JACS
JOURNAL OF THE AMERICAN CHEMICAL SOCIETY
Article
pubs.acs.org/JACS

Unexpected Position-Dependent Effects of Ribose G-Quartets in G-Quadruplexes

Jun Zhou,[*,†,‡] Samir Amrane,[‡] Frédéric Rosu,[#] Gilmar F. Salgado,[‡] Yunqiang Bian,[⊥] Hisae Tateishi-Karimata,[¶] Eric Largy,[‡] Dursun Nizam Korkut,[‡] Anne Bourdoncle,[‡] Daisuke Miyoshi,[¶] Jian Zhang,[∥] Huangxian Ju,[†] Wei Wang,[*,∥] Naoki Sugimoto,[¶] Valérie Gabelica,[‡] and Jean-Louis Mergny[*,‡,§]

[†]State Key Laboratory of Analytical Chemistry for Life Science, School of Chemistry & Chemical Engineering, Nanjing University, Nanjing 210023, China

[‡]Université de Bordeaux, INSERM U1212, CNRS UMR 5320, ARNA Laboratory, IECB, F-33600 Pessac, France

[#]Université de Bordeaux, CNRS UMS 3033, INSERM US001, IECB, F-33600 Pessac, France

[⊥]Shandong Provincial Key Laboratory of Functional Macromolecular Biophysics, Institute of Biophysics, Dezhou University, Dezhou 253023, China

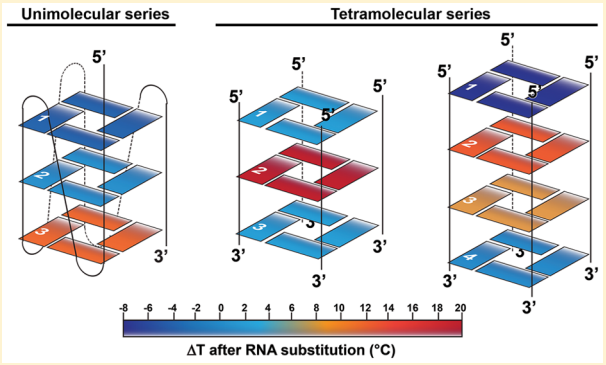
[¶]Frontier Institute for Biomolecular Engineering Research (FIBER) and Graduate School of Frontiers of Innovative Research in Science and Technology (FIRST), Konan University, 7-1-20 Minatojima-minamimachi, Chuo-ku, Kobe 650-0047, Japan

[∥]Collaborative Innovation Center of Advanced Microstructures and Department of Physics, Nanjing University, Nanjing 210093, China

[§]Institute of Biophysics of the CAS, v.v.i., Královopolská 135, 612 65 Brno, Czech Republic

Supporting Information

ABSTRACT: To understand the role of ribose G-quartets and how they affect the properties of G-quadruplex structures, we studied three systems in which one, two, three, or four deoxyribose G-quartets were substituted with ribose G-quartets. These systems were a parallel DNA intramolecular G-quadruplex, d(TTGGGTGGGTTGGGTGGGTT), and two tetramolecular G-quadruplexes, d(TGGGT) and d(TGGGGT). Thermal denaturation experiments revealed that ribose G-quartets have position-dependent and cumulative effects on G-quadruplex stability. An unexpected *destabilization* was observed when rG quartets were presented at the 5′-end of the G stack. This observation challenges the general belief that RNA residues stabilize G-quadruplexes. Furthermore, in contrast to past proposals, hydration is not the main factor determining the stability of our RNA/DNA chimeric G-quadruplexes. Interestingly, the presence of rG residues in a central G-quartet facilitated the formation of additional tetramolecular G-quadruplex topologies showing positive circular dichroism signals at 295 nm. 2D NMR analysis of the tetramolecular TGgGGT (lowercase letter indicates ribose) indicates that Gs in the 5′-most G-quartet adopt the *syn* conformation. These analyses highlight several new aspects of the role of ribose G-quartets on G-quadruplex structure and stability, and demonstrate that the positions of ribose residues are critical for tuning G-quadruplex properties.



INTRODUCTION

G-quadruplexes are four-stranded nucleic acid structures formed by guanine-rich nucleic acids in the presence of certain cations (typically Na^+ and K^+).[1−3] Both DNA and RNA G-quadruplexes possess physiological significance in eukaryotes and prokaryotes; as regulators of replication and gene expression, these structures are potential therapeutic targets.[4−13] Compared with DNA G-quadruplexes which have been studied in great detail, for RNA G-quadruplexes the information available is less extensive. The main differences between DNA and RNA are that (i) the sugar present in DNA is a deoxyribose, whereas the sugar present in an RNA molecule is a ribose, which means that there is a hydroxyl group at the C2′ position (2′-OH) of the monosaccharide in RNA, and (ii) uracil (U) is found in RNA, whereas thymine (T), which has a methyl group, is present in DNA. The effects of these changes on G-quadruplexes are not fully elucidated.

Our work and other reports show that RNA G-quadruplex structures are generally more stable than their DNA counter-

Received: January 19, 2017
Published: May 19, 2017

DOI: 10.1021/jacs.7b00648
J. Am. Chem. Soc. 2017, 139, 7768−7779

parts.[14−17] Uracil and thymine tetrads have different structural characteristics that contribute to the differences between DNA and RNA G-quadruplexes.[15,18−20] Furthermore, the stabilization may result from a lower hydration of RNA compared to DNA G-quadruplexes,[16] although recent crystallographic data revealed subtle differences between hydrogen-bonding interactions and water structure in DNA and RNA G-quadruplexes.[8,21] Therefore, the exact role of the 2′-OH and how it affects G-quadruplexes is still unclear.

RNA G-quadruplexes are generally less polymorphic than their DNA counterparts. For example, the two-repeat human telomeric RNA sequence r(UAGGGUUAGGGU) forms a parallel-stranded bimolecular G-quadruplex in both Na^+ and K^+ solutions,[23,24] whereas its DNA counterpart adopts multiple conformations in solution.[7,8] Furthermore, the recent crystal structures of both RNA and DNA parallel G-quadruplexes formed by r(UAGGGUUAGGGU) and d(TAGGGTTAGGGT), respectively, revealed that these two structures exhibit a high degree of similarity. The additional hydroxyl group in RNA imposes a steric hindrance favoring the *anti*-glycosidic torsion angles for the RNA bases, making the parallel structure the predominant conformation.[8,21] This conclusion agrees with previous results obtained by selective substitution of rG for dG.[25] Thus, the consensus is that RNA G-quadruplexes are more stable than their DNA counterparts[8,14−17] and adopt parallel-stranded configurations because the rG residues favor the *anti* glycosidic conformation.[5−8,16,17,21,23−25]

To investigate the role of 2′-OH groups on the properties of G-quadruplex structures, we designed both intra- and intermolecular parallel G-quadruplexes, in which we selectively substituted the DNA G-quartets with one, two, three, or four RNA quartets. U-quartets were not considered here because they have been previously demonstrated to stabilize G-quadruplexes.[15,18−20]

The G-quadruplex formed by d(TTGGGTGGGTTGGGTGGGTT) is expected to be intramolecular monomorphic, due to the terminal thymines,[26] and fully parallel (all guanines in *anti*) because of the two single-thymine loops (underlined).[27] The sequence chosen is expected to be monomorphic; changes in salt composition or sugar nature are unlikely to affect its parallel fold, in contrast with other motifs.[28] We selectively substituted one, two, or three G-quartets with RNA (quartets are numbered as shown in Figure 1A) in order to obtain all possible chimeric analogues, referred to as the "RQ" series (Table 1; each individual quartet being all RNA or all DNA). The insertion of ribose G-quartets was not expected to affect G-quadruplex topology because the 2′-OH favors the *anti* glycosidic conformation,[5−8,16,17,21,23−25] and chimeric analogues can thus be used to study the effect of the 2′-OH's on a given structure.

We also studied RNA-to-DNA substitutions in two tetramolecular G-quadruplex structures involving either three or four G-quartets (Figure 1B). We named the chimeric analogues with three and four quartets as the "T" series and "F" series for three and four guanine strands, respectively. We selectively substituted one or more dGs with rG to obtain all possible analogues containing one, two, three, or four rG residues (Table 2). T0 and F0 (all DNA) predominantly form parallel structures with all guanine residues in *anti* glycosidic conformation, but a minor species with the 5′-terminal quartet in *syn* conformation has also been reported.[29−32] The introduction of hydroxyl groups in the chimeric analogues was not expected to induce a change in topology for the reasons described above.

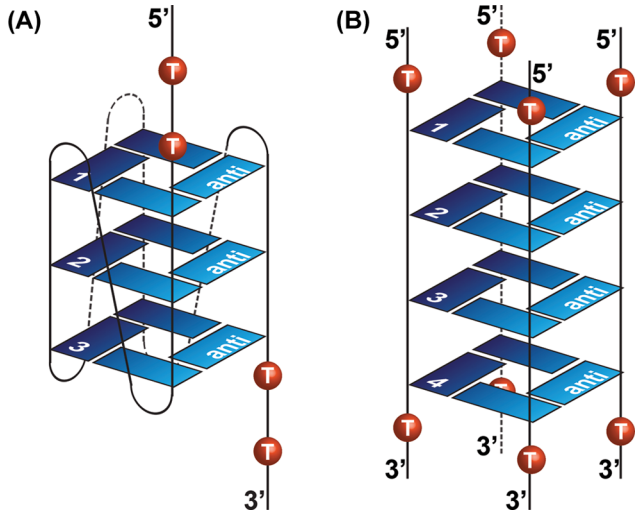


Figure 1. Designed (A) intramolecular and (B) tetramolecular G-quadruplex structures, in which all strands are parallel and all guanine residues are in the *anti* glycosidic conformation. The selective substitution of DNA G-quartets (numbered 1, 2, 3, or 4) with RNA was not expected to change the topology. Two T residues were added to both ends of the intramolecular G-quadruplex to favor the formation of monomeric G-quadruplex structures.

