VOLUME 117, NUMBER 4
FEBRUARY 1, 1995

JOURNAL OF THE AMERICAN CHEMICAL SOCIETY

Energetics of Formation of Sixteen Triple Helical Complexes Which Vary at a Single Position within a Pyrimidine Motif

George C. Best and Peter B. Dervan*

Contribution from the Division of Chemistry and Chemical Engineering, California Institute of Technology, Pasadena, California 91125

Received September 12, 1994[⊗]

Abstract: The association constants for the formation of 16 pyrimidine · purine · pyrimidine triple helical complexes which vary at a single common position (Z · XY where Z = A, G, ^{Me}C, T and XY = AT, GC, CG, TA) were determined by quantitative affinity cleavage titration. The sequence context in which the 16 pairings Z · XY were studied is 5′-d(AAAAAGAGAXAGAGA)-3′ within a 242 base pair DNA fragment. The association constants of four pyrimidine rich oligodeoxyribonucleotides, 5′-d(TTTTT^{Me}CT^{Me}CTZT^{Me}CT^{Me}CT)-3′, ranged from 3.8×10^4 M^{-1} (Z · XY = A · CG) to 5.6×10^7 M^{-1} (Z · XY = T · AT) at 22 °C, pH 7.0, 10 mM bis-tris, 100 mM NaCl, 250 μM spermine. For the sequences studied, T · AT is more stable than T · GC, T · CG, or T · TA by ≥2.3 kcal mol^{-1}; ^{Me}C+GC is more stable than ^{Me}C · AT, ^{Me}C · CG, or ^{Me}C · TA by ≥1.4 kcal mol^{-1}; G · TA is more stable than G · AT, G · CG, or G · GC by ≥0.8 kcal mol^{-1}.

Oligonucleotide-directed triple helix formation is a versatile method for accomplishing the sequence specific recognition of double helical DNA.[1] Triple helices can be classified into two structural motifs: those in which the third strand is primarily composed of pyrimidine bases, and those in which the third strand is primarily composed of purine bases.[1-3] Remarkably, triple helices composed of combinations of these two motifs can also be formed within some sequence contexts.[4] Triple helix formation is sensitive to the length of the third strand,[1a,5] single base mismatches,[1a,5,6] pH,[7] cation concentration and valence,[8] temperature,[9] and backbone composition (DNA or RNA) of the three strands.[10] Oligonucleotide-directed triple helix formation has been shown to inhibit sequence specific DNA binding proteins[11] and has been used to mediate single site cleavage of human chromosomal DNA.[12] The ability to target a broad range of DNA sequences[1,13] and the high specificity and stability of the resulting local triple helical structures make this a powerful

[⊗] Abstract published in *Advance ACS Abstracts*, January 1, 1995.

(1) (a) Moser, H. E.; Dervan, P. B. *Science* **1987**, *238*, 645–650. (b) LeDoan, T.; Perrouault, L.; Praseuth, D.; Habhoub, N.; Decout, J.-L.; Thuong, N. T.; Lhomme, J.; Hélène, C. *Nucleic Acids Res.* **1987**, *15*, 7749–7760.

(2) (a) Cooney, M.; Czernuszewicz, G.; Postel, E. H.; Flint, S. J.; Hogan, M. E. *Science* **1988**, *241*, 456–459. (b) Beal, P. A.; Dervan, P. B. *Science* **1991**, *251*, 1360–1363. (c) Beal, P. A.; Dervan, P. B. *Nucleic Acids Res.* **1992**, *20*, 2773–2776. (d) Stilz, H. U.; Dervan, P. B. *Biochemistry* **1993**, *32*, 2177–2185. (e) Malkov, V. A.; Voloshin, O. N.; Soyfer, V. N.; Frank-Kamenetskii, M. D. *Nucleic Acids Res.* **1993**, *21*, 585–591.

(3) For early work on three stranded RNA polymers, see: Felsenfeld, G.; Davies, D.; Rich, A. *J. Am. Chem. Soc.* **1957**, *79*, 2023–2024.

(4) For alternate strand binding, see: (a) Jayasena, S. D.; Johnston, B. H. *Biochemistry* **1992**, *31*, 320–327. (b) Beal, P. A.; Dervan, P. B. *J. Am. Chem. Soc.* **1992**, *114*, 4976–4982. (c) Jayasena, S. D.; Johnston, B. H. *Nucleic Acids Res.* **1992**, *20*, 5279–5288. (d) Jayasena, S. D.; Johnston, B. H. *Biochemistry* **1993**, *32*, 2800–2807.

(5) Singleton, S. F.; Dervan, P. B. *J. Am. Chem. Soc.* **1992**, *114*, 6957–6965.

(6) (a) Mergny, J.-L.; Sun, J.-S.; Rougée, M.; Montenay-Garestier, T.; Barcelo, F.; Chomilier, J.; Hélène, C. *Biochemistry* **1991**, *30*, 9791–9798. (b) Rougée, M.; Faucon, B.; Mergny, J. L.; Barcelo, F.; Giovannageli, C.; Garestier, T.; Hélène, C. *Biochemistry* **1992**, *31*, 9269–9278.

(7) (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.

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

(9) Singleton, S. F.; Dervan, P. B. *J. Am. Chem. Soc.* **1994**, in press.

