Use of pK_a Differences To Enhance the Formation of Base Triplets Involving C−G and G−C Base Pairs

Dong Li Chen and Larry W. McLaughlin*

Department of Chemistry, Merkert Chemistry Center, 2609 Beacon Street, Boston College, Chestnut Hill, Massachusetts 02467

larry.mclaughlin@bc.edu

Received May 17, 2000

Two C-nucleosides are employed for the recognition of dC-dG base pairs. Both derivatives are related to dC but lack the O2-carbonyl. The absence of the carbonyl should eliminate any unfavorable steric interactions at this site. One of the derivatives contains a 2-aminopyridine heterocycle (d2APy) while the second contains a 2-aminopyrimidine heterocycle (d2APm). The former with a pK_a of 6.8 functions better for the recognition of dG-dC base pairs than it does in the binding to dC-dG base pairs. The d2APm derivative with a pK_a of 3.3 functions better to form base triplets with dC-dG base pairs than with dG-dC targets. Triplex T_m's in both cases are compared with the sequence containing the native dC residue. The dC analogues appear to make two hydrogen bonds to a target dG base residue, one of which requires protonation of the ring nitrogen. Recognition of a target dC residue appears to require the formation of a single hydrogen bond to the C-nucleoside and having that nitrogen largely in the unprotonated state facilitates its formation.

Introduction

The recognition of double-stranded DNA using a third single-stranded pyrimidine-rich oligonucleotide to form DNA triplexes occurs most effectively when the target sequence is composed largely or solely of purine residues. The interruption of a polypurine target sequence by one or more pyrimidine bases usually results in significant losses in triplex stability[1−6] presumably in part due to the inability of the Hoogsteen face of the pyrimidines to form one or more hydrogen bonding interactions with the third strand. The pyrimidine can be bypassed with an abasic residue, but such complexes are also reduced in stability.[7] A number of native and analogue nucleosides have been examined for their ability to interact with either dT-dA or dC-dG base pair targets. In the simplest case, dG can interact with a dT-dA base pair[1,8] and form a single interstrand hydrogen bond.[9] All of the possible base triplets involving the common nucleosides have been examined[1−6,10−16] but only the dG-dT-dA (parallel)[16] and dT-dC-dG (antiparallel)[17] base triplets exhibit moderately stabilizing effects. In the latter case, a single hydrogen bond appears to form between the O2-carbonyl of the dT in the third strand and the N4-amino group of the target dC. This concept has been improved upon with the related 5-methyl-pyrimidin-2-one, which also appears capable of forming a single hydrogen bond involving the N4-amino group.[18]

A number of analogue residues in the third strand have also been examined for their ability to target pyrimidine-purine base pairs.[19−25] The most common approach here is to introduce added functionality such that hydrogen-bonding interactions could occur with the target pyrimidine as well as its partner purine and thus recognize the base pair as a unit.[26,27] Modeling studies with native nucleosides have indicated that when the dC residue of a dC-dG base pair is located in the purine-rich target

* To whom correspondence should be addressed. Tel: (617) 552 3622. Fax: (617) 552 2705.

(1) Griffin, L. C.; Dervan, P. B. *Science* **1989**, *245*, 967−971.
(2) Mergny, J.; Sun, J.; Rougee, M.; Montenay-Garestier, T.; Barcelo, F.; Chomilier, J.; Helene, C. *Biochemistry* **1991**, *30*, 9791−9798.
(3) Macaya, R. F.; Gilbert, D. E.; Malek, S.; Sinsheimer, J. S.; Feigon, J. *Science* **1991**, *254*, 270−274.
(4) Radhakrishnan, I.; Patel, D. J. *J. Am. Chem. Soc.* **1992**, *114*, 6913−6915.
(5) Yoon, K.; Hobbs, C. A.; Koch, J.; Sardaro, M.; Kutny, R.; Weis, A. L. *Proc. Natl. Acad. Sci. U.S.A.* **1992**, *89*, 3840−3844.
(6) Miller, P. S.; Cushman, C. D. *Biochemistry* **1993**, *32*, 2999−3004.
(7) Horne, D. A.; Dervan, P. B. *Nucleic Acids Res.* **1991**, *19*, 4963−4965.
(8) Radhakrishnan, I.; Gao, X.; de los Santos, C.; Live, D.; Patel, D. J. *Biochemistry* **1991**, *30*, 9022−9030.
(9) Wang, E.; Malek, S.; Feigon, J. *Biochemistry* **1992**, *31*, 4838−4846.
(10) Belotserkovskii, B. P.; Veselkov, A. G.; Filippov, S. A.; Dobrynin, V. N.; Mirkin, S. M.; Frankkamenetskii, M. D. *Nucleic Acids Res.* **1990**, *18*, 6621−6624.
(11) Pei, D.; Ulrich, H. D.; Schultz, P. G. *Science* **1991**, *253*, 1408−1411.
(12) Roberts, R. W.; Crothers, D. M. *Science* **1992**, *258*, 1463−1466.
(13) Rougee, M.; Faucon, B.; Mergny, J. L.; Barcelo, F.; Giovannangeli, C.; Garestier, T.; Helene, C. *Biochemistry* **1992**, *31*, 9269−9278.
(14) Fossella, J. A.; Kim, Y. J.; Shih, H.; Richards, E. G.; Fresco, J. R. *Nucleic Acids Res.* **1993**, *21*, 4511−4515.
(15) Chandler, S. P.; Fox, K. R. *Febs Lett.* **1993**, *332*, 189−192.
(16) Best, G. C.; Dervan, P. B. *J. Am. Chem. Soc.* **1995**, *117*, 1187−1193.
(17) Dittrich, K.; Gu, J.; Tinder, R.; Hogan, M.; Gao, X. *Biochemistry* **1994**, *33*, 4111−4120.
(18) Prevot-Halter, I.; Leumann, C. J. *Bioorg. Med. Chem. Lett.* **1999**, *9*, 2657−2660.
(19) Stilz, H. U.; Dervan, P. B. *Biochemistry* **1993**, *32*, 2177−2185.
(20) Koshlap, K. M.; Gillepsie, P.; Dervan, P. B.; Feigon, J. *J. Am. Chem. Soc.* **1993**, *115*, 7908−7909.
(21) Huang, C.; Cushman, C. D.; Miller, P. S. *J. Org. Chem.* **1993**, *58*, 5048−5049.
(22) Huang, C.; Miller, P. S. *J. Am. Chem. Soc.* **1993**, *115*, 10456−10457.
(23) Griffin, L. C.; Kiessling, L. L.; Beal, P.; Gillespie, P.; Dervan, P. B. *J. Am. Chem. Soc.* **1992**, *114*, 7976−7909.
(24) Lehman, T. E.; Greenberg, W. A.; Liberles, D. A.; Wada, C. K.; Dervan, P. B. *Helv. Chim. Acta* **1997**, *80*, 2002−2022.
(25) Orson, F. M.; Klysik, J.; Bergstrom, D. E.; Ward, B.; Glass, G. A.; Hua, P.; Kinsey, B. M. *Nucl. Acids Res.* **1999**, *27*, 810−816.
(26) Guzzo-Pernell, N.; Tregear, G. W.; Haralambidis, J.; Lawlor, J. M. *Nucleosides Nucleotides* **1998**, *17*, 1191−1207.
(27) Gianolio, D. A.; McLaughlin, L. W. *Nucleosides Nucleotides* **1999**, *18*, 1751−1769.

