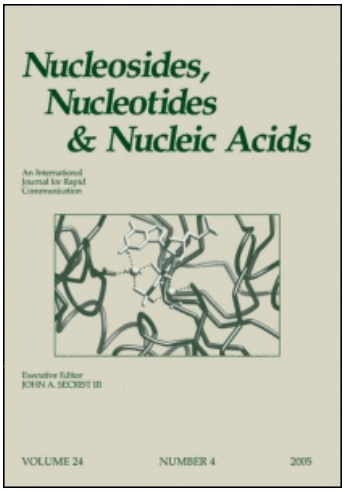


Nucleosides, Nucleotides and Nucleic Acids
Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t713597286

The Design and Synthesis of N^4-Anthraniloyl-2'-dC, the Improved Syntheses of N^4-Carbamoyl-and $N⁴$- Ureidocarbamoyl-2'-dC, Incorporation into Oligonucleotides and Triplex Formation Testing

Nancy Guzzo-Pernell[a]; Geoff W. Tregear[a]; Jim Haralambidis[a]; John M. Lawlor[b]
[a] Howard Florey Institute of Experimental Physiology and Medicine, Parkville, Victoria, Australia [b] School of Chemistry, University of Melbourne, Parkville, Victoria, Australia

To cite this Article Guzzo-Pernell, Nancy , Tregear, Geoff W. , Haralambidis, Jim and Lawlor, John M.(1998) 'The Design and Synthesis of N^4-Anthraniloyl-2'-dC, the Improved Syntheses of N^4-Carbamoyl-and $N⁴$-Ureidocarbamoyl-2'-dC, Incorporation into Oligonucleotides and Triplex Formation Testing', Nucleosides, Nucleotides and Nucleic Acids, 17: 7, 1191 — 1207
To link to this Article: DOI: 10.1080/07328319808004232
URL: http://dx.doi.org/10.1080/07328319808004232

NUCLEOSIDES & NUCLEOTIDES, 17(7), 1191-1207 (1998)

THE DESIGN AND SYNTHESIS OF N^4-ANTHRANILOYL-2'-dC, THE IMPROVED SYNTHESES OF N^4-CARBAMOYL- AND N^4-UREIDOCARBAMOYL-2'-dC, INCORPORATION INTO OLIGONUCLEOTIDES AND TRIPLEX FORMATION TESTING

Nancy Guzzo-Pernell,[a]* Geoff W Tregear,[a] Jim Haralambidis,[a] and John M Lawlor.[b]

[a]Howard Florey Institute of Experimental Physiology and Medicine, Parkville, Victoria, 3052, Australia. [b]School of Chemistry, University of Melbourne, Parkville, Victoria, 3052, Australia.

Abstract

Three modified nucleosides were designed with the aim of achieving triplet formation with the **CG** base pair of duplex DNA. Direct anthraniloylation of 2'-deoxycytidine, using isatoic anhydride, afforded the novel N^4-anthraniloyl-2'-deoxycytidine. Much improved preparations of N^4-carbamoyl-2'-deoxycytidine and of N^4-ureidocarbonyl-2'-deoxycytidine were accomplished. The modified nucleosides were incorporated into oligonucleotides. Thermal denaturation studies and gel mobility shift analysis suggest that these nucleosides do not form base triplets with any of the four base pairs of DNA.

Introduction

The sequence-specific binding of an oligonucleotide to double-stranded DNA offers a novel and exciting technique for controlling the expression of specific genes.[1-3] This technology is known as the antigene strategy, and its application to rational drug discovery may lead to a new class of pharmacological agents. In addition, the technology may also be useful in diagnostic procedures based on the use of specific DNA probes.[4]

The most stable triple helical DNA involves Hoogsteen hydrogen bonding of pyrimidine oligonucleotides to parallel purine tracts in the major groove of the Watson-Crick duplex. Target sequences are limited to thymine (T) recognition of adenine-thymine (**AT**) to give the **T:AT** triplet, and protonated cytosine (C^+) recognition of guanine-cytosine to give the C^+**:GC** triplet as shown in Figure 1.[2,3,5-7] The preference for triplexes to be formed by the **T:AT** and C^+**:GC** triplets is attributed to the ability of these "canonical" triplets to accommodate the third-strand backbone without significant torsional strain.[7] Triplexes containing other triplets (**G:TA**, **T:CG**, **A:AT**, and **G:GC**) have been demonstrated but these triplexes are weaker and less selective than triplexes composed of

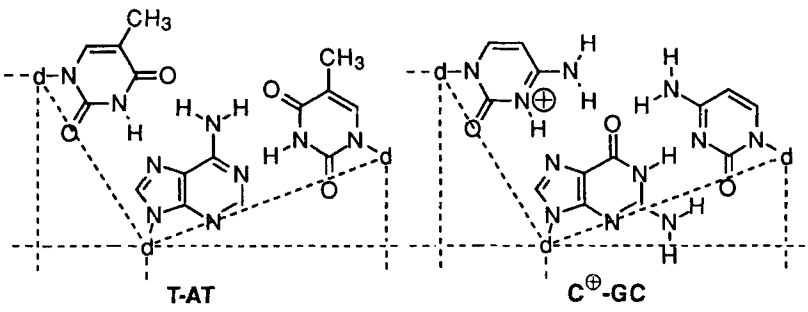


Figure 1

the canonical triplets.[7,8] Even triplexes comprised of **T:AT** and C^+**:GC** triplets are rather weak and slow in forming and so are of little practical value (quite apart from the very limited range of targets possible); weaker, and less selective, triplets are unlikely to be of any use.

Within the canonical triplets, C^+**:GC** can be used only in mildly acidic solution because the triplet is deprotonated at physiological pH. This problem has been addressed by replacement of the protonated cytosine with unnatural bases[9] that can form two hydrogen bonds with **GC** while the triplet is essentially isomorphic with the canonical triplets.

Of the two base pairs that cannot as yet be recognised with sufficient selectivity and strength, the **CG** pair seems the easiest target because nucleosides designed to target this pair appeared to be accessible by simple modification of 2'-deoxycytidine. [The 5-methyl group of thymine appears to preclude the use of the natural nucleosides as precursors of ligands for **TA**]. The major criteria for the modifications of 2'-deoxycytidine were the provision of appropriate hydrogen bonding sites and simplicity of synthesis. There was also the need for the Hoogsteen hydrogen bonding to be accomplished without too much distortion of the third-strand backbone, though the stringency of this requirement had not been established when this project was started. There is also the possibility that stacking between neighbouring third-strand bases may be required, though such stacking must be much weaker than that between adjacent base pairs of DNA, because the overlap between the third-strand base is much smaller.

