EurJOC European Journal of Organic Chemistry

MICROREVIEW

DOI: 10.1002/ejoc.201101821

Towards the Sequence-Selective Recognition of Double-Stranded DNA Containing Pyrimidine-Purine Interruptions by Triplex-Forming Oligonucleotides

Yoshiyuki Hari,*[a] Satoshi Obika,[a] and Takeshi Imanishi*[a,b]

Keywords: DNA / DNA recognition / Nucleobases / Nucleotides / Oligonucleotides / Carbohydrates / Triple helix

Triplex formation with double-stranded DNA (dsDNA) by oligonucleotides has potential for applications in attractive technologies such as gene therapy and genetic diagnosis. However, triplex-forming oligonucleotides (TFOs) can only recognize homopurine strands in homopurine-homopyrimidine regions in dsDNA, either through Hoogsteen or through reverse-Hoogsteen hydrogen bonds. A straightforward and powerful approach to overcoming this sequence limitation is the development of artificial nucleic acids capable of recognizing specific pyrimidine-purine interruptions (i.e., a CG or TA base pair) in triplex formation. This review describes artificial nucleic acids, especially those containing non-natural nucleobases, developed to recognize CG or TA base pairs in dsDNA targets.

Introduction

Triplex formation between double-stranded DNA (dsDNA) and an oligonucleotide, called a triplex-forming oligonucleotide (TFO), is an attractive principle applicable to various nucleic acid technologies.[1] It would, for example, allow the down/up-regulation of the transcription of target genes and the induction of mutations of target genes for in vivo applications. Many in vitro applications of TFOs have already been performed.

The recognition of dsDNA by TFOs to form triplexes is classified into two types: parallel and antiparallel motifs (Figure 1). In a parallel motif, a pyrimidine-rich TFO binds to the purine strand of dsDNA to form T-AT and protonated C (C^+H)-GC triplets through Hoogsteen hydrogen bonds. G can also form a G-GC triplet with a GC base pair, although the stability strongly depends on the sequence.[2] In an antiparallel motif, A (or T) and G in a TFO recognize AT and GC base pairs, respectively, in dsDNA through reverse-Hoogsteen hydrogen bonds. Although triplexes in both motifs are constructed through the formation of strong and sequence-selective hydrogen bonds, there are several inherent problems of triplex formation with TFOs consisting of natural nucleic acids.

The first problem is the low stabilities of the resulting triplexes under physiological conditions. In a parallel motif, for instance, because C in a TFO has to be protonated to have affinity towards a GC base pair, stable triplex formation requires a low pH. To overcome this problem, nucleobases that serve as substitutes for C have been developed. In particular, 5-methylcytosine (^{Me}C),[3] 2-aminopyridine (P)[4] and 2-amino-3-methylpyridine (^{Me}P)[4] nucleobases (Figure 2), which enable the stable recognition of a GC base pair at neutral pH, are widely used. In addition, intercalators or sugar modifications of nucleotides are also effective in increasing the stability of parallel triplex DNA.[5]

In an antiparallel motif, depending on the situation, a G forming a base triplet with a GC base pair easily causes aggregation in the presence of a physiological amount of monovalent cations, so the development of G analogues to prevent the formation of aggregates such as G-quartets has also been studied.[6] It has also been shown that insertion of the bulge pyrenylethynylphenyl moieties (twisted intercalating nucleic acids, TINAs) centrally in G-runs of TFOs disrupts formation of G aggregates and promotes the formation of stable antiparallel triplexes at physiological pH values and monovalent ion concentrations.[7]

The second problem is the limitations of the target sequence of dsDNA. In each motif, recognition by the TFO is generally restricted to only the homopurine region in the dsDNA, as shown in Figure 1. This sequence limitation is considered fatal from the point of view of practical use in various applications based on triplex formation. A straightforward and powerful approach towards overcoming this limitation is the development of nucleobases capable of recognizing CG or TA base pairs in a sequence-selective manner. Indeed, some natural nucleobases can interact with CG or TA base pairs. In a parallel motif triplex, T is known to interact with a CG base pair and also forms a stable T-AT

[a] Graduate School of Pharmaceutical Sciences, Osaka University, 1-6 Yamadaoka, Suita, Osaka 565-0871, Japan
[b] BNA Inc., Saito Bio-Innovation Center, 7-7-20 Saito-Asagi, Ibaraki, Osaka 567-0085, Japan
Fax: +81-6-6879-8201 (to Y. H.)
+81-72-646-5763 (to T. I.)
E-mail: hari@phs.osaka-u.ac.jp
imanishi@bna.jp
imanishi@phs.osaka-u.ac.jp

triplet.[8] A single hydrogen bond between the 2-carbonyl group of T and the 4-amino group of C is formed, as shown in Figure 3 (a).[9] G also selectively recognizes a TA base pair,[8,10] and the structure of a G-TA base triplet, as shown in Figure 3 (b), has been elucidated by NMR studies of an intramolecular triplex.[11] Any positive interaction between G and the 5-methyl group of T, such as shape complementarity, can also be considered, because a G-UA triplet is not stable.[12] In an antiparallel triplex, T has a weak interaction with a CG base pair,[13,14] based on the formation of a hydrogen bond between the 4-carbonyl group of T and the 4-amino group of C (Figure 3, c).[15] On the subject of the recognition of pyrimidine-purine interruptions, a large number of nucleobases derived from these natural base triplets have been developed. An alternative approach based on use of molecular modelling to design nucleobases containing hydrogen donors and/or acceptors at suitable positions for hydrogen bond formation with interruptions has also been pursued (Figure 4). Furthermore, the problem of the quite slow nature of triplex formation has also been addressed; the use of TFOs bearing positively charged moieties to overcome this problem has been investigated. Many reviews on TFOs including artificial nucleosides to form triplex DNA have been published.[5,16] In this microreview we focus on the artificial nucleobases so far developed that directly recognize CG or TA base pairs in triplex formation.

Approaches Based on a Parallel Motif

In 1992, Dervan's group developed 4-(3-benzamidophenyl)imidazole (D3, Figure 5) as a nucleobase designed to recognize a CG base pair through two hydrogen bonds.[17,18] This approach pioneered the use of artificial nucleobases for pyrimidine-purine recognition. However, D3 showed affinity (by affinity cleavage assay) not only towards a CG base pair but also towards a TA base pair. From ^{1}H NMR studies the recognition of pyrimidine-purine base pairs by D3 was found to be the result of intercalation by D3 at the 3′-site of the target base pair.[19]

The capabilities of carbocyclic nucleosides L1 or L2 (Figure 5) as partners for a CG base pair in triplex formation were also evaluated.[20] Through quantitative DNase I footprinting analysis, it was concluded that L1 and L2, like D3, acted as intercalators in order to show affinities towards pyrimidine-purine base pairs. In addition, imidazolo[4,5-*h*]quinoline nucleobases (e.g., ^{NH_2}Q, Figure 5) or the debutylated analogue (Z) of the nucleobase developed by Zimmerman's group[21] as a receptor for a CG base pair at the nucleoside level showed low affinities with slight preferences for a GC base pair.[20]

With the aid of ^{1}H NMR studies in organic solvent with nucleoside monomers,[22] Zimmerman's group designed a [6-(3-ethylureido)-1-oxoisoindolin-2-yl]methyl nucleobase for a CG base pair.[23] However, UV melting experiments indicated that the nucleobase showed little affinity for any base pairs.

Weisz's group found by NMR measurements in dichloromethane at the nucleoside monomer level that 4-(3-ureidophenyl)imidazole, with a similar structure to D3, was a key receptor for a CG base pair.[24] 4-[3-(3-Butylureido)phenyl]-imidazol-1-yl nucleobase (D4, Figure 6) was tested in TFOs for affinity and selectivity for base pairs by UV melting ex-

Yoshiyuki Hari was born in Tokushima, Japan, in 1974. He received his M.Sc. degree from Osaka University in 1999 and his Ph.D. degree from Osaka University in 2002, supervised by Professor Takeshi Imanishi. He worked as an assistant professor at Nagoya City University from 2002 to 2008. During that period he visited the group of Professor Floyd E. Romesberg at the Scripps Research Institute as a visiting scientist for one year. He was promoted to lecturer at Nagoya City University in 2008. He moved to Osaka University as an associate professor in 2009. His research interests include the development of functional nucleic acids for nucleic-acid-based technologies and the development of synthetic methods for heterocycles.

Satoshi Obika is Professor at the Graduate School of Pharmaceutical Sciences, Osaka University. He was born in 1967 and graduated from Osaka University in 1990. After obtaining his M.Sc. degree in 1992 under the supervision of Professor Takeshi Imanishi, he began his career as an Assistant Professor at Osaka University. In 1998 he earned his Ph.D. degree from Osaka University. After working as a visiting researcher with Professor Thomas C. Bruice at the University of California at Santa Barbara for one year from 2002, he returned to Osaka University and became the PRESTO researcher at the Japan Science and Technology Agency (JST). He was promoted to Associate Professor in 2006 and full Professor in 2008. He received the Pharmaceutical Society of Japan Kinki Branch Award for Young Scientists in 1999 and the Pharmaceutical Society of Japan Award for Young Scientists in 2004. His research interests are in the areas of bioorganic chemistry and nucleic acid chemistry.

Takeshi Imanishi was born in Osaka, Japan in 1944. He studied synthetic organic chemistry at Osaka University (1963–1972), obtaining his Ph.D. degree in pharmaceutical chemistry from Osaka University in 1972, spent some years at the Faculty of Pharmaceutical Sciences, Osaka University, as Research Associate, and then moved to Kanazawa University as Associate Professor at the Faculty of Pharmaceutical Sciences (1978–1982). He returned to Osaka University in 1982 and was promoted to Full Professor of Organic & Bioorganic Chemistry at the Faculty & Graduate School of Pharmaceutical Sciences in 1990. He received the Pharmaceutical Society of Japan Award for Young Scientists in 1983 and the First Prize for The Third Bio Business Competition Japan in 2004. His research interests include precise synthetic organic, bioorganic, and nucleic acid chemistry. In 2008 he retired and now is a Professor Emeritus of Osaka University, and President & CEO of BNA Inc.

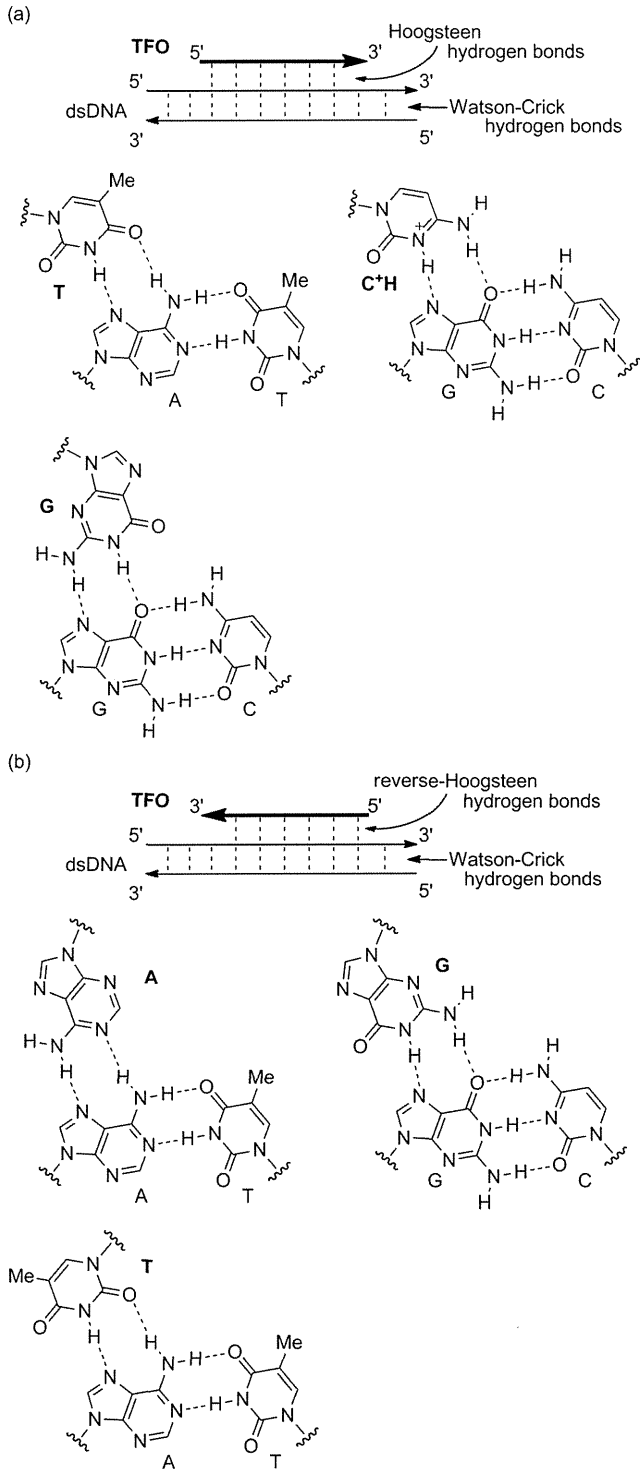


Figure 1. Structures of canonical base triplets and orientations of individual strands in triplex DNA. a) Parallel motif triplex. b) Antiparallel motif triplex. Abbreviations of nucleobases in TFOs are shown in bold.

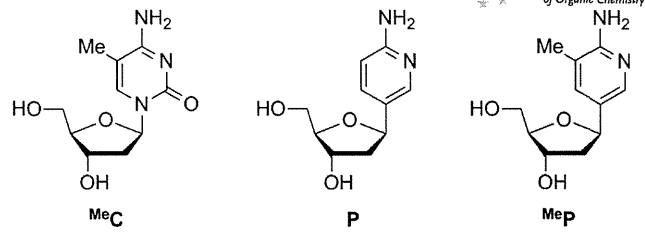


Figure 2. Structures of nucleosides containing 5-methylcytosine (^{Me}C), 2-aminopyridine (P) and 2-amino-3-methylpyridine (^{Me}P) nucleobases.

