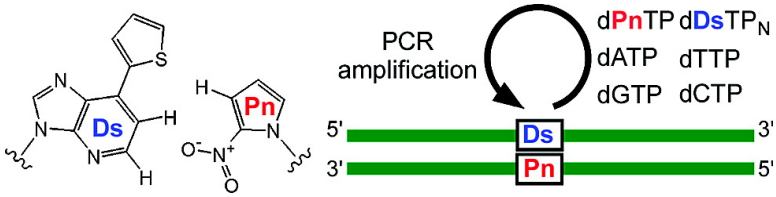
Article

An Efficient Unnatural Base Pair for PCR Amplification

Ichiro Hirao, Tsuneo Mitsui, Michiko Kimoto, and Shigeyuki Yokoyama

J. Am. Chem. Soc., **2007**, 129 (50), 15549-15555 • DOI: 10.1021/ja073830m



More About This Article

Additional resources and features associated with this article are available within the HTML version:

- Supporting Information
- Links to the 3 articles that cite this article, as of the time of this article download
- Access to high resolution figures
- Links to articles and content related to this article
- Copyright permission to reproduce figures and/or text from this article

View the Full Text HTML

J|A|C|S
ARTICLES
Published on Web 11/21/2007

An Efficient Unnatural Base Pair for PCR Amplification

Ichiro Hirao,*,† Tsuneo Mitsui,† Michiko Kimoto,† and Shigeyuki Yokoyama*,†,‡,§

Contribution from the Protein Research Group, RIKEN Genomic Sciences Center, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan, Department of Biophysics and Biochemistry, Graduate School of Science, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan, and RIKEN Harima Institute at SPring-8, 1-1-1 Kouto, Mikazuki-cho, Sayo, Hyogo 679-5148, Japan

Received May 28, 2007; E-mail: ihirao@riken.jp; yokoyama@biochem.s.u-tokyo.ac.jp

Abstract: Expansion of the genetic alphabet by an unnatural base pair system provides a powerful tool for modern biotechnology. As an alternative to previous unnatural base pairs, we have developed a new pair between 7-(2-thienyl)imidazo[4,5-*b*]pyridine (**Ds**) and 2-nitropyrrole (**Pn**), which functions in DNA amplification. **Pn** more selectively pairs with **Ds** in replication than another previously reported pairing partner, pyrrole-2-carbaldehyde (**Pa**). The nitro group of **Pn** efficiently prevented the mispairing with A. High efficiency and selectivity of the **Ds**−**Pn** pair in PCR amplification were achieved by using a substrate mixture of the γ-amidotriphosphate of **Ds** and the usual triphosphates of **Pn** and the natural bases, with Vent DNA polymerase as a 3′ to 5′ exonuclease-proficient polymerase. After 20 cycles of PCR, the total mutation rate of the **Ds**−**Pn** site in an amplified DNA fragment was ∼1%. PCR amplification of DNA fragments containing the unnatural **Ds**−**Pn** pair would be useful for expanded genetic systems in DNA-based biotechnology.

Introduction

Expansion of the genetic alphabet by an unnatural base pair system has great potential in providing a variety of novel nucleic acids and proteins with functional components of interest.[1] This can be achieved by creating an unnatural base pair that is compatible with the natural A−T and G−C base pairs in replication, transcription, and translation. Many unnatural base pairs have been synthesized and tested in replication.[2] Among them, a hydrophobic, unnatural base pair between 7-(2-thienyl)-imidazo[4,5-*b*]pyridine (**Ds**) and pyrrole-2-carbaldehyde (**Pa**) with specific shape complementation exhibits high selectivity in PCR amplification, with the combination of the usual triphosphate substrates for G, C, T, and **Pa**, and modified γ-amidotriphosphates for A ($dATP_N$) and **Ds** ($d\mathbf{Ds}TP_N$) (Figure 1).[3] Both γ-amidotriphosphates increase the selectivity of the **Ds**−**Pa** pairing in replication; $dATP_N$ and $d\mathbf{Ds}TP_N$ reduce the mispairings of A−**Pa** and **Ds**−**Ds**, respectively.

Although the **Ds**−**Pa** pair can be used practically for *in vitro* applications, the use of the γ-amidotriphosphates, especially $dATP_N$, decreases the efficiency of PCR amplification and restricts *in vivo* applications. Since the number of A bases in DNA fragments is typically much larger than that of **Ds** bases, the use of $dATP_N$ significantly reduces the amplification efficiency. Thus, further development of another unnatural base, instead of **Pa**, which more selectively pairs with **Ds** than A, would bypass the need for $dATP_N$.

Here, we describe a novel unnatural base, 2-nitropyrrole (**Pn**), as an efficient pairing partner of the hydrophobic **Ds** base (Figure 1A). To prevent mispairing with A, we replaced the aldehyde group of **Pa** with the nitro group of **Pn**. The nitro group was expected to electrostatically repel the 1-nitrogen of A and, thus, increase the selectivity of the **Ds**−**Pn** pairing in replication. In addition, nitro groups have been used as components of universal bases, such as 3-nitropyrrole and 5-nitroindole, since they enhance the stacking interactions with neighboring bases, but are not strong hydrogen-bond acceptors with pairing bases.[4] Thus, the nitro group might be a useful component for the development of unnatural, non-hydrogen-bonded base pairs. We chemically synthesized the **Pn** triphos-

† RIKEN Genomic Sciences Center.
‡ The University of Tokyo.
§ RIKEN Harima Institute at SPring-8.

