Organic & Biomolecular Chemistry

RSC Publishing

PAPER

View Article Online
View Journal | View Issue

Cite this: *Org. Biomol. Chem.*, 2013, **11**, 3918

N-(Guanidinoethyl)-2′-deoxy-5-methylisocytidine exhibits selective recognition of a CG interrupting site for the formation of anti-parallel triplexes†

Hidenori Okamura, Yosuke Taniguchi* and Shigeki Sasaki*

Received 7th March 2013,
Accepted 16th April 2013

DOI: 10.1039/c3ob40472b

www.rsc.org/obc

The development of novel nucleoside analogues for the formation of triplex DNA containing pyrimidine–purine inversion sites has been a challenging field. In this paper, we describe the design and synthesis of non-natural nucleoside analogues, *N*-substituted-2′-deoxy-5-methylisocytidine derivatives, and their evaluation for triplex formation. It has been shown that *N*-(guanidinoethyl)-2′-deoxy-5-methylisocytidine exhibits selective recognition of a CG interrupting site and potentiates the formation of anti-parallel triplexes.

Introduction

Sequence specific recognition of duplex DNA by forming triplex DNA with oligonucleotides (TFOs) is a powerful tool for gene targeting agents and gene diagnosis.[1] There are two types of triplexes: one is a parallel triplex formed with homopyrimidine TFO and the other is an anti-parallel triplex formed with homopurine TFO. The TFO forms Hoogsteen base pairs in the parallel triplex or reverse Hoogsteen base pairs in the anti-parallel triplex.[2] In both cases, a pyrimidine insertion in the homopurine strand of the duplex DNA prevents stable triplex DNA, as a Hoogsteen or reverse Hoogsteen base pair is not formed.

A family of non-natural nucleoside analogues has been proposed to overcome this problem, especially for parallel-type triplexes.[3] In contrast, only a few non-natural nucleosides have been designed for anti-parallel triplexes.[4] Previously, we developed W-shaped nucleoside analogues (WNA analogues) and demonstrated WNA-βT and WNA-βC for the recognition of a TA or a CG interrupting site, respectively.[5] Subsequent studies synthesised a variety of WNA-βT analogues to examine the sequence dependence of TA base pair recognition,[6] and WNA-βT was applied to oncogene-targeting TFOs to achieve antiproliferative effects in tumour cells.[7] However, recognition of a CG base pair remains elusive,[8] and thus the development of non-natural nucleoside analogues for a CG interrupting site is still an active area of research.

Graduate School of Pharmaceutical Sciences, Kyushu University, Fukuoka, Higashiku Maidashi 3-1-1, Japan. E-mail: taniguch@phar.kyushu-u.ac.jp, sasaki@phar.kyushu-u.ac.jp; Fax: +81 92 642 6615; Tel: +81 92 642 6615

†Electronic supplementary information (ESI) available: HPLC, MALDI-TOF mass and NMR data. See DOI: 10.1039/c3ob40472b

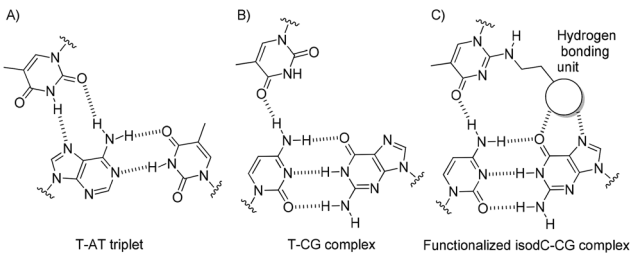
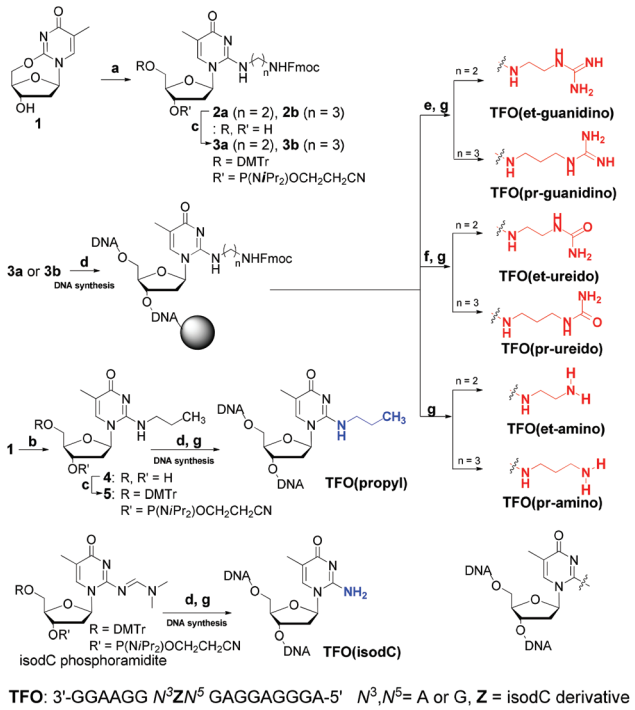


Fig. 1 The schematic structure of triplex DNA. (A) T–AT triplet, (B) T–CG complex, (C) *N*-substituted-isodC–CG complex.

A thymidine forms a reverse Hoogsteen base pair with an adenine base of an AT base pair to stabilize anti-parallel triplexes (Fig. 1A). An NMR study showed that a single hydrogen bond formed between the 4-carbonyl of the thymine and the 4-amino of the cytosine of a CG base pair resulted in a relative stabilization of the anti-parallel triplex (Fig. 1B).[9] Thus, we attempted to design new nucleoside analogues for recognition of a CG base pair based on a T–CG complex. A 2′-deoxy-5-methyl isocytidine (isodC) unit was adopted to enhance the selectivity for CG sites over AT sites. An additional hydrogen bonding site was designed to attach to the 2-amino of isodC through an alkyl linker (Fig. 1C). It was expected that hydrogen bond formation at a Hoogsteen face of a guanine base of a CG base pair would provide selective stabilization for a CG pair. In this paper, we describe the synthesis of a variety of *N*-substituted-isodC derivatives and the evaluation of triplex formation using TFOs containing these derivatives.

