"Paper-Clip" Type Triple Helix Formation by 5′-d-$(TC)_3T_a(CT)_3C_b(AG)_3$ (a and b = 0−4) as a Function of Loop Size with and without the Pseudoisocytosine Base in the Hoogsteen Strand†

Tsung-Mei Chin,*,‡ Shwu-Bin Lin,§ Szu-Yin Lee,‡ Mu-Li Chang,‡ Alice Yu-Yu Cheng,‡,|| Fu-Chao Chang,‡,|| Laura Pasternack,⊥ Dee-Hua Huang,⊥ and Lou-Sing Kan*,||

Institute of Applied Chemistry, Chinese Culture University, Graduate Institute of Medical Technology, National Taiwan University, and Institute of Chemistry, Academia Sinica, Taipei, Taiwan, and Department of Chemistry, The Scripps Research Institute, La Jolla, California 92037

Received February 23, 2000; Revised Manuscript Received July 17, 2000

ABSTRACT: The formation of a DNA "paper-clip" type triple helix (triplex) with a common sequence 5′-d-$(TC)_3T_a(CT)_3C_b(AG)_3$ (a and b = 0−4) was studied by UV thermal melting experiments and CD spectra. These DNA oligomers form triplexes and duplexes under slightly acidic and neutral conditions, respectively. The stability of the formed triplexes (at pH 4.5) or duplexes (at pH 7.0 or 8.0) does not vary significantly with the size of the loops (a and b = 1−4). At pH 6.0, the triplex stability is, however, a function of a and b. It is also interesting to note that the oligomer 5′-d-$(TC)_3(CT)_3(AG)_3$ (a and b = 0) forms a stable triplex at pH 4.5 with a slightly lower T_m value, due to dissociation of a base triad at one end and a distorted base triad at the other, observed by ^{1}H NMR. Thus, we have here a model system, 5′-d-$(TC)_3T_a(CT)_3C_b(AG)_3$, that could form a triplex effectively with (a and b = 1−4) and without (a and b = 0) loops under acidic conditions. In addition, the triplex formation of oligomers with replacement of one, two, or three 2′-deoxycytidine in the Hoogsteen strand by either 2′-deoxypseudoisocytidine (D) or 2′-*O*-methylpseudoisocytidine (M) was also studied in the sequence 5′-d-$(TX)_3T_2(CT)_3C_2(AG)_3$ (where X is C, D, or M). Both CD spectra and UV melting results showed that only D3 [$(TX)_3 = (TD)_3$] and M3 [$(TX)_3 = (TM)_3$] were able to form the paper-clip structure under both neutral and acidic conditions. This is because the N_3H of a pseudoisocytosine base can serve as a proton donor without protonation. We hereby proved that the 2′-deoxypseudoisocytidine, similar to 2′-*O*-methylpseudoisocytidine, could replace 2′-deoxycytidine in the Hoogsteen strand to provide triplex formation at neutral pH.

The study of oligodeoxyribonucleotide-directed triple helix (triplex) formation has become increasingly popular. This is because triplex formation could manipulate gene expression through several possible mechanisms (*1*), compete with DNA binding protein (*2*), and map chromosome via the linking of cleaving reagents at the ends of the Hoogsteen strand (*3*). To date, two classes of triple-helical binding motifs, pyrimidine·purine·pyrimidine and purine·purine·pyrimidine, have been studied extensively (*4*). Triplex formation of oligodeoxyribonucleotides with duplex DNA in the pyrimidine·purine·pyrimidine binding motif requires protonation of cytosine units at N_3 in the Hoogsteen strand (*5*). The N_3-protonated cytosine is, however, stable only under slightly acidic conditions, and hence, this kind of triplex formation is not favorable under physiological conditions. To overcome this obstacle, we have designed non-natural C-nucleosides carrying the charge-neutral pseudoisocytosine base ($^{\Psi}$C),[1] which could display the hydrogen-bonding pattern of a protonated cytosine (Figure 1). We have demonstrated that oligodeoxyribonucleotides containing the pseudoisocytosine base as a substitute for the cytosine base can indeed form triplexes under neutral conditions (*6*).

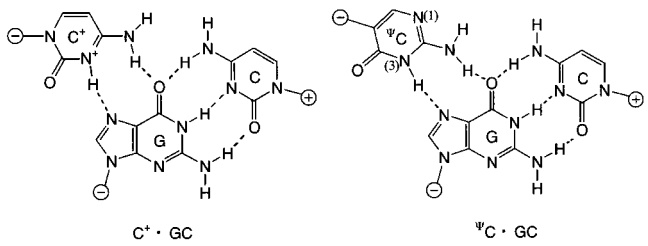

FIGURE 1: Proposed hydrogen-bonding schemes of the C$^+$·GC and $^{\Psi}$C·GC base triads.

In this paper, we present a simple DNA system that shows a robust propensity for triplex formation. This system is an extension of our previous hairpin type to a "paper-clip" type

† This work is supported by NSC Grants 89-2113-M-034-002 and 89-2113-M-001-043 to T.-M.C. and L.-S.K., respectively.

* To whom correspondence should be addressed. Telephone: 886-2-28610511, ext 585, or 886-2-27898550. Fax: 886-2-28616219 or 886-2-2783-1237. E-mail: tmchin@faculty.pccu.edu.tw or lskan@chem.sinica.edu.tw.

‡ Chinese Culture University.

§ National Taiwan University.

|| Academia Sinica.

⊥ The Scripps Research Institute.

[1] Abbreviations: $^{\Psi}$C, pseudoisocytosine; D, 2′-deoxypseudoisocytidine; M, 2′-*O*-methylpseudoisocytidine; $TIPDSCl_2$, 1,3-dichloro-1,1,3,3-tetraisopropyldisiloxane; DMF, *N*,*N*-dimethylformamide; THF, tetrahydrofuran; DMTrCl, 4,4′-dimethoxytrityl chloride.

10.1021/bi0004201 CCC: $19.00
Published on Web 09/13/2000

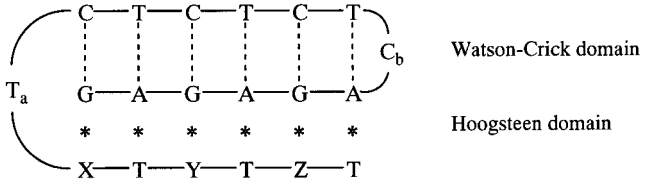

FIGURE 2: Folded conformation of a paper-clip type triplex with the common sequence 5′-d-$(TC)_3T_a(CT)_3C_b(AG)_3$ (a and $b = 0-4$) with and without the pseudoisocytosine base replacement in the Hoogsteen strand. X, Y, and Z are 2′-deoxycytidine (C), 2′-*O*-methylpseudoisocytidine (M), or 2′-deoxypseudoisocytidine (D).

