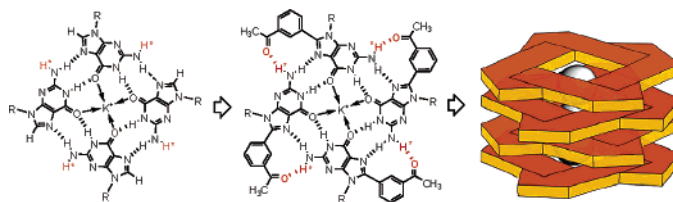# Expanding the Hoogsteen Edge of 2′-Deoxyguanosine: Consequences for G-Quadruplex Formation

Vladimir Gubala, José E. Betancourt, and José M. Rivera*

Department of Chemistry, University of Puerto Rico, Río Piedras Campus, Río Piedras, Puerto Rico 00931

jrivera@cnnet.upr.edu

The synthesis and self-assembling properties of 8-aryl-2′-deoxyguanosine derivatives are described. Our studies suggest that a properly placed acetyl group can increase the stability and specificity of the resulting G-quadruplex supramolecules by enhancing noncovalent interactions such as hydrogen bonds and π-stacking.

Advances in supramolecular chemistry and related fields rely on the availability of recognition motifs with increased selectivity, specificity, and ease of synthesis. Nucleosides in general, and guanosine (G) in particular, stand out as excellent candidates for the elaboration of a variety of supramolecular structures. Guanosine can form tetrameric structures (G-tetrads) that self-assemble in the presence of a variety of cations to form higher ordered structures known as G-quadruplexes. We describe the self-assembling properties of lipophilic 8-aryl-*d*G analogues (8Ar*d*G); in particular, we demonstrate that such properties can be modulated by the presence of an aryl group attached to C^8 at the guanine base, as well as the nature and stereochemical arrangement of functional groups capable of increasing the number of hydrogen bonds within a tetrad. We show that, in CDCl$_3$ or CD$_3$CN and in the presence of KI, 8-(3-acetylphenyl)-*d*G

forms assemblies of increased thermal stability and specificity when compared with the parent *d*G or a related compound that lack an expanded Hoogsteen edge.

G-Quadruplexes are relevant to areas as diverse as cancer research and nanotechnology.[1a] The G-rich sequence of telomeres, ends of chromosomes, have a propensity to form

(1) (a) Davis, J. T. *Angew. Chem., Int. Ed.* **2004**, *43*, 668−698. (b) Mergny, J.-L.; Riou, J.-F.; Mailliet, P.; Teulade-Fichou, M.-P.; Gilson, E. *Nucl. Acids Res.* **2002**, *30*, 839−865.

(2) (a) Shi, X.; Mullaugh, K. M.; Fettinger, J. C.; Hofstadler, S. A.; Davis, J. T. *J. Am. Chem. Soc.* **2003**, *125*, 10830−10841. (b) Rinaldi, R.; Maruccio, G.; Biasco, A.; Arima, V.; Cingolani, R.; Giorgi, T.; Masiero, S.; Spada, G. P.; Gottarelli, G. *Nanotechnology* **2002**, *13*, 398−403. (c) Shi, X.; Fettinger, J. C.; Davis, J. T. *Angew. Chem., Int. Ed.* **2001**, *40*, 2827−2831. (d) Mezzina, E.; Mariani, P.; Itri, R.; Masiero, S.; Pieraccini, S.; Spada, G. P.; Spinozzi, F.; Davis, J. T.; Gottarelli, G. *Chem. Eur. J.* **2001**, *7*, 388−395.

(3) Two notable examples for the modulation of the supramolecular behavior of G-analogues is illustrated by the work of the Gottarelli and Davis groups. The former evaluated the impact of long alkyl chains attached to the sugar ring of *d*G as a means of controlling the formation of liquid crystalline phases, and the latter demonstrated how to control the kinetics of self-assembly by changing the anion bound at the surface of the quadruplex (See refs 2d and 2a, respectively).

(4) Kettani, A.; Gorin, A.; Majumdar, A.; Hermann, T.; Skripkin, E.; Zhao, H.; Jones, R. *J. Mol. Biol.* **2000**, *297*, 627−644.

(5) Molecular modeling was performed using HyperChem 7.5 for Windows (from Hypercube, Inc.) with the Amber 99 force field.

(6) Western, E. C.; Daft, J. R.; Johnson, E. M.; Gannett, P. M.; Shaughnessy, K. H. *J. Org. Chem.* **2003**, *68*, 6767−6774.

