A Cytosine Analogue Capable of Clamp-Like Binding to a Guanine in Helical Nucleic Acids

Kuei-Ying Lin and Mark D. Matteucci*

Gilead Sciences, 333 Lakeside Drive
Foster City, California 94404

Received April 16, 1998

The Watson−Crick pairing of heterocycles within duplex DNA is the foundation of biomolecular recognition.[1] The cytosine−guanine interaction is formed by three hydrogen bonds. The guanine base contains two unused hydrogen bond acceptors in the major groove at the O6 and N7 of the Hoogsteen binding face. A cytosine analogue termed G-clamp, when incorporated into oligonucleotides(ODNs), simultaneously recognizes both the Watson−Crick and Hoogsteen faces of a complementary guanine within a helix. A single G-clamp analogue substitution within an ODN results in substantially enhanced helical thermal stability and mismatch discrimination when hybridized to complementary DNA and RNA. These properties of enhanced affinity and specificity are of interest in the fields of nucleic acid based diagnostics[2,3] and the sequence-specific targeting of RNA by the antisense approach.[4,5]

We previously reported the synthesis and binding properties of ODNs containing the tricyclic 2′-deoxycytidine analogue, phenoxazine (structure **1**, Figure 1).[6] This heterocyclic modification provided a rigid scaffold for appending groups designed to interact with the Hoogsteen binding face of a complementary base-paired guanine. Model building studies suggested that the protonated amino group of the G-clamp (structure **2**, Figure 1) could make a specific hydrogen bonded contact with the O6 of the targeted guanine within a helix as shown in Figure 2.

The syntheses of the monomer synthons required for ODN synthesis are reported elsewhere.[7] The ODNs shown in Table 1 were synthesized and purified by standard methods[8] and characterized by MALDI-TOF mass spectrometry. The ODNs were hybridized to a complementary ODN and T_m measurements recorded. The data are shown in Table 1. Only the G-clamp containing ODN showed dramatically enhanced affinity relative to a 5-methyl cytosine control. The tricyclic phenoxazine **1** bearing no arm showed affinity enhancement consistent with the improved stacking interactions observed previously.[6] The virtually isosteric tricyclic analogue **3** bearing the weakly hydrogen-bond-donating hydroxyl group showed no enhanced affinity.[9] The acyclic derivative **4** lacking the conformational restriction of G-clamp demonstrated no enhanced affinity.[11] The G-clamp's dramatic affinity for the complementary guanine depended on the appropriate positioning of the strong hydrogen bond donor.

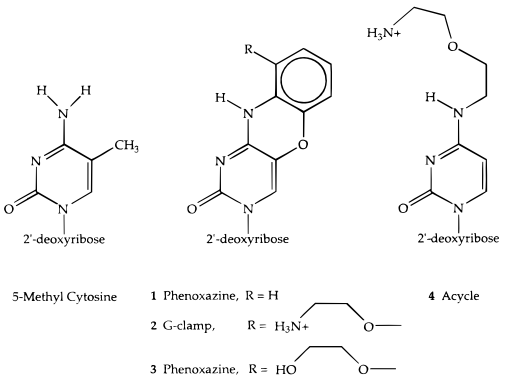

Figure 1. Chemical structures of cytosine analogues.

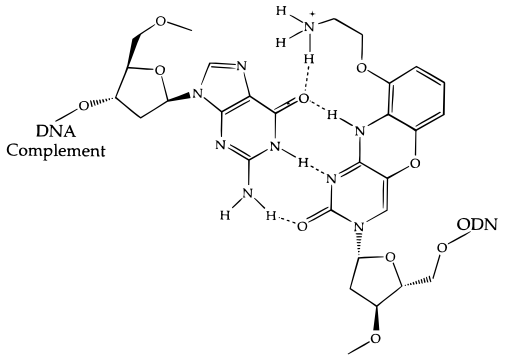

Figure 2. Model of G clamp−guanine interaction within a DNA−DNA helix.

Table 1. Effect of Structural Variation on T_m Values[a]

X	Tm (°C)	ΔTm Relative to 5-MeC
5-Methyl Cytosine (5-MeC)	50.5	—
1 Phenoxazine, R = H	57.0	+6.5
2 G-clamp, R = H_3N^+ O—	68.5	+18.0
3 Phenoxazine, R = HO O—	51.5	+1.0
4 Acycle	49.5	-1.0

[a] Buffer conditions were 140 mM KCl, 5 mM Na_2HPO_4, 1 mM $MgCl_2$, pH 7.2. The error in T_m values was ± 0.5 C. Target DNA: 3′-AGAGGGAGAGA$_5$. Test ODN: 5′-TCTC**X**CTCTC.

