Cation-Templated Self-Assembly of a Lipophilic Deoxyguanosine: Solution Structure of a K^+−dG_8 Octamer

Allison L. Marlow,[†] Elisabetta Mezzina,[‡] Gian Piero Spada,[‡] Stefano Masiero,[‡] Jeffery T. Davis,*[,†] and Giovanni Gottarelli*[,‡]

Department of Chemistry and Biochemistry, University of Maryland, College Park, Maryland 20742, and Dipartimento di Chimica Organica "A Mangini", Universita di Bologna, Via S. Donato 15, I-40127 Bologna, Italy

Received January 26, 1999

The lipophilic nucleoside 3′,5′-didecanoyl-2′-deoxyguanosine, dG **1**, extracts potassium salts from water into organic solvents. The K^+ extraction drives the self-association of dG **1** to give G-quartet structures. A series of ^{1}H NMR experiments indicates that the identity of the assembled species in $CDCl_3$ is modulated by the amount of K^+ extracted by dG **1**. At an 8:1 dG **1** to K^+ picrate ratio, the octamer (dG **1**)$_8$−K^+ predominates in solution. The (dG **1**)$_8$−K^+ supramolecular complex, formed by coordination of a single K^+ ion by eight dG **1** monomers, is robust and structurally unique. The ^{1}H NMR chemical shifts for both the exchangeable and nonexchangeable protons of (dG **1**)$_8$−KI in $CDCl_3$ were assigned from a combination of 2D ^{1}H−^{1}H and ^{13}C−^{1}H correlation experiments. One set of ^{1}H NMR signals corresponds to a dG **1** nucleoside with an *anti* conformation about the C(1′)−N(9) glycosidic bond, whereas the other set of signals is due to 50% of the didecanoyl dG **1** adopting a *syn* conformation. Although the possible arrangements of an octamer containing a 1:1 ratio of *anti* dG **1** to *syn* dG **1** are many, the present NMR analysis leads to a defined single species composed of two G-quartets. In one tetramer, all of the dG **1** components have a *syn* conformation about the C(1′)−N(9) glycosidic bond, while the other tetramer has an "all-*anti*" conformation. Moreover, intertetramer NOEs are consistent with stacking of the "all-*anti*" tetramer in a "head-to-tail" orientation on top of the "all-*syn*" tetramer, thus sandwiching a central K^+ ion. This solution structure is, to our knowledge, different from all of the assembled structures described so far for guanine aggregates. Presumably, the K^+-bound octamer represents the first observable stage of the assembly process in the aggregation of dG **1**.

Introduction

Ion complexation is important in chemistry and biology. Ionophores such as cyclic peptides, crown ethers, and calixarenes have preorganized covalent structures that facilitate ion binding.[1,2] There have been fewer ionophores built using noncovalent interactions.[3−7] For several years we have been studying the self-assembly and liquid crystal formation of guanosine analogues and folates.[8] These processes, which occur in water, lead to columnar aggregates. The columnar aggregates contain guanine tetramers (G-quartets) as the fundamental structural unit (Figure 1). The G-quartet is formed by hydrogen bonding of four G nucleobases. This cyclic structure has four carbonyl oxygens pointing in toward a central cavity. The cavity is well-organized for coordinating cations, particularly K^+. The presence of K^+ increases the stability and length of the columnar aggregates and lowers the critical concentrations at which the liquid crystalline phases appear.

Guanosine derivatives have long been known to self-associate in water.[9] Formation of helical stacks of G-quartets was proposed to explain why guanosine forms gels in water.[10] Later, X-ray fiber diffraction of G analogues showed stacks of quartets.[11−14] Whereas G nucleotides form gels under acidic conditions, soluble and ordered aggregates were detected under neutral and basic conditions. In the 1970s, Miles provided IR evidence for existence of the G-quartet in water, and Pinnavaia used ^{1}H NMR spectroscopy to demonstrate the presence of G-quartets in water.[11,12] Later, an important finding was that monovalent cations, especially Na^+ and K^+, stabilized discrete G_8−M^+ octamers for 5′-GMP.[13−19]

* To whom correspondence should be addressed. E-mail: jd140@umail.umd.edu and gottarel@alma.unibo.it.
[†] University of Maryland.
[‡] Universita di Bologna.

(1) Lehn, J. M. *Angew. Chem., Int. Ed. Engl.* **1988**, *27*, 89−112.
(2) Cram, D. J.; Cram, J. M. *Container Molecules and Their Guests*; Royal Society of Chemistry: Cambridge, 1994.
(3) Gottarelli, G.; Masiero, S.; Spada, G. P. *J. Chem. Soc., Chem. Commun.* **1995**, 2555−2557.
(4) Davis, J. T.; Tirumala, S.; Jenssen, J. R.; Radler, E.; Fabris, D. *J. Org. Chem.* **1995**, *60*, 4167−4176.
(5) Davis, J. T.; Tirumala, S.; Marlow, A. L. *J. Am. Chem. Soc.* **1997**, *119*, 5271−5272.
(6) Tirumala, S.; Davis, J. T. *J. Am. Chem. Soc.* **1997**, *119*, 2769−2776.
(7) Marlow, A. L.; Davis, J. T. *Tetrahedron Lett.* **1999**, *40*, 3539−3542.
(8) Gottarelli, G.; Spada, G. P.; Garbesi, A. In *Comprehensive Supramolecular Chemistry, vol. 9*; Sauvage, J. P., Hosseini, M. W., Eds.; Pergamon: Elmsford, 1996; Vol. 9; pp 483−506.
(9) For a review, see: Guschlbauer, W.; Chantot, J. F.; Thiele, D. *J. Biomol. Struct. Dynam.* **1990**, *8*, 491−511.
(10) Gellert, M.; Lipsett, M. N.; Davies, D. R. *Proc. Natl. Acad. Sci., U.S.A.* **1962**, *48*, 2013−2018.
(11) Miles, H. T.; Frazier, J. *Biochem. Biophys. Res. Comm.* **1972**, *49*, 199−204.
(12) Pinnavaia, T. J.; Miles, H. T.; Becker, E. D. *J. Am. Chem. Soc.* **1975**, *97*, 7198−7200.
(13) Pinnavaia, T. J.; Marshall, C. L.; Mettler, C. M.; Fisk, C. L.; Miles, H. T.; Becker, E. D. *J. Am. Chem. Soc.* **1978**, *100*, 3625−3627.
(14) Detellier, C.; Laszlo, P. *Helv. Chim. Acta* **1979**, *62*, 1559−1565.
(15) Borzo, M.; Detellier, C.; Laszlo, P.; Paris, A. *J. Am. Chem. Soc.* **1980**, *102*, 1124−1134.
(16) Detellier, C.; Laszlo, P. *J. Am. Chem. Soc.* **1980**, *102*, 1135−1141.

