Stacking of G-quadruplexes: NMR structure of a G-rich oligonucleotide with potential anti-HIV and anticancer activity[†]

Ngoc Quang Do[1,2], Kah Wai Lim[1,2], Ming Hoon Teo[1], Brahim Heddi[1] and Anh Tuân Phan[1,*]

[1]School of Physical and Mathematical Sciences and [2]School of Biological Sciences, Nanyang Technological University, Singapore

ABSTRACT

G-rich oligonucleotides *T30695* (or *T30923*), with the sequence of $(GGGT)_4$, and *T40214*, with the sequence of $(GGGC)_4$, have been reported to exhibit anti-HIV and anticancer activity. Here we report on the structure of a dimeric G-quadruplex adopted by a derivative of these sequences in K^+ solution. It comprises two identical propeller-type parallel-stranded G-quadruplex subunits each containing three G-tetrad layers that are stacked via the $5'-5'$ interface. We demonstrated control over the stacking of the two monomeric subunits by sequence modifications. Our analysis of possible structures at the stacking interface provides a general principle for stacking of G-quadruplexes, which could have implications for the assembly and recognition of higher-order G-quadruplex structures.

INTRODUCTION

G-rich oligonucleotides are capable of forming four-stranded helical structures called G-quadruplexes (1–6), built from the stacking of multiple G•G•G•G tetrads (or G-tetrads) (7). There is a diversity of G-quadruplex structures with different G-tetrad cores and loop arrangements (1–6). In particular, there are four types of G-tetrad cores based on relative strand orientations [parallel-stranded; (3 + 1); antiparallel-stranded up–up–down–down; antiparallel-stranded up–down–up–down] and three main types of loops (edgewise; diagonal; double-chain-reversal) (6). This folding principle has been supported by a number of reported high-resolution structures of G-quadruplexes (1–6). The possibility of G-quadruplex stacking (8–13) and interlocking (14–16) has been proposed and could be an element of higher-order nucleic acid structures (17,18). For instance, various stacking modes (5′–5′, 3′–3′ and 3′–5′) have been proposed as a means for higher-order packaging of parallel-stranded G-quadruplexes in telomeric DNA and RNA (6). However, to date, the details of G-quadruplex stacking have not been fully understood.

Different experimental techniques may provide different views on the formation and structure of G-quadruplexes. For example, gel shift (19) and mass spectrometry (20–22) give indications on the molecular sizes of G-quadruplexes, while CD spectra are empirically used to speculate on the relative orientation of the four strands constituting the G-tetrad core. Parallel-stranded (characterized by a positive peak at 260 nm) and antiparallel-stranded (characterized by a positive peak at 290–295 nm) G-tetrad cores provide distinct CD signatures (23–25). However, CD alone might be insufficient to conclude on the folding topology of G-quadruplexes, as loop residues could also contribute to the CD signals (26,27). So far, only NMR and X-ray crystallography can provide atomic-resolution structures of G-quadruplexes. However, even these high-resolution techniques could suffer incorrect structural interpretations (28–30). Some X-ray crystallographic structures of G-quadruplexes were incorrectly interpreted at low-resolution diffractions and have been later corrected in the literature [see (28) and (29) for discussion]. For NMR structural analysis of G-quadruplexes, unambiguous and model-independent spectral assignment approaches are required (31). There were examples where model-dependent spectral assignments led to incorrect conclusions about G-quadruplex topologies [as pointed out and discussed in (30)].

A number of synthetic G-rich DNA oligonucleotides have been reported to exhibit anticoagulant (32,33), anti-HIV (34–40) and anticancer activities (14,41–44). These include *T30695* (or *T30923*), with the sequence of $(GGGT)_4$ (41,42) and *T40214*, with the sequence of $(GGGC)_4$ (43), which have been reported as inhibitors of HIV integrase (IN), a viral enzyme responsible for the

integration of viral DNA into the host-cell genome (45). In addition, *T40214* has also been reported to possess anticancer properties through the inhibition of STAT3 (36). However, there remains ambiguity in the literature regarding the structure adopted by these sequences. Jing *et al*. (41) studied *T30695* in K$^+$ solution by NMR spectroscopy and reported an antiparallel-stranded G-quadruplex, which consists of two G•G•G•G tetrads and a G•T•G•T tetrad, where all three loops are edgewise. However, spectral overlap could have hampered the structural interpretation, as the authors reported contrasting CD spectra of *T30695* with a positive peak at 260 nm (41,42), which is a signature of parallel-stranded G-quadruplexes (23–25). This structure was thought to be an exception to the rules regarding CD spectra of G-quadruplexes (46). A related G-rich oligomer HIV IN inhibitor d(GGGGTGGGAGGAGGGT), named *93del*, was later revealed to consist of an interlocked dimeric G-quadruplex with two parallel-stranded subunits (14), and prompted questions on the stoichiometry and strand orientations of *T30695*. Studies from different groups have concurred that single-residue loop is most favorable to adopt the double-chain-reversal configuration (47–51). Moreover, mass spectrometry studies have detected dimer formation for *T30695* and similar oligonucleotides (12,52), and a propeller-type parallel-stranded G-quadruplex of *T30695* was found to be stable in a recent molecular dynamics simulations study (53).

In this work, we present a NMR structural study on the G-quadruplex formed by *T30695* and its derivative sequences in K$^+$ solution. We could significantly improve NMR spectra with a single guanine-to-inosine substitution. With unambiguous resonance assignments and stoichiometry determination, we showed that this sequence adopts a dimeric G-quadruplex, formed by the stacking of two propeller-type parallel-stranded G-quadruplex subunits at their 5′-ends. We present an analysis of possible structures at the stacking interface, as well as the conditions controlling this stacking.

MATERIALS AND METHODS

Sample preparation

Unlabeled and site-specific labelled DNA oligonucleotides were chemically prepared using products from Glen Research and Cambridge Isotope Laboratories. Samples were purified following Glen Research protocol and then were dialyzed successively against KCl solution and water. DNA oligonucleodites were dissolved in solution containing 70 mM potassium chloride and 20 mM potassium phosphate (pH 7.0). DNA concentration was expressed in strand molarity using a nearest-neighbor approximation for the absorption coefficients of the unfolded species (54).

Gel electrophoresis

Molecular sizes of different G-quadruplexes were characterized in electrophoresis experiments, performed at 120 V on native gels containing 20% polyacrylamide (Acrylamide:Bis–acrylamide = 37.5:1) in TBE buffer (89 mM Tris–borate, 2 mM EDTA, pH 8.3) supplemented with 3 mM KCl. Each sample contained 5 μg DNA. Gels were viewed by UV shadowing.

Disintegration assay

The disintegration assay was performed essentially as described previously (14,55). The reaction mixture contained 20 mM HEPES (pH 7.5), 10 mM MnCl$_2$, 30 mM NaCl, 10 mM DTT, 0.05% Nonidet-P40, 600 nM HIV-1 integrase, 200 nM DB-Y1. The DNA substrate DB-Y1 (5′-TGCTAGTTCTAGCAGGCCCTTGGGCCGGCGCTT GCGCC) used in the reaction was labeled with 6-FAM^{TM} fluorescein at the 5′-end (1st BASE, Singapore). After incubating at 37°C in 2 h, the reaction mixture was mixed with equal volume of 99.5% deionized formamide (Sigma), 10 mM EDTA (pH 8.0) and heated at 90°C for 3 min. For inhibition test, the inhibitors were added into the mixture and incubated in 30 min before adding DB-Y1. The reaction products were monitored by electrophoresis on 20% polyacrylamide denaturing gels with 7 M urea.