Table 1. Sequences of RQ Series Oligonucleotides and Apparent Melting Temperatures of the Intramolecular Structures Adopted in 10 mM Cacodylic Acid Buffer (pH 7.0) Supplemented with 5 mM K^+, 100 mM Na^+, or 100 mM NH_4^+[a]

		T_m[c] (°C)		
name	sequence[b]	K^+[d]	Na^+	NH_4^+
RQ0	TTGGGTGGGTTGGGTGGGTT	72	53	57
RQ1	TTgGGTgGGTTgGGTgGGTT	68	51.5	52
RQ2	TTGgGTGgGTTGgGTGgGTT	73	54.5	60
RQ3	TTGGgTGGgTTGGgTGGgTT	83	66	71
RQ12	TTggGTggGTTggGTggGTT	73	56	61
RQ13	TTgGgTgGgTTgGgTgGgTT	79	65	67
RQ23	TTGggTGggTTGggTGggTT	85	72	77
RQ123	TTgggTgggTTgggTgggTT	80	71	70

[a]Strand concentration, 5 μM; heating rate, 0.5 °C/min; absorbance monitored at 295 nm. [b]Lower case indicates ribonucleotides. [c]Standard deviation is ≤1 °C. [d]T_ms determined in 5 mM K^+ as RQ3 and RQ23 do not fully unfold at 99 °C in 100 mM K^+.

As expected, we found that all these chimeras formed parallel G-quadruplex structures. Surprisingly, we found that a 5′-ribose G-quartet *destabilized* the G-quadruplexes. These observations challenge the simplistic view that 2′-OH's of ribose residues always have a stabilizing effect.[5−8,14−17,21−25] Furthermore, in some tetramolecular structures, the presence of rG residues in the central G-tracts leads to the formation of additional G-quadruplex structures containing the 5′-most guanines in *syn* conformation. Our studies indicate that a simple hydroxyl group may play a critical and position-dependent role in the G-quadruplex structure and stability and that the effects of the ribose on the stability of G-quadruplex structures are cumulative.

DOI: 10.1021/jacs.7b00648
J. Am. Chem. Soc. 2017, 139, 7768−7779

Table 2. Sequences of T and F Series Oligonucleotides, Apparent Melting Temperatures ($T_{1/2}$; Hysteresis Was Observed) of the Tetramolecular G-Quadruplex Structures Formed, Measured by UV and CD in the Presence of 100 mM K^+ (T series) or 100 mM Na^+ (F series), and the Differences in $T_{1/2}$ Values between the Chimeric and Control Samples[a]

name	sequence[b]	$T_{1/2}$[c] (°C) UV	$T_{1/2}$[c] (°C) CD	$\Delta T_{1/2}$[d] (°C)
		T Series in 100 mM KCl		
T0	TGGGT[e]	50	47	0 (0)
T1	TgGGT	52	48	2 (1)
T2	TGgGT	69	63	19 (16)
T3	TGGgT	52	49	2 (2)
T12	TggGT	60	55	10 (8)
T13	TGgGT	56	51	6 (4)
T23	TGggT	74	67	24 (20)
T123	TgggT	67	61	17 (14)
UG3U	ugggu[f]	52, >90[g,h]	48, >90[g,h]	2, >40 (1, >40)[g,h]
		F Series in 100 mM NaCl		
F0	TGGGGT[e]	59	54	0 (0)
F1	TgGGGT	51	46	−8 (−8)
F2	TGgGGT	70	66	11 (12)
F3	TGGgGT	66	60	7 (6)
F4	TGGGgT	59	54	0 (0)
F12	TggGGT	64	59	5 (5)
F13	TgGgGT	59	53	0 (−1)
F14	TgGGgT	53	48	−6 (−6)
F23	TGggGT	77	72	18 (18)
F24	TGgGgT	73	69	14 (15)
F34	TGGggT	72	67	13 (13)
F123	TgggGT	74	69	15 (15)
F124	TggGgT	67	62	8 (8)
F134	TgGggT	66	60	7 (6)
F234	TGgggT	81	76	22 (22)
F1234	TggggT	62, 82[g,h]	58, 76[g,h]	3, 23 (4, 22)
UG4U	uggggu[f]	>90[h]	89	>31 (35)

[a]Strand concentration, 10 μM; heating rates, 0.5 and 0.25 °C/min for UV and CD, respectively. See text for details. [b]Lowercase letters indicate RNA bases. [c]Standard deviation ≤1 °C. [d]The differences in $T_{1/2}$ values between the chimeric samples and respective control structures, T0 or F0, as measured by UV and CD (in parentheses). [e]All DNA sequence. [f]All RNA sequence. [g]The melting curve showed two transitions; $T_{1/2}$ values for each transition are given. [h]Due to high melting temperature, sample was not fully dissociated at high temperature.

RESULTS, PART I: INTRAMOLECULAR CHIMERIC G-QUADRUPLEX STRUCTURES

Intramolecular Chimeric G-Quadruplex Structures Are Parallel and Monomorphic. To assess the molecularity and topology of structures formed by oligonucleotides, several biophysical methods were combined.[33] Circular dichroism (CD) spectra of all samples acquired in the presence of 100 mM K^+, Na^+, or NH_4^+ exhibit positive and negative peaks around 260 and 240 nm, respectively (Supporting Information, Figure S1), in agreement with the formation of parallel (type I) G-quadruplexes for all sequences under all conditions.[34−36] The presence of two very short loops successfully constrained the topology of the RQ series oligonucleotides to fully parallel folds in all cases. All the absorbance thermal difference spectra (TDS) have two positive peaks around 243 and 273 nm and a negative peak around 295 nm, which are typical G-quadruplex features (Figure S2).[37] All the chimeric analogues thus form type I G-quadruplex structures in our conditions.

NMR spectroscopy confirmed monomorphism: spectra of all samples show a single set of 12 protons signals in the 10.8−12.2 ppm region, characteristic of Hoogsteen hydrogen bonds, suggesting that these chimeras form a single G-quadruplex structure with three G-quartets (spectra obtained in K^+ are shown in Figure 2; spectra in Na^+ and NH_4^+ are similar). Spectra are similar to those previously described for intramolecular G-quadruplexes.[38]

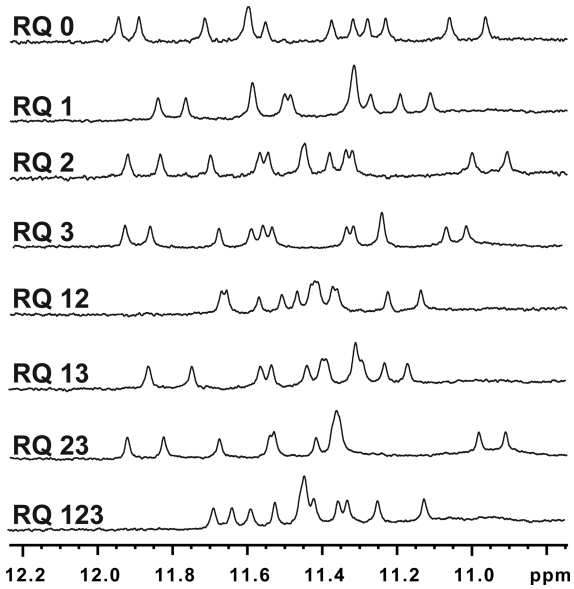


Figure 2. ^{1}H NMR imino regions of RQ series oligonucleotides (sequences shown in Table 1) in 10 mM phosphate buffer containing 100 mM K^+ (pH 6.9).

The strand molecularity was determined by electrospray ionization mass spectrometry (ESI-MS) in ammonium acetate[39,40] and from relative elution volumes on size-exclusion high-performance liquid chromatography (SE-HPLC).[41] The ESI-MS spectra revealed that the monomer peaks predominate (charge states 5− and 4−) in all chimeric sequences (Figure S3). A zoom on the 5− charge state demonstrated that the oligonucleotides retain two ammonium cations, indicating the formation of three G-quartets with tight coordination of two ammonium cations between them. The SE-HPLC chromatograms also have one main peak with reduced elution volumes indicating a monomeric (intramolecular) form. Only one sequence, RQ12, displayed a peak corresponding to a dimer; abundance was less than 10% (Figure S4). These results demonstrate that the RQ chimeric sequences predominately fold into monomeric G-quadruplex structures.

Ribose G-Quartets Have Position-Dependent and Cumulative Effects on G-Quadruplex Thermal Stability. Because all chimeric sequences derived from RQ0 fold into monomeric, parallel-stranded G-quadruplexes with all guanines in the *anti* conformation, thermal stabilities can be directly compared by ultraviolet (UV) melting experiments as previously described[42] to study the effect of 2′-OH groups on G-quadruplex stability. The apparent melting temperatures (T_m) in 5 mM K^+, 100 mM Na^+, or 100 mM NH_4^+ are summarized in Table 1 (melting profiles shown in Figure S5).

DOI: 10.1021/jacs.7b00648
J. Am. Chem. Soc. 2017, 139, 7768−7779

A potassium concentration of 100 mM led to very high melting temperatures, and under this condition, T_ms could not be determined for RQ3 and RQ23 (not shown). The stability of chimeric G-quadruplexes depends strongly on the location of the ribose. Strikingly, as compared to the all-DNA parent sequence, RQ1 is *destabilized* (by −4, −1.5, and −5 °C in K^+, Na^+, and NH_4^+, respectively), whereas RQ3 is *stabilized* (by +11, + 13, and +14 °C in K^+, Na^+, and NH_4^+, respectively); RQ2 has an intermediate behavior (Figure 3 and Table 1). The

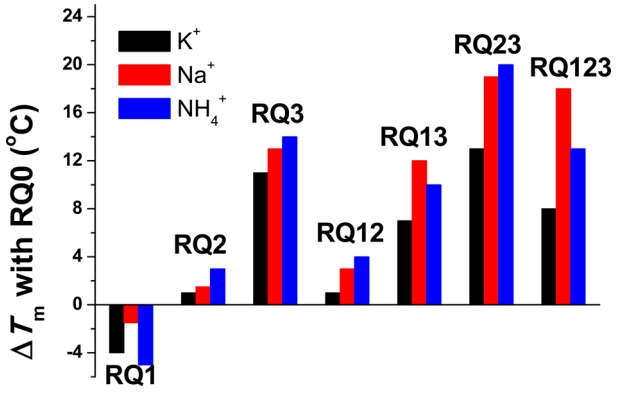


Figure 3. Relative T_m values of chimeric G-quadruplex structures (RQ1−RQ123) as compared to the all-DNA RQ0 sequence in the presence of 5 mM K^+ (black), 100 mM Na^+ (red), or 100 mM NH_4^+ (blue).

usual assumption that a 2′-OH stabilizes RNA G-quadruplexes is therefore clearly an oversimplification, as a *destabilization* effect was observed when the 5′-most G-quartet residues were ribose.

