Supplementary Information for:

DNA amplification and functionalization PCR by an unnatural base pair system

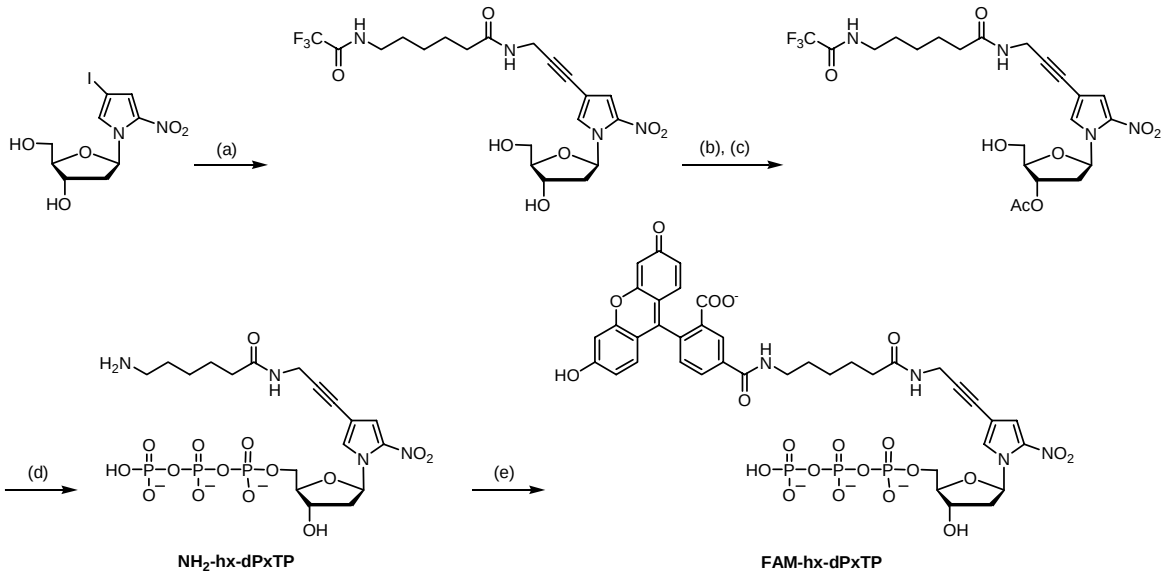
Michiko Kimoto[1,2], Rie Kawai[1], Tsuneo Mitsui[2], Shigeyuki Yokoyama[1,3], Ichiro Hirao[1,2*]

[1]Systems and Structural Biology Center, RIKEN, and [2]TagCyx Biotechnologies, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan, [3]Department of Biophysics and Biochemistry, Graduate School of Science, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan

CONTENTS

Supplementary Methods

Chemical synthesis of NH_2-hx-dPxTP and FAM-hx-dPxTP



Conditions: (a) CuI, $Pd[P(C_6H_5)_3]_4$, DMF, triethylamine, room temperature, then *N*-(2-propynyl)-6-trifluoroacetamidohexanamide; (b) DMTr-Cl, pyridine, room temperature; (c) acetic anhydride, pyridine, room temperature, then dichloroacetic acid, dichloromethane, 0 °C; (d) 2-chloro-4*H*-1,3,2-benzodioxaphosphorin-4-one/dioxane, pyridine, tri-*n*-butylamine, bis(tri-*n*-butylammonium)pyrophosphate, DMF, then I_2/pyridine/H_2O, H_2O, NH_4OH, room temperature; (e) FAM-*N*-hydroxysuccinimidyl ester/DMF, 0.1 M $NaHCO_3$-Na_2CO_3 buffer (pH 8.5), room temperature, then NH_4OH.

1-(2-Deoxy-β-D-ribofuranosyl)-4-[3-(6-trifluoroacetamidohexanamido)-1-propynyl]-2-nitropyrrole.

1-(2-Deoxy-β-D-ribofuranosyl)-4-iodo-2-nitropyrrole (354 mg, 1 mmol) was co-evaporated with dry CH_3CN twice. The residue was dissolved in DMF (5.0 ml) with CuI(I) (31 mg, 160 μmol) and $Pd[P(C_6H_5)_3]_4$ (58 mg, 50 μmol), followed by triethylamine (210 μl, 1.5 mmol), and the mixture was stirred in the dark at room temperature. To the solution,

N-(2-propynyl)-6-trifluoroacetamidohexanamide (396 mg, 1.5 mmol) in DMF (4 ml) was added dropwise, and the solution was stirred overnight at room temperature. The reaction mixture was concentrated *in vacuo*. The product was purified from the residue by silica gel column chromatography (5-10% CH_3OH in CH_2Cl_2) and C18-HPLC (39-41% CH_3CN in H_2O, 15 min) to give the title compound (408 mg, 83%) as the β isomer.

^{1}H NMR (300 MHz, DMSO-d6) δ 9.39 (brs, 1H), 8.29 (t, 1H, J = 5.4 Hz), 7.96 (d, 1H, J = 2.2 Hz), 7.31 (d, 1H, J = 2.2 Hz), 6.55 (t, 1H, J = 5.8 Hz), 5.29 (d, 1H, J = 4.5 Hz), 5.12 (t, 1H, J = 5.1 Hz), 4.24 (m, 1H), 4.06 (d, 2H, J = 5.4 Hz), 3.85 (m, 1H), 3.70-3.53 (m, 2H), 3.16 (m, 2H), 2.48-2.40 (m, 1H), 2.28-2.20 (m, 1H), 2.09 (t, 2H, J = 7.5 Hz), 1.55-1.42 (m, 4H), 1.28-1.18 (m, 2H). HRMS (FAB, 3-NBA matrix) for $C_{20}H_{26}F_3N_4O_7$ $[M + H]^+$: calcd, 491.1754; found, 491.1761.

