26–38 *Nucleic Acids Research, 2017, Vol. 45, No. 1* *Published online 24 November 2016*
doi: 10.1093/nar/gkw1049

Assembly of supramolecular DNA complexes containing both G-quadruplexes and i-motifs by enhancing the G-repeat-bearing capacity of i-motifs

Yanwei Cao[1], Shang Gao[1], Yuting Yan[1], Michael F. Bruist[2,*], Bing Wang[1] and Xinhua Guo[1,*]

[1]College of Chemistry, Jilin University, 2699 Qianjin Street, Changchun 130012, P. R. China and [2]Department of Chemistry & Biochemistry, University of the Sciences, 600 South 43rd Street, Philadelphia, PA 19104, USA

Received September 6, 2016; Revised October 17, 2016; Editorial Decision October 18, 2016; Accepted October 20, 2016

ABSTRACT

The single-step assembly of supramolecular complexes containing both i-motifs and G-quadruplexes (G4s) is demonstrated. This can be achieved because the formation of four-stranded i-motifs appears to be little affected by certain terminal residues: a five-cytosine tetrameric i-motif can bear ten-base flanking residues. However, things become complex when different lengths of guanine-repeats are added at the 3′ or 5′ ends of the cytosine-repeats. Here, a series of oligomers d(XG_iXC_5X) and d(XC_5XG_iX) (X = A, T or none; i < 5) are designed to study the impact of G-repeats on the formation of tetrameric i-motifs. Our data demonstrate that tetramolecular i-motif structure can tolerate specific flanking G-repeats. Assemblies of these oligonucleotides are polymorphic, but may be controlled by solution pH and counter ion species. Importantly, we find that the sequences d(TG_iAC_5) can form the tetrameric i-motif in large quantities. This leads to the design of two oligonucleotides d(TG_4AC_7) and d($TG^{Br}GG^{Br}GAC_7$) that self-assemble to form quadruplex supramolecules under certain conditions. d(TG_4AC_7) forms supramolecules under acidic conditions in the presence of K^+ that are mainly V-shaped or ring-like containing parallel G4s and antiparallel i-motifs. d($TG^{Br}GG^{Br}GAC_7$) forms long linear quadruplex wires under acidic conditions in the presence of Na^+ that consist of both antiparallel G4s and i-motifs.

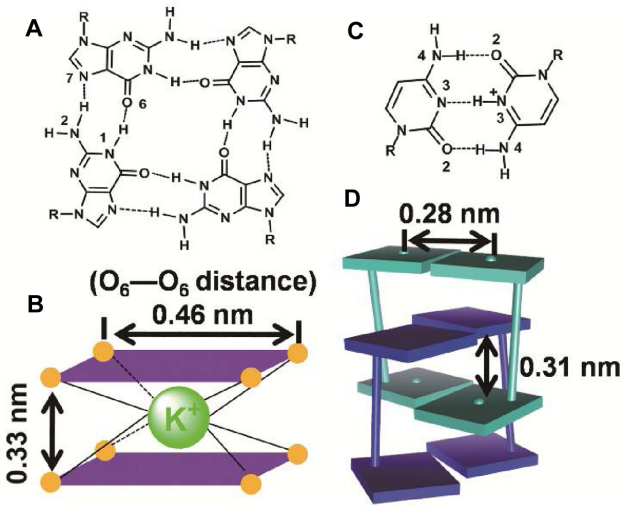
INTRODUCTION

The G-rich oligonucleotide sequences containing tandem G-stretches can fold into tetra-stranded DNA structures called G-quadruplexes (G4s) that may play an important role in biological processes such as immunoglobulin gene rearrangements, promoter activation and telomere maintenance (1). Many factors such as sequence, cationic species and concentration, small molecule ligands, molecular crowding and more affect the stability of G4s and induce the formation of polymorphs (2–4). Furthermore, the existence of complementary C-rich strands has significant impact on the G4 DNA formation (5) and under slightly acidic conditions such C-rich sequences can form their own tetra-stranded non-canonical secondary structure known as an i-motif (6). Such C-rich strands with the potential to fabricate i-motif scaffolds are frequent in the human genome (7) and may also have crucial importance in biological functions (8).

In a G4, Hoogsteen hydrogen bonds connect guanines within quartets, which stack with mono- or divalent cations in the central channel to enhance its stability (Scheme 1A and B) (9–11). For unimolecular and bimolecular G4s the paths of various loops connecting stacked G's above grooves must also be considered. In contrast, an i-motif is made up of two parallel duplexes held together by hemiprotonated cytosine-cytosine ($C{\bullet}CH^+$) pairs (Scheme 1C and D) (12,13). These duplexes are alternatively intercalated in an antiparallel orientation to form a stable and long-lived four-stranded nucleic acid secondary structure that is stabilized by π–π and electrostatic interactions (6,8,14). Compared to antiparallel B-form Watson–Crick (WC) duplex structures, quadruplex structures may be more flexible and functional due to their richer base stacking and abundance of ligand binding sites (8,9).

The simple fact that sequential runs of G's or C's occur more frequently in the genome than expected by random chance, indicates that they are likely to have biological significance (15). Previous studies demonstrate that poly[d(G + C)] oligonucleotides can easily form different types of duplexes, for example, B-form, A-form or Z-form (16,17) and sometime even quadruplexes (18–20). In addition, studies have revealed that the C-runs or cytosine residues in (G + C)-rich oligonucleotides significantly affect the formation (21) and stability of G4s (22,23). Mao and his research

*To whom correspondence should be addressed. Tel: +86 431 8922 8949; Fax: +86 431 8922 8949; Email: guoxh@jlu.edu.cn
Correspondence may also be addressed to Prof. Michael F. Bruist. Tel: +1 215 596 8530; Fax: +1 215 596 8543; Email: m.bruist@usciences.edu



Scheme 1. Schematic descriptions of (**A**) a G-quartet linked by Hoogsteen hydrogen bonds, (**B**) the chelation cage formed by eight O_6 atoms of guanines and K^+; (**C**) a C•CH^+ base pair, (**D**) two pairs of maximally intercalated C•CH^+ base pairs.

group show that for certain double-stranded sequences, the formation of G4s and i-motifs is mutually exclusive even under a favorable condition, which is mainly governed by steric hindrance in the double helix structure (24,25). Earlier studies also show that tetra-stranded i-motif structures formed by five consecutive cytosines have comparable stability to ten-base-pair-long duplexes (26). However, as yet, there are no systematic investigations into the influence of guanine residues or complementary G-runs on the formation of i-motif structures. Here, we show that in the right context this i-motif remains stable in the presence of a run of three or more G's; thus, intermolecular i-motif structures can outcompete intermolecular WC pairs with the right sequence and conditions.

DNA scaffolds with specific structural features and remarkable molecular recognition properties are one of the most promising materials for building nanoscale structures. Two principles for DNA macromolecular design have emerged over the past three decades: first, DNA secondary structures can serve as monomeric units; second, these monomers must have flexibility in order to further assemble into larger structures (27,28). The principles have allowed the assemblage of DNA supramolecular complexes into precise topological structures (29,30) and crossover DNA tiles (31,32). We believe that i-motif assemblies should be excellent construction modules to facilitate supramolecular DNA complexes due to their ability to assemble specifically and the flexibility of the motif (33). For example, oligonucleotide hybrid supramolecules composed of parallel or antiparallel duplexes and tetrameric i-motifs have been reported by Mei *et al.* (34). Others have shown that complementary DNA fragments (26) and C-rich repeats (35,36) located at the terminal of tetra-stranded i-motifs also act as flexible building blocks to allow the assembly supramolecular DNA structures via the formation of double helices and tetra-stranded i-motif structures. However, none of these systems have both C's and G's in the same oligonucleotide that each contribute to distinct tetraplexes. There is a special problem that G's and C's present in the same oligonucleotide may complement each other if the strand folds on itself or contribute to the formation of competing duplexes between strands. We postulate that with appropriate restrictions tetra-stranded i-motif structures can bear terminal G-repeats capable of folding into four-stranded G4s that contribute to the assembly of DNA nanowires containing both G4s and i-motifs.

In this work, short DNA strands $d(TG_iTC_5T)$ and $d(TC_5TG_iT)$ ($i < 5$) are explored as the primary sequences to evaluate the above postulates. Thymines are used as the 'spacers' because they stabilize i-motif structures either in the middle or at the end of C-rich strands (37,38), and they are less prone to stacking or pairing with guanines (39). The rational investigation of C-runs with spacer and flanking residues has directed us to model i-motifs that can bear longer G-runs for the formation of G4s. To our knowledge, this is the first systematic study of DNA sequences composed of C- and G-runs and their assembly simultaneously into both i-motifs and G4s in a designed fashion.

