Cooperative Triple-Helix Formation at Adjacent DNA Sites: Sequence Composition Dependence at the Junction

Natalia Colocci and Peter B. Dervan*

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

Received December 8, 1994[⊗]

Abstract: The energetics of cooperative binding by oligodeoxyribonucleotides to adjacent sites by triple helix formation have been determined as a function of sequence composition at the junction. The binding affinity of an 11-mer in the presence of a neighboring bound oligonucleotide is enhanced by factors of 12, 17, 61, and 127 when 5′-TT-3′, 5′-^{m}C^{m}C-3′, 5′-T^{m}C-3′, and 5′-^{m}CT-3′ stacks, respectively, are formed at the junction (10 mM Bis-Tris·HCl at pH 7.0, 10 mM NaCl, 250 μM spermine, 24 °C). These binding enhancements correspond to interaction energies between the two oligonucleotides of 1.5, 1.7, 2.5, and 2.9 kcal·mol^{-1}, respectively. The energetic penalties for a single-base mismatch differ depending on sequence and the location of the mismatch with respect to the 5′- or 3′-side of the junction. In the case of a 5′-TT-3′ stack, a T·GC mismatch on the 5′-side of the junction decreases the interaction energy from 1.5 to 0.6 kcal·mol^{-1}, whereas a T·GC mismatch on the 3′-side destroys cooperativity. For a 5′-^{m}CT-3′ stack, a ^{m}C·AT mismatch on the 5′-side of the junction decreases the cooperative interaction energy from 2.9 to 1.7 kcal·mol^{-1}, whereas a T·GC mismatch on the 3′-side of the junction destroys cooperativity. Two 11-mer oligonucleotides interacting through a 5′-TT-3′ stack binding to adjacent sites on DNA are significantly more sensitive to single-base mismatches than the corresponding 22-mer binding to the same two abutting sites.

Cooperative interactions between proteins bound to DNA are essential for the regulation of gene expression. These cooperative interactions serve to increase the sequence specificity of DNA-binding proteins as well as the sensitivity of the binding equilibrium to concentration changes.[1] When cooperative interactions occur between two (or more) oligonucleotides bound at neighboring sites on double-helical DNA, the specific binding of each oligonucleotide is enhanced.[2,3] Pyrimidine oligodeoxyribonucleotides 11 nucleotides (nt) in length are known to bind cooperatively to abutting sites on double-helical DNA by triple-helix formation, resulting in a 1.8 kcal·mol^{-1} interaction energy between the oligonucleotides.[2b] This interaction energy likely arises from π-stacking between the thymines at the triple-helical junction. An increase in cooperativity between triple-helix-forming oligonucleotides was achieved by the use of modified bases such as 5-(1-propynyl)-2′-deoxyuridine which allow greater stacking interactions in the third strand.[2c,4] Even stronger cooperative interactions were obtained by the addition of dimerization domains to oligonucleotides such as those capable of forming a short Watson−Crick helix or intercalators.[3,5]

The nature of the cooperative interactions between triple-helix-forming oligonucleotides binding to adjacent sites on DNA is still not well understood. A one-base-pair (1-bp) gap between the adjacent triple helix sites *abolishes cooperativity*, suggesting that cooperative interactions largely arise from *stacking* between the bases at the triplex junction.[2] On the basis of these results, the interaction energy between cooperatively binding oligonucleotides should depend on the sequence of the base stack at the junction. Here we report the energetics of cooperative binding of pyrimidine oligodeoxyribonucleotides to adjacent sites on DNA by triple-helix formation as a function of sequence composition at the junction (base stacks 5′-TT-3′, 5′-^{m}CT-3′, 5′-T^{m}C-3′, and 5′-^{m}C^{m}C-3′). The equilibrium association constants of four oligonucleotides binding to four 11-bp sites paired as four different neighbors are measured in the absence and presence of a second 11-mer oligonucleotide occupying an adjacent, directly abutting site by quantitative affinity cleavage titrations (QACT).[6] The effect of a *single-base mismatch at the junction* on the cooperative energy between two triple-helix-forming oligonucleotides reveals that hydrogen bonding between the third strand and the target site is important for proper base stacking in the third strand. Finally, we determine the equilibrium association constant of a 22-mer oligonucleotide targeted to two abutting 11-bp sites in the presence and absence of a single-base mismatch. This data allows us to compare the sensitivity to single-base mismatches of oligonucleotides binding cooperatively to DNA with that of a longer oligonucleotide binding noncooperatively to the same two abutting sites on DNA.

Results

A system was designed to allow determination of the interaction energies between two oligonucleotides when a 5′-

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

(1) (a) Ptashne, M. *A Genetic Switch*; Blackwell Scientific Publications and Cell Press: Palo Alto, CA, 1986. (b) Cantor, C. R.; Schimmel, P. R. *Biophysical Chemistry Part III: The Behavior of Biological Macromolecules*; W. H. Freeman and Co.: New York, NY, 1980. (c) Hill, T. L. *Cooperativity Theory in Biochemistry: Steady State and Equilibrium Systems*; Springer-Verlag: New York, NY, 1985. (d) Adhya, S. *Annu. Rev. Genet.* **1989**, *23*, 227.

(2) (a) Strobel, S. A.; Dervan, P. B. *J. Am. Chem. Soc.* **1989**, *111*, 6956. (b) Colocci, N.; Distefano, M. D.; Dervan, P. B. *J. Am. Chem. Soc.* **1993**, *115*, 4468. (c) Colocci, N.; Dervan, P. B. *J. Am. Chem. Soc.* **1994**, *116*, 785.

(3) (a) Distefano, M. D.; Shin, J. A.; Dervan, P. B. *J. Am. Chem. Soc.* **1991**, *113*, 5901. (b) Distefano, M. D.; Dervan, P. B. *J. Am. Chem. Soc.* **1992**, *114*, 11006. (c) Distefano, M. D.; Dervan, P. B. *Proc. Natl. Acad. Sci. U.S.A.* **1993**, *90*, 1179.

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

(5) Mergny, J. L.; Duval-Valentin, G.; Nguyen, C. H.; Perrouault, L.; Faucon, B.; Rougée, M.; Montenay-Garestier, T.; Bisagni, E.; Helène, C. *Science* **1992**, *256*, 1681.

(6) (a) Singleton, S. F.; Dervan, P. B. *J. Am. Chem. Soc.* **1992**, *114*, 6956. (b) Singleton, S. F.; Dervan, P. B. *Biochemistry* **1992**, *31*, 10995. (c) Singleton, S. F.; Dervan, P. B. *Biochemistry* **1993**, *32*, 13171. (d) Singleton, S. F.; Dervan, P. B. *J. Am. Chem. Soc.* **1994**, *116*, 10376.