(10) (a) Roberts, R. W.; Crothers, D. M. *Science* **1992**, *258*, 1463–1466. (b) Han, H.; Dervan, P. B. *Proc. Natl. Acad. Sci. U.S.A.* **1993**, *90*, 3806–3810. (c) Escude, C.; François, J. C.; Sun, J. S.; Ott, G.; Sprinzl, M.; Garestier, T.; Hélène, C. *Nucleic Acids Res.* **1993**, *21*, 5547–5553. (d) Han, H.; Dervan, P. B. *Nucleic Acids Res.* **1994**, *22*, 2837.

0002-7863/95/1517-1187$09.00/0

technique for the recognition of single sites within megabase segments of double helical DNA.

In an effort to understand the interactions which give rise to recognition in the major groove of DNA by oligonucleotides, the specificity afforded by the natural bases in triple helix formation has been characterized qualitatively by affinity cleavage,[14] optical melting,[15] gel mobility shift,[16] and intramolecular triple helix formation.[17] Specific interactions in triple helix formation implicated by these studies have been characterized by NMR spectroscopy.[18,19] Hoogsteen type hydrogen bonds have been observed between thymine (T) residues in the third strand and adenine−thymine (AT) base pairs in the duplex,[18] and between N3 protonated cytosine (C+) or 5-methylcytosine (^{Me}C+) in the third strand and guanine−cytosine (GC) base pairs in the duplex.[19] Recent studies have extended the pyrimidine motif through the incorporation of both purine and non-natural bases.[20] The best characterized of these are the G·TA, D_3·TA, and P_1·GC triplets.[21,22,23]

The recent introduction of quantitative methods has allowed the determination of the equilibrium association constants for the binding of an oligonucleotide at a single site on a DNA plasmid fragment.[5] Quantitative affinity cleavage titration (QACT) was used to determine the effects of pH, cation valency and concentration, and temperature on the energetics of triple helix formation for one sequence context.[7c,8,9] Although reports in the literature suggest that three to four triplets are particularly stable, considerably less is known about the quantitative difference among all 16 natural triplets. Here we report the equilibrium association constants and free energies of formation of 16 triple helical complexes which vary at a single position (Z·XY where Z = A, G, ^{Me}C, and T and Watson−Crick XY = AT, GC, CG, and TA) (Figure 1). We are primarily interested in comparing the differences of free energy values ($\Delta\Delta G°$) which are relevant to the issue of specificity.

Results and Discussion

Methods. A detailed description of the theoretical and experimental basis of the quantitative affinity cleavage titration technique has been presented elsewhere;[5] hence, the protocol is outlined briefly in the Experimental Section.[24] A 3′-^{32}P end labeled 242 base pair restriction fragment containing a single purine rich 15 base pair target site 5′-d(AAAAAGAGAX-AGAGA)-3′ was allowed to equilibrated at 22 °C (pH 7.0) with a series of concentrations of third strand pyrimidine rich oligonucleotide-EDTA·Fe, 5′-d(TTTTT^{Me}CT^{Me}CTZT^{Me}CT^{Me}CT)-3′, that ranged from 10 μM to 250 pM (Figure 1). After a 24 h equilibration, oxidative cleavage reactions were initiated with the addition of 4 mM dithiothreitol (DTT). These reactions were quenched after 6 h by ethanol precipitation of the DNA. 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 site-specific cleavage (I_{site}) for each oligonucleotide concentration ([O_{tot}]). Equation 1 (see Experimental Section) was fit to the pairs of I_{site}, [O_{tot}] data by nonlinear least-squares regression, using the association constant, K_T, and maximal cleavage, I_{sat}, as adjustable parameters (Figure 2). K_T values from three independent titrations were averaged to obtain the association constants reported in Table 1.

Affinity. Experimental conditions, such as temperature, salt concentration, and pH, are chosen such that the lowest and highest equilibrium binding constants for the 16 triple helical complexes can be measured. The cleavage titration experiments were carried out at 22 °C, pH 7.0, 10 mM bis-tris, 100 mM NaCl, 250 μM spermine. The values of the 16 association constants (K_T) range from 4×10^4 to 6×10^7 M^{-1} (Table 1). An examination of the data confirms that three base triplets (T·AT, ^{Me}C+GC, and G·TA) are particularly stable. However, these interactions are not all of the same strength and other "mismatched" triplets are close in value (Table 2). Triple helical complexes containing the T·AT and ^{Me}C+GC triplets at the Z·XY position are the most stable interactions in the pyrimidine·purine·pyrimdine triple helical complexes. The total standard free energy of the helix forming reactions is $\Delta G° = -10.5$ and -10.3 kcal mol^{-1}, respectively. In contrast, the G·TA base triplet is of intermediate stability in this sequence context. The total standard free energy of the helix forming reactions is $\Delta G° = -8.8$ kcal mol^{-1}. Triple helices containing the T·AT and ^{Me}C+GC base triplets at the variable position are approximately 15-fold more stable than those which contain the G·TA base triplet at that position.

Specificity. If specific local triple helices are to be useful structures for accomplishing the recognition of single sites in megabase DNA, it is important that the complex be stabilized via specific interactions and be sensitive to single base mismatches. The results presented here demonstrate that there is a unique best choice of base Z in the third strand to accomplish the sequence specific recognition of three of the

(11) (a) Maher, L. J.; Dervan, P. B.; Wold, B. *Science* **1989**, *245*, 725−730. (b) Collier, D. A.; Thuong, N. T.; Hélène, C. *J. Am. Chem. Soc.* **1991**, *113*, 1457−1458. (c) Maher, L. J.; Dervan, P. B.; Wold, B. *Biochemistry* **1992**, *31*, 70−81. (d) Grigoriev, M.; Praseuth, D.; Robin, P.; Hemar, A.; Saison-Behmoaras, T.; Dautry-Varsat, A.; Thuong, N. T.; Hélène, C.; Harel-Bellan, A. *J. Biol. Chem.* **1992**, *267*, 3389−3395.

