Structure and Conformational Dynamics of a Stacked Dimeric G-Quadruplex Formed by the Human CEB1 Minisatellite

Michael Adrian,[†] Ding Jie Ang,[†] Christopher J. Lech,[†] Brahim Heddi,[†] Alain Nicolas,[‡] and Anh Tuân Phan*,[†]

[†]School of Physical and Mathematical Sciences, Nanyang Technological University, Singapore 637371, Singapore
[‡]Institut Curie, Centre de Recherche, UMR3244 CNRS, Université Pierre et Marie Curie, Paris 75248, France

Ⓢ Supporting Information

ABSTRACT: CEB1 is a highly polymorphic human minisatellite. In yeast, the size variation of CEB1 tandem arrays has been associated with the capacity of the motif to form G-quadruplexes. Here we report on the NMR solution structure of a G-quadruplex formed by the CEB1 DNA G-rich fragment d(AGGGGGGAGGGAGGGTGG), harboring several G-tracts including one with six continuous guanines. This sequence forms a dimeric G-quadruplex involving the stacking of two subunits, each being a unique snapback parallel-stranded scaffold with three G-tetrad layers, three double-chain-reversal loops, and a V-shaped loop. The two subunits are stacked at their 5′-end tetrads, and multiple stacking rotamers may be present due to a high symmetry at the stacking interface. There is a conformational exchange in the millisecond time scale involving a swapping motion between two bases of the six-guanine tract. Our results not only add to the understanding of how the G-quadruplex formation in human minisatellite leads to genetic instability but also address the fundamental questions regarding stacking of G-quadruplexes and how a long continuous G-tract participates in the structure and conformational dynamics of G-quadruplexes.
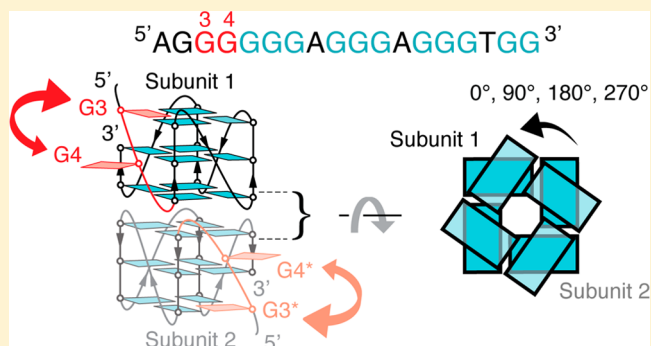
INTRODUCTION

G-quadruplex is a noncanonical four-stranded structure of nucleic acids formed by G-rich DNA or RNA oligonucleotides in the presence of stabilizing cations such as K^+ or Na^+. A G-quadruplex comprises multiple stacked tetrads, each being a planar association of four guanines interconnected through Hoogsteen hydrogen bonds.[1−4] Nucleic acid sequences prone to G-quadruplex formation were found to be evolutionary conserved, suggesting putative roles of such structures in cellular processes such as replication and transcription.[5] A general rule of predicting intramolecular G-quadruplex-forming sequences $d(G_{3+}N_{1−7}G_{3+}N_{1−7}G_{3+}N_{1−7}G_{3+})$, which requires four tracts of at least three continuous guanines separated by linkers of 1−7 residues, has been used to identify G-quadruplex-prone sequences in human genome.[6,7] For instance, there are several genomic regions characterized by high G-quadruplex-forming potential including telomeres,[8] proto-oncogenes,[9,10] and mutational hot spot sequences.[11−13] The formation of G-quadruplexes in these regions has been associated with the inhibition of cancer and other human diseases.[9,14−16] Consequently, G-quadruplexes have become promising therapeutic targets for drug design.[17,18]

Different strand orientations, glycosidic conformations of guanine bases, groove widths, intervening loops, and molecularities constitute the polymorphism of G-quadruplex folding topologies.[19] Generally, an intramolecular G-quadruplex can be adopted by a DNA sequence with four G-tracts, which would form four columns supporting the G-tetrad core. However, some oncogenic promoter G-quadruplexes display irregular folding topologies: the c-myc promoter sequence of five G-tracts folds into a snapback configuration involving a guanine from the last G-tract of two residues,[20] and the c-kit promoter sequence forms a unique parallel snapback scaffold with an isolated guanine participating in tetrad formation.[21] Sequences with interrupted G-tracts can form G-quadruplexes containing bulges.[22] A number of sequences including human chl1 intron,[23] a locked nucleic acid modified,[24] and other engineered oligomers[25,26] exhibit exceptional V-shaped loops within the scaffolds spanning two to three tetrad planes. Additionally, high DNA and cation concentrations can cause structural transition from unimolecular to bimolecular G-quadruplexes.[27−30] High-resolution NMR structures of stacked dimeric G-quadruplexes have been previously reported,[28,31] showing different tetrad overlap patterns at the dimerization interface compared to that observed in interlocked G-quadruplexes.[32−35]

Many human sequences have the potential to form G-quadruplexes,[36] including numerous subtelomeric minisatellites consisting of tandem repetitions of 10−100 nt motifs which are

assembled into large arrays of 0.5 to >15 kb. CEB1 (D2S90) is one of the most polymorphic human minisatellites with an average size of 2−3 kb encompassing repeat units of 37−43 nucleotides.[37] Additionally, CEB1 is also a hot spot of germline recombination which leads to a high rate of size variation.[38,39] Mechanistically, CEB1 contains a binding motif for the Prdm9 protein that plays a role in the selection of the meiotic recombination initiation sites.[40] However, the CEB1 motif carries another layer of information. In addition to duplex, it can adopt noncanonical G-quadruplex structures.[13] The genetic instability of the CEB1 minisatellite has been studied in the model organism *Saccharomyces cerevisiae*. It is triggered in wild-type cells by the introduction of G-quadruplex stabilizing compound Phen-DC$_3$ or the inactivation of G-quadruplex resolving helicase Pif1, while a mutated CEB1 devoid of G-quadruplex-forming motifs remains stable.[13,41] Additional genetic and molecular analyses demonstrated that CEB1 instability is dependent on the formation of G-quadruplexes, replication, and homologous recombination.[42]

Herein, we investigate the structure of a G-quadruplex in the CEB1 minisatellite. In K$^+$ solution, the CEB1 G-rich fragment d(AGGGGGGAGGGAGGGTGG) forms a stacked dimeric G-quadruplex, where each subunit adopts a parallel snapback scaffold. The subunit is characterized with a unique conformational exchange at the 5′-end whereby a guanine in a V-shaped loop could swap position with a neighboring tetrad-bound guanine. The two subunits stack at their 5′-end tetrads, and multiple stacking rotamers may be present due to a high symmetry of the stacking interface. The structure of a CEB1 G-quadruplex reported in this work should add to the understanding of how the formation of G-quadruplexes in human minisatellite leads to genetic instability.

■ METHODS

DNA Sample Preparation. Unlabeled and site-specific labeled DNA oligonucleotides (Tables 1, S1, and S2) were chemically

Table 1. List of DNA Sequences Analyzed in This Work

Name	Sequence (5'-3')							
CEB1	A	GGGGGG	A	GGG	A	GGG	T	GG
A1ΔG2Δ		GGGGG	A	GGG	A	GGG	T	GG
G2T	A	TGGGGG	A	GGG	A	GGG	T	GG
G3T	A	GTGGGG	A	GGG	A	GGG	T	GG
G4T	A	GGTGGG	A	GGG	A	GGG	T	GG
G(2,3)T	A	TTGGGG	A	GGG	A	GGG	T	GG
G(2,4)T	A	TGTGGG	A	GGG	A	GGG	T	GG
G(3,4)T	A	GTTGGG	A	GGG	A	GGG	T	GG
G(17,18)T	A	GGGGGG	A	GGG	A	GGG	T	TT

synthesized on an ABI 394 DNA/RNA synthesizer. Oligonucleotides were purified and dialyzed successively against 25 mM KCl solution and water. Unless otherwise stated, samples were dissolved in solution containing 70 mM KCl 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.[43]

Gel Electrophoresis. The molecular size of G-quadruplexes was visualized by nondenaturing polyacrylamide gel electrophoresis (PAGE).[44] Samples were incubated in 20 mM potassium phosphate buffer (pH 7) before loaded on 20% polyacrylamide gels supplemented with variable concentrations of KCl and run at 26 °C. 40% (v/v) sucrose was added before loading. The gels were revealed by UV-shadowing.

