Design of an N^7-Glycosylated Purine Nucleoside for Recognition of GC Base Pairs by Triple Helix Formation

Jürg Hunziker, E. Scott Priestley, Helmut Brunar, and Peter B. Dervan*

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

Received November 23, 1994

Pyrimidine oligodeoxyribonucleotides bind in the major groove of DNA *parallel* to the purine Watson−Crick strand by formation of specific hydrogen bonds between thymine and adenine (T•AT triplet) and protonated cytosine and guanine (C+GC triplet) on the Hoogsteen face of the purine base.[1] Alternatively, purine oligodeoxyribonucleotides bind in an *antiparallel* orientation relative to the purine Watson−Crick strand by formation of G•GC and A•AT triplets.[2] The prerequisite protonation of cytosine in C+GC triplets leads to a considerable pH dependence in the binding affinity of C-containing oligodeoxyribonucleotides (Figure 1).[1,3] Substitution of 5-methylcytosine (^{m}**C**) for cytosine results in increased binding affinities near physiological pH.[3a,d,e] In an attempt to eliminate the necessity for protonation, recent efforts have been directed toward the synthesis of *nonnatural* nucleosides which display the hydrogen bonding functionality of protonated cytosine.[4,5]

We report that an N^7-glycosylated purine, 7-(2-deoxy-β-D-*erythro*-pentofuranosyl) guanine (**d^{7}G**), when incorporated in a pyrimidine oligonucleotide, binds with remarkable specificity the Watson−Crick guanine−cytosine (GC) base pair by triple helix formation. By attaching the deoxyribose moiety at the N^7-position of a guanine base, *the third strand orientation in a G•GC base triplet is reversed* and becomes *parallel* to the purine Watson−Crick strand (Figure 1).

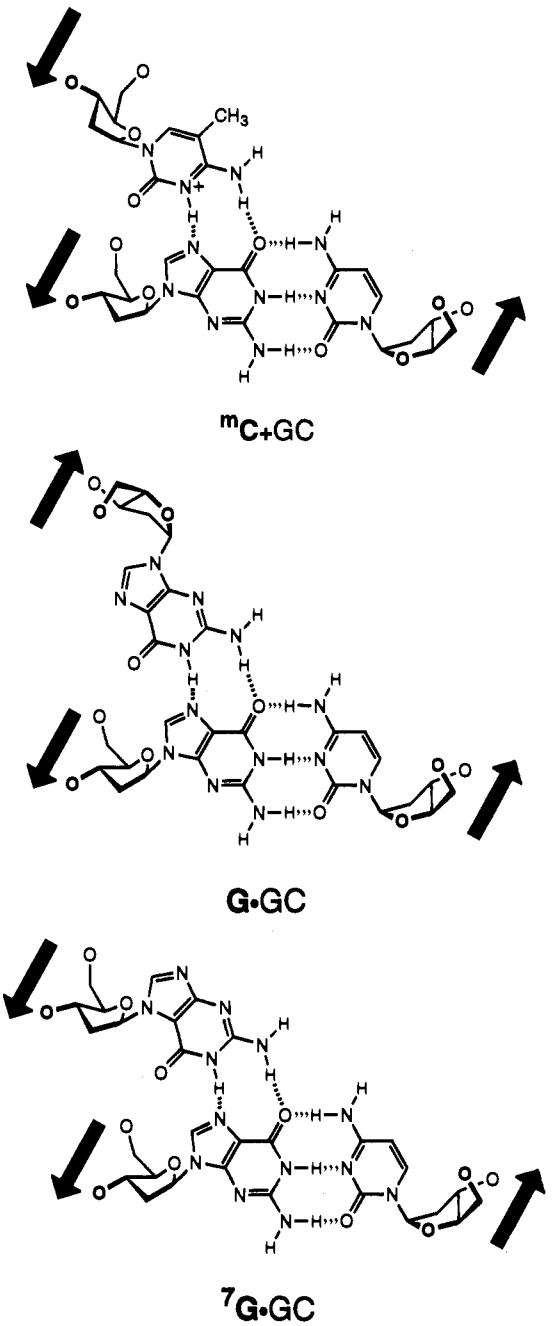

Figure 1. Schematic representations of the base triplets ^{m}C+GC (top), G•GC (middle), and ^{7}G•GC (bottom). Arrows indicate 5′ → 3′ direction of the phosphodiester backbone.

