7190 *Biochemistry* **2010,** *49,* 7190–7201
DOI: 10.1021/bi101093a

BIOCHEMISTRY
including biophysical chemistry & molecular biology
Article

Sole and Stable RNA Duplexes of G-Rich Sequences Located in the 5′-Untranslated Region of Protooncogenes[†]

Sarika Saxena,[‡] Daisuke Miyoshi,*[,‡,§] and Naoki Sugimoto*[,‡,§]

[‡]*Frontier Institute for Biomolecular Engineering Research* (*FIBER*) *and* [§]*Faculty of Frontiers of Innovative Research in Science and Technology* (*FIRST*), *7-1-20 Minatojima-Minamimachi, Chuo-ku, Kobe 650-0047, Japan*

Received July 9, 2010

ABSTRACT: Guanine- (G-) rich nucleic acid sequences can form four-stranded structures called G-quadruplexes. It is widely held that the formation of a G-quadruplex in RNA is more feasible than in DNA because of the lack of a complementary strand in mRNA. Here, we analyzed sequences of 5′-untranslated regions of protooncogenes and surprisingly found that these regions showed an enrichment of not only guanine (G) but also cytosine (C) nucleotides. Since neighboring cytosine- (C-) rich regions can affect the formation and stability of a G-quadruplex structure, we further investigated the properties of DNA and RNA structures of G-rich and GC-rich regions. We selected typical GC-rich RNA sequences from protooncogenes and corresponding DNA sequences and investigated their structures. It was found that the GC-rich RNA sequences formed stable A-form duplexes as their major structure independent of the surrounding conditions, including the presence of different cations (Na^+, K^+, or Li^+) or molecular crowding with 40 wt % poly(ethylene glycol) with an average molecular mass of 200 Da although there are a few exceptions in which only a combination of K^+ and molecular crowding induced a G-quadruplex structure of an extremely G-rich RNA sequence. In contrast, structural polymorphisms involving duplexes, G-quadruplexes, and i-motifs were observed for GC-rich DNA sequences depending on the surrounding factors. These results demonstrate the considerable structural and functional differences in GC-rich sequences of the genome (DNA) and transcriptosome (mRNA) with respect to the nucleic acid backbone. Moreover, it was suggested that structural study for a G-rich RNA sequence should be carried out under cell-mimicking condition where K^+ and crowding cosolutes exist.

DNA is known to form duplexes with Watson−Crick base pairs (*1*). However, polymorphic structures of DNA have attracted great attention over the past 2 decades (*2*, *3*). In particular, it is well-known that G-rich and C-rich nucleic acid sequences can form four-stranded structures, the G-quadruplex, and i-motif, respectively (*4*, *5*). Sequences with high potential to form DNA G-quadruplexes have been found in many regions of the genome such as telomeres (*6*), promoter regions of protooncogenes (*7*), growth factors (*8*), immunoglobulin switch regions (*3*), insulin regulatory sequences (*9*), and the region responsible for fragile X syndrome (*10*). Recent publications have demonstrated the presence of G-quadruplex-forming sequences throughout the human genome and their enrichment adjacent to transcription start sites, thus emphasizing the biological significance of this structure (*11*−*20*). Structural studies using X-ray (*21*−*23*) crystallography and NMR (*24*−*27*) have further revealed that naturally occurring G-rich DNA sequences form G-quadruplexes with various combinations of strand directions. Moreover, it has been demonstrated that since genomic DNA, except for its telomeres, is double stranded, G-quadruplex formation by G-rich sequences requires the canonical Watson−Crick base pairs to open to form Hoogsteen base pairs (*28*, *29*). These structural transitions among secondary structures modulating gene function are influenced by pH, temperature, cations, and molecular crowding, which are essential as chemical stimuli inside cells (*30*−*33*).

Not only for DNA, but also for RNA, bioinformatic analysis has revealed numerous G-quadruplex-forming sequences in transcribed mRNAs, especially in their 5′-untranslated regions (UTR)[1] (*11*). For example, RNA G-quadruplex formation has been proposed in insulin-like growth factor-II (IGF-II) (*34*), fragile X mental retardation protein (FMR1) (*35*), fibroblast growth factor 2 (FGF-2) (*36*), matrix metalloproteinase (MT3MMP) (*37*), human neuroblastoma RAS viral oncogene (v-ras) (*38*), and zinc-finger protein (Zic-1) (*39*). Moreover, RNA G-quadruplex structures are thought to be involved in the regulation of translation both *in vitro* and *in vivo* (*40*−*42*) and in various other biological functions such as structural roles (*43*), intron splicing (*44*), and protein binding (*37*, *45*).

Recently, the thermodynamic stability of the RNA G-quadruplex at its natural position within the 5′-UTR of NRAS has

[†]This work was supported in part by Grants-in-Aid for Scientific Research, the "Core Research" project (2009−2014), and the "Academic Frontier" project (2004−2009) from the Ministry of Education, Culture, Sports, Science, and Technology, Japan, the Hirao Taro Foundation of the Konan University Association for Academic Research, and the Long-Range Research Initiative Project of Japan Chemical Industry Association.

*To whom correspondence should be addressed. D.M.: phone, +81-78-303-1426; fax, +81-78-303-1495; e-mail, miyoshi@center.konan-u.ac.jp, N.S.: phone, +81-78-303-1416; fax, +81-78-303-1495; e-mail, sugimoto@konan-u.ac.jp.

[1]Abbreviations: UTR, untranslated region; CD, circular dichroism; PAGE, polyacrylamide gel electrophoresis; T_m, melting temperature; PEG, poly(ethylene glycol); PAGE, polyacrylamide gel electrophoresis; UV, ultravisible.

 Published on Web 07/20/2010

Table 1: Sequences of DNA and RNA Used in This Study[a]

abbreviation	sequence	gene	length
	RNA Sequences		
R1	5′-rCCGGGGCUCCGGGCCCUCCCUGCCGGCGGCC-3′	bcl-2	31-mer
R2	5′-rCCCACCGCCCGCCCCCCUUGGGGCGCAGGGCAUGGUGUGAAAGG-3′	THRA	44-mer
R2-G	5′-rGGGGCGCAGGGCAUGGUGUGAAAGG-3′	THRA	25-mer
R3	5′-rCCGCUGGGCGGCGGGCGCCGGGGGCCGGAGGGGC-3′	CRK-II	34-mer
	DNA Sequences		
D1	5′-dCCGGGGCTCCGGGCCCTCCCTGCCGGCGGCC-3′	bcl-2	31-mer
D2	5′-dCCCACCGCCCGCCCCCCTTGGGGCGCAGGGCATGGUGTGAAAGG-3′	THRA	44-mer

[a]Sequences are located in the 5′-untranslated region (UTR) of THRA, bcl-2, and CRK-II genes. DNA and RNA sequences differ with respect to the backbone and the presence of base (U ribonucleotide in RNA and T deoxynucleotide in DNA). R2-G is the truncated G-rich fragment of R2.

been reported as an important factor to repress translation (*46*). Systematic studies of RNA G-quadruplexes have shown recently (*47*−*51*) that naturally occurring sequences around G-rich regions are heterogeneous in their arrangement such that they can affect the structure and stability of G-quadruplexes (*52*). Single-stranded mRNA occurs naturally without a complementary strand, indicating that G-quadruplex formation has no competition from duplex structures. However, the G-rich sequences in mRNA can form a variety of structures with their neighboring regions, e.g., hairpin loop structures with C- (cytosine-) rich sequences, which result in strong competition against G-quadruplex formation. Therefore, the neighboring sequences of the G-rich sequences should be taken into account when studying the structural functions of mRNA.

Here, we focused on G-rich mRNA sequences located in the 5′-UTR of human protooncogenes to test for structures other than G-quadruplexes. Interestingly, sequence analysis revealed that regulatory regions of human protooncogenes were enriched in G and C, generating GC-rich regions in which C-rich regions are located either adjacent to or mixed with a G-rich regions. This sequence analysis indicated that we have to expand the investigation of DNA and RNA structures from G-rich regions to GC-rich regions. We investigated the structural features of two typical GC-rich RNA sequences found in mRNAs as well as their corresponding DNA sequences as references. We found that these GC-rich RNA sequences formed a stable A-form duplex as a major structure in all of the conditions tested, even more G-rich and less C-rich RNA sequence also forms the A-form duplex, although only the combination of K^+ and molecular crowding can induce G-quadruplex structure for a more G-rich and less C-rich RNA sequence. In contrast, corresponding GC-rich DNA sequences showed structural polymorphisms including duplexes, G-quadruplexes, and i-motifs, depending on the surrounding conditions. The results obtained here demonstrated monomorphism of RNA structures and polymorphism of DNA structures of biologically important GC-rich sequences. The role of the 2′-OH group and the biological significance of the monomorphic RNA and polymorphic DNA structures are discussed.

MATERIALS AND METHODS

Sample Preparation. High-performance liquid chromatography (HPLC) purification grade unlabeled and fluorescent-labeled RNA sequences and DNA sequences (Table 1) were purchased from Hokkaido System Science (Sapporo, Japan). Single-strand concentrations of the RNA and DNA sequences were determined by measuring absorbance at 260 nm at a high temperature using a spectrophotometer (1700; Shimadzu, Kyoto, Japan) connected to a thermoprogammer. Single-strand extinction coefficients were calculated from mononucleotide and dinucleotide data using the nearest neighbor approximation (*53*, *54*).

Circular Dichroism (CD) Measurements. CD experiments utilizing a spectropolarimeter (J-820; Jasco, Hachioji, Japan) were measured at 4 °C in a 0.1 cm path length cuvette for 1, 2, or 4 μM total strand concentration of RNA and DNA. Samples were prepared in 30 mM MES buffer (pH 7.0) or Tris−acetate buffer (pH 4.5) containing 100 mM NaCl, KCl, or LiCl and 0.5 mM Na_2EDTA, K_2EDTA, or Li_2EDTA with and without 40 wt % poly(ethylene glycol) with an average molecular mass of 200 Da (PEG 200). The CD spectrum was the average of at least three scans from 200 to 350 nm. The temperature of the cell holder was regulated by a temperature controller (PTC-348, Jasco), and the cuvette-holding chamber was flushed with a constant stream of dry N_2 gas to avoid condensation of water on the cuvette exterior. Before measurement, the sample was heated to 95 °C, gently cooled at a rate of 0.5 °C min^{-1}, and incubated at 4 °C overnight.