(12) Strobel, S. A.; Doucette-Stamm, L. A.; Riba, L.; Housman, D. E.; Dervan, P. B. *Science* **1991**, *254*, 1639−1642.

(13) Horne, D. A.; Dervan, P. B. *J. Am. Chem. Soc.* **1990**, *112*, 2435−2437.

(14) (a) Griffin, L. C.; Dervan, P. B. *Science* **1989**, *245*, 967−971. (b) Beal, P. A.; Dervan, P. B. *Nucleic Acids Res.* **1992**, *20*, 2773−2776. (c) Shimizu, M.; Inoue, H.; Ohtsuka, E. *Biochemistry* **1994**, *33*, 606−613.

(15) (a) Mergny, J.-L.; Sun, J.-S.; Rougée, M.; Montenay-Garestier, T.; Barcelo, F.; Chomilier, J.; Hélène, C. *Biochemistry* **1991**, *30*, 9791−9798. (b) Miller, P. S.; Cushman, C. D. *Biochemistry* **1993**, *32*, 2999. (c) Fossella, J. A.; Kim, Y. J.; Shih, H.; Richards, E. G.; Fresco, J. R. *Nucleic Acids Res.* **1993**, *21*, 4511−4515.

(16) Yoon, K.; Hobbs, C. A.; Koch, J.; Sardaro, M.; Kutny, R.; Weis, A. *Proc. Natl. Acad. Sci. U.S.A.* **1992**, *89*, 3840−3844.

(17) (a) 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. (b) Xodo, L. E.; Alunni-Fabbroni, M.; Manzini, G.; Quadrifoglio, F. *Eur. J. Biochem.* **1993**, *212*, 395−401.

(18) (a) de los Santos, C.; Rosen, M.; Patel, D. J. *Biochemistry* **1989**, *28*, 7282−7289. (b) Rajagopal, P.; Feigon, J. *Nature* **1989**, *339*, 637−640. (c) Rajagopal, P.; Feigon, J. *Biochemistry* **1989**, *28*, 7859−7870.

(19) Live, D. H.; Radhakrishnan, I.; Misra, V.; Patel, D. J. *J. Am. Chem. Soc.* **1991**, *113*, 4687−4688.

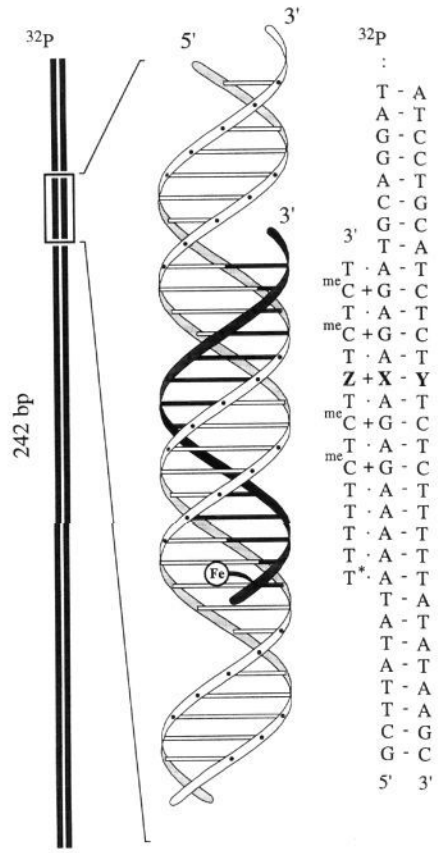
(20) (a) Ono, A.; Ts'o, P. O. P.; Kan, L. *J. Am. Chem. Soc.* **1991**, *113*, 4032−4033. (b) Griffin,L. C.; Kiessling, L. L.; Beal, P. A.; Gillespie, P.; Dervan, P. B. *J. Am. Chem. Soc.* **1992**, *114*, 7976−7982. (c) Miller, P. S.; Bhan, P.; Cushman, C. D.; Trapane, T. L. *Biochemistry* **1992**, *31*, 6788−6793. (d) 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.

(21) G·TA triplet references: (a) Griffin, L. C.; Dervan, P. B. *Science* **1989**, *245*, 967−971. (b) Radhakrishnan, I.; Gao, X.; de los Santos, C.; Live, D.; Patel, D. J. *Biochemistry* **1991**, *30*, 9022−9030. (c) Wang, E.; Malek, S.; Feigon, J. *Biochemistry* **1992**, *31*, 4838−4846. (d) Radhakrishnan, I.; Patel, D. J.; Veal, J. M.; Gao, X. *J. Am. Chem. Soc.* **1992**, *114*, 6913−6915.

(22) P_1·GC triplet references: (a) Koh, J. S.; Dervan, P. B. *J. Am. Chem. Soc.* **1992**, *114*, 1470−1478. (b) Radhakrishnan, I.; Patel, D. J.; Priestly, E. S.; Nash, H.; Dervan, P. B. *Biochemistry* **1993**, *32*, 11228−11234.

(23) D_3·TA and D_3·CG references: (a) Griffin, L. C.; Kiessling, L. L.; Beal, P. A.; Gillespie, P.; Dervan, P. B. *J. Am. Chem. Soc.* **1992**, *114*, 7976−7982. (b) Koshlap, K. M.; Gillespie, P.; Dervan, P. B.; Feigon, J. *J. Am. Chem. Soc.* **1993**, *115*, 7908−7909.

