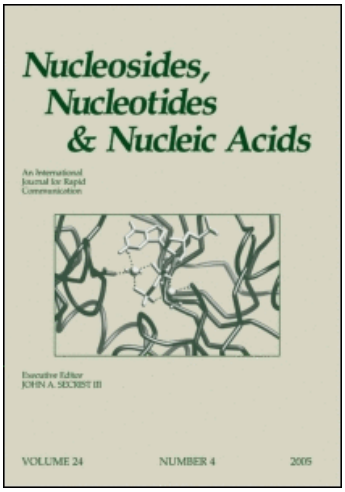


Nucleosides, Nucleotides and Nucleic Acids
Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t713597286

Synthesis and Triplex Forming Properties of Pyrimidine Derivative Containing Extended Functionality

Diego A. Gianolio[a]; Larry W. McLaughlin[a]
[a] Department of Chemistry, Merkert Chemistry Center, Chestnut Hill, MA

To cite this Article Gianolio, Diego A. and McLaughlin, Larry W.(1999) 'Synthesis and Triplex Forming Properties of Pyrimidine Derivative Containing Extended Functionality', Nucleosides, Nucleotides and Nucleic Acids, 18: 8, 1751 — 1769
To link to this Article: DOI: 10.1080/07328319908044841
URL: http://dx.doi.org/10.1080/07328319908044841

NUCLEOSIDES & NUCLEOTIDES, 18(8), 1751–1769 (1999)

SYNTHESIS AND TRIPLEX FORMING PROPERTIES OF PYRIMIDINE DERIVATIVE CONTAINING EXTENDED FUNCTIONALITY

Diego A. Gianolio and Larry W. McLaughlin*
Department of Chemistry
Merkert Chemistry Center
2609 Beacon St.
Chestnut Hill, MA 02167

ABSTRACT

Two pyrimidine nucleosides have been synthesized containing extended hydrogen bonding functionality. In one case the side chain is based upon semicarbazide and in the second monoacetylated carbohydrazide was employed. DNA sequences could be prepared using both analogue nucleosides in a reverse coupling protocol, and provided that the normal capping step was eliminated and that the iodine-based oxidizing solution was replaced with one based upon 10-camphorsulfonyl oxaziridine. Both derivatives exhibited moderate effects in targeting selectively C-G base pairs embedded within a polypurine target sequence.

INTRODUCTION

Oligodeoxynucleotides can interact with double stranded DNA forming Hoogsteen-type hydrogen bonds within the major groove and result in triple helices.[1-3] Once the third strand has bound the duplex DNA target, it can then interfere with biological processes such as gene expression[1] and selective cleavage of DNA,[4] among the most studied processes. Oligodeoxynucleotides can be viewed therefore as potential pharmaceuticals and offer bright prospects for medicinal developments. It is of considerable interest for these applications to recognize selectively a chosen sequence of DNA, and to date homopurine tracts are the only viable target sites.

The most studied cases involve third strands containing pyrimidines that bind to the purines in a duplex in a parallel orientation following the pyr-pur-pyr motif and resulting in bidentate Hoogsteen-type hydrogen bonds (T-A-T and C+-G-C). The third strand looses it's ability to form stable triple helices when the DNA target sequence is a combination of

pyrimidine and purine bases, presumably due to the pyrimidine inability to offer two Hoogsteen type hydrogen bonds. However, triplexes have been observed[5] when the nucleobase G is matched to a single T-A base pair interrupting an otherwise homo-pu-py base pair motif forming the triad G-T-A. Likewise triple helices are possible with the T- or U-C-G triads. This case may involve the 4-carbonyl of T or U serving as a hydrogen bond acceptor for the N^4-amino group of C. These triplexes exhibited only moderate stability, thus the synthesis of non-natural bases that can selectively and effectively bind to the C-G and/or T-A base pairs is of particular interest.

Recent attempts to generate such analogs have included carbocyclic non-natural nucleosides with nucleobases attached via amide bonds containing aminothiazole and pyrrole moieties.[6] These analogs exhibited selectivity for: C-G ~ T-A > G-C ~ A-T, suggesting that binding affinity was derived from sequence specific intercalation with preferential stacking ability over py-pu base pairs, these analyses did not distinguish between C-G and T-A base pairs. In other attempts to synthesize non-natural nucleotides, the original pyrimidine structure of C has been retained and a functionality has been added at the N^4 position.

Using 6-amino pyridine as an extended functionality for C,[7] triplexes can be formed when a target base pair is embedded in the usual py-pu-py motif. The more stable triplexes are: X-A-T and X-C-G, where additional hydrogen bonding interactions may be involved invoking an unusual imino tautomeric form, which provides a pattern of hydrogen bond donor and acceptor groups that would permit the formation of two hydrogen bonds with the C-G base pair and three hydrogen bonds with the A-T base pair. Only a single hydrogen bond can be drawn between the pyridinyl amino group and the G-C or T-A base pair, accounting for the differential stabilities of these complexes.

Aliphatic amino chains instead of an aromatic ring have also been added as extended functionality at the N^4 position of cytosine. An unnatural nucleoside has been synthesized using spermine as an extended functionality, having also a methyl group in the C5 position of C.[8] This approach offers two positive attributes in promoting the triple helix formation for both moieties: reduction of the net negative charge of the system by means of the zwitterionic character imparted by the spermine conjugation[9-11] and increasing the hydrophobicity by the aliphatic methyl group.[12,13] In this case the observed target selectivity was: G-C ~ A-T > C-G > T-A.[12]

In this paper we examine the effect of a potentially increased hydrogen bonding ability of two modified nucleosides resulting from the addition of an extension rich in hydrogen bonding acceptors and donor sites (C=O and N-H) to m^5C. These derivatives differ from a related series of compounds described recently,[13] which were reported to inhibit triple-helix formation. Our derivatives add functionality with some flexibility,

achieved by alternating facile rotating NH-NH bonds with planar amidic bonds, allowing the extension to accommodate better to the DNA base pair in the major groove. We also decided to keep the methyl group in the C5 position to favor hydrophobic interactions. We report here the synthesis of two unnatural nucleobases, their specific incorporation into DNA and their ability to form triple-helices with target sequences of DNA.

EXPERIMENTAL

NMR spectra were obtained on Varian spectrophotometers (400 MHz). Mass spectra were obtained using FAB ionization from the Mass Spectrometry Laboratory, School of Chemical Sciences, University of Illinois, Urbana, IL. Rotary evaporations were performed under reduced pressure with Buchi systems. Thin layer chromatography (TLC) was performed on Silica Gel 60 F_{254} precoated on aluminum sheets (EM Separations Technology), in CH_3OH:CH_2Cl_2 (1:19) unless otherwise specified. Anhydrous solvents and starting materials were purchased from Aldrich Chemical Company and used without further purification unless otherwise specified. UV scans and absorbances were obtained using a Beckman DU 640 spectrophotometer. Oligodeoxynucleotides were synthesized on an Applied Biosystems 381A DNA Synthesizer. 2'-Deoxynucleotide phosphoramidites and 3' terminal nucleoside controlled pore glass support (CPG) were purchased from Glenn Research (Sterling, VA). High performance liquid chromatography (HPLC) was performed on a Beckman HPLC system using C-8 reversed-phase columns (ODS-Hypersil, 5mm particle size, 120A pore) with detection at 260 nm. Oligonucleotides were desalted with Econo-Pac 10DG disposable chromatography Columns (Bio Rad, Hercules, CA). T_m measurements were performed on an AVIV Spectrophotometer, Model 14DS UV/Vis. MALDI-TOF mass spectra of the oligodeoxynucleotides were performed by Dr. D. J. Fu, Sequenom Inc., San Diego, CA.