(1) (a) Bain, J. D.; Switzer, C.; Chamberlin, A. R.; Benner, S. A. *Nature* **1992**, *356*, 537−539. (b) Benner, S. A.; Burgstaller, P.; Battersby, T. R.; Jurczyk, S. In *The RNA World*, 2nd ed.; Gesteland, R. F., Cech, T. R., Atkins, J. F., Eds.; Cold Spring Harbor Laboratory Press: Cold Spring Harbor, NY, 1999; p 163. (c) Hirao, I.; Ohtsuki, T.; Fujiwara, T.; Mitsui, T.; Yokogawa, T.; Okuni, T.; Nakayama, H.; Takio, K.; Yabuki, T.; Kigawa, T.; Kodama, K.; Yokogawa, T.; Nishikawa, K.; Yokoyama, S. *Nat. Biotechnol.* **2002**, *20*, 177−182. (d) Benner, S. A.; Sismour, A. M. *Nat. Rev. Genet.* **2005**, *6*, 533−543. (e) Hunziker, J.; Mathis, G. *Chimia* **2005**, *59*, 780−784. (f) Kawai, R.; Kimoto, M.; Ikeda, S.; Mitsui, T.; Endo, M.; Yokoyama, S.; Hirao, I. *J. Am. Chem. Soc.* **2005**, *127*, 17286−17295. (g) Prudent, J. R. *Expert Rev. Mol. Diagn.* **2006**, *6*, 245−252. (h) Hirao, I. *Biotechniques* **2006**, *40*, 711−717.
(2) (a) Kool, E. T. *Curr. Opin. Chem. Biol.* **2000**, *4*, 602−608. (b) Henry, A. A.; Romesberg, F. E. *Curr. Opin. Chem. Biol.* **2003**, *7*, 727−733. (c) Hirao, I. *Curr. Opin. Chem. Biol.* **2006**, *10*, 622−627.
(3) Hirao, I.; Kimoto, M.; Mitsui. T.; Fujiwara, T.; Kawai, R.; Sato, A.; Harada, Y.; Yokoyama, S. *Nat. Methods* **2006**, *3*, 729−735.
(4) (a) Nichols, R.; Andrews, P. C.; Zhang, P.; Bergstrom, D. E. *Nature* **1994**, *369*, 492−493. (b) Loakes, D.; Brown, D. M. *Nucleic Acids Res.* **1994**, *22*, 4039−4043. (c) Aerschot, A. V.; Rozenski, J.; Loakes, D.; Pillet, N.; Schepers, G.; Herdewijn, P. *Nucleic Acids Res.* **1995**, *23*, 4363−4370. (d) Loakes, D.; Brown, D. M.; Linde, S.; Hill, F. *Nucleic Acids Res.* **1995**, *23*, 2361−2366. (e) Bergstrom, D. E.; Zhang, P.; Toma, P. H.; Andrews, P. C.; Nichols, R. *J. Am. Chem. Soc.* **1995**, *117*, 1201−1209. (f) Amosova, O.; George, J.; Fresco, J. R. *Nucleic Acids Res.* **1997**, *25*, 1930−1934.

10.1021/ja073830m CCC: $37.00

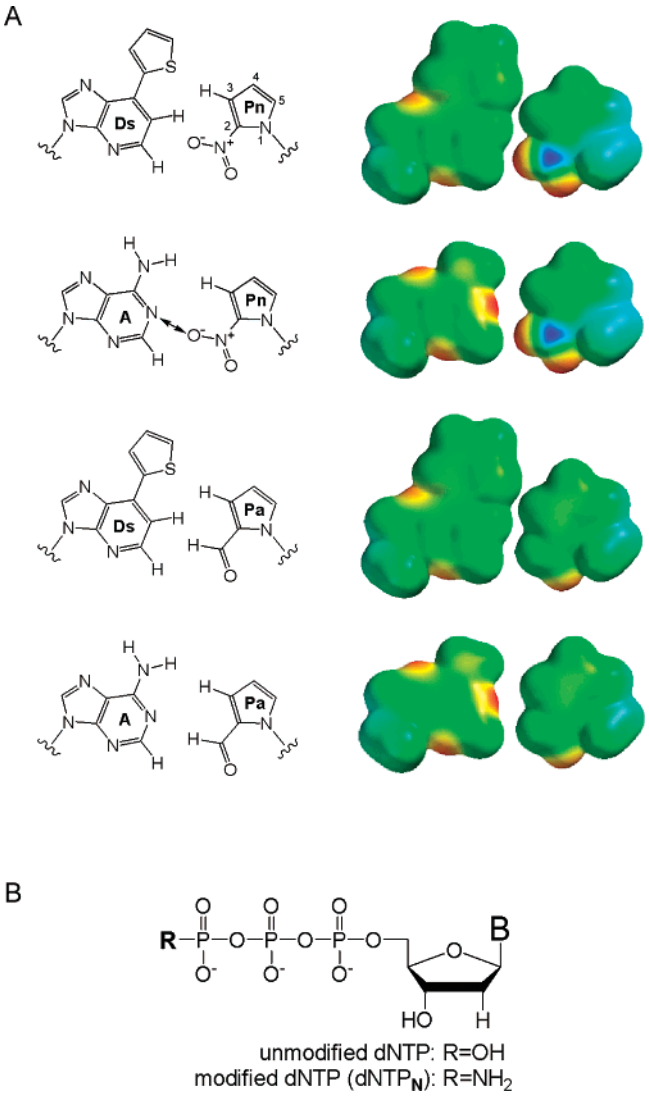


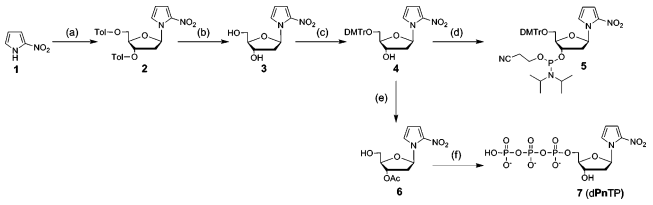
Figure 1. (A) Structures of the unnatural **Ds**−**Pn** and **Ds**−**Pa** pairs and A−**Pn** and A−**Pa** mispairs. Space-filling models of bases alone (with methyl in place of deoxyribose) are shown, with electrostatic potentials mapped on van der Waals surfaces (PM3 calculations, Spartan '06, Wavefunction Inc.). The arrow in the A−**Pn** pair indicates the electrostatic repulsion between the 1-nitrogen of A and the nitro group of **Pn**. (B) The structures of usual (R = OH) and modified γ-amidotriphosphate (R = NH_2) substrates.

phate and DNA fragments containing **Pn** and examined the ability of the **Ds**−**Pn** pairing in replication, in comparison to that of the **Ds**−**Pa** pairing. We found that PCR amplification using the **Ds**−**Pn** pair did not require $dATP_N$ for higher efficiency and fidelity, as compared to the amplification using the **Ds**−**Pa** pair.

Results and Discussion

Chemical Synthesis. The 2′-deoxyribonucleoside of **Pn** (**3**) was synthesized by the coupling of a fully protected ribofuranosyl chloride with the sodium salt of 2-nitropyrrole (**1**),[5] followed by deprotection (Scheme 1). The structure of **3** was confirmed by NMR and high-resolution mass spectroscopy (see Supporting Information 1 and 2). The HMQC and HMBC spectra revealed the *N*-glycoside bond between the sugar C1′ and the N-1 position of the 2-nitropyrrole base moiety. The anomeric configuration of **3** was identified as the β-form by the ^{1}H/^{1}H NOESY spectrum. The NOESY spectrum showed cross-peaks between H1′ and H4′ and between H5 and H3′/H2″.

