HETEROCYCLES, Vol. 103, No. 1, 2021, pp. 322 - 330. © 2021 The Japan Institute of Heterocyclic Chemistry
Received, 29th May, 2020, Accepted, 31st July, 2020, Published online, 2nd September, 2020
DOI: 10.3987/COM-20-S(K)16

THE EFFECT OF THE BASE TRIPLETS ADJACENT TO A T•CG OR 5-METHYLC•CG TRIPLET IN THE TRIPLEX DNA

Yuta Ito,[1] Chisa Hama,[1] Takashi Osawa,[1,2] and Yoshiyuki Hari[1]*

[1]Faculty of Pharmaceutical Sciences, Tokushima Bunri University, Nishihama, Yamashiro-cho, Tokushima 770-8514, Japan. [2]Graduate School of Pharmaceutical Sciences, Osaka University, 1-6 Yamadaoka, Suita, Osaka 565-0871, Japan. Corresponding author. E-mail: hari@ph.bunri-u.ac.jp

Dedicated to Professor Dr. Yasuyuki Kita on the occasion of his 77th birthday

Abstract – The effect of the base triplets adjacent to an X•CG triplet [X = T, 5-methylcytosine (C̲), or their LNA and ENA counterparts] in a triplex DNA was examined through the UV-melting experiments. Whereas the triplex containing a T•CG triplet, sandwiched by T•AT triplets, was stable, the replacement of T•AT with C̲•GC at the 5′-adjacent site significantly destabilized the stability. When one or both of the adjacent T•AT triplets were replaced with C̲•GC, the triplexes containing a C̲•CG triplet were generally more stable than those containing a T•CG triplet. Further increases in the thermal stability of the triplexes were observed by LNA and ENA modifications. These results suggested that a C̲•CG triplet rather than a T•CG triplet would be suitable for triplex formation when C̲•GC triplets existed at its adjacent site.

INTRODUCTION

Triplex-forming oligonucleotide (TFO) consisting of homopyrimidine sequence is able to form the stable triplex, so-called parallel triplex, with the homopurine-homopyrimidine region in the double-stranded DNA (dsDNA) because T and protonated C of TFOs recognize A and G bases of AT and GC base pairs *via* two Hoogsteen hydrogen bonds, respectively (Figure 1a). Therefore, such TFO is expected to be a promising candidate for antigene technology.[1,2] T or G bases in the TFO can also recognize C and T bases of CG or TA base pairs in the dsDNA, respectively;[3-6] however, the triplexes containing T•CG and G•TA triplets exhibited lower thermal stability than those containing T•AT or C•GC ones because T•CG and

G•TA form only one hydrogen bond (Figure 1b). To overcome this sequence limitation, variety of chemical modifications of nucleobase and sugar moieties have been investigated.[7,8] We have also investigated the chemically modified nucleobases for the recognition of a CG base pair,[9-15] and developed 2′-*O*,4′-*C*-methylene-bridged nucleic acids (2′,4′-BNAs)/locked nucleic acids (LNAs) bearing 2-pyridone (P^B),[9-11] 2-pyridine (Py^B),[12] and 4-[(3*S*)-3-guanidinopyrrolidino]-5-methylpyrimidin-2-one (GP^B).[14] For example, the triplex **T1** containing a GP^B•CG triplet, sandwiched by T•AT triplets, was highly stable (Figure 2). However, when one or both of the adjacent T•AT triplets were replaced with C•GC (C = 5-methylcytosine), the stability of the triplexes **T2-4** decreased drastically. Notably, the melting temperatures (T_ms) of the triplexes **T3** and **T4**, where a C•GC triplet is the 5′-adjacent site of a GP^B•CG triplet, decreased by approximately 20 °C compared to that of **T1** with T•AT triplets at both the sites of a GP^B•CG triplet. Under such a background, exploring a suitable nucleobase for the recognition of a CG base pair even in the presence of C•GC triplets at the neighboring sites is needed.