0002-7863/95/1517-4781$09.00/0

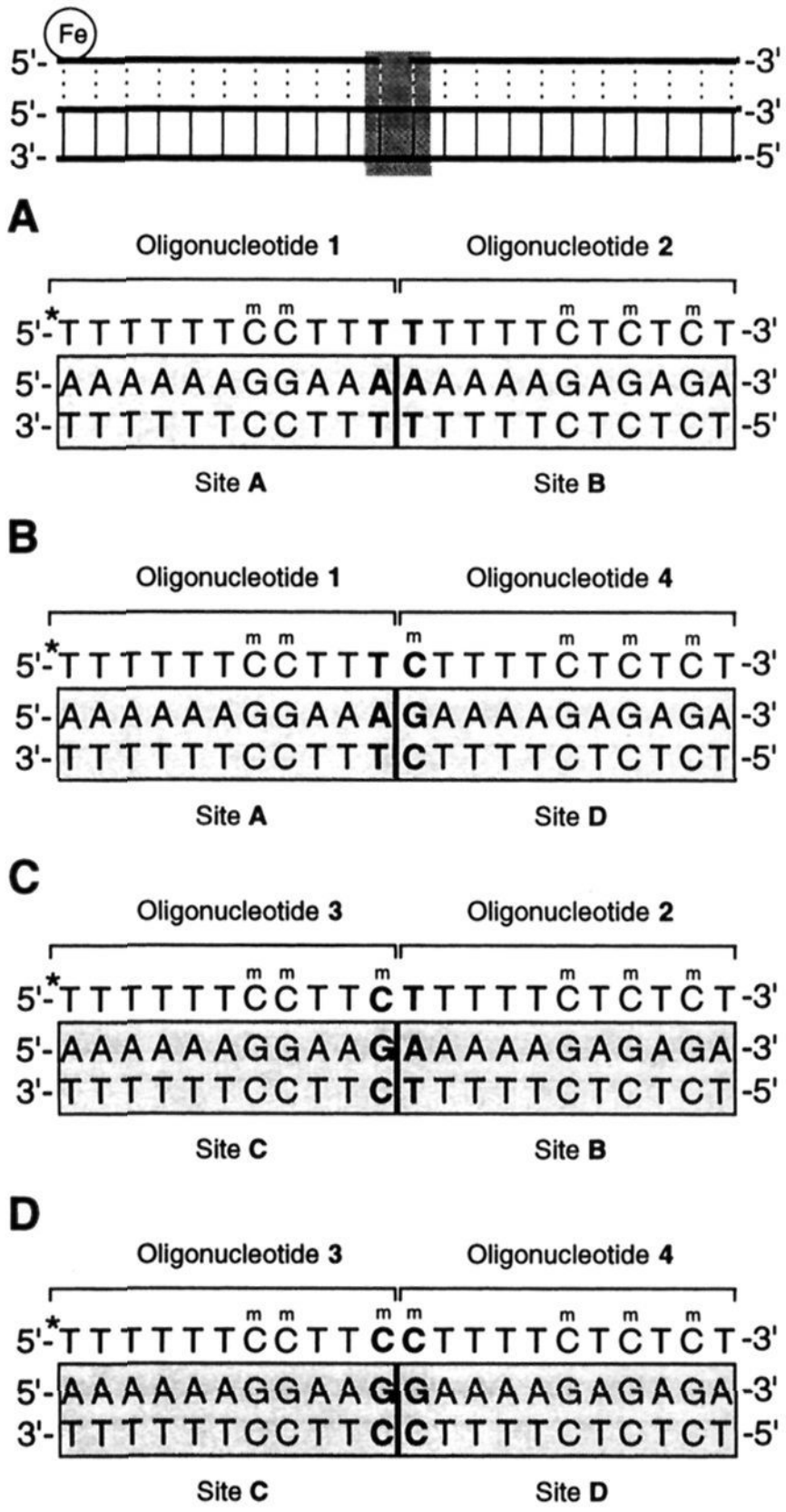

Figure 1. Schematic representation of a complex composed of two triple-helix-forming oligonucleotides binding at adjacent sites on double-helical DNA wherein a 5′-TT-3′ (A), 5′-T^{m}C-3′ (B), 5′-^{m}CT-3′ (C), or a 5′-^{m}C^{m}C-3′ (D) stack is formed. Thick solid lines represent the DNA backbone of the target site and associated oligonucleotides. Thin solid lines represent Watson−Crick hydrogen bonds while dashed lines indicate Hoogsteen hydrogen bonds. Binding of each oligonucleotide is assessed by affinity cleavage.[6]

TT-3′, 5′-T^{m}C-3′, 5′-^{m}CT-3′, or 5′-^{m}C^{m}C-3′ stack is formed at the junction (Figure 1). Two 11-nt oligodeoxyribonucleotides, 5′-d(*TTTTTT^{m}C^{m}CTTT)-3′ (**1**) or 5′-d(*TTTTTT^{m}C^{m}CTT^{m}C)-3′ (**3**) and 5′-d(TTTTT^{m}CT^{m}CT^{m}CT)-3′ (**2**) or 5′-d(^{m}CTTTT^{m}-CT^{m}CT^{m}CT)-3′ (**4**) bind site-specifically to adjacent sites A (or C) and B (or D) on an 850-bp 3′-^{32}P-end-labeled duplex DNA fragment (24 °C, pH 7.0). Thymidine−EDTA (T*)[7,8] was incorporated at the 5′-terminus of oligonucleotides **1** and **3** to allow determination of their equilibrium association constants to sites A and C, respectively, in the absence or presence of **2** or **4**, using the QACT method.[6] A model which relates these experimental data to the cooperative interaction energy between oligonucleotides is described below.[2b]

Heterodimeric Cooperative Systems. Hill has derived a grand partition function (ξ) for a heterodimeric cooperative system in which two different ligands, L_A and L_B, bind uniquely to two distinct sites and each ligand perturbs the binding equilibrium at the noncognate site:[1c]

$$\xi = 1 + K_A[L_A] + K_B[L_B] + yK_A[L_A]K_B[L_B] \quad (1)$$

where K_A and K_B are equilibrium association constants for L_A and L_B binding in the absence of each other and y is related to the interaction energy, E_{coop}, between L_A and L_B by

$$y = \exp(-E_{coop}/RT) \quad (2)$$

The fractional occupancy of site A by L_A is given by

$$\theta_A = [L_A]\frac{\partial \ln \xi}{\partial [L_A]} = \frac{\Phi K_A[L_A]}{1 + \Phi K_A[L_A]} \quad (3)$$

where

$$\Phi = \frac{1 + yK_B[L_B]}{1 + K_B[L_B]} \quad (4)$$

Equations 2 and 4 can be combined and rearranged to give the following expression for the interaction energy:

$$E_{coop} = -RT \ln \frac{\Phi(1 + K_B[L_B]) - 1}{K_B[L_B]} \quad (5)$$

This model requires that two conditions be satisfied: the ligands do not interact in solution, and under the experimental conditions chosen, L_A does not bind to site B and L_B does not bind to site A. These conditions were satisfied under the conditions chosen, and therefore, the model could be reliably used to determine E_{coop}. For the cooperatively binding oligonucleotides described here, L_A is oligonucleotide **1** or **3** and K_A is the association constant for **1** or **3** to its respective site. Similarly, L_B is oligonucleotide **2** or **4** and K_B is the association constant for **2** or **4** to its respective site. Comparison of the Langmuir expression (see the Experimental Section, eq 7) and eq 3 shows that the association constant of ligand A in the presence of a constant concentration of neighboring ligand B, K_{A+B}, in the former equals ΦK_A in the latter. Thus, $\Phi = K_{A+B}/K_A$. For ease of discussion, duplex DNA containing a 5′-AA-3′, 5′-AG-3′, 5′-GA-3′, or 5′-GG-3′ sequence at the triplex junction site will be referred to as the AA, AG, GA, or GG DNA fragment, respectively.