10.1021/jo9901440 CCC: $18.00
Published on Web 06/16/1999

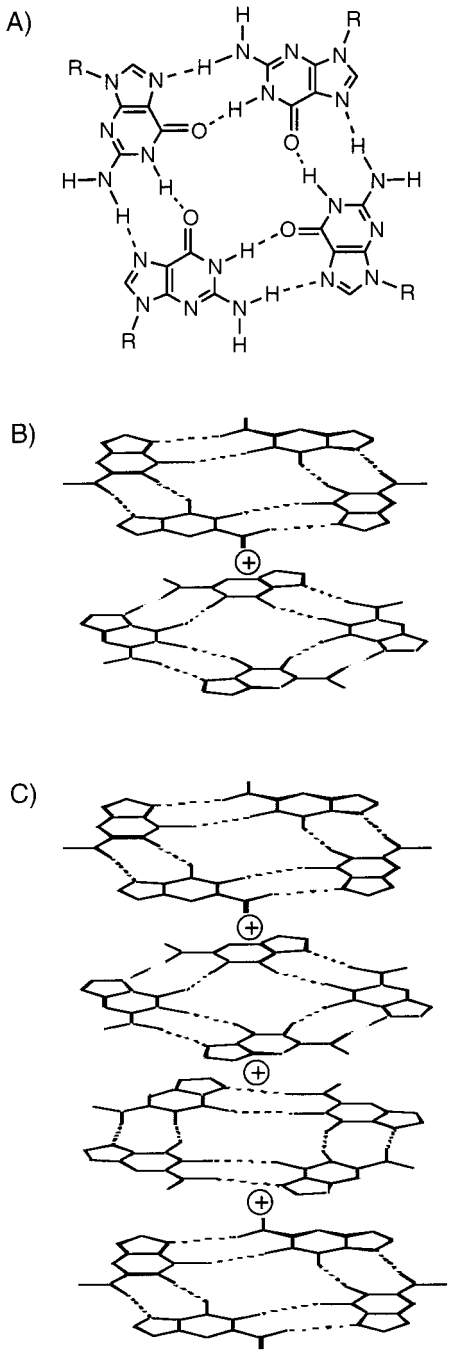
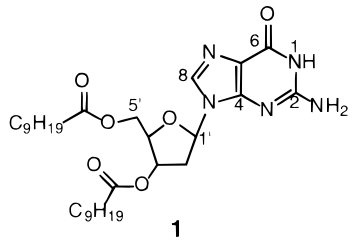


Figure 1. (A) The G-quartet; (B) an octamer, G_8-K^+; (C) a stack of G-quartets held together by K^+ ions.

Although recent attention has focused on the G-quartet motif in oligonucleotides,[20] further studies on the self-assembly of mononucleotide G-quartets have appeared.[21−23] Although there is little doubt that mononucleotides can form G-quartets and M^+-bound octamers in solution, there have been few studies aimed at determining glycosidic conformation and stacking orientation in these soluble aggregates.[23] Information about the structures of G_8-M^+ octamers formed by mononucleosides may provide insight into the structure of the larger aggregates, such as those in gels, liquid crystals, and nucleic acids.

To date, characterization of the G-quartet in organic solvents has not been described.[24] This lack of information may be due to difficulties in manipulating organic-soluble G derivatives and the notion that base-stacking is not favored in organic solvents. Certainly organic solvents are excellent media for studying isolated electrostatic interactions, such as hydrogen bonds and cation−ligand coordination. Many classical studies of nucleobase pairing have been done in organic solvents.[25−28] Also, replacement of phosphorylated G analogues with uncharged derivatives eliminates potential complications in the structural analysis of G aggregates caused by nonselective cation binding to phosphates in the G-quartet's backbone.[16,17] Without the charged phosphates, the use of organic-soluble G analogues ensures that M^+ coordination occurs at the G-quartet's primary binding site, its central ring of four carbonyl oxygens.

Recently, we have explored whether guanosine self-assembly might also occur in organic solvents.[3,4] We found that 3′,5′-didecanoyl-2′-deoxyguanosine, dG **1**, a lipophilic analogue soluble in chlorinated solvents, can transfer alkali metal salts from the aqueous to the organic phase.[3] This salt transfer presumably occurs via a G-quartet assembly process driven by the extracted alkali cations.



1

In this paper, we describe an NMR study of the "octamer" (dG **1**)$_8-K^+$ in $CDCl_3$. The (dG **1**)$_8-K^+$ supramolecular complex, formed by coordination of a single K^+ ion by eight dG **1** monomers, is robust and structurally unique. The K^+-bound octamer represents the first observable stage in the aggregation of dG **1**. Although the possible octamer structures are many, this NMR analysis leads us to propose a defined species composed of two G-quartets. In one tetramer, all of the dG **1** components have a *syn* conformation around the C(1′)−N(9) glycosidic bond, while the other tetramer has an "all-*anti*" conformation. Moreover, the all-*anti* tetramer stacks in a head-to-tail orientation on top of the "all-*syn*" tetramer, presumably sandwiching a central K^+ ion. This

(17) Fisk, C. L.; Becker, E. D.; Miles, H. T.; Pinnavaia, T. J. *J. Am. Chem. Soc.* **1982**, *104*, 3307−3314.
(18) Bouhoutsos-Brown, E.; Marshall, C. L.; Pinnavaia, T. J. *J. Am. Chem. Soc.* **1982**, *104*, 6576−6584.
(19) Walmsley, J. A.; Barr, R. G.; Bouhoutsos-Brown, E.; Pinnavaia, T. J. *J. Phys. Chem.* **1984**, *88*, 2599−2605.
(20) For a review on G-quartets in oligonucleotides, see: Williamson, J. R. *Annu. Rev. Biophys. Biomol. Struct.* **1994**, *23*, 703−730.
(21) Lee, J. S. *Nucleic Acids Res.* **1990**, *18*, 6057−6060.
(22) Eimer, W.; Dorfmüller, T. *J. Phys. Chem.* **1992**, *96*, 6790−6800.
(23) Jiang, F. S.; Makinen, M. W. *Inorg. Chem.* **1995**, *34*, 1736−1744.
(24) Walmsley has shown that 5′-GMP can undergo self-association in DMSO, although specific evidence for G-quartet formation was lacking. See: Winchester, W. R.; Walmsley, J. A. *Nucleic Acids Res.* **1994**, *22*, 5128−5134.
(25) Iwahashi, H.; Kyogoku, Y. *J. Am. Chem. Soc.* **1977**, *99*, 7761−7765.
(26) Iwahashi, H.; Sugeta, H.; Kyogoku, Y. *Biochemistry* **1982**, *21*, 631−638.
(27) Williams, N. G.; Williams, L. D.; Shaw, B. R. *J. Am. Chem. Soc.* **1989**, *111*, 7205−7209.
(28) Williams, L. D.; Williams, N. G.; Shaw, B. R. *J. Am. Chem. Soc.* **1990**, *112*, 829−833.

structure for (dG **1**)$_8$−K$^+$, containing an all-*syn* tetramer, is different from all of the quartet structures described so far for G-rich oligonucleotides.

Results and Discussion

Characterization of the Species at Different Nucleoside/K$^+$ Picrate Ratios. Addition of K$^+$ salts to a $CDCl_3$ solution of 3′,5′-didecanoyl dG **1** causes diagnostic changes in the ^{1}H NMR spectrum (Supporting Information, Figure 1). The region of the ^{1}H NMR spectrum between 7.0 and 13.0 ppm, which includes resonances for the H8 aromatic proton and the imino NH1 proton, is used to characterize the assembled species in solution.[3] In the absence of K$^+$ picrate, the broad resonance for NH1 in dG **1** is near $\delta = 12.20$ ppm, and the H8 resonance is at $\delta = 7.70$ ppm. The downfield chemical shift for NH1 indicates a hydrogen-bonded species. Even in the absence of added K$^+$, dG **1** likely exists as a dimer at NMR concentrations in a noncompetitive solvent.[29]

Typically, K$^+$ picrate is insoluble in $CDCl_3$. However, the lipophilic dG **1** can extract K$^+$ picrate into $CDCl_3$, resulting in changes in the ^{1}H NMR spectrum of dG **1**. The picrate's signal at $\delta = 8.80$ ppm enables the dG **1** to K$^+$ stoichiometry to be determined by peak integration. As the K$^+$ picrate concentration that is extracted increases, the original NMR signals for dG **1** decrease and two sets of new resonances appear. These new ^{1}H NMR resonances for dG **1** reflect formation of an assembled species. The first observable species in the assembly process predominates after addition of one-eighth of an equivalent of K$^+$ picrate to a $CDCl_3$ solution of dG **1**. The spectrum of this species is characterized by two new H8 signals at 8.01 and 7.38 ppm (in a 1:1 ratio) and two new sharp NH1 imino resonances near 12.05 ppm.[3,29] These closely spaced, but separate, NH1 imino signals are shifted upfield of the original NH1 signal. Upon further addition of K$^+$ picrate, the intensity of the ^{1}H NMR signals for (dG **1**)$_8$−K$^+$ decrease and a new set of broad signals appears. Lowering the temperature of this sample causes significant line-broadening, consistent with a high molecular weight species in solution.

