Sequence Composition Effects on the Energetics of Triple Helix Formation by Oligonucleotides Containing a Designed Mimic of Protonated Cytosine

E. Scott Priestley and Peter B. Dervan*

Contribution from the Arnold and Mabel Beckman Laboratories of Chemical Synthesis, California Institute of Technology, Pasadena, California 91125

Received December 1, 1994[⊗]

Abstract: A nonnatural nucleoside, 1-(2-deoxy-β-D-ribofuranosyl)-3-methyl-5-amino-1*H*-pyrazolo[4,3-*d*]pyrimidin-7-one (P), mimics protonated cytosine and specifically binds GC base pairs within a pyrimidine·purine·pyrimidine triple helix. Quantitative footprint titration experiments at neutral pH (22 °C, 100 mM NaCl, 10 mM bis-tris, 250 μM spermine) reveal dramatic sequence composition effects on the energetics of triple helix formation by oligonucleotides containing P or 5-methylcytosine (^{m}C). Purine tracts of sequence composition 5′-d(AAAAAGAGAGAGAGA)-3′ are bound by oligonucleotide 5′-d(TTTTT^{m}CT^{m}CT^{m}CT^{m}CT^{m}CT)-3′ 4 orders of magnitude more strongly than by 5′-d(TTTTTPTPTPTPTPT)-3′ ($K_T \approx 3 \times 10^9$ M^{-1} and $K_T = 1 \times 10^5$ M^{-1}, respectively). Conversely, purine tracts of sequence composition 5′-d(AAAAGAAAAGGGGGGA)-3′ are bound by oligonucleotide 5′-d(TTTT^{m}CTTTT^{m}C^{m}C^{m}C^{m}C^{m}C^{m}CT)-3′ 5 orders of magnitude less strongly than by 5′-d(TTTT^{m}CTTTTPPPPPPT)-3′ ($K_T < 5 \times 10^4$ M^{-1} and $K_T \approx 4 \times 10^9$ M^{-1}, respectively). The complementary nature of P and ^{m}C expands the repertoire of G-rich sequences which may be targeted by triple helix formation.

Introduction

The thermodynamic stability of pyrimidine·purine·pyrimidine triple helices decreases with increasing pH due to the requirement of protonating cytosine bases to form C+GC triplets.[1,2] Within the context of pyrimidine oligonucleotide-directed recognition of double helical DNA, there are serious sequence composition limitations with regard to targeting *contiguous* G-rich purine tracts near physiological pH, presumably due to electrostatic repulsion between protonated cytosines in adjacent C+GC triplets. Replacement of cytosine by 5-methylcytosine (^{m}C) increases the stability of pyrimidine·purine·pyrimidine triple helices, but does not alleviate the pH dependence (Figure 1).[2] Development of oligonucleotides whose energetics of triple helix formation are less sensitive to pH would benefit applications which require near physiologically relevant conditions.

A series of non-natural nucleosides have been shown to mimic protonated cytosine and bind GC base pairs with significantly less pH dependence, including pseudoisocytosine,[3] 8-oxoadenine,[4] 1-(2-deoxy-β-D-ribofuranosyl)-3-methyl-5-amino-1*H*-pyrazolo[4,3-*d*]pyrimidin-7-one (P),[5] and most recently 4-amino-5-methyl-2,6-pyrimidione.[6] In a qualitative study, we demonstrated that pyrimidine-rich oligonucleotides containing P

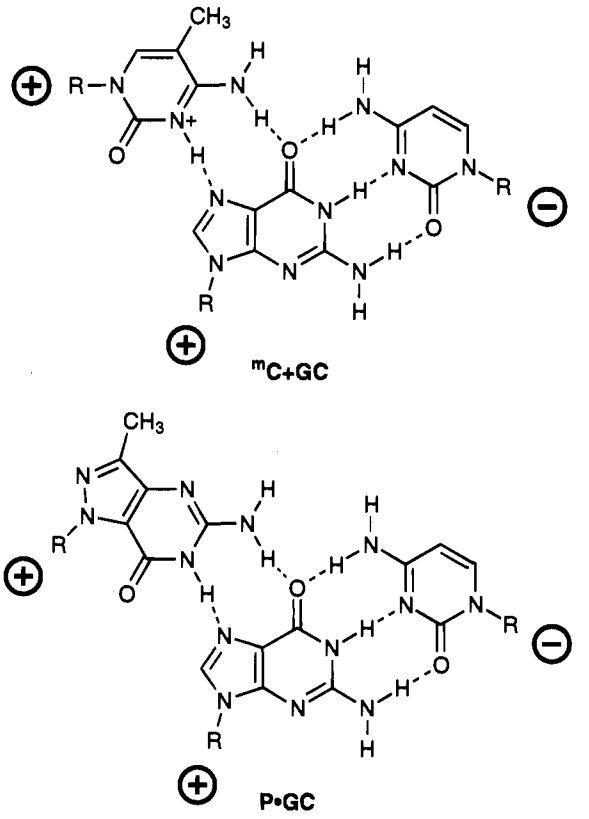

Figure 1. Two-dimensional representations of ^{m}C+GC and P·GC triplets.[5] In each case, the third strand base forms two hydrogen bonds with the Hoogsteen face of the Watson−Crick purine base.

[⊗] Abstract published in *Advance ACS Abstracts,* March 15, 1995.