MATERIALS AND METHODS

Materials and sample preparation

The oligonucleotides (OPC grade for ordinary sequences and HPLC grade for the modified sequence) were synthesized by TaKaRa Biotechnology Co., Ltd. (Dalian, China) and used as received. The DNA fragments studied here are listed in Table 1. Stock solutions of the oligonucleotides were prepared by directly dissolving the lyophilized powder in Milli-Q water to approximately 1 mM and were stored at −20°C. Accurate concentration of the oligonucleotides were obtained from their UV absorbances at 260 nm at 90°C using molar absorptivities tabulated at the website (http://scitools.idtdna.com/scitools/Applications/oligoAnalyzer).

For the supramolecular DNA assembly, 40–100 μM of each oligomer was placed in pH 4.5 (acetic acid–acetate) or pH 9.0 (acetate–ammonia-water) buffer solution with different counter ions (50–100 mM NH_4^+, Li^+, Na^+ or K^+ ions). Samples were heated in a 90°C water bath for 15 min and slowly cooled to room temperature (about 10 h), followed by equilibration at 4°C for more than 4 days.

Mass spectrometry

Mass spectrometric analysis was performed in ESI negative-ion mode on an ESI-Q-TOF (micrOTOF-Q II, Bruker, Breman, Germany) mass spectrometer controlled by Bruker *ESI Compass Data Analysis 4.0* software. This system can measure m/z in the range 50–3000. The samples were infused directly into ion source at a flow rate of 3 μl min^{-1}. Optimal soft ionization conditions for quadruplex structures were found using the oligomer $d(G_4T_4G_4)$ to tune the instrument to produce abundant $[(dG_4T_4G_4)_2 + 3NH_4]^{5-}$ ions as described previously (37,40). These conditions also produce good spectra for i-motifs (37). Before the MS analysis samples were diluted with an equal volume of methanol solution 60% in water in order to obtain a good spray (41). The data collection time was 0.6 min for each spectrum. For

Table 1. Sequences of oligodeoxynucleotides studied here

DNA sequence	MW	DNA sequence	MW
TC_5T	1991.37	TG_3AC_5T	3291.58
$TGTC_5T$	2624.47	TG_3AC_5	2987.54
TG_2TC_5T	2953.52	G_3AC_5T	2987.54
TG_3TC_5T	3282.57	TG_4AC_6	3605.64
TC_5TGT	2624.47	$TG_4A_2C_6$	3918.69
TC_5TG_2T	2953.52	TG_4AC_7	3894.68
TC_5TG_3T	3282.57	$TG^{Br}GG^{Br}GAC_7$	4050.52
TG_3C_5T	2978.53		

MS/MS experiments, precursor ions were selected and isolated in the quadrupole and then fragmented in the collision cell. The in-source collision induced dissociation (ISCID) energy was set at 50 eV. The collision energy was varied over the range of 10–30 eV.

MS ion distribution analysis

The gas phase ion abundances and theoretical isotope distributions were obtained by using Bruker *ESI Compass Data Analysis 4.0* software. The percentage (TP) of each ion species is calculated using Equation (1):

$$TP = \frac{I_{nM} \times \mathrm{n}}{I_M + I_{2M} \times 2 + I_{3M} \times 3 + I_{4M} \times 4} \times 100\% \quad (1)$$

where the numerator is the sum of all strands in one ion species, and the denominator is the sum of all strands in all ions. Each I is the intensity of detected ion in different charge states. For simplicity, the ion species such as fragment ions in the gas phase and hydrolyzed ions in solution were not counted here due to their low abundance or low influence on the overall results. Additionally, when a sequence generated a series of $[4M + yNH_4]^{x-}$ ions ($y = 0 \sim 3$), only the strongest ion peak was counted. When the peaks from monomers and dimers overlap with the tetramer, their intensities are subtracted from the tetramer intensities based on the calculated isotopic profiles. Similarly, isotopic profiles were also used to estimate intensities of peaks with m/z out of range.

Cd spectroscopy

CD spectra were recorded on a J-810 CD spectrometer (JASCO, Tokyo, Japan) using 0.2 mm path length Hellma cell at the room temperature. Each sample of 40 μM DNA strand was dissolved in 50 mM NH_4OAc or LiOAc (pH 4.5 or 9.0). Spectra were recorded from 320 to 200 nm as the averages of three scans. Each trace was measured at 100 nm/min of scanning speed, with 2 s response time, 1 nm data pitch and 1 nm bandwidth. The background spectra corresponding to the buffer alone were subtracted from all DNA spectra. Characteristic peaks of various nucleic acid structures are listed in Supplementary Table S1.

UV absorption spectrophotometry

UV absorbance versus temperature melting profiles were obtained using a UV-2550 spectrophotometer (SHIMADZU, Kyoto, Japan) equipped with a S-1700 temperature controller and measured at 260, 265 and 295 nm, respectively, according to previous reports (42,43). The samples with final DNA concentrations of 40 μM dissolved in 50 mM LiOAc at pH 4.5 or 9.0 were used for UV experiments. In each step, the temperature was increased by 0.5°C and equilibrated for 0.5 min before recording the absorbance. The absorption cell was sealed to avoid solvent evaporation, and the gas bubbles generated in heating process were removed by stirring.

Native gel electrophoresis

Gel electrophoresis was run with an 8% poly/acrylamide gel at 4°C controlled by a condenser circulating water system. 2-(N-morpholino) ethanesulfonic acid monohydrate (MES, pH 4.5, 50 mM) buffer was used to prepare the gel as well as the running buffer, which was further supplemented with 100 mM KOAc. Each sample used for Native Gel Electrophoresis contained 100 μM DNA strand and 100 mM counter ions (either Li^+, Na^+ or K^+ ions) at pH 4.5 or 9.0. The electrophoresis was performed for 3 h at 150 V. Results were visualized by Ultrapower™ visible light transilluminator (Bioteke, Beijing, China) after using Gelgreen DNA staining agent (Biotium) in 100 mM NaCl. The gel image was enhanced using a grayness transform by using Adobe Photoshop CS6 software to reduce background interference.

Atomic force microscopy

Atomic force microscopy (AFM) was performed on mica surfaces that bind nucleic acids with bridging magnesium ions (44). Here, the DNA samples (100 μM of each DNA) were annealed in different buffers containing 100 mM cations at 4°C for 1 week. Then aliquots were diluted with 2 mM $MgCl_2$ aqueous solution to give a 20 μl analyte containing 1.5 μM DNA. The analytes were incubated on freshly cleaved mica plates for 5–8 min. After that, excess fluid was removed. Then the mica plates were washed carefully and repeatedly with Milli-Q water, and left to air-dry, in a clean container to prevent dust settling. AFM analysis was in tapping mode with a Nanoscope IIIa scanning probe microscope (Bruker) from Digital Instruments utilizing NANOSENSORS™ PPP-NCHR AFM probes. If necessary, the AFM images were processed by flattening to remove the background slope and adjusted for contrast and brightness.

RESULTS AND DISCUSSION

We seek to create a supramolecular structure in which one oligonucleotide contributes a run of C's to an i-motif and a run of G's to a G4. We use mass spectrometry to observe multimeric complexes of the oligonucleotide. I-motif anions will have a mass that is simply the sum of the four strands with the appropriate number of hydrogens to give the observed charge. G4 anions require *n*-1 cations (here NH_4^+) for *n* G-quartets assembled. The structural arrangement is confirmed by CD. Since i-motifs only form under acid conditions, the CD difference spectrum resulting from a pH 9.0 spectrum subtracted from a pH 4.5 spectrum reveals the i-motif contribution. For G4s, we use cation dependence: K^+ supports parallel quartet stacking, while Na^+ shows preference for antiparallel stacking and NH_4^+ falls in between these two cations based on the nature of sequences, but Li^+ does not support any stacking of G-quartets (3,45,46). Thus, the CD difference spectrum of a sample in the presence of NH_4^+ minus the spectrum in Li^+, gives the G4 contribution. See SI for a table of characteristic CD peaks for different DNA structures.

The influence of the location of guanine residues or G-stretches on the tetramolecular i-motif formation

Suitable solution conditions for the formation of tetrameric i-motifs were determined by studying the oligonucleotide d(TC_5T). MS spectra showed abundant tetramolecular ions of d(TC_5T) (including $[4M]^{5-}$ at m/z 1592.69 and $[4M]^{4-}$ at m/z 1991.12) in pH 4.5 NH_4OAc buffer, containing 50 mM NH_4^+. These tetrameric species disappeared when the solution pH was adjusted to 9.0, indicating the formation and dissociation of tetrameric i-motifs (Supplementary Figure S1a and b) in accordance with our previous results (37).