Scheme 1. Synthesis of the Nucleoside Derivatives of 2-Nitropyrrole[a]



[a] Conditions: (a) NaH, 1-chloro-2-deoxy-3,5-di-*O*-toluoyl-α-D-*erythro*-pentofuranose, CH_3CN; (b) NH_3, methanol; (c) 4,4′-dimethoxytrityl chloride, pyridine; (d) 2-cyanoethyl-*N*,*N*-diisopropylamino chloro phosphoramidite, diisopropylethylamine, THF; (e) acetic anhydride, pyridine, then dichloroacetic acid, dichloromethane; (f) 2-chloro-4*H*-1,3,2-benzodioxaphosphorin-4-one, dioxane, pyridine, tri-*n*-butylamine, bis(tributylammonium)pyrophosphate, DMF, then I_2/pyridine, water, NH_4OH. Tol: toluoyl. DMTr: 4,4′-dimethoxytrityl. Ac: acetyl.

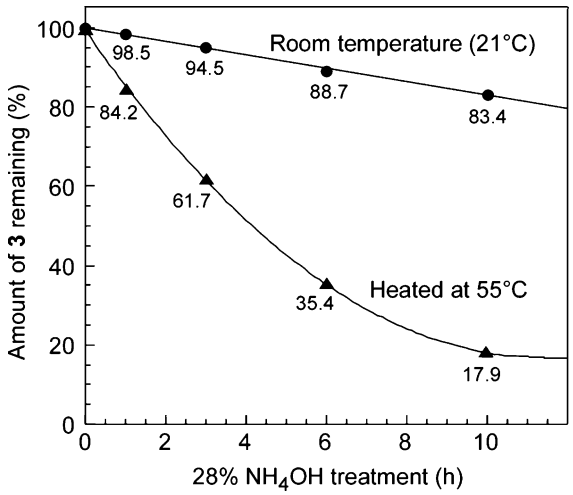


Figure 2. Stability of the 2′-deoxyribonucleoside of **Pn** (**3**) under basic conditions. Compound **3** was treated with concentrated ammonia at 55 °C (▲) or room temperature (●) for 1, 3, 6, and 10 h. The amounts (%) of **3** remaining were determined by HPLC using an internal standard, thymidine.

The nucleoside was converted to the amidite (**5**) for DNA chemical synthesis by conventional methods and to the triphosphate (d**Pn**TP, **7**) for enzymatic incorporation. The UV and visible spectra of d**Pn**TP in 10 mM phosphate buffer (pH 7.0) showed absorption maxima (λ_{max}) of 232 nm (ϵ 3100) and 352 nm (ϵ 12 600), which were similar to those of 2-nitropyrrole in methanol [λ_{max} = 231 nm (ϵ 4074) and 335 nm (ϵ 17 000)],[5] rather than those of 3-nitropyrrole [λ_{max}=268 nm (ϵ 7244) and 315 nm (ϵ 5370)][5] and 1-(2-deoxy-β-D-ribofuranosyl)-3-nitropyrrole [λ_{max} = 284 nm (ϵ 4850)].[4e]

We found that the nucleoside derivatives of **Pn** were considerably decomposed by treatments with concentrated ammonia at high temperature. The decomposition of **3** was analyzed and quantified by HPLC (Figure 2). After the treatment of **3** with concentrated ammonia at 55 °C for 6 h, 65% of **3** was decomposed. Under these conditions, the removal of 2-nitropyrrole from a short DNA fragment containing **Pn**, d(TT**Pn**T), was observed by mass spectrometry (data not shown). Thus, for DNA synthesis involving **Pn**, a base-labile protection method, using phenoxyacetyl for A, *p*-isopropylphe-

(5) Morgan, K. J.; Morrey, D. P. *Tetrahedron* **1966**, *22*, 57−62.

Table 1. Steady-State Kinetic Parameters for Insertion of Single Nucleotides into a Template−Primer Duplex[a]

primer temp35N-1	5′ACTCACTATAGGGAGGAAGA 3′-TATTATGCTGAGTGATATCCCTCCTTCT*N*TCTCGA				
entry	template (*N*)	nucleoside triphosphate	K_M (μM)	V_{max} (% min^{-1})	efficiency (V_{max}/K_M)[c]
1	**Ds**	d**Pn**TP	91 (6)[b]	34 (5)	3.7×10^5
2	A	d**Pn**TP	130 (60)	2.7 (1.2)	2.1×10^4
3	G	d**Pn**TP	80 (56)	1.1 (0.4)	1.4×10^4
4	C	d**Pn**TP	140 (130)	0.10 (0.05)	7.1×10^2
5	T	d**Pn**TP	140 (40)	2.9 (0.6)	2.1×10^4
6[d]	**Ds**	d**Pa**TP	340 (150)	21 (3)	6.2×10^4
7[d]	A	d**Pa**TP	330 (160)	17 (7)	5.2×10^4
8[d]	G	d**Pa**TP	140 (20)	0.061 (0.006)	4.4×10^2
9[d]	C	d**Pa**TP	n.d[e]	n.d	-
10[d]	T	d**Pa**TP	170 (60)	0.053 (0.016)	3.1×10^2
11[d]	**Ds**	d**Ds**TP	8.0 (3.9)	1.6 (0.1)	2.0×10^5
12[d]	**Ds**	dATP	150 (40)	0.36 (0.09)	2.4×10^3
13[d]	**Ds**	dGTP	n.d.	n.d.	-
14[d]	**Ds**	dCTP	410 (190)	0.34 (0.05)	8.3×10^2
15[d]	**Ds**	dTTP	220 (20)	0.41 (0.17)	1.9×10^3
16[d]	**Ds**	d**Ds**TP_N	79 (13)	0.78 (0.12)	9.9×10^3
17[d]	A	dTTP	0.70 (0.40)	2.8 (1.5)	4.0×10^6
18[d]	A	dCTP	1200 (600)	2.2 (0.9)	1.8×10^3
19[d]	G	dCTP	0.24 (0.18)	5.5 (1.7)	2.3×10^7
20[d]	G	dTTP	140 (70)	0.29 (0.12)	2.1×10^3

[a] Assays were carried out at 37 °C for 1.4−20 min, using 5 μM template−primer duplex, 5−50 nM enzyme, and 6−1500 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 four data sets. [b]Standard deviations are given in parentheses. [c]The units of this term are % min^{-1} M^{-1}. [d]These parameters were referred from *Nat. Methods* **2006**, *3*, 729−735. [e]Not determined. Minimal insert products (<2%) were detected after an incubation for 20 min with 1500 μM nucleoside triphosphate and 50 nM enzyme.

noxyacetyl for G, and acetyl groups for C,[6] was used, and deprotection was carried out at room temperature for 2−3 h with concentrated ammonia. Under these conditions, the decomposition of DNA fragments containing **Pn** was reduced to less than 10%. No decomposition of either the **Pn** nucleoside in DNA fragments or d**Pn**TP was observed under the conditions corresponding to 40 cycles of PCR amplification in a pH 7.5 buffer (data not shown).

