BIOCHEMISTRY
including biophysical chemistry & molecular biology

Article

pubs.acs.org/biochemistry

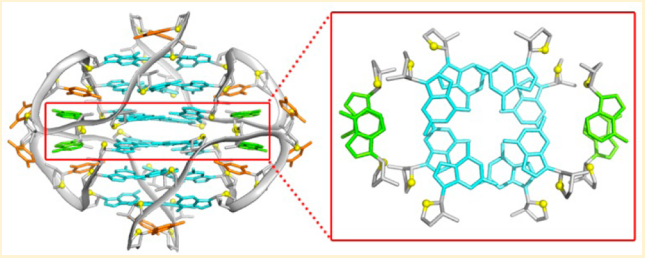
Structure of Human Telomeric RNA (TERRA): Stacking of Two G-Quadruplex Blocks in K^+ Solution

Herry Martadinata[†,‡] and Anh Tuân Phan*[,†]

[†]School of Physical and Mathematical Sciences and [‡]School of Biological Sciences, Nanyang Technological University, Singapore

Supporting Information

ABSTRACT: Telomeric repeat-containing RNAs (TERRA) are transcription products of the telomeres. Human TERRA sequences containing UUAGGG repeats can form parallel-stranded G-quadruplexes. The stacking interaction of such structures was shown to be important for ligand targeting and higher-order arrangement of G-quadruplexes in long TERRA sequences. Here we report on the first high-resolution structure of a stacked G-quadruplex formed by the 10-nucleotide human TERRA sequence r(GGGUUAGGGU) in potassium solution. This structure comprises two dimeric three-layer parallel-stranded G-quadruplex blocks, which stack on each other at their 5′-ends. The adenine in each UUA loop is nearly coplanar with the 5′-end G-tetrad forming an A·(G·G·G·G)·A hexad, thereby increasing the stacking contacts between the two blocks. Interestingly, this stacking and loop conformation is different from all structures previously reported for the free human TERRA but resembles the structure previously determined for a complex between a human TERRA sequence and an acridine ligand. This stacking conformation is a potential target for drugs that recognize or induce the stacking interface.



Telomeres, the nucleoprotein complexes at the ends of linear eukaryotic chromosomes, protect chromosomes against genome instability.[1] Besides the protective role, telomeres also act as internal timers for a cell before it eventually undergoes senescence or apoptosis.[2,3] Telomeric DNA in vertebrates consists of tandem repeats of the TTAGGG sequence[4] with a 3′-end single-stranded overhang.[5]

Telomeres were thought to be transcriptionally silent until the recent discovery of various telomeric transcripts,[6−9] including the RNAs named TERRA that contain telomeric UUAGGG repeats in vertebrates.[6,7,10,11] Various regulatory functions have been assigned to TERRA, such as heterochromatin regulation, telomerase inhibition, telomere length regulation, and telomere protection.[6,7,10−19]

Because of their G-rich nature, telomeric DNA and RNA can fold into G-quadruplexes, a type of nucleic acid structures formed by stacking of G·G·G·G tetrads.[20−22] Formation of G-quadruplex structures by telomeric DNA and RNA oligonucleotides has been demonstrated.[23−47] In contrast to the conformational diversity of telomeric DNA G-quadruplexes,[23−35] TERRA was found to form only parallel-stranded G-quadruplexes.[33,36−47]

To date, three atomic-resolution structures have been reported for G-quadruplexes formed by human TERRA sequences.[37,44,46] The first structure (PDB entry 2KBP) was determined by using NMR spectroscopy for the 12-nucleotide sequence r(UAGGGUUAGGGU), which forms a dimeric three-layer parallel-stranded propeller-type G-quadruplex in potassium solution.[37] Subsequently, the crystal structure (PDB entry 3IBK) of a modified version of this sequence, r(^{Br}UAGGGUUAGGGU), also revealed the same G-quadruplex fold.[44] In addition, two such G-quadruplex blocks were observed to interlock with each other at their 5′-ends through the formation of two U·A·U·A tetrads between the two G-quadruplex blocks by the first two residues of the RNA sequence. The third structure (PDB entry 3MIJ) was reported for the r(UAGGGUUAGGGU) sequence cocrystallized with an acridine ligand.[46] Similar to the previous crystal structure, two parallel-stranded G-quadruplex blocks were observed to associate with each other at their 5′-ends. However, there are two major differences compared to the previous crystal structure. First, two acridine molecules are inserted and stacked between the two G-quadruplex blocks. Second, the 5′-ends (UA) and the loops (UUA) are rearranged, and all the adenine bases are positioned in a coplanar arrangement with the 5′-end G-tetrad forming an (A:G:A:G:A:G:A:G) octad. Thus, the 5′-end of the TERRA parallel-stranded propeller-type G-quadruplex represents an attractive surface for molecular recognition. An important question is whether such a conformation preexists in free RNA (being recognized and trapped by the acridine) or is induced only in the presence of this ligand.

These structures and other studies indicate that parallel-stranded G-quadruplexes have a strong propensity for stacking.[37,44−46,48−52] Studies have also suggested that the stacking interaction is important for higher-order arrangements of G-quadruplexes in long TERRA sequences.[37,47] It has been

Received: November 30, 2012
Revised: February 25, 2013
Published: February 28, 2013

proposed that TERRA G-quadruplexes might be arranged as "beads-on-a-string", in which each bead is a single G-quadruplex or two-block stacked G-quadruplexes.[25,30,32,47]

Previously, we showed that the 10-nucleotide human TERRA sequence r(GGGUUAGGGU) formed a structure in which two blocks of parallel-stranded propeller-type G-quadruplexes stacked at their 5′-ends (Figure 1).[37] However, the structural details of the stacking interface and loop conformations are not yet available. Here we present the first high-resolution structure of this stacked RNA G-quadruplex in potassium solution. The stacking of two G-quadruplex blocks induces considerable loop rearrangements. Interestingly, the stacking and loop conformation is different from those of all structures previously reported for free human TERRA but similar to that observed in the crystal structure of a complex between a human TERRA sequence and an acridine ligand.[46] This stacking conformation might serve as an attractive target for drugs that recognize the stacking interface or induce the stacking conformation.