10.1021/jo000754w CCC: $19.00
Published on Web 09/15/2000

strand a pyrimidine is preferred in the third strand on the basis of minimized steric contacts.[2,28] A dC residue tethering a naphthalene diimide intercalator has been shown to be selective for dC-dG base pair targets.[29]

Although a target dC residue offers only a single Hoogsteen hydrogen bonding site (the N4-amino group), it appeared that a dC residue in the third strand could be positioned to form a base triplet essentially isomorphic with dT-dA-dT or dC+-dG-dC. In this orientation, its N3-nitrogen could function as a hydrogen bond acceptor. Its effectiveness in this role will depend in part upon its pK_a value. Steric constraints, possibly involving the O2-carbonyl might reduce the stability of the simple C−C-G base triplet, consistent with reported observations.[1−6,10−16] We therefore decided to examine the use of related analogue nucleosides, those lacking the O2-carbonyl, as possible triplex forming residues for C−G target base pairs. We report here on the synthesis of a new pyrimidine C-nucleosides and triplex forming abilities of this derivative and a related pyridine C-nucleoside that exhibit widely varying pK_a values.

Experimental Section

Materials. ^{1}H NMR and ^{13}C NMR used TMS as the internal standard; ^{31}P NMR used phosphoric acid as the external standard. Flash chromatography was performed using silica gel 60 with particle size 35−75 μm (EM Science, Germany). Thin-layer chromatography was performed on Silica Gel 60 F254 precoated TLC aluminum sheet (EM Science, Germany). Visualization by UV (254 nm) lamp and/or by treating with a solution of 10% sulfuric acid, followed by heating. All reactions were carried out under nitrogen, using purified and distilled solvents. Pyridine, acetonitrile and THF were dried with CaH_2 and distilled. Other dry solvents were purchased from Aldrich Chemical Co and used without further purification. The four common 2′-deoxynucleotide phosphoramidites and 3′ terminal nucleoside-bound controlled pore glass (CPG) supports were purchased from Glen Research (Sterling, VA).

Methods. 2-Amino-5-iodopyrimidine (2). To a solution of 2-aminopyrimidine (**1**) (19 g, 0.20 mol) in acetic acid (200 mL), sulfuric acid (2.5 mL), and water (30 mL) were added iodine (21.7 g, 0.086 mol) and periodic acid dihydrate (6.5 g, 0.029 mol). The mixture was stirred at 90 °C for 24 h and poured into 10% $Na_2S_2O_3$ aqueous solution to remove unreacted iodine, and then it was extracted with methylene chloride. Removing methylene chloride gave solid crude compound. Recrystallization from water gave pure compound **2** (39 g, 90% yield): R_f = 0.26 (CH_2Cl_2/CH_3OH 95/5) UV−vis λ_{max}(CH_3OH) 241 (ϵ = 26950 l/Mcm), 317 (ϵ = 3090 l/Mcm) nm; IR (KBr) 3320 (sb), 3180 (s), 1636 (s), 1569 (s), 1541 (s), 1491 (s), 797 (s), 568 (sb), 517 (s), 478 (s) cm^{-1}; ^{1}H NMR ($CDCl_3$) δ 8.41 (s, 2H), 5.06 (sb, 2H) ppm; ^{13}C NMR ($CDCl_3$) δ 162.3, 161.5, 75.9 ppm; mp 179−180 °C; HRMS calcd for $C_4H_4IN_3$ (M + 1) 221.9528, found 221.9527.

5-(β-D-Glyceropentofuran-3′-ulos-1′-yl)-2-aminopyrimidine (5). A mixture of bis(dibenzylideneacetone)palladium (288 mg, 0.5 mmol) and tris(pentaflorophenyl)phosphine (532 mg, 1.0 mmol) in acetonitrile (250 mL) was stirred under nitrogen at room temperature for 1 h. *N*,*N*-Diisopropylethylamine (8 mL, 45 mmol), 1,4-anhydro-2-deoxy-3-*O*-[(1,1-dimethylethyl)diphenylsilyl]-D-erythro-1-enitol (**3**)[1] (3.55 g, 10 mmol), and **2** (2.22 g, 10 mmol) were then added, and the mixture was refluxed under nitrogen for 66 h. After the volatiles were removed in vacuo, the residue was purified by flash chromatography on silica gel using a solvent gradient (CH_2Cl_2/CH_3OH from 27:1 to 10:1) to give compound **4** (4.1 g, 92% yield) as a white solid slightly contaminated by trace amounts of **5** and *N*,*N*-diisopropylethylamine. This material was sufficiently pure for the next transformation.

Characterization of crude **4**: R_f = 0.33 (CH_2Cl_2/CH_3OH 9:1); ^{1}H NMR ($CDCl_3$) δ = 7.88 (s, 2H), 7.83 (d, 2H, J = 6.8 Hz), 7.68 (d, 2H, J = 6.4 Hz), 7.44 (t, 2H, J = 7.2 Hz), 7.38−7.32 (m, 4H), 5.61 (sb, 3H), 5.34 (d, 1H, J = 3.2 Hz), 4.64−4.62 (m, 1H), 4.07−4.05 (m, 1H), 3.90−3.82 (m, 2H) ppm.

To a solution of compound **4** (4.1 g, 9.2 mmol) in THF (15 mL) at 0 °C was added acetic acid (2.5 mL, 46 mmol) followed by 11 mL of a 1 M solution of tetra-*n*-butylammonium fluoride in THF. The desilylation reaction was completed in 1 h, based on TLC analysis. The volatiles were removed in vacuo, and the residue was separated by flash chromatography on silica gel using a solvent gradient (CH_2Cl_2/CH_3OH from 30:1 to 1:1) to yield compound **5** slightly contaminated by N,N-diisopropylethylamine. Crystallization from dichloromethane/methanol (∼500/1) gave crystals of **5** (1.54 g, yield over two steps 74%).

