

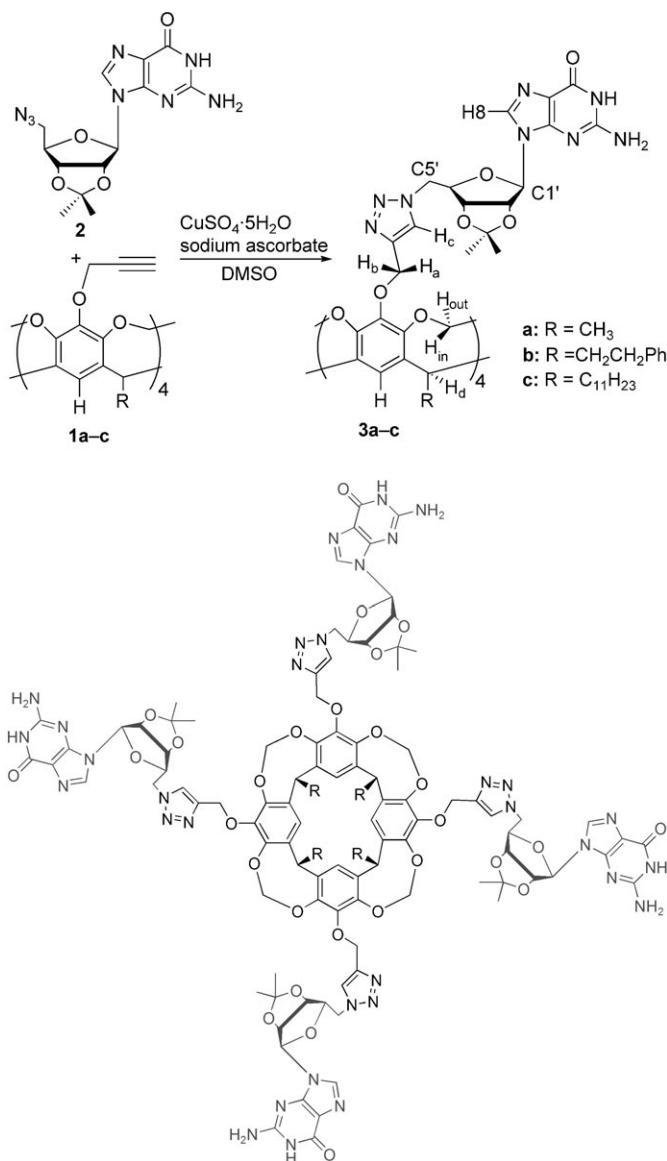
Template-Assembled Synthetic G-Quartets (TASQs)**

Mehran Nikan and John C. Sherman*

Organization of small molecules into well-defined assemblies is one of the challenges of supramolecular chemistry.^[1] A biologically relevant assembly that lends itself well to synthetic supramolecular study is the G-quartet, which is a H-bonded structure composed of four Hoogsteen-paired guanine bases.^[2] Guanine-rich sequences are abundant in telomeric ends of chromosomes and promoter regions of DNA, and are capable of forming G-quartets in vitro.^[3]

Guanine self-assembly in lipophilic systems has been the focus of much research in the past and has been reviewed in detail.^[4] Guanines have been linked to calixarenes^[5] for structural or recognition purposes, and synthetic hydrophilic unimolecular G-quartet assemblies have been reported.^[6] G-quartets are typically templated and stabilized by cations,^[7] whereas guanine aggregation in the absence of cations generally results in the formation of ribbonlike structures.^[8] Cation-free G-quartets are rare due to the repulsion of the coplanar carbonyl groups and the high stability of less-ordered polymeric ribbons.^[9] Herein, we introduce a new class of compounds, guanine-linked cavitanes, and propose a general term for them, template-assembled synthetic G-quartets (TASQs), analogous to the term template-assembled synthetic proteins (TASPs) created by Mutter.^[10] The lipophilic TASQs reported herein were synthesized by click chemistry,^[11] and manifest unusual cation-independent stability. This stability is likely due to the preorganization afforded by the cavitanne scaffold, thus exemplifying one of the hallmarks of supramolecular chemistry.^[12] These TASQs link the chemistry of G-quartets to that of cavitanes and offer potential opportunities, including the creation of singular G-quartet baskets that are stable at low concentrations and in the absence of cations.

