J|A|C|S COMMUNICATIONS

Published on Web 01/29/2009

Helical Structure of Disodium 5′-Guanosine Monophosphate Self-Assembly in Neutral Solution

Gang Wu* and Irene C. M. Kwan

Department of Chemistry, Queen's University, 90 Bader Lane, Kingston, Ontario, Canada K7L 3N6

Received November 26, 2008; E-mail: gang.wu@chem.queensu.ca

Among the fundamental subunits of nucleic acids, guanosine 5′-monophosphate (5′-GMP) has the unique ability to self-associate spontaneously in either acidic[1,2] or neutral[3−11] aqueous solutions, forming ordered helical structures. Although it was suggested in 1962 that a hydrogen-bonded guanine tetramer known as the G-quartet[2] (Figure 1a) is the basic building block of such helices, their exact structures have remained unsolved for more than four decades. Here we use nuclear magnetic resonance (NMR) spectroscopic methods to investigate the detailed structures formed by Na_2(5′-GMP) self-assembly in neutral solution. We have found that three types of 5′-GMP aggregates generally coexist in solution, using monomers, dimers, and G-quartets as basic building blocks, respectively. The dimer formation is based on the centrosymmetric structure denoted as GG3[2] (Figure 1a) in the Jeffrey and Saenger notation.[12] Most interestingly, the G-quartets stack on top of each other forming a right-handed helix where alternating C2′-*endo* and C3′-*endo* sugar puckers are found along the helical strand.

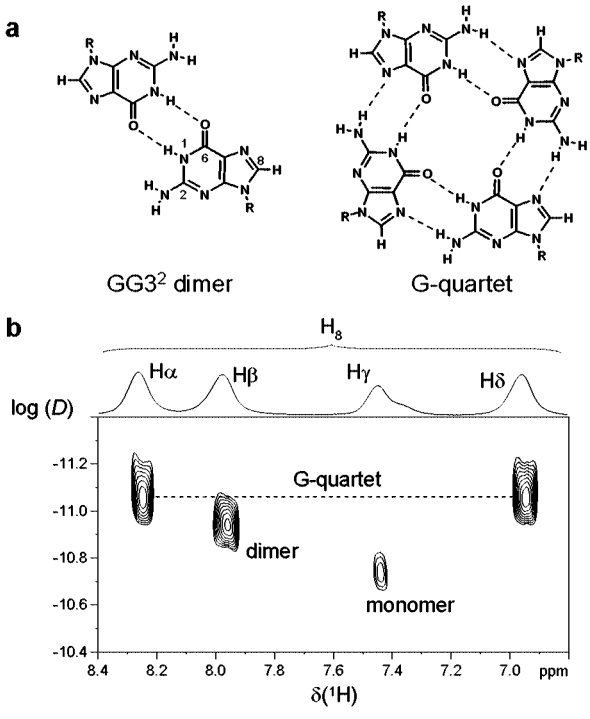

Figure 1. (a) Structures of GG3[2] dimer and G-quartet. (b) H_8 region of a ^{1}H 2D DOSY NMR spectrum of 1.0 M Na_2(5′-GMP) in D_2O at 278 K.