The phosphoramidite of **d^{7}G** was synthesized in four steps from methyl 3,5-di-*O*-benzoyl-α,β-D-*erythro*-pentafuranoside and N^2-isobutyrylguanine (Figure 2).[7] The **d^{7}G** nucleotide was incorporated at a single position within 15-mer pyrimidine oligodeoxyribonucleotide **1** by automated methods.[8] To test the affinity and specificity of **d^{7}G**, the association constants for the formation of 20 triple helical complexes which vary at a single

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

(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) For NMR studies, see: Radhakrishnan, I.; de los Santos, C.; Patel, D. J. *J. Mol. Biol.* **1993**, *234*, 188−197.

(3) (a) Povsic, T. J.; Dervan, P. B. *J. Am. Chem. Soc.* **1989**, *111*, 3059−3061. (b) Plum, G. E.; Park, Y.-W.; Singleton, S. F.; Dervan, P. B.; Breslauer, K. J. *Proc. Natl. Acad. Sci. U.S.A.* **1990**, *87*, 9436−9440. (c) Maher, L. J., III; Dervan, P. B.; Wold, B. J. *Biochemistry* **1990**, *29*, 8820−8826. (d) Xodo, L. E.; Manzini, G.; Quadrifoglio, F.; van der Marel, G. A.; van Boom, J. H. *Nucleic Acids Res.* **1991**, *19*, 5625−5631. (e) Singleton, S. F.; Dervan, P. B. *Biochemistry* **1992**, *31*, 10995−11003.

(4) Earlier work in our group focused on the novel nucleoside P1,[5d] which forms specific hydrogen bonds between its exocyclic amino and imino groups and the Hoogsteen face of guanine[6] and recognizes GC base pairs over an extended pH range relative to ^{m}C or C.

(5) (a) Ono, A.; Ts'o, P. O. P.; Kan, L.-S. *J. Am. Chem. Soc.* **1991**, *113*, 4032−4033. (b) Ono, A.; Ts'o, P. O. P.; Kan, L.-s. *J. Org. Chem.* **1992**, *57*, 3225−3230. (c) Young, S. L.; Krawczyk, S. H.; Matteucci, M. D.; Toole, J. J. *Proc. Natl. Acad. Sci. U.S.A.* **1991**, *88*, 10023−10026. (d) Koh, J. S.; Dervan, P. B. *J. Am. Chem. Soc.* **1992**, *114*, 1470−1478. (e) 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. (f) Miller, P. S.; Bhan, P.; Cushman, C. D.; Trapane, T. L. *Biochemistry* **1992**, *31*, 6788−6793. (g) Jetter, M. C.; Hobbs, F. W. *Biochemistry* **1993**, *32*, 3249−3254. (h) Davison, E. C.; Johnsson, K. *Nucleosides Nucleotides* **1993**, *12*, 237−243.

(6) Radhakrishnan, I.; Patel, D. J.; Priestley, E. S.; Nash, H. N.; Dervan, P. B. *Biochemistry* **1993**, *32*, 11228−11234.

(7) (a) For a recent synthesis, see: Rao, T. S.; Durland, R. H.; Revankar, G. R. *J. Heterocycl. Chem.* **1994**, *31*, 935−940. (b) Methyl 3,5-di-*O*-benzoyl-α,β-D-*erythro*-pentofuranoside was synthesized according to: Blair, M. G.; Lipkin, D.; Sowden, J. C.; Strobach, J. C. *J. Org. Chem.* **1960**, *25*, 1679−1680. (c) N^2-Isobutyrylguanine was prepared in analogy to: Hrebabecky, H.; Farkas, J. Acetylation of an Amino Group on a Purine Ring with Acetic Anhydride in an Aprotic, Dipolar Solvent. In Townsend, L. B.; Tipson, R. S. *Nucleic Acid Chemistry*; J. Wiley & Sons: New York, 1978; Vol. 1.