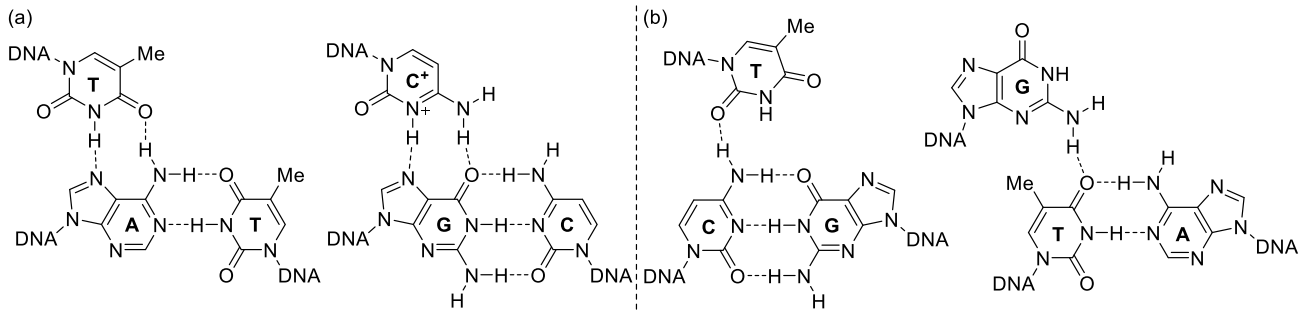


Figure 1. (a) T•AT and C•GC triplets and (b) T•CG and G•TA triplets

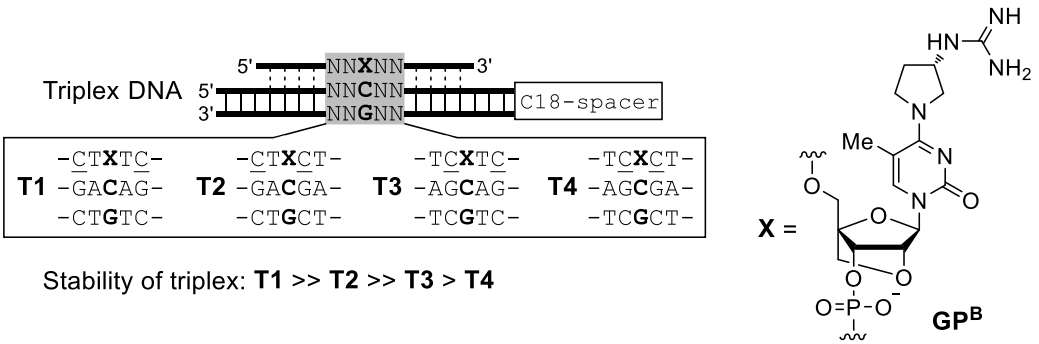


Figure 2. The effect of the base triplets adjacent to a GP^B•CG triplet in the triplex DNA

It is known that C, besides T, can moderately interact with a CG base pair.[16] However, to the best of our knowledge, there is no report on the effect of the base triplets, adjacent to a C•CG triplet, on the triplex stability. Since C may recognize a CG base pair via not only a conventional hydrogen bond but also a CH-O interaction (Figure 3),[17] it is anticipated that a C•CG triplet can stabilize the triplex compared to a

T•CG one. On the other hand, it has been reported that the LNA and 2′-*O*,4′-*C*-ethylene-bridged nucleic acid (ENA) modifications in the TFO dramatically improved the triplex-forming ability with dsDNA.[18,19] Thus, the combination of C and LNA or ENA modification is expected to further stabilize the triplex. Herein, the effect of base triplets at the neighboring sites of T•CG, C•CG, and their LNA and ENA counterparts in the triplex DNA is described.

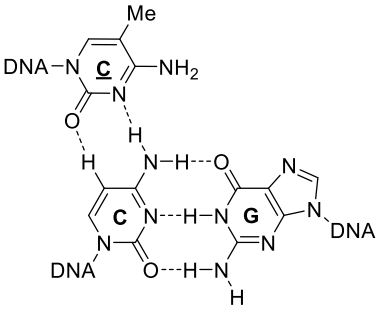


Figure 3. The proposed recognition mode of C to a CG base pair

RESULTS AND DISCUSSION

All the TFOs and hairpin dsDNAs used in this study are shown in Figure 4.[20] TFOs containing ENA-T (T^E), and ENA-C (C^E) were synthesized using an automated DNA synthesizer by using the standard phosphoramidite chemistry with a prolonged coupling time of 10 min for the incorporation of T^E and C^E. They were purified by reversed-phase HPLC (RP-HPLC) and characterized by electrospray ionization-time-of-flight (ESI-TOF) mass spectrometry. TFOs containing LNA-T (T^L), LNA-C (C^L) and hairpin dsDNAs were purchased from GeneDesign, Inc.