Replication Studies. We first compared the efficiency and selectivity of the **Ds**−**Pn** pairing in replication with those of the **Ds**−**Pa** pairing, by analyzing the steady-state kinetics of single-nucleotide insertion experiments[7] using the exonuclease-deficient Klenow fragment (KF exo$^-$). The insertion experiments using KF exo$^-$ are reliable for assessing the replication ability of newly developed, unnatural base pairs.[8] In addition, we previously reported that the kinetics data using KF exo$^-$ are informative for the application of unnatural base pairs to PCR amplification.[3] The experiments were carried out using a template strand (35-mer, Tables 1 and 2) containing **Pn**, **Pa**, or **Ds**, a primer labeled with 6-carboxyfluorescein (20-mer), and each triphosphate, such as d**Pn**TP, d**Pa**TP, d**Ds**TP, d**Ds**TP_N, $dATP_N$, and the natural triphosphates. The insertion products were analyzed with an automated ABI 377 DNA sequencer with the *GeneScan* software.[9]

The incorporation efficiency of d**Pn**TP into DNA opposite **Ds** in a template ($V_{max}/K_M = 3.7 \times 10^5$ %·min^{-1}·M^{-1}) was 6-fold higher than that of d**Pa**TP opposite **Ds** ($V_{max}/K_M = 6.2 \times 10^4$ %·min^{-1}·M^{-1}). In contrast, the misincorporation of d**Pn**TP opposite A ($V_{max}/K_M = 2.1 \times 10^4$ %·min^{-1}·M^{-1}) decreased by 2.5-fold, as compared to that of d**Pa**TP opposite A ($V_{max}/K_M = 5.2 \times 10^4$ %·min^{-1}·M^{-1}). Similarly, the incorporation efficiency of dATP opposite **Pn** ($V_{max}/K_M = 5.5 \times 10^3$ %·min^{-1}·M^{-1}) was also 7.8-fold lower than that of dATP opposite **Pa** ($V_{max}/K_M = 4.3 \times 10^4$ %·min^{-1}·M^{-1}). As a whole, the incorporation efficiency of d**Pn**TP opposite **Ds** was more than 154 times higher than those of the natural base substrates opposite **Ds** (Table 1, entries 12−15). In addition, the incorporation efficiency of d**Ds**TP opposite **Pn** was more than 39 times higher than those of the natural base substrates opposite **Pn** (Table 2, entries 3 and 5−7).

The combination of the **Ds** γ-amidotriphosphate (d**Ds**TP_N) and **Pn** also increased the selectivity of the **Ds**−**Pn** pairing in replication, relative to that of the **Ds**−**Pa** pairing (Tables 1 and 2). Although the incorporation efficiency of d**Ds**TP opposite **Pn** ($V_{max}/K_M = 8.5 \times 10^5$ %·min^{-1}·M^{-1}) was slightly lower than that of d**Ds**TP opposite **Pa** ($V_{max}/K_M = 1.1 \times 10^6$ %·min^{-1}·M^{-1}), the efficiency of d**Ds**TP_N incorporation opposite **Pn** ($V_{max}/K_M = 8.6 \times 10^4$ %·min^{-1}·M^{-1}) was slightly higher than that opposite **Pa** ($V_{max}/K_M = 6.7 \times 10^4$ %·min^{-1}·M^{-1}). Since the **Pn** template efficiently excluded the misincorporation of dATP, the selectivity of **Pn** between d**Ds**TP_N and dATP increased by 16-fold, as compared to that of **Pa** between d**Ds**TP_N and dATP (1.6-fold). This improvement was mainly caused by the decrease in the V_{max} value of the dATP incorporation

(6) (a) Schulhof, J. C.; Molko, D.; Teoule, R. *Nucleic Acids Res.* **1987**, *15*, 397−416. (b) Chaix, C.; Duplaa, A. M.; Molko, D.; Teoule, R. *Nucleic Acids Res.* **1989**, *18*, 7381−7393.
(7) (a) Petruska, J.; Goodman, M. F.; Boosalis, M. S.; Sowers, L. C.; Cheong, C.; Tinoco, I. *Proc. Natl. Acad. Sci. U.S.A.* **1988**, *85*, 6252−6256. (b) Goodman, M. F.; Creighton, S.; Bloom, L. B.; Petruska, J. *Crit. Rev. Biochem. Mol. Biol.* **1993**, *28*, 83−126.
(8) (a) Morales, J. C.; Kool, E. T. *Nat. Struct. Biol.* **1998**, *5*, 950−954. (b) Matray, T. J.; Kool, E. T. *Nature* **1999**, *399*, 704−708. (c) McMinn, D. L.; Ogawa, A. K.; Wu, Y.; Liu, J.; Schultz, P. G.; Romesberg, F. E. *J. Am. Chem. Soc.* **1999**, *121*, 11585−11586. (d) Wu, Y.; Ogawa, A. K.; Berger, M.; McMinn, D. L.; Schultz, P. G.; Romesberg, F. E. *J. Am. Chem. Soc.* **2000**, *122*, 7621−7632. (e) Mitsui, T.; Kimoto, M.; Harada, Y.; Yokoyama, S.; Hirao, I. *J. Am. Chem. Soc.* **2005**, *127*, 8652−8658. (f) Matsuda, S.; Henry, A. A.; Romesberg, F. E. *J. Am. Chem. Soc.* **2006**, *128*, 6369−6375.
(9) Kimoto, M.; Yokoyama, S.; Hirao, I. *Biotechnol. Lett.* **2004**, *26*, 999−1005.

Table 2. Steady-State Kinetic Parameters for Insertion of Single Nucleotides into a Template−Primer Duplex[a]

Primer Temp35N-2	5′-ACTCACTATAGGGAGCTTCT 3′-TATTATGCTGAGTGATATCCCTCGAAGA***N***AGAGCT				
entry	template (*N*)	nucleoside triphosphate	K_M (μM)	V_{max} (% min^{-1})	efficiency (V_{max}/K_M)[c]
1	**Pn**	d**Ds**TP	20 (7)[b]	17 (7)	8.5×10^5
2	**Pn**	d**Ds**TP$_N$	110 (20)	9.5 (2.1)	8.6×10^4
3	**Pn**	dATP	510 (210)	2.8 (1.6)	5.5×10^3
4	**Pn**	dATP$_N$	760 (330)	0.39 (0.14)	5.1×10^2
5	**Pn**	dGTP	220 (30)	4.8 (1.2)	2.2×10^4
6	**Pn**	dCTP	n.d.[e]	n.d.	-
7	**Pn**	dTTP	630 (80)	1.3 (0.4)	2.1×10^3
8	**Pn**	dPnTP	110 (60)	0.74 (0.32)	6.7×10^3
9[d]	**Pa**	d**Ds**TP	26 (12)	28 (5)	1.1×10^6
10[d]	**Pa**	d**Ds**TP$_N$	180 (20)	12 (1)	6.7×10^4
11[d]	**Pa**	dATP	490 (260)	21 (6)	4.3×10^4
12[d]	**Pa**	dATP$_N$	1200 (400)	2.2 (1.3)	1.8×10^3
13[d]	**Pa**	dGTP	480 (140)	0.42 (0.09)	8.8×10^2
14[d]	**Pa**	dCTP	n.d.	n.d.	-
15[d]	**Pa**	dTTP	880 (530)	0.097 (0.025)	1.1×10^2
16[d]	T	dATP	0.81 (0.44)	3.3 (1.8)	4.1×10^6
17[d]	C	dATP	500 (90)	2.3 (0.8)	4.6×10^3
18[d]	C	dGTP	2.3 (0.1)	16 (4)	7.0×10^6
19[d]	T	dGTP	420 (20)	1.2 (0.1)	2.9×10^3