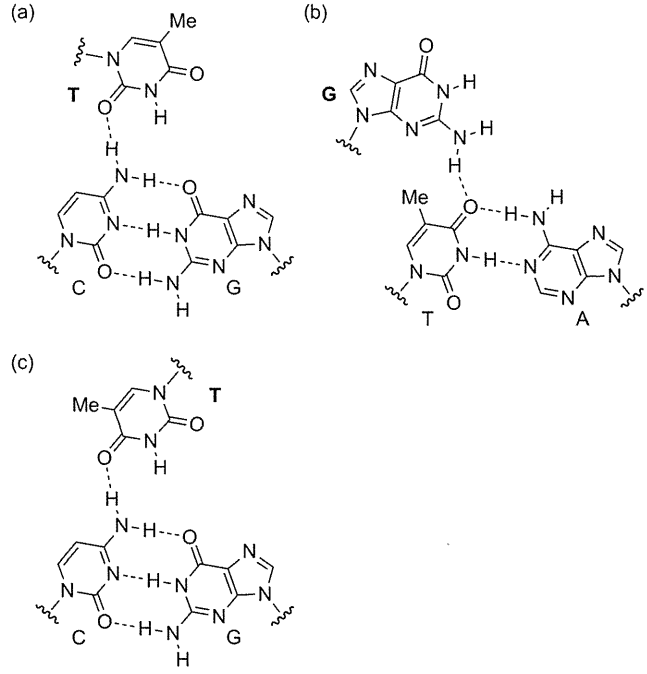


Figure 3. Structures of non-canonical base triplets. a) T-CG triplet in parallel motif triplex. b) G-TA triplet in parallel motif triplex. c) T-CG triplet in antiparallel motif triplex. Abbreviations of nucleobases in TFOs are shown in bold.

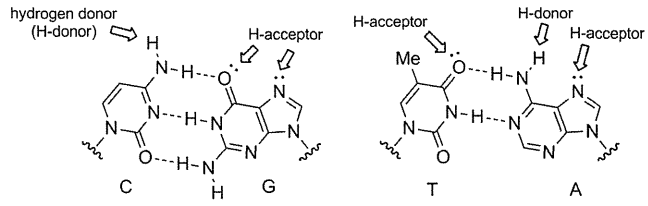


Figure 4. Recognition sites based on hydrogen bonds in pyrimidine-purine base pairs.

periments.[25] Although D4 had a strong preference for AT and GC base pairs over a CG base pair, a triplex containing a D4-CG triplet exhibited a melting temperature (T_m) value approximately 3 °C higher than that of a T-CG triplet at physiological pH. The *N*-alkylated urocanamide nucleoside (U2) was also explored for CG base pair recognition as a result[26] of NMR experiments performed with nucleoside monomers (Figure 6).[27] However, the TFO including U2 had only a weak affinity for dsDNA with a CG base pair.

The 2′,4′-BNA/LNA modification (Figure 7), independently developed by Wengel's group[28] and by ourselves,[29] in a TFO stabilizes the canonical triplex with dsDNA without loss of sequence-selectivity.[30] We thus evaluated a combination of 2′,4′-BNA/LNA and five-membered heteroaro-

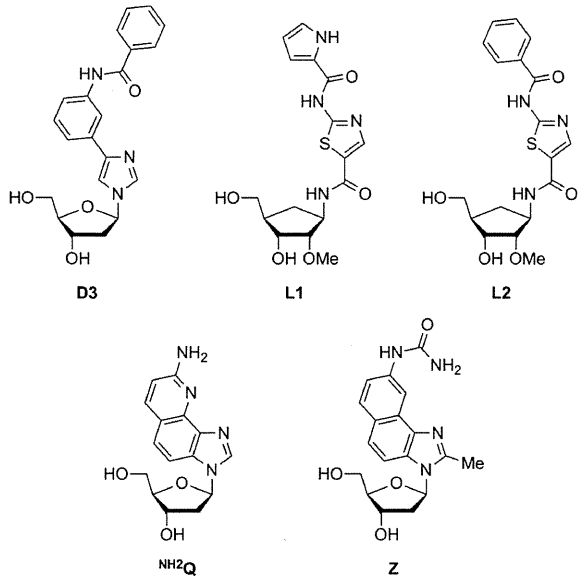


Figure 5. Structures of nucleosides containing artificial nucleobases.

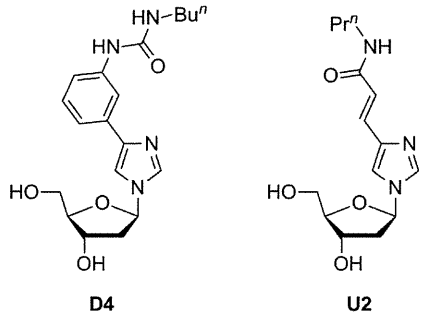


Figure 6. Structures of nucleosides containing artificial nucleobases.

matic nucleobases such as oxazoles and imidazoles in the recognition of pyrimidine-purine interruption in dsDNA.[31,32] TFOs incorporating 2′,4′-BNA/LNAs with 2-phenyloxazol-5-yl (pO^B) or imidazol-1-yl (I^B) nucleobases showed no stabilization of the triplexes formed with dsDNA containing a CG base pair whereas a 2′,4′-BNA/LNA with the oxazol-5-yl nucleobase (O^B) did show an apparent interaction with a CG base pair (Figure 7, a). In general, a heteroatom at the α-position – but not the β-position – of the glycosyl bond in a five-membered nucleobase might play a role in the recognition of the CG base pair (Figure 8). It was reported by Dervan's group that a 4-phenylimidazol-1-yl nucleobase had weak affinity for all base pairs, which is comparable to what has been observed for natural mismatch base triplets.[17] This might also support our idea shown in Figure 8. In contrast, the affinity and selectivity of a 2′,4′-BNA/LNA with 2-(*N*-methylbenzamido)thiazole (Tz^B) towards a TA base pair were found to be significantly superior to those observed with a G-TA triplet (Figure 7, a).[33] Examination of the effect on the adjoining triplets suggested that the recognition of a TA base pair by Tz^B did not seem to be based on intercalation, but further studies will be required to clarify the mode of recognition. Recently, we also screened 18 variants of 1-substituted 1,2,3-triazole nucleobases, easily obtainable through Cu^I-catalysed alkyne/azide cycloaddition reactions between an oligonucleotide bearing an ethynyl unit and azide compounds.[34] The triazole-based nucleobases generally appeared to have affinities in the order CG > TA > AT ≈ GC base pairs at neutral pH. In particular, a 1-(4-ureidophenyl)triazole nucleobase showed high CG-selectivity with stability similar to that of T with a CG base pair, because the 4-ureido group contributed to a significant decrease in binding affinity with a TA base pair (Figure 7, b). To elucidate the mode of recognition of a CG base pair by the triazole moiety and the role of the ureido group, further detailed investigations including NMR studies will be needed.

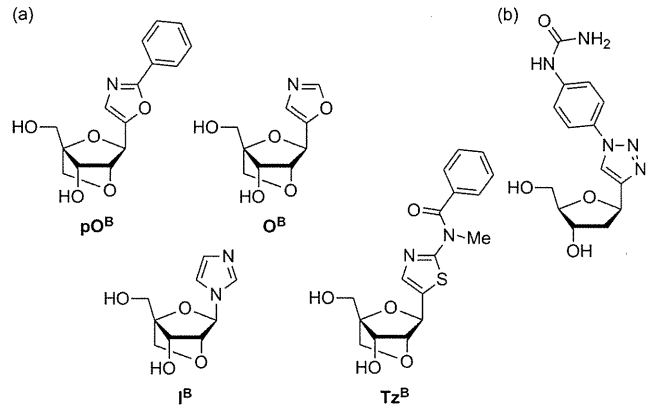


Figure 7. Structures of a) 2′,4′-BNA/LNA monomers, and b) a nucleoside with an artificial nucleobase.

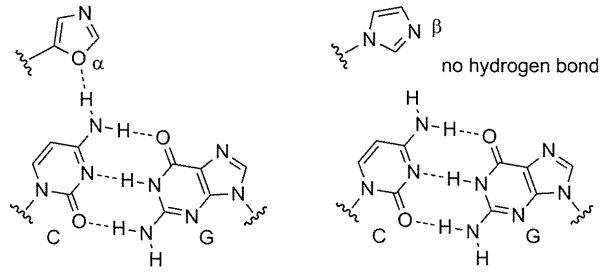


Figure 8. Proposed schemes for recognition of a CG base pair by oxazol-5-yl and imidazol-1-yl nucleobases as examples.

Designs of nucleobases based on the T-CG base triplet[8] have also been developed. The first examples are the N^4-modified cytosines reported by Miller's group in 1993 (Figure 9, a).[35] N^4-(3-Acetamidopropyl)cytosine (R = NHCOMe), for example, could interact selectively with a CG base pair in the triplex DNA at pH 7.0. The T_m value and ΔT_m, the difference between the T_m value with a CG base pair and those with other base pairs, were 19 °C and ≥11 °C, respectively, whereas a canonical triplex showed a T_m value of 32 °C. The synthesis was very distinctive and was achieved by transamination of cytosine in the TFO with propane-1,3-diamine in the presence of sodium bisulfite followed by acetylation. N^4-(6-Amino-2-pyridinyl)-

cytosine nucleobase was also developed; its affinity towards a CG base pair in UV melting experiments was comparable to that observed in a canonical C^+H-GC triplet, although the nucleobase also showed high affinity towards an AT base pair.[36,37] The recognition seemed to involve an unusual imino tautomerization of the nucleobase as shown in Figure 10. The 6-amino group was shown to play an essential role in the interaction with a CG base pair by comparison with the result obtained with an analogue without the amino group, which had no affinity towards a CG base pair. Moreover, it was reported by Dervan's group that the 6-benzamido derivative showed CG- and TA-selectivity as an intercalator.[20] To target an TA base pair, Miller's group also designed cytosine derivatives in which carboxylic acid groups were predicted by molecular modelling to interact with the 6-amino group in A of the TA base pair (Figure 9, b).[38] UV melting experiments showed that N^4-(3-carboxypropynyl)cytosine nucleobase in a 15-mer TFO selectively recognized a TA base pair of dsDNA (15 bps), with a T_m value (Tris buffer, pH 7.0) 7 °C higher than that of a triplex containing the corresponding G-TA triplet. The nucleobase did not seem to interact with a UA base pair at all. In contrast, when MOPS buffer (pH 7.0) was employed, high GC-selectivity was observed. These results suggested that formation of an ion pair between the carboxylic acid group of the nucleobase and a Tris cation was necessary to induce the TA-selectivity. N^4-(3-Carboxytriazolyl)cytosine nucleobase had affinity in a gel shift assay towards TA, UA and CG base pairs in addition to a GC base pair.

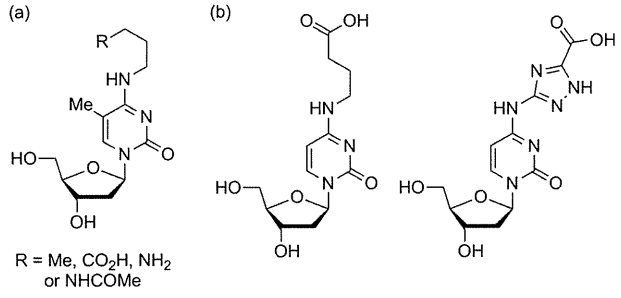


Figure 9. Structures of nucleosides contining modified cytosine nucleobases.

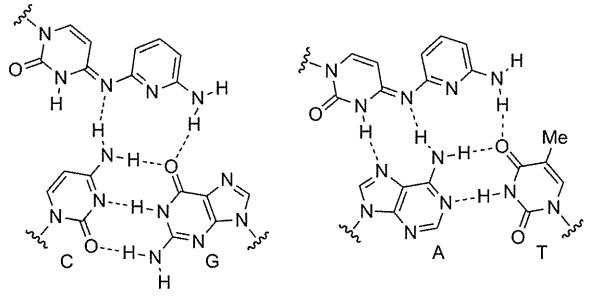


Figure 10. Proposed schemes for recognition of CG and AT base pairs by N^4-(6-amino-2-pyridinyl)cytosine nucleobase.

The modified cytosine nucleobases shown in parts a and b of Figure 11 were developed for a CG base pair by Guzzo-Pernell's group[39] and by McLaughlin's group,[40] respectively. The semicarbazide derivative (R = H) had moderate CG-selectivity in triplex DNA at pH 7.0 (ΔT_m = 3 °C), although the stability was less than that of the ^{Me}C-CG triplet (Figure 11, b).

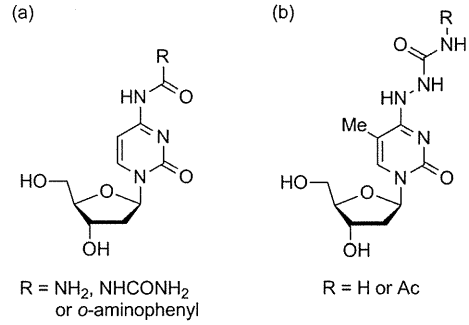


Figure 11. Structures of nucleosides containing modified cytosine nucleobases.

