HETEROCYCLES, Vol. 90, No. 1, 2015, pp. 432 - 441. © 2015 The Japan Institute of Heterocyclic Chemistry
Received, 23rd June, 2014, Accepted, 16th July, 2014, Published online, 23rd July, 2014
DOI: 10.3987/COM-14-S(K)41

BASE PAIR RECOGNITION ABILITY OF 2-(METHYLAMINO)PYRIMIDIN-4-YL NUCLEOBASE IN PARALLEL TRIPLEX DNA

Yoshiyuki Hari,* Satoshi Kashima, Yuya Matsuda, Akihiro Sakata, Ryutaro Takamine, Shin Ijitsu, and Satoshi Obika*

Graduate School of Pharmaceutical Sciences, Osaka University, 1-6 Yamadaoka, Suita, Osaka 565-0871, Japan; E-mail: hari@phs.osaka-u.ac.jp and obika@phs.osaka-u.ac.jp

Dedicated to Professor Dr. Isao Kuwajima on the occasion of his 77th birthday

Abstract – A phosphoramidite bearing a 2-(methylamino)pyrimidin-4-yl nucleobase was synthesized and the modified oligonucleotide (triplex-forming oligonucleotide, TFO) was successfully obtained using the phosphoramidite on an automated DNA synthesizer. UV-melting experiment of triplex DNA between the synthesized TFO and duplex DNA targets indicated that the 2-aminopyrimidin-4-yl unit in TFO could be a useful core structure for recognition of a TA base pair in duplex DNA.

INTRODUCTION

Triplex DNA formation is a basic strategy for targeting duplex DNA by an oligonucleotide (triplex-forming oligonucleotide, TFO).[1] The targeting is achieved by Hoogsteen or reverse-Hoogsteen hydrogen bonding between duplex DNA and TFO. However, the duplex DNA sequences capable of forming triplex DNA are limited to homopurine regions without any pyrimidine interruption. For example, in the case of triplex DNA formation via Hoogsteen hydrogen bonds, T and the protonated C in TFO can specifically recognize A and G in the homopurine regions (Figure 1). Therefore, many researchers have tried to develop modified nucleobases that can recognize pyrimidine interruption, namely pyrimidine-purine base pairs such as CG and TA base pairs, within duplex DNA.[2] However, it is still hard to target an arbitrary sequence of duplex DNA with sequence selectivity and strong affinity.

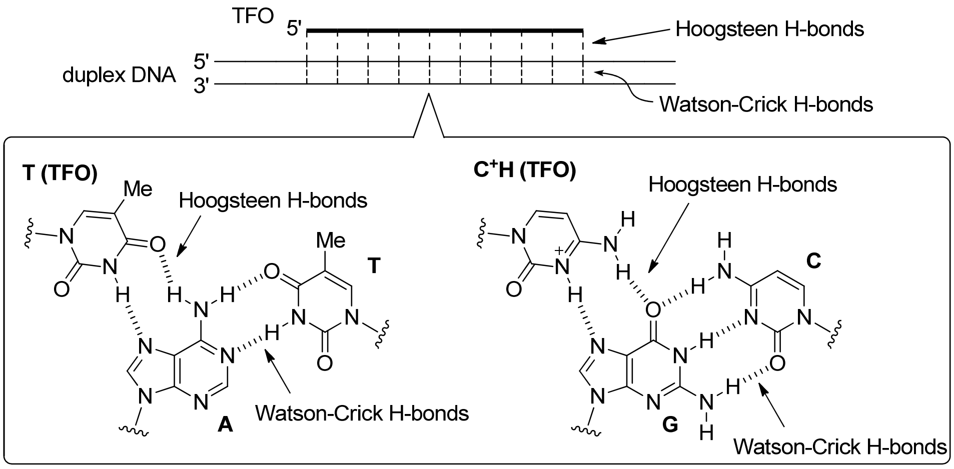


Figure 1. Structures of base triplets of triplex DNA via Hoogsteen hydrogen bonds (H-bonds)

We have also engaged in the development of modified nucleobases recognizing CG and TA base pairs,[3-5] and very recently, a 4-[(3*S*)-3-guanidinopyrrolidino]-5-methylpyrimidin-2-one nucleobase (GP) was developed as a counterpart for the CG base pair.[3] On the other hand, concerning the recognition of the TA base pair, it was found that the 3-hydroxybenzene nucleobase (3H) might recognize T in the TA base pair in such a way that avoids sterically the steric hindrance of the 5-methyl group of T (Figure 2a).[4] However, base pair recognition by 3H was non-selective, probably because the hydroxyl group of 3H acted not only as a hydrogen donor but also as a hydrogen acceptor.[6] Herein, the 2-(methylamino)pyrimidin-4-yl nucleobase (MeAPm) was designed as a new modified nucleobase on the basis of the findings on 3H (Figure 2b). It is considered that the 2-amino group works only as a hydrogen donor, suitable for recogning the 4-carbonyl group of T, and the electron-deficient pyrimidine ring increases the hydrogen-donating ability of the 2-amino group. Moreover, replacing the *N*-methyl group in MeAPm with various substituents would be an effective approach to search for the structure (R in Figure 2b) capable of recognizing A in the TA base pair. We considered comparing the ability of MeAPm with that of 3H in recognizing the TA base pair. In this study, we describe the synthesis of TFO containing MeAPm and the base pair recognition ability of the modified nucleobase in the triplex DNA formed.