We investigated the influence of guanine residues or G-runs on the formation of i-motif structures and the competition between WC duplex and i-motif structures using the sequences d(TG_iTC_5T) and d(TC_5TG_iT) (i = 1, 2, 3). The spectra in Supplementary Figure S2 indicate that both d($TGTC_5T$) and d(TC_5TGT) (i = 1) form four-stranded i-motifs under acidic conditions, but that the 3′ terminal GT produces less. Specifically, the tetrameric ions ($[4M]^{7-}$ at m/z 1499.36, $[4M]^{6-}$ at m/z 1749.36 and $[4M]^{5-}$ at m/z 2099.42) in the MS spectra (Supplementary Figure S2a and b) and the CD maximum at 288 nm and minimum at 265 nm (Supplementary Figure S2c) all indicate the presence of i-motif tetramers (47) and that they are more prominent in d($TGTC_5T$). We have previously demonstrated that short C-rich strands facilely form duplexes with C•CH^+ pairs under acidic condition and can be observed along with tetrameric i-motifs (37). Consequently, the low-abundance duplex ions observed here are mainly C•CH^+-containing duplexes (C-duplex) since there are insufficient complementary G and C bases between two oligonucleotides to create WC duplexes (Supplementary Figure S2a and b).

The two oligonucleotides d(TG_2TC_5T) and d(TC_5TG_2T) show the same tendency, with an even greater inhibition of i-motif formation by a 3′ tailing G_2T. The d(TG_2TC_5T) sequence assembles into a tetrameric i-motif structure in acidic solution, as supported by the observation of tetrameric ions ($[4M]^{7-}$ at m/z 1687.57 and $[4M]^{6-}$ at m/z 1968.99) and the characteristic CD signature of the i-motif structure (Figure 1A and E). However, d(TC_5TG_2T) only produces the duplex ions ($[2M]^{5-}$ at m/z 1181.01, $[2M]^{4-}$ at m/z 1476.26 and $[2M]^{3-}$ at m/z 1968.68) and no i-motif tetrameric ions were observed (Figure 1B). We propose that the bimolecular ions observed under acidic conditions may correspond to two types of duplex structures: a C•CH^+-containing-duplex as well as a bulged WC duplex (the C-duplex with five non-intercalated C•CH^+ base pairs will lose another five hydrogen cations to give the same m/z as the bulged WC duplex). In contrast, the bimolecular ions formed under alkaline conditions mainly correspond to the bulged WC duplex, which has no hemi-protonation of cytosines (Supplementary Figure S3). The percentage of $[TC_5TG_2T]_2$ ions decreased from 42.56 to 9.99%, when the solution pH was adjusted from 4.5 to 9.0 (Table 2). If it is assumed that the same fraction of the samples ionize, we can conclude that the proportion of C-duplex and bulge duplex is 3.3 : 1 (32.57% : 9.99%) under acidic condition when the influence of solution pH on the formation of WC duplexes is ignored. Since CD signal of the non-intercalative C•CH^+ base pair stacking may be weak (48), the CD spectrum of d(TC_5TG_2T) only displays a single positive band around 290 nm. However, the CD difference spectrum obtained by subtracting the pH 9.0 spectrum from the pH 4.5 spectrum (Figure 1F, inset) shows the positive band at 288 nm and the negative band at 265 nm, suggestive of parallel C-duplex.

In contrast, i-motif structures were not formed effectively in acidic solutions when the G-stretches extended to three guanines. Figure 1C and D show that only bimolecular ions (including $[2M]^{5-}$ at m/z 1312.45, $[2M]^{4-}$ at m/z 1640.81 and $[2M]^{3-}$ at m/z 2188.35) were detected in mass spectra both for the sequences d(TG_3TC_5T) and d(TC_5TG_3T). Scheme 2B and C propose two possible structures for dimers of d(TG_3TC_5T): a parallel C-duplex or a bulged double helix. Data of Table 2 indicates that the proportions of these two structures are 20.39% : 15.04% (≈ 1.4

Table 2. The percent abundance of various ion species formed at different pH's as measured by mass spectrometry for the various oligonucleotides studied

Sequence	pH value	M (%)	2M (%)	3M (%)	4M (%)
TG2TC5T	4.5	67.04	8.95	1.02	22.99
	9.0	88.58	11.42		
TC5TG2T	4.5	57.44	42.56		
	9.0	90.01	9.99		
TC3TG5T	4.5	64.57	35.43		
	9.0	84.96	15.04		
TC5TG3T	4.5	76.61	23.39		
	9.0	89.50	10.50		
TG3C5T	4.5	67.41	22.21		10.38
	9.0	41.73	54.98	3.29	
TG3AC5T	4.5	42.04	18.74		39.22
	9.0	86.79	13.21		
TG3AC5	4.5	18.86	1.19		79.95
	9.0	87.34	12.66		
G3AC5T	4.5	48.63	22.26		29.11
	9.0	79.88	17.38	2.74	

*Each data obtained is the average of two measurements.

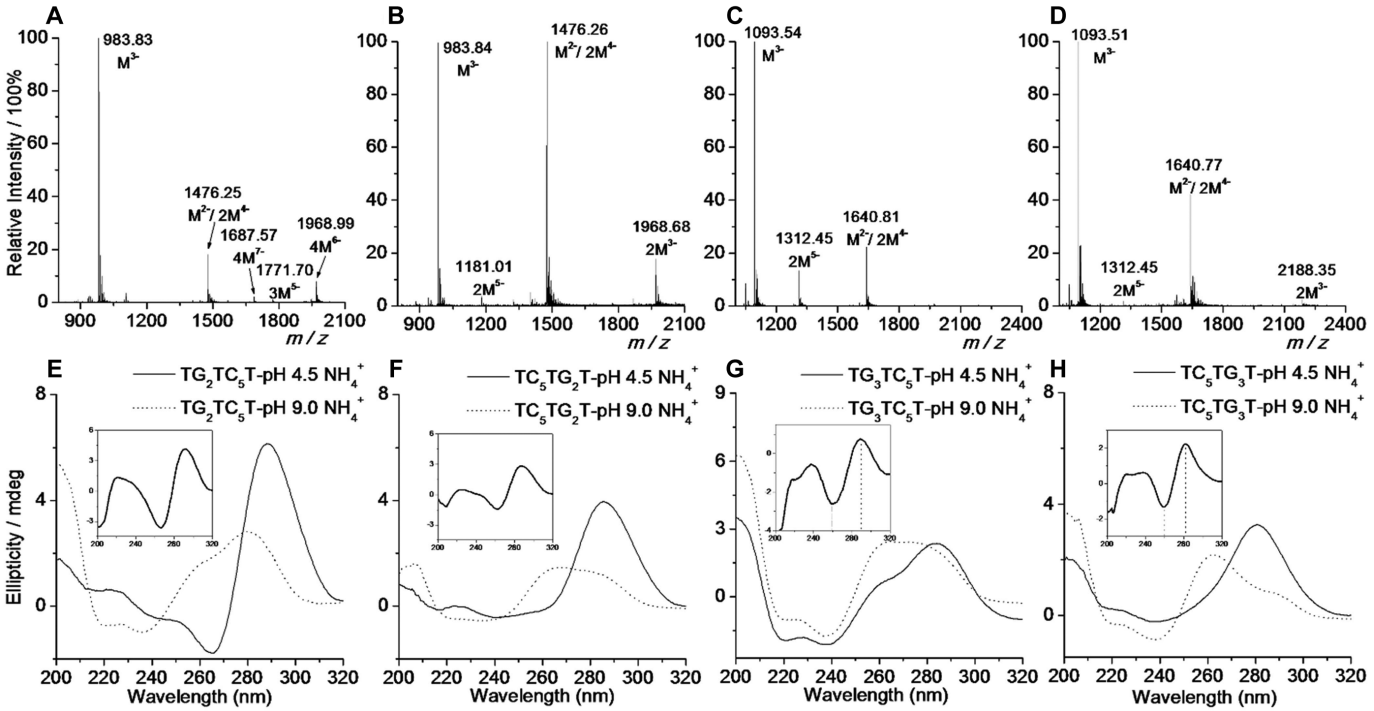
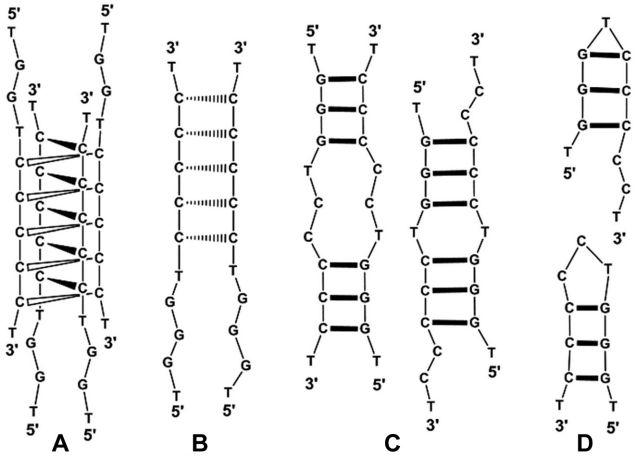


Figure 1. Mass spectra of assemblies of oligonucleotides with sequences (**A**) d(TG_2TC_5T), (**B**) d(TC_5TG_2T), (**C**) d(TG_3TC_5T) and (**D**) d(TC_5TG_3T) at 60 μM in 50 mM NH_4OAc (pH 4.5) after incubating the samples at 4°C for 4 days; (**E**–**H**) CD spectra of the four oligonucleotides in the acidic (solid line) and alkaline solutions (dotted line); inset: subtracted CD spectrum obtained by subtracting the CD spectrum in alkaline buffer solution from the CD spectrum in acidic buffer solution.