It has been well-known that the H_8 region of the ^{1}H NMR spectrum of a concentrated Na_2(5′-GMP) solution exhibits four major signals (H_α, H_β, H_γ, and H_δ as shown in Figure 1b). Pinnavaia and colleagues[7,8,11] interpreted these signals as being due to the presence of C_4 and D_4 stereoisomers of a G-octamer. However, we recently showed that the size of Na_2(5′-GMP) self-aggregates is on the nanometer scale, much larger than that of a G-octamer.[13] This finding immediately called for a new spectral interpretation and ultimately led to the complete structural determination reported herein. To establish the exact identities of these ^{1}H NMR signals, we first used diffusion-ordered spectroscopy (DOSY)[14] NMR experiments. As seen in Figure 1b, the H_α and H_δ signals are associated with the same molecular aggregate that has a much smaller translational diffusion coefficient (D), $(8.8 \pm 0.5) \times 10^{-12}$ m^2/s at 278 K, than those giving rise to the H_β and H_γ signals. The D value observed for the H_β signal, $(11.6 \pm 0.5) \times 10^{-12}$ m^2/s, suggests that this aggregate is due to the stacking of 5′-GMP dimers (*vide infra*). The H_γ signal exhibits the largest D, $(17.2 \pm 0.5) \times 10^{-12}$ m^2/s, and arises from the stacking of monomers.[13] Because of the presence of four sets of very similar signals, ^{1}H NMR spectra of Na_2(5′-GMP) are extremely overcrowded in the region containing sugar proton resonances, making it difficult to use a conventional NMR approach for structural determination. To gain structural information about the H_α/H_δ aggregates, we employed a combined DOSY and NOESY approach.[15] Figure 2a shows parts of the 2D ^{1}H DOSY-NOESY spectrum of 1.0 M Na_2(5′-GMP) in D_2O. Because the signals from 5′-GMP dimers and monomers are diffusively "filtered out" in this spectrum, two sets of ^{1}H resonances can be clearly identified. Complete spectral assignment for sugar ^{1}H resonances was established using ^{1}H DQF-COSY and DOSY-NOESY spectra. Homonuclear ($^3J_{HH}$) and heteronuclear ($^1J_{CH}$ and $^3J_{PH}$) indirect spin−spin coupling constants were measured from ^{1}H DQF-COSY, ^{1}H−^{13}C HSQC, and ^{1}H−^{31}P COSY spectra, respectively. Stereochemical assignment of the H5′ and H5" signals was achieved using $^1J(C_{5'}, H_{5'}) > {}^1J(C_{5'},H_{5''})$.[16] Our density functional theory (DFT) calculations at the B3LYP/6−311++G(d,p) level also confirmed this trend of J-coupling constants; see Supporting Information. All ^{13}C resonances were unambiguously assigned from ^{1}H−^{13}C HSQC and HMBC spectra. Because the values of $^4J(H_8, H_{1'})$ are too small to be useful, we used a combination of $H_{1'}$−C_8 HMBC and C_8−H_8 HSQC experiments[17] to establish the connectivity between H_8 signals (H_α and H_δ) and $H_{1'}$ resonances. Resonances for exchangeable imino protons, N_1H, were assigned using refocused ^{1}H−^{13}C HMBC spectra[18] in which $H_8 \rightarrow C_5 \rightarrow N_1H$ connectivity was established. A complete list of ^{1}H, ^{13}C, and ^{31}P chemical shifts and J-coupling constants are provided in the Supporting Information.

To further establish that G-quartet formation is responsible for the presence of both sets of ^{1}H resonances, we obtained 2D ^{1}H NOESY spectra of 1.0 M Na_2(5′-GMP) in D_2O/H_2O (1:1) at 278 K allowing detection of exchangeable protons. As seen in Figure 2b, the spectral signatures of G-quartet formation, H_8/N_2H^A and N_2H^A/N_1H NOE cross peaks, are clearly observed. Figure 2b also shows that only the imino proton from the 5′-GMP dimer, N_1H(d), is involved in hydrogen bonding. This observation suggests that the two guanine bases are held together by two $N_1H\cdots O_6$=C