Characterization of **5**: R_f = 0.47 (CH_2Cl_2/CH_3OH 4:1); UV−vis λ_{max} (CH_3OH) 232 (ϵ = 19900 l/Mcm), 301 (ϵ = 2880 l/Mcm) nm; IR (KBr) 3393 (s), 3208 (sb), 3024 (m), 2967 (m), 2934 (m), 2890 (m), 2839 (m), 1754 (s), 1647 (s), 1608 (s), 1558 (s), 1496 (s) cm^{-1}; ^{1}H NMR (D_2O) δ 8.28 (s, 2H), 5.10 (dd, 1H, J_1 = 5.6 Hz, J_2 = 19.8 Hz), 4.05 (t, 1H, J = 2.8 Hz), 3.74 (dd, 1H, J_1 = 2.8 Hz, J_2 = 18.4 Hz), 3.71 (dd, 1H, J_1 = 2.8 Hz, J_2 = 18.4 Hz), 2.81 (dd, 1H, J_1 = 5.6 Hz, J_2 = 18.4 Hz), 2.51 (dd, 1H, J_1 = 10.8 Hz, J_2 = 18.4 Hz) ppm; ^{13}C NMR (DMSO-d_6) δ 214.4, 163.6, 157.2, 121.7, 82.7, 73.2, 60.5, 44.1 ppm; mp 148−149 °C; HRMS calcd for $C_9H_{11}N_3O_3$ (M + 1) 210.0879, found 210.0880.

2-Amino-5-(2-deoxy-β-D-ribofuranosyl)pyrimidine (6). To a solution of compound **5** (1.4 g, 6.7 mmol) in acetonitrile (50 mL) and acetic acid (60 mL) at 0 °C was added sodium triacetoxyborohydride (4 g, 18.9 mmol). The reaction was complete within 1 h, realized by TLC analysis. Volatiles were removed in vacuo. Sodium carbonate (3 g) was added into the solution of residue in methanol/dichloromethane (1:1, 20 mL). The mixture was stirred at ambient temperature for 1 h, and then the solvent was removed in vacuo. The solid residue was isolated with flash chromatography on silica gel (CH_2Cl_2/CH_3OH from 10:1 to 4:1) to give compound **6** (1.3 g, yield 90%). Compound **6** was recrystallized from mixing solvent (CH_3OH/CH_2Cl_2 1:400): R_f = 0.28 (CH_2Cl_2/CH_3OH 4:1); UV−vis λ_{max} (CH_3OH) 230 (ϵ = 32600 l/Mcm), 297 (ϵ = 5470 l/Mcm) nm; IR (KBr) 3354 (s), 3169 (s), 3024 (s), 2979 (w), 2957 (w), 2929 (w), 2912 (w), 2901 (w), 2856 (w), 1659 (s), 1569 (m), 1508 (s) cm^{-1}; ^{1}H NMR (CD_3OD) δ 8.30 (s, 1H), 4.98 (dd, 1H, J_1 = 5.2 Hz, J_2 = 10.4 Hz), 4.34−4.33 (m, 1H), 3.91−3.88 (m, 1H), 3.66 (dd, 1H, J_1 = 4.4 Hz, J_2 = 12.4 Hz), 3.63 (dd, 1H, J_1 = 5.2 Hz, J_2 = 12.4 Hz), 2.14 (ddd, 1H, J_1 = 1.6 Hz, J_2 = 5.2 Hz, J_3 = 13.2 Hz), 1.99 (ddd, 1H, J_1 = 6.0 Hz, J_2 = 10.4 Hz, J_3 = 13.2 Hz) ppm; ^{13}C NMR (CD_3OD) δ 167.95, 161.85, 128.53, 92.81, 81.16, 78.00, 67.47, 47.59 ppm; mp 138−139 °C; HRMS calcd for $C_9H_{13}N_3O_3$ (M + 1) 212.1035, found 212.1036.

2-[*N*-[1-(Dimethylamino)ethylidine]amino]-5-(2′-deoxy-β-D-ribofuranosyl)pyrimidine (7). To a solution of compound **6** (1.06 g, 5 mmol) in methanol (200 mL) was added *N*,*N*-dimethylacetamide dimethyl acetal (2.0 g, 15 mmol), and the reaction mixture was stirred at 78 °C for 24 h. Volatiles were removed in vacuo, and the resulting residue was separated by column chromatography on silica gel (CH_2Cl_2/CH_3OH 20:1) to afford compound **7** (1.3 g, 94% yield), a highly hygroscopic solid: R_f = 0.40 (CH_2Cl_2/CH_3OH 4:1); UV−vis λ_{max} (CH_3OH) 206 (ϵ = 7490 l/Mcm), 240 (ϵ = 6990 l/Mcm), 276 (ϵ = 14240 l/Mcm) nm; IR (KBr) 3364 (sb), 3011 (w), 2923 (s), 2873 (m), 1620 (s), 1590 (s), 1532 (s), 1400 (s), 1104 (s), 1060 (s), 1023 (s), 947 (m), 866 (m), 828 (m) cm^{-1}; ^{1}H NMR ($CDCl_3$) δ 8.50 (s, 2H), 5.08 (dd, 1H, J_1 = 5.6 Hz, J_2 = 10.0 Hz), 4.46−4.44 (m, 1H), 4.03 (sb, 2H), 4.01−3.97 (m, 1H), 3.72 (dd, 1H, J_1 = 4.4 Hz, J_2 = 11.6 Hz), 3.68 (dd, 1H, J_1 = 4.8 Hz, $J_2$11.6 Hz), 3.08 (s, 6H), 2.25 (ddd, 1H, J_1 = 2.0 Hz, J_2 = 5.6 Hz, J_3 = 12.8 Hz), 2.06 (s, 3H), 1.99 (ddd, 1H, J_1 = 6.0 Hz, J_2 = 10.0 Hz, J_3 = 12.8 Hz) ppm; ^{13}C NMR ($CDCl_3$) δ 166.44, 161.32, 156.62, 126.67, 87.77, 76.09, 73.12, 62.94, 43.09, 38.40, 16.81 ppm; HRMS calcd for $C_{13}H_{21}N_4O_3$ 281.1614, found 281.1616.