Scheme 2. Schematic drawings of (**A**) 3′ E (end) fully intercalated tetramolecular i-motif formed by d(TG_2TC_5T); (**B**) C•CH^+ base pair duplex; (**C**) and (**D**) WC double helix and hairpin structure with different lengths of flanking residues formed by d(TG_3TC_5T); solid wedged bond represents intercalation C•CH^+ base pair; dashed wedged bond represents non-intercalation; solid stick bond represents WC pairing.

: 1) for d(TG_3TC_5T) and 12.89% : 10.5% (≈ 1.2 : 1) for d(TC_5TG_3T) under acidic conditions. The CD difference spectrum of d(TG_3TC_5T) (Figure 1G, inset) displays the characteristic CD signature of protonated C-duplex with an approximate 5 nm blue shift for the negative band at 265 nm, while d(TC_5TG_3T) shows the two typical peaks of protonated C-duplex, each blue shifted by approximately 5 nm in the CD difference spectrum (Figure 1H, inset). Presumably the increase in complementarity allows the formation of more bulged WC duplexes. In addition, CD spectrum of d(TG_3TC_5T) shows that a small negative band around 210 nm coexists with the positive band around 260 nm (Figure 1G), which implies that the oligonucleotide of d(TG_3TC_5T) not only forms the B-form double helix but also forms the A-form hairpin structure (Scheme 2D) as has been described by others (49). In summary, we conclude that runs of three G's are detrimental to i-motif formation and that the introduction of G-repeats at the 3′ end hinders the formation of i-DNA tetramers more significantly than their introduction at the 5′ end.

We assume that the i-motif structure formed by d(TG_2TC_5T) sequence will have two sets consecutive G's at each terminus (Scheme 2A) available for interchain interactions. The formation of a parallel G4 goes through dimer and trimer intermediates (50,51). However, two parallel dangling 5′-GGT's are unlikely to pair with one another since the tetrameric G4 with only two G-quartets is known to be very unstable (52).

The effect of central and terminal thymines on tetramer formation

The oligonucleotide d(TG_3TC_5T) has little tendency to form i-motif structure in our experimental conditions. We investigated whether or not the spacer nucleotide at the junction between G- and C-repeats influenced i-motif formation. First we studied the strand d(TG_3C_5T), which lacks a spacer base. Supplementary Figure S4 shows that the

tetrameric ions of d(TG_3C_5T) (including $[4M]^{7-}$ at m/z 1701.90 and $[4M]^{6-}$ at m/z 1985.40) are observed at low abundance under acidic condition and the pattern of the CD difference spectrum clearly shows the characteristic CD signature of i-motif without blue shift, suggesting that the lack of a spacer supports formation of i-motifs. However, it is still the case that bimolecular ions (including $[2M]^{5-}$ at m/z 1190.84, $[2M]^{4-}$ at m/z 1488.80 and $[2M]^{3-}$ at m/z 1985.40) corresponding to the WC duplex were abundant under basic condition for d(TG_3C_5T) sequence (Supplementary Figure S3e and Data of Table 2). These results demonstrate that the WC duplex formation also becomes more favorable by removing the spacer base. This is likely due to eliminating the mismatch in the middle of double-helix structures.

We then tried d(TG_3AC_5T), which substitutes the spacer thymine with adenine, the only other base that does not extend either the G- or C-run. Figure 2A shows that both dimeric and tetrameric ions (including $[2M]^{5-}$ at m/z 1316.08, $[2M]^{4-}$ at m/z 1645.59, $[4M]^{7-}$ at m/z 1880.81 and $[4M]^{6-}$ at m/z 2194.45) are observed in MS spectrum, indicating the coexistence of duplex and i-motif assemblies. In addition, CD spectrum of d(TG_3AC_5T) measured in acidic solution exhibits the characteristic peaks of the i-motif with a reasonable shift for the negative peak around 265 nm (26), while the characteristic peaks of A-type-duplex-containing hairpins are not observed (Figure 2D). We propose that d(TG_3AC_5T) forms tetrameric i-motifs because the adenine purine supports stacking over bending into the loop of a single-strand hairpin structure (53). Furthermore, the steric hindrance of paired adenines may be unfavorable to the formation of a bulged WC duplex.

It has been known that the flanking residues of DNA tetramers at both ends have significant effect on the strand association and structural stability (54,55). Hence the 3′ end thymine of d(TG_3AC_5T) was removed. Figure 2B and D show that the tetramolecular ions of d(TG_3AC_5) are more abundant than that of d(TG_3AC_5T) (the percentage of tetramolecular ions changes from 39.22 to 79.95%). The CD spectra in Figure 2D indicate a sharp enhancement of the i-motif peaks under acidic conditions when the 3′ thymine is removed. In comparison, the removal of the 5′ thymine has little effect on the i-motif formation: both the abundance of tetramolecular ions in Figure 2C (the percentage of tetramolecular ions is 29.11%) and the CD signals of i-motif structure in Figure 2D have a slight decrease compared with the sequence d(TG_3AC_5T).

This indicates that d(TG_3AC_5) is most amenable to tetrameric i-motifs formation. I-motif structures can be classified into two types, 3′ E and 5′ E, determined by the outside terminus of the intercalated C•CH^+ base pairs (8,13). 3′ E and 5′ E i-motif structures with or without 3′ thymines of d(TG_3AC_5T) are depicted in Scheme 3. In both 3′ E and 5′ E, the C•CH^+ base pairs from two orthogonal C-duplexes stack to place the 3′ thymine residues far from the central adenine residues. The distance between them is about 0.31 nm, which is similar to the space between two WC base pairs (about 0.34 nm) (13,56). Consequently, the 3′ thymine and central adenine residues do not easily form A•T WC base pairs to further stabilize the i-motif structure. Furthermore, the 3′ thymine residues crowd the terminal regions of i-motif structures. A recent kinetics investigation demonstrated that the association rates of tetrameric i-motifs with partial (out of register) or full intercalation topology are comparable (14). By reducing the 3′ thymine residues, various i-motifs with partial registry will form along with the fully intercalated (registered) species without restrictions. This dynamically favors the formation of i-motifs when they compete with WC duplex structures. For these reasons, removing the 3′ thymine can promote the formation of i-motif structure dramatically.

Controlled assembly of TG_mAC_n quadruplex DNA nanoarchitectures

In our experiments, d(TG_3AC_5) forms a significant amount of i-motif with two parallel G-runs at each end. These parallel G-runs should serve as flexible branches to link two i-motif scaffolds through the formation of interstrand antiparallel or parallel G4s, as depicted in Scheme 4. Antiparallel association should give DNA quadruplex nanowires with alternating G4s and i-motifs, while parallel association would create G-quartet hubs surrounded i-motifs. Since both i-motifs and G4s have slow assembly kinetics (14,57), it is possible that single i-motifs and G4s might be observed simultaneously as the oligonucleotide assembles into a supramolecular structure. These may be independent assemblies or partially associated supramolecular complexes that break down during ionization.

No tetrameric ions with the appropriate number of cations corresponding to G4s were observed in mass spectra of d(TG_3AC_5). It could be that the three-nucleotide G-run is too short to support a stable tetrameric G4 in this context. Hence, we extended the length of G- and C-runs to d(TG_4AC_6). Supplementary Figure S5a shows that under acidic conditions allowing i-motifs, this strand produced tetrameric ions with three NH_4^+ ions (including $[4M + 3NH_4^+]^{7-}$ at m/z 2067.62 and $[4M + 3NH_4^+]^{6-}$ at m/z 2412.40), indicative of G4 structures, while no tetrameric ions are observed for d(TG_4AC_6) under alkaline conditions (Supplementary Figure S5b). Unfortunately, these tetramer species only represent a small proportion of the sum of all strands (17.57%). In particular, there are still strong peaks from monomers and dimers, indicating that a significant part of the sample is not in a tetraplex or more complex assemblage. Supplementary Figure S5c shows that NH_4^+ can enhance the CD signatures at 220, 245 and 265 nm obviously in comparison with Li^+, suggesting the formation of parallel G4 in NH_4^+ solution. The difference spectrum of Li^+ subtracted from NH_4^+ containing solutions, both taken at pH 4.5, does not give a typical profile of a parallel G4, which may be due to less abundant G4s formed (Supplementary Figure S5c, inset).

We added a central adenine to give d($TG_4A_2C_6$). Supplementary Figures S5d and e show that no tetrameric ions are observed for this strand in the presence NH_4^+ under both pH conditions. Solutions of this strand also produced no change in the CD signal when Li^+ was substituted for NH_4^+ (Supplementary Figure S5f). The additional adenine is detrimental to tetramer formation.

