Selective Recognition of a dC−dG Base Pair by Oligonucleotide-Directed Triplex Formation Using a dC Residue Tethering an Intercalator

Diego A. Gianolio and Larry W. McLaughlin*

Department of Chemistry, Merkert Chemistry Center
Boston College, Chestnut Hill, Massachusetts 02167

Received March 12, 1999

The recognition of double-stranded DNA using a third single-stranded pyrimidine (or purine) oligonucleotide to form a DNA triplex occurs most effectively when the target sequence is a polypurine tract. The interruption of a polypurine target sequence by one or more pyrimidine bases has proven to be detrimental to triplex stability[1−6] presumably in part due to the inability of the Hoogsteen face of pyrimidines to form a bidentate hydrogen-bonding interaction with the third strand. A number of native and analogue nucleosides have been examined for their ability to interact with either T−dA or dC−dG base pair targets. In the simplest case, dG can interact with a T−dA base pair[1,7] and form a single interstrand hydrogen bond.[8] All of the possible base triplets involving the common nucleosides have been examined, but only the dG−T−dA (parallel)[9] and T−dC−dG (antiparallel)[10] base triplets exhibit moderately stabilizing effects. A number of analogue residues in the third strand have also been examined for their ability to target pyrimidine−purine base pairs.[11−16] The most common approach here is to introduce an added functionality such that hydrogen-bonding interactions could occur with the target pyrimidine as well as its partner purine and thus recognize the base pair as a unit.[14,17,18] While these modifications have resulted in moderate selectivity, it is often achieved by destabilization of the complex at nontarget base pairs. In the present work we have examined the use of a simple analogue which can be used to form a single hydrogen bond and through intercalation provide additional hydrophobic interactions.

The design of the analogue began with a dm^5C residue since its N^3-nitrogen appeared to be positioned such that it could take part in hydrogen bond formation with the available N^4−H of the target dC−dG base pair (Figure 1a). To the dm^5C residue was covalently attached a naphthalene-based intercalator through the N^4-amino group. Naphthalene diimides have recently been shown

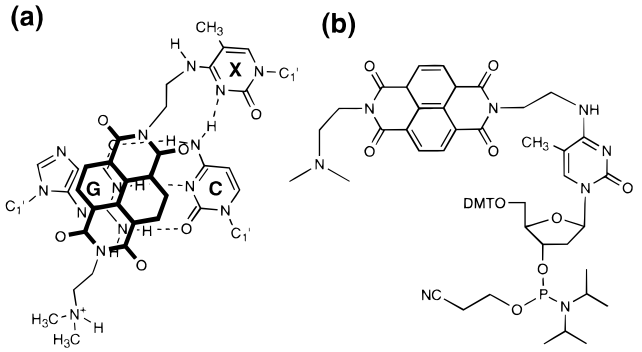

Figure 1. (a) Potential hydrogen bonding and intercalating interactions between the analogue nucleoside in the third strand and the target C−G base pair. (b) Structure of the phosphoramidite derivative of the analogue suitable for DNA synthesis.

to bind effectively to nucleic acids by an intercalation mode.[19−21] The synthesis of the building block was straightforward, beginning with the sulfonation of thymidine essentially as described previously.[22] The unsymmetrical naphthalene diimide, containing both 2-aminoethyl and 2-(*N*,*N*-dimethyl)aminoethyl substituents, was prepared[21] and reacted with the sulfonated thymidine. This reaction generated the critical linkage between the N^4-amino group and the naphthalene derivative. The modified nucleoside was then converted to a phosphoramidite derivative by conventional procedures to yield a compound suitable for DNA synthesis (Figure 1b).