These NMR titrations indicate that the identity of the assembled species in $CDCl_3$ is modulated by the amount of K$^+$ picrate extracted by dG **1**. At an 8:1 dG **1** to K$^+$ picrate ratio, the data is consistent with an octamer as the predominate species in solution. The K$^+$-templated aggregation of the lipophilic dG **1** in $CDCl_3$ to give an octamer, (dG **1**)$_8$−K$^+$, parallels the Na$^+$/K$^+$-mediated self-assembly of 5′-GMP to form octamers in water.[18,19]

NMR Studies of (dG 1)$_8$−KI in $CDCl_3$. Whereas K$^+$ picrate was used to determine the stoichiometry of (dG **1**)$_8$−K$^+$ picrate, KI was used to generate (dG **1**)$_8$−KI for 2D NMR characterization. Extraction of KI from water by a $CDCl_3$ solution of dG **1** gave a complex with the same ^{1}H and ^{13}C NMR chemical shifts as (dG **1**)$_8$−K$^+$ picrate.

The ^{1}H NMR spectrum of (dG **1**)$_8$−KI in $CDCl_3$ has two sets of signals in a 1:1 ratio (Figure 2). As described below, one set of NMR signals corresponds to a dG **1** nucleoside with an *anti* conformation about the C(1′)−N(9) glycosidic bond, whereas the other set of signals is due to 50% of the didecanoyl dG **1** adopting a *syn* conformation (Scheme 1).

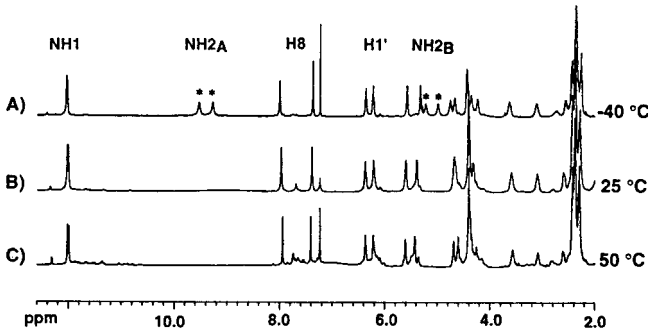
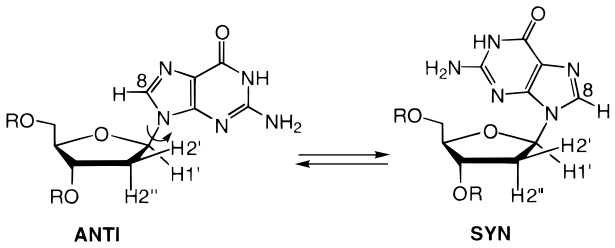


Figure 2. A region of the 500 MHz ^{1}H NMR spectra of (dG **1**)$_8$−KI (20 mM) in $CDCl_3$ at (A) −40 °C; (B) 25 °C; (C) 50 °C. Resonances for the two sets of separate amino protons, $NH2_A$ and $NH2_B$, observable at −40 °C, are labeled with asterisks.

Scheme 1



The ^{1}H NMR chemical shifts for the exchangeable and nonexchangeable protons of (dG **1**)$_8$−KI in $CDCl_3$ at −40 °C are listed in Table 1. The nonexchangeable protons were assigned from combined 2D COSY, TOCSY, and NOESY experiments. Exchangeable protons for *syn* and *anti* conformers were assigned using combinations of ^{1}H−^{1}H NOEs and ^{13}C−^{1}H correlations. The ^{13}C chemical shifts (Supporting Information, Table 1), some assigned to either *syn* or *anti* conformers, were obtained from HMBC and HMQC experiments.

The stability of (dG **1**)$_8$−KI is impressive for a noncovalent assembly. Variable temperature ^{1}H NMR showed that the K$^+$ octamer's structure is robust (Figure 2). The ^{1}H NMR spectra of (dG **1**)$_8$−KI in $CDCl_3$ are essentially temperature-independent over a 110 degree range, from −60 to 50 °C. Even at 50 °C, the K$^+$ octamer remains intact, with minimal formation of other signals. The chemical shifts of the NH1 resonances near $\delta = 12.05$ ppm are relatively invariant over this temperature range, consistent with involvement of the imino protons in strong intermolecular hydrogen bonds.

The major variation in the NMR spectra of (dG **1**)$_8$−KI at different temperatures is due to the signals for the exocyclic amino protons, $NH2_A$ and $NH2_B$. Above −10 °C the amino proton signals are not observed, being exchange-broadened into the baseline. Below −10 °C, however, the dG **1** exocyclic amino groups give two sets of well-separated ^{1}H NMR signals.

Assignment of *anti* and *syn* dG 1. The (dG **1**)$_8$−KI octamer has two sets of signals in a 1:1 ratio that do not interconvert on the NMR time scale (Supporting Information, Figure 2). The separate signals are due to two distinct conformations about the C(1′)−N(9) glycosidic bond. One species is an *anti* conformer (with χ O4′−C1′−N9−C4 between −90° and −170°), while the other species adopts a *syn* conformation (with χ between 40° and 90°) as shown in Scheme 1.[30] The presence of rigid *anti* and *syn* conformers of dG **1** is remarkable. Rotation about the

(29) Gottarelli, G.; Mariani, P.; Masiero, S.; Mezzina, E.; Spada, G. P.; Recanatini, M. *Helv. Chim. Acta* **1998**, *81*, 2078−2092.

Table 1. ^{1}H NMR (500 MHz) Chemical Shifts for (dG 1)$_8$−KI[a]

^{1}H	NH1	NH2A	NH2B	H8	H1′	H2′	H2″	H3′	H4′	H5′[b]	H5″[b]
syn	12.05	9.28	5.23	7.38	6.23	3.63	2.36	5.58	4.44	4.68	4.36
anti	12.04	9.54	4.99	8.01	6.37	3.11	2.55	5.33	4.42	4.76	4.23

[a] For a 20 mM solution of (dG **1**)$_8$−KI in $CDCl_3$ at −40 °C. [b] Stereospecific assignment of *pro*S and *pro*R protons have not been made.