(28) Vanvlijmen, H. W. T.; Rame, G. L.; Pettitt, B. M. *Biopolymers* **1990**, *30*, 517−532.
(29) Gianolio, D. A.; McLaughlin, L. W. *J. Am. Chem. Soc.* **1999**, *121*, 6334−6335.

5-[2′-Deoxy-5′-*O*-(4,4-dimethoxytrityl)-*β*-D-ribofuranosyl]-2-[*N*-[1-(dimethylamino)ethylidene]amino]-pyrimidine (8). To a solution of compound **7** (1.2 g, 4.3 mmol) in anhydrous pyridine (150 mL) at 0 °C was added 4,4-dimethoxytrityl chloride (1.8 g, 5.4 mmol). Then the reaction mixture was stirred at ambient temperature. The reaction was completed in 22 h, realized by TLC analysis. The mixture was concentrated under reduced pressure, dissolved in dichloromethane (60 mL), and washed with saturated sodium bicarbonate water solution (3 × 30 mL) and water (3 × 30 mL). The organic portion was dried over sodium sulfate overnight. The drying reagent was removed by filtration and the solution was concentrated in vacuo. The residue was purified by flash chromatography on silica gel (CH_2Cl_2/CH_3-OH 27:1) to give compound 8 (2.1 g, 82% yield) as colorless foam: R_f = 0.58 (CH_2Cl_2/CH_3OH 9:1); UV−vis λ_{max} (CH_3OH) 204 (ϵ = 48700), 233 (ϵ = 20700), 276 (ϵ = 15800) nm; IR (KBr) 3415 (b), 3057 (w), 3034 (w), 3001 (w), 2929 (w), 2906 (w), 2867 (w), 2828 (w), 1608 (s), 1580 (s), 1518 (s), 1446 (s), 1395 (s), 1256 (s) cm^{-1}; ^1H NMR ($CDCl_3$) δ 8.51 (s, 2H), 7.43 (d, 2H, J = 7.2 Hz), 7.32 (d, 4H, J = 8.8 Hz), 7.26 (t, 2H, J = 7.2 Hz), 7.18 (t, 1H, J = 7.2 Hz), 6.81 (d, 4H, J = 8.8 Hz), 5.09 (dd, 1H, J_1 = 5.2 Hz, J_2 = 10.4 Hz), 4.44−4.45 (m, 1H), 4.07−4.04 (m, 1H), 3.77 (s, 6H), 3.32 (dd, 1H, J_1 = 4.4 Hz, J_2 = 9.6 Hz), 3.22 (dd, 1H, J_1 = 5.2 Hz, J_2 = 9.6 Hz), 3.08 (s, 6H), 2.46 (s, 1H), 2.22 (ddd, 1H, J_1 = 1.6 Hz, J_2 = 5.2 Hz, J_3 = 12.8 Hz), 2.07 (s, 3H), 2.03 (ddd, 1H, J_1 = 6.4 Hz, J_2 = 10.4 Hz, J_3 = 12.8 Hz) ppm; ^{13}C NMR ($CDCl_3$) δ 166.88, 161.04, 158.38, 158.36, 156.62, 144.71, 135.97, 135.88, 130.02, 129.96, 128.09, 127.77, 126.71, 126.59, 113.08, 86.59, 86.10, 76.12, 73.96, 64.50, 55.14, 43.41, 38.13, 16.43 ppm; mp 86 °C dec; HRMS calcd for $C_{34}H_{38}N_4O_5$ (M + 1) 583.2921, found 583.2921.

5-[2′-Deoxy-3′-*O*-(2-cyanoethyl-*N*,*N*-diisopropylphosphino)-5′-*O*-(4,4-dimethoxytrityl)-*β*-D-ribofuranosyl]-2-[*N*-[1-(dimethylamino)ethylidene]amino]pyrimidine (9). To a solution of compound **8** (1.0 g, 1.7 mmol) and *N*,*N*-diisopropylethylamine (1.13 mL, 6.8 mmol) in anhydrous dichloromethane (80 mL) was added dropwise 2-cyanoethyl *N*,*N*-diisopropylchlorophosphoramidite (0.53 mL, 2.4 mmol) at 0 °C. The reaction was stirred at ambient temperature for 2 h, methanol (1 mL) was added and stirring was continued for 20 min. The mixture was washed with saturated sodium bicarbonate (3 × 30 mL) and water (3×) and dried over sodium sulfate. Volatiles were removed by rotary evaporation in vacuo, and the residue was purified by precipitation from hexane to give product **9** (0.93 g, 70% yield) as a colorless foam consisting of a 1:1 mixture of diastereoisomers: R_f = 0.44 (CH_2Cl_2/CH_3OH 95:5); ^{31}P NMR ($CDCl_3$) δ 145.8, 145.7 ppm; HRMS calcd for $C_{43}H_{55}N_6O_6P$ (M + 1) 783.3999, found 783.3999.

Oligonucleotide Syntheses. Oligonucleotides were synthesized from 1 μmol of bound nucleoside on 500 Å silica supports using conventional automated standard phosphoramidite chemistry (trityl on mode) on an Applied Biosystems 381A DNA synthesizer. Oligodeoxynucleotides were created of the sequence: 5′-TTTC^{5Me}TTTTC^{5Me}T **X** TC^{5Me}TT-3′ [where C^{5Me} is 5-methyl-dC and **X** is d2APy (**I**) or d2APm (**II**) or alternatively is simply dC]. Controlled pore glass and protecting groups were removed from the oligomers using concentrated ammonium hydroxide (25%) at 55 °C for 12 h. Purification of these oligonucleotides was accomplished by HPLC (trityl on) using 50 mM triethylammonium acetate (pH 7.0) with a linear gradient of 0−35% acetonitrile over 10 min at a flow rate of 7 mL/min through the Oligo R3 reversed phase column. The collected DMT-protected oligonucleotides were reduced in volume, detritylated with 80% aqueous acetic acid (30 mL, 2 h, 0 °C), and desalted using a Sephadex G-10 size exclusion column eluted with water. The purity of the oligomers were determined by reversed phase (C-18) HPLC using 20 mM potassium dihydrogenphosphate buffer solution (pH = 5.5) with a linear gradient of 0−70% methanol over 60 min at a flow rate of 1.2 mL/min.

MALDI-TOF: calcd for oligomer **I** (M + 1) 4464.8, found 4464.9; calcd for **II** (M + 1) 4465.8, found 4465.8.