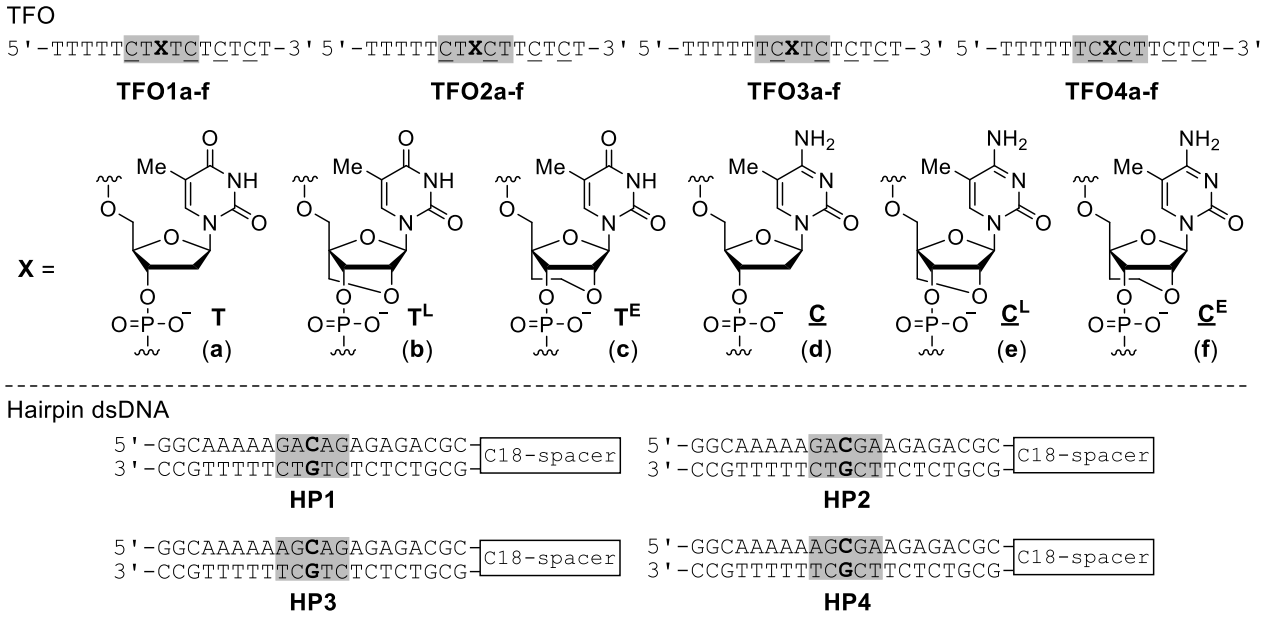
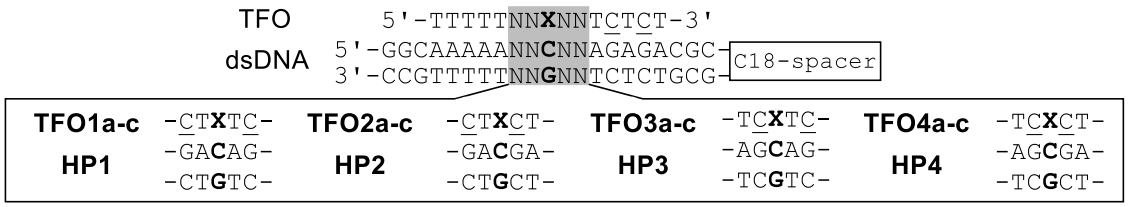


