Pergamon

SCIENCE DIRECT®

Bioorganic & Medicinal Chemistry Letters 13 (2003) 4515–4518

BIOORGANIC & MEDICINAL CHEMISTRY LETTERS

An Unnatural Hydrophobic Base, 4-Propynylpyrrole-2-carbaldehyde, as an Efficient Pairing Partner of 9-Methylimidazo[(4,5)-*b*]pyridine

Tsuneo Mitsui,[a] Michiko Kimoto,[b] Akira Sato,[b] Shigeyuki Yokoyama[b,c,d,*] and Ichiro Hirao[a,b,*]

[a]*Research Center for Advanced Science and Technology, The University of Tokyo, 4-6-1 Komaba, Meguro-ku, Tokyo 153-8904, Japan*

[b]*Protein Synthesis Technology Team, Protein Research Group, RIKEN Genomic Sciences Center (GSC), 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan*

[c]*Department of Biophysics and Biochemistry, Graduate School of Science, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan*

[d]*RIKEN Harima Institute at Spring-8, 1-1-1 Kouto, Mikazuki-cho, Sayo, Hyogo 679-5148, Japan*

Received 7 July 2003; accepted 31 July 2003

Abstract—To develop unnatural base pairs that function in replication, we designed 4-propynylpyrrole-2-carbaldehyde (designated as **Pa**′) and synthesized the nucleoside derivatives of **Pa**′. The base pairing of **Pa**′ with the partner, 9-methylimidazo[(4,5)-*b*]pyridine (**Q**), was compared to that of pyrrole-2-carbaldehyde (**Pa**), which was previously developed as a specific pairing partner of **Q**. The thermal stability of a DNA duplex containing the **Q**–**Pa**′ pair and the incorporation efficiency of the **Pa**′ substrate (d**Pa**′TP) into DNA opposite **Q** by the Klenow fragment of *Escherichia coli* DNA polymerase I were improved, in comparison with those of the **Q**–**Pa** pair. These improvements result from the increased hydrophobicity and stacking stability of **Pa**′ by the introduction of the propynyl group to **Pa**, providing valuable information for the further development of unnatural base pairs toward the expansion of the genetic alphabet.

The creation of unnatural base pairs that can work with the natural A–T and G–C base pairs in the cell expands the genetic alphabet, leading to novel biotechnology. In addition, the unnatural base pairs are useful for understanding the mechanisms of replication, transcription, and translation. Recent advancements in the development of unnatural base pairs have been achieved by the use of hydrophilic[1–4] and hydrophobic[5–8] base analogues. One of the hydrophobic base pairs, 9-methylimidazo[(4,5)-*b*]pyridine (**Q**) and 2,4-difluorotoluene (**F**), was developed as the isostere of the A–T pair, and **Q** and **F** can work in replication as A and T analogues, respectively[6] (Fig. 1a and 1b). In particular, the A–**F** pair functions as well as the A–T and **Q**–**F** pairs. Thus, the **Q**–**F** pair cannot be used as the third base pair for the expansion of the genetic alphabet. However, the hydrophobic base pair revealed the importance of the shape fitting between pairing bases in replication.[9]

To confer exclusive selectivity to the **Q**–**F** pair, we developed 4-propynylpyrrole-2-carbaldehyde (**Pa**′), to replace **F** (Fig. 1c). Previously, we developed pyrrole-2-carbaldehyde (**Pa**) as a pairing partner of **Q**[10] (Fig. 1d). The five-membered ring of **Pa** fits well with **Q**, but not with A, and the recognition of the aldehyde group of **Pa** by polymerases is superior to that of the fluoro group of **F**. However, the stacking stability and the hydrophobicity of **Pa** are inferior to those of **F**, such that a DNA duplex containing the **Q**–**Pa** pair is less stable than that containing the **Q**–**F** pair, and the enzymatic incorporation of d**Pa**TP into DNA opposite **Q** is less effective than that of d**F**TP opposite **Q**.[10] To address these shortcomings, we designed the unnatural base, **Pa**′, in which the propynyl group was introduced at position 4 of **Pa**. Here, we report the synthesis of the nucleoside derivatives of **Pa**′, the thermal stability of a DNA duplex containing the **Q**–**Pa**′ pair, and the enzymatic incorporation of d**Pa**′TP into DNA.