Circular Dichroism. Circular dichroism (CD) spectra were recorded on a JASCO-815 spectropolarimeter using 1 cm path length quartz cuvettes with reaction volume of 600 μL. The DNA oligonucleotides (∼5 μM) were prepared in solution containing 20 mM potassium phosphate (pH 7) and 70 mM KCl. For each experiment, an average of three scans was taken, the spectrum of the buffer was subtracted, and the data were zero-corrected at 320 nm.

Thermal Difference Spectra. The thermal difference spectra (TDS) were obtained by taking the difference between the absorbance spectra of the unfolded and folded oligonucleotides that were respectively recorded much above and below the melting temperature. TDS provide specific signatures of different DNA structural conformations.[45] Spectra were recorded between 220 and 320 nm on a JASCO V-650 UV−Vis spectrophotometer using 1 cm path length quartz cuvettes. The DNA oligonucleotides (∼5 μM) were prepared in solution containing 20 mM potassium phosphate (pH 7) and 70 mM KCl. For each experiment, an average of three scans was taken, and the data were zero-corrected at 320 nm.

UV-Melting Experiments. The thermal stability of DNA G-quadruplexes at different salt and DNA strand concentrations was characterized in heating/cooling experiments by recording the UV absorbance at 295 nm as a function of temperature using a JASCO V-650 UV−Vis spectrophotometer.[46] UV-melting experiments at variable concentrations of potassium chloride and DNA were conducted as previously described.[47] The heating and cooling rates were 0.2 °C per minute. Experiments were performed with 1, 0.2, and 0.05 cm path length quartz cuvettes.

NMR Spectroscopy. NMR experiments were performed on 600 and 700 MHz Bruker spectrometers at 25 °C, unless otherwise specified. The DNA strand concentration of the NMR samples was typically 0.2−1.5 mM. Resonances for guanine residues were assigned unambiguously by using site-specific low-enrichment ^{15}N labeling,[48] site-specific ^{2}H labeling,[49] and through-bond correlations at natural abundance.[50] Spectral assignments were completed by NOESY, TOCSY, {^{13}C−^{1}H}-HMBC, and {^{13}C−^{1}H}-HSQC as previously described.[51] Interproton distances were deduced from NOESY experiments at various mixing times. All spectral analyses were performed using the FELIX (Felix NMR, Inc.) program.

Conformational exchange rate between the major and minor forms was calculated from the ratio between the exchange cross-peak and the diagonal peak intensities for a particular proton in NOESY spectra at different mixing times.[52] The relative population of the major and minor forms was estimated from their diagonal peak intensities.

Structure Calculation. Structure computations were performed using the XPLOR-NIH program in two general steps:[53] (i) distance geometry simulated annealing and (ii) restrained molecular dynamics refinement. The 10 lowest-energy structures were then solvated and further refined with 1 ns restrained molecular dynamics simulations. The restrained simulations included hydrogen-bond restraints, interproton distance restraints, dihedral restraints, and planarity restraints for structure calculations. Interproton distances were deduced from NOESY experiments performed in H$_2$O (mixing times, 300 ms) and ^{2}H$_2$O (mixing times, 100, 150, 200, 250, and 300 ms). The glycosidic dihedral restraints were based on intraresidues NOE of H1′-H6/8 cross-peak intensities. Two intermolecular proton distances were each derived from NOESY spectra of the native and modified sequences (mixing times, 300 ms) and were used to guide the initial orientation of stacking of the two subunits during simulated annealing and refinement in vacuum by XPLOR-NIH program. Only one intermolecular proton distance (observed in the NOESY spectrum of the native sequence) was used during further MD refinement in explicit solvent. Structures were displayed using the PyMOL program.

Molecular Dynamics Simulation. A total of 5 K$^+$ ions were manually added between tetrads within a G-quadruplex and between stacked subunits. Additional 29 K$^+$ ions were then added to neutralize the system. Systems were then solvated in a TIP3P truncated octahedral box containing 10000 water molecules.[54] Systems were simulated with the periodic boundary conditions under constant pressure (1 bar) and temperature (300 K) using the Berendsen algorithm.[55] A 2 fs time step was employed with covalent bonds

involving hydrogen atoms constrained using the SHAKE algorithm.[56] Long-range electrostatic interactions were computed using the Ewald method with a nonbonded cutoff of 9 Å.[57] The list of nonbonded pairs was updated every 10 ps with the center-of-mass motion also being removed at these intervals.

Systems initially underwent a series of restrained minimization and dynamics. First, systems were minimized with a harmonic potential position restraint of 25 kcal mol^{-1} Å^{-2} placed on DNA atoms and interior K$^+$ ions. Systems were then heated from 100 to 300 K and equilibrated under constant volume while maintaining position restraints. Subsequent steps of minimization and equilibrium were performed with position restraints gradually reduced to 5, 4, 3, 2, 1, and 0.5 kcal mol^{-1} Å^{-2}.

All explicit solvent simulations were performed using the AMBER 10 software with the Parmbsc0 modification to the Cornell et al. force field.[58−60] Molecular dynamics refinement of the 10 lowest-energy structures produced in XPLOR calculations was performed using 1 ns simulations inclusive of NOE-derived distance restraints. Unrestrained simulations of 200 ns length were performed for different rotamers of the *CEB1* dimer. The starting structure for one rotamer was derived from the NMR models presented herein. Starting structures for three other rotamers were generated by rotating one of the *CEB1* subunits from the NMR-based structure by 90°, 180°, and 270°. Molecular dynamics trajectories were processed and analyzed using the PTRAJ program package in AMBER. Analysis of G-tetrad stacking geometry was performed using a program developed by our laboratory.

Data Deposition. The coordinates of the *CEB1* G-quadruplex have been deposited in the Protein Data Bank under accession code 2MB4.

■ RESULTS

G-Quadruplex Formation by a Short G-Rich Sequence of the Human CEB1 Minisatellite. The 18-nt fragment d(AGGGGGGGAGGGGAGGGTGG) (designated *CEB1*), taken from the natural 39-nt CEB1 minisatellite repeating unit, comprises G-tracts of various lengths (Table 1 and Figure 1A). Twelve major imino protons at 10.5−11.5 ppm were observed in the NMR spectrum of *CEB1* in K$^+$ solution (Figure 1B), indicating the formation of three G-tetrad layers in a G-
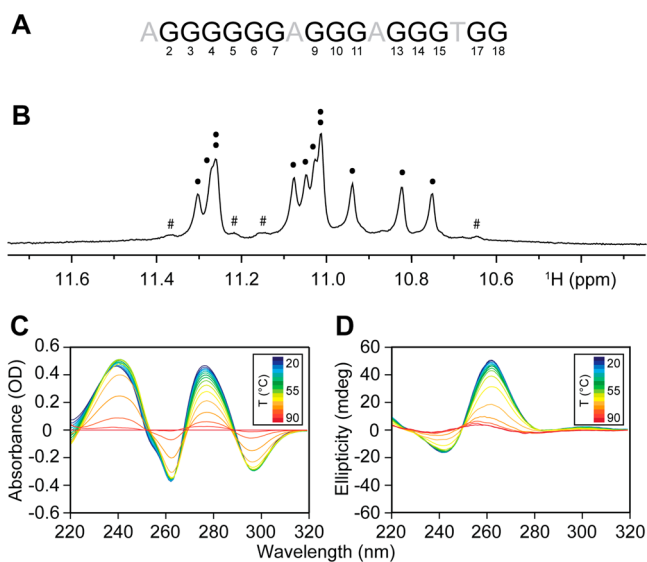