Cooperative Binding of Two 11-mers in the Presence of a 5′-TT-3′ Stack. Analysis of cleavage data performed on the AA DNA fragment yielded equilibrium association constants of 1.3×10^6 M^{-1} for **1** binding alone ($K_{1,AA}$) and 1.5×10^7 M^{-1} for **1** binding in the presence of 5.0 μM **2** ($K_{1+2,AA}$) (Table 1, Figure 1A). The equilibrium association constant of **5**, an oligonucleotide with the same sequence as **2** but containing T* at the 3′-end targeted to the 11-bp site B, was found to be 2.8×10^6 M^{-1}. The binding enhancement $K_{1+2,AA}/K_{1,AA}$ is a factor of 12. Substitution of the values 12, 2.8×10^6 M^{-1}, and 5×10^{-6} M^{-1} for Φ, K_B, and $[L_B]$, respectively, in eq 5 affords a value of 1.5 kcal·mol^{-1} for E_{coop} between oligonucleotides **1** and **2**.

(7) For syntheses of several modified thymidines with a short linker EDTA at the 5 position, see: (a) Dreyer, G. B.; Dervan, P. B. *Proc. Natl. Acad. Sci. U.S.A.* **1985**, *82*, 968. (b) Han, H.; Dervan, P. B. *Nucleic Acids Res.* **1994**, *22*, 2837. (c) Glen Research Corp., 44901 Falcon Place, Sterling, VA 22170.

(8) Thymidine-EDTA is designated T*. For details of linker composition see the Experimental Section.[7]

Table 1. Equilibrium Association Constants for Two Triple-Helix-Forming Oligonucleotides Binding at Adjacent Sites on DNA as a Function of Four Different Base Stacks[a]

stack	oligonucleotide	K (M^{-1})	K_{11+11}/K_{11}	E_{coop} (kcal·mol^{-1})
5′-TT-3′	**1**	1.3 (±0.2) × 10^6	12 (±2)	1.5 (±0.1)
	1 + **2** (5 μM)	1.5 (±0.2) × 10^7		
5′-T<u>C</u>-3′	**1**	4.6 (±0.8) × 10^6	61 (±12)	2.5 (±0.2)
	1 + **4** (5 μM)	2.8 (±0.3) × 10^8		
5′-<u>C</u>T-3′	**3**	1.1 (±0.1) × 10^6	127 (±21)	2.9 (±0.1)
	3 + **2** (5 μM)	1.4 (±0.2) × 10^8		
5′-<u>CC</u>-3′	**3**	1.9 (±0.7) × 10^6	17 (±6)	1.7 (±0.2)
	3 + **4** (5 μM)	3.2 (±0.3) × 10^7		

[a] Values reported in the table are mean values measured from affinity cleavage titration experiments performed in association buffer (10 mM Bis-Tris·HCl, 10 mM NaCl, 250 μM spermine, pH 7.0, 24 °C). <u>C</u> designates 5-methyl-2′-deoxycytidine.

Cooperative Binding of Two 11-mers in the Presence of a 5′-T^{m}C-3′ Stack. Analysis of cleavage data performed on the AG DNA fragment yielded equilibrium association constants of 4.6 × 10^6 M^{-1} for **1** binding alone ($K_{1,AG}$) and 2.8 × 10^8 M^{-1} for **1** binding in the presence of 5.0 μM **4** ($K_{1+4,AG}$) (Table 1, Figure 1B). The equilibrium association constant of **6**, an oligonucleotide with the same sequence as **4** but containing T* at the 3′-end targeted to the 11-bp site B, was found to be 5.4 × 10^6 M^{-1}. The binding enhancement $K_{1+4,AG}/K_{1,AG}$ is a factor of 61. Substitution of the values 61, 5.4 × 10^6 M^{-1}, and 5 × 10^{-6} M^{-1} for Φ, K_B, and [L_B], respectively, in eq 5 affords a value of 2.5 kcal·mol^{-1} for E_{coop} between oligonucleotides **1** and **4**.

Cooperative Binding of Two 11-mers in the Presence of a 5′-^{m}CT-3′ Stack. Analysis of cleavage data performed on the GA DNA fragment yielded equilibrium association constants of 1.1 × 10^6 M^{-1} for **3** binding alone ($K_{3,GA}$) and 1.4 × 10^8 M^{-1} for **3** binding in the presence of 5.0 μM **2** ($K_{3+2,GA}$) (Table 1, Figure 1C). The equilibrium association constant of **5** was found to be 4.3 × 10^6 M^{-1}. The binding enhancement $K_{3+2,GA}/K_{3,GA}$ is a factor of 127. Substitution of the values 127, 4.3 × 10^6 M^{-1}, and 5 × 10^{-6} M^{-1} for Φ, K_B, and [L_B], respectively, in eq 5 affords a value of 2.9 kcal·mol^{-1} for E_{coop} between oligonucleotides **3** and **2**.

Cooperative Binding of Two 11-mers in the Presence of a 5′-^{m}C^{m}C-3′ Stack. Analysis of cleavage data performed on the GG DNA fragment yielded equilibrium association constants of 1.9 × 10^6 M^{-1} for **3** binding alone ($K_{3,GG}$) and 3.2 × 10^7 M^{-1} for **3** binding in the presence of 5.0 μM **4** ($K_{3,4,GG}$) (Table 1, Figure 1D). The equilibrium association constant of **6** was found to be 7.1 × 10^6 M^{-1}. The binding enhancement $K_{3+4,GG}/K_{3,GG}$ is a factor of 17. Substitution of the values 17, 7.1 × 10^6 M^{-1}, and 5 × 10^{-6} M^{-1} for Φ, K_B, and [L_B], respectively, in eq 5 affords a value of 1.7 kcal·mol^{-1} for E_{coop} between oligonucleotides **3** and **4**.

Binding of a 22-mer Oligonucleotide Targeted to Two Abutting 11-bp Sites on DNA. The equilibrium association constant of a 22-mer oligonucleotide with sequence 5′-d(TTTTTT^{m}C^{m}CTTTTTTTT^{m}CT^{m}CT^{m}CT*)-3′ targeted to sites A and B (Figure 1A) was found to be 1.4 × 10^8 M^{-1}. The affinity of the 22-mer is therefore ca. 1 order of magnitude higher than that of 11-mer oligonucleotide **1** binding to site A in the presence of **2** bound to site B.

Effect of Mismatches on a 5′-TT-3′ Stack. Due to the right-handed nature of DNA, we would expect a hydrogen-bond mismatch to affect a stack at the triplex junction differently depending on whether the mismatch is on the 5′- or 3′-side of the junction. We investigated the effect of mismatches on both sides of the junction on cooperativity. To determine the effect