DNA synthesis with the analogue nucleoside proceeded by essentially standard protocols,[23] using thymidine and N^4-acetyl-protected 2′-deoxycytidine building blocks in addition to the analogue. Deprotection of the oligonucleotide was performed in 0.05 M potassium carbonate in methanol. These conditions avoided transamination of the analogue dm^5C residue that was observed in the presence of concentrated ammonia and previously observed for other N^4-substituted cytidine derivatives. However, under these conditions some cleavage of the linker was still apparent. Control experiments indicated that the diimide ring itself was stable to the conditions of deprotection. Loss of the linker presumably occurred as a result of an addition/elimination reaction at C_4. After an initial HPLC isolation based upon a DMT-mediated mobility shift, a second isolation was required to resolve the 15-mer containing the naphthalene diimide linker (62%) from the 15-mer that had lost the linker during deprotection (38%). In a typical synthesis based upon 1 μmol of bound nucleoside, approximately 50 A_{260} units of purified analogue-containing 15-mer could be isolated—roughly 60% of what would be expected in the typical 15-mer synthesis. After purification of the 15-mer sequence containing the analogue in this manner, S1 nuclease/alkaline phosphatase digestion resulted in three peaks. Two early eluting peaks corresponded to dC and T, while the third and later eluting peak corresponded to the dimer 5′-d(XpT). This latter identification was the result of MALDI-TOF mass spectrometry analysis and independent synthesis of the dimer. Analysis of the

(1) Griffin, L. C.; Dervan, P. B. *Science* **1989**, *245*, 967−971.
(2) Mergny, J.; Sun, J.; Rougee, M.; Montenay-Garestier, T.; Barcelo, F.; Chomilier, J.; Helene, C. *Biochemistry* **1991**, *30*, 9791−9798.
(3) Macaya, R. F.; Gilbert, D. E.; Malek, S.; Sinsheimer, J. S.; Feigon, J. *Science* **1991**, *254*, 270−274.
(4) Radhakrishnan, I.; Patel, D. J. *J. Am. Chem. Soc.* **1992**, *114*, 6913−6915.
(5) Yoon, K.; Hobbs, C. A.; Koch, J.; Sardaro, M.; Kutny, R.; Weis, A. L. *Proc. Natl. Acad. Sci. U.S.A.* **1992**, *89*, 3840−3844.
(6) Miller, P. S.; Cushman, C. D. *Biochemistry* **1993**, *32*, 2999−3004.
(7) Radhakrishnan, I.; Gao, X.; de los Santos, C.; Live, D.; Patel, D. J. *Biochemistry* **1991**, *30*, 9022−9030.
(8) Wang, E.; Malek, S.; Feigon, J. *Biochemistry* **1992**, *31*, 4838−4846.
(9) Best, G. C.; Dervan, P. B. *J. Am. Chem. Soc.* **1995**, *117*, 1187−1193.
(10) Dittrich, K.; Gu, J.; Tinder, R.; Hogan, M.; Gao, X. *Biochemistry* **1994**, *33*, 4111−4120.
(11) Stilz, H. U.; Dervan, P. B. *Biochemistry* **1993**, *32*, 2177−2185.
(12) Koshlap, K. M.; Gillepsie, P.; Dervan, P. B.; Feigon, J. *J. Am. Chem. Soc.* **1993**, *115*, 7908−7909.
(13) Huang, C.; Cushman, C. D.; Miller, P. S. *J. Org. Chem.* **1993**, *58*, 5048−5049.
(14) Huang, C.; Miller, P. S. *J. Am. Chem. Soc.* **1993**, *115*, 10456−10457.
(15) Lehman, T. E.; Greenberg, W. A.; Liberles, D. A.; Wada, C. K.; Dervan, P. B. *Helv. Chim. Acta* **1997**, *80*, 2002−2022.
(16) Orson, F. M.; Klysik, J.; Bergstrom, D. E.; Ward, B.; Glass, G. A.; Hua, P.; Kinsey, B. M. *Nucleic Acids Res.* **1999**, *27*, 810−816.
(17) Guzzo-Pernell, N.; Tregear, G. W.; Haralambidis, J.; Lawlor, J. M. *Nucleosides Nucleotides* **1998**, *17*, 1191−1207.
(18) Gianolio, D. A.; McLaughlin, L. W. *Nucleosides Nucleotides* **1999**, in press.
(19) (a) Liu, Z. R.; Hecker, K. H.; Rill, R. L. *J. Biomol. Struct. Dynam.* **1996**, *14*, 331−339. (b) Takenaka, S.; Yokoyama, M.; Kondo, H. *Chem. Commun.* **1997**, 115−116.
(20) Bevers, S.; O.Dea, T. P.; McLaughlin, L. W. *J. Am. Chem. Soc.* **1998**, *120*, 11004−11005.
(21) Lokey, R. S.; Kwok, Y.; Guelev, V.; Pursell, C. J.; Hurley, L. H.; Iverson, B. L. *J. Am. Chem. Soc.* **1997**, *119*, 7202−7210.
(22) Gildea, B.; McLaughlin, L. W. *Nucleic Acids Res.* **1989**, *17*, 2261−2281.
(23) Matteucci, M. D.; Caruthers, M. H. *J. Am. Chem. Soc.* **1981**, *103*, 3185−3191.
(24) For the cooling vs absorbance analysis, the X−C−G-containing triplex exhibited a hysteresis of greater magnitude than, for example, the C^+−G−C-containing triplex, suggesting a slower associate rate.