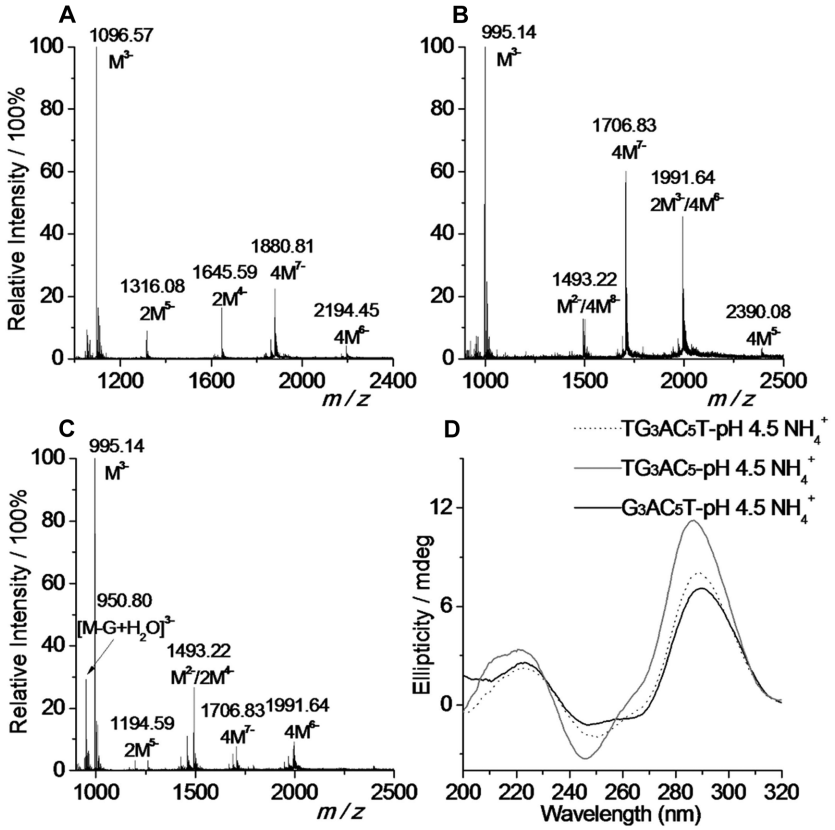
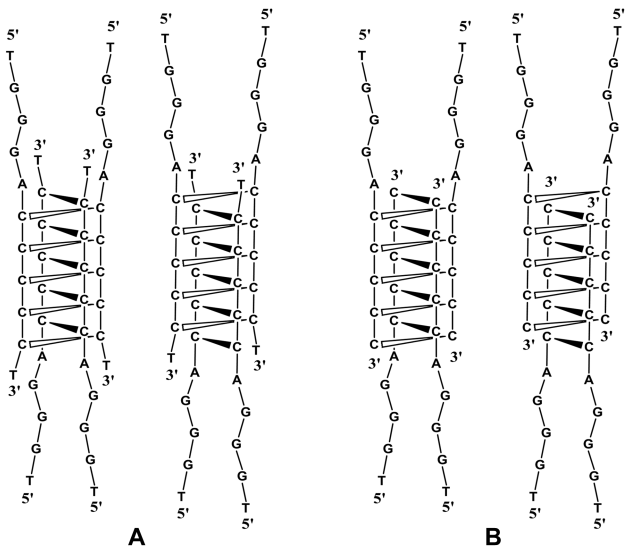


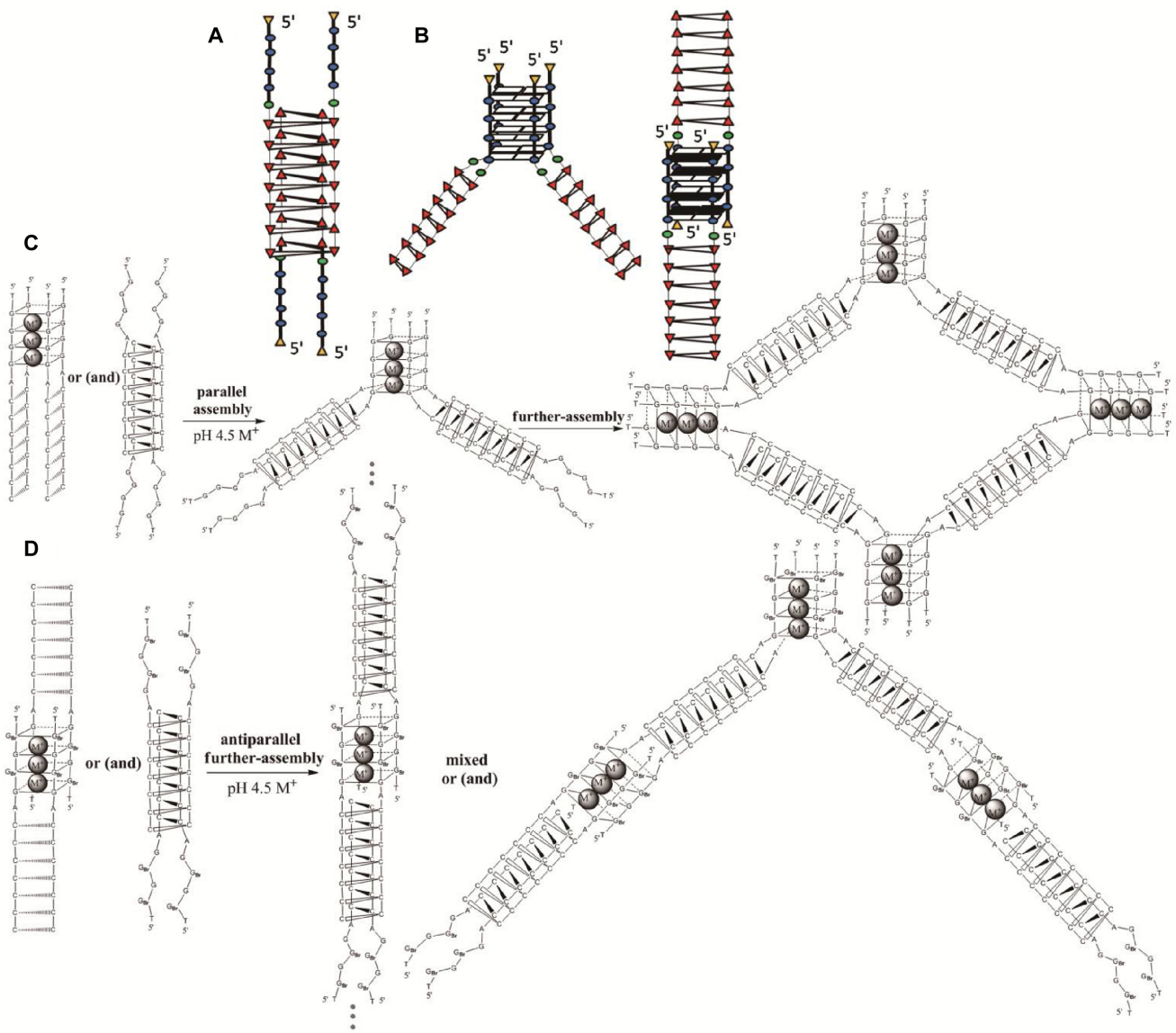
Figure 2. Mass spectra of assemblies of oligonucleotides with sequences (**A**) d(TG_3AC_5T), (**B**) d(TG_3AC_5) and (**C**) d(G_3AC_5T) at 60 μM in 50 mM NH_4OAc (pH 4.5) after incubating the samples at 4°C for 4 days; (**D**) CD spectra of the three DNAs at pH 4.5.



Scheme 3. Schematic drawings of 3′ E and 5′ E fully intercalated tetramolecular i-motifs (**A**) with 3′ thymines [d(TG_3AC_5T)] and (**B**) without 3′ thymines [d(TG_3AC_5)].

The oligonucleotide with only one spacer A and one more C, d(TG_4AC_7), produces an abundance of tetrameric ions (Figure 3A: $[4M + 3NH_4]^{6-}$ at m/z 2605.23; $[4M + 3NH_4]^{7-}$ at m/z 2232.77; $[4M + 3NH_4]^{8-}$ at m/z 1953.56) that decrease significantly under alkaline conditions (from 49.07 to 0%; Supplementary Figure S6a). NH_4^+-adduct-free tetrameric i-motif ions were not observed under this condition; however, the CD difference spectrum obtained by subtracting the pH 9.0 spectrum from the pH 4.5 spectrum indicates the presence of C•CH^+ base pairs (Supplementary Figure S6c, inset). Under acid conditions decreasing NH_4^+ ions from 50 to 30 mM, the lower abundant ions corresponding to i-motif tetrameric ions were observed (the ions with m/z at 2596.71 for 6– and at 2225.60 for 7–) along with G4 tetrameric ions; continuously decreasing NH_4^+ ion concentrations, at 20 mM NH_4^+, the abundance of tetramolecular ions containing three NH_4^+ ions decreased drastically; at 10 mM NH_4^+, the G4s were even more unstable and only bimolecular ions were observed (data no shown). The G4 monomer units of d(TG_4AC_7) are more easily observed than i-motif monomer units, which may be due to a number of factors. G4s with protonated C-duplexes may be more stable than i-motifs with G-rich overhangs (Scheme 4A versus B), perhaps because the C's sequestered in C•CH^+ pairs, and were unable to form WC pairs with G's (22,23). When alkaline conditions release these C's they readily form WC pairs with the G's, preventing quadruplex formation.