1-(2-Deoxy-3-*O*-acetyl-β-D-ribofuranosyl)-4-[3-(6-trifluoroacetamidohexanamido)-1-propynyl]-2-nitropyrrole.

1-(2-Deoxy-β-D-ribofuranosyl)-4-[3-(6-trifluoroacetamidohexanamido)-1-propynyl]-2-nitropyrrole (394 mg, 803 μmol) was co-evaporated with dry pyridine three times. The residue was dissolved in pyridine (4 ml) with 4,4'-dimethoxytrityl chloride (286 mg, 844 μmol), and the solution was stirred at room temperature for 2.5 h. The solution was partitioned with EtOAc and H_2O. The organic layer was washed with a saturated $NaHCO_3$ solution, dried with $MgSO_4$, and evaporated *in vacuo*. The residue was purified by silica gel column chromatography (0-0.5% CH_3OH in CH_2Cl_2) to obtain the dimethoxytrityl derivative (543 mg). 1-(2-Deoxy-5-*O*-dimethoxytrityl-β-D-ribofuranosyl)-4-[(3-(6-trifluoroacetamidohexanamido)-1-propynyl]-2-nitropyrrole (542 mg, 684 μmol) was co-evaporated with dry pyridine three times. To the residue in pyridine (7 ml) was added acetic anhydride (169 μl, 1.79 mmol). The mixture was stirred overnight at room temperature. The solution was partitioned with EtOAc and a saturated $NaHCO_3$ solution. The organic layer was washed with brine, dried with

$MgSO_4$, and evaporated *in vacuo*. To the residue in dichloromethane (68 ml) was added dichloroacetic acid (680 μl), and the solution was stirred at 0 °C for 15 min. The reaction mixture was poured into a saturated $NaHCO_3$ solution, and was extracted with CH_2Cl_2. The combined organic layer was dried with $MgSO_4$ and evaporated *in vacuo*. The product (328 mg, 77%, 2 steps) was purified by silica gel column chromatography (2% CH_3OH in CH_2Cl_2).

^{1}H NMR (300 MHz, DMSO-*d*6) δ 9.39 (brs, 1H), 8.30 (t, 1H, J = 5.4 Hz), 7.94 (d, 1H, J = 2.2 Hz), 7.33 (d, 1H, J = 2.2 Hz), 6.61 (t, 1H, J = 6.4 Hz), 5.26 (brs, 1H), 5.22 (m, 1H), 4.13 (m, 1H), 4.06 (d, 2H, J = 5.4 Hz), 3.64 (m, 2H), 3.16 (t, 2H, J = 7.0 Hz), 2.62 (ddd, 1H, J = 3.0, 6.0, 14.2 Hz), 2.43 (m, 1H), 2.10 (m, 2H), 2.07 (s, 3H), 1.49 (m, 4H), 1.24 (m, 2H). HRMS (FAB, 3-NBA matrix) for $C_{22}H_{28}F_3N_4O_8$ $[M + H]^+$: calcd, 533.1859; found, 533.1907.

1-(2-Deoxy-β-D-ribofuranosyl)-4-[3-(6-aminohexanamido)-1-propynyl]-2-nitropyrrole 5'-triphosphoric acid (NH_2-hx-dPxTP).

1-(2-Deoxy-3-*O*-acetyl-β-D-ribofuranosyl)-4-[3-(6-trifluoroacetamidohexanamido)-1-propyn yl]-2-nitropyrrole (53 mg, 100 μmol) was co-evaporated with dry pyridine three times. The residue was dissolved in pyridine (100 μl) and dioxane (300 μl). A 1 M solution of 2-chloro-4*H*-1,3,2-benzodioxaphosphorin-4-one in dioxane (110 μl, 110 μmol) was added to the solution, and then the reaction mixture was stirred at room temperature for 10 min. Tri-*n*-butylamine (100 μl) was added to the reaction mixture, followed by a 0.5 M solution of bis(tri-*n*-butylammonium)pyrophosphate in DMF (300 μl). After 10 min, 1% iodine in pyridine/H_2O (98/2, v/v) (2.0 ml) was then added. After stirring for 15 min, the oxidation was quenched by the addition of a 5% aqueous solution of $NaHSO_3$ (150 μl). The reaction mixture was then evaporated *in vacuo*, and the residue was dissolved in H_2O (5 ml). After stirring for 30 min, concentrated ammonia (20 ml) was added to the solution, and the reaction mixture was stirred for 8 h and lyophilized. The product (18 μmol, 18%) was purified by DEAE Sephadex A-25 column chromatography (1.5 cm x 30 cm, eluted by a linear gradient from 50

mM to 1 M TEAB) and by C18 HPLC (eluted by a linear gradient of CH_3CN in 100 mM TEAA, pH 7.0).

^{1}H NMR (300 MHz, D_2O) δ 7.79 (s, 1H), 7.31 (s, 1H), 6.68 (t, 1H, J = 5.7 Hz), 4.56 (m, 1H), 4.16 (m, 3H), 4.07 (s, 2H), 3.13 (q, 18H, J = 7.3 Hz), 2.90 (t, 2H, J = 7.6 Hz), 2.58 (m, 1H), 2.42 (m, 1H), 2.24 (t, 2H, J = 7.0 Hz), 1.60 (m, 4H), 1.33 (m, 2H), 1.20 (t, 27H, J = 7.3 Hz).

^{31}P NMR (121 MHz, D_2O) δ -8.26 (1P), -10.51 (1P), -22.05 (1P).