Figure 4. TFOs and hairpin dsDNAs used in this study

Initially, the triplex-forming ability of the TFOs containing natural T, T^L, and T^E with the hairpin dsDNA targets was evaluated using the UV-melting experiments (Table 1).[21] For the triplexes containing a T•CG triplet, the T_m of the triplex **TFO2a/HP2** with a C•GC triplet at the 3′-adjacent site of a T•CG triplet was 38 °C, which was almost same as that of **TFO1a/HP1** sandwiched by T•AT triplets. Meanwhile, the T_ms of the triplexes **TFO3a/HP3** and **TFO4a/HP4** where a C•GC triplet is at the 5′-adjacent site of a T•CG triplet were significantly lower than that of **TFO1a/HP1**. These results are consistent with the previous report that the triplexes containing C•GC triplets at either 5′-flanking site or both 3′- and 5′-flanking sites of a T•CG triplet were remarkably destabilized.[9,10] The T_m of the triplex **TFO1b/HP1** containing a T^L•CG triplet sandwiched by T•AT triplets was 48 °C, which showed a dramatic increase compared to that of the unmodified **TFO1a/HP1**. However, the triplexes **TFO3b/HP3** and **TFO4b/HP4** where the 5′-adjacent site is a C•GC triplet were significantly destabilized and their ΔT_ms were –10 °C and –14 °C, respectively. The effect of C•GC triplets at the neighboring sites of a T^E•CG triplet on the triplex stability showed similar trends to those of the T^L•CG triplet. The T_m of the triplex **TFO4c/HP4** decreased to 28 °C, which was almost same as that of **TFO4a/HP4** (T_m = 29 °C) containing a T•CG triplet. The UV-melting profiles of the triplexes **TFO4a-c/HP4** are shown in Figure 5.

Table 1. T_ms (°C) of the triplexes containing T•CG, T^L•CG, and T^E•CG triplets



Triplex	**a** (**X** = T)	**b** (**X** = T^L)	**c** (**X** = T^E)
TFO1a-c/HP1	39	48	43
TFO2a-c/HP2	38 (–1)	46 (–2)	40 (–3)
TFO3a-c/HP3	31 (–8)	38 (–10)	33 (–10)
TFO4a-c/HP4	29 (–10)	34 (–14)	28 (–15)

Conditions: 10 mM sodium phosphate buffer (pH 6.0), 100 mM KCl, 5 mM $MgCl_2$, and 1.5 μM of each oligonucleotide. The change in the T_m value (ΔT_m) in parentheses compared with that of the corresponding triplexes formed between **TFO1a-c** and **HP1**.

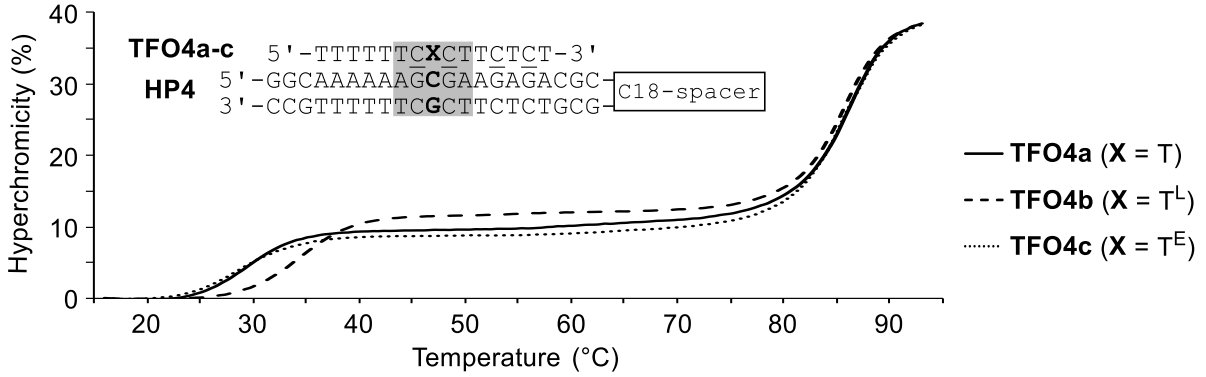
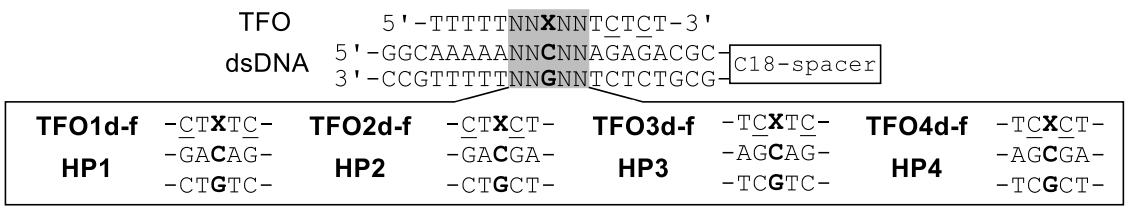