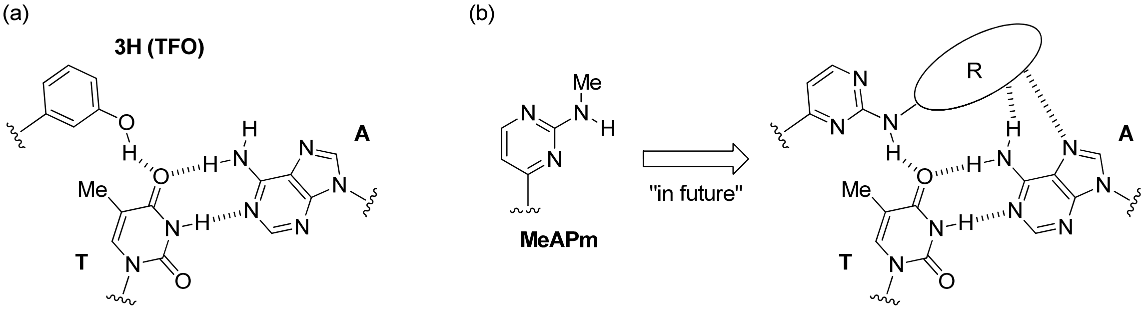
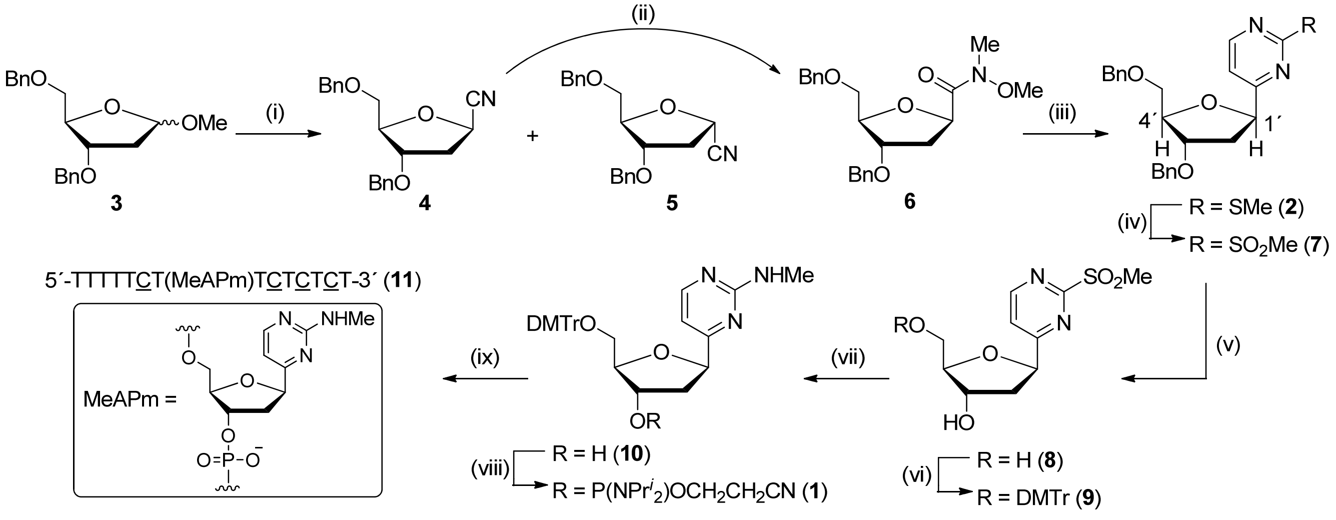


Figure 2. (a) Proposed structure of 3H-TA base triplet and (b) design of nucleobase for the TA base pair

RESULTS AND DISCUSSION

The synthetic route of the desired phosphoramidite **1** bearing 2-(methylamino)pyrimidin-4-yl nucleobase was shown in Scheme 1. Initially, compound **2** bearing the 2-(methylthio)pyrimidin-4-yl nucleobase was synthesized from compound **3**[7] in reference to the letter reported by Baldwin's group.[8] Compound **3** underwent a reaction with TMSCN and TMSOTf to give β-isomer **4** and α-isomer **5** in 25% and 67% yield, respectively. Weinreb amide **6** was obtained by acid treatment of **4** followed by condensation with MeNHOMe. Reaction of **6** with TMSC≡CLi and successive annulation with *S*-methylisothiourea gave desired **2** in 35% yield for two steps. The stereochemistry of **2** at the 1′-position was confirmed to be β-configuration by NOE correlation between 1′-H and 4′-H. After oxidation of **2** into sulfone **7** in quantitative yield, debenzylation of **7** by BCl_3 resulted in diol **8**, the dimethoxytritylation of which led to compound **9** in 93% yield. Substitution of the methylsulfonyl moiety in **9** with a methylamino group yielded **10**, the phosphitylation of which resulted in the desired **1**, a suitable building block for oligonucleotide synthesis. On an automated DNA synthesizer using a common phosphoramidite chemistry for oligonucleotide synthesis, the desired TFO **11** containing the 2-(methylamino)pyrimidin-4-yl nucleobase (MeAPm) was successfully synthesized.