Two derivatives of cytosine, N^4-ureidocarbonyl-2'-dC (**4**) and N^4-carbamoyl-2'-dC (**6**), were selected because of their potential for forming at least two hydrogen bonds with the target **CG** pair, and because facile preparation had been reported. Two possible modes of the hydrogen bonding of **6** to **CG** without gross deviation from the canonical structure are shown in Figure 2. For **4**, the same interactions are available, and there are other potential bonding modes using the terminal carbamoyl group; one such possibility is shown in Figure 2.

It is, of course, possible that intramolecular hydrogen bonding of the carbamoyl moieties of the 4-substituents of **4** and **6** to N^3 would favour conformations incapable of triplet formation in the intended manner. There are other negative considerations. The 4-substituents are very hydrophilic and the cost of de-solvating them on entering the major groove[10] might be too high. And there is no provision for intra third-strand stacking.[11]

Both of the modified deoxycytidines **4** and **6** were reported[12] to have been prepared via a triazinetrione derivative **3** of 3',5'-di-*O*-acetyldeoxycytidine (**2**) made by treating **2** with chlorocarbonyl isocyanate. We found, however, this to be an extremely inefficient and tedious method, so we devised new syntheses which were much simpler than the reported methods and gave higher yields.

When it was becoming clear that **4** and **6** were not capable of forming triplets, we thought that we should test a substituent on the 4-amino group which is less hydrophilic and at least offers the possibility of intra third-strand stacking. The anthraniloyl group is such a substituent, and N^4-anthraniloyl-2'-deoxycytidine (**9**, Figure 2) was readily accessible simply by warming 2'-deoxycytidine with isatoic anhydride in DMF.

Particularly well-characterised examples of thermal denaturation profiles and gel mobility shift patterns were available in a report from Bischofberger et al,[13] and we chose to use their oligomers in the control experiments and as test-beds for our nucleosides. Because triplex stability is relatively insensitive to base mismatches near the ends of oligonucleosides,[14] the substitutions of natural nucleosides by our unnatural ones were made in the middle of the oligomers.

Results

Kumar and Leonard described two ways of protecting the hydroxyl groups of the 2'-deoxycytidine: as acetyl esters (**2**), and as the 1,1,3,3-tetraisopropyldisiloxanyl ethers. Since both of these starting materials were reported to give the same product, we selected **2** for our preparation of **4** and **6**.

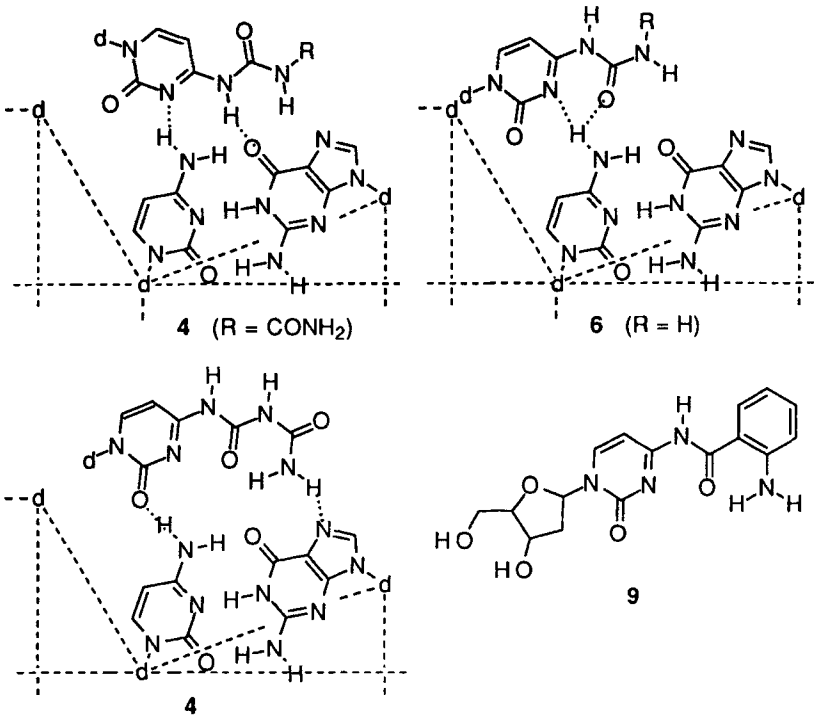


Figure 2

We used two reported methods to prepare **2**. Partial hydrazinolysis of triacetyldeoxycytidine (**1**)[15] gave mixtures of **1**, **2**, and di-deacetylated materials separable by chromatography; recycling the non-target materials gave a good yield (80%) eventually. Zinc-ion catalysed methanolysis of **1**[16] gave a mixture of **1** and **2** only, but much of the nucleoside material could not be extracted from the reaction mixture so, although this was the less tedious of the two methods, the yields (40%) were modest.

The annelation of *O*-protected cytidine with chlorocarbonyl isocyanate to give the triazinetrione **3** was carried out as reported by Kumar and Leonard.[12] The triazinetrione ring was very easily cleaved by alcohols and moisture, and considerable care was needed in obtaining the spectra required for the characterisation of **3**.

According to the report, ammonolysis of **3** afforded **4**, and methanolysis gave the corresponding methyl ester **5**, and ammonolysis of the latter provided **6**. In our hands, ammonolysis of the trione **3** and of the ester **5** both gave reaction products apparently

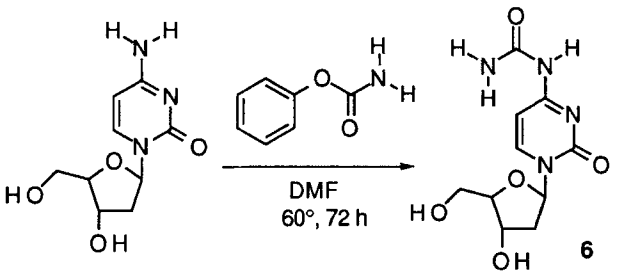
homogeneous by TLC. However, ^{1}H NMR spectra showed each of the reaction products to be complex mixtures. Consistent with the TLC analyses, preparative chromatography on normal silica gel was unsuccessful. We were able to separate the mixture by RP-HPLC and it became evident that instead of usefully regioselective syntheses of **4** and **6** we had obtained merely somewhat different mixtures of **4**, **6**, and 2'-dC.

While RP-HPLC did enable us to accumulate sufficient amounts of the compounds for characterisation, the method of Kumar and Leonard proved much too inefficient in our hands to provide enough material for the oligonucleotide experiments, and new, regio-specific, methods were adopted.