Figure 5. UV-Melting profiles of the triplexes formed between **TFO4a-c** and **HP4**

Then, the triplex-forming ability of the TFOs containing C, C^L, and C^E was examined.[21] The results are summarized in Table 2, and the representative UV-melting profiles are shown in Figure 6. The T_ms of the triplexes **TFO2d–f/HP2**, where the 3′-flanking site of an X•CG triplet is a C•GC triplet, were almost equal to or slightly superior to those of the corresponding triplexes **TFO1d–f/HP1** sandwiched by T•AT triplets. Although the triplexes **TFO3d/HP3** (ΔT_m = –4 °C) and **TFO4d/HP4** (ΔT_m = –5 °C) containing the 5′-adjacent C•GC triplet exhibited decreased stability relative to **TFO1d/HP1**, their ΔT_ms were significantly smaller than those observed in **TFO3a/HP3** (ΔT_m = –8 °C) and **TFO4a/HP4** (ΔT_m = –10 °C) containing a T•CG triplet. Previously, a solution structure of a DNA triplex containing a T•CG triplet has been investigated by Radhakrishnan and Patel.[4] The T of a T•CG triplet is partially overlapped with the 5′-adjacent T in the TFO and a stacking interaction may occur. Thus, C of the 5′-C•GC triplet in place of the T•AT triplet is considered to weaken the interaction with the T of a T•CG triplet though the reason is unclear. On the other hand, in the case of triplexes containing the C•CG triplet, because a C•CG triplet slightly differs from a T•CG triplet in the recognition mode (Figures 1 and 3), the structural difference may possibly reduce the destabilization by the 5′-C•GC triplet. Although the LNA-modified **TFO2e–4e** and ENA-modified **TFO2f–4f** had the same tendency as **TFO2d–4d**, their T_ms was clearly high in all the cases.

Table 2. T_ms (°C) of the triplexes containing C•CG, C^L•CG, and C^E•CG triplets



Triplex	**d** (**X** = C)	**e** (**X** = C^L)	**f** (**X** = C^E)

TFO1d-f/HP1	37	45	41
TFO2d-f/HP2	38 (+1)	47 (+2)	42 (+1)
TFO3d-f/HP3	33 (–4)	41 (–4)	38 (–3)
TFO4d-f/HP4	32 (–5)	38 (–7)	35 (–6)

Conditions: 10 mM sodium phosphate buffer (pH 6.0), 100 mM KCl, 5 mM $MgCl_2$, and 1.5 μM of each oligonucleotide. The change in the T_m (ΔT_m) in parentheses compared with that of the corresponding triplexes formed between **TFO1d-f** and **HP1**.

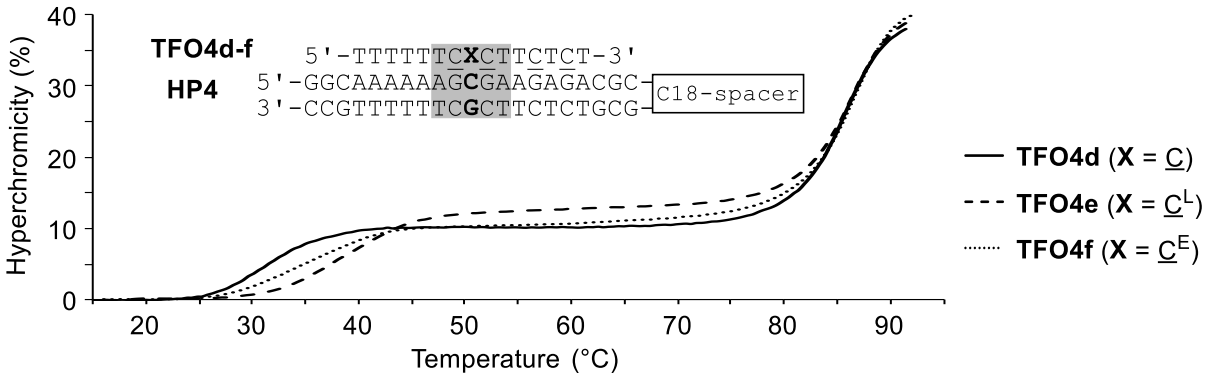


Figure 6. UV-Melting profiles of the triplexes formed between **TFO4d-f** and **HP4**