Scheme 1. Reagents and conditions: (i) TMSCN, TMSOTf, CH_2Cl_2, rt, 25% (**4**) and 67% (**5**); (ii) 35% HCl aq., 1,4-dioxane, reflux, then MeNHOMe•HCl, EDC, DMAP, CH_2Cl_2, rt, 66% for two steps; (iii) TMSC≡CLi, THF, -78 °C, then (*S*-methylisothiourea)$_2$•H_2SO_4, K_2CO_3, MeCN-H_2O, reflux, 35% for two steps; (iv) *m*CPBA, CH_2Cl_2, rt, quant.; (v) BCl_3, CH_2Cl_2-hexane, rt, 57%; (vi) DMTrCl, pyridine, rt, 93%; (vii) $MeNH_2$, THF-H_2O, rt, 93%; (viii) *i*-$Pr_2NP(Cl)OCH_2CH_2CN$, *i*-Pr_2NEt, CH_2Cl_2, rt, 76%; (ix) oligonucleotide synthesis.

Next, the base pair recognition ability of MeAPm in the synthesized TFO was measured by UV-melting experiment and was compared with that of 3H. The results are summarized in Table 1. The affinity of

MeAPm to the TA base pair was same as that of 3H, and the T_{m} value of the triplexes (YZ = TA) was 26 °C. These results indicate that like 3H, MeAPm can form a hydrogen bond with T in a TA base pair in duplex DNA. On the contrary, MeAPm led to a significant decrease in the stability of triplex DNA with other duplex DNA targets (YZ = CG, AT and GC) compared with that shown by 3H. As a result, it was found that the MeAPm in TFO was able to selectively recognize a TA base pair within duplex DNA although further improvement of the affinity to the TA base pair is required.[9] The selective TA base pair recognition by MeAPm occurred as MeAPm behaved only as a hydrogen donor by using the N-H of the methylamino group, as we expected. In addition, it was shown that the methyl group in MeAPm would not prevent the hydrogen bond formation between the N-H in MeAPm and the 4-carbonyl group of T in the TA base pair. This observation strongly suggests that the recognition of A in the TA base pair via a suitable substituent in place of the methyl group of MeAPm might be feasible in accordance with our plan shown in Figure 2b.

Table 1. T_{m} values (°C) of triplexes between TFOs and hairpin-loop duplex DNA[a]

TFO 5'-TTTTTCT**X**TCTCTCT-3'
hairpin-loop duplex DNA 5'-GGCAAAAAGA**Y**AGAGAGACGC—C18spacer
3'-CCGTTTTTCT**Z**TCTCTCTGCG—

X	YZ			
	TA	CG	AT	GC
3H[b]	26	27	28	25
MeAPm	26	22	19	21

[a] Condition : 10 mM sodium cacodylate buffer (pH 6.8) containing 100 mM KCl, 50 mM $MgCl_2$ and 1.89 μM of each oligonucleotide. C : 2′-Deoxy-5-methylcytidine. [b] The data are shown in ref 4.

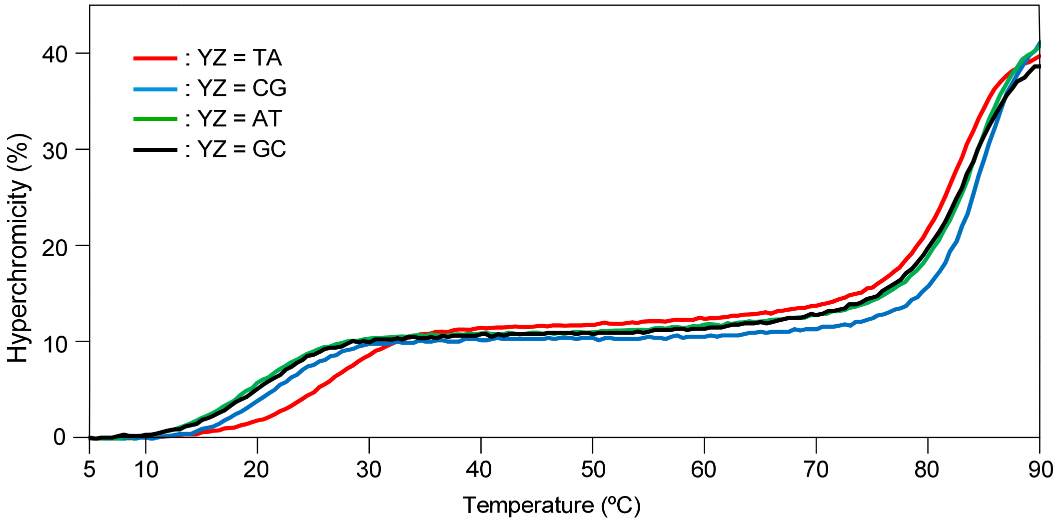


Figure 3. UV-melting curves of triplex DNA between TFO **11** (X = MeAPm) and duplex DNA targets

In conclusion, we designed a new 2-(methylamino)pyrimidin-4-yl nucleobase (MeAPm) for the recognition of the TA base pair in triplex DNA, and we succeeded in the synthesis of TFO containing MeAPm. UV-melting experiments of triplex DNA demonstrated that MeAPm could recognize a TA base pair in a sequence-selective fashion. Therefore, the 2-aminopyrimidin-4-yl structure is promising as a core unit for TA base pair recognition. In future, we plan to explore a variety of substituents on the 2-amino group to develop a nucleobase that can strongly and selectively recognize the TA base pair through hydrogen bonding with both nucleobases.