Figure 1. G-quadruplex formation by a G-rich sequence of the CEB1 minisatellite in K$^+$ solution, as monitored by spectroscopic techniques: (A) The DNA sequence of *CEB1*, d(AGGGGGGGAGGGGAGGGTGG). (B) NMR imino proton spectrum. (C) TDS and (D) CD spectra of *CEB1*. Imino peaks from major and minor form are marked with a dot and hash sign, respectively.

quadruplex. Along with this predominant G-quadruplex, a minor form (<10%) associated with low-intensity imino proton peaks was also observed (see further discussion below). The G-quadruplex formation was supported by the observation of a typical UV absorption TDS pattern with a negative trough around 295 nm (Figure 1C).[45] CD spectra of *CEB1* show a positive peak at 260 nm and a negative trough at 240 nm, characteristic of a parallel G-quadruplex structure (Figure 1D).[61] These CD intensities approached near zero at 95 °C, indicating that the G-quadruplex structure was unfolded at high temperature.

***CEB1* G-Quadruplex Is a Dimer.** The multimeric nature of *CEB1* was revealed from studying its thermal stability as a function of DNA strand concentration in UV melting experiments. The melting temperature of *CEB1* increased from 64 °C at 2 μM to around 68 °C at 44 μM (Figure 2A).[46]
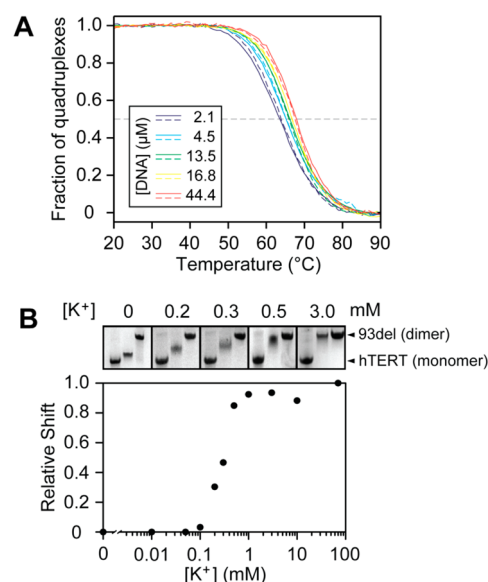

Figure 2. Formation of a dimer by *CEB1*. (A) Melting curves of *CEB1* at different DNA concentrations in 20 mM potassium phosphate buffer (pH 7). Solid and dashed lines represent cooling and heating curves, respectively. (B) Molecular size of *CEB1* at different K$^+$ concentrations, as observed by nondenaturing PAGE. Images of preformed G-quadruplexes migrated in gel supplemented with different K$^+$ concentrations were detected by UV shadowing. *hTERT* and *93del* were used as references for monomeric and dimeric G-quadruplexes, respectively. Relative mobility shift of *CEB1* as a function of K$^+$ concentration is shown in the bottom graph.

Plotting the concentration of the folded species versus that of the unfolded one at different temperatures[22] suggested that *CEB1* was formed from two subunits (Figure S1). Note that a hysteresis between the cooling and heating curves was observed at high DNA concentrations suggesting higher-ordered structure aggregation.

The stochiometry of *CEB1* at different K$^+$ concentrations was also probed by nondenaturing gel electrophoresis experiments (Figure 2B). At low K$^+$ concentration (0 to 0.1 mM), *CEB1* migrated as fast as the monomeric G-quadruplex reference formed by the *hTERT* promoter sequence.[62] At K$^+$ concentration above 3 mM, *CEB1* migrated at a similar rate as the dimeric G-quadruplex reference, *93del*.[32] The transition of *CEB1* from the monomer size to the dimer size was observed for the K$^+$ concentration between 0.1 and 3 mM. The relative

mobility shift of *CEB1*, normalized to that at 0 and 70 mM K$^+$, shows the time-averaged value for an interconversion between the dimeric and monomeric G-quadruplex structures. Compared to *93del*, which is a robust interlocked dimeric G-quadruplex, the *CEB1* can interconvert faster between dimeric and monomeric forms in gel electrophoresis experiments at very low K$^+$ concentration.

CEB1 Adopts a Stacked Dimeric G-Quadruplex Structure. Guanine imino (H1) and aromatic (H8) protons of *CEB1* were unambiguously assigned using site-specific low-enrichment ^{15}N labeling,[48] site-specific substitution ^{2}H labeling,[49] and through-bond correlations at natural abundance in a {^{13}C−^{1}H}-JRHMBC experiment (Figure 3 and Table S1).[50]
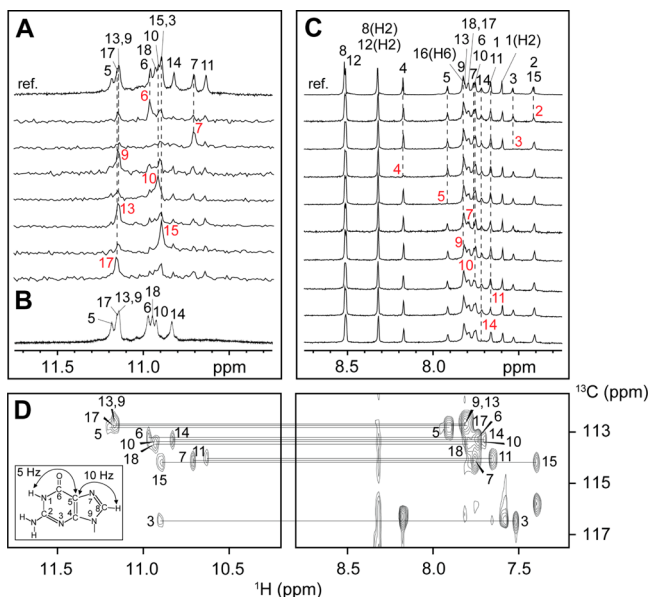

Figure 3. Spectral assignments of *CEB1*. (A) Imino proton assignments from ^{15}N-filtered spectra of samples, 2% ^{15}N-enriched at the indicated positions. (B) Imino proton spectrum after 15 min exposure in ^{2}H$_2$O solvent at 25 °C. (C) H8 proton assignments by site-specific ^{2}H labeling at the indicated positions. The reference spectra of imino and aromatic protons are shown at the top of the corresponding assignment spectra. (D) Through-bond correlations between guanine imino and H8 protons via ^{13}C5 at natural abundance, using long-range *J*-couplings shown in the inset. Assignments are labeled with residue numbers.

These unambiguous assignments were used cooperatively with other through-bond correlation experiments, such as TOCSY and {^{13}C−^{1}H}-HSQC, to trace the H8/H6−H1′ NOE sequential connectivity from A1 to G18 and to assign cross-peaks in NOESY spectra (Figure 4A,B).