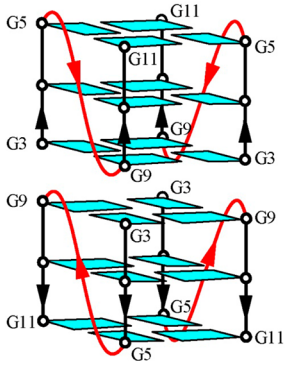


Figure 1. Schematic structure of a stacked parallel-stranded propeller-type G-quadruplex formed by the 10-nucleotide human TERRA sequence r(GGGUUAGGGU) in potassium solution. *anti* guanines are colored cyan, and UUA linkers are colored red.

MATERIALS AND METHODS

Sample Preparation. The oligoribonucleotides were either purchased from Research Instrument Pte Ltd. (Singapore) with PAGE purification or synthesized on an ABI 394 DNA/RNA synthesizer and purified according to the manufacturer's protocol (Glen Research). The sequences used in this study are listed in Table 1. The samples were dissolved in a buffer containing 70 mM KCl and 20 mM potassium phosphate (pH 7). The samples were heated at 100 °C and slowly cooled to room temperature overnight.

Nuclear Magnetic Resonance (NMR) Spectroscopy. All NMR experiments were performed on a 600 MHz Bruker NMR spectrometer at 25 °C. The oligonucleotide strand concentrations were in the range of 0.1−1.3 mM. JR-type water suppression was used for experiments in H_2O.[53,54] Site-specific ribose-to-deoxyribose substitutions were performed for the purposes of assignment.[37,55] Heteronuclear through-bond correlation experiments at the ^{13}C natural abundance, [$^{13}C-^{1}H$] JR-HMBC and [$^{13}C-^{1}H$] HSQC, and homonuclear experiments, TOCSY and COSY, were used to assist spectral assignments.[54,56] NOESY experiments were performed with mixing times of 120 and 300 ms in D_2O and 200 ms in H_2O.

XPLOR Structure Calculation. The structures of the G-quadruplex formed by the 10-nucleotide human TERRA sequence r(GGGUUAGGGU) were first calculated using the XPLOR-NIH program, version 2.27.[57,58] NMR-restrained computations were performed as described previously.[37] Structures were displayed using PyMOL.[59] The numbers of NOE and hydrogen-bond distance restraints and torsion angle restraints were quadrupled for all four strands.

Topology and Parameter Files. The topology (nucleic-2.0.top) and parameter (nucleic-2.0.param) files[58] were used for all computations.

NOE Distance Restraints. The distances between non-exchangeable protons of the 10-nucleotide human TERRA sequence r(GGGUUAGGGU) were obtained from NOESY cross-peaks at various mixing times (120 and 300 ms). Buildup measurements were performed to derive the distance restraints for each peak. H5−H6 NOE peaks of uracil bases were used to calibrate the buildup distance measurements. The upper and lower boundary limits were ±30% of the distance value obtained from the buildup distance measurements. Overlapping peaks were classified manually as strong (strong peaks at 120 ms), medium (peaks that were weak but observable at 120 ms), weak (peaks that were observable only at 300 ms), and very weak (observable weak peaks at 300 ms). The distance restraints corresponding to these classifications are 2 ± 1, 3.75 ± 1.25, 5.5 ± 1.5, and 7 ± 1.5 Å, respectively. The peaks from exchangeable protons were also classified manually with a NOESY experiment in H_2O with a mixing time of 200 ms. The peaks were classified as medium, weak, or very weak, with restraints of 3.75 ± 1.25 Å for medium, 5.5 ± 1.5 Å for weak, and 7 ± 2 Å for very weak peaks.

Distance restraints between two G-quadruplex blocks are listed in Table 2. These NOE restraints were manually determined to be interblock NOEs, because rational structural analyses show that such distances cannot occur within each block. NOEs observed for G3(H1)−G3(H1′) and G9(H1)−G9(H1′) cross-peaks were classified as ambiguous distance restraints, as these distances could satisfy both interblock and intrablock distance ranges between the corresponding proton pairs.

Table 1. RNA Oligonucleotides Used in This Study[a]

sequence	residue 1	residue 2	residue 3	residue 4	residue 5	residue 6	residue 7	residue 8	residue 9	residue 10	residue 11	residue 12
10-nucleotide human TERRA			G	G	G	U	U	A	G	G	G	U
10-nucleotide human TERRA (dT6)			G	G	G	**dT**	U	A	G	G	G	U
10-nucleotide human TERRA (dT7)			G	G	G	U	**dT**	A	G	G	G	U

[a]Residue numbering starts from position 3 to facilitate structural comparison with previously reported structures of 12-nucleotide human TERRA sequences.[37,44,46]

dx.doi.org/10.1021/bi301606u | *Biochemistry* 2013, 52, 2176−2183

Table 2. List of NOE Cross-Peaks Observed between the Two G-Quadruplex Blocks

block 1	block 2	NOE classification
G3(H1)	G3(H8)	medium
G3(H1′)	A8(H1′)	weak
G3(H8)	A8(H1′)	medium
G3(H8)	G9(H1′)	weak
G3(H8)	G9(H4′)	weak
G3(H8)	G9(H8)	weak
G4(H8)	G9(H8)	very weak
G9(H1)	G9(H8)	weak
A8(H2)	A8(H1′)	weak
A8(H2)	A8(H8)	weak

Table 3. List of NOE Cross-Peaks Observed between A8(H2) and the G-Tetrad Core Residues

peak 1	peak 2	NOE classification
A8(H2)	G3(H1′)	medium
A8(H2)	G4(H1′)	medium
A8(H2)	G3(H2′)	medium
A8(H2)	G4(H3′)	weak
A8(H2)	G3(H3′)	weak
A8(H2)	G3(H4′)	weak
A8(H2)	G4(H2′)	weak
A8(H2)	G4(H5′/5″)	weak
A8(H2)	G4(H4′)	medium
A8(H2)	G3(H8)	weak
A8(H2)	G4(H8)	weak
A8(H2)	G9(H8)	weak

Table 3 lists the NOEs observed between A8(H2) and the core residues (Figure 2b–d). Initial structure calculations were performed with these distance restraints omitted. The structure calculation revealed the preferred intrablock positioning of unrestrained A8(H2) protons pointing toward the G-tetrad core of their own blocks (Figure S3 of the Supporting Information). On the basis of this structure calculation result, these NOEs were determined to be intrablock NOEs and incorporated into the final structure calculation.