10.1021/ja990799i CCC: $18.00
Published on Web 06/19/1999

Table 1. T_m Values for 15-mer DNA Triplexes for Various N−R−Z Base Triplets[a]

5′- TTTCTTTTCTN**N**TCTT - 3′
5′ - GCGCGAAAGAAAAGA**R**AGAACCCGG - 3′
3′ - CGCGCTTTCTTTTCT**Z**TCTTGGGCC - 5′

	pH Value		
base triplet	pH 6.4	pH 7.0	pH 7.5
R−Z = G−C			
N = C	30	22	15[b]
N = X	28	22	—[c]
R−Z = C−G			
N = C	13	–	–
N = X	36	30	27
R−Z = A−T			
N = C	9	–	–
N = X	26	19	11
R−Z = T−A			
N = C	–	–	–
N = X	28	22	b[d]

[a] For assay conditions see caption to Figure 2. Values reported are the result of at least two independent assays. [b] T_m characterized by poor hyperchromicity, suggesting a poorly formed triplex. [c] No triplex → duplex transition observed. [d] Broad transition observed below 20 °C.

latter material confirmed that the phosphodiester linkage 3′ to the analogue nucleoside was refractory to enzymatic cleavage.

A number of assays of triplex T_m values were made using four 25-mer double-stranded target sequences, and two 15-mers, one containing the modified nucleoside at a preselected position and the other containing dC at this site. To test for selectivity, the four 25-mer target duplexes each contained one of the possible four DNA base pairs (dC−dG, dG−dC, T−dA, dA−T) at a preselected site (see Table 1). T_m values were obtained at three different pH values, including mildly acidic (pH 6.4), neutral (pH 7.0), and mildly basic (pH 7.5) conditions. Absorbance vs temperature plots generally resulted in two transitions, with the higher transition occurring near 74 °C. Analysis of the duplex alone resulted in this same transition. The temperature of the lower transition was dependent on the nature of the triplex formed and on the pH of the solution and was interpreted to reflect cooperative denaturation of the third strand from the duplex. An example of the two transitions observed for a triplex with an X−C−G base triplet at pH 6.4 is illustrated in Figure 2.