EXPERIMENTAL

All chemicals were purchased from chemical suppliers. For column chromatography, Fuji Silysia silica gel PSQ-60B and FL-60D were used. ^{1}H NMR, ^{13}C NMR, and ^{31}P NMR spectra were recorded on JEOL AL300, JEOL ECS400 and JEOL GX-500 spectrometers. Optical rotations were recorded on a JASCO P-2200 polarimeter. IR spectra were recorded on a JASCO FT/IR-4200 spectrometer. Mass spectra were obtained using a JEOL JMS-700, JEOL JMS-S3000, and Bruker Daltonics Autoflex II TOF/TOF mass spectrometer.

1β-Cyano-3,5-di-*O*-benzyl-1,2-dideoxy-D-ribofuranose (4) and 1α-cyano-3,5-di-*O*-benzyl-1,2-dideoxy-D-ribofuranose (5)

Under a nitrogen atmosphere, TMSOTf (8.10 mL, 29.0 mmol) in anhydrous CH_2Cl_2 (200 mL) was added to a solution of 2-deoxy-3,5-di-*O*-benzyl-D-ribofuranose **3**[7] (6.34 g, 19.3 mmol) and TMSCN (4.70 mL, 58.0 mmol) in anhydrous CH_2Cl_2 (200 mL), and the mixture was stirred at room temperature for 2 h. After addition of sat. $NaHCO_3$ aq., the mixture was extracted with CH_2Cl_2. The extracts were washed with water and brine, dried over Na_2SO_4, and concentrated under reduced pressure. The residue was purified by flash silica gel column chromatography (*n*-hexane/AcOEt = 5:1) to give compounds **4** (1.57 g, 25%) and **5** (4.17 g, 67%), respectively. Compound **4**: A colorless oil. $[\alpha]_D$ 30.1 (*c* 0.65, $CHCl_3$, 29 ºC). IR ν_{max} (KBr) 3031, 2925, 2856, 2245, 1455, 1098 cm^{-1}. ^{1}H NMR (400 MHz, $CDCl_3$) δ 2.40–2.43 (2H, m), 3.50 (1H, dd, *J* = 5.0, 10.5 Hz), 3.57 (1H, dd, *J* = 4.0, 10.5 Hz), 4.18–4.22 (1H, m), 4.23–4.26 (1H, m), 4.48 (1H, d, *J* = 12.0 Hz), 4.49 (1H, d, *J* = 12.0 Hz), 4.53 (1H, d, *J* = 12.0 Hz), 4.58 (1H, d, *J* = 12.0 Hz), 4.78 (1H, dd, *J* = 7.5, 8.0 Hz), 7.25–7.37 (10H, m). ^{13}C NMR (101 MHz, $CDCl_3$) δ 37.6, 65.7, 70.1, 71.4, 73.6, 80.0, 84.7, 118.8, 127.7, 127.7, 127.8, 128.0, 128.4, 128.5, 137.3, 137.7. HRMS (MALDI) *m/z* calcd for $C_{20}H_{21}NNaO_3$ $[M+Na]^+$: 346.1414; found, 346.1414. Compound **5**: A colorless oil. $[\alpha]_D$ 64.7 (*c* 1.13, $CHCl_3$, 29 ºC). IR ν_{max} (KBr) 3031, 2864, 2239, 1454, 1092 cm^{-1}. ^{1}H NMR (400 MHz, $CDCl_3$) δ 2.38–2.41 (2H, m), 3.51 (2H, d, *J* = 4.0 Hz), 4.15–4.18 (1H, m), 4.36–4.39 (1H, m), 4.49 (1H, d, *J* = 12.0 Hz), 4.50 (1H, d, *J* = 12.0 Hz), 4.51 (1H, d, *J* = 12.0 Hz), 4.58 (1H, d, *J* = 12.0 Hz), 4.87 (1H, dd, *J* = 4.0, 4.0 Hz), 7.25–7.38 (10H, m). ^{13}C NMR (101

MHz, $CDCl_3$) δ 37.2, 66.1, 69.8, 71.3, 73.5, 79.3, 84.8, 119.0, 127.6, 127.6, 127.8, 128.4, 128.5, 137.4, 137.6. HRMS (MALDI) *m/z* calcd for $C_{20}H_{21}NNaO_3$ $[M+Na]^+$: 346.1414; found, 346.1411.