Table 1: List of Synthetic Oligodeoxyribonucleotides[a]

sequence (5′−3′)	symbol
$TCTCTC(T)_0CTCTCT(C)_0AGAGAG$	T0C0
$TCTCTC(T)_1CTCTCT(C)_1AGAGAG$	T1C1
$TCTCTC(T)_2CTCTCT(C)_2AGAGAG$	T2C2
$TCTCTC(T)_3CTCTCT(C)_3AGAGAG$	T3C3
$TCTCTC(T)_3CTCTCT(C)_4AGAGAG$	T3C4
$TCTCTC(T)_4CTCTCT(C)_3AGAGAG$	T4C3
$TCTCTC(T)_4CTCTCT(C)_4AGAGAG$	T4C4
$TMTCTC(T)_2CTCTCT(C)_2AGAGAG$	M1
$TMTMTC(T)_2CTCTCT(C)_2AGAGAG$	M2
$TMTMTM(T)_2CTCTCT(C)_2AGAGAG$	M3
$TDTCTC(T)_2CTCTCT(C)_2AGAGAG$	D1
$TDTDTC(T)_2CTCTCT(C)_2AGAGAG$	D2
$TDTDTD(T)_2CTCTCT(C)_2AGAGAG$	D3

[a] M and D represent 2′-*O*-methylpseudoisocytidine and 2′-deoxypseudoisocytidine, respectively.

triplex (Figure 2) through a connection of the target sequence 5′-d-$(AG)_3$ with the probe 5′-d-$(TC)_3T_a(CT)_3$ by a d-C_b loop to form 5′-d-$(TC)_3T_a(CT)_3C_b(AG)_3$ (*7*). In searching for the most appropriate conditions under which the oligodeoxyribonucleotide can fold into a paper-clip type triplex by itself, we have synthesized and characterized a group of oligomers with a varying chain length of the loop (a and $b = 0-4$) as shown in Table 1.

An important variation of the system described above is replacement of 2′-deoxycytidine in the Hoogsteen strand by either 2′-deoxypseudoisocytidine (D) or 2′-*O*-methylpseudoisocytidine (M). Oligomers with one, two, and all three 2′-deoxycytidines substituted with D or M were synthesized, as summarized in Table 1. The physical properties of these oligomers were studied by measurement of UV thermal melting experiments and CD spectra. The results showed that the pseudoisocytosine base was able to form Hoogsteen hydrogen bonds to the G base and, thus, helped triplex formation under neutral conditions. The triplex containing M in the Hoogsteen strand gave higher stability than that with D at pH 7.0.

A preliminary structural study of 5′-d-$(TC)_3(CT)_3(AG)_3$ (namely, a and $b = 0$) by ^{1}H NMR at pH 4.5 is reported. We found that the four internal base triads formed typical pyrimidine·purine·pyrimidine hydrogen-bonding configurations. At the C^+·GC end of the triplex (Figure 2), three bases formed a distorted but stable triad. In contrast, at the other T·AT end of the triplex (Figure 2), no base triad was formed except a hairpin turn of TA associated with terminal T.

EXPERIMENTAL PROCEDURES

Materials. Pseudouridine was purchased from Kyowa Hakko Kogyo Co., Ltd. (Tokyo, Japan). Other chemicals were obtained from Aldrich or Sigma. Thin-layer chromatography was performed on MN silica gel 60 F_{254} plates and column chromatography on MN silica gel G60. All reagents were used as obtained commercially. The following buffers were used in CD and UV melting experiments. Distilled, deionized water was used for all aqueous solutions. Buffer A was 20 mM acetate buffer, 0.01 mM EDTA, and 150 mM NaCl (pH range of 4.5−6.0). Buffer B was 20 mM Tris buffer, 0.01 mM EDTA, and 150 mM NaCl (pH range of 7.0−8.0). Buffer C was 20 mM acetate buffer and 150 mM NaCl (pH 4.5). Buffer D was 20 mM Tris buffer and 150 mM NaCl (pH 7.0).

General Physical Methods for Synthesis of Phosphoramidite Synthons. IR spectra were recorded on a Perkin-Elmer 882 spectrophotometer. Standard spectra were regularly obtained for polystyrene to verify the accuracy. ^{1}H NMR and ^{13}C NMR spectra were obtained using Bruker AC-200 and AC-300 instruments. Chemical shifts were reported in parts per million relative to the internal standard of solvent residues (for ^{1}H and ^{13}C NMR in organic solutions) or the external standard of 85% phosphoric acid (for ^{31}P NMR). Mass spectra were recorded with a JEOL JMS-SX/SX102A or a JEOL JMS-HX110 spectrometer. The melting point was recorded without correction on a melting point apparatus (MEL-TEMP II Laboratory Devices).

UV Thermal Melting Experiment. The UV absorbance versus temperature profiles were measured by a JASCO V-560 UV/VIS spectrophotometer at 260 nm. Cells with 0.1, 1, 10, or 100 mm path lengths were used for concentration-dependent experiments. The temperature of three former cells was controlled by a JASCO TPU-436 temperature programmer and with a Peltier type thermostatic cell holder (EHC-441). The pH values of samples were measured by a JENCO pH meter (model 6071, Taiwan). The UV melting curves were recorded with heating and cooling rates of 0.5 °C/min. To keep the thermostatic holder working accurately, an aluminum space filler was made and inserted into a heat sinker with the thin cell (0.1 or 1 mm) during thermal melting experiments. A homemade water jacket was used for the temperature control of a 100 mm cell. A temperature-controlled water flow was slowly pumped into one side of the jacket from the bottom, and then flew out from top of the other side. Temperatures were monitored on both in and out flows. The temperature of the 100 mm cell was taken as the average of two readings (±1 °C).

CD Spectroscopy. CD spectra were obtained on a JASCO-720 circular dichroism spectropolarimeter. The temperature was controlled by a water-circulated jacketed cell. Dry nitrogen gas was blown through the cell compartment to expel oxygen and to prevent moisture condensation. Cells with 0.1, 1, 10, or 100 mm path lengths were used.

Two-Dimensional ^{1}H NMR Spectroscopy and Structure Calculation. The two-dimensional NOESY ^{1}H NMR experiments were carried out on a Bruker DRX-600 NMR spectrometer. The sample was dissolved in H_2O (9:1 H_2O:D_2O ratio) and 100% D_2O solvents containing 0.1 M NaCl and 0.01 M acetate buffer (pH 4.5). The final concentration of the T0C0 is 80 OD/0.5 mL. The spectrum was taken with a mixing time of 200 ms and 96 scans. The assignment of proton resonances was carried out by standard sequential analysis procedures (*8*) via two-dimensional spectroscopy