(1) (a) Moser, H. E.; Dervan, P. B. *Science* **1987**, *238*, 645−650. (b) For a review, see Thuong, N. T.; Helene, C. *Angew Chem. Int. Ed. Engl.* **1993**, *32*, 666−690.

(2) (a) Povsic, T. J.; Dervan, P. B. *J. Am. Chem. Soc.* **1989**, *111*, 3059−3061. (b) Xodo, L. E.; Manzini, G.; Quadrifoglio, F.; van der Marel, G. A.; van Boom, J. H. *Nucleic Acids Res.* **1991**, *19*, 5625−5631. (c) Singleton, S. F.; Dervan, P. B. *Biochemistry* **1992**, *31*, 10995−11003.

(3) (a) Ono, A.; Ts'o, P. O. P.; Kan, L. *J. Am. Chem. Soc.* **1991**, *113*, 4032−4033. (b) Ono, A.; Ts'o, P. O. P.; Kan, L. *J. Org. Chem.* **1992**, *57*, 3225−3230.

(4) (a) Young, S. L.; Krawczyk, S. H.; Matteucci, M. D.; Toole, J. J. *Proc. Natl. Acad. Sci. U.S.A.* **1991**, *88*, 10023−10026. (b) Krawczyk, S. H.; Milligan, J. F.; Wadwani, S.; Moulds, C.; Froehler, B. C.; Matteucci, M. D. *Proc. Natl. Acad. Sci. U.S.A.* **1992**, *89*, 3761−3764. (c) Miller, P. S.; Bahn, P.; Cushman, C. D.; Trapane, T. L. *Biochemistry* **1992**, *31*, 6788−6793. (d) Jetter, M. C.; Hobbs, F. W. *Biochemistry* **1993**, *32*, 3249−3254. (e) Davison, E. C.; Johnsson, K. *Nuclesides Nucleotides* **1993**, *12*, 237−243.

(5) (a) Koh, J. S.; Dervan, P. B. *J. Am. Chem. Soc.* **1992**, *114*, 1470−1478. (b) Note that nucleoside P is labeled P1 in ref 5a and 7.

(6) Xiang, G.; Soussou, W.; McLaughlin, L. W. *J. Am. Chem. Soc.*, in press.

0002-7863/95/1517-4761$09.00/0

could bind target sites containing multiple GC base pairs (Figure 1).[5] However, no conclusions regarding the energetic differences, if any, between sites of sequence composition 5′-d(AAAAAGAGAGAGAGA)-3′ and 5′-d(AAAAGAAAAGGGGGGA)-3′ could be drawn.[5] Although NMR studies do not reveal any major backbone distortion for P·GC triplets, the P·GC triplet is not isomorphous with adjacent T·AT triplets.[7] Therefore, the issue remains whether triple helices containing *multiple* P residues are energetically disfavored relative to ^{m}C at neutral pH. We report here the results of quantitative DNase footprint titration experiments which demonstrate a dramatic *effect of sequence composition* on triple helix formation by oligonucleotides containing P and ^{m}C (Figure 1).

Results and Discussion

Methods. A description of the quantitative footprint titration method has been presented elsewhere for protein−DNA complexation[8] and has been applied to small molecule−DNA interactions.[9] However, the method has not been used extensively to study triple helix formation;[10] therefore, a protocol for this purpose is outlined in the Experimental Section. In separate experiments, 3′-^{32}P-end-labeled 242, 253, or 314 base pair restriction fragments containing purine-rich 15 or 16 base pair target sites were allowed to equilibrate at 22 °C, pH 7.0, with a series of concentrations of third strand oligonucleotide that ranged from 40 μM to 80 pM (Figures 2 and 3). Following a 120 h equilibration, DNase I was added and digestion was allowed to proceed for 6 min at room temperature. After quenching, the reaction mixtures were separated by denaturing polyacrylamide gel electrophoresis and the resulting gels were imaged by storage phosphor autoradiography. Integration of site and reference blocks allowed the determination of the apparent fractional occupancy of the site at each oligonucleotide concentration. A binding isotherm was fit to the resulting pairs of (θ_{app}, $[O]_{tot}$) values (see Experimental Section), and the equilibrium association constant (K_T) was calculated. K_T values from three independent titrations were averaged to obtain each of the association constants reported in Tables 1−3.

Affinity and Specificity of P. In order to test the affinity and specificity of P, the association constants for the formation of 20 triple helical complexes which vary at a single common position, Z·XY (where Z = P, ^{m}C, T, G, and A and XY = AT, CG, GC, and TA), were determined within a pyrimidine motif by quantitative footprint titration at 22 °C, pH 7.0 (Figure 2). The sequence context in which the Z·XY pairings were studied is 5′-d(AAGAGA**X**AGAGGAAA)-3′ within a 314 base pair DNA restriction fragment.[11] The affinity of P relative to ^{m}C is determined by comparison of oligonucleotides **1** and **2** (Z = P and ^{m}C, respectively) opposite XY = GC. The sequence specificity of P is confirmed by comparison of oligonucleotide **1** (Z = P) opposite each of the four Watson−Crick base pairs (XY = AT, CG, GC, and TA). The stability of triple helical