(8) Enzymatic degradation and base composition analysis by HPLC confirmed the composition of oligonucleotide **1**. Analysis by MALDI TOF mass spectrometry gave excellent agreement between the experimental (4519.7) and calculated (4519.1) masses (M − 1) of this oligonucleotide.

(9) For a detailed description of the quantitative DNase footprint titration method, see: (a) Brenowitz, M.; Senear, D. F.; Shea, M. A.; Ackers, G. K. *Methods Enzymol.* **1986**, *130*, 132−181. (b) Singleton, S. F.; Dervan, P. B. *J. Am. Chem. Soc.* **1992**, *114*, 6957−6965. (c) Priestley, E. S; Dervan, P. B. *J. Am. Chem. Soc.*, in press.

(10) For examples of a similar assay, see: Jones, R. J.; Swaminathan, S.; Milligan, J. F.; Wadwani, S.; Froehler, B. C.; Matteucci, M. D. *J. Am. Chem. Soc.* **1993**, *115*, 9816−9817. Jones, R. J.; Lin, K.-Y.; Milligan, J. F.; Wadwani, S.; Matteucci, M. D. *J. Org Chem.* **1993**, *58*, 2983−2991.

0002-7863/95/1517-2661$09.00/0

Table 1. Equilibrium Association Constants (M^{-1}) for 20 Triple Helical Complexes Which Vary at a Single Position[a]

		XY			
Z	oligo	AT	CG	GC	TA
^{7}G	**1**	$<10^4$	$4.6 (\pm 1.9) \times 10^5$	$\geq$**1.6 × 10^8**	$<10^4$
^{m}C	**2**	$8.2 (\pm 1.9) \times 10^4$	$1.7 (\pm 0.6) \times 10^5$	$\geq$**1.4 × 10^8**	$<10^4$
T	**3**	$\geq$**2.1 × 10^8**	$5.4 (\pm 0.7) \times 10^5$	$<10^4$	$<10^4$
G	**4**	$<10^4$	$<10^4$	$<5 \times 10^4$	$1.1 (\pm 0.3) \times 10^6$
A	**5**	$1.6 (\pm 0.6) \times 10^5$	$<10^4$	$<10^4$	$<10^4$

[a] Data reported are the mean values of three footprint titration experiments ± SEM. The titrations were performed at 22 °C, pH 7.0, in the presence of 100 mM NaCl, 10 mM Bis-Tris·HCl, and 0.25 mM spermine·4HCl and allowed to equilibrate for 120 h prior to footprinting. Association constants greater than 1×10^8 M^{-1} are reported as lower limits due to the slow kinetics of triple helix formation under these conditions.