When more than one quartet was substituted with RNA, the effects were cumulative. Among the structures with two ribose quartets, RQ23 was more stable than RQ12 and RQ13. The ribose in the 5′-most G-quartet destabilizes the G-quadruplexes formed by RQ12 and RQ13 (Figure 3).

To further evaluate destabilization and cumulative effects of ribose substitutions, we determined the standard free energies of unfolding at 25 °C ($\Delta G°_{25}$) from the melting profiles (Figure S5) assuming a two-state process (Table S1). The destabilization and cumulative effects described above were thus maintained at 25 °C. Furthermore, in K^+, RQ2, RQ3, and RQ23 were more stable than RQ12, RQ13, and RQ123, respectively, which indicates that a ribose G-quartet located at the 5′ position is destabilizing (Figure 1A). Although results in sodium are more complex (see Supporting Information), we conclude that the position-dependent and cumulative effects of ribose G-quartets are general.

Hydration Is Not the Sole Factor Determining the Stability of RNA G-Quadruplexes. Previous work demonstrated that hydration greatly affects the stability of DNA and RNA G-quadruplexes.[16] The position-dependent T_m values observed here for chimeric G-quadruplexes prove that the position of the ribose G-quartet is very important for the stability of G-quadruplex structures. To get insight into the role of the 2′-OH, we investigated the hydration effect using the osmotic stressing method described previously.[16,43−45] UV melting profiles (Figure S6) and thermodynamic data (Table S2) of these chimeric analogues in the presence of 100 mM Na^+ in a buffer containing varying amounts of PEG 200 (0%−20%, v/v) demonstrated that the stability ($\ln(K_{obs})$) of chimeric G-quadruplexes decreases linearly with an increase in $\ln(a_w)$ (Figure S7). The slope allows one to estimate the number of water molecules released during G-quadruplex formation.[43−45] We calculated the number of water molecules released per nucleotide upon folding in Na^+ using the equations described previously.[43−45]

All oligonucleotides released water molecules upon G-quadruplex formation (Figure 4), in agreement with the

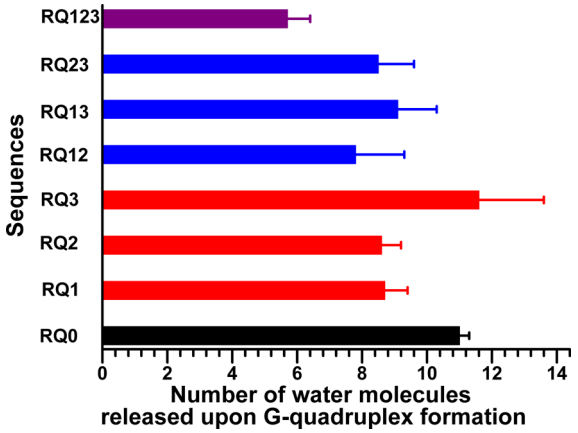


Figure 4. Number of water molecules released per nucleotide upon formation of intramolecular G-quadruplex in 100 mM Na^+.

literature.[16,44−49] All chimeric G-quadruplexes except RQ3 released fewer water molecules than their fully DNA counterpart RQ0. Dehydration cannot, however, explain the *differences* in stability observed for our chimeric sequences. For substitutions of single G-quartets with ribose, the number of water molecules released during G-quadruplex formation was in the order RQ3 > RQ1 ≈ RQ2 (Figure 4). This order is not in agreement with thermal stability, as RQ2 was more stable than RQ1 in 100 mM Na^+ (Table 1 and Figure 3). Furthermore, for double-quartet ribose substitutions (RQ12, RQ13, and RQ23), similar numbers of water molecules were released upon G-quadruplex formation, but their stabilities differed significantly. Finally, the fully RNA RQ123 released fewer water molecules than any other structure, although it was more stable than nearly all other G-quadruplexes on the basis of T_m values. These results clearly demonstrate that factors other than hydration affect the stability of G-quadruplexes containing one or more ribo-G-quartets.

RESULTS, PART II: TETRAMOLECULAR CHIMERIC G-QUADRUPLEX STRUCTURES

CD Reveals the Presence of *Syn* Guanines in Several Chimeric DNA/RNA Sequences. The CD spectra of chimeric oligonucleotides containing a single rG in the T and F series, in the presence of K^+ or Na^+, are shown in Figure 5A,B, and spectra of oligonucleotides containing several rG's are shown in Figure S8. Surprisingly, although the spectra of all-DNA and all-RNA oligonucleotides exhibit a perfect type-I CD profile,[34−36] spectra of T2 and F2, and to a lesser extent T3 and F3, have shoulder peaks around 295 nm (Figure 5A,B). Spectra of T23, F23, F24, and F234 also present the same shoulder (Figure S8). A positive CD signal around 295 nm typically indicates the presence of guanine residues in a *syn* glycosidic conformation that enables alternate stacking between quartets.[50] The low intensity of this band suggests that only a fraction of the

DOI: 10.1021/jacs.7b00648
J. Am. Chem. Soc. 2017, 139, 7768−7779

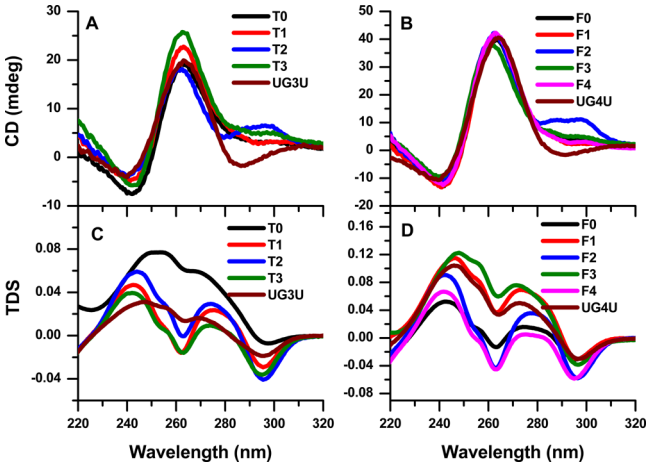


Figure 5. (A,B) CD spectra and (C,D) TDS of the chimeric oligonucleotides from T (in K^+) and F (in Na^+) series and their corresponding natural DNA and RNA controls. Spectra of chimeric oligonucleotides containing more than one rG are shown in Figure S8. CD measurements were performed at 20 °C in 10 mM cacodylate buffer (pH 7.0) with 100 mM K^+ or Na^+. The concentration of each sample was 10 μM, and sequences are provided in Table 2.

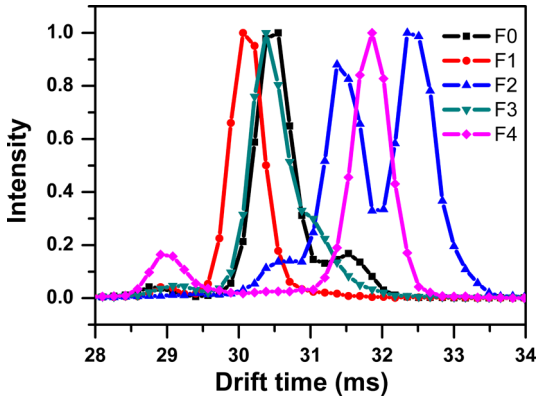


Figure 6. Arrival time distribution of the tetramolecular G-quadruplexes F0, F1, F2, F3, and F4 in K^+ conditions, monitored by ion mobility spectrometry. The samples were prepared in 100 mM trimethylammonium acetate with 1 mM KCl.

population has *syn* guanines. We observed this phenomenon when rG was in position 2, but not position 1 (Figure 1B). The situation was slightly different in the presence of NH_4^+. In the F series, the shoulder peak at 295 nm for F2 is similar to that in Na^+ solution (Figure S9A). However, shoulder peaks are insignificant in spectra of F3 and F4 (Figure S9A), whereas F34 shows the 295 nm peak in NH_4^+ but not in Na^+ (Figures S8B and S9B). The nature of the cations (Na^+ and NH_4^+) subtly modulates the structure of chimeric F series samples, and precautions were taken when comparing data. The TDS of all tetramolecular complexes were similar except for some differences in intensity (Figures 5C,D, S8 and S9D). The profiles are all typical of G-quadruplex structures.[37]

All Structures Are Tetramolecular. To confirm tetramolecular G-quadruplex formation, we measured the elution volumes of all samples by SE-HPLC. The F series samples in Na^+ eluted in the 1.52—1.57 V_e/V_0 range, in agreement with the formation of tetramolecular assemblies (Figure S10).[41] The minor peaks that eluted at ca. 1.70 V_e/V_0 correspond to the monomers. Monomer abundance as measured by SE-HPLC was inversely correlated with G-quadruplex stability (e.g., F0, F1, F4, F13, and F14 had $T_{1/2}$ < 60 °C; Table 2). We did not detect dimers or octamers in Na^+. The presence of the CD shoulder at 295 nm was not, therefore, the result of the formation of G-quadruplexes with different molecularities. We also performed SE-HPLC measurements in NH_4^+ solution, and the results were similar to those in Na^+ (Figure S11).

ESI-MS confirmed the formation of tetramolecular complexes, both in ammonium (100 mM NH_4OAc) and in potassium (1 mM KCl in 100 mM trimethylammonium acetate, TMAA).[51,52] All the chimeric sequences with one rG quartet had spectra similar to that of the F0 control sequence, with three main m/z peaks attributed to tetramolecular G-quadruplexes with 6, 5, or 4 negative charges (Figure S12). All tetramers of the F series were detected with three monovalent cations bound, confirming the formation of four quartets.

Several Tetramolecular Topologies Coexist when the Second G-Quartet Is rG. Ion mobility spectrometry separation was performed on the F0—F4 tetramers (Figure 6). The ions of a given m/z (here corresponding to the respective tetramers with three monovalent cations bound) drift in the mobility tube with a velocity dependent on their compactness. Interestingly, the F2 tetramer had a bimodal arrival time distribution, indicating that at least two conformations are adopted by this oligonucleotide and that these conformations do not interconvert in the drift tube.