(24) In controls, the QACT assay has been checked against a different method, quantitative footprint titration, and the agreement with QACT is very good (see ref 5).



Oligonucleotide	Z
1	A
2	^{Me}C
3	G
4	T

Figure 1. (Left) Ribbon model of the triple helix formed upon binding of the 15mer oligonucleotide probe to the 15 bp target site within a 242 bp restriction fragment. The Watson−Crick duplex strands are depicted as white ribbons while the oligonucleotide third strand is depicted as a dark ribbon. (Right) The sequence modeled by the ribbons is shown. The position of the variable triplet is indicated as Z·XY, where Z = A, ^{Me}C, G, T and XY = AT, CG, GC, TA.

four Watson−Crick base pairs XY in double helical DNA (comparison across rows of Table 2) (Figures 3 and 4). For example, T·AT is more stable than ^{Me}C·AT, A·AT, and G·AT by ≥1.6 kcal/mol. ^{Me}C+GC is more stable than G·GC, T·GC, and A·GC by ≥2.3 kcal/mol. G·TA is more stable than T·TA, A·TA, and ^{Me}C·TA by >1.6 kcal/mol. These interactions are also specific in the reciprocal sense. For three of the four bases Z studied (Z = T, ^{Me}C, G), one of the four Z·XY triplets formed is particularly stable (Table 2). The T·AT triplet is 2.3 kcal mol^{-1} more stable than the T·CG triplet, the next most stable T·XY triplet. Similarly, ^{Me}C+GC is 1.4 kcal mol^{-1} more stable than any other ^{Me}C·XY. The G·TA triplet is 1.5 kcal mol^{-1} more stable than any other Z·TA triplet.

These results confirm that G does offer a relatively specific means of recognizing a *single* TA base pair within this sequence context.[25] The lower free energy value of the G·TA base triplet relative to the T·AT base triplet may result from the combination of the formation of only one hydrogen bond, rather than the two seen in the T·AT and ^{Me}C+GC base triplets, and the distortion away from planarity of and loss of continuous stacking caused by the G·TA triplet[21d] (Figures 3 and 4). Since the sequence used in this study provides the ideal flanking base pairs for the G·TA triplet,[25] the relative weakness of the G·TA interaction suggests that a novel base that can recognize TA base pairs with higher affinity than G is needed, particularly if sequences containing more than one TA base pair are to be targeted.

The data in Tables 1 and 2 reveal that, under these conditions and in this sequence context, there is no particularly favorable Z·CG triplet. While a comparison across the CG row of Table 2 reveals that both T and ^{Me}C form triplets with CG that are of intermediate stability, these interactions are not specific, since both T and ^{Me}C can form other energetically preferred base triplets (comparison down the columns of Table 2). This suggests that, while either of T or ^{Me}C might be used to tolerate the presence of one CG base pair in a target site, use of these bases would result in a decrease in the overall specificity of the oligonucleotide. Thus the specific recognition of the CG base pair remains a design-synthesis challenge. We would anticipate that the relative rankings of the mismatched base triplets may differ from one study to the next due to the differences in sequence studied, as well as differences in pH and salt concentration.

A major issue that remains to be solved in a systematic manner is the effect of sequence composition on the stability of various base triplets. It has been established qualitatively that the identity of the neighboring triplets does have an effect on the stability of a given triplet.[25] This effect is, in part, due to the importance of stacking interactions in stabilizing the triple helix. The need to protonate C or ^{Me}C residues suggests that binding contiguous GC base pairs, which requires protonation at adjacent residues in the third strand, is energetically disfavored.[22a]

Conclusion. These results demonstrate that the T·AT and ^{Me}C+GC base triplets are significantly more stable than the 14 other triplets studied. The G·TA base triplet possesses a stability between that of the T·AT, ^{Me}C+GC base triplets and the worst mismatches. Importantly, the T·AT, ^{Me}C+GC, and G·TA base triplets are highly specific interactions. This means that oligonucleotides containing T, ^{Me}C, and G are capable of recognizing target sites composed of Watson−Crick AT, GC, and TA base pairs in a specific manner; the energetic penalty for such an oligonucleotide binding to a site with a single mismatch is in the range of 0.8−2.3 kcal/mol. We would emphasize that our results are for one sequence composition context and that the dependence of the energetics of triple helix formation on sequence composition remains to be elucidated.

Experimental Section

General. Sonicated, phenol extracted calf thymus DNA (Pharmacia) was dissolved in H_2O to a final concentration of 2.0 mM in base pairs and was stored at 0 °C. Glycogen (20 mg/mL in water) and deoxynucleotide triphosphates were supplied by Boehringer Mannheim. α-^{32}P nucleotide triphosphates (3000 Ci/mmol) were purchased from Amersham or DuPont/NEN. α-^{35}S dATP for Sanger sequencing was obtained from Amersham. Cerenkov radioactivity was counted on a Beckman LS 2801 scintillation counter. UV−visible spectroscopy was performed on a Hewlett-Packard 8452A diode array spectrophotometer. Restriction enzymes and the Klenow fragment of DNA polymerase were purchased from Boehringer−Mannheim and used in the buffers provided. 5-Methyl-2′-deoxycytidine phosphoramidite was purchased from Cruachem. Other phosphoramidites were purchased from ABI. Bis-tris was purchased from Sigma. All other chemicals were of reagent grade and were used as supplied.

Oligonucleotide Preparation. T* phosphoramidite was prepared

(25) Kiessling, L. L.; Griffin, L. C.; Dervan. P. B. *Biochemistry* **1992**, *31*, 2829−2834.