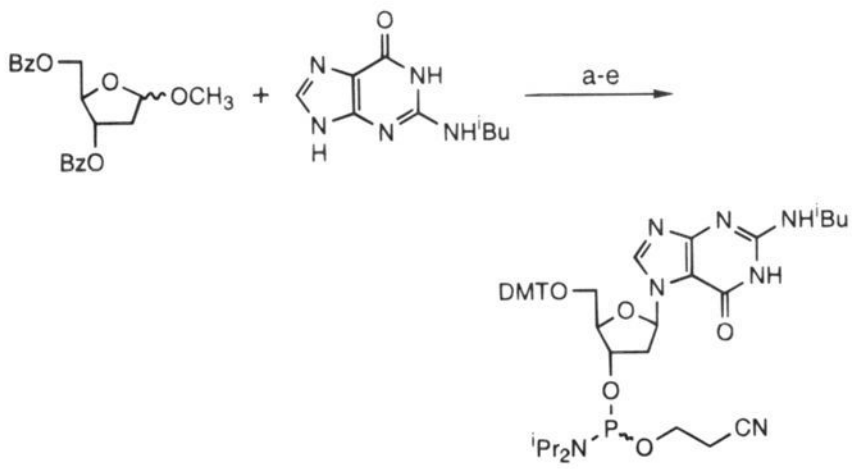

Figure 2. Synthesis of the DMT protected phosphoramidite of **d^{7}G**: (a) *N,O*-bis(trimethylsilyl)acetamide, $SnCl_4$, CH_3CN, room temperature, 2 h; (b) NaOH, THF, MeOH, H_2O, 0 °C, 25 min; (c) DMTCl, pyridine, room temperature, 4 h; (d) column chromatography; (e) *N,N*-diisopropylethylamine, 2-cyanoethyl *N,N*-diisopropylchlorophosphoramidite, THF, room temperature, 1 h.

common position Z·XY (where Z = ^{7}G, ^{m}C, T, G, and A, and XY = AT, GC, CG, and TA) were determined within a pyrimidine motif by quantitative DNase I footprint titration (100 mM NaCl, 10 mM Bis-Tris·HCl, 0.25 mM spermine·4HCl, pH 7.0, 22 °C) (Table 1).[9] The sequence context in which the Z·XY pairings were studied is 5′-d(AAGAGAXAGAGGAAA)-3′ within a 314 base pair DNA restriction fragment.[10] The affinity of ^{7}G is determined by comparison of oligonucleotides **1** and **2** (Z = ^{7}G, ^{m}C) opposite XY = GC (Figure 3). The sequence specificity of ^{7}G is determined by comparison of oligonucleotide **1** (Z = ^{7}G) opposite each of the four Watson−Crick base pairs (XY = AT, GC, CG, and TA) (Figure 3).

We find that oligonucleotide **1** (Z = ^{7}G) binds the target site with an equilibrium association constant $\geq 1.6 \times 10^8$ M^{-1}, an affinity equal to that of oligonucleotide **2** (Z = ^{m}C) (Table 1). Importantly, **d^{7}G** is specific for GC base pairs; the stabilities of triple helical complexes containing **d^{7}G** decrease in the order d^{7}G·GC ≫ d^{7}G·CG > d^{7}G·AT, d^{7}G·TA (Table 1). The high affinity and sequence specificity observed are consistent with a model where **^{7}G** within the pyrimidine motif forms two specific hydrogen bonds to the purine strand of the Watson−Crick GC base pair (Figure 1). On the basis of these results, we envision that *a new parallel-stranded motif* comprising wholly N^7 purines might be possible for DNA recognition by triple helix formation.

Acknowledgment. We are grateful to the Office of Naval Research for support, the Schweizerischer Nationalfonds for a Postdoctoral Fellowship to J.H., the Austrian Science Foundation for an Erwin Schrödinger Postdoctoral Fellowship to H.B., and the National Science Foundation for a Predoctoral Fellowship to ESP.