UV Melting Analysis. UV absorbance was measured with the spectrophotometer equipped with the temperature controller. Melting curves of RNA structures were obtained by measuring the UV absorbance at 260 or 295 nm. Samples were prepared in 30 mM MES buffer (pH 7.0) containing 100 mM NaCl, KCl, or LiCl and 0.5 mM Na_2EDTA, K_2EDTA, or Li_2EDTA at 0 wt % PEG 200 or 40 wt % PEG 200. Before measurement, the sample was heated to 95 °C, gently cooled at a rate of 0.5 °C min^{-1}, and incubated at 4 °C for several hours. Measurement was performed using a 1 cm path length cuvette, and for higher concentrations, a 0.1 cm path length cuvette was also used. The melting temperature (T_m) values for RNA structures were obtained from the UV melting curves as described previously (*53*, *54*). The heating rate was 0.5 °C min^{-1}.

Nondenaturing Gel Electrophoresis. Nondenaturing gel electrophoresis was carried out with 15% polyacrylamide gels (19:1 acrylamide/bisacrylamide). RNA and DNA samples of 6 μM were mixed with ice-cold loading buffer. Samples were prepared in 30 mM MES buffer (pH 7.0) containing 100 mM NaCl, KCl, or LiCl and 0.5 mM Na_2EDTA, K_2EDTA, or Li_2EDTA with 40 wt % PEG 200. A 4 μL aliquot of the mixed solution was loaded and run at 5 V cm^{-1} at 4 °C. The gel was stained using SYBR Gold (Invitrogen, Ltd., U.K.) and imaged using an FLA-5100 (Fujifilm Co., Ltd., Tokyo, Japan).

RNase T1 Footprinting. The 5′-fluorescence-labeled sequence of R2 was used for RNase T1 (Applied Biosystems/Ambion, Texas) footprinting experiments. Before RNase T1 digestion, samples

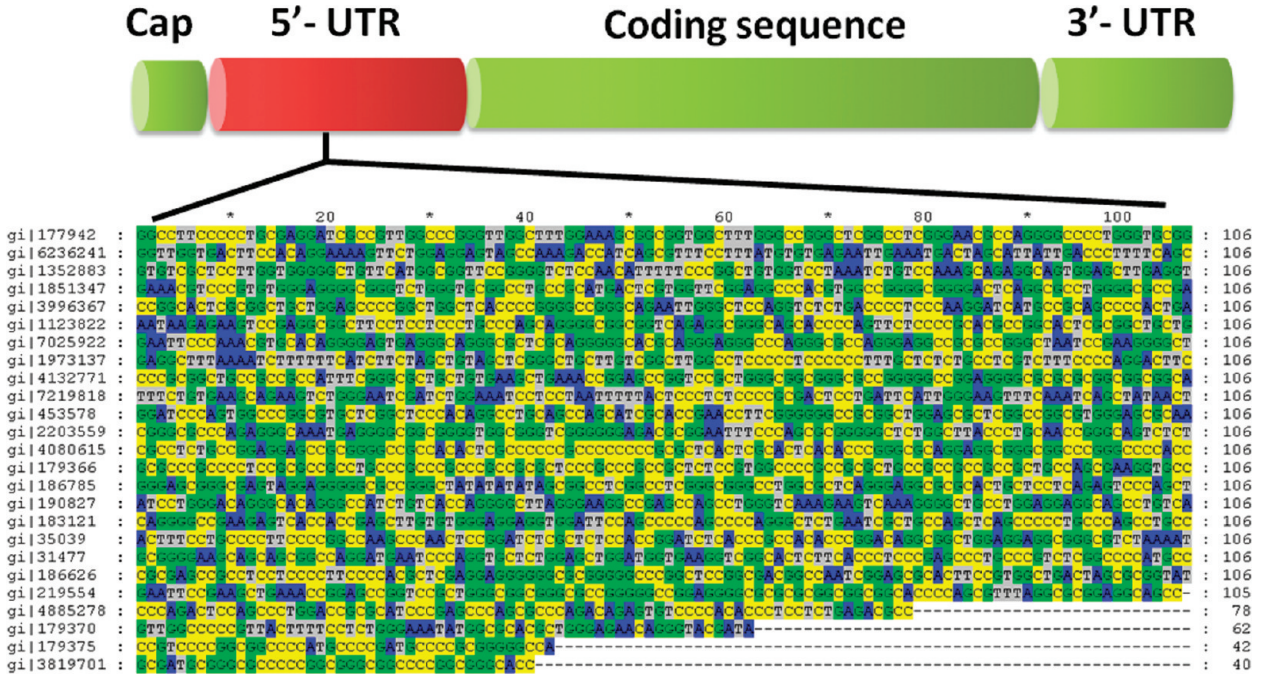

FIGURE 1: Multiple sequence alignment of the segment of 5′-UTR of human protooncogenes. Sequences selected for study were taken from the bcl2 gene (> gi|179366|) and THRA gene (> gi|40806158|) and CRK-II (gi|41327711). Gene ids (gi) are shown along with gene names. Guanine, cytosine, adenine, and thymine are indicated by green, yellow, blue, and gray, respectively.

were prepared in 30 mM MES buffer (pH 7.0) in the presence of 100 mM NaCl, KCl, or LiCl at 0 wt % PEG 200 or 40 wt % PEG 200. Before cleavage, samples were heated at 95 °C for 5 min and then allowed to cool slowly at a rate of 0.5 °C min^{-1}. After annealing, RNA samples were incubated for 1 h at 4 °C. RNase T1 footprinting was performed by incubating the RNA samples with 0.20 unit of RNase T1 at 37 °C for 15 min. To assign each cleaved base, alkaline hydrolysis was performed by incubation at 95 °C for 5 min in alkaline hydrolysis buffer (Applied Biosystems/Ambion) containing 50 mM sodium carbonate (pH 9.2) and 1 mM Na_2EDTA. In both native and alkaline hydrolysis conditions, reactions were stopped by adding gel loading buffer (Applied Biosystems/Ambion). Samples were heated at 95 °C for 2 min before denaturing polyacrylamide gel electrophoresis. An equal volume from each sample was loaded on a 15% polyacrylamide gel containing 7 M urea. The gel was run at 23 V cm^{-1} at room temperature for 4 h. The gel image was visualized using an FLA-5100 (Fujifilm Co., Ltd., Tokyo, Japan).

RESULTS AND DISCUSSION

Design of GC-Rich Sequences. We carried out sequence analysis for 25 human protooncogene RNA sequences (Figure 1). Gene names were verified from the HUGO Gene Nomenclature Committee 2006. These mRNA sequences were taken from NCBI build (36.3). The 5′-UTR was extracted by considering the transcript of each gene separately before the coding sequence. To trace out the abundance of each nucleotide near the G-rich region, we aligned selected mRNA sequences of human protooncogenes together using the Genedoc multiple sequence alignment editor and shading utility (version 2.7.000) (*55*). Figure 1 shows each gene transcript from base numbers 1 to 106. There are a total of 2447 nucleotides in the sequences involving 849 G nucleotides, 846 C nucleotides, 367 adenines (A) nucleotides, and 385 thymines (T) nucleotides. The frequency of the occurrence of G is approximately equal to that of C and more than twice that of A and T. The results clearly showed that the regulatory regions of the human protooncogenes were enriched with both G and C. Close examination of the sequence arrangement of the regulatory region of the protooncogenes revealed that GC-rich sequences are more abundant than only G-rich sequences. Thus, GC-rich sequences are the typical conserved sequences in the 5′-UTR of human protooncogenes. Moreover, the distribution of the G- and C-rich regions can be categorized two ways, with the C-rich region located either adjacent to or mixed with the G-rich region. For example, in the gene region from bases 1 to 40 in gene id (gi) 3819701, G- and C-rich regions are mixed with each other. On the other hand, in the gene region from 7 to 43 in gi 186626, the C-rich region is located next to the G-rich region. Such C-rich regions could form intra- or intermolecular duplexes with the G-rich regions leading to competition against G-quadruplex formation. In addition, it has been reported that flanking sequences affect the thermodynamic stability of the G-quadruplex (*56*−*58*). These results and the sequence analysis in this study raise the question of whether G-rich RNA sequences existing in UTR regions form stable RNA G-quadruplexes but not duplexes in the absence of a complementary strand (*9*). Therefore, it is necessary to investigate G-rich sequences capable of forming G-quadruplex structures from a wider region of the genome, especially when a C-rich region is located in close proximity to a G-rich region.

In order to investigate the possible structures of GC-rich RNA sequences, we selected two typical GC-rich sequences from the 5′-UTR of human protooncogenes (Table 1) and their corresponding GC-rich DNA sequences. One sequence contained mixed and evenly distributed G and C (named R1 for RNA and D1 for DNA) and the other sequence contained a long C-rich region followed by a G-rich region (named R2 for RNA and D2 for DNA). Since the length of the C-rich region may have a significant effect on the structure and on the stability of the G-rich region, we selected a short fragment of R2 composed only of a G-rich region (named R2-G). Moreover, R3 have more number

