Development of artificial nucleic acids to selectively recognize CG inversion site for the triplex DNA formation and its antigene application to efficient inhibition of hTERT gene

王, 磊

Development of artificial nucleic acids to selectively recognize CG inversion site for the triplex DNA formation and its antigene application to efficient inhibition of hTERT gene

Division of Bioorganic and Synthetic Chemistry 3PS16010E Lei Wang

[Introduction]

The antigene strategy is based on the sequence-specific recognition of duplex DNA by triplex-forming oligonucleotides (TFOs) at the major groove side, which can modulate gene expression at the transcriptional level (**Fig. 1**). TFOs have been exploited in a wide range of biological activities, including gene expression regulation, site-specific DNA damage, DNA repair, recombination and mutagenesis.[1] Triplex DNA is formed by the interaction of TFOs with the homo-purine strand in the target duplex DNA. According to the binding orientation of the phosphate backbone of the TFOs, triplex DNA formations are divided into parallel type triplex DNA and anti-parallel type triplex DNA. In an anti-parallel type triplex DNA, a purine-rich TFO binds to the homo-purine strand in the reverse direction to the phosphate backbone to form the A/AT, G/GC and T/AT triplets through two reverse Hoogsteen hydrogen bonds under neutral conditions (**Fig. 2A**). However, there is no natural nucleoside can that form stable hydrogen bonds with dT and dC in TA and CG base pairs from the major groove side, which are known as inversion sites (**Fig. 2B**). Their presence in the target purine DNA sequences drastically restricts the formation of stable triplex DNA. Hence, the design of an artificial nucleoside, which can selectively recognize these inversion sites with a high affinity, should be of great significance.

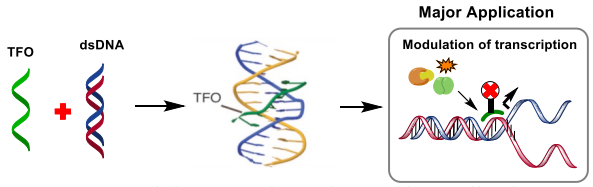


Figure 1 Triplex DNA formation and its application

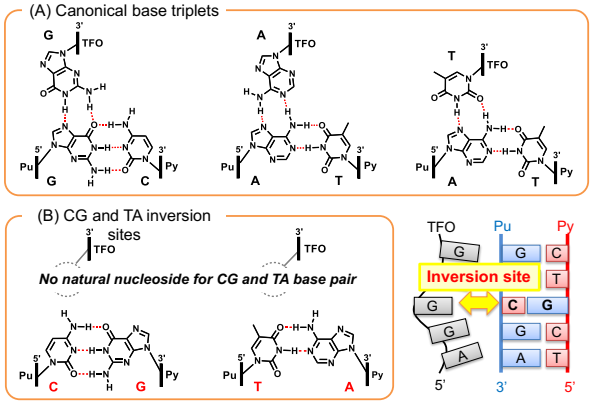


Figure 2 (A) Canonical base triplets and (B) Base inversion sites

[Method and Result]

① Development of 3-halogenated AP-ΨdC derivatives:

Recently, we determined that 2-amino-3-methylpyridinyl pseudo-dC (^{3Me}AP-ΨdC, **Fig. 3A**) possessed significant affinity and selectivity toward a CG inversion site and showed effective inhibition of gene expression.[2] It was anticipated that the 1-*N* position of 2-aminopyridine within ^{3Me}AP-ΨdC was protonated to form three hydrogen bonds with the CG base pair. To affirm this hypothesis, I designed AP-ΨdC derivatives in which the 3-methyl group of the aminopyridine unit was replaced by a halogen atom (Cl, Br or I, **Fig. 3B**) that caused a reduction of the pK_a value of the 1-*N* position. Therefore, it was expected that the replacement would change the triplex stabilization property of AP-ΨdC to gain more information about the electronic effect on the triplex stability.[3]