UV Melting Experiments. UV melting experiments were performed on an AVIV 14DS UV−vis spectrophotometer.

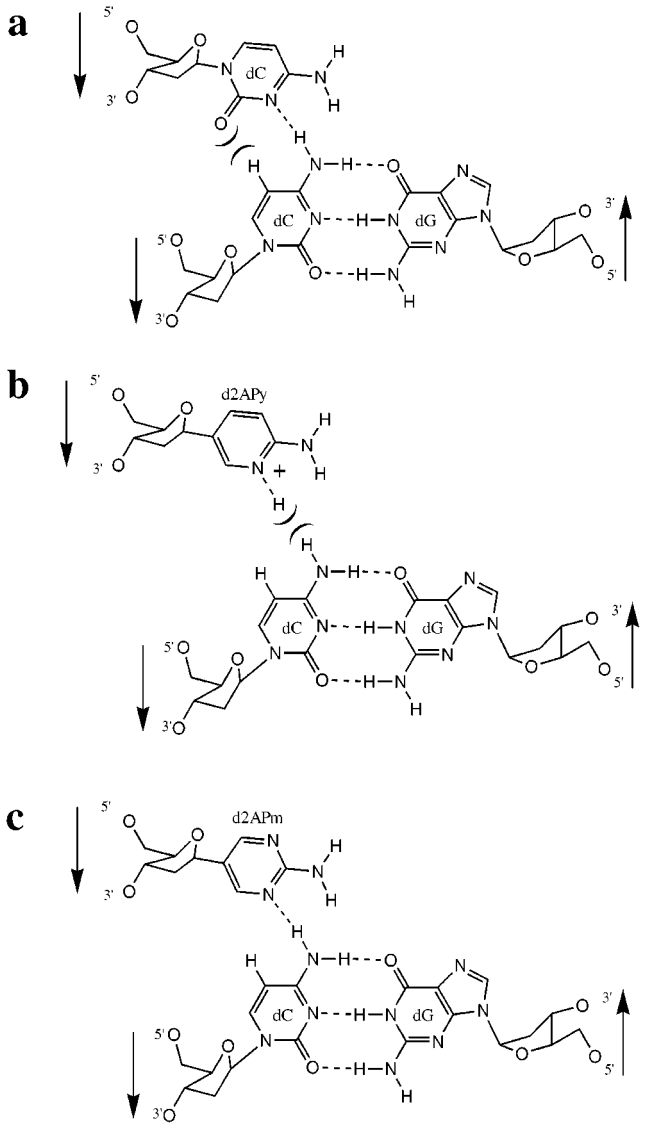


Figure 1. Three possible base triplets involving a dC-dG target base pair and a third strand containing (a) dC, (b) d2APy, and (c) d2APm.

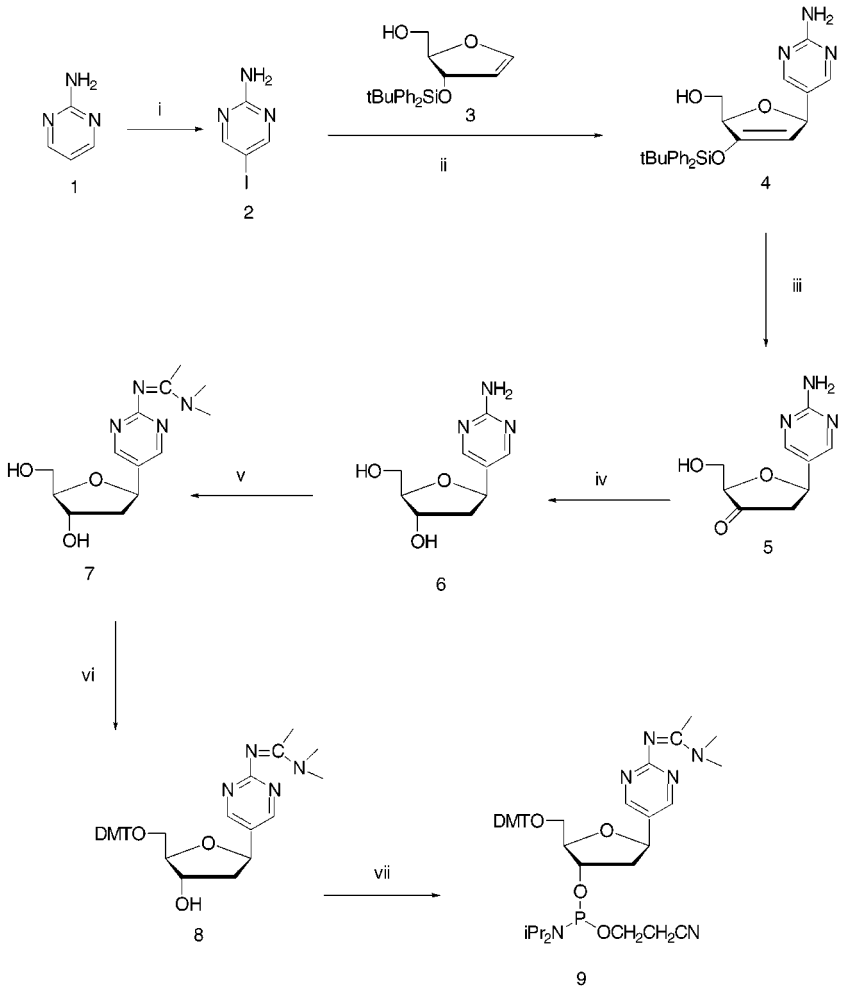
Solutions were adjusted using PIPES (25 mM) for pH values 5.9 and 7.0 and using PEPES (25 mM) for pH values 7.5 and 8.0. All solutions contained 50 mM NaCl and 10 mM $MgCl_2$. Melting curves were recorded at 0.5 °C steps (0−95 °C) with monitoring at 260 nm.

Measurement of p*K*$_a$ of 2-Amino-5-(2′-deoxy-*β*-D-ribofuranosyl)pyrimidine (6). The pK_a of 2-amino-5-(2-deoxy-*β*-D-ribofuranosyl)pyrimidine (**6**) was determined by carrying out a pH titration and monitoring the chemical shift of ^1H NMR of two protons (H4 and H6) of pyrimidine heterocycle. The D_2O solution of **6** (5.7 mM) was titrated with NaOD and DCl solution, and pK_a was calculated from the plot of chemical shift (ppm) vs the pH value.