Results and discussion

The synthetic scheme for TFOs with *N*-substituted-isodC derivatives is shown in Scheme 1. A cyclic derivative of



Scheme 1 Synthesis of the isodC unit and the TFOs. *Reagents and conditions*: (a) (1) ethylenediamine (or propylenediamine), 1 h (80% or 94%), (2) FmocCl, dioxane, 10% Na_2CO_3 solution, 2.5 h (78% for **2a** and 62% for **2b**). (b) Propylamine, 15 h (97%). (c) (1) DMTrCl, pyridine, 1.5–2 h (65–97%), (2) $iPr_2NP(Cl)OCH_2CH_2CN$, DIPEA, CH_2Cl_2, 0 °C, 1–2 h (67–97%). (d) DNA synthesis. (e) 1*H*-Pyrazole-1-carboxamidine hydrochloride, DBU, CH_3CN. (f) TMSNCO, DBU, CH_3CN, 55 °C. (g) (1) 28% NH_3 solution, 55 °C, (2) HPLC purification.

thymine (**1**)[10] was treated with ethylenediamine or propylenediamine to introduce the linker unit, and the resulting primary amino group was Fmoc protected to produce 5-methyl-*N*-substituted-isodC skeleton **2a** having an ethylene linker and **2b** having a propylene linker. To obtain the control compound (**4**) without the terminal amino unit, compound **1** was reacted with aminopropane. The diol compounds (**2a**, **2b** and **4**) were then converted to the corresponding phosphoramidite precursors (**3a**, **3b** and **5**, respectively). The amino-protected isodC derivative was incorporated into **TFO** with flanking nucleotides around the isodC using an automated DNA synthesizer (Scheme 1).

A functional group was introduced to the amino group of the isodC of the synthesized oligodeoxynucleotides (ODNs) before cleaving from the CPG column. First, the Fmoc group was selectively deprotected using DBU in CH_3CN, and the resulting primary amino group was converted to the corresponding guanidino and ureido groups by treatment with 1*H*-pyrazole-1-carboxamidine[11] and TMSNCO, respectively. The modified ODNs were then cleaved from the CPG resin and the protecting groups of the nucleotides were removed in 28% NH_4OH at 55 °C. The crude ODNs were purified using HPLC. The main peak around 16 min was separated, and the structures of the purified ODNs were determined by MALDI-TOF MS measurement (Fig. S1 and S2, Table S1, ESI†). Each TFO was named according to its modification: **TFO**(guanidino), **TFO**(ureido) and **TFO**(amino). Consequently, we successfully demonstrated the efficient synthesis of modified ODNs by applying the reaction onto the synthesized ODN attached to the CPG column. **TFO** bearing isodC, propyl-isodC and natural thymidine were synthesized for comparison.

The target duplex was designed for the formation of anti-parallel triplex with TFO to have four base pairs, XY = GC, CG, AT and TA. In addition, flanking base pairs around XY of duplex were complementary to N^3 and N^5 of TFO. Triplex stability of parallel type triplexes is generally evaluated by UV melting experiments. However, the thermal denaturing profile of anti-parallel triplexes is complicated and often troublesome for the determination of T_m values.[8b] Therefore triplex stabilities of the synthesized TFOs were evaluated by gel shift analysis in this study. The TFOs were incubated with a duplex with the pyrimidine strand labelled with fluorescent FAM in the presence of 20 mM Tris-HCl, 20 mM $MgCl_2$, 2.5 mM spermidine and 10% sucrose buffer at pH 7.5 for 12 h at 30 °C, and the formed triplexes were confirmed to be low-mobility bands within the non-denatured polyacrylamide gel. Examples of gel shift analysis using **TFO**(et-guanidino) and **TFO**(propyl) with N^3ZN^5 = AZG are shown in Fig. 2. The triplexes were observed as low mobility bands in the gel in the case of **TFO**(et-guanidino) (Fig. 2A), and **TFO**(propyl) having a propylene linker unit did not show remarkable triplex formation (Fig. 2B). The equilibrium association constants (K_s) were calculated from the band intensities of the triplex and duplex DNA[7] (Table 1). The natural **TFO**(T) formed stable triplexes selectively with AT base pairs (Table 1, runs 1, 10, 16 and 22). These results were in accord with the previous results reported for the formation of an anti-parallel triplex including a T–AT base triplet (Fig. 1A).[2] Triplexes were also formed with a CG pair, although the K_s values were smaller, indicating the possible formation of a single hydrogen bond (Fig. 1B).[12,13] Additionally, relatively stable triplexes were formed with isodC at GC sites (Table 1, runs 2, 11 and 23). These data suggested the formation of a single hydrogen bond between isodC and dG such as that shown in the triplet structure in Fig. 3A. It should be noted that **TFO**(et-guanidino) showed binding preference for a CG base pair (Table 1, runs 3 and 12). As the 2-amino group of isodC was alkylated, the hydrogen bond expected to form between isodC and dG would probably not be formed for guanidinoethyl-isodC (Fig. 3B). In contrast, ureido-, amino- or propyl-isodC did not significantly improve the stability of triplex formation (Table 1, runs 4–6 and 13–15). To check the effects of the length of the linker unit on the triplex stability, K_s values of the TFO incorporating isodC having a propylene-linked guanidino, ureido and amino group in the sequence of the 3′A–Z–G-5′ context were compared with those obtained with the corresponding ethylene-linked derivatives. The K_s values indicated that the propylene linker caused a decrease in stability and selectivity (Table 1, runs 3 and 4 for an ethylene linker *vs.* 7 and 8 for a propylene linker). These results indicate that the ethylene linker is suitable for the guanidino group to interact with the target CG site. The amino group did not