10.1021/ja809258y CCC: $40.75

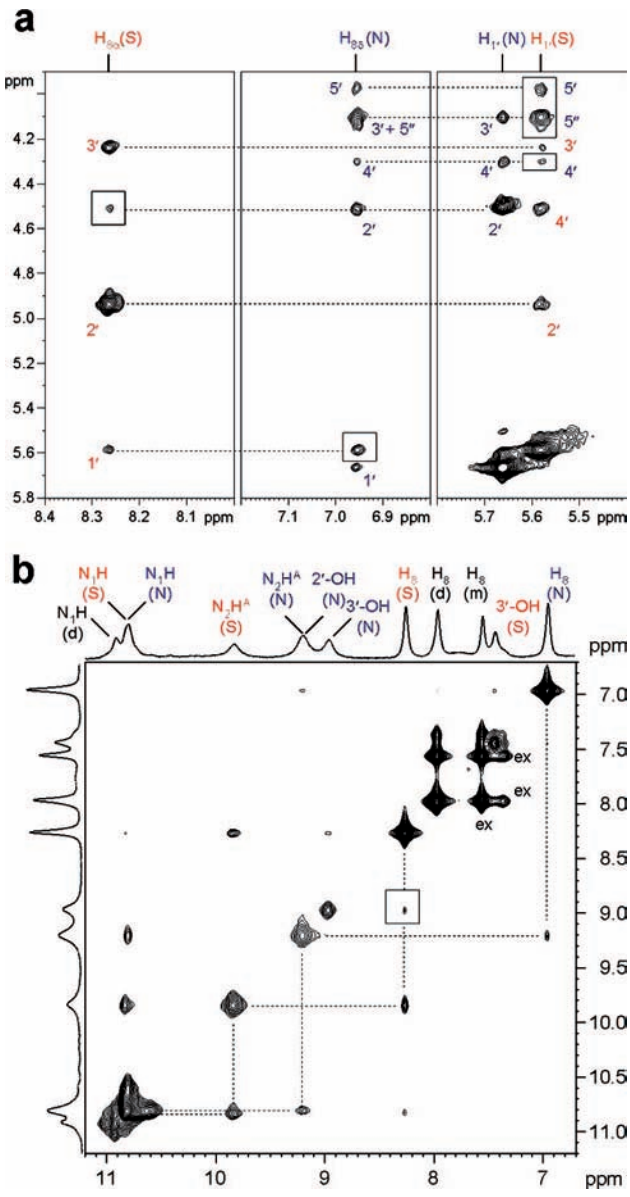

Figure 2. (a) Parts of the ^{1}H DOSY-NOESY NMR spectrum of 1.0 M Na_2(5′-GMP) in D_2O at 298 K. A mixing time of 100 ms was used. Interquartet NOE cross peaks are boxed in. (b) A region of the 2D ^{1}H NOESY spectrum of 1.0 M Na_2(5′-GMP) in D_2O/H_2O (1:1) at 278 K. A mixing time of 50 ms was used. Cross peaks due to chemical exchange are labeled as "ex".

hydrogen bonds giving rise to a centrosymmetric dimer, GG3[2] (Figure 1a). Although *ab initio* calculations have long predicted GG3[2] to be the most stable homo base pair,[19] to our knowledge, this is the first time that GG3[2] is observed in the condense phase. It is also worth noting that chemical exchange cross peaks (confirmed from ROESY spectra) are observed between H_β and H_γ, suggesting that 5′-GMP dimers and monomers undergo exchange on the 10^{-3} s time scale. But the exchange between G-quartets and dimer/monomer, if any, is too slow to be detected under the present condition.[20]

Most surprisingly, the observed values of $^3J(H_{1'}, H_{2'})$ for the two sets of ^{1}H resonances, 9 and 2 ± 1 Hz, immediately suggest that one set of ^{1}H signals are related to a sugar pucker conformation of 100% C2′-*endo* (S) and the other of 100% C3′-*endo* (N). The observed intrasugar NOE cross peaks indeed support this conclusion. Torsion angles $\chi(O_{4'}-C_{1'}-N_9-C_4)$, $\gamma(O_{5'}-C_{5'}-C_{4'}-C_{3'})$, and $\beta(P-O_{5'}-C_{5'}-C_{4'})$ were estimated from NOE cross peaks,[3] $^3J_{HH}$ and $^3J_{HP}$ in the usual fashion,[16] and then refined using the experimental interquartet NOE cross peaks mentioned earlier. The final glycosidic bond of the C3′-*endo* sugar pucker is in the normal *anti* conformation ($\chi = -130°$), but the C2′-*endo* sugar pucker displays a quite unusual high-*anti* conformation ($\chi = -60°$). The two sugar puckers also display distinct exocyclic $C_{4'}-C_{5'}$ bond orientations: (+)*gauche* (or +*sc*) for the C2′-*endo* sugar ($\gamma = 40°$) and *trans* (or *ap*) for the C3′-*endo* sugar ($\gamma = -140°$). It is also important to point out that the NOESY results (Figure 2b) also suggest that each of the two G-quartets is composed of the same sugar puckers, i.e., G(N)·G(N)·G(N)·G(N) and G(S)·G(S)·G(S)·G(S), rather than of mixed sugar puckers such as G(N)·G(S)·G(N)·G(S) or G(N)·G(N)·G(S)·G(S). Otherwise, NOE cross peaks such as $H_8(N)\rightarrow N_2H^A(S)$ or $H_8(S)\rightarrow N_2H^A(N)$ would have been observed.