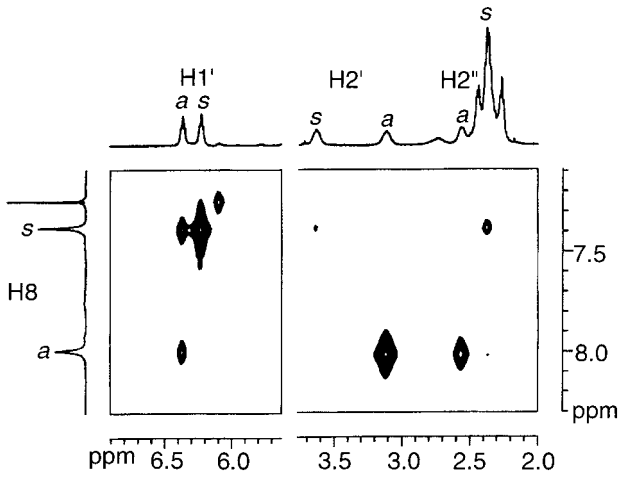


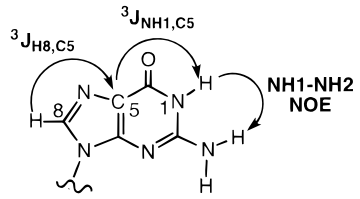
Figure 3. The H8−H1′, H2′, H2″ regions of the 2D NOESY (τ_m = 100 ms) NMR spectrum for (dG **1**)$_8$−KI in $CDCl_3$ at −40 °C. The *anti* conformer has a strong H8−H2′ NOE and medium H8−H2″ and H8−H1′ NOEs, while the *syn* conformer has a strong H8−H1′ NOE cross-peak and weak H8−H2′, H2″ NOEs.

C(1′)−N(9) glycosidic bond of (deoxy)nucleosides is typically fast on the NMR time scale, and the observed ^{1}H NMR signals are time averages of rapidly equilibrating *anti* and *syn* rotamers.[31] Moreover, the *syn* conformer is usually only predominant in purines with C(8) substituents that are large or capable of intramolecular hydrogen bonding.[31] Clearly, K^+ both templates and stabilizes the structure of the octamer (dG **1**)$_8$−KI in $CDCl_3$.

Two major criteria ($^3J_{CH}$ couplings and ^{1}H−^{1}H NOEs) were used to assign NMR peaks to either the *anti*-dG **1** or *syn*-dG **1** family. First, the magnitude of $^3J_{C(4),\,H(1')}$ and $^3J_{C(8),H(1')}$ coupling constants, evident from HMBC cross-peak intensities, allowed stereospecific assignments of *anti* and *syn* conformers. In determining glycosidic bond conformation in an oligonucleotide G-tetrad, Bax showed that purine $^3J_{CH}$ couplings follow a Karplus equation relationship.[32] Thus, the intensity of a ^{1}H−^{13}C correlation in an HMBC spectrum is proportional to the $^3J_{CH}$ coupling constant, and the magnitude of $^3J_{CH}$ is directly proportional to the glycosidic torsion angle. In the 2D HMBC spectrum (Supporting Information, Figure 3), the downfield-shifted H1′ resonance (δ 6.37 ppm) for (dG **1**)$_8$−KI in $CDCl_3$ at −40 °C has a relatively strong coupling to C(8) and a weaker ^3J coupling to C(4), consistent with an *anti* C(1′)−N(9) conformation. Conversely, the H1′ resonance at δ = 6.23 ppm was assigned as *syn* on the basis of its relatively strong $^3J_{C(4),H(1')}$ correlation and its weaker $^3J_{C(8),(H1')}$ cross-peak.[33]

Second, relative ^{1}H−^{1}H NOE cross-peak intensities further defined glycosidic bond conformations. The H8−H1′, H2′, H2″ regions of the NOESY spectrum (τ_m = 100 ms), shown in Figure 3, clearly distinguish the two H8 resonances. The *anti* conformer (H8 = δ 8.01 ppm) has a strong H8−H2′ NOE and medium strength H8−H2″ and H8−H1′ NOEs, while the *syn* conformer (H8 = δ 7.38 ppm) has a strong H8−H1′ NOE cross-peak and weak H8−H2′ and H8−H2″ NOE cross-peaks.

Scheme 2



Assignment of Exchangeable Imino NH and Exocyclic Amino Protons. The octamer (dG **1**)$_8$−K^+ has two sets of distinct signals for the exchangeable imino (NH1) and amino (NH_2) protons at −40 °C in $CDCl_3$. Although ^{1}H−^{1}H NOEs correlated both imino NH1 resonances with their respective amino signals, the exchangeable protons could not be unequivocally assigned as either *anti* or *syn* from the NOE data. Assignment of imino NH1 and amino protons in G-quartet structures is complicated by the lack of through-bond connectivities and intraresidue NOEs to the assigned sugar protons or H8 proton.[34] In this case, a ^{1}H−^{13}C HMBC experiment enabled the key assignments of exchangeable protons. As illustrated in Scheme 2, the HMBC experiment correlated the known *anti* H8 and *syn* H8 signals with their corresponding *anti* NH1 and *syn* NH1 resonances via mutual $^3J_{CH}$ coupling to C5. By using these ^{13}C−^{1}H connectivities, each exchangeable proton resonance was unequivocally assigned.

The two C5 ^{13}C resonances of (dG **1**)$_8$−K^+ were well separated (δ_{anti} C5 = 115.6 ppm and δ_{syn} C5 = 116.8 ppm). The C5 resonances were assigned from $^3J_{H8,C5}$ correlations with the known *anti* and *syn* H8 resonances. Assignment of the imino resonances as either *anti* NH1 or *syn* NH1 via $^3J_{NH1,C5}$ coupling was then possible, with δ_{anti} NH1 = 12.04 ppm and δ_{syn} NH1 = 12.05 ppm. As shown in Figure 4, the $^3J_{NH1,C5}$ cross-peaks in the 2D HMBC spectrum provided sufficient resolution to identify the closely spaced imino NH1 resonances. Having identified the NH1 resonances from the HMBC experiment, the anti and syn amino protons were then readily assigned on the basis of their strong intraresidue NH1−NH_2 NOEs.

A 2D ^{1}H−^{1}H ROESY experiment,[35] showed that the NH1−NH_2 NOE cross-peaks were indeed due to through-space magnetization transfer and not due to chemical exchange between the imino and amino protons.

(30) Wüthrich, K. *NMR of Proteins and Nucleic Acids*; John Wiley & Sons: New York, 1986; p 208.
(31) Davies, D. B. *Prog. Nucl. Magn. Reson. Spectrosc.* **1978**, *12*, 135−225.
(32) Zhu, G.; Live, D.; Bax, A. *J. Am. Chem. Soc.* **1994**, *116*, 8370−8371.
(33) Davies, D. B.; Rajani, P.; MacCross, M.; Danyluk, S. S. *Magn. Res. Chem.* **1985**, *23*, 72−77.
(34) Feigon, J.; Koshlop, K. M.; Smith, F. W. In *Methods in Enzymology*; James, T. L., Ed.; Academic Press: San Diego, 1995; Vol. 261; pp 225−255.
(35) Bothner-By, A. A.; Stephens, K. L.; Lee, J.; Warren, C. D.; Jeanloz, R. W. *J. Am. Chem. Soc.* **1984**, *106*, 811−812.

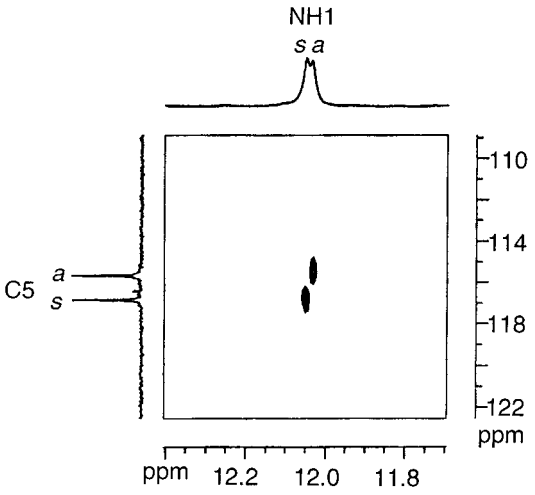
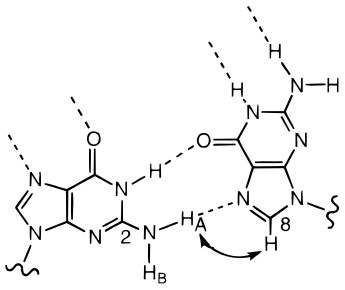


Figure 4. A region of the $^{13}C-^{1}H$ HMBC spectrum for (dG **1**)$_8$−KI in $CDCl_3$ at −40 °C showing $^3J_{C(5),NH1}$ cross-peaks. This experiment resolves the *anti* and *syn* NH1 resonances.