Dihedral Restraints. All the glycosidic bonds were restrained to 240 ± 40° corresponding to their *anti* conformation.

Hydrogen-Bond and Planarity Restraints. The G3·G9·G3·G9, G4·G10·G4·G10, and G5·G11·G5·G11 tetrads within each dimeric RNA G-quadruplex block were restrained with O6–N1 and N7–N2 distances, which were set to 2.95 ± 0.1 and 2.90 ± 0.1 Å, respectively. The force constant of hydrogen bonds was kept at 100 kcal mol^{-1} Å^{-2} throughout the computation. Restraints were applied to the bases of these G-tetrads to enforce their planarity.

Distance Geometry and Simulated Annealing. The calculation started with generation of four extended RNA strands. Initial distance geometry-simulated annealing calculation was performed by incorporating hydrogen-bond, dihedral angle, planarity, and NOE restraints. One hundred structures were generated and subjected to further refinement. As NMR spectra indicated the symmetry among the four strands, noncrystallographic symmetry restraints were also applied in the calculation with a force constant of 500 kcal mol^{-1} Å^{-2}.

Refinement. The 100 structures calculated from the simulated annealing step were further refined. The refinement process followed the distance-restrained molecular dynamics protocol. The molecule was initially heated from 300 to 1000 K in 5 ps. Equilibration for 1 ps was performed with the force constant for nonexchangeable and exchangeable NOE restraints being kept at 2 kcal mol^{-1} Å^{-2}. The force constants for both types of NOE restraints were then increased to 100 kcal mol^{-1} Å^{-2} in 26 ps. The system was then cooled to 300 K in 14 ps, followed by a 10 ps equilibration. Coordinates of the molecule were saved every 0.5 ps during the last 4.0 ps and then averaged. The structure produced was then subjected to minimization until the energy gradient was less than 0.1 kcal/mol. The energy terms used for the refinement were van der Waals, electrostatic, bond angle, bond length, dihedral angle, improper angle, NOE, noncrystallographic symmetry, and planarity energies. Dihedral (20 kcal mol^{-1} Å^{-2}) and planarity (1 kcal mol^{-1} Å^{-2}) restraints were maintained throughout the refinement process. The 10 lowest-energy structures were selected.

AMBER Molecular Dynamics Refinement. The 10 lowest-energy structures subsequently underwent molecular dynamics (MD) refinement in explicit solvent using the AMBER 10.0 program[60] with the reparametrized ff9bsc0 force field.[61] K^+ cations were added between G-tetrad layers (five K^+ cations in total). The system was first energy-minimized in vacuo with 1000 steps (500 steps of steepest descent followed by 500 steps of conjugate gradient) to remove any steric clashes. The system was then neutralized with 31 additional external K^+ cations (the system contained 36 K^+ cations in total) and solvated with ∼7800 water molecules (TIP3P) in a truncated octahedral box. Simulations were then performed according to the previously published protocol.[49] Hydrogen-bond and NOE distance constraints were applied throughout the simulation (1 ns) with a force constant of 100 kcal/mol.

Structures were taken from the MD simulations every picosecond for further analysis. The lowest-energy structures were selected. The water molecules and K^+ cations were stripped from the structures, and the structures were further minimized and deposited in the Protein Data Bank (PDB).

Accession Number. Coordinates of the 10 lowest-energy G-quadruplex structures of r(GGGUUAGGGU) have been deposited in the PDB with accession code 2M18 (the residue numbering starts from 1, which is different from that used in Table 1).

RESULTS AND DISCUSSION

Stacking of Two G-Quadruplex Blocks in K^+ Solution Characterized by NMR. Our previous data[37] showed that in K^+ solution the 10-nucleotide human TERRA r(GGGUUAGGGU) sequence formed a structure involving two dimeric parallel-stranded propeller-type G-quadruplex blocks (Figure 1), which stacked on each other at their 5′-ends. The COSY spectrum of r(GGGUUAGGGU) (Figure 2a) indicates the presence of both C2′-*endo* and C3′-*endo* sugar puckering conformations in the G-tetrad core. Site-specific ribose-to-deoxyribose substitutions (Table 1) were used to assist spectral assignments, as the upfield-shifted H2′ and H2″ protons of the substituted residue could be easily recognized (Figures S1 and S2 of the Supporting Information). Complete spectral assignments were achieved with the help of [^{13}C–^{1}H] JRHMBC, [^{13}C–^{1}H] HSQC, TOCSY, COSY, and NOESY experiments (Figure 2).

Compared to the spectra of one G-quadruplex block formed by the 12-nucleotide human TERRA r(UAGGGUUAGGGU)