To make an independent evaluation of the 5′-GMP molecular structures determined from NOE and *J*-coupling constrains, we decided to perform quantum chemical calculations of $\delta(^1H)$, $\delta(^{13}C)$, and $^1J(^1H,^{13}C)$. Because these NMR parameters are primarily determined by the molecular geometry, we chose to calculate two isolated 5′-GMP molecules. The computational results provide strong confirmation of the derived 5′-GMP molecular structures (see Supporting Information, Table S4 and Figure S8).

Now how are the two G-quartets, $G_4(N)$ and $G_4(S)$, stacked on top of each other to form a helix? The interquartet NOE cross peaks highlighted in Figure 2a provide critical clues. The $H_8(N)\rightarrow H_{1'}(S)$ cross peak suggests that the head face[21] of $G_4(S)$ points to the tail face of $G_4(N)$, i.e., head-to-tail stacking. Furthermore, $H_8(S)\rightarrow H_{2'}(N)$ and $H_{1'}(S)\rightarrow H_{5'/5''}(N)$ cross peaks are consistent with the Zimmerman model[5] in which the two G-quartets are twisted by 30° and stacked in a right-handed fashion with an axial rise of 3.4 Å. Interestingly, Gellert et al.[2] proposed an octamer model for the 3′-GMP helix. But the difference is that the G-quartets in the 3′-GMP octamer are stacked head-to-head (or tail-to-tail) with the same sugar pucker. We have also identified all four hydroxyl ^{1}H resonances from the two sugar conformers. Interestingly, judging from the observed $\delta(^1H)$ values, three of the four hydroxyl groups, $O_{2'}H(N)$, $O_{3'}H(N)$, and $O_{3'}H(S)$ are involved in strong hydrogen bonding. This observation provides additional hints about the final helical structure. Figure 3a displays a single "strand" of the 5′-GMP quadruple helix to highlight how individual 5′-GMP molecules are "stitched" together via $P-O^-\cdots H-O$ hydrogen bonds with C2′-*endo* and C3′-*endo* sugar puckers alternating along the helical strand. Moreover, an additional $[P(S)-O^-]_i\cdots[H-O_3'(N)]_{i+3}$ hydrogen bond interlocks the helical structure. Figure 3b displays the hydrogen bond linkage along the 5′-GMP helix in a conventional fashion used for polynucleotides.[19] It is striking to notice that the arrangement of adjacent 5′-GMP molecules is such that they are perfectly positioned for phosphodiester bond formation. Such a self-organized structure of 5′-GMP may provide a clue for formation of RNA oligomers under prebiotic conditions. Another notable feature of the 5′-GMP helix is that, within the $G_4(S)/G_4(N)$ octamer, $[P(S)]_i$ and $[P(N)]_{i+1}$ are separated by 6.7 Å, making it possible or even necessary for a Na^+ ion to bridge the two negatively charged groups. This mode of Na^+ binding to the phosphate groups in 5′-GMP self-assembly was first proposed by Detellier and Laszlo.[10] In comparison, $[P(N)]_{i+1}$ and $[P(S)]_{i+2}$ are separated by a longer distance, 7.2 Å. Such a $P-O^-\cdots Na^+\cdots O^-P$ interaction plays a crucial role in the 5′-GMP helix formation, because replacement of Na^+ by K^+ or Rb^+ would lead to a different yet known ordered structure.[6] Another important structural role that Na^+, K^+, and Rb^+ ions share is to occupy the central channel of the 5′-GMP helix.[22−24]