Finally, the difference (ΔT_m) from the T_m of a T•CG triplet in each triplex is summarized in Figure 7. In the triplexes where both the sites of an X•CG triplet are T•AT triplets, the TFOs containing T (**TFO1a**), T^L (**TFO1b**), and T^E (**TFO1c**) formed stable triplexes with **HP1** compared to those containing the corresponding C congeners (**TFO1d–f**). In contrast, when the triplexes contained a C•GC triplet at the 3′-adjacent site of an X•CG triplet, the triplex-forming abilities of **TFO2d-f** containing C, C^L, and C^E were similar to or slightly higher than those of the corresponding T congeners (**TFO2a-c**). Furthermore, for the triplexes with C•GC triplets at either the 5′-adjacent or both 3′- and 5′-adjacent sites of an X•CG triplet, the TFOs containing C, C^L, and C^E as a CG recognition base efficiently stabilized the triplexes compared to those containing the corresponding T congeners. For instance, the ΔT_ms of **TFO3f/HP3** and **TFO4f/HP4** containing C^E were higher (5 °C and 7 °C, respectively) than those of the triplexes containing T^E (**TFO3c/HP3** and **TFO4c/HP4**, respectively). These results suggested that C could be a promising nucleobase for the recognition of CG base pairs when a C•GC triplet existed at the adjacent sites of an X•CG triplet in the triplex DNA. In particular, it was found that the use of C^L would allow the formation of a stable triplex.

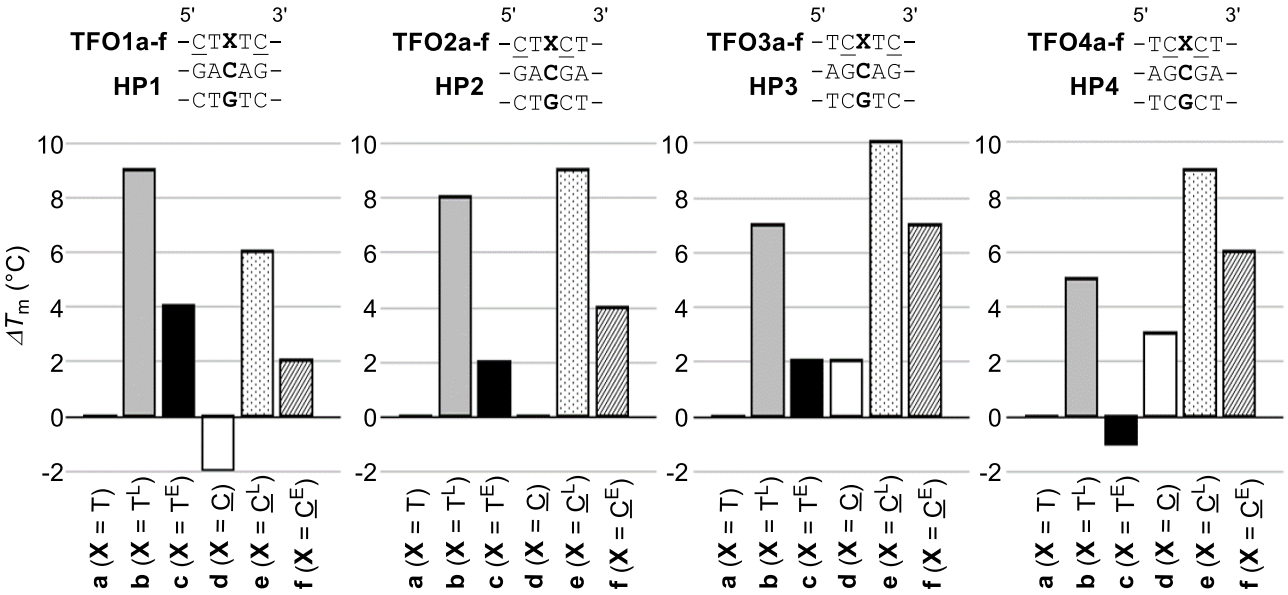


Figure 7. Summary of the effect of the base triplets adjacent to an X•CG triplet. The ΔT_m showed the difference from that of the corresponding triplexes containing T•CG triplets.