(7) Sessler's group reported the formation of G-tetrads in the absence of a templating metal cation, both in the solid state and in solution: (a) Sessler, J. L.; Sathiosatham, M.; Doerr, K.; Lynch, V.; Abboud, K. A. *Angew. Chem., Int. Ed.* **2000**, *39*, 1300−1303. For a recent related study, see: (b) Giorgi, T.; Lena, S.; Mariani, P.; Cremonini, M. A.; Masiero, S.; Pieraccini, S.; Rabe, J. P.; Samori, P.; Spada, G. P.; Gottarelli, G. *J. Am. Chem. Soc.* **2003**, *125*, 14741−14749.

(8) (a) Borzo, M.; Detellier, C.; Laszlo, P.; Paris, A. *J. Am. Chem. Soc.* **1980**, *102*, 1124−1134. (b) Bouhoutsos-Brown, E.; Marshall, C. L.; Pinnavaia, T. J. *J. Am. Chem. Soc.* **1982**, *104*, 6576−6584. (c) Fisk, C. L.; Becker, E. D.; Miles, H. T.; Pinnavaia, T. J. *J. Am. Chem. Soc.* **1982**, *104*, 3307−3314.

G-quadruplex structures in vitro. Molecules capable of binding to and stabilizing G-quadruplexes have shown promising anticancer activity because of their inhibitory effect on telomerase.[1b] In the areas of supramolecular chemistry and nanotechnology, the Davis and Gottarelli groups have made various lipophilic G-analogues for the construction of self-assembled ionophores, self-assembled liquid crystals, and other molecular devices.[1a,2] The stated importance of G-quadruplexes makes it appropriate to determine structure–function relationships with G-analogues. In particular, G-analogues that are modified at the guanine base could be used to modulate the properties of oligonucleotides or other self-assembled structures containing them.[3]

Each guanine base in a G-tetrad has an unpaired hydrogen bond donor (H*) at N^2 that is not hydrogen bound within the structure (Figure 1a). Because H* is located at the

grooves of a G-quadruplex, it is solvent-exposed, and in aqueous solutions it is most likely H-bound to water molecules. In related structures such as the DNA-hexad, two of the four H*s are hydrogen bound to adenine bases.[4] Molecular modeling shows that a G-analogue with a hydrogen bond acceptor located near H* could make an additional H-bond.[5] A specific example of a G-analogue with an expanded Hoogsteen edge is shown in Figure 1b. The oxygen, shown in red, in the 8-(3-acetylphenyl) group could hydrogen bond to N^2H^*, forming a tetrad with up to four extra hydrogen bonds. The resulting G-quadruplexes formed from the stacking of such tetrads are expected to have additional noncovalent interactions (π-stacking, CH–π) and also to be more specific, i.e., forming mostly one type of supramolecular structure, and more stable than the naturally occurring G-quadruplexes.

The key step in the synthesis of 8ArdG analogues is a Suzuki–Miyaura cross-coupling reaction using the methodology recently published by the Shaughnessy group (Scheme 1).[6] Esterification of the 5′- and 3′-hydroxyl groups in **1**, **3**,
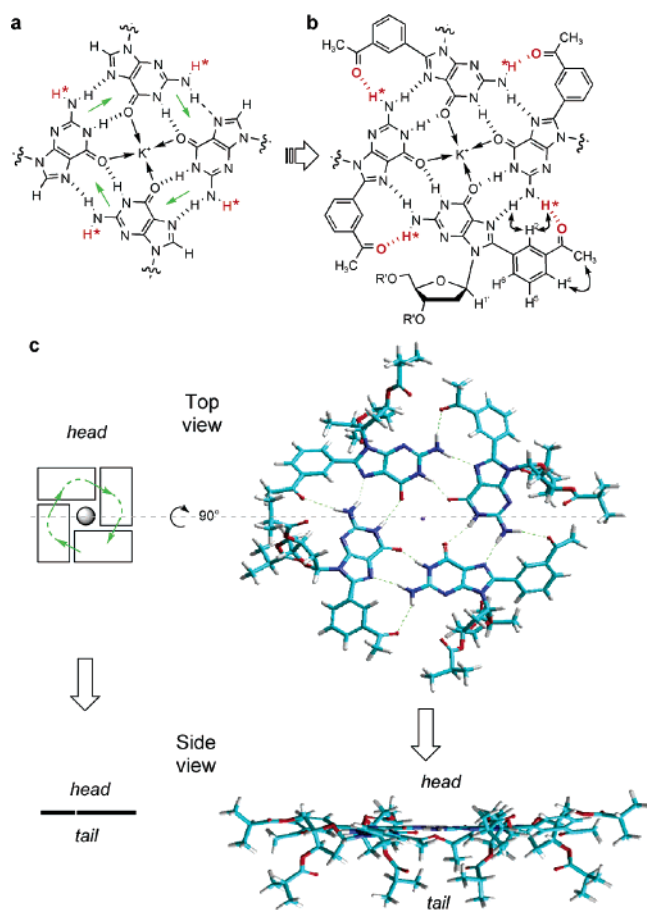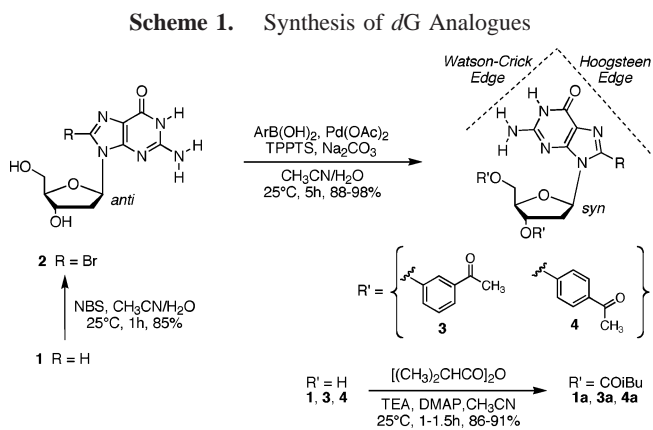