Circular dichroism

Circular dichroism CD spectra were recorded on a Jasco-815 spectropolarimeter using 1-cm path-length quartz cuvette in a reaction volume of 600 μl at 20°C. Scans from 220 to 320 nm were performed with 200 nm/min, 1-nm pitch and 1-nm bandwidth. DNA concentration was 6 μM.

NMR spectroscopy

NMR experiments were performed on 600 and 700 MHz NMR Bruker spectrometers equipped with a cryoprobe at 25°C, unless otherwise specified. Guanine resonances were unambiguously assigned by using site-specific low-level ^{15}N labeling (56), site-specific ^{1}H-to-^{2}H substitutions (57), and through-bond correlations at natural abundance (58). Spectra assignments were completed by COSY, TOCSY, HSQC and NOESY experiments. Interproton distances were measured by NOESY experiments at various mixing times.

Structure calculation

Inter-proton distances for *J19* (Table 1) were classified based on NOESY experiments performed in H$_2$O (mixing time, 200 ms) and D$_2$O (mixing times, 100, 200 and 300 ms), and were duplicated for the two monomers. In vacuum, models were generated using the XPLOR-NIH program (59) in two general steps: (i) distance geometry simulated annealing and (ii) distance-restrained molecular dynamics refinement. Hydrogen-bond restraints, inter-proton distance restraints, dihedral restraints, planarity restraints, and non-crystallographic

Table 1. DNA sequences used for structural study

Name	Sequence (5′–3′)							
T30695	GGG	T	GGG	T	GGG	T	GGG	T
J19	GIG	T	GGG	T	GGG	T	GGG	T
T40214	GGG	C	GGG	C	GGG	C	GGG	C

symmetry restraints were imposed during structure calculations. Ten lowest-energy structures were then subjected to distance-restrained molecular dynamics refinement in explicit solvent using the AMBER program (60), in which the dihedral, planarity and noncrystallographic symmetry restraints were removed. Detailed procedures are described in the Supplementary Data. Structures were displayed using the PyMOL program (61).

Data deposition

The coordinates of ten lowest-energy d(GIGTGGGTGG GTGGGT) (*J19*) dimeric G-quadruplex structures upon distance-restrained molecular dynamics refinement in explicit solvent have been deposited in the Protein Data Bank (accession code 2LE6).

RESULTS AND DISCUSSION

G-quadruplexes of *T30695* and *T40214* in K$^+$ solution

NMR spectra including 1D spectra (Figure 1) and 2D NOESY (Supplementary Figure S1 and Supplementary Data) indicated that the *T30695* and *T40214* sequences (Table 1) form similar G-quadruplex structures in K$^+$ solution. In our hands, *T30695* showed similar 1D imino proton spectra to the previously reported ones by Jing *et al.* (41), but very different 2D NOESY spectra. In particular, we did not observe strong intraresidue H8-H1′ NOE cross-peaks for guanines, which would be indicative of *syn* glycosidic conformations. *T30695* and *T40214* exhibit similar CD spectra with a positive band at 260 nm (Figure 2), a characteristic signature of parallel-stranded G-quadruplexes (23).

A single guanine-to-inosine substitution improves NMR spectra of *T30695* in K$^+$ solution

Proton resonances in the NMR spectra of *T30695* were heavily overlapped (Figure 1b; Supplementary Figure S1b and Supplementary Data), presumably due to the quasi-symmetry of the structure arising from the repetitive nature of the sequence, which could have hampered detailed structural analysis by NMR (41). We found that a single guanine-to-inosine substitution at position 2 of *T30695* greatly improved the spectral resolution (Figure 1c; Supplementary Figure S1c and Supplementary Data). This modification, resulting in the sequence GIGT(GGGT)$_3$ (henceforth designated *J19*), most likely broke up some symmetry of the structure (see NMR data and structure below) and, consequently, remove the degeneracy of the NMR spectra.

NMR spectra of *J19* in K$^+$ solution (Figure 1c) showed eleven sharp guanine imino protons at 10.8–11.6 ppm and an inosine imino proton at 13.8 ppm, corresponding to three G-tetrads. Similar spectral characteristics, including chemical shifts and NOESY patterns (Figure 1b; Supplementary Figure S1b and Supplementary Data), indicated the same conformation for *J19* and *T30695* (and *T40214*). These were corroborated by their similar CD spectra, which showed a positive band at 260 nm (Figure 2), consistent with the formation of parallel-stranded G-quadruplexes (23). The correspondence between *J19*, *T30695* and *T40214* was also observed in gel electrophoresis experiments (see below).

J19 forms a stacked dimeric G-quadruplex in K$^+$ solution

Guanine imino protons of *J19* were unambiguously assigned using site-specific 2% ^{15}N-labeled samples (56) (Figure 3a). The downfield-shifted peak at 13.8 ppm was assigned to the imino proton of I2 (62). Guanine H8 protons were assigned by through-bond correlations to the respective imino protons via ^{13}C5 at natural abundance (58) (Figure 3b). Some of these assignments were independently confirmed by site-specific ^{2}H substitutions at the H8 position (57) (Supplementary Figure S2 and Supplementary Data). Assignments of other protons were assisted by through-bond (COSY, TOCSY and HSQC), and through-space (NOESY) correlations (31). The H8/H6-H1′ sequential NOE connectivities could be observed from G1 to G3, G5 to G7, G9 to G11 and G13 to T16, while interruptions were observed at T4, T8 and T12 (Figure 4a), which adopt the double-chain-reversal configurations (see structure below). The moderate intensity of intraresidue H8/H6-H1′ NOEs indicated that all residues adopt *anti* glycosidic conformations.
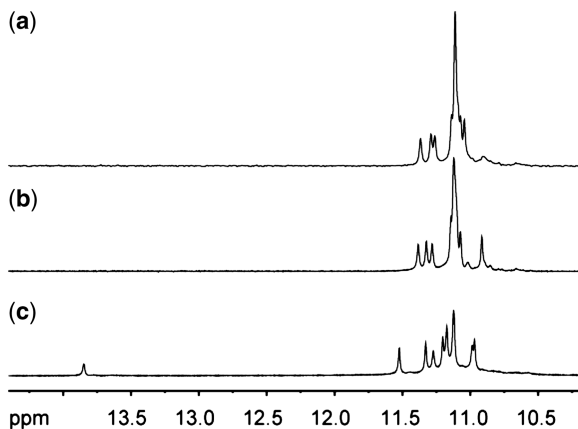

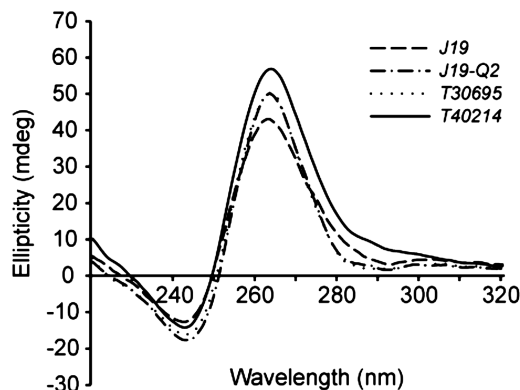**Figure 1.** Imino proton NMR spectra of (**a**) *T40214*, (**b**) *T30695* and (**c**) *J19* in K$^+$ solution at 25°C.