Leumann's group designed the simple 5-methylpyrimidinone nucleobase ^{4H}T (Figure 12), lacking the 4-amino group of ^{Me}C, for a CG base pair.[41] In UV melting experiments at pH 6.5, the 15-mer TFO containing a single ^{4H}T component with a 2′-deoxyribose sugar unit was found to form the most stable triplex with dsDNA containing a CG base pair, with a ΔT_m value of more than 4 °C, although the stability of the ^{4H}T-CG triplet (T_m = 28.3 °C) was almost the same as that of ^{Me}C-CG (T_m = 27.1 °C). Replacement of the surrounding ^{Me}C by P (Figure 2) for a GC base pair led to an increase in stability and selectivity of ^{4H}T towards a CG base pair. A proposed binding scheme of ^{4H}T with a CG base pair with an additional weak C–H···O interaction[42] is shown in Figure 13. A combination of the ^{4H}T nucleobase and a 2′-*O*-aminoethyl sugar modification (2′-AE)[43] – 2′-AE-^{4H}T (Figure 12) – increased the stability of the resulting triplex without loss of sequence selectiv-

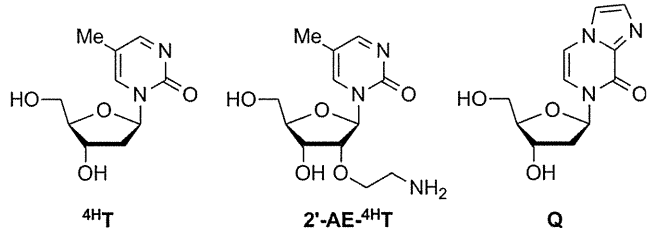


Figure 12. Structures of nucleosides and of a 2′-AE monomer combined with an artificial nucleobase.

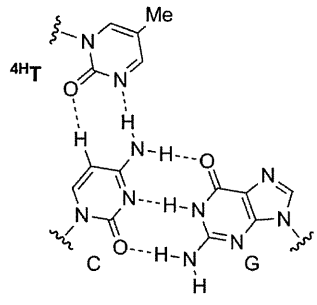
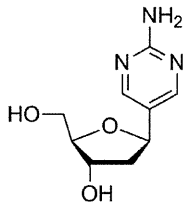


Figure 13. Proposed recognition scheme for a ^{4H}T-CG triplet.

ity.[43–46] All 2′-AE-modified TFOs (15-mers) containing five ^{4H}T nucleobases could stably form triplexes with the corresponding dsDNA containing five CG base pairs.[44] The Q nucleobase shown in Figure 12 was also designed and evaluated.[47] It recognized a GC base pair rather than a CG base pair in triplex DNA.

McLaughlin's group compared the affinities of P (shown in Figure 2) and of 2-aminopyrimidine (2APm, Figure 14) towards CG and GC base pairs with that of C by UV melting experiments.[48] Of these, 2APm formed the most stable triplet with a CG base pair at pH 7.0. In addition, 2APm led to a drastic destabilization (relative to C or P) of the triplex formed with dsDNA containing a GC base pair, probably due to the prevention of protonation of 2APm by its low pK_a (pK_a = 3.3).



2APm

Figure 14. Structure of 2APm.

We designed a simple 2-pyridone (PD) as a nucleobase without the functional groups either of T or of C to recognize an AT or a GC base pair, respectively.[49–51] In UV melting experiments at physiological pH, the affinity and selectivity of PD towards a CG base pair appeared to be comparable with those of ^{4H}T, although the CG recognition scheme for PD (Figure 15) was different from the ^{4H}T-CG triplet as shown in Figure 13. Moreover, a 2-pyridone moiety attached to a 2′,4′-BNA/LNA (P^B, Figure 16) led to an increased T_m value of 9 °C without loss in CG selectivity ($\Delta T_m \geq 9$ °C).[49] The presence of the 5-methyl congener ($^{Me}P^B$, Figure 16) in a TFO gave a triplex formation result similar to that seen with P^B.[50] With use of a 15-mer TFO containing four 2′,4′-BNA/LNA components bearing natural nucleobases and three P^B units, the efficient formation of a triplex containing three P^B-CG triplets with the target dsDNA was observed under physiological conditions. However, by comparison with 2′,4′-BNA/LNA without a nucleobase (H^B, Figure 16), it was demonstrated that a 2-pyridone nucleobase should be capable of recognition both of a CG base pair and of an AT base pair. A 2′,4′-BNA/LNA moiety bearing a 1-isoquinolone nucleobase (iQ^B, Figure 16) was thus designed to produce steric repulsion against an AT base pair and the base pair recognition capability of iQ^B in a TFO was investigated (Figure 17).[52,53] As expected, results of T_m measurements with a 15-mer TFO showed that the affinity of iQ^B towards an AT base pair was greatly decreased – by 8 °C relative to that of P^B and almost the same as that of H^B.[52] The stability of the iQ^B-CG triplet, as measured by the T_m value, however, was 4 °C less than that of the P^B-CG triplet, probably due to steric hindrance by the benzene ring in the 1-isoquinolone nucleobase. On the other hand, a pyridone derivative containing three aromatic rings (bP^B, Figure 16) displayed a slight TA-selectivity and the bP^B-TA triplet was significantly more stable than a G-TA triplet.[33] Examination of the neighbouring group effect suggested that bP^B acts, like D3, as an intercalator to produce affinity towards a TA base pair. We have recently developed a 2′,4′-BNA/LNA system bearing a simple pyridine nucleobase (Py^B, Figure 16) for a CG base pair.[54] The ability of Py^B to recognise a CG base pair in triplex formation was the same as or greater than that of P^B. A pyridine moiety as the core structure of a nucleobase for a CG base pair might thus be attractive for allowing further modifications to improve the recognition ability.

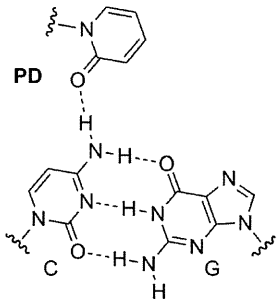


Figure 15. Proposed recognition scheme for a CG base pair by PD.

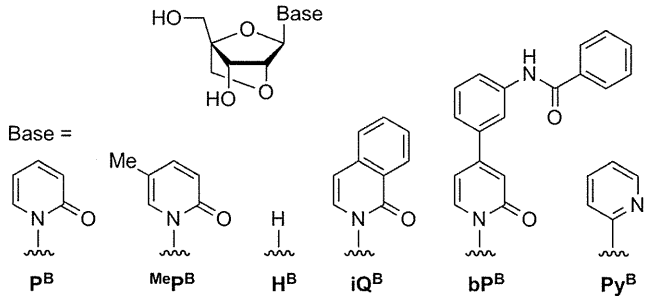


Figure 16. Structures of 2′,4′-BNA/LNA monomers bearing (or not bearing) artificial nucleobases.

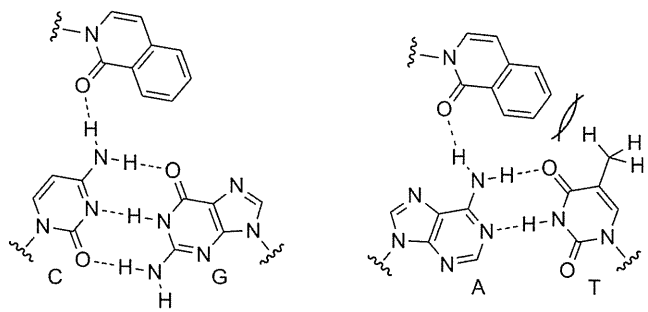


Figure 17. Design concept behind a 1-isoquinolone nucleobase.

Brown's group developed *N*-methylpyrrolo[2,3-*d*]pyrimidin-2(7*H*)-one, which possesses functional groups of ^{4H}T for CG base pair recognition, as a core nucleobase structure (Figure 18).[55–57] It is also easy to introduce further functionality on the 6-carbon of the core to improve the recognition ability. Evaluation of ^{M}P, ^{A}EP and ^{A}PP (Figure 18, a) was performed by means of T_m experiments and

quantitative DNase I footprinting analysis.[55] Holistic interpretation of the results indicated that the order of affinity for a CG base pair was ^{A}PP ≈ ^{M}P > ^{A}EP ≈ T and that these artificial nucleobases had CG-selectivity. This also implies that the 3-aminopropyl side chain at the 6-position is better than the 2-aminoethyl side chain for affinity to a CG base pair.

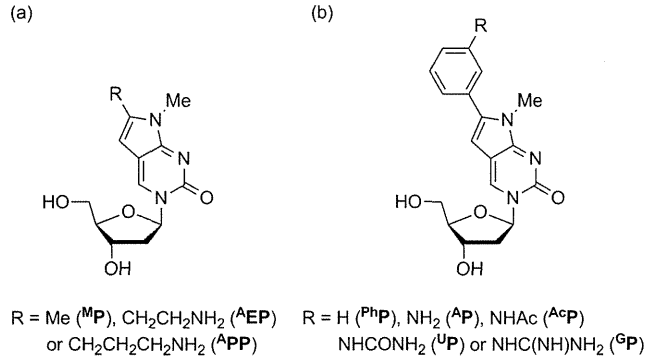


Figure 18. Structures of nucleosides containing 6-substituted *N*-methylpyrrolo[2,3-*d*]pyrimidin-2(7*H*)-one nucleobases.

^{Ph}P, ^{A}P, ^{Ac}P, ^{U}P and ^{G}P (Figure 18, b) were further designed as ring-fused analogues of ^{A}PP with potential for additional stacking interactions.[56,57] However, their recognition properties in triplex formation seemed to depend on the TFO sequence [e.g., a sequence rich in 5-(3-aminoprop-1-ynyl)-dU (pdU, Figure 19)[58] in place of T as a partner to an AT base pair or a sequence with few pdU modifications]. Eventually it was concluded that ^{A}P was the optimum analogue for the recognition of a CG base pair (Figure 20). As a partner for a CG base pair, ^{A}P in conjunction with a sugar modification such as 2′-AE might be promising.

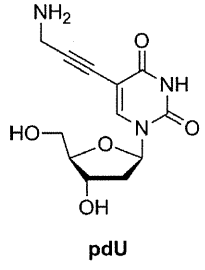


Figure 19. Structure of pdU.

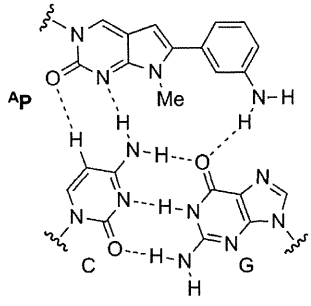


Figure 20. Proposed recognition scheme for an ^{A}P-CG triplet.

Seidman's group investigated the triplex-forming abilities of TFOs containing 11 kinds of N^4-monosubstituted cytosine nucleobases for CG base pair recognition to target the *Hprt* gene, which has a single CG interruption.[59] The nucleobases were typically prepared from triazolylated U by treatment with the appropriate primary amines. This synthetic approach can allow the preparation of various derivatives by modification after the construction of the TFO sequence (i.e., post-elongation modifications). Binding assays and UV melting experiments with the fully 2′-*O*-methyl-modified TFOs indicated that a N^4-(2-guanidinylethyl)-5-methylcytosine nucleobase (Figure 21) had the highest affinity and selectivity for a CG base pair. The T_m value and ΔT_m with a CG base pair were 54 °C and ≥16 °C, respectively, whereas those of T with an AT base pair to form a canonical T-AT triplet were 59 °C and ≥12 °C, respectively. In the case of a 2′-*O*-methyl-modified TFO containing a 2′-AE cluster,[60] the ΔT_m value with a CG base pair was significantly decreased to ≥8 °C, although the ΔT_m value observed for the T-AT triplet was ≥11 °C. However, this nucleobase would be one of the most suitable partners of a CG base pair. The recognition scheme provided by molecular dynamics (MD) simulations is shown in Figure 21. The guanidine moiety appeared to have no interaction with any phosphate backbone.

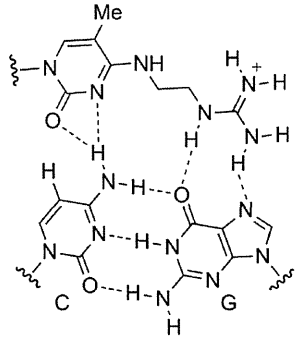


Figure 21. MD-simulation-based recognition scheme for a CG base pair by a N^4-(2-guanidinylethyl)-5-methylcytosine nucleobase.

We performed systematic investigations based on UV melting experiments on the recognition properties of 16 variations of *N*,*N*-disubstituted cytosine nucleobases, synthesized by post-elongation modifications, in triplex DNA formation.[61] Screening of azacycloalkane derivatives of different ring sizes showed the affinities towards a CG base pair decreasing in the following order of ring size: four ≈ five > six ≈ seven. This is presumably due to the bulkiness imparted by the large six- or seven-membered rings (Figure 22). The CG-recognition abilities of azacycloalkane derivatives containing four- or five-membered rings were almost the same as those of T or C. Moreover, investigations on ring-substituted azacycloalkane cytosine nucleobases as shown in Figure 23 demonstrated that the (3*S*)-hydroxypyrrolidino derivative and its C_2-symmetric (3*R*,4*R*)-dihydroxypyrrolidino analogue stabilized the triplet with a

CG base pair without any stabilization with other base pairs relative to an unsubstituted pyrrolidino derivative. In contrast, further stabilization by other nucleobases was not observed. Finally, 4-[(3*R*,4*R*)-dihydroxypyrrolidino]pyrimidin-2-one nucleobase gave the best result, in the form of a T_m of 33 °C, which was 4 °C higher than that observed with a T-CG triplet, and a ΔT_m value of ≥15 °C with a CG base pair, which was equal to that of T with an AT base pair (Figure 24). This combination with sugar modifications might further improve affinity and selectivity towards a CG base pair.

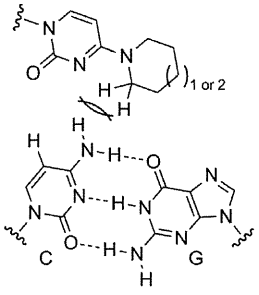


Figure 22. Putative steric repulsion between artificial cytosine nucleobases and a CG base pair.

