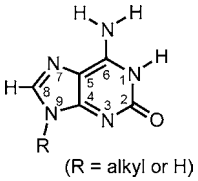
DNA Pentaplex Structure

DOI: 10.1002/ange.201203459

Solution Structure of a Parallel-Stranded Oligoisoguanine DNA Pentaplex Formed by d(T(iG)$_4$T) in the Presence of Cs^+ Ions**

Mijeong Kang, Ben Heuberger, John C. Chaput, Christopher Switzer, and Juli Feigon**

Isoguanine (iG) is a naturally occurring isomer of guanine, where the positions of the C2 amino group and C6 carbonyl group in guanine are switched to positions 6 and 2, respectively (Scheme 1). In the presence of K^+ or NH_4^+ ions, iG residues self-assemble to form tetrameric complexes that are related to the well-known G quadruplexes.[1–4] In the presence of Cs^+ ions, however, iG-containing DNA forms a pentaplex,[5] a structure that is unprecedented in nature. The self-assembly properties of iG, including pentaplex formation, have found use in building DNA nanostructures[6] and in biological applications.[7,8] Additionally, the affinity of iG for Cs^+ ions has generated interest in the possibility of selectively extracting radioactive cesium from nuclear waste.[9,10]



Scheme 1. Isoguanine(iG).

Despite increasing interest in iG multimers, little detail is known about their structures. Models for iG complexes include a non-planar bowl for a mononucleoside tetramer,[11] a four-stranded helix with C_2 symmetry for polynucleotides in aqueous NaCl solution at neutral pH,[12] and planar C_4 and C_5 symmetric structures for both monomers and oligomers.[10] While unequivocal evidence for iG DNA pentaplex formation with Cs^+ ions was provided by a combined gel electrophoretic and radiometric study,[5] and both ion-exchange chromatography[13] and mass spectroscopy[14] studies also support formation of exclusively pentaplexes by Cs^+ ions, no 3D structure is available for a multiplex of iG DNA oligomers. Indeed, the only atomic resolution structure to date is from the X-ray crystallographic analysis of a pentamer of iG mononucleosides.[10] Herein, we present NMR spectroscopy structural studies of an iG-containing DNA oligomer in the presence of different cations (Na^+, K^+, NH_4^+, Rb^+, Cs^+). We also present the first solution structure of a DNA pentaplex: Cs^+-[d(T(iG)$_4$T)]$_5$.

When CsCl was added to a solution of the d(T(iG)$_4$T) DNA oligomer, the ^{1}H NMR spectra showed four imino proton resonances that grew in intensity with increasing Cs^+ ion concentration (Figure 1 A), and a corresponding set of hydrogen-bonded NH6-1 and H8 proton resonances (Figure 3 A). These resonance signals represent hydrogen bonding by the four isoguanine bases contained within a given d(T(iG)$_4$T) strand. The appearance of only one imino resonance for each iG residue in the strand indicates that each strand is symmetric relative to the others in the multiplex formed. This also indicates that all strands are parallel, as this is the only arrangement that could give rise to a single set of resonances. In contrast, addition of K^+ ions to a solution of the d(T(iG)$_4$T) DNA oligomer resulted in the appearance of a set of broad imino proton resonances (Figure 1 B) indicating the existence of a mixture of different conformations, which exchange on the NMR time scale. This is somewhat unexpected considering that previous studies reported a quadruplex of d(T$_4$(iG)$_4$T) and of d(T(iG)$_4$T$_4$) as the major products in the presence of K^+ ions, as analyzed by gel electrophoresis[5] and mass spectrometry,[14] respectively, in addition to finding pentaplex formation with Cs^+ ions. At the same time, however, the past gel electrophoretic studies did provide clear evidence for environment-dependent formation of multiple structures in the presence of K^+ ions, since approximately 5–10% pentaplex was detected along with the balance of quadruplex when the complex was incubated at 25 °C before electrophoresis, whereas no pentaplex and only

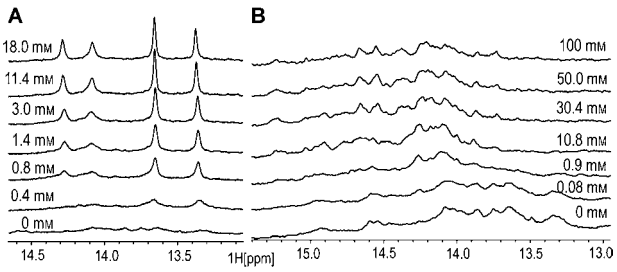


Figure 1. A) Cs^+ ion and B) K^+ ion titration into 0.2 mM d(T(iG)$_4$T), showing the imino proton region of the 1D ^{1}H NMR spectra in H_2O. The 0 mM salt spectra in (A) and (B) are the same, but are plotted at different scales so that the broader signals in (B) are more visible. The concentration of CsCl or KCl is indicated on each spectrum.