Results and Discussion

The design of a pyrimidine analogue to interact with a dC-dG base pair target by oligonucleotide directed triply helix formation maintained the six-membered nature of the ring, but eliminated the O2-carbonyl—a site of potentially unfavorable steric interactions with the target pyrimidine base. One possible analogue residue that fits this criterion is the 2-aminopyridine C-nucleoside (d2APy,

Scheme 1[a]



[a] Key: (i) $I_2/H_5IO_6/H_2SO_4/CH_3COOH$, 90 °C, 24 h; (ii) $(dba)_2Pd^0/Ph_3P/iP_2EtN/CH_3CN$, 95 °C, 66 h; (iii) $nBu_4N^+F^-/CH_3COOH/THF$, 0 °C, 1 h; (iv) $NaB(O_2CCH_3)_3H/CH_3COOH/CH_3CN$, 0 °C, 1 h; (v) $(CH_3)_2NC(OCH_3)_2CH_3/CH_3OH$, 78 °C, 24 h; (vi) $DMTCl/C_5H_5N$, rt, 22 h; (vii) $[(CH_3)_2CH_2NP(Cl)OCH_2CH_2CN/iP_2EtN/CH_2Cl_2$, rt, 2 h.

see Figure 1b) and a second analogue is the related 2-aminopyrimidine C-nucleoside (d2APm, see Figure 1c). We prepared the former by a palladium-mediated coupling[30] between the iodinated pyridine heterocycle and the appropriate glycal,[31,32] and we have used this same procedure for the preparation of the 2-aminopyrimidine derivative (Scheme). The use of the *tert*-butyldiphenylsilyl protecting group blocked on side of the sugar glycan such that the Heck coupling occurred only from the "upper" side as illustrated in the Scheme. Using this approach only the β-C-nucleoside was obtained. Removal of the silyl protecting group and reduction of the ketone generated the native nucleoside (**6**) illustrated in Scheme 1. For both nucleosides we have employed a ethylidene protecting group for the exocyclic amino groups and then converted each analogue to the requisite DMT-protected phosphoramidite derivative for solid-phase DNA synthesis.

During DNA sequence assembly, coupling yields with the analogue nucleosides were comparable with those of the common nucleosides. After deprotection and purification, the presence of the analogues could be verified both by HPLC analyses of enzyme digests (see Figure 2). The d2APm nucleoside exhibits a maximum absorbance at 301 nm and consequently when analyzed at 260 nm does not appear in the nucleoside digest (Figure 2, left panel). However, when the digest is reanalyzed at 301 nm the presence of the d2Apm residue could be confirmed. MALDI-TOF analyses of the intact oligonucleotides also confirmed the presence of the analogue residues.

The d2APy analogue has been reported to function well in DNA triplex formation for targeting dG-dC base pairs[33] owing to its increased pK_a value (~6.8) relative to dC (~4.3). This increase in pK_a value permits the analogue to be more effectively protonated at higher pH values, which then provides two hydrogen-bonding interactions for the targeting of the Hoogsteen face of the dG residues in dG-dC base pairs. In the present design, such protonation might prove to be detrimental by eliminating the N3-nitrogen as a hydrogen bond acceptor for the N4-amino group of the target dC residues (Figure 1b). For comparison of protonation effects, we determined the pK_a value of the d2APm analogue by titration with NMR detection and observed a pK_a value of approximately 3.3 (Figure 3).

(30) Daves, G. D. *Acc. Chem. Res.* **1990**, *23*, 201−206.
(31) Hsieh, H. P.; McLaughlin, L. W. *J. Org. Chem.* **1995**, *60*, 5356−5359.
(32) Searls, T.; Chen, D.-L.; Lan, T.; McLaughlin, L. W. *Biochemistry* **2000**, *39*, 4375−4382.
(33) Hildbrand, S.; Blaser, A.; Parel, S. P.; Leumann, C. J. *J. Am. Chem. Soc.* **1997**, *119*, 5499−5511.

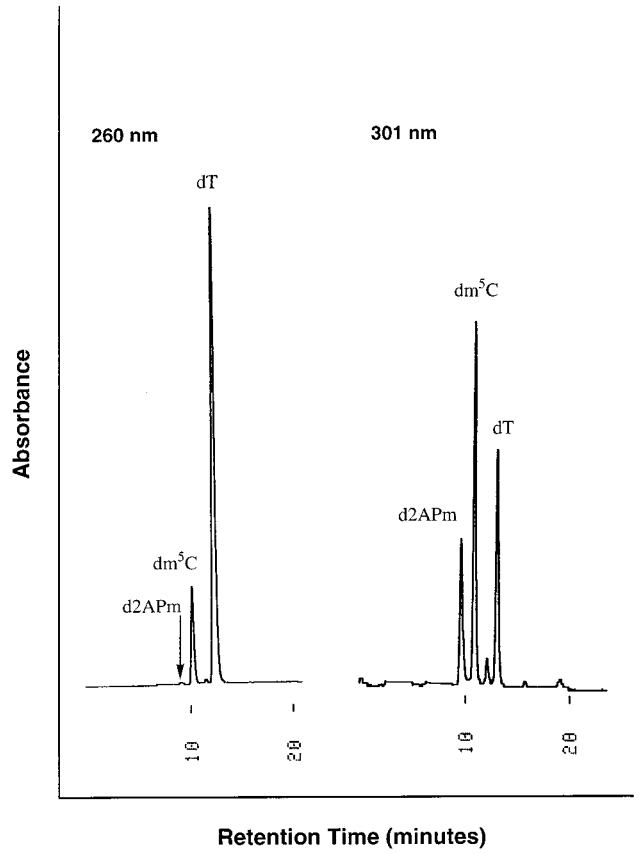


Figure 2. Nucleoside analysis of the digest obtained from the 15-mer sequence containing the analogue d2APm. Left panel: Analysis with detection at 260 nm (d2 APm does not absorb signficantly at this wavelength). Right panel: Analysis with detection at 301 nm.