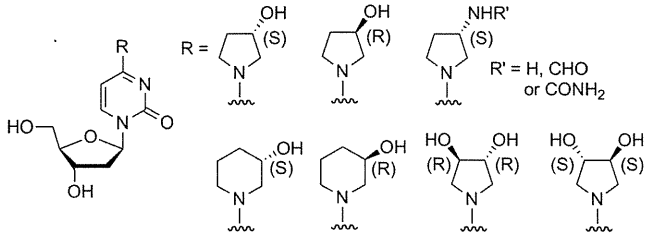


Figure 23. Structures of nucleosides containing *N*,*N*-disubstituted cytosine nucleobases.

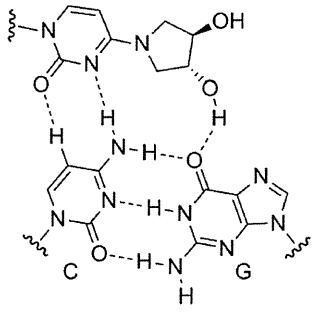


Figure 24. Proposed recognition scheme for a CG base pair by the 4-[(3*R*,4*R*)-dihydroxypyrrolidino]pyrimidin-2-one nucleobase.

For recognition of a TA base pair, a nucleobase S (Figure 25), containing two unfused aromatic rings and an acetamide unit directly attached to 2′-deoxyribose, was designed by Sun's group.[62] S could recognize a TA base pair with some discrimination relative to other base pairs and the S-TA triplet was significantly more stable than a G-TA triplet. Replacement of the aminothiazole part of S with aniline drastically reduced the affinity towards a TA base pair, clearly indicating that the aminothiazole component was integral to TA recognition. It was considered that TA recognition by S, unlike in the case of D3, was through hydrogen bonds, and not through intercalation. As well as S, various analogues such as B_t, S_1, S_a and S_{a1} were also investigated (Figure 25).[63,64] S was found to be the best nucleoside for a TA base pair.[63]

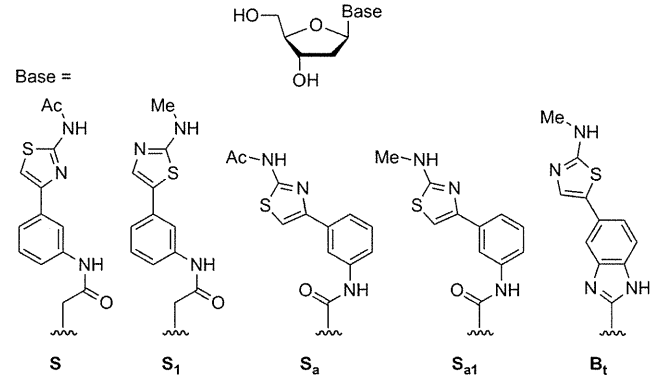


Figure 25. Structures of nucleosides containing artificial nucleobases.

Fox's group attempted triplex formation with dsDNA containing all four base pairs by TFOs containing BAU (Figure 26),[65] ^{Me}P (Figure 2), ^{A}PP (Figure 18, a) and S for recognition of AT, GC, CG and TA base pairs, respectively.[66] Fluorescence melting experiments and a DNase I footprinting assay showed that a stable triplex could be formed with 19 bps of dsDNA containing two CG and two TA base pairs at physiological pH. However, no sequence-selectivity was observed. The S appeared to lose TA selectivity and bound to a CG base pair when TFOs containing BAU in place of T were used. On the other hand, Fox's group demonstrated that 2′-AE-modified S (2′-AE-S) was superior to S for recognition of a TA base pair through triplex formation in a pH range from 5.0 to 6.0 (Figure 27).[67] Further investigations of 2′-AE-S for TA base

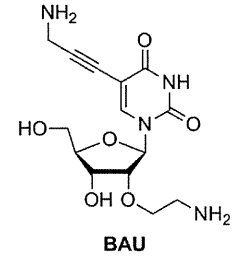


Figure 26. Structure of BAU.

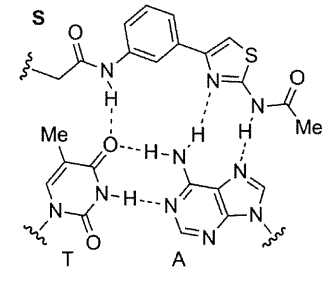


Figure 27. Proposed recognition scheme for an S-TA triplet.

pair recognition should be very interesting. Moreover, a 2′-methoxyethyl S (2′-MOE-S) monomer was recently developed by Brown's group and its triplex stability in conjunction with 2′-substituted analogues of ^{Me}P shown in Figure 2 was evaluated.[68] In the future, more detailed results of base pair recognition abilities of 2′-MOE-S may also be expected.

Fresco's group used deoxyribonucleosides bearing 2-aminopurine and 2,6-diaminopurine as G analogues.[69] The affinities of these nucleobases towards a TA base pair were significantly less than that of G and they did not show TA selectivity. This result, based on NMR studies, suggested that some additional factor contributing to the stability of the G-TA triplet though only the 2-amino group of G was essential for the recognition of a TA base pair. Moreover, 2′-AE derivatives with G or 2-aminopurine as nucleobase components seemed to produce decreases in the stabilities with a TA base pair.[44]

Nielsen's group reported the interesting result that a 3-oxo-2,3-dihydropyrazine nucleobase (E, Figure 28) attached to a peptide nucleic acid (PNA) could recognize a TA base pair with high selectivity in a PNA-DNA-PNA triplex, though the stability appears to be significantly less than that of a canonical triplet in the triplex (Figure 28).[70] The binding affinity and selectivity of a PNA containing E towards dsDNA containing a TA base pair could be promising.

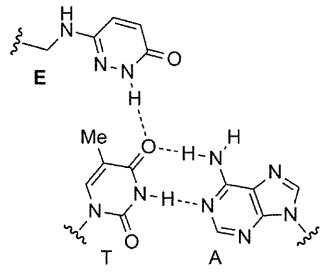


Figure 28. Proposed structure of an E-TA triplet.

Approaches Based on Antiparallel Motifs

Jayaraman's group evaluated the base pair recognition capabilities, based on a T-CG triplet (Figure 2, c), of 2′-deoxynucleosides containing T, U and 5-halogenated U as nucleobases by gel shift assay.[14] Although all of the nucleobases displayed AT selectivity in the manner shown in Figure 1 (b), 5-fluoroU bound most strongly to a CG base pair in the triplex formation. In the case of pyridin-4-one as a nucleobase simplified for CG base pair recognition, the affinity towards a CG base pair was substantially decreased relative to that by T.[14]

Dervan's group studied the affinities of 2′-deoxynebularine (N, Figure 29) towards different base pairs in dsDNA with formation of triplexes at physiological pH values and temperatures by affinity cleavage assays.[71] The affinities decreased in the order CG ≈ AT >> GC ≈ TA; both the N-CG and the N-AT triplets probably included single hydrogen bonds to N as shown in Figure 29. The stability of the N-CG triplet was comparable to that of a T-CG triplet but the selectivity towards a CG base pair was poor.

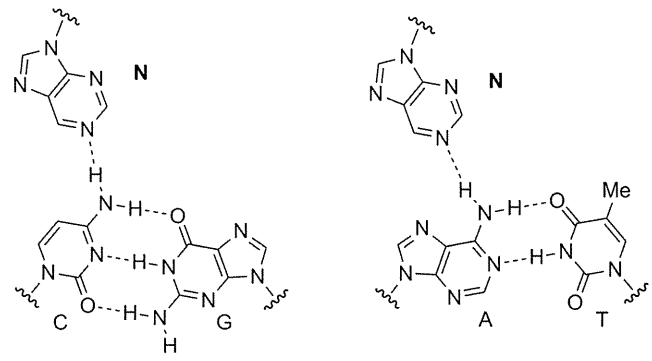


Figure 29. Proposed structures of N-CG and N-AT triplets.

Revanker's group designed 2′-deoxyformycin A (F) for a CG base pair, the proposed recognition mode of which is shown in Figure 30.[72] Although the F-CG triplet appeared to distort the canonical triplex structure, a gel shift assay with a 36-mer TFO found that the triplex containing three F-CG triplets was 10 times more stable than that containing three G-CG triplets. Evaluation of the binding affinities of F towards other base pairs will be needed to estimate the capability of F as a partner for a CG base pair. Screening of simple azoles such as pyrazole, imidazole, 1,2,4-triazole and tetrazole was also carried out.[73,74] The results indicate that they had affinities towards a TA base pair and that the pyrazole nucleobase also had TA selectivity. The mode of recognition might be shape-fitting between a TA base pair and the small azole ring, but not through hydrogen bond formation. However, the base triplets appear to be slightly weaker than a T-CG triplet.

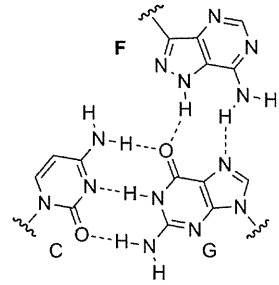


Figure 30. Proposed structure of an F-CG triplet.

Sasaki's group developed a W-shaped nucleoside analogue (WNA) containing additional tetrahydrofuran and aromatic rings (Figure 31, a).[75–79] The tetrahydrofuran ring placed the nucleobase in the correct position for recognition of a pyrimidine-purine interruption and the aromatic ring provided an increase in the stability of the triplex through stacking. The base pair recognition ability of the WNA appeared to depend considerably on the sequence contexts. When a G-GC triplet was formed at the 5′-site of the WNA moiety in a TFO, WNA-βT (base = T, R = H)

was able to bind to a TA base pair in a strong and sequence-selective way.[75,76] The presence, however, of an A-AT triplet at the 5′-site of WNA-βT apparently destroys the affinity of WNA-βT for any base pairs.

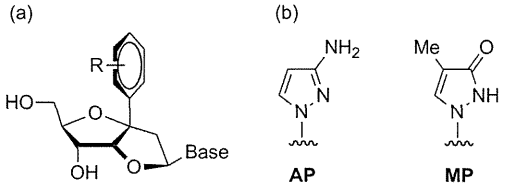


Figure 31. a) General structure of a WNA. b) Structures of AP and MP nucleobases.

WNA-βC (base = C, R = H) only effectively recognized a CG base pair in the case of G-GC and A-AT triplets at the 5′- and 3′-sites of WNA, respectively.

In addition, the affinities of the WNA derivatives towards each base pair in the triplex were considerably influenced by the substituent on the aromatic ring. The *m*-bromo derivative of WNA-βT (base = T, R = *m*-Br) was able to target a CG base pair in dsDNA in G-GC and A-AT triplets at the 5′- and 3′-sites of WNA, respectively, whereas the *p*-bromo derivative selectively formed the base triplet with a TA base pair only when both sides were A-AT triplets.[76] *p*-Amino derivatives of WNA-βC and WNA-βT showed decreased selectivities towards a CG and a TA base pair, respectively, perhaps due to non-selective interaction of their amino groups with the phosphate backbone of the target dsDNA.[77] Investigation of WNAs bearing 3-aminopyrazole (AP) or 4-methylpyrazol-3-one (MP) nucleobases showed that WNA-βAP was capable of recognizing a CG base pair in G-GC triplets at both 5′- and 3′-sites of WNA (Figure 31, b).[78]

Approaches in which a purine base of the alternate strand at a pyrimidine-purine interruption site is recognized have also been implemented. One such approach, independently proposed by Hélène's and Behr's groups, is based on the α- or β-configurations of the nucleotides, as shown in Figure 32.[5a,16a] Behr's group designed 4-guanidinopyrimidin-2-one as a nucleobase for CG and GC base pairs,[80] but no triplex formation was detectable. Timofeev's group tried an approach based on the use of natural nucleobases.[81,82] α,β-Oligodeoxycytidylates generally produced self-associates such as the *i*-motif. However, when a long (30-mer), random and non-alternating α,β-oligodeoxycytidylate was used as a TFO, α- and β-protonated C (C^+H) was able to bind to TA and AT base pairs, respectively.[81] α-T was also shown to recognize single and quintuple TA base pairs in the chimeric α,β-TFOs, although the presence of three dispersed TA base pairs prevented triplex formation, probably due to the lost stacking interactions between the base triplets.[82]

Leumann's group, on the other hand, examined hypoxanthin-7-yl, 2-aminopurin-7-yl or 2-aminopurin-9-yl nucleobases in a similar approach.[83,84] When 15-mer TFOs containing 2-aminopurin-7-yl nucleobases (α-7AP or β-7AP, Figure 33) with α- or β-configuration at a single position were used, α-7AP and β-7AP were able to bind selectively to TA and AT base pairs, respectively. Consequently, in an antiparallel motif triplex α-7AP was a promising candidate for a TA base pair, although the affinity of α-7AP was approximately four times less than that of T towards an AT base pair to form a canonical triplet.

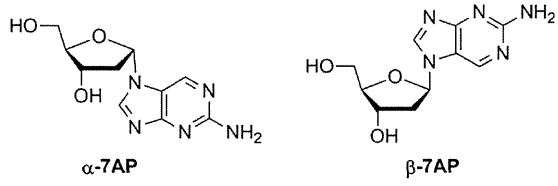


Figure 33. Structures of α-7AP and β-7AP.