Figure 1. (a) Top view of a G-tetrad showing the directions of the hydrogen bond donors (green arrows) on the Watson–Crick edge. (b) Top view of the **3b**-tetrad in which the oxygen from the 3-acetyl group expands the Hoogsteen edge of the guanine base, allowing the formation of up to four additional hydrogen bonds per tetrad. The double-headed arrows indicate selected two-dimensional NOESY cross-peaks. (c) Top view and side view of a molecular model of the **3b**-tetrad and a cartoon representation of it. The former shows the direction of the hydrogen bond donors on the Watson–Crick edge that define the head (*h*) of the tetrad when the rotation is clockwise, as shown, and the tail (*t*) when the direction is counterclockwise.

Scheme 1. Synthesis of *d*G Analogues

and **4** affords analogues **1a**, **3a**, and **4a**, respectively, which are soluble in organic solvents.

Compounds **1a**, **3a**, and **4a** show sharp and well-defined signals in DMSO-d_6 that are indicative of the presence of mostly monomeric species (Figure 2a). In CDCl$_3$, the peaks are also sharp and well-resolved with the exception of the N^1H peaks, which are slightly exchange-broadened, indicating the formation of loosely bound aggregates (Figure 2b, Scheme 2, **M**).[7] The spectra of **1a**, **3a**, and **4a** in CDCl$_3$ changes significantly upon the addition of KI. Metal cations such as K$^+$ promote the formation of octamers, hexadecamers, and higher ordered aggregates by lipophilic guanosine analogues in organic solvents[2] and by guanosine nucleotides and oligonucleotides in aqueous environments.[8]

Titration experiments of **1a**, **3a**, and **4a** in CDCl$_3$ with KI reveal the ratio of monomers to metal cation in the quadruplex and the specificity for such quadruplexes as a function of [K$^+$] (Figure 2). Below a 1:8 ratio of K$^+$ to monomer, **1a** shows a preference for the formation of a head-to-tail (*ht*)[9] octamer (O), but beyond a 1:8 ratio, the spectrum becomes more complex with the emergence of several species in
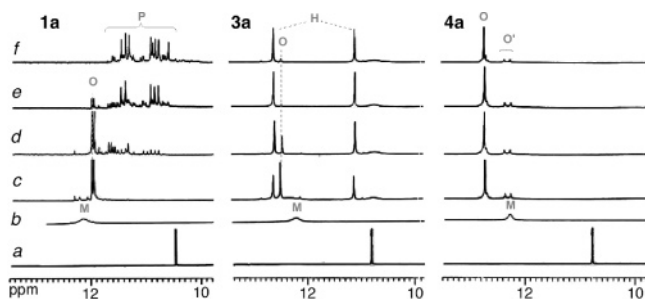
Figure 2. Titration experiments (295 K, 500 MHz) for compounds **1a**, **3a**, and **4a** in (a) DMSO-d_6 (36, 32, 35 mM) and (b–f) CDCl$_3$ (58, 57, 52 mM) with the following equivalents of KI relative to monomer: (b) 0, (c) 1/32, (d) 1/8, (e) 1/6, (f) 1/2. The portion of the ^{1}H NMR spectra shows the region of the N^1–H peaks. Key: **M** (loosely organized monomers), **O** (octamer), **H** (hexadecamers), **P** (stacked polymers), as shown in Scheme 2.