Scheme 3. The intraquartet NH2$_A$−H8 NOE in a G-quartet base-pair



The asterisks in Figure 2A show the two sets of dG **1** amino protons at −40 °C. The *anti* nucleoside has resonances at δ 9.54 and 4.99 ppm, while the *syn* rotamer has resonances at δ 9.28 and 5.23 ppm. The large difference for the amino protons' shifts ($\Delta\delta_{anti} = 4.55$ ppm and $\Delta\delta_{syn} = 4.05$ ppm) indicates that one of the amino protons in each pair is in a hydrogen bond, while the other is solvent-exposed.[34] The separate signals also indicate slow rotation on the NMR chemical shift time scale about the C(2)−N(2) bond for both the *anti* and *syn* rotamers, consistent with intermolecular hydrogen bonding.[27,28] The downfield-shifted resonances at δ 9.54 and 9.28 ppm were assigned as the hydrogen-bonded amino protons (NH2$_A$), and the upfield-shifted resonances at δ 4.99 and 5.23 ppm were assigned as the non-hydrogen-bonded amino protons (NH2$_B$). Assignment of the two NH1 imino protons and four exocyclic amino protons was crucial for structure determination of the (dG **1**)$_8$−KI octamer, particularly for determining the relative arrangement of the *anti* and *syn* conformers within the two G-quartets that make up (dG **1**)$_8$−K$^+$.

Structural Models for the K$^+$ Octamer. On the basis of analogy with the literature,[13,16−19] we propose a central K$^+$ cation sandwiched by two G-quartets. With a 1:1 ratio of *anti* dG **1** and *syn* dG **1**, there are many possible arrangements of the two deoxynucleoside rotamers within the octamer. Alternative arrangements include an all-*anti* dG **1** tetramer stacked on an all-*syn* dG **1** tetramer, two stacked tetramers with alternating *syn* and *anti* dG **1** residues (the alternating structure), and two stacked tetramers both containing a *syn-syn-anti-anti* dG **1** arrangement (the adjacent structure). In addition, the two tetramers can be arranged in four different relative orientations, either head-to-tail, tail-to-tail, tail-to-head, or head-to-head.[36] Some of these structural possibilities were dismissed on the basis of symmetry. Because there are only two sets of ^1H NMR signals observed for (dG **1**)$_8$−K$^+$, even at −60 °C, those arrangements that would give more than two sets of ^1H NMR signals were eliminated from further consideration.

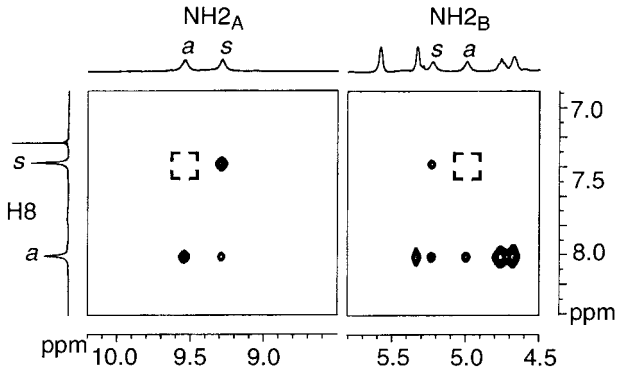


Figure 5. The H8−NH2$_{A,B}$ regions of the 2D NOESY NMR spectrum (τ_m = 100 ms) for (dG **1**)$_8$−KI in $CDCl_3$ at −40 °C. For both the hydrogen-bonded amino proton, NH2$_A$, and the non-hydrogen-bonded amino proton, NH2$_B$, (dG **1**)$_8$−KI has *anti* NH$_2$−*anti* H8, *syn* NH$_2$−*syn* H8, and *syn* NH$_2$−*anti* H8 NOEs, but no *anti* NH$_2$−*syn* H8 NOEs (illustrated by the empty boxes). The NH$_2$−H8 NOEs are consistent with one all-*anti* quartet and one all-*syn* quartet. The absence of the *anti* NH$_2$−*syn* H8 NOE rules out the alternating and adjacent *anti-syn* arrangements within an individual G-quartet.

The observed NH2$_A$−H8 NOE pattern rules out G-quartets that have either the alternating structure (*syn-anti-syn-anti*) or the adjacent structure (*syn-syn-anti-anti*). As shown in Scheme 3, the distance from the hydrogen-bonded amino proton (NH2$_A$) of one dG **1** residue to the H8 of the neighboring dG **1** in an individual base-pair is ca. 3.1−3.3 Å in the G-quartet, well within $^1H-^1H$ NOE range.[37] Thus, intraquartet NH2$_A$−H8 NOEs can be used to unambiguously define the hydrogen bonding pattern within a G-quartet.

Figure 5 shows the NH2$_A$−H8 and NH2$_B$−H8 regions of the $^1H-^1H$ NOESY spectrum (τ_m = 100 ms) for (dG **1**)$_8$−KI. The spectrum has strong *anti* NH2$_A$−*anti* H8 and strong *syn* NH2$_A$−*syn* H8 NOEs of equal intensity and a relatively weak *syn* NH2$_A$−*anti* H8 NOE. The absence of a crucial NOE, however, was as important as these three observed NOEs. The empty box in Figure 5 signifies the lack of an *anti* NH2$_A$−*syn* H8 NOE. This *anti* NH2$_A$−*syn* H8 NOE was missing even when longer mixing times of τ_m = 300 or 400 ms were used in the NOESY pulse sequence. The missing *anti* NH2$_A$−*syn* H8 NOE rules out any quartet arrangement wherein an *anti* dG **1** base-pairs with a neighboring *syn* dG **1**. The same patterns are observed for the NH2$_B$−H8 NOEs. In short, the NOE data for (dG **1**)$_8$−KI is not consistent with either the alternating structure (*syn-anti-syn-anti*) or the adjacent structure (*syn-syn-anti-anti*) within an individual G-quartet.

Figure 6 illustrates NOE patterns that would be expected if a G-quartet were to exist in either the

(36) Smith, F. W.; Lau, F. W.; Feigon, J. *Proc. Natl. Acad. Sci., U.S.A.* **1994**, *91*, 10546−10550.
(37) Phillips, K.; Dauter, Z.; Murchie, A. I. H.; Lilley, D. M. J.; Luisi, B. *J. Mol. Biol.* **1997**, *273*, 171−182.