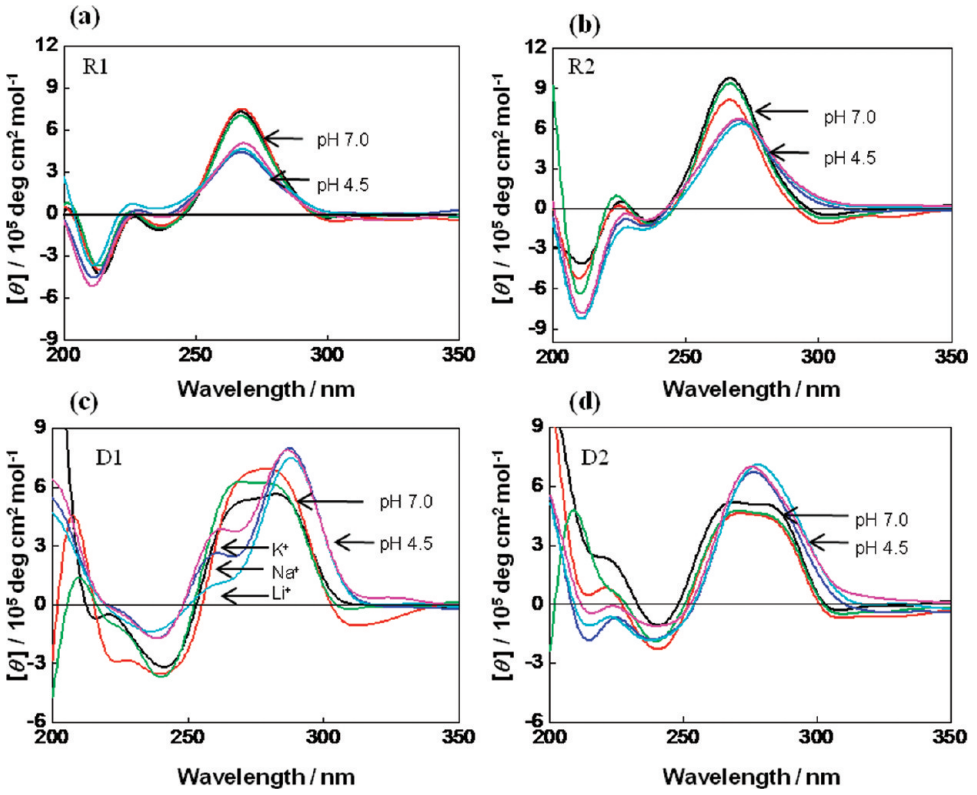

FIGURE 2: CD spectra of 2 μM R1 (a), 2 μM R2 (b), 4 μM D1 (c), and 4 μM D2 (d). Measurements were carried out at 4 °C in 30 mM MES buffer (pH 7.0) containing 100 mM NaCl and 0.5 mM Na_2EDTA (black); 100 mM KCl and 0.5 mM K_2EDTA (red); or 100 mM LiCl and 0.5 mM Li_2EDTA (green); and in 30 mM Tris−acetate buffer (pH 4.5) containing 100 mM NaCl and 0.5 mM Na_2EDTA (blue); 100 mM KCl and 0.5 mM K_2EDTA (pink); or 100 mM LiCl and 0.5 mM Li_2EDTA (cyan).

of Gs and less number of Cs as compared with R1 and R2 forming a typical G-quadruplex forming sequence. Thus, R3 can be useful to confirm whether the results with the other sequences are general in the typical putative G-quadruplex forming sequences.

Structural Analysis of GC-Rich RNA and DNA Sequences. First, we studied the structure of R1 using CD spectroscopy. Figure 2a shows CD spectra of 2 μM R1 in MES buffer (pH 7.0) and Tris−acetate buffer (pH 4.5) in the presence of 100 mM Na^+, K^+, or Li^+ at 4 °C. CD spectra of R1 had a positive peak at 260 nm and negative peaks at 210 and 240 nm under all conditions. It has been reported that the CD spectrum of the A-form DNA duplex has a positive peak around 260 nm and negative peaks at 210 and 240 nm (*59*, *60*), while the RNA duplex displays a positive peak at 260 nm and negative peak near 210 nm (*61*). This indicates that R1 forms the A-form duplex. Of note, there was no significant change in the CD spectra and intensity with different monovalent cation species, indicating that the structure of R1 is mostly independent of the type of cation. The isodichroic points near 258 and 286 nm in the CD spectra of R1 suggest that the structure of R1 is in a two-state transition between some structures as discussed later. CD spectra of 4 μM R2 showed similar signatures suggesting the formation of A-form duplexes under all conditions (Figure 2b).

Next, we studied the effect of molecular crowding with 40 wt % PEG 200 on the conformation of R1 and R2 (Supporting Information Figure S1a). CD spectra of 4 μM R1 and R2 in the presence of 100 mM Na^+, K^+, or Li^+ with 40 wt % PEG 200 showed CD signatures characteristic of an A-form duplex indicating duplex formation as the major structure under molecular crowding conditions, which mimic the intracellular environment. The strong propensity of R1 and R2 to form A-form duplexes may be due to the presence of the 2′-OH group which arranges the ribose into a 3′-endo chair conformation and generates an *anti*-glycosidic conformation characteristic of the A-form conformation (*62*). Moreover, the independence of the CD spectra of R1 and R2 with different cations also supports the formation of an A-form duplex, because the stability of the duplex and G-quadruplex are independent and dependent on the cation species, respectively (*57*, *58*, *63*).

In contrast to R1 and R2, CD spectra of D1 and D2 were observed to be highly polymorphic. The CD spectra of 4 μM D1 at pH 7.0 had a broad positive peak covering the region from 260 to 290 nm along with negative peaks at 210 and 240 nm (Figure 2c). It has been reported that B-form DNA displays characteristic positive peaks between 250 and 280 nm and two negative peaks at 210 and 240 nm, while the A-form duplex is mostly defined by a positive peak at 260 nm and negative peaks at 210 and 240 nm (*59*, *60*). This suggests that D1 folded into a mixture of A- and B-form duplexes. The mixed duplex structure of D1 is consistent with a previous report on the self-complementary G-rich sequence d(GGGGCCCC) (*64*). It was shown that the CD spectrum of d(GGGGCCCC) indicated an A-form duplex with stacking of half of the bases, whereas the other half stack in a B-form-like fashion. The CD spectra of D1 at pH 4.5 had a positive peak at 285 nm with a shoulder at 260 nm and negative peaks at 240 and 265 nm. The positive peak at 285 nm and negative peak at 265 nm are typical for the CD spectrum of an i-motif or an antiparallel G-quadruplex (*65*). In addition, the shoulder at 260 nm increased in the order $Li^+ < Na^+ < K^+$, which is in good agreement with the order of the stabilization effect of cations on G-quadruplexes (*38*). Thus, it is reasonable to assign the shoulder at 260 nm to a parallel G-quadruplex, which has a positive peak

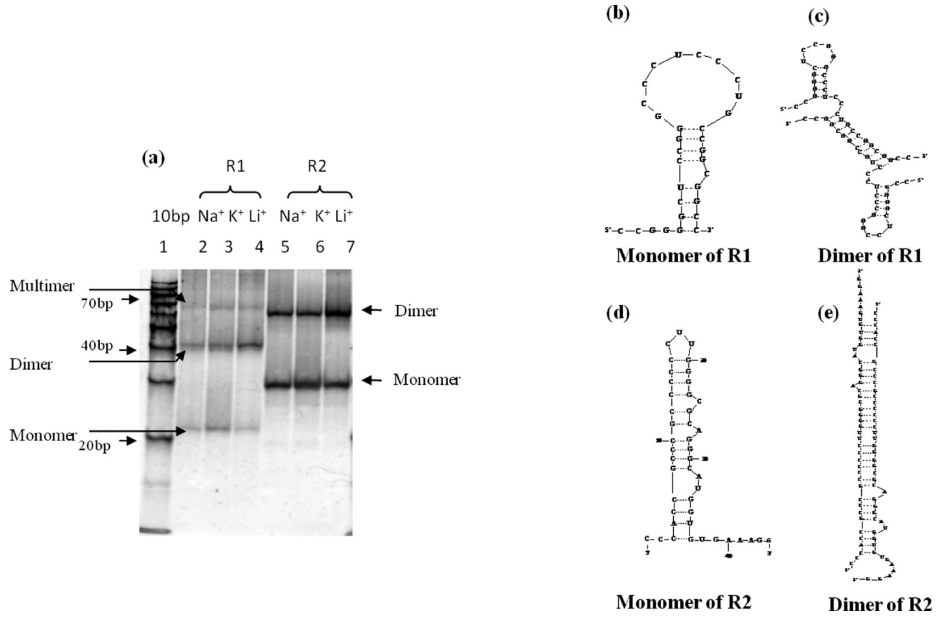

FIGURE 3: (a) Nondenaturing 15% polyacrylamide gel of 6 μM R1 (lanes 2−4) and 6 μM R2 (lanes 5−7). Electrophoresis was performed at 4 °C with 0 wt % PEG 200 in 30 mM MES buffer (pH 7.0) containing 100 mM NaCl and 0.5 mM Na_2EDTA (lanes 2 and 5); 100 mM KCl and 0.5 mM K_2EDTA (lanes 3 and 6); or 100 mM LiCl and 0.5 mM Li_2EDTA (lanes 4 and 7). (b−e) Schematic representation using Mfold of predicted monomers (b and d) and dimers (c and e) of R1 and R2, respectively.

around 260 nm (*66*). Similar to D1, CD spectra of D2 at pH 7.0 also showed broad positive peaks covering the region from 260 to 286 nm along with negative peaks at 210 and 240 nm, indicating a mixture of A- and B-form duplexes under all conditions (Figure 2d). At pH 4.5, the presence of a positive peak at 276 nm indicated predominance of the B-form duplex. Therefore, it may be possible that the structures of D1 and D2 are in equilibrium between A- and B-form duplexes, i-motifs, and G-quadruplexes. The fact that there was no isodichroic point in the CD spectra of D1 and D2 further supported the coexistence of several structures.

Next, we examined the effect of cosolutes on the conformation of D1 and D2 in the presence of 100 mM Na^+, K^+, or Li^+ with 40 wt % PEG 200 (Supporting Information Figure S1b). CD spectra of D1 and D2 showed polymorphic structures characterized by positive peaks at 286 and 260 nm and a negative peak at 240 nm under molecular crowding conditions. Similar to their CD spectra under the dilute conditions, the CD spectra of D1 and D2 may be due to antiparallel, parallel G-quadruplexes, and i-motif formations. This polymorphic character in D1 and D2 can be attributed to the deoxyribonucleotide backbone that can adopt both *anti* and *syn* conformations depending on the solution conditions, which is in contrast to the all *anti* conformation of RNA.