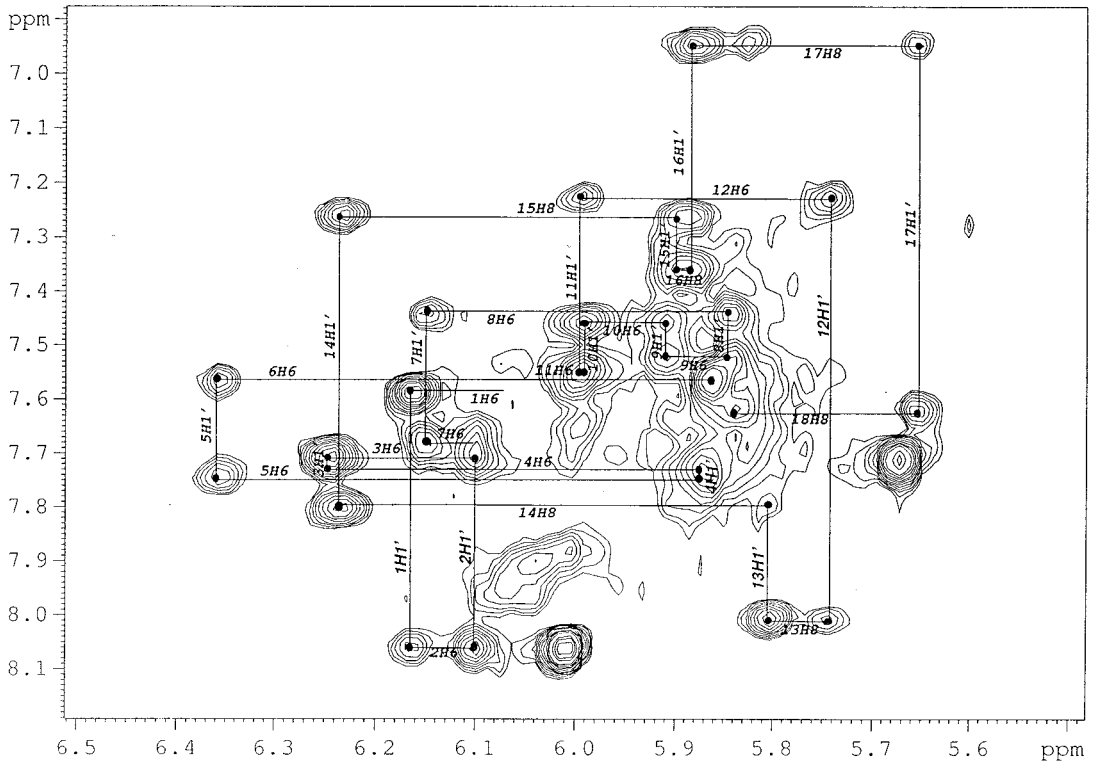

FIGURE 3: Two-dimensional NOESY spectrum of T0C0 in H_2O (see the text).

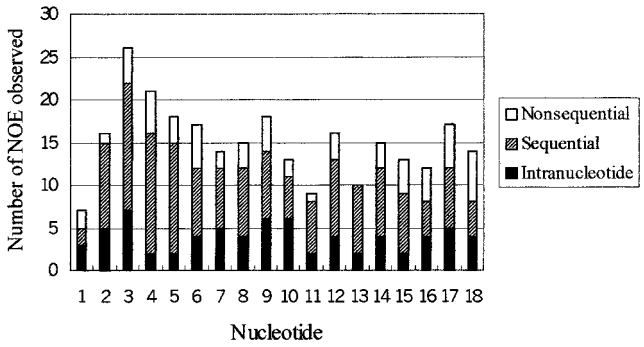

FIGURE 4: Graph of NOEs by residues of the triplex T0C0. Trivial intranucleotide NOEs are not shown.

at 1 °C. Only the $H_{6/8}-H_{1'}$ region of T0C0 is displayed in Figure 3. Distance restraints were determined by analysis of two-dimensional NOESY data (data not shown here). For nonexchangeable protons, the cross-peak intensities were converted to distances by comparison to the geminal protons at $C_{2'}$ (1.9 Å), as well as grouped as strong (s, 1.8−2.4 Å), medium (m, ≤3.5 Å), and weak (w, ≤5 Å). For exchangeable protons, a uniform restraint of 1.8−5 Å was used since exchange with water made peak integrals independent of the proton−proton distance. Figure 4 shows the total NOEs observed from each nucleotide. The model was assembled as a single continuous strand using the MSI software. All calculations were performed using Discover and the Amber force field. Calculations occurred in a two-step process: (i) simulated annealing without explicit water using a distant-dependent dielectric constant of 4*r* to establish the backbone conformation and base orientation and (ii) 1000-step minimization in explicit water.

Synthesis of Oligodeoxyribonucleotides. All oligodeoxyribonucleotides were synthesized on a DNA synthesizer (Applied Biosystems model 391) using solid-support phosphoramidite chemistry (*9*). The detail synthetic procedures for 2′-deoxypseudoisocytidine and 2′-*O*-methylpseudoisocytidine phosphoramidite synthons are described in the Supporting Information. The coupling time during incorporation of 2′-*O*-methylpseudoisocytidine or 2′-deoxypseudoisocytidine into oligodeoxyribonucleotides was extended to provide warranty for the yields. The oligodeoxyribonucleotides, after deblocked and cleaved from the solid support, were purified by reverse-phase cartridges (Poly-Pak, Glen Research Corp.), followed by HPLC. HPLC was performed on a JASCO PU-980 (JASCO) with Hypersil BDS-C18 columns (4.6 mm and 10 mm × 250 mm). The chain length and purity of each oligodeoxyribonucleotide were verified with gel electrophoresis. The concentrations of single-stranded oligodeoxyribonucleotides were determined at 260 nm, using the following molar extinction coefficients based on the Fasman method (*10*): 165 680 cm^{-1} M^{-1} (T0C0), 180 040 cm^{-1} M^{-1} (T1C1), 196 000 cm^{-1} M^{-1} (T2C2, D1, D2, D3, M1, M2, and M3), 211 960 cm^{-1} M^{-1} (T3C3), 219 400 cm^{-1} M^{-1} (T3C4), 220 480 cm^{-1} M^{-1} (T4C3), and 227 820 cm^{-1} M^{-1} (T4C4).

RESULTS AND DISCUSSION

Synthesis of 2′-O-Methylpseudoisocytidine and 2′-Deoxypseudoisocytidine Phosphoramidites. Both of our targets, 2′-deoxy and 2′-*O*-methyl derivatives of pseudoisocytidine, can be commonly converted from pseudouridine by reactions with guanidine (*11*). A concern, however, is that pseudouridine undergoes anomerization easily when it is heated in acid or in alkali (*12*). Accordingly, it is important to perform the chemical reactions under neutral conditions and protect the ring nitrogens. The synthetic routes to phosphoramidite synthons of 2′-deoxy- and 2′-*O*-methylpseudoisocytidine are shown in Figure 5.