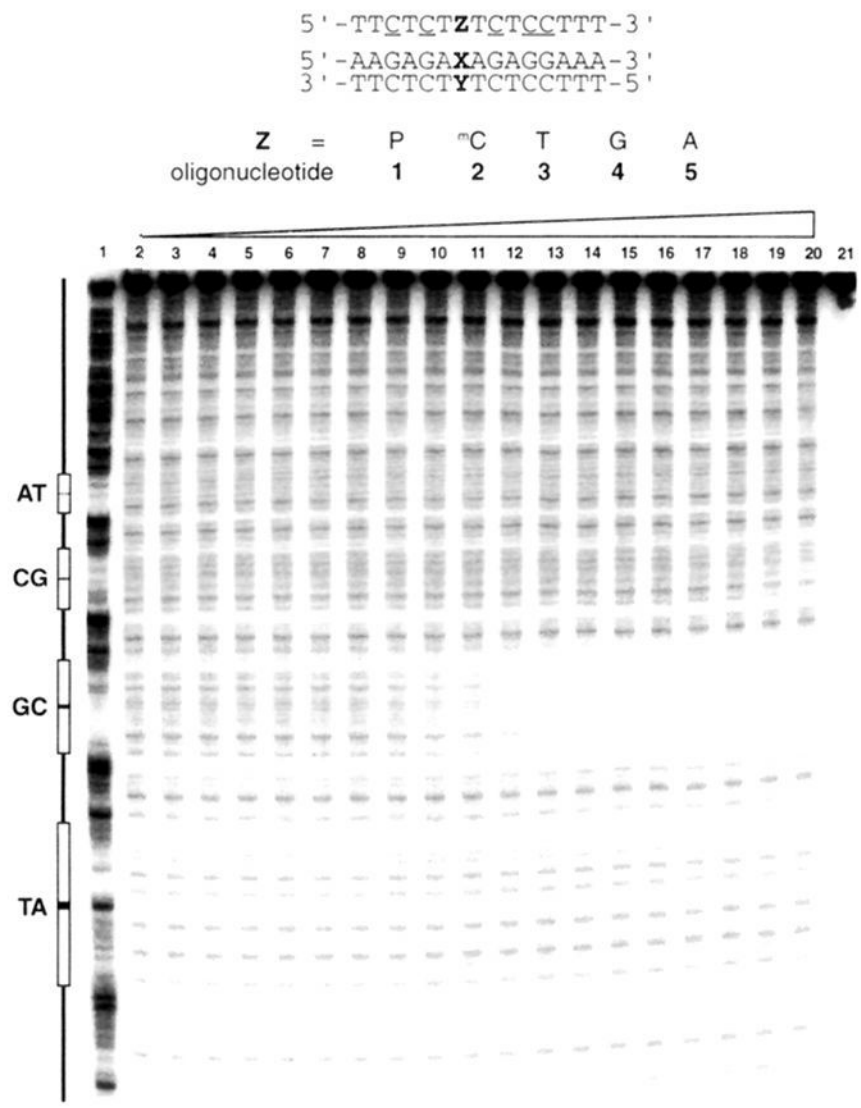

Figure 2. (Top) Sequences of the oligonucleotides and the four purine-rich target sites for the specificity experiments. The duplex target sites, which differ only at a single position labeled XY, are arranged sequentially on the 314 bp *Afl*II/*Fsp*I restriction fragment of pYSPEC2, each separated by 15 bp of intervening sequence. (C indicates 5-methyl-2′-deoxycytidine, and P indicates 1-(2-deoxy-β-D-ribofuranosyl)-3-methyl-5-amino-1*H*-pyrazolo[4,3-*d*]pyrimidin-7-one.) (Bottom) Storage phosphor autoradiogram of an 8% denaturing polyacrylamide gel used to separate the fragments generated by DNase I digestion in a quantitative footprint titration experiment: (lane 1) products of an adenine specific sequencing reaction; (lanes 2−20) DNase I digestion products obtained in the presence of varying concentrations of oligonucleotide **1**, no oligonucleotide (lane 2), 80 pM (lane 3), 200 pM (lane 4), 400 pM (lane 5), 800 pM (lane 6), 2 nM (lane 7), 4 nM (lane 8), 8 nM (lane 9), 20 nM (lane 10), 40 nM (lane 11), 80 nM (lane 12), 200 nM (lane 13), 400 nM (lane 14), 800 nM (lane 15), 2 μM (lane 16), 4 μM (lane 17), 8 μM (lane 18), 20 μM (lane 19), 40 μM (lane 20); (lane 21) intact 3′-labeled DNA after incubation in the absence of the third strand oligonucleotide. The bar drawn to the left of the autoradiogram indicates the positions of the 15 bp binding sites within the 314 bp restriction fragment.

complexes is very sensitive to solution conditions such as pH,[2c] cation valence and concentration,[12] and temperature.[13] Therefore, we have chosen experimental conditions (pH 7.0, 22 °C, 100 mM NaCl, 10 mM bis-tris, 250 μM spermine) such that the association constants for the triple helical complexes span the range which can be measured by quantitative DNase I footprinting.