A specific interaction with the Hoogsteen face of the targeted guanine should result in enhanced specificity. The G-clamp possessed greater discrimination between the perfect match with guanine and mismatches with adenine, thymine, and cytosine. The matrix of matched and mismatched T_m values are shown in Table 2. The G-clamp conferred enhanced specificity in all cases relative to 5-methyl cytosine and the parent phenoxazine **1**. Only the specific hybridization with a targeted guanine resulted in enhanced affinity.

A specific interaction with the Hoogsteen face of the targeted guanine should be relatively insensitive to the flanking sequence. The context of the G-clamp was changed as shown in Table 3. The placement of the targeted guanine between two adenines still

* Corresponding author: Mark Matteucci, Ph.D., Gilead Sciences, Inc., 333 Lakeside Drive, Foster City, CA 94404. Phone: 650 573 4749. FAX: 650 573 4890. E-mail: mark_ matteucci@gilead.com.

(1) Watson, J. D.; Crick, F. H. C. *Nature* **1953**, *171*, 737−738.
(2) Whelen, A. C.; Persing, D. H. *Annu. Rev. Microbiol.* **1996**, *50*, 349−373.
(3) Sidransky, D. *Science* **1997**, *278*, 1054−1059.
(4) Monia, B. P.; Johnston, J. F.; Geiger, T.; Muller, M.; Fabbro, D. *Nat. Med.* **1996**, *2*, 668−675.
(5) Matteucci, M. D.; Wagner, R. W. *Nature* **1996**, *384*, 20−22.
(6) Lin, K. Y.; Jones, R. J.; Matteucci, M. D. *J. Am. Chem. Soc.* **1995**, *117*, 3873−3874.
(7) See Supporting Information.
(8) Froehler, B. C.; Ng, P. G.; Matteucci, M. D. *Nucleic Acids Res.* **1986**, *14*, 5399−5407.
(9) An alternative explanation for the relatively poor binding of **3** is that in the unbound state, the oxygen of the terminal OH is hydrogen (H) bonded to the H of the N4 (cytosine numbering). This intramolecular H bond would need to be broken upon hybridization, resulting in an enthalpic penalty and no entropic benefit.[10]
(10) Dunitz, J. D. *Chem. Biol.* **1995**, *2*, 709−712.
(11) This derivative is not preordered to resemble its bound conformation and likely exists in a number of conformations. This disorder would have to be ordered upon binding with a consequent entropic penalty.[10]

S0002-7863(98)01286-4 CCC: $15.00
Published on Web 08/07/1998

Table 2. Comparison of Matched versus Mismatched Hybridization by T_m Analysis[a]

	T_m (°C)			
X =	Y = G	A	T	C
5-methyl cytosine	50.5	32.0 (18.5)	30.0 (20.5)	29.0 (21.5)
1 phenoxazine, R = H	57.0	44.5 (12.5)	42.0 (15.0)	33.0 (24.0)
2 G-clamp	68.5	45.5 (23.0)	41.0 (27.5)	40.0 (28.5)

[a] Values in parentheses are the difference between the matched T_m with the guanine-containing target and the T_m from the mismatched base shown. The buffer conditions and the error value were the same as those in Table 1. Target DNA: 3′-AGAG**Y**GAGAGA$_5$. Test ODN: 5′-TCTC**X**CTCTC.

Table 3. Comparison of T_m Values with Different Flanking Sequence[a]

X	T_m (°C)	ΔT_m relative to 5-MeC
5-methyl cytosine	50.5	
2 G-clamp	65.0	+14.5

[a] The buffer conditions and error limits were the same as those in Table 1. Target DNA: 3′-AGAGGGAGAGA$_5$. Test ODN: 5′-TCTC-CCT**X**TC.

Table 4. Ionic Strength Dependence of the T_m Values[a]

	T_m (°C)		
X =	1.4 M KC1	0.14 M KC1	0.014 M KC1
5-methyl cytosine	57.0	50.5	45.0
2 G-clamp	73.0	68.5	62.5
	ΔT_m = 16.0	ΔT_m = 18.0	ΔT_m = 17.5

[a] ΔT_m values are the difference between the T_m values of the G clamp and the 5-MeC hybrids at the stated KC1 concentrations. Buffer conditions are the indicated KC1 concentration plus 5 mM Na_2HPO_4, 1 mM $MgC1_2$, pH 7.2. The error in T_m values was ± 0.5 °C. Target DNA: 3′-AGAGGGAGAGA$_5$. Test ODN: 5′-TCTC**X**CTCTC.

resulted in substantially enhanced affinity relative to 5-methyl cytosine.