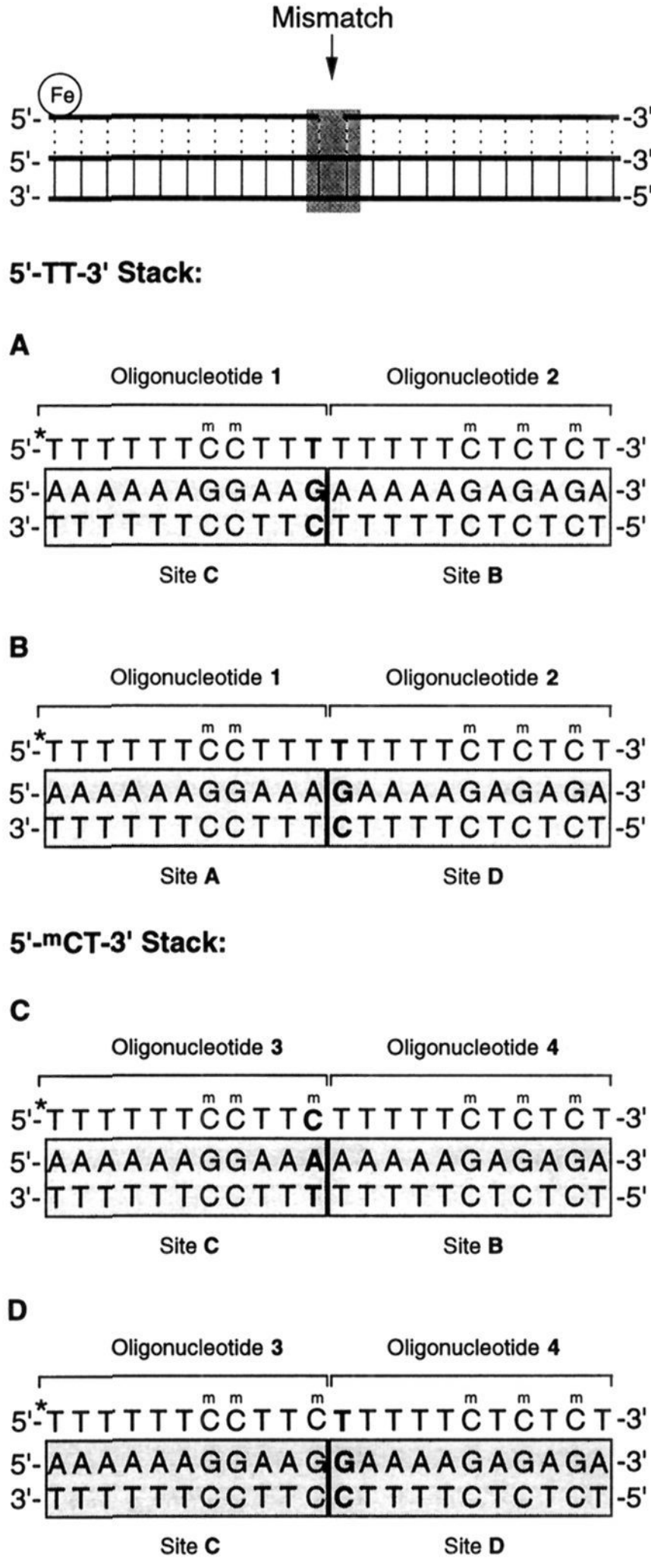

Figure 2. Schematic representation of a complex composed of two triple-helix-forming oligonucleotides binding at adjacent sites on double-helical DNA in the presence of a single-base mismatch at the triplex junction. (A) A 5′-TT-3′ stack is formed at the junction in the presence of a T·GC mismatch on the 5′-side of the junction. (B) A 5′-TT-3′ stack is formed at the junction in the presence of a T·GC mismatch on the 3′-side of the junction. (C) A 5′-^{m}CT-3′ stack is formed at the junction in the presence of a ^{m}C·AT mismatch on the 5′-side of the junction. (D) A 5′-^{m}CT-3′ stack is formed at the junction in the presence of a T·GC mismatch on the 3′-side of the junction. Thick solid lines represent the DNA backbone of the target site and associated oligonucleotides. Thin solid lines represent Watson–Crick hydrogen bonds while dashed lines indicate Hoogsteen hydrogen bonds. Binding of the oligonucleotides is assessed by affinity cleavage.[6]

Table 2. Effect of Mismatches on a 5′-TT-3′ Stack: Equilibrium Association Constants for Two Triple-Helix-Forming Oligonucleotides Binding at Adjacent Sites on DNA in the Presence of a Mismatch at the Junction[a]

mismatch	oligonucleotide	K (M^{-1})	K_{11+11}/K_{11}	E_{coop} (kcal·mol^{-1})
5′-**T** T-3′	**1**	2.1 (±0.2) × 10^5	2.5 (±0.8)	0.6 (±0.2)
5′-GA-3′	**1** + **2** (5 μM)	5.2 (±1.2) × 10^5		
3′-CT-5′				
5′-TT-3′	**1**	4.6 (±0.8) × 10^6	1.0 (±0.9)	0
5′-AG-3′	**1** + **2** (5 μM)	5.3 (±0.6) × 10^6		
3′-TC-5′				

[a] Values reported in the table are mean values measured from affinity cleavage titration experiments performed in association buffer (10 mM Bis-Tris·HCl, 10 mM NaCl, 250 μM spermine, pH 7.0, 24 °C).

Table 3. Effect of Mismatches on a 5′-^{m}CT-3′ Stack: Equilibrium Association Constants for Two Triple-Helix-Forming Oligonucleotides Binding at Adjacent Sites on DNA in the Presence of a Mismatch at the Junction[a]

mismatch	oligonucleotide	K (M^{-1})	K_{11+11}/K_{11}	E_{coop} (kcal·mol^{-1})
5′-CT-3′	**3**	1.3 (±0.1) × 10^5	17 (±3)	1.7 (±0.1)
5′-AA-3′	**3** + **2** (5 μM)	2.2 (±0.4) × 10^6		
3′-TT-5′				
5′-CT-3′	**3**	1.9 (±0.7) × 10^6	1.0 (± 0.5)	0
5′-GG-3′	**3** + **2** (5 μM)	1.3 (±0.1) × 10^6		
3′-CC-5′				

[a] Values reported in the table are mean values measured from affinity cleavage titration experiments performed in association buffer (10 mM Bis-Tris·HCl, 10 mM NaCl, 250 μM spermine, pH 7.0, 24 °C). C designates 5-methyl-2′-deoxycytidine.

of a mismatch on the 5′-side of the junction on cooperative interactions involving a 5′-TT-3′ stack, we measured the association constants for oligonucleotide **1** binding to site A alone ($K_{1,GA}$) and in the presence of 5.0 μM **2** ($K_{1+2,GA}$) on the noncomplementary GA DNA fragment (Figure 2A). In this case, a T·GC mismatch is formed at the 3′-end of oligonucleotide **1**. We found that the binding enhancement $K_{1+2,GA}/K_{1,GA}$ is a factor of 2.5 and little cooperativity was observed (Table 2). Analysis of Table 2 shows that the association constant of **1** binding in the presence of **2** is lowered by a factor of 29 from 1.5 × 10^7 M^{-1} in the absence of any mismatch to 5.2 × 10^5 M^{-1} in the presence of a mismatch.[9a] To determine the effect of a mismatch on the 3′-side of the junction on a 5′-TT-3′ stack, we measured the association constants for oligonucleotide **1** binding to site A alone ($K_{1,AG}$) and in the presence of 5.0 μM **2** ($K_{1+2,AG}$) on the noncomplementary AG DNA fragment (Figure 2B). In this case, a T·GC mismatch is formed at the 5′-end of oligonucleotide **3** and no cooperativity was observed (Table 2).[9b]

Effect of Mismatches on a 5′-^{m}CT-3′ Stack. We also investigated the effect of mismatches on cooperative interactions involving the very strong 5′-^{m}CT-3′ stack. To determine the effect of a mismatch on the 5′-side of the junction, we measured the association constants for oligonucleotide **3** binding to site A alone ($K_{3,AG}$) and in the presence of 5.0 μM **4** ($K_{3+4,AG}$) on the noncomplementary AG DNA fragment (Figure 2C). A ^{m}C·AT mismatch is formed at the 3′-end of oligonucleotide **3**. The binding enhancement $K_{3+4,AG}/K_{3,AG}$ is a factor of 17, corresponding to an E_{coop} of 1.7 kcal·mol^{-1} (Table 3).[9c] To determine the effect of a mismatch on the 3′-side of the junction on a 5′-^{m}CT-3′ stack, we measured the association constants for oligonucleotide **3** binding to site A alone ($K_{3,GG}$) and in the presence of 5.0 μM **4** ($K_{3+4,GG}$) on the noncomplementary GG DNA fragment (Figure 2D). In this case, a T·GC mismatch is formed at the 5′-end of oligonucleotide **4** and no cooperativity was observed.[9d] Analysis of Table 3 reveals that the association constant of **3** binding in the presence of **2** is lowered by a factor of 64 from 1.4 × 10^8 M^{-1} in the absence of any mismatch to 2.2 × 10^6 M^{-1} in the presence of the mismatch.