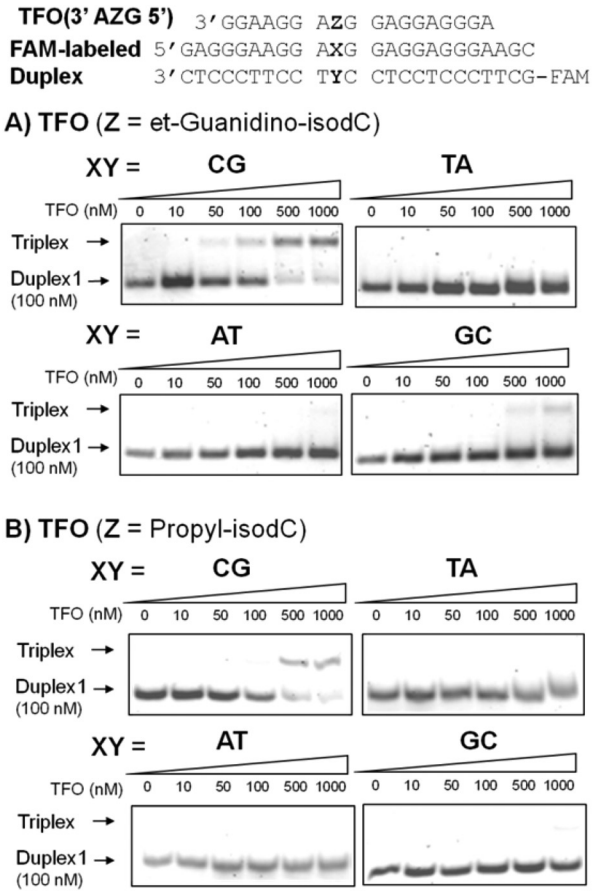


Fig. 2 Gel-shift analysis of the triplex formation ability. Triplex was formed using 100 nM duplex containing a FAM-labelled pyrimidine strand in the presence of TFO at various concentrations. A mixture was incubated for 12 h at 30 °C in 20 mM Tris-HCl, 20 mM $MgCl_2$, 2.5 mM spermidine and 10% sucrose buffer at pH 7.5. Electrophoresis was performed at 10 °C with a non-denatured polyacrylamide gel. (A) **TFO**(et-guanidino-isodC) consisting of 3′-AZG-5′, (B) **TFO**(Propyl-isodC) consisting of 3′-AZG-5′.

stabilize triplexes either in an ethylene linker or in a propylene linker (Table 1, runs 5 and 9), although its protonated form was expected to form a hydrogen bond with a guanine of a CG pair. These results have suggested the formation of hydrogen bonds between the guanidino group at the end of the ethylene linker and the target CG pair. The optimized structure between guanidinoethyl-isodC and a CG base pair and its schematic illustration are shown in Fig. 4, and indicate that the guanidino group can interact with dG from the major groove to contribute to the selectivity for recognition of a CG base pair in an anti-parallel triplex formation. These results indicated that the CG recognition ability of guanidino-isodC was potentiated by the hydrogen bonding ability of the terminal guanidino group.

In this study, TFOs with A–Z–A or G–Z–A in the N^3ZN^5 context did not form stable triplexes, except in the case when Z was the natural base T (Table 1). A similar sequence dependence in anti-parallel triplex formation with TFO having A–Z–A or G–Z–A contexts was reported, although an explanation was not immediately apparent.[4a,6,8c]

Table 1 Equilibrium association constants of triplexes (K_s, 10^6, M^{-1})[a]

			Target base pair (XY)			
Run	N^3ZN^5	Z	GC	CG	AT	TA
1	A–Z–G	T	2.30	8.00	**29.2**	0.98
2		isodC	**10.9**	8.40	<0.01	0.56
3		et-Guanidino	0.28	**5.16**	<0.01	<0.01
4		et-Ureido	0.13	3.70	<0.01	<0.01
5		et-Amino	0.04	0.23	<0.01	0.11
6		Propyl	<0.01	1.84	<0.01	<0.01
7		pr-Guanidino	1.94	0.80	<0.01	<0.01
8		pr-Ureido	<0.01	2.68	<0.01	<0.01
9		pr-Amino	1.97	0.91	<0.01	<0.01
10	G–Z–G	T	0.23	6.78	**95.1**	0.38
11		isodC	**15.2**	3.92	5.30	7.45
12		et-Guanidino	1.55	**12.4**	0.54	1.36
13		et-Ureido	1.55	3.95	0.54	1.54
14		et-Amino	1.79	2.96	1.06	1.59
15		Propyl	0.14	1.16	1.19	1.19
16	A–Z–A	T	0.13	0.47	**32.6**	<0.01
17		isodC	0.45	<0.01	<0.01	<0.01
18		et-Guanidino	<0.01	0.07	<0.01	<0.01
19		et-Ureido	<0.01	<0.01	<0.01	<0.01
20		et-Amino	<0.01	<0.01	<0.01	<0.01
21		Propyl	<0.01	<0.01	<0.01	<0.01
22	G–Z–A	T	6.92	6.73	**39.2**	0.48
23		isodC	**11.1**	<0.01	<0.01	<0.01
24		et-Guanidino	0.20	0.60	<0.01	<0.01
25		et-Ureido	2.21	<0.01	<0.01	<0.01
26		et-Amino	0.67	0.74	<0.01	<0.01
27		Propyl	1.18	1.50	<0.01	<0.01

[a] The K_s values were calculated using the following equation: K_s = [Triplex]/[Duplex][TFO]. Errors were estimated to be within 10% for triplicate experiments.