Gold's group developed 2′-deoxynucleosides based on a set of four heterocyclic nucleobases (TRIPsides, Figure 34) with use of positional differentiation of the glycosyl bond instead of α- or β-configurations.[85–88] The nucleobases in the TRIPsides preferentially adopted *anti* conformations due to steric hindrance between the nucleobases and the sugar moieties. Although triplexes formed by TFOs consisting of TRIPsides – oligoTRIPs – strictly belong neither to parallel or to antiparallel motifs, they are named on the basis of the antiparallel type (e.g., as antiTA). It was shown that the four TRIPsides could recognize their target base pairs in the formation of intramolecular triplexes. With respect to intermolecular triplexes between oligoTRIPs and dsDNA, the oligoTRIPs examined, such as $(antiCG)_6$-$(antiGC)_6$, $(antiGC)_6(antiCG)_6$ and $(antiGC)_5(antiTA)_2$-

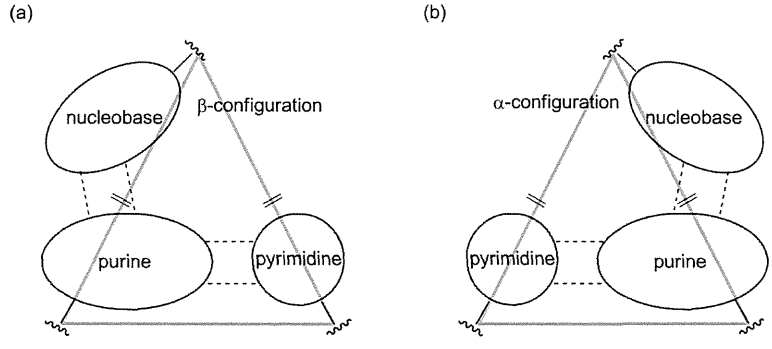


Figure 32. Approach based on α- or β-configurations of nucleotides. a) For a purine-pyrimidine base pair. b) For the pyrimidine-purine interruption. Dashes indicate hydrogen bond formation.

(antiGC)$_7$(antiTA)$_3$(antiGC)$_2$, also formed sequence-selective triplexes with the corresponding dsDNA at neutral pH.[85] Triplex formation of oligoTRIPs with any dsDNA sequence will be very interesting in terms of their practical uses.

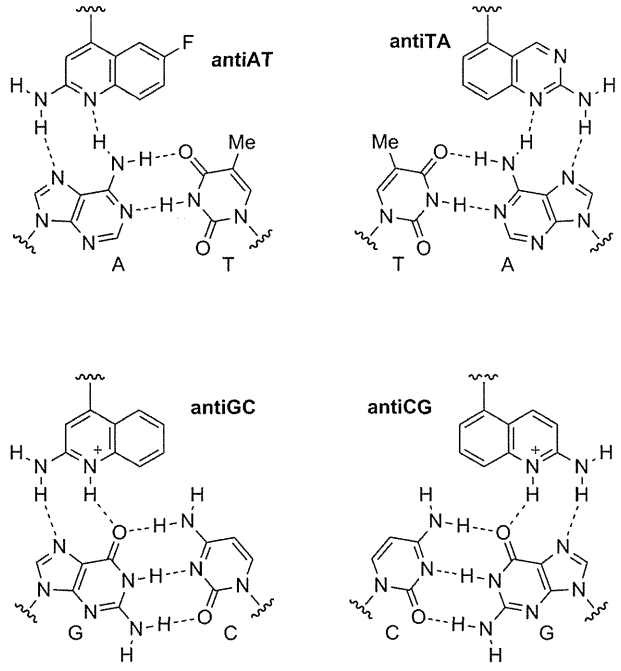


Figure 34. Recognition of base pairs by TRIPsides (antiAT, antiTA, antiGC, and antiCG).

Conclusions

To overcome the sequence limitation to target dsDNA through triplex formation, many researchers have engaged in the development of artificial nucleic acids, in particular nucleobases. As a result, several promising candidates have been found. However, there are still problems to be solved. Selective recognition of a TA base pair in a parallel triplex remains insufficient, for example; the bulkiness of the 5-methyl group in T is considered to lead to the difficulty in nucleobase design for TA base pair recognition. In addition, the neighbouring base pairs of pyrimidine-purine interruptions both in parallel and in antiparallel motifs have significant effects on the recognition abilities of interruption by nucleobases and the affinities towards interruptions often decrease significantly. However, much useful information has been obtained to date and should contribute to the discovery of further improved nucleobases. We believe that targeting of dsDNA by TFOs will be a truly practical and general strategy in the future.

Acknowledgments

The authors are grateful to the Ministry of Education, Culture, Sports, Science and Technology, Japan, for the Program to Disseminate Tenure Tracking System (to Y. H.).

[1] Reviews: a) D. Praseuth, A. L. Guieysse, C. Hélène, *Biochim. Biophys. Acta Gene Struct. Expression Biochim. Biophys. Acta* **1999**, *1489*, 181–206; b) M. Faria, C. Giovannangeli, *J. Gene Med.* **2001**, *3*, 299–310; c) M. P. Knauert, P. M. Glazer, *Hum. Mol. Genet.* **2001**, *10*, 2243–2251; d) M. M. Seidman, P. M. Glazer, *J. Clin. Invest.* **2003**, *112*, 487–494; e) J. Y. Kuan, P. M. Glazer, *Methods Mol. Biol.* **2004**, *262*, 173–194; f) R. I. Mahato, K. Chen, R. V. Guntaka, *Expert Opin. Drug Delivery* **2005**, *2*, 3–28; g) S. Sasaki, F. Nagatsugi, *Curr. Opin. Chem. Biol.* **2006**, *10*, 615–621; h) J. Y. Chin, E. B. Schleifman, P. M. Glazer, *Front. Biosci.* **2007**, *12*, 4288–4297; i) P. Simon, F. Cannata, J.-P. Concordet, C. Giovannangeli, *Biochimie* **2008**, *90*, 1109–1116; j) M. Duca, P. Vekhoff, K. Oussedik, L. Halby, P. B. Arimondo, *Nucleic Acids Res.* **2008**, *36*, 5123–5138; k) J. Y. Chin, P. M. Glazer, *Mol. Carcinog.* **2009**, *48*, 389–399; l) A. Mukherjee, K. M. Vasquez, *Biochimie* **2011**, *93*, 1197–1208.

[2] a) J.-S. Sun, T. de Bizemont, G. Duval-Valentin, T. Montenay-Garestier, C. Hélène, *C. R. Acad. Sci., Ser. III* **1991**, *313*, 585–590; b) C. Giovannangéli, M. Rougée, T. Garestier, N. T. Thuong, C. Hélène, *Proc. Natl. Acad. Sci. USA* **1992**, *89*, 8631–8635.

[3] a) J. S. Lee, M. L. Woodsworth, L. J. P. Latimer, A. R. Morgan, *Nucleic Acids Res.* **1984**, *12*, 6603–6614; b) T. J. Povsic, P. B. Dervan, *J. Am. Chem. Soc.* **1989**, *111*, 3059–3061; c) L. E. Xodo, G. Manzini, F. Quadrifoglio, G. A. van der Marel, J. H. van Boom, *Nucleic Acids Res.* **1991**, *19*, 5625–5631.

[4] a) S. Hildbrand, C. Leumann, *Angew. Chem.* **1996**, *108*, 2100; *Angew. Chem. Int. Ed. Engl.* **1996**, *35*, 1968–1970; b) P. J. Bates, C. A. Laughton, T. C. Jenkins, D. C. Capaldi, P. D. Roselt, C. B. Reese, S. Neidle, *Nucleic Acids Res.* **1996**, *24*, 4176–4184; c) S. Hildbrand, A. Blaser, S. P. Parel, C. J. Leumann, *J. Am. Chem. Soc.* **1997**, *119*, 5499–5511.

[5] For reviews, see: a) N. T. Thoung, C. Hélène, *Angew. Chem.* **1993**, *105*, 697; *Angew. Chem. Int. Ed. Engl.* **1993**, *32*, 666–690; b) J.-S. Sun, C. Hélène, *Curr. Opin. Struct. Biol.* **1993**, *3*, 345–356; c) K. R. Fox, *Curr. Med. Chem.* **2000**, *7*, 17–37; d) J. Robles, A. Grandas, E. Pedroso, F. J. Luque, R. Eritja, M. Orozco, *Curr. Org. Chem.* **2002**, *6*, 1333–1368; e) S. Buchini, C. J. Leumann, *Curr. Opin. Chem. Biol.* **2003**, *7*, 717–726; f) D. A. Rusling, V. J. Broughton-Head, T. Brown, K. R. Fox, *Curr. Chem. Biol.* **2008**, *2*, 1–10.

[6] a) J. F. Milligan, S. H. Krawczyk, S. Wadwani, M. D. Matteucci, *Nucleic Acids Res.* **1993**, *21*, 327–333; b) T. S. Rao, A. F. Lewis, R. H. Durland, G. R. Revankar, *Tetrahedron Lett.* **1993**, *34*, 6709–6712; c) T. S. Rao, A. F. Lewis, T. S. Hill, G. R. Revankar, *Nucleosides Nucleotides* **1995**, *14*, 1–12; d) T. S. Rao, R. H. Durland, D. M. Seth, M. A. Myrick, V. Bodepudi, G. R. Revankar, *Biochemistry* **1995**, *34*, 765–772; e) J. E. Gee, G. R. Revankar, T. S. Rao, M. E. Hogan, *Biochemistry* **1995**, *34*, 2042–2048; f) W. M. Olivas, L. J. Maher III, *Nucleic Acids Res.* **1995**, *23*, 1936–1941; g) Y. Aubert, L. Perrouault, C. Hélène, C. Giovannangeli, U. Asseline, *Bioorg. Med. Chem.* **2001**, *9*, 1617–1624.

[7] a) M. Paramasivam, S. Cogoi, V. V. Filichev, N. Bomholt, E. B. Pedersen, L. E. Xodo, *Nucleic Acids Res.* **2008**, *36*, 3494–3507; b) O. Doluca, A. S. Boutorine, V. V. Filichev, *ChemBioChem* **2011**, *12*, 2365–2374.

[8] K. Yoon, C. A. Hobbs, J. Koch, M. Sardaro, R. Kutny, A. L. Weis, *Proc. Natl. Acad. Sci. USA* **1992**, *89*, 3840–3844.

[9] I. Radhakrishnan, D. J. Patel, *J. Mol. Biol.* **1994**, *241*, 600–619.

[10] L. C. Griffin, P. B. Dervan, *Science* **1989**, *245*, 967–971.

[11] a) I. Radhakrishnan, X. Gao, C. de los Santos, D. Live, D. J. Patel, *Biochemistry* **1991**, *30*, 9022–9030; b) I. Radhakrishnan, D. J. Patel, *Biochemistry* **1992**, *31*, 2514–2523; c) E. Wang, S. Malek, J. Feigon, *Biochemistry* **1992**, *31*, 4838–4846; d) I. Radhakrishnan, D. J. Patel, *Structure* **1994**, *2*, 17–32; e) L. Jiang, I. M. Russu, *Nucleic Acids Res.* **2001**, *29*, 4231–4237.

[12] P. S. Miller, C. D. Cushman, *Biochemistry* **1993**, *32*, 2999–3004.