MS (ESI) for $C_{18}H_{28}N_4O_{15}P_3$, $[M-H]^-$: calcd, 633.08; found, 632.99. UV (10 mM sodium phosphate buffer, pH 7.0) ε 368 = 11,100.

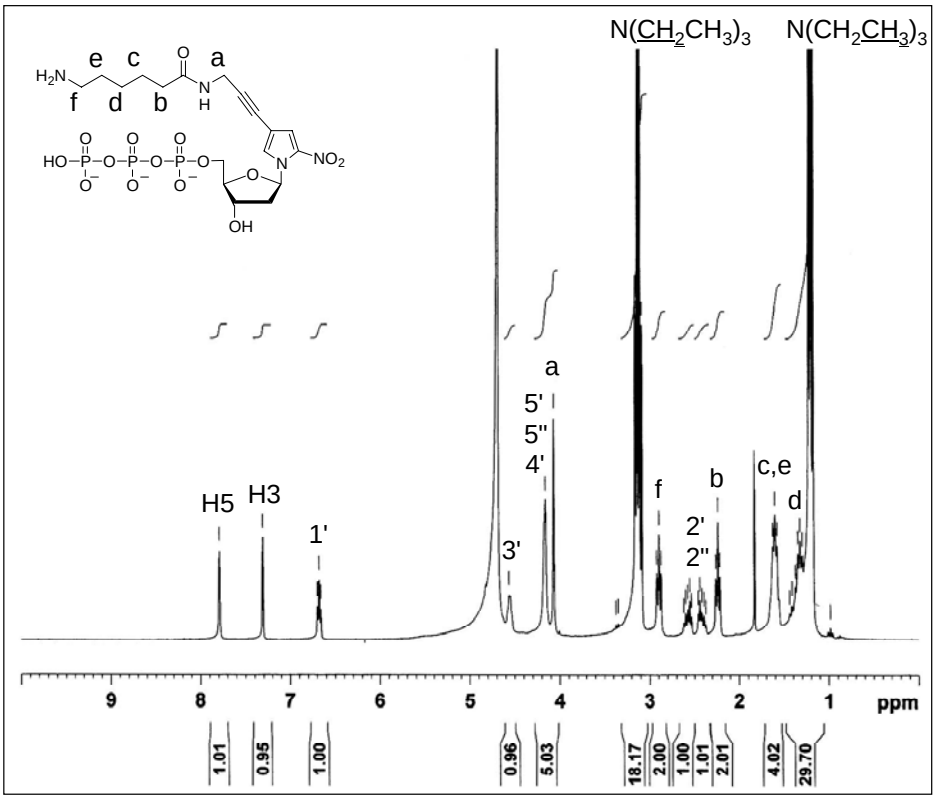
1-(2-Deoxy-β-D-ribofuranosyl)-4-[3-(6-fluorescein-5-carboxamidohexanamido)-1-propynyl] -2-nitropyrrole 5'-triphosphoric acid (FAM-hx-dPxTP).

A 0.1 M $NaHCO_3$-Na_2CO_3 buffer solution (pH 8.5, 0.72 ml) of NH_2-hx-d**Px**TP (6 μmol) was reacted with 5-carboxyfluorescein *N*-hydroxysuccinimidyl ester (FAM-SE) (3.7 mg, 7.8 μmol) in DMF (500 μl) in the dark at room temperature. After 8 h, the reaction mixture was treated with concentrated ammonia (0.5 ml) for 5 min. The product (3.2 μmol, 53%) was purified by DEAE Sephadex A-25 column chromatography (1.5 cm x 30 cm, eluted by a linear gradient from 50 mM to 1 M TEAB) and C18 HPLC (eluted by a linear gradient of CH_3CN in 100 mM TEAA, pH 7.0).

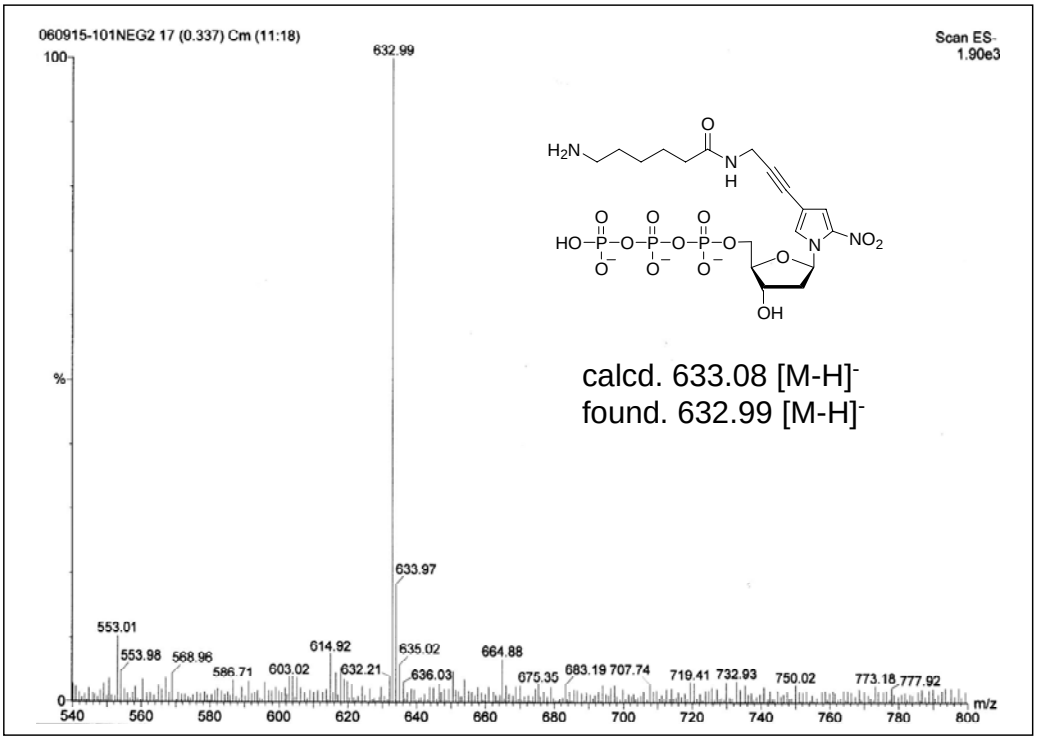
^{1}H NMR (300 MHz, D_2O) δ 8.29 (s, 1H), 8.04 (d, 1H, J = 7.4 Hz), 7.58 (brs, 1H), 7.26 (d, 1H, 7.5 Hz), 7.00-6.69 (m, 7H), 6.39 (brs, 1H), 4.40 (m, 1H), 4.03 (m, 3H), 3.94 (s, 2H), 3.41 (m, 2H), 3.12 (q, 18H, J = 7.3 Hz), 2.41 (m, 1H), 2.24 (m, 2H), 2.13 (m, 1H), 1.63 (m, 4H), 1.38 (m, 2H), 1.20 (t, 27H, J = 7.3 Hz).