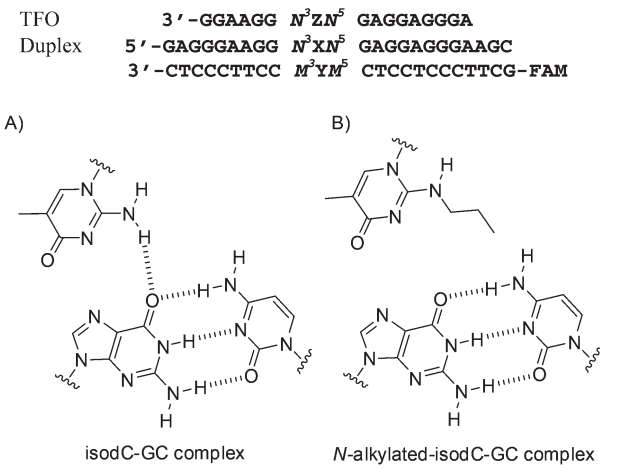


Fig. 3 Speculated recognition structures. (A) isodC–GC complex and (B) *N*-alkylated-isodC–GC complex.

To corroborate the recognition ability of **TFO**(et-guanidino) for a CG base pair in comparison with the selectivity of **TFO**(T), DNase I foot-printing analysis was performed using long duplex DNA including the AT and CG target sites in one sequence (Fig. 5). The bands corresponding to both the AT target site and the CG site disappeared in the presence of

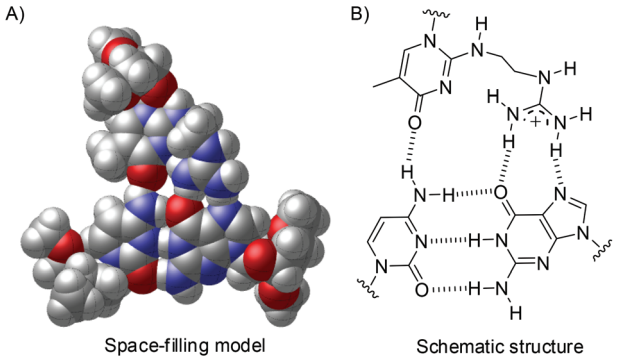


Fig. 4 Optimized structure of the *N*-guanidinoethyl-isodC–CG complex using Gaussian09 B3LYP/6-31G* *in vacuo*. (A) CPK model. (B) Schematic structure.

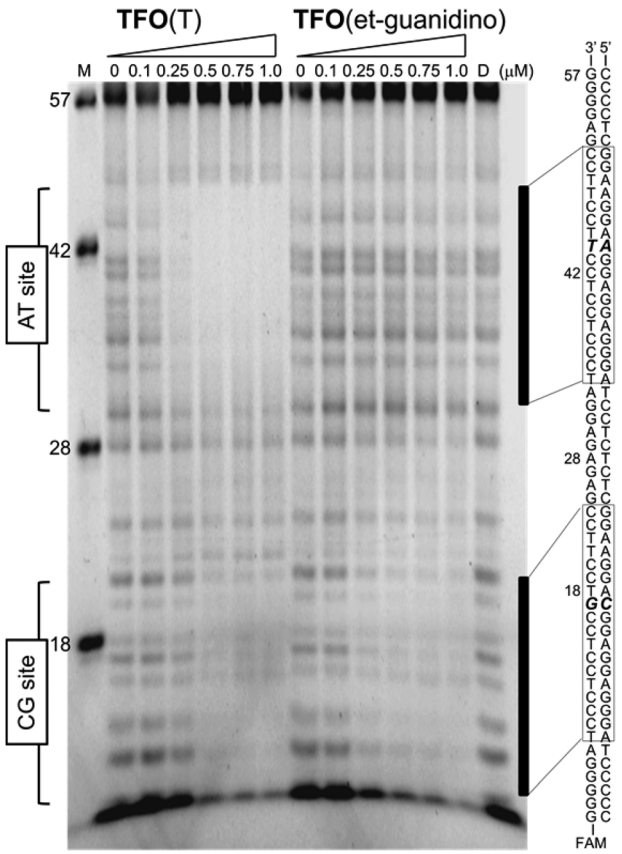


Fig. 5 DNase I foot-printing analysis of the selectivity of **TFO**(T) and **TFO**-(et-guanidino). M is a marker of cleavage sites. D is an absence of TFOs. Black bars show the triplex formation sites including the target AT or CG site.

0.25–1.0 μM **TFO**(T), indicating that **TFO**(T) formed triplexes with these regions and protected them from enzymatic digestion. In the presence of **TFO**(et-guanidino), the bands corresponding to the AT site were present and only bands for the CG site disappeared. These results indicated that **TFO**(et-guanidino) interacted selectively with the CG target site.

Conclusions

In conclusion, we have demonstrated that the newly designed *N*-guanidinoethyl-2′-deoxy-5-methyl-isodC is a selective, non-natural base analogue for a CG base pair and potentiates the formation of anti-parallel triplexes. This isodC has not previously been used as a platform to design a non-natural nucleoside analogue for the formation of triplexes; thus, the present study has revealed that isodC is a potential unit for the formation of anti-parallel triplexes.

Experimental

General

The ^{1}H-NMR (400 MHz) and ^{13}C-NMR (100, 125 MHz) spectra were recorded by a VarianUNITY-400 and an INOVA-500 spectrometer. ^{31}P-NMR (162 MHz) spectrum was recorded using 10% phosphoric acid in D_2O as the internal standard at 0 ppm. The IR spectra were obtained using a Perkin Elmer FTIR-SpectrumOne. The high-resolution mass spectra were recorded by an Applied Biosystems Mariner System 5299 spectrometer. The MALDI-TOF mass spectra were recorded by a BRUKER DALTONICS Microflex.