3,5-Di-*O*-benzyl-1,2-dideoxy-1β-(*N*-methoxy-*N*-methylcarbamoyl)-D-ribofuranose (6)

A solution of compound **4** (1.10 g, 3.40 mmol) in 1,4-dioxane (60 mL) and 35% HCl aq. (6 mL) was refluxed for 4 h. After evaporation, the residue was extracted with CH_2Cl_2. The extracts were washed with brine, dried over Na_2SO_4, and concentrated under reduced pressure. The residue (1.2 g) was dissolved in anhydrous CH_2Cl_2 (40 mL); MeNHOMe•HCl (497 mg, 5.10 mmol), EDC•HCl (977 mg, 5.10 mmol), and DMAP (20 mg, 0.16 mmol) were added at room temperature. After being stirred at room temperature for 6 h, water was added and the mixture was extracted with AcOEt. The extracts were washed with water and brine, dried over Na_2SO_4, and concentrated under reduced pressure. The residue was purified by flash silica gel column chromatography ($CHCl_3$/MeOH = 50:1) to give compound **6** (869 mg, 66% from **4**). A colorless oil. $[\alpha]_D$ 10.3 (*c* 1.12, $CHCl_3$, 29 ºC). IR ν_{max} (KBr) 1661, 1454 cm^{-1}. ^1H NMR (400 MHz, $CDCl_3$) δ 2.23–2.25 (2H, m), 3.20 (3H, s), 3.45 (1H, dd, *J* = 7.0, 10.0 Hz), 3.64 (1H, dd, *J* = 5.5, 10.0 Hz), 3.71 (3H, s), 4.13–4.16 (1H, m), 4.26–4.31 (1H, m), 4.52 (1H, d, *J* = 12.0 Hz), 4.54–4.58 (3H, m), 4.98–5.02 (1H, m), 7.25–7.35 (10H, m). ^{13}C NMR (126 MHz, $CDCl_3$) δ 32.2, 34.8, 61.7, 70.7, 71.1, 73.3, 74.7, 80.8, 83.7, 127.5, 127.6, 127.6, 128.3, 128.4, 137.9, 138.1, 172.0. HRMS (MALDI) *m/z* calcd for $C_{22}H_{28}NO_5$ $[M+H]^+$: 386.1962; found, 386.1958.

3,5-Di-*O*-benzyl-1,2-dideoxy-1β-(2-methylthiopyrimidin-4-yl)-D-ribofuranose (2)

Under a nitrogen atmosphere, trimethylsilylacetylene (0.78 mL, 5.64 mmol) and *n*-BuLi (1.6 M in *n*-hexane, 3.52 mL, 5.64 mmol) were added to a solution of compound **6** (1.45 g, 3.76 mmol) in anhydrous THF (60 mL) at -78 ºC, and the mixture was stirred at -78 ºC for 2 h. After addition of sat. NH_4Cl aq., the mixture was extracted with AcOEt. The extracts were washed with water and brine, dried over Na_2SO_4, and concentrated under reduced pressure. The residue (1.6 g) was dissolved in MeOH (31 mL) and MeCN (31 mL); K_2CO_3 (2.07 g, 15.0 mmol) and (*S*-methylisothiourea)$_2$•H_2SO_4 (1.04 g, 3.74 mmol) were added. After being refluxed for 17 h, the mixture was extracted with AcOEt. The extracts were washed with water and brine, dried over Na_2SO_4, and concentrated under reduced pressure. The residue was purified by flash silica gel column chromatography (*n*-hexane/AcOEt = 10:1) to give compound **2** (569 mg, 35% from **6**). A colorless oil. $[\alpha]_D$ 77.5 (*c* 1.60, $CHCl_3$, 28 ºC). IR ν_{max} (KBr) 3028, 2862, 1567, 1355, 1094 cm^{-1}. ^1H NMR (300 MHz, $CDCl_3$) δ 2.00 (1H, ddd, *J* = 6.0, 10.0, 13.0 Hz), 2.53 (3H, s), 2.53–2.60 (1H, m), 3.56 (1H, dd, *J* = 5.0, 11.0 Hz), 3.61 (1H, dd, *J* = 5.0, 11.0 Hz), 4.13 (1H, ddd, *J* = 2.0, 2.5, 6.0 Hz), 4.34 (1H, ddd, *J* = 2.5, 5.0, 5.0 Hz), 4.53 (1H, d, *J* = 12.0 Hz), 4.56 (2H, s), 4.57 (1H, d, *J* = 12.0 Hz), 5.12 (1H, dd, *J* = 6.0, 10.0 Hz), 7.18 (1H, d, *J* = 5.0 Hz), 7.27–7.34 (10H, m), 8.42 (1H, d, *J* = 5.0 Hz). ^{13}C NMR (76 MHz, $CDCl_3$) δ 14.1, 38.8, 70.8, 71.1, 73.4, 80.1, 80.5, 84.3, 112.3, 127.6, 127.6, 127.7, 127.7, 128.4, 128.4, 137.9, 137.9, 157.6, 171.0. HRMS (MALDI) *m/z* calcd for $C_{24}H_{27}N_2O_3S$ $[M+H]^+$: 423.1737; found, 423.1736.