[*] Dr. M. Kang, Prof. Dr. J. Feigon
Department of Chemistry and Biochemistry
University of California, Los Angeles
Los Angeles, CA 90095-1569 (USA)
E-mail: feigon@mbi.ucla.edu
Dr. B. Heuberger,[+] Dr. J. C. Chaput,[++] Prof. Dr. C. Switzer
Department of Chemistry, University of California
Riverside, CA 92521 (USA)
E-mail: christopher.switzer@ucr.edu
[+] Current address: Department of Molecular Biology, Massachusetts General Hospital
Boston, MA 02114 (USA)
[++] Current address: Department of Chemistry, Arizona State University (USA)

[**] Work in the group of J.F. is supported by the U.S. Department of Energy. Work in the group of C.S. is supported by NASA.
Supporting information for this article (experimental details) is available on the WWW under http://dx.doi.org/10.1002/anie.201203459.

quadruplex was seen when the complex was incubated at 0 °C instead.[5]

To compare the effects of different ions on the association of the d(T(iG)$_4$T) DNA oligomer, Na^+ and Rb^+ ions were also tested, as they had been reported as pentaplex forming ions.[14] Four distinctive H8 proton signals were identified, but imino proton signals were much broader in comparison to the Cs^+ ion case (Figure 2), implying that the predominant conformation is a symmetric multiplex that is less stable than the Cs^+ ion multiplex or in exchange with a minor

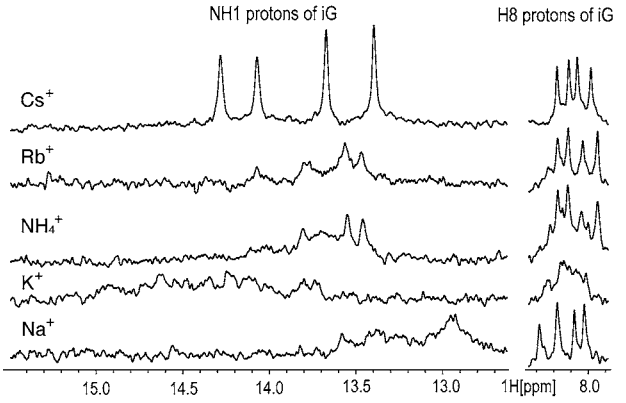


Figure 2. 1D ^{1}H NMR spectra of 0.08 mM d(T(iG)$_4$T) oligomer in the presence of 200 mM of Na^+, K^+, NH_4^+, Rb^+, or Cs^+ ions.

conformation. With NH_4^+ ions, which had been predicted to form a mixture of quadruplex and pentaplex,[14] more than four signals were detected for H8 protons, implying more than one conformation exists (Figure 2). None of the investigated metal complexes with d(T(iG)$_4$T) were found to pass through a membrane filter of 3000 molecular weight cut off (MWCO) based on UV absorption at A_{260} (data not shown). Since the size of this filter is much larger than the monomer, but at least two to three times smaller than quadruplex and pentaplex, these results indicate multiplex formation in the case of all metal complexes. In summary, the NMR spectra indicate that iG DNA oligomers can form multiple different cation-dependent complexes and that Cs^+ ions specifically and selectively stabilize iG DNA oligomer pentaplexes. In contrast to iG oligomers, where both quadruplexes and pentaplexes can form, only quadruplexes form with guanine oligomers.[15]

Only one set of resonances was detected in the 2D NOESY NMR spectra of the Cs^+ ion complex (Figure S1), which corresponds to a single hydrogen-bonded d(T(iG)$_4$T) multiplex with *n*-fold symmetry. Because of the symmetry, each signal can originate from intrastrand or interstrand residue interactions; both possibilities were carefully tested during NOE-based calculations (see Supporting Information for details) by giving unique numbers to symmetrically identical residues from different strands (Figure 3C). Cs^+ ions were not included in the structure calculation because information on the distance between Cs^+ and other atoms cannot be obtained from proton NMR data. Finally, NOE constrained structural calculations were