N^1-(4-methoxytrityl)carbohydrazide (1)

1.17 g (13.0 mmol) of carbohydrazide were suspended in 25 ml of a mixture 1:1 of freshly distilled pyridine and anhydrous DMF. The mixture was then gradually heated by means of an oil bath and stirred briskly until the carbohydrazide was completely dissolved (120 °C). 1 g (3.23 mmol) of monomethoxytrityl chloride (MMT-Cl) was dissolved separately in 8 ml of freshly distilled pyridine and added dropwise to the reaction flask by means of a dripping column during 1 hour of time. After the addition the reaction was stirred an additional 15 minutes, was allowed to cool below 80 °C and then quenched with 10 ml of water. Once the reaction flask reached room temperature, the solvent was evaporated *in vacuo* until a red colored residue remained. The desired product was isolated by silica gel column chromatography using a gradient of methanol in dichloromethane

(3%), resulting in 0.389 g (1.07 mmol) of **1** (33 % yield), as a yellowish colored solid. The remaining material consisted largely of carbohydrazide not reacted and its bis-MMT derivative.

Rf (methanol/dichloromethane, 1/19): 0.25; mp : 122-126 °C ;

U.V.(methanol) : λ_{max} = 203 (ε =53.15x10^3 cm^{-1}M^{-1}), 255 (sh) nm.

^{1}H NMR (DMSO$_{d6}$) : δ = 3.70 (s, 3H, -OCH$_3$); 3.85 (bs, 2H, NH$_2$); 5.55 (bs, 1H, N^{1}H); 6.82-7.38 (m, 16H, aromatics + N^{2}H + N^{4}H) ppm.

LRMS 331 (MMT-NH-NH-CO$^+$); 289 (MMT-NH$^+$ +H$^+$); 273 (MMT$^+$).

N^1-acetyl-N^5-(4-methoxytrityl)carbohydrazide (2)

1.51 g (4.16 mmol) of **1**, were suspended in 10 ml of freshly distilled pyridine and the reaction flask was cooled to 0°C under nitrogen. 1.97 ml (20.79 mmol) of acetic anhydride were added dropwise to the mixture and after 10 minutes the reaction was allowed to stir at ambient temperature until the mixture turned clear. After a total time of 45 minutes, the reaction was quenched with water and the solvent was evaporated *in vacuo* until a solid residue was obtained. The desired product was isolated by silica gel column chromatography using a gradient of methanol in dichloromethane (2%), resulting in 1.55 g (3.83 mmol) of **2** (92 % yield) as a white solid.

Rf (methanol/dichloromethane, 1/19): 0.22; mp : 129-132 °C ;

U.V.(methanol) : λ_{max} = 202 (ε = 51.57x10^3 cm^{-1}M^{-1}), 227 (sh) nm.

^{1}H NMR (DMSO$_{d6}$) : δ = 1.76 (s, 3H, acetyl-CH$_3$); 3.71 (s, 3H, -OCH$_3$); 5.74 (bs, 1H, N^{5}H); 6.82-7.40 (m, 16H, aromatics + N^{2}H + N^{4}H); 9.46 (s, 1H, N^{1}H) ppm.

HRMS calculated for C$_{23}$H$_{25}$N$_4$O$_3$ (M+H$^+$) : 405.1927, found : 405.1929.

N^1-acetyl-N^5-hydrocarbohydrazide *p*-toluene sulfonate (3)

1.68 g (4.16 mmol) of **2** were dissolved in 15 ml of a mixture 10% methanol in dichloromethane and then cooled to 0 °C. 1.977 g (10.40 mmol) of *p*-toluenesulfonic acid monohydrate were added to the solution and the reaction was allowed to stir for 2.5 hours. The reaction was then allowed to sit on dry ice for 20 minutes to favor the formation of the product as a white precipitate. The solid was then filtered and washed with cold 10% methanol in dichloromethane, dried under vacuum to give 1.22 g (3.99 mmol) of **3** (96% yield).

mp : 166-168 °C ;

U.V.(methanol) : λ_{max} = 220 nm (ε = 1.3x10^3 cm^{-1}M^{-1}).

^{1}H NMR (DMSO$_{d6}$) : δ = 1.82 (s, 3H, acetyl-CH$_3$); 2.26 (s, 3H, toluene-CH$_3$); 7.09 (d, J = 7.6 Hz, 2H, aromatics); 7.45 (d, J = 8.0 Hz, 2H, aromatics); 8.86, 9.29, 9.67 (s, 3H, NHs); 9.90 (bs, 3H, NH$_3^+$) ppm.

^{13}C NMR (CD_3OD) : δ = 21.4, 22.2, 127.8, 130.7, 142.7, 144.1, 160.0, 174.2 ppm.
LRMS : 155 (CH_3-ph-SO_2^+); 135 (CH_3CO-NHNH-CO-$NHNH_3^+$+ $2H^+$).

5'-O-4,4'-dimethoxytrityl-3'-O-tertbutyldimethylsilyl thymidine (4)

To 1.71 g (3.1 mmol) of 5'-O-4,4'-dimethoxytrityl thymidine dissolved in 15 ml of freshly distilled pyridine was added 0.44 ml (3.1 mmol) of anhydrous triethylamine dropwise into the reaction flask and the mixture was stirred at ambient temperature for 10 minutes. To this mixture was added 1.42 g (9.4 mmol) of *t*-butyldimethylsilyl chloride and 0.86 g (12.6 mmol) of imidazole and the mixture stirred under a nitrogen atmosphere for 12 hours. The reaction was then quenched with 5 ml of methanol and the solvent was evaporated *in vacuo* until a solid residue was left. The desired product was then purified by silica gel column chromatography using a gradient of methanol in dichloromethane (3%). 1.96 g (2.9 mmol) of **4** were obtained. Yield 95 %.
Rf (methanol/dichloromethane, 1/19): 0.48; mp : 92-94 °C ;
U.V. (methanol) : λ_{max} = 203, 233 (ε = 39.5 x 10^3 $cm^{-1}M^{-1}$), λ_{min} = 267 nm.
^1H NMR ($DMSO_{d6}$) : δ = 0.02 (s, 6H, Si-CH_3); 0.76 (s, 9H, *t*-but.CH_3); 1.49 (s, 3H, base-CH_3); 2.10 - 2.31 (m, 2H, H2'2''); 3.11 - 3.24 (m, 2H, H5'5''); 3.71 (s, 6H, -OCH_3); 3.79 (m, 1H, H4'); 4.43 (m, 1H, H3'); 6.13 (t, J = 6.8 Hz, 1H, H1'); 6.86 - 7.38 (m, 13H, Ar-H); 7.52 (s, 1H, Ar-H); 11.33 (s, 1H, N-H) ppm.
^{13}C NMR ($CDCl_3$) : δ = -4.9, -4.7, 11.9, 20.7, 25.7, 41.5, 55.2, 62.9, 72.1, 84.9, 86.8, 111.0, 113.2, 113.2, 127.1, 128.0, 128.1, 130.0, 130.0, 135.4, 135.6, 144.3, 150.1, 159.7 ppm.
HRMS calculated for $C_{37}H_{46}N_2O_7Si$ ($M + H^+$): 658.3074, found : 658.3073.