Compounds **3a–c** were synthesized in 62–66 % yield from cavitanes **1a–c**^[13] and 5'-azido-2',3'-O-isopropylidene-guanosine (**2**)^[14] (Scheme 1). ¹H NMR spectroscopic data for **3c** in [D₆]DMSO and CDCl₃ are given in Figure 1 and Table 1 (see the Supporting Information for complete assignments). The sugar protons were identified by their correlations to adjacent protons starting from H1'. The diastereotopic H5', H_a/H_b, and



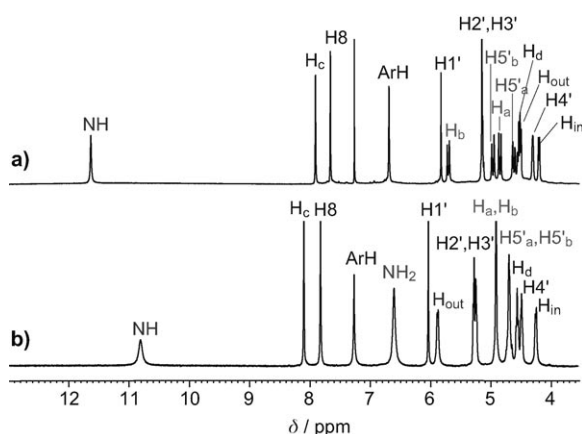


Figure 1. 400 MHz ^1H NMR spectra of **3c** at 25 °C a) in CDCl_3 , b) in $[\text{D}_6]\text{DMSO}$.

Table 1: Spectral assignments of **3c** in $[\text{D}_6]\text{DMSO}$ and CDCl_3 at ambient temperature.

Proton	$\delta_{\text{H}}([\text{D}_6]\text{DMSO})^{[a]}$	$\delta_{\text{H}}(\text{CDCl}_3)$	$^1\text{H}-^1\text{H}$ COSY ^[b,c]	$^1\text{H}-^{13}\text{C}$ HMQC ^[b]
H1'	6.03	5.82	H2'*	89.17
H2'	5.27	5.14	H3', H1'*	78.44
H3'	5.24	5.13	H2', H4'	83.57
H4'	4.48	4.30	H3', H5'a*	82.90
H5'a	4.69	4.61	H5'b, H4'*	49.19
H5'b	4.69	4.96	H5'a	49.19
NH	10.78	11.63	—	—
NH _{2a}	6.59	broad	—	—
NH _{2b}	6.59	broad	—	—
H8	7.82	7.65	—	136.51
ArH	7.26	6.68	—	113.89
H _{in}	4.24	4.19	H _{out}	98.40
H _{out}	5.88	4.50	H _{in}	98.40
H _a	4.91	4.84	H _b	70.60
H _b	4.91	5.70	H _a	70.60
H _c	8.10	7.90	—	126.13
H _d	4.55	4.53	CH ₂ (feet)	37.17

[a] The signals of diastereotopic protons overlap in $[\text{D}_6]\text{DMSO}$. [b] 2D data acquired for a 2×10^{-2} M solution of the sample in CDCl_3 . [c] The asterisk indicates weak COSY cross-peaks.

A G-quartet appears to form when **3c** is dissolved in CDCl_3 , even in the absence of cations. Such a species is highly unusual, and thus a detailed account is in order. The imino (NH) signal shifts downfield to 11.63 ppm, indicating a H-bonded system.^[16] At low temperatures, the NH_2 signal appears as two distinct singlets, one at $\delta = 9$ and one at $\delta = 4.9$ ppm (Figure 2a and Supporting Information), corresponding to H-bonded and non-H-bonded protons, respectively.^[17] At -40°C , a 2D NOESY spectrum yields a cross-peak between H8 and NH_{2b} (Figure 2a, top), a correlation that has been used to authenticate a G-quartet assembly.^[18] Moreover, a strong (i.e., intraresidue) NOE between H8 and H1' was observed, which is indicative of a *syn* conformation along the glycosidic bond (Figure 2b).^[19] *Syn* conformations are known to prevent the formation of G-ribbons.^[9a] NOEs between amino and imino protons (Figure 2a, bottom) indicate Hoogsteen-paired guanine bases.^[17,20] Taken