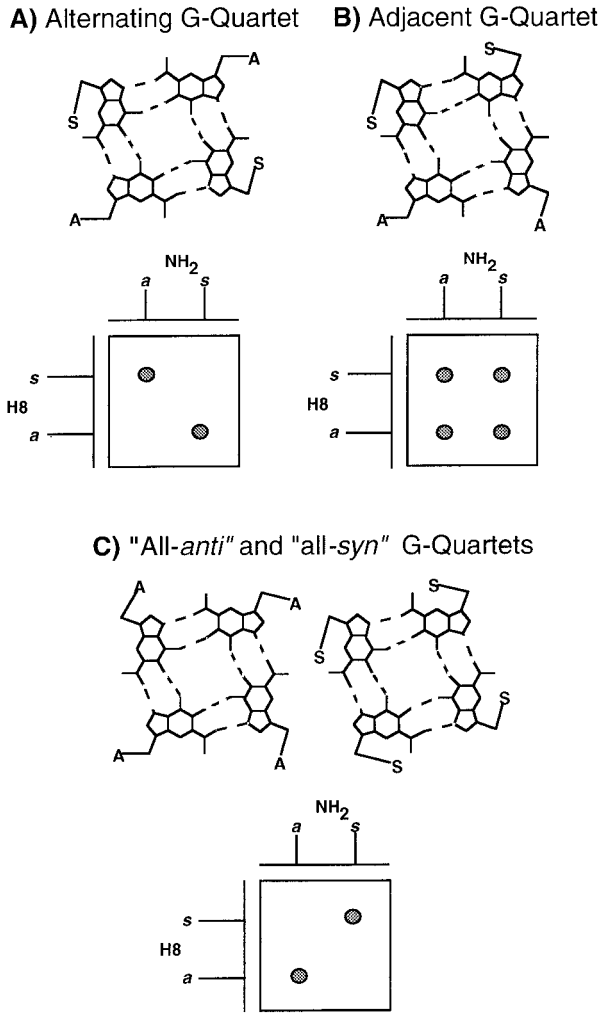


Figure 6. Theoretical $^1H-^1H$ NOE patterns for the H8−NH_{2A} region for the (A) alternating G-quartet structure, *syn-anti-syn-anti*; (B) adjacent G-quartet structure, *syn-syn-anti-anti*; and (C) combination of all-*syn* and all-*anti* G-quartet structures.

alternating structure (Figure 6A) or the adjacent structure (Figure 6B). An alternating structure should give two NOE cross-peaks (*anti* $NH2_A$−*syn* H8 and *syn* $NH2_A$−*anti* H8) of similar intensity. The NOE data in Figure 5, with its missing *anti* $NH2_A$−*syn* H8 NOE, clearly shows that the alternating structure is not present in (dG **1**)$_8$−KI. The other possible arrangement with two *syn* and two *anti* nucleosides in an individual G-quartet, the adjacent structure, should give all four $NH2_A$−H8 NOE cross-peaks of equal intensity. Again, the absence of the *anti* $NH2_A$−*syn* H8 NOE in Figure 5 discounts the adjacent structure, as does the weaker *syn* $NH2_A$−*anti* H8 NOE. Instead, the observed strong *anti* $NH2_A$−*anti* H8 and strong *syn* $NH2_A$−*syn* H8 NOEs for (dG **1**)$_8$−KI are most consistent with one G-quartet containing four *anti* dG **1** units and the other G-quartet containing four *syn* dG **1** nucleosides (Figure 6C). As discussed below, the weaker *syn* $NH2_A$−*anti* H8 NOE for (dG **1**)$_8$−KI must then arise from interquartet interactions in the stacking of the all-*anti* G-quartet and the all-*syn* G-quartet. On the basis of these $NH2_A$−H8 NOEs we conclude that the most likely structure for (dG **1**)$_8$−K^+ has an all-*anti* tetramer stacked on an all-*syn* tetramer.

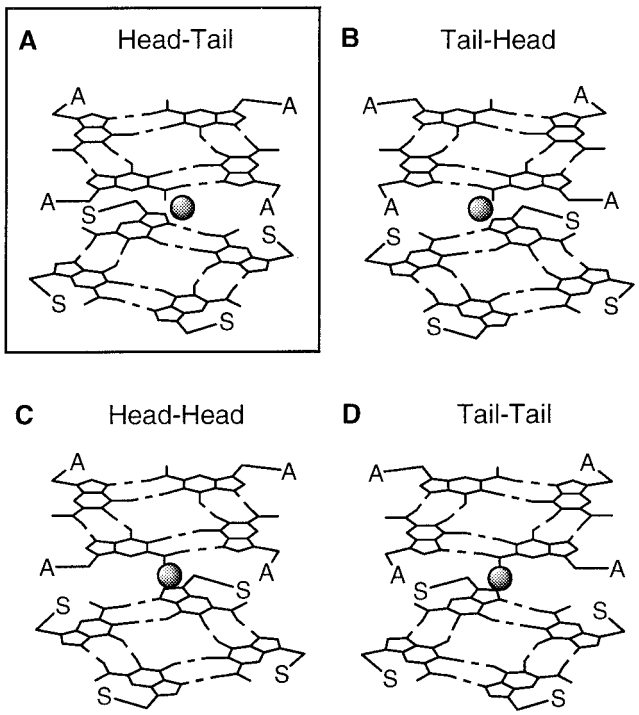


Figure 7. The four possible arrangements of (dG **1**)$_8$−K^+ containing all-*anti* and all-*syn* G-quartets. Definition of the "head" and "tail" sides of a G-quartet is as defined by Feigon.[36] The head, relative to the central K^+, has a clockwise rotation of the N−H···O=C hydrogen-bonding pattern (i.e., from the donors to the acceptors), whereas the tail has a counterclockwise rotation.

Whereas parallel-stranded DNA quadruplexes contain all-*anti* G-quartets, unimolecular and bimolecular quadruplexes always have some *syn* dG present, either as *syn-anti-syn-anti* or *syn-syn-anti-anti* tetramers.[34] In these "fold-back" G-quartet structures, some of the dG residues must adopt a *syn* glycosidic bond to accommodate an antiparallel strand orientation. Because no covalent bonds connect didecanoyl dG residues in (dG **1**)$_8$−KI, it is not obvious what factors control the all-*syn* grouping in one of the tetramers. Although an all-*syn* tetramer is less favorable for entropic reasons than an all-*anti* tetramer, it is the energy of the octamer as a whole that must be considered. One possibility is that an all-*anti* on all-*syn* arrangement for (dG **1**)$_8$−K^+ maximizes base-stacking and electrostatic interactions between the tetramers while minimizing steric repulsions. Future molecular dynamics calculations may provide the reasons for formation of this particular (dG **1**)$_8$−K^+ octamer.

The $^1H-^1H$ NOESY data for the K^+−dG octamer is most consistent with a head-to-tail arrangement of the all-*anti* and all-*syn* quartets. Unlike DNA, the supramolecular organization of a nucleoside like dG **1** is not constrained by a phosphate backbone. As noted by Gellert et al. in their study of G-quartet formation by 5′-GMP: "*When no covalent backbone exists it is no longer obligatory to stack the bases in a regular manner such that the face of one base is in contact with the back of the next base along the helix.*" [10] Theoretically, the all-*anti* dG **1** quartet and all-*syn* dG **1** quartets in (dG **1**)$_8$−K^+ can be arranged in four possible relative orientations, either head-to-tail, tail-to-tail, tail-to-head, or head-to-head.[36] Each of these arrangements should give rise to two sets

Table 2. Predicted Interquartet NOEs for the Four Possible Octamer Arrangements A−D[a]

		structure			
no.[b]	NOE	**A** (head−tail)	**B** (tail−head)	**C** (head−head)	**D** (tail−tail)
1	*a*H8−*s*NH2	+[c]		− (*s*H8−*a*H8)	− (*s*H8−*a*H8)
2	*a*H1′−*s*H8	+	− (*s*H8−*a*H2′)[b]	− (*s*H8-*a*H8)	+
3	*a*H4′−*s*H8	+	− (*s*H8−*a*H2′)	?	− (*s*H8−*a*H2′)
4	aNH$_2$−*s*H5′,5″ (both aNH$_2$ protons)	+	?	− (aNH$_2$−*s*H2′)	− (aNH$_2$−*s*H2′)
5	*a*H8−*s*H5′,5″	+	− (*a*H8−*s*H1′)		− (*a*H8−*s*H2′)