^{31}P NMR (121 MHz, D_2O) δ -10.86 (2P), -23.13 (1P).

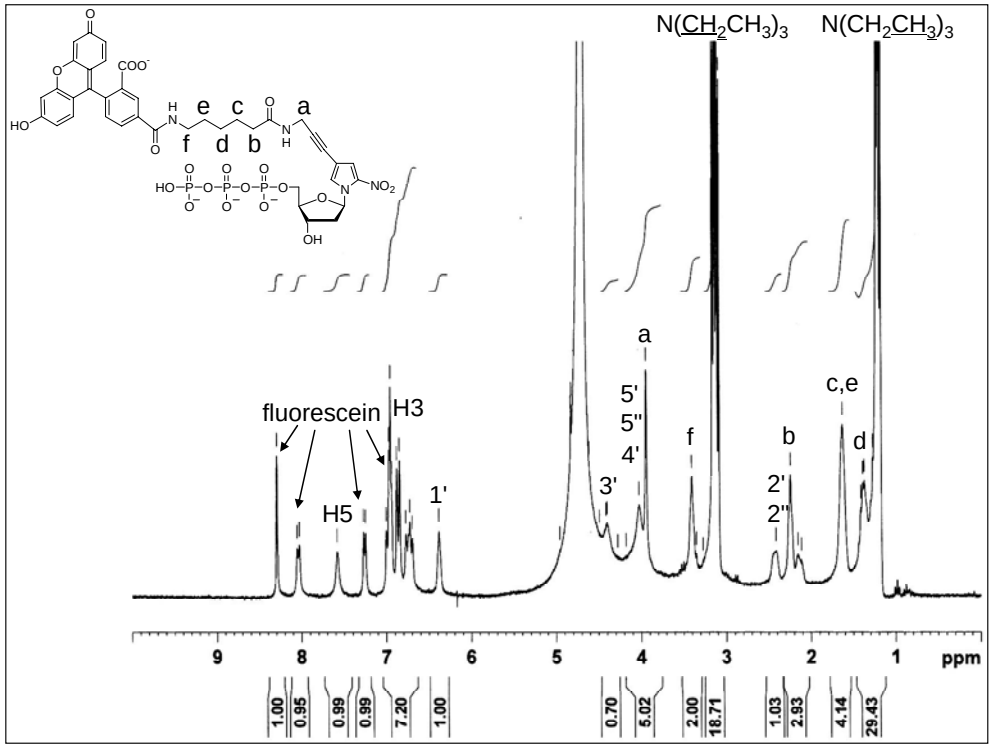
MS (ESI) for $C_{39}H_{38}N_4O_{21}P_3$ $[M-H]^-$: calcd, 991.12; found, 990.58. UV (10 mM sodium phosphate buffer, pH 7.0) ε 493 = 64,400.



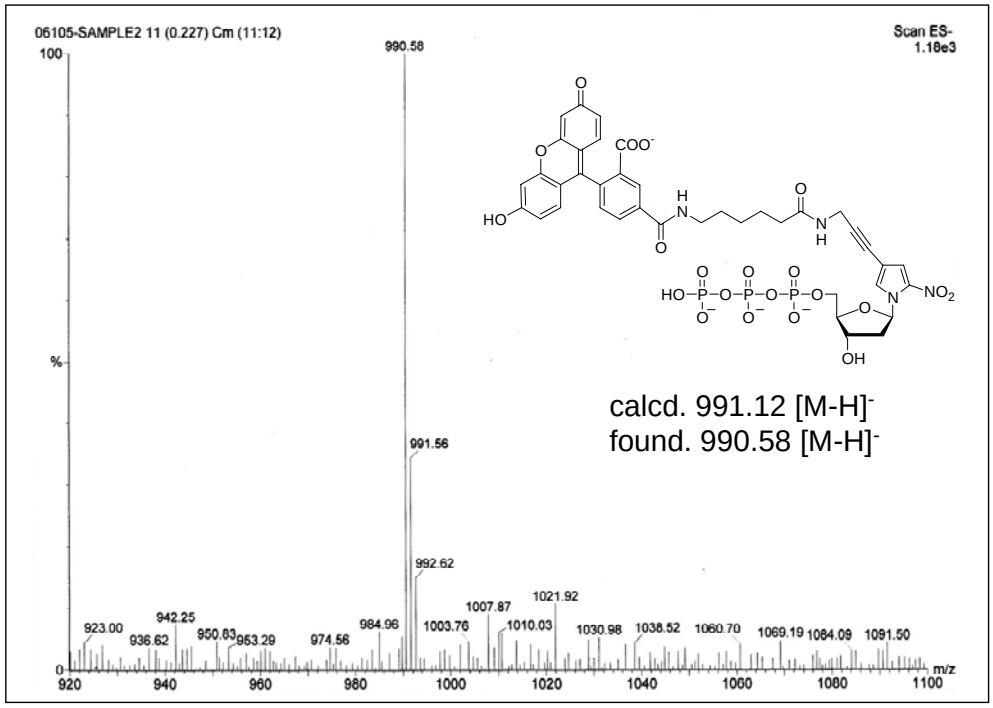
^{1}H NMR (300 MHz, D_2O) spectrum of NH_2-hx-dPxTP