***N*^{2}-[2-[(9-Fluorenylmethyloxycarbonyl)amino]ethyl]-2′-deoxy-5-methylisocytidine (2a).** 2,5′-Anhydrothymidine (500 mg, 2.2 mmol) in 20 ml of ethylenediamine was stirred at room temperature for 1 h. The solution was concentrated, and the residue was purified by column chromatography ($CHCl_3$–MeOH–NH_3 aq. = 10 : 5 : 1) to give a white foam (80%, 509 mg). To a solution of this white foam (300 mg, 1.1 mmol) in 7 ml of dioxane and 4 ml of a 10% aqueous Na_2CO_3 solution was added 9-fluorenylmethyl chloroformate (545 mg, 2.1 mmol), and then the mixture was stirred at room temperature for 2.5 h. The reaction mixture was concentrated, and the residue was purified by column chromatography ($CHCl_3$–MeOH = 10 : 1 to 6 : 1) to give **2a** as a white powder (78%, 415 mg). Mp = 135–137 °C; n_{max} (neat) 3296, 1698, 1663 cm^{-1}; ^{1}H NMR (400 MHz, CD_3OD) δ 7.77 (2H, d, J = 7.6 Hz), 7.60 (2H, d, J = 7.6 Hz), 7.53 (1H, s), 7.36 (2H, t, J = 7.4 Hz), 7.28 (2H, t, J = 7.4 Hz), 5.78–5.74 (1H, m), 4.42–4.41 (1H, m), 4.35–4.26 (2H, m), 4.16 (1H, t, J = 7.0 Hz), 3.90 (1H, q, J = 2.8 Hz), 3.74 (2H, ddd, J = 2.4, 12.0, 20.0 Hz), 3.54–3.44 (2H, m), 3.34–3.29 (2H, m), 2.47–2.40 (1H, m), 2.18–2.13 (1H, m), 1.86 (3H, s); ^{13}C NMR (125 MHz, CD_3OD) δ 174.6, 159.2, 154.7, 145.4, 145.3, 142.6, 138.8, 128.7, 128.1, 126.2, 126.1, 120.9, 114.7, 92.0, 89.2, 71.8, 67.8, 64.3, 62.0, 42.2, 41.4, 39.8, 13.5; HRMS (ESI-TOF) Calcd for $C_{27}H_{31}N_4O_6$: $[M + H]^+$, 507.2238; Found: 507.2212.

***N*^{2}-[2-[(9-Fluorenylmethyloxycarbonyl)amino]propyl]-2′-deoxy-5-methylisocytidine (2b).** 2,5′-Anhydrothymidine (500 mg, 2.2 mmol) in 20 ml of propylenediamine was stirred at room temperature for 12 h. The solution was concentrated, and the residue was purified by column chromatography ($CHCl_3$–MeOH–NH_3 aq. = 10 : 5 : 1) to give a white foam (94%, 626 mg). To a solution of this white foam (400 mg, 1.3 mmol) in 9 ml of dioxane and 4.5 ml of a 10% aqueous Na_2CO_3 solution was

added 9-fluorenylmethyl chloroformate (746 mg, 2.9 mmol), and then the mixture was stirred at room temperature for 2.5 h. The reaction mixture was concentrated, and the residue was purified by column chromatography ($CHCl_3$–MeOH = 10 : 1 to 6 : 1) to give **2b** as a white powder (62%, 430 mg). Mp = 132–134 °C; n_{max} (neat) 3317, 2923, 1702, 1663, 1603 cm^{-1}; ^1H NMR (400 MHz, CD_3OD) δ 7.78 (2H, d, J = 8.0 Hz), 7.64 (2H, d, J = 8.0 Hz), 7.53 (1H, s), 7.37 (2H, t, J = 7.2 Hz), 7.29 (2H, t, J = 7.2 Hz), 5.81–5.78 (1H, m), 4.47–4.44 (1H, m), 4.35 (2H, d, J = 6.8), 4.19 (1H, t, J = 6.8 Hz), 3.94 (2H, t, J = 2.8 Hz), 3.79 (2H, ddd, J = 2.4, 12.0, 19.6 Hz), 3.15 (2H, t, J = 6.8 Hz), 2.51–2.44 (1H, m), 2.24–2.19 (1H, m), 1.88 (3H, s), 1.75 (2H, t, J = 6.8 Hz); ^{13}C NMR (100 MHz, CD_3OD) δ 174.6, 159.0, 154.5, 145.3, 142.6, 138.9, 128.7, 128.1, 126.1, 120.9, 114.5, 92.0, 89.1, 71.8, 67.6, 62.0, 39.8, 39.3, 30.3, 13.6; HRMS (ESI-TOF) Calcd for $C_{28}H_{33}N_4O_6$: $[M + H]^+$: 521.2395, Found: 521.2437.