Tetramolecular ions adducted with NH_4^+ observed in our mass spectrum show good agreement with the G4s formed by d(TG_4AC_7) that specifically have three buried



Scheme 4. Postulated quadruplex structures for d(TG_4AC_7) and d($TG^{Br}GG^{Br}GAC_7$): (**A**) the tetrameric i-motif with G-rich overhangs at each terminal; (**B**) tetra-stranded parallel and antiparallel G-quadruplexes (G4s) with two pairs of protonated C-duplexes formed by d(TG_4AC_7) and d($TG^{Br}GG^{Br}GAC_7$) respectively under acidic conditions; circle: purine; triangle: pyrimidine; in a G-tetrad, *anti* guanosines are indicated in white and *syn* guanosines are indicated in black; Supramolecular DNA structure that could assemble from oligonucleotide with sequences (**C**) d(TG_4AC_7) constructed by parallel tetra-strand G4s and antiparallel i-motifs; and (**D**) d($TG^{Br}GG^{Br}GAC_7$) formed by antiparallel tetra-strand G4s, antiparallel i-motifs and (or) parallel G4s.

NH_4^+ ions among four stacked G-quartets (Figure 3A). In order to show that these NH_4^+ ions are not the ions adsorbed via electrostatic interactions with the phosphate backbone, MS/MS experiments for the d(TG_4AC_7) $[4M + 3NH_4]^{x-}$ ions were performed. Figure 4 shows product-ion spectra of the d(TG_4AC_7) $[4M + 3NH_4]^{7-}$ ion (m/z = 2232.77) resulting from collision energies of 10, 20 and 30 eV. The separation of NH_4^+ ions is not significant until the collision energy reaches 30 eV. This energy is sufficient to disrupt interchain interactions (37); hence, the NH_4^+ ions can only dissociate upon disruption of the tetraplex.

I-motif formation and stability were investigated with UV melting experiments at 265 and 295 nm for the sequence d(TG_4AC_7) in pH 4.5 LiOAc buffer. Both melting curves give two apparent melting points around 50°C and 75°C, which are considered to be the dissociations of duplexes (including WC duplex and C-duplex) and i-motif structures, respectively (Supplementary Figure S7a). This conclusion is further supported by the UV melting experiment at pH 9.0 in LiOAc buffer solution, Supplementary Figure S7b. It shows only a single transition point at 260 and 265 nm, corresponding to the dissociation of WC duplex and the melting temperature (T_m) of WC duplex obtained from the UV melting curve (at 260 and 265 nm) is about 60°C, while no obvious transition point is observed for the measurement at 295 nm. It is unwise to further increase the length of C-repeats since this may induce the formation of an intramolecular C^+•GC triplex (58).

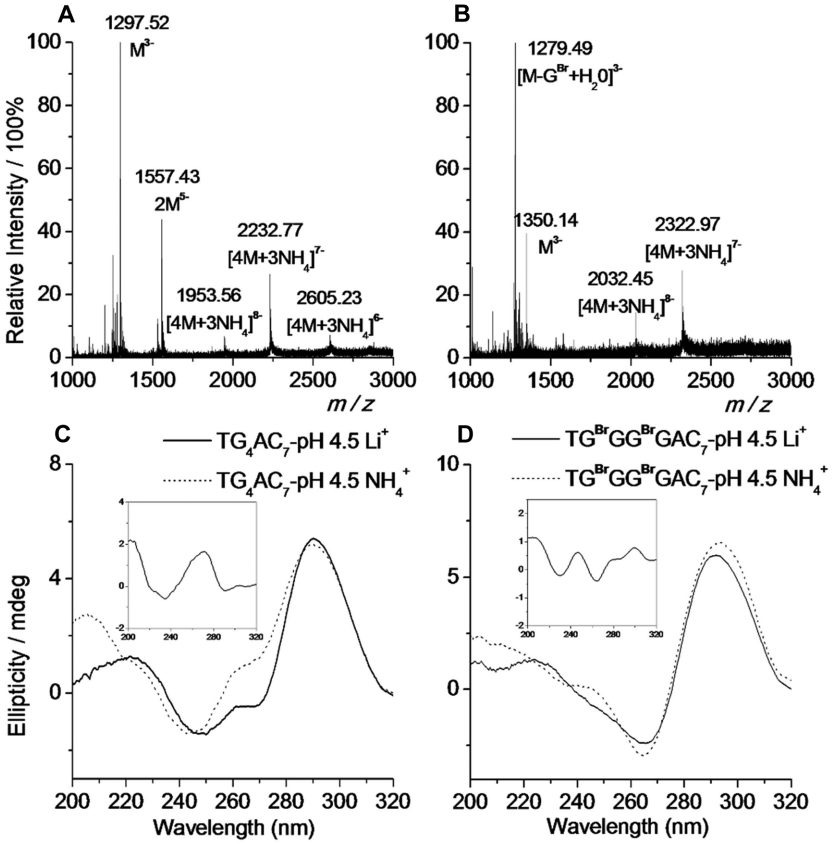


Figure 3. Mass spectra of assemblies of oligonucleotides with sequences (**A**) d(TG_4AC_7) and (**B**) d($TG^{Br}GG^{Br}GAC_7$) at 60 μM in 50 mM NH_4OAc (pH 4.5) after incubating each sample at 4°C for 4 days; (**C** and **D**) CD spectra of the two sequences in pH 4.5 NH_4OAc and pH 4.5 LiOAc buffers; inset: CD difference spectrum obtained by subtracting the CD spectrum with the LiOAc solution from the CD spectrum with the NH_4OAc solution.

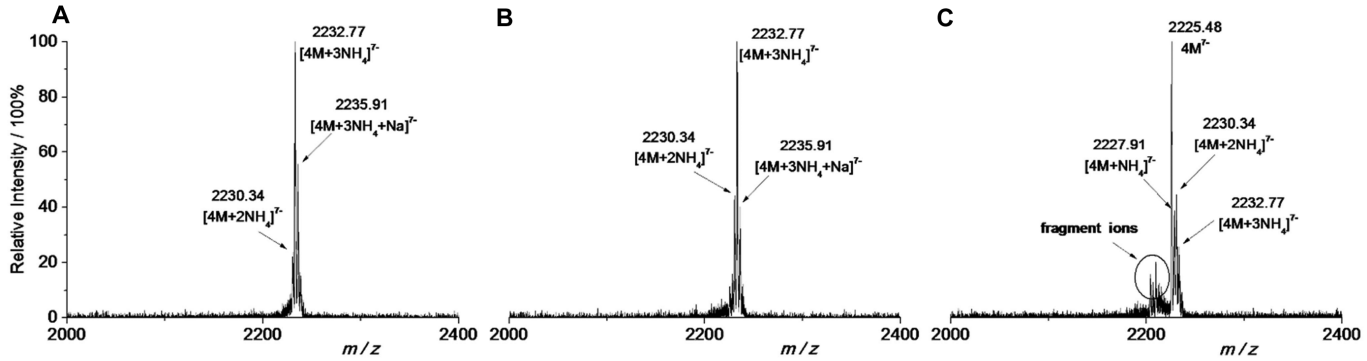


Figure 4. Product-ion spectra of d(TG_4AC_7) $[4M+3NH_4]^{7-}$ ion ($m/z = 2232.77$) for collision energies of (**A**) 10 eV, (**B**) 20 eV and (**C**) 30 eV.