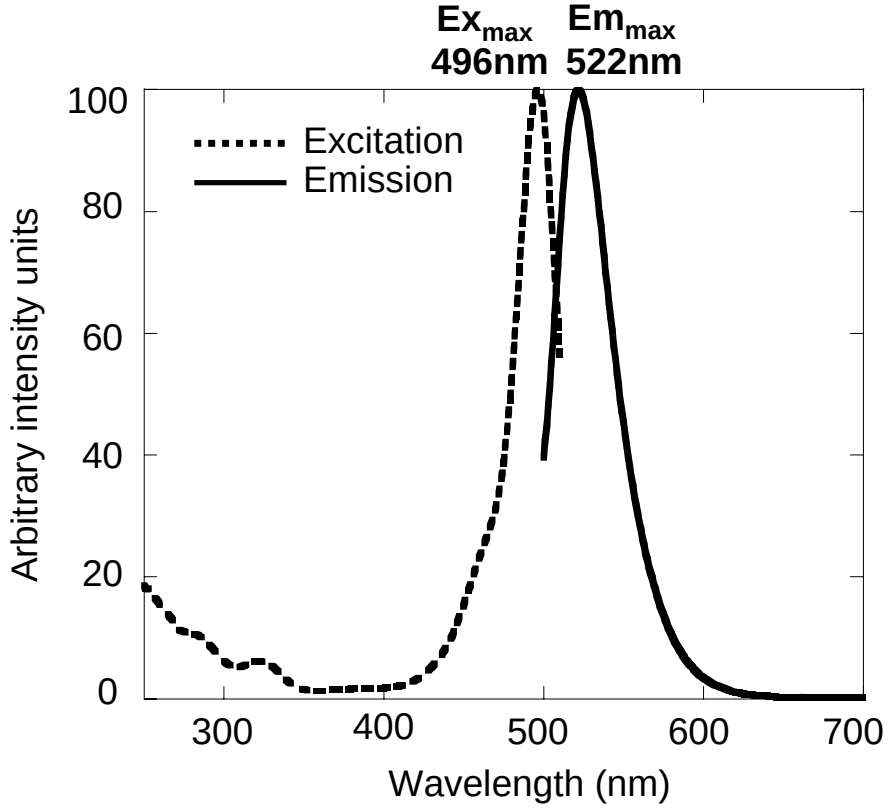
ESI-mass spectrum of NH_2-hx-dPxTP



^{1}H NMR (300 MHz, D_2O) spectrum of FAM-hx-dPxTP



ESI-mass spectrum of FAM-hx-dPxTP



Excitaion and emission spectra of FAM-hx-dPxTP

Single-nucleotide insertion experiments of the Ds-Px pairing

A primer (20-mer) labeled with 6-carboxyfluorescein at the 5'-end was annealed with a template (35-mer), in 100 mM Tris-HCl (pH 7.5) buffer containing 20 mM $MgCl_2$, 2 mM DTT, and 0.1 mg/ml bovine serum albumin. The primer-template duplex solution (2 or 10 μM, 5 μl) was mixed with 2 μl of an enzyme solution containing the exonuclease-deficient Klenow fragment, KF exo– (Amersham USB). The mixture was incubated for more than 2 min, and then the reactions were initiated by adding each dNTP solution (3 μl) to the duplex-enzyme mixture at 37°C. The amount of enzyme used (2–20 nM), the reaction time (1–28.2 min), and the gradient concentration of dNTP (6–600 μM) were adjusted to give reaction extents of 25% or less. The reactions were quenched with 10 μl of stop solution (95% formamide and 20 mM EDTA), and the mixtures were immediately heated at 75°C for 3 min. The diluted products

were analyzed on an automated ABI 377 DNA sequencer equipped with the *GeneScan* software (version 3.0). Relative velocities (v_0) were calculated as the extents of the reaction divided by the reaction time, and were normalized to the enzyme concentration (20 nM) and to the duplex concentration (5 μM) for the various enzyme and duplex concentrations used. The kinetic parameters (K_M and V_{max}) were obtained from Hanes-Woolf plots of [dNTP]/v_0 against [dNTP]. The kinetic parameters are listed in Supplementary Table 1.