Effect of Mismatches on the Binding of a 22-mer Targeted to Sites A and B. The sensitivity to single-base mismatches of a noncooperative system compared to that of a cooperative one was examined. The sensitivity to single-base mismatches of the cooperative binding of two 11-mers **1** and **2** targeted to sites C and B, respectively (Figure 2A), was compared to the binding of the corresponding 22-mer oligonucleotide with sequence 5′-d(TTTTTT^{m}C^{m}CTTTTTTTT^{m}CT^{m}CT^{m}CT*)-3′ targeted to the 22-bp site (B + C). In both cases, a T·GC mismatch is present at the 3′-end of site C. The association constants of the 22-mer in the absence and presence of the T·GC mismatch are 1.8 × 10^8 and 7.2 × 10^7 M^{-1}, respectively. The association constants of oligonucleotide **1** binding cooperatively in the presence of **2** in the absence and presence of the same mismatch are 1.5 × 10^7 and 5.2 × 10^5 M^{-1}, respectively, as mentioned above.

Discussion

Two pyrimidine oligonucleotides bind cooperatively to adjacent sites on double-helical DNA with an interaction energy ranging from 1.5 to 2.9 kcal·mol^{-1}, depending on the sequence composition at the triplex junction (Table 1). Cooperative interactions may arise from propagation of a conformational transition between adjacent sites and from base stacking of the terminal bases. Previous experiments show that, when the two target sites are separated by one base pair, no cooperativity is observed, suggesting that base stacking is the major contributor to the cooperative interaction energy.[2b] The experimental results reported here are consistent with this model.

Base stacking is dominated by electrostatic, polarization, and dispersion forces.[10] According to optimized potential calculations, dispersion and polarization interactions determine the magnitude of the total interaction energy, whereas electrostatic forces determine orientational preferences.[10d] In an attempt to explain the differences in interaction energies observed when cooperatively binding oligonucleotides interact through different base stacks, the electrostatic potentials of protonated N^1,5-dimethylcytosine (^{Nm}C) and N^1-methylthymine (^{Nm}T), model compounds of the protonated 5-methylcytosine and thymine, respectively, were analyzed using the program Spartan.[11] These calculations do not take into account the change in charge distribution of the bases due to hydrogen bonding to the purine strand. The AM1 electrostatic potentials were calculated at the

(9) (a) The association constant of oligonucleotide **5** binding to site B on the GA plasmid was determined to be 4.3 (± 1.1) × 10^6 M^{-1}. (b) The association constant of oligonucleotide **5** binding to site D on the AG plasmid was determined to be 5.3 (± 0.3) × 10^5 M^{-1}. (c) The association constant of oligonucleotide **5** binding to site B on the AA plasmid was determined to be 2.8 (± 0.1) x 10^6 M^{-1}. (d) The association constant of oligonucleotide **5** binding to site D on the GG plasmid was determined to be 7.1 (± 0.4) x 10^5 M^{-1}.

(10) (a) Sowers, L. C.; Shaw, B. R.; Sedwick, W. D. *Biochem. Biophys. Res. Commun.* **1987**, *148*, 790. (b) DeVoe, H.; Tinoco, I. *J. Mol. Biol.* **1962**, *4*, 500. (c) Sanyal, N. K.; Roychoudhury, M.; Ruhela, K. R.; Tiwari, S. N. *J. Comput. Chem.* **1986**, *8*, 604. (d) Ornstein, R. L.; Rein, R.; Breen, D. L.; MacElroy, R. D. *Biopolymers* **1978**, *17*, 2341. (e) Broom, A. D.; Schweizer, M. P.; Ts'o, P. O. P. *J. Am. Chem. Soc.* **1967**, *89*, 3612. (f) Ts'o, P. O. P. *Molecular Associations in Biology*; Academic Press: New York, NY, 1968; pp 39–75. (g) Nakano, N. I.; Igarishi, S. J. *Biochemistry* **1970**, *9*, 577. (h) Breslauer, K. J.; Frank, R.; Blocker, H.; Marky, L. A. *Proc. Natl. Acad. Sci. U.S.A.* **1986**, *83*, 3746. (i) Delcourt, S. G.; Blake, R. D. *J. Biol. Chem.* **1991**, *266*, 15160. (j) Herskovitz, T. T. *Arch. Biochem. Biophys.* **1962**, *97*, 474. (k) Herskovitz, T. T. *Biochemistry* **1963**, *2*, 335.

(11) Spartan Version 2.0, Wavefunction, Inc., 18401 Von Karman Ave., Suite 210, Irvine, CA 92715.