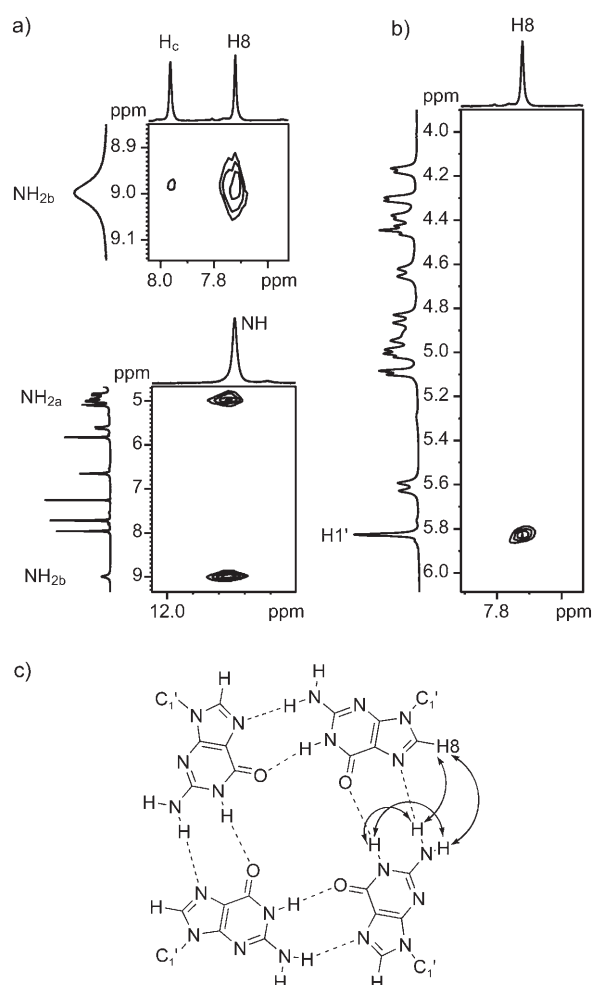


Figure 2. NOE effects indicative of a) the formation of a G-quartet, b) the *syn* conformation at 400 MHz in CDCl_3 at -40°C . c) Inter- and intrabase NOE effects in a G-quartet.

together, these results suggest that **3c** spontaneously forms a G-quartet in CDCl_3 .

The non-exchangeable protons also exhibit changes consistent with the formation of a G-quartet. The H_{out} proton of the cavatand undergoes a significant upfield shift in CDCl_3 relative to DMSO ($\Delta\delta = -1.38$ ppm; Figure 1), which suggests a crowding of the upper rim of the cavatand by the aromatic guanine residues in CDCl_3 . Diastereotopic protons H_a and H_b , which appear as one broad signal in DMSO , give a set of doublets in CDCl_3 , one of which exhibits a considerable downfield shift ($\Delta\delta = 0.79$). Examination of CPK molecular models suggests that this proton (H_b) is relegated to outside of the anisotropic current of the aromatic rings upon formation of a G-quartet.

As to kinetic stability, above 30°C , the rate of rotation about the amino C–N bond is fast on the ^1H NMR time scale, as an average signal is apparent for the NH_2 protons (see the Supporting Information).^[21] This kinetic stability for cation-free TASQ **3c** is comparable to that of some cation-bound structures.^[16] As to thermodynamic stability, there is only a small change ($\Delta\delta = -0.2$ ppm) in the chemical shift of the imino (NH) signal over a 100 K temperature range (-50 to

+50°C). This indicates that the H-bonding remains largely intact even at 50°C in CDCl₃.