JA943798T

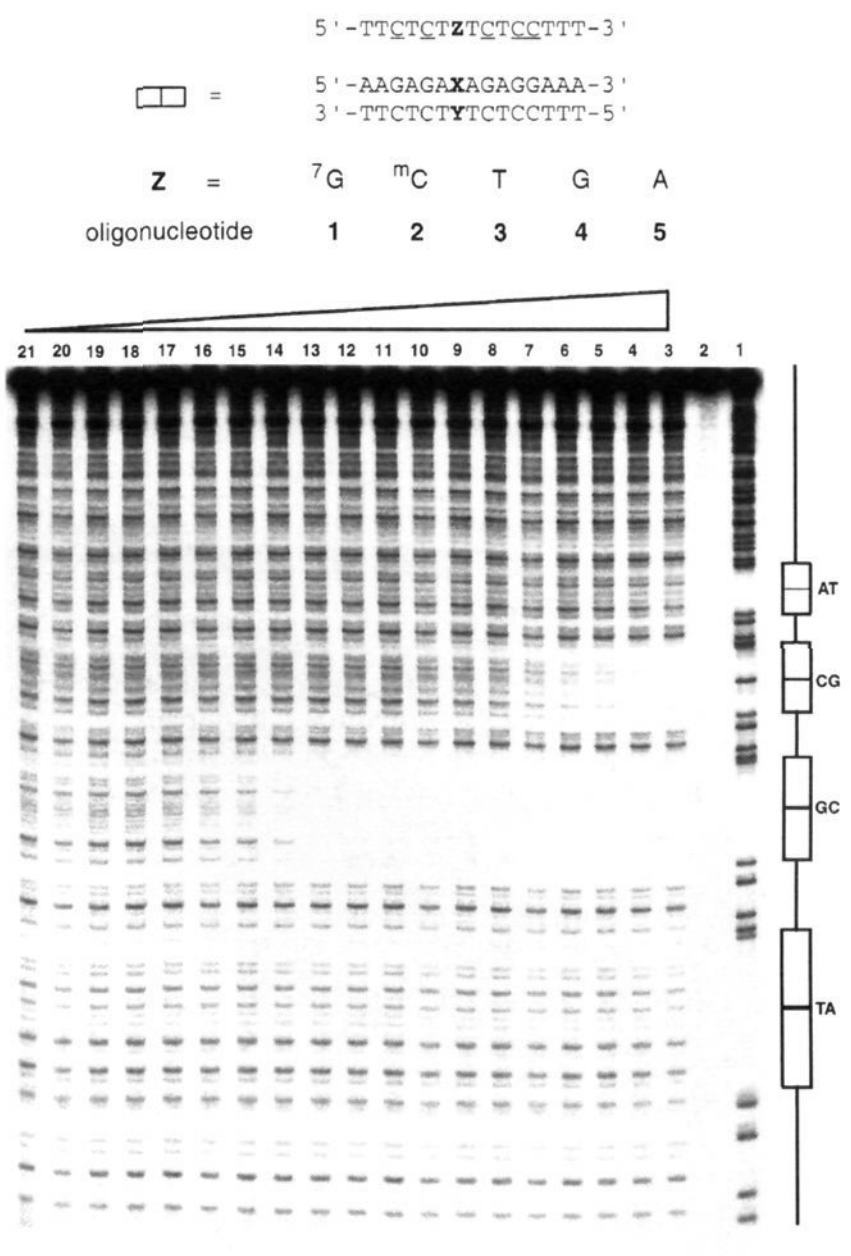

Figure 3. (top) Schematic representation of the oligonucleotides and duplex target sites used in the quantitative DNase footprint titrations (C indicates ^{m}C). (bottom) Storage phosphor autoradiogram of an 8% denaturing polyacrylamide gel used to separate the fragments generated by DNase I digestion in a quantitative footprint titration experiment. The bars drawn to the right of the autoradiogram indicate the positions of the 15-mer binding sites within the 314-bp restriction fragment. Lane 1, products of a guanine-specific sequencing reaction; lane 2, intact 3′ labeled duplex after incubation in the absence of a third strand oligonucleotide; lanes 3−20, DNA cleavage products produced by oligonucleotide **1** at various concentrations (40 μM, lane 3; 20 μM, lane 4; 8 μM, lane 5; 4 μM, lane 6; 2 μM, lane 7; 800 nM, lane 8; 400 nM, lane 9; 200 nM, lane 10; 80 nM, lane 11; 40 nM, lane 12; 20 nM, lane 13; 8 nM, lane 14; 4 nM, lane 15; 2 nM, lane 16; 800 pM, lane 17; 400 pM, lane 18; 200 pM, lane 19; and 80 pM, lane 20, and lane 21, DNA cleavage products produced in the absence of a third strand oligonucleotide.