*Corresponding authors. Tel.: +81-45-503-9644; fax: +81-45-503-9645; e-mail: ihirao@postman.riken.go.jp (I. Hirao); tel.: +81-3-5841-4413; fax: +81-3-5841-8057; e-mail: yokoyama@biochem.s.u-tokyo.ac.jp (S. Yokoyama).

0960-894X/$ - see front matter © 2003 Elsevier Ltd. All rights reserved.
doi:10.1016/j.bmcl.2003.09.059

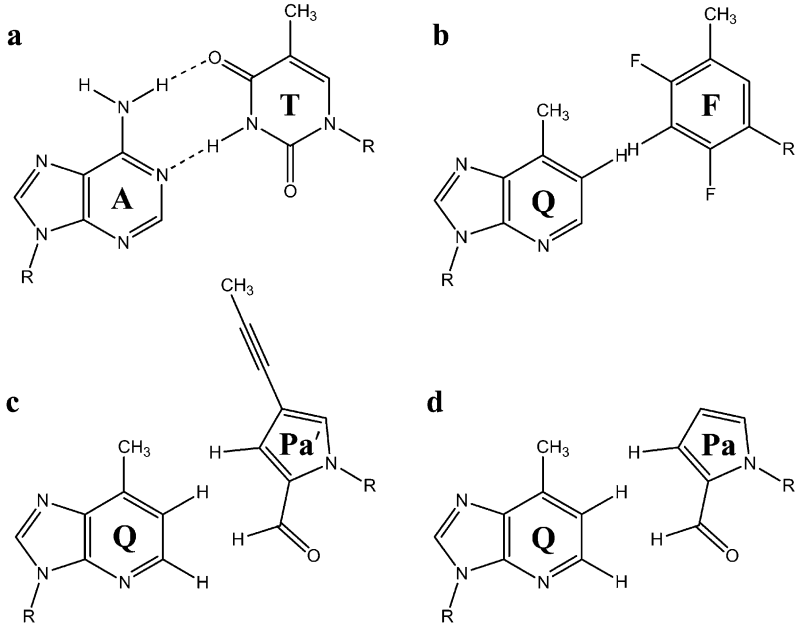
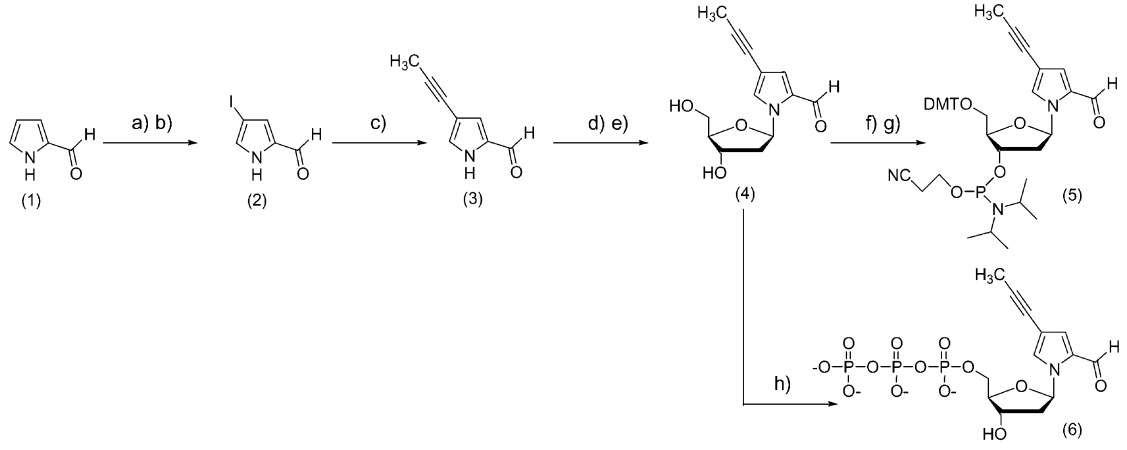


Figure 1. Base-pair structures: the natural A–T pair (a) and the unnatural Q–F (b), Q–Pa′ (c), and Q–Pa (d) pairs. R denotes a ribose moiety.