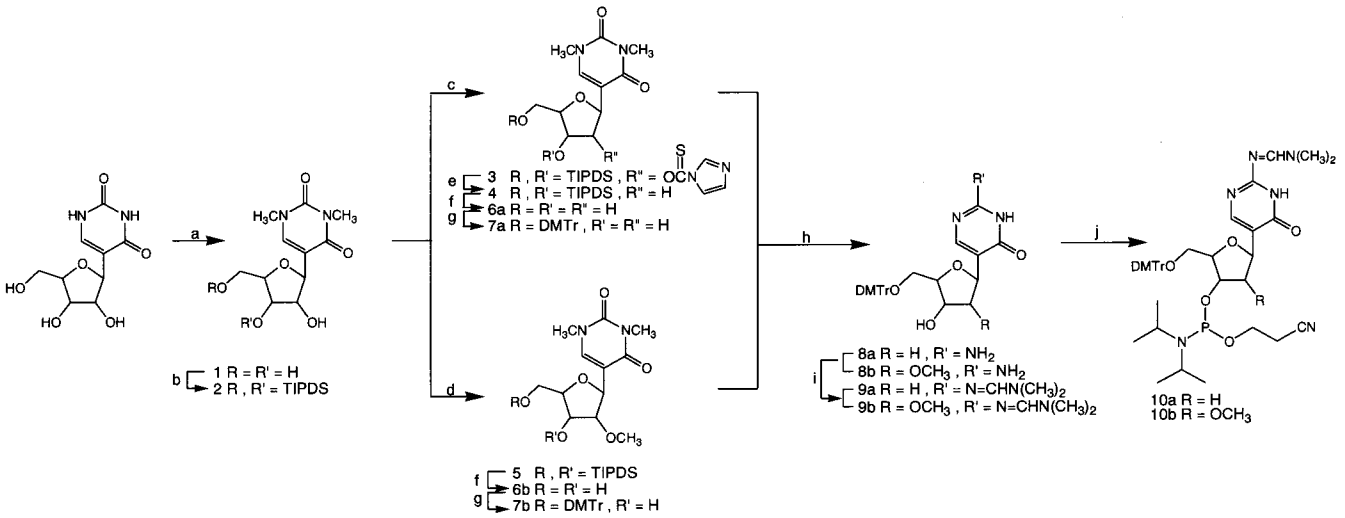

FIGURE 5: Synthetic routes for phosphoramidite synthons of 2′-deoxy- and 2′-*O*-methylpseudoisocytidine: (a) *N*,*N*-dimethylformamide dimethyl acetal, refluxing; (b) $TIPDSCl_2$, pyridine, −15 °C; (c) 1,1′-thiocarbonyldiimidazole, DMF, 60 °C; (d) CH_3I, Ag_2O, benzene, 35 °C; (e) 2,2′-azobis(2-methylpropionitrile), *n*-Bu_3SnH, toluene, refluxing; (f) *n*-Bu_4NF, THF, 0 °C; (g) DMTrCl, pyridine, room temperature; (h) guanidine hydrochloride, sodium ethoxide, ethanol, refluxing under N_2, room temperature; (i) *N*,*N*-dimethylformamide dimethyl acetal, DMF, room temperature; and (j) 2-cyanoethyl *N*,*N*-diisopropylphosphoramidous chloride, diisopropylethylamine, CH_2Cl_2, room temperature.

To prevent undesirable α,*β*-anomerization in the following reactions, the ring nitrogens N_1 and N_3 of pseudouridine were first methylated by refluxing with *N*,*N*-dimethylformamide dimethyl acetal to yield the 1,3-dimethyl derivative **1**. Subsequently, the 3′- and 5′-hydroxyl groups of **1** were regioselectively silylated by treatment with 1,3-dichloro-1,1,3,3-tetraisopropyldisiloxane in pyridine to give the 3′,5′-protected nucleoside **2** in which only the 2′-hydroxyl group was available for modification. Nucleoside **2** was, then, converted into 2′-deoxy derivative **4** and 2′-*O*-methyl derivative **5**. The 2′-*O*-methyl derivative **5** was prepared via methylation of the appropriately protected nucleoside **2** with methyl iodide in the presence of silver(I) oxide. On the other hand, the deoxygenation was performed in two steps. Conversion of **2** with 1,1′-thiocarbonyldiimidazole, followed by tri-*n*-butyltin hydride treatment, afforded the deoxygenated nucleoside **4**. It is interesting to note that under the condition employed for the synthesis of **3**, no formation of anhydro nucleoside or dinucleosidyl thiocarbonate occurs as they are rather common byproducts in thiocarbonylation (*13*). Moreover, the subsequent reduction of thioester **3** proceeded slowly and required a large amount of tri-*n*-butyltin hydride. It was observed that the yields of reduction increased with dilution. For optimum conversion, 2′-deoxynucleoside **4** could be obtained in an overall yield of 70% from **2** under the conditions employed here. This improved procedure for the synthesis of 2′-deoxy C-nucleosides should have wide applicability in the studies of N- and C-nucleosides. Desilylation of **4** and **5** was affected by treatment with *n*-Bu_4NF in THF at 0 °C and afforded quantitatively 1,3-dimethyl-2′-deoxypseudouridine **6a** and 1,3,2′-*O*-trimethylpseudouridine **6b**, respectively. The results indicated that anomerization did not occur during steps of reduction and deprotection.

α,*β*-Anomerization is possibly associated with conversion of pseudouridine to pseudoisocytidine by reaction with guanidine under alkali conditions (*11*). It was observed that the large, hydrophobic group at position 5′ could stabilize the *β*-anomer and improve the yield (*14*). Thus, before conversion of the pyrimidine moiety, **6a** and **6b** were protected at position 5′ with the 4,4′-dimethoxytrityl group, and subsequent treatment of **7a** and **7b** with guanidine produced the corresponding pseudoisocytidine analogues **8a** and **8b**, respectively. It was found that 2′-*O*-methyl derivative **7b** was converted predominantly to the *β*-anomer of **8b** and only a trace of the α-anomer was detected during the reaction as checked by TLC, while anomerization took place in the case of **7a** to give an approximately 3:11 mixture of α- and *β*-anomers. The mixture of α and *β*-anomers was readily separated into pure anomers by column chromatography. The desirable *β*-anomer **8a**, as evidenced by ^1H NMR, was isolated in 55% yield. Sequential protection of the 2-amino group of **8a** and **8b**, followed by phosphitylation of the 3′-hydroxyl group, afforded phosphoramidites **10a** and **10b**, respectively, which were ready to be introduced into oligodeoxyribonucleotides by a DNA synthesizer.

Triplex Formation of Oligodeoxyribonucleotides without the Pseudoisocytosine Base in the Hoogsteen Strand. The structural stability of oligodeoxyribonucleotides with a sequence 5′-d-$(TC)_3T_a(CT)_3C_b(AG)_3$ (*a* and *b* = 0−4) was examined by UV thermal melting experiments and CD spectra. Both physical measurements were taken in a pH range from 4.5 to 8.0 in buffer A (for acidic conditions) or buffer B (for neutral to basic conditions). The T_m data are collected in Table 2. Representative CD spectra of T2C2 are depicted in Figure 6. The UV melting studies carried out on oligomers T*a*C*b* (*a* and *b* = 0−4) indicated that two transitions could be deduced from the absorbance versus temperature profiles at pH 6.0 while only a single transition was observed at pH 4.5, 7.0, or 8.0 (data not shown here). The biphasic curve is indicative of triplex formation. The Hoogsteen domain is less stable than the Watson−Crick domain at pH 6.0. Thus, the lower T_m (T_{m1}) corresponds to dissociation of the Hoogsteen strand and the higher one (T_{m2}) to dissociation of the Watson−Crick duplex. In contrast to the two T_m values at pH 6.0, a single T_m was observed at other pH values (Table 2). The stability of a C^+·GC base triad arises from N_3 protonation of the cytosine base in the Hoogsteen strand. Consequently, the Hoogsteen domain will

Table 2: Melting Temperatures (°C) of Oligodeoxyribonucleotides with a Common Sequence 5′-d-$(TC)_3T_a(CT)_3C_b(AG)_3$ (a and b = 0−4) as a Function of pH[a]

	pH 4.5	pH 6.0		pH 7.0	pH 8.0
oligomer	T_{m1}/T_{m2}	T_{m1}	T_{m2}	T_{m2}	T_{m2}
T0C0	56	32	55	57	57
T1C1	66	43	67	66	66
T2C2	60	15	64	63	64
T3C3	63	29	66	63	62
T3C4	63	25	64	63	62
T4C3	64	22	64	64	64
T4C4	63	17	65	65	63

[a] The strand concentration of each oligomer is 4.0 μM. Buffers A and B are used for pH 4.5−6.0 and 7.0−8.0, respectively.