The values of the 20 association constants (K_T) ranged from $<10^4$ to $\geq 2.1 \times 10^8$ M^{-1} (Table 1). An examination of the data confirms that two base triplets, T·AT and ^{m}C+GC, are particularly stable and that the G·TA triplet is of intermediate stability.[14,15] We find that oligonucleotide **1** (Z = P) binds the target site with an equilibrium association constant of 3×10^7 M^{-1}, a factor of 5 lower affinity than that of oligonucleotide

(7) Radhakrishnan, I.; Patel, D. J.; Priestley, E. S.; Nash, H. M.; Dervan, P. B. *Biochemistry* **1993**, *32*, 11228−11234.
(8) Brenowitz, M.; Senear, D. F.; Shea, M. A.; Ackers, G. K. *Methods Enzymol.* **1986**, *130*, 132−181.
(9) (a) Stankus, A.; Goodisman, J.; Dabrowiak, J. C. *Biochemistry* **1992**, *31*, 9310−9318. (b) Mrksich, M.; Parks, M. E.; Dervan, P. B. *J. Am. Chem. Soc.* **1994**, *116*, 7983−7988.
(10) (a) Singleton, S. F.; Dervan, P. B. *J. Am. Chem. Soc.* **1992**, *114*, 6957−6965. (b) Beal, P. A.; Dervan, P. B. *J. Am. Chem. Soc.* **1992**, *114*, 4976−4982.
(11) For previous reports of a similar four-site assay, see: (a) Jones, R. J.; Lin, K.-Y.; Milligan, J. F.; Wadwani, S.; Matteucci, M. D. *J. Org. Chem.* **1993**, *58*, 2983−2991. (b) Jones, R. J.; Swaminathan, S.; Milligan, J. F.; Wadwani, S.; Froehler, B. C.; Matteucci, M. D. *J. Am. Chem. Soc.* **1993**, *115*, 9816−9817.
(12) Singleton, S. F.; Dervan, P. B. *Biochemistry* **1993**, *32*, 13171−13179.
(13) Singleton, S. F.; Dervan, P. B. *J. Am. Chem. Soc.* **1994**, in press.

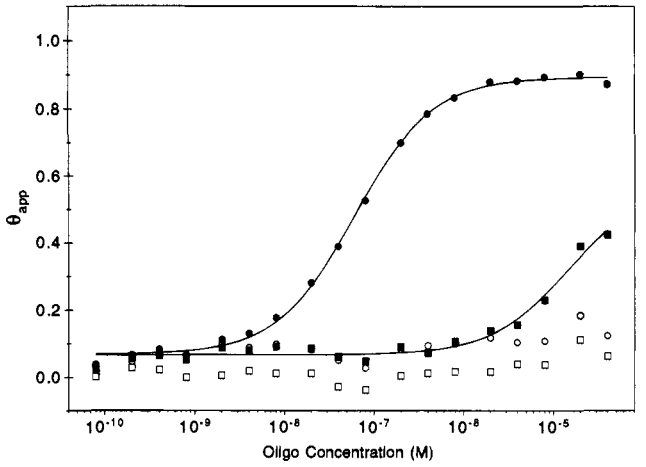

Figure 3. (θ_{app}, $[O]_{tot}$) data derived from the DNase footprinting gel shown in Figure 2 for binding of oligonucleotide **1** to the target sites where XY = AT (○), CG (■), GC (●), or TA (□). For the sites at which binding was observed (CG and GC), theoretical binding isotherms were fit to the data points by nonlinear least squares regression as described in the Experimental Section.

Table 1. Equilibrium Association Constants for 20 Triple Helical Complexes which Vary at a Single Position[a]

oligo	Z[b]	XY	K_T	$\Delta G°$ [c]
1	P	AT	$<10^4$	>-5.4
		CG	$<5 \times 10^4$	>-6.3
		GC	$(3.1 \pm 1.0) \times 10^7$	-10.1 ± 0.2
		TA	$<10^4$	>-5.4
2	^{m}C	AT	$(8.2 \pm 1.9) \times 10^4$	-6.6 ± 0.1
		CG	$(1.7 \pm 0.6) \times 10^5$	-7.1 ± 0.2
		GC	$\geq 1.4 \times 10^8$	<-11.0
		TA	$<10^4$	>-5.4
3	T	AT	$\geq 2.1 \times 10^8$	<-11.2
		CG	$(5.4 \pm 0.7) \times 10^5$	-7.7 ± 0.1
		GC	$<10^4$	>-5.4
		TA	$<10^4$	>-5.4
4	G	AT	$<10^4$	>-5.4
		CG	$<10^4$	>-5.4
		GC	$<5 \times 10^4$	>-6.3
		TA	$(1.1 \pm 0.3) \times 10^6$	-8.2 ± 0.1
5	A	AT	$(1.6 \pm 0.6) \times 10^5$	-7.0 ± 0.2
		CG	$<10^4$	>-5.4
		GC	$<10^4$	>-5.4
		TA	$<10^4$	>-5.4

[a] Each reported K_T value is the mean ±SEM of three independent measurements which were performed in 100 mM NaCl, 10 mM bis-tris, and 250 μM spermine at pH 7.0 and 22 °C. Values for K_T and $\Delta G°$ are reported in units of M^{-1} and kcal·mol^{-1}, respectively. [b] Z refers to the position in the third strand oligonucleotide opposite the variable position (XY) in the duplex. [c] $\Delta G°$ values were obtained from the relationship $\Delta G° = -RT \ln K$ (T = 295 K), and error estimates were obtained by propagating the SEM for K_T through this formula.

2 (Z = ^{m}C). Importantly, P is specific for GC base pairs. The stabilities of triple helical complexes containing P·XY decrease in the order P·GC > P·CG > P·AT, P·TA (Table 1). The high affinity and sequence specificity observed are consistent with a model where P forms two specific hydrogen bonds to the purine strand of the Watson–Crick GC base pair.

