A 2′,4′-Bridged Nucleic Acid Containing 2-Pyridone as a Nucleobase: Efficient Recognition of a C·G Interruption by Triplex Formation with a Pyrimidine Motif**

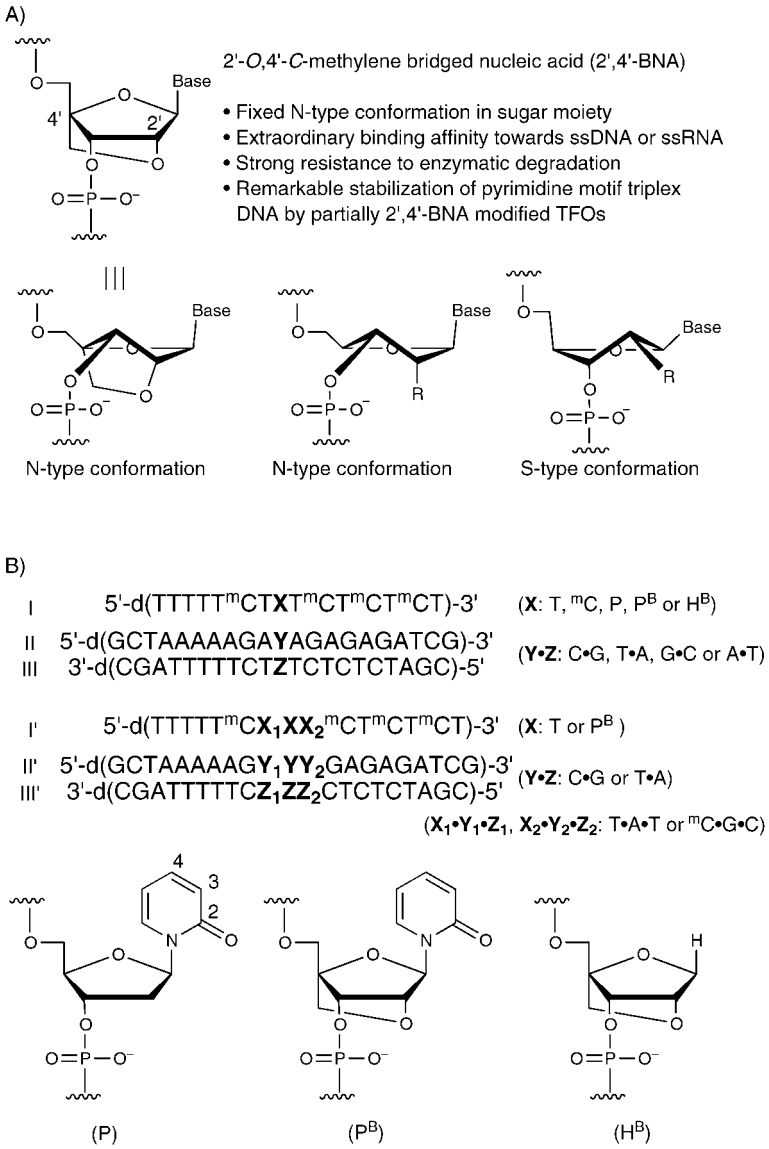
Satoshi Obika, Yoshiyuki Hari, Mitsuaki Sekiguchi, and Takeshi Imanishi*

Triplex formation between duplex DNA and a triplex-forming oligonucleotide (TFO) has been noted in practical applications of antigene methodology and in genomic analysis. In the triplex DNA formed with the pyrimidine motif, the homopyrimidine TFO binds to the homopurine tract of the target duplex DNA in a sequence-specific manner through the formation of Hoogsteen hydrogen bonds to form T·A·T and C^+·G·C triads. Interruption of the homopurine sequence of the target duplex DNA by a pyrimidine nucleotide causes destabilization of the triplex. Chemical modification of the TFOs in the nucleobase moiety has been reported to give efficient recognition of a pyrimidine·purine base pair, such as a C·G or T·A base pair.[1, 2] However, development of practical TFOs for recognition of general duplex DNA sequences is still pending.