Table 1. Association Constants (K_T) for the Formation of 16 Triple Helical Complexes Containing the Z·XY Triplets at 22 °C, 100 mM NaCl, 250 μM Spermine, 10 mM Bis-tris, and pH 7.0[a,b]

	Z =			
XY	G	T	^{Me}C	A
AT	2.1 (±0.7) × 10^5	5.6 (±0.3) × 10^7	3.8 (±0.4) × 10^6	4.8 (±2.4) × 10^5
GC	8.0 (±2.8) × 10^5	4.3 (±0.6) × 10^5	4.1 (±0.7) × 10^7	2.7 (±0.8) × 10^5
CG	4.6 (±2.5) × 10^5	1.2 (±0.4) × 10^6	1.3 (±0.7) × 10^6	3.8 (±0.8) × 10^4
TA	3.2 (±0.8) × 10^6	2.0 (±0.9) × 10^5	1.5 (±0.6) × 10^5	2.1 (±1.2) × 10^5

[a] K_T values are reported as the mean (± the standard error of the mean) of three measurements. The K_T values are reported in units of M^{-1}. [b] The identity of the base Z is indicated across the column headss; the identity of the Watson–Crick base pair XY is indicated in the first column.

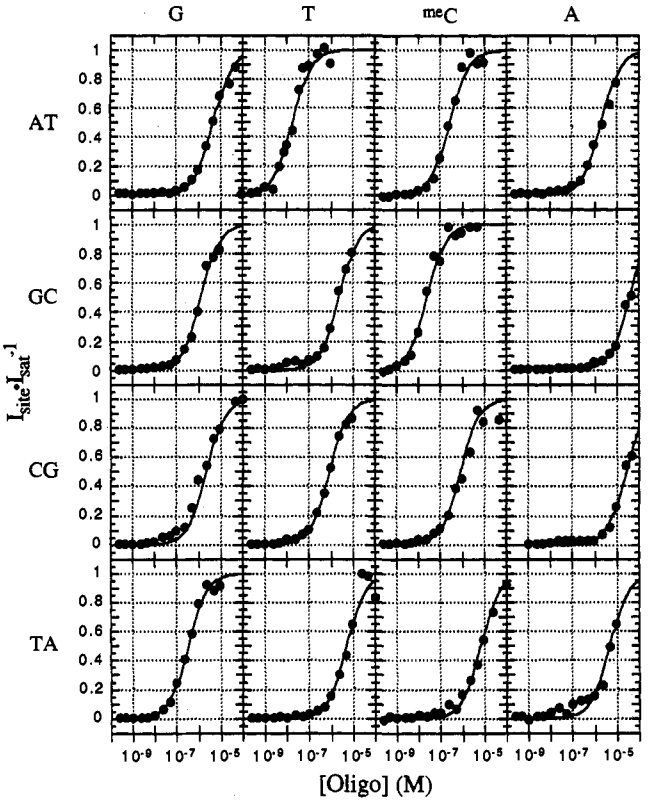


Figure 2. Binding isotherms for the 16 Z·XY triplets studied. The base Z is indicated across at the top of the columns. The base pair XY is indicated to the left of each row. Each isotherm represents the average of three experiments conducted at 22 °C, 10 mM bis-tris, pH 7.0, 100 mM NaCl, 250 μM spermine.

Table 2. Free Energy (ΔG°) of Formation of 16 Triple Helical Complexes Containing the Z·XY Triplets[a,b]

	G	T	^{Me}C	A
AT	−7.2 ± 0.2	−10.5 ± 0.1	−8.9 ± 0.1	−7.7 ± 0.3
GC	−8.0 ± 0.2	−7.6 ± 0.1	−10.3 ± 0.1	−6.0 ± 0.2
CG	−7.6 ± 0.3	−8.2 ± 0.2	−8.2 ± 0.3	−6.2 ± 0.2
TA	−8.8 ± 0.1	−7.2 ± 0.3	−7.0 ± 0.2	−7.2 ± 0.5

[a] Free energy values are calculated from the measured association constants at 22 °C and are reported in kcal mol^{-1}. [b] The identity of the base in the third strand, Z, is indicated across the column heads; the identity of the Watson–Crick base pair, XY, is indicated in the first column.

following established procedures.[26] Oligonucleotides were synthesized on a 1 μmol scale using an ABI model 380B DNA synthesizer and the DNA synthesis protocols supplied by the manufacturer. After the final coupling, the 5′-dimethoxytrityl group was left on the oligonucleotide; deprotection and cleavage from the support was accomplished by treatment with 0.1 N NaOH at 55 °C for 24 h. The oligonucleotides were desalted on NAP 25 columns (Pharmacia), frozen, and lyophilized. The residue was dissolved in 2 mL of 100 mM triethylammonium acetate (TEAA) buffer, pH 7.0, filtered through a 0.45 μm cellulose acetate filter, and purified by FPLC on a ProRPC HR 16/10 reversed phase column using a linear gradient from 0 to 40% acetonitrile in 100 mM TEAA, pH 7.0. Fractions containing the fully deprotected oligonucleotide were combined and lyophilized. The residue was dissolved in water, frozen, and lyophilized. Detritylation was accomplished by treatment of the residue with 1.5 mL of 40% acetic acid in H_2O (20 min, room temperature). The acetic acid was removed *in vacuo* and the residue dissolved in 2 mL of 100 mM TEAA, pH 7.0, filtered through a 0.45 μm filter, and FPLC purified as above. Fractions containing the oligonucleotide were pooled, frozen, and lyophilized. The residue was then lyophilized twice more from water. Prior to packaging, the oligonucleotides were desalted on NAP 25 columns. Oligonucleotides were packaged in 5 nmol aliquots according to the UV absorbance at 260 nm. Extinction coefficients were determined by adding the extinction coefficients of the individual monomers: T and T*, 8700 M^{-1}; ^{Me}C, 5700 M^{-1} dG, 11 500 M^{-1}; dA, 15 400 M^{-1}. The resulting aliquots were lyophilized to dryness and stored at −20 °C.