5'-O-4,4'-dimethoxytrityl-3'-O-tertbutyldimethylsilyl-4-O-triisopropylphenylsulfonyl thymidine (5)

To 1.0 g (1.52 mmol) of **4** that had been dried overnight in a flask under high vacuum was added 15 ml of anhydrous dichloromethane under an atmosphere of nitrogen. To this mixture was added dropwise 0.85 ml (6.07 mmol) of anhydrous triethylamine and the flask was wrapped with aluminum foil prior to the addition of 0.827 g (2.73 mmol) of 2,4,6-triisopropylbenzenesulfonyl chloride and 46 mg (0.38 mmol) of dimethylamino-pyridine. The mixture was then stirred at ambient temperature for 18 hours. The solution (now a crimson color) was quenched with 5 ml of methanol and evaporated *in vacuo* until an oily residue remained. The desired product was purified by silica gel column chromatography, using a gradient of hexane in dicholoromethane. A yellowish colored solid was obtained that was dried for few hours under high vacuum and immediately stored under nitrogen at 0 °C yielding 0.73 g (0.79 mmol) of **5**, 52 % yield.

R_f (methanol/dichloromethane, 1/19): 0.86; mp : 87-91 °C ;
U.V.(methanol) : λ_{max} = 202, 233 (ε =3.94x10 ^{3} $cm^{-1}M^{-1}$), λ_{min} = 282, 303 nm.
^1H NMR ($DMSO_{d6}$) : δ = 0.12 - 0.24 (m, 6H, Si-CH_3); 0.72 (s, 9H, t-but.CH_3); 1.18 - 1.22 (m, 18H, isopr.CH_3); 1.59 (s, 3H, base-CH_3); 2.24 (t, J = 6 Hz, 2H, H2'2''); 2.94 (m, 1H, CH); 3.15 - 3.42 (m, 2H, H5'5''); 3.71 (s, 6H, -OCH_3); 3.86 (m, 1H, H4'); 4.15 (m, 2H, CH); 4.36 (m, 1H, H3'); 5.96 (t, J = 6.0 Hz, 1H, H1'); 6.84 - 7.30 (m, 13H, Ar-H); 7.32 (s, 2H, Ar-H); 8.17 (s, 1H, Ar-H) ppm.
^{13}C NMR ($CDCl_3$) : δ = -5.0, -4.7, 11.5, 17.8, 23.4, 23.5, 24.3, 24.6, 25.6, 29.6, 34.2, 42.1, 55.2, 61.8, 70.4, 86.7, 86.8, 87.1, 103.6, 113.2, 113.2, 124.0, 127.2, 128.0, 128.2, 130.0, 135.2, 143.2, 144.1, 151.1, 153.7, 154.3, 158.7, 166.3 ppm.
LRMS : 925 (M^+), 303 (DMT^+).

N^4-ureido-5'-O-4,4'-dimethoxytrityl-3'-O-tertbutyldimethylsilyl-5-methyl-2'-deoxycytidine (6)

To 0.181 g (1.62 mmol) of semicarbazide chloride evaporated three times from freshly distilled pyridine and under nitrogen atmosphere was added 10 ml of freshly distilled pyridine, followed by 0.76 ml (5.40 mmol) of anhydrous triethylamine dropwise and the mixture was stirred for 15 minutes at ambient temperature. To this suspension was added 1 g (1.08 mmol) of **5** followed by 66 mg (0.54 mmol) of dimethylaminopyridine and the flask was then kept under nitrogen and stirred for 12 hours. The solvent was removed *in vacuo* until an oily residue was obtained, the desired product was subsequently purified by silica gel column chromatography, using a gradient of methanol in chloroform (3%) resulting in 0.408 g (0.57 mmol) of **6**, 53% yield, as a white solid.
Rf (methanol/dichloromethane, 1/19): 0.41;
U.V.(methanol) : λ_{max} = 203, 233 (ε = 12.99 x 10 ^{3} $cm^{-1}M^{-1}$), λ_{min} = 282 nm.
^1H NMR ($DMSO_{d6}$) : δ = 0.02 (s, 6H, Si-CH_3); 0.77 (s, 9H, t-but.CH_3); 1.50 (s, 3H, base-CH_3); 2.03 - 2.23 (m, 2H, H2'2''); 3.03 - 3.19 (m, 2H, H5'5''); 3.71 (s, 6H, -OCH_3); 3.75 (m, 1H, H4'); 4.41 (m, 1H, H3'); 6.10 - 6.14 (m, 3H, H1' + NH_2); 6.83 (s, 1H, Ar-H); 6.86 - 7.38 (m, 13H, Ar-H); 8.87 (s, 1H, N^1-H); 10.01 (s, 1H, N^4-H) ppm.
HRMS calculated for $C_{38}H_{50}N_5O_7Si$ (M + H^+) : 716.3479, found : 716.3474.

N^4-(N^2-acetamidoureido)-5'-O-4,4'-dimethoxytrityl-3'-O-tertbutyldimethylsilyl-5-methyl-2'-deoxycytidine (7)

To 66 mg (0.22 mmol) of **3** evaporated three times from freshly distilled pyridine was added 4 ml of freshly distilled pyridine under a nitrogen atmosphere followed by 0.15 ml (1.08 mmol) of anhydrous triethylamine added dropwise. The mixture was stirred for

15 minutes at ambient temperature at which point 200 mg (0.22 mmol) of **5** were added followed by 13 mg (0.11 mmol) of dimethylaminopyridine. The reaction mixture was kept under a nitrogen atmosphere and the mixture stirred for 12 hours. The solvent was removed *in vacuo* until an oily residue remained. The desired product was purified by silica gel preparative thin layer chromatography, using 8% methanol in dichloromethane as eluent producing 20 mg (0.03 mmol) of **7**, 14 % yield, as a white solid.

Rf (methanol/dichloromethane, 1/9): 0.71;

U.V.(methanol) : λ_{max} = 201, 231 (ε =10.12 x 10^3 $cm^{-1}M^{-1}$), λ_{min} = 277 nm.

^1H NMR ($DMSO_{d6}$) : δ = -0.03 - 0.01 (m, 6H, Si-CH_3); 0.77 (s, 9H, t-but.CH_3); 1.51 (s, 3H, base-CH_3); 1,80 (s, 3H, acetyl-CH_3); 2.05 - 2.26 (m, 2H, H2'2''); 3.15 - 3.20 (m, 2H, H5'5''); 3.70 (s, 6H, -OCH_3); 3.75 (m, 1H, H4'); 4.42 (m, 1H, H3'); 6.12 (t, J = 6.8 Hz, 1H, H1'); 6.86 - 7.38 (m, 14H, Ar-H); 8.18 (s, 1H, N^3-H); 9.22 (s, 1H, N^1-H); 9.51 (s, 1H, N^2-H); 10.02 (s, 1H, N^4-H) ppm.