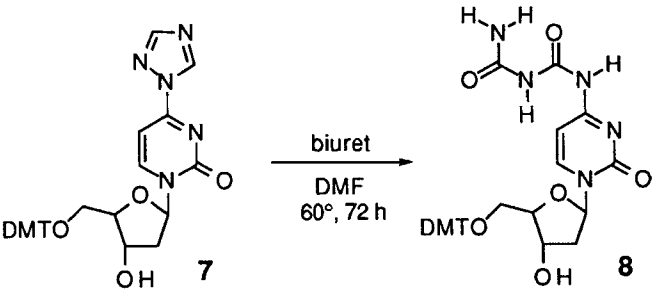
The direct N^4-carbamoylation of 2'-deoxycytidine in aqueous DMF with potassium cyanate and acetic acid failed, but N^4-carbamoyl-2'-dC (**6**) was able to be prepared by warming the unprotected 2'-deoxycytidine in dry DMF with a small excess of phenyl carbamate for several days (Scheme 1). Perhaps the near absence of water allows the cyanic acid eliminated from the phenyl carbamate the opportunity to react with the weakly nucleophilic cytidine 4-amino group. The only nucleosidic materials detected in the reaction mixture were 2'-dC and **6**, and the pure product could be isolated easily by flash chromatography on silica.

The improved synthesis of N^4-ureidocarbonyl-2'-dC (**4**) was found through nucleophilic displacement of the triazolyl group from 4-(1,2,4-triazol-1-yl)-1-[5'-*O*-di(*p*-anisyl)phenylmethyl-2'-deoxy-β-D-ribofuranosyl]-2-pyrimidone (**7**) by biuret (Scheme 2). The triazolyl nucleoside was prepared from uracil;[17] and the tritylation was done before the displacement reaction in order to facilitate the isolation of the otherwise hydrophilic product from the reaction mixture. The tritylated triazolyl intermediate was reacted with biuret in DMF at 50°C for 72 h. Only the target compound and starting material were detected in the product mixture, and the tritylated ureidocarbonyl-2'-dC (**8**) was readily isolated by flash chromatography in 64% yield. Use of pyridine instead of DMF, gave poorer yields.

The oligonucleotides reported (Figure 3) by Bischofberger and co-workers were prepared to provide the datum for oligonucleotides containing our unnatural bases. We obtained T_m values close to those in the literature for oligomers **O1a:O1b:O1c**, **O1a:O1b:O2c**, and **O1a:O1b:O3c**, and confirmed the finding that if a base in the middle of a third-strand did not match the target base-pair no triplex was observed.



Scheme 1



Scheme 2

Triplex formation testing of oligonucleotides containing N^4-carbamoyl-, N^4-ureidocarbonyl-, or N^4-anthraniloyl-2'-dC was accomplished by thermal denaturation studies (Table 1) and gel mobility shift analysis.

Our thermodynamic data showed only one transition, corresponding to the dissociation of the modified Watson and Crick duplex. The modified oligonucleotides were investigated further by repeating the thermal denaturation studies at higher salt concentrations.[18] The NaCl concentration in the buffer used for the thermal denaturation studies was increased from 100 mM to 500 mM and 1M, but in each case only the duplex transition was apparent. The increase in the T_m value with increasing NaCl concentration observed is consistent with an increased stability for the Watson and Crick duplex. The lack of a second transition at lower temperatures even in the presence of high NaCl concentrations suggests that the modified nucleosides **4**, and **6** do not form a triplet with the CG base pair. Further, our gel mobility shifts analysis showed that the duplex and triplex mixtures migrate at the same time, thus supporting our thermodynamic data.

Triplex-1 parallel (control)

O1c 5'-TTTCTTT**C**TTTTCTT-3'
O1b 5'-TGAGTGAGTAAAGAAA**G**AAAAGAATGAGTGCCAA-3'
O1a 3'-ACTCACTCATTTCTTT**C**TTTTCTTACTCACGGTT-5'

Antiparallel triplex (control)

O2c 3'-TTTCTTT**C**TTTTCTT-5'
O1b 5'-TGAGTGAGTAAAGAAA**G**AAAAGAATGAGTGCCAA-3'
O1a 3'-ACTCACTCATTTCTTT**C**TTTTCTTACTCACGGTT-5'

Parallel natural triplex with mismatch (control)

O3c 5'-TTTCTTT**T**TTTTCTT-3'
O1b 5'-TGAGTGAGTAAAGAAA**G**AAAAGAATGAGTGCCAA-3'
O1a 3'-ACTCACTCATTTCTTT**C**TTTTCTTACTCACGGTT-5'

Parallel triplex with modified base

O4c 5'-TTTCTTT**6**TTTTCTT-3'
O4b 5'-TGAGTGAGTAAAGAAA**C**AAAAGAATGAGTGCCAA-3'
O4a 3'-ACTCACTCATTTCTTT**G**TTTTCTTACTCACGGTT-5'

Parallel triplex with modifed base

O5c 5'-TTTCTTT**4**TTTTCTT-3'
O4b 5'-TGAGTGAGTAAAGAAA**C**AAAAGAATGAGTGCCAA-3'
O4a 3'-ACTCACTCATTTCTTT**G**TTTTCTTACTCACGGTT-5'

Parallel triplex modified base

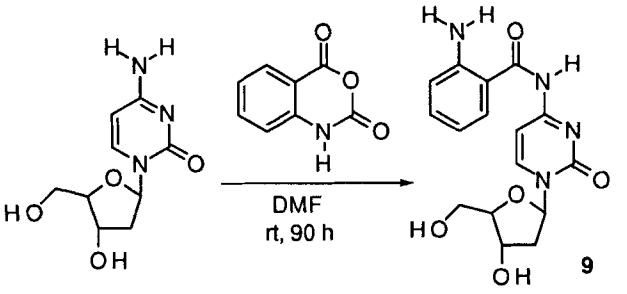
O6c 5'-TTTCTTT**9**TTTTCTT-3'
O4b 5'-TGAGTGAGTAAAGAAA**C**AAAAGAATGAGTGCCAA-3'
O4a 3'-ACTCACTCATTTCTTT**G**TTTTCTTACTCACGGTT-5'

Figure 3

We then synthesized the novel nucleoside N^4-anthraniloyl-2'-dC (**9**). This compound was designed on the premise that the nucleosides **4** and **6** did not form base triplets with the target **CG** base pair because they may be so hydrophilic that they prefer the solvent water to the major groove. Nucleoside **9** was synthesized via anthraniloylation of 2'-deoxycytidine with isatoic anhydride;[19] it is not necessary to protect the 3'- and 5'-hydroxyl groups (Scheme 3). The N^4-anthraniloyl-2'-dC was then converted into its 5'-

Table 1

Complex mixture		T_m (°C) 100 mM NaCl		T_m (°C) 500 mM NaCl		T_m (°C) 1 M NaCl	
O1a:O1b:O1c	**(C^+:GC)**	37.2	63.9	44.0	74.0	55.2	76.8
O1a:O1b			63.5		72.4		77.0
O1a:O1b:O2c	**(C^+:GC)**		64.3				
O1a:O1b:O3c	**(C^+:GC)**		64.6				
O4a:O4b:O4c	**(6:CG)**	<5	64.9	<5	71.9	<5	77.0
O4a:O4b			64.4		72.8		76.6
O4a:O4b:O5c	**(4:CG)**	<5	65.5	<5	73.0	<5	77.1
O4a:O4b:O6c	**(9:GC)**		64.4		72.8		76.6



Scheme 3

dimethoxytrityl-3'-phosphoramidite and incorporated into oligonucleotides which were then tested for triplex formation by thermal denaturation studies (using NaCl concentrations from 100 mM to 1M) and by gel mobility shift analysis. Our results show that N^4-anthraniloyl-2'-dC does not form a triplet with any of the four base pairs of DNA.