[13] P. A. Beal, P. B. Dervan, *Nucleic Acids Res.* **1992**, *20*, 2773–2776.
[14] R. H. Durland, T. S. Rao, G. R. Revankar, J. H. Tinsley, M. A. Myrick, D. M. Seth, J. Rayford, P. Singh, K. Jayaraman, *Nucleic Acids Res.* **1994**, *22*, 3233–3240.
[15] a) K. Dittrich, J. Gu, R. Tinder, M. Hogan, X. Gao, *Biochemistry* **1994**, *33*, 4111–4120; b) J. Ji, M. E. Hogan, X. Gao, *Structure* **1996**, *4*, 425–435.
[16] Reviews: a) S. O. Doronina, J.-P. Behr, *Chem. Soc. Rev.* **1997**, *26*, 63–71; b) I. Luyten, P. Herdewijn, *Eur. J. Med. Chem.* **1998**, *33*, 515–576; c) D. M. Gowers, K. R. Fox, *Nucleic Acids Res.* **1999**, *27*, 1569–1577; d) M. G. M. Purwanto, K. Weisz, *Curr. Org. Chem.* **2003**, *7*, 427–446; e) K. R. Fox, T. Brown, *Quart. Rev. Biophys.* **2005**, *38*, 311–320; f) V. Malnuit, M. Duca, R. Benhida, *Org. Biomol. Chem.* **2011**, *9*, 326–336; g) K. R. Fox, T. Brown, *Biochem. Soc. Trans.* **2011**, *39*, 629–634.
[17] L. C. Griffin, L. L. Kiessling, P. A. Beal, P. Gillespie, P. B. Dervan, *J. Am. Chem. Soc.* **1992**, *114*, 7976–7982.
[18] L. L. Kiessling, L. C. Griffin, P. B. Dervan, *Biochemistry* **1992**, *31*, 2829–2834.
[19] a) K. M. Koshlap, P. Gillespie, P. B. Dervan, J. Feigon, *J. Am. Chem. Soc.* **1993**, *115*, 7908–7909; b) E. Wang, K. M. Koshlap, P. Gillespie, P. B. Dervan, J. Feigon, *J. Mol. Biol.* **1996**, *257*, 1052–1069.
[20] T. E. Lehmann, W. A. Greenberg, D. A. Liberles, C. K. Wada, P. B. Dervan, *Helv. Chim. Acta* **1997**, *80*, 2002–2022.
[21] S. C. Zimmerman, P. Schmitt, *J. Am. Chem. Soc.* **1995**, *117*, 10769–10770.
[22] E. Mertz, S. Mattei, S. C. Zimmerman, *Org. Lett.* **2000**, *2*, 2931–2934.
[23] E. Mertz, S. Mattei, S. C. Zimmerman, *Bioorg. Med. Chem.* **2004**, *12*, 1517–1526.
[24] W. Wang, M. G. M. Purwanto, K. Weisz, *Org. Biomol. Chem.* **2004**, *2*, 1194–1198.
[25] F. Wachowius, M. Rettig, G. Palm, K. Weisz, *Tetrahedron Lett.* **2008**, *49*, 7264–7267.
[26] a) M. G. M. Purwanto, D. Lengeler, K. Weisz, *Tetrahedron Lett.* **2002**, *43*, 61–64; b) M. G. M. Purwanto, K. Weisz, *J. Org. Chem.* **2004**, *69*, 195–197.
[27] M. G. M. Purwanto, K. Weisz, *Tetrahedron Lett.* **2006**, *47*, 3849–3852.
[28] S. K. Singh, P. Nielsen, A. A. Koshkin, J. Wengel, *Chem. Commun.* **1998**, 455–456.
[29] a) S. Obika, D. Nanbu, Y. Hari, K. Morio, Y. In, T. Ishida, T. Imanishi, *Tetrahedron Lett.* **1997**, *38*, 8735–8738; b) S. Obika, D. Nanbu, Y. Hari, J. Andoh, K. Morio, T. Doi, T. Imanishi, *Tetrahedron Lett.* **1998**, *39*, 5401–5404.
[30] a) S. Obika, Y. Hari, T. Sugimoto, M. Sekiguchi, T. Imanishi, *Tetrahedron Lett.* **2000**, *41*, 8923–8927; b) H. Torigoe, Y. Hari, M. Sekiguchi, S. Obika, T. Imanishi, *J. Biol. Chem.* **2001**, *276*, 2354–2360; c) S. Obika, T. Uneda, T. Sugimoto, D. Nanbu, T. Minami, T. Doi, T. Imanishi, *Bioorg. Med. Chem.* **2001**, *9*, 1001–1011; d) B.-W. Sun, B. R. Babu, M. D. Sørensen, K. Zakrzewska, J. Wengel, J.-S. Sun, *Biochemistry* **2004**, *43*, 4160–4169.
[31] S. Obika, Y. Hari, K. Morio, T. Imanishi, *Tetrahedron Lett.* **2000**, *41*, 221–224.
[32] Y. Hari, S. Obika, H. Inohara, M. Ikejiri, D. Une, T. Imanishi, *Chem. Pharm. Bull.* **2005**, *53*, 843–846.
[33] S. Obika, H. Inohara, Y. Hari, T. Imanishi, *Bioorg. Med. Chem.* **2008**, *16*, 2945–2954.
[34] Y. Hari, M. Nakahara, J. Pang, M. Akabane, T. Kuboyama, S. Obika, *Bioorg. Med. Chem.* **2011**, *19*, 1162–1166.
[35] C.-Y. Huang, C. D. Cushman, P. S. Miller, *J. Org. Chem.* **1993**, *58*, 5048–5049.
[36] C.-Y. Huang, P. S. Miller, *J. Am. Chem. Soc.* **1993**, *115*, 10456–10457.
[37] C.-Y. Huang, G. Bi, P. S. Miller, *Nucleic Acids Res.* **1996**, *24*, 2606–2613.
[38] S. Verma, P. S. Miller, *Bioconjugate Chem.* **1996**, *7*, 600–605.
[39] N. Guzzo-Pernell, G. W. Tregear, J. Haralambidis, J. W. Lawlor, *Nucleosides Nucleotides* **1998**, *17*, 1191–1207.
[40] D. A. Gianolio, L. W. McLaughlin, *Nucleosides Nucleotides* **1999**, *18*, 1751–1769.
[41] I. Prévot-Halter, C. J. Leumann, *Bioorg. Med. Chem. Lett.* **1999**, *9*, 2657–2660.
[42] J. Marfurt, C. J. Leumann, *Angew. Chem.* **1998**, *110*, 184; *Angew. Chem. Int. Ed.* **1998**, *37*, 175–177.
[43] a) B. Cuenoud, F. Cassert, D. Hüsken, F. Natt, R. M. Wolf, K.-H. Altmann, P. Martin, H. E. Moser, *Angew. Chem.* **1998**, *110*, 1350; *Angew. Chem. Int. Ed.* **1998**, *37*, 1288–1291; b) M. J. J. Blommers, F. Natt, W. Jahnke, B. Cuenoud, *Biochemistry* **1998**, *37*, 17714–17725.
[44] S. Buchini, C. J. Leumann, *Eur. J. Org. Chem.* **2006**, 3152–3168.
[45] S. Buchini, C. J. Leumann, *Tetrahedron Lett.* **2003**, *44*, 5065–5068.
[46] S. Buchini, C. J. Leumann, *Angew. Chem.* **2004**, *116*, 4015; *Angew. Chem. Int. Ed.* **2004**, *43*, 3925–3928.
[47] I. Prévot, C. J. Leumann, *Helv. Chim. Acta* **2002**, *85*, 502–515.
[48] D. L. Chen, L. W. McLaughlin, *J. Org. Chem.* **2000**, *65*, 7468–7474.
[49] S. Obika, Y. Hari, M. Sekiguchi, T. Imanishi, *Angew. Chem.* **2001**, *113*, 2149; *Angew. Chem. Int. Ed.* **2001**, *40*, 2079–2081.
[50] S. Obika, Y. Hari, M. Sekiguchi, T. Imanishi, *Chem. Eur. J.* **2002**, *8*, 4796–4802.
[51] H. Torigoe, Y. Hari, S. Obika, T. Imanishi, *Nucleosides Nucleotides Nucleic Acids* **2003**, *22*, 1097–1099.
[52] Y. Hari, S. Obika, M. Sekiguchi, T. Imanishi, *Tetrahedron* **2003**, *59*, 5123–5128.
[53] H. Torigoe, Y. Hari, S. Obika, T. Imanishi, *Nucleosides Nucleotides Nucleic Acids* **2003**, *22*, 1571–1573.
[54] Y. Hari, S. Matsugu, H. Inohara, Y. Hatanaka, M. Akabane, T. Imanishi, S. Obika, *Org. Biomol. Chem.* **2010**, *8*, 4176–4180.
[55] R. T. Ranasinghe, D. A. Rusling, V. E. C. Powers, K. R. Fox, T. Brown, *Chem. Commun.* **2005**, 2555–2557.
[56] S. R. Gerrard, N. Srinivasan, K. R. Fox, T. Brown, *Nucleosides Nucleotides Nucleic Acids* **2007**, *26*, 1363–1367.
[57] S. R. Gerrard, M. M. Edrees, I. Bouamaied, K. R. Fox, T. Brown, *Org. Biomol. Chem.* **2010**, *8*, 5087–5096.
[58] J. Bijapur, M. D. Keppler, S. Bergqvist, T. Brown, K. R. Fox, *Nucleic Acids Res.* **1999**, *27*, 1802–1809.
[59] A. Semenyuk, E. Dariann, J. Liu, A. Majumdar, B. Cuenoud, P. S. Miller, A. D. MacKerell Jr, M. M. Seidman, *Biochemistry* **2010**, *49*, 7867–7878.
[60] a) N. Puri, A. Majumdar, B. Cuenoud, F. Natt, P. Martin, A. Boyd, P. S. Miller, M. M. Seidman, *Biochemistry* **2002**, *41*, 7716–7724; b) N. Puri, A. Majumdar, B. Cuenoud, P. S. Miller, M. M. Seidman, *Biochemistry* **2004**, *43*, 1343–1351.
[61] Y. Hari, M. Akabane, Y. Hatanaka, M. Nakahara, S. Obika, *Chem. Commun.* **2011**, *47*, 4424–4426.
[62] D. Guianvarc'h, R. Benhida, J.-L. Fourrey, R. Maurisse, J.-S. Sun, *Chem. Commun.* **2001**, 1814–1815.
[63] D. Guianvarc'h, J.-L. Fourrey, R. Maurisse, J.-S. Sun, R. Benhida, *Bioorg. Med. Chem.* **2003**, *11*, 2751–2759.
[64] D. Guianvarc'h, J.-L. Fourrey, R. Maurisse, J.-S. Sun, R. Benhida, *Org. Lett.* **2002**, *4*, 4209–4212.
[65] a) M. Sollogoub, R. A. J. Darby, B. Cuenoud, T. Brown, K. R. Fox, *Biochemistry* **2002**, *41*, 7224–7231; b) S. D. Osborne, V. E. C. Powers, D. A. Rusling, O. Lack, K. R. Fox, T. Brown, *Nucleic Acids Res.* **2004**, *32*, 4439–4447.
[66] D. A. Rusling, V. E. C. Powers, R. T. Ranasinghe, Y. Wang, S. D. Osborne, T. Brown, K. R. Fox, *Nucleic Acids Res.* **2005**, *33*, 3025–3032.
[67] Y. Wang, D. A. Rusling, V. E. C. Powers, O. Lack, S. D. Osborne, K. R. Fox, T. Brown, *Biochemistry* **2005**, *44*, 5884–5892.
[68] a) C. Lou, Q. Xiao, R. R. Tailor, N. B. Gaied, N. Gale, M. E. Light, K. R. Fox, T. Brown, *Med. Chem. Commun.* **2011**, *2*, 550–558; b) C. Lou, M. Shelbourne, K. R. Fox, T. Brown, *Chem. Eur. J.* **2011**, *17*, 14851–14856.

[69] O. A. Amosova, J. R. Fresco, *Nucleic Acids Res.* **1999**, *27*, 4632–4635.
[70] A. B. Eldrup, O. Dahl, P. E. Nielsen, *J. Am. Chem. Soc.* **1997**, *119*, 11116–11117.
[71] H. U. Stilz, P. B. Dervan, *Biochemistry* **1993**, *32*, 2177–2185.
[72] T. S. Rao, M. E. Hogan, G. R. Revankar, *Nucleosides Nucleotides* **1994**, *13*, 95–107.
[73] R. H. Durland, T. S. Rao, V. Bodepudi, D. M. Seth, K. Jayaraman, G. R. Revankar, *Nucleic Acids Res.* **1995**, *23*, 647–653.
[74] K. Jayaraman, R. H. Durland, T. S. Rao, G. R. Revankar, V. Bodepudi, N. Chaudhary, J. Guy-Caffey, *Nucleosides Nucleotides* **1995**, *14*, 951–955.
[75] S. Sasaki, Y. Taniguchi, R. Takahashi, Y. Senko, K. Kodama, F. Nagatsugi, M. Maeda, *J. Am. Chem. Soc.* **2004**, *126*, 516–528.
[76] Y. Taniguchi, A. Nakamura, Y. Senko, F. Nagatsugi, S. Sasaki, *J. Org. Chem.* **2006**, *71*, 2115–2122.
[77] Y. Taniguchi, M. Togo, E. Aoki, Y. Uchida, S. Sasaki, *Tetrahedron* **2008**, *64*, 7164–7170.
[78] Y. Taniguchi, Y. Uchida, T. Takaki, E. Aoki, S. Sasaki, *Bioorg. Med. Chem.* **2009**, *17*, 6803–6810.
[79] S. Sasaki, H. Yamaguchi, F. Nagatsugi, R. Takahashi, Y. Taniguchi, M. Maeda, *Tetrahedron Lett.* **2001**, *42*, 6915–6918.
[80] S. O. Doronina, S. Blanalt-Feidt, J.-P. Behr, *Nucleosides Nucleotides* **1999**, *18*, 1617–1618.
[81] a) E. N. Timofeev, O. F. Borisova, A. K. Shchyolkina, *J. Biomol. Struct. Dyn.* **2000**, *17*, 655–664; b) E. N. Timofeev, S. V. Kochetkova, V. L. Florentiev, *Mol. Biol. (Russian)* **2004**, *38*, 547–552.
[82] E. N. Timofeev, B. V. Goryaeva, V. L. Florentiev, *J. Biomol. Struct. Dyn.* **2006**, *24*, 183–188.
[83] S. P. Parel, C. J. Leumann, *Nucleic Acids Res.* **2001**, *29*, 2260–2267.
[84] S. P. Parel, J. Marfurt, C. J. Leumann, *Nucleosides Nucleotides Nucleic Acids* **2001**, *20*, 411–417.
[85] J.-S. Li, F.-X. Chen, R. Shikiya, L. A. Marky, B. Gold, *J. Am. Chem. Soc.* **2005**, *127*, 12657–12665.
[86] J.-S. Li, Y.-H. Fan, Y. Zhang, L. A. Marky, B. Gold, *J. Am. Chem. Soc.* **2003**, *125*, 2084–2093.
[87] J.-S. Li, R. Shikiya, L. A. Marky, B. Gold, *Biochemistry* **2004**, *43*, 1440–1448.
[88] J.-S. Li, B. Gold, *J. Org. Chem.* **2005**, *70*, 8764–8771.