Table 2. Equilibrium Association Constants for Triple Helix Formation at the $(GA)_5$ Site[a]

oligo	pH	K_T (M^{-1})
6	7.0	$\approx 3 \times 10^9$
	7.5	$(9.4 \pm 1.6) \times 10^7$
7	7.0	$(1.1 \pm 0.2) \times 10^5$
	7.5	$(1.3 \pm 0.2) \times 10^5$

[a] Each reported K_T value is the mean ± SEM of three independent measurements which were performed in 100 mM NaCl, 10 mM bis-tris, and 250 μM spermine at the indicated pH and 22 °C.

Table 3. Equilibrium Association Constants for Triple Helix Formation at the G_6 Site[a]

oligo	pH	K_T (M^{-1})
8	7.0	$<5 \times 10^4$
	7.5	$<10^4$
9	7.0	$\approx 4 \times 10^9$
	7.5	$\approx 2 \times 10^9$

[a] Each reported K_T value is the mean ±SEM of three independent measurements which were performed in 100 mM NaCl, 10 mM bis-tris, and 250 μM spermine at the indicated pH and 22 °C.

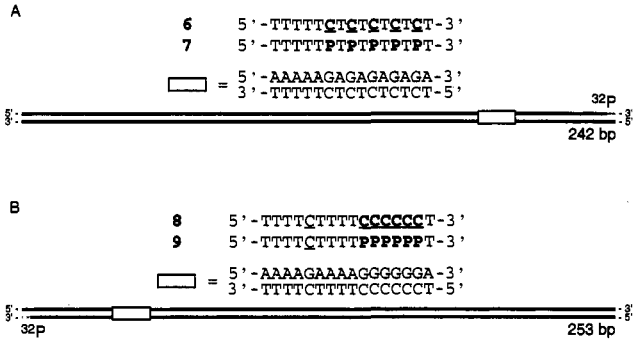

Figure 4. Sequence composition experiments: (A) sequences of oligonucleotides **6** and **7** and the $(GA)_5$ target site; the target site is located within the 242 bp *Eco*RI/*Pvu*II restriction fragment of pGCBGC; (B) sequences of oligonucleotides **8** and **9** and the G_6 target site; The target site is located within the 253 bp EcoRI/PvuII restriction fragment of pSPHIV. (C indicates 5-methyl-2′-deoxycytidine, and P indicates 1-(2-deoxy-β-D-ribofuranosyl)-3-methyl-5-amino-1*H*-pyrazolo[4,3-*d*]-pyrimidin-7-one.)

Sequence Composition Effects. Sequence composition effects on the energetics of triple helix formation by oligonucleotides containing multiple P moieties were examined for two different purine target sites within 242 and 253 base pair restriction fragments at pH 7.0 and 7.5 (Figure 4). The sequence of one site is the purine tract 5′-d(AAAAAGAGAGAGAGA)-3′ in a 242 base pair (bp) restriction fragment (which we will refer to as the $(GA)_5$ site). The other purine site is derived from the LTR region of the HIV genome and has the sequence 5′-d(AAAAGAAAAGGGGGGA)-3′ in a 253 bp restriction fragment (which we will refer to as the G_6 site) (Figure 4).

The energetics of association of oligonucleotides **6** and **7** of sequence composition 5′-d(TTTTT^{m}CT^{m}CT^{m}CT^{m}CT^{m}CT)-3′ and 5′-d(TTTTTPTPTPTPTPT)-3′, respectively, allow comparison of the ability of ^{m}C and P to bind multiple GC base pairs in a target sequence composed of alternating G and A nucleotides (the $(GA)_5$ site). Similarly, the energetics of binding oligonucleotides **8** and **9** of sequence composition 5′-d(TTTT-^{m}CTTTT^{m}C^{m}C^{m}C^{m}C^{m}C^{m}CT)-3′ and 5′-d(TTTT^{m}CTTTTPPPP-PPT)-3′ allow this comparison for a target sequence containing contiguous GC base pairs (the G_6 site).

(14) For early qualitative work, see: (a) Griffin, L. C.; Dervan, P. B. *Science* **1989**, *245*, 967–971. (b) Belotserkovskii, B. P.; Veselkov, A. G.; Filippov, S. A.; Dobrynin, V. N.; Mirkin, S. M.; Frank-Kamenetskii, M. D. *Nucleic Acids Res.* **1990**, *18*, 6621–6624. (c) Mergny, J.-L.; Sun, J.-S.; Rougee, M.; Montenay-Garestier, T.; Barcelo, F.; Chomilier, J.; Helene, C. *Biochemistry* **1991**, *30*, 9791–9798. (d) 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. (e) Fossella, J. A.; Kim, Y. J.; Shih, H.; Richards, E. G.; Fresco, J. R. *Nucleic Acids Res.* **1993**, *21*, 4511–4515.

(15) For a quantitative affinity cleavage titration study, see: Best, G. C.; Dervan, P. B. *J. Am. Chem. Soc.* **1995**, *117*, 1187.