Based on the H1−H8 proton cyclic NOE patterns, the alignment of three guanine tetrads, G3·G7·G11·G15, G18·G6·G10·G14 and G17·G5·G9·G13, and the overall folding topology of a stacked dimeric G-quadruplex of *CEB1* could be established (Figure 4C). For each subunit, three backbone columns (G5-G6-G7, G9-G10-G11, and G13-G14-G15) and the G17-G18 segment assemble in a parallel arrangement. The single guanine G3 completes the last column with opposite local strand directionality and participates in the (3 + 1) G3·G7·G11·G15 tetrad with G3 adopting *syn* glycosidic conformation. Consequently, G4 is placed in a V-shaped loop configuration bridging the first and third tetrads. The three
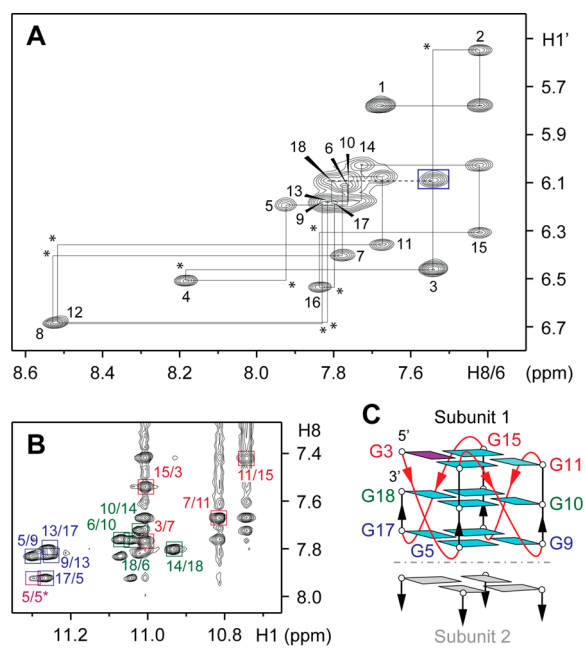

Figure 4. Folding topology of *CEB1* determined by NMR. (A) H8/H6−H1′ sequential connectivities in NOESY spectrum (mixing time, 300 ms). Intraresidue H8/6-H1′ cross-peaks are labeled with residue numbers. Missing cross peaks are marked with asterisks. The cross-peak corresponding to the NOE between G18(H1′) and G3(H8) is framed. (B) Imino-H8 cyclic connectivities in NOESY spectrum (mixing time, 300 ms). The tetrad arrangements were identified from cross-peaks which are framed and labeled with the residue number of the imino proton in the first position and that of the H8 proton in the second position. Intermolecular cross-peak is colored in magenta. (C) Schematic representation of stacked dimeric *CEB1* satisfying the connectivities shown in (A) and (B). *Anti* and *syn* tetrad guanines are colored cyan and magenta, respectively. The backbones of the core and loops are in black and red, correspondingly. For clarity, only one tetrad of subunit 2 (colored in grey) is shown.

remaining loops (A8, A12, and T16) are single-residue double-chain-reversal loops.

Broken H8/H6−H1′ NOE connectivities at various positions are consistent with the formation of the V-shaped (G4) and double-chain-reversal (A8, A12, and T16) loops (Figure 4A). NOE cross-peak between G18(H1′) and G3(H8) was observed, consistent with a snapback configuration of the 3′-end that places the terminal residue G18 in the close proximity with G3 (Figure 4C). The intensity of intraresidue H8−H1′ NOE cross-peaks indicated *syn* glycosidic conformation for A1 and G3 and *anti* conformation for all other residues (Figure S2).

The formation of a stacked dimeric G-quadruplex was supported by solvent exchange experiments (Figures 3B and S3), which showed that imino protons of the middle and the stacking interface G-tetrad layers (G18·G6·G10·G14 and G17·G5·G9·G13) were protected from the exchange with solvent, and by intermolecular NOE across the stacking dimerization interface (see below).

Effect of Mutations on the *CEB1* Stacked Dimeric G-Quadruplex. Several variant sequences (Table 1) were synthesized based on the folding topology of *CEB1* to assess structural features of the molecule. The role of A1 and G2 in capping the outer tetrads turned out to be important to presumably hinder possible aggregation, which happened for the mutant *A1ΔG2Δ* with both A1 and G2 being removed
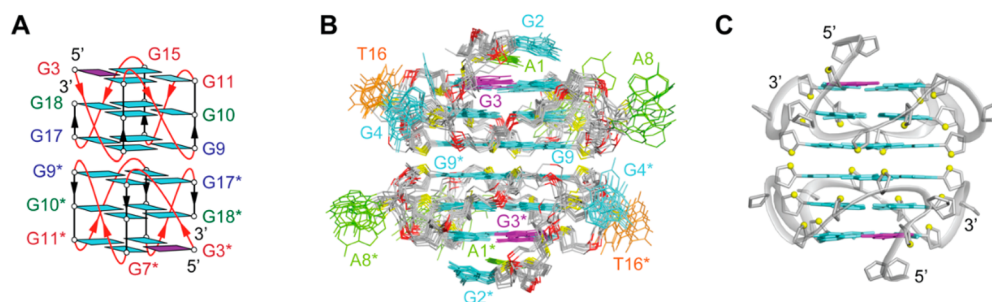
Figure 5. Structure of a stacked dimeric G-quadruplex formed by *CEB1*. (A) Schematic structure. (B) Ten superimposed structures following distance-restrained molecular dynamics refinement in explicit solvent. (C) Ribbon view of a representative structure. *Anti* and *syn* guanines are colored cyan and magenta, respectively; adenines, green; thymines, orange; backbone and sugar, gray; O4′ atoms, yellow; phosphorus atoms, red.

(Figure S4). The role of the 3′-end G17-G18 fragment was confirmed, as the mutant *G(17,18)T* with G17 and G18 being substituted by T, completely distorted the NMR spectrum (Figure S4). The question of how a tract of six guanines participates in a G-tetrad core is interesting in general genomic contexts, and in the case of *CEB1*, four guanine residues (G3, G5, G6 and G7) from the six-guanine tract are used in the G-tetrad core. While the G5-G6-G7 fragment forms a column supporting the G-tetrad core, only a single guanine (G3) out of the remaining three guanines is required to complete another column with the snapback fragment G17-G18. The fact that the G3 residue is at the best position to accomplish this role is consistent with the observation that a single G-to-T mutation at position 2 and 4 did not result in significant changes to the NMR spectrum, while a mutation at position 3 showed multiple G-quadruplex conformations (Figure S5). Analysis of NMR spectra of sequences containing two G-to-T mutations among the three positions (2, 3, and 4) established the formation of the same G-quadruplex fold (Figure S6−S11), demonstrating that any guanine among these three can complete the snapback motif like G3 does in *CEB1* (see below).

Stacking Orientation between Two Subunits of *CEB1*. The two subunits of *CEB1* are coupled at the dimerization interface by π−π stacking interactions between the G17·G5·G9·G13 tetrads of each subunit. At such a 5′-5′ stacking interface, the stacking tetrads should be in an opposite-polarity arrangement.[28,63] Due to the symmetry of a G-tetrad, there could be four possible isomers (or rotamers), differing by every 90° rotation of one G-quadruplex block with respect to the other about the common helical axis (Figure S11).

The observed NOE cross-peak between G5(H1) and G5(H8) (Figure 4B) could only occur across the stacking interface, between the guanines G5 of the two different subunits. This NOE cross-peak determines the *CEB1* isomer, in which the G5 residue of one subunit is directly on top of the G5* residue of the other subunit (Figure 5A). All the other three rotation isomers would result in the G5(H1)-G5*(H8) distance larger than 8 Å. Due to the spectral overlap, no other NOEs could be distinctly confirmed across the stacking interface of *CEB1*. Therefore, we could not confirm or preclude the presence of other stacking orientations in *CEB1*.