Molecularity of R1 and R2. In order to confirm the structures of R1 and R2, we performed nondenaturing polyacrylamide gel electrophoresis (PAGE) in the presence of 100 mM Na^+, K^+, or Li^+ at pH 7.0 with 0 wt % PEG 200 (Figure 3a). In the case of R1, one major band around the 40-bp DNA marker and two minor ones around the 20- and 70-bp markers were observed. The lower band around 20 bp corresponds to an intramolecular structure of R1 (31-mer) since we confirm that RNA migrates slower than DNA with the same length (data not shown). The major and middle bands around 40 bp correspond to intermolecular structures. In addition, the top band may correspond to a multimer. There was no difference in any of the band intensities of R1 with different monovalent ions, suggesting that the R1 structures are based on duplexes but not G-quadruplexes. These results indicate that the major structure of R1 is a duplex as observed in CD spectra, although the duplex can be a monomer, dimer, or multimer. PAGE results of R2 (44-mer) showed two bands. The lower bands around 30 bp correspond to a monomer while the top bands near 70 bp correspond to a dimer. Similar to R1, these band intensities did not vary depending on the coexisting cations, further supporting the hypothesis that the structures were duplexes. We observed almost identical band patterns of R1 and R2 with 40 wt % PEG 200 under molecular crowding conditions (Supporting Information Figure S2). In order to better understand the monomer and dimer structures of R1 and R2, secondary structures were predicted using Mfold (*67*, *68*) (Figure 3b−e). The monomer of R1 showed a short complementary stem with a varying number of bases in loops along with short dangling ends at the 5′-terminus (Figure 3b), while the dimer of R1 showed a hairpin loop structure at both termini (Figure 3c). Thus, it is possible that the hairpin loop structure is opened and can then associate with another strand, leading to the formation of the multimeric structure as observed in native PAGE (Figure 3a). On the other hand, the monomer of R2 has a short complementary stem followed by dangling ends at both termini (Figure 3d). These dangling ends can provide sites to open the monomeric structure of R2 to coexist with the dimer form as observed in PAGE. In addition, the dimer form of R2 does not involve a hairpin loop at the terminus (Figure 3e). This may inhibit the multimer formation that was observed in the case of R1.

UV Melting Studies of R1 and R2. We next examined the thermal denaturation of R1 and R2. The denaturation curves of 4 μM R1 were traced by UV absorbance at 260 and 295 nm in the presence of various cations with 0 wt % PEG 200 (Figure 4a,b). Similar melting curves were observed in the presence of Na^+, K^+, and Li^+. It is well-known that the thermostability of the duplex is independent of coexisting cations, while that of a G-quadruplex is dependent on them (*69*−*73*). Therefore, the UV melting curves of R1 show duplex formation. Furthermore, it is known that the dissociation of a G-quadruplex leads to a hypochromic melting transition which was not observed at 295 nm for R1. In addition, almost identical melting curves were observed at 40 wt % PEG

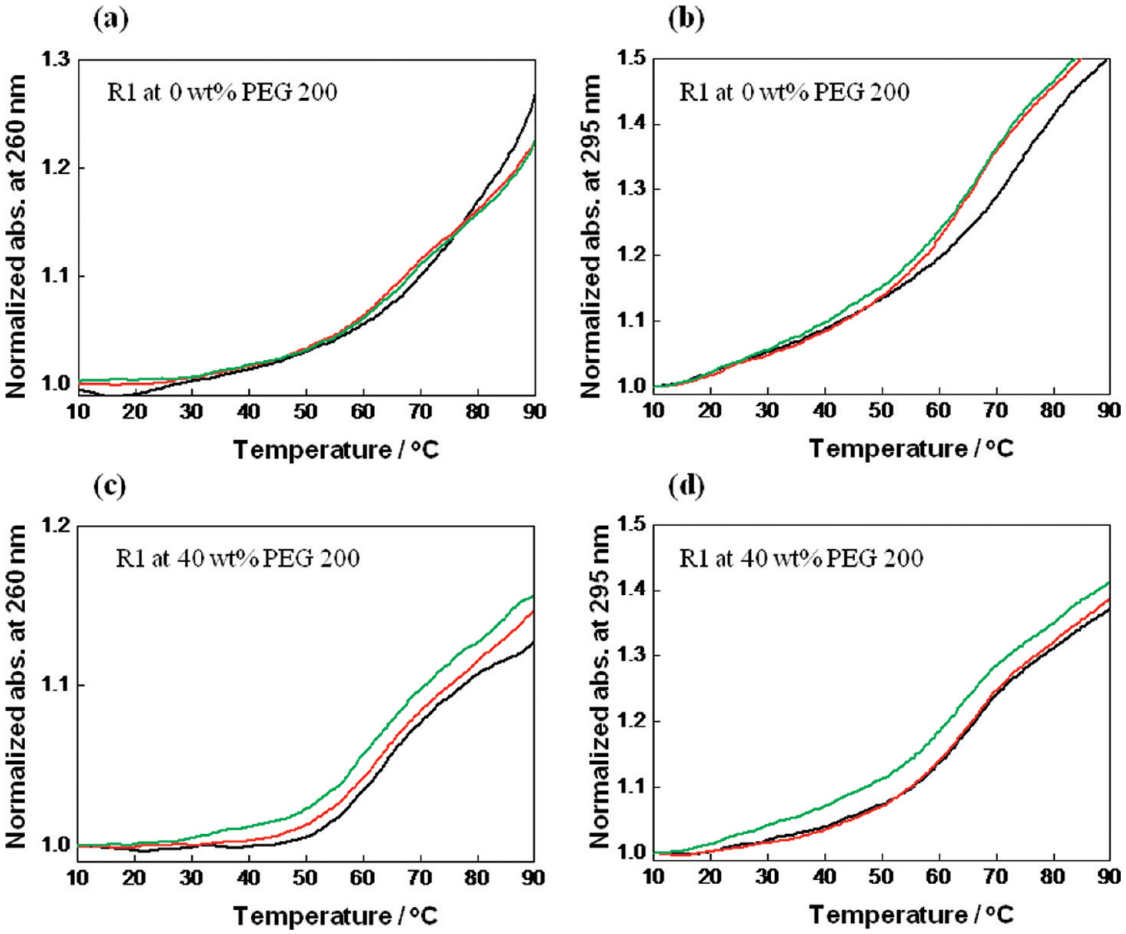

FIGURE 4: Normalized UV melting curves of 4 μM R1 with 0 wt % (a and b) or 40 wt % (c and d) PEG 200. Measurements were carried out in 30 mM MES buffer (pH 7.0) containing 100 mM NaCl and 0.5 mM Na_2EDTA (black); 100 mM KCl and 0.5 mM K_2EDTA (red); or 100 mM LiCl and 0.5 mM Li_2EDTA (green). Melting was assessed by measuring the UV absorbance at 260 nm (a and c) or 295 nm (b and d).

200 (Figure 4c,d), except that the melting temperatures of R1 structures decreased by the addition of PEG 200. This is consistent with a previous report showing that molecular crowding conditions destabilized the duplex (*74*). These results as well as the CD and PAGE results lead us to conclude that R1 forms A-form duplexes although they can be intramolecular and intermolecular structures. Similar to R1, R2 showed clear hyperchromicity in the thermal denaturation curves both at 260 and at 295 nm, respectively (Supporting Information Figure S3) indicating the duplex formation.

Effect of C-Rich Regions on the Structure and Stability of R2. Next, we examined the G-rich-only fragment of R2 (named R2-G) and recorded CD spectra in the presence of Na^+, K^+, or Li^+ with 0 wt % PEG 200 (Figure 5a) or 40 wt % PEG 200 (Figure 5b). The CD spectra of R2-G showed a positive band around 260 nm and a negative band near 240 nm, with a positive band at 210 nm with 0 wt % PEG 200 and 40 wt % PEG 200. These features are typical for a parallel-oriented G-quadruplex structure, revealing that R2-G forms a parallel G-quadruplex (*66*). In addition, the shoulder at 290 nm in the presence of K^+ is characteristic of an antiparallel conformation, indicating the formation of a hybrid parallel/antiparallel G-quadruplex structure in the presence of K^+ (*71*, *73*). We further studied the melting behaviors of R2-G in the presence of 100 mM Na^+, K^+, and Li^+ with 0 wt % PEG 200 (Figure 6a) or 40 wt % PEG 200 (Figure 6b) at 295 nm. The UV melting curve at 295 nm showed a hypochromic melting transition with T_m values in the presence of 100 mM Na^+ or 100 mM K^+ of 38.1 and 52.6 °C, respectively, whereas no hypochromic melting transition was observed in the presence of Li^+. This behavior (i.e., dependence on coexisting cations) is characteristic of RNA G-quadruplex formation. Komiyama et al. reported T_m values of 41.0 and 51.0 °C for 12 and 24 nucleotide telomeric sequences that formed parallel RNA G-quadruplexes in 100 mM Na^+ (*75*). Subramanian et al. also reported that an RNA G-quadruplex motif (5′-GGGAGGG-GCGGGUCUGGG-3′) located in the 5′-UTR of the NRAS protooncogene folded into a stable intramolecular RNA G-quadruplex structure with a T_m of 43.0 °C in the absence of cations and 63.0 °C in the presence of 1 mM K^+, which is indicative of a very stable G-quadruplex (*38*). In addition, an extra-stable RNA G-quadruplex has been reported in the 5′-UTR of MT3 matrix metalloproteinase mRNA, for which the addition of only 1 mM K^+ resulted in an increase in melting temperature of 30.0 °C (*37*). R2-G contains only two planes for intramolecular G-quadruplex formation, and the resultant stability based on their T_m values in accordance with previous reports is reasonable. With 40 wt % PEG 200, hypochromic melting transitions were observed with T_m values of 37.0, 54.1, and 10.7 °C in the presence of Na^+, K^+, and Li^+, respectively (Figure 6b). The increase in T_m values of R2-G under molecular crowding conditions suggests that crowding conditions favorably stabilized the structure compared to dilute conditions as shown in the previous report in which molecular crowding stabilizes G-quadruplex (*72*). Moreover, it was recently reported that the RNA G-quadruplex was

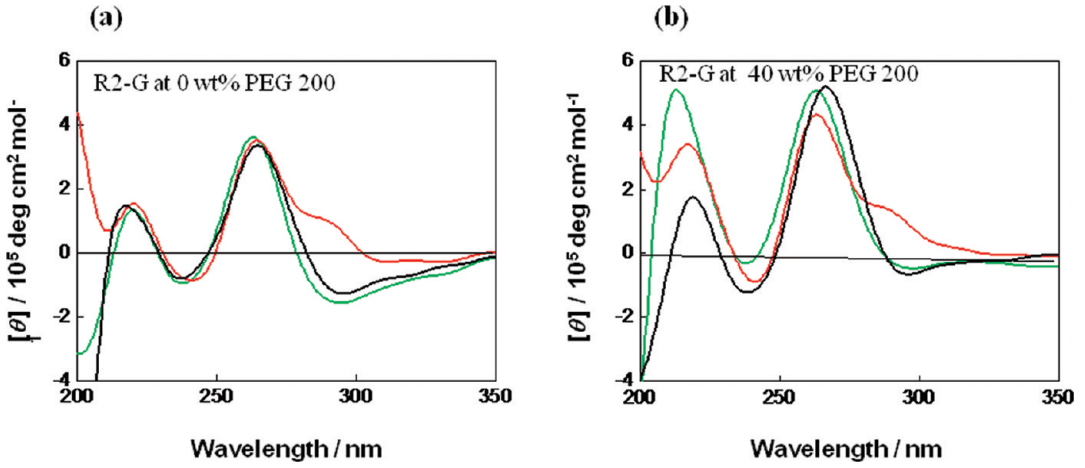

FIGURE 5: CD spectra of 2 μM R2-G with 0 wt % (a) or 40 wt % (b) PEG 200. Measurements with R2-G were carried out at 4 °C in 30 mM MES buffer (pH 7.0) containing 100 mM NaCl and 0.5 mM Na_2EDTA (black); 100 mM KCl and 0.5 mM K_2EDTA (red); or 100 mM LiCl and 0.5 mM Li_2EDTA (green).