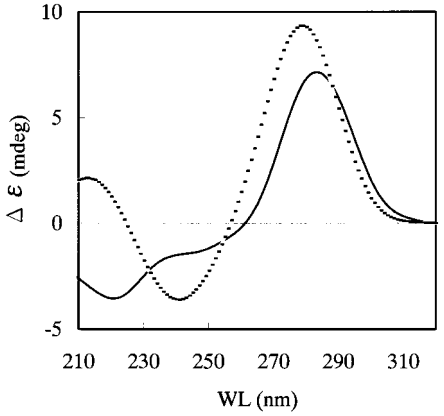

FIGURE 6: CD spectra of T2C2 at 15 °C: pH 4.5 (solid line) and pH 8.0 (dotted line). The strand concentration of T2C2 is 4.0 μM. Buffers A and B are used for pH 4.5 and 8.0, respectively.

be destabilized as the pH is increased. As can be seen in Table 2, the melting temperatures for dissociation of the Hoogsteen strand at pH 6.0 are much lower than those at pH 4.5, which demonstrates that the decreasing stability is closely related to the extent of protonation of cytosine bases in the Hoogsteen domain (*7*). Thus, the triplex exists predominantly at low pH (4.5). The observed single melting temperature reflects the fact that the melting of Watson−Hoogsteen strands occurs at the same temperature as in Watson−Crick strands at pH 4.5. In previous studies, triplexes in different sequences have been melted to single strands directly under similar conditions (*15*). The T_m at high pH (7.0 and 8.0) represents dissociation of the Watson−Crick duplex (*7*). These results can be verified by CD spectra as shown in Figure 6.

In the study presented here, we focus on the sequence 5′-d-$(TC)_3T_a(CT)_3C_b(AG)_3$ (a and b = 0−4), which is strikingly similar to the hairpin type triplex composed of the probe 5′-$(TC)_3T_4(CT)_3$ and the target 5′-$(AG)_3$ (*7*). The CD spectrum of T2C2 at pH 4.5 exhibited a characteristic band at 220 nm (Figure 6) similar to the absorbance of triplexes with similar sequences (*7*). Therefore, the similarity of CD spectra of these two systems strongly implies formation of the triplex here. Furthermore, the band at 220 nm diminished and shifted to 240 nm when the pH increased from 4.5 to 8.0 at 15 °C (Figure 6). This change in the CD pattern indicated dissociation of the Hoogsteen strand from the Watson−Crick duplex. Similar results were also observed in other oligomers (data not shown here). Thus, oligomers with the common sequence 5′-d-$(TC)_3T_a(CT)_3C_b(AG)_3$ (a and b = 0−4) formed triple helices under slightly acidic conditions, but gradually dissociated to double helices as the pH was increased.

Table 3: Melting Temperatures of Oligodeoxyribonucleotides with a Common Sequence 5′-d-$(TC)_3T_a(CT)_3C_b(AG)_3$ (a and b = 0−4) as a Function of the Strand Concentration[a]

concentration	T_m (°C)						
(μM)	T0C0	T1C1	T2C2	T3C3	T3C4	T4C3	T4C4
4.0×10^{-1}	52	64	61	64	66	66	62
4.0×10^{0}	56	66	60	63	63	64	63
4.0×10	55	66	60	60	60	65	62
4.0×10^{2}	−	69	58	59	60	64	60

[a] Buffer A at pH 4.5 is used as the solvent.

Formation of a Paper-Clip Type Triplex. Do the above oligodeoxyribonucleotides form interstrand as well as intrastrand triplexes? This question can be answered by a concentration-dependent study of T_m. Over the range of concentrations studied here [from 0.4 μM (with a 100 mm cell) to 400 μM (with a 0.1 mm cell)], the T_m values for each of the triple-helical oligodeoxyribonucleotide at pH 4.5 fluctuated only within a small range (maximum of ca. 6 °C) which is smaller than the effect caused by the concentration difference in the first-order reaction. These T_m data are summarized in Table 3. These results are consistent with a single-stranded oligodeoxyribonucleotide forming an intrastrand triple-helical structure. In other words, oligomers with the sequence 5′-d-$(TC)_3T_a(CT)_3C_b(AG)_3$ (a and b = 0−4) adopt a paper-clip type triplex by self-folding (Figure 2) (*16*). Although we cannot rule out the possibility of some intermolecular association, we believe that, by analogy to the hairpin type triplex formation in our previous work, intramolecular association is highly favored with our present oligomers.

Effect of Loop Size. Crothers and co-workers reported that the favorable loop size was three to five in small RNA hairpin loops (*17*). Shimizu et al. tested the "loop size" effect in triplex formation in supercoiled plasmids and found that the size of the loop could be four, six, eight, and ten bases without affecting the stability of triplex formation (*18*). Recently, the hairpin type duplex formation with only one base in the loop has also been reported (*19*). The results of our study presented here revealed that the T_m values of T1C1, T2C2, T3C3, T3C4, T4C3, and T4C4 were very similar under the same condition of pH 7.0 or 8.0 (Table 2). Thus, the loop size has no major influence on the stability of a hairpin type duplex. On the other hand, T_m values of T1C1, T3C3, T3C4, T4C3, and T4C4 were also very similar at pH 4.5 but about 4 °C on average higher than that of T2C2. Thus, the loop size did not contribute a significant effect to paper-clip type triplex formation. The slightly lower T_m of T2C2 may reflect a fine-tuning of the structural stability in 5′-d-$(TC)_3T_a(CT)_3C_b(AG)_3$ (a and b = 1−4). Since T2C2 is the most sensitive one among all others to triplex formation, we used it as a model for studying triplexes containing modified bases.