HRMS calculated for $C_{40}H_{52}N_6O_8Si$: 772.3616, found : 772.3613.

N^4-ureido-5'-O-4,4'-dimethoxytrityl-5-methyl-2'-deoxycytidine (8)

To 0.408 g (0.57 mmol) of **6** dissolved in 5 ml of anhydrous THF was added 0.74 ml of a 1M solution of tetrabutylammonium fluoride (TBAF) dropwise. The mixture was allowed to stir at ambient temperature and after 2.5 hours the reaction was complete as monitored by TLC. The solvent was removed *in vacuo* and the desired product was purified by silica gel column chromatography, using a gradient of methanol in chloroform (5%) resulting in 0.296 g (0.49 mmol) of **8**, 86% yield, as a white solid.

Rf (methanol/dichloromethane, 1/9): 0.33; mp : 98 - 100 °C ;

U.V.(methanol) : λ_{max} = 203, 230 (ε = 49.2x10^2 $cm^{-1}M^{-1}$), λ_{min} = 280 nm.

^1H NMR ($DMSO_{d6}$) : δ = 1.41 (s, 3H, base-CH_3); 2.02 - 2.18 (m, 2H, H2'2''); 3.14 - 3.15 (m, 2H, H5'5''); 3.71 (s, 6H, -OCH_3); 3.81 (m, 1H, H4'); 4.27 (m, 1H, H3'); 5.26 (d, J = 4.0 Hz, 1H, C3'-OH); 6.16 (m, 3H, H1' + NH_2); 6.82 (s, 1H, Ar-H); 6.84 - 7.38 (m, 13H, Ar-H); 8.87 (s, 1H, N^1-H); 9.98 (s, 1H, N^4-H) ppm.

HRMS calculated for $C_{32}H_{36}N_5O_7$ ($M+H^+$): 602.2615, found: 602.2617.

N^4-(N^2-acetamidoureido)-5'-O-4,4'-dimethoxytrityl-5-methyl-2'-deoxycytidine (9)

To 25 mg (0.032 mmol) of **7** dissolved in 1 ml of anhydrous THF was added 0.042 ml of a 1M solution of TBAF dropwise. The mixture was allowed to stir at room temperature and after 3.5 hours the reaction was complete as monitored by TLC. The solvent was removed *in vacuo* and the desired product was purified by silica gel preparative thin layer chromatography, using 10% methanol in dichloromethane as eluent, resulting in 11 mg (0.017 mmol) of **9**, 53% yield, as a white solid.

Rf (methanol/dichloromethane, 1/9): 0.30; mp : 128-130 °C ;
U.V.(methanol) : λ_{max} = 201, 226 (ε = 8.04x10 ^{2} $cm^{-1}M^{-1}$), λ_{min} = 282 nm.
^{1}H NMR ($DMSO_{d6}$) : δ = 1.42 (s, 3H, base-CH_3); 1.81 (s, 3H, acetyl-CH_3); 2.05 - 2.22 (m, 2H, H2'2''); 3.14 - 3.16 (m, 2H, H5'5''); 3.71 (s, 6H, -OCH_3); 3.82 (m, 1H, H4'); 4.28 (m, 1H, H3'); 5.27 (d, J = 4.0 Hz, 1H, C3'-OH); 6.17 (t, J = 7.6 Hz, 1H, H1'); 6.86 - 7.38 (m, 14H, Ar-H); 8.17 (s, 1H, N^3-H); 9.22 (s, 1H, N^1-H); 9.50 (s, 1H, N^2-H); 9.99 (s, 1H, N^4-H) ppm.
HRMS calculated for $C_{34}H_{39}N_6O_8$ ($M+H^+$) : 659.2829, found : 659.2829.

N^4-ureido-5'-O-4,4'-dimethoxytrityl-3'-O-(2-cyanoethoxy)diisopropyl aminophosphino-5-methyl-2'-deoxycytidine (10)

To 50 mg (0.083 mmol) of **8** dissolved in 1 ml of anhydrous dichloromethane was added 0.024 ml (0.075 mmol) of 2-cyanoethyl tetraisopropylphosphoramidite dropwise, followed by the addition of 3 mg (0.042 mmol) of 1H-tetrazole. The mixture was allowed to stir at room temperature for 15 hours and then concentrated *in vacuo*. The desired product was precipitated from the solution by adding cold hexane. The product was filtered and washed with additional cold hexane to give 62 mg (0.077 mmol) of **10**, yield 93%.
Rf (methanol/dichloromethane, 1/9, alumina): 0.77;
^{31}P NMR ($CDCl_3$): δ = 147.34, 147.55 ppm.
HRMS calculated for $C_{41}H_{52}N_7O_8P$: 802.3693, found: 802.3694.

DNA synthesis

The native 25-mers were prepared by solid phase DNA synthesis under standard conditions.[15] The 15-mers containing the modified nucleobases were synthesized following standard conditions until the 4th natural nucleobase was coupled and the DMT protecting group was removed. Subsequently 135 µl (0.425 mmol) of 2-cyanoethyl tetraisopropylphosphoramidite diluted in 300 ml of freshly distilled acetonitrile was delivered to the column to react with the free 5'-OH of the 4-mer over 1 hour of time. Then, 15 mg (25 µmol) of **8** or 17 mg (25 µmol) of **9**, dissolved in 0.25 ml of freshly distilled acetonitrile were delivered together a solution of sublimed 1H-tetrazole in acetonitrile as activating agent and coupled to the phosphoramidite 4 times over a 2 hour period. The solid support was then rinsed with acetonitrile, oxidized with a 0.5 M solution of (1S)-(+)-(10-camphorsulfonyl)-oxaziridine (CSO) in acetonitrile and the DMT protecting group was removed in acidic conditions. The automatic DNA synthesis procedure was then restarted, eliminating the capping step, using CSO as an oxidizing reagent and allowing 30 extra minutes for the coupling of the first base following the modified nucleobase. The terminal protecting group was left on and the oligonucleotides were deprotected in

concentrated ammonia hydroxide at room temperature during a 5 hour period. Purification of the oligonucleotides was accomplished by HPLC (trityl on) using a linear gradient 20 - 60% B over 40 minutes (A : 50 mM triethylammonium acetate, adjusted to pH = 7.0 with glacial CH_3COOH, B : 50 mM triethylammonium acetate in 70% acetonitrile). The DMT protected 15-mer oligonucleotides had a retention time about 25 minutes, the collected oligonucleotides were reduced in volume, detrytilated with 80% aqueous acetic acid (45 min, 0 °C), desalted (Sephadex G-10), and stored at -20 °C. The resulting oligonucleotides were analyzed by HPLC (4.6 x 250 mm column) using a linear gradient 0 - 100% B over 60 minutes (A : 20 mM KH_2PO_4, adjusted to pH = 5.5 with KOH 2M, B : 20 mM KH_2PO_4 in 70% methanol). The 15-mer oligonucleotides had retention times of about 25 minutes and eluted as single peaks. The unnatural nucleosides could be incorporated into the DNA strands with coupling efficiencies largely sufficient to conduct analyses of characterization and triple helices formation studies.