dx.doi.org/10.1021/bi301606u | *Biochemistry* 2013, 52, 2176–2183

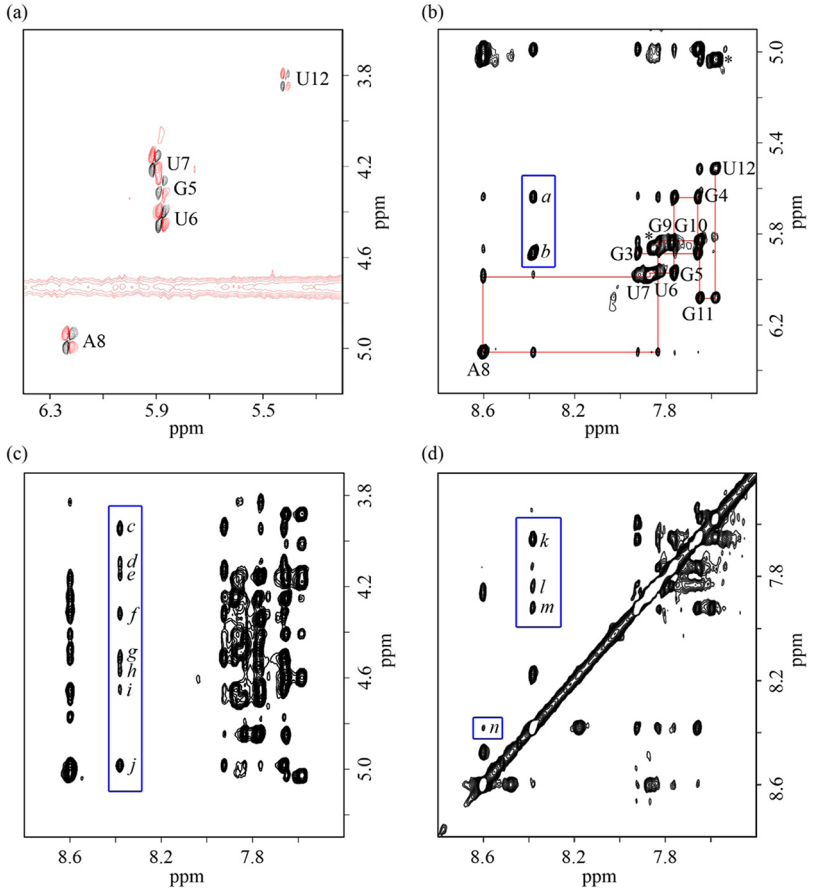


Figure 2. Two-dimensional NMR spectra of the 10-nucleotide human TERRA r(GGGUUAGGGU) sequence in K^+ solution. (a) COSY spectrum. Strong H1′−H2′ peaks indicate the C2′-*endo* sugar puckering conformation. Positive and negative levels are colored black and red, respectively. Errors on the previous assignments of G5 and U7[37] have been corrected. (b−d) NOESY spectrum (mixing time, 300 ms). NOE cross-peaks from A8(H2) to G3, G4, and G9, indicating loop rearrangements, are labeled as follows: *a*, A8(H2)−G4(H1′); *b*, A8(H2)−G3(H1′); *c*, A8(H2)−G4(H4′); *d*, A8(H2)−G4(H5′/5″); *e*, A8(H2)−G4(H2′); *f*, A8(H2)−G3(H4′); *g*, A8(H2)−G3(H3′); *h*, A8(H2)−G4(H5′/5″); *i*, A8(H2)−G4(H3′); *j*, A8(H2)−G3(H2′); *k*, A8(H2)−G4(H8); *l*, A8(H2)−G9(H8); *m*, A8(H2)−G3(H8); *n*, A8(H8)−A8(H2).

sequence,[37] the 10-nucleotide r(GGGUUAGGGU) sequence shows additional NOE cross-peaks for protons of the loops and the 5′-end G-tetrad at the stacking interface between the two blocks, indicating significant structural rearrangements of the loops upon stacking of the two G-quadruplex blocks. In particular, the observation of NOEs between the H2 proton of adenine A8 in the UUA loops with protons of G3, G4, and G9 (Figure 2b−d and Table 3) indicates that the adenine H2 protons of the UUA loops are pointing inward toward the G-tetrad core. This contrasts with the situation of one G-quadruplex block reported previously.[37] We also observed an NOE cross-peak between A8(H8) and A8(H2) (peak *n*, Figure 2d), indicating loop−loop interaction across the two stacked G-quadruplex blocks.

Solution Structure of a Stacked G-Quadruplex Formed by Human TERRA in Potassium Solution. The G-quadruplex structure of the 10-nucleotide human TERRA r(GGGUUAGGGU) sequence was calculated on the basis of NMR restraints (Table 4) using the XPLOR-NIH program.[57,58] For NOEs at the stacking interface between the two G-quadruplex blocks, the distance restraints were classified as intrablock, interblock, or ambiguous (see Materials and Methods).

The calculated structure is composed of two individual G-quadruplex blocks (Figure 3). Each block adopts a dimeric parallel-stranded propeller-type fold with four medium-sized grooves, two of which are occupied by a UUA double-chain-reversal loop. The adenines in the UUA loops are facing toward the G-tetrad core, consistent with the NOESY data (Figure 2b−d). The U6 uracils are also facing toward the core, while the U7 uracils are pointing outside. This arrangement results in a very compact structure. The U12 bases are stacked on the 3′-end of the G-tetrad core.

The two G-quadruplex blocks are stacked on their 5′-ends, consistent with the NOE cross-peaks observed between the blocks (Table 2) and the results of a solvent exchange experiment showing the protection of the imino protons at the stacking interface.[37] We also observed stacking of the adenines across the two opposite blocks (Figure 4a,b), consistent with the NOE cross-peak between A8(H8) and A8(H2) (peak *n*, Figure 2d). At the interface between the two G-quadruplex blocks, guanines are stacked with a partial overlap of their five-membered rings (Figure 4b). The positioning of A8 in the same plane of the 5′-end G-tetrad forming an A·(G·G·G·G)·A hexad with A8(H2) pointing toward the G-tetrad core is also held by hydrogen bonding between A8(N1) and G3(OH2′) (Figure 4c).

The base arrangement and base overlap observed at the stacking interface here are different from those reported previously for some related structures containing

dx.doi.org/10.1021/bi301606u | *Biochemistry* 2013, 52, 2176−2183

Table 4. NMR Restraints and Structure Statistics

(A) NMR Restraints		
distance restraints	nonexchangeable	exchangeable
intraresidue distance restraints	340	12
sequential (i, i + 1) distance restraints	248	24
long-range (i, $\geq i$ + 1) distance restraints	128	64
interblock distance restraints	32	8
ambiguous distance restraints	0	8
other restraints		
hydrogen-bond restraints	96	
dihedral angle restraints	40	
(B) Statistics for 10 Structures following Refinement		
NOE violations		
number (>0.2 Å)	0.7 ± 0.949	
maximal violation (Å)	0.186 ± 0.021	
rmsd of violations (Å)	0.024 ± 0.002	
deviations from the ideal covalent geometry		
bond lengths (Å)	0.01 ± 0.00	
bond angles (deg)	0.90 ± 0.06	
impropers (deg)	0.52 ± 0.03	
pairwise all heavy atom rmsd values (Å)		
all heavy atoms except U6, U7, A8, and U12	0.92 ± 0.19	
all heavy atoms	1.09 ± 0.2	

A·(G·G·G·G)·A hexads at the stacking interface,[62−65] where A(H2) protons are pointing outward from the G-tetrad core and the guanine bases show either full overlaps between five-membered rings or overlaps between five- and six-membered rings (Figure S4 of the Supporting Information).