Received: December 19, 2011
Published Online: March 12, 2012

医薬・診断薬への応用に向けた化学修飾核酸アプタマーの開発

栗 原 正 靖[*1]・小比賀 聡[*2]

1. はじめに

人工核酸を用いた転写および逆転写反応は，人工核酸をランダムスクリーニング法へ応用する上で重要な技術となる．核酸のランダムスクリーニング法には，SELEX 法(Systematic Evolution of Ligand by Exponential Enrichment)[1,2]やキャピラリー電気泳動(CE)-SELEX法[3,4]，Cell-SELEX法[5,6]などがあり，抗体と類似の機能をもつ核酸アプタマーを任意の標的分子(ターゲット)に対して創製することができる(図1)．核酸アプタマーは，生物を用いることなく創出することができる点や，化学合成によってmg～gスケールで安価に製造できる点，乾燥状態で常温保存できる点などの特長があるため，研究試薬や医薬・診断薬への応用が期待されている．しかし，核酸はヌクレアーゼ(核酸分解酵素)によって生体内で容易に分解されてしまうことが実用化の妨げになっている[7]．本稿では，医薬・診断薬等への応用を目指して，核酸塩基部や糖部などを化学修飾することでヌクレアーゼ耐性を向上させた人工核酸アプタマーの創製法開発への取り組みについて紹介する．

[*1] Masayasu Kuwahara 群馬大学大学院工学研究科 准教授 博士(学術)
[*2] Satoshi Obika 大阪大学大学院薬学研究科 教授 博士(薬学)
Development of Chemically Modified Nucleic-Acid Aptamers Towards the Application to Medicine and Diagnostic Agents

2. 化学修飾核酸の酵素的合成法

核酸は，ヌクレオチドといわれる基本単位がリン酸ジエステル結合で直列に繋がった分子である(図2)．ヌクレオチドは塩基部位，糖部位，リン酸部位からなり，アンチセンス医薬や核酸プローブなどへの応用や，二重鎖形成機構の解明などを目的として，これらの部位を化学的に修飾・改変

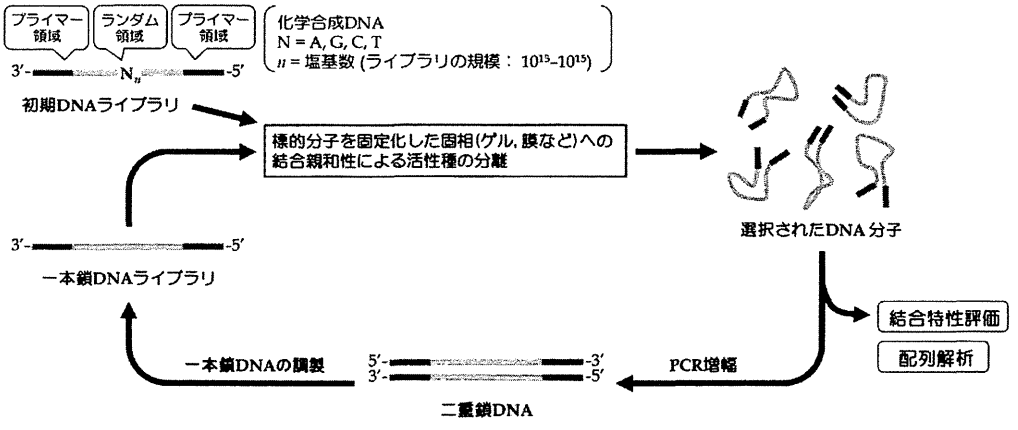


図1 SELEX 法による DNA アプタマーの作製スキーム

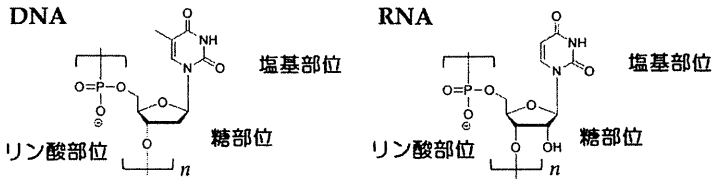


図2　DNA および RNA の化学構造

した種々の人工核酸が報告されている．筆者らは，これらの部位への修飾がポリメラーゼ反応に及ぼす効果に着目し，ポリメラーゼによる人工核酸の転写および逆転写について精査した系統的なデータに基づいて研究開発を進めており，これまでに塩基部位修飾人工核酸を用いたポリメラーゼ反応については，特に詳細に検討してきた[8~12]．

塩基部位への修飾は古くから研究されており，チミンやシトシンなどのピリミジン塩基では5位，アデニンやグアニンなどのプリン塩基では7位への置換基導入に対して，ポリメラーゼは比較的寛容であることが知られている．したがって，これらの部位に蛍光基を導入したヌクレオシド三リン酸誘導体は，核酸の蛍光標識化試薬として市販されている(**図3**)．これらの試薬は，一般に天然型の基質三リン酸(dATP，dGTP，dCTP，dTTP)の共存下で用いるため，蛍光基は DNA 鎖中の適当な箇所に挿入されることになる．例えば，蛍光標識した dTTP の誘導体を4種類の天然型 dNTP と共に用いた場合，T が挿入されるべきところには，天然型の T が入ったり修飾型の T が入ったりすることになる．蛍光標識化という点では，蛍光基の挿入箇所が定まっていなくても全く問題はないが，この技術を SELEX 法へ応用する場合には，用いる修飾型基質三リン酸に対応する天然型基質三リン酸は抜いてポリメラーゼ反応を行う必要がある．つまり，dTTP の誘導体を使用するときは天然型の dTTP を，dATP の誘導体を使用するときは，天然型の dATP を抜いて反応を行う．そうすれば，T が挿入されるべきところには全て修飾型の T が，A が挿入されるべきところには全て修飾型の A が挿入されることになる．しかしながら，この場合，4種類の天然型基質三リン酸に修飾型基質三リン酸を添加する前述の方法に比べて，ポリメラーゼ反応の効率が大きく低下する．そこで，修飾基の異なる種々の修飾型基質三リン

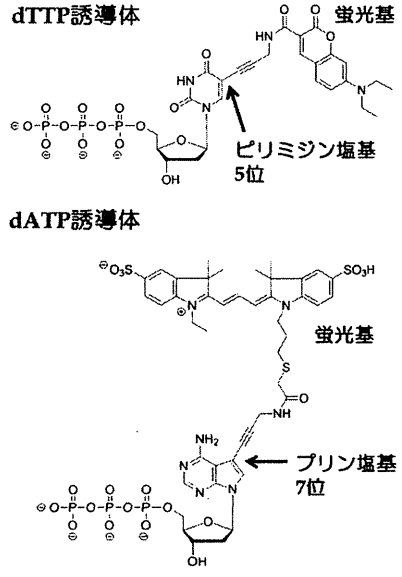


図3　蛍光基を導入したヌクレオシド三リン酸誘導体の例

酸を合成し，それらを用いたポリメラーゼ反応を系統的に検討することによって反応の最適化を試みた．その結果，修飾基の化学構造やポリメラーゼの種類，またそれらの組み合わせが反応の収率や忠実度を大きく左右することが分かった．特に，ポリメラーゼの種類は，古細菌(*Thermococcus kodakaraensis*)由来の耐熱性 DNA ポリメラーゼ(*KOD Dash*)が，ヌクレオチドの化学修飾に対して最も寛容であり，かつヌクレオチド取り込みの正確さ(忠実度)が高いことを見出した[8]．

3．糖部位修飾ヌクレオチドの酵素的導入およびその効果

一般に，糖の2'位を化学修飾したヌクレオチドの核酸鎖への導入は，ヌクレアーゼ耐性の向上効果をもたらすことが知られている．これまでに筆者らは，糖部位修飾核酸の中でも，2'位と4'位を化学的に架橋した人工核酸(BNAs；bridged

nucleic acids)が，高いヌクレアーゼ耐性を示すことを明らかにしている[13~15]．そこで修飾ヌクレオチドに対して優れた触媒特性を示す *KOD* DNA ポリメラーゼを遺伝子工学的に改変した 8 種類の変異体について，糖部位修飾ヌクレオシド三リン酸を伴うポリメラーゼ反応を新たに検証した[11]．まず，プライマー伸長反応における一ヌクレオチド取り込み反応を測定することで，キネティックパラメーター(V_{max}, K_m)を求め，さらに取り込みの正確さを算出した(**表 1**)．表 1 において基質 1 および 2 は，それぞれ**図 4** (a)に示した修飾ヌクレオシド三リン酸(JTP および MTP)である．また，実験には 3',5' エキソヌクレアーゼ活性を欠失させた変異体 *KOD*1，2，3 および 8 を用いた．興味深いことに，*Vent*(*exo*-) DNA リメラーゼの K_m 値は，塩基対において誤取り込み(ミスインコーポレーション)となる dCTP と正しい取り込み(コレクトインコーポレーション)となる dTTP および 1，2 との間に非常に大きな差異が見られるのに対し，*KOD* 変異体ではいずれもそれほど大きな差異は見られなかった．対照的に *Vent*(*exo*-) DNA ポリメラーゼの V_{max} 値は取り込み基質によって大きく変わらないのに対し，*KOD* 変異体ではかなりの差異が見られた．つまり，*KOD* およびその改変体が基質の化学修飾に寛容であり，かつ取り込みの正確さが高いという一見相反する

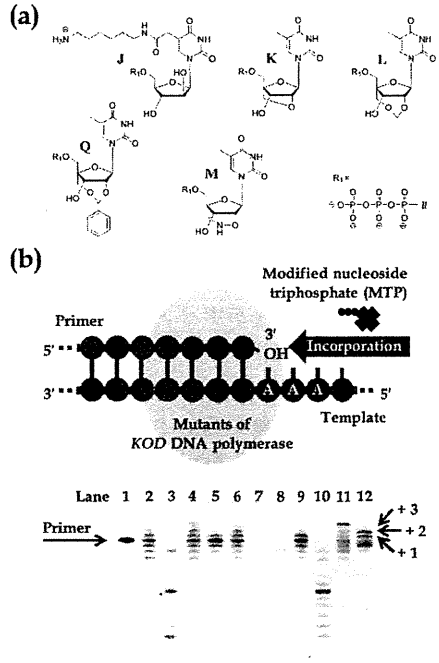


図 4　糖部位修飾ヌクレオチドの酵素的取り込み：合成した種々の糖部位修飾ヌクレオシド三リン酸(a)，修飾ヌクレオシド三リン酸(MTP)および耐熱性 DNA ポリメラーゼ(レーン 1：プライマー鎖のみ，レーン 2：*KOD Dash* を用いた反応液，レーン 3：野生型 *KOD* を用いた反応液，レーン 4：変異体 *KOD*1 を用いた反応液，レーン 5：変異体 *KOD*2 を用いた反応液，レーン 6：変異体 *KOD*3 を用いた反応液，レーン 7：変異体 *KOD*4 を用いた反応液，レーン 8：変異体 *KOD*5 を用いた反応液，レーン 9：変異体 *KOD*6 を用いた反応液，レーン 10：変異体 *KOD*7 を用いた反応液，レーン 11：変異体 *KOD*8 を用いた反応液，レーン 12：変異体 *Vent*(*exo*-)を用いた反応液)を用いた架橋型ヌクレオチドの連続取込み実験(b)

表 1　一ヌクレオチド取り込み反応の速度論的解析

DNA ポリメラーゼ	基質 dNTP	最大反応速度 V_{max} (min^{-1})	ミカエリス定数 K_m (mM)	取り込み効率 V_{max}/K_m	取り込みの正確さ
*KOD*1	T	1.8	0.021	88	1
	1	11	0.14	81	0.92
	2	0.10	0.26	3.8	0.043
	C	0.074	0.66	0.11	0.0013
*KOD*2	T	2.1	0.035	60	1
	1	4.0	0.083	48	0.80
	2	0.32	0.099	3.3	0.055
	C	0.019	0.11	0.17	0.0029
*KOD*3	T	2.3	0.047	49	1
	1	4.4	0.081	55	1.1
	2	0.27	0.099	2.7	0.055
	C	0.033	0.70	0.048	0.00097
*KOD*8	T	9.5	0.063	150	1
	1	53	0.31	170	1.1
	2	2.4	0.38	6.3	0.041
	C	0.11	1.0	0.11	0.00073
Vent(*exo*-)	T	0.34	0.0030	110	1
	1	0.29	0.0027	110	0.95
	2	0.049	0.0044	11	0.099
	C	0.14	1.7	0.083	0.00073

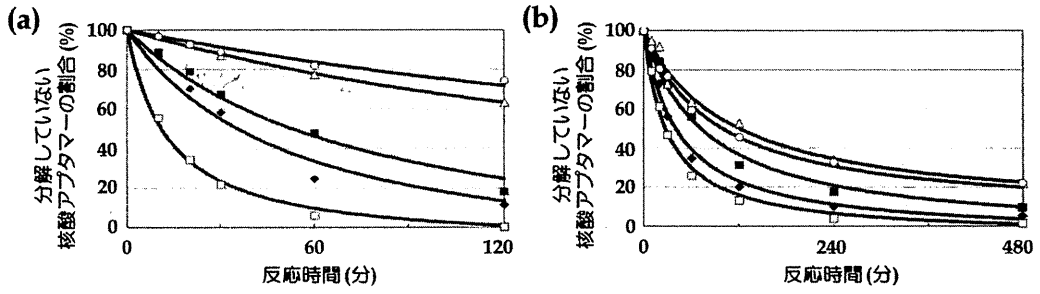


図5 架橋型ヌクレオチドの導入がトロンビン結合性DNAアプタマーの生体内安定性向上に及ぼす効果：蛇毒ホスホジエステラーゼ溶液中(a)，ヒト血清中(b)

性質を示すのは，触媒効率において V_{max} 値が支配的であることに起因するものと考えられる．次に架橋型ヌクレオチドMの連続取り込みについて検証したところ，*KOD*1および8を用いた場合，効率的な連続取り込みが確認された(図4(b))．

さらに，高いヌクレアーゼ耐性が期待される架橋型ヌクレオチドについては，核酸鎖へのそれらの導入が生体内安定性に及ぼす効果についても検証した(**図5**(a)および5(b)[16,17])．例えば，トロンビン結合性DNAアプタマー(TBA)の3'末端に架橋型ヌクレオチドK，L，M，Qを導入した場合，ヌクレアーゼである蛇毒ホスホジエステラーゼに対する抵抗性は大きく向上し，Q修飾TBA(○)＞M修飾TBA(△)＞L修飾TBA(■)＞K修飾TBA(◆)＞未修飾のTBA(□)の順であった(図5(a))．最も高い耐性を示したQ修飾TBAの分解速度は，未修飾のTBAのそれよりも27倍も遅いことが分かった．同様にヒト血清中における安定性を検証したところ，M修飾TBA(△)＞Q修飾TBA(○)＞L修飾TBA(■)＞K修飾TBA(◆)＞未修飾のTBA(□)の順に安定であった(図5(b))．ヒト血清には，エキソヌクレアーゼの他にエンドヌクレアーゼも存在するため，3'末端への架橋型ヌクレオチド導入が分解抵抗性に及ぼす効果は減少したが，それでも最も高い安定性を示したM修飾TBAの分解速度は，未修飾のTBAのそれよりも4倍も遅いという結果が得られた．