[a] Space-filling models were used to build structures **A**−**D**, the four structures for an all-*anti* quartet stacked on an all-*syn* quartet with a 45° twist between tetramers. [b] Signifies number for particular NOE in a contour plot shown in Figure 3 of the Supplemental material. [c] The designation (+) indicates that the observed NOE is consistent with a particular structure, while (−) indicates that the observed NOE is not possible for that arrangement. In parentheses after some of the (−) symbols, is the predicted NOE that should exist for that particular structure. The symbol (?) indicates that we were unsure from examining space-filling models whether the observed NOE would be expected.

of NMR signals for each resonance. These four arrangements, labeled as models **A**−**D**, are shown in Figure 7. For consistency, we always show the all-*anti* quartet stacked above the all-*syn* quartet. Space-filling models of octamers **A**−**D** were built with the two quartets oriented approximately 45° relative to each other. This twist about the central axis minimizes steric clashes between the sugars in the upper and lower quartets and maximizes base-stacking interactions. A 30−45° rotation of one G-quartet relative to the G-quartet stacked upon it has been observed by X-ray fiber diffraction of the gels of guanine derivatives.[38−40]

The relative orientation of the two quartets in (dG **1**)$_8$−KI was deduced from a number of important interquartet NOEs listed in Table 2 (a contour plot showing these NOEs is provided in the Supporting Information, Figure 4). Examination of the space-filling models **A**−**D** revealed whether any of the four different octamers were consistent with the observed interquartet NOEs. As described below, the interquartet NOEs for (dG **1**)$_8$−KI are consistent only with model **A**, an all-*anti* quartet stacked above an all-*syn* quartet in a head-to-tail arrangement.

As previously discussed, the relatively weak *syn* NH2$_A$−*anti* H8 NOE shown in Figure 5 is due to interquartet interactions and not because of proximity of these protons within a Hoogsteen hydrogen-bonded base-pair in an individual tetramer. This *syn* NH2$_A$−*anti* H8 NOE is consistent with a head-to-tail orientation. As shown in Figure 8A the *anti* dG base-pair is stacked over an exocyclic amino group of the *syn* tetramer. Thus, the *anti* H8 proton shows intraquartet NOEs to the *anti* amino group of its base-pair (intra-quartet) and interquartet NOEs to the *syn* amino group over which it is stacked. Figure 8A also shows that in a head-to-tail stacking of an all-*anti* quartet on all-*syn* quartet, the *syn* H8 and the *anti* NH2$_A$ are indeed too far away (d = 5.4 Å) to give an NOE, as confirmed by the empty box in Figure 5. As illustrated in Figure 8B, the other interquartet sugar−base interactions are also consistent with a head-to-tail stacking arrangement. Thus, interquartet *anti* H1′-*syn* H8 and *anti* H4′−*syn* H8 NOEs arise from interactions of protons on the bottom face of the *anti* sugar with *syn* H8. These NOEs are consistent only with model **A**, where the bottom face of the upper *anti* sugar directly faces the *syn* H8 in the lower tetramer. Other interquartet base−sugar NOEs support the proposed 45° twist angle about the octamer's central axis. The *anti* NH2$_{A,B}$−*syn* H5′,5″ and *anti* H8−*syn* H5′,5″ NOEs must arise from dipolar interactions of the *syn* H5′,5″ pair with two separate dG **1** monomers in the upper all-*anti* quartet.

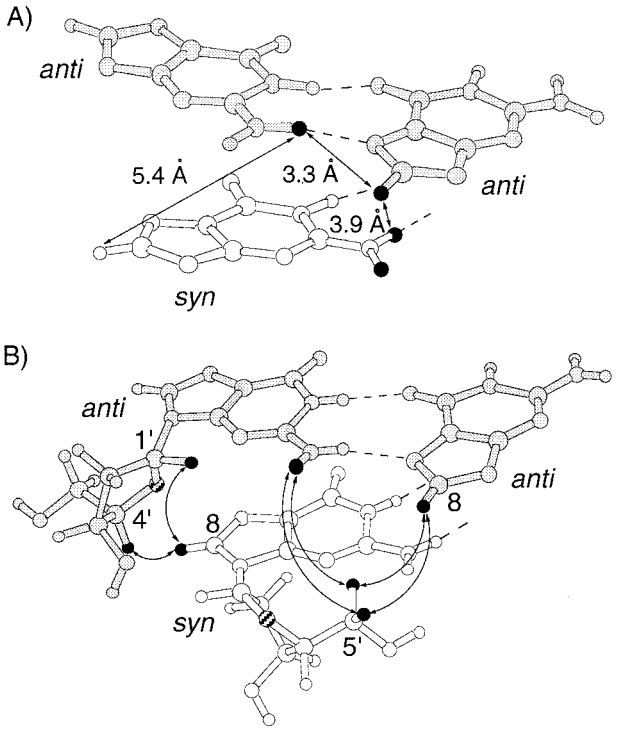


Figure 8. (A) The bases in an *anti* dG base-pair stacked over a *syn* dG residue when the quartets are stacked head-to-tail with a 45° twist. Interproton H8−NH2 distances are derived from the X-ray crystal structure of 5′-tBDMS-2′,3′-isopropylidene G, also shown to form head-to-tail octamers in the presence of K^+ picrate.[41] (B) Some interquartet ^{1}H−^{1}H NOEs for Model A from the 2D NOESY NMR spectrum (τ_m = 100 ms) for (dG **1**)$_8$−KI in $CDCl_3$ at −40 °C.

In summary, the interquartet NOEs for (dG **1**)$_8$−KI are consistent only with model **A**, an all-*anti* quartet stacked above an all-*syn* quartet in a head-to-tail arrangement. The observed interquartet NOEs are not possible for models **B**−**D**. Moreover, as listed in Table 2, NOEs other than those actually observed would be predicted for models **B**−**D**.

Conclusion

The lipophilic nucleoside 3′,5′-didecanoyl-2′-deoxyguanosine, dG **1**, had been previously shown to extract alkali metal ions from water into $CHCl_3$.[3] Nucleoside dG **1** is

(38) Tougard, P.; Chantot, J.-F.; Guschlbauer, W. *Biochim. Biophys. Acta* **1973**, *1973*, 9−16.
(39) Zimmerman, S. B.; Cohen, G. H.; Davies, D. R. *J. Mol. Biol.* **1975**, *92*, 181−192.
(40) Zimmerman, S. B. *J. Mol. Biol.* **1976**, *106*, 663−673.

also known to form ribbon aggregates in organic solvents.[29] In this work, we have used NMR data to determine a qualitative structure of the minimal soluble aggregate formed by dG **1** upon K^+ coordination in $CDCl_3$. This soluble aggregate is likely the building block for the formation of higher ordered structures. We found that dG **1** can extract salts such as K^+ picrate and KI into $CDCl_3$ to give an octamer, (dG **1**)$_8$−K^+, a remarkably stable structure for the noncovalent assembly of eight unconnected deoxynucleosides. The templating K^+ leads to the cooperative association of eight dG **1** monomers to give a 1:1 mixture of *syn* and *anti* glycosidic rotamers. The octamer's distinctive ^1H and ^{13}C NMR spectra enabled us to carry out a detailed conformational analysis on this stable aggregate. Careful 2D NMR analysis of the octamer in $CDCl_3$ has allowed us to assign the base conformation for the individual quartets. Our NMR studies indicate that the octamer (dG **1**)$_8$−K^+ is consistent with an all-*anti* G-quartet stacked on top of an all-*syn* G-quartet. Finally, intertetramer NOEs show that the two tetramers are likely to stack in a head-to-tail orientation with a ca. 45° twist about the central axis.