We recently achieved the first synthesis of a novel nucleoside with a fixed N-type conformation (C-3′-*endo*),[3] namely, 2′-*O*,4′-*C*-methyleneribonucleic acid (2′-*O*,4′-*C*-methylene *B*ridged *N*ucleic *A*cid 2′,4′-BNA; Scheme 1 A)[4, 5] and found that pyrimidine oligonucleotides with partial 2′,4′-BNA modification exhibited strong triplex-forming abilities at neutral pH values.[6–9]

From the fact that T or C moderately interacts with a C·G base pair in homopurine·homopyrimidine double-stranded DNA (dsDNA),[10–12] we considered that the 2-carbonyl oxygen atom of T or C plays an important role in the recognition of a C·G base pair. Therefore, we selected the 2-pyridone group,[13, 14] which has only a 2-carbonyl group and lacks the 3-nitrogen atom and a 4-carbonyl or amino group of thymine or cytosine (Scheme 1 B), as a nucleobase to interact with a C·G base pair. Here, we report the synthesis of the novel 2′,4′-BNA monomer **1** bearing a 2-pyridone group, and its application to the efficient recognition of C·G interruption in homopurine·homopyrimidine dsDNA.

The synthetic route to **1** is shown in Scheme 2. The starting material **2**,[15] which was easily prepared from D-glucose, was treated with 2-pyridone, *N*,*O*-bis(trimethylsilyl)acetamide (BSA), and trimethylsilyl trifluoromethanesulfonate (TMSOTf) in dichloroethane to give **3**. Bicyclic nucleoside **4** was obtained upon exposure of **3** to potassium carbonate in methanol. Next, hydrogenolysis of **4** afforded the desired compound **1**. The structure of 2′,4′-BNA monomer **1** was confirmed by X-ray crystallographic analysis (Figure 1),[16] which revealed that the sugar moiety in **1** was fixed in an N-type conformation (pseudorotation phase angle *P* is 16.7°), the same as that in the 2′,4′-BNA uracil monomer (*P* is 17.4°).[4] The phosphoramidite **6**, the suitable building block for DNA synthesis, was obtained by dimethoxytritylation (**1**→**5**) and phosphitylation (**5**→**6**), and then incorporated into TFOs I and I′ (Scheme 1 B) by a standard phosphoramidite protocol on a DNA synthesizer. The purity of the modified TFOs was verified by using reversed-phase HPLC, and the compositions were determined by matrix-assisted laser desorption/ionization time-of-flight (MALDI-TOF) MS.[17]

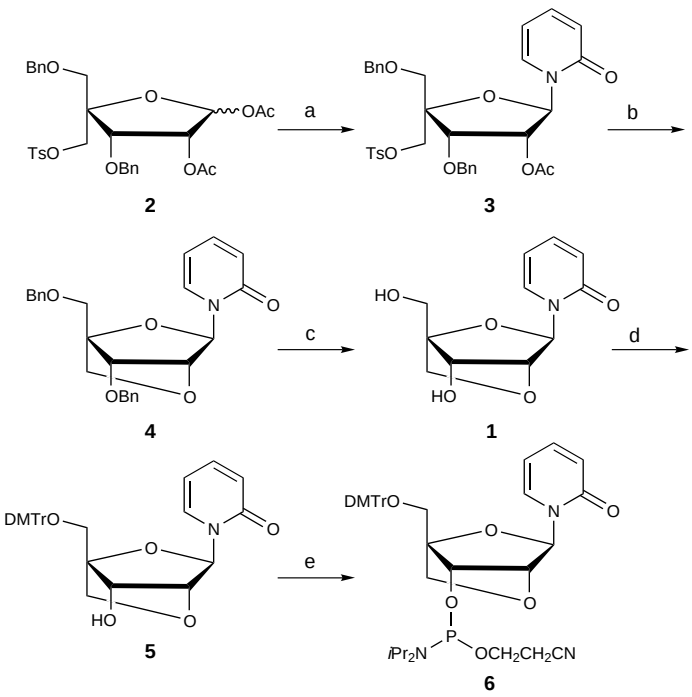


Scheme 1. A) Structure and characteristics of 2′,4′-BNA. B) The oligonucleotides used in this study.

[*] Prof. Dr. T. Imanishi, Dr. S. Obika, Y. Hari, M. Sekiguchi
Graduate School of Pharmaceutical Sciences
Osaka University
1-6 Yamadaoka, Suita, Osaka 565-0871 (Japan)
Fax: (+81) 6-6879-8204
E-mail: imanishi@phs.osaka-u.ac.jp

[**] Part of this work was supported by a Grant-in-Aid for Scientific Research (B) (No. 12557201) from the Japan Society for the Promotion of Science.