Triplex Formation. Triplex formation was analyzed by comparative T_m values obtained for either nucleoside analogue relative to native dC. A 25-mer duplex was prepared that contained a 15-mer polypurine target sequence. The target sequence contained eleven dA residues, three dG residues and at a preselected site, also contained either another dG residue or a single dC (Table 1 and Chart 1). The third strand contained eleven dT residues for targeting the dA-dT base pairs, three 5-methyl-2′-deoxycytidine residues for targeting the dG-dC base pairs, and at the site matching the preselected site

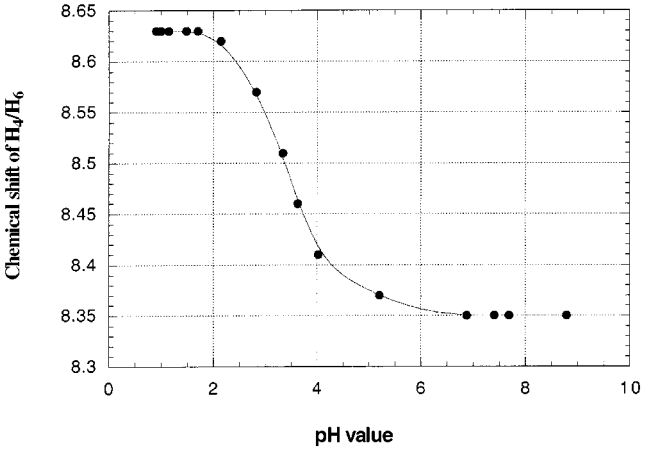


Figure 3. Graph of chemical shift vs pH for the free nucleoside d2APm used to estimate the pK_a value of the analogue nucleoside.

Table 1. T_m Values for 15-Mer Triplexes Containing dC-dC-dG and dC-dG-dC Base Triplets (°C)[a]

```
        5'     TTTMT TTTMT C TMTT
5'    GCGCG AAAGA AAAGA Y AGAA CCCGG
3'    CGCGC TTTCT TTTCT Z TCTT GGGCC
```

	pH value			
	5.9	6.4	7.0	7.5
Y-Z = G-C	53.6	43.8	37.1	25.1
Y-Z = C-G	35.1	25.0	18.2	12.4

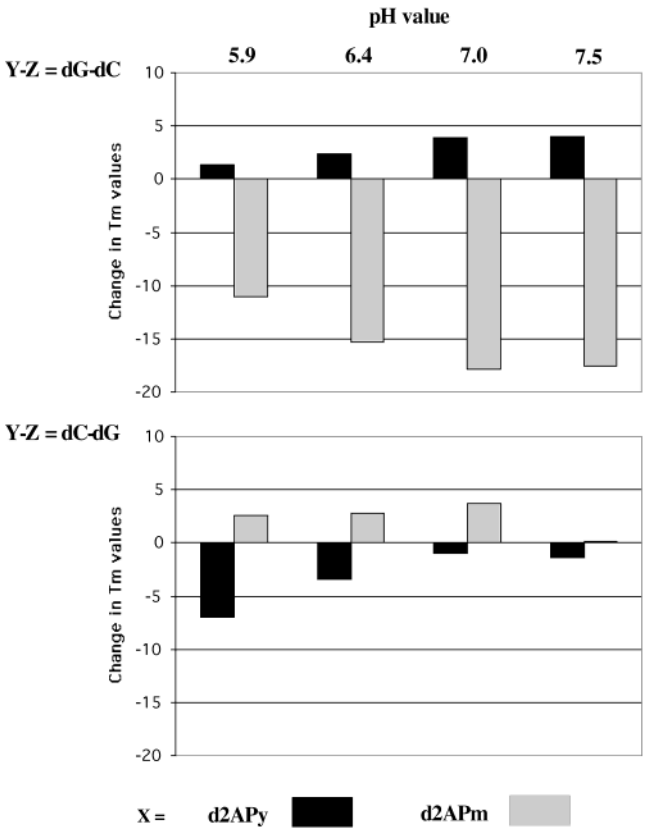
[a] Reported values are the average of at least two independent assays.

containing a dG or dC residue in the target strand, we incorporated a single dC, d2APy or d2APm residue. The absorbance vs temperature plots for these complexes typically exhibited two transitions (Figure 4). The second transition was present in all samples, and was also present in the absence of the third strand and is interpreted to represent the duplex to random coil transition of the target duplex. The position of the earlier transition varied depending upon the nature of the third strand and the target duplex and is interpreted to represent the triplex to duplex transition. The hyperchromicity of the early transition is reduced relative to the later for two reasons, one is that the third strand is 10 residues shorter than the target duplex, and second, the second transition involves the unstacking of two strands, a total of fifty base residues.

With dC in the third strand at the test site, the T_m values were higher with a dG-dC target base pair than

Chart 1. ΔT_m Balues for Triplexes Containing dG-dC or dC-dG Base Pair Targets (°C)

```
        5'          TTTMT TTTMT X TMTT
5'    GCGCG AAAGA AAAGA Y AGAA CCCGG
3'    CGCGC TTTCT TTTCT Z TCTT GGGCC
```



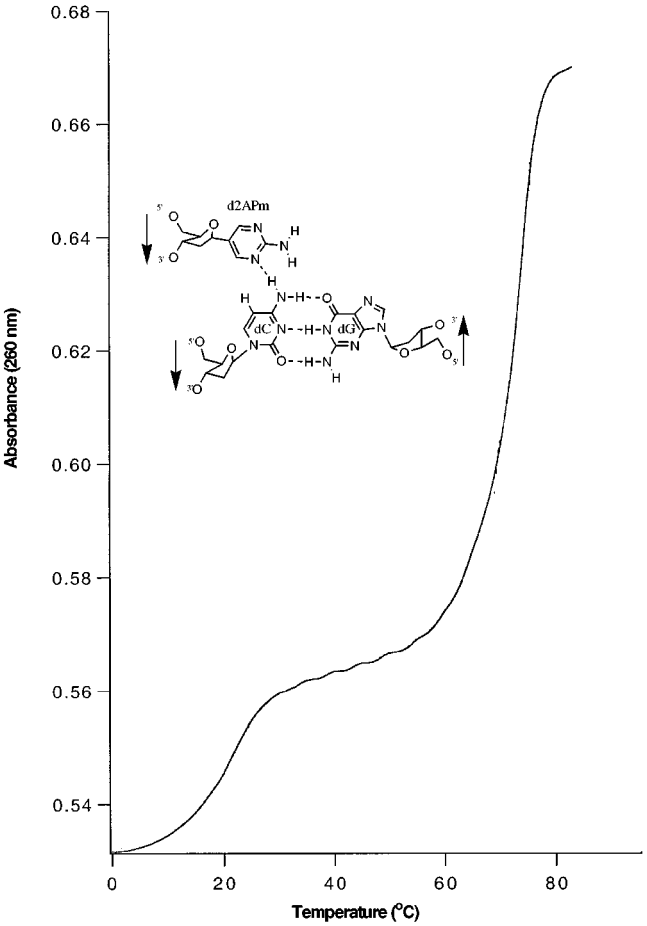


Figure 4. Absorbance vs temperature plot for the 15-mer triplex containing the base triplet d2APm-dC-dG at pH 7.0.