Figure 2. CD spectra of *T30695*, *J19*, *J19-Q2* and *T40214* in K$^+$ solution.
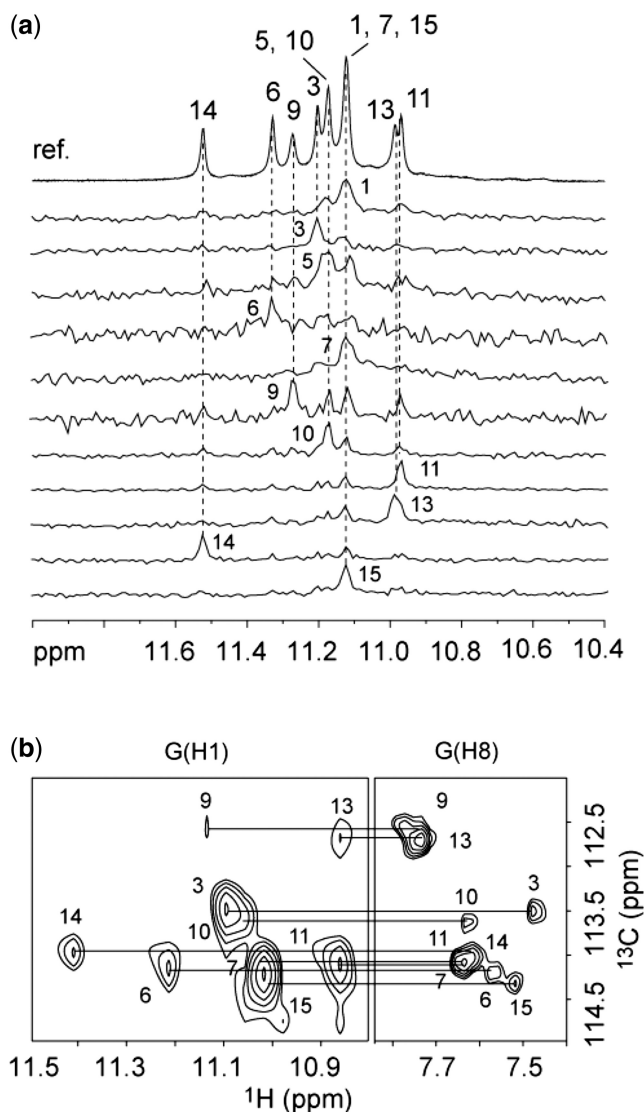
(a)

(b)

Figure 3. Proton resonance assignments of *J19* in K$^+$ solution. (**a**) Imino protons were assigned by 2% ^{15}N-labeling at the indicated positions. (**b**) H8 proton were assigned by through-bond correlations between imino and H8 protons via ^{13}C5 at natural abundance, using long-range *J* couplings.

Characteristic NOE patterns between imino and H8 protons pointed to the formation of three tetrads: G1•G5•G9•G13, I2•G6•G10•G14 and G3•G7•G11•G15 (Figure 4b and c). These tetrad alignments, combined with other data (see below), indicated that *J19* forms a stacked dimeric G-quadruplex, in which each half of the structure is a propeller-type parallel-stranded G-quadruplex with three tetrad layers and three double-chain-reversal loops (Figure 4d). Data supporting the 5′-end stacking and dimeric nature of *J19* include: (i) gel electrophoresis showing the migration rate of *J19* to be similar to that of the interlocked dimeric G-quadruplex *93del* and slower than that of *J19-Q2*, a monomeric propeller-type G-quadruplex containing three tetrad layers (see below); (ii) solvent-exchange data showing that the

imino protons of guanines in the G1•G5•G9•G13 tetrad at the stacking interface are protected from the exchange with D$_2$O (Supplementary Figure S3); (iii) the observations of several NOE cross-peaks across the interface between two monomeric subunits (see below); and (iv) the disruption of stacking interaction when additional bases were extended from the 5′-end of *J19* (see below).

The solution structure of *J19* was calculated on the basis of NMR restraints (Table 2), including NOEs at the interface between the two subunits (discussed below), which were imposed as ambiguous distance restraints. Ten lowest-energy structures after distance-restrained molecular dynamics refinement in explicit solvent were superimposed and presented in Figure 5. The core of the G-quadruplex is well converged and there is tight packing across the interface. The hydrogen-bond directionalities of the tetrads are the same within each subunit, but opposite between the two subunits. The thymine bases T4, T8 and T12 project outwards of the core, typical of single-nucleotide double-chain-reversal loops (14).

Structure of the stacking interface

Stacking between G-quadruplexes has been proposed (8–13), but so far there have been limited discussions regarding the structural details at the stacking interface. For a pair of G-quadruplex monomers that stack through planar tetrads, we can envisage a range of isomers, wherein the two subunits are rotated with respect to each other about the common helical axis. We might suppose that certain orientations of the two subunits are most favorable, corresponding to states of lowest energies, which will be determined by many factors, including van der Waals contacts, orbital overlap, ion coordination and steric hindrance. For the current case involving two parallel-stranded G-quadruplexes, one major mode of stacking was observed: the two monomers stack in a 5′–5′ manner, with G1, G5, G9, and G13 from one subunit directly above G9*, G5*, G1* and G13*, respectively, from the other subunit (Figure 5; Supplementary Figure S4 and Supplementary Data). The sugars from the two tetrads are immediately adjacent to one another (with the faces containing the H1′ and H4′ atoms facing each other), the backbones of the two subunits align in a staggered manner, and the five- and six-membered rings of the guanines display partial overlap across the interface (Supplementary Figure S5iii, iii′ and Supplementary Data). This dimeric arrangement was supported by the observation of NOEs among base and sugar protons of the G1•G5•G9•G13 tetrad, including those between imino-H8 of G5–G5* and G13–G13*, as well as reciprocal NOEs between H8–H1′, H1′–H1′ and H1′–H3′ of G1 and G9* (or G9 and G1*). Some of these NOEs were unambiguously confirmed in samples where G(H8) were site-specifically substituted by deuterium (Figure 6, see figure legend; Supplementary Figure S6 and Supplementary Data). In principle, there are three other possible arrangements which would produce almost identical stacking pattern, each being related to this major mode by successive 90° rotation(s) of one subunit. Due to spectral overlap, evidence for their existence or absence is limited, and it
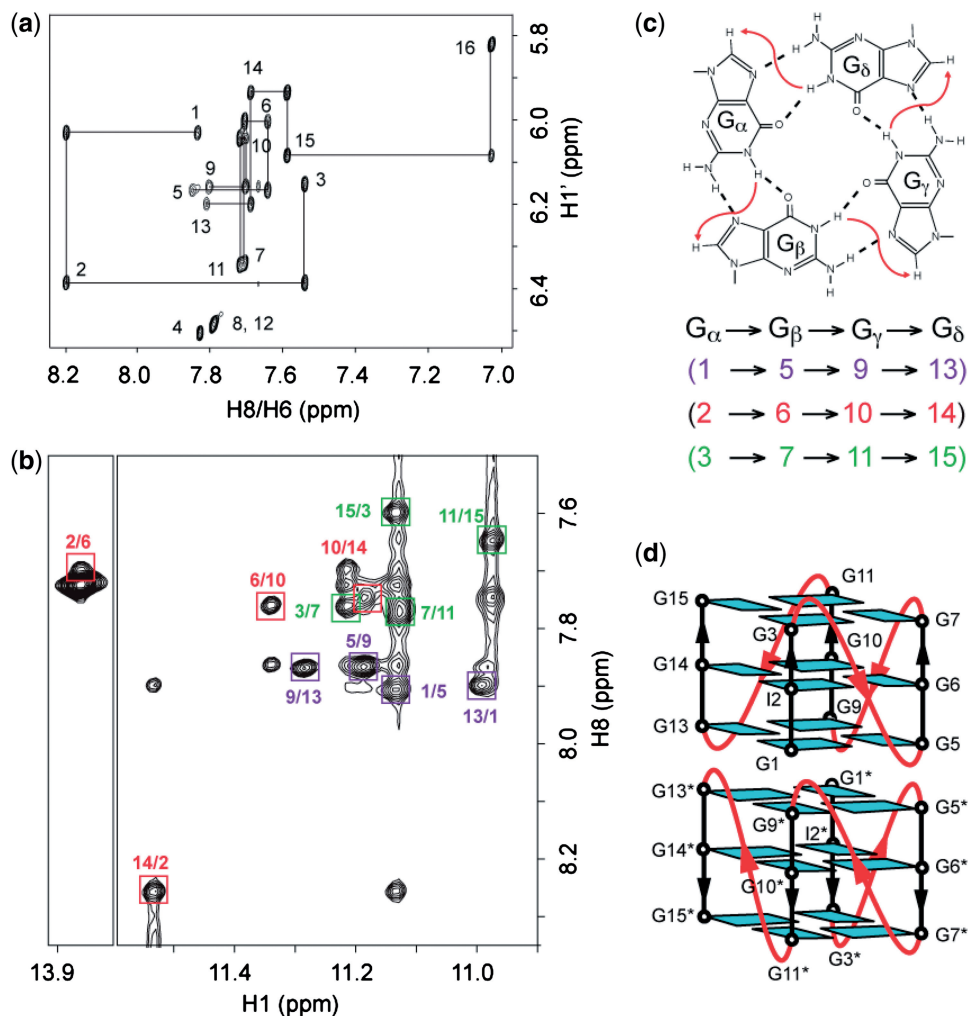
Figure 4. Determination of G-quadruplex folding topology of *J19* in K⁺ solution. (**a**) H8/H6-H1′ region of NOESY spectrum (mixing time, 300 ms) at 25°C. The assignments and NOE sequential connectivities are shown. (**b**) Imino-H8 region of NOESY spectrum (mixing time, 400 ms) at 25°C. The characteristic guanine imino-H8 cross-peaks for G-tetrads are framed and labeled with the imino proton assignment in the first position and that of the H8 proton in the second position. (**c**) Specific imino-H8 connectivity pattern around a G-tetrad (G$_\alpha$•G$_\beta$•G$_\gamma$•G$_\delta$) indicated with arrows (connectivity between G$_\delta$ and G$_\alpha$ implied). The connectivities observed for G1•G5•G9•G13 (purple), I2•G6•G10•G14 (red) and G3•G7•G11•G15 (green) tetrads are shown below. (**d**) The schematic structure of *J19* dimeric parallel-stranded G-quadruplex.