[a] Assays were carried at 37 °C for 1.2−21.8 min, using 5 μM template-primer duplex, 10−50 nM enzyme, and 15−1500 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 four data sets. [b]Standard deviations are given in parentheses. [c]The units of this term are % min^{-1} M^{-1}. [d]These parameters were referred from *Nat. Methods* **2006**, *3*, 729−735. [e]Not determined. Minimal insert products (<2%) were detected after an incubation for 20 min with 1500 μM nucleoside triphosphate and 50 nM enzyme.

opposite **Pn**. In addition, the misincorporation of d**Ds**TP$_N$ opposite **Ds** ($V_{max}/K_M = 9.9 \times 10^3$ %·min^{-1}·M^{-1}) was 37-fold lower than that of d**Pn**TP opposite **Ds** ($V_{max}/K_M = 3.7 \times 10^5$ %·min^{-1}·M^{-1}).

Primer extension involving the **Ds**−**Pn** pair by the exonuclease-proficient Klenow fragment (KF exo$^+$) also exhibited higher selectivity, relative to that involving the **Ds**−**Pa** pair (Figure 3). The elongation of the **Ds**−**Pn** pairing occurred efficiently, and elongation products were obtained (33-mer in Figure 3A, lane 2 and 34−35-mer in Figure 3B, lane 2). The A−**Pn** mispairing (Figures 3A, lane 4 and 3B, lane 5) was effectively prevented, relative to the A−**Pa** mispairing (Figures 3A, lane 10 and 3B, lane 6). In addition, the efficiency of the primer extension with the combination of d**Ds**TP$_N$ and the **Pn** template (Figure 3A, lane 3) was as high as that of d**Ds**TP and the **Pn** template (Figure 3A, lane 2). Although a small amount of **Pn** self-pairing was observed (Figure 3A, lane 6), on the whole, the selectivity of the **Ds**−**Pn** pair in replication was improved.

PCR Amplification. We next performed PCR amplification of a DNA fragment (174-mer) containing one **Ds**−**Pn** pair (see the Supporting Information 3), using Vent DNA polymerase and a substrate mixture of d**Ds**TP$_N$, d**Pn**TP, dATP, dGTP, dCTP, and dTTP. The single-nucleotide insertion experiments using KF exo$^-$ showed that the efficiency and selectivity of the **Ds**−**Pn** pairing were still lower than those of the natural A−T and G−C pairings. In our experiments, the selectivity of the cognate A−T and G−C pairings were 890−11 000-fold higher than that of the noncognate A−C and G−T pairings (Table 1, entries 17−20 and Table 2, entries 16−19). However, we previously demonstrated that the 3′ to 5′ exonuclease activity of Vent DNA polymerase increased the selectivity of the **Ds**−**Pa** pair, for the use of the unnatural base pair in PCR amplification.[3]

The amplification efficiency after 20 cycles of PCR with the DNA fragment containing a single **Ds**−**Pn** pair (Figure 4A, lane 3) was higher than that of the DNA fragment containing a single **Ds**−**Pa** pair using dATP$_N$ instead of dATP (Figure 4A, lane 6). The DNA fragment containing a single **Ds**−**Pn** pair was approximately 500-fold amplified after 20 cycles of PCR. Although the amplification efficiency of the DNA fragment containing the **Ds**−**Pa** using dATP was relatively high (Figure 4A, lanes 8 and 9), the mutation rate of the **Ds**−**Pa** site also increased, as mentioned below (Figure 4F and 4G).

The selectivity of the **Ds**−**Pn** pair in PCR was assessed by sequencing the amplified DNA fragments. As we previously reported, DNA fragments containing **Ds** can be sequenced by a conventional dideoxy-dye-terminator method with or without a modified **Pa** substrate, 4-propynylpyrrole-2-carbaldehyde (d**Pa′**TP).[3] In the sequencing with d**Pa′**TP, only the peak corresponding to the base opposite **Ds** in the templates disappears. In addition, in the sequencing without d**Pa′**TP, the sequencing terminated at the position opposite **Ds** and the read-through peaks were significantly reduced, if the DNA fragment was faithfully amplified at the **Ds**−**Pa** position.[3]

The high selectivity of the **Ds**−**Pn** pair was proven by a comparison of the sequencing patterns of the 20-cycle PCR products (Figure 4D and 4E) with those of the original DNA fragment containing the **Ds**−**Pn** pair (Figure 4B and 4C). By analyzing the read-through peak heights of the sequencing without d**Pa′**TP (Figure 4E), the mutation rate of the **Ds**−**Pn** pair was assessed by comparison to the sequencing patterns of the control DNA fragments (see the Supporting Information 4).[3] Since the control DNA fragments contain 1−10% of the A−T pair in place of the unnatural base pair, in the sequencing without d**Pa′**TP, the read-through peak heights depend on the replacement rate (1−10%) of the A−T pair in the control DNA fragments. The reliability of this method has been confirmed by comparison to the mutation rate determined by transcription using PCR-amplified DNA templates containing the **Ds**−**Pa** pair with the ribonucleoside triphosphate of biotin-linked **Pa**.[3]