  1433-7851/01/4011-2079 $ 17.50+.50/0



Scheme 2. Synthesis of 2′,4′-BNA-pyridone monomer **1** and its phosphoramidite derivative **6**. a) 2-pyridone, BSA, TMSOTf, dichloroethane, reflux, 74%; b) K_2CO_3, MeOH, RT, 100%; c) 20% $Pd(OH)_2/C$, cyclohexene, EtOH, reflux, 95%; d) DMTrCl, pyridine, RT, 96%; e) 2-cyanoethyl *N,N,N′,N′*-tetraisopropylphosphorodiamidite, diisopropylammonium tetrazolide, MeCN/THF, RT, 98%. Bn = benzyl, Ac = acetyl, Ts = toluene-4-sulfonyl, DMTr = dimethoxytrityl.

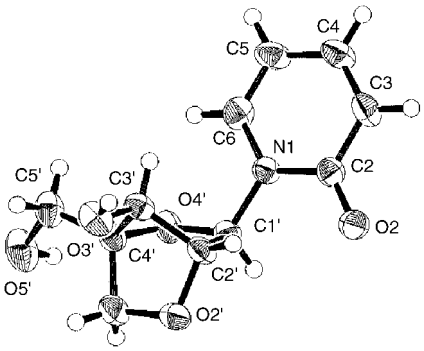


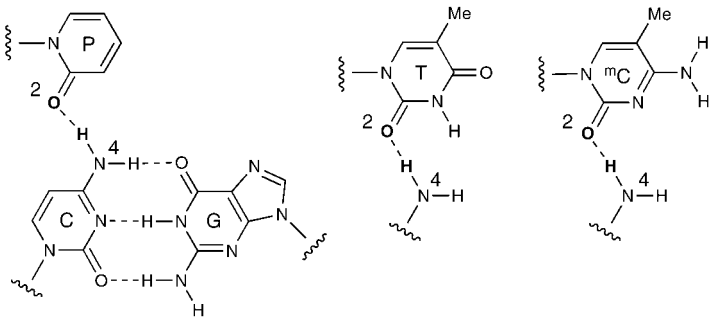
Figure 1. X-ray structure of 2′,4′-BNA-pyridone monomer **1**.

The melting temperatures (T_m) for the triplexes formed by TFO I containing the 2′,4′-BNA-2-pyridone monomer (P^B) were compared with those of triplexes where TFO I contained a natural nucleobase (T or ^mC), a DNA-2-pyridone monomer (P),[14] or the 2′,4′-BNA abasic monomer (H^B, Table 1).[7, 18] At first, it was found that the triplex I · II · III containing a P · C · G triad was as stable as the triplexes I · II · III containing a T · C · G or ^mC · C · G triad, which indicates the great importance of the 2-carbonyl oxygen atom of T, ^mC, and P in the recognition of a C · G base pair.[19] Furthermore, as we expected, the TFO I containing P showed high sequence selectivity, while that containing T or ^mC preferentially recognizes an A · T or G · C base pair. The proposed hydrogen-bonding pattern for the P · C · G triad is shown in Scheme 3.

Table 1. T_m values [°C] of the triplex formed between TFO I and the target duplex II · III.[a]

	Y · Z			
X	C · G	G · C	T · A	A · T
T	25	20	17	44
^mC	25	43	16	18
P	**24**	16	15	15
P^B	**33**	19	14	23
H^B	24	20	20	16

[a] UV melting profiles were measured in 7 mM sodium phosphate buffer (pH 7.0) containing 140 mM KCl and 10 mM $MgCl_2$ at a scan rate of 0.5 K min^{-1} at 260 nm. The oligonucleotide concentration used was 1.5 μM for each strand.



Scheme 3. Proposed structures of the P · C · G, T · C · G, and ^mC · C · G triads.

Triplex formation between TFO I (X = P^B) and the target duplex II · III was also found to be sequence selective, with the triplex I · II · III containing a P^B · C · G triad being the most stable in the triplexes I · II · III with a P^B · X · Y triad (Figure 2A). Moreover, the thermal stability of the triplex I · II · III containing a P^B · C · G triad is preferable to that of the triplex I · II · III containing a T · C · G triad (Figure 2B). Comparison of the T_m values of the triplexes with P^B · C · G (33 °C) and P · C · G triad (24 °C) demonstates that the level of triplex stabilization induced by the 2′,4′-BNA modification of the 2-pyridone derivative for recognition of the C · G interruption ($\Delta T_m = +9$ K) is comparable to that caused by the 2′,4′-BNA modification of thymidine or 5-methylcytidine in the fully matched triplexes.[7]

To confirm the general applicability of this result, the ability of P^B to recognize the C · G interruption in other dsDNA targets was also studied (Table 2). Replacement of an A · T base pair by a G · C base pair at the neighboring sites of the C · G interruption caused a decrease in the T_m values as a result of the low stability of a ^mC · G · C triad under neutral conditions. However, the TFOs I′ (X = P^B) preserved the reasonable triplex-forming ability towards the corresponding target duplex (Y · Z = C · G). Furthermore, the triplexes I′ · II′ · III′ containing a P^B · C · G triad had much superior thermal stability than any other triplexes I′ · II′ · III′ having a P^B · T · A, T · C · G, or T · T · A triad.