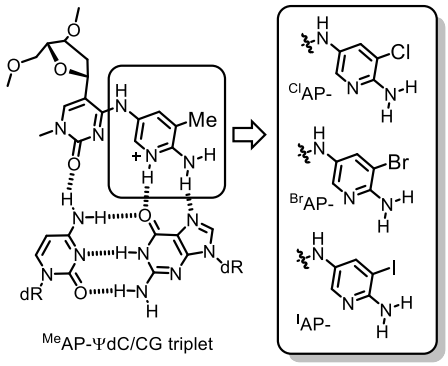
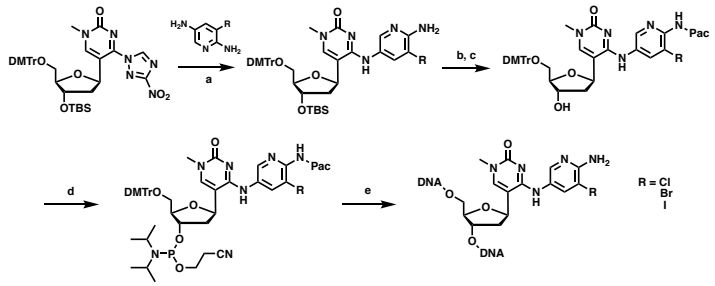


Figure 3 (A) ^{3Me}AP-ΨdC/CG triplets and (B) design of 3-halogenated ΨdC derivatives

The synthesis of the corresponding phosphoramidite compounds of the 3-halogenated AP-ΨdC derivatives is shown in **Scheme 1**. Finally, those corresponding phosphoramidite compounds were incorporated into the middle of the 18 mer

TFOs using an automated DNA synthesizer. The synthesized TFOs were cleaved from the resin and the protecting groups were removed by heating in 28 % ammonium hydroxide at 55 °C, then purified by HPLC. The DMTr group was removed in an aqueous AcOH solution, then the final target product was purified by HPLC. The structure and purity of the TFOs bearing different combinations of the flanking bases for each derivative were confirmed by MALDI-TOF MS measurements. To evaluate the triplex forming ability of the synthesized TFO, I used FAM labeled duplex DNAs for the gel-electrophoretic mobility shift assays. The equilibrium association constants (K_s) were also obtained and summarized in **Tab. 1**. Although its stability is lower than the original ^{3Me}AP-ΨdC, the TFO containing the 3'-G-(^{I}AP-ΨdC)-A-5' context showed a selective triplex forming ability with the corresponding target duplex DNA including a CG base pair. These results have shown that the hydrogen bond formation by the protonation at 1-*N* of the aiminopyridine on the AP-ΨdC derivatives mainly contributes to the formation of the stable and selective triplex DNA for the CG inversion site.



Reagent and condition: (a) DIPEA, MeCN, reflux; (b) PacCl, pyridine, overnight; (c) 3HF-TEA, THF, overnight; (d) DIPEA, CH_2Cl_2, $iPr_2NPCl(OCH_2CH_2CN)$, 0 °C; (e) DNA synthesizer; 28 %, NH_4OH, 55 °C, overnight, then HPLC; 5 % AcOH, 30 min, then HPLC.

Scheme 1 Synthesis of TFO incorporated 3-halo AP-ΨdC derivatives

Table 1 Association constant of the synthesized TFO

TFO: 3′ –GGAAGG **NZN′** GAGGAGGGA
Duplex DNA: 5′ –GAGGGAAGG **NXN′** GAGGAGGGAAGC
3′ –CTCCCTTCC **MYM′** CTCCTCCCTTCG–FAM
Conditions : 20 mM Tris-HCl buffer, 20 mM $MgCl_2$, pH 7.5, 37 °C
K_s (10^{-6} M^{-1}) = [Triplex] / ([TFO][Duplex])

3′NZN5′	Z	K_s [x 10^6 M^{-1}] for XY			
		AT	TA	GC	CG
3′AZA5′	^{Me}AP-ΨdC	n.d.	n.d.	0.2 ± 0.1	20.8 ± 0.9
	^{Cl}AP-ΨdC	n.d.	n.d.	n.d.	n.d.
	^{Br}AP-ΨdC	n.d.	n.d.	n.d.	n.d.
	^{I}AP-ΨdC	n.d.	n.d.	n.d.	n.d.
3′AZG5′	^{Me}AP-ΨdC	n.d.	0.2 ± 0.1	1.8 ± 0.6	19.4 ± 1.8
	^{Cl}AP-ΨdC	1.4 ± 0.1	n.d.	1.0 ± 0.3	3.1 ± 0.2
	^{Br}AP-ΨdC	n.d.	n.d.	n.d.	2.8 ± 0.4
	^{I}AP-ΨdC	n.d.	n.d.	0.11 ± 0.06	1.6 ± 0.6
3′GZA5′	^{Me}AP-ΨdC	n.d.	n.d.	1.8 ± 0.5	32.6 ± 0.5
	^{Cl}AP-ΨdC	n.d.	n.d.	0.4 ± 0.1	10.4 ± 0.6
	^{Br}AP-ΨdC	n.d.	n.d.	0.1 ± 0.1	8.9 ± 1.2
	^{I}AP-ΨdC	n.d.	n.d.	0.4 ± 0.1	18.9 ± 1.8
3′GZG5′	^{Me}AP-ΨdC	0.8 ± 0.1	2.6 ± 0.5	5.3 ± 0.4	16.6 ± 0.5
	^{Cl}AP-ΨdC	2.2 ± 0.1	6.0 ± 1.4	4.3 ± 0.6	6.5 ± 0.6
	^{Br}AP-ΨdC	0.6 ± 0.2	3.9 ± 0.4	3.1 ± 1.1	5.9 ± 0.2
	^{I}AP-ΨdC	0.3 ± 0.1	4.9 ± 0.1	4.6 ± 0.4	6.6 ± 0.3