will be prudent not to preclude them entirely. In a recent RNA G-quadruplex structure consisting of a dimer of dimers, a major mode of stacking was also observed (11). We note that in these two examples, the terminal 5′-OH groups of the component strands are positioned furthest away from each other in the major conformation (Supplementary Figure S7 and Supplementary Data), which might explain why this isomer is more stable than the rest.

Patterns of overlap at the interface for different stacking orientations

Various patterns of overlap of the guanine bases at the interface have been observed across a number of G-quadruplex structures involving a pair of tetrads with opposing hydrogen-bond directionalities: (i) maximum overlap of the five-membered rings of guanines (9,14,63,64); (ii) maximum overlap of the six-membered rings of guanines

(65,66); and (iii) partial overlap of the five- and six-membered rings of guanines (67–69). Comparing against the current case, i.e. pattern (iii) (Supplementary Figure S5 and Supplementary Data), (i) is related by a ~25° clockwise rotation of the bottom subunit with respect to the top, while (ii) is related by a ~15° anticlockwise rotation of the bottom subunit with respect to the top. Pattern (i) was observed at the interface of the interlocked dimeric G-quadruplex *93del* (14), and is typically observed within an antiparallel G-quadruplex. Pattern (ii) was seen in crystal structures of tetrameric quadruplex assemblies where the interface involves the intersection of eight strands, with octad formation at the sandwiching layers immediately adjacent to the interface (65,66). Pattern (iii) was observed across a number of dimeric structures that involve stacking of intramolecular G-quadruplexes (67–69), including the present case of *J19*. It is interesting to note that regardless if *J19* is twisted into orientation

Table 2. Statistics of the computed structures of the dimeric parallel-stranded G-quadruplex *J19*[a]

NMR restraints[b]		
Distance restraints	D$_2$O	H$_2$O
Intra-residue distance restraints	574	20
Sequential (*i*, *i*+1) distance restraints	322	36
Long-range (*i*, ≥ *i*+2) distance restraints	22	92
Inter-subunit distance restraints[c]	16	4
Other restraints		
Hydrogen-bond restraints	92	
Dihedral restraints	64	
Structure statistics for 10 molecules following distance-restrained molecular dynamics refinement in explicit solvent		
NOE violations		
Number (>0.2 Å)	0.100 ± 0.300	
Maximum violation (Å)	0.156 ± 0.031	
RMSD of violations (Å)	0.014 ± 0.001	
Deviations from the ideal covalent geometry		
Bond lengths (Å)	0.002 ± 0.000	
Bond angles (°)	0.603 ± 0.010	
Impropers (°)	0.795 ± 0.078	
Pairwise all heavy atom RMSD values (Å)		
All heavy atoms except T4, T8, T12 and T16	0.87 ± 0.12	
All heavy atoms	1.50 ± 0.25	

[a]PDB ID: 2LE6.
[b]Restraints were duplicated for the two monomers, and include those within each subunit, as well as those at the interface between the two subunits.
[c]These distance restraints were imposed as ambiguous restraints and specified with the sum-averaging option.

(i) or (ii), it reverted to (iii) upon unrestrained molecular dynamics simulations in explicit solvent (data not shown).