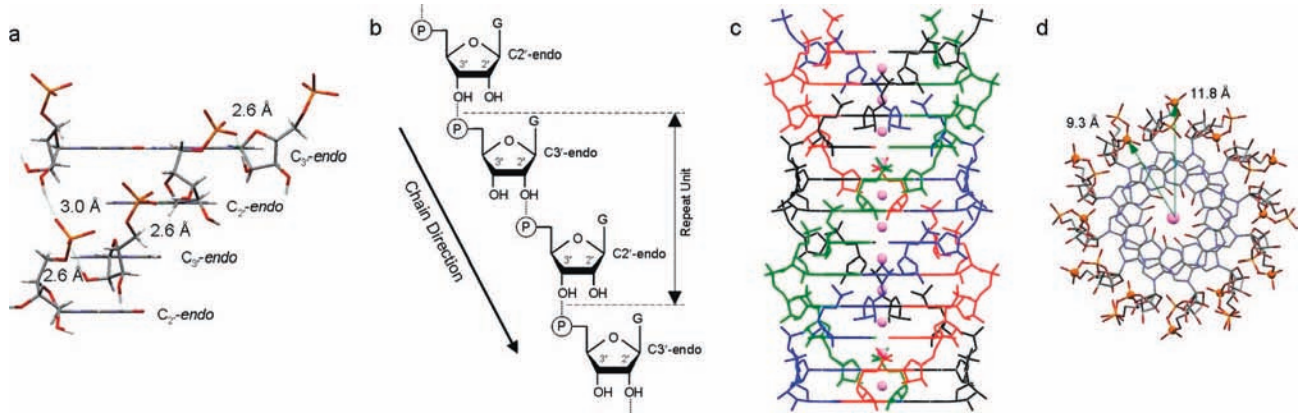

Figure 3. (a) A partial structure of the 5′-GMP helix showing key hydrogen bonds. The O···O hydrogen bond distances are given. (b) Scheme of the hydrogen bond linkage along the 5′-GMP helix following the convention for polynucleotides. (c) A full turn of the right-handed 5′-GMP quadruple helix in which the central channel is filled with Na^+ ions (purple balls). (d) The top view of the 5′-GMP helix highlighting the two different types of phosphorus atoms (gold balls). In (c) and (d), hydrogen atoms are omitted for clarity.

Figure 3c displays a full turn of this 5′-GMP quadruple helix with the central channel filled with Na^+ ions.

A comparison of common nucleic acid helices shows that the 5′-GMP quadruple helix represents a new class of nucleic acid structure; see Supporting Information, Table S4. In particular, while right-handed B- and A-form helices consist of exclusively C2′-*endo* sugar puckers, respectively, the 5′-GMP helix displays alternating C2′-*endo* and C3′-*endo* sugar puckers along the helical strand. It is interesting to note that the alternating C2′-*endo* and C3′-*endo* sugar puckers in the right-handed 5′-GMP helix are very similar to those seen in left-handed Z-DNA. Also similar to the situation in Z-DNA, the two types of phosphate groups in the 5′-GMP helix have different helical radii, 11.8 and 9.3 Å (Figure 3d). Consequently, two ^{31}P NMR signals in a 1:1 ratio are observed for the 5′-GMP helix with a chemical shift difference of ∼2.3 ppm, comparable to that found in Z-DNA.[25,26]

In summary, we have obtained new structural information about the molecular aggregates formed by spontaneous self-association of Na_2(5′-GMP) in a neutral solution. The 5′-GMP quadruple helical structure is particularly remarkable in that individual 5′-GMP molecules utilize *all* available weak molecular interactions including hydrogen bonding (base−base and hydroxyl-phosphate), ion−dipole coordination, ion−phosphate, and base stacking interactions to form an extraordinarily stable helix. This is perhaps the best example of molecular self-assembly. Because the ^1H resonances of the 5′-GMP helix show little temperature dependence, we believe that the same helical structure exists in the solid state. Because many guanosine derivatives are known to self-assemble into ordered structures,[21,27] it is possible that similar structures also exist in these compounds or, more generally, in other nucleic acid systems.