Discussion

There have been numerous attempts to devise new bases to recognise the CG base-pair during the past few years. The uniform experience is that those bases that have some ability to bind to any base-pair do so too weakly and with insufficient selectivity to be of practical use. Perhaps the most successful is Miller's modified cytosine, N^4-(6-aminopyridin-2-yl)-2'-dC,[20] which binds to CG with a strength comparable to the

C^+:GC canonical triplets, but it also recognises the other base pairs, though rather more weakly.

One feature that does emerge from the accumulated experience is the necessity for new bases to conform with the canonical structure - all the most successful of the new bases do so. The failure of the carbamoyl and ureidocarbamoyl derivatives of cytosine to show any triplet-forming ability is consistent with this observation in that strong hydrogen bonding of **4** and **6** to the target base-pair requires a substantial departure from the canonical structure.

It is also possible that the highly hydrophilic N^4-substitutents of **4** and **6** may repel the hydrophobic region presented by the next base of the third strand. The base-base stacking interaction of polynucleotides has been studied, using simplified models of di- and trinucleoside phosphates, by UV absorption, optical rotatory dispersion, CD and Raman laser.[21] These studies suggest that stacking interactions may be the major driving force in the formation of double helical DNA. Stacking appears to be important in triplex formation also. Xodo has shown that higher triplex stability can be achieved by substituting cytosines with 5-methylcytosines. Apparently, the increased stability is due to the hydrophobic methyl groups which form a hydrophobic helical spine within the major groove resulting from the elimination of hydrating water molecules from the Watson-Crick duplex, and this release of water molecules results in an entropically favourable complexation.[22] Froehler has also found that a 5-propynyl analog has a similar effect.[23]

It is, therefore, plausible that the nucleosides **4** and **6** are so hydrophilic that they prefer the solvent water to the major groove and, as a consequence, have a limited accessibility of the exocylic binding moiety to the **CG** base pair. However, the introduction of the anthraniloyl group which is less hydrophilic did not appear to promote triplet formation.

Experimental

General Procedures. DMF was distilled from barium oxide under reduced pressure (43-44 °C, 13 mmHg), and stored over 3Å sieves, or purchased from Auspep Melbourne already distilled and stored over 3Å sieves. Pyridine and diisopropylamine were distilled over CaH_2 at atmospheric pressure, and stored over 5Å molecular sieves. Dichloromethane was distilled over CaH_2, and stored over 3Å sieves.

NMR spectra were recorded by a JEOL GX-400 or Varian 300 spectrometer at 399.9 or 299.9 MHz (^1H) respectively. ^{13}C NMR spectra were recorded at 99.98 MHz by a

JEOL GX-400. DEPT experiments were performed with a 135° ^1H selection pulse. The internal references were residual DMSO-d_5 and DMSO-d_6. ^{31}P NMR spectra were obtained on a Varian 300 spectrometer at 121.42 MHz, and the external reference was 85% H_3PO_4/H_2O. UV spectra were recorded on a Varian Cary 1 spectrophotometer. Elemental analyses were obtained from CMAS Pty. Ltd. Melbourne. Low resolution FAB mass spectra were obtained in a thioglycerol matrix on a VG Micromass spectrometer.

HPLC was carried out on a Waters 600 multisolvent delivery system connected to a Waters 484 absorbance detector. Preparative RP-HPLC purification of oligonucleotides was carried out on a Synchrom (pore size 300Å, particle size 6.5mm) C_{18} column (250 x 10 mm). A Phenomenex ULTRACARB 10 ODS 20 (pore size 90Å, particle size 10 mm) C_{18} column (250 x 22.5 mm) was used for preparative purification of nucleosides. Analyses of nucleosides were carried out on a RP-HPLC Phenomenex ULTRACARB 5 ODS 30 (pore size 60Å, particle size 5mm) C_{18} column (250 x 4.6 mm).

Oligonucleotides were synthesized on an Applied Biosystems 380A DNA synthesizer. Thermal denaturation data was obtained by a Varian Cary 1 UV-Visible spectrophotometer connected to a temperature controller. Thermal denaturation calculations were carried out using a Cary Thermal Easy program. Melting points were recorded on a Gallenkamp melting point apparatus and are uncorrected.

For those compounds that are recorded in the literature, we report only new spectroscopic data to permit more complete characterisation of the compounds.

***N*^{4},3',5'-Triacetyl-2'-deoxycytidine**[24] **(1).** UV λ_{max} 248, 299 nm (MeOH); ^1H NMR (400 MHz, $CDCl_3$/TMS) δ 8.01 (d, J = 7.6 Hz, 1 H, H-6), 7.49 (d, J = 2.0 Hz, 1 H, H-5), 6.22 (t, J = 6.6 Hz, 1 H, H-1'), 5.2 (d, J = 7.3 Hz, 1 H, H-3'), 4.33 (m, 3 H, H-4', CH_2-5'), 2.8 (m, 1 H, CH_2-2'), 2.44 (s, 3 H, *N*-Acetyl), 2.06, 2.10 (2s, 6 H, 2 x AcO); ^{13}C NMR (100 MHz, $CDCl_3$/TMS) δ 170.61, 170.28, 170.13 (3 x C=O of AcO), 162.7 (C4), 154.5 (C2), 143.6 (C6), 96.4 (C5), 87.4 (C1'), 83.0 (C4'), 74.0 (C3'), 63.5 (C5'), 39.0 (C2'), 30.8, 24.8, 20.7 (3 x CH_3 of AcO); low-resolution FAB-MS, 354 (M + H), 376 (M + Na).

3',5'-Diacetyl-2'-deoxycytidine[15,16] **(2).** UV λ_{max} 239, 272 nm (MeOH); ^1H NMR (400 MHz, DMSO-d_6) δ 7.62 (d, J = 8 Hz, 1 H, H-6), 7.25 (bd, 2 H, NH_2), 6.17 (t, J = 6.9 Hz, 1 H, H-1'), 5.7 (d, J = 8 Hz, 1 H, H-5), 5.16 (m, 1 H, H-3'), 4.2 (d, J = 8 Hz, 2 H, CH_2-5'), 4.1 (m, 1 H, H-4'), 2.09, 2.04 (2s, 6 H, 2 x AcO); low-resolution FAB-MS 312 (M + H), 334 (M + Na).