We find that oligonucleotide **6**, containing ^{m}C and T, binds tightly to the $(GA)_5$ site ($K_T \approx 3 \times 10^9$ M^{-1}) at pH 7.0, and this affinity drops by at least a factor of 30 as the pH is increased to 7.5. This is consistent with previous data which have shown that triple helix formation by oligonucleotides containing cytosine or ^{m}C is pH dependent.[2] At pH 7.0, oligonucleotide **7**, containing P in place of ^{m}C, binds 4 orders of magnitude weaker than oligonucleotide **6**, albeit pH independently over the range studied. *The situation changes entirely for the G_6 target sequence containing contiguous GC base pairs.* Oligonucleotide **8**, which contains $(^{m}C)_6$, binds the G_6 site very weakly ($K_T < 5 \times 10^4$) at pH 7.0. In the case of oligonucleotide **9**, which contains P_6, the association constant at pH 7.0 is very high ($K_T \approx 4 \times 10^9$ M^{-1}). Remarkably, the high affinity oligonucleotide **9** decreases by only a factor of 2 at pH 7.5.

The $(GA)_5$ site contains ten 5′-AG-3′ or 5′-GA-3′ junctions, while the G_6 site contains only two such junctions in the region of interest. The 4 orders of magnitude difference in the affinity of P-containing oligonucleotides for these two sites is probably due to the lack of structural isomorphism in the P•GC and T•AT triplets. The location of the third strand deoxyribose−phosphate backbone is not identical when the two triplets are overlaid. Thus, the 5′-AG-3′ and 5′-GA-3′ junctions could generate energetically unfavorable distortions in the backbone in triple helical complexes relative to contiguous A or G sequences. In the case of ^{m}C -containing oligonucleotides, the large difference in affinity for the two sites is probably due to electrostatics, since the T•AT and ^{m}C+GC triplets appear structurally isomorphous. We would conclude from this limited study that, in ^{m}C- and, T-containing oligonucleotides, the composition in the third strand of sequences 5′-TT-3′, 5′- ^{m}CT-3′, and 5′-T ^{m}C-3′ makes the most favorable interactions with the double helix, while a 5′- ^{m}C^{m}C-3′ dinucleotide repeat is unfavorable. Sequences of contiguous ^{m}C+GC triplets may be energetically unfavorable due to repulsion between adjacent protonated ^{m}C bases.

In conclusion, a quantitative study of the energetics of triple helix formation for different purine tracts has revealed that ^{m}C and P provide complementary solutions to the recognition of GC base pairs by triple helix formation. At neutral pH, ^{m}C binds isolated GC base pairs with higher affinity than P, while P binds contiguous GC base pairs with higher affinity than ^{m}C. Remarkably, both types of G-rich tracts can be targeted by choice of the appropriate oligonucleotide composition. In addition, this work emphasizes the importance of studying the effects of sequence composition on energetics as design−synthesis of non-natural bases expands the repertoire of sites which may be targeted by triple helix formation.

Experimental Section

Materials. Aqueous solutions for DNA manipulations were prepared with water purified by a Milli-Q Water System (Millipore). *Escherichia coli* XL-1 Blue competent cells were obtained from Stratagene. Deoxynucleotide triphosphates (Ultrapure grade) and calf thymus DNA (sonicated and phenol extracted) were purchased from Pharmacia, diluted to the appropriate concentration, and stored at −20 °C. [α-^{32}P]-thymidine-5′-triphosphate (≥3000 Ci/mmol) and [α-^{32}P]deoxyadenosine-5′-triphosphate (≥6000 Ci/mmol) were obtained from Du Pont NEN. Restriction endonucleases were purchased from either New England Biolabs or Boehringer Mannheim and used according to the supplier's protocol in the buffer provided. Sequenase (version 2.0) was obtained from United States Biochemicals, and deoxyribonuclease I (FPLCpure) was purchased from Pharmacia. Snake venom phosphodiesterase and alkaline phosphatase were obtained from Boehringer Mannheim. Buffers and salts for preparation of association buffer were Fluka Microselect grade. 5-methyl-2′-deoxycytidine phosphoramidite was purchased from Cruachem. All other DNA synthesis reagents were provided by Glen Research.

Oligonucleotide Preparation. Oligonucleotides were synthesized by standard automated solid-support chemistry on an Applied Biosystems 380B DNA synthesizer using *O*-(cyanoethyl)-*N,N*-diisopropylphosphoramidites. P phosphoramidite was synthesized as described previously.[5] Oligonucleotides **2**−**5** were synthesized with the 5′-OH unprotected and deblocked for 20−24 h in concentrated aqueous NH_3. These oligomers were initially purified by ion exchange FPLC (Pharmacia) on a Mono Q HR 10/10 column (Pharmacia) using a gradient of 0.1−1.0 M NaCl in 10 mM bis-trisCl, pH 7.0. Oligonucleotides **1** and **6**−**9** were synthesized with a 5′-DMT protecting group and deblocked for 24−36 h in concentrated aqueous NH_3. These oligonucleotides were initially purified by reversed phase FPLC on a ProRPC HR 10/10 column (Pharmacia) using a linear gradient of 0−40% CH_3CN in 0.1 M triethylammonium acetate, pH 7.0, and then detritylated with 80% acetic acid/H_2O. Both sets of oligonucleotides were chromatographed a second time by reversed phase FPLC using either 0−20% CH_3CN in 10 mM NH_4HCO_3, pH 8.1 (oligonucleotides **2**−**5**), or 0−20% CH_3CN in 0.1 M triethylammonium acetate, pH 7.0 (oligonucleotides **1** and **6**−**9**). A final desalting step on a NAP-10 column (Pharmacia) was performed for all oligonucleotides. Concentrations of the oligonucleotide solutions were determined by UV absorbance at 260 nm, using extinction coefficients of 8800 (M^{-1}·cm^{-1}) for thymidine, 5700 (M^{-1}·cm^{-1}) for 5-methyl-2′-deoxycytidine, 11 700 (M^{-1}·cm^{-1}) for deoxyguanosine, 15 400 (M^{-1}·cm^{-1}) for deoxyadenosine, and 2500 (M^{-1}·cm^{-1}) for P. The oligonucleotides were stored dry as 5 or 10 nmol aliquots in 0.5 mL microcentrifuge tubes at −20 °C.