We next investigated the effect of modification of guanines on the formation and stability of tetramolecular G4s. It has been reported that the inclusion of two 8-methyl-2′-deoxyguanosine (M) residues in d(TMGMGT) promotes the formation of an unprecedented antiparallel tetra-stranded G4 in which two adjacent strands run in the same direction (59). Since a bromine atom has a similar size as a methyl group and the 8-bromoguanine residue has the same tendency to promote a *syn* glycosidic conformation (7,60–62), we investigated the assembly properties of d($TG^{Br}GG^{Br}GAC_7$), where G^{Br} indicates an 8-bromoguanine residue. Figure 3B shows that under acidic conditions d($TG^{Br}GG^{Br}GAC_7$) forms tetrameric ions with three NH_4^+ ions corresponding to tetramolecular G4 structures ($[4M + 3NH_4]^{8-}$ at m/z 2032.45; $[4M + 3NH_4]^{7-}$ at m/z 2322.97). The CD difference spectrum (Figure 3D, inset) suggests that the tetrameric G4 is antiparallel as depicted in Scheme 4B based on previous reports (59,63). As with d(TG_4AC_7), the tetramers were dissociated under alkaline conditions (Supplementary Figure S6b) and the CD difference spectrum obtained by subtracting the pH 9.0 spectrum from the pH 4.5 spectrum indicated the presence

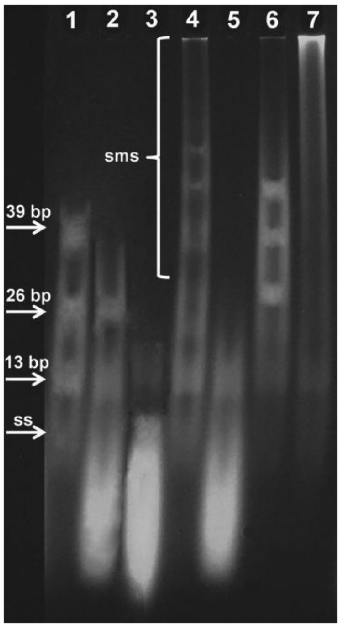


Figure 5. Native electrophoretic mobility on 8% polyacrylamide at pH 4.5 of oligonucleotides d(TG_4AC_7) and d($TG^{Br}GG^{Br}GAC_7$) in different buffers containing 100 mM cations after annealing DNA samples at 4°C for 1 week: Lane 1: a molecular mass ladder: 13, 26 and 39 bp; Lane 2: d(TG_4AC_7) with pH 4.5 LiOAc; Lane 3: d(TG_4AC_7) with pH 9.0 KOAc; Lane 4: d(TG_4AC_7) with pH 4.5 KOAc; Lane 5: d($TG^{Br}GG^{Br}GAC_7$) with pH 9.0 KOAc; Lane 6: d($TG^{Br}GG^{Br}GAC_7$) with pH 4.5 NaOAc; Lane 7: d($TG^{Br}GG^{Br}GAC_7$) with pH 4.5 KOAc; ss, single-stranded structures; sms, supramolecular structures.

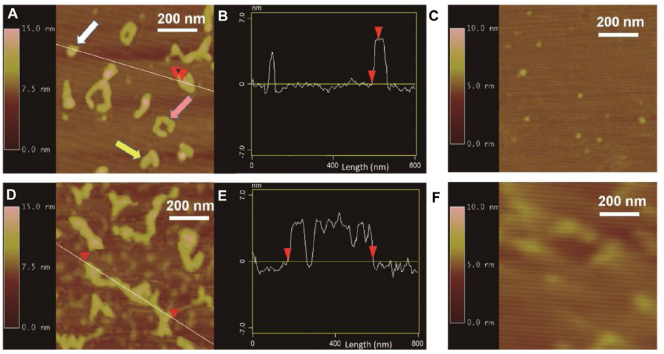


Figure 6. AFM images with the scale bar of 200 nm for DNA quadruplex wires self-assembled from oligonucleotide sequences d(TG_4AC_7) (**A** and **C**) and d($TG^{Br}GG^{Br}GAC_7$) (**D** and **F**) with different solutions and height scales: (A and D) pH 4.5 KOAc buffer; (C and F) pH 9.0 NaOAc solution. The height and length profiles of the aggregates formed by d(TG_4AC_7) (**B**) and d($TG^{Br}GG^{Br}GAC_7$) (**E**) deposited on the freshly cleaved mica substrate follows the lines on (**A**) and (**C**) respectively. The triangles indicate the same points in the image and the height analysis. The arrows in (A) are used to point to DNA structures with different shapes and sizes.

of C•CH^+ base pairs under acidic conditions (Supplementary Figure S6d, inset). We are not able to observe larger species formed by d(TG_4AC_7) or d($TG^{Br}GG^{Br}GAC_7$) due to our instrument's m/z range limit.

The formation of quadruplex DNA nanoarchitectures was further indicated by non-denaturing polyacrylamide gel electrophoresis (PAGE). K^+ ions can stabilize parallel G4s effectively, while Na^+ and K^+ ions have distinct effects on various types of antiparallel G4s (3). Consequently, d(TG_4AC_7) was mainly annealed in KOAc buffer solutions at pH 4.5 and pH 9.0 to ensure the desirable stability of the parallel G4 formed by GGGG-fragment in d(TG_4AC_7), and d($TG^{Br}GG^{Br}GAC_7$) was annealed in both KOAc and NaOAc buffer at the two pH's for the PAGE experiment since both parallel and antiparallel G4s may be formed by $G^{Br}GG^{Br}$G-fragment in d($TG^{Br}GG^{Br}GAC_7$). These gels run at pH 4.5 and 4°C in order to stabilize i-motifs. We believe that samples from higher pH's form only a small amount or no supramolecular DNA structures due to the pH drop upon loading and running in the gel, because the rate of formation of such supramolecular DNA structures is very slow at low temperature (35).

In accordance with the UV melting study, Figure 5 shows that d(TG_4AC_7) incubated in pH 4.5 LiOAc solution has three obvious bands corresponding to single-, double-stranded structures and tetra-stranded i-motifs. However, d(TG_4AC_7) incubated in pH 9.0 KOAc solution has only two obvious bands corresponding to single- and double-stranded structures. The various smeared bands migrating slowly are believed to be supramolecular structures assembled by the oligonucleotide at pH 4.5 in the presence of K^+. Similar results were also obtained for d($TG^{Br}GG^{Br}GAC_7$). The bands of single-stranded structures are much broader than the double-stranded structures under alkaline conditions. They are likely to contain two components: the open random coil conformation and the hairpin conformation, which may be exchanging. d($TG^{Br}GG^{Br}GAC_7$) forms larger structures at pH 4.5 with either Na^+ or K^+ ions. When K^+ ions are present the structures stay at the top of the gel. Under this condition both parallel and antiparallel G4s can form. The dendritic clusters from parallel G-quartet assembly (Scheme 4D, right) may not enter the gel. It may also be that the supramolecular structures formed at pH 4.5 with K^+ are more stable.

Lastly, we provide evidence for the formation of DNA supramolecular structures directly using AFM. Figure 6A shows several small prolate ellipsoids [blobs] and V-shaped and ring-like aggregates formed at pH 4.5 in KOAc buffer by d(TG_4AC_7). The ellipsoids have diameters (height) around 3.2 nm and lengths of up to 60 nm. These are consistent with the dimensions of individual tetra-strand G4s or i-motifs, while the V-shaped and ring-like aggregates (4.5 nm in height) are consistent with larger multimers of tetra-strand G4s and i-motifs (Figure 6B) such as those depicted in Scheme 4C (64,65). As expected, these structures were not observed at pH 9.0, suggesting the dissociation of these high-order structures and possible inhibition of tetrameric structures by competing WC interactions within and between strands (Figure 6C). We propose that the nanostructures formed by d(TG_4AC_7) are mainly constructed by parallel G4s and antiparallel i-motifs. In this case, the tetramers will assemble to be a 'V-shaped' structure in first step of polymerization. Meanwhile, the 'V-shaped' structures will further assemble to be a ring-like texture.

When the counter ion is Na^+, the structures formed at pH 4.5 are less regular and only some small ellipsoids and V-shaped aggregates, which may be catenated, were ob-

served (Supplementary Data, Supplementary Figure S8a). This was in agreement with the reports that the K^+ stabilizes parallel G4s better than Na^+ (3,66).

Even larger aggregates are seen by AFM with $d(TG^{Br}GG^{Br}GAC_7)$ (Figure 6D), including V-shaped and rod-like shape aggregates with an average height of 4.0 nm. These may be quadruplex nanowires (64,67,68) with alternating G4s and i-motifs. Figure 6E indicates that the lengths of these rod-like shape aggregates can get as long as 550 nm. Just as before, the quadruplex nanowires dissociate under alkaline conditions (Figure 6F). Both the parallel C-duplexes and the two G-repeats at each end are excellent junctions to further propagate the linear supramolecular DNA structures (Figure 6D). It should be noted that the V-shaped aggregates are also observed for $d(TG^{Br}GG^{Br}GAC_7)$ in pH 4.5 KOAc solution, indicating that both parallel and antiparallel four-stranded G4s form in the presence of K^+ (Scheme 4D, right). However, only some small prolate ellipsoids and rod-like aggregates were observed in AFM image when the $d(TG^{Br}GG^{Br}GAC_7)$ strand was annealed in pH 4.5 NaOAc buffer (Supplementary Figure S8b).