N^4-Ureidocarbonyl-2'-deoxycytidine[12] **(4).** ^{13}C NMR (400 MHz, DMSO-d_6) δ 161.9 (C-4), 152.0, 153.2, 153.8 (3 x C=O), 144.7 (C-6), 94.7 (C-5), 87.8 (C-4'), 86 (C-1'), 69.9 (C-3'), 60.9 (C-5'), 40.7 (C-2').

N^4-Carbamoyl-2'-deoxycytidine[12] **(6).** 2'-Deoxycytidine (0.1 g, 0.44 mmole) was dissolved in anhydrous DMF (0.5 mL) at 60 °C. To this solution was added phenyl carbamate (0.24 g, 1.76 mmole) and the reaction mixture was stirred at 60 °C until TLC indicated that the 2'-dC had been consumed (72 h). The mixture was evaporated to dryness and then purified by silica gel chromatography using a gradient of $CHCl_3$/MeOH to give 80 mg (67%) of **6**, homogeneous by RP-HPLC and NMR; Rf 0.66 ($CHCl_3$/MeOH, 7:3); mp 180-181 °C, lit. 179-181 °C; ^{13}C NMR (100 MHz, DMSO-d_6) δ 162.36 (C-4), 154.2, 153.59 (2 x C=O), 143.34 (C-6), 95 (C-5), 87.7 (C-4'), 85.7 (C-1'), 69.9 (C-3'), 60.98 (C-5'), 40.65 (C-2')

N^4-Anthraniloyl-2'-deoxycytidine (9). 2'-Deoxycytidine (3.0 g, 13.16 mmole) was dissolved in anhydrous DMF (8 mL) with warming. To this solution was added a two-fold excess of isatoic anhydride (4.29 g, 26.33 mmole) which was predissolved in dry DMF (8 mL) with heating. The mixture was then stirred at rt. The reaction was monitored by TLC and the formation of a complex mixture was observed. The reaction was continued until no further formation of the required compound was observed (90 h), and the reaction mixture was then rotary evaporated to dryness. The crude residue was purified by silica gel flash chromatography using a gradient of $CHCl_3$/MeOH (99:1, 98:2, 95:5, 90:10) as the eluant. The relevant fractions were pooled and the solvent removed in *vacuo*. The purified product was then recrystallized from ethanol to give 2.1 g (46%) of **9** as yellow crystals, homogeneous by RP-HPLC and NMR; mp 186.5-187.0 °C; Rf 0.23 ($CHCl_3$/MeOH, 9:1); ^1H NMR (400 MHz, DMSO-d_6) δ 10.84 (bs, 1 H, NH ex), 8.3 (d, J = 7.5 Hz, 1 H, H-6), 7.7 (d, J = 7.4 Hz, 1 H, H-c), 7.2 (dd, J = 1.89, 1.9 Hz, 2 H, H-d & H-e), 6.75 (d, J = 7.4 Hz, 1 H, H-f), 6.59 (bs, 2 H, NH_2), 6.5 (d, 1 H, H-5), 6.12 (t, J = 6.1 Hz, 1 H, H-1'), 5.25 (d, 1 H, OH ex), 5.03 (t, 1 H, OH ex), 4.2 (m, 1 H, H-3'), 3.8 (m, 1 H, H-4'), 3.6 (m, 2 H, 5'-CH_2), 2.0, 2.25 (m, 2 H, 2'-CH_2); ^{13}C NMR (100 MHz, DMSO-d_6) δ 169.0 (C-a), 163.0 (C-4), 154.49, 150.93 (2 x C=O), 144.4 (C-6), 133.5, 129.8 (C-c & C-f), 116.57, 114.6 (C-d & C-e), 112.6, (C-b), 96.0 (C-5), 97.8 (C-4'), 86.0 (C-1'), 69.9 (C-3'), 60.9 (C-5'), 40.87 (C-2'); low-resolution FAB-MS, 347 (M + H), 369 (M + Na). Anal. Calcd for $C_{16}H_{18}O_5N_4$: C 55.50, H 5.24, N 16.18. Found: C 55.32, H 5.30, N 16.11.

N^4-Ureidocarbonyl-5'-*O*-di(*p*-anisyl)phenylmethyl-2'-deoxycytidine. 4-(1,2,4-Triazol-1-yl)-1-[2'-deoxy-β-D-ribofuranosyl]-2-pyrimidone[17] was tritylated according to the method of Koh and Dervan[25] with 4,4'-dimethoxytrityl chloride. 4-(1,2,4-Triazol-1-yl)-1-[5'-*O*-di(*p*-anisyl)phenylmethyl-2'-deoxy-β-D-ribofuranosyl]-2-pyrimidone (**7**) (0.075 g, 0.128 mmole) was dissolved in anhydrous DMF (1 mL) at 50 °C. Biuret (0.133 g, 1.28 mmole) was added and the mixture was stirred at 60 °C until TLC indicated that, after 72 h, **7** had been nearly all consumed and by-products (including 2'-dC) were distinct. The reaction mixture was rotary evaporated to dryness and the residue dissolved in $CHCl_3$/MeOH/Et_3N (90:9:1) and then purified by silica gel chromatography using a gradient of $CHCl_3$/MeOH/Et_3N (99:0:1, 98:1:1, 97:2:1, 95:4:1, 92:7:1, 90:9:1). The fractions with the R_f 0.49 ($CHCl_3$/MeOH/Et_3N, 90:9:1) were pooled and the solvent removed in *vacuo* to give 0.05 g (63%) of **8** homogeneous by RP-HPLC and NMR; ^{1}H NMR (400 MHz, DMSO-d_6) δ 10.6 (bs, 1 H, NH ex), 10.4 (bs, 1 H, NH ex), 8.1 (d, J = 10 Hz, 1 H, H-6), 7.3-7.2 (m, 9 H, Ph and ortho CH of $PhOCH_3$), 6.87 (m, 4 H, meta CH of $PhOCH_3$), 6.4 (bd, 1 H, H-5), 6.09 (t, J = 5.8 Hz, 1 H, H-1'), 5.3 (d, J = 4.76 Hz, 1 H, OH ex), 4.2 (m, 1 H, H-3'), 3.9 (m, 1 H, H-4'), 3.3 (m, 2 H, CH_2-5'), 3.17 (bs, 1 H, NH ex), 2.2, 2.0 (m, 2 H, CH_2-2'); low-resolution FAB-MS, 616 (M + H), 638 (M + Na).

The nucleosides **4** and **6** were also tritylated with 4,4'-dimethoxytrityl chloride by Koh and Dervan's method[25], and the products purified by chromatography on silica.