In 1D ^{1}H NMR spectra, the imino proton region (which is a distinctive feature of G-quartets[38]), the aromatic region (where we observe the H8 from guanines and H6 from thymines), and the sugar region (also with characteristic signatures[40]) are used to determine whether the oligonucleotide adopts a unique structure, and to provide insight into its symmetry. In agreement with previous reports,[29,31] ^{1}H spectra of the unmodified samples, T0 and F0, show respectively three and four main signals in the 10.8—11.7 ppm region and six main singlets in the 7—8 ppm region, consistent with highly symmetric G-quadruplex structures containing three (for T0) and four (for F0) G-quartets (Figure 7).

NMR spectra for T and F series oligonucleotides with single rG quartets are presented in Figure 7. Additional chimeric sequences with several rG quartets are shown in Figure S13. The T1, F1, and F4 samples have spectra similar to those of their corresponding control DNA samples, T0 or F0, indicating that these sequences adopt a single predominant G-quadruplex conformation. However, we found that the ^{1}H NMR spectra of T2 and F2 have additional imino peaks in this region (six and eight for T2 and F2, respectively; Figure 7). The spectra of oligonucleotides of the T series containing at least two rGs are even more complex: these spectra exhibit significantly more peaks in the imino region (Figure S13A). The same holds for the F23, F24, F34, and F234 sequences: several topologies coexist.

Interestingly, the samples exhibiting multiple topologies (T2, T12, T13, T23, T123, F2, F23, F24, and F234) all have the rG in the second position of the G tract. Introducing a hydroxyl group at this position has the greatest effect on structure. This observation is correlated with the CD results, in which a shoulder peak around 295 nm was generally observed for the oligonucleotides with an rG at the second position (Figures 5 and S13; exceptions are T12, T13, and T123).

DOI: 10.1021/jacs.7b00648
J. Am. Chem. Soc. 2017, 139, 7768−7779

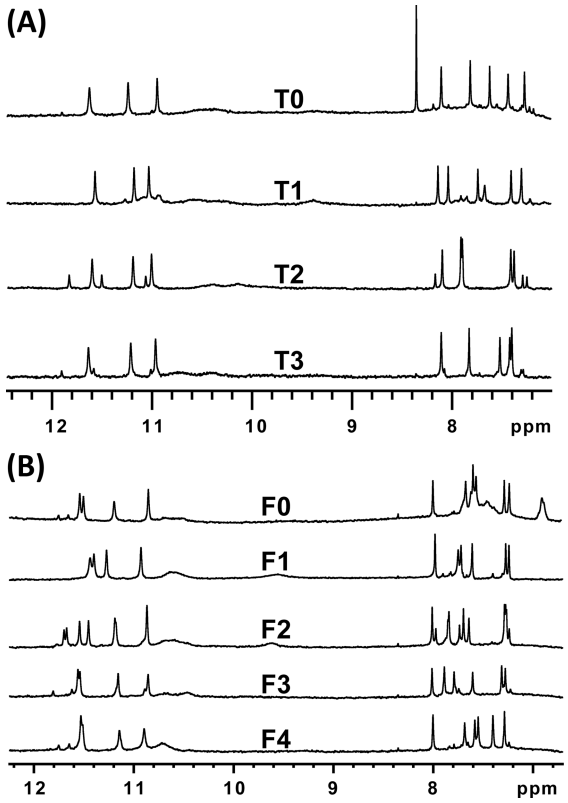


Figure 7. ^{1}H NMR spectra of the imino regions of chimeric oligonucleotides containing one rG base of (A) T series in K^+ (pH 6.9) and (B) F series in Na^+ (pH 6.8) and their corresponding natural DNA controls (sequences in Table 2). Samples were prepared in 10 mM phosphate buffer containing 100 mM K^+ or Na^+. The detection temperature was 25 °C. Spectra of chimeric oligonucleotides containing more rGs are shown in Figure S13.

CD and NMR experiments were also carried out for the F series in NH_4^+ (Figures S9 and S14). The F34 CD spectrum exhibits the 295 nm shoulder peak in NH_4^+ solution but not in Na^+ (Figures S13B and S14B; data reconstructed in Figure S15A). Interestingly, the imino region NMR spectra of the same oligonucleotide exhibited eight peaks in NH_4^+ but only four in Na^+ (Figures S13B and S14B; data reconstructed in Figure S15B).

To summarize, all samples that have a positive CD peak around 295 nm showed an increased number of imino peaks in their NMR spectra, indicating the presence of several structures. Those oligonucleotides analyzed by ion mobility spectrometry also showed several conformers. This polymorphism is particularly favored when the rG residue is located at the second G-quartet from the 5′ end. Conversely, no shoulder peak and single topologies were found for all samples in which the first guanine is an rG (including oligonucleotides with single or multiple substitutions).

Two-Dimensional NMR of F2 Reveals That 5′ G-Quartet Nucleotides Adopt the *Syn* Conformation. Recently, the Galeone and Plavec groups demonstrated that the G-quartet at the 5′-end of $[d(TGGGT)]_4$ can adopt an all-*syn* form, although this population is minor.[31,32] The possibility of alternative tetrameric topologies involving a *syn* G-quartet was unexpected, because it contrasts to what is usually observed in parallel tetramolecular G-quadruplexes and with the fact that riboses favor *anti* conformations.[5−8,16,17,21,23−25] Adoption of the *syn* conformation would explain the presence of a shoulder around 295 nm in the CD spectra and extra imino peaks observed by NMR in the oligonucleotides in our series with rG in the second G-quartet.

To understand the role of the hydroxyl group in chimeric G-quadruplexes, we therefore carried out a detailed two-dimensional NMR study on the F2 G-quadruplex. This oligonucleotide was chosen because (i) its CD spectrum exhibits the clearest shoulder peak at 295 nm, (ii) ion mobility spectrometry demonstrated that two tetramolecular, four-quartet topologies coexist and have similar abundances, (iii) the number of imino protons suggests a mixture of at least two conformers, and (iv) the F2 substitution leads to a strong stabilizing effect (Table 2).

The NMR analysis indicates that two stable conformational states coexist for F2 (**A** and **B** in red and gray, respectively, in Figure 8). The sequence of F2 is TGgGGT; for all NMR studies, bases are numbered from 5′ to 3′, starting from the 5′ thymine T1. According to the relative NMR cross peak intensities, **A** and **B** are 45% and 55% of the population, respectively. For both conformers, we followed the two sets of sequential ^{1}H−^{1}H NOE connectivities of the aromatic H8/H6 protons with the sugar rings protons (H1′, H2′, and H2″), the H1 imino protons, and the methyl groups (Figures 8 and S16).

For conformer **B**, in the aromatic/anomeric (H8−H6/H1′) region of the NOESY spectrum, we were able to follow all the sequential steps from T1b at the 5′ end to T6b at the 3′ end. In the context of a classical G-quadruplex structure core, these kind of correlations are usually observable for either the 5′-*syn*-*anti*-3′ or 5′-*anti*-*anti*-3′ sequential steps but never for the 5′-*anti*-*syn*-3′ steps for which the distance between two successive H1′ and H8 exceeds 6 Å. The uninterrupted sequential walk suggests there is no 5′-*anti*-*syn*-3′ step in conformer **B**. Likewise, in the H8−H6/H2′-H2″ region, we can follow a complete sequential walk from T1b to T6b. The T1b-H2′-H2″/G2b-H8 sequential correlation indicates an extensive stacking of the first thymine on top of the G2b base. The strong intraresidue H2′-H2″/H8 correlations for G2b indicate a short-distance interaction between these two protons, suggesting an *anti* conformation for this residue. We draw the same conclusions from the well-resolved G3b-H2′/H8 and G5b-H2′-H2″/H8 peaks. The sequential correlations between G5b and T6b suggest an extensive stacking of T6b on top of G5b. This was confirmed in the H8−H6/H7 region by the presence of a cross peak between the H7 from the methyl group of the T6b and the H8 aromatic proton from G5b. All these correlations demonstrate that conformer **B** has the same topology as the TGGGGT structure with all guanines in *anti*.

In the case of conformer **A**, the chemical shifts of all protons (imino, aromatic, sugars, and methyl) from G4a, G5a, and T6a are strikingly close to those of conformer **B**, with a perfect superimposition for T6a and T6b. If these protons have the same chemical environment, it is likely that the two conformers share similar structural features for the G4G5T6-3′ stretch. Indeed, from the H1′ of G3a through T6a at the 3′ end, we can clearly follow a continuous sequential walk, again suggesting the absence of 5′-*anti*-*syn*-3′ steps for this part of conformer **A**.

In contrast, for the 5′-T1G2G3 stretch, the proton frequencies of the two conformers are very distinct. In the aromatic/anomeric region, we observe that G2a-H8/H1′ and G3a-H8/H1′ cross peak intensities were weaker than those of G4a-H8/H1′ and G5a-H8/H1′, whereas for conformer **B**, intensities were comparable for all the G residues. This difference suggests that the 5′ region of conformer **A** is more