Enzymatic Digestion and HPLC Analysis. Oligonucleotide aliquots (10 nmol) were digested with three units of calf alkaline phosphatase and 0.009 unit of snake venom phosphodiesterase in 50 μL of 50 mM trisHCl and 10 mM $MgCl_2$, pH 8.0, for 2−20 h at 37 °C. Aliquots (25 μL) of these solutions were analyzed by reversed phase HPLC on a C18 column (VYDAC No. 201HS54) with a gradient of 0−50% CH_3CN in 20 mM NH_4OAc, pH 5.0. Comparison with chromatograms of nucleoside standards and coinjections with authentic materials established the base composition of all oligonucleotides.

Plasmid Construction. pYSPEC2 and pSPHIV were prepared by standard methods.[16] Briefly, pYSPEC2 was prepared by hybridizing three complementary sets of synthetic oligonucleotides: 5′-CCGGC-CGAAGTCTTGAGGCTTTCCTCTTTCTCTTCCGAAGTCTTGA-GGCTTTCCTCTGTCTCTCCGAA-3′ and 5′-AGACTTCGGAAGAGA-CAGAGGAAAGCCTCAAGACTTCGGAAGAGAAAA-GAGGAAAGCCTCAAGACTTCGG-3′, 5′-GTCTTGAGGCTTTC-CTCTCTCTCTTCCGAAGTCTTGAGGCTTTCCTCTATCTCTTCC-GAAGTCTTGAGGC-3′ and 5′-AGCTGCCTCAAGACTTCGGAA-GAGATAGAGGAAAGCCTCAAGACTTCGGAAGAGAGA-GAGAGAGAGGAAAGCCTCA-3′, 5′-AGCTTGGCGTAATCATG-GTCCTTAAGTTCGAAG-3′ and 5′-TCGACTTCGAACTTAAGGAC-CATGATTACGCCA-3′. These self-complementary oligonucleotide duplexes were phosphorylated with adenosine 5′-triphosphate and T4 polynucleotide kinase and then ligated into the large *AvaI*/*SalI* restriction fragment of pUC19. *E. coli* XL-1 Blue competent cells were transformed by the ligated plasmid, and plasmid DNA from ampicillin-resistant white colonies was isolated. The presence of the desired insert was determined by restriction analysis and dideoxy sequencing. The plasmid was isolated on a preparative scale using a Qiagen Maxi Prep kit. pSPHIV was prepared in a similar manner, except that the phosphorylation step was omitted and the following synthetic oligonucleotides were cloned into the large *Bam*HI/*Hind*III restriction fragment of pUC19, 5′-GATCTTAGCCACTTTTTAAAAGAAAA-GGGGGGACTGGAAGGGCTAATTA-3′ and 5′-AGCTTAATTAGC-CCTTCCAGTCCCCCCTTTTCTTTTAAAAAGTGGCTAA-3′.

3′ End Labeling of Restriction Fragments. In a typical end labeling, 2 μg of plasmid was digested with the appropriate restriction enzymes (*Afl*II/*Fsp*I for pYSPEC2 and *Eco*RI/*Pvu*II for pGCBGC and pSPHIV), phenol/$CHCl_3$ extracted, and EtOH precipitated. The resulting fragments were 3′ end labeled with [α-^{32}P]thymidine-5′-triphosphate and [α-^{32}P]deoxyadenosine-5′-triphosphate, using Sequenase (version 2.0). Excess unlabeled nucleotide triphosphates were added after labeling to ensure complete fill in. Unincorporated nucleotides were

(16) Sambrook, J.; Fritsch, E. F.; Maniatis, T. *Molecular Cloning*; Cold Spring Harbor Laboratory Press: Plainview, NY, 1989.

removed on a gel filtration column (Microspin S-200 HR, Pharmacia), and the labeled restriction fragments were separated on a 5% non-denaturing polyacrylamide gel. The desired band was excised from the gel, crushed, and eluted overnight at 37 °C into 25 mM trisHCl, 250 mM NaCl, and 1 mM EDTA, pH 8.0. This solution was filtered, and the restriction fragment was precipitated with isopropyl alcohol. The labeled DNA was resuspended in TE, pH 7.5, phenol/$CHCl_3$ extracted three times, and EtOH precipitated with 2 M NaCl. The purified restriction fragment was then resuspended in TE, pH 7.5, to an activity of approximately 30 000 cpm/μL. A typical yield from 2 μg of plasmid DNA was a total Cerenkov radioactivity of 3 000 000 cpm.