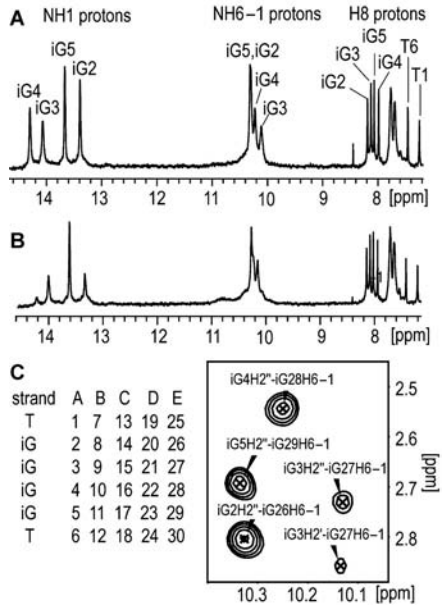


Figure 3. A) Region of the 1D ^{1}H NMR spectrum of d(T(iG)$_4$T) in the presence of 50 mM CsCl at 25 °C in H_2O showing the exchangeable and aromatic proton resonances. B) The same region of the 1D ^{1}H NMR spectrum of a sample where the solvent was exchanged back to H_2O after 12 days stored at 4 °C in D_2O. C) Numbering of nucleotides in [d(T(iG)$_4$T)]$_5$ (left) used to differentiate equivalent residues in each strand and (right) a region of the 2D NOESY 600 MHz ^{1}H NMR spectrum of [d(T(iG)$_4$T)]$_5$ showing the cross-peaks between amino and H2′, H2″ resonances in the presence of 50 mM CsCl at 25 °C in H_2O.

tested on both quadruplex and pentaplex models. However, only the pentaplex model fit the NOE data successfully.

The structure of the [d(T(iG)$_4$T)]$_5$ pentaplex in the presence of Cs^+ ions was determined as the root-mean-square deviation (RMSD) to the mean of 0.40 ± 0.02 Å for all heavy atoms (Figure 4, Table S1, and Supporting Information). In the structure, all strands are parallel. The glycosidic torsion angles fall into the *anti* range (about −83° to −130°) and all sugars adopt C2′-endo (S-type) pucker conformation. The fivefold symmetric (C_5) iG pentads are associated by counterclockwise base pairing between the Watson–Crick edge and the sugar edge of the nucleosides (Figure 4B). Each NH1 proton is hydrogen-bonded to the O2 atom of the neighboring iG residue, and each NH6-1 proton is hydrogen-bonded to the N3 atom of the neighboring iG residue within the same pentad. This arrangement places the NH6-1 protons close to the sugar protons of the iG residue to which it is hydrogen-bonded (Figure 4B), resulting in strong NOE cross-peaks between them (Figure 3C and Figure S2). All of the pentads have non-planar structures as predicted in the DFT-based quantum chemical studies of the isoguanine quintet complexes,[16–18] and each layer has a different direction and depth of the curve. The first pentad containing iG2 is nearly planar with slight convexity. The second pentad with iG3 is a spiral. The third and fourth pentads bearing iG4 and iG5, respectively, are concave curves, but the depth of the fourth pentad is much deeper than the third.