Cations contribute to the stability and polymorphism of G-quadruplex structures. They can induce structural changes or trigger conformational transitions.^[22] Similar observations have been made in lipophilic systems.^[23] Thus, we investigated the recognition of TASQ **3c** with different cations. Extraction of solid sodium picrate by a CHCl₃ solution of **3c**, for example, induced changes in the ¹H NMR spectrum of **3c**.^[24] At low temperature the signal for the H-bonded amino group shifted from δ =9 ppm (Na⁺-free) to δ =10 ppm in the presence of Na⁺ (see the Supporting Information). This observation supports the notion that the former system is indeed cation-free and that **3c** recognizes common G-quartet stabilizing cations.^[25]

In biological systems, cation templation is the key stabilizing element of a G-quadruplex. H-bonding, hydrophobic interactions and the phosphodiester backbone are other factors important for stabilizing a G-quadruplex. Likewise, in lipophilic systems, cation templation overcomes the repulsive interaction of the carbonyl oxygen atoms in the central core of a G-quartet. Little attention has been paid to the role of an external backbone or templating scaffold. This study provides a model system of how a lipophilic G-quartet can be designed and synthesized with the help of an external template. This is an unusually stable cation-free G-quartet whose scaffold-induced unimolecularity provides structural integrity even at low concentrations. These findings suggest potential applications for future TASQs, for example as G-quartet aptamers or as G-quartet recognizing protein screens. Current efforts include exploration of cation-bound morphologies of lipophilic TASQs, and creation of hydrophilic TASQs.

Received: September 11, 2007

Revised: April 24, 2008

Published online: May 21, 2008

Keywords: cavitands · G-quartets · nucleosides · supramolecular chemistry · template synthesis