Comparison with Previously Reported Structures of TERRA G-Quadruplexes. In comparison with the previously reported structure of a single G-quadruplex block formed by the 12-nucleotide human TERRA r(UAGGGUUAGGGU) sequence in K^+ solution[37] (PDB entry 2KBP), the structure determined in this work for the 10-nucleotide r(GGGUUAGGGU) sequence shows significant loop rearrangements (Figure 5a,b). The overall structure of each block of the latter is more compact than the former structure. In the latter structure, the adenines in the loops fling inward, positioned coplanar with the 5′-end G-tetrad (Figure 5c), thereby increasing the stacking contacts and possibly contributing to a greater stabilization of the stacking conformation; the U6 bases also fling inward facing the core, in contrast to the outward-pointing configuration of U6 in the former structure (Figure 5c). The U12 bases at the 3′-ends of both structures are stacked on the 3′-end G-tetrad, with only a small shift observed (Figure 5d).

The conformation of loops in the current structure is also different from that found in the crystal structure (PDB entry 3IBK)[44] of the 12-nucleotide human TERRA r(^{Br}UAGGGUUAGGGU) sequence (Figure 6a,b). It appears that a slight flattening of the UUA loops of the crystal structure would result in a conformation similar to that observed here (Figure 6c). At the 3′-end, the positioning of the U12 residues in the current structure is very similar to that of one U residue in the crystal structure (Figure 6d).

Intriguingly, the current NMR structure closely resembles to the crystal structure of the 12-nucleotide human TERRA r(UAGGGUUAGGGU) sequence bound to an acridine ligand[46] (PDB entry 3MIJ). The similarity can be observed both for the core and for the loop backbone (Figure 7a,b). In both cases, the adenines from the UUA loops are positioned coplanar with the 5′-end G-tetrad plane (Figure 7c). Hydrogen bonds formed between N1 of the A8 adenines and OH2′ of the G3 guanines of the 5′-end G-tetrad might contribute to holding the adenines in their position (Figure 4c and Figure S5 of the Supporting Information). The positioning of the U7 (but not U6) bases in the UUA loops is similar for the NMR and crystal structures (Figure 7d).

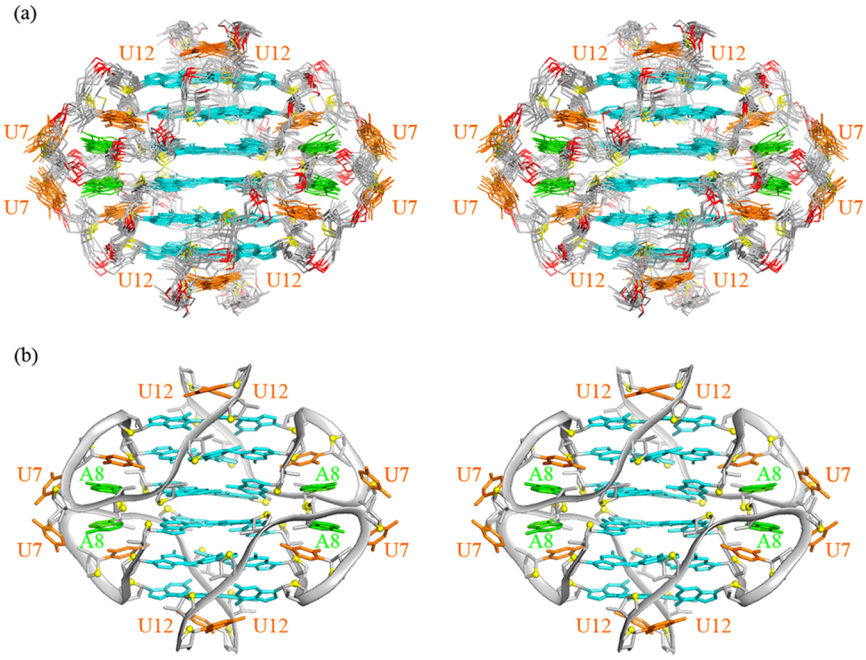


Figure 3. Stereoviews of the G-quadruplex structures formed by the 10-nucleotide human TERRA r(GGGUUAGGGU) sequence in K^+ solution. (a) Superposition of the 10 lowest-energy structures. (b) Ribbon view of the lowest-energy structure. Bases of guanines are colored cyan, adenines green, uracils orange, backbones gray, O4′ atoms yellow, and P atoms red.

dx.doi.org/10.1021/bi301606u | *Biochemistry* 2013, 52, 2176−2183

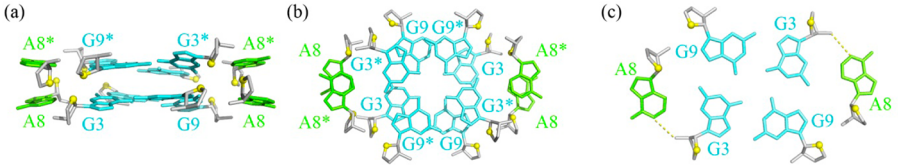


Figure 4. Stacking interface between two subunits of the G-quadruplex formed by the 10-nucleotide human TERRA r(GGGUUAGGGU) sequence in K^+ solution. (a) Side view and (b) top view of the stacking interface between the two blocks. (c) Top view of the stacking interface of one block. The proposed hydrogen bonds between A8(N1) and G3(OH2′) are shown as yellow dashed lines.

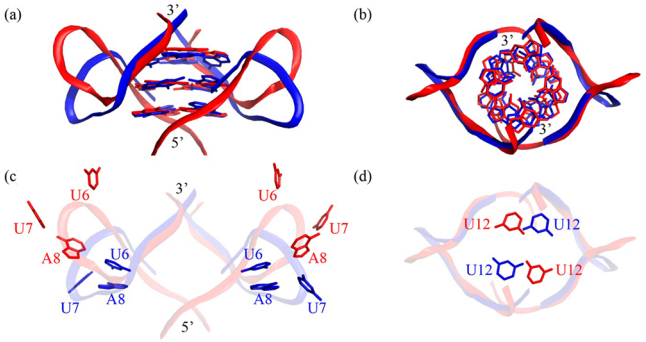


Figure 5. Comparison between single G-quadruplex blocks of the structure formed by the 10-nucleotide human TERRA r(GGGUUAGGGU) sequence (determined in this work, blue) and the structure formed by the 12-nucleotide human TERRA r(UAGGGUUAGGGU) sequence (PDB entry 2KBP, red), both in K^+ solution. Superposition of the two structures was based on the best fit of the G-tetrad core. (a) Side view and (b) top view of the two superimposed structures with the bases of the G-tetrad core shown. (c) Side view and (d) top view of the two superimposed structures with the bases of the UUA loop (c) and terminal U bases (d) shown.