On the other hand, the stability of the triple-helical structure at pH 6.0 was found to be a function of a and b (a and b = 1−4). T1C1 showed a slightly higher T_{m1} among all oligomers at this pH. When a and b were greater than 2,

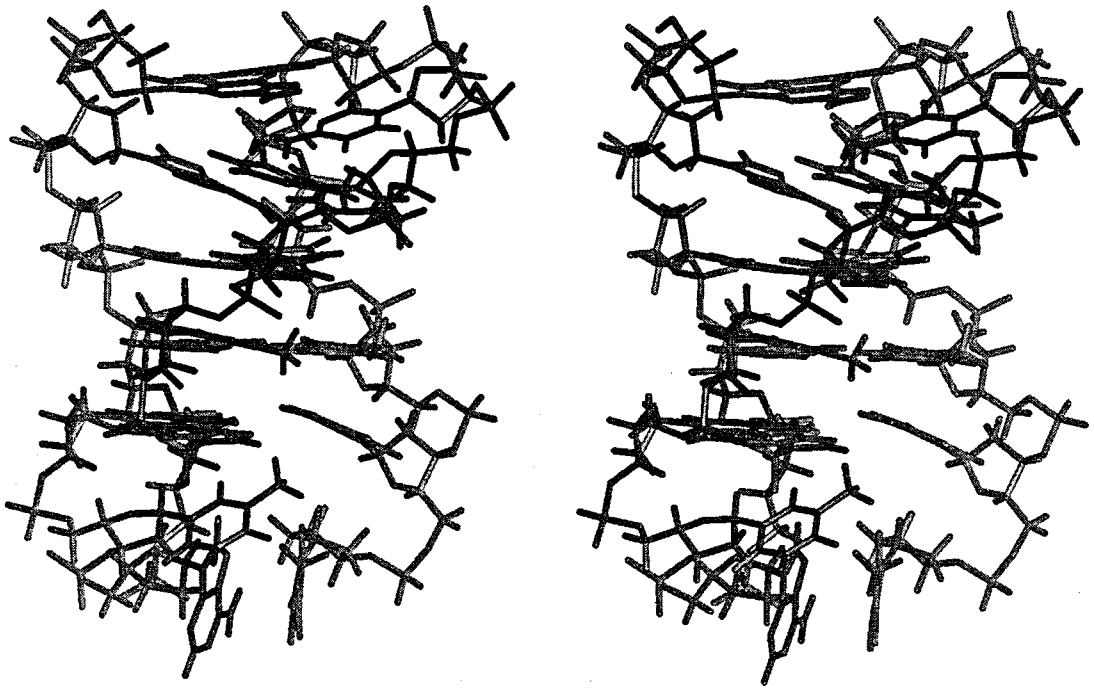

FIGURE 7: Stereoview of the T0C0 triplex structure. 5′-T1 is located at the bottom right corner of this view with the G18-3′ at the upper left corner (see the text for the detailed description of the model).

the stability of the Hoogsteen domain became sensitive to loop length. When the loop length of either d-C_b (connecting Watson and Crick strands) or d-T_a (connecting Crick and Hoogsteen strands) was increased by one nucleotide, we found that the T_{m1} values were 4 or 7 °C lower than that of T3C3 for T3C4 and T4C3, respectively (Table 2). Obviously, this means that association of the Hoogsteen base pairs is more sensitive to loop length and/or loop conformation than that of the Watson−Crick base pairs, reflecting the stability of the Watson−Crick domain over that of the Hoogsteen domain. As expected, the melting temperature of the Hoogsteen strand of T4C4 appeared to be even lower than that of T3C4 or T4C3.

Structure of the Triplex Formed by T0C0. For the sake of clarity and simplicity in the description of the structure of T0C0, we print and number the bases from 1 to 18: 5′-T1C2T3C4T5C6C7T8C9T10C11T12A13G14A15G16A17G18. The correspondence of the numbering and base position in triplex can be viewed in Figure 2. That is to say that the T0C0 would have formed six base triads (Figure 2). Four of them are internal base triads ($C2^+$·G14C11, T3·A15T10, $C4^+$·G16C9, and T5·A17T8) and two ends (T1·A13T12 and $C6^+$·G18C7). The preliminary results of NMR studies indicate that three of four internal base triads (T3·A15T10, $C4^+$·G16C9, and T5·A17T8) form a typical pyrimidine·purine·pyrimidine hydrogen-bonding configuration (*15*). In the $C2^+$·G14C11 triad, G14 and C11 form a typical Watson−Crick base pairing while the $C2^+$ residue is slightly out of plane. Presumably, this is due to the fraying of the 5′ end. The triad bases at both ends function as turns in the DNA and, therefore, do not form typical pyrimidine·purine·pyrimidine base triads. Qualitatively, the turn composed of T12 and A13 forms in such a way as to displace the A13 base from the neighboring base triad $C2^+$·G14C11. The T12 base stacks on its 5′ end neighbor C11. This overall loop structure and base orientation are very similar to the hairpin turn of GTTA reported by Hilbers and co-workers (*20*). The phosphate−sugar backbone between T12 and A13 appears to be unusually extended. There are NOEs observed between C_6H of T1 and $C_{2'}H_a$, $C_{2'}H_b$, and CH_3 of T12. This indicates that T1 is in the vicinity of the T12−A13 loop. Thus, while T1, T12, and A13 do not form a typical T·AT base triad with hydrogen bonds, they do cluster together and obviously have a stable conformation. While the bases $C6^+$, C7, and G18 do not form a typical Hoogsteen triad, they clearly form a stable triad. The presence of N_3H^+ of $C6^+$ indicates the presence of hydrogen bonding involving the N_3H^+ proton. Initial calculations indicate that the hydrogen bond formed is between N_3H^+ of $C6^+$ and C_6O of G18. The C7 base appears to stack onto the T8 base. A stereoview of the resulting preliminary model is shown in Figure 7.

Thermodynamic Analysis of Triplex T0C0. The T_m values of T0C0 at pH 4.5 with various concentrations (Table 3) were ∼10 °C lower than those of other oligomers under the same conditions. The UV thermal melting profiles of T0C0 and T1C1 at pH 4.5 are depicted in Figure 8. The absorbance versus temperature curve of T0C0 was very similar to that of T1C1, except for the lower T_m value and the broader breadth. These results indicated that the dissociation behavior of T0C0 was similar to that of T1C1.

The changes in enthalpy and entropy for the intramolecular triplex formation of T0C0 and T1C1 at pH 4.5 were obtained by shape analysis of their thermal profiles using the method of Marky and Breslauer (*21*). The ΔH and ΔS values were calculated from the Arrhenius correlation between the association constant ln K and $-\Delta H/RT + \Delta S/R$. The free energy values (ΔG) were deduced using the thermodynamic relationship $\Delta G = \Delta H - T\Delta S$. These thermodynamic parameters are listed in Table 4. The ΔH value of T0C0 is 76 kJ/mol less than that of T1C1. This destabilization may result from the distorted conformation of the triplex T0C0,

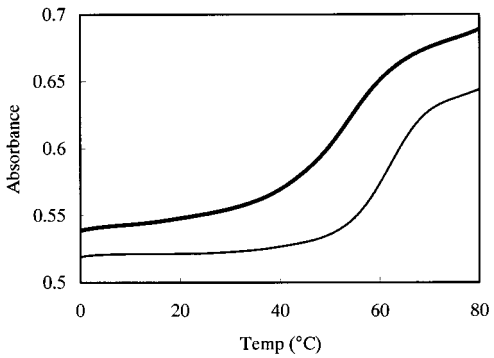

FIGURE 8: UV thermal melting curves of T0C0 (thick line) and T1C1 (thin line) in buffer A at pH 4.5. The strand concentration of T0C0 or T1C1 is 4.0 μM.