② Newly designed ΨdC derivatives: bicyclic restricted structure and 4-electron-donating group introduction:

According to the results of last section, I have confirmed the hypothesis of the contribution of the protonated 1-*N* position to triplex formation with the AP-ΨdC derivatives. Therefore, this study was initiated to improve the hydrogen binding affinity, by increasing the pK_a of the 1-*N* position of aminopyridine within AP-ΨdC derivatives. According to the molecular design concept, we devised two types of new AP-ΨdC derivatives to confirm the hydrogen bonding effect of the 1-*N* position and substituent effect of the aminopyridine unit. One series includes bicyclic compounds; i.e. 2,3-dihydro-7-azaindole (DHAID-ΨdC), 1,2,3,4-tetrahydro-1,8-naphthyridine (THNTD-ΨdC) and 7-azaindole (AID-ΨdC) (**Fig. 4**). In another series, the methyl or methoxy group was introduced at the 4-position of the 2-aminopyridine unit, 2-amino-4-methylpyridinyl-ΨdC (^{4Me}AP-ΨdC) and 2-amino-4-methoxypyridinyl-ΨdC (^{4OMe}AP-ΨdC) (**Fig. 4**), respectively.[4] The synthesis of the corresponding phosphoramidite compounds of the newly ΨdC derivatives is shown in **Scheme 2**. Finally, those corresponding phosphoramidite compounds were incorporated in to the middle of the 18 mer TFOs using an automated DNA synthesizer. The synthesis and purification of TFO incorporating the newly ΨdC derivatives were carried out in the

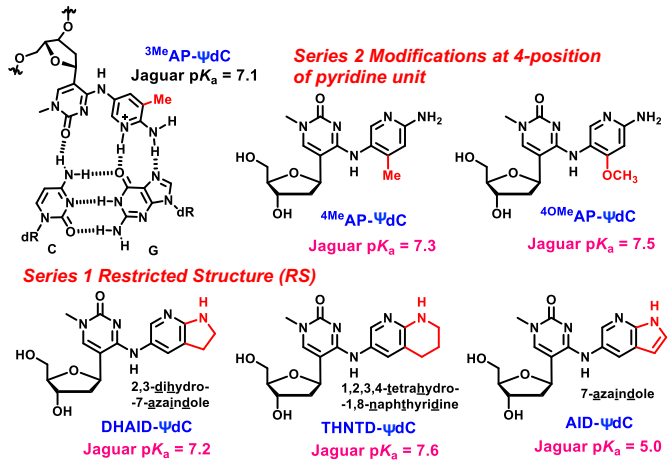
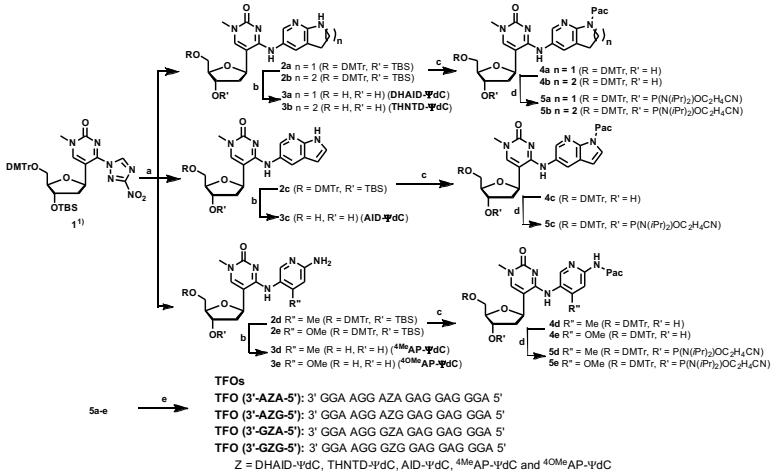


Figure 4 Design of pseudo-dC derivatives to enhance the recognition effect for CG base pair

same manner as described above.