Control of stacking between G-quadruplexes

Gel electrophoresis, which can provide information on the molecular sizes of G-quadruplexes, was used to examine the possibility of stacking of parallel-stranded G-quadruplex blocks to form dimeric structures. Sequences beginning with G at the 5′-end, such as *T30695* and *J19*, were found to favor stacking of G-quadruplexes at this end. The migration of these stacked dimeric G-quadruplexes is comparable to that of the reference interlocked dimeric G-quadruplex *93del* (14) (Figure 7), which also consists of six G-tetrad layers. It is conceivable that the addition of terminal non-G residue(s) would reduce the stacking propensity of these G-quadruplexes (8,10,11). Indeed, the extension of two thymines at the 5′-end of *J19* (giving *J19-Q2*; Table 3) led to the sequence migrating faster than *93del* (Figure 7). We have shown that this sequence forms in a similar condition a propeller-type G-quadruplex with three G-tetrad layers (unpublished results). The extension of three thymines instead of two only marginally increased the mobility of the structure (Supplementary Figure S8 and Supplementary Data). On the other hand, sequences with only one additional T at the 5′-end migrated faster than the original sequences (*T30695* and *J19*), but slower
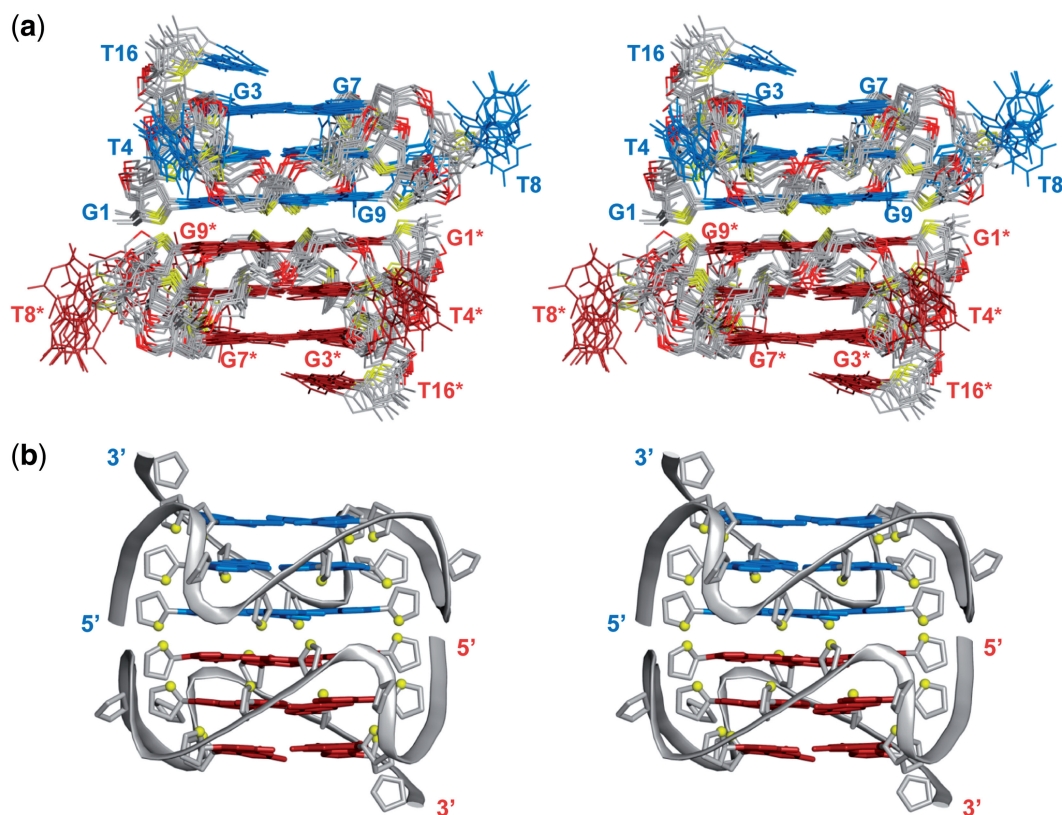

Figure 5. Stereoviews of the dimeric parallel-stranded G-quadruplex structure of *J19* in K$^+$ solution. (**a**) Ten superimposed structures after distance-restrained molecular dynamics refinement in explicit solvent. (**b**) Ribbon view of a representative structure. Bases from the top monomer are colored blue while those from the bottom monomer are colored red. Backbone and sugar atoms are colored gray, with O4′ atoms in yellow and phosphorus atoms in red.
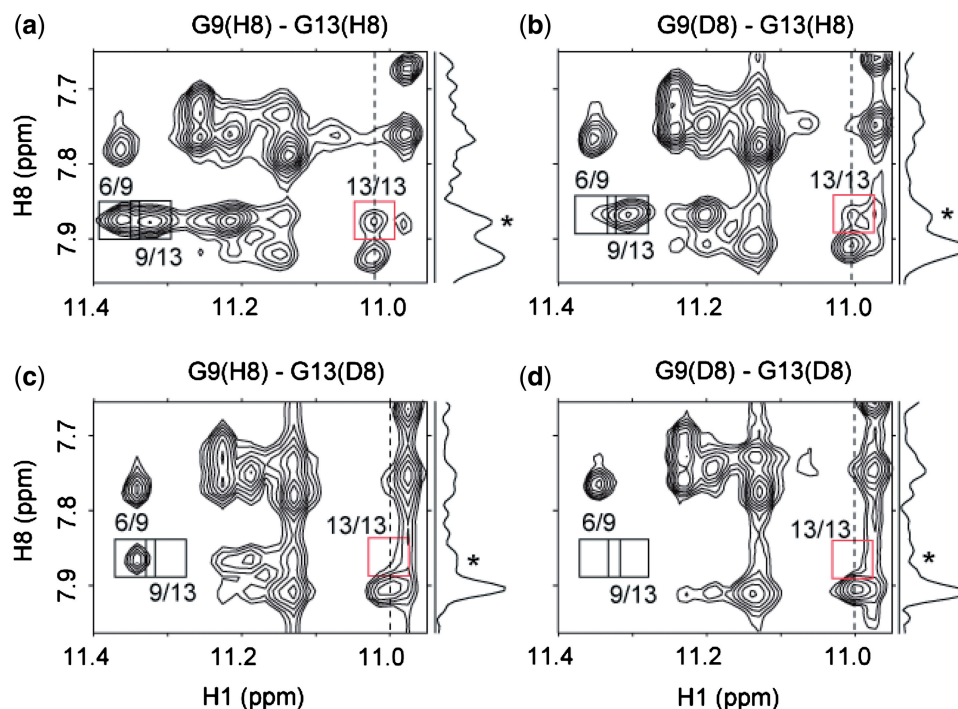
Figure 6. Determination of the stacking pattern of two *J19* subunits. Imino-H8 proton region of NOESY spectra (mixing time, 400 ms) of (**a**) *J19* and (**b–d**) modified sequences with site-specific ^{2}H substitution at position H8 of (b) G9, (c) G13 and (d) both G9/G13 are shown. The peaks involving substituted guanines (framed) are missing for modified sequences. The inter-subunit NOE, whose intensity is significantly decreased for the G13-substituted sequence (marked with asterisk on the projection), is framed in red.

than those containing two thymines (Figure 7). There could be equilibrium between the monomeric and stacked dimeric forms, resulting in the intermediate migration rate. In contrast to the 5′-end, there could be minimal stacking over the 3′-end surfaces: each additional T at the 3′-end only slightly increased the mobility of the parent sequences (Figure 7; Supplementary Figure S8 and Supplementary Data). Similar NMR (Supplementary Figures S1 and S9; Supplementary Data) and gel electrophoresis (Supplementary Figure S10) results were obtained for *T40214*, consistent with the formation of a similar stacked dimeric G-quadruplex for this sequence.

HIV-1 IN inhibition activity of various G-rich sequences

In order to probe if the various G-rich oligonucleotides retain anti-HIV IN activity, we performed assays on the reverse 'disintegration' reaction by HIV-1 IN (14,55). The results are presented in Figure 8. *T30695*, *J19* and their derivative sequences with thymines at the 3′-end exhibited comparable inhibition activity as *93del*, whereas derivative sequences containing thymines at the 5′-end were somewhat less active, which might be attributed to their lower potentials to form stacked dimeric structures. Combined with other studies (14,70), this work suggested dimeric parallel-stranded G-quadruplexes comprising a total of six G-tetrad layers as potential HIV-1 IN inhibitors. Note however that the conclusion from the assay would be more qualitative than quantitative, as the concentrations here are in the micromolar range, which are around two orders of magnitude larger than the reported
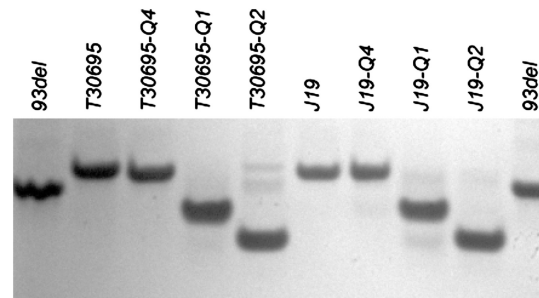

Figure 7. Breaking of the stacked dimer. Gel electrophoresis experiments showing the effect of adding T residues at the 3′ and 5′-ends on the formation of higher-order structures for *T30695* and *J19*. The interlocked dimeric G-quadruplex *93del* was used as a reference. Sequence name is labeled on the top of each lane.