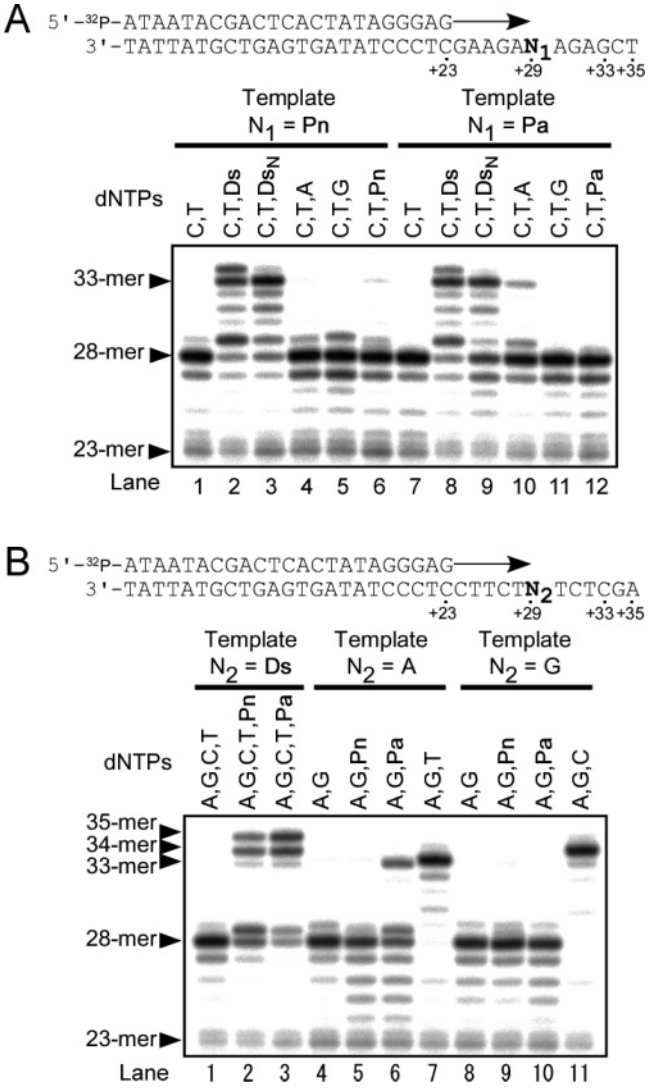


Figure 3. Primer extensions using natural and unnatural nucleotides in the templates containing **Pn** or **Pa** (A) or **Ds**, A, or G (B). The reactions were performed by the 3′ → 5′ exonuclease-proficient Klenow fragment (0.1 unit/μL) with 10 μM substrates and 200 nM templates at 37 °C for 5 min. Ds_N means d**Ds**TP_N.

The total mutation rate of the **Ds**−**Pn** site in the amplified DNA fragment after 20 cycles was ∼1%, when the PCR was performed with the usual triphosphates of the natural and **Pn** bases and the γ-amidotriphosphate of **Ds**. In the PCR amplification of the **Ds**−**Pa** pairing, the total mutation rate of the **Ds**−**Pa** site was 3−4%, even when the PCR was carried out with both the γ-amidotriphosphates of **Ds** and A.[3] When dATP was used instead of $dATP_N$ for PCR amplification, the mutation rate of the **Ds**−**Pa** site further increased by ∼10% (Figure 4F and 4G). Thus, the **Ds**−**Pn** pair was significantly improved for PCR amplification, as compared with the **Ds**−**Pa** pair.

In summary, we have developed the unnatural **Ds**−**Pn** base pair, which exhibits high efficiency and selectivity in *in vitro* replication. Even though the nitro group of **Pn** prevents the mispairing with A, **Pn** efficiently pairs with **Ds**, indicating that the nitro group of **Pn** also functions as a hydrogen acceptor residue for recognition by DNA polymerases. Although the **Pn** nucleoside is unstable under harshly basic conditions, it can withstand the conditions in at least 40 cycles of PCR amplification. We are studying the decomposition mechanism and working to improve the stability of **Pn**. As compared with the **Ds**−**Pa** pair, the PCR amplification involving the **Ds**−**Pn** pair requires no $dATP_N$ and, thus, increases the amplification efficiency. The **Ds**−**Pn** pair therefore promises to be useful in a variety of technologies employing genetic expansion systems.

Materials and Methods

General. Reagents and solvents were purchased from standard suppliers and used without further purification. Reactions were monitored by thin-layer chromatography (TLC) using 0.25 mm silica gel 60 plates impregnated with 254 nm fluorescent indicator (Merck). ^{1}H, ^{13}C, and ^{31}P NMR spectra were recorded on JEOL EX270 and BRUKER (300-AVM) magnetic resonance spectrometers. Nucleoside purification was performed on a Gilson HPLC system with a preparative C18 column (Waters Microbond Sphere, 150 mm × 19 mm). Triphosphate derivatives were purified with a DEAE-Sephadex A-25 column (300 mm × 15 mm) and a C18 column (Synchropak RPP, 250 mm × 4.6 mm, Eichrom Technologies). High-resolution mass spectra (HRMS) and electrospray ionization mass spectra (ESI-MS) were recorded on a JEOL JM 700 mass spectrometer and a Waters micromass ZMD 4000 equipped with a Waters 2690 LC system, respectively. The 2-nitropyrrole (**1**) was synthesized according to the previously reported method.[5]

1-(2-Deoxy-β-D-ribofuranosyl)-2-nitropyrrole (3). To a mixture of **1**[5] (224 mg, 2.0 mmol) and CH_3CN (20 mL) was added NaH (80 mg, 2.0 mmol, 60% dispersion in mineral oil). The resulting mixture was stirred for 30 min at room temperature. To the reaction mixture was added 1-chloro-2-deoxy-3,5-di-*O*-toluoyl-α-D-*erythro*-pentofuranose[10] (855 mg, 2.2 mmol). After stirring for 2 h at room temperature, the reaction mixture was separated by ethyl acetate and water. The organic phase was washed three times with saturated NaCl, dried with Na_2SO_4, and evaporated *in vacuo*. The product was purified by silica gel column chromatography (10% hexane in CH_2Cl_2) to give **2** (722 mg, 78%) as a white foam. To 722 mg of **2** was added methanolic ammonia (50 mL) that was saturated at 0 °C. The solution was stirred for 12 h at room temperature. The solvent was evaporated *in vacuo*. The residue was purified by silica gel column chromatography (2% methanol in CH_2Cl_2) and RP-HPLC to give 291 mg (82%) of compound **3**.

1-[2-Deoxy-5-*O*-(4,4′-dimethoxytrityl)-β-D-ribofuranosyl]-2-nitropyrrole 2-Cyanoethyl-*N*,*N*-diisopropylphosphoramidite (5). A portion (228 mg, 1.0 mmol) of **3** was coevaporated with dry pyridine three times. To the residue in dry pyridine (10 mL) was added 4,4′-dimethoxytrityl chloride (373 mg, 1.1 mmol). The resulting mixture was stirred for 1 h at room temperature. The mixture was separated by ethyl acetate and water. The organic phase was washed three times with 5% $NaHCO_3$ and saturated NaCl, dried with Na_2SO_4, and evaporated *in vacuo*. The product was purified by silica gel column chromatography (10% ethyl acetate in CH_2Cl_2) to give **4** (493 mg, 93%). After evaporation of compound **4** (265 mg, 0.5 mmol) with pyridine three times, the residue was dissolved in THF (2.5 mL). To the mixture was added diisopropylethylamine (131 μL, 0.75 mmol) and 2-cyanoethyl-*N*,*N*-diisopropylamino chlorophosphoramidite (123 μL, 0.55 mmol). The reaction mixture was stirred at room temperature for 1 h. After addition of methanol (50 μL), the mixture was diluted with ethyl acetate/triethylamine (20:1, v/v) and then washed with 5% $NaHCO_3$ and saturated NaCl. The organic phase was dried with Na_2SO_4 and evaporated *in vacuo*. The product was purified by silica gel column chromatography (20% CH_2Cl_2 in hexane containing 2% triethylamine) to give compound **5** (315 mg, 86%).