N^4-Carbamoyl-5'-*O*-di(*p*-anisyl)phenylmethyl-2'-deoxycytidine; ^{1}H NMR (400 MHz, DMSO-d_6) δ 9.7 (s, 1 H, NH ex), 7.95 (d, J = 7.3, 1 H, 6-H), 7.3-7.2 (m, 9 H, Ph and ortho CH of $PhOCH_3$), 6.85 (d, J = 8.79 Hz, 4 H, meta CH of $PhOCH_3$), 6.16 (bd, 1 H, 5-H), 6.1 (t, J = 5.8 Hz, 1 H, 1'-H), 5.3 (d, J = 4.76 Hz, 1 H, OH ex), 4.2 (m, 1 H, 3'-H), 3.9 (m, 1 H, 4'-H), 3.73 (s, 6 H, 2 x OCH_3), 3.2 (d, 2 H, 5'-CH_2), 2.92 (bs, 1 H, NH ex), 2.22, 2.1 (m, 2 H, 2'-CH_2); low-resolution FAB-MS, 573 (M + H), 595 (M + Na). **N^4-Anthraniloyl-5'-*O*-di(*p*-anisyl)phenylmethyl-2'-deoxycytidine**; ^{1}H NMR (400 MHz, DMSO-d_6) δ 11.96 (s, 1H, NH), 8.59 (dd, J = 1.84, 1.83 Hz, 2 H, H-d & H-e), 8.13 (d, J = 14.29 Hz, 1 H, H-6), 8.02 (d, J = 9.0 Hz, 1H, H-f,7.2-7.4 (m, 9 H, Ph and ortho CH of $PhOCH_3$), 7.15 (d, J = 7.33, 1 H, H-c), 7.08 (s, 1 H, NH_2), 6.88 (m, 4H, meta CH of $PhOCH_3$), 6.1 (t, J = 7.22, 1 H, H-1'), 5.87 (d, J = 14.29 Hz, 1H, H-5), 5.27 (d, 1H, OH ex), 4.2 (m, 1 H, H-3'), 4.75 (m, 1 H, H-4'), 3.12 (m, 2 H, H-5'), 1.9, 2.0 (m, 2 H, H-2'); low-resolution FAB-MS, 649 (M + H), 671 (M + Na). Anal. Calcd for $C_{37}H_{36}O_7N_4$: C, 66.5; H, 5.8; N, 8.64. Found: C, 66.7; H, 6.0; N, 8.7.

β-Cyanoethyl *N,N*-diisopropyl-phosphoramidites were prepared as reported by Koh and Dervan,[25] and were homogeneous by ^1H and ^{31}P NMR (only the ureidocarbonyl derivative required chromatography to achieve this homogeneity). **N^4-Ureidocarbonyl-5'-*O*-di(*p*-anisyl)phenylmethyl-2'-deoxycytidine-3'-(β-cyanoethyl *N,N*-diisopropylphosphoramidite)** (49%); ^1H NMR (300 MHz, CD_2Cl_2) δ 10.6 (bs 1 H, NH ex), 10.4 (bs, 1 H, NH ex), 8.1 (d, *J* = 9 Hz, 1 H, H-6), 7.3-7.2 (m, 9 H, Ph and ortho CH of $PhOCH_3$), 6.9 (m, 4 H, meta CH of $PhOCH_3$), 6.43 (d, *J* = 8.8 Hz, 1 H, H-5), 6.09 (t, 1 H, H-1') 4.29 (m, 1 H, H-3'), 3.85 (m, 1 H, H-4'), 3.80 (m, 6 H, DMT OCH_3), 3.4-3.6 (m, 2H, isopropyl CH), 3.29 (m, 2 H, CH_2-5'), 2.60, 2.47 (2 x t, 4 H, *J* = 6.4 Hz, OCH_2CH_2CN), 2.2, 2.39 (m, 2 H, CH_2-2'), 1.3, 1.4 (2 x d, *J* = 6.8 Hz, 2 x isopropyl CH_3), 1.0, 1.2 (2 x d, *J* = 6.8 Hz, 6 H, 2 x isopropyl CH_3); ^{31}P NMR (121.42 MHz, CD_2Cl_2) d 149.2, 149.5 (2s, 2 diastereoisomers); low-resolution FAB-MS, 817 (M + H), 839 (M + Na). **N^4-Carbamoyl-5'-*O*-di(*p*-anisyl)phenylmethyl-2'-deoxycytidine 3'-(β-cyano-ethyl *N,N*-diisopropylphosphoramidite)**; ^1H NMR (300 MHz, CD_2Cl_2) δ 10.6 (bs, 1 H, NH ex), 10.4 H, H-6), 7.3-7.2 (m, 9 H, Ph and ortho CH of $PhOCH_3$), 6.9 (m, 4 H, meta CH of $PhOCH_3$), 6.43 (d, *J* = 8.8 Hz, 1 H, H-5), 6.09 (t, 1 H, H-1'), 4.26 (m, 1 H, H-3'), 3.85 (m, 1 H, H-4'), 3.8 (m, 6 H, DMT OCH_3), 3.45-3.6 (m, 2H, 2 x isopropyl CH), 3.29 (m, 2 H, CH_2-5'), 2.6, 2.47 (2t, *J* = 6.4 Hz, 4 H, OCH_2CH_2CN), 2.39 (m, 2 H, CH_2-2'), 1.2, 1.35 (2d, *J* = 6.9 Hz, 6 H, isopropyl CH_3), 1.1, 1.2 (2d, *J* = 7.0 Hz, 6 H, isopropyl CH_3); ^{31}P NMR (121.42 MHz, CD_2Cl_2) δ 149.5, 148.0 (2s, 2 diastereoisomers); low-resolution FAB-MS, 773 (M + H), 795 (M + Na). **N^4-Anthraniloyl-5'-*O*-di(*p*-anisyl)phenylmethyl-2'-deoxycytidine 3'-(β-cyanoethyl *N,N*-diisopropylphosphoramidite)**; ^1H NMR(400 MHz, CD_2Cl_2) δ 8.70, 8.74 (dd, *J* = 7.69, 7.7 Hz, 2H, H-d & H-e), 8.13 (d, *J* = 8.0 Hz, 1 H, H-6), 7.72 (d, *J* = 7.0 Hz, 1 H, H-f), 7.57 (d, *J* = 7.0 Hz, 1 H, H-c), 7.2-7.5 (m, 9 H, ortho CH of $PhOCH_3$), 6.85 (m, 4 H, meta, CH of $PhOCH_3$), 6.01 (t, 1 H, H-1'), 5.98 (d, 1 H, H-5), 4.85 (m, 1 H, H-3'), 4.1-4.13 (m, 1 H, H-4'), 3.85 (m, 1 H, isopropyl CH), 3.8 (2s, 6 H, 2 x DMT OCH_3), 3.6 (m, 1H, isopropyl CH), 2.77 (m, 2 H, CH_2-5'), 2.6, 2.7 (2t, *J* = 6.6 Hz, 4 H, OCH_2CH_2CN), 2.4 (m, 2 H, CH_2-2'), 1.2,1.3 (2d, *J* = 7.0 Hz, 6 H, isopropyl CH_3), 1.1, 1.2 (2d, *J* = 6.0 Hz, isopropyl CH_3); ^{31}P NMR (121.42 MHz, CD_2Cl_2) δ 147.65, 147.66 (2s, 2 diastereoisomers); low-resolution FAB-MS, 849 (M + H), 871 (M + Na), 979 (M + Na + TG).