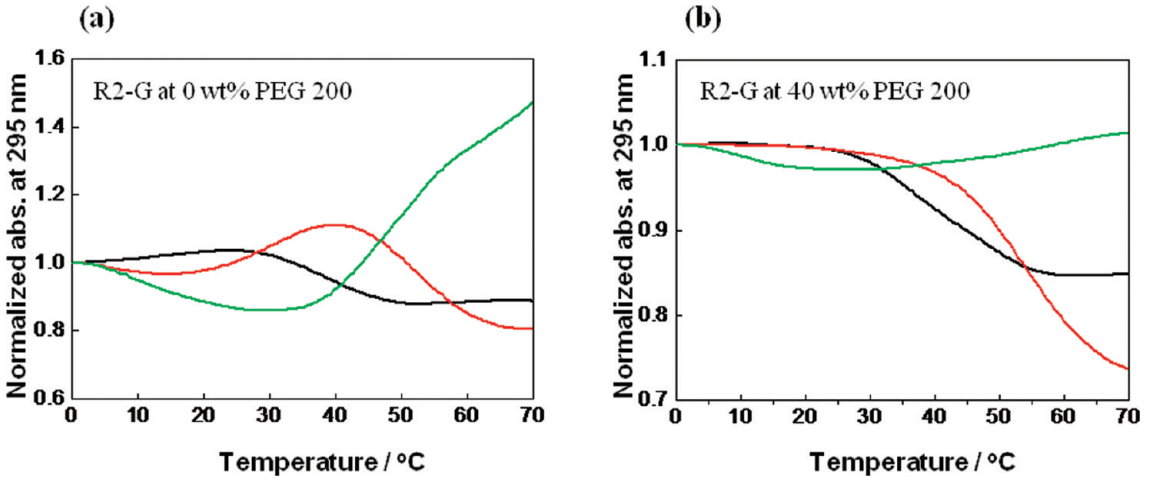

FIGURE 6: Normalized UV melting curves of 4 μM R2-G at 0 wt % (a) or 40 wt % (b) PEG 200. Melting was carried out in 30 mM MES buffer (pH 7.0) containing 100 mM NaCl and 0.5 mM Na_2EDTA (black); 100 mM KCl and 0.5 mM K_2EDTA (red); and 100 mM LiCl and Li_2EDTA (green). Melting was assessed by measuring the UV absorbance at 295 nm.

stabilized by molecular crowding but not altered its whole structure (*72*). Therefore, it is reasonable to conclude that R2-G forms an intramolecular G-quadruplex structure under both dilute and molecular crowding conditions. Formation of the G-quadruplex structure by R2-G and duplex structure by R2 confirmed that nucleic acid sequences present in close vicinity around a G-rich region critically affect the structure and stability of the G-quadruplex.

In order to confirm this point, we further studied the structure and stability of R3 which has more number of G and less number of C. Figure 7a shows CD spectra of 4 μM R3 in MES buffer (pH 7.0) in the presence of 100 mM Na^+, K^+, or Li^+ at 0 wt % PEG 200. As shown in R1 and R2, CD spectra of R3 had a positive peak at 260 nm and negative peaks at 210 and 240 nm. However, it should be noted that CD intensity at 210 nm decreased in the order $Li^+ > Na^+ > K^+$, indicating that the structure of R3 depends on the cation species. In addition, the parallel G-quadruplex has positive and negative peaks at 260 and 240 nm, respectively (*72*). Thus, it is not possible to determine whether R3 forms a duplex or G-quadruplex in the conditions. Thus, we carried out UV melting experiments for R3 in the presence of cations at 0 wt % PEG 200 at 260 nm (Figure 7b) and at 295 nm (Figure 7c). Although R3 showed hyperchromic transitions at 260 nm in the presence of cations, a hypochromic transition at 295 nm was observed only in the presence of K^+. Since a hypochromic transition at 295 nm is the characteristic for nucleic acid structures with Hoogsteen base pairs, i.e., G-quadruplex (*76*), the hypochromic transition observed around 90 °C indicates that R3 forms G-quadruplex. This hypothesis is further supported by the CD spectra and UV melting curves of R3 under the molecular crowding conditions where G-quadruplex and duplex were stabilized and destabilized, respectively (*72*, *74*). Although CD spectra of R3 at 40 wt % PEG 200 (Supporting Information Figure S4a) have almost identical shape with these at 0 wt % PEG 200, the decrement in the CD intensity at 210 nm becomes more significant in the presence of K^+, indicating the formation of G-quadruplex induced by the combination of PEG 200 and K^+. This is consistent with our recent paper showing that molecular crowding with PEG 200 and K^+ stabilizes G-quadruplex significantly (*72*). Supporting Information Figure S4b,c shows UV melting curves of R3 at 260 and 295 nm at 40 wt % PEG 200. The UV melting curves of R3 at 40 wt % PEG 200 in the presence of K^+ showed that the melting temperature of R3 in the condition is higher than 90 °C. Moreover, the melting curve in the presence of Na^+ also showed hypochromicity due to the stabilization of the G-quadruplex by molecular crowding. Thus, these results suggest

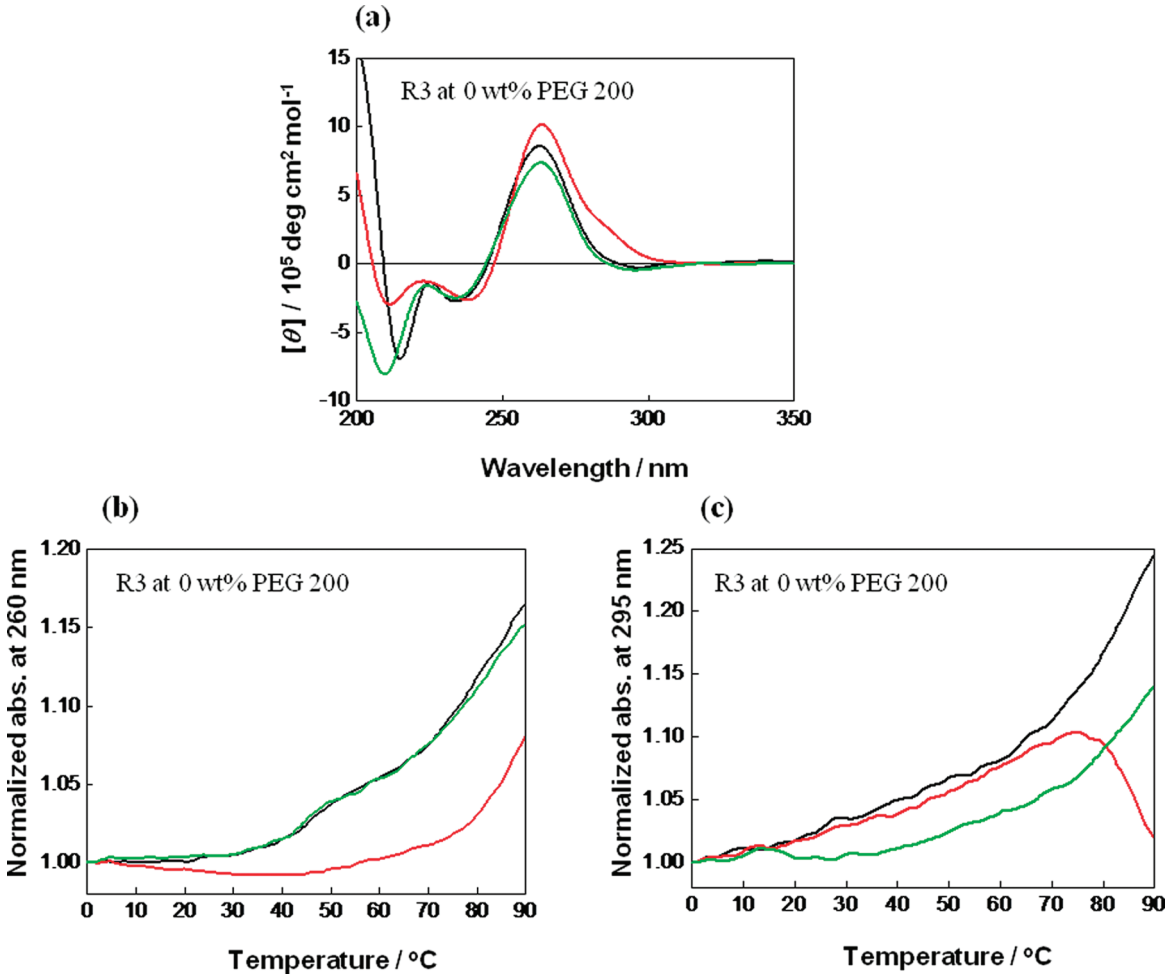

FIGURE 7: CD spectra of 4 μM R3 at 0 wt % PEG 200 (a). Measurements were carried out at 4 °C in 30 mM MES buffer (pH 7.0) containing 100 mM NaCl and 0.5 mM Na_2EDTA (black); 100 mM KCl and 0.5 mM K_2EDTA (red); or 100 mM LiCl and 0.5 mM Li_2EDTA (green). (b, c) Normalized UV melting curves of 4 μM R3 with 0 wt % PEG 200 in 30 mM MES buffer (pH 7.0) containing 100 mM NaCl and 0.5 mM Na_2EDTA (black); 100 mM KCl and 0.5 mM K_2EDTA (red); or 100 mM LiCl and Li_2EDTA (green). Melting was assessed by UV absorbance at 260 nm (b) or 295 nm (c).

that R3 which includes the typical G-quadruplex-forming sequence with more number of G mainly forms the A-form duplex and G-quadruplex under the dilute and molecular crowding conditions, respectively.