Plasmid Preparation. The plasmids for the work reported in this paper were prepared by established procedures.[27] Briefly, oligonucleotides were synthesized, deprotected, and purified by reversed phase FPLC. The oligonucleotides were then annealed to give the insert duplex and added to *Bam* HI/*Hin*d III cleaved pUC 19. The resulting mixture was treated overnight with T4 DNA ligase and the ligation mixture was then used to transform Epicurean Coli XL1-Blue Supercompetent cells (Stratagene) according to the manufacturer's protocol. Transformants containing a plasmid with an insert were identified by α complementation and grown overnight in 5 mL of liquid LB/ampicillin. The plasmids were isolated using a Qiagen miniprep kit and sequenced using a Sequenase 2.0 sequencing kit (U. S. Biochem). Large-scale plasmid preparation was accomplished by the Qiagen maxiprep procedure. The sequence of the isolated plasmids was confirmed by repeating the sequencing protocol. The isolated plasmids were diluted to a concentration of ≈2.5 μg/μL and stored at −20 °C.

DNA Labeling. The procedure used to prepare 3′ end labeled DNA was as follows: The initial restriction digest and 3′ end labeling were performed simultaneously. To 20 μg of plasmid was added the following: 10–20 units of *Eco* RI, 10 units of DNA polymerase, Klenow fragment, 10 μL (≈100 μCi) of α-^{32}P dATP, 10 μL (≈100 μCi) of α-^{32}P TTP, 10 μL of 10× *Eco* RI reaction buffer, and sufficient water for a total reaction volume of 90 μL. After a 2.5 h incubation at 37 °C, 10 μL of a solution of cold dNTPs (10 mM in each) was added, and the incubation was allowed to continue for another 0.5–1 h. Unincorporated radioactivity was removed using a NICK column (Pharmacia) and the linearized plasmid was ethanol precipitated. After drying, the pellet was resuspended and then digested with 20–30 units of Pvu II (final reaction volume 100 μL). After a 4 h incubation at 37 °C, 20 μL of 15% Ficoll loading buffer containing only bromphenol blue was added. The labeled fragment was purified by preparative polyacrylamide gel electrophoresis (5% nondenaturing gel, 1:29 cross linked, 200 V, 1 h). The desired 242 bp fragment was located by autoradiography and the band was excised from the gel and crushed. After addition of 1 mL of elution buffer (10 mM Tris HCl, pH 7.0, 10 mM EDTA, 250 mM NaCl, 0.1% SDS), the DNA was incubated overnight at 37 °C. The elution buffer was filtered to remove polyacrylamide fragments and 700 μL of 2-propanol was added to

(26) Dreyer, J. B.; Dervan, P. B. *Proc. Natl. Acad. Sci. U.S.A.* **1985**, *82*, 968–972.

(27) Sambrook, J.; Fritsch, E. F.; Maniatis, T. *Molecular Cloning: A LAboratory Manual*; Cold Spring Harbor Press: Cold Spring Harbor, NY, 1989.