DOI: 10.1021/jacs.7b00648
J. Am. Chem. Soc. 2017, 139, 7768−7779

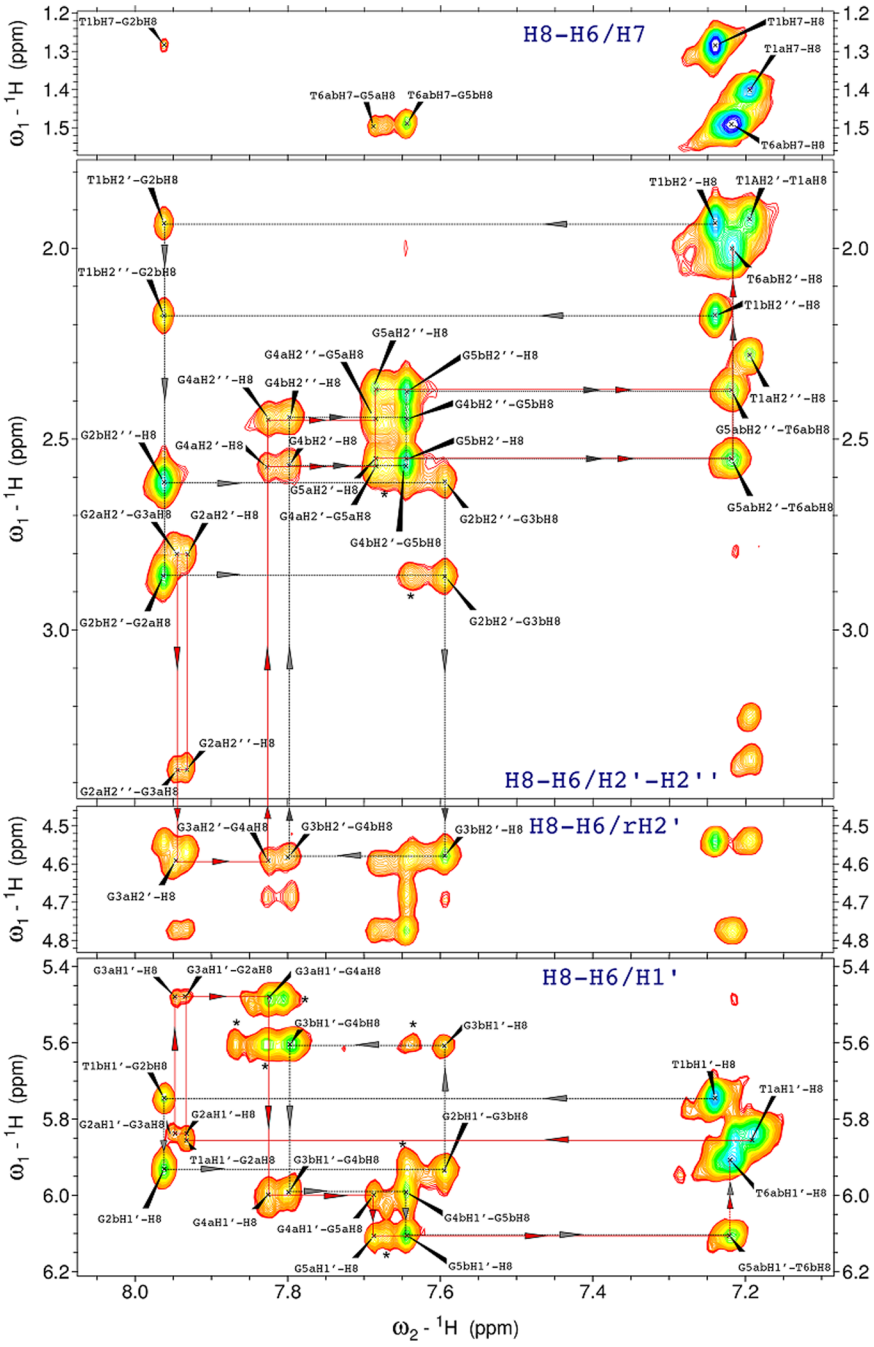


Figure 8. ^{1}H−^{1}H NOESY spectra (280 ms mixing time, T = 293 K) of F2 depicting NOE cross peaks from (ω_1) the imino (H1), sugar (H1′, H2′, H2″), methyl (H7), and (ω_2) guanine aromatic H8/thymine aromatic H6 regions. F2 was at a strand concentration of 2 mM and dissolved in 20 mM sodium phosphate buffer and 70 mM NaCl. Red lines and arrows correspond to the sequential walk of conformer **A**; gray dashed lines and arrows correspond to conformer **B**. Peaks from a third minor species are marked with a star (*).

dynamic than the 3′ region. Interestingly, we observed a squared connectivity pattern between G2aH8/H1′ and G3aH8/H1′, which is a convincing fingerprint of the 5′-*syn*-*anti*-3′ step as previously noted,[53] with G2a adopting a *syn* conformation and G3a an *anti* conformation. However, we do not observe a clear difference in intensity between G2aH8 and G3aH8. Likely, the dynamic properties of this region of the structure might be the reasons why we do not observe a strong signal for G2aH8/H1′ cross peak despite the *syn* conformation.

To confirm this result, we performed a ^{1}H−^{13}C HSQC experiment at natural abundance focused on the H1′C1′ region. G2a-C1′ resonated at around 87 ppm, while all the other C1′ signals from the guanines of conformers **A** and **B** resonated between 81 and 84 ppm (Figure S17A). This clear 3−6 ppm upfield shift again argues for a *syn* conformation of G2a base as suggested by Greene et al.[54] The sequential connectivity between T1a-H1′ and G2a-H8 are indicative of extensive stacking between the two residues. This was confirmed by analysis of the aromatic/imino region (H8−H6/H1): the T1a aromatic H6 provides a cross peak with G2a H1 imino (Figure S16B). In contrast with conformer **B**, conformer **A** lacks sequential connectivity between T1a-H2′-H2″ protons and G2a-H8. The fact that T1a and G2a clearly stack but lack proton connectivity in this region confirm the *syn*

DOI: 10.1021/jacs.7b00648
J. Am. Chem. Soc. 2017, 139, 7768−7779

conformation of G2a as the distance between T1a-H2′-H2″ and G2a-H8 exceeds 6 Å in this case.

We then performed a molecular dynamics modeling of the G-quadruplex formed by F2 to gain some insight into the effect of 2′-OH on the G-quadruplex structures. Two structures (called "I" and "II" hereafter) of F2 were generated and are schematically presented in Figure 9. In conformer I, all four G-

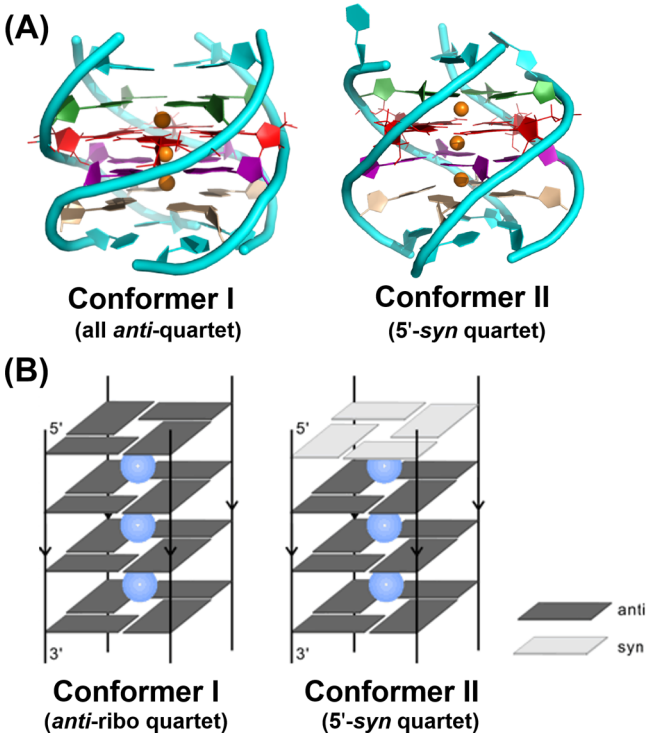


Figure 9. (A) Possible structures formed by F2 (rGs are shown in red) after 450 ns molecular dynamics modeling. In conformer I, all the G-quartet residues are in the same topology (configuration). In conformer II, the residues in the 5′-most G-quartet are flipped into the *syn* conformation at the glycosidic bond. (B) Schematic structures of two conformations.

quartets have the same topology, with the guanosines in *anti*-conformation (conformer **B** in 2D NMR). In conformer II, the 5′-end quartet residues are in the *syn* conformation and nucleotides in the other three quartets are in the *anti* conformation (conformer **A** in 2D NMR). Both conformers are very stable, as their energies do not change significantly in the molecular dynamics simulation (Figure S18), which may be the reason for similar thermal stabilities we observe as described in the next section.

Thermal Stability of Chimeric Tetramolecular G-Quadruplexes. In order to evaluate the thermal stabilities of the G-quadruplex structures, we performed CD and UV melting experiments. The apparent melting temperatures ($T_{1/2}$) in 100 mM K^+ or Na^+ (T and F series, respectively) are summarized in Table 2, and the melting profiles are shown in Figures S19 and S20. It should be noted that, due to experimental constraints, the CD and UV melting rates were different (0.25 °C/min and 0.5 °C/min, respectively). As we and others have demonstrated that the apparent melting temperature of a tetramolecular G-quadruplex depends on the rate of heating,[14,51] the $T_{1/2}$ values of the same samples are therefore slightly different. To circumvent this problem, we plotted the difference between $T_{1/2}$ for the modified samples and their all-DNA controls T0 and F0 (Figure 10); these $\Delta T_{1/2}$ values obtained by CD and UV are in reasonable agreement.

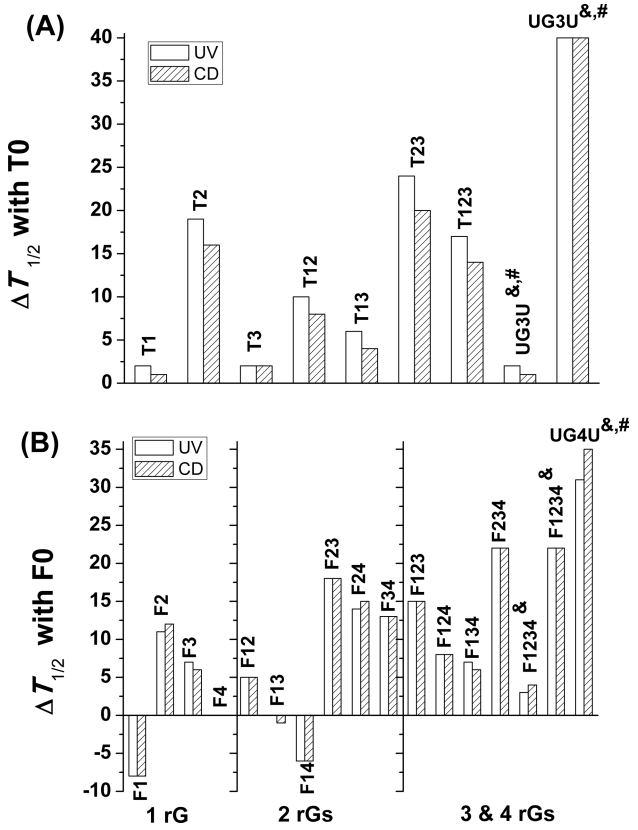


Figure 10. Effect of rG substitutions on thermal stabilities for the (A) T and (B) F series. $\Delta T_{1/2}$ values are provided compared to the all-DNA sequences d(TGGGT) (T0) and d(TGGGGT) (F0) and were determined by UV absorbance (white bars) or CD (hatched bars) melting experiments. [&]The melting curve shows two transitions; $T_{1/2}$ of both transitions are shown here. [#]These values are estimated, as we did not observe the fully unfolded state at high temperature.