Table 4: Thermodynamic Parameters of Triplex Formation of Oligomers T0C0 and T1C1 in Buffer A at pH 4.5

oligomer	T_m (°C)	ΔH (kJ/mol)	ΔS (J K^{-1} mol^{-1})	ΔG (298 K) (kJ/mol)
T0C0	56	−119.788	−0.368	−10.124
T1C1	66	−195.569	−0.586	−20.941

Table 5: T_m Values (°C) of D1−D3 and M1−M3 in Buffer C (pH 4.5) or Buffer D (pH 7.0) with and without 20 mM $MgCl_2$[a]

			D1	M1	D2	M2	D3	M3
	pH 4.5	T_{m1}	60	60	53	30	29	25
		T_{m2}				65	63	67
	pH 7.0	T_{m1}					5	18
		T_{m2}	65	65	66	63	67	67
$MgCl_2$	pH 4.5	T_{m1}	−	−	−	−	32	26
		T_{m2}	−	−	−	−	63	67
	pH 7.0	T_{m1}	−	−	−	−	10	21
		T_{m2}	−	−	−	−	65	67

[a] The strand concentrations of all the oligomers are 4.0 μM, except that for M2 (3.3 μM).

as evidenced by ^{1}H NMR. Though T0C0 exhibited less enthalpy gained upon triplex formation, ΔH was offset by a decrease in the entropy penalty on going from the single strand to the more structured helical state (Table 4).

Inspection of ΔG values in comparison with the corresponding T_m values revealed the expected correlation; larger T_m values corresponded to more negative ΔG values. It is interesting to note that the calculated free energies of triplex formation near room temperature, ΔG(298 K), are −10 and −21 kJ/mol for T0C0 and T1C1, respectively, more favorable for the triple-helical structure than for the single-stranded coil. Therefore, one base in the loop will thermodynamically stabilize the triplex.

Effect of 2′-Deoxypseudoisocytidine (D) and 2′-O-Methylpseudoisocytidine (M) on Triplex Formation. We have also studied the formation of triplexes containing up to three D's or M's in a basic sequence of T2C2. Oligomers M1 (or D1) to M3 (or D3) are formed by replacing one to three 2′-deoxycytidines in the Hoogsteen strand of T2C2 with the pseudoisocytidine derivative M (or D). The exact sequences are shown in Table 1. The UV thermal melting studies of these oligomers were performed in buffer C (pH 4.5) or buffer D (pH 7.0) with and without 20 mM $MgCl_2$. The results are summarized in Table 5. Two T_m values of M3 or D3 were observed at pH 7.0. These observations can be interpreted in terms of triplex and duplex dissociation with the aid of CD spectra at low (5 °C) and high (40 °C) temperatures (data not shown here). These results indicated that formation of the $^{\Psi}$C·GC base triad stabilized the triple-helical structure of oligodeoxyribonucleotides at neutral pH.

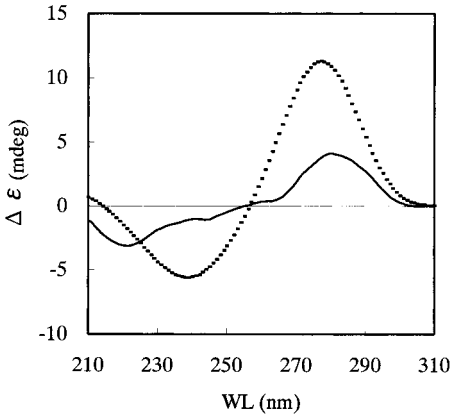

FIGURE 9: CD spectra of D1 at 5 °C: pH 4.5 (solid line) and pH 7.0 (dotted line). The strand concentration of D1 is 4.0 μM. Buffers C and D are used for pH 4.5 and 7.0, respectively.

However, there is only one T_m (∼65 °C) observed in D1, D2, M1, and M2 at pH 7.0. This T_m is very close to the T_{m2} (∼67 °C) of D3 or M3 under the same condition. The CD patterns of D1, D2, M1, and M2 at pH 7.0 exhibited a characteristic band at 240 nm (Figure 9), similar to the absorbance of their parent oligomer T2C2 at pH 8.0 (Figure 6). Thus, this T_m corresponds to duplex dissociation of D1, D2, M1, and M2. Obviously, the triplex cannot be formed at pH 7.0 without all three $^{\Psi}$C·GC base triads present in these oligomers. The T_m (5 °C) of D3 is lower than that of M3 (18 °C) under the same condition. This may be due to the fact that the conformation of the Hoogsteen domain in M3 is a mixture of A and B forms (*22*) as the 2′-*O*-methyl derivative prefers A-DNA (*23*).

Comparison of the Stability of $^{\Psi}$C·GC versus C^+·GC. Although only a single transition in the UV melting curve was detected for D1, D2, or M1 at pH 4.5, the analysis of their CD spectra confirmed formation of the triple-helical structure (Figure 9). The shoulder at 240 nm in the CD spectrum of D1 at pH 7.0 disappeared and shifted to 220 nm as the pH was decreased to 4.5. Thus, oligomers D1, D2, and M1 formed stable triplexes and dissociated directly to single strands, rather than going through an intermediate partial duplex state at acidic pH. This observation is not surprising, given the influence of the strong C^+·GC base triad at low pH. In addition, the thermal stability of triplexes at pH 4.5 reverses the relationship with the number of D or M derivative. This implies that $^{\Psi}$C·GC is less stable than C^+·GC (*7*). However, there is only a small population of C^+·GC existing at pH 7.0. This is why pseudoisocytidine derivatives are useful for triplex formation.

$^{\Psi}$C·GC Stablized by Protonation of $^{\Psi}$C. It is interesting to note that the T_{m1} values of D3 and M3 are raised dramatically (from 5 to 24 °C) when the pH is decreased from 7.0 to 4.5 (Table 5). The stability of triplexes containing D or M is pH-dependent. This is most likely primarily due to protonation of the pseudoisocytosine base. Pseudoisocytidine has two tautomers (i.e., the imino proton can be located at either N_1 or N_3) (*24*). The N_3H tautomer favors formation of the $^{\Psi}$C·GC base triad (Figure 1). Thus, triplex formation

at neutral pH is a competition between N_1H and N_3H tautomers. The triplex stability of D3 or M3 is enhanced under acidic conditions, because the protonated pseudoisocytosine base possesses protons at both N_1 and N_3 sites all the time (*24*).

Effect of $MgCl_2$. The facilitation of triplex formation in oligodeoxyribonucleotides by polyvalent cations is well-known. Therefore, solutions containing additional 20 mM $MgCl_2$ were also studied in the cases of D3 and M3. The results summarized in Table 5 show that the T_{m1} values of both D3 and M3 with or without Mg^{2+} are very close. This may indicate that Mg^{2+} has an insignificant effect on the structural stabilization of oligodeoxyribonucleotides containing D or M.