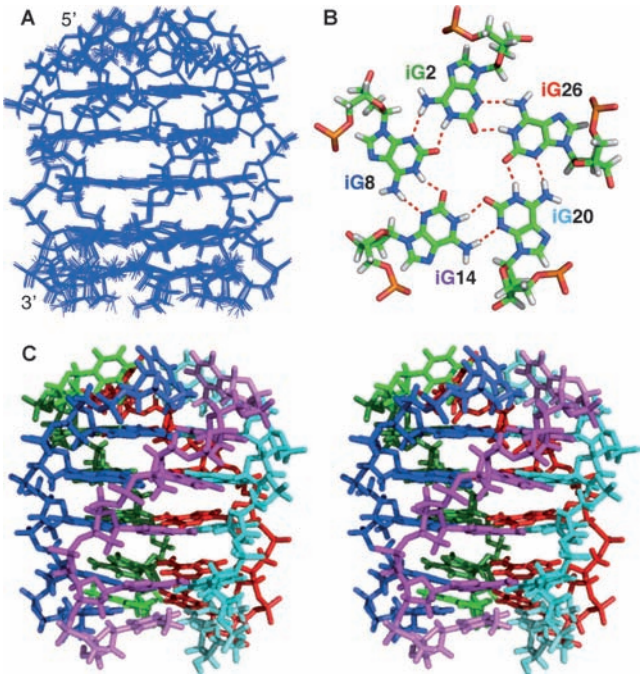


Figure 4. A) An ensemble of the 16 lowest energy structures of $[d(T(iG)_4T)]_5$ with Cs^+.[24] B) Stick representations of the top view of the iG2 pentad with red dashed lines representing hydrogen bonds between iGs in the pentad; H = gray, C = green, O = red, N = blue, P = orange. C) Stereoview of the lowest energy structure of the $[d(T(iG)_4T)]_5$ pentaplex formed with Cs^+ ions. Each strand is represented by a different color.

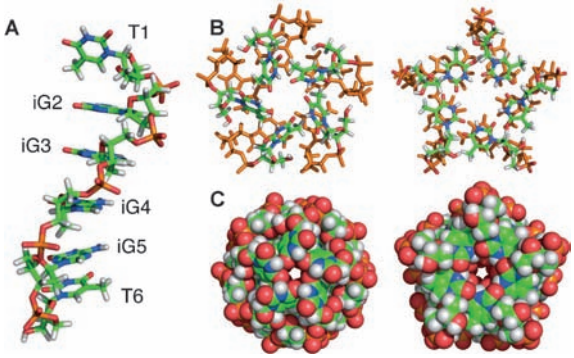


Figure 5. A) Stick representation of one strand of $d(T(iG)_4T)$. B) 5′-thymines with the iG2 pentad underneath in orange (left), and 3′-thymines with the iG5 pentad underneath in orange (right). C) CPK representations of top view (left), and bottom view (right) of the lowest energy structure of the $[d(T(iG)_4T)]_5$ pentaplex formed with Cs^+ ions; H = gray, C = green, O = red, N = blue, P = orange.

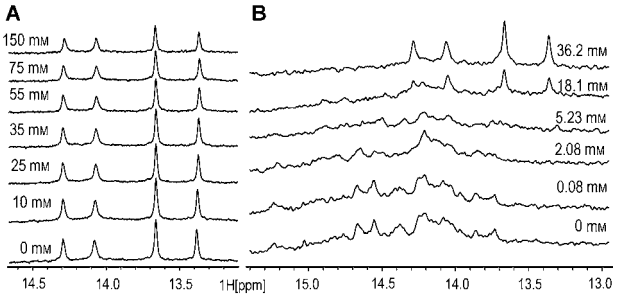


Figure 6. 1D ^1H NMR spectra of A) K^+ ion titration of a mixture containing 30 mM CsCl and 0.2 mM $d(T(iG)_4T)$ and B) Cs^+ ion titration of a mixture containing 100 mM KCl and 0.2 mM $d(T(iG)_4T)$. The concentration of the added ion is indicated on each spectrum.

The five 5'-thymines at the top of the pentaplex do not show any evidence of the formation of hydrogen bonds. The unusually strong cross-peak between T1H3′–iG2H8 and the weak cross-peak between T1H1′ and iG2H8 protons (Figure S1) explains the positions of the base plane of the thymines at the 5′ end, which are tilted up at the O2-atom side by approximately 56°, resulting in poor base stacking with the iG2 pentad (Figure 5A,B left). All O4 atoms in the pentad point to the center (Figure 5C left) leading to a negative charge density in this region. The 5′-thymines are arranged in a half-bowl shape with sufficient interior space to possibly support binding of a Cs^+ ion (Figure 4C). On the other hand, the 3′-end thymines stack directly under the iG5 pentad, with the O4 atom side tilted slightly up (Figure 5A,B right). The N3H and O2 of neighboring thymine bases are within hydrogen-bonding distance, but no imino proton signals are observed, perhaps because the N3H and C2 carbonyl are not co-planar.