3′-*O*-[2-Cyanoethoxy(diisopropylamino)phosphino]-5′-*O*-(4,4′-dimethoxytrityl)-N^2-(2-((9-fluorenylmethyloxycarbonyl)amino)-ethyl) 2′-deoxy-5-methylisocytidine (3a). 4,4′-Dimethoxytrityl chloride (300 mg, 0.9 mmol) was added to a solution of **2** (300 mg, 0.6 mmol) in 6 ml of pyridine, and the mixture was stirred at room temperature for 1.5 h. The solution was diluted with CH_2Cl_2 and successively washed with a saturated aqueous $NaHCO_3$ solution and brine, dried over Na_2SO_4 and concentrated. The residue was purified by flash chromatography ($CHCl_3$–MeOH = 1 : 0 to 20 : 1) to give a white powder (84%, 402 mg). To a solution of this white powder (330 mg, 0.4 mmol) in 8.2 ml of CH_2Cl_2 were added DIPEA (430 µl, 2.5 mmol) and 2-cyanoethyl-*N*,*N*-diisopropylchlorophosphoramidite (275 µl, 1.2 mmol), and then the mixture was stirred at 0 °C for 1 h. The solution was diluted with CH_2Cl_2 and successively washed with a saturated aqueous $NaHCO_3$ solution and brine, dried over Na_2SO_4 and concentrated. The residue was purified by flash chromatography ($CHCl_3$–MeOH = 100 : 1 to 50 : 1) to give **3a** as a white powder (78%, 321 mg). ^1H NMR (400 MHz, $CDCl_3$) δ 7.72 (2H, d, J = 7.6 Hz), 7.54 (2H, m), 7.34 (4H, d, J = 7.2 Hz), 7.29–7.16 (10H, m), 6.80–6.77 (4H, m), 6.06 (0.5H, br), 5.95 (0.5H, s), 5.78–5.70 (1H, m), 5.53–5.46 (1H, m), 4.61–4.53 (1H, m), 4.31–4.30 (2H, m), 4.17–4.12 (2H, m), 3.75 (6H, s), 3.60–3.24 (10H, m), 2.68–2.26 (4H, m), 1.72 (1.5H, s), 1.71 (1.5H, s), 1.24–1.00 (12H, m); ^{31}P NMR (162 MHz, $CDCl_3$) δ 148.9, 147.9; HRMS (ESI-TOF) Calcd for $C_{57}H_{66}N_6O_9P$: $[M + H]^+$, 1009.4623; Found: 1009.4613.

3′-*O*-[2-Cyanoethoxy(diisopropylamino)phosphino]-5′-*O*-(4,4′-dimethoxytrityl)-N^2-(2-((9-fluorenylmethyloxycarbonyl)amino)-propyl) 2′-deoxy-5-methylisocytidine (3b). 4,4′-Dimethoxytrityl chloride (521 mg, 1.5 mmol) was added to a solution of **2b** (400 mg, 0.8 mmol) in 7.7 ml of pyridine, and the mixture was stirred at room temperature for 1 h. The solution was diluted with CH_2Cl_2 and successively washed with a saturated aqueous $NaHCO_3$ solution and brine, dried over Na_2SO_4 and concentrated. The residue was purified by flash chromatography ($CHCl_3$–MeOH = 1 : 0 to 20 : 1) to give a white powder (65%, 410 mg). To a solution of this white powder (200 mg, 0.2 mmol) in 4.9 ml of CH_2Cl_2 were added DIPEA (255 µl, 1.5 mmol) and 2-cyanoethyl-*N*,*N*-diisopropyl-chlorophosphoramidite (163 µl, 0.7 mmol), and then the mixture was stirred at 0 °C for 1 h. The solution was diluted with CH_2Cl_2 and successively washed with a saturated aqueous $NaHCO_3$ solution and brine, dried over Na_2SO_4 and concentrated. The residue was purified by flash chromatography ($CHCl_3$–MeOH = 1 : 0 to 100 : 1) to give **3b** as a white powder (97%, 242 mg). ^1H NMR (400 MHz, $CDCl_3$) δ 7.73 (2H, d, J = 7.2 Hz), 7.59 (2H, t, J = 7.2 Hz), 7.36 (4H, t, J = 7.2 Hz), 7.30–7.17 (10H, m), 6.81–6.77 (4H, m), 6.07 (0.5H, s), 5.96 (0.5H, s), 5.79 (1H, dd, J = 8.0, 13.6 Hz), 5.52–5.50 (0.5H, m), 5.46–5.44 (0.5H, m), 4.63–4.56 (1H, m), 4.35–4.30 (2H, m), 4.22–4.07 (2H, m), 3.75 (3H, s), 3.74 (3H,s), 3.62–3.25 (8H, m), 3.18–3.14 (2H, m), 2.60–2.35 (4H, m), 1.73 (1.5H, s), 1.72 (1.5H, s), 1.67–1.62 (2H, m), 1.28–0.99 (12H, m); ^{31}P NMR (162 MHz, $CDCl_3$) δ 148.9, 148.4; HRMS (ESI-TOF) Calcd for $C_{58}H_{68}N_6O_9P$ $[M + H]^+$: 1023.4780, Found: 1023.4753.

N^2-Propyl-2′-deoxy-5-methylisocytidine (4). 2,5′-Anhydrothymidine (100 mg, 0.5 mmol) in 5 ml of propylamine was stirred at room temperature for 15 h. The solution was concentrated, and the residue was purified by column chromatography ($CHCl_3$–MeOH = 20 : 1 to 5 : 1) to give **4** as a white foam (97%, 122 mg). n_{max} (neat) 3312, 2926, 1663 cm^{-1}; ^1H NMR (400 MHz, $CDCl_3$) δ 7.51 (1H, s), 5.79 (1H, dd, J = 8.0, 6.0 Hz), 4.48–4.45 (1H, m), 3.94 (1H, qr, J = 2.8 Hz), 3.79 (2H, ddd, J = 12.0, 6.8, 2.8 Hz), 3.41–3.32 (2H, m), 2.51–2.44 (1H, m), 2.22–2.16 (1H, m), 1.87 (3H, s), 1.60 (2H, sext, J = 7.2 Hz), 0.92 (3H, t, J = 7.2 Hz); ^{13}C NMR (125 MHz, CD_3OD) δ; 174.7, 154.6, 139.0, 114.4, 92.3, 89.1, 71.9, 62.0, 44.2, 39.6, 23.4, 13.5, 11.6; HRMS (ESI-TOF) Calcd for $C_{13}H_{22}N_3O_4$: $[M + H]^+$, 284.1605; Found: 284.1633.