1-(2-Deoxy-β-D-ribofuranosyl)-2-nitropyrrole 5′-Triphosphate (7, dPnTP). A portion (159 mg, 0.30 mmol) of **4** was coevaporated with dry pyridine three times. The residue was added into dry pyridine (3.0 mL) and was combined with acetic anhydride (57 μL, 0.6 mmol). The

(10) Rolland, V.; Kotera, M.; Lhomme, J. *Synth. Commun.* **1997**, *27*, 3505−3511.

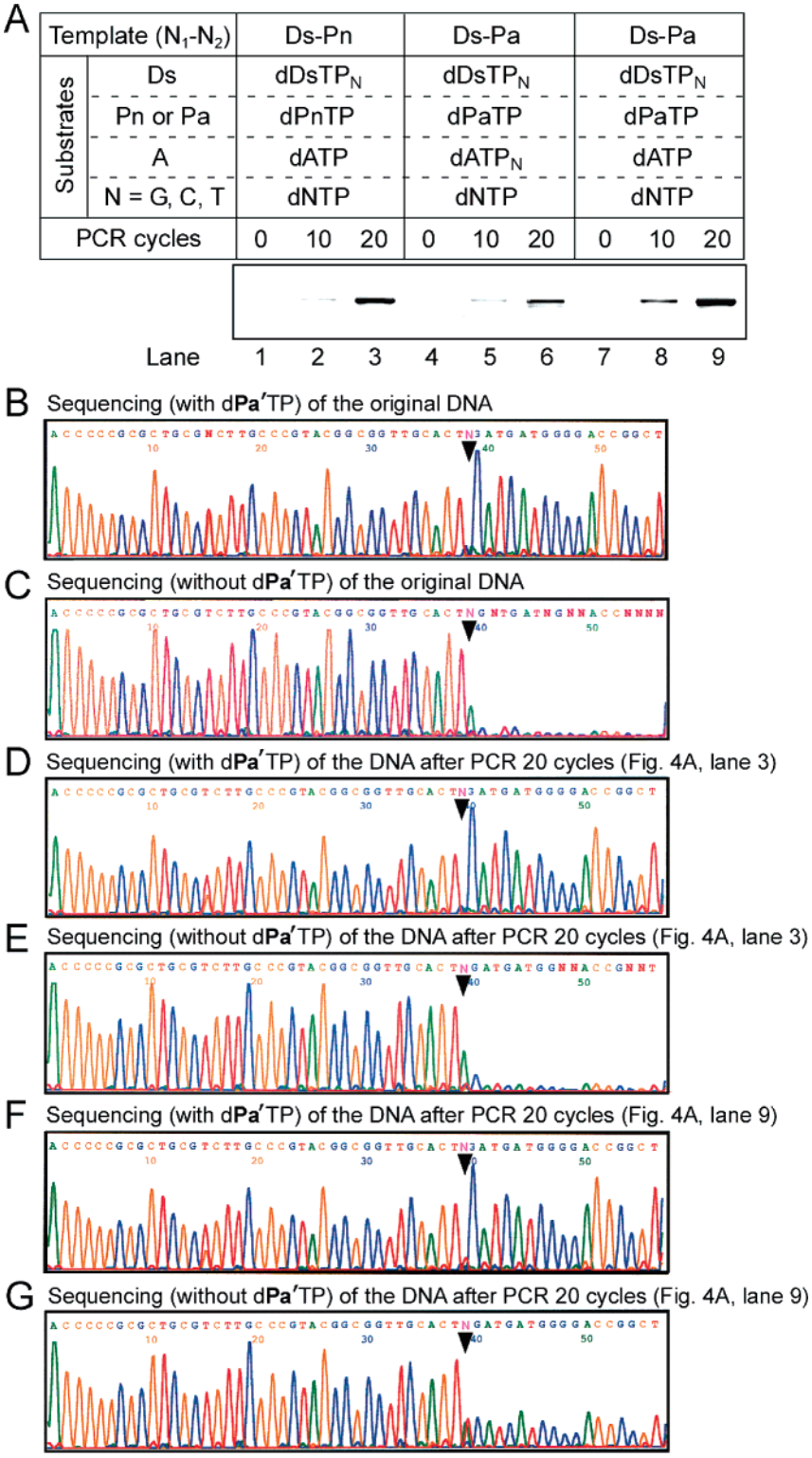


Figure 4. PCR amplification of a DNA fragment (174-mer) containing the **Ds**−**Pn** or **Ds**−**Pa** base pair. (A) Agarose-gel analysis of original fragments and PCR products containing one **Ds**−**Pn** pair with ATP (lanes 1−3), one **Ds**−**Pa** pair with ATP_N (lanes 4−6), or one **Ds**−**Pa** pair with ATP (lanes 7−9). (B−G) Sequencing of the original DNA (B, C), 20-cycle amplified DNA fragments containing the **Ds**−**Pn** pair using the PCR conditions in Figure 4A lane 3 (D, E), and 20-cycle amplified DNA fragments containing the **Ds**−**Pa** pair using the PCR conditions in Figure 4A lane 9 (F, G). The sequencing was performed in the presence (B, D, and F) or absence (C, E, and G) of d**Pa′**TP, using a BigDye Terminator v1.1 Cycle Sequencing kit. The arrow indicates the unnatural base position.