In the mutated sequences *G(2,4)T*, *G(2,3)T*, and *G(3,4)T*, where G-to-T substitutions were performed at two out of three possible positions (indicated by numbers in the sample names), we showed by NMR the adoption of a similar stacked dimeric G-quadruplex (Figures S6−S11). Furthermore, NOESY spectra of (i) *G(2,4)T* showed G5(H1)-G5*(H8), G17(H1)-G17*(H8) and [G17(H1)-G9*(H8) and/or G17(H1)-

G13*(H8)] cross-peaks across the stacking interface; (ii) *G(2,3)T* showed G5(H1)-G5*(H8), G17(H1)-G13*(H8), and [G17(H1)-G9*(H8) and/or G17(H1)-G17*(H8)] cross peaks across the stacking interface; (iii) *G(3,4)T* showed G5(H1)-G5*(H8), G9(H1)-G5*(H8), G13(H1)-G5*(H8), G17(H1)-G17*(H8), and [G17(H1)-G9*(H8) and/or G17(H1)-G13*(H8)] cross-peaks across the stacking interface. While the G5(H1)-G5*(H8) and G17(H1)-G9*(H8) cross-peaks determine the same stacking isomer described in *CEB1*, the G17(H1)-G17*(H8) and G13(H1)-G5*(H8) cross-peaks establish a stacking arrangement where G17 of one subunit is on top of G17* of the other subunit (180° rotation), and the G9(H1)-G5*(H8) and G17(H1)-G13*(H8) cross-peaks determine a stacking arrangement where G17 (G5) of one subunit is on top of G13* (G9*) of the other subunit (270° rotation) (Figure S11). The "90° rotation" stacking arrangement could not be observed in these sequences due to spectral overlap.

Solution Structure of a *CEB1* Stacked Dimeric G-Quadruplex. We performed structure calculation for the first stacking isomer of *CEB1*, which could be experimentally confirmed for this sequence. The solution structure of *CEB1* was initially computed using the XPLOR program according to the set of restraints derived from NMR experiments (Table 2).

Table 2. Statistics of the Computed Structures of the Stacked Dimeric *CEB1* G-Quadruplex[a]

(A) NMR Restraints[b]		
Distance Restraints	^2H_2O	H_2O
intraresidue	586	12
sequential (i, $i + 1$)	232	42
long-range (i, $\geq i + 2$)	52	168
Other Restraints		
hydrogen bond	96	
dihedral angle	62	
(B) Structure Statistics		
NOE Violations		
number (>0.2 Å)	0	
Deviations from the Ideal Covalent Geometry		
bond lengths (Å)	0.004 ± 0.000	
bond angles (°)	0.720 ± 0.010	
impropers (°)	0.370 ± 0.007	
Pairwise All Heavy Atom rmsd Values (Å)		
G-quadruplex core, A1 and G2	0.84 ± 0.29	
all residues	1.55 ± 0.47	

[a]PDB ID: 2MB4. [b]Restraints were duplicated for the two subunits.

Nevertheless, as only a limited number of NOEs could be distinctly confirmed across the stacking interface, we subsequently performed a molecular dynamics (MD) refinement on stacking orientation between the two subunits of *CEB1*.[64] Distance-restrained MD refinement of the ten structures generated in XPLOR revealed a convergence in interface stacking geometry within the first nanosecond of simulation (Figure S12).

The computed structures (Figure 5) are well converged with the overall root-mean-square deviation (rmsd) of 1.55 Å. The base of A1 sits on top of the Hoogsteen hydrogen bonds between G11 and G15, whereas the sugar moiety of A1 is located on top of the G15 base, consistent with the observation of numerous NOE cross-peaks between A1(H1′/2′/2″/4′/5′/5″) and G15(H1). Being constrained by NOE cross-peaks from G2(H8) and sugar protons of both A1 and G2, the base of G2 is converged above G3 and G7. The loop residues G4, A8, A12, and T16 are less well-defined, possibly due to the dynamics of these loops. The converged stacking arrangement of *CEB1* is described by a relative rotation angle (θ_{rot}) of about 55−60° corresponding to a "5/6-ring" stacking orientation with a partial overlap of the five- and six-membered rings (Figure 6).[63]

MD Study on the Preferential Stacking Orientation between Subunits of *CEB1*. Further MD simulations were performed in order to explore the geometry of the *CEB1*

Figure 6. Structure of *CEB1* stacking interface. (A) Top view along helical axis showing a partial overlap of the five- and six-membered rings of stacked guanine bases. (B) Side view showing intermolecular distances between subunits derived from observed NOE cross peaks. α represents 5(H1)-5*(H8); α*, 5*(H1)-5(H8); β, 9(H8)-17*(H1); β*, 17*(H1)-9(H8) distances. β or β* cross peak was observed in NOESY spectra of mutated sequences. Subunits are colored in blue and red, respectively.

dimeric interface on a longer time scale and to investigate the geometries adopted by other rotamers of stacked *CEB1* subunits. Starting from the reported NMR solution structure of *CEB1*, one subunit of the dimer was manually rotated by 90°, 180°, and 270°. Unrestrained MD simulations of these four rotamers were run to a length of 200 ns, and their trajectories were analyzed. All four *CEB1* rotamers maintained stable stacked orientations over the course of the simulation (Figure S13), in agreement with experimental observation of multiple rotamers being adopted by mutated *CEB1* sequences. Additionally, a similar geometry was adopted and maintained for all four rotamers over the entire 200 ns simulation, with a 5/6-ring orientation and a separation distance of 3.5 Å between interface tetrads being observed (Figures S14 and S15). These results suggest that the dimeric interface geometry observed in the NMR solution structure of *CEB1* is a stable configuration.

Conformational Exchange in *CEB1*: A Base Swapping Motion at the 5′-End. The observation of low-intensity peaks along with the major peaks in the imino proton NMR spectrum of *CEB1* indicated the presence of a minor G-quadruplex conformation (Figure 1B). A number of exchange cross-peaks pertaining to minor imino and aromatic protons were identified in NOESY spectra and further confirmed using a ROESY experiment (Figures 7 and S16). This low-populated conformation (∼7% of the total population) was found to be in conformational exchange at the millisecond time scale (estimated exchange time ∼600 ms) with the major stacked dimeric G-quadruplex form (Figure S17).

The exchange dynamics was localized at the 5′ end of the strand, mostly involving A1, G2, G3, and G4. The largest chemical shift changes was observed for the G4(H8) and G3(H1′) protons. We propose a model for the minor conformation, in which G4[#] (instead of G3 in the major conformation) participates in the formation of the G-tetrad core (Figure 7C), based on the following: (i) upfield shifting of the G4(H8) proton from a frequency characteristic of a propeller loop of a purine residue (such as G4, A8, and A12) in the major form to a frequency characteristic of a tetrad-bound *syn* guanine in the minor form (Figure 7C); (ii) upfield shifting of the G3(H1′) proton to a frequency characteristic for a residue situated above the G-tetrad core (such as A1 and G2) in the minor form (Figure 7C); (iii) a relatively strong intensity of the intraresidue H8−H1′ NOE cross-peak for G4 in the minor form indicating a *syn* glycosidic conformation (Figure S18); and (iv) mutated sequences *G4T* and *G(2,4)T* (but not *G2T*) showing the exchange cross-peaks suppressed (see below).

The *G2T*, *G4T*, and *G(2,4)T* sequences containing G-to-T substitutions at positions 2 and 4 or both positions were analyzed to investigate the conformational exchange at the 5′-end of the *CEB1* G-quadruplex (Figure 7C). Exchange cross-peaks were observed in NOESY spectra of *CEB1* and *G2T* but not in the corresponding spectra of *G4T* and *G(2,4)T* (Figure 8), consistent with the model for a swapping dynamics between G3 and G4 bases, which were suppressed when G4 was mutated.