3′-*O*-[2-Cyanoethoxy(diisopropylamino)phosphino]-5′-*O*-(4,4′-dimethoxytrityl)-N^2-propyl-2′-deoxy-5-methylisocytidine (5). 4,4′-Dimethoxytrityl chloride (263 mg, 0.8 mmol) was added to a solution of **4** (110 mg, 0.4 mmol) in 4 ml of pyridine, and the mixture was stirred at room temperature for 2 h. The solution was diluted with CH_2Cl_2 and successively washed with a saturated aqueous $NaHCO_3$ solution and brine, dried over Na_2SO_4 and concentrated. The residue was purified by flash chromatography ($CHCl_3$–MeOH = 50 : 1 to 10 : 1) to give a white powder (97%, 211 mg). To a solution of a white powder (155 mg, 0.3 mmol) in 2.7 ml of CH_2Cl_2 were added DIPEA (280 µl, 1.6 mmol) and 2-cyanoethyl-*N*,*N*-diisopropylchlorophosphoramidite (180 µl, 0.8 mmol), and then the mixture was stirred at 0 °C for 2 h. The solution was diluted with CH_2Cl_2 and successively washed with a saturated aqueous $NaHCO_3$ solution and brine, dried over Na_2SO_4 and concentrated. The residue was purified by flash chromatography ($CHCl_3$–MeOH = 100 : 1 to 80 : 1) to give **5** as a white powder (67%, 140 mg). ^1H NMR (400 MHz, $CDCl_3$) δ 7.36–7.20 (9H, m), 7.12 (0.5H, s), 7.01 (0.5H, s), 6.81–6.78 (4H, m), 5.72–5.55 (2H, m), 4.61–4.55 (1H, m), 4.20 (0.5H, m), 4.15 (0.5H, m), 3.77 (3H, s), 3.76 (3H, s), 3.70–3.23 (8H, m), 2.75–2.26 (4H, m), 1.79 (1.5H, s), 1.76 (1.5H, s), 1.57–1.48 (2H, m), 1.24–1.04 (12H, m), 0.85–0.80 (3H, m); ^{31}P NMR (162 MHz, $CDCl_3$) δ 149.63, 148.93; HRMS (ESI-TOF) Calcd for $C_{43}H_{57}N_5O_7P$: $[M + H]^+$, 786.3990; Found: 786.4010.

Synthesis of TFO containing aminoethyl-isodC or propylethyl-isodC derivatives

The TFO was cleaved from the CPG resin by heating in 28% aqueous ammonia overnight at 55 °C, followed by HPLC purification (HPLC conditions: column: Nacalai tesque COSMOSIL 5C18-ARII, 10 × 250 mm, solvents: A: 0.1 M TEAA buffer, B: CH_3CN, B: 10% to 40%/20 min, 40% to 100%/30 min, flow rate: 3.0 ml min^{-1}, UV: 254 nm). The DMTr group was deprotected in 5% aqueous acetic acid at room temperature for 15 min, followed by HPLC purification (HPLC conditions: column: SHISEIDO C18, 4.6 × 250 mm, solvents: A: 0.1 M TEAA buffer, B: CH_3CN, B: 10% to 30%/20 min, 30% to 100%/30 min, flow rate: 1.0 ml min^{-1}, UV: 254 nm).

Synthesis of TFO containing guanidinoethyl-isodC derivative

A suspension of the support-bound TFO in 250 μl of CH_3CN was treated with 1*H*-pyrazole-1-carboxamidine hydrochloride (1.0 mg) and DBU (50 μl) overnight at room temperature. The support was washed two times with CH_3CN, and the resulting TFO was treated with 28% aqueous ammonia overnight at 55 °C, followed by HPLC purification (HPLC conditions: column: Nacalai tesque COSMOSIL 5C18-ARII, 10 × 250 mm, solvents: A: 0.1 M TEAA buffer, B: CH_3CN, B: 10% to 40%/20 min, 40% to 100%/30 min, flow rate: 3.0 ml min^{-1}, UV: 254 nm). The DMTr group was deprotected in 5% aqueous acetic acid at room temperature for 15 min, followed by HPLC purification (HPLC conditions: column: SHISEIDO C18, 4.6 × 250 mm, solvents: A: 0.1 M TEAA buffer, B: CH_3CN, B: 10% to 30%/20 min, 30% to 100%/30 min, flow rate: 1.0 ml min^{-1}, UV: 254 nm).

Synthesis of TFO containing ureidoethyl-isodC derivative

A suspension of the support-bound TFO in 250 μl of CH_3CN was treated with trimethylsilyl isocyanate (50 μl) and DBU (50 μl) overnight at 55 °C. The support was washed two times with CH_3CN, and the resulting TFO was treated with 28% aqueous ammonia overnight at 55 °C, followed by HPLC purification (HPLC conditions: column: Nacalai tesque COSMOSIL 5C18-ARII, 10 × 250 mm, solvents: A: 0.1 M TEAA buffer, B: CH_3CN, B: 10% to 40%/20 min, 40% to 100%/30 min, flow rate: 3.0 ml min^{-1}, UV: 254 nm). The DMTr group was deprotected in 5% aqueous acetic acid at room temperature for 15 min, followed by HPLC purification (HPLC conditions: column: SHISEIDO C18, 4.6 × 250 mm, solvents: A: 0.1 M TEAA buffer, B: CH_3CN, B: 10% to 30%/20 min, 30% to 100%/30 min, flow rate: 1.0 ml min^{-1}, UV: 254 nm).