Compared to T0 and F0, the corresponding all-RNA structures UG3U and UG4U are very stable ($T_{1/2} \geq 90$ °C), as reported previously.[14] In mixed DNA/RNA constructs, we observed a significant position-dependent effect of rG substitution on stability (details in Supporting Information), in line with the results described for intramolecular chimeras. Similar observations were made upon analysis of the melting data of F series oligonucleotides in NH_4^+ solution (Figure S21), revealing a common trend that can be generalized.

These results highlight interesting position-dependent effects of the presence of rG residues. The hydroxyl group *destabilizes* the canonical all-parallel tetramolecular G-quadruplex when present in the 5′-most quartet. Yet, the canonical topology remains predominant. When the second G-quartet is ribose, we observed formation of a second conformer with a 5′-*syn* DNA quartet in addition to the canonical all-*anti* G-quadruplex.

We tried to resolve the stability difference between these two main conformations by performing CD melting at 260 and 295 nm (the shoulder peak arises only in the presence of *syn* quartets, hence only for conformer **B**). We observed no wavelength-dependent differences in $T_{1/2}$ for F2, F23, F24, or F234, implying that the coexisting conformations have very similar thermal stabilities (Figure 11). These data suggest that

DOI: 10.1021/jacs.7b00648
J. Am. Chem. Soc. 2017, 139, 7768−7779

the presence of a ribose G-quartet at the second position stabilizes both conformers.

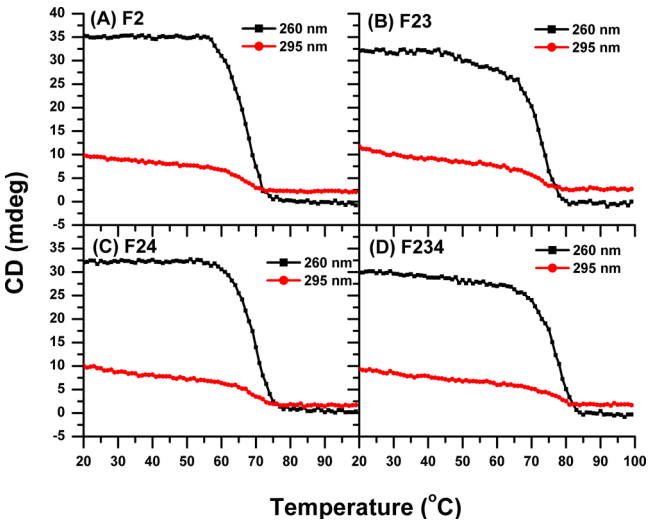


Figure 11. CD melting profiles of (A) F2, (B) F23, (C) F24, and (D) F234 at 260 nm (black) or 295 nm (red) in 10 mM cacodylate buffer (pH 7.0) in the presence of 100 mM Na^+.

DISCUSSION

Here we investigated how ribose G-quartets affect the structures and stabilities of G-quadruplexes. dG and rG residues are naturally occurring bases, and the chimeras formed by dG/rG can be regarded as modifications of native DNA or RNA sequences. DNA/RNA chimeras occur during DNA replication of the lagging strand, and individual ribonucleotides may be found in DNA. Chemically modified G-quadruplexes are finding applications in a variety of fields, as discussed in several recent excellent reviews.[40,55−57] With a few exceptions,[58−60] the guanines in parallel G-quadruplex structures are almost exclusively in the *anti* glycosidic bond conformation. *Anti* glycosidic angles are favored by ribose nucleotides. The topologies of the unimolecular antiparallel type II thrombin binding aptamer (TBA) and hybrid type III DNA human telomere sequences are converted to parallel G-quadruplex conformations by selective replacement of *syn*-dG positions by rG.[25,61] A recent work showed that rG substitutions at *anti*-dG positions left the structure unaltered, suggesting a reason why RNA G-quadruplexes prefer parallel topologies.[62] Here we performed systematic substitutions of the G-quartets in a DNA oligonucleotide that adopts an intramolecular all-parallel, all-*anti* conformation G-quadruplex. As expected, the intramolecular parallel topology was unaltered. However, the situation was more complex for all-parallel tetramolecular complexes: rG substitutions caused some bases to flip from *anti* to *syn*. Interestingly, upon incorporation of the bulky bromine substituent in the C-8 position of guanine (8-bromoguanosine, ^{Br}G), which favors the *syn* conformation, an exceptional *syn*-quartet was observed in the intramolecular parallel G-quadruplex.[63] These results indicate that G-quadruplexes can maintain a parallel orientation when one G-quartet flips into *syn*.[31,32,50,58−60,63]

RNA G-quadruplexes are generally more stable than their DNA counterparts. For example, for the antiparallel, type II TBA and hybrid type III human telomeric G-quadruplexes, inserting rG at positions with *anti* glycosidic conformations led to greater thermal stability.[25,61] The presence of 2′-OH's has been shown in several studies to increase the stability of G-quadruplexes.[8,15−17,62] A possible reason is an extra O···H−C hydrogen bond between the 2′-OH and the H8 atom of the guanine immediately 3′ of the rG.[62]

In contrast with all previous reports, we found a *destabilization* effect of 2′-OH's when located on the 5′-terminal G-quartet (RQ1, T1, and F1), although the overall topology was maintained. This position-dependence was unexpected, as all guanines are usually all in the *anti* conformation for all-DNA intramolecular and tetramolecular complexes. Therefore, previous explanations for position-dependent effects of dG to rG substitution reported for intramolecular G-quadruplexes cannot apply here.[25,61,62] Moreover, these effects are cumulative.

G-quadruplex stability is affected by multiple forces, such as stacking interactions, hydrogen bonding, solvation, loop length, and cation binding.[64] We explored solvation effects for intramolecular G-quadruplexes (Figure 4) and found that hydration does not account for the order of stability we observe. Hydrogen bonding may be the reason for ribose G-quadruplex stabilization.[21,62] Determining whether this is the case will require a high-resolution structure of our system, an endeavor beyond the scope of the current paper.

The preferred conformation of the ribose sugar is C3′-endo; however, riboses can adopt C2′-endo sugar puckers in RNA G-quadruplexes.[20,21,23,24,54] The unfavorable conformation of riboses may confer certain energy penalties on RNA G-quadruplex structures.[21,23,24,54] Recently, Petersen et al. demonstrated that sugar rings do not participate in G-quartet formation, and that sugar conformation is governed by the arrangement of the sugar−phosphate backbone in the G-quadruplex.[65] The *destabilizing* effect of the ribose G-quartet that we observed at certain positions within otherwise deoxy G-quadruplexes implies that this sugar pucker is not always well accommodated.

CONCLUSIONS

The chimeric DNA/RNA oligonucleotides studied here, in which one or more G-quartets were ribose, formed well-defined parallel G-quadruplexes. Strong position-dependent and cumulative effects of 2′-OH on stability were observed. Interestingly, a *destabilization* effect was observed when 2′-OH groups were introduced in the 5′-most G-quartet (RQ1, T1, and F1), demonstrating that the assumption that ribose residues stabilize G-quadruplexes is an oversimplification. Furthermore, hydration measurements of intramolecular chimeric G-quadruplex structures indicated that hydration was not the sole factor affecting the stability of RNA G-quadruplexes.[13,30−32] Moreover, introducing a ribose in the second quartet from the 5′-end of tetramolecular G-quadruplexes stabilized the tetramer, but unexpectedly induced the formation of a second structural ensemble with the 5′-terminal quartet in *syn*. Overall, our results on intra- and tetramolecular G-quadruplexes broaden our understanding of the roles of 2′-OHs on RNA G-quadruplexes and highlight that the positions of ribose residues are critical to the properties of RNA G-quadruplexes.

MATERIALS AND METHODS

Sample Preparation. All desalted oligonucleotides listed in Tables 1 and 2 were purchased from Integrated DNA Technologies without

DOI: 10.1021/jacs.7b00648
J. Am. Chem. Soc. 2017, 139, 7768−7779

further purification. Concentrations were determined by UV absorption at 260 nm using the extinction coefficients provided by the manufacturer. All the chemicals used were purchased from Sigma unless otherwise stated.

Intramolecular RQ Series Samples. The samples were heated to 95 °C for 5 min and cooled to 20 °C at a rate of 0.5 °C/min before use unless otherwise stated. The concentration of each strand was 5 μM for CD and UV experiments, and the buffer was 10 mM cacodylic acid buffer (pH 7.0) containing 100 mM K^+, Na^+, or NH_4^+. For melting experiments in K^+, potassium concentration was 5 mM, as some samples are not fully unfolded at 100 °C in 100 mM K^+. For 1D NMR experiments, the concentration of each strand was 50 μM, and the buffer was 10 mM KH_2PO_4/K_2HPO_4 (pH 6.9) containing 100 mM K^+. For SE-HPLC and ESI-MS experiments, the samples were prepared in 100 mM ammonium acetate solution (pH ≈ 7) (BioUltra, Sigma-Aldrich).

Intermolecular T and F Series Samples. In order to guarantee that all the samples adopted a major structure in solution, the samples were prepared at high concentrations of oligonucleotides and salts. The concentrated samples (typically 100 μM strand concentration) were heated to 95 °C for 5 min, cooled slowly to room temperature, and then stored at 4 °C for at least 2 weeks in the presence of 100 mM Na^+, K^+, or NH_4^+ before use unless otherwise stated. For UV and CD experiments, the stock solutions were diluted to 10 μM with cacodylic acid buffer (pH adjusted to 7.0 with LiOH) supplemented with the desired concentration of Na^+, K^+ or NH_4^+. For 1D NMR experiments, samples were prepared at a concentration of 50 μM for F series and 100 μM for T series oligonucleotides. The buffers were 20 mM NaH_2PO_4/Na_2HPO_4 (pH 6.8), KH_2PO_4/K_2HPO_4 (pH 6.9), or $NH_4H_2PO_4/(NH_4)_2HPO_4$ (pH 7.0) for F series and T series, respectively, and also contained 100 mM Na^+, K^+, or NH_4^+ as indicated (H_2O/D_2O 9:1 v/v). 2D NMR samples of F2 were prepared at a concentration of 1−3 mM in 20 mM NaH_2PO_4/Na_2HPO_4 buffer (pH 7.0) and 70 mM NaCl.