Scheme 2. Hierarchical Self-Assembly for G-Quadruplexes

loosely organized monomers (M)   octamer (O)   hexadecamer (H)   stacked polymers (P)
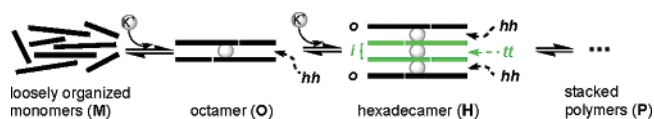
different proportions (Figure 2d: **1a**). This behavior is consistent with a shift in the equilibrium from an octamer toward stacked polymers (Scheme 2, **P**), which agrees with the behavior of related *d*G analogues reported by the Gottarelli group.[2d]

Upon titration of **3a** with KI, two new species appear, a species with one set of peaks and a second species with two sets of peaks (Figure 2: **3a**). When [KI] ≥ 3/16 equiv, the main species (>95%) is the one with two sets of peaks. This behavior is consistent with the formation of a hexadecamer (H) with a formula (**3a**)$_{16}$·3K$^+$ and an octamer (O) with a formula (**3a**)$_8$·K$^+$. On the other hand, as we increase the concentration of KI, **4a** shows the formation of primarily a head-to-head (*hh*) octamer (O) (**4a**)$_8$·K$^+$ with a small amount of a head-to-tail (*ht*) octamer (O′). Molecular modeling and two-dimensional NOESY experiments of (**3a**)$_{16}$·3K$^+$ and (**4a**)$_8$·K$^+$ reveal that the head (*h*) of the tetrad is less crowded than its tail (*t*) because the monomers **3a** and **4a** stay in the syn conformation around the glycosidic bond (Scheme 2).[7,10] Those experiments also indicate that the interfaces between the tetrads are *hh–tt–hh* for (**3a**)$_{16}$·3K$^+$ and *hh* for (**4a**)$_8$·K$^+$ as shown in Scheme 2.

The increased stability of the quadruplex formed by **3a** over those formed by **1a** and **4a** was determined using ^{1}H NMR through variable-temperature (VT) and dilution experi-

(9) See Figure 1c for the definition of the head and the tail of a tetrad. Also, see: Smith, F. W.; Lau, F. W.; Feignon, J. *Proc. Natl. Acad. Sci. U.S.A.* **1994**, *91*, 10546−10550.

(10) (a) Michelson, A. M.; Monny, C.; Kapuler, A. M. *Biochim. Biophys. Acta* **1970**, *217*, 7. (b) Rao, S. N.; Kollman, P. A. *J. Am. Chem. Soc.* **1986**, *108*, 3048−3053. (c) Dias, E.; Battiste, J. L.; Williamson, J. R. *J. Am. Chem. Soc.* **1994**, *116*, 4479−4480.

ments. VT experiments over a 112 K temperature range reveal the increased thermal stability of (**3a**)$_{16}$·3K$^+$ when compared with (**1a**)$_n$·mK$^+$ and (**4a**)$_8$·K$^+$ (Figure 3a−c). At
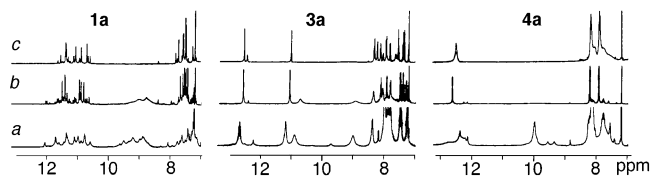

Figure 3. Variable-temperature experiments for compounds **1a**, **3a**, and **4a** in CDCl$_3$ with 1/6 equiv of KI. The temperatures were (a) 221, (b) 288, and (c) 333 K.

221 and 333 K, the main species is still (**3a**)$_{16}$·3K$^+$, whose ^{1}H NMR peaks broaden at 221 K, but remain sharp and show all the multiplicities at 333 K. The spectra of (**4a**)$_8$·K$^+$, in which the 4-acetyl group cannot form additional hydrogen bonds, shows the collapse of the main species at 221 K; at 333 K, the peaks are broad and show no multiplicities, both of which are indicative of a more dynamic system (Figure 3a−c: **4a**).