A G-quadruplex fold equivalent to the minor form could be constructed by the *G(2,3)T* mutated sequence. Unambiguous assignments of NMR spectra of *G(2,3)T* rigorously established the existence of a G-quadruplex with zero-nucleotide V-shaped loop (Figures S7 and S9). Furthermore, similar G-quadruplex fold utilizing guanine at position 2 and thus involving two-
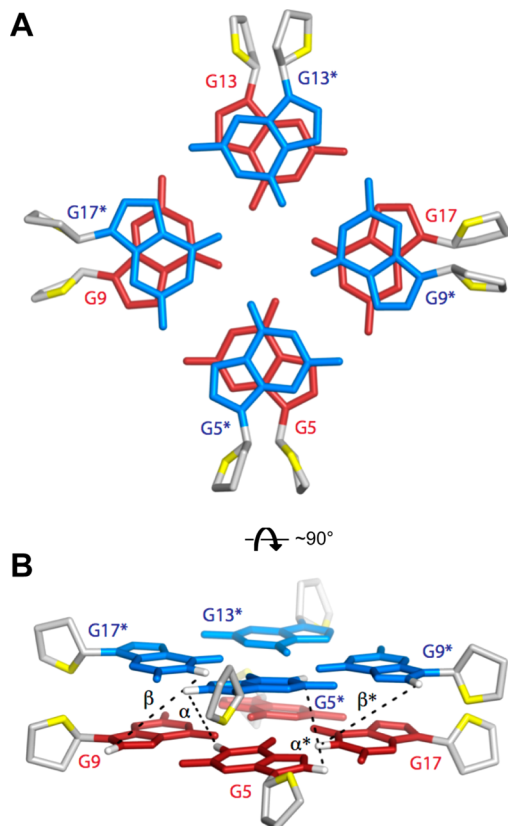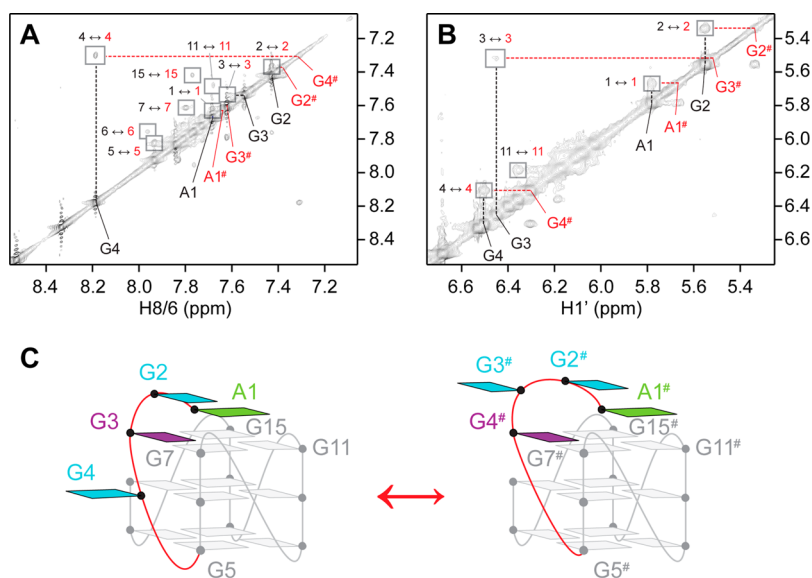
Figure 7. Conformational exchange in *CEB1*. (A,B) ROESY spectra (mixing time, 200 ms): diagonal peaks are labeled with residues from major (black) and minor (red) forms. ROE cross-peaks resulted from the conformational exchange are framed. (C) The schematic representation of a *CEB1* subunit: the major and minor forms are shown on the left and right, respectively. Dynamic residues and loops are drawn in colors. *Syn* tetrad guanines are in magenta; free guanines, cyan; adenines, green; loop backbones, red.
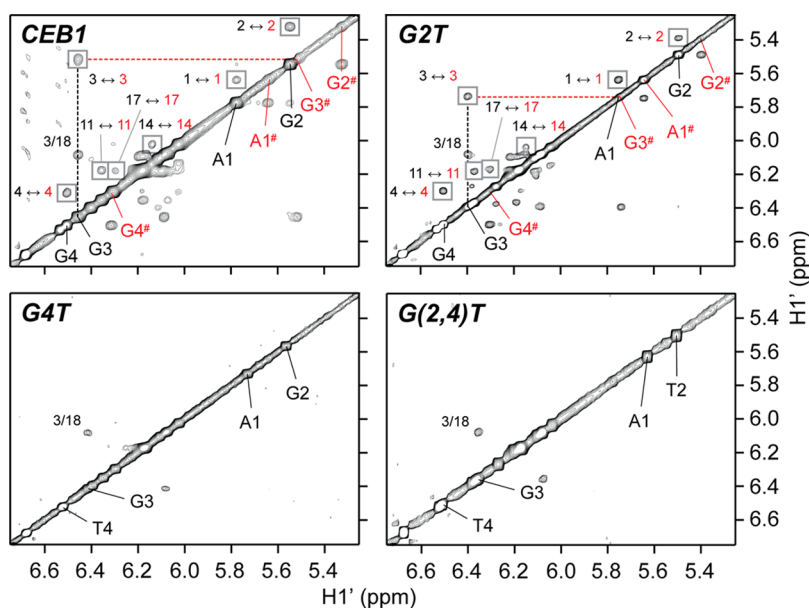

Figure 8. Conformational exchange cancelled by residue-specific mutations. NOESY spectra (mixing time, 300 ms) of *CEB1*, *G2T*, *G4T*, and *G(2,4)T*. The exchange cross-peaks (framed) present in the spectra of *CEB1* and *G2T* disappeared in those of *G4T* and *G(2,4)T*. The signature NOE cross-peaks between the G3(H1′) and G8(H1′) protons are labeled "3/18".

nucleotide V-shaped loop was also shown to be formed by the *G(3,4)T* mutated sequence (Figures S8 and S10).

■ DISCUSSION

We have shown that the truncated human CEB1 minisatellite sequence d(AGGGGGGAGGGAGGGTGG), containing a 6-G, two 3-G, and a 2-G tracts, forms a stacked dimeric G-quadruplex. In each subunit, the *CEB1* sequence folds into a parallel snapback G-quadruplex with four Gs out of the 6-G tract and all Gs of the remaining G-tracts participating in the G-tetrad core. The *CEB1* G-quadruplex has some robust structural elements as previously observed including the parallel-stranded arrangement and single-residue double-chain-reversal (or propeller) loops. It also has some particular structural motifs including the snapback fold and a V-shaped loop. Different snapback parallel G-quadruplexes have been observed in several oncogenic promoters including *c-kit*,[21] *c-myc*,[20] and *PDGFR-β*,[65] while V-shaped loop was observed in human *chl1* intronic G-quadruplex scaffold.[23] Another particular feature of the *CEB1* G-quadruplex is the presence of a long 6-G tract. The involvement of a long G-tract (≥4 nt) in a G-quadruplex has been previously reported.[33,66] In the crystal structure of an interlocked dimeric G-quadruplex formed by a *B-raf* gene promoter sequence, G-tracts of four and five nucleotides were observed to take part in a G-quadruplex structure involving seven G-tetrad layers.[33] A five-G tract in a VEGF aptamer

sequence was shown to facilitate the folding of a unique three-layered G-quadruplex involving a zero-nucleotide propeller loop within the G-tract.[66]

Previous study on the folding pathways of intramolecular G-quadruplexes suggested the involvement of few transient species such as hairpin and G-triplex intermediates.[67−69] In the structure of the *CEB1* G-quadruplex, the exchange position between G3 and G4 requires the tetrad-bound G3 to first detach from the tetrad core before G4 can occupy now vacant slot on the outermost tetrad. During this swapping process, the existence of an intermediate guanine triad is expected to maintain the overall scaffold unchanged.