To the best of our knowledge, P^B is one of the best nucleic acid analogues for recognizing a C · G base pair since it gives a significant increase in binding affinity without loss of selectivity. Considering that the TFO I containing the 2′,4′-BNA abasic monomer (H^B) had only a slight effect on the triplex

  1433-7851/01/4011-2080 $ 17.50+.50/0

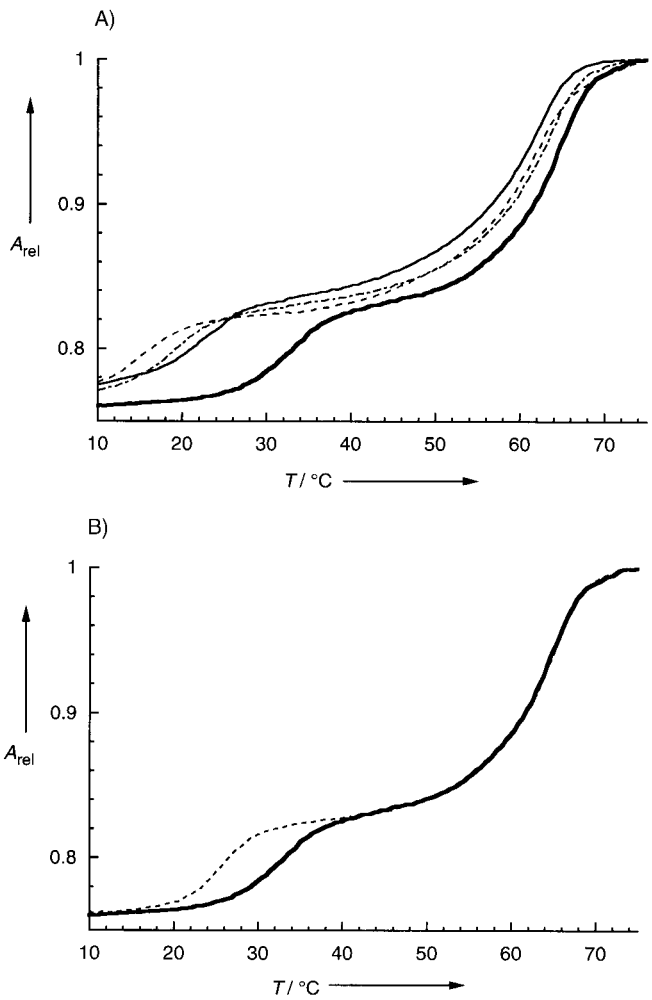


Figure 2. T_m profiles of triplex I·II·III containing a A) $P^B \cdot C \cdot G$ (——), $P^B \cdot T \cdot A$ (- - - -), $P^B \cdot G \cdot C$ (•–•–), or $P^B \cdot A \cdot T$ (——) triad and B) a $P^B \cdot C \cdot G$ (——) or T·C·G (- - - -) triad.

Table 2. T_m values [°C] of the triplex formed from the TFO I′ and the target duplex II′·III′.[a]

	Y·Z			Y·Z	
X_1-X-X_2	C·G	T·A	X_1-X-X_2	C·G	T·A
T-T-^{m}C	16	ca. 10	^{m}C-P^B-T	**21**	–[b]
T-$P^{B\text{-}m}$C	**27**	ca. 10	^{m}C-T-^{m}C	–[b]	–[b]
^{m}C-T-T	13	ca. 10	^{m}C-$P^{B\text{-}m}$C	**16**	–[b]

[a] See captions in Table 1. [b] Typical hyperchromicity for the triplex dissociation was not observed.

stabilization, the unprecedented C·G recognition ability of P^B would be attributable to the 2′-*O*,4′-*C*-methylene bridged ribofuranose moiety coupled with an appropriate hydrogen bond between the 2-carbonyl oxygen atom in P^B and the 4-amino group in C.

The results presented here demonstrate that the 2-carbonyl oxygen atoms of pyridone and pyrimidine nucleobases play an important role in the recognition of a C·G base pair. In contrast to T or ^{m}C, the 2-pyridone derivative (P) has a reasonable C·G selectivity because of the lack of a 3-nitrogen atom and a 4-carbonyl or amino group which are crucial for hydrogen bonding with other base pairs. The combination of 2-pyridone and the 2′-*O*,4′-*C*-methylene bridged sugar moiety significantly enhances the binding affinity with a C·G base pair, without loss of selectivity. The application of the present findings to the regulation of gene expression is currently under way.