IC$_{50}$ for the inhibition activity of *T30695* against HIV-1 IN (71).

CONCLUSION

A DNA sequence derived from the oligonucleotides *T30695* and *T40214* was shown to adopt a dimeric G-quadruplex, formed by the stacking of two identical propeller-type parallel-stranded G-quadruplex subunits at their 5′-ends. The structure of the stacking interface was analyzed in detail. We showed that stacking interaction between the two subunits was disfavored by extension of non-G residues from the terminals.

Table 3. DNA sequences used for control of stacking

| Name	Sequence (5′–3′)
T30695	GGG T GGG T GGG T GGG T
T30695-Q1	T GGG T GGG T GGG T GGG T
T30695-Q2	TT GGG T GGG T GGG T GGG T
T30695-Q3	TTT GGG T GGG T GGG T GGG T
T30695-Q4	GGG T GGG T GGG T GGG TT
T30695-Q5	TT GGG T GGG T GGG T GGG
T30695-Q6	TT GGG T GGG T GGG T GGG TT
T30695-Q7	TT GGG T GGG T GGG T GGG TTT
J19	GIG T GGG T GGG T GGG T
J19-Q1	T GIG T GGG T GGG T GGG T
J19-Q2	TT GIG T GGG T GGG T GGG T
J19-Q3	TTT GIG T GGG T GGG T GGG T
J19-Q4	GIG T GGG T GGG T GGG TT
J19-Q5	TT GIG T GGG T GGG T GGG
J19-Q6	TT GIG T GGG T GGG T GGG TT
J19-Q7	TT GIG T GGG T GGG T GGG TTT
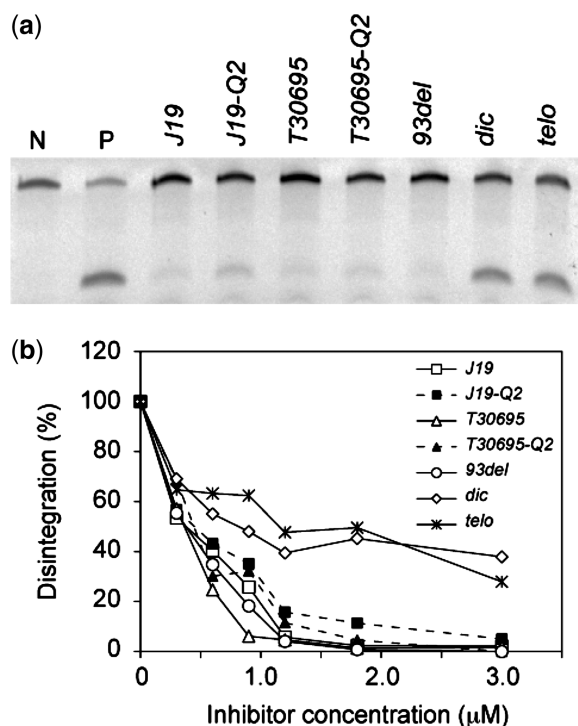
Figure 8. HIV-1 IN disintegration inhibition assays. (**a**) Lanes N and P are controls without and with HIV-1 integrase alone, respectively. Other lanes are inhibition assays with added inhibitors as indicated, where *dic* is the Dickerson d(CGCGAATTCGCG) dodecamer and *telo* is the human telomeric sequence d[TT(GGGTTA)$_3$GGGA]. The concentrations of the enzyme, substrate and inhibitors were 0.6, 0.2 and 3.0 μM, respectively. (**b**) The assay was performed with different concentration of inhibitors.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

FUNDING

Conflict of interest statement. None declared.

ACCOMPANYING MANUSCRIPT

This article should be read in conjunction with article: *Nucl. Acids Res.* (2011), **39**(20): 8984–8991, doi:10.1093/nar/gkr540.

REFERENCES

1. Davis,J.T. (2004) G-quartets 40 years later: from 5′-GMP to molecular biology and supramolecular chemistry. *Angew. Chem. Int. Ed. Engl.*, **43**, 668–698.
2. Burge,S., Parkinson,G.N., Hazel,P., Todd,A.K. and Neidle,S. (2006) Quadruplex DNA: sequence, topology and structure. *Nucleic Acids Res.*, **34**, 5402–5415.
3. Phan,A.T., Kuryavyi,V. and Patel,D.J. (2006) DNA architecture: from G to Z. *Curr. Opin. Struct. Biol.*, **16**, 288–298.
4. Patel,D.J., Phan,A.T. and Kuryavyi,V. (2007) Human telomere, oncogenic promoter and 5'-UTR G-quadruplexes: diverse higher order DNA and RNA targets for cancer therapeutics. *Nucleic Acids Res.*, **35**, 7429–7455.
5. Neidle,S. (2009) The structures of quadruplex nucleic acids and their drug complexes. *Curr. Opin. Struct. Biol.*, **19**, 239–250.
6. Phan,A.T. (2010) Human telomeric G-quadruplex: structures of DNA and RNA sequences. *FEBS J.*, **277**, 1107–1117.
7. Gelleert,M., Lipsett,M.N. and Davies,D.R. (1962) Helix formation by guanylic acid. *Proc. Natl Acad. Sci. USA*, **48**, 2013–2018.
8. Wang,Y. and Patel,D.J. (1992) Guanine residues in d(T2AG3) and d(T2G4) form parallel-stranded potassium cation stabilized G-quadruplexes with anti glycosidic torsion angles in solution. *Biochemistry*, **31**, 8112–8119.
9. Kettani,A., Gorin,A., Majumdar,A., Hermann,T., Skripkin,E., Zhao,H., Jones,R. and Patel,D.J. (2000) A dimeric DNA interface stabilized by stacked A.(G.G.G.G).A hexads and coordinated monovalent cations. *J. Mol. Biol.*, **297**, 627–644.
10. Kato,Y., Ohyama,T., Mita,H. and Yamamoto,Y. (2005) Dynamics and thermodynamics of dimerization of parallel G-quadruplexed DNA formed from d(TTAGn) (n = 3-5). *J. Am. Chem. Soc.*, **127**, 9980–9981.
11. Martadinata,H. and Phan,A.T. (2009) Structure of propeller-type parallel-stranded RNA G-quadruplexes, formed by human telomeric RNA sequences in K$^+$ solution. *J. Am. Chem. Soc.*, **131**, 2570–2578.
12. Smargiasso,N., Rosu,F., Hsia,W., Colson,P., Baker,E.S., Bowers,M.T., De Pauw,E. and Gabelica,V. (2008) G-quadruplex DNA assemblies: Loop length, cation identity, and multimer formation. *J. Am. Chem. Soc.*, **130**, 10208–10216.
13. Šket,P. and Plavec,J. (2010) Tetramolecular DNA quadruplexes in solution: insights into structural diversity and cation movement. *J. Am. Chem. Soc.*, **132**, 12724–12732.
14. Phan,A.T., Kuryavyi,V., Ma,J.-B., Faure,A., Andréola,M.-L. and Patel,D.J. (2005) An interlocked dimeric parallel-stranded DNA quadruplex: a potent inhibitor of HIV-1 integrase. *Proc. Natl Acad. Sci. USA*, **102**, 634–639.
15. Krishnan-Ghosh,Y., Liu,D.S. and Balasubramanian,S. (2004) Formation of an interlocked quadruplex dimer by d(GGGT). *J. Am. Chem. Soc.*, **126**, 11009–11016.
16. Zhang,N., Gorin,A., Majumdar,A., Kettani,A., Chernichenko,N., Skripkin,E. and Patel,D.J. (2001) V-shaped scaffold: a new architectural motif identified in an A x (G x G x G x G) pentad-containing dimeric DNA quadruplex involving stacked G(anti) x G(anti) x G(anti) x G(syn) tetrads. *J. Mol. Biol.*, **311**, 1063–1079.