4．キャピラリー電気泳動(CE)-SELEX法によるスクリーニング

迅速かつ簡便に核酸アプタマーを取得する方法としてCE-SELEX法およびいくつかの改良法が開発された．通常のSELEX法では，選別の工程で固相(アフィニティ・ゲルやニトロセルロース膜など)が汎用されるのに対し，これらの方法はいずれもCEを用いることを特徴とする(**図6**)．固相を用いない溶液中での結合解離反応と高い分離能をもつ自動化されたCE装置の使用によって，スクリーニング操作の繰り返し(ラウンド数)が大幅に低減され，簡便かつ迅速な核酸アプタマーの取得が可能となった．一般にSELEX法のラウンド数は10回前後にも及ぶのに対し，CE-SELEX法はわずか2～4回で完結する．これまでにCE-SELEX法によってヒト免疫不全ウイルス逆転写酵素や免疫グロブリンE(IgE)，MutSミスマッチ結合タンパク質，リシンなどさまざまな標的に対して核酸アプタマーが得られている．得られた核酸アプタマーの結合親和性(K_d 値)は，それぞれ0.18 nM，23 nM，3.6 nM，58 nMであり，分子認識分子として十分に高い活性を示す．

一般に，CE-SELEXでは，フューズドシリカキャピラリーが用いられる．CEによる生体分子の分析では，生体分子とキャピラリー内壁との相互作用によるピークのテーリングを抑えるために，しばしばキャピラリー内壁がコーティングされたものが用いられるが，核酸もキャピラリー内壁も共に負電荷を帯びているので，特別にコーティングしていないフューズドシリカキャピラリーを用いることができる．また，ライブラリに用いるオリゴヌクレオチドは，30～40 merのランダム領域を含む70～90 merの長さであり，その5'末端は，結合活性をもつものを感度良く検出するために，FAMなどで蛍光標識化されている．

タンパク質結合性DNAアプタマーのスクリー

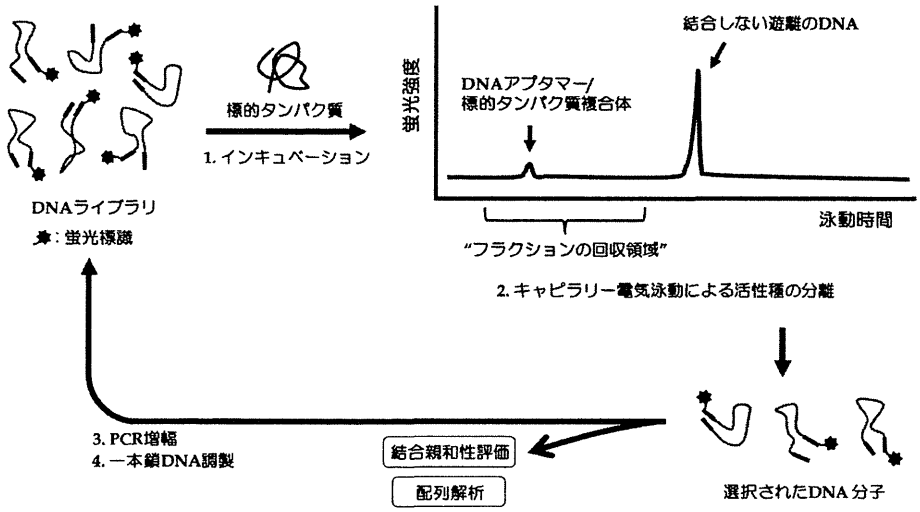


図6　CE-SELEX 法による DNA アプタマーの作製スキーム

ニングでは，一般に DNA ライブラリは標的分子と混合した後，非平衡条件下でのキャピラリーゾーン電気泳動法により，電気泳動する(図7)．電圧を印加するとキャピラリー中では陽極から陰極に向けて電気浸透流(EOF)が流れるので，サンプルは陽極側に検出器は陰極側にセットする．キャピラリー中の物質の泳動速度は，分子量が小さく正に帯電しているものほど速い．多くの場合，標的タンパク質と複合体を形成する DNA は，結合しない遊離の DNA よりも早く陰極側に到達し，複合体と遊離の DNA のピークは明瞭に分離できる．しかし，標的タンパク質の分子量が大きい場合，複合体のピークと遊離の DNA のピークが近くなり，両者を十分に分離できないことがある．その場合には，ポリビニルアルコールなどで内壁がコートされたキャピラリーを用いて分離を行う．コートされたキャピラリーでは EOF が生じないので，標的タンパク質の等電点より高い pH の緩衝液中で電気泳動を行うと，サンプルは陰極側から陽極側に向かって流れる．したがって，フューズドシリカキャピラリーを用いる分離とは反対に，サンプルは陰極側に検出器は陽極側にセットする．遊離の DNA は複合体よりも先に検出器に到達する．

泳動中，DNA と標的タンパク質の結合解離は非平衡状態であり，DNA は徐々に標的タンパク質から解離する．したがって，キャピラリーを長くすることにより，解離速度の遅い DNA アプタマーが得られると考えられる．ただし，長いキャピラリーを用いると泳動に要する時間も長くなってしまう．キャピラリーの内径を大きくすれば，一度にロードするサンプル量を多くすることがで

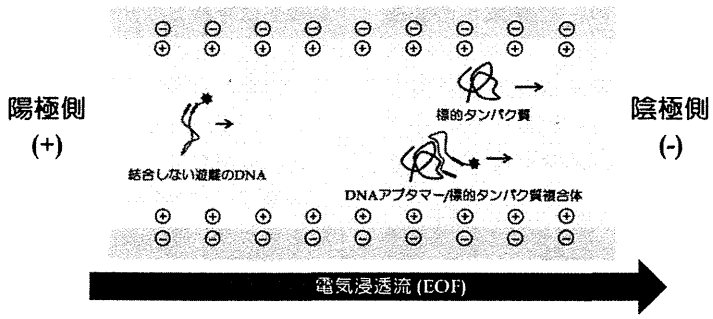


図7　フリーゾーンキャピラリー電気泳動(FZCE)法を用いた DNA アプタマー分離のしくみ

きる．しかし，同時に電気泳動で生じるジュール熱の放散効率が低下するため，DNA アプタマーの活性構造が熱変性によって不安定化する可能性が高まる．したがって，用いるキャピラリー長さと直径の最適化は，目的の DNA アプタマーを効率的に得る上で大変重要である．多くの場合，直径 50 もしくは 75 μm，長さ 31～80 cm のキャピラリーが使用される．

CE-SELEX の欠点として，用いるライブラリの規模に制約があることが挙げられる．つまり，一回の泳動に，数 nL しかサンプル液を注入できないからである．従来の SELEX 法では，通常，10^{13}～10^{15} 通りの配列を含むオリゴヌクレオチドを初期ライブラリとするが，CE-SELEX で同等の規模のスクリーニングを行おうとすると，数十～数百 mM のライブラリ溶液を用いなければならない計算となる．このように濃い DNA 濃度では，キャピラリー電気泳動で DNA アプタマー/標的タンパク質複合体と遊離の DNA とを分離することは非常に困難である．したがって，CE-SELEX の場合，初期ライブラリの規模は概ね 10^{10}～10^{13} である．

このような制約にも関わらず，CE-SELEX はネガティブセレクションを必要としないスクリーニング操作より，少ないラウンド数で目的の核酸アプタマーを高効率で得ることができる．例えば，CE-SELEX による IgE 結合性 DNA アプタマーのスクリーニングでは，わずか 2 回のラウンド数で，濃縮されたライブラリ中のほとんど配列が結合活性を示し，ほぼ 100%の濃縮効率が達成されている[18]．このような速い活性種の濃縮は，従来の SELEX 法では見られない．典型的な SELEX 法によるスクリーニングでは，最初の数ラウンドで濃縮効率が 50%を超えるのは稀である．さらに CE-SELEX の発展型として，スクリーニング操作の繰り返しを必要としない Non-SELEX 法も開発されている[4]．

このように，高い分離能をもつキャピラリー電気泳動法を SELEX 法に応用することにより，核酸アプタマーの作製を劇的に速めることが可能となった．筆者らは，架橋型ヌクレオチド K を含む人工核酸ライブラリを用いた同法によるスクリーニングで，血液凝固因子であるトロンビンに特異的に結合する人工核酸アプタマーを得た．標的分子に対する結合親和性（K_d 値）は，非平衡キャピラリー電気泳動法によって測定することが可能である．キャピラリー電気泳動のエレクトログラムにおいて，得られた人工核酸アプタマーとトロンビンの複合体の顕著なピークが観察されたことから，標的分子に対し高い結合親和性を示すことが分かった（図８）．K_d 値は 1 nM 未満（サブナノモラー）であった．

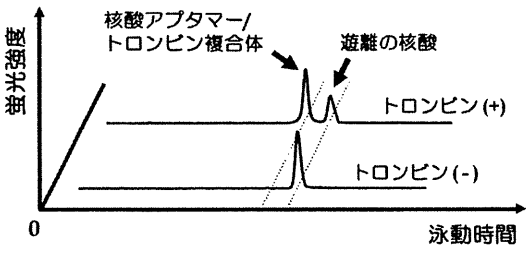


図８　得られた人工核酸アプタマーのエレクトログラム：トロンビンあり（上），トロンビンなし（下）

５．おわりに

自動化されたキャピラリー電気泳動装置を用いるスクリーニング法によって，高い結合親和性を示す種々の人工核酸アプタマーを簡便かつ迅速に作製することができると考えられる．また，架橋型ヌクレオチドの導入は，劇的に核酸アプタマーの生体内安定性を高めることが確認された．さらに，架橋構造差異によってその効果は大きく変わることも分かった．同時に，ポリメラーゼによる架橋型ヌクレオチドの取り込みでは，架橋構造がコンパクトな方が有利であるという知見も得られたため，今後，分子設計においてそれらの両立が鍵となると考えられる．一方でポリメラーゼの改変も有効であることが示唆された．したがって，改変部位と修飾基質に対する触媒活性との相関データから，人工核酸合成に最適な改変体を創製することによって，人工核酸ライブラリの化学レパートリを大幅に拡張されることが期待される[19～21]．現在，架橋型ヌクレオチドや塩基部位修飾ヌクレオチドの導入により，高い生体内安定性と結合親和性・多様性をもつ人工核酸アプタマーのスクリーニングを行っている．

謝　辞

本研究は独立行政法人新エネルギー・産業技術総合開発機構(NEDO)平成20年度産業技術研究助成事業より研究費の支援を受けています．また，*KOD* DNA ポリメラーゼ変異体は東洋紡積株式会社より提供されたものです．

文　献

1) A.D. Ellington, J.W. Szostak, *Nature*, **346**, 818-822 (1990)
2) C. Tuerk, L. Gold, *Science*, **249**, 505-510 (1990)
3) S.D. Mendonsa, M.T. Bowser, *J. Am. Chem. Soc.*, **127**, 9382-9383 (2005)
4) M. Berezovski, M. Musheev, A. Drabovich, S.N. Krylov, *J. Am. Chem. Soc.*, **128**, 1410-1411 (2006)
5) D.A. Daniels, H. Chen, B.J. Hicke, K.M. Swiderek, L. Gold, *Proc. Natl. Acad. Sci. USA.*, **100**, 15416-15421 (2003)
6) D. Shangguan, L. Meng, Z.C. Cao, Z. Xiao, X. Fang, Y. Li, D. Cardona, R.P. Witek, C. Liu, W. Tan, *Anal Chem.*, **80**, 721-728 (2008)
7) 菜原正靖，CSJ カレントレビュー：核酸化学のニュートレンド，日本化学会編，(化学同人)，78-85 (2011)
8) M. Kuwahara, J. Nagashima, M. Hasegawa, T. Tamura, R. Kitagata, K. Hanawa, S. Hososhima, T. Kasamatsu, H. Ozaki, H. Sawai, *Nucleic Acids Res.*, **34**, 5383-5394 (2006)
9) M. Kuwahara, S. Obika, J. Nagashima, Y. Ohta, Y. Suto, H. Ozaki, H. Sawai, and T. Imanishi, *Nucleic Acids Res.*, **36**, 4257-4265 (2008)
10) M. Kuwahara, H. Takeshima, J. Nagashima, S. Minezaki, H. Ozaki, H. Sawai, *Bioorg. Med. Chem.*, **17**, 3782-3788 (2009)
11) M. Kuwahara, Y. Takano, Y. Kasahara, H. Nara, H. Ozaki, H. Sawai, A. Sugiyama, S. Obika, *Molecules*, **15**, 8229-8240 (2010)
12) T. Kajiyama, M. Kuwahara, M. Goto, H. Kambara, *Anal. Biochem.*, **416**, 8-17 (2011)
13) T. Imanishi, S. Obika, *J. Syn. Org. Chem., Jpn.*, **57**, 969-980 (1999)
14) K. Miyashita, S.M.A. Rahman, S. Seki, S. Obika, T. Imanishi, *Chem. Commun.*, 3765-3767 (2007)
15) M. Nishida, T. Baba, T. Kodama, A. Yahara, T. Imanishi, S. Obika, *Chem. Commun.*, **46**, 5283-5285 (2011)
16) M. Kuwahara, S. Obika, H. Takeshima, Y. Hagiwara, J. Nagashima, H. Ozaki, H. Sawai, T. Imanishi, *Bioorg. Med. Chem. Let.*, **19**(11), 2941-2943 (2009)
17) Y. Kasahara, S. Kitadume, M. Kuwahara, H. Ozaki, H. Sawai, K. Morihiro, T. Imanishi, S. Obika, *Bioorg. Med. Chem. Lett.*, **20**, 1626-1629 (2010)
18) S.D. Mendonsa, M.T. Bowser, *Anal. Chem.*, **76**, 5387-5392 (2004)
19) M.M. Masud, M. Kuwahara, H. Ozaki, H. Sawai, *Bioorg. Med. Chem.*, **12**, 1111-1120 (2004)
20) A. Shoji, M. Kuwahara, H. Ozaki, H. Sawai, *J. Am. Chem. Soc.*, **129**, 1456-1464 (2007)
21) K. Ohsawa, T. Kasamatsu, J. Nagashima, K. Hanawa, M. Kuwahara, H. Ozaki, H. Sawai, *Anal. Sci.*, **24**, 167-172 (2008)