From the results for R1, R2, and R3 at 0 wt % and 40 wt % PEG 200, it is possible to conclude two important points. First, the main structure of GC-rich RNA sequences found in the protooncogenes is the A-form duplex. Second, only the combination of K^+ and molecular crowding can induce G-quadruplex structure, while this is the case only for the ratio of G which is much more than that of C.

RNase T1 Cleavage Assay. We further investigated the structure of R2 using RNase T1, an endoribonuclease that specifically cleaves RNA at the 3′-end of free G nucleotides but not at G involved in secondary and tertiary structures (*77*, *78*). The cleavage reaction was performed in the presence of 100 mM Na^+, K^+, or Li^+ at 0 wt % PEG 200 (Figure 8). Interestingly, an identical cleavage pattern was observed in the presence of different cation species. If G-quadruplex is stabilized by K^+, there should be a structure change depending on the coexisting cation, leading to a change in the cleavage pattern. Thus, the identical cleavage pattern independent of the type of cation showed that R2 folded into a duplex, consistent with the CD spectra indicating an A-form duplex. Cleavage results showed that the G nucleotides located at positions 7, 11, 20, 25, 35, 37, 39, 43, and 44 were completely cleaved. These G nucleotides were not involved in base pair formation. Conversely, G nucleotides at positions 21, 22, 23, 29, 30, and 34 appeared to be protected from RNase T1 cleavage, indicating that they were involved in base pair formation. In addition, G nucleotides at position 28 appeared to be partially protected from RNase T1 cleavage. On the basis of the cleavage pattern, we propose a possible secondary structure of R2 in the monomeric and dimeric forms (Figure 9). Two stretches with G nucleotides from 21 to 23 and from 28 to 30 are involved in base pairing. Therefore, these G nucleotides form duplex stems with C nucleotides. On the other hand, most of the G nucleotides near the 3′-end are not involved in base pairing. In the proposed intramolecular structure, G7 and G11 are not involved in base pairs. However, G11 is at the junction connecting the two stems. Thus, the accessibility of RNase T1 is reduced. In the case of G7, this nucleotide is in the terminal base pair with a very long dangling end, which may result in flanking of the terminal base pair. Moreover, the intermolecular structure shown in Figure 9 is consistent with the cleavage pattern. In addition, there is a possibility of noncanonical unstable base pair formation which may protect from the cleavage. Although details of the structure still remain unclear, the cleavage assay further confirmed duplex formation of R2 in monomer or dimer forms, whereas

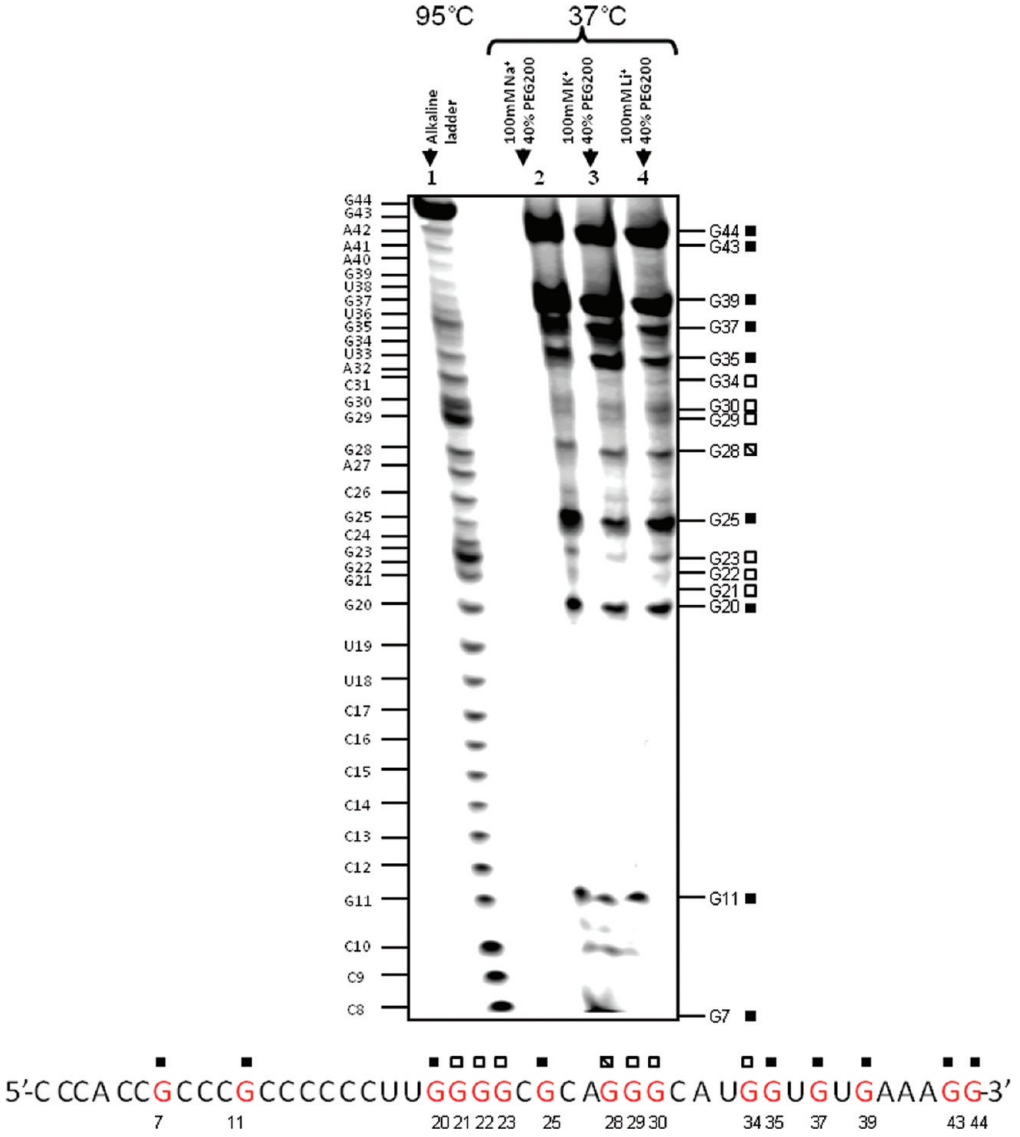

FIGURE 8: RNase T1 cleavage of R2 assay with 0 wt % PEG 200 in 30 mM MES buffer (pH 7.0) containing 100 mM NaCl and 0.5 mM Na_2EDTA (lane 2); 100 mM KCl and 0.5 mM K_2EDTA (lane 3); or 100 mM LiCl and 0.5 mM Li_2EDTA (lane 4). Cleaved products were compared with alkaline ladder bands (lane 1) of R2 to assign the positions of free cleaved G. Open squares indicate the bases involved in base pair formation. Filled squares indicate bases involved in loop formation. Crossed squares indicate partially cleaved bases.

it is not possible to draw a G-quadruplex structure from these cleavage patterns because at least four continuous G-stretches are required for the formation of G-quadruplex. In addition, the RNase T1 cleavage pattern for R2 in the presence of 100 mM Na^+, K^+, or, Li^+ at 40 wt % PEG 200 (Supporting Information Figure S5) was similar to the cleavage pattern at 0 wt % PEG 200 (Figure 8), although a few positions were different from them at 0 wt % PEG 200. On the basis of the cleavage result, we proposed a possible secondary structure of R2 in the monomeric and dimeric forms under the molecular crowding conditions (Supporting Information Figure S6) which slightly differs from the proposed model under dilute condition (Figure 9), for example, at position G20. This difference may be induced by the molecular crowding which destabilizes Watson−Crick base pairs and stabilizes non-Watson−Crick base pairs as well as non base pair regions (*79*). From this RNase T1 cleavage assay for R2, it is therefore possible to conclude that R2 forms the A-form duplex which is not altered by the coexisting cations.

Conclusion and Biological Meaning of the RNA and DNA Structures. Our present study revealed the enrichment of not only G but also C nucleotides in the regulatory regions of protooncogenes. We further demonstrated that such GC-rich RNAs formed stable and monomorphic A-form duplexes, although only the combination of K^+ and molecular crowding can induce G-quadruplex structure of the GC-rich RNA sequence with more G and less C nucleotides. On the contrary to GC-rich RNA sequences, corresponding DNA sequences showed structural polymorphisms depending on the cellular crowding factors (like cations and molecular crowding). Considerable structural differences in DNA and RNA sequences highlight their possible differential functional roles during gene expression. Unlike DNA-mediated regulatory signals, whose activity is essentially mediated by their primary structure, the biological activity of regulatory motifs at the RNA level relies on a combination of both primary and secondary structures. mRNAs encoding proteins involved in many biological processes, such as growth factors, transcription factors, or protooncogenes, all need to be strongly and finely regulated and form stable secondary structures that affect translation efficiency. In this regard, our present study on GC-rich RNA sequences capable of forming robust and stable A-form duplexes insensitive to the surrounding conditions can provide valuable information about their roles in biological