MALDI-TOF MS for 5'TTTMTTTTMTXTMTT3' (for X, see Figure 1) : calculated for (15-mer^+) : 4556, found : 4557; calculated for ($15\text{-mer} + Na^+$) : 4580, found : 4579.

MALDI-TOF MS for 5'TTTMTTTTMTYTMTT3' (for Y, see Figure 1) : calculated for (15-mer^+) : 4613, found : 4614; calculated for ($15\text{-mer} + Na^+$) : 4636, found : 4635.

Nucleoside Analysis

Oligomers were effectively digested into monomeric units with a combination of snake venom phosphodiesterase and alkaline phosphatase: a 20µl reaction mixture containing 0.5 A_{260} unit of oligomer, 50 mM tris·HCl, pH 8.0, 100 mM $MgCl_2$, 1 unit of snake venom phosphodiesterase, 1 unit of alkaline phosphatase was incubated overnight at 37 °C. An aliquot of this mixture was analyzed by HPLC (4.6 x 250 mm column) using a linear gradient 0 - 100% B over 60 minutes (A : 20 mM KH_2PO_4, adjusted to pH = 5.5 with KOH 2M, B : 20 mM KH_2PO_4 in 70% methanol). Nucleosides analogs X and Y eluted with a retention time of 10.6 (see Figure 3b) and 7.8 min respectively.

Thermal Denaturation Studies

Thermal Denaturation Studies were performed in 25 mM PIPES pH 6.4 or 7.0, 25 mM HEPES pH 7.5 and 10 mM magnesium chloride, 50 mM sodium chloride at strands ratios of 1:1:1 at a concentration of 1 µM. Absorbance and temperature values were measured with an AVIV 14DS UV/Visible spectrophotometer equipped with digital temperature control. The temperature of the cell compartment was increased in 1.0 °C steps (from 0 to 85 °C) and when thermal equilibrium was reached, temperature and absorbance data were collected. T_m values were determined from first-order derivatives of the Absorbance vs Temperature plots (see Figure 4) or graphically from the same plots.

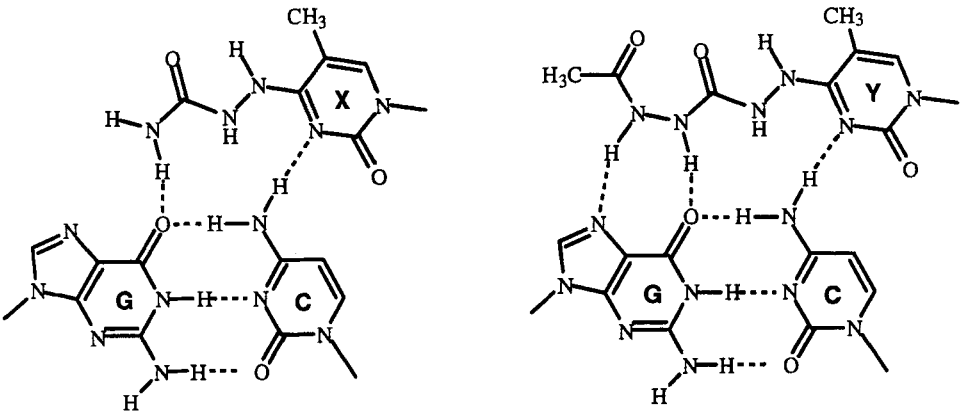
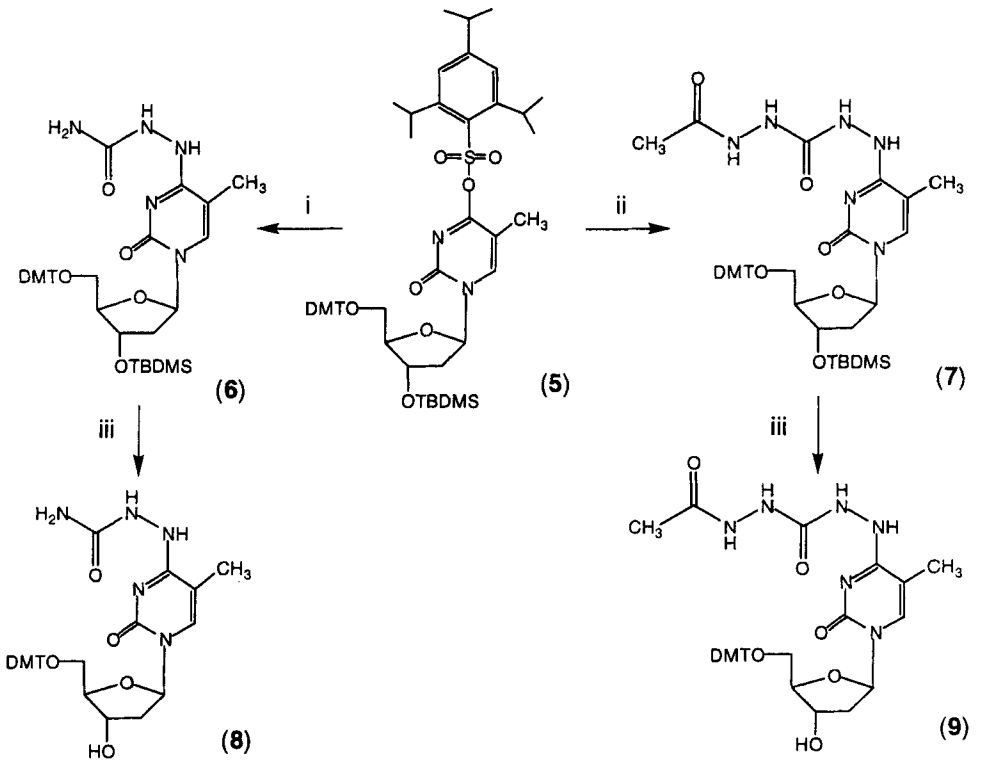


Figure 1. X-C-G and Y-C-G base triplets where X and Y contain extended functionality based upon a semicarbazide or carbohydrazide side chain.

RESULTS AND DISCUSSION

Recognition of DNA duplexes by a third DNA strand results in the formation of a DNA triplex. In the most common form of the DNA triplex the third strand makes a series of bidentate contacts with the purine residues of one of the duplex strands in the DNA target. Owing to this general requirement, DNA triplex formation occurs most effectively at homopurine sequences. One of the problems in recognizing mixed purine/pyrimidine sequences on the target strand is the difficulty in making a stable base-pyrimidine-purine base triplet. While in the purine-pyrimdine target, the Hoogsteen face of the purine residue provides two functional groups for hydrogen bonding interactions with the third base residue, the complementary pyrimidine-purine target requires that the third base bind to the pyrimidine. Since pyrimidines lack a Hoogsteen face, the only functional group available for hydrogen bonding interactions is the amino (dC) or carbonyl (dT) at the C4 position, both already involved in Watson-Crick hydrogen bonding. In order to extend the sequence recognition capabilities of the third strand in targeting DNA duplexes, we designed a pyrimidine base residue with extended functionality. In the ideal case, the N^3-nitrogen of the pyrimidine would hydrogen bond to the amino group at the N^4-position of the target dC residue, and the added functionality would make one or more hydrogen bonds with functional groups present in the complementary purine. Two such analogues are illustrated in Figure 1.