17. Marsh,T.C. and Henderson,E. (1994) G-wires-self-assembly of a telomeric oligonucleotide, d(GGGGTTGGGG), into large superstructures. *Biochemistry*, **33**, 10718–10724.

18. Pisano,S., Savino,M., Micheli,E., Varra,M., Coppola,T., Mayol,L. and De Santis,P. (2008) Superstructural self-assembly of the G-quadruplex structure formed by the homopurine strand in a DNA tract of human telomerase gene promoter. *Biophys. Chem.*, **136**, 159–163.

19. Sen,D. and Gilbert,W. (1988) Formation of parallel four-stranded complexes by guanine-rich motifs in DNA and its implications for meiosis. *Nature*, **334**, 364–366.

20. Rosu,F., Gabelica,V., Houssier,C., Colson,P. and De Pauw,E. (2002) Triplex and quadruplex DNA structures studied by electrospray mass spectrometry. *Rapid Commun. Mass Spectrom.*, **16**, 1729–1736.

21. Xu,Y., Kaminaga,K. and Komiyama,M. (2008) G-quadruplex formation by human telomeric repeats-containing RNA in Na$^+$ solution. *J. Am. Chem. Soc.*, **130**, 11179–11184.

22. Collie,G.W., Parkinson,G.N., Neidle,S., Rosu,F., De Pauw,E. and Gabelica,V. (2010) Electrospray mass spectrometry of telomeric RNA (TERRA) reveals the formation of stable multimeric G-quadruplex structures. *J. Am. Chem. Soc.*, **132**, 9328–9334.

23. Balagurumoorthy,P., Brahmachari,S.K., Mohanty,D., Bansal,M. and Sasisekharan,V. (1992) Hairpin and parallel quartet structures for telomeric sequences. *Nucleic Acids Res.*, **20**, 4061–4067.

24. Gray,D.M., Wen,J.D., Gray,C.W., Repges,R., Repges,C., Raabe,G. and Fleischhauer,J. (2008) Measured and calculated CD spectra of G-quartets stacked with the same or opposite polarities. *Chirality*, **20**, 431–440.

25. Renciuk,D., Kejnovska,I., Skolakova,P., Bednarova,K., Motlova,J. and Vorlickova,M. (2009) Arrangements of human telomere DNA quadruplex in physiologically relevant K+ solutions. *Nucleic Acids Res.*, **37**, 6625–6634.

26. Hu,L., Lim,K.W., Bouaziz,S. and Phan,A.T. (2009) Giardia telomeric sequence d(TAGGG)$_4$ forms two intramolecular G-quadruplexes in K$^+$ solution: effect of loop length and sequence on the folding topology. *J. Am. Chem. Soc.*, **131**, 16824–16831.

27. Lim,K.W., Lacroix,L., Yue,D.J.E., Lim,J.K.C., Lim,J.M.W. and Phan,A.T. (2010) Coexistence of two distinct G-quadruplex conformations in the hTERT promoter. *J. Am. Chem. Soc.*, **132**, 12331–12342.

28. Kelly,J.A., Feigon,J. and Yeates,T.O. (1996) Reconciliation of the X-ray and NMR structures of the thrombin-binding aptamer d(GGTTGGTGTGGTTGG). *J. Mol. Biol.*, **256**, 417–422.

29. Haider,S., Parkinson,G.N. and Neidle,S. (2002) Crystal structure of the potassium form of an Oxytricha nova G-quadruplex. *J. Mol. Biol.*, **320**, 189–200.

30. Wang,Y. and Patel,D.J. (1995) Solution structure of the Oxytricha telomeric repeat d[G4(T4G4)3] G-tetraplex. *J. Mol. Biol.*, **251**, 76–94.

31. Phan,A.T., Guéron,M. and Leroy,J.L. (2001) Investigation of unusual DNA motifs. *Methods Enzymol.*, **338**, 341–371.

32. Bock,L.C., Griffin,L.C., Latham,J.A., Vermaas,E.H. and Toole,J.J. (1992) Selection of single-stranded-DNA molecules that bind and inhibit human thrombin. *Nature*, **355**, 564–566.

33. Schultze,P., Macaya,R.F. and Feigon,J. (1994) Three-dimensional solution structure of the thrombin-binding DNA aptamer d(GGTTGGTGTGGTTGG). *J. Mol. Biol.*, **235**, 1532–1547.

34. Bates,P.J., Kahlon,J.B., Thomas,S.D., Trent,J.O. and Miller,D.M. (1999) Antiproliferative activity of G-rich oligonucleotides correlates with protein binding. *J. Biol. Chem.*, **274**, 26369–26377.

35. Simonsson,T. and Henriksson,M. (2002) c-myc suppression in Burkitt's lymphoma cells. *Biochem. Biophys. Res. Commun.*, **290**, 11–15.

36. Jing,N., Li,Y., Xiong,W., Sha,W., Jing,L. and Tweardy,D.J. (2004) G-quartet oligonucleotides: a new class of signal transducer and activator of transcription 3 inhibitors that suppresses growth of prostate and breast tumors through induction of apoptosis. *Cancer Res.*, **64**, 6603–6609.

37. Qi,H., Lin,C.-P., Fu,X., Wood,L.M., Liu,A.A., Tsai,Y.-C., Chen,Y., Barbieri,C.M., Pilch,D.S. and Liu,L.F. (2006) G-quadruplexes induce apoptosis in tumor cells. *Cancer Res.*, **66**, 11808–11816.

38. Choi,E.W., Nayak,L.V. and Bates,P.J. (2010) Cancer-selective antiproliferative activity is a general property of some G-rich oligodeoxynucleotides. *Nucleic Acids Res.*, **38**, 1623–1635.

39. Wyatt,J.R., Vickers,T.A., Roberson,J.L., Buckheit,R.W. Jr, Klimkait,T., DeBaets,E., Davis,P.W., Rayner,B., Imbach,J.L. and Ecker,D.J. (1994) Combinatorially selected guanosine-quartet structure is a potent inhibitor of human immunodeficiency virus envelope-mediated cell fusion. *Proc. Natl Acad. Sci. USA*, **91**, 1356–1360.

40. Rando,R.F., Ojwang,J., Elbaggari,A., Reyes,G.R., Tinder,R., Mcgrath,M.S. and Hogan,M.E. (1995) Suppression of human-immunodeficiency-virus type-1 activity in-vitro by oligonucleotides which form intramolecular tetrads. *J. Biol. Chem.*, **270**, 1754–1760.

41. Jing,N., Gao,X., Rando,R.F. and Hogan,M.E. (1997) Potassium-induced loop conformational transition of a potent anti-HIV oligonucleotide. *J. Biomol. Struct. Dyn.*, **15**, 573–585.

42. Jing,N.J., Marchand,C., Liu,J., Mitra,R., Hogan,M.E. and Pommier,Y. (2000) Mechanism of inhibition of HIV-1 integrase by G-tetrad-forming oligonucleotides in vitro. *J. Biol. Chem.*, **275**, 21460–21467.

