–H]^-$ 723.1.

**5-*N*-(6-[Biotinyl]aminohexyl)carbamoylmethyl-2′-deoxyuridine 5′-triphospahte (6).** The yield of **6** was 19% from **1**. ESI-Mass (negative mode) *m*/*z*: found 849.3; calcd for $[M–H]^-$ 849.2.

**5-*N*-(6-[3-(2-Pyridyldithio)propionyl]aminohexyl)carbamoylmethyl-2′-deoxyuridine 5′-triphospahte (7).** The yield of the nucleotide **7** was 59% from 1. ESI-Mass (negative mode) *m*/*z*: found 820.0; calcd for $[M–H]^-$ 820.1.

**5-*N*-(6-[3-Mercaptopropinonyl]aminohexyl)carbamoylmethyl-2′-deoxyuridine 5′-triphospahte (8).** The yield of the nucleotide **8** was 66%. ESI-Mass (negative mode) *m*/*z*: found 711.0; calcd for $[M–H]^-$ 711.1.

**5-*N*-(6-[ε-Trifulouroacetylamidohexyl]aminohexynyl)carbamoylmethyl-2′-deoxyuridine 5′-triphospahte (9).** The yield of the nucleotide **9** was 88% from **1**. ESI-Mass (negative mode) *m*/*z*: found 832.2; calcd for $[M–H]^-$ 832.2.

**5-*N*-(6-[Aminohexynyl]aminohexyl)carbamoylmethyl-2′-deoxyuridine 5′-triphospahte (10).** The nucleotide **10** was obtained in 93% yield from **1**. ESI-Mass (negative mode) *m*/*z*: found 736.4; calcd for $[M–H]^-$ 736.2.

**5-*N*-(6-Succinylaminohexyl)carbamoylmethyl-2′-deoxyuridine 5′-triphospahte (11).** The nucleotide **11** was obtained in 70% yield from **1**. ESI-Mass (negative mode) *m*/*z*: found 723.1; calcd for $[M–H]^-$ 723.1.

**5-*N*-(6-[9-Phenanthrylcarbonyl]aminohexyl)carbamoylmethyl-2′-deoxyuridine 5′-triphosphate (12).** The yield of **12** was 45% yield from **1**. ESI-Mass (negative mode) *m*/*z*: found 829.1; calcd for $[M–H]^-$ 8291.

**5-*N*-(6-[5-(2,9-Dimethylphenanthroline)thiouryl]aminohexyl)-carbamoyl-methyl-2′-deoxyuridine 5′-triphosphate (13).** The yield of **13** was 54% from **1**. ESI-Mass (negative mode) *m*/*z*: found 888.1; calcd for $[M–H]^-$ 888.2.

**5-*N*-(4-[Fluoresceinthiouryl]aminohexyl)carbamoyl-methyl-2′-deoxyuridine 5′-triphosphate (14).** The yield of **14** was 27% yield from **1**. ESI-Mass (negative mode) *m*/*z*: found 1012.3; calcd for $[M–H]^-$ 1012.3.

**5-*N*-(4-[Fluoresceinylcarbonyl]aminohexyl)carbamoyl-methyl-2′-deoxyuridine 5′-triphosphate (16).** The yield of **16** was 6% yield from **1**. ESI-Mass (negative mode) *m*/*z*: found 981.0; calcd for $[M–H]^-$ 981.1.

**Synthesis of modified DNAs by PCR with modified TTP using KOD Dash DNA polymerase.** The reaction mixture (20 μl) contained 0.5 ng $\mu l^{-1}$ DNA template (pUC18 2686 bp), 0.2 μM of each primer, 0.2 mM of natural dNTPs (dATP + dGTP + dCTP + TTP) or modified dNTP mix (dATP + dGTP + dCTP + modified TTP) and 0.05 unit per 10 μl or 0.5 unit per 10 μl of DNA polymerase, respectively, in the buffer supplied by the maker for the DNA polymerase reaction. PCR assays were carried out at 94 °C for 1 min, 30 cycles of 94 °C for 30 sec,

52 °C for 30 sec, 74 °C for 1 min and 74 °C for 5 min. The reaction mixture was quenched by the addition of a formamide-dye solution (90% formamide : 1% bromophenolblue : 1% xylenencyanol) and the PCR products were analyzed by 2% agarose gel electrophoresis after staining with ethidium bromide. The band intensity of the resulting modified DNA on the gel was quantified with a Bio-Rad Molecular Imager. Relative yield of the modified DNA from each thymidine analogue was estimated from the band intensity.