with a dC-dG base pair. This result is expected since two hydrogen bonds can be formed with the target dG residue in the former case, while only one possible hydrogen bond is possible in the latter case (Table 1). As expected for both complexes, increasing the pH value of the solution resulted in a decreases in T_m values largely as the result of the need for protonation to form the three dM-dG-dC base triplets present in the complex (dM = 2′-deoxy-5-methylcytidine) as well as the dC-dG-dC when present in the test site. The destabilization of the dC-dC-dG containing complex relative to the dC-dG-dC containing complex likely argues against the presence of a ═O···H−C≡ hydrogen bond in this complex that provides any significant complex stabilizing interactions. By comparison, recent analyses of such bonds in protein structures suggests that they may be roughly half the magnitude of the corresponding ═O···H−N═ interaction.[34]

When third strands containing either the d2APy, d2APm, or dC residues used to bind a duplex containing a dC in the target sequence there likely exists the potential for but a single interstrand hydrogen bond (Figure 1). Relative to complexes containing dG in the target site, the complexes containing dX-dC-dG base triplets exhibit T_m valued reduced by roughly 18 °C. This is consistent with the observation that bidentate interactions involving the Hoogsteen face of purines result in much more stable triplexes. The experiments here are designed to examine potential design parameters for third strands targeting dC residues and whether alterations in pK_a values can alter selectivities, given that dX-dC-dG base triplets are inherently less stable than the dX-dG-dC counterparts.

In the first set of experiments we compared T_m values for the target strand containing the dG-dC base pair at the preselected site within the target duplex (four containing d2APy in the third strand and four containing d2APm in the third strand) with those containing dC in the third strand. In each case the T_m values were obtained at the pH values 5.9, 6.4, 7.0, and 7.5. In all cases we observed two transitions. The early transition varied with the pH value as would be expected for the triplex to duplex transition. The changes in the T_m values for the transitions (relative to the complex containing dC in the third strand) are plotted as a bar graph in the Table 2 (upper panel, black bars). At pH 5.9 the d2APy analogue results in a T_m value very similar to that for dC, while at pH 7.5 the ΔT_m is +4 °C. This observation is consistent with expectation that the d2APy analogue, with the higher pK_a value, will exist to a greater extent in the protonated form relative to dC. It is also in agreement with a previous report[33] suggesting a 3.9 °C increase in T_m for each d2APy analogue relative to dC.

By comparison, the triplex formed with the d2APm derivative is significantly reduced in stability and even at pH 5.9 exhibits a ΔT_m of −11.1 °C as illustrated in the Table 2 (upper panel, gray bars). This reduction in T_m is consistent with the reduced pK_a value for this derivative (pK_a ∼ 3.3). In fact, owing to the presence of two equivalent nitrogens in the d2APm base, the effective pK_a for a selected nitrogen is further reduced by a factor of 2. Even at the lowest pH value employed in this study, the d2APm derivative will only be partially protonated and unable to effectively form two hydrogens bond with the Hoogsteen face of the dG target.

As noted above, when these same third strands containing either the d2APy or d2APm derivative are used to bind a duplex containing a dC in the target sequence there likely exists the potential for but a single interstrand hydrogen bond (see Figure 1c). Relative to complexes containing the d2APy+-dG-dC base triplet, the complexes containing d2APm-dC-dG base triplets exhibit overall reduced T_m values. However, the results for selectivity by d2APy and d2APm (relative to dC) for a dC-dG target are essentially reversed from that observed for dG-dC targets (Table 2, lower panel). In fact, dC itself is not very effective in promoting triplex formation when opposite to a target dC residue and this may in part be the result of steric influences introduced by the presence of the O2-carbonyl (see Figure 1a). Sequences containing d2APy generally exhibited reduced T_m values relative to dC (black bars, lower panel, Table 2), although these differences largely disappear at higher pH values. Even in the absence of the O2-carbonyl, the higher pK_a value for d2APy will result in more effective protonation of the ring nitrogen and interfere its ability to function effectively as a hydrogen bond acceptor (see Figure 1b). This effect should be more significant at lower pH values, consistent with the observed trend.

By comparison, the d2APm analogue with a target dC residue results in higher T_m values than those obtained with dC (gray bars, lower panel, Table 2). At neutral and slightly acidic pH values the observed ΔT_m values vary between +3 and +4 °C. Two factors might explain this

(34) Garza, V. R.; Dixon, D. A.; Hay, B. P. *J. Am. Chem. Soc.* **2000**, *122*, 4750−4755.

difference, one with two ring nitrogens available in the heterocycle, and in the absence of the O2-carbonyl the base residue can always provide an unprotonated nitrogen to function as a hydrogen bond acceptor. In fact by rotation about the glycosidic bond, it can even do so while one of the ring nitrogens remains protonated. Second, in the absence of the O2-carbonyl, the steric effects between the third strand residue and the target dC are minimized (compare Figures 1a and 1c). At pH 7.5 the ΔT_m value for the d2APm vs dC essentially disappears (lower panel, Table 2). This observation may simply reflect that as the pH value increases the three dM-dG-dC base triplets become less effectively protonated, destabilizing the triplex and masking any difference at the test site between dC and d2APm.

Conclusions

The results presented here suggest that the two C-nucleosides d2APy and d2APm, although similar in structure, exhibit two different selectivities based upon their pK_a values, 6.8 and 3.3, respectively. The d2APy residue with a greater pK_a value, even higher than dC (pK_a = 4.3) is more effective than dC in targeting dG-dC base particularly at slightly basic pH values. By comparison, the d2APm residue with a pK_a lower than dC is more effective than dC in targeting dC-dG base pairs. The overall T_m values for sequences containing the d2APm-dC-dG base triplet are significantly reduced ($\Delta T_m \sim -18$ °C) from those complexes containing d2APy-dG-dC base triplets in part because the latter triplet is formed from a bidentate interaction while the former can only form a single interstrand hydrogen bond (Figure 1c). Nonetheless, the selectivity observed by d2APm for dC-dG, resulting largely from shifting its pK_a value relative to d2APy or dC, suggests that it could function as a scaffold for the introduction of additional stabilizing interactions.

Acknowledgment. This work was supported by a grant from the NIH (GM 53201).

Supporting Information Available: Proton, carbon, and phosphorus NMR spectra as appropriate for all new compounds. This material is available free of charge via the Internet at http://pubs.acs.org.

JO000754W