Supplementary Table 1. Steady-state kinetic parameters of single-nucleotide insertion experiments of the Ds-Px pairing

Primer	5'-ACTCACTATAGGGAGGAAGA				
Template	3'-TATTATGCTGAGTGATATCCCTCCTTCT**N**TCTCGA				
Entry	Template (**N**)	Nucleoside triphosphate	K_M (μM)	V_{max} (% min^{-1})	Efficiency (V_{max}/K_M)[c]
1	**Ds**	NH_2-hx-d**Px**TP	58 (20)[b]	43 (10)	7.4×10^5
2	A	NH_2-hx-d**Px**TP	100 (21)	2.2 (0.9)	2.2×10^4
3	G	NH_2-hx-d**Px**TP	69 (14)	0.24 (0.09)	3.5×10^3
4	C	NH_2-hx-d**Px**TP	160 (90)	0.13 (0.05)	8.1×10^2
5	T	NH_2-hx-d**Px**TP	250 (110)	10 (4)	4.0×10^4
6[a]	**Ds**	d**Pn**TP	91 (6)	34 (5)	3.7×10^5
7[a]	A	d**Pn**TP	130 (60)	2.7 (1.2)	2.1×10^4
8[a]	G	d**Pn**TP	80 (56)	1.1 (0.4)	1.4×10^4
9[a]	C	d**Pn**TP	140 (130)	0.10 (0.05)	7.1×10^2
10[a]	T	d**Pn**TP	140 (40)	2.9 (0.6)	2.1×10^4
11[a]	**Ds**	d**Pa**TP	340 (150)	21 (3)	6.2×10^4
12[a]	**Ds**	d**Ds**TP	8.0 (3.9)	1.6 (0.1)	2.0×10^5
13[a]	**Ds**	dATP	150 (40)	0.36 (0.09)	2.4×10^3
14[a]	**Ds**	dGTP	n.d.[d]	n.d.[d]	-[d]
15[a]	**Ds**	dCTP	410 (190)	0.34 (0.05)	8.3×10^2
16[a]	**Ds**	dTTP	220 (20)	0.41 (0.17)	1.9×10^3
17[a]	A	dTTP	0.70 (0.40)	2.8 (1.5)	4.0×10^6
18[a]	A	dCTP	1200 (600)	2.2 (0.9)	1.8×10^3
19[a]	G	dCTP	0.24 (0.18)	5.5 (1.7)	2.3×10^7
20[a]	G	dTTP	140 (70)	0.29 (0.12)	2.1×10^3

Assays (Entries 1-5) were carried at 37ºC for 1-28.2 min using 1 or 5 μM template-primer duplex, 2-20 nM enzyme, and 6-600 μM nucleoside triphosphate in a solution (10 μl) containing 50 mM Tris-HCl (pH 7.5), 10 mM $MgCl_2$, 1 mM DTT, and 0.05 mg/ml bovine serum albumin. Each parameter was averaged from three to five data sets. [a]The parameters were derived from the previous data (Entries 6-19, *Nature Methods*, **3**, 729-735, 2006 & *J. Am. Chem. Soc.*, **129**, 15549-15555, 2007). [b]Standard deviations are given in parentheses. [c]The units of this term are % min^{-1} M^{-1}. [d]Reaction was too slow to calculate parameters (V_{max} <0.05).

Results

These results indicate the high efficiency and selectivity of the **Ds-Px** pairing in replication. The incorporation efficiency of NH_2-hx-**Px** opposite **Ds** ($V_{max}/K_M = 7.4 \times 10^5$ %·min^{-1}·M^{-1}) was higher than those of **Pa** ($V_{max}/K_M = 6.2 \times 10^4$ %·min^{-1}·M^{-1}) and **Pn** ($V_{max}/K_M = 3.7 \times 10^5$ %·min^{-1}·M^{-1}) opposite **Ds**, and the efficiency of the **Ds-Px** pairing was also superior to those of the non-cognate pairings, especially the **Ds-Ds** pairing ($V_{max}/K_M = 2.0 \times 10^5$ %·min^{-1}·M^{-1}).

Supplementary Table 2. Sequences of chemically synthesized DNA fragments

Primers for Sequencing and PCR
5'-Primer for *in vitro* selection (20-mer)
5'-GATAATACGACTCACTATAG-3'

5'-Primer for PCR (40-mer)
5'-CGTTGTAAAACGACGGCCAGGATAATACGACTCACTATAG-3'

Primer for sequencing of PCR products (20-mer)
5'-CGTTGTAAAACGACGGCCAG-3'

3'-Primer for *in vitro* selection and PCR (24-mer)
5'-TTTCACACAGGAAACAGCTATGAC-3'

DNA fragments (55-mer, the primer regions are underlined.)
Ds-Library
5'-TTTCACACAGGAAACAGCTATGACGGNNN**Ds**NNNCCCTATAGTGAGTCGTATTA
TC-3'

DNA S1
5'-TTTCACACAGGAAACAGCTATGACGGCAC**Ds**TTGCCCTATAGTGAGTCGTATTA
TC-3'

DNA S2
5'-TTTCACACAGGAAACAGCTATGACGGCCC**Ds**TTGCCCTATAGTGAGTCGTATTA
TC-3'

DNA S3
5'-TTTCACACAGGAAACAGCTATGACGGCGC**Ds**TTGCCCTATAGTGAGTCGTATTA
TC-3'

DNA S4
5'-TTTCACACAGGAAACAGCTATGACGGTAC**Ds**TTGCCCTATAGTGAGTCGTATTA
TC-3'

DNA S5
5'-TTTCACACAGGAAACAGCTATGACGGCAC**Ds**TCGCCCTATAGTGAGTCGTATTA
TC-3'

DNA S6
5'-TTTCACACAGGAAACAGCTATGACGGTGC**Ds**TTGCCCTATAGTGAGTCGTATTA
TC-3'

DNA S7
5'-TTTCACACAGGAAACAGCTATGACGGTAC**Ds**TCGCCCTATAGTGAGTCGTATTA
TC-3'