Circular Dichroism Spectroscopy. CD spectra were recorded on a Jasco J-815 instrument equipped with a Peltier temperature control accessory. Each spectrum was obtained by averaging five scans recorded at a speed of 100 nm/min at 20 °C. A background CD spectrum of corresponding buffer solution was subtracted from the average scan for each sample. CD melting curves of tetramolecular G-quadruplexes were obtained by measuring the CD spectra of samples from low temperature to high temperature with 1 °C increments. The ellipticity at 260 nm was then plotted against temperature. It should be noted that collecting data at each temperature requires about 3 min: taking into account the time required to raise temperature by 1 °C, the average temperature gradient was ∼0.25 °C/min. Nearly identical protocols were used for all samples, as the temperature gradient affects the apparent melting temperature of tetramolecular G-quadruplexes.[14,50]

Ultraviolet Melting Curves, Thermal Difference Spectra, and Thermodynamic Analysis. UV-absorbance measurements were performed on a Cary100 UV−vis spectrophotometer (Agilent Technologies) equipped with a Peltier temperature control accessory or on a Uvikon XS (Secomam) spectrophotometer and an external circulating water bath. An inert glass sensor was immersed into the quartz cell. The absorbance was monitored at 295 nm with a rate of 0.5 °C/min as previously described,[42] and experiments were repeated at least three times. For each sample, the TDS was obtained by subtracting the absorbance spectrum at 20 °C from the one at 95 or 99 °C as previously described.[37]

The thermodynamic parameters listed in Tables S1 and S2 were calculated by fitting the melting profiles to a theoretical equation for an intramolecular association as described previously.[42] Small amounts of dimer were formed in some cases, so $T_{1/2}$ values are reported.

Electrospray Ionization Mass Spectrometry (ESI-MS) and Ion Mobility Spectrometry (IMS). For ammonium ion conditions, the samples were prepared in 100 mM ammonium acetate solution (BioUltra, Sigma-Aldrich) at pH 7.0. The strand concentration was 50 μM, and the oligonucleotides were left overnight at 4 °C to allow the complex formation. The samples were diluted to 5 μM with 100 mM NH_4OAc prior to injection. In K^+, the samples were prepared at 50 μM strand concentration in 100 mM trimethylammonium acetate (TMAA) and 1 mM KCl. The samples were left overnight at 4 °C. The oligonucleotides were diluted to 10 μM with a mixture of 100 mM TMAA/1 mM KCl prior to mass analysis. Mass spectrometry experiments were performed on the Thermo Exactive orbitrap instrument, tuned for soft ion transfer to avoid gas-phase dissociation. IMS coupled to ESI-MS experiments were performed on a Thermo Exactive orbitrap instrument, tuned for soft ion transfer to avoid gas-phase dissociation of the tetramolecular structures, and an Agilent 6560 DTIMS-Q-TOF spectrometer (Agilent Technologies) operated in negative ion mode. The sample preparation was the same as above. Ion mobility experiments were performed in helium (P = 3.89 Torr). The funnel's radio frequency amplitude was set to 180 V, and the trap entrance grid delta was set to 4 V. The data were analyzed using Agilent MassHunter software (B.07) and IM-MS Browser B.06.

Nuclear Magnetic Resonance Spectroscopy. NMR experiments were performed on a 700 or 800 MHz Bruker spectrometer equipped with TXI probes. 1D ^{1}H NMR experiments were acquired at 25 °C using a pulse sequence with spin−echo water suppression. To guarantee that all the samples adopted the thermodynamically favored major structure in solution, NMR studies were performed at least 2 weeks after sample preparation. The pulse program used to perform the ^{1}H-^{1}H NOESY experiments was noesyesgpphzs,[66] and dipsi2esgpph[67] was the pulse program for the ^{1}H-^{1}H TOCSY experiment, which was done with a 40 ms mixing time (Figure S17B). The ^{1}H-^{1}H COSY experiment was recorded using the pulse cosydfesgpph.[68] All the 2D experiment pulse programs employed excitation sculpting with gradients for water suppression.[69] Usually a spectral width of 2048 × 512 complex points in F2 and F1 dimensions, respectively, was used in 2D homonuclear experiments.

Size-Exclusion High Performance Liquid Chromatography. SE-HPLC experiments were performed on an UltiMate 3000 UHPLC system (Thermo Scientific Dionex) equipped with an autosampler, a diode array detector, and a Thermo Acclaim SEC-300 column as described recently.[41] The temperature of the column was set at 20 °C. Thorough equilibration (with at least 3 column volumes; 15 mL with our experimental setup) was performed before the first injection. Eluted species were monitored by absorbance measurements (210 to 600 nm; 1 nm bandwidth, 5-Hz data collection rate, 2-s response time), and the chromatograms (normalized to [0,1]) were plotted from the absorbance at 260 nm, using the relative elution volume V_e/V_0 as x-axis, where V_e is the elution volume and V_0 the dead volume (typically 1.86 mL).[41] For each peak, the nature of the species was confirmed by its characteristic absorbance spectrum.

The structure index (SI) of a given peak is an arbitrary score, calculated as follows: $SI = V_e/V_0 \times \log_{10}(MW_{strand})$, where V_e/V_0 is the relative elution volume and MW_{strand} the molecular weight of the oligonucleotide. The reference values for monomeric and dimeric G-quadruplex structures, noted SI_{ref}, are calculated from the chromatograms of 28 and 12 sequences, respectively, taken from ref 41.

Water Activity Measurements and Calculation of Water Molecules Released upon Formation of G-Quadruplex. The water activity was determined by the osmotic stressing method via vapor phase osmometry using a Wescor 5520XR pressure osmometer. To calculate the number of water molecules released upon formation of chimeric G-quadruplex, we used the methods described previously.[16,22,43]

Molecular Modeling. Using the biopolymer building tool of Discover, the initial chimeric G-quadruplexes formed by F2 were generated by modifying the G-quadruplex [d(TGGGGT)4] (PDB code: 2O4F).[70] The first thymine nucleotide was deleted and the second dG mutated to rG in each of the four strands. Three internal sodium ions were positioned in the central channel. Following these steps, we obtained the G-quadruplex structure in which all the glycosidic conformations were in *anti*-form. Then, the bases of the rGs were rotated 180° around the glycosidic bonds to yield another G-quadruplex structure in which the rGs were in *syn*-form. After the two initial G-quadruplex structures were built, two 450 ns length MD simulations were performed to study their stability with Gromacs

DOI: 10.1021/jacs.7b00648
J. Am. Chem. Soc. 2017, 139, 7768−7779

(version 4.5.5) using force field amber99sb_parmbsc0.[71−74] The initial structures were solvated in a periodic box with explicit TIP3P water molecules, and the Na^+ and Cl^- ions were added to neutralize the system and maintain a salt concentration of 100 mM. The time step was 2 fs as all the bond lengths are constrained with LINCS algorithm. The particle mesh Ewald (PME) method, which calculates the electrostatic potential energy, was used to treat the electrostatic interactions with a cutoff of 1.0 nm. The cutoff of nonbonded van der Waals interactions was also set to 1.0 nm. The two systems were first optimized by a 1000-step energy minimization, followed by a gradual increase of temperature to 300 K. After that 2 ns MD runs were performed with NPT ensemble at 1 atm and 300 K to equilibrate the systems. The conformations at the end of the 2 ns runs were used as the initial structures for the 450 ns length simulations.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/jacs.7b00648.

Additional tables and figures (PDF)

■ AUTHOR INFORMATION

Corresponding Authors

*jun.zhou@nju.edu.cn
*wangwei@nju.edu.cn
*jean-louis.mergny@inserm.fr

ORCID

Dursun Nizam Korkut: 0000-0001-9905-5762
Jian Zhang: 0000-0003-2338-0395
Naoki Sugimoto: 0000-0003-2928-7028
Valérie Gabelica: 0000-0001-9496-0165
Jean-Louis Mergny: 0000-0003-3043-8401

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

This work was supported in part by National Natural Science Foundation of China (No. 21503229 to J.Z.; No. 11334004 to W.W.), Fundamental Research Funds for the Central Universities (No. 020514380070 to J.Z.), and National Basic Research Program of China (973 Program) (2013CB834100 to W.W.), self-project fund by State Key Laboratory of Analytical Chemistry for Life Science, Nanjing University (to J.Z.), Grants-in-Aid for Scientific Research and MEXT-Supported Program for the Strategic Research Foundation at Private Universities (2014−2019), the Hirao Taro Foundation of KONAN GAKUEN for Academic Research, Japan, the Inserm (ATIP-Avenir Grant no. R12086GS to V.G.), the *Conseil Régional d'Aquitaine* (Grant no. 20121304005 to V.G.), and the EU (FP7-PEOPLE-2012-CIG-333611 to V.G.). This work has benefited from the facilities and expertises of the Biophysical and Structural Chemistry Platform (BPCS) at IECB, CNRS UMS3033, Inserm US001, Bordeaux University. J.-L.M. acknowledges funding from *Conseil Régional d'Aquitaine* and *Agence Nationale de la Recherche* (OligoSwitch [ANR-12-IS07-0001], "Quarpdiems" [ANR-12-BSV8-0008-01], and "VIBBnano" [ANR-10-NANO-04-03]) as well as the SYMBIT project (reg. no. CZ.02.1.01/0.0/0.0/15_003/0000477) financed by the ERDF. J.-L.M. dedicates this article to the memory of Luc Cappa (1959−2016).