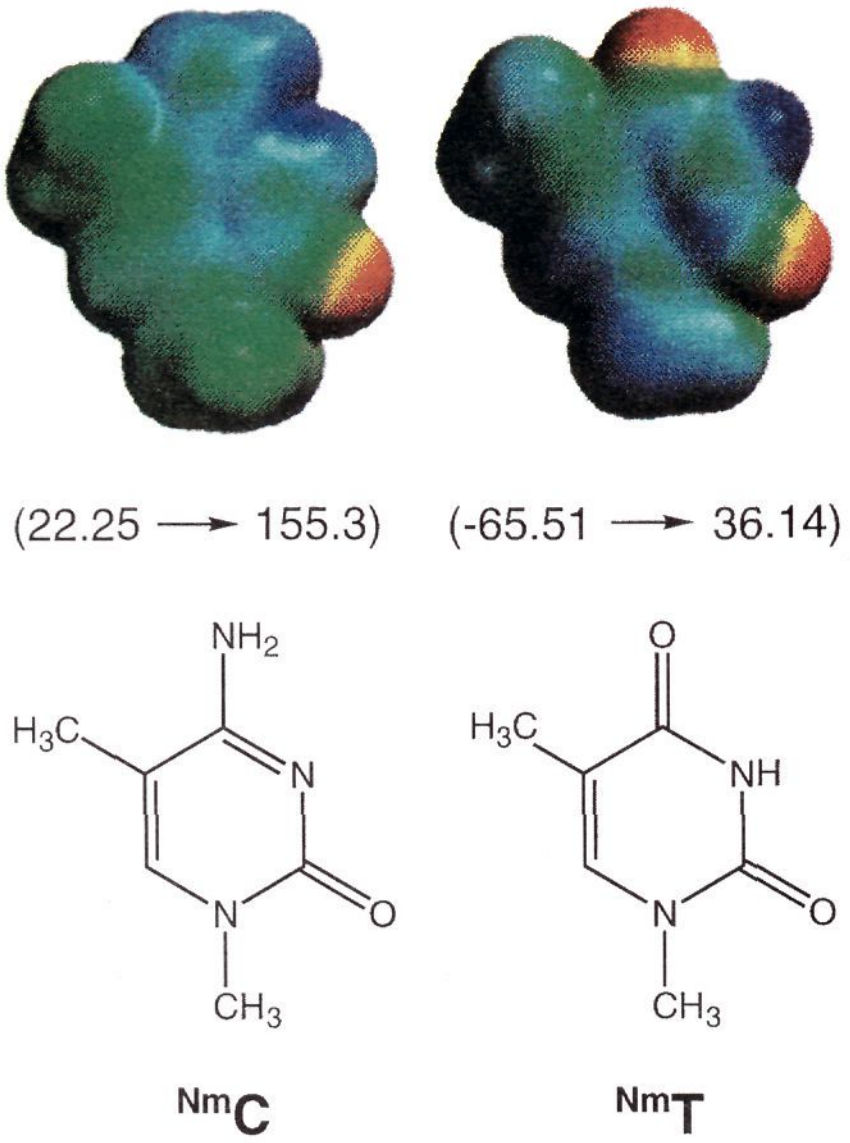

Figure 3. Representation of the calculated AM1 electrostatic potential surfaces of model compounds N^1,5-dimethylcytidine (^{Nm}C) and N^1-methylthymidine (^{Nm}T). The range of electrostatic potential values calculated for each molecule are shown under the surfaces. Electrostatic potentials values are shown colorimetrically and range from 22.25 (red) to 155.3 (blue) kcal·mol^{-1} for ^{Nm}C and from −65.51 (red) to 36.14 (blue) kcal·mol^{-1} for ^{Nm}T.

semiempirical level and mapped onto the electron-density surface of each molecule (Figure 3). Regions of electrostatic potential greater than or equal to 155.3 kcal·mol^{-1} for ^{Nm}C and 36.14 kcal·mol^{-1} for ^{Nm}T are shown in blue, and regions of electrostatic potential less than or equal to 22.25 kcal·mol^{-1} for ^{Nm}C and −65.51 kcal·mol^{-1} for ^{Nm}T are shown in red. The lower and higher ends of the color scale were chosen such as to highlight the charge distributions in the two molecules, and therefore, the scale values in the two molecules are different. On the basis of these scale values, ^{Nm}C (22.25 → 155.3) is more polarized than ^{Nm}T (−65.51 → 36.14). ^{Nm}C contains a region of low electrostatic potential (electron-deficient) localized around the protonated nitrogen and a region of high positive electrostatic potential (electron-rich) localized on the carbonyl oxygen. ^{Nm}T contains an electron-deficient region extending over one-half of its ring opposite to the two carbonyl groups.

The four base stacking configurations 5′-TT-3′, 5′-T^{m}C-3′, 5′-^{m}CT-3′, and 5′-^{m}C^{m}C-3′ drawn on the basis of the structure proposed for DNA triple helix (T·AT)$_n$ are shown in Figure 4.[12] These base stacks are characterized by four different charge−charge interactions, depending on the electrostatic distribution of the individual bases and the relative positioning of the bases with respect to each other. Our data suggests that the weakest interactions are a 5′-^{m}C^{m}C-3′ and a 5′-TT-3′ stack, which are comparable (Table 1). Although we usually depict ^{m}C as protonated when bound to DNA in a triple-helical complex,

(12) (a) Arnott, S.; Bond, P. J.; Selsing, E.; Smith, P. J. C. *Nucleic Acids Res.* **1976**, *3*, 2459. (b) Saenger, W. *Principles of Nucleic Acid Structure*; Springer-Verlag, New York.

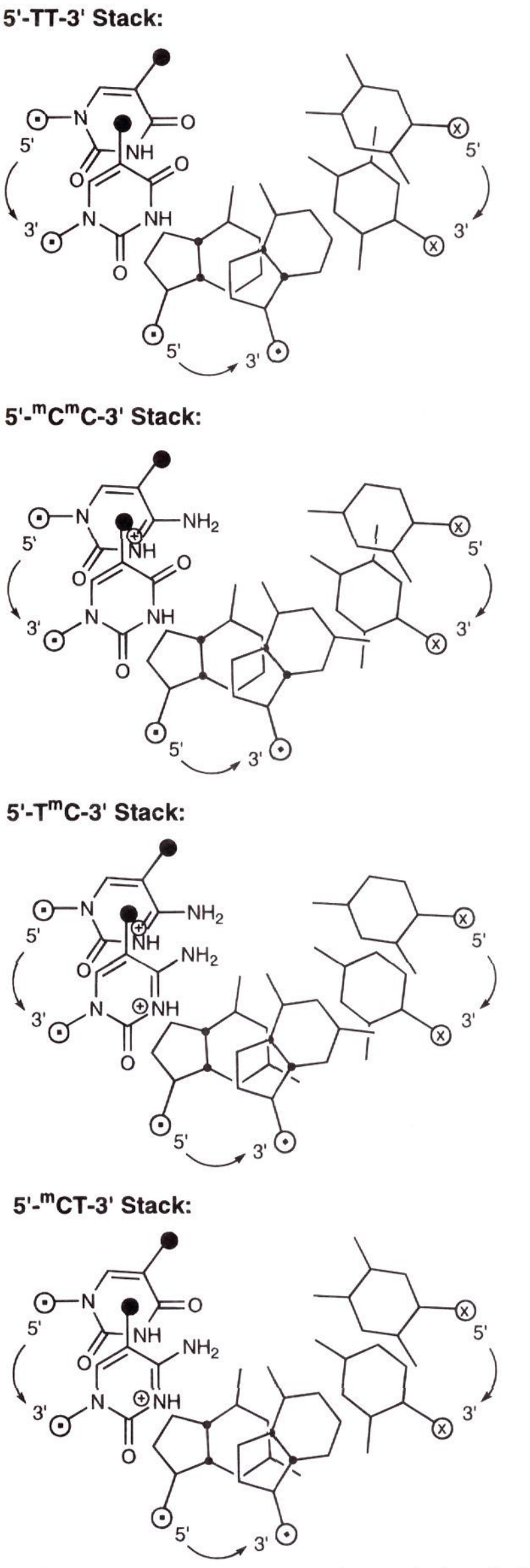

Figure 4. Base-stacking configurations of 5′-^{m}C^{m}C-3′, 5′-TT-3′, 5′-T^{m}C-3′, and 5′-^{m}CT-3′ triples drawn on the basis of the structure proposed for DNA triple-helix (T·AT)$_n$.[12] The C1′ atoms of the deoxyribose sugars are represented by the open circles. The 5′ → 3′ polarity of the strands is indicated by the ⊗ symbol (into the plane of the page) and ⊙ (out of the plane of the page).

under the conditions used (pH 7.0), it may not be protonated at every position. An analysis of the electronic distribution of ^{Nm}C and ^{Nm}T reveals that there are no favorable charge−charge interactions in a 5′-^{m}C^{m}C-3′ or 5′-TT-3′ stack. The 1.5−1.7 kcal·mol^{-1} cooperative energy observed when either a 5′-^{m}C^{m}C-3′ or a 5′-TT-3′ stack is formed may reflect generally favorable polarization interactions between the two bases across the junction.