DNA S8
5'-TTTCACACAGGAAACAGCTATGACGGATC**Ds**TATCCCTATAGTGAGTCGTATTA
TC-3'

DNA N9
5'-TTTCACACAGGAAACAGCTATGACGGTAC**Ds**TTCCCCTATAGTGAGTCGTATTA
TC-3'

DNA N10
5'-TTTCACACAGGAAACAGCTATGACGGTAG**Ds**TTGCCCTATAGTGAGTCGTATTA

TC-3'

DNA N11
5'-TTTCACACAGGAAACAGCTATGACGGATG**Ds**AACCCCTATAGTGAGTCGTATTA
TC-3'

DNA N12
5'-TTTCACACAGGAAACAGCTATGACGGTAC**Ds**GTGCCCTATAGTGAGTCGTATTA
TC-3'

DNA Cont
5'-TTTCACACAGGAAACAGCTATGACGGATCCATTCCCTATAGTGAGTCGTATTAT
C-3'

DNA Cont2
5'-TTTCACACAGGAAACAGCTATGACACATGGAACTGCTATAGTGAGTCGTATTAT
C-3'

DNA Cont3
5'-TTTCACACAGGAAACAGCTATGACCATGATGCAGACTATAGTGAGTCGTATTAT
C-3'

DNA Cont4
5'-TTTCACACAGGAAACAGCTATGACTTGATCCGTATCTATAGTGAGTCGTATTAT
C-3'

DNA fragments containing two Ds bases: (the primer regions are underlined.)
DNA2-60 (60-mer)
5'-TTTCACACAGGAAACAGCTATGACGGCCCDsTTACDsGTGCCCTATAGTGAGTCG
TATTATC-3'
DNA2-62 (62-mer)
5'-TTTCACACAGGAAACAGCTATGACGGCCCDsTTGTACDsGTGCCCTATAGTGAGT
CGTATTATC-3'

DNA2-65 (65-mer)
5'-TTTCACACAGGAAACAGCTATGACGGCCCDsTTGTAATACDsGTGCCCTATAGTG
AGTCGTATTATC-3'

DNA2-68 (68-mer)
5'-TTTCACACAGGAAACAGCTATGACGGCCCDsTTGTAACGATACDsGTGCCCTATA
GTGAGTCGTATTATC-3'

Supplementary Table 3. PCR amplification of DNA fragments involving the Ds-Px pair

DNA	3'-nnn **Ds** nnn-5'[a]	Number of clones isolated from the 5th round of selection	Fold amplification	
			PCR: 15 cycles Input: 15 fmol[b]	PCR: 30 cycles Input: 0.15 amol[c]
S1	GTT **Ds** CAC	3	3.3×10^2	-
S2	GTT **Ds** CCC	4	3.4×10^2	2.1×10^7
S3	GTT **Ds** CGC	2	3.1×10^2	-
S4	GTT **Ds** CAT	3	3.2×10^2	1.7×10^7
S5	GCT **Ds** CAC	2	3.2×10^2	-
S6	GTT **Ds** CGT	3	3.3×10^2	-
S7	GCT **Ds** CAT	1	2.9×10^2	8.0×10^6
S8	TAT **Ds** CTA	2	2.9×10^2	-
N9	CTT **Ds** CAT	0	2.0×10^2	-
N10	GTT **Ds** GAT	0	1.3×10^2	-
N11	CAA **Ds** GTA	0	33	-
N12	GTG **Ds** CAT	0	3.3×10^2	-
Library	NNN **Ds** NNN	-	1.9×10^2	-
Cont	TTA C CTA	0	7.1×10^2	3.9×10^7

a) Bases with conserved ratios of more than 35% in the sequences of the 66 clones are colored red.

b) The fold amplification of each single-stranded DNA fragment (initial amount = 15 fmol) after 15 cycles of PCR using FAM-hx-d**Px**TP and d**Ds**TP.

c) The fold amplification of each single-stranded DNA fragment (initial amount = 0.15 amol) after 30 cycles of PCR using FAM-hx-d**Px**TP and d**Ds**TP.

Supplementary Figure 1. Selected sequences of 66 clones obtained from the library after five rounds of *in vitro* selection