The nucleoside derivatives of **Pa′** (**4**) were synthesized as shown in Scheme 1. We first synthesized 4-iodopyrrole-2-carbaldehyde (**2**)[11] to introduce the propynyl group at position 4 of pyrrole-2-carbaldehyde (**1**). During the process, the aldehyde group was protected with an iminium salt[12] for the efficient *meta*-directed electrophilic substitution of **1**. Then, the iminium salt was converted to **2** using *N*-iodosuccinimide, followed by a treatment with $NaHCO_3$ (57%: two-step total yield).[12,13] Compound **2** was converted to 4-propynylpyrrole-2-carbaldehyde (**3**) by the palladium-mediated coupling with tributyl(1-propynyl)tin (90%).[13] By the same procedure used for the synthesis of **Pa**,[10] the glycosidation of **Pa′** was carried out using 1-chloro-2-deoxy-3,5-di-*O*-toluoyl-α-D-erythro-pentofuranose and **3**, activated with NaH, followed by deprotection of the toluoyl groups with methanolic ammonia to give **4** with a 36% yield. The nucleoside of **Pa′** (**4**)[14] was characterized by ^1H and ^{13}C NMR spectra and mass spectrometry. The β-configuration of **4** was confirmed by the NOE conectivities between H1′ and H4′ and between H1′ and H2″ (data not shown). The nucleoside of **Pa′** (**4**) was converted to the phosphoramidite (**5**) for the chemical synthesis of DNA fragments, and to the nucleoside triphosphate (**6**, d**Pa′**TP) to function as a substrate in replication. DNA fragments were synthesized using a DNA synthesizer (PE Applied Biosystems), and the coupling efficiency of **5** was more than 98%.

Assessments of the thermal stabilities of DNA duplexes containing **Pa′**, **Pa**, **F**, or **Q**, were performed using the DNA fragments, 5′-CGCAT$\mathbf{N_1}$GTTACC ($\mathbf{N_1}$ = **Pa′**, **Pa**, **F**, or the natural bases) and 5′-GGTAAC$\mathbf{N_2}$ATGCG ($\mathbf{N_2}$ = **Q** or the natural bases), (each 5 mM) in 10 mM sodium phosphate (pH 7.0) containing 100 mM NaCl and 0.1 mM EDTA. The stability of the DNA duplex containing the **Q**–**Pa′** pair (T_m = 42.2 °C) was increased, in comparison to those containing the **Q**–**Pa** pair



Scheme 1. Conditions: (a) pyrrolidinium perchlorate, pyrrolidine in benzene, 92%; (b) *N*-iodosuccinimide in acetonitrile, then $NaHCO_3$, 62%; (c) tributyl(1-propynyl)tin, $Pd(PPh_3)_4$, CuI, in CH_3CN, 90%; (d) NaH, 1-chloro-2-deoxy3,5-di-*O*-toluoyl-α-D-erythropentofuranose in CH_3CN; (e) NH_3 in MeOH, 36% (two-step yield); (f) DMT-Cl in pyridine; (g) $CIP(N\text{-}iPr_2)(OCH_2CH_2CN)$, iPr_2EtN in THF, 82% (two-step yield); (h) $POCl_3$, proton sponge in $PO(OCH_3)_3$ then tri-*n*-butylamine, bis(tributylammonium)pyrophosphate in DMF. Abbreviations: DMT, 4,4′-dimethoxytrityl; *i*Pr, isopropyl.

Table 1. Steady-state kinetic parameters for insertion of single nucleotides into a template-primer duplex by the exonuclease-deficient Klenow fragment[a]