(i) semicarbazide hydrochloride, TEA, DMAP; (ii) acetylhydrocarbohydrazide *p*-tolunesulfonate, TEA, DMAP; (iii) TBAF 1M in THF.

Scheme 1

Synthesis of the Nucleoside

The two nucleosides were prepared from a common intermediate derived from sulfonation[14,20] of the O^4-carbonyl of dT (Scheme 1). The sulfonyl derivative undergoes addition/elimination reaction with nucleophiles, and the semicarbazide reaction proceeded without difficulty. The corresponding carbohydrazide reaction was complicated by unknown side reactions - possibly related to the terminal hydrazino nitrogen. Certainly the parent compound hydrazine, reacts with the 5,6-double bond of pyrimidines since this reaction is the basis for the Maxam-Gilbert sequencing technique. To limit such possibilities, we prepared it as the mono-acetyl derivative, and with this compound, the subsequent reaction with the sulfonylated nucleoside proceeded without incident.

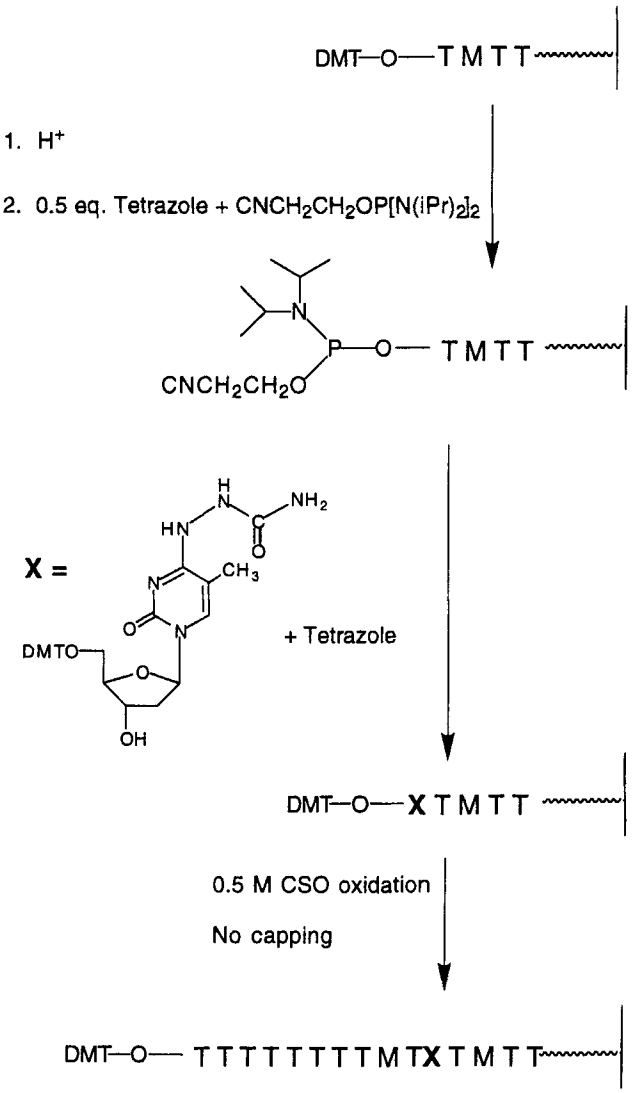
In both analogues we were concerned about side reactions involving the nitrogen attached to the C4-carbon, but found this site to be of low reactivity such that we could not effectively protect it. Nevertheless, analysis of the stability of both derivatives under the conditions of DNA assembly and deprotection indicated that neither analogue withstood the conditions of capping or oxidation. In the former case it is likely that this same nitrogen undergoes partial acetylation, in the later case we expect that oxidation to the diazo derivative is the likely side reaction. We could eliminate both potential problems by (i) eliminating the capping step from the synthesis, and (ii) replacing the I_2-based oxidizing solution. Although other oxidizing agents[16-18] will work for the P(III) to P(V) oxidation, in our hands most resulted in the same oxidative side reaction. Only(1S)-(+)-(10-camphorsulfonyl)-oxaziridine (CSO)[19] appeared to facilitate this oxidation without formation of the putative diazo compound.

Unfortunately, in addition to the noted problems with the capping and oxidation steps, the coupling reaction for the X analogue (as phosphoramidite derivatives) occurred with only minimal success for reasons that are not clear. This low yield was confirmed by the preparation and analysis of small dimers and trimers using the X phosphoramidite. To avoid further complications, we did not prepare the phosphoramidite derivative of Y, and introduced the analogues by a procedure that is essentially a reversed coupling protocol (Scheme 2).

In this procedure, at the sequence position of the analogue, the DMT protecting group was removed and the sequence phosphitylated to generate the phosphoramidite derivative on the 5'-terminus of the support-bound oligonucleotide. The DMT-protected analogue (nucleoside X or Y with a free 3'-hydroxyl) was then added in the presence of tetrazole. After oxidation (CSO) and removal of the DMT group (acid), the DNA synthesis was continued in the absence of the capping step.

The need for anhydrous solvents for this procedure is noteworthy. In a normal phoshoramidite coupling procedure, the active phosphoramidite species is the reagent in solution, added in excess. Trace quantities of water reduce the excess of the incoming phosphoramidite but do not necessarily result in a failed coupling. In the reversed coupling procedure the active phosphoramidite species is the support bound oligonucleotide, and is the limiting reagent. Trace quantities of water can drastically reduce the coupling yield by reacting with the phosphoramidite.

Since the reversed coupling procedure does not proceed with the high yields characteristic of phosphoramidite couplings, and the capping step is eliminated after the introduction of either X or Y, it is to be expected that some elongation will occur of sequences that failed to take part in the initial phosphitylation reaction. It then becomes necessary to resolve such products after completion of the assembly and deprotection procedures.



Scheme 2

The effectiveness of the reversed coupling protocol could be estimated from HPLC traces during the isolation procedure (Figure 2). The major peak (A, Figure 2) represents the desired 15-mer product with the terminal DMT-group. Minor products B and C are likely to be failed sequences that resulted from a less than quantitative reversed coupling procedure and the absence of the capping step. They would also still carry a terminal DMT group and elute relatively late from the column. Earlier eluting peaks represent failed sequences (those capped prior to the reversed coupling procedure and those that may have

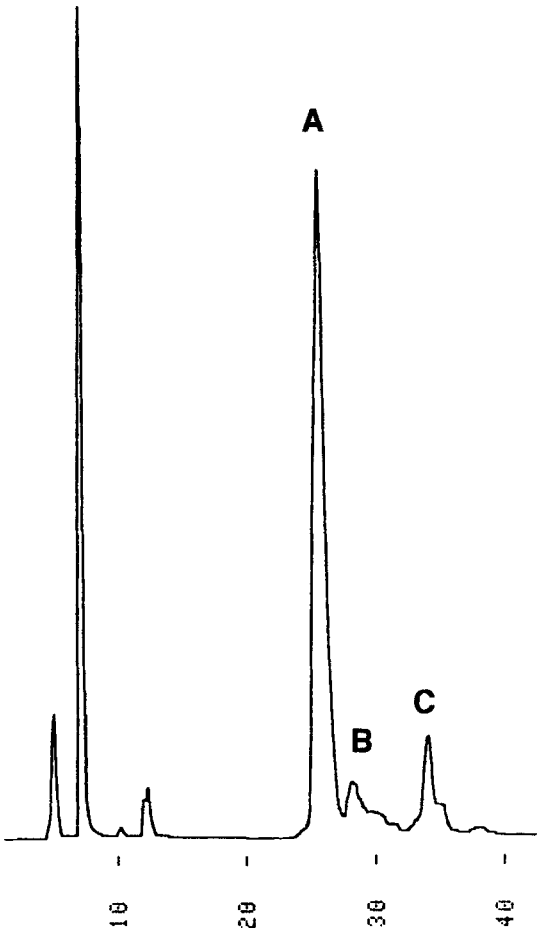


Figure 2. HPLC isolation (reversed phase C-8 column) of the 15-mer 5'TTTMTTTTMTXTMTT with DMT at the 5'-terminus.