3,5-Di-*O*-benzyl-1,2-dideoxy-1β-(2-methylsulfonylpyrimidin-4-yl)-D-ribofuranose (7)

75% *m*CPBA (1.19 g, 5.17 mmol) was added to a solution of compound **2** (950 mg, 2.25 mmol) in CH_2Cl_2 (20 mL), and the mixture was stirred at room temperature for 21 h. After addition of sat. $Na_2S_2O_3$ aq. and sat. $NaHCO_3$ aq., the mixture was extracted with AcOEt. The extracts were washed with sat. $Na_2S_2O_3$ aq., sat. $NaHCO_3$ aq., water and brine, dried over Na_2SO_4, and concentrated under reduced pressure. The residue was purified by flash silica gel column chromatography (n-hexane/AcOEt = 3:1) to give compound **7** (1.02 g, quant.). An orange oil. $[\alpha]_D$ 62.0 (*c* 0.92, $CHCl_3$, 23 ºC). IR ν_{max} (KBr) 1579, 1318, 1135, 1096 cm^{-1}. ^1H NMR (400 MHz, $CDCl_3$) δ 2.02 (1H, ddd, *J* = 6.0, 9.5, 13.0 Hz), 2.68 (1H, ddd, *J* = 2.5, 6.5. 13.0 Hz), 3.34 (3H, s), 3.61 (1H, dd, *J* = 4.0, 10.0 Hz), 3.64 (1H, dd, *J* = 4.5, 10.0 Hz), 4.17 (1H, ddd, *J* = 2.5, 3.0, 6.0 Hz), 4.37 (1H, ddd, *J* = 3.0, 4.0, 4.5 Hz), 4.51 (1H, d, *J* = 12.0 Hz), 4.54 (2H, s), 4.59 (1H, d, *J* = 12.0 Hz), 5.30 (1H, dd, *J* = 6.5, 9.5 Hz), 7.26–7.38 (10H, m), 7.89 (1H, d, *J* = 5.5 Hz), 8.74 (1H, d, *J* = 5.5 Hz). ^{13}C NMR (101 MHz, $CDCl_3$) δ 39.0, 39.1, 70.5, 71.2, 73.4, 79.8, 80.3, 84.6, 120.0, 127.6, 127.8, 128.4, 128.4, 137.6, 137.7, 158.6, 165.2, 173.9. HRMS (FAB) *m/z* calcd for $C_{24}H_{27}N_2O_5S$ $[M+H]^+$: 455.1635; Found 455.1647.

1,2-Dideoxy-1β-(2-methylsulfonylpyrimidin-4-yl)-D-ribofuranose (8)

Under a nitrogen atmosphere, BCl_3 (1.0 M in *n*-hexane, 0.31 mL, 0.31 mmol) was added to a solution of compound **7** (65 mg, 0.14 mmol) in anhydrous CH_2Cl_2 (2 mL) at -78 ºC and the mixture was stirred at room temperature for 1 h. After addition of $NaHCO_3$ powder (35 mg) and anhydrous MeOH (2 mL), the mixture was concentrated under reduced pressure. The residue was purified by flash silica gel column chromatography ($CHCl_3$/MeOH = 15:1) to give compound **8** (23 mg, 57%). A colorless oil. $[\alpha]_D$ 66.4 (*c* 1.06, MeOH, 28 ºC). IR ν_{max} (KBr) 3287, 2928, 1583, 1308, 1134 cm^{-1}. ^1H NMR (400 MHz, CD_3OD) δ 2.08 (1H, ddd, *J* = 5.5, 9.5, 13.5 Hz), 2.44 (1H, ddd, *J* = 2.5, 6.5, 13.5 Hz), 3.62 (3H, s), 3.66 (1H, dd, *J* = 5.0, 12.0 Hz), 3.68 (1H, dd, *J* = 5.0, 12.0 Hz), 4.05 (1H, ddd, *J* = 2.5, 5.0, 5.0 Hz), 4.33 (1H, ddd, *J* = 2.5, 2.5, 5.5 Hz), 5.26 (1H, dd, *J* = 6.5, 9.5 Hz), 7.95 (1H, d, *J* = 5.5 Hz), 8.95 (1H, d, *J* = 5.5 Hz). ^{13}C NMR (101 MHz, CD_3OD) δ 39.4, 42.9, 63.7, 73.8, 80.7, 89.9, 121.5, 160.5, 166.6, 175.2. HRMS (MALDI) *m/z* calcd for $C_{10}H_{14}N_2NaO_5S$ $[M+Na]^+$: 297.0516; found, 297.0515.

1,2-Dideoxy-5-*O*-(4,4′-dimethoxytrityl)-1β-(2-methylsulfonylpyrimidin-4-yl)-D-ribofuranose (9)