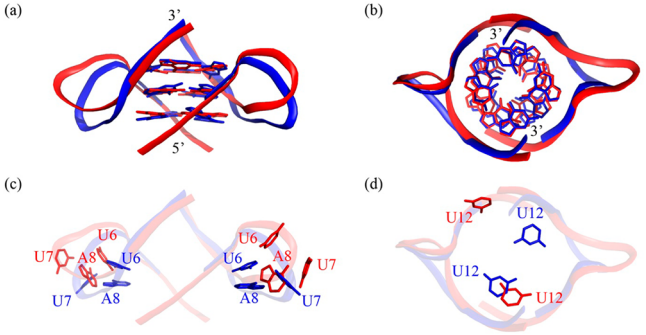


Figure 6. Comparison between single G-quadruplex blocks of the structure formed by the 10-nucleotide human TERRA r(GGGUUAGGGU) sequence in K^+ solution (determined in this work, blue) and the crystal structure of the 12-nucleotide human TERRA sequence r(^{Br}UAGGGUUAGGGU) (PDB entry 3IBK, red). Superposition of the two structures was based on the best fit of the G-tetrad core. (a) Side view and (b) top view of the two superimposed structures with the bases of the G-tetrad core shown. (c) Side view and (d) top view of the two superimposed structures with the bases of the UUA loop (c) and terminal U bases (d) shown.

CONCLUSION

We have determined the first high-resolution structure of a stacked G-quadruplex formed by the human TERRA sequence in potassium solution. The adenine in the UUA loops was observed to be positioned in the plane of the 5′-end G-tetrad forming an A·(G·G·G·G)·A hexad, stabilized by hydrogen

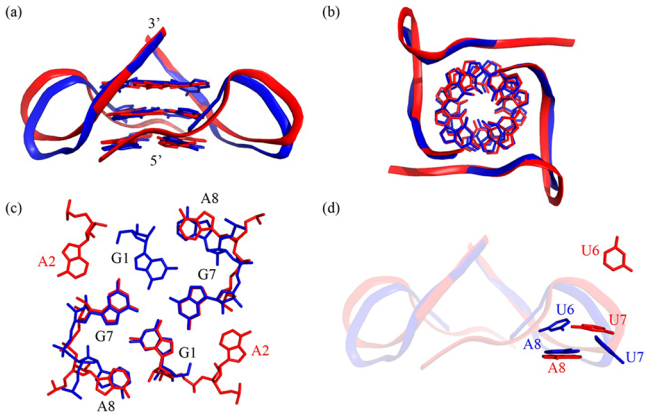


Figure 7. Comparison between single G-quadruplex blocks of the structure formed by the 10-nucleotide human TERRA r(GGGUUAGGGU) sequence in K^+ solution (determined in this work, blue) and the crystal structure of the 12-nucleotide human TERRA r(UAGGGUUAGGGU) sequence bound to acridine (PDB entry 3MIJ, red). Superposition of the two structures was based on the best fit of the G-tetrad core. (a) Side view and (b) top view of the two superimposed structures with the bases of the G-tetrad core shown. (c) Top view of the superimposed stacking interfaces of the two structures and (d) side view of the two superimposed structures with the bases of a UUA loop shown.

bonding between A8(N1) and G3(OH2′). Adenine−adenine stacking across the two G-quadruplex blocks was also observed.

ASSOCIATED CONTENT

Supporting Information

Figures S1−S5 showing one-dimensional and NOESY NMR spectra of site-specific ribose-to-deoxyribose substituted sequences, computed G-quadruplex structures with restraints in Table 3 omitted, comparison of hexad interfaces of published structures, and stacking interfaces in the crystal structure of the 12-nucleotide human TERRA r(UAGGGUUAGGGU) sequence bound to an acridine ligand (PDB entry 3MIJ). This material is available free of charge via the Internet at http://pubs.acs.org.

AUTHOR INFORMATION

Corresponding Author

*E-mail: phantuan@ntu.edu.sg. Telephone: +65 6514 1915. Fax: +65 6795 7981.

Funding

This research was supported by Singapore Ministry of Education and Nanyang Technological University grants to A.T.P.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

We thank Dr. Brahim Heddi for helpful discussions and his assistance with structure calculation.

■ ABBREVIATIONS

TERRA, telomeric RNA; NMR, nuclear magnetic resonance; rmsd, root-mean-square deviation; PAGE, polyacrylamide gel electrophoresis; PDB, Protein Data Bank; JR, jump-and-return; HMBC, heteronuclear multiple-bond correlation; HSQC, heteronuclear single-quantum correlation; TOCSY, total correlation spectroscopy; COSY, correlation spectroscopy; NOESY, nuclear Overhauser effect spectroscopy; NOE, nuclear Overhauser effect.