A specific interaction with the Hoogsteen face of guanine would not be ionic in nature. The protonated amino group could potentially be interacting with an anionic phosphate on the complementary strand. Such an ionic interaction would be sensitive to the salt concentration of the T_m analysis.[12,13] An ionic interaction would be most significant at low ionic strength. The salt concentration was varied from approximately 14 mM to 1.4 M, and the data are shown in Table 4. All T_m values increased with increasing ionic strength because of the salt screening of the anion−anion repulsion of the phosphates.[14] However, the enhanced affinity of the G-clamp relative to the 5-methyl cytosine control was not dependent on salt concentration. Additionally, specific phosphate anions in the target were removed by replacement one at a time with the neutral methyl phosphonate analogue.[15] The enhanced T_m value of G-clamp was not reduced by the specific anion eliminations in the target as shown in Table 5. Taken together, these results strongly suggest that the origin of this specific interaction is not ionic and is that of hydrogen bonding to the Hoogsteen face of the guanine.

Table 5. Dependence of the T_m Values on Methyl Phosphonate Substitutions[a]

	T_m (°C) methyl phosphonate position				
X =	none	1	2	3	4
5-methyl cytosine	50.5	50.0	50.5	51.0	48.0
2 G-clamp	68.5	69.0	69.0	69.0	66.0
	ΔT_m = 18.0	ΔT_m = 19.0	ΔT_m = 18.5	ΔT_m = 18.0	ΔT_m = 18.0

[a] The positions of the single methyl phosphonate substitutions in the target DNA are shown by bold superscript numbers. ΔT_m values are the difference between the T_m values of the G clamp and the 5-MeC hybrids. The buffer conditions and error limits were the same as those in Table 1. Target DNA: 3′-AGA$^{\mathbf{1}}$G$^{\mathbf{2}}$G$^{\mathbf{3}}$G$^{\mathbf{4}}$AGAGA$_5$. Test ODN: 5′-TCTC**X**CTCTC.

A specific interaction with the Hoogsteen face should be insensitive to changing the helix structure from a B form DNA−DNA hybrid to an A form DNA−RNA helix. The ΔT_m analysis was performed with the complementary RNA target with the sequences being those shown in Table 1. The G-clamp ODN again showed a ΔT_m of +16.0 °C relative to the 5-methyl cytosine control. This enhanced affinity to complementary RNA is of particular interest to the field of antisense targeting of RNA.[4,5]

To our knowledge this is the first report of a rationally designed nucleobase analogue capable of an additional hydrogen bond to a complementary base within a helix. The monomeric guanine heterocycle has been previously recognized by four hydrogen bonds in organic solvents using a synthetic receptor.[16] This receptor, however, is not amenable to incorporation into ODNs. Cationic amines have been appended from the 5 position of 2′-deoxyuridine and incorporated into ODNs.[17−19] These derivatives have resulted in modest enhancement of affinity through non-specific ionic interactions.

The data presented strongly suggest a hydrogen bond interaction between the ammonium group on the G-clamp and the Hoogsteen face of guanine. Model building suggests the interaction is with the O6 group on guanine. The tether arm appears not to be long enough to reach the N7. The interaction of protonated amino groups with the O6 of guanine in duplex DNA has precedence in protein−DNA interactions. X-ray crystallography of the λ repressor-DNA operator complex revealed lysine residues making specific contacts with the O6 of guanines in the major groove of the DNA.[20] Definitive structural characterization of the G clamp interaction awaits X-ray crystallography or high-field NMR analysis.

Acknowledgment. We thank S. Swaminathan for helpful discussions, T. Terhorst for ODN synthesis, and U. von Krosigk for early synthetic investigations.

Supporting Information Available: Preparation and characterization of monomer synthons for ODN synthesis (6 pages, print/PDF). See any current masthead page for ordering information and Web access instructions.

JA981286Z

(12) Dempcy, R. O.; Browne, K. A.; Bruice, T. C. *J. Am. Chem. Soc.* **1995**, *117*, 6140−6141.
(13) Smith, J. S.; Scholtz, J. M. *Biochemistry* **1998**, *37*, 33−40.
(14) Saenger, W. *Principles of Nucleic Acid Structure*; Springer-Verlag: New York, 1984; p 145.
(15) Smith, S. A.; McLaughlin, L. W. *Biochemistry* **1997**, *36*, 6046−6058.
(16) Bell, T. W.; Hou, Z.; Zimmerman, S. C.; Thiessen, P. A. *Angew. Chem., Int. Ed. Engl.* **1995**, *34*, 2163-2165.
(17) Dande, P.; Liang, G.; Chen, F.-X.; Roberts, C.; Nelson, M. G.; Hashimoto, H.; Switzer, C.; Gold, B. *Biochemistry* **1997**, *36*, 6024−6030.
(18) Ueno, Y.; Kumagai, I.; Haginoya, N.; Matsuda, A. *Nucleic Acids Res.* **1997**, *25*, 3777−3782.
(19) Haginoya, N.; Ono, A.; Nomura, Y.; Ueno, Y.; Matsuda, A. *Bioconjugate Chem.* **1997**, *8*, 271.
(20) Clarke, N. D.; Beamer, L. J.; Goldberg, H. R.; Berkower, C.; Pabo, C. O. *Science* **1991**, *254*, 267−270.