Received: December 4, 2000
Revised: March 2, 2001 [Z16220]

[1] I. Luyten, P. Herdewijn, *Eur. J. Med. Chem.* **1998**, *33*, 515–576, and references therein.
[2] a) S. O. Doronina, J.-P. Behr, *Chem. Soc. Rev.* **1997**, 63–71, and references therein; b) I. Prévot-Halter, C. J. Leumann, *Bioorg. Med. Chem. Lett.* **1999**, *9*, 2657–2660.
[3] W. Saenger, *Principles of Nucleic Acid Structure*, Springer, Berlin, **1984**, p. 17.
[4] S. Obika, D. Nanbu, Y. Hari, K. Morio, Y. In, T. Ishida, T. Imanishi, *Tetrahedron Lett.* **1997**, *38*, 8735–8738.
[5] We defined BNA as a novel class of nucleic acid analogues containing the 2′-*O*,4′-*C*- or 3′-*O*,4′-*C*-methylene bridged structure (2′,4′-BNA or 3′,4′-BNA, respectively). The 2′,4′-BNA has also been called "LNA"; see S. K. Singh, P. Nielsen, A. A. Koshkin, J. Wengel, *Chem. Commun.* **1998**, 455–456.
[6] T. Imanishi, S. Obika, *J. Synth. Org. Chem. Jpn.* **1999**, *57*, 969–980.
[7] S. Obika, Y. Hari, T. Sugimoto, M. Sekiguchi, T. Imanishi, *Tetrahedron Lett.* **2000**, *41*, 8923–8927.
[8] H. Torigoe, Y. Hari, M. Sekiguchi, S. Obika, T. Imanishi, *J. Biol. Chem.* **2001**, *276*, 2354–2360.
[9] S. Obika, T. Uneda, T. Sugimoto, D. Nanbu, T. Minami, T. Doi, T. Imanishi, *Bioorg. Med. Chem.*, in press.
[10] K. Yoon, C. A. Hobbs, J. Koch, M. Sardaro, R. Kutny, A. L. Weis, *Proc. Natl. Acad. Sci. USA* **1992**, *89*, 3840–3844.
[11] J. L. Mergny, D. Collier, M. Rougée, T. Montenay-Garestier, C. Hélène, *Nucleic Acids Res.* **1991**, *19*, 1521–1526.
[12] B. P. Belotserkovskii, A. G. Veselkov, S. A. Filippov, V. N. Dobrynin, S. M. Mirkin, M. D. Frank-Kamenetskii, *Nucleic Acids Res.* **1990**, *18*, 6621–6624.
[13] U. Niedballa, H. Vorbrüggen in *Nucleic Acid Chemistry*, *Part 1* (Eds.: L. B. Townsend, R. S. Tipson), Wiley, New York, **1978**, pp. 481–484.
[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. A. Koshkin, V. K. Rajwanshi, J. Wengel, *Tetrahedron Lett.* **1998**, *39*, 4381–4384.
[16] Crystallographic data (excluding structure factors) for the structure reported in this paper have been deposited with the Cambridge Crystallographic Data Centre as supplementary publication no. CCDC-152458. Copies of the data can be obtained free of charge on application to CCDC, 12 Union Road, Cambridge CB2 1EZ, UK (fax: (+44) 1223-336-033; e-mail: deposit@ccdc.cam.ac.uk).
[17] MALDI-TOF-MS data: *m/z* (calcd): TFO I (X = P^B) $[M-H]^-$ 4493.37 (4493.05); TFO I (H^B) $[M-H]^-$ 4400.41 (4399.97); TFO I′ (X_1-X-X_2 = T-$P^{B\text{-}m}$C) $[M-H]^-$ 4491.60 (4492.07); TFO I′ (^{m}C-P^B-T) $[M-H]^-$ 4492.83 (4492.07); TFO I′ (^{m}C-$P^{B\text{-}m}$C) $[M-H]^-$ 4490.48 (4491.08).
[18] L. Kværnø, J. Wengel, *Chem. Commun.* **1999**, 657–658.
[19] Recently, Prévot-Halter and Leumann proposed that the 3-nitrogen atom in pyrimidine contributed to the hydrogen bonding of the 4-amino group in C. See ref. [2b].

  1433-7851/01/4011-2081 $ 17.50+.50/0