Spectral overlap did not allow us to definitively confirm or preclude the presence of four-fold stacking orientations in *CEB1*. Nonetheless, better-resolved NOESY spectra of mutant samples lead to the observation of other stacking orientations. In previously reported structures of stacked G-quadruplexes formed by a telomeric RNA sequence[70] and anti-HIV G-rich oligunucleotide,[28] a major mode of stacking was suggested to result from repulsive interactions between terminal 5′-OH groups located on each stacking plane. In *pilE* stacked dimeric G-quadruplex, the observed stacking arrangement was presumably favored by even distribution of extruding loops along the axis of rotation.[31] In our study, MD computations show that each rotamer adopts a stable "5/6-ring" orientation which is maintained over the length of the 200 ns simulation. A highly symmetric dimerization-stacking interface and similar propeller loop sizes of *CEB1* may result in different rotamers being more equally populated.

Our preliminary mutagenesis study shows that the 6-G tract located at the 5′-end of the *CEB1* G-quadruplex plays an important role in CEB1 genomic instability. Due to the difficulty in sequencing CEB1 arrays, our knowledge on the polymorphisms between the motifs of the same array or different arrays of the human population is still limited. However, considering the CEB1−1.8 allele (42 repeats) that we succeeded to sequence,[71] this 6-G tract is evolutionary extremely well conserved: 41/42 motifs carry either the same 6-G tract (33 motifs) or a single-G substitution leaving a 5-G, 4-G, or 3-G tract. Our structural data show that ±1 nt mutations in this G-tract do not affect the G-quadruplex folding, as sufficient guanines remain to allow the same G-quadruplex fold.

As compared with G-quadruplexes formed by a nonrepetitive promoter sequence, the stacked dimeric *CEB1* G-quadruplex, if formed in a repetitive array of the CEB1 minisatellite, would occur between different repeats. In addition to G-quadruplex structures formed within individual repeats, as observed for the CEB25 minisatellite,[72] structures formed between different repeats would give additional structural features in the chromatin organization and might synergistically contribute to modulate the genomic instability in minisatellites.

CEB1 is located in the promoter region of the human subtelomeric NEU4 gene encoding a sialidase enzyme.[73] NEU4 is associated with diseases including galactosialidosis and clear-cell adenocarcinoma.[74] The questions of whether the length variation of CEB1 alleles and/or their capacity to fold into G quadruplexes affects NEU4 expression in human cells, perturbs replication, or stimulates chromosomal rearrangements as observed in yeast[75] remain to be examined.

CONCLUSION

We report on the structure of a stacked dimeric G-quadruplex (*CEB1*) formed by the G-rich fragment of the human CEB1 minisatellite in K^+ solution. Each subunit of *CEB1* forms a parallel snapback scaffold characterized with a unique conformational exchange in the millisecond time scale whereby a guanine residue in a V-shaped loop configuration could swap position with a tetrad-bound guanine in proximity. Multiple stacking rotamers may be present due to a high symmetry of the 5′-tetrad dimerization-stacking interface.

ASSOCIATED CONTENT

Ⓢ Supporting Information

Additional experimental data and analyses (Tables S1−S3 and Figures S1−S18). This material is available free of charge via the Internet at http://pubs.acs.org.

AUTHOR INFORMATION

Corresponding Author

phantuan@ntu.edu.sg

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

REFERENCES

(1) Sen, D.; Gilbert, W. *Nature* **1988**, *334*, 364.

(2) Davis, J. T. *Angew. Chem., Int. Ed.* **2004**, *43*, 668.

(3) Patel, D. J.; Phan, A. T.; Kuryavyi, V. *Nucleic Acids Res.* **2007**, *35*, 7429.

(4) Neidle, S. *Curr. Opin. Struct. Biol.* **2009**, *19*, 239.

(5) Bochman, M. L.; Paeschke, K.; Zakian, V. A. *Nat. Rev. Genet.* **2012**, *13*, 770.

(6) Todd, A. K.; Johnston, M.; Neidle, S. *Nucleic Acids Res.* **2005**, *33*, 2901.

(7) Huppert, J. L.; Balasubramanian, S. *Nucleic Acids Res.* **2005**, *33*, 2908.

(8) De Cian, A.; Lacroix, L.; Douarre, C.; Temime-Smaali, N.; Trentesaux, C.; Riou, J. F.; Mergny, J. L. *Biochimie* **2008**, *90*, 131.

(9) Siddiqui-Jain, A.; Grand, C. L.; Bearss, D. J.; Hurley, L. H. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 11593.

(10) Simonsson, T.; Pecinka, P.; Kubista, M. *Nucleic Acids Res.* **1998**, *26*, 1167.

(11) Vallur, A. C.; Maizels, N. *Proc. Natl. Acad. Sci. U.S.A.* **2008**, *105*, 16508.

(12) Maizels, N. *Nat. Struct. Mol. Biol.* **2006**, *13*, 1055.

(13) Ribeyre, C.; Lopes, J.; Boule, J. B.; Piazza, A.; Guedin, A.; Zakian, V. A.; Mergny, J. L.; Nicolas, A. *PLoS. Genet.* **2009**, *5*, e1000475.

(14) Mergny, J. L.; Helene, C. *Nat. Med.* **1998**, *4*, 1366.

(15) Rankin, S.; Reszka, A. P.; Huppert, J.; Zloh, M.; Parkinson, G. N.; Todd, A. K.; Ladame, S.; Balasubramanian, S.; Neidle, S. *J. Am. Chem. Soc.* **2005**, *127*, 10584.

(16) Wu, Y.; Brosh, R. M., Jr. *FEBS J.* **2010**, *277*, 3470.

(17) Brooks, T. A.; Hurley, L. H. *Nat. Rev. Cancer* **2009**, *9*, 849.

(18) Balasubramanian, S.; Hurley, L. H.; Neidle, S. *Nat. Rev. Drug Discov.* **2011**, *10*, 261.

(19) Phan, A. T. *FEBS J.* **2010**, *277*, 1107.

6304

dx.doi.org/10.1021/ja4125274 | *J. Am. Chem. Soc.* 2014, 136, 6297−6305

(20) Phan, A. T.; Kuryavyi, V.; Gaw, H. Y.; Patel, D. J. *Nat. Chem. Biol.* **2005**, *1*, 167.

(21) Phan, A. T.; Kuryavyi, V.; Burge, S.; Neidle, S.; Patel, D. J. *J. Am. Chem. Soc.* **2007**, *129*, 4386.

(22) Mukundan, V. T.; Phan, A. T. *J. Am. Chem. Soc.* **2013**, *135*, 5017.

(23) Kuryavyi, V.; Patel, D. J. *Structure* **2010**, *18*, 73.

(24) Nielsen, J. T.; Arar, K.; Petersen, M. *Angew. Chem., Int. Ed.* **2009**, *48*, 3099.

(25) Zhang, N.; Gorin, A.; Majumdar, A.; Kettani, A.; Chernichenko, N.; Skripkin, E.; Patel, D. J. *J. Mol. Biol.* **2001**, *311*, 1063.

(26) Crnugelj, M.; Sket, P.; Plavec, J. *J. Am. Chem. Soc.* **2003**, *125*, 7866.

(27) Kuryavyi, V.; Phan, A. T.; Patel, D. J. *Nucleic Acids Res.* **2010**, *38*, 6757.

(28) Do, N. Q.; Lim, K. W.; Teo, M. H.; Heddi, B.; Phan, A. T. *Nucleic Acids Res.* **2011**, *39*, 9448.