**Dimer formation from 8 under the PCR conditions.** The reaction mixture (20 μl) containing **8** (0.2 mM) in the buffer for PCR was kept under the PCR conditions. The mixture was analyzed by HPLC on an ODS-silica gel column (4 mm × 250 mm) with a linear gradient elution of acetonitrile (2–37%) in 50 mM triethylammonium acetate (pH 7.2) for 35 min at a flow rate of 1.0 ml $min^{-1}$. The peak due to **8** (retention time, 18.3 min) completely disappeared and one major peak (retention time, 20.6 min) and three minor peaks appeared in the HPLC. The major peak was collected and analyzed by ESI-mass to confirm the dimer. ESI-Mass (negative mode) *m/z*: found 710.1 and 1421.3; calcd for $[M–2H]^{2-}$ 2 710.1 and $[M–H]^{-}$ 1421.2.

**Sequencing of the modified DNA from the 5′-triphosphate of thymidine analogue bearing a guanidinium group.** Correct incorporation of the modified thymidine analogue in place of TTP was checked by sequencing of the resulting modified DNA. The modified DNA as obtained by PCR was converted in to the natural type DNA by PCR using natural type substrates and KOD Dash DNA polymerase. The DNA was purified by 2% agarose gel electrophoresis. The collected DNA was further applied to PCR using natural substrates and Taq DNA polymerase to give the DNA with 3′-terminal deoxyadenosine, which was reacted with T-vector using a DNA ligation kit according to the maker's protocol. The ligated plasmid DNA was treated with *E. coli* competent cells for transformation. The cells were cultured on the cultural medium containing ampicillin and the transformed cells were selected by the difference in colour after treating with X-gal. The transformed cells were applied to PCR using Taq DNA polymerase and primers **C** and **D**, 5′-GCCAAGCTTGGTACCGA-3′ and 5′-TAATACGACTCACTCACTATAGGG-3′, respectively, to check the insertion of the 108 mer DNA into the plasmid DNA. The formation of 258 mer DNA by this PCR was confirmed by 2% agarose gel electrophoresis. The colony of the cells in which the presence of the inserted plasmid DNA was confirmed was further cultured. The plasmid DNA was extracted with a GenElute Plasmid Purification Kit. The sequencing of the inserted DNA was carried out using a BigDye Terminator Cycle Sequencing Kit (Applied Biosystems) and an ABI Prism 310NT Genetic Analyser, according to the protocol.

## Acknowledgements

This work is partially supported by a Grant-in-Aid for Scientific Research from The Ministry of Education, Science and Technology, Japan.