CONCLUSION

We have undertaken an extensive study of the model system 5′-d-$(TC)_3T_a(CT)_3C_b(AG)_3$ (a and $b = 0-4$) by UV thermal melting experiments and CD spectra. In addition, the preliminary results of ^{1}H NMR studies of 5′-d-$(TC)_3(CT)_3(AG)_3$ (a and $b = 0$) are also presented. We conclude that (i) DNA oligomers containing only three T·AT and C^+·GC base triads are an effective model for triplex formation, especially as base analogues are incorporated in the sequence; (ii) the length of a loop (a and $b = 1-4$) has no major effect on the triplex stability in this system if the Hoogsteen domain is associated tightly; (iii) the easily accessible C-nucleosides, 2′-deoxypseudoisocytidine and 2′-*O*-methylpseudoisocytidine, can be used to replace 2′-deoxycytidine for triplex formation under physiological conditions; (iv) the ability of 2′-*O*-methylpseudoisocytidine to form triplexes is superior than that of 2′-deoxypseudoisocytidine at neutral pH; and (v) a distorted triplex can also be formed in oligomer 5′-d-$(TC)_3(CT)_3(AG)_3$ with no loop (a and $b = 0$). This novel triple-helical structure may have biological significance. A study of the monoclonal antibody induction with this structure is under way.

SUPPORTING INFORMATION AVAILABLE

UV thermal melting curves and the first-differential curves (dA/dT vs T) of D3 in buffer D at pH 7.0 with 20 mM $MgCl_2$ and of T4C4 in buffer A at pH 6.0 as well as the step-by-step synthetic procedure and characterization of 2′-*O*-methylpseudoisocytidine and 2′-deoxypseudoisocytidine. This material is available free of charge via the Internet at http://pubs.acs.org.

REFERENCES

1. Panyutin, I. G., and Neumann, R. D. (1997) *Nucleic Acids Res. 25*, 883−887. (b) Maher, L. J., III (1992) *BioEssays 14*, 807−815. (c) Duval-Valentin, G., Thuong, N. T., and Helene, C. (1992) *Proc. Natl. Acad. Sci. U.S.A. 89*, 504−508.
2. Francois, J. C., Saison-Behmoaras, T., Thuong, N. T., and Helene, C. (1989) *Biochemistry 28*, 9617−9619. (b) Maher, L. J., III, Wold, B., and Dervan, P. B. (1989) *Science 245*, 725−730.
3. Strobel, S. A., Doucette-Stamm, L. A., Riba, L., Housman, D. E., and Dervan, P. B. (1991) *Science 254*, 1639−1642. (b) Ts'o, P. O. P., Aurelian, L., Cheng, E., and Miller, P. S. (1992) *Ann. N.Y. Acad. Sci. 660*, 159−177. (c) Helene, C. (1993) *Curr. Opin. Biotechnol. 4*, 29−36.
4. Singleton, S. F., and Dervan, P. B. (1992) *J. Am. Chem. Soc. 114*, 6957−6965.
5. Singleton, S. F., and Dervan, P. B. (1992) *Biochemistry 31*, 10995−11003.
6. Ono, A., Ts'o, P. O. P., and Kan, L.-S. (1991) *J. Am. Chem. Soc. 113*, 4032−4033. (b) Ono, A., Ts'o, P. O. P., and Kan, L.-S. (1992) *J. Org. Chem. 57*, 3225−3230. (c) Ono, A., Ts'o, P. O. P., and Kan, L.-S. (1997) *J. Chin. Chem. Soc. 44*, 601−607.
7. Tsay, L.-M., Lin, S.-B., Tsay, H.-T., Chen, H.-H., and Kan, L.-S. (1995) *J. Biomol. Struct. Dyn. 12*, 1235−1245. (b) Tsay, H.-T., Lin, S.-B., Tsay, L.-M., and Kan, L.-S. (1995) *Chemistry (Taipei) 54*, 24−32. (c) Lin, S.-B., Kao, C.-F., Lee, S.-C., and Kan, L.-S. (1994) *Anti-Cancer Drug Des. 9*, 1−8.
8. Wijmenga, S. S., Mooren, M. M. W., and Hilbers, C. W. (1993) in *NMR of Macromolecules: A Practical Approach* (Roberts, G. C. K., Ed.) pp 217−288, Oxford University Press, New York.
9. Atkinson, T., and Smith, M. (1984) in *Oligonucleotide Synthesis: A Practical Approach* (Gait, M. J., Ed.) p 35, IRL Press, Oxford. (b) Yuhasz, S. C., Kanhouwa, N., Shida, T., and Kan, L.-S. (1987) *Biotechniques 5*, 54−59.
10. Fasman, G. D., Ed. (1975) *CRC Handbook of Biochemistry and Molecular Biology*, 3rd ed., Vol. I, p 589, CRC Press, Cleveland, OH.
11. Hirota, K., Watanabe, K. A., and Fox, J. J. (1978) *J. Org. Chem. 43*, 1193−1197.
12. Chambers, R. W., Kurkov, V., and Shapiro, R. (1963) *Biochemistry 2*, 1192−1203.
13. Pankiewicz, K., Matsuda, A., and Watanabe, K. A. (1982) *J. Org. Chem. 47*, 485−488.
14. Matsuda, A., Chu, C. K., Pankiewicz, K., Watanabe, K. A., and Fox, J. J. (1981) *J. Org. Chem. 46*, 3603−3609.
15. Kan, L.-S., Callahan, D. E., Trapane, T. L., Miller, P. S., Ts'o, P. O. P., and Huang, D. H. (1991) *J. Biomol. Struct. Dyn. 8*, 911−933. (b) Kan, L.-S., and Ono, A. (1994) *J. Chin. Chem. Soc. 41*, 865−869.
16. Macaya, R., Wang, E., Schultz, P., Sklenar, V., and Feigon, J. (1992) *J. Mol. Biol. 225*, 755−773.
17. Gralla, J., and Crothers, D. M. (1973) *J. Mol. Biol. 73*, 497−511. (b) Ravetch, J., Gralla, J., and Crothers, D. M. (1974) *Nucleic Acids Res. 1*, 109−127.
18. Shimizu, M., Hanvey, J. C., and Wells, R. D. (1989) *J. Biol. Chem. 264*, 5944−5949.
19. Chou, S.-H., Tseng, Y.-Y., and Wang, S.-W. (1999) *J. Mol. Biol. 287*, 301−313. (b) Chou, S.-H., Tseng, Y.-Y., and Chu, B.-Y. (1999) *J. Mol. Biol. 292*, 309−320.
20. van Dongen, M. J. P., Mooren, M. M. W., Willems, E. F. A., van der Marel, G. A., van Boom, J. H., Wijmenga, S. S., and Hilbers, C. W. (1997) *Nucleic Acids Res. 25*, 1537−1547.
21. Marky, L. A., and Breslauer, K. J. (1987) *Biopolymers 26*, 1601−1620.
22. Liu, M., Yang, L., Deng, W., Su, M., Wang, C., Lin, S.-B., Kan, L.-S., and Bai, C. (1998) *J. Biomol. Struct. Dyn. 15*, 895−903.
23. Shimizu, M., Konishi, A., Shimada, Y., Inoue, H., and Ohtsuka, E. (1992) *FEBS Lett. 302*, 155−158. (b) Uesugi, S., Kaneyasu, T., Imura, J., Ikehara, M., Cheng, D. M., Kan, L.-S., and Ts'o, P. O. P. (1983) *Biopolymers 22*, 1189−1202.
24. Kan, L.-S., Lin, W.-C., Yadav, R. D., Shih, J. H., and Chao, I. (1999) *Nucleosides Nucleotides 18*, 1091−1093.

BI0004201