resulting mixture was stirred for 12 h at room temperature. The mixture was poured into 5% $NaHCO_3$ and extracted with ethyl acetate. After drying over Na_2SO_4, the solvent was evaporated *in vacuo*. The residue in dry methylene chloride (30 mL) was mixed with dichloroacetic acid (300 μL) at 0 °C. The resulting mixture was stirred for 15 min at 0 °C. The mixture was poured into 5% $NaHCO_3$ and extracted with methylene chloride. After drying over Na_2SO_4, the solvent was evaporated *in vacuo*. The residue was subjected to silica gel chromatography, using methylene chloride/ethyl acetate (9:1, v/v) as an eluent, to afford 74 mg of **6** in a 91% yield. Compound **6** (41 mg, 0.15 mmol) was dissolved

in pyridine and evaporated to dryness *in vacuo*. The residue was dissolved in pyridine (150 μL) and dioxane (450 μL). A 1 M solution of 2-chloro-4*H*-1,3,2-benzodioxaphosphorin-4-one in dioxane (180 μL, 0.18 mmol) was added.[11] After 10 min, tri-*n*-butylamine (150 μL) and 0.5 M bis(tributylammonium)pyrophosphate in DMF (450 μL) were added to the reaction mixture. The mixture was stirred at room temperature for 10 min. A solution of 1% iodine in pyridine/water (98:2, v/v) (3.0 mL) was then added. After 15 min, a 5% aqueous solution (230 μL) of $NaHSO_3$, followed by water (7.5 mL), was added to the reaction mixture. The solution was stirred at room temperature for 30 min, and then concentrated ammonia (30 mL) was added. Hydrolysis was carried out at room temperature for 2 h. After the reaction mixture was concentrated *in vacuo*, the product was purified by DEAE Sephadex (A-25) column chromatography (eluted by a linear gradient of 50 mM to 1 M TEAB) and then by C18-HPLC (eluted by a gradient of 0% to 30% CH_3CN in 100 mM triethylammonium acetate) to give the nucleoside 5′-triphosphate (**7**, d**Pn**TP) in a 49% yield.

Stability of the Pn Nucleoside. A mixture of the **Pn** nucleoside (5.7 mM) and thymidine (2.0 mM), which was added as an internal standard, in H_2O (100 μL) was treated with 28% NH_4OH (2.0 mL) at room temperature or heated at 55 °C for 1−10 h in a sample vial (Wheaton sample vial, 4 mL, with a white-rubber lined cap). The treated solutions were evaporated *in vacuo*, and the residues were dissolved in H_2O. The solutions were analyzed by RP-HPLC (0−30% CH_3CN in 100 mM TEAA, 10 min at a flow rate of 1 mL/min, detected at 280 nm, Gilson HPLC system, Synchropak RPP, 250 mm × 4.6 mm, Eichrom Technologies). The amounts (%) of the **Pn** nucleoside remaining intact were calculated by normalization to the peak area of the thymidine.

Steady-State Kinetics for the Single-Nucleotide Insertion Experiments with KF exo$^-$. 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 (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 (5−50 nM), the reaction time (1.2−21.8 min), and the gradient concentration of dNTP (6−1500 μM) were adjusted to give reaction extents of 25% or less. The reactions were quenched with 10 μL of a 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).[9] 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) for the various enzyme concentrations used. The kinetic parameters (K_M and V_{max}) were obtained from Hanes−Woolf plots of [dNTP]/v_0 against [dNTP].

Primer Extension Experiments with KF exo$^+$. A primer (23-mer) labeled with ^{32}P at the 5′-end was annealed to a 35-mer template in an annealing buffer (20 mM Tris-HCl (pH 7.5), 14 mM $MgCl_2$, and 0.2 mM DTT), by heating at 95 °C and slow cooling to 4 °C. The duplex solution (400 nM, 5 μL) was mixed with 2 μL of dNTP solution (50 μM each) on ice, and the reactions were started by adding 3 μL of a solution containing the exonuclease-proficient Klenow fragment (1U, TaKaRa, Tokyo) diluted in distilled water. The reaction mixture was incubated at 37 °C for 5 min and was terminated by adding 10 μL of a dye solution (10 M urea and 0.05% BPB in 1 × TBE) and heating at 75 °C for 3 min. The products were analyzed on a 15% polyacrylamide gel containing 7 M urea. The reaction extents were measured with a bio-imaging analyzer (Fuji model BAS2500).

PCR Amplification of the 174-mer Templates. PCR was performed in 20 mM Tris-HCl buffer (pH 8.8), with 10 mM KCl, 10 mM $(NH_4)_2SO_4$, 2 mM $MgSO_4$, 0.1% Triton X-100, 0.3 mM d**Pn**TP or d**Pa**TP, dATP or $dATP_N$, dGTP, dCTP, dTTP, and d**Ds**TP_N, 1 μM each Primer 1 and Primer 2, approximately 0.6 nM double-stranded DNA templates (N_1-N_2 = **Ds**−**Pn** or **Ds**−**Pa**; see Supporting Information 3), and 0.04 unit/μL Vent DNA polymerase (NEB) on a PTC-100 Programmable Thermal Controller (MJ Research). The PCR cycle was as follows: 94 °C, 0.5 min; 45 °C, 0.5 min; 65 °C, 4 min.[3] The products were analyzed on a 4% agarose gel stained with EtBr and were quantified by using a Bio-Rad molecular imager FX. For sequencing analysis, the PCR products were purified by gel electrophoresis (7% polyacrylamide-7 M urea gel) or filtration using Microcon YM-30 (Millipore) and Micropure-EZ (Millipore) filters.

Dye Terminator Sequencing of the 174-mer DNA Fragments. The cycle sequencing reactions (20 μL) were performed on the PTC-100 controller (MJ Research) with the Cycle Sequencing Mix (8 μL) from the BigDye Terminator v1.1 Cycle Sequencing Kit (Applied Biosystems), containing approximately 0.3 pmol of template and 4 pmol of Primer 2, in the presence or absence of 1 nmol of d**Pa′**TP.[3] After 20 cycles of PCR (96 °C, 10 s; 50 °C, 5 s; 60 °C, 4 min), the residual dye terminators were removed from the reaction with CENTRI-SEP columns (Princeton Separations), and the solutions were dried. The residues were resuspended in a formamide solution (4 μL) and were fractionated on the ABI 377 DNA sequencer, using a 6% polyacrylamide−6 M urea gel. The sequence data were analyzed with the Applied Biosystems PRISM sequencing analysis v3.2 software.

Acknowledgment. We thank Akira Sato and Tsuyoshi Fujiwara for synthesizing the nucleoside derivatives. This work was supported by a Grant-in-Aid for Scientific Research (KAKENHI 15350097 (I.H.), 19201046 (I.H.), and 18710197 (M.K.)) from the Ministry of Education, Culture, Sports, Science, and Technology and by the RIKEN Structural Genomics/Proteomics Initiative (RSGI), the National Project on Protein Structural and Functional Analyses, Ministry of Education, Culture, Sports, Science and Technology of Japan.

Supporting Information Available: NMR and MS data for the nucleoside derivatives of **Pn**. Sequences of DNA fragments for the PCR amplification and sequencing experiments. Sequencing patterns of the control DNA fragments containing 1−10% of the A−T pair in place of the unnatural base pair. This material is available free of charge via the Internet at http://pubs.acs.org.

JA073830M

(11) Ludwig, J.; Eckstein, F. *J. Org. Chem.* **1989**, *54*, 631−635.