■ REFERENCES

(1) Zakian, V. A. (1995) Telomeres: Beginning to understand the end. *Science 270*, 1601−1607.
(2) Olovnikov, A. M. (1973) A theory of marginotomy. The incomplete copying of template margin in enzymic synthesis of polynucleotides and biological significance of the phenomenon. *J. Theor. Biol. 41*, 181−190.
(3) Bodnar, A. G., Ouellette, M., Frolkis, M., Holt, S. E., Chiu, C. P., Morin, G. B., Harley, C. B., Shay, J. W., Lichtsteiner, S., and Wright, W. E. (1998) Extension of life-span by introduction of telomerase into normal human cells. *Science 279*, 349−352.
(4) Moyzis, R. K., Buckingham, J. M., Cram, L. S., Dani, M., Deaven, L. L., Jones, M. D., Meyne, J., Ratliff, R. L., and Wu, J. R. (1988) A highly conserved repetitive DNA-sequence, $(TTAGGG)_n$, present at the telomeres of human-chromosomes. *Proc. Natl. Acad. Sci. U.S.A. 85*, 6622−6626.
(5) Makarov, V. L., Hirose, Y., and Langmore, J. P. (1997) Long G tails at both ends of human chromosomes suggest a C strand degradation mechanism for telomere shortening. *Cell 88*, 657−666.
(6) Azzalin, C. M., Reichenbach, P., Khoriauli, L., Giulotto, E., and Lingner, J. (2007) Telomeric repeat containing RNA and RNA surveillance factors at mammalian chromosome ends. *Science 318*, 798−801.
(7) Schoeftner, S., and Blasco, M. A. (2008) Developmentally regulated transcription of mammalian telomeres by DNA-dependent RNA polymerase II. *Nat. Cell Biol. 10*, 228−236.
(8) Bah, A., Wischnewski, H., Shchepachev, V., and Azzalin, C. M. (2012) The telomeric transcriptome of *Schizosaccharomyces pombe*. *Nucleic Acids Res. 40*, 2995−3005.
(9) Greenwood, J., and Cooper, J. P. (2012) Non-coding telomeric and subtelomeric transcripts are differentially regulated by telomeric and heterochromatin assembly factors in fission yeast. *Nucleic Acids Res. 40*, 2956−2963.
(10) Feuerhahn, S., Iglesias, N., Panza, A., Porro, A., and Lingner, J. (2010) TERRA biogenesis, turnover and implications for function. *FEBS Lett. 584*, 3812−3818.
(11) Porro, A., Feuerhahn, S., Reichenbach, P., and Lingner, J. (2010) Molecular dissection of telomeric repeat-containing RNA biogenesis unveils the presence of distinct and multiple regulatory pathways. *Mol. Cell. Biol. 30*, 4808−4817.
(12) Horard, B., and Gilson, E. (2008) Telomeric RNA enters the game. *Nat. Cell Biol. 10*, 113−115.
(13) Luke, B., Panza, A., Redon, S., Iglesias, N., Li, Z. J., and Lingner, J. (2008) The Rat1p 5′ to 3′ exonuclease degrades telomeric repeat-containing RNA and promotes telomere elongation in *Saccharomyces cerevisiae*. *Mol. Cell 32*, 465−477.
(14) Caslini, C., Connelly, J. A., Serna, A., Broccoli, D., and Hess, J. L. (2009) MLL associates with telomeres and regulates telomeric repeat-containing RNA transcription. *Mol. Cell. Biol. 29*, 4519−4526.
(15) Deng, Z., Norseen, J., Wiedmer, A., Riethman, H., and Lieberman, P. M. (2009) TERRA RNA binding to TRF2 facilitates heterochromatin formation and ORC recruitment at telomeres. *Mol. Cell 35*, 403−413.
(16) Luke, B., and Lingner, J. (2009) TERRA: Telomeric repeat-containing RNA. *EMBO J. 28*, 2503−2510.
(17) Schoeftner, S., and Blasco, M. A. (2009) A 'higher order' of telomere regulation: Telomere heterochromatin and telomeric RNAs. *EMBO J. 28*, 2323−2336.
(18) Deng, Z., Campbell, A. E., and Lieberman, P. M. (2010) TERRA, CpG methylation and telomere heterochromatin: Lessons from ICF syndrome cells. *Cell Cycle 9*, 69−74.
(19) Redon, S., Reichenbach, P., and Lingner, J. (2010) The non-coding RNA TERRA is a natural ligand and direct inhibitor of human telomerase. *Nucleic Acids Res. 38*, 5797−5806.
(20) Davis, J. T. (2004) G-quartets 40 years later: From 5′-GMP to molecular biology and supramolecular chemistry. *Angew. Chem., Int. Ed. 43*, 668−698.
(21) 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.
(22) 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.
(23) Wang, Y., and Patel, D. J. (1993) Solution structure of the human telomeric repeat $d[AG_3(T_2AG_3)_3]$ G-tetraplex. *Structure 1*, 263−282.
(24) Parkinson, G. N., Lee, M. P., and Neidle, S. (2002) Crystal structure of parallel quadruplexes from human telomeric DNA. *Nature 417*, 876−880.
(25) Vorlickova, M., Chladkova, J., Kejnovska, I., Fialova, M., and Kypr, J. (2005) Guanine tetraplex topology of human telomere DNA is governed by the number of (TTAGGG) repeats. *Nucleic Acids Res. 33*, 5851−5860.
(26) Xu, Y., Noguchi, Y., and Sugiyama, H. (2006) The new models of the human telomere $d[AGGG(TTAGGG)_3]$ in K^+ solution. *Bioorg. Med. Chem. 14*, 5584−5591.
(27) Ambrus, A., Chen, D., Dai, J., Bialis, T., Jones, R. A., and Yang, D. (2006) Human telomeric sequence forms a hybrid-type intramolecular G-quadruplex structure with mixed parallel/antiparallel strands in potassium solution. *Nucleic Acids Res. 34*, 2723−2735.
(28) Luu, K. N., Phan, A. T., Kuryavyi, V., Lacroix, L., and Patel, D. J. (2006) Structure of the human telomere in K^+ solution: An intramolecular (3 + 1) G-quadruplex scaffold. *J. Am. Chem. Soc. 128*, 9963−9970.
(29) Phan, A. T., Luu, K. N., and Patel, D. J. (2006) Different loop arrangements of intramolecular human telomeric (3 + 1) G-quadruplexes in K^+ solution. *Nucleic Acids Res. 34*, 5715−5719.
(30) Yu, H., Miyoshi, D., and Sugimoto, N. (2006) Characterization of structure and stability of long telomeric DNA G-quadruplexes. *J. Am. Chem. Soc. 128*, 15461−15468.
(31) Lim, K. W., Amrane, S., Bouaziz, S., Xu, W., Mu, Y., Patel, D. J., Luu, K. N., and Phan, A. T. (2009) Structure of the human telomere in K^+ solution: A stable basket-type G-quadruplex with only two G-tetrad layers. *J. Am. Chem. Soc. 131*, 4301−4309.