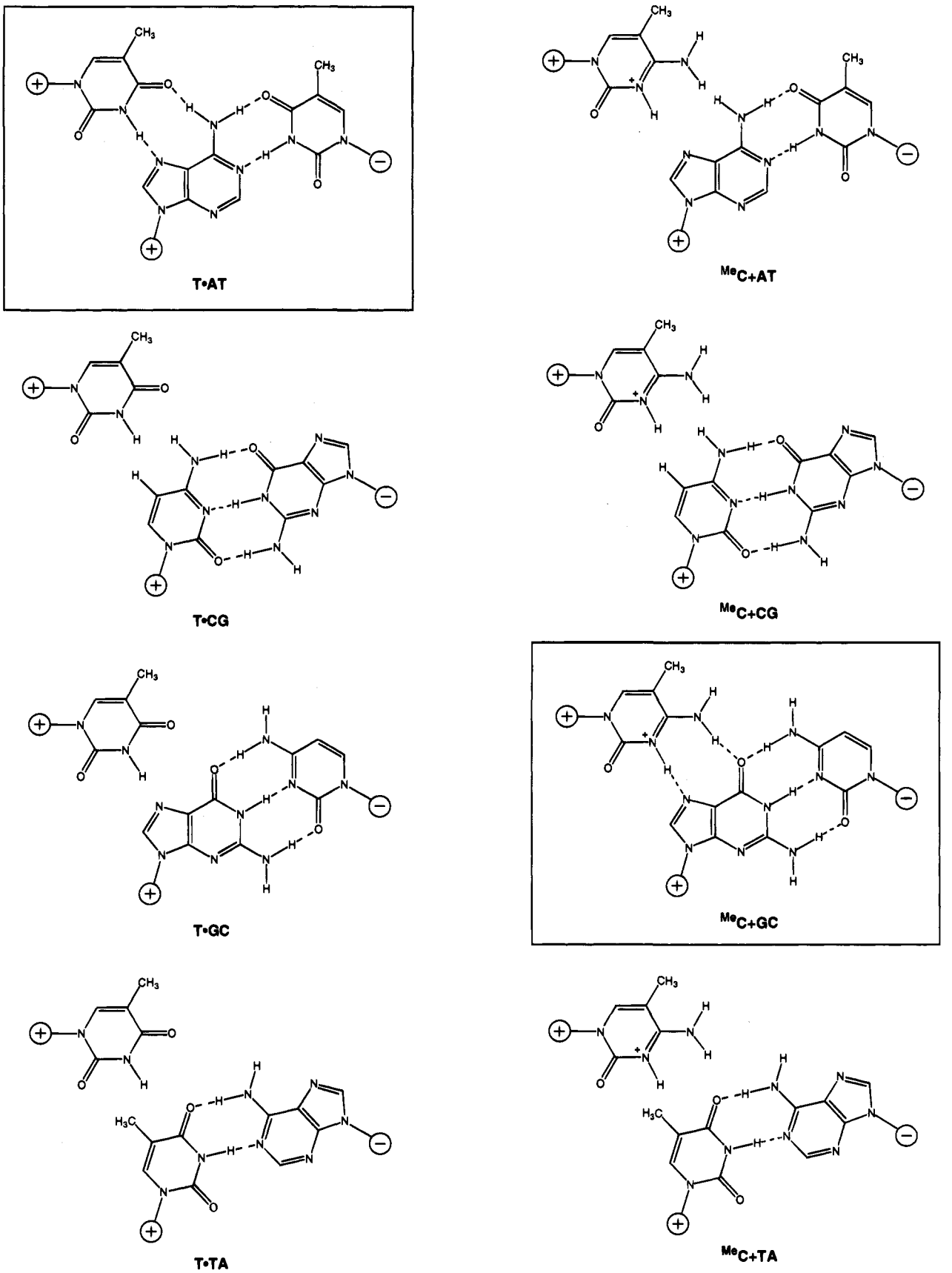


Figure 3. Structures of the T · XY and ^{Me}C · XY base triplets studied. The plus and minus signs indicate the relative strand polarity.

precipitate the DNA. The pellet was resuspended in 100 μL of 0.5× TE, pH 8.0. The resulting solution was extracted three times with phenol and once with 24:1 chloroform–isoamyl alcohol, and desalted on a NICK column using 0.5× TE. The DNA was precipitated with 50 μL of 3 M NaCl, pH 4.0, and 1000 μL of ethanol, resuspended in 100 μL of 0.5× TE and desalted again with a NICK column in 0.5× TE. The DNA was counted and stored at −20 °C.

Quantitative Affinity Cleavage Titrations.[24,29,30] Cleavage experiments were performed in sets of 18 microcentrifuge tubes. For each set of cleavage reactions, a cocktail containing 160 μL of 5× reaction