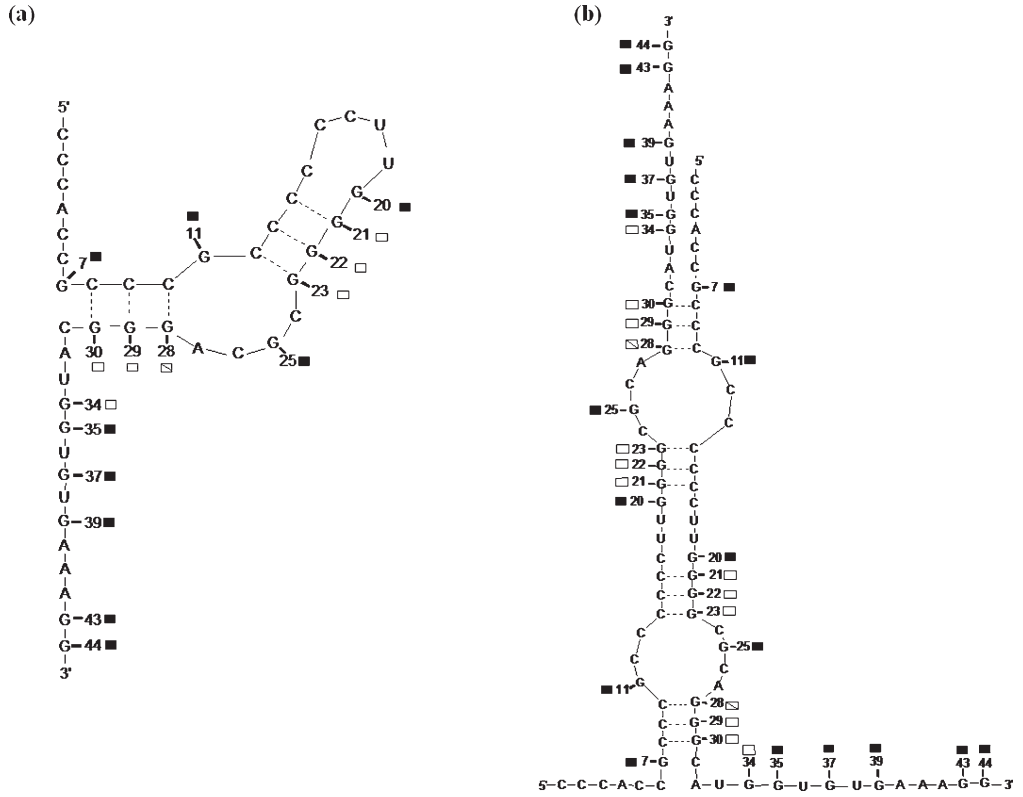

FIGURE 9: Proposed model of R2 in monomer form (a) and in dimer form (b) based on the cleavage results of the RNase T1 assay shown in Figure 8.

functions. The major difference in the backbone of RNA and DNA is the presence of a 2′-OH group in the ribose ring. This is chemically subtle but has a large effect on the corresponding structures. The presence of a 2′-OH group in RNA locks the glycosidic torsion angles to the *anti* conformation. Therefore, in the random coil state, both *syn* and *anti* conformations are possible in DNA, while only the *anti* conformation is possible in RNA. This difference in the initial state of folding can lead to the monomorphic structure of RNA and polymorphic structure of DNA. Moreover, it has been reported that water molecules form bridges between the 2′-OH of ribose and phosphate oxygens on the 3′ side in a crystal structure of tRNA (*80*). It is also well-known that under dehydration conditions B-form duplexes switch reversibly to A-form duplexes in DNA (*81*). These coordinated water molecules may affect the structure and stability of RNA and DNA. Further quantitative studies on the hydration of DNA and RNA structures, including duplexes and G-quadruplexes, are required to understand how the chemically subtle difference in the 2′-OH of ribose results in the large difference between RNA and DNA structures. Monomorphic RNA structures and polymorphic DNA structures responding to cellular environmental factors imply possible roles of DNA to regulate gene expression and of RNA to maintain its function in living cells depending on the cell cycle and conditions.

SUPPORTING INFORMATION AVAILABLE

CD spectra of R1, R2, D1, and D2 in the presence of different cation species at 40 wt % PEG 200; nondenaturing 15% polyacrylamide gel of R1 and R2 in the presence of different cation species at 40 wt % PEG 200; normalized UV melting curves of the R2 in the presence of different cation species at 0 wt % PEG 200 and 40 wt % PEG 200 at 260 and 295 nm; CD spectra of R3 at 40 wt % PEG 200 and normalized UV melting curves of the R3 in the presence of different cation species at 40 wt % PEG 200 at 260 and 295 nm; RNase T1 cleavage assay of R2 in the presence of different cation species at 40 wt % PEG 200; proposed model of R2 in monomer and dimer form based on RNase T1 cleavage results at 40 wt% PEG 200. This material is available free of charge via the Internet at http://pubs.acs.org.

REFERENCES

1. Watson, J. D., and Crick, F. H. C. (1953) A structure for deoxyribose nucleic acid. *Nature 171*, 737–738.
2. Turner, D. H. (2000) Conformational changes (Bloomfield, V. A., Crothers, D. M., and Tinoco, I., Jr., Eds.) pp 259−334, University Science Books, Mill Valley, CA.
3. Sen, D., and Gilbert, W. (1990) A sodium-potassium switch in the formation of four-stranded G4-DNA. *Nature 344*, 410–414.
4. Sundquist, W. I., and Klug, A. (1989) Telomeric DNA dimerizes by formation of guanine tetrads between hairpin loops. *Nature 342*, 825–829.
5. Gueron, M., and Leroy, J. L. (2000) The i-motif in nucleic acids. *Curr. Opin. Struct. Biol. 10*, 326–331.
6. Blackburn, E. H. (1991) Structure and function of telomeres. *Nature 350*, 569–573.
7. Simonsson, T., Pecinka, P., and Kubista, M. (1998) DNA tetraplex formation in the control region of c-myc. *Nucleic Acids Res. 26*, 1167–1172.
8. Kexiao, G., Vijay, G., Laurence, H. H., and Daekyu, S. (2008) Intramolecularly folded G-quadruplex and i-motif structures in the proximal promoter of the vascular endothelial growth factor gene. *Nucleic Acids Res. 36*, 4598–4608.
9. Michael, C. U., Kosack, H., Dobrinski., B., Lurz, R., Docherty, K., and Kilpatrick, M. W. (1992) The human insulin gene linked polymorphic region exhibits an altered DNA structure. *Nucleic Acids Res. 20*, 231–236.
10. Fry, M., and Loeb, L. A. (1994) The fragile X syndrome d(CGG)n nucleotide repeats form a stable tetrahelical structure. *Proc. Natl. Acad. Sci. U.S.A. 91*, 4950–4954.
11. Huppert, J. L., and Balasubramanian, S. (2007) G-quadruplexes in promoters throughout the human genome. *Nucleic Acids Res. 35*, 406–413.
12. Huppert, J. L., and Balasubramanian, S. (2005) Prevalence of quadruplexes in the human genome. *Nucleic Acids Res. 33*, 2908–2916.