Under a nitrogen atmosphere, DMTrCl (272 mg, 0.804 mmol) was added to a solution of compound **8** (210 mg, 0.766 mmol) in anhydrous pyridine (8 mL) at room temperature and the mixture was stirred for 14 h. After addition of sat. $NaHCO_3$ aq., the mixture was extracted with AcOEt. The extracts were washed with water and brine, dried over Na_2SO_4, and concentrated under reduced pressure. The residue was purified by flash silica gel column chromatography (*n*-hexane/AcOEt = 1:2 to 1:3) to give compound **9** (411 mg, 93%). A white powder. $[\alpha]_D$ 41.5 (*c* 1.06, $CHCl_3$, 29 ºC). IR ν_{max} (KBr) 3490, 2931, 1607, 1579, 1506, 1318, 1254 cm^{-1}. ^{1}H NMR (300 MHz, $CDCl_3$) δ 1.81 (1H, s), 2.18 (1H, ddd, *J* = 6.0, 8.5, 13.0 Hz), 2.54 (1H, ddd, *J* = 4.0, 6.5, 13.0 Hz), 3.32–3.36 (5H, m), 3.79 (6H, s), 4.17 (1H, ddd, *J* = 4.0, 4.0, 4.0 Hz), 4.43 (1H, ddd, *J* = 4.0, 4.0, 6.0 Hz), 5.32 (1H, dd, *J* = 6.5, 8.5 Hz), 6.82 (4H, d, *J* = 8.0 Hz), 7.20–7.31 (7H, m), 7.40 (2H, d, *J* = 8.0 Hz), 7.79 (1H, d, *J* = 5.0 Hz), 8.77 (1H, d, *J* = 5.0 Hz). ^{13}C NMR (76 MHz, $CDCl_3$) δ 39.1, 41.8, 55.2, 63.7, 73.4, 79.2, 86.4, 86.7, 113.1, 120.0, 126.9, 127.9, 128.1, 130.0, 135.6, 144.5, 158.5, 158.8, 165.3, 173.9. HRMS (FAB) *m/z* calcd for $C_{31}H_{32}N_2NaO_7S$ $[M+Na]^+$: 599.1822; Found 599.1841.

1,2-Dideoxy-5-*O*-(4,4′-dimethoxytrityl)-1β-[2-(methylamino)pyrimidin-4-yl]-D-ribofuranose (10)

40% MeNH aq. (2.0 mL) was added to a solution of compound **9** (220 mg, 0.44 mmol) in THF (6 mL) at room temperature and the mixture was stirred for 20 min. After addition of AcOEt and water, the mixture was extracted with AcOEt. The extracts were washed with water and brine, dried over Na_2SO_4, and concentrated under reduced pressure. The residue was purified by flash silica gel column chromatography (*n*-hexane/AcOEt = 1:1) to give compound **10** (216 mg, 93%). A pale yellow powder. $[\alpha]_D$ 39.6 (*c* 1.24, $CHCl_3$, 27 ºC). IR ν_{max} (KBr) 3395, 3289, 2933, 1583, 1508, 1252 cm^{-1}. ^{1}H NMR (300 MHz, $CDCl_3$) δ 2.15 (1H, ddd, *J* = 6.0, 9.0, 13.0 Hz), 2.34 (1H, ddd, *J* = 3.0, 6.5, 13.0 Hz), 2.79 (1H, d, *J* = 5.0 Hz), 2.94 (3H, d, *J* = 5.0 Hz), 3.25 (1H, dd, *J* = 4.5, 10.0 Hz), 3.32 (1H, dd, *J* = 4.5, 10.0 Hz), 3.78 (6H, s), 4.07-4.10 (1H, m), 4.35-4.39 (1H, m), 5.00 (1H, dd, *J* = 6.5, 9.0 Hz), 5.08 (1H, q, *J* = 5.0 Hz), 6.72 (1H, d, *J* = 5.0 Hz), 6.81 (4H, d, *J* = 9.0 Hz), 7.20–7.45 (9H, m), 8.21 (1H, d, *J* = 5.0 Hz). ^{13}C NMR (126 MHz, $CDCl_3$) δ 28.3, 41.2, 55.2, 64.2, 73.9, 79.8, 86.2, 106.3, 113.1, 126.8, 127.8, 128.1, 130.0, 135.9, 144.7, 158.4, 158.5, 162.6, 171.3. HRMS (FAB) *m/z* calcd for $C_{31}H_{33}N_3NaO_5$ $[M+Na]^+$: 550.2312; Found 550.2328.

3-*O*-[2-Cyanoethoxy(diisopropylamino)phosphino]-1,2-dideoxy-5-*O*-(4,4′-dimethoxytrityl)-1β-[2-(methylamino)pyrimidin-4-yl]-D-ribofuranose (1)

Under a nitrogen atmosphere, *i*-$Pr_2NP(Cl)OCH_2CH_2CN$ (60 μL, 0.27 mmol) was added to a solution of compound **10** (101 mg, 0.19 mmol) and *i*-Pr_2NEt (0.10 mL, 0.58 mmol) in anhydrous CH_2Cl_2 (2 mL) at 0 ºC and the mixture was stirred at room temperature for 1 h. After addition of sat. $NaHCO_3$ aq., the mixture was extracted with AcOEt. The extracts were washed with water and brine, dried over Na_2SO_4, and concentrated under reduced pressure. The residue was purified by flash silica gel column chromatography (*n*-hexane/AcOEt = 1:2) to give compound **1** (105 mg, 76%) as a white solid. ^{1}H NMR (300 MHz, $CDCl_3$) δ 1.08 (3H, d, *J* = 7.0 Hz), 1.15–1.31 (9H, m), 2.07–2.18 (1H, m), 2.38–2.52 (2H, m), 2.59–2.63 (1H, m), 2.96 (1H, d, *J* = 5.0 Hz), 3.20–3.34 (1H, m), 3.50–3.92 (4H, m), 3.78 (6H, s), 4.23–4.26 (1H, m), 4.43–4.48 (1H, m), 4.97–5.01 (2H, m), 6.75–6.83 (5H, m), 7.19–7.46 (9H, m), 8.23 (1H, d, *J* = 5.0 Hz). ^{31}P NMR (122 MHz, $CDCl_3$) δ 147.8, 148.2. HRMS (FAB) m/z calcd for $C_{40}H_{50}N_5NaO_6P$ [M+Na]+: 750.3391; found 750.3395.