Acknowledgment. This work was supported by the Natural Sciences and Engineering Research Council (NSERC) of Canada. All quantum chemical calculations were performed at the High Performance Computing Virtual Laboratory (HPCVL) at Queen's University. I.C.M.K. thanks NSERC of Canada for an Alexander Graham Bell Canada Graduate Scholarship (CGS).

Supporting Information Available: Materials and methods, NMR spectra, and computational results (4 tables, and 8 figures). This material is available free of charge via the Internet at http://pubs.acs.org.

References

(1) Bang, I. *Biochem. Z.* **1910**, *26*, 293–311.
(2) Gellert, M.; Lipsett, M. N.; Davies, D. R. *Proc. Natl. Acad. Sci. U.S.A.* **1962**, *48*, 2013–2018.
(3) Miles, H. T.; Frazier, J. *Biochem. Biophys. Res. Commun.* **1972**, *49*, 199–204.
(4) Pinnavaia, T. J.; Miles, H. T.; Becker, E. D. *J. Am. Chem. Soc.* **1975**, *97*, 7198–7200.
(5) Zimmerman, S. B. *J. Mol. Biol.* **1976**, *106*, 663–672.
(6) 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.
(7) Fisk, C. L.; Becker, E. D.; Miles, H. T.; Pinnavaia, T. J. *J. Am. Chem. Soc.* **1982**, *104*, 3307–3314.
(8) Bouhoutsos-Brown, E.; Marshall, C. L.; Pinnavaia, T. J. *J. Am. Chem. Soc.* **1982**, *104*, 6576–6584.
(9) Borzo, M.; Detellier, C.; Laszlo, P.; Paris, A. *J. Am. Chem. Soc.* **1980**, *102*, 1124–1134.
(10) Detellier, C.; Laszlo, P. *J. Am. Chem. Soc.* **1980**, *102*, 1135–1141.
(11) Walmsley, J. A.; Barr, R. G.; Bouhoutsos-Brown, E.; Pinnavaia, T. J. *J. Phys. Chem.* **1984**, *88*, 2599–2605.
(12) Jeffrey, G. A.; Saenger, W. *Hydrogen Bonding in Biological Structures*; Springer-Verlag: Berlin, 1991.
(13) Wong, A.; Ida, R.; Spindler, L.; Wu, G. *J. Am. Chem. Soc.* **2005**, *127*, 6990–6998.
(14) Johnson, C. S. *Prog. Nucl. Magn. Reson. Spectrosc.* **1999**, *34*, 203–256.
(15) Gozansky, E. K.; Gorenstein, D. G. *J. Magn. Reson., Ser. B* **1996**, *111*, 94–96.
(16) Wijmenga, S. S.; van Buuren, B. N. M. *Prog. Nucl. Magn. Reson. Spectrosc.* **1998**, *32*, 287–387.
(17) Phan, A. T. *J. Magn. Reson.* **2001**, *153*, 223–226.
(18) Phan, A. T. *J. Biomol. NMR* **2000**, *16*, 175–178.
(19) Saenger, W. *Principles of Nucleic Acid Structure*; Springer: New York, 1984.
(20) Led, J. J.; Gesmar, H. *J. Phys. Chem.* **1985**, *89*, 583–588.
(21) Davis, J. T. *Angew. Chem., Int. Ed.* **2004**, *43*, 668–698.
(22) Wong, A.; Wu, G. *J. Am. Chem. Soc.* **2003**, *125*, 13895–13905.
(23) Wu, G.; Wong, A.; Gan, Z.; Davis, J. T. *J. Am. Chem. Soc.* **2003**, *125*, 7182–7183.
(24) Ida, R.; Wu, G. *Chem. Commun.* **2005**, 4294–4296.
(25) Patel, D. J.; Kozlowski, S. A.; Nordheim, A.; Rich, A. *Proc. Natl. Acad. Sci. U.S.A.* **1982**, *79*, 1413–1417.
(26) Sklenar, V.; Bax, A.; Zon, G. *J. Am. Chem. Soc.* **1987**, *109*, 2221–2222.
(27) Guschlbauer, W.; Chantot, J. F.; Thiele, D. *J. Biomol. Struct. Dyn.* **1990**, *8*, 491–511.

JA809258Y