The T_m values for the lower transition generally decreased as the pH increased, reflecting the requisite protonation of C in the C^+−G−C base triplets. With C present in the preselected site in the third strand, the most stable complex resulted with a target G−C base pair as expected from previous studies. Of the remaining three targets, both C−G and A−T exhibited a thermally induced transition at pH 6.4, while with the T−dA target no triplex appeared to form at any pH value. The corresponding results with the analogue nucleoside present in the preselected site were quite different. DNA triplexes with dG−dC, dA−T and T−dA target base pairs exhibited similar T_m values of 26−28 °C at pH 6.4, and were also similar (19−22 °C) at pH 7.0. The complex with the dC−dG base pair target exhibited increased values of 36 and 30 °C at pH values of 6.4 and 7.0, respectively. These results indicate that the analogue residue is quite selective for the dC−dG base pair target relative to the other three possibilities. This selectivity could be explained in that by design the complex formed at dC−dG base pair targets benefits both from intercalation by the naphthalene adduct as well as from the potential

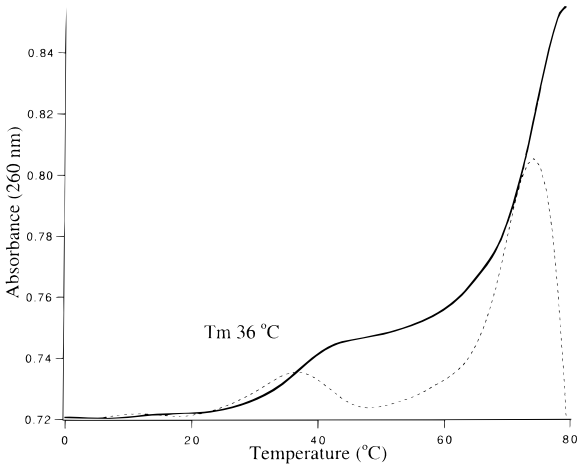

Figure 2. Absorbance (260 nm) vs temperature plot for the DNA triplex containing a X−C−G base triplet at DNA strand concentrations of 1 μM in 10 mM PIPES, pH 6.4, 10 mM $MgCl_2$, 50 mM NaCl.[24] Solutions were incubated overnight at 4 °C prior to analysis. First order differential of the absorbance vs temperature plot is illustrated as a dashed line.

formation of an interstrand hydrogen bond (Figure 1a). The remaining three base pair targets likely benefit from the intercalation mode of binding but are less likely to form an interresidue hydrogen bond. In this respect the three complexes with non-dC−dG base pair targets have very similar, but reduced T_m values. With the dA−T and T−dA base pair targets, the analogue complex was more stable than that containing simply dC, an expected result if little interresidue hydrogen bonding is present for either target and intercalative binding occurs for both. With the dG−dC base pair target the two complexes exhibited similar T_m values. With dC in the third strand two interresidue hydrogen bonds can form, while the analogue intercalation likely rotates the N^4-H residue away from the target such that with intercalation at least one if not both interresidue hydrogen bonds are lost.

To confirm that the naphthalene adduct is likely bound by intercalation, we conducted the thermal melting studies at 383 nm, the λ_{max} for the naphthalene diimide ring system. For the dX−dC−dG containing complex (T_{m260} = 36 °C, pH 6.4), a cooperative transition was observed with a midpoint of 43 °C. We interpret this result as suggesting a cooperative unstacking of the naphthalene diimide adduct during the denaturation process. That the midpoint of this transition occurs at a higher temperature than that observed for the transition monitored at 260 nm suggests that the unstacking of the naphthalene diimide might occur relatively late in the overall triplex to duplex transition.

Although a number of analogue nucleosides have been reported for the recognition of either dC−dG or T−dA base pairs, the present analogue exhibits significant selectively for the dC−dG target. Additionally, the stability as measured by T_m values indicates that the dX−dC−dG base triplet exhibits stability that is better than that observed for the well-characterized bidentate dC^+−dG−dC base triplet under the conditions of the assay.

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

Supporting Information Available: Analogue synthesis procedures, conditions for thermal denaturation experiments, sample absorbance vs temperature plots are available (PDF). This material is available free of charge via the Internet at http://pubs.acs.org.

JA990799I