Synthesis of TFO 11

The synthesis of **11** was performed on a 0.2-μmol scale on an automated DNA synthesizer (Gene Design nS-8) using the common phosphoramidite protocol. TFO synthesized on DMTr-ON mode was cleaved from the CPG resin and all the protecting groups on TFO were removed by treatment with 28% NH_3 aq. at room temperature for 4 h. The obtained crude TFO was purified on Sep-Pak® Plus C18 cartridges (Waters) followed by reversed-phase HPLC (Waters XBridge® MS C18 2.5 μm, 10 mm × 50 mm). The composition of the TFO was confirmed by MALDI-TOF-MS analysis. MALDI-TOF-MS data ($[M–H]^-$) for **11**: Found 4480.47 (calcd 4480.01).

UV-melting experiments

UV-melting experiments were carried out on SHIMADZU UV-1650 and SHIMADZU UV-1800 spectrometers equipped with T_m analysis accessory. Equimolecular amounts of the target hairpin-loop duplex DNA and TFO were dissolved in 10 mM sodium cacodylate buffer (pH 6.8) containing 100 mM KCl and 50 mM $MgCl_2$ to give a final concentration of 1.89 μM for each strand. The samples were annealed by heating at 100 ºC followed by slow cooling to 5 ºC. The melting profiles were recorded at 260 nm from 5 ºC to 90 ºC at a scan rate of 0.5 ºC/min. The two-point average method was used to obtain the Tm values and the final values were determined by averaging three independent measurements, which were accurate to within 1 ºC.

ACKNOWLEDGEMENTS

This work was supported in part by a Grant-in-Aid for Scientific Research on Innovative Areas (No. 22136006) from the Ministry of Education, Culture, Sports, Science and Technology, Japan, and Mochida Memorial Foundation for Medical and Pharmaceutical Research.

REFERENCES AND NOTES

1. Reviews: D. Praseuth, A. L. Guieysse, and C. Hélène, *Biochim. Biophys. Acta Gene Struct. Expression*, 1999, **1489**, 181; M. P. Knauert and P. M. Glazer, *Hum. Mol. Genet.*, 2001, **10**, 2243; M. Duca, P. Vekhoff, K. Oussedik, L. Halby, and P. B. Arimondo, *Nucleic Acids Res.*, 2008, **36**, 5123.
2. Reviews: S. O. Doronina and J.-P. Behr, *Chem. Soc. Rev.*, 1997, **26**, 63; D. M. Gowers and K. R. Fox, *Nucleic Acids Res.*, 1999, **27**, 1569; M. G. M. Purwanto and K. Weisz, *Curr. Org. Chem.*, 2003, **7**, 427; V. Malnuit, M. Duca, and R. Benhida, *Org. Biomol. Chem.*, 2011, **9**, 326; Y. Hari, S. Obika, and T. Imanishi, *Eur. J. Org. Chem.*, 2012, 2875; Y. Hari, *Yakugaku Zasshi*, 2013, **133**, 1201.
3. Y. Hari, M. Akabane, and S. Obika, *Chem. Commun*, 2013, **49**, 7421.
4. Y. Hari, S. Kashima, H. Inohara, S. Ijitsu, T. Imanishi, and S. Obika, *Tetrahedron*, 2013, **69**, 6381.
5. Recent papers. Y. Hari, M. Nakahara, J. Pang, M. Akabane, T. Kuboyama, and S. Obika, *Bioorg. Med. Chem.*, 2011, **19**, 1162; Y. Hari, M. Akabane, Y. Hatanaka, M. Nakahara, and S. Obika, *Chem.*

Commun, 2011, **47**, 4424; Y. Hari, M. Nakahara, S. Ijitsu, and S. Obika, *Bioorg. Med. Chem.*, 2013, **21**, 5583; Y. Hari, M. Nakahara, S. Ijitsu, and S. Obika, *Heterocycles*, 2014, **88**, 377.

6. For example, the oxygen of the hydroxyl group in 3H as a hydrogen acceptor could recognize the N-H of C in the CG base pair.
7. W. Wierenga and H. I. Skulnick, *Carbohydr. Res.*, 1981, **90**, 41.
8. R. M. Adlington, J. E. Baldwin, G. J. Pritchard, and K. C. Spencer, *Tetrahedron Lett.*, 2000, **41**, 575.
9. The T_m value of the corresponding full-match triplex (X-YZ = T-AT) was 44 ºC, and the affinity of MeAPm to the TA base pair was significantly lower than that of T to the AT base pair. This is probably because MeAPm interacts with the TA base pair via a single hydrogen bond in contrast with interaction of T to the AT base pair via two hydrogen bonds. Therefore, to improve the recognition ability to the TA base pair, we made the plan shown in Figure 2b.