(32) Xu, Y., Ishizuka, T., Kurabayashi, K., and Komiyama, M. (2009) Consecutive formation of G-quadruplexes in human telomeric-overhang DNA: A protective capping structure for telomere ends. *Angew. Chem., Int. Ed. 48*, 7833−7836.
(33) Phan, A. T. (2010) Human telomeric G-quadruplex: structures of DNA and RNA sequences. *FEBS J. 277*, 1107−1117.
(34) Heddi, B., and Phan, A. T. (2011) Structure of human telomeric DNA in crowded solution. *J. Am. Chem. Soc. 133*, 9824−9833.
(35) Yu, H., Gu, X., Nakano, S. I., Miyoshi, D., and Sugimoto, N. (2012) Beads-on-a-string structure of long telomeric DNAs under molecular crowding conditions. *J. Am. Chem. Soc. 134*, 20060−20069.
(36) 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.
(37) 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.
(38) Randall, A., and Griffith, J. D. (2009) Structure of long telomeric RNA transcripts: The G-rich RNA forms a compact repeating structure containing G-quartets. *J. Biol. Chem. 284*, 13980−13986.
(39) Arora, A., and Maiti, S. (2009) Differential biophysical behavior of human telomeric RNA and DNA quadruplex. *J. Phys. Chem. B 113*, 10515−10520.
(40) Joachimi, A., Benz, A., and Hartig, J. S. (2009) A comparison of DNA and RNA quadruplex structures and stabilities. *Bioorg. Med. Chem. 17*, 6811−6815.
(41) Xu, Y., Ishizuka, T., Kimura, T., and Komiyama, M. (2010) A U-tetrad stabilizes human telomeric RNA G-quadruplex structure. *J. Am. Chem. Soc. 132*, 7231−7233.
(42) Xu, Y., Suzuki, Y., Ito, K., and Komiyama, M. (2010) Telomeric repeat-containing RNA structure in living cells. *Proc. Natl. Acad. Sci. U.S.A. 107*, 14579.
(43) Zhang, D. H., Fujimoto, T., Saxena, S., Yu, H. Q., Miyoshi, D., and Sugimoto, N. (2010) Monomorphic RNA G-quadruplex and polymorphic DNA G-quadruplex structures responding to cellular environmental factors. *Biochemistry 49*, 4554−4563.
(44) Collie, G. W., Haider, S. M., Neidle, S., and Parkinson, G. N. (2010) A crystallographic and modelling study of a human telomeric RNA (TERRA) quadruplex. *Nucleic Acids Res. 38*, 5569−5580.
(45) 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.
(46) Collie, G. W., Sparapani, S., Parkinson, G. N., and Neidle, S. (2011) Structural basis of telomeric RNA quadruplex-acridine ligand recognition. *J. Am. Chem. Soc. 133*, 2721−2728.
(47) Martadinata, H., Heddi, B., Lim, K. W., and Phan, A. T. (2011) Structure of long human telomeric RNA (TERRA): G-quadruplexes formed by four and eight UUAGGG repeats are stable building blocks. *Biochemistry 50*, 6455−6461.
(48) Sket, P., and Plavec, J. (2010) Tetramolecular DNA quadruplexes in solution: Insights into structural diversity and cation movement. *J. Am. Chem. Soc. 132*, 12724−12732.
(49) Do, N. Q., Lim, K. W., Teo, M. H., Heddi, B., and Phan, A. T. (2011) Stacking of G-quadruplexes: NMR structure of a G-rich oligonucleotide with potential anti-HIV and anticancer activity. *Nucleic Acids Res. 39*, 9448−9457.
(50) 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.
(51) Do, N. Q., and Phan, A. T. (2012) Monomer-dimer equilibrium for the 5′-5′ stacking of propeller-type parallel-stranded G-quadruplexes: NMR structural study. *Chem.—Eur. J. 18*, 14752−14759.
(52) Kuryavyi, V., Cahoon, L. A., Seifert, H. S., and Patel, D. J. (2012) RecA-binding pilE G4 sequence essential for Pilin antigenic variation forms monomeric and 5′ end-stacked dimeric parallel G-quadruplexes. *Structure 20*, 2090−2102.
(53) Plateau, P., and Guéron, M. (1982) Exchangeable proton NMR without base-line distortion, using new strong-pulse sequences. *J. Am. Chem. Soc. 104*, 7310−7311.
(54) Phan, A. T., Guéron, M., and Leroy, J. L. (2001) Investigation of unusual DNA motifs. *Methods Enzymol. 338*, 341−371.
(55) Phan, A. T., Kuryavyi, V., Darnell, J. C., Serganov, A., Majumdar, A., Ilin, S., Raslin, T., Polonskaia, A., Chen, C., Clain, D., Darnell, R. B., and Patel, D. J. (2011) Structure-function studies of FMRP RGG peptide recognition of an RNA duplex-quadruplex junction. *Nat. Struct. Mol. Biol. 18*, 796−804.
(56) 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.
(57) Brunger, A. T. (1992) *X-PLOR: A system for X-ray crystallography and NMR*, Yale University Press, New Haven, CT.
(58) 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.
(59) DeLano, W. L. (2002) *The PyMOL User's Manual*, DeLano Scientific, Palo Alto, CA.
(60) Case, D. A., Cheatham, T. E., III, Darden, T., Gohlke, H., Luo, R., Merz, K. M., Jr., Onufriev, A., Simmerling, C., Wang, B., and Woods, R. J. (2005) The Amber biomolecular simulation programs. *J. Comput. Chem. 26*, 1668−1688.
(61) Zgarbova, M., Otyepka, M., Sponer, J., Mladek, A., Banas, P., Cheatham, T. E., III, and Jurecka, P. (2011) Refinement of the Cornell et al. nucleic acids force field based on reference quantum chemical calculations of glycosidic torsion profiles. *J. Chem. Theory Comput. 7*, 2886−2902.
(62) 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.
(63) Liu, H., Matsugami, A., Katahira, M., and Uesugi, S. (2002) A dimeric RNA quadruplex architecture comprised of two G:G(:A):G:G-(:A) hexads, G:G:G:G tetrads and UUUU loops. *J. Mol. Biol. 322*, 955−970.
(64) Parkinson, G. N., Cuenca, F., and Neidle, S. (2008) Topology conservation and loop flexibility in quadruplex-drug recognition: Crystal structures of inter- and intramolecular telomeric DNA quadruplex-drug complexes. *J. Mol. Biol. 381*, 1145−1156.
(65) Mashima, T., Matsugami, A., Nishikawa, F., Nishikawa, S., and Katahira, M. (2009) Unique quadruplex structure and interaction of an RNA aptamer against bovine prion protein. *Nucleic Acids Res. 37*, 6249−6258.