To estimate the relative stability of the multiplexes, the possibility of exchanging the coordinating cation from an initially formed multiplex was tested by adding an excess of a different cation (Figure 6). Even when a fivefold excess of K^+ ions was added to a Cs^+-containing multiplex, a significant amount of the Cs^+ ion complex remained (Figure 6A). However, addition of about one third that amount of Cs^+ ions appeared to replace K^+ with Cs^+ in the multiplex (Figure 6B). This implies that coordination of the $d(T(iG)_4T)$ oligomer with the Cs^+ ion is stable enough to maintain the pentaplex even in the presence of an excess of K^+, while it is the opposite case in the multiplex formed by K^+ ions. These data are consistent with previous electrophoretic studies, which showed that running a Cs^+ ion pentaplex in a K^+-containing native gel resulted in most of the Cs^+ ion pentaplex persisting and relatively little conversion to quadruplex.[5] In addition, the thermal stability of the Cs^+-coordinated pentaplex was monitored by NMR spectroscopy. Only minor changes in the imino signals from the pentaplex were observed in spectra up to 50 °C (instrument limit), indicating that the pentaplex remains intact (Figure S3).

Among the four imino protons, the iG4 imino proton was recovered slowest when the solvent of the Cs^+ ion pentaplex was changed from D_2O to H_2O (Figure 3A). This result indicates that the imino protons in the iG4 pentad are the least solvent exchangeable. Unusually slow-exchanging imino protons have previously been observed in spectra of G-quadruplex DNA.[2] Although central pentads would be expected to exchange more slowly than terminal pentads, the iG4 pentad is adjacent to the most rapidly exchanging pentad iG5 while the iG2 and iG3 pentads exchange at about the same rate. A reasonable explanation for this is that the iG4 pentad has the highest frequency of bound Cs^+ ions.

In the present study, we determined the solution structure of the parallel-stranded iG pentaplex formed in the presence of Cs^+ ions. Within the iG-iG-iG-iG segment of this structure, the helical twist and rise for each step varied from 27.5–11.2° and 4.0–4.6 Å. These values are notably smaller than the average twist of approximately 29.2° and greater than the average rise of approximately 3.3 Å within parallel G-quadruplex structures,[19–21] which vary as a function of coordinating metal ion. Nevertheless, the iG pentaplex appears to be underwound in relation to G quadruplexes, likely a manifestation of Cs^+ ion coordination.

The positions of metal ions in the interior of G quadruplexes depend on their size, their charge, and a balance between attractive interactions with carbonyl groups (ion–dipole) and repulsive interactions with adjacent cations (Coulombic) and nucleobases (van der Waals). In the case of Na^+ ions,[19] the radius (0.95 Å) is small enough to enable its positioning within the plane of a quartet. However, as a result of the balance between attractive and repulsive forces, Na^+ ions are found both within planes and between planes, with near equidistance maintained between the ions. In the case of K^+ ions,[20] the radius (1.33 Å) is too large to be coordinated within the G-quartet plane. Consequently, the K^+ ion is always located between two stacked-quartet planes.[22,23] For the Cs^+-pentaplex solution structure refined in the absence of Cs^+ ions, the average distance between a point in the center of a pentad and the O2 atoms was determined as 2.7 Å over all four pentads (Figure S5). In contrast, the average distance between a point equidistant between two pentad layers and the O2 atoms was found to be 3.6 Å over all three possible pentad sandwiches (Figure S5). In general, Cs–ligand distances fall in the range of the latter sandwich model, and not the former in-plane model. More specifically, the average Cs–O distances reported in the X-ray structure of two stacked self-assembled pentads of a modified iG were 3.40 Å,[10] a value that is congruent with the sandwich-model measurements. Moreover, two previous computational studies on pentads place the Cs^+ ion 1.15–1.93 Å above the pentad plane,[16,17] with the Cs–O distances of 3.06–3.10 Å. The one computational study on a Cs^+-pentad sandwich (where the cesium coordination number is higher than for a simple pentad and the bond length is correspondingly greater) gave a Cs–O distance of 3.5 Å.[5] These are reasonable distances considering that the van der Waals radius of an oxygen atom is 1.52 Å and the ionic radius of a cesium ion is 1.67 Å. Taken together, these data support the occurrence of a Cs^+ ion between pentad planes in the solution structure. Such placement maintains a balance between favorable electrostatic interactions of Cs^+ ions with O2 atoms and repulsive interactions from the relatively large ionic radius of the Cs^+ ion. Indeed, the balance of forces that point to a Cs^+-pentad sandwich structure for the iG pentaplex is analogous to the K^+-quartet sandwich interactions within G quadruplexes—each comprise the most stable arrangement in their respective series.