Primer Template	5′-ACTCACTATAGGGAGGAAGA 3′-TATTATGCTGAGTGATATCCCTCCTTCT**N**TCTCGA			
Template (**N**)	Nucleside triphosphate	K_M (μM)	V_{max} (% min^{-1})[d]	Efficiency, V_{max}/K_M (% min^{-1} M^{-1})
Q	Pa′	97 (44)[b]	33 (11)	3.4×10^5
A	Pa′	130 (40)	22 (4)	1.7×10^5
G	Pa′	67 (31)	0.27 (0.11)	4.0×10^3
C	Pa′	110 (30)	0.20 (0.05)	1.8×10^3
T	Pa′	180 (130)	0.37 (0.06)	2.1×10^3
Q[e]	Pa	240 (80)	30 (7)	1.3×10^5
A[e]	Pa	520 (210)	29 (11)	5.6×10^4
G[e]	Pa	460 (160)	0.23 (0.06)	5.0×10^2
C[e]	Pa	n.d.[c]	n.d.[c]	
T[e]	Pa	430 (270)	0.12 (0.05)	2.8×10^2
Q	Q	74 (30)	8.2 (1.2)	1.1×10^5
A	T	1.6 (0.7)	2.1 (1.3)	1.3×10^6

[a]Assays were carried out at 37 °C for 1–20 min using 5 μM template-primer duplex, 10–50 nM enzyme, and 0.6–2100 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.
[b]Standard deviations are given in parentheses.
[c]No inserted products were detected after an incubation for 20 min with 1500 or 2100 μM nucleoside triphosphate and 50 nM enzyme.
[d]The values were normalized to the enzymatic concentration (20 nM) for the various enzyme concentrations used.
[e]Data taken from ref 10.

($T_m = 39.6$ °C) or the **Q**–**F** pair ($T_m = 41.1$ °C). Interestingly, the selectivity of the **Q**–**Pa**′ pair was also improved, relative to the non-cognate A–**Pa**′ and G–**Pa**′ pairs; the stability of the **Q**–**Pa**′ duplex ($T_m = 42.2$ °C) was higher than those of the A–**Pa**′ ($T_m = 35.2$ °C) and G–**Pa**′ ($T_m = 36.2$ °C) duplexes. The stability of **Pa**′ pairing with A or G was not remarkably increased in comparison to that of **Pa** pairing with A or G; the T_m values of the **Q**–**Pa**, A–**Pa**, and G–**Pa** duplexes were 39.6, 34.2, and, 36.0 °C, respectively.[10]

The incorporation efficiency of d**Pa**′TP into DNA was assessed by single-nucleotide insertion experiments, using the exonuclease-deficient Klenow fragment[15,16] (Table 1). Similar to the thermal stability of **Pa**′, the incorporation efficiency of d**Pa**′TP opposite **Q** ($V_{max}/K_M = 3.4\times10^5$% min^{-1} M^{-1}) was increased, in comparison to that of d**Pa**TP opposite **Q** ($V_{max}/K_M = 1.3\times10^5$) or d**F**TP opposite **Q** ($V_{max}/K_M = 2.1\times10^5$). Although the incorporation efficiencies of d**Pa**′TP opposite the natural bases were also increased, the **Q**–**Pa**′ pairing was still more efficient than the **Pa**′ pairing with the natural A, G, C, and T bases ($V_{max}/K_M = 1.7\times10^5$, 4.0×10^3, 1.8×10^3, and 2.1×10^3, respectively). On the other hand, d**Q**TP was incorporated into DNA in a self-complementary manner by the Klenow fragment, and the incorporation efficiency of d**Q**TP opposite **Q** ($V_{max}/K_M = 1.1\times10^5$) was as high as that of d**Pa**TP opposite **Q**. However, this **Q**–**Q** pairing was inferior to the **Q**–**Pa**′ pairing.

By the introduction of the propynyl group to **Pa**, the **Q**–**Pa**′ pair was endowed with efficiency and specificity in replication. The increase in the incorporation efficiencies of d**Pa**′TP by the Klenow fragment results from the decrease in the K_M values (67–180 μM), in comparison to those of d**Pa**TP (240–520 μM, except for opposite C) (Table 1). In contrast, the difference of the V_{max} values between the incorporations of d**Pa**′TP (0.20–33% min^{-1}) and d**Pa**TP (0.13–31% min^{-1}, except for opposite C) is relatively small. These results suggest that the hydrophobicity and stacking stability of the base affect the affinity of the substrate for the open complex of the Klenow fragment, providing insights into the further development of the unnatural base pairs. Application and further improvements of the **Q**–**Pa**′ pair are being investigated.