Reagents and conditions: (a) the corresponding aromatic amines monomers, DIPEA, MeCN, reflux; (b) 0.5 M HCl in MeOH, r.t., overnight; (c) i) Pac_2O, pyridine, rt, overnight; ii) 3HF-TEA, Et_3N, THF, rt, overnight; (d) DIPEA, CH_2Cl_2, $iPr_2NPCl(OCH_2CH_2CN)$, 0 ºC; (e) DNA synthesizer and HPLC purification.

Scheme 2 Synthesis of TFO incorporated newly designed AP-ΨdC derivatives

The triplex-forming ability of the synthesized TFOs was evaluated by gel-electrophoretic mobility shift assays using FAM-labeled duplex DNAs of an antiparallel DNA type. The equilibrium association constants (K_s) were also obtained and summarized in **Fig. 5**. In the case of TFOs incorporating the DHAID-ΨdC, THNTD-ΨdC and AID-ΨdC derivatives, which were designed as bicyclic structures, THNTD-ΨdC showed the highest affinity for the CG base pair in all four 3′-NZN'-5′ combinations. In particular, THNTD-ΨdC exhibited a relatively high affinity for the CG base pair in the 3′-AZA-5′ and 3′-GZG-5′ sequences compared to the original ^{3Me}AP-ΨdC, whereas DHAID-ΨdC showed only a weak interaction with the CG base pair. In the 3'-AZG 5'or 3'-GZA-5' sequence, in which ^{3Me}AP-ΨdC selectively stabilized triplexes with the CG base pair, both THNTD-ΨdC and DHAID-ΨdC could not recognize the CG base pair with a sufficient stability compared to the ^{3Me}AP-ΨdC derivative. The ^{4Me}AP-ΨdC TFO formed a stable triplex only when it was incorporated into the 3'-GZA-5' sequence. In remarkable contrast, TFOs containing the ^{4OMe}AP-ΨdC derivative demonstrated triplex-forming ability and selectivity with the CG base pair in all sequences.

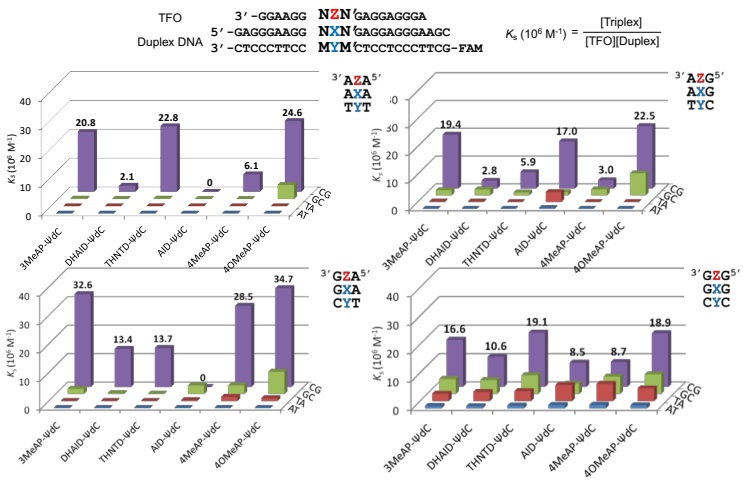


Figure 5 Association constant of the synthesized TFO incorporating newly designed derivatives

Taking into consideration that the CG inversion sites in the promoter regions of the human gene are multiple and consecutive, the recognition ability of TFOs containing AP-ΨdC derivatives for multiple CG base pairs was further demonstrated by triplex formation against human hTERT and Cyclin D1 gene (**Fig. 6A** and **6B**, respectively) sequences. The triplex-forming ability was evaluated by a gel shift assay, and the gel results and K_s values of the targeting for the promoter region of the hTERT and Cyclin D1 gene are depicted in **Fig. 6C** and **6D**, respectively. In the case of targeting the hTERT sequence, the TFO containing THNTD-ΨdC showed a moderate triplex-forming ability compared to ^{3Me}AP-ΨdC. In contrast, the TFO containing ^{4OMe}AP-ΨdC could form a stable triplex DNA with high affinity even at a low concentration compared to the original ^{3Me}AP-ΨdC. On the other hand, for the target Cyclin D1 sequence with CG inversion sites, the TFO containing ^{4OMe}AP-ΨdC derivative showed a moderate triplex-forming ability compared to ^{3Me}AP-ΨdC.