Computational calculation of the *N*-guanidinoethyl-isodC–CG complex

Geometry optimization was performed using the B3LYP density functional with 6-31G* basis set for *N*-guanidinoethyl-isodC–CG complex *in vacuo* as implemented in the Gaussian09 program.

DNase I foot printing analysis

Each TFO (**TFO**(T): GGAAGGATGGAGGAGGGA or **TFO**(et-guanidino): GGAAGGAZGGAGGAGGGA (Z = et-guanidine-isodC), 0 to 1 μM) was incubated with 100 nM of the 5′ FAM-labeled duplex (5′ CCCCTCGGAAGGAAGGAGGAGGGATCCTCTCTCG-GAAGGACGGAGGAGGGATCCCCC 3′ and 3′ GGGGAGCCTTC-CTTCCTCCTCCCTAGGAGAGAGCCTTCCTGCCTCCTCCCTAGG-GGG-FAM 5′, 100 nM) in a buffer consisting of 20 mM Tris-HCl, 5 mM $MgCl_2$, 2.5 mM spermidine and 10% sucrose at pH 7.5 (total volume of 8 μl) for 12 h at 30 °C. 2 μl of a DNase I solution (0.1 unit in a buffer containing 400 mM of Tris-HCl, 80 mM of $MgCl_2$ and 50 mM of DTT at pH 7.5) was added and the solution was digested for 1 min at 30 °C. The reaction was stopped by adding formamide containing 20 mM EDTA at 90 °C, and electrophoresis was carried out with an 18% denaturing polyacrylamide gel containing 8 M urea. FAM-bands were visualized by LAS-4000.

Acknowledgements

This study was supported by a Grant-in-Aid for Scientific Research (S) (S.S.; Grant Number 21229002) and a Grant-in-Aid for Young Scientists (A) (Y.T.; Grant Number 24689006) from the Japan Society for the Promotion of Science.

Notes and references

1 A. Mukherjee and K. M. Vasquez, *Biochimie*, 2011, **93**, 1197.
2 V. N. Soyfer and V. N. Potaman, *Triple-Helical Nucleic Acids*, Springer-Verlag New York, Inc., 1996.
3 (*a*) Y. Hari, S. Obika and T. Imanishi, *Eur. J. Org. Chem.*, 2012, 2875; (*b*) Y. Hari, M. Akabane, Y. Hatanaka, M. Nakahara and S. Obika, *Chem. Commun.*, 2011, **47**, 4424; (*c*) A. Semenyuk, E. Darian, J. Liu, A. Majumdar, B. Cuenoud, P. S. Miller, A. D. MacKerell Jr. and M. M. Seidman, *Biochemistry*, 2010, **49**, 7867; (*d*) D. A. Rusling, V. E. C. Powers, R. T. Ranasinghe, Y. Wang, S. D. Osborne, T. Brown and K. R. Fox, *Nucleic Acids Res.*, 2005, **33**, 3025.
4 (*a*) N. A. Kolganova, A. K. Shchyolkina, A. V. Chudinov, A. S. Zasedatelev, V. L. Florentiev and E. N. Timofeev, *Nucleic Acids Res.*, 2012, **40**, 8175; (*b*) S. P. Parel and C. J. Leumann, *Nucleic Acids Res.*, 2001, **29**, 2260.
5 (*a*) S. Sasaki, Y. Taniguchi, R. Takahashi, Y. Senko, K. Kodama, F. Nagatsugi and M. Maeda, *J. Am. Chem. Soc.*, 2004, **126**, 516; (*b*) S. Sasaki, H. Yamauchi, F. Nagatsugi, R. Takahashi, Y. Taniguchi and M. Maeda, *Tetrahedron Lett.*, 2001, **42**, 6915.
6 Y. Taniguchi, A. Nakamura, Y. Senko, F. Nagatsugi and S. Sasaki, *J. Org. Chem.*, 2006, **71**, 2115.
7 Y. Taniguchi and S. Sasaki, *Org. Biomol. Chem.*, 2012, **10**, 8336.
8 (*a*) Y. Taniguchi, H. Okamura, N. Fujino and S. Sasaki, *Tetrahedron*, 2013, **69**, 600; (*b*) E. Aoki, Y. Taniguchi, Y. Wada and S. Sasaki, *ChemBioChem*, 2012, **13**, 1152; (*c*) Y. Taniguchi, Y. Uchida, T. Takaki, E. Aoki and S. Sasaki,

Bioorg. Med. Chem., 2009, **17**, 6803; (*d*) Y. Taniguchi, M. Togo, E. Aoki, Y. Uchida and S. Sasaki, *Tetrahedron*, 2008, **64**, 7164.

9 K. Dittrich, J. Gu, R. Tinder, M. Hogan and X. Gao, *Biochemistry*, 1994, **33**, 4111.

10 (*a*) S. C. Jurczyk, J. T. Kodra, J. D. Rozzell, S. A. Benner and T. R. Battersby, *Helv. Chim. Acta*, 1998, **81**, 793; (*b*) Y. Tor and P. B. Dervan, *J. Am. Chem. Soc.*, 1993, **115**, 461.

11 A. Semenyuk, E. Darian, J. Liu, A. Majumdar, B. Cuenoud, P. S. Miller, A. D. Mackerell Jr. and M. M. Seidman, *Biochemistry*, 2010, **49**, 7867.

12 P. A. Beal and P. B. Dervan, *Nucleic Acids Res.*, 1992, **20**, 2773.

13 R. H. Durland, T. S. Rao, G. R. Revankar, J. H. Tinsley, M. A. Myrick, D. M. Seth, J. Rayford, P. Singh and K. Jayaraman, *Nucleic Acids Res.*, 1994, **22**, 3233.