■ REFERENCES

(1) Sen, D.; Gilbert, W. *Nature* **1988**, *334*, 364.
(2) Williamson, J. R. *Annu. Rev. Biophys. Biomol. Struct.* **1994**, *23*, 703.
(3) Huppert, J. L. *Chem. Soc. Rev.* **2008**, *37*, 1375.
(4) Oganesian, L.; Bryan, T. M. *BioEssays* **2007**, *29*, 155.
(5) Shafer, R. H.; Smirnov, I. *Biopolymers* **2000**, *56*, 209.
(6) Xu, Y. *Chem. Soc. Rev.* **2011**, *40*, 2719.
(7) Phan, A. T. *FEBS J.* **2010**, *277*, 1107.
(8) Collie, G. W.; Parkinson, G. N. *Chem. Soc. Rev.* **2011**, *40*, 5867.
(9) Bugaut, A.; Balasubramanian, S. *Nucleic Acids Res.* **2012**, *40*, 4727.
(10) Kumari, S.; Bugaut, A.; Huppert, J. L.; Balasubramanian, S. *Nat. Chem. Biol.* **2007**, *3*, 218.
(11) Huppert, J. L.; Bugaut, A.; Kumari, S.; Balasubramanian, S. *Nucleic Acids Res.* **2008**, *36*, 6260.
(12) Wieland, M.; Hartig, J. S. *Nat. Protoc.* **2009**, *4*, 1632.
(13) Halder, K.; Wieland, M.; Hartig, J. S. *Nucleic Acids Res.* **2009**, *37*, 6811.
(14) Mergny, J. L.; De Cian, A.; Ghelab, A.; Saccá, B.; Lacroix, L. *Nucleic Acids Res.* **2005**, *33*, 81.
(15) Saccá, B.; Lacroix, L.; Mergny, J. L. *Nucleic Acids Res.* **2005**, *33*, 1182.
(16) Arora, A.; Maiti, S. *J. Phys. Chem. B* **2009**, *113*, 10515.
(17) Joachimi, A.; Benz, A.; Hartig, J. S. *Bioorg. Med. Chem.* **2009**, *17*, 6811.
(18) Cheong, C.; Moore, P. B. *Biochemistry* **1992**, *31*, 8406.
(19) Deng, J.; Xiong, Y.; Sundaralingam, M. *Proc. Natl. Acad. Sci. U. S. A.* **2001**, *98*, 13665.
(20) Xu, Y.; Ishizuka, T.; Kimura, T.; Komiyama, M. *J. Am. Chem. Soc.* **2010**, *132*, 7231.
(21) Collie, G. W.; Haider, S. M.; Neidle, S.; Parkinson, G. N. *Nucleic Acids Res.* **2010**, *38*, 5569.
(22) Miyoshi, D.; Karimata, H.; Sugimoto, N. *J. Am. Chem. Soc.* **2006**, *128*, 7957.
(23) Xu, Y.; Kaminaga, K.; Komiyama, M. *J. Am. Chem. Soc.* **2008**, *130*, 11179.
(24) Martadinata, H.; Phan, A. T. *J. Am. Chem. Soc.* **2009**, *131*, 2570.
(25) Tang, C. F.; Shafer, R. H. *J. Am. Chem. Soc.* **2006**, *128*, 5966.
(26) Smargiasso, N.; Rosu, F.; Hsia, W.; Colson, P.; Baker, E. S.; Bowers, M. T.; De Pauw, E.; Gabelica, V. *J. Am. Chem. Soc.* **2008**, *130*, 10208.
(27) Guédin, A.; Gros, J.; Alberti, P.; Mergny, J. L. *Nucleic Acids Res.* **2010**, *38*, 7858.
(28) Largy, E.; Marchand, A.; Amrane, S.; Gabelica, V.; Mergny, J. L. *J. Am. Chem. Soc.* **2016**, *138*, 2780.
(29) Aboul-ela, F.; Murchie, A. I. H.; Lilley, D. M. J. *Nature* **1992**, *360*, 280.
(30) Aboul-ela, F.; Murchie, A. I. H.; Norman, D. G.; Lilley, D. M. J. *J. Mol. Biol.* **1994**, *243*, 458.
(31) Šket, P.; Virgilio, A.; Esposito, V.; Galeone, A.; Plavec, J. *Nucleic Acids Res.* **2012**, *40*, 11047.
(32) Esposito, V.; Randazzo, A.; Piccialli, G.; Petraccone, L.; Giancola, C.; Mayol, L. *Org. Biomol. Chem.* **2004**, *2*, 313.
(33) Bedrat, A.; Lacroix, L.; Mergny, J. L. *Nucleic Acids Res.* **2016**, *44*, 1746.
(34) Randazzo, A.; Spada, G. P.; da Silva, M. W. *Top. Curr. Chem.* **2012**, *330*, 67.
(35) Paramasivan, S.; Rujan, I.; Bolton, P. H. *Methods* **2007**, *43*, 324.
(36) Kypr, J.; Kejnovská, I.; Renčiuk, D.; Vorlíčková, M. *Nucleic Acids Res.* **2009**, *37*, 1713.
(37) Mergny, J.-L.; Li, J.; Lacroix, L.; Amrane, S.; Chaires, J. B. *Nucleic Acids Res.* **2005**, *33*, e138.
(38) Feigon, J.; Koshlap, K. M.; Smith, F. W. *Methods Enzymol.* **1995**, *261*, 225.
(39) Rosu, F.; Gabelica, V.; Houssier, C.; Colson, P.; De Pauw, E. *Rapid Commun. Mass Spectrom.* **2002**, *16*, 1729.
(40) Zhou, J.; Rosu, F.; Amrane, S.; Korkut, D. N.; Gabelica, V.; Mergny, J. L. *Methods* **2014**, *67*, 159.
(41) Largy, E.; Mergny, J. L. *Nucleic Acids Res.* **2014**, *42*, e149.
(42) Mergny, J. L.; Phan, A. T.; Lacroix, L. *FEBS Lett.* **1998**, *435*, 74.

DOI: 10.1021/jacs.7b00648
J. Am. Chem. Soc. 2017, 139, 7768−7779

(43) Nakano, S.; Karimata, H.; Ohmichi, T.; Kawakami, J.; Sugimoto, N. *J. Am. Chem. Soc.* **2004**, *126*, 14330.
(44) Miyoshi, D.; Nakamura, K.; Tateishi-Karimata, H.; Ohmichi, T.; Sugimoto, N. *J. Am. Chem. Soc.* **2009**, *131*, 3522.
(45) Miller, M. C.; Buscaglia, R.; Chaires, J. B.; Lane, A. N.; Trent, J. O. *J. Am. Chem. Soc.* **2010**, *132*, 17105.
(46) Macaya, R.F.; Schultze, P.; Smith, F. W.; Roe, J. A.; Feigon, J. *Proc. Natl. Acad. Sci. U. S. A.* **1993**, *90*, 3745.
(47) Hud, N. V.; Smith, F. W.; Anet, F. A.; Feigon, J. *Biochemistry* **1996**, *35*, 15383.
(48) Kankia, B. I.; Marky, L. A. *J. Am. Chem. Soc.* **2001**, *123*, 10799.
(49) Nakano, S.; Miyoshi, D.; Sugimoto, N. *Chem. Rev.* **2014**, *114*, 2733.
(50) Masiero, S.; Trotta, R.; Pieraccini, S.; De Tito, S.; Perone, R.; Randazzo, A.; Spada, G. P. *Org. Biomol. Chem.* **2010**, *8*, 2683.
(51) Marchand, A.; Gabelica, V. *J. Am. Soc. Mass Spectrom.* **2014**, *25*, 1146.
(52) Merkina, E. E.; Fox, K. R. *Biophys. J.* **2005**, *89*, 365.
(53) Adrian, M.; Heddi, B.; Phan, A. T. *Methods* **2012**, *57*, 11.
(54) Greene, K. L.; Wang, Y.; Live, D. *J. Biomol. NMR* **1995**, *5*, 333.
(55) Avino, A.; Fabrega, A. C.; Tintore, M.; Eritja, R. *Curr. Pharm. Des.* **2012**, *18*, 2036.
(56) Doluca, O.; Withers, J. M.; Filichev, V. V. *Chem. Rev.* **2013**, *113*, 3044.
(57) Sagi, J. *J. Biomol. Struct. Dyn.* **2014**, *32*, 477.
(58) Zhou, J.; Murayama, K.; Amrane, S.; Rosu, F.; Kashida, H.; Bourdoncle, A.; Asanuma, H.; Mergny, J. L. *Chem. Sci.* **2013**, *4*, 3693.
(59) Esposito, V.; Virgilio, A.; Randazzo, A.; Galeone, A.; Mayol, L. *Chem. Commun.* **2005**, *31*, 3953.
(60) Esposito, V.; Virgilio, A.; Pepe, A.; Oliviero, G.; Mayol, L.; Galeone, A. *Bioorg. Med. Chem.* **2009**, *17*, 1997.
(61) Qi, J.; Shafer, R. H. *Biochemistry* **2007**, *46*, 7599.
(62) Dickerhoff, J.; Appel, B.; Muller, S.; Weisz, K. *Angew. Chem., Int. Ed.* **2016**, *55*, 15162.
(63) Karg, B.; Haase, L.; Funke, A.; Dickerhoff, J.; Weisz, K. *Biochemistry* **2016**, *55*, 6949.
(64) Lane, A. N.; Chaires, J. B.; Gray, R. D.; Trent, J. O. *Nucleic Acids Res.* **2008**, *36*, 5482.
(65) Pradhan, D.; Hansen, L. H.; Vester, B.; Petersen, M. *Chem. - Eur. J.* **2011**, *17*, 2405.
(66) Thrippleton, M. J.; Keeler, J. *Angew. Chem., Int. Ed.* **2003**, *42*, 3938.
(67) Shaka, A. J.; Lee, C. J.; Pines, A. *J. Magn. Reson.* **1988**, *77*, 274.
(68) Howe, P. W. *Magn. Reson. Chem.* **2014**, *52*, 329.
(69) Hwang, T. L.; Shaka, A. J. *J. Magn. Reson., Ser. A* **1995**, *112*, 275.
(70) Creze, C.; Rinaldi, B.; Haser, R.; Bouvet, P.; Gouet, P. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2007**, *63*, 682.
(71) Pagano, B.; Mattia, C. A.; Cavallo, L.; Uesugi, S.; Giancola, C.; Fraternali, F. *J. Phys. Chem. B* **2008**, *112*, 12115.
(72) Hess, B.; Kutzner, C.; van der Spoel, D.; Lindahl, E. *J. Chem. Theory Comput.* **2008**, *4*, 435.
(73) Guy, A. T.; Piggot, T. J.; Khalid, S. *Biophys. J.* **2012**, *103*, 1028.
(74) Pérez, A.; Marchán, I.; Svozil, D.; Sponer, J.; Cheatham, T. E., III; Laughton, C. A.; Orozco, M. *Biophys. J.* **2007**, *92*, 3817.

DOI: 10.1021/jacs.7b00648
J. Am. Chem. Soc. 2017, 139, 7768−7779