The sequences of the **Px**-containing DNA strands are listed. The number of clones is shown in parentheses, and the sequences for more than one clone are underlined. The DNA sequences shown in Table S3 are indicated in brackets.

```
5′-GATAATACGACTCACTATAGGG-NNN Px NNN–CCGTCATAGCTGTTTCCTGTGTGAAA-3′
(total 66 clones)

(4) GATAATACGACTCACTATAGGG CAA X GGG CCGTCATAGCTGTTTCCTGTGTGAAA      [S2]
(1) GATAATACGACTCACTATAGGG TAA X GGG CCGTCATAGCTGTTTCCTGTGTGAAA
(1) GATAATACGACTCACTATAGGG CAA X GGT CCGTCATAGCTGTTTCCTGTGTGAAA
(3) GATAATACGACTCACTATAGGG CAA X GTG CCGTCATAGCTGTTTCCTGTGTGAAA      [S1]
(2) GATAATACGACTCACTATAGGG CAA X GCG CCGTCATAGCTGTTTCCTGTGTGAAA      [S3]
(1) GATAATACGACTCACTATAGGG CAA X GTC CCGTCATAGCTGTTTCCTGTGTGAAA
(3) GATAATACGACTCACTATAGGG CAA X GTA CCGTCATAGCTGTTTCCTGTGTGAAA      [S4]
(3) GATAATACGACTCACTATAGGG CAA X GCA CCGTCATAGCTGTTTCCTGTGTGAAA      [S6]
(1) GATAATACGACTCACTATAGGG GAA X GAG CCGTCATAGCTGTTTCCTGTGTGAAA
(1) GATAATACGACTCACTATAGGG CCA X GTG CCGTCATAGCTGTTTCCTGTGTGAAA
(1) GATAATACGACTCACTATAGGG CCA X GAG CCGTCATAGCTGTTTCCTGTGTGAAA
(1) GATAATACGACTCACTATAGGG ACA X GGG CCGTCATAGCTGTTTCCTGTGTGAAA
(1) GATAATACGACTCACTATAGGG CCA X GAC CCGTCATAGCTGTTTCCTGTGTGAAA
(1) GATAATACGACTCACTATAGGG GCA X GAC CCGTCATAGCTGTTTCCTGTGTGAAA
(1) GATAATACGACTCACTATAGGG CCA X GCA CTGTCATAGCTGTTTCCTGTGTGAAA
(1) GATAATACGACTCACTATAGGG AGA X GCA CCGTCATAGCTGTTTCCTGTGTGAAA
(1) GATAATACGACTCACTATAGGG TGA X GCA CCGTCATAGCTGTTTCCTGTGTGAAA
(2) GATAATACGACTCACTATAGGG CGA X GTG CCGTCATAGCTGTTTCCTGTGTGAAA      [S5]
(1) GATAATACGACTCACTATAGGG TGA X GTG CCGTCATAGCTGTTTCCTGTGTGAAA
(1) GATAATACGACTCACTATAGGG GGA X GTG CCGTCATAGCTGTTTCCTGTGTGAAA
(1) GATAATACGACTCACTATAGGG AGA X GTG CCGTCATAGCTGTTTCCTGTGTGAAA
(1) GATAATACGACTCACTATAGGG CGA X GTA CCGTCATAGCTGTTTCCTGTGTGAAA      [S7]
(2) GATAATACGACTCACTATAGGG CGA X GGG CCGTCATAGCTGTTTCCTGTGTGAAA
(1) GATAATACGACTCACTATAGGG AGA X GTA CCGTCATAGCTGTTTCCTGTGTGAAA
(1) GATAATACGACTCACTATAGGG CGA X GCG CCGTCATAGCTGTTTCCTGTGTGAAA
(1) GATAATACGACTCACTATAGGG GGA X GCGCCCGTCATAGCTGTTTCCTGTGTGAAA
(2) GATAATACGACTCACTATAGGG GGA X GGG CCGTCATAGCTGTTTCCTGTGTGAAA
(1) GATAATACGACTCACTATAGGG AGA X GGG CCGTCATAGCTGTTTCCTGTGTGAAA
(1) GATAATACGACTCACTATAGGG GGA X GAG CCGTCATAGCTGTTTCCTGTGTGAAA
(1) GATAATACGACTCACTATAGGG CGA X GAG CCGTCATAGCTGTTTCCTGTGTGAAA
(1) GATAATACGACTCACTATAGGG CGA X AGG CCGTCATAGCTGTTTCCTGTGTGAAA
(2) GATAATACGACTCACTATAGGG ATA X GAT CCGTCATAGCTGTTTCCTGTGTGAAA      [S8]
(1) GATAATACGACTCACTATAGGG GTA X GAT CCGTCATAGCTGTTTCCTGTGTGAAA
(1) GATAATACGACTCACTATAGGG ATA X GCG CCGTCATAGCTGTTTCCTGTGTGAAA
(1) GATAATACGACTCACTATAGGG ATA X GTG CCGTCATAGCTGTTTCCTGTGTGAAA
(1) GATAATACGACTCACTATAGGG ATA X GGC CCGTCATAGCTGTTTCCTGTGTGAAA
(1) GATAATACGACTCACTATAGGG ATA X GGA CCGTCATAGCTGTTTCCTGTGTGAAA
(1) GATAATACGACTCACTATAGGG TTA X GCT CCGTCATAGCTGTTTCCTGTGTGAAA
(1) GATAATACGACTCACTATAGGG ATA X AAC CCGTCATAGCTGTTTCCTGTGTGAAA
(1) GATAATACGACTCACTATAGGG AAA X AAC CCGTCATAGCTGTTTCCTGTGTGAAA
(1) GATAATACGACTCACTATAGGG AAA X ACT CCGTCATAGCTGTTTCCTGTGTGAAA
(2) GATAATACGACTCACTATAGGG GGA X ACG CCGTCATAGCTGTTTCCTGTGTGAAA
(1) GATAATACGACTCACTATAGGG TGA X ATG CCGTCATAGCTGTTTCCTGTGTGAAA
(1) GATAATACGACTCACTATAGGG CCA X ACG CCGTCATAGCTGTTTCCTGTGTGAAA
(1) GATAATACGACTCACTATAGGG TGA X CAT CCGTCATAGCTGTTTCCTGTGTGAAA
(1) GATAATACGACTCACTATAGGG CAA X TAA CCGTCATAGCTGTTTCCTGTGTGAAA
(1) GATAATACGACTCACTATAGGG TAC X ACG CCGTCATAGCTGTTTCCTGTGTGAAA
(1) GATAATACGACTCACTATAGGG TCC X GAC CCGTCATAGCTGTTTCCTGTGTGAAA
(1) GATAATACGACTCACTATAGGG CCC X TTC CCGTCATAGCTGTTTCCTGTGTGAAA
(1) GATAATACGACTCACTATAGGG GAT X ACA CCGTCATAGCTGTTTCCTGTGTGAAA
(1) GATAATACGACTCACTATAGGG CAT X TAT CCGTCATAGCTGTTTCCTGTGTGAAA
```