Stronger interaction energies between two oligonucleotides binding to adjacent sites on DNA are observed when a 5′-T^{m}C-3′ or a 5′-^{m}CT-3′ stack is formed at the junction (Table 1). An analysis of the electronic distribution of ^{Nm}T and ^{Nm}C reveals that, in both 5′-T^{m}C-3′and 5′-^{m}CT-3′ stacks, there may be a favorable charge−charge interaction between the electron-deficient exocyclic amino group on ^{m}C and the electron-rich C4 carbonyl group on T. This interaction may explain why cooperative interactions may be stronger when a 5′-T^{m}C-3′or a 5′-^{m}CT-3′ stack is formed rather than a 5′-TT-3′ or 5′-^{m}C^{m}C-3′ stack. A comparison between the 5′-T^{m}C-3′ and 5′-^{m}CT-3′ stacking configurations themselves in terms of the charge distribution of the individual bases and their interactions reveals information about the relative strength of the two stacks. The data suggest that a 5′-^{m}CT-3′ stack is significantly stronger than a 5′-T^{m}C-3′ stack. In the latter stack, the highly polarized region of ^{m}C faces away from T and no strong charge−charge interactions occur. In contrast, in a 5′-^{m}CT-3′ stack, the highly polarized region of ^{m}C is located directly above T, allowing the electron-deficient region of T to interact favorably with the electron-rich oxygen of ^{m}C. In addition, there may be a favorable interaction between the positively charged ring nitrogen of ^{m}C and the electron-rich C4 carbonyl group of T.

Effect of Mismatches on Cooperativity. On the basis of experimental data, the cooperative interaction energy between two triple-helix-forming oligonucleotides likely arises from base stacking at the triplex junction. The question arises as to what degree base stacking is sensitive to hydrogen-bond mismatches at the junction. A mismatch could presumably disrupt base stacking either by preventing a base from aligning favorably with respect to the other base across the junction, disrupting favorable electrostatic interactions, or by changing the electronic distribution of the mismatched base itself. Experiments were carried out to determine the degree to which cooperativity is disrupted by a single-base mismatch on either side of the triplex junction. The effect of a mismatch on cooperativity was examined in the cases where the moderately strong 5′-TT-3′ stack and the very strong 5′-^{m}CT-3′ stack are formed. In the case of a 5′-TT-3′ stack, a T·GC mismatch on the 3′-side of the junction (Figure 2A) significantly reduces cooperative interactions, whereas a T·GC mismatch on the 5′-side virtually destroys cooperativity (Figure 2B). The disruption of cooperativity may indicate that either the electron distribution or the relative position of T with respect to the adjacent T across the junction does not allow stacking between the two bases. In the case of the strong 5′-^{m}CT-3′ stack, a ^{m}C·AT mismatch on the 5′-side of the junction (Figure 2C) decreases the cooperative interaction energy between oligonucleotides **3** and **2** by 1.7 kcal·mol^{-1}, whereas a T·GC mismatch on the 3′-side of the junction completely destroys cooperativity (Figure 2D). Therefore, in this system, cooperativity is affected differently depending on the nature of the mismatch and/or the location of the mismatch with respect to the junction.

Sensitivity of a Cooperative vs Noncooperative System to Single-Base Mismatches. Analysis of the data shows that the equilibrium association constants of a 22-mer of sequence 5′-d(TTTTTT^{m}C^{m}CTTTTTTTT^{m}CT^{m}CT^{m}CT*)-3′ in the absence and presence of a T·GC mismatch are 1.8×10^8 and 7.2×10^7 M^{-1}, respectively. The association constants of 11-mer oligonucleotide **1** binding cooperatively in the presence of **2** to the same sites in the absence and presence of the same mismatch are 1.5×10^7 and 5.2×10^5 M^{-1}, respectively. Therefore, the same mismatch decreases the binding constant of the 22-mer by a factor of 2, whereas it lowers the association constant of cooperatively binding 11-mer **1** by a factor of 29. This observation suggests that the cooperative system is more specific than the noncooperative one.

Summary. Pyrimidine oligonucleotides bind cooperatively to abutting sites on double-helical DNA with an interaction energy of 1.5−2.9 kcal·mol^{-1} depending on the sequence of the base stack at the triple helix junction. Base stacking appears to be the major source of the positive cooperativity. Knowledge of the relative strength of base stacks in the third strand of a triple helix may serve as a baseline for the design of artificial cooperative systems containing novel structural domains for stabilizing interactions.

Experimental Section

General Procedures. *E. coli* JM110 was obtained from Stratagene. Qiagen plasmid kits were purchased from Qiagen Inc. Sonicated, deproteinized calf thymus DNA was purchased from Pharmacia. Nucleotide triphosphates were obtained from Pharmacia. Nucleotide triphosphates labeled with ^{32}P (≥3000 Ci/mmol) were purchased from Amersham. Restriction endonucleases were purchased from New England Biolabs or Boehringer Mannheim and used according to the suggested protocol in the provided buffer. Thymidine and 5-methyl-2′-deoxycytidine phosphoramidites were obtained from Glen Research. Sephadex resins were obtained from Pharmacia. Sequenase Version 2.0 was obtained from United States Biochemical.

Construction of pMD5152, pNCAG, pNCGA, and pNCGG. These plasmids were prepared by standard methods.[13] Briefly, pMD5152 was prepared by hybridization of two synthetic oligonucleotides, 5′-d(ACGTTCCTAAAAAAGGAAAAAAAAGAGAGAGAGATCTG)-3′ and 5′-d(GATCCAGATCTCTCTCTCTTTTTTTTCCTTTTTTAGGA)-3′, followed by ligation of the resulting duplex with pUC18 DNA previously digested with Hind III and BamH I; this ligation mixture was used to transform *E. coli* JM110 (Stratagene). Plasmid DNA from ampicillin-resistant white colonies was isolated and the presence of the desired insert confirmed by restriction analysis and Sanger sequencing. Preparative isolation of plasmid DNA was performed using a Qiagen plasmid kit. Similar procedures were used to prepare plasmid pNCAG with the oligonucleotides 5′-d(AGCTTCCTAAAAAAGGAAAGAAAAAGAGAGAGAGATCTG)-3′ and 5′-d(GATCCAGATCTCTCTCTCTTTTTCTTTCCTTTTTTAGGA)-3′, plasmid pNCGA with the oligonucleotides 5′-d(AGCTTCCTAAAAAAGGAAGAAAAAAGAGAGAGAGATCTG)-3′ and 5′-d(GATCCAGATCTCTCTCTCTTTTTTCTTCCTTTTTTAGGA)-3′, and plasmid pNCGG with the oligonucleotides 5′-d(AGCTTCCTAAAAAAGGAAGGAAAAAGAGAGAGAGATCTG)-3′ and 5′-d(GATCCAGATCTCTCTCTCTTTTCCTTTCCTTTTTTAGGA)-3′.

Oligonucleotide Preparation. Oligonucleotides were synthesized by standard automated solid-support chemistry on an Applied Biosystems Model 380B DNA synthesizer using *O*-(cyanoethyl) *N,N*-diisopropylphosphoramidites.[14] The phosphoramidite of the Amino-Modifier-dT [5′-(dimethoxytrityl)-5-[[[*N*-(trifluoroacetyl)amino]ethyl]-3-acrylimido]-2′-deoxyuridine, 3′-*O*-(2-cyanoethyl)-*N,N*-diisopropylphosphoramidite] purchased from Glen Research was incorporated at the 5′-terminus of oligonucleotides **1** and **3** during oligonucleotide synthesis. Controlled pore glass derivatized with thymidine-EDTA (T*) was prepared as described by Strobel.[15] All oligonucleotides were deprotected with concentrated ammonium hydroxide at 55 °C for 24 h

(13) Sambrook, J.; Fritsch, E. F.; Maniatis, T. *Molecular Cloning*, 2nd ed.; Cold Spring Harbor Laboratory Press: New York, NY, 1989.