43. Jing,N., Xiong,W., Guan,Y., Pallansch,L. and Wang,S. (2002) Potassium-dependent folding: a key to intracellular delivery of G-quartet oligonucleotides as HIV inhibitors. *Biochemistry*, **41**, 5397–5403.

44. de Soultrait,V.R., Lozach,P.Y., Altmeyer,R., Tarrago-Litvak,L., Litvak,S. and Andréola,M.L. (2002) DNA aptamers derived from HIV-1 RNase H inhibitors are strong anti-integrase agents. *J. Mol. Biol.*, **324**, 195–203.

45. Craigie,R. (2001) HIV integrase, a brief overview from chemistry to therapeutics. *J. Biol. Chem.*, **276**, 23213–23216.

46. Dapić,V., Abdomerović,V., Marrington,R., Peberdy,J., Rodger,A., Trent,J.O. and Bates,P.J. (2003) Biophysical and biological properties of quadruplex oligodeoxyribonucleotides. *Nucleic Acids Res.*, **31**, 2097–2107.

47. Phan,A.T., Modi,Y.S. and Patel,D.J. (2004) Propeller-type parallel-stranded G-quadruplexes in the human c-myc promoter. *J. Am. Chem. Soc.*, **126**, 8710–8716.

48. Hazel,P., Huppert,J., Balasubramanian,S. and Neidle,S. (2004) Loop-length-dependent folding of G-quadruplexes. *J. Am. Chem. Soc.*, **126**, 16405–16415.

49. Rachwal,P.A., Findlow,I.S., Werner,J.M., Brown,T. and Fox,K.R. (2007) Intramolecular DNA quadruplexes with different arrangements of short and long loops. *Nucleic Acids Res.*, **35**, 4214–4222.

50. Bugaut,A. and Balasubramanian,S. (2008) A sequence-independent study of the influence of short loop lengths on the stability and topology of intramolecular DNA G-quadruplexes. *Biochemistry*, **47**, 689–697.

51. Guédin,A., De Cian,A., Gros,J., Lacroix,L. and Mergny,J.-L. (2008) Sequence effects in single-base loops for quadruplexes. *Biochimie*, **90**, 686–696.

52. Li,H., Yuan,G. and Du,D. (2008) Investigation of formation, recognition, stabilization, and conversion of dimeric G-quadruplexes of HIV-1 integrase inhibitors by electrospray ionization mass spectrometry. *J. Am. Soc. Mass Spectrom.*, **19**, 550–559.

53. Li,M.H., Zhou,Y.H., Luo,Q. and Li,Z.S. (2010) The 3D structures of G-Quadruplexes of HIV-1 integrase inhibitors: molecular dynamics simulations in aqueous solution and in the gas phase. *J. Mol. Model.*, **16**, 645–657.

54. Cantor,C.R., Warshaw,M.M. and Shapiro,H. (1970) Oligonucleotide interactions. 3. Circular dichroism studies of the conformation of deoxyoligonucleotides. *Biopolymers*, **9**, 1059–1077.

55. Chow,S.A., Vincent,K.A., Ellison,V. and Brown,P.O. (1992) Reversal of integration and DNA splicing mediated by integrase of human immunodeficiency virus. *Science*, **255**, 723–726.

56. Phan,A.T. and Patel,D.J. (2002) A site-specific low-enrichment (15)N,(13)C isotope-labeling approach to unambiguous NMR

spectral assignments in nucleic acids. *J. Am. Chem. Soc.*, **124**, 1160–1161.

57. Huang,X.N., Yu,P.L., LeProust,E. and Gao,X.L. (1997) An efficient and economic site-specific deuteration strategy for NMR studies of homologous oligonucleotide repeat sequences. *Nucleic Acids Res.*, **25**, 4758–4763.

58. Phan,A.T. (2000) Long-range imino proton-^{13}C J-couplings and the through-bond correlation of imino and non-exchangeable protons in unlabeled DNA. *J. Biomol. NMR*, **16**, 175–178.

59. Schwieters,C.D., Kuszewski,J.J., Tjandra,N. and Clore,G.M. (2003) The Xplor-NIH NMR molecular structure determination package. *J. Magn. Reson.*, **160**, 65–73.

60. Case,D.A., Cheatham,T.E., Darden,T., Gohlke,H., Luo,R., Merz,K.M., Onufriev,A., Simmerling,C., Wang,B. and Woods,R.J. (2005) The Amber biomolecular simulation programs. *J. Comput. Chem.*, **26**, 1668–1688.

61. DeLano,W.L. (2002) *The PyMOL User's Manual*. DeLano Scientific, Palo Alto, CA.

62. Jiang,F., Patel,D.J., Zhang,X., Zhao,H. and Jones,R.A. (1997) Specific labeling approaches to guanine and adenine imino and amino proton assignments in the AMP-RNA aptamer complex. *J. Biomol. NMR*, **9**, 55–62.

63. Caceres,C., Wright,G., Gouyette,C., Parkinson,G. and Subirana,J.A. (2004) A thymine tetrad in d(TGGGGT) quadruplexes stabilized with Tl$^+$/Na$^+$ ions. *Nucleic Acids Res.*, **32**, 1097–1102.

64. Creze,C., Rinaldi,B., Haser,R., Bouvet,P. and Gouet,P. (2007) Structure of a d(TGGGGT) quadruplex crystallized in the presence of Li+ ions. *Acta Crystallogr. D Biol. Crystallogr.*, **63**, 682–688.

65. Deng,J.P., Xiong,Y. and Sundaralingam,M. (2001) X-ray analysis of an RNA tetraplex (UGGGU)$_4$ with divalent Sr2$^+$ ions at subatomic resolution (0.61 angstrom). *Proc. Natl Acad. Sci. USA*, **98**, 13665–13670.

66. Pan,B.C., Xiong,Y., Shi,K., Deng,J.P. and Sundaralingam,M. (2003) Crystal structure of an RNA purine-rich tetraplex containing adenine tetrads: Implications for specific binding in RNA tetraplexes. *Structure*, **11**, 815–823.

67. Parkinson,G.N., Lee,M.P.H. and Neidle,S. (2002) Crystal structure of parallel quadruplexes from human telomeric DNA. *Nature*, **417**, 876–880.

68. Matsugami,A., Ouhashi,K., Kanagawa,M., Liu,H., Kanagawa,S., Uesugi,S. and Katahira,M. (2001) An intramolecular quadruplex of (GGA)$_4$ triplet repeat DNA with a G:G:G:G tetrad and a G(:A):G(:A):G(:A):G heptad, and its dimeric interaction. *J. Mol. Biol.*, **313**, 255–269.

69. Matsugami,A., Okuizumi,T., Uesugi,S. and Katahira,M. (2003) Intramolecular higher order packing of parallel quadruplexes comprising a G:G:G:G tetrad and a G(:A):G(:A):G(:A):G heptad of GGA triplet repeat DNA. *J. Biol. Chem.*, **278**, 28147–28153.

70. Mukundan,V.T., Do,N.Q. and Phan,A.T. (2011) HIV-1 integrase inhibitor T30177 forms a stacked dimeric G-quadruplex structure containing bulges. *Nucleic Acids Res.*, **39**, 8984–8991.

71. Jing,N. and Hogan,M.E. (1998) Structure-activity of tetrad-forming oligonucleotides as a potent anti-HIV therapeutic drug. *J. Biol. Chem.*, **273**, 34992–34999.