The observation that iG multiplexes form irrespective of ion size is a particularly notable outcome of the NMR titration studies (Figure 1 and Figure 2). In contrast, only ions of particular sizes form multiplexes with guanine oligomers.[14] These data suggest that preformed isoguanine multiplexes exist even at low cation concentrations, and that adding salt stabilizes the multiplexes to differing extents over this base line, with Cs^+ ions stabilizing the pentaplex to the greatest extent. Accordingly, Cs^+-specific separation (for example, in the selective extraction of radioactive cesium from nuclear waste)[9,10] may prove feasible on the basis of the unique structure and stability properties of iG DNA pentaplexes.

Received: May 4, 2012
Published online: July 5, 2012

Keywords: DNA structures · isoguanines · NMR spectroscopy · pentaplexes · self-assembled DNA

[1] C. Kang, X. Zhang, R. Ratliff, R. Moyzis, A. Rich, *Nature* **1992**, *356*, 126 – 131.
[2] F. W. Smith, J. Feigon, *Nature* **1992**, *356*, 164 – 168.
[3] A. N. Lane, J. B. Chaires, R. D. Gray, J. O. Trent, *Nucleic Acids Res.* **2008**, *36*, 5482 – 5515.
[4] J. L. Huppert, *FEBS J.* **2010**, *277*, 3452 – 3458.
[5] J. C. Chaput, C. Switzer, *Proc. Natl. Acad. Sci. USA* **1999**, *96*, 10614 – 10619.
[6] P. K. Lo, K. L. Metera, H. F. Sleiman, *Curr. Opin. Chem. Biol.* **2010**, *14*, 597 – 607.
[7] B. A. Rosenzweig, N. T. Ross, D. M. Tagore, J. Jayawickramarajah, I. Saraogi, A. D. Hamilton, *J. Am. Chem. Soc.* **2009**, *131*, 5020 – 5021.
[8] S. R. Nallagatla, B. Heuberger, A. Haque, C. Switzer, *J. Comb. Chem.* **2009**, *11*, 364 – 369.
[9] J. T. Davis, S. K. Tirumala, A. L. Marlow, *J. Am. Chem. Soc.* **1997**, *119*, 5271 – 5272.
[10] M. Cai, A. L. Marlow, J. C. Fettinger, D. Fabris, T. J. Haverlock, B. A. Moyer, J. T. Davis, *Angew. Chem.* **2000**, *112*, 1339 – 1341; *Angew. Chem. Int. Ed.* **2000**, *39*, 1283 – 1285.
[11] S. Tirumala, J. T. Davis, *J. Am. Chem. Soc.* **1997**, *119*, 2769 – 2776.
[12] T. Golas, M. Fikus, Z. Kazimierczuk, D. Shugar, *Eur. J. Biochem.* **1976**, *65*, 183 – 192.
[13] F. Seela, R. Kroschel, *Bioconjugate Chem.* **2001**, *12*, 1043 – 1050.
[14] S. E. Pierce, J. Wang, J. Jayawickramarajah, A. D. Hamilton, J. S. Brodbelt, *Chem. Eur. J.* **2009**, *15*, 11244 – 11255.
[15] F. Aboul-ela, A. I. Murchie, D. G. Norman, D. M. Lilley, *J. Mol. Biol.* **1994**, *243*, 458 – 471.
[16] M. Meyer, J. Suhnel, *J. Phys. Chem. A* **2003**, *107*, 1025 – 1031.
[17] J. Gu, J. Wang, J. Leszczynski, *J. Comput. Chem.* **2007**, *28*, 1790 – 1795.
[18] M. Meyer, T. Steinke, J. Suhnel, *J. Mol. Model.* **2007**, *13*, 335 – 345.
[19] K. Phillips, Z. Dauter, A. I. Murchie, D. M. Lilley, B. Luisi, *J. Mol. Biol.* **1997**, *273*, 171 – 182.
[20] G. N. Parkinson, M. P. H. Lee, S. Neidle, *Nature* **2002**, *417*, 876 – 886.
[21] M. P. H. Lee, G. N. Parkinson, P. Hazel, S. Neidle, *J. Am. Chem. Soc.* **2007**, *129*, 10106 – 10107.
[22] P. Schultze, N. V. Hud, F. W. Smith, J. Feigon, *Nucleic Acids Res.* **1999**, *27*, 3018 – 3028.
[23] S. Haider, G. N. Parkinson, S. Neidle, *J. Mol. Biol.* **2002**, *320*, 189 – 200.
[24] Coordinates for the 16 lowest energy structures of $[d(T(iG)_4T)]_5$ have been deposited in the PDB under accession code 2LUJ. Chemical shifts have been deposited in the BMRB under accession code 18524.