General method for the synthesis of oligonucleotides. Normal trityl-on oligonucleotides were synthesized on an Applied Biosystems 380Å DNA synthesizer using standard β-cyanoethyl nucleoside phosphoramidites on a 0.2 μmole scale. Modified oligonucleotides were synthesized as trityl-off oligonucleotides using Expedites instead of

standard phosphoramidites (with *t*-butylphenoxyacetic anhydride/methylimidazole for the capping step), on a 0.2 μmole scale.

Purification of oligonucleotides. (i) *Preparative RP-HPLC purification of trityl-on oligonucleotides:* The 5'-dimethoxytrityl-protected oligonucleotides were purified by RP-HPLC using a triethylammonium acetate buffer pH 7.0 and acetonitrile. UV detection was at 260 nm. Fractions were collected and stored with 5 mg/mL of Tris base at 0 °C if purification was to be carried out over more than one day. The relevant fractions were pooled and checked for homogeneity by analytical RP-HPLC. The oligonucleotides were then detritylated by the addition of an equal volume of AR acetic acid and leaving to react for 15 min. The solvent was then removed by rotary evaporation or butanol extraction. The oligonucleotides were redissolved in 10 mL of water, and 20 mL of diethyl ether was added. The flask was shaken for 10 sec and then left standing at rt for 5-10 min to allow the layers to separate. The ether was then removed and the extraction repeated twice. The aqueous solution containing the oligonucleotides was then lyophilized.

(ii) *Preparative PAGE purification of oligonucleotides.*
The oligonucleotides were dissolved in 0.1 mM EDTA and the electrophoreses run on a 20% denaturing polyacrylamide gel using a TBE buffer; and the bands detected by Methylene Blue staining. Relevant electrophoretic mobility bands were collected and then left in 0.1 mM EDTA (~2-4 mL) for 72 h. The aqueous solution containing the oligonucleotides was then dialysed against H_2O in a 2000 MW cutoff tubing with at least three changes of water every 3 h. The amount of oligonucleotide was then determined spectrophotometrically and the oligonucleotide solutions lyophilized.

Characterization of modified oligonucleotides. Modified oligonucleotides were characterized by deoxynucleoside composition analysis as described by Connolly.[26]

Gel mobility shift analyses of duplexes and triplexes
Samples (10 μg of each duplex strand and 8.9 μg of the third strand) were dissolved in an annealing buffer (30 μL), [10 mM potassium phosphate buffer, pH 6.1, 1 mM EDTA, 1 or 2 mM spermine. The mixture was then heated to 70 °C, returned to rt and then left standing at 4 °C for 72-120 h. A 20% non-denaturing polyacrylamide gel were made up to a width of 0.8 mm. The gel was pre-electrophoresed for ~2.5 h at 4 °C with constant current at 15W (300-500V), using a Tris buffer (89 mM Tris, 2 mM EDTA, 1 mM spermine, pH 6.1 adjusted with phosphoric acid). The samples were then loaded onto the gel in 15% Ficoll and after electrophoresis the gel was stained with methylene blue and photographed.

Method for determining thermal denaturation/naturation profiles and thermodynamics. Melting studies. Thermal transitions were recorded at 260 nm using a Varian Cary 1 UV-Visible spectrophotometer equipped with the Thermal Easy program and connected to a Cary 1 temperature controller. In addition to using teflon stoppers the rims of the quartz cuvettes were coated with Sigmacoat to minimise evaporation via a capillary effect. Degassed (helium) solutions, 2 μmolar each strand, were prepared in an annealing buffer [(10 mM potassium phosphate, pH 6.1), 1 mM EDTA,] containing NaCl at various concentrations. The mixtures were then heated to 70 °C, returned to rt, filtered using a Millipore filter, and then left standing at rt overnight. Melting data (absorbance *vs* temperature) was recorded automatically using a Cary Themal Easy program, wavelength 260 nm, signal averaging time 1.0 sec, data interval 0.10 °C, slit band width 2.0 nm, and ramp rate 0.25 °C/min (from 5-50 °C), and 1.0 °C (from 50-90 °C) followed by identical negative ramping. During the first 15 min of (+) ramping (5-20 °C) the cell compartment was flushed with dry N_2, in order to reduce condensation of the quartz cuvettes. Multiple runs of duplicate samples were conducted, and T_ms were reproducible to ± 2.0 °C. Hyperchromicity calculations were performed using two different methods: the first involved the Varian Cary Thermal Easy program which employs a modified version of the Marky method.[27]

REFERENCES

(1) Cohen, J. S.; Hogan, M. E. *Scientific American.* **1994**, *Dec.* 50-55.

(2) Le Doan, T.; Perrouault, L.; Praseuth, D.; Habhoub, N.; Decout, J. L.; Lhomme, J.; Helene, C. *Nucleic Acids Res.* **1987**, *15*, 7749-7760.

(3) Moser, H. E.; Dervan, B. P. *Science* **1987**, *238*, 645-650.

(4) (a) Helene, C. *Anti-cancer Drug Des.* **1991**, *6*, 569-584. (b) Helene, C. *Eur. J. Cancer.* **1991**, *27*, 1466-1471. (c) Helene, C.; Thuong, N. T.; Harel-Ballan, A. *Annals. N. Y. Academy of Science* **1992**, *660*, 27-36.

(5) Frank-Kamenetskii, M. D.; Mirkin, S. M. *Annu. Rev. Biochem.* **1995**, *64*, 65-95.

(6) (a) Maher, L. J.; Wold, B. J.; Dervan, P. B. *Science* **1989**, *245*, 725-730. (b) Maher, L. J.; Dervan, P. B.; Wold, B. *Biochemistry* **1990**, *29*, 8820-8826. (c) Griffin, L. C.; Dervan, P. B. *Science* **1989**, *245*, 967-971.

(7) Thuong, N. T.; Helene, C.; *Angew. Chem. Int. Engl.* **1993**, *32*, 666-690.

(8) Zhou, B-W.; Marchard, C.; Asseline, V.; Thuong, N. T.; Sun, J-S.; Garestier, T.; Helene, C. *Bioconjugate Chem.* **1995**, *6*, 516-523.