References and Notes

1. Piccirilli, J. A.; Krauch, T.; Moroney, S. E.; Benner, S. A. *Nature* **1990**, *343*, 33.
2. Bain, J. D.; Switzer, C.; Chamberlin, A. R.; Benner, S. A. *Nature* **1992**, *356*, 537.
3. Otsuki, T.; Kimoto, M.; Ishikawa, M.; Mitsui, T.; Hirao, I.; Yokoyama, S. *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98*, 4992.
4. Hirao, I.; Otsuki, 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.
5. Morales, J. C.; Kool, E. T. *Nat. Struct. Biol.* **1998**, *5*, 950.
6. Morales, J. C.; Kool, E. T. *J. Am. Chem. Soc.* **1999**, *121*, 2323.
7. McMinn, D. L.; Ogawa, A. K.; Wu, Y.; Liu, J.; Schultz, P. G.; Romesberg, F. E. *J. Am. Chem. Soc.* **1999**, *121*, 11585.
8. Ogawa, A. K.; Wu, Y.; McMinn, D. L.; Liu, J.; Schultz, P. G.; Romesberg, F. E. *J. Am. Chem. Soc.* **2000**, *122*, 3274.
9. Kool, E. T. *Annu. Rev. Biophys. Biomol. Struct.* **2001**, *30*, 1.
10. Mitsui, T.; Kitamura, A.; Kimoto, M.; To, T.; Sato, A.; Hirao, I.; Yokoyama, S. *J. Am. Chem. Soc.* **2003**, *125*, 5298.
11. 4-Iodopyrrole-2-carbaldehyde (**2**): ^{1}H NMR (270 MHz, DMSO-d_6): δ 7.12 (m, 1H), 7.34 (m, 1H), 9.42 (d, 1H, $J = 0.8$ Hz), 12.40 (bs, 1H); HRMS (FAB, 3-NBA matrix) calcd for C_5H_5NOI (M + 1) 221.9416, found 221.9428; 4-Propynylpyrrole-2-carbaldehyde (**3**): ^{1}H NMR (270 MHz, $CDCl_3$): δ 2.00 (s, 3H), 6.94 (m, 1H), 7.16 (m, 1H), 9.46 (d, 1H, $J = 1.1$ Hz); HRMS (FAB, 3-NBA matrix) calcd for C_8H_8NO (M + 1) 134.0606, found 134.0617.
12. Sonnet, P. E. *J. Org. Chem.* **1971**, *36*, 1005.
13. Alvarez, A.; Guzman, A.; Ruiz, A.; Velarde, E.; Muchowski, J. M. *J. Org. Chem.* **1992**, *57*, 1653.

14. 1-(2′-Deoxy-β-D-ribofuranosyl)-4-propynylpyrrole-2-carbaldehyde (**4**): ^{1}H NMR (270 MHz, DMSO-d_6): δ 1.97 (s, 3H), 2.10 (m, 1H), 2.27 (m, 1H), 3.55 (m, 2H), 3.80 (m, 1H), 4.23 (m, 1H), 4.99 (t, 1H, J = 5.5 Hz), 5.22 (d, 1H, J = 4.0 Hz), 6.65 (t, 1H, J = 6.3 Hz), 7.09 (d, 1H, J = 1.8 Hz), 7.86 (s, 1H), 9.45 (d, 1H, J = 1.0 Hz); ^{13}C NMR (67 MHz, DMSO-d_6): δ 3.90, 41.95, 61.07, 69.80, 73.21, 85.11, 86.21, 87.53, 105.93, 126.52, 129.99, 130.28, 179.44: HRMS (FAB, 3-NBA matrix) calcd for $C_{13}H_{16}NO_4$ (M + 1) 250.1079, found 250.1040; UV–vis (in EtOH): λ_{max} 311 nm (7.12×10^3), λ_{max} 259 nm (9.03×10^3), λ_{max} 228 nm (15.63×10^3), λ_{min} 282 nm (3.50×10^3), λ_{min} 248 nm (7.03×10^3); TLC: R_f = 0.21 (CH_2Cl_2/MeOH = 20:1, v/v).

15. 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.

16. Goodman, M. F.; Creighton, S.; Bloom, L. B.; Petruska, J. *Crit. Rev. Biochem. Mol. Biol.* **1993**, *28*, 83.