## References

1 I. Luyten and P. Herdewijin, *Eur. J. Med. Chem.*, 1998, **33**, 515.
2 S. Verma and F. Eckstein, *Ann. Rev. Biochem.*, 1998, **67**, 99.
3 J. L. Ruth, in *Oligonucleotides and Analogues*, ed. F. Eckstein, IRL Press, Oxford, 1991, p. 255.
4 H. Sawai, A. Nakamura, S. Sekiguchi, K. Yumoto, M. Endoh and H. Ozaki, *J. Chem. Soc., Chem. Commun.*, 1994, 1997.
5 H. Ozaki, A. Nakamura, M. Arai, M. Endoh and H. Sawai, *Bull. Chem. Soc. Jpn.*, 1995, **68**, 1981.
6 S. W. Santoro and G. F. Joyce, *Proc. Natl. Acad. Sci. USA*, 1997, **94**, 4262.
7 D. S. Wilson and J. W. Szostak, *Ann. Rev. Biochem.*, 1999, **68**, 611.
8 R. Breaker, *Chem. Rev.*, 1997, **97**, 371.
9 M. Famulok, G. Mayer and M. Blind, *Acc. Chem. Res.*, 2000, **33**, 591.
10 S. E. Osborne and A. D. Ellington, *Chem. Rev.*, 1997, **97**, 349.
11 L. Gold, B. Polisky, O. Uhlenbeck and M. Yarus, *Ann. Rev. Biochem.*, 1995, **64**, 763.
12 W. Kusser, *Rev. Mol. Biotechnol.*, 2000, **74**, 27.
13 S. W. Santoro, G. F. Joyce, K. Sakthivel, S. Gramatikova and C. F. Barbas, III, *J. Am. Chem. Soc.*, 2000, **122**, 2433.
14 L. Lerner, Y. Roupioz, R. Ting and D. M. Perrin, *J. Am. Chem. Soc.*, 2002, **124**, 9960.
15 A. V. Sidrov, J. A. Grasby and D. M. Wiiliams, *Nucleic Acids Res.*, 2004, **32**, 1591.
16 R. P. Langer, A. A. Waldrop and D. C. Ward, *Proc. Natl. Acad. Sci. USA*, 1981, **78**, 6633.
17 K. Sakthivel and C. F. Barbas, III, *Angew. Chem., Int. Ed.*, 1998, **37**, 2872.
18 T. S. Godovikova, D. M. Kolpashchikov, T. N. Orlova, V. A. Richter, T. M. Ivanova, S. L. Grochovsky, T. V. Nasedkina, L. S. Victorova and A. I. Poletaev, *Bioconjugate Chem.*, 1999, **10**, 529.
19 J. A. Latham, R. Johnson and J. J. Toole, *Nucleic Acids Res.*, 1994, **22**, 2817.
20 T. R. Battersby, D. N. Ang, P. Burgstaller, S. C. Jurczyk, M. T. Bowser, D. D. Buchanan, R. T. Kennedy and S. A. Benner, *J. Am. Chem. Soc.*, 1999, **121**, 9781.
21 S. E. Lee, A. Sidrov, T. Gourlain, N. Mignet, S. J. Thorpe, J. A. Brazier, M. J. Dickman, D. P. Hornby, J. A. Grasby and D. M. Williams, *Nucleic Acids Res.*, 2001, **29**, 1565.
22 H. A. Held and S. A. Benner, *Nucleic Acid Res.*, 2002, **30**, 3857.
23 D. M. Perrin, T. Garestier and C. Helen, *Nucleosides Nucleotides*, 1999, **18**, 377.
24 T. Gourlain, A. Sidrov, N. Mignet, S. J. Thorpe, S. E. Lee, J. A. Grasby and D. M. Williams, *Nucleic Acids Res.*, 2001, **29**, 1898.
25 O. Thum, S. Jaeger and M. Falmulok, *Angew. Chem., Int. Ed.*, 2001, **40**, 3990.
26 S. Jaeger and M. Falmulok, *Angew. Chem., Int. Ed.*, 2004, **43**, 3337.
27 H. Sawai, A. Ozki-Nakamura, M. Mine and H. Ozaki, *Bioconjugate Chem.*, 2002, **13**, 309.
28 H. Sawai, A. N. Ozaki, F. Satoh, T. Ohbayashi, M. M. Masud and H. Ozaki, *Chem. Commun.*, 2001, 2604.
29 M. Kuwahara, Y. Takahata, A. Shoji, A. N. Ozaki, H. Ozaki and H. Sawai, *Bioorg. Med. Chem. Lett.*, 2003, **13**, 3735.
30 M. Takagi, M. Nishioka, H. Kakihara, M. Kita, H. Inoue, B. Kawakami, M. Oka and T. Imanaka, *Appl. Environ. Microbiol.*, 1997, **63**, 4504.
31 V. Roig and U. Asseline, *J. Am. Chem. Soc.*, 2003, **125**, 4416.
32 T. Prakash, A. Plischl, E. Lesiak, V. Mohan, V. Tereskho, M. Egli and M. Manoharan, *Org. Lett.*, 2004, **6**, 1971.
33 T. Ohbayashi, M. M. Masud, A. N. Ozaki, H. Ozaki, M. Kuwahara and H. Sawai, *Bioorg. Med. Chem. Lett.*, 2002, **12**, 1167.
34 H. Hashimoto, M. Nishioka, S. Fujiwara, H. Takagi, T. Imanaka, T. Inou and Y. Kai, *J. Mol. Biol.*, 2001, **306**, 469.
35 E. Brand and F. C. Brand, *Org. Synth.*, 1955, **3**, 440.