In conclusion, the effect of the base triplets adjacent to an X•CG triplet (X = T, C, and their LNA and ENA counterparts) in the triplex DNA was investigated in this study. The UV-melting experiments demonstrated that C could be a suitable base for a CG recognition in the triplex formation when a C•GC triplet existed at the adjacent sites of an X•CG triplet. In addition, the introduction of C^L and C^E into TFO improved thermal stability of the triplex compared to that of unmodified C. Although T^L can recognize a CG base pair, these results demonstrated that C^L could also be a powerful material for recognition of a CG base pair depending on the sequences.

EXPERIMENTAL

General Methods. The synthesis of the TFOs was performed on an automated DNA synthesizer (GeneDesign nS-8II). For high-performance liquid chromatography (HPLC), a JASCO EXTREMA (PU-4180, CO-4060, UV-4075, and AS-4050) instrument with a CHF122SC (ADVANTEC) fraction collector was used. High-resolution mass spectrometry was performed on a Waters SYNAPT G2-Si mass spectrometer (Quadrupole/TOF). The UV-melting experiments were carried out using a JASCO V-730 UV/VIS spectrophotometer equipped with a T_m analysis accessory.

Synthesis of TFOs. dT-phosphoramidite (Glen Research), Ac-5-Me-dC-phosphoramidite (Sigma), and T^E phosphoramidite (KNC laboratories) were used. The C^E phosphoramidite was synthesized as previously described.[19] Syntheses of the TFOs were performed on a 0.2 μmol scale using a standard phosphoramidite protocol (DMTr-ON mode), except for a prolonged coupling time of 10 min for T^E and

$\underline{C}^E$. The cleavage from the CPG support and the removal of the protecting groups were accomplished by 28% NH_3 aq. at room temperature for 2 h. After the removal of ammonia in vacuo, the crude TFOs were purified with Sep-Pak Plus C18 cartridges (Waters), followed by reversed-phase HPLC (Waters XBridge Prep Shield RP18 5 μm, 10 × 50 mm) using 0.1 M triethylammonium acetate buffer (pH 7.0) as an ion-pairing mobile phase. The compositions of the modified TFOs were confirmed by ESI-TOF-MS analysis. The deconvoluted ESI-TOF-MS data [M] for **TFO1c–4c** and **TFO1f–4f** were as follows: **TFO1c**, found 4539.00 (calcd 4539.06); **TFO2c**, found 4539.00 (calcd 4539.06); **TFO3c**, found 4539.50 (calcd 4539.06); **TFO4c**, found 4538.70 (calcd 4539.06); **TFO1f**, found 4538.40 (calcd 4538.08); **TFO2f**, found 4538.30 (calcd 4538.08); **TFO3f**, found 4538.30 (calcd 4538.08); **TFO4f**, found 4538.20 (calcd 4538.08).

UV-Melting Experiments. For UV-melting experiments using the triplexes formed by TFO and hairpin dsDNA, the oligonucleotides were dissolved in 10 mM sodium phosphate buffer (pH 6.0) containing 100 mM KCl and 5 mM $MgCl_2$ to give a final concentration of 1.5 μM for each strand. The hairpin dsDNA targets linked to a hexa(ethylene glycol) unit (C18-spacer) were used to stabilize the duplex and to prevent the transition of the duplex into the single strands from overlapping with that of the triplex into the TFO and the duplex. The samples were annealed in boiling water, followed by slow cooling to room temperature. The melting profiles were recorded at 260 nm from 15 to 95 °C for at a scan rate of 0.5 °C/min. The two-point average method was employed to obtain T_m, and the final values were determined by averaging three independent measurements, which were accurate within a 1 °C range.

ACKNOWLEDGEMENTS

This work was partially supported by the Uehara Memorial Foundation.