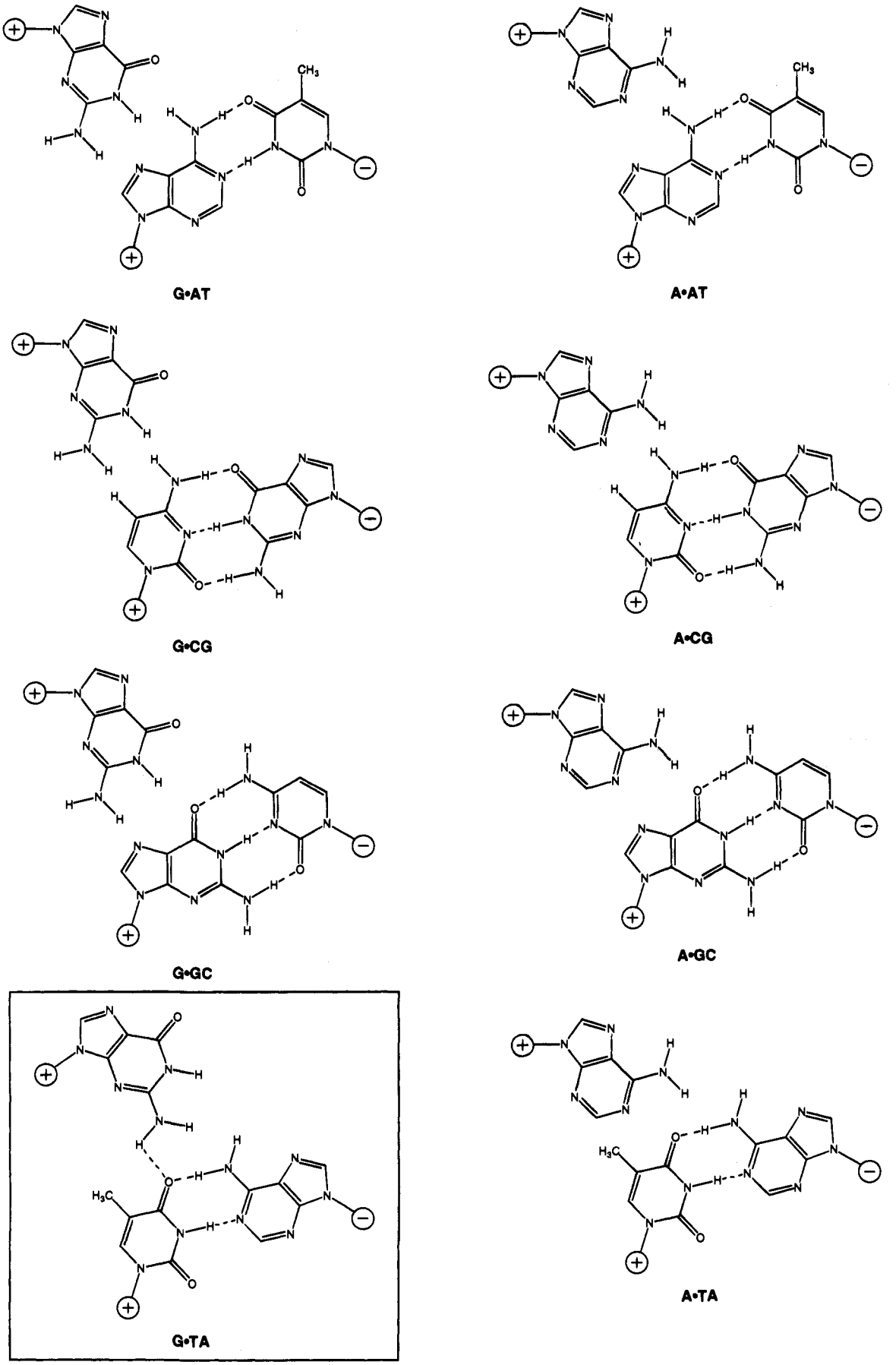


Figure 4. Structure of the G·XY and A·XY base triplets studied. The plus and minus signs indicate the relative strand polarity.

buffer (500 mM NaCl; 50 mM Bis-tris, pH 7.0; 5 mM Spermine), 40 μL of calf thymus DNA (2.0 mM bp), labeled restriction fragment (≈600 000 cpm), and sufficient water to make the total volume 640 μL was prepared. The cocktail was distributed 32 μL per tube, and

necessary additional water added to the tubes. The dried aliquot of oligonucleotide was resuspended in 50 μL of 250 mM $Fe(NH_4)_2(SO_4)_2 \cdot 6H_2O$ and allowed 15 min for equilibration. The oligonucleotide was then serially diluted and distributed to give final probe concentrations of 10 μM, 5 μM, 2.5 μM, 1 μM, 500 nM, 250 nM, 100 nM, 50 nM, 25 nM, 10 nM, 5 nM, 2.5 nM, 1 nM, 500 pM, and 250 pM. The final equilibration volume was 40 μL. Final conditions were 100 mM NaCl, 25 mM Fe(II), 10 mM bis-tris · HCl, pH 7.0, 250 μM spermine. After a 24 h equilibration, cleavage reactions were initiated by the addition of 4 μL of 40 mM dithiothreitol (DTT). Cleavage reactions were allowed to proceed for 6 h and then terminated by the addition of 2 μL of glycogen (2 mg/mL), 4 μL of 3 M NaOAc, 0.1 M $MgCl_2$, pH 5.0, and 115 μL of ethanol. The DNA was precipitated at −20 °C for 30 min and centrifuged for 30 min at 4°C. The supernatant was removed and 60 μL of 50 °C water was added. The samples were then frozen and lyophilized. The DNA was resuspended in 5 μL of formamide loading buffer (4:1 formamide/10× TE) and counted. The samples were heat denatured (90 °C for 3 min) and loaded on a pre-electrophoresed polyacrylamide gel (8% polyacrylamide, 1:19 cross linked, 7 M urea, 0.4 mm thick). After loading, residual radioactivity in the microfuge tubes was counted. The polyacrylamide sequencing gels were run at ≈40 V/cm. When completed, the gel was transferred to filter paper, dried, and exposed to a storage phosphor screen.

Construction of Equilibrium Titration Isotherms. The autoradiography data were analyzed as described previously[5] to obtain a set of $[\mathbf{O}]_{tot}$, I_{site} points. A theoretical binding curve was fit to these data using eq 1, with K_T and I_{sat} (the apparent maximum cleavage) as adjustable parameters:

$$I_{fit} = I_{sat}\left(\frac{K_T[\mathbf{O}]_{tot}}{1 + K_T[\mathbf{O}]_{tot}}\right) \quad (1)$$

where $[\mathbf{O}]_{tot}$ is the total concentration of the oligonucleotide probe. All data from a gel were used unless visual inspection revealed a flaw in the gel at either the target or reference sites, or the I_{site} value for a single data point was more than two standard errors away from the data points on either side. Data from experiments for which less than 80% of the lanes were usable were discarded. The goodness of fit of the binding curve was judged by the reduced χ^2 criterion.[28] Fits with values of $\chi^2_n > 1.5$ were discarded. Fits which met this criterion had correlation coefficients greater than 0.95.

Repeat experiments for a particular triplet used different serial dilutions of oligonucleotide prepared from a unique aliquot. Uniquely prepared batches of 3′- end labeled DNA, buffer, and calf thymus DNA were used. All K_T values reported in the text or tables are the means of three experimental observations plus or minus the standard error of the mean.

Acknowledgment. We are grateful to the National Institutes of Health (GM 35724) for support and the Fannie and John Hertz Foundation for a predoctoral fellowship to G.C.B. We thank S. Priestley for careful reading of the manuscript.

JA9430107

(28) Bevington, P. R. *Data Reduction and Error Analysis for the Physical Sciences*; McGraw-Hill: New York, 1969.

(29) For published representative examples of primary data see refs 5 and 7c.

(30) For triple helical studies we typically add calf thymus DNA as carrier. It is assumed that there is no significant interaction of the oligonucleotide with the "carrier" DNA which is present in large excess. We have explicitly tested this assumption and demonstrated in controls that the differences in the association constants obtained in the presence and absence of carrier DNA are within experimental error.[9] This is likely due to the remarkably high specificity of a 15 base pair binding ligand.