- [1] J. M. Lehn, *Angew. Chem.* **1990**, *102*, 1347–1362; *Angew. Chem. Int. Ed. Engl.* **1990**, *29*, 1304–1319.
- [2] M. Gellert, M. N. Lipsett, D. R. Davies, *Proc. Natl. Acad. Sci. USA* **1962**, *48*, 2013–2018.
- [3] E. Henderson, C. C. Hardin, S. K. Walk, I. Tinoco, E. H. Blackburn, *Cell* **1987**, *51*, 899–908.
- [4] a) J. T. Davis, *Angew. Chem.* **2004**, *116*, 684–716; *Angew. Chem. Int. Ed.* **2004**, *43*, 668–698; b) J. T. Davis, G. P. Spada, *Chem. Soc. Rev.* **2007**, *36*, 296–313.
- [5] a) G. M. L. Consoli, G. Granata, E. Galante, F. Cunsolo, C. Geraci, *Tetrahedron Lett.* **2006**, *47*, 3245–3249; b) C. C. Zeng, Q. Y. Zheng, Y. L. Tang, Z. T. Huang, *Tetrahedron* **2003**, *59*, 2539–2548; c) S. J. Kim, B. H. Kim, *Nucleic Acids Res.* **2003**, *31*, 2725–2734; d) V. Sidorov, F. W. Kitch, M. El-Khouedi, J. T. Davis, *Chem. Commun.* **2000**, 2369–2370.
- [6] a) G. Oliviero, N. Borbone, A. Galeone, M. Varra, G. Piccialli, L. Mayol, *Tetrahedron Lett.* **2004**, *45*, 4869–4872; b) M. S. Kaucher, W. A. Harrell, J. T. Davis, *J. Am. Chem. Soc.* **2006**, *128*, 38–39.
- [7] T. J. Pinnavaia, C. L. Marshall, C. M. Mettler, C. I. Fisk, H. T. Miles, E. D. Becker, *J. Am. Chem. Soc.* **1978**, *100*, 3625–3627.
- [8] T. Giorgi, F. Grepioni, I. Manet, P. Mariani, S. Masiero, E. Mezzina, S. Pieraccini, L. Saturni, G. P. Spada, G. Gottarelli, *Chem. Eur. J.* **2002**, *8*, 2143–2152.
- [9] a) J. L. Sessler, M. Sathiosatham, K. Doerr, V. Lynch, K. A. Abboud, *Angew. Chem.* **2000**, *112*, 1356–1359; *Angew. Chem. Int. Ed.* **2000**, *39*, 1300–1303; b) F. W. Kitch, V. Sidorov, Y. F. Lam, K. J. Kayser, H. Li, M. S. Kaucher, J. T. Davis, *J. Am. Chem. Soc.* **2003**, *125*, 15140–15150; c) R. Otero, M. Schock, L. M. Molina, E. Laegsgaard, I. Stensgaard, B. Hammer, F. Besenbacher, *Angew. Chem.* **2005**, *117*, 2310–2315; *Angew. Chem. Int. Ed.* **2005**, *44*, 2270–2275.
- [10] M. Mutter, *Angew. Chem.* **1985**, *97*, 639–654; *Angew. Chem. Int. Ed. Engl.* **1985**, *24*, 639–653.
- [11] V. V. Rostovtsev, L. G. Green, V. V. Fokin, K. B. Sharpless, *Angew. Chem.* **2002**, *114*, 2708–2711; *Angew. Chem. Int. Ed.* **2002**, *41*, 2596–2599.
- [12] a) R. C. Helgeson, B. J. Selle, I. Goldberg, C. B. Knobler, D. J. Cram, *J. Am. Chem. Soc.* **1993**, *115*, 11506–11511; b) G. M. Whitesides, E. E. Simanek, J. P. Mathias, C. T. Seto, D. N. Chin, M. Mammen, D. M. Gordon, *Acc. Chem. Res.* **1995**, *28*, 37–44.
- [13] a) E. S. Barrett, J. L. Irwin, K. Picker, M. S. Sherburn, *Aust. J. Chem.* **2002**, *55*, 319–325; b) D. J. Cram, R. Jaeger, K. Deshayes, *J. Am. Chem. Soc.* **1993**, *115*, 10111–10116.
- [14] M. G. Stout, M. J. Robins, R. K. Olsen, R. K. Robins, *J. Med. Chem.* **1969**, *12*, 658–662.
- [15] R. A. Newmark, C. R. Cantor, *J. Am. Chem. Soc.* **1968**, *90*, 5010–5017.
- [16] F. W. Smith, J. Feigon, *Nature* **1992**, *356*, 164–168.
- [17] A. L. Marlow, E. Mezzina, G. P. Spada, S. Masiero, J. T. Davis, G. Gottarelli, *J. Org. Chem.* **1999**, *64*, 5116–5123.
- [18] a) Y. Wang, D. J. Patel, *Biochemistry* **1992**, *31*, 8112–8119; b) X. Y. Liu, I. C. M. Kwan, S. N. Wang, G. Wu, *Org. Lett.* **2006**, *8*, 3685–3688.
- [19] D. J. Patel, S. A. Kozlowski, A. Nordheim, A. Rich, *Proc. Natl. Acad. Sci. USA* **1982**, *79*, 1413–1417.
- [20] F. Aboul-ela, A. I. H. Murchie, D. M. J. Lilley, *Nature* **1992**, *360*, 280–282.
- [21] L. D. Williams, N. G. Williams, B. R. Shaw, *J. Am. Chem. Soc.* **1990**, *112*, 829–833.
- [22] a) S. Neidle, S. Balasubramanian, *Quadruplex Nucleic Acids*, RSC Publishing, Cambridge, **2006**; b) D. Sen, W. Gilbert, *Nature* **1990**, *344*, 410–414.
- [23] G. Gottarelli, S. Masiero, G. P. Spada, *J. Chem. Soc. Chem. Commun.* **1995**, 2555–2557.
- [24] After analogous extraction of K⁺, Sr²⁺, and Cs⁺ picrates, distinct signals (from **3c** alone) were observed.
- [25] Addition of [2.2.2]cryptand to the solution of **3c**·Na⁺ resulted in a ¹H NMR spectrum identical to the spectrum of the cation-free species. Addition of [2.2.2]cryptand to cation-free **3c** resulted in no change (see Figure S15). Atomic absorption experiments indicate that there are 250 ppm of Na⁺ present in **3c**, which is less than 3 mol%. This agrees well with the NMR results (see Figure S2), and confirms cation-free **3c** as a different entity from the sodium-bound system (see the Supporting Information).