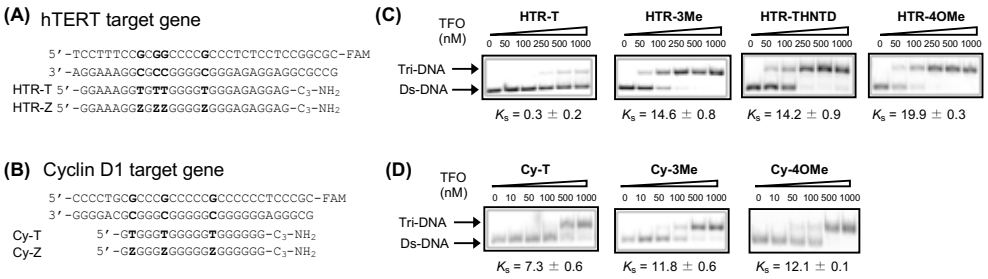


Figure 6 Evaluation of triplex forming ability of TFOs containing modified AP-ΨdC derivatives against multiple CG base pairs.

③ Antigene application to efficient inhibition of hTERT gene:

Formation of triplex DNA to the gene promoter region inhibits the interaction between transcription factor and dsDNA, leading to gene expression inhibition. Therefore, gene expression inhibition experiments on hTERT gene were carried out using newly designed dC derivative capable of selectively and stably recognizing CG inversion sites. In order to protect the TFO from digestion by nuclease in the cytoplasm, the Guard-duplex protection strategy was performed. It is reported that the high-ordered structure of oligo cannot be digested by nuclease in the cytoplasm rapidly. According to this phenomenon, the Guard-oligo containing different number and position of dU base was performed to protect TFO from digestion by nuclease in the form of Guard-duplex. Once the Guard-duplex reaches the nucleus, the nuclear repair enzyme, UDG, could recognize and excise them to release the single-stranded TFO, followed by binding target gene to control the gene expression (**Fig. 7**). In order to verify the feasibility of this strategy, the nuclease resistance assay of single-stranded TFO and Guard-duplex, TFO release assay from Guard-duplex with the treatment of UDG *in vitro*, fluorescence microscopy measurement using Hela cells, hTERT gene expression inhibition assay by RT-PCR and Hela cell proliferation inhibition assay by MTS method were performed. The detail results were shown in the presentation.

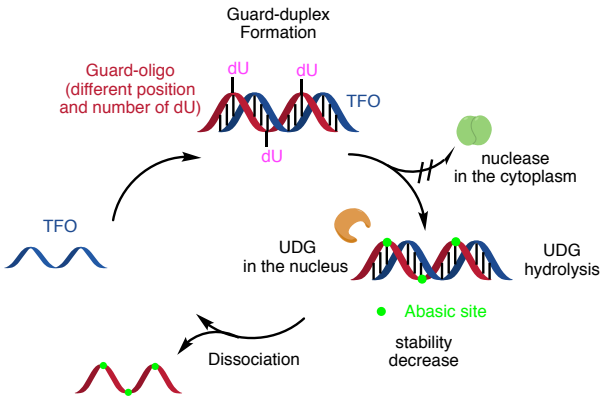


Figure 7 New strategy for the Guard-oligo protection and dissociation

[Conclusion]

In this study, we succeeded in developing 3-halogenated AP-ΨdC derivatives to confirm the contribution of protonation of 1-*N* of pyridine unit to the stable triplex DNA formation, and developing bicyclic restricted structure and 4-electron-donating group introduced derivatives capable of selectively recognizing CG base pair. In particular, CG base pair recognition by ^{4OMe}AP is not affected by adjacent bases, and it can recognize the multiple CG inversion sites in the promoter region of the hTERT gene. Furthermore, the Guard-duplex protection strategy was performed to protect TFO from digestion by nuclease. Efficient inhibition of hTERT gene was achieved using this strategy compared with the treatment of TFO only according to the results of RT-PCR. We hope for future development as a nucleic acid medicine.

[Reference]

1) A. Mukherjee and K. M. Vasquez, *Biochimie*, **2011**, *93*, 1197-1208.
2) H. Okamura, Y. Taniguchi and S. Sasaki, *Angew. Chem. Int. Ed.,* **2016**, *55*, 12445-12449.
3) L. Wang, Y. Taniguchi, H. Okamura and S. Sasaki, *Bioorg. Med. Chem.*, **2017**, *25*, 3853-3860.
4) L. Wang, Y. Taniguchi, H. Okamura and S. Sasaki, *Nucleic Acids Res.*, **2018**, *46*, 8679-8688.