been phosphitylated but failed to react with the analogue nucleoside) and protecting groups. The analysis illustrated suggests that the reversed coupling of the X nucleoside occurred with a minimum yield > 50% less than that obtained by traditional methods, but sufficient to generate the material needed.

After removal of the terminal DMT-group from the isolated material, HPLC analysis indicated the presence of a single peak (Figure 3a). Digestion of a small amount of material yielded a major peak (T) a secondary peak (m^5C = M) and one additional small peak (Figure 3b). The same analysis run in the presence of an authentic standard for the analogue X, indicated that the unmodified analogue nucleoside had been incorporated into the sequence (Figure 3c).

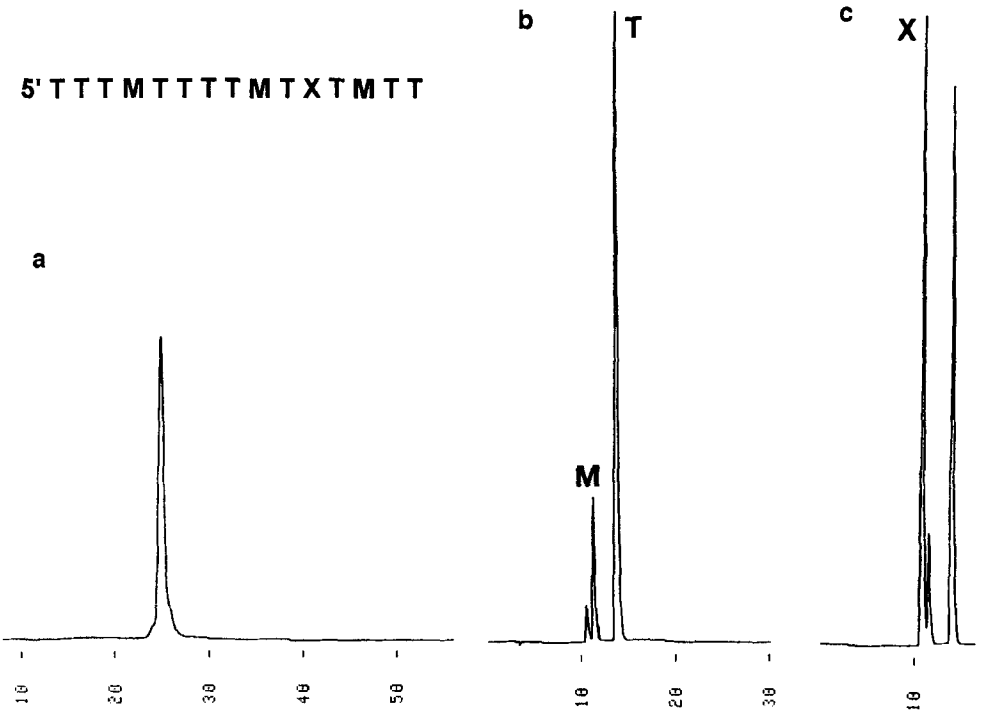


Figure 3. (a) HPLC analysis of the purified 15-mer containing the X-nucleoside. (b) Snake venom phosphodiesterase/alkaline phosphatase digestion of a small quantity of the X-containing 15-mer and its resolution on a reversed phase C-8 column. (c) Analysis as in (b) with co-injection of an authentic sample of the X-nucleoside.

Similar procedures resulted in the preparation of sequences containing the carhohydrazide-based Y analogue.

MALDI-TOF mass spectroscopy analysis conducted on the 15-mers, confirmed as well the presence of analogues (X or Y) along the strands.

DNA Triplex Studies

We prepared 15-mer third strands for the targeting of a 15-residue homopurine sequence embedded within a 25-mer duplex. At one site a single base pair target was altered such that four different sequences resulted; each of the four possible base pairs were available for testing as targets for the analogue residues. T_m values were obtained from thermal denaturation studies monitored at 260 nm. All samples exhibited two cooperative transitions, an early one whose position varied depending upon the triplex formed, and a

later transition that occurred consistently near 70 °C (see Figure 4). The early T_m was interpreted as reflecting the triplex to duplex transition while the latter reflected the duplex to random coil transition. Hyperchromicity values for the duplex to triplex transitions varied and were more significant for the more stable structures at lower pH values (compare Figures 4a and 4b). This difference suggests that the less stable triplexes may not be fully formed - consistent with disruptions in hydrogen bonding at the site of the analogue residue and/or loss of protonation effects at higher pH values. However, even those transitions occurring at low temperatures were characterized by cooperative changes in the absorbance vs. temperature plots (e.g., Figure 4b).

The results of these experiments are summarized in Table 1. Both analogue bases form the most stable triplexes with target duplexes containing a C-G base pair as designed (see Figure 1), suggesting some ability to discriminate between the four possible target sequences. The complex containing the X-C-G base triplet exhibits a T_m of 26 °C at pH 6.4 and 19 °C at pH 7.0. These values are 4 - 7 °C higher than the values obtained with G-C, T-A or A-T base pairs at the target site (see Table 1).

In all cases examined, the extension of the multifunctional linker (compare X with Y in Figure 1) did not significantly contribute to the triplex forming properties of the sequences containing the analogue nucleosides.

The control sequence containing m^5C in place of X (or Y) exhibits essentially the same T_m value with the C-G target as does either analogue sequence. This observation suggests that additional hydrogen bonding interactions afforded by the analogue do not appear to enhance the stability of the complex beyond that present with the simple m^5C sequence. The sequence discrimination observed for these derivatives occurs more as a result of complex destabilization at non-cognate sequences rather than enhanced stabilization at the target site.