This all-*syn* tetramer is a unique motif for G-quartet conformation, being different from all the G-quartet structures previously described for guanine-rich oligonucleotides. Presumably, the structurally unique octamer (dG **1**)$_8$−K^+ can then further aggregate in the presence of excess K^+ to give a polymeric stack of G-quartets. Our future plans include obtaining a more quantitative solution structure of (dG **1**)$_8$−K^+ using NOE-constrained molecular dynamics. Besides providing insight into the formation of aggregates in organic solvents by dG **1**, a more detailed structural model may clarify the factors that lead to formation of this unusual all-*syn* G-quartet.

Experimental Procedures

Sample Preparation for NMR Experiments. A solution of (dG **1**)$_8$−KI was prepared by dissolving 40 mg (0.069 mmol) of 3′,5′-didecanoyl-2′-deoxyG **1**[29] in 1 mL of freshly distilled $CDCl_3$ and washing this with 3 × 1 mL of a 10 mM KI solution. The layers were separated, and the $CDCl_3$ was centrifuged to remove residual water. The sample had identical ^1H NMR chemical shifts as (dG **1**)$_8$−K^+ picrate, consistent with formation of (dG **1**)$_8$−KI. The $CDCl_3$ was evaporated under a stream of N_2 to 0.5 mL so that the final concentration of the octamer (dG **1**)$_8$−KI was 20 mM.

NMR Spectroscopy. 1D NMR data were collected on either a Bruker DRX-400 or Bruker DRX-500 spectrometer. All 2D data were collected on a DRX-500 spectrometer (500 MHz for ^1H and 125.77 MHz for ^{13}C). All spectra were taken in $CDCl_3$, and the chemical shifts are reported in parts per million (δ) relative to the nondeuterated solvent peak. The temperature was controlled to ±0.1 °C. The data were processed using XWINNMR version 2.0.

(A) 1D NMR. The 1D ^1H and ^{13}C NMR spectra were taken at various temperatures with a sweep width of 14.0 ppm for ^1H and 190.0 ppm for ^{13}C.

(B) COSY. A homonuclear ^1H−^1H COSY spectrum was recorded at −40 °C using the COSY-90 pulse sequence.[42,43] The data were collected using a 90° pulse of 9.5 μs and a relaxation delay of 3 s. The spectral width was 6983 Hz in each dimension. A total of 48 scans were collected for each time increment. A total of 512 serial files were collected, resulting in a data matrix of 512 × 2048.

(C) TOCSY. A homonuclear Hartmann−Hahn transfer spectrum was obtained at −40 °C using the MLEV-17 pulse sequence.[44] The experiment was conducted in the phase-sensitive mode using the TPPI method. The data were collected using a 90° pulse of 10.1 μs. The relaxation delay was 3 s, and the mixing time was 75 ms. The spectral width was 6944 Hz in each dimension. A total of 32 scans were collected for each time increment. A total of 512 serial files were collected, resulting in a data matrix of 512 × 4096.

(D) NOESY. The NOESY experiments were recorded at −40 °C with mixing times of 50, 100, and 200 ms using the NOESYGPTP (Bruker XWINNMR Version 2.0) pulse sequence.[45] The experiments were conducted in the phase-sensitive mode using the TPPI method. The data were collected using a 90° pulse of 10.1 μs and a relaxation delay of 2 s. The spectral width was 6983 Hz in each dimension. A total of 64 scans were collected for each time increment. A total of 512 serial files were collected, resulting in a data matrix of 512 × 2048.

(E) ROESY. The 2D ROESY experiments were done using a 200 ms spin-locking pulse with a 2.5 kHz field strength. The r_f carrier frequency was placed at 4.9 ppm to avoid problems from Hartmann−Hahn correlations. The experiments were conducted in the phase-sensitive mode using the TPPI method. The data were collected using a 90° pulse of 10.1 μs and a relaxation delay of 2 s. The spectral width was 6983 Hz in each dimension. A total of 64 scans were collected, for each time increment. A total of 512 serial files were collected, resulting in a data matrix of 512 × 2048.

(F) HMQC. A 2D ^1H−^{13}C correlation via heteronuclear zero and double quantum coherence transfer was obtained using the INVBTP pulse sequence with the BIRD sequence.[46] The data were collected in the phase-sensitive mode using the TPPI method. The ^{13}C nuclei were decoupled during the data sampling period. The data were collected using a 90° pulse of 9.5 μs for ^1H and 10.0 μs for ^{13}C. The relaxation delay was 2 s. The delay used in the BIRD sequence was optimized to 0.3 s. The spectral width was 6983 Hz in the F2 dimension (^1H) and 23 810 Hz in the F1 dimension (^{13}C). A total of 128 scans were collected for each time increment. A total of 243 serial files were collected in the F1 dimension, resulting in a data matrix of 243 × 4096.

(G) HMBC. A 2D ^1H−^{13}C correlation via heteronuclear zero and double quantum coherence transfer was recorded using the INV4LPLRND pulse sequence.[47] A low-pass J-filter was used to suppress one-bond correlations. The delay for evolution of long-range couplings was 100 ms. There was no ^{13}C decoupling during the data sampling period. The data were collected using a 90° pulse of 9.5 μs for ^1H and 10.0 μs for ^{13}C. The relaxation delay was 2 s. The spectral width was 6983 Hz in the F2 dimension (^1H) and 23 810 Hz in the F1 dimension (^{13}C). A total of 400 scans were collected for each time increment. A total of 154 serial files were collected in the F1 dimension, resulting in a data matrix of 154 × 4096.

Acknowledgment. This work was supported by a NATO Collaborative Grant to J.D. (Maryland) and G.G. (Bologna). J.D. also thanks the Dreyfus Foundation for a Teacher-Scholar Award. G.G. thanks CNR and the University of Bologna for financial support. A.L.M. thanks the Center of Biomolecular Structure and Organization at the University of Maryland for a fellowship. We thank Dr. Yui-Fai Lam and Professor Wolfgang von Phillipsborn for helpful discussions.

Supporting Information Available: 1D ^1H NMR spectra of dG **1** with varying amounts of K^+ picrate and regions of the 2D ^1H−^1H COSY, ^{13}C−^1H HMBC, and ^1H−^1H NOESY NMR spectra. This material is available free of charge via the Internet at http://pubs.acs.org.

JO9901440

(41) Davis, J. T.; Fettinger, J. F.; Forman, S., unpublished results.
(42) Aue, W. P.; Bartholdi, E.; Ernst, R. R. *J. Chem. Phys.* **1976**, *64*, 2229−2247.
(43) Nagayama, K.; Kumar, A.; Wuthrich, K.; Ernst, R. R. *J. Magn. Reson.* **1980**, *40*, 321−334.
(44) Bax, A.; Davis, D. G. *J. Magn. Reson.* **1985**, *65*, 355−360.
(45) Wagner, R.; Berger, S. *J. Magn. Reson.* **1996**, *123A*, 119−121.
(46) Bax, A.; Subramanian, S. *J. Magn. Reson.* **1986**, *67*, 565−569.
(47) Bax, A.; Summers, M. F. *J. Am. Chem. Soc.* **1986**, *108*, 2093−2094.