DNase I Footprint Titrations. These experiments were performed essentially according to the published protocol.[8] Typically, a 10 nmol aliquot of the third strand oligonucleotide was resuspended in H_2O to a concentration of 100 μM, and this solution was serially diluted to a total of 18 concentrations spanning the desired range for the experiment. The minimum amount of solution transferred in these dilutions was 20 μL, in order to minimize pipetting error. A stock solution was prepared by combining 5X association buffer (50 mM bis-trisHCl, 500 mM NaCl, 1.25 mM spermine, previously adjusted to the appropriate pH with HCl), 50 μM base pairs calf thymus DNA, the labeled restriction fragment solution, and H_2O in appropriate amounts to give a total volume of 567 μL. This stock solution was then distributed in 27 μl aliquots among 20 labeled 1.7 mL microcentrifuge tubes. The appropriate oligonucleotide solution or H_2O (for reference lanes) (18 μL) was added to each tube. The tubes were then vortexed, centrifuged, and then allowed to equilibrate for either 3 (sequence composition experiments) or 5 (specificity experiments) days at 22 °C (final solution conditions: 100 mM NaCl, 10 mM bis-tris, 250 μM spermine, 5 μM bp calf thymus DNA, ~20 000 cpm labeled restriction fragment). After equilibration, 5 μL of a solution containing DNase I (at the appropriate concentration to give ~50% intact labeled DNA; this was determined by titration of the DNase activity), 50 mM $CaCl_2$, 50 mM $MgCl_2$, 10 μM non-specific oligonucleotide (5′-d(CGACCGTCGA)-3′), 10 mM bis-trisHCl at the appropriate pH, and 5% glycerol was added, and digestion was allowed to proceed for 6 min at room temperature. The DNase reactions were quenched by the addition of 8.3 μL of 1.4 M NaCl, 0.14 M EDTA, pH 8.0, and 0.35 μg/μL glycogen, followed by 120 μL of EtOH. The reactions were stored at −20 °C for several hours and then centrifuged at 14000*g* for 30 min at 0 °C. The supernatant was decanted and the DNA pellet washed with cold 75% EtOH. The pellet was resuspended in 20 μL H_2O, frozen, and lyophilized to dryness. The pellet was then resuspended in 7 μl 80% formamide loading buffer containing 1X TBE. The samples were denatured at 85 °C for 10 min, chilled on ice, and immediately loaded onto a preelectrophoresed 8% denaturing polyacrylamide gel. The reaction products were separated by electrophoresis in 1X TBE buffer at 1800 V. The gel was dried for 1 h at 80 °C on a slab drier, and exposed to a storage phosphor screen (Molecular Dynamics) overnight at ambient temperature.

Quantitation and Data Analysis. Data from the footprint titrations were obtained by quantitating the sequencing gels using ImageQuant software (Molecular Dynamics) running on an AST Premium 386/33 computer. Volume integration of rectangles encompassing the footprint site(s) and a reference site at which DNase reactivity was invariant across the titration generated values for site intensity (I_{site}) and reference intensity (I_{ref}). These I_{site} and I_{ref} values were corrected by subtracting out background intensity as described in the detailed protocol.[8] The apparent fractional occupancy (θ_{app}) of the site at each oligonucleotide concentration ($[O]_{tot}$) was calculated from the corrected I_{site} and I_{ref} values using the following equation:

$$\theta_{app} = 1 - \frac{I_{site}/I_{ref}}{I^{o}_{site}/I^{o}_{ref}}$$

where I^{o}_{site} and I^{o}_{ref} are the site and reference intensities, respectively, from a control lane to which no third strand oligonucleotide has been added. The resulting pairs of (θ_{app}, $[O]_{tot}$) values were plotted on a semilog scale. The following binding isotherm was fit to the experimental data using a nonlinear least squares algorithm in the program Kaleidagraph 3.0.1 running on a Macintosh IIfx or IIci:

$$\theta_{app} = \theta_{min} + (\theta_{max} - \theta_{min})\frac{K_T[O]_{tot}}{1 + K_T[O]_{tot}}$$

where θ_{min} is the apparent fractional occupancy at the lowest oligonucleotide concentrations, θ_{max} is the apparent fractional occupancy at saturation, and K_T is the equilibrium association constant. Data reported in the paper are the average of three determinations ± SEM. Each titration within a set of three measurements was prepared from a unique aliquot and serial dilution of oligonucleotide, a unique preparation of labeled restriction fragment, and unique batches of association buffer and calf thymus DNA.

The experimental conditions employed in this study are quite stringent, and in numerous cases, triple helix formation was not observed. The association constant is reported as $<10^4$ in these cases, as this is the limit below which binding is not observed in this assay. For those cases in which a footprint is only observed at the highest oligonucleotide concentrations, and reliable curve fit could not be obtained, the association constant is reported as $<5 \times 10^4$. In those cases where the apparent binding constant is above 5×10^8, it is reported as an approximate value. This is due to the fact that the necessary assumption that the concentration of labeled DNA is at least 2 orders of magnitude lower than the triple helix dissociation constant[8] is no longer strictly valid. (We estimate that the concentration of labeled restriction fragment in these experiments is approximately 20 pM.) Finally, in the specificity experiments, the triplet combinations with the highest binding affinities (T·AT and ^{m}C+GC) were not at equilibrium even when the equilibration times were extended to 120 h; these values are therefore reported as lower limits. The true association constants are probably 2−4-fold greater than the reported values.

Acknowledgment. We are grateful for financial support from the Office of Naval Research and for a National Science Foundation predoctoral fellowship to E.S.P.

JA943900M