(14) Gait, M. J., Ed. *Oligonucleotide Synthesis: A Practical Approach*; IRL Press: Oxford, U.K., 1984.

(15) Strobel, S. A. Dissertation, California Institute of Technology, Pasadena, CA, 1992.

and dried *in vacuo*. The crude 5′-terminal-dimethoxytrityl-protected oligonucleotides were purified by reverse phase FPLC on a ProRPC 10/10 (C_2–C_8) column (Pharmacia LKB) and a gradient of 0–40% CH_3CN in 0.1 M triethylammonium acetate (pH 7.0) detritylated in 80% AcOH and chromatographed a second time. Oligonucleotides **1** and **3** containing the Amino-Modifier-dT were dissolved in a 0.4 M $NaHCO_3$ solution and added to an Eppendorf tube containing 13.6 mg (50 μmol) of EDTA monoanhydride. The reaction mixtures were vortexed at room temperature for 30 min, filtered through Centrex, and purified by ion-exchange FPLC on a Mono Q 10/10 column (Pharmacia LKB) and a gradient of 0.2–2 M NH_4OAc (pH 7.0). Oligonucleotides **1** and **3** were desalted on Pharmacia NAP-25 columns and dried *in vacuo*. All other oligonucleotides were desalted via repeated resuspensions of the oligonucleotides in water followed by lyophilization. Oligonucleotide concentrations were determined by UV absorbance at 260 nm using extinction coefficients (M^{-1} cm^{-1}) of 8800 (T and T* derivative from Glen Research's Amino-Modifier-dT) and 5700 (^{m}C).

Preparation of Labeled Restriction Fragment. The 3′-^{32}P-labeled duplex was prepared by digestion of the desired plasmid with EcoRI, followed by treatment with [α-^{32}P]dATP and [α-^{32}P]TTP in the presence of Sequenase. To remove nonincorporated radioactivity, the fragment was passed through a Pharmacia NICK column. The DNA was then digested with XmnI, followed by separation of the resulting products on a 5% nondenaturing polyacrylamide gel (19:1 monomer/bis). The gel band corresponding to the desired 850-bp fragment was visualized by autoradiography, excised, crushed, and soaked in 10 mM Tris (pH 8.0), and 20 mM EDTA, at 37 °C for 15–18 h. The resulting suspension was filtered through a 0.45 μm filter, and the eluted DNA present in the supernatant was precipitated with NaOAc/EtOH. The DNA pellet was washed with 70% EtOH, dried *in vacuo*, resuspended in TE buffer (10 mM Tris (pH 8.0), 1 mM EDTA), extracted four times with phenol and twice with 24:1 $CHCl_3$/isoamyl alcohol, and reprecipitated with NaOAc/EtOH. The DNA pellet was washed with 70% EtOH, dried *in vacuo*, resuspended in 10 mM Bis-Tris·HCl and 100 mM NaCl (pH 7.0) at a final activity of 100 000 cpm/μL, and stored at −20 °C.

Quantitative Affinity Cleavage Titrations. In a typical experiment, a 5 nmol aliquot of the desired oligonucleotide–EDTA was dissolved in 50 μL of aqueous 200 μM $Fe(NH_4)_2(SO_4)_2{\cdot}6H_2O$ to produce a solution that was 100 μM in oligonucleotide. This solution was then diluted serially to yield 100 μM, 10 μM, 1 μM, 100 nM, 10 nM, 1 nM, and 100 pM solutions. The appropriate volume of these solutions was then distributed among 14–16 1.5 mL microcentrifuge tubes at the appropriate concentrations. A stock solution containing labeled target DNA, Bis-Tris·HCl buffer, NaCl, calf thymus DNA, the second oligonucleotide where needed, and water was distributed to each reaction tube in 12 μL aliquots, and the appropriate volume of water and 4 μL of 250 μM spermine were added to each tube to bring the total volume to 36 μL. The oligonucleotide–EDTA·Fe(II) and the DNA were allowed to equilibrate for 24 h at 24 °C. After 21 h, 2 μL of 100 μM $Fe(NH_4)_2(SO_4)_2{\cdot}6H_2O$ was added to each tube. The cleavage reactions were initiated by the addition of 2 μL of a 80 mM aqueous DTT solution to each tube. Final reaction conditions in 40 μL of association buffer were 10 mM Bis-Tris·HCl at pH 7.0, 10 mM NaCl, 250 μM spermine, 100 μM bp calf thymus DNA, 4 mM DTT, and approximately 10 000 cpm labeled DNA. The reactions were incubated for 2 h at 24 °C. Cleavage reactions were terminated by NaOAc/EtOH precipitation. The precipitated DNA cleavage products were dissolved in 35 μL of water and dried *in vacuo*. The DNA was then resuspended in 10 μL of formamide-TBE loading buffer and heated at 55 °C for 10 min to effect dissolution. The Cerenkov radioactivity of the solutions was measured with a scintillation counter. The DNA was denatured at 95 °C for 3 min, and 5 μL from each sample was loaded onto a 8% denaturing polyacrylamide gel.

Construction of Titration Binding Isotherms. Gels were exposed to photostimulable storage phosphor imaging plates (Kodak Storage Phosphor Screen S0230 obtained from Molecular Dynamics) in the dark at 24 °C for 8–24 h.[16] A Molecular Dynamics 400S PhosphorImager was used to obtain data from the phosphorimaging screens. Rectangles of the same dimensions were drawn around the cleavage bands at the target and at the references sites. The ImageQuant v. 3.0 program running on an AST Premium 386/33 computer was used to integrate the volume of each rectangle.

Affinity Cleavage Titration Fitting Procedure. A detailed description of the affinity cleavage titration procedure has been published.[6] The relative cleavage efficiencies at the target site for each oligonucleotide concentration were determined by using the following equation:

$$I_{site} = I_{tot} - \lambda I_{ref} \tag{6}$$

where I_{tot} and I_{ref} are the intensities of the cleavage bands at the target site and at the reference site, respectively, and λ is a scaling parameter defined as the ratio I_{tot}/I_{ref} at the lowest oligonucleotide concentration. A theoretical binding curve, represented by eq 7, where I_{sat} is the apparent maximum cleavage, K_i the equilibrium association constant for oligonucleotide i, and $[O]_{tot}$ the oligonucleotide–EDTA concentration, was used to fit the experimental data using I_{sat} and K_i as adjustable parameters:

$$I_{fit} = I_{sat}\frac{K_i[O]_{tot}}{1 + K_i[O]_{tot}} \tag{7}$$

KaleidaGraph software (version 2.1, Abelweck Software) running on a Macintosh IIfx computer was used to minimize the difference between I_{fit} and I_{site} for all data points. All values reported in the text are the means of three to six experimental observations ± SEM, unless otherwise indicated. For graphical representation and comparison, I_{site} values were converted to apparent fractional occupancies (θ_{app}) by dividing I_{site} by I_{sat}.

Electrostatic Potential Calculations. AM1 electrostatic potential surfaces calculations were performed at the semiempirical level using the program Spartan running on a Silicon Graphics computer.

Acknowledgment. We are greatful to the National Institutes of Health for grant support (GM 51747) and for a National Research Service Award (Predoctoral Training in Genome Analysis HG00021-03) to N.C. We thank Patrick C. Kearney of the Dougherty group at Caltech for generous help with electrostatic potential calculations.

JA943969U

(16) Johnston, R. F.; Pickett, S. C.; Barker, D. L. *Electrophoresis* **1990**, *11*, 355.