CONCLUSIONS

Nucleoside analogs containing added functionality at the C4-position of dC can be prepared quite easily from the native nucleoside and simple derivatives such as semicarbazide and carbohydrazide. These multifunctional derivatives cannot be easily converted to phosphoramidites and used in standard DNA assembly protocols, at least not without further attention to protection/deprotection options. However, the use of a reversed coupling protocol was effective in introducing the desired analogues into DNA sequences. The triplexes examined that contained either the X or Y nucleoside both exhibited the most favorable Tm's with target C-G base pairs. Additionally, they significantly destabilized triplexes containing a target G-C base pair (where the unmodified

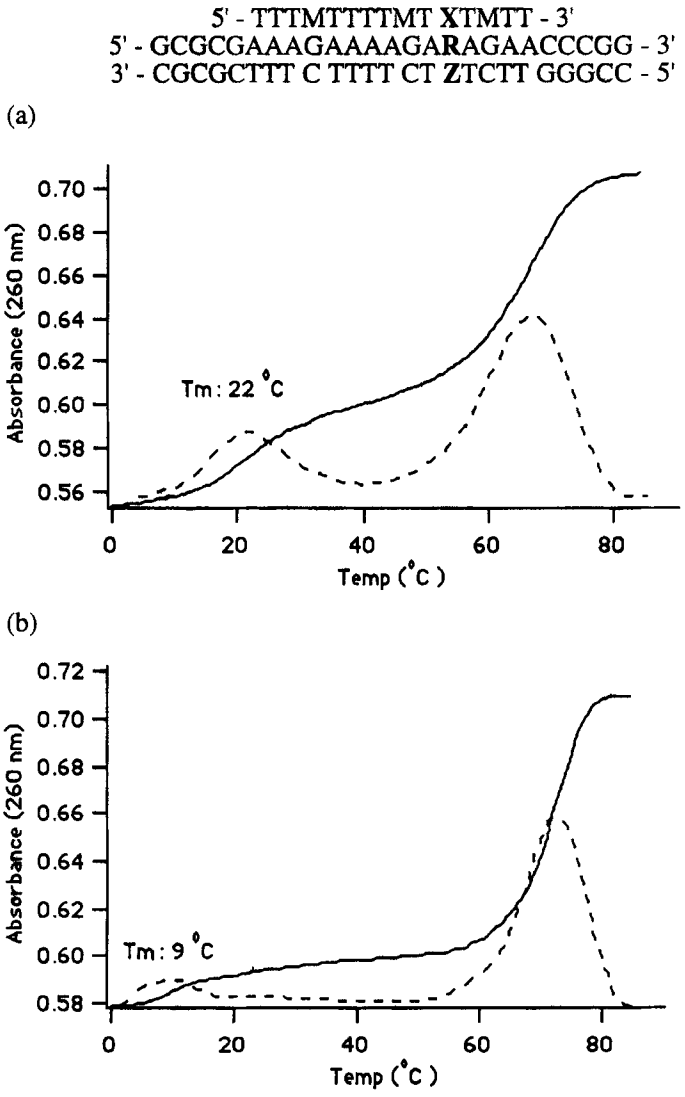


Figure 4. Two absorbance vs. temperature plots (solid lines) and the calculated first derivative plots (dotted lines). X = nucleoside containing semicarbazide functionality, M = m^5C, (a) R = A, Z = T, pH = 6.4; (b) R = G, Z = C, pH = 7.5.

Table 1. Tm values for Analogue Triplexes*

5'- TTTMTTTTMT**N**TMTT - 3'
5' - GCGCGAAAGAAAAGA**R**AGAACCCGG - 3'
3' - CGCG CTTTCTTTT C T**Z**TCTTGGG CC - 5'

R-Z = C-G

Analogue, N =	pH 6.4	pH 7.0	pH 7.5
X	26	19	14
Y	26	19	14
M	27	22	14

R-Z = G-C

Analogue, N =	pH 6.4	pH 7.0	pH 7.5
X	19	12	9
Y	19	10	8
M	49	38	32

R-Z = T-A

Analogue, N =	pH 6.4	pH 7.0	pH 7.5
X	21	12	7
Y	21	14	7
M	23	17	8

R-Z = A-T

Analogue, N =	pH 6.4	pH 7.0	pH 7.5
X	22	16	13
Y	27	19	12
M	26	20	8

*M = m^5C, X = nucleoside containing semicarbazide functionality, Y = nucleoside containing acetylcarbohydrazide functionality.

m^5C is most effective). Such selectivity is an important criteria for effective DNA triplex targeting, but the observation that neither X nor Y were more effective than m^5C in targeting the C-G base pair suggests that the goal of using multifunctional derivatives to create additional hydrogen bonding interactions with the target duplex was not reached with these two derivatives.

ACKNOWLEDGMENTS

This work was supported by a grant from the NIH (GM53201).

REFERENCES

(1) Cooney, M.; Czernuszewicz, G.; Postel, E. H.; Flint, S. J.; Hogan, M. E. *Science* **1988**, *241*, 456-9.
(2) Moser, H. E.; Dervan, P. B. *Science* **1987**, *238*, 645-650.
(3) LeDoan, T.; Perrouault, L.; Praseuth, D.; Habhoub, N.; Decout, J. L.; Thoung, N. T.; Lhomme, J.; Helene, C. *Nucleic Acids Res.* **1987**, *15*, 7749-7760.
(4) Strobel, S. A.; Doucette-Stamm, L. A.; Riba, L.; Housman, D. E.; Dervan, P. B. *Science* **1991**, *254*, 1639-42.
(5) Miller, P. S.; Cushman, C. D. *Biochemistry* **1993**, *32*, 2999-3004.
(6) Lehman, T. E.; Greenberg, W. A.; Liberles, D. A.; Wada, C. K.; Dervan, P. B. *Helvetica Chimica Acta* **1997**, *80*, 2002-2022.
(7) Chin-Yi, H.; Miller, P. S. *J. Am. Chem. Soc.* **1993**, *115*, 10456-10457.
(8) Prakash, T. P.; Barawkar, D. A.; Kumar, V.; Ganesh, K. N. *Bioorg. & Med. Chem. Letters* **1994**, *4*, 1733-1738.
(9) Hampel, K. J.; Crosson, P.; Lee, J. S. *Biochemistry* **1991**, *30*, 4455-4459.
(10) Thomas, T.; Thomas, T. J. *Biochemistry* **1993**, *32*, 14069-14074.
(11) Tung, C. H.; Breslauer, K. J.; Stein, S. *Nucleic Acids Res.* **1993**, *21*, 5489-5494.
(12) Barawkar, D. A.; Rajeev, K. G.; Kumar, V. A.; Ganesh, K. N. *Nucleic Acids Research* **1996**, *24*, 1229-1237.
(13) Guzzo-Pernell, N.; Tregear, G. W.; Haralambidis, J.; Lawlor, J. M. *Nucleosides and Nucleotides* **1998**, *17*, 1191-1207.
(14) Gildea, B.; McLaughlin, L. W. *Nucl. Acids Res.* **1989**, *17*, 2261-2281.
(15) Matteucci, M. D.; Caruthers, M.H. *J. Am. Chem. Soc.* **1981**, *103*, 3185-3191.
(16) Sproat, B.; al., e. *Nucleosides and Nucleotides* **1995**, *14*, 255-273.
(17) Hayakawa, Y.; Uchiyama, M.; Noyori, R. *Tetrahedron Lett.* **1986**, *27*, 4191-4194.
(18) Tanaka, T.; Letsinger, R. L. *Nucl. Acids Res.* **1982**, *10*, 3249-3260.
(19) Ugi, I.; al., e. *Nucleosides and Nucleotides* **1988**, *7*, 605-608.
(20) Gaffney, B. L.; Marky, L. A.; Jones, R. A. *Tetrahedron* **1983**, *40*, 3-13.

Received 10/8/98
Accepted 2/26/99