13. Zhang, R., Lin, Y., and Zhang, C. T. (2008) Greglist: a database listing potential G-quadruplex regulated genes. *Nucleic Acids Res. 36*, D372–eD376.
14. Todd, A. K., Johnston, M., and Neidle, S. (2005) Highly prevalent putative quadruplex sequence motifs in human DNA. *Nucleic Acids Res. 33*, 2901–2907.
15. Hershman, S. G., Chen, Q., Lee, J. Y., Kozak, M. L., Yue, P., Wang, L.-S., and Johnson, F. B. (2008) Genomic distribution and functional analyses of potential G-quadruplex-forming sequences in *Saccharomyces cerevisiae*. *Nucleic Acids Res. 36*, 144–156.
16. Kostadinov, R., Malhotra, N., Viotti, M., Shine, R., D'Antonio, L., and Bagga, P. S. (2006) GRSDB: a database of quadruplex forming G-rich sequences in alternatively processed mammalian pre-mRNA sequences. *Nucleic Acids Res. 34*, D119–D124.
17. D'Antonio, L., Bagga, P. S. (2004) IEEE Computational Systems Bioinformatics Conference, Computational methods for predicting intramolecular G-quadruplexes in nucleotide sequences, pp 561−562.
18. Kikin, O., D'Antonio, L., and Bagga, P. S. (2006) A web-based server for predicting G-quadruplexes in nucleotide sequences. *Nucleic Acids Res. 34*, W676–W682.
19. Kikin, O., Zappala, Z., D'Antonio, L., and Bagga, P. S. (2008) GRSDB2 and GRS_UTRdb: databases of quadruplex forming G-rich sequences in pre-mRNAs and mRNAs. *Nucleic Acids Res. 36*, D141–D148.
20. Eddy, J., and Maizels, N. (2006) Gene function correlates with potential for G4 DNA formation in the human genome. *Nucleic Acids Res. 34*, 3887–3896.
21. Phillips, K., Dauter, Z., Murchie, A. I., Lilley, D. M. J., and Luisi, B. (1997) The crystal structure of a parallel-stranded guanine tetraplex at 0.95 angstrom resolution. *J. Mol. Biol. 273*, 171–182.
22. Parkinson, G. N., Lee, M. P. H., and Neidle, S. (2002) Crystal structure of parallel quadruplexes from human telomeric DNA. *Nature 417*, 876–880.
23. 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.
24. Keniry, M. A., Strahan, G. D., Owen, E. A., and Shafer, R. H. (1995) Solution structure of the Na^+ form of the dimeric quadruplex. *Eur. J. Biochem. 233*, 631–643.
25. Macaya, R. F., Schultze, P., Smith, F. W., Roe, J. A., and Feigon, J. (1993) Thrombin-binding DNA aptamer forms a unimolecular quadruplex structure in solution. *Proc. Natl. Acad. Sci. U.S.A. 90*, 3745–3749.
26. Wang, Y., and Patel, D. J. (1993) Solution structure of the human telomeric repeat d[AG3(T2AG3)3] G-tetraplex. *Structure 1*, 76–94.
27. Wang, Y., and Patel, D. J. (1995) Solution structure of the *Oxytricha* telomeric repeat d G-tetraplex. *J. Mol. Biol. 251*, 263–282.
28. Phan, A. T., and Mergny, J. L. (2002) Human telomeric DNA: G-quadruplex, i-motif and Watson-Crick double helix. *Nucleic Acids Res. 30*, 4618–4625.
29. Li, W., Miyoshi, D., Nakano, S., and Sugimoto, N. (2003) Structural competition involving G-quadruplex DNA and its complement. *Biochemistry 42*, 11736–11744.
30. Miyoshi, D., Nakano, A., Toda, T., and Sugimoto, N. (2001) Effect of divalent cations on antiparallel G-quartet structure of $d(G_4T_4G_4)$. *FEBS Lett. 496*, 128–133.
31. Miyoshi, D., Matsumura, S., Li, W., and Sugimoto, N. (2003) Structural polymorphism of telomeric DNA regulated by pH and divalent cation. *Nucleosides, Nucleotides, Nucleic Acids 22*, 203–221.
32. Risitano, A., and Fox, K. R. (2003) Stability of intramolecular DNA quadruplexes: comparison with DNA duplexes. *Biochemistry 42*, 6507–6513.
33. Miyoshi, D., Nakano, A., and Sugimoto, N. (2002) Molecular crowding regulates the structural switch of the DNA G-quadruplex. *Biochemistry 41*, 15017–15024.
34. Christiansen, J., Kofod, M., and Nielsen, F. C. (1994) A guanosine quadruplex and two stable hairpins flank a major cleavage site in insulin-like growth factor II mRNA. *Nucleic Acids Res. 22*, 5709–5716.
35. Darnell, J. C., Jensen, K. B., Jin, P., Brown, V., Warren, S. T., and Darnell, R. B. (2001) Fragile X mental retardation protein targets G quartet mRNAs important for neuronal function. *Cell 107*, 489–499.
36. Bonnal, S., Schaeffer, C., Creancier, L., Clamens, S., Moine, H., Prats, A. C., and Vagner, S. (2003) A single internal ribosome entry site containing a G quartet RNA structure drives fibroblast growth factor 2 gene expression at four alternative translation initiation codons. *J. Biol. Chem. 278*, 39330–39336.
37. Morris, M. J., and Basu, S. (2009) An unusually stable G-quadruplex within the 5′-UTR of the MT3 matrix metalloproteinase mRNA represses translation in eukaryotic cells. *Biochemistry 48*, 5313–5319.
38. Kumari, S., Bugaut, A., Huppert, J. L., and Balasubramanian, S. (2007) An RNA G-quadruplex in the 5′ UTR of the NRAS proto-oncogene modulates translation. *Nat. Chem. Biol. 4*, 218–221.
39. Arora, A., Dutkiewicz, M., Scaria, V., Hariharan, M., Maiti, S., and Kurreck, J. (2008) Inhibition of translation in living eukaryotic cells by an RNA G-quadruplex motif. *RNA 14*, 1290–1296.
40. Wieland, M., and Hartig, J. S. (2007) RNA quadruplex-based modulation of gene expression. *Chem. Biol. 14*, 757–763.
41. Khateb, S., Weisman, S. P., Hershco, S. I., Ludwig, A. L., and Fry, M. (2007) The tetraplex $(CGG)_n$ destabilizing proteins hnRNP A2 and CBF-A enhance the *in vivo* translation of fragile X premutation mRNA. *Nucleic Acids Res. 35*, 5775–5788.
42. Awang, G., and Sen, D. (1993) Mode of dimerization of HIV-1 genomic RNA. *Biochemistry 32*, 11453–11457.
43. Bonnal, S., Schaeffer, C., Créancier, L., Clamens, S., Moine, H., Prats, A. C., and Vagner, S. (2003) A single internal ribosome entry site containing a G quartet RNA structure drives fibroblast growth factor 2 gene expression at four alternative translation initiation codons. *J. Biol. Chem. 278*, 39330–39336.
44. Gomez, D., Lemarteleur, T., Lacroix, L., Mailliet, P., Mergny, J. L., and Riou, J. F. (2004) Telomerase downregulation induced by the G-quadruplex ligand 12459 in A549 cells is mediated by hTERT RNA alternative splicing. *Nucleic Acids Res. 32*, 371–379.
45. Zanotti, K. J., Lackey, P. E., Evans, G. L., and Mihailescu, M. R. (2006) Thermodynamics of the fragile X mental retardation protein RGG box interactions with G quartet forming RNA. *Biochemistry 45*, 8319–8330.
46. Kumari, S., Bugaut, A., and Balasubramanian, S. (2008) Position and stability are determining factors for translation repression by an RNA G-quadruplex-forming sequence within the 5′ UTR of the NRAS proto-oncogene. *Biochemistry 47*, 12664–12669.
47. Jianying, Q. I., and Shafer, R. H. (2007) Human telomere quadruplex: refolding and selection of individual conformers via RNA/DNA chimeric editing. *Biochemistry 46*, 7599–7606.
48. 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.
49. Olsen, C. M., and Marky, L. A. (2009) Energetic and hydration contributions of the removal of methyl groups from thymine to form uracil in G-quadruplexes. *J. Phys. Chem. B. 113*, 9–11.
50. Olsen, C. M., Lee, H. T., and Marky, L. A. (2009) Unfolding thermodynamics of intramolecular G-quadruplexes: base sequence contributions of the loops. *J. Phys. Chem. B 113*, 2587–2595.
51. Arora, A., and Maiti, S. (2009) Stability and molecular recognition of quadruplexes with different loop length in absence and presence of molecular crowding agent. *J. Phys. Chem. B 113*, 10515–20.
52. Kumar, N., and Maiti, S. (2008) A thermodynamic overview of naturally occurring intramolecular DNA quadruplexes. *Nucleic Acids Res. 36*, 5610–5622.
53. Sugimoto, N., Nakano, S., Kotah, S., Matsumura, A., Nakamuta, H., Ohmichi, T., Yoneyama, M., and Sasaki, M. (1995) Thermodynamic parameters to predict stability of RNA/DNA hybrid duplexes. *Biochemistry 34*, 11211–11216.
54. Nakano, S., Kanzaki, T., and Sugimoto, N. (2004) Influences of ribonucleotide on a duplex conformation and its thermal stability: study with the chimeric RNA-DNA strands. *J. Am. Chem. Soc. 126*, 1088–1095.
55. Nicholas, K. B., Nicholas, H. B., Jr., and Deerfield, D. W., II (1997) GeneDoc: analysis and visualization of genetic variation. *EMBNEW. NEWS 4*, 14.
56. Guo, Q., Lu, M., and Kallenbach, N. R. (1993) Effect of thymine tract length on the structure and stability of model telomeric sequences. *Biochemistry 32*, 3596–3603.
57. Lu, M., Guo, Q., and Kallenbach, N. R. (1992) Structure and stability of sodium and potassium complexes of dT4G4 and dT4G4T. *Biochemistry 31*, 2455–2459.
58. Merkina, E. E., and Fox, K. R. (2005) Kinetic stability of intermolecular DNA quadruplexes. *Biophys. J. 89*, 365–373.
59. Johnson, W. C. (1994) Circular Dichroism: Principles and Applications, pp 523−540, VCH, New York, NY.
60. Fedor, S., Monique, M., Aurelie, M., Claude, M., and Serge, F. (1995) The high stability of the triple helices formed between short purine oligonucleotides and SIV/HIV-2 vpx genes is determined by the targeted DNA structure. *Nucleic Acids Res. 23*, 3831–3836.
61. Hung, S. H., Yu, Q., Gray, D. M., and Ratliff, R. L. (1994) Evidence from CD spectra that d(purine)-r(pyrimidine) and r(purine)-d(pyrimidine)

hybrids are in different structural classes. *Nucleic Acids Res. 22*, 4326–4334.
62. Markus, C. W., and Sundaralingam, M. (2000) B-form to A-form conversion by a 3′-terminal ribose: crystal structure of the chimera d(CCACTAGTG)r(G). *Nucleic Acids Res. 28*, 4356–4363.
63. Yu, H. Q., Miyoshi, D., and Sugimoto, N. (2006) Characterization of structure and stability of long telomeric DNA G-quadruplexes. *J. Am. Chem. Soc. 128*, 15461–15468.
64. Stefl, R., Trantirek, L., Vorlickova, M., Koca, J., Skelenar, V., and Kypr, J. (2001) A-like guanine-guanine stacking in the aqeuous DNA duplex of d(GGGGCCCC). *J. Mol. Biol. 307*, 513–524.
65. Pataskar, S. S., Dash, D., and Brahmachari, S. K. (2001) Intramolecular i-motif structure at acidic pH for progressive myoclonus epilepsy (EPM1) repeat d(CCCCGCCCCGCG)n. *J. Biomol. Struct. Dyn. 2*, 307–313.
66. 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.
67. Zuker, M. (2003) Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Res. 31*, 3406–3415.
68. Mathews, D. H., Disney, M. D., Childs, J. L., Schroeder, S. J., Zuker, M., and Turner, D. H. (2004) Incorporating chemical modification constraints into a dynamic programming algorithm for prediction of RNA secondary structure. *Proc. Natl. Acad. Sci. U.S.A. 101*, 7287–7292.
69. Manning., G. S. (1972) On the application of polyelectrolyte "limiting laws" to the helix-coil transition of DNA. I. Excess univalent cations. *Biopolymers 11*, 937–949.
70. Record, M. T., Jr., Anderson, C. F., and Lohman, T. M. (1978) Thermodynamic analysis of ion effects on the binding and conformational equilibria of proteins and nucleic acids: the roles of ion association or release, screening, and ion effects on water activity. *Q. Rev. Biophys. 11*, 103–178.
71. Miyoshi, D., Karimata, H., and Sugimoto, N. (2006) Characterization of structure and stability of long telomeric DNA G-quadruplexes. *J. Am. Chem. Soc. 128*, 7957–7963.
72. 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*, 4554–63.
73. 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.
74. Miyoshi, D., Matsumura, S., Nakano, S., and Sugimoto, N. (2004) Duplex dissociation of telomere DNAs induced by molecular crowding. *J. Am. Chem. Soc. 126*, 165–169.
75. 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.
76. Mergny, J. L., Phan, A. T., and Lacroix, L. (1998) Following G-quartet formation by UV-spectroscopy. *FEBS Lett. 435*, 74–78.
77. Darnell, J. C., Jensen, K. B., Jin, P., Brown, V., Warren, S. T., and Darnell, R. B. (2001) Fragile X mental retardation protein targets G quartet mRNAs important for neuronal function. *Cell 107*, 489–499.
78. Christiansen, J., Kofod, M., and Nielsen, F. C. (1994) A guanosine quadruplex and two stable hairpins flank a major cleavage site in insulin-like growth factor II mRNA. *Nucleic Acids Res. 22*, 5709–5716.
79. Miyoshi, D., Karimata, H., and Sugimoto, N. (2007) Hydration regulates the thermodynamic stability of DNA structures under molecular crowding conditions. *Nucleosides, Nucleotides, Nucleic Acids 26*, 589–595.
80. Jovine, L., Djordjevic, S., and Rhodes., D. (2000) The crystal structure of yeast phenylalanine tRNA at 2.0 Å resolution: cleavage by Mg(2+) in 15-year old crystals. *J. Mol. Biol. 301*, 401–414.
81. Paster, N. (2005) The B- to A-DNA transition and the reorganization of solvent at the DNA surface. *Biophys. J. 88*, 3262–3275.