(29) Trajkovski, M.; da Silva, M. W.; Plavec, J. *J. Am. Chem. Soc.* **2012**, *134*, 4132.

(30) Do, N. Q.; Phan, A. T. *Chem.—Eur. J.* **2012**, *18*, 14752.

(31) Kuryavyi, V.; Cahoon, L. A.; Seifert, H. S.; Patel, D. J. *Structure* **2012**, *20*, 2090.

(32) Phan, A. T.; Kuryavyi, V.; Ma, J. B.; Faure, A.; Andreola, M. L.; Patel, D. J. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 634.

(33) Wei, D.; Todd, A. K.; Zloh, M.; Gunaratnam, M.; Parkinson, G. N.; Neidle, S. *J. Am. Chem. Soc.* **2013**, *135*, 19319.

(34) Deng, J.; Xiong, Y.; Sundaralingam, M. *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98*, 13665.

(35) Pan, B.; Xiong, Y.; Shi, K.; Deng, J.; Sundaralingam, M. *Structure* **2003**, *11*, 815.

(36) Maizels, N.; Gray, L. T. *PLoS Genet.* **2013**, *9*, e1003468.

(37) Vergnaud, G.; Mariat, D.; Apiou, F.; Aurias, A.; Lathrop, M.; Lauthier, V. *Genomics* **1991**, *11*, 135.

(38) Buard, J.; Vergnaud, G. *EMBO J.* **1994**, *13*, 3203.

(39) Debrauwere, H.; Buard, J.; Tessier, J.; Aubert, D.; Vergnaud, G.; Nicolas, A. *Nat. Genet.* **1999**, *23*, 367.

(40) Berg, I. L.; Neumann, R.; Lam, K. W.; Sarbajna, S.; Odenthal-Hesse, L.; May, C. A.; Jeffreys, A. J. *Nat. Genet.* **2010**, *42*, 859.

(41) Piazza, A.; Boule, J. B.; Lopes, J.; Mingo, K.; Largy, E.; Teulade-Fichou, M. P.; Nicolas, A. *Nucleic Acids Res.* **2010**, *38*, 4337.

(42) Lopes, J.; Piazza, A.; Bermejo, R.; Kriegsman, B.; Colosio, A.; Teulade-Fichou, M. P.; Foiani, M.; Nicolas, A. *EMBO J.* **2011**, *30*, 4033.

(43) Cantor, C. R.; Warshaw, M. M.; Shapiro, H. *Biopolymers* **1970**, *9*, 1059.

(44) Guedin, A.; De Cian, A.; Gros, J.; Lacroix, L.; Mergny, J. L. *Biochimie* **2008**, *90*, 686.

(45) Mergny, J. L.; Li, J.; Lacroix, L.; Amrane, S.; Chaires, J. B. *Nucleic Acids Res.* **2005**, *33*, e138.

(46) Mergny, J. L.; Phan, A. T.; Lacroix, L. *FEBS Lett.* **1998**, *435*, 74.

(47) Mergny, J. L.; Lacroix, L. *Oligonucleotides* **2003**, *13*, 515.

(48) Phan, A. T.; Patel, D. J. *J. Am. Chem. Soc.* **2002**, *124*, 1160.

(49) Huang, X.; Yu, P.; LeProust, E.; Gao, X. *Nucleic Acids Res.* **1997**, *25*, 4758.

(50) Phan, A. T. *J. Biomol. NMR* **2000**, *16*, 175.

(51) Phan, A. T.; Gueron, M.; Leroy, J. L. *Methods Enzymol.* **2001**, *338*, 341.

(52) Jeener, J.; Meier, B. H.; Bachmann, P.; Ernst, R. R. *J. Chem. Phys.* **1979**, *71*, 4546.

(53) Schwieters, C. D.; Kuszewski, J. J.; Tjandra, N.; Clore, G. M. *J. Magn. Reson.* **2003**, *160*, 65.

(54) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. *J. Chem. Phys.* **1983**, *79*, 926.

(55) Berendsen, H. J. C.; Postma, J. P. M.; Vangunsteren, W. F.; Dinola, A.; Haak, J. R. *J. Chem. Phys.* **1984**, *81*, 3684.

(56) Vangunsteren, W. F.; Berendsen, H. J. C. *Mol. Phys.* **1977**, *34*, 1311.

(57) Darden, T.; York, D.; Pedersen, L. *J. Chem. Phys.* **1993**, *98*, 10089.

(58) Case, D. A.; Cheatham, T. E., III; Darden, T.; Gohlke, H.; Luo, R.; Merz, K. M., Jr.; Onufriev, A.; Simmerling, C.; Wang, B.; Woods, R. J. *J. Comput. Chem.* **2005**, *26*, 1668.

(59) Perez, A.; Marchan, I.; Svozil, D.; Sponer, J.; Cheatham, T. E., III; Laughton, C. A.; Orozco, M. *Biophys. J.* **2007**, *92*, 3817.

(60) Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz, K. M.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P. A. *J. Am. Chem. Soc.* **1996**, *118*, 2309.

(61) Gray, D. M.; Wen, J. D.; Gray, C. W.; Repges, R.; Repges, C.; Raabe, G.; Fleischhauer, J. *Chirality* **2008**, *20*, 431.

(62) Lim, K. W.; Lacroix, L.; Yue, D. J.; Lim, J. K.; Lim, J. M.; Phan, A. T. *J. Am. Chem. Soc.* **2010**, *132*, 12331.

(63) Lech, C. J.; Heddi, B.; Phan, A. T. *Nucleic Acids Res.* **2013**, *41*, 2034.

(64) Sponer, J.; Spackova, N. *Methods* **2007**, *43*, 278.

(65) Chen, Y.; Agrawal, P.; Brown, R. V.; Hatzakis, E.; Hurley, L.; Yang, D. *J. Am. Chem. Soc.* **2012**, *134*, 13220.

(66) Marusic, M.; Veedu, R. N.; Wengel, J.; Plavec, J. *Nucleic Acids Res.* **2013**, *41*, 9524.

(67) Zhang, A. Y.; Balasubramanian, S. *J. Am. Chem. Soc.* **2012**, *134*, 19297.

(68) Mashimo, T.; Yagi, H.; Sannohe, Y.; Rajendran, A.; Sugiyama, H. *J. Am. Chem. Soc.* **2010**, *132*, 14910.

(69) Limongelli, V.; De Tito, S.; Cerofolini, L.; Fragai, M.; Pagano, B.; Trotta, R.; Cosconati, S.; Marinelli, L.; Novellino, E.; Bertini, I.; Randazzo, A.; Luchinat, C.; Parrinello, M. *Angew. Chem., Int. Ed.* **2013**, *52*, 2269.

(70) Martadinata, H.; Phan, A. T. *J. Am. Chem. Soc.* **2009**, *131*, 2570.

(71) Lopes, J.; Ribeyre, C.; Nicolas, A. *Mol. Cell. Biol.* **2006**, *26*, 6675.

(72) Amrane, S.; Adrian, M.; Heddi, B.; Serero, A.; Nicolas, A.; Mergny, J. L.; Phan, A. T. *J. Am. Chem. Soc.* **2012**, *134*, 5807.

(73) Seyrantepe, V.; Landry, K.; Trudel, S.; Hassan, J. A.; Morales, C. R.; Pshezhetsky, A. V. *J. Biol. Chem.* **2004**, *279*, 37021.

(74) Miyagi, T. *Proc. Jpn. Acad. Ser. B Phys. Biol. Sci.* **2008**, *84*, 407.

(75) Piazza, A.; Serero, A.; Boulé, J. B.; Legoix-Né, P.; Lopes, J.; Nicolas, A. *PLoS Genet.* **2012**, *8*, e1003033.