The dilution experiments in CD$_3$CN allowed us to evaluate the ratio (*f*) of ordered species formed by **3a**, **4a**, and **1a**, which is useful to make comparisons between the stabilities of (**3a**)$_{16}$·3K$^+$, (**4a**)$_8$·K$^+$, and (**1a**)$_n$·mK$^+$ (Table 1). In the

Table 1. Fraction of Ordered Species (*f*)[a] as a Function of Total Concentration (G_T) of Nucleosides **1a**, **3a**, or **4a** in CD$_3$CN[b] at 295 K

G_T	*f* **1a**	*f* **3a**	*f* **4a**
44	0.84	0.96	0.94
25	0.84	0.95	0.92
20	0.81	0.95	0.90
15	0.81	0.94	0.88
10	0.63	0.92	0.88
5	0.50	0.87	0.79
2	0.15	0.73	0.56
1	0	0.50	0.31
0.5	0	0.24	0.06

[a] Ratio $f = [X_O]/[X_T]$, where $[X_O]$ and $[X_T]$ represent the concentration in ordered form (i.e., quadruplexes) and the total concentration of **1a**, **3a**, or **4a**, respectively. [b] In CD$_3$CN, both monomeric and assembled species can be clearly distinguished in the ^{1}H NMR spectra.

concentration range between 44 and 10 mM, *f* remains effectively constant for both **3a** and **4a**, but below 10 mM it is significantly larger for **3a**. In contrast, *f* is lower for **1a** throughout the entire concentration range examined. These results indicate that an aryl group at C^8 and a properly placed functional group (capable of increasing the number of hydrogen bonds within a tetrad) impart stability to the resulting quadruplexes when compared to the parent compound.

The visibility of separate peaks for N^2H^{WC} and N^2H^* at temperatures of up to 300 K suggests slower rotation around the C^2-N^2 bond caused by the formation of additional hydrogen bonds.[11] Cross-peaks in the two-dimensional NOESY spectrum of $(3a)_{16} \cdot 3K^+$ measured at 295 K between H^2 and both N^2H^{WC} and N^2H^* and between the acetyl-CH_3 and H^4 (Figure 1b) confirm that the monomers in the outer (o) tetrads are in a suitable conformation to make the extra hydrogen bonds with NH* (Scheme 2, **H**).[12] According to the same two-dimensional NOESY spectrum, the monomers in the inner (i) tetrads are prevented from making such hydrogen bonds because the 8-aryl group is deviated significantly from the plane of the tetrad (Scheme 2, **H**). The increased steric crowding at the *tt* interface must be forcing the 8-aryl group to adopt this alternative conformation as a means of minimizing steric repulsion and maximizing other noncovalent interactions. The chemical shifts for the N^2H^* signals in $(3a)_{16} \cdot 3K^+$ (outer tetrad, inner tetrad) and $(4a)_8 \cdot K^+$ are 7.8, 5.5, and 5.7 ppm, respectively, at 278 K. The significant downfield chemical shift between the former ($\Delta \delta > 2$ ppm) and the latter two is also a strong indicator for the formation of the putative hydrogen bond.

The 8Ar*d*Gs described in this article offer an attractive way of obtaining easily prepared recognition motifs that can be used to modulate the supramolecular properties (stoichiometry, specificity, stability) of G-quadruplexes. The evidence presented suggests that such modulation results from a combination of the following factors: (i) the preorganization of the 8Ar*d*Gs in the syn conformation caused by substitution at C^8; (ii) the larger surface area of the modified base, which should increase π-stacking interactions; (iii) the expansion of the Hoogsteen edge obtained by strategically placing functional groups capable of forming extra hydrogen bonds. All of these effects combine to increase the stability and specificity of the resulting quadruplexes. More detailed studies with other 8Ar*d*Gs are underway to determine why the 3-acetyl group in **3a** shifts the equilibrium toward a hexadecamer as well as the scope and limitations of this strategy. The results of such studies will be presented soon.

Supporting Information Available: Synthetic procedures and titration, VT, and two-dimensional NOESY spectra. This material is available free of charge via the Internet at http://pubs.acs.org.

OL048013V

(11) Williams, L. D.; Williams, N. G.; Shaw, R. *J. Am. Chem. Soc.* **1990**, *112*, 829−833.

(12) Assignment of outer (o) and inner (i) tetrads is based on the large upfield chemical shift displacement ($\Delta \delta$ up to 1.5 ppm for some peaks, Figure 2d: **3a**) of the latter. This large $\Delta \delta$ is most likely the result of the anisotropic shielding created by the aromatic rings of the outer (o) tetrads. See also Supporting Information, p S9, for more details.