REFERENCES AND NOTES

1. S. Buchini and C. J. Leumann, *Curr. Opin. Chem. Biol.*, 2003, **64**, 717.
2. D. Praseuth, A. L. Guieysse, and C. Hélène, *Biochim. Biophys. Acta Gene Struct. Expr.*, 1999, **1489**, 181.
3. K. Yoon, C. A. Hobbs, J. Koch, M. Sardaro, R. Kutny, and A. L. Weis, *Proc. Natl. Acad. Sci. USA*, 1992, **89**, 3840.
4. I. Radhakrishnan and D. J. Patel, *J. Mol. Biol.*, 1994, **241**, 600.
5. L. C. Griffin and P. B. Dervan, *Science*, 1989, **245**, 967.
6. I. Radhakrishnan, X. Gao, C. de los Santos, D. Live, and D. J. Patel, *Biochemistry*, 1991, **30**, 9022.
7. Reviews: S. O. Doronina and J.-P. Behr, *Chem. Soc. Rev.*, 1997, **26**, 63; I. Luyten and P. Herdewijn,

Eur. J. Med. Chem., 1998, **33**, 515; D. M. Gowers and K. R. Fox, *Nucleic Acids Res.*, 1999, **27**, 1569; K. R. Fox, *Curr. Med. Chem.*, 2000, **7**, 17; M. G. M. Purwanto and K. Weisz, *Curr. Org. Chem.*, 2003, **7**, 427; K. R. Fox and T. Brown, *Quart. Rev. Biophys.*, 2005, **38**, 311; 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.

8. Recent papers: A. Ohkubo, T. Ohnishi, S. Nishizawa, Y. Nishimura, and S. Hisamatsu, *Bioorg. Med. Chem.*, 2020, **28**, 115350; A. Ohkubo, K. Yamada, Y. Ito, K. Yoshimura, K. Miyauchi, T. Kanamori, Y. Masaki, K. Seio, H. Yuasa, and M. Sekine, *Nucleic Acid Res.*, 2015, **43**, 5675.
9. S. Obika, Y. Hari, M. Sekiguchi, and T. Imanishi, *Angew. Chem. Int. Ed.*, 2001, **40**, 2079.
10. S. Obika, Y. Hari, M. Sekiguchi, and T. Imanishi, *Chem. Eur. J.*, 2002, **8**, 4796.
11. H. Torigoe, Y. Hari, S. Obika, and T. Imanishi, *Nucleosides Nucleotides Nucleic Acids*, 2003, **22**, 1571.
12. Y. Hari, S. Matsugu, H. Inohara, Y. Hatanaka, M. Akabane, T. Imanishi, and S. Obika, *Org. Biomol. Chem.*, 2010, **8**, 4176.
13. Y. Hari, M. Akabane, Y. Hatanaka, M. Nakahara, and S. Obika, *Chem. Commun.*, 2011, **47**, 4424.
14. Y. Hari, M. Akabane, and S. Obika, *Chem. Commun.*, 2013, **49**, 7421.
15. M. Akabane-Nakata, S. Obika, and Y. Hari, *Org. Biomol. Chem.*, 2014, **12**, 9011.
16. I. Prévot-Halter and C. J. Leumann, *Bioorg. Med. Chem. Lett.*, 1999, **9**, 2657.
17. S. Buchini and C. J. Leumann, *Tetrahedron Lett.*, 2003, **44**, 5065; S. Buchini and C. J. Leumann, *Angew. Chem. Int. Ed.*, 2004, **43**, 3925.
18. S. Obika, Y. Hari, K. Morio, and T. Imanishi, *Tetrahedron Lett.*, 2000, **41**, 221.
19. M. Koizumi, K. Morita, M. Daigo, S. Tsutsumi, K. Abe, S. Obika, and T. Imanishi, *Nucleic Acids Res.*, 2003, **31**, 3267.
20. The sequences were designed not to change the composition/number of T•AT and C•GC triplets among triplexes, which was considered to effectively elucidate the neighboring effect of an X•CG triplet.
21. The triplex-forming abilities of **TFO1a**, **1b**, **1d**, and **1e** with **HP1** were measured in the pH 7.0 solution of the same salt concentration. Because C•GC triplets requiring protonation were destabilized by the pH change from 6.0 to 7.0, the stabilities of all triplexes were largely decreased. The stabilities of the triplexes **TFO1a/HP1** containing a T•CG triplet and **TFO1d/HP1** containing a C•CG triplet were too low (T_ms < 20 °C). The T_ms of LNA-modified triplexes **TFO1b/HP1** and **TFO1e/HP1** were 27 °C and 24 °C, respectively, and the T_m difference was 3 °C, which was equal to that (3 °C) of pH 6.0. These results suggested that the C of a C•CG triplet was not affected by protonation under this condition.