(9) (a) Froehler, B. C.; Ricca, D. J. *J. Am. Chem. Soc.* **1992**, *114*, 8320-8322. (b) Miller, P. S.; Bhan, P.; Cushman, C. D.; Trapane, T. L. *Biochemistry* **1992**, *31*, 6788-

6793. (c) Hunziker, J.; Priestley, S.; Brunar, H.; Dervan, P. B. *J. Am. Chem. Soc.* **1995**, *117*, 2661-2662. (d) Barawkar, D. A., Rajeev, K. G., Kumar, V. A., Ganesh, K. N. Nucleic Acid Research **1996**, *24*, 1229-1237. (e) Ono, A.; Ts'o P. O. P.; Kan. L. *J. Org. Chem.* **1992**, *57*, 3225-3230. (f) Miller, P. S., Bi, G., Kipp, S. A., Fok, V., Dehong, R. K. *Nucleic Acids Research* **1996**, *24*, 730-736.

(10) (a) Xodo, E. L.; Manzini, G.; Quadrifoglio, G. A.; van der Marel, G.; van Boom, J. H. *Nucleic Acids Res.* **1991**, *19*, 5625-5631. (b) Plum, G. E.; Park, Y-W.; Singleton. S. F.; Dervan, P. B., Breslauer, K. J. *Proc. Natl. Acad. Sci. USA* **1990**, *87*, 9436-9440.

(11) (a) Froehler, B. C.; Wadwani, S.; Terhorst, T. J.; Gerrard, S. R. *Tetrahedron Lett.* **1992**, *33*, 5307-5310. (b) Xodo, E. L.; Manzini, G.; Quadrifoglio, G. A.; van der Marel, G.; van Boom, J. H. *Nucleic Acids Res.* **1991**, *19*, 5625-563112

(12) Kumar, S.; Leonard, N. J. *J. Org. Chem.* **1988**, *53*, 3959-3967.

(13) Shea, R.G.; Ng, P.; Bischofberger, N. *Nucleic Acids Res.* **1990**, *18*, 4859-4866.

(14) (a)Wang, E.; Malek, S.; Feigon, J. *Biochemistry* **1992**, *31*, 4838-4846. (b) Macaya, R. F.; Gilbert, D. E.; Malek, S.; Sinsheimer, J. S.; Feigon, J. *Science* **1991**, *254*, 270-274. (c) Miller, P. S.; Cushman, C. D. *Biochemistry* **1993**, *32*, 2999-3004. (d) Griffin, L. C., Dervan, P. B. *Science* **1989**, *245*, 967-971.

(15) Ishido, Y.; Nakazaki, N.; Sakairi, N. *J. Chem. Soc. Perkin Trans.* **1979**, *1*, 2088-2098.

(16) (a) Kierzek, R.; Ito, H.; Bhatt, R.; Itakura, K.*Tetrahedron Lett.* **1981**, *2*, 3761-3764. (b) Corey, E.J.; Gras, J-L.; Ulrich, P. *Tetrahedron lett.* **1976**, *11*, 309-312.

(17) Webb, T. R.; Matteucci, M. D. *Nucleic Acids Res.* **1986**, *14*, 7661-7675.

(18) Singleton, S. F.; Dervan, P. B. *Biochemistry* **1993**, *32*, 13171-13179.

(19) (a) Meyer, J. F.; Wagner, E. C. *J. Org. Chem.* **1943**, *8*, 239-252. (b) Kolbe. *J. Prakt. Chem.* **1884**, *30*, 469-476. (c) Feldman, J. R.; Wagner, E. C. *J. Org. Chem.* **1942**, *7*, 31-47.

(20) Huang, C-Y.; Miller, P.S.; *J.Am. Chem. Soc.* **1993**, *115*, 10456-10457.

(21) (a) Kang, H.; Chou, P-J.; Johnson, C. W., Jr.; Weller, D.; Huang, S-B.; Summerton, J. E. *Biopolymers* **1992**, *32*, 1351-1363. (b) Brahms, J.; Maurizot, J. C.; Michelson, A. M. *J. Mol. Biol.* **1967**, *25*, 481-495. (c) Powell, J. T.; Richards, E. G.; Gratzer, W. B. *Biopolymers* **1972**, *11*, 235-250. (d) Cantor, C. R.; Tinoco, I., Jr. *J. Mol. Biol.* **1965**, *13*, 65-77. (e) Warshaw, M. M.; Cantor, C. R. *Biopolymers* **1970**, *9*, 1079-1103. (f) Lowe, M. J.; Schellman, J. A. *J. Mol. Biol.* **1972**, *65*, 91-109. (g) Vournakis, J. N.; Poland, D.; Scherager, H. A. *Biopolymers* **1967**, *5*, 103-122. (h) Dewey, T. G.; Turner, D. H. *Biochemistry* **1979**, *18*, 5757-5762. (i) Dewey, T. G.;

Turner, D. H. *Biochemistry* **1980**, *19*, 1681-1685. (j) Freier, S. M.; Hill, K. O.; Dewey, T. G.; Marky, L. A.; Breslauer, K. J.; Turner, D. H. *Biochemistry* **1981**, *20*, 1419-1426. (k) Olsthoorn, C. S. M.; Doornbos, J.; De Leeuw, H. P. M, Altona, C. *Eur. J. Biochem.* **1982**, *125*, 67-382. (l) Kondo, N. S.; Danyluk, S. S. *Biochemistry* **1976**, *15*, 756-768. (m) Lee, C. H.; Ezra, F. S.; Kondo, N. S.; Sarma, R. H.; Danyluk, S. S. *Biochemistry* **1976**, *15*, 3627-3639.

(22) Xodo, E. L.; Manzini, G.; Quadrifoglio, G. A.; van der Marel, G.; van Boom, J. H. *Nucleic Acids Res.* **1991**, *19*, 5625-5631.

(23) Froehler, B. C.; Wadwani, S.; Terhorst, T. J.; Gerrard, S. R. *Tetrahedron Lett.* **1992**, *33*, 5307-5310.

(24) (a) Robins, M. J.; MacCross, M.; Naik, S. R.; Ramani, G. *J. Am. Chem. Soc.* **1976**, *98*, 7381-7389. (b) Koster, H.; Kulikowski, K.; Liese, T.; Herkens, W.; Kholi, V. *Tetrahedron* **1981**, *37*, 363-369.

(25) Koh, J. S.; Dervan, P. B. *J. Am. Chem. Soc.* **1992**, *114*, 1470-1478.

(26) Connolly, B. A. In *Oligonucleotides and Analogues a Practical Approach* **1991**, ed. F. Eckstein. Oxford University Press, NewYork. pp.179-180.

(27) (a) Marky, L. A.; Breslauer, K. J. *Biopolymers*, **1987**, *26*, 1601-1620.(b) Marky, L. A.; Kallenbach, N. R.; McDonough, K. A.; Seeman, N. C.; Breslauer, K. J. *Biopolymers* **1987**, *26*, 1621-1634.