We also investigated the solution-state behavior of three sequences $d(TC_5T)$, $d(TG_3AC_5)$ and $d(TG_4AC_6)$ observed by AFM. Supplementary Figure S9a–c show that only small prolate ellipsoids are formed by $d(TC_5T)$ and small amounts of V-shaped aggregates are produced by $d(TG_3AC_5)$ and $d(TG_4AC_6)$ respectively in pH 4.5 KOAc solutions, indicating that these sequences are unfavorable to form larger aggregates. Additionally, we investigated the effect of cation species and pH on the control of assembly of the supramolecular structures formed by $d(TG_4AC_7)$ and $d(TG^{Br}GG^{Br}GAC_7)$. Supplementary Figure S9d and g show that some prolate ellipsoids formed by $d(TG_4AC_7)$ and $d(TG^{Br}GG^{Br}GAC_7)$ respectively at pH 4.5 in LiOAc buffer solutions correspond to tetra-stranded i-motifs. These prolate ellipsoids can further assemble to give V-shaped or rod-like supramolecules when K^+ ions are added into the pH 4.5 LiOAc solutions as shown in Supplementary Figure S9e and h. Similar results are also obtained when $d(TG_4AC_7)$ and $d(TG^{Br}GG^{Br}GAC_7)$ are annealed in pH 9.0 KOAc solution, and then the solution pH is adjusted to 4.5 (Supplementary Figure S9f and i). Hence, we propose that the assembly of quadruplex supramolecules can be controlled or reproduced; however, the structural transitions are less efficient in comparison with the assembling of them under optimum conditions, which may indicate a slow shift process.

CONCLUSIONS

A sequence with a long C-run and a shorter G-run can simultaneously form various DNA structures, including a hairpin, a WC duplex, a G4 and an i-motif under different conditions. This polymorphic property gives such oligonucleotides the potentiality to form supramolecular nanostructures. We found sequences that support the stable formation of tetra-stranded i-motifs with G-repeats at two ends by altering the position of G-repeats, by changing the linker between G- and C-repeats, and removing the terminal base of the C-repeat. Ultimately, we find that the sequences $d(TG_mAC_n)$ ($m < n$) have the ability to form four-stranded i-motif structures with relative long G-repeats at each terminus. Two sequences $d(TG_4AC_7)$ and $d(TG^{Br}GG^{Br}GAC_7)$ were investigated under various conditions, utilizing the ability of each G-repeat to form stable parallel and antiparallel tetramolecular G4s, respectively. What is noteworthy is that both of two sequences form quadruplex nanowire structures under acidic conditions in the presence of specific metal cations. The nanowire structures formed by $d(TG_4AC_7)$ are composed of antiparallel tetrameric i-motifs and parallel tetrameric G4s, which take on a V-shaped conformation (most likely two i-motif tetra-strand units and one G4 tetra-strand unit; Scheme 4C middle) and ring-like conformations (including additional i-motif and G4 tetra-strand units; Scheme 4C right). These aggregates resemble the products of DNA origami. On the other hand, the nanowire structures formed by $d(TG^{Br}GG^{Br}GAC_7)$ are made up of antiparallel tetra-strand i-motifs and antiparallel G4s to give linear DNA pillars stabilized by H^+ and Na^+ ions. Studies have shown that a single G-strand consisting of a number of 4*n* G-repeats can form higher-ordered structures containing *n* G4s (69–71). These require the synthesis of very long DNA pieces. Here, our results demonstrated that four-stranded G4s or (and) i-motifs formed from relatively short (G + C)-rich oligonucleotide sequence could also self-assemble to drive supramolecular structures under specific conditions. Furthermore, the association and dissociation of these nanostructures can be modulated by solution pH and metal cations, which may provide a mechanism for controllable molecular machines.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

FUNDING

National Natural Science Foundation of China [21675060, 51273080]; International Collaboration Project of Jilin Province [20160414010GH]; Open Project of State Key Laboratory for Supramolecular Structure and Materials [SKLSSM201610]. Funding for open access charge: National Natural Science Foundation of China [21675060 and 51273080].
Conflict of interest statement. None declared.

REFERENCES

1. Wu,Y. and Brosh,R.M. (2010) G-quadruplex nucleic acids and human disease. *FEBS J.*, **277**, 3470–3488.
2. Zhou,J. and Yuan,G. (2007) Specific recognition of human telomeric G-quadruplex DNA with small molecules and the conformational analysis by ESI mass spectrometry and circular dichroism spectropolarimetry. *Chemistry*, **13**, 5018–5023.
3. Smargiasso,N., Rosu,F., Hsia,W., Colson,P., Baker,E.S., Bowers,M.T., De Pauw,E. and Gabelica,V. (2008) G-quadruplex DNA assemblies: loop length, cation identity, and multimer formation. *J. Am. Chem. Soc.*, **130**, 10208–10216.
4. Arora,A. and Maiti,S. (2009) Stability and molecular recognition of quadruplexes with different loop length in the absence and presence of molecular crowding agents. *J. Phys. Chem. B*, **113**, 8784–8792.
5. 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.

6. Guéron,M. and Leroy,J.-L. (2000) The i-motif in nucleic acids. *Curr. Opin. Struct. Biol.*, **10**, 326–331.
7. Xu,Y. and Sugiyama,H. (2006) Formation of the G-quadruplex and i-motif structures in retinoblastoma susceptibility genes (Rb). *Nucleic Acids Res.*, **34**, 949–954.
8. Benabou,S., Aviñó,A., Eritja,R., González,C. and Gargallo,R. (2014) Fundamental aspects of the nucleic acid i-motif structures. *RSC Adv.*, **4**, 26956–26980.
9. Huppert,J.L. (2008) Four-stranded nucleic acids: structure, function and targeting of G-quadruplexes. *Chem. Soc. Rev.*, **37**, 1375–1384.
10. Williamson,J.R. (1994) G-quartet structures in telomeric DNA. *Annu. Rev. Biophys. Biomol. Struct.*, **23**, 703–730.
11. Kotch,F.W., Fettinger,J.C. and Davis,J.T. (2000) A lead-filled G-quadruplex: insight into the G-quartet's selectivity for Pb^{2+} over K^{+}. *Org. Lett.*, **2**, 3277–3280.
12. Berger,I., Kang,C.H., Fredian,A., Ratliff,R., Moyzis,R. and Rich,A. (1995) Extension of the four-stranded intercalated cytosine motif by adenine•adenine base pairing in the crystal structure of d(CCCAAT). *Nat. Struct. Biol.*, **2**, 416–425.
13. Leroy,J.-L., Snoussi,K. and Guéron,M. (2001) Investigation of the energetics of C-H···O hydrogen bonds in the DNA i-motif via the equilibrium between alternative intercalation topologies. *Magn. Reson. Chem.*, **39**, S171–S176.
14. Leroy,J.L. (2009) The formation pathway of i-motif tetramers. *Nucleic Acids Res.*, **37**, 4127–4134.
15. MacLeod,M.C., Johnston,D.A., LaBate,M. and White,R.A. (1996) The probability of occurrence of oligomer motifs in the human genome and genomic microheterogeneity. *J. Theor. Biol.*, **181**, 311–318.
16. Drew,H., Takano,T., Tanaka,S., Itakura,K. and Dickerson,R.E. (1980) High-salt d(CpGpCpG), a left-handed Z′ DNA double helix. *Nature*, **286**, 567–573.
17. Bansal,A., Prasad,M., Roy,K. and Kukreti,S. (2012) A short GC-rich palindrome of human mannose receptor gene coding region displays a conformational switch. *Biopolymers*, **97**, 950–962.
18. Deng,H. and Braunlin,W.H. (1995) Duplex to quadruplex equilibrium of the self-complementary oligonucleotide d (GGGGCCCC). *Biopolymers*, **35**, 677–681.
19. Penázová,H. and Vorlickova,M. (1997) Guanine tetraplex formation by short DNA fragments containing runs of guanine and cytosine. *Biophys. J.*, **73**, 2054–2063.
20. Zamiri,B., Mirceta,M., Bomsztyk,K., Macgregor,R.B. Jr and Pearson,C.E. (2015) Quadruplex formation by both G-rich and C-rich DNA strands of the C9orf72 $(GGGGCC)_8$•$(GGCCCC)_8$ repeat: effect of CpG methylation. *Nucleic Acids Res.*, **43**, 10055–10064.
21. Kypr,J., Fialová,M., Chládková,J., Tůmová,M. and Vorlíčková,M. (2001) Conserved guanine-guanine stacking in tetraplex and duplex DNA. *Eur. Biophys. J.*, **30**, 555–558.
22. Hardin,C.C., Watson,T., Corregan,M. and Bailey,C. (1992) Cation-dependent transition between the quadruplex and Watson-Crick hairpin forms of $d(CGCG_3GCG)$. *Biochemistry*, **31**, 833–841.
23. Hardin,C.C., Corregan,M., Brown,B.A. and Frederick,L.N. (1993) Cytosine-cytosine^{+} base pairing stabilizes DNA quadruplexes and cytosine methylation greatly enhances the effect. *Biochemistry*, **32**, 5870–5880.
24. Dhakal,S., Yu,Z., Konik,R., Cui,Y., Koirala,D. and Mao,H. (2012) G-quadruplex and i-motif are mutually exclusive in ILPR double-stranded DNA. *Biophys. J.*, **102**, 2575–2584.
25. Cui,Y., Kong,D., Ghimire,C., Xu,C. and Mao,H. (2016) Mutually exclusive formation of G-quadruplex and i-motif is a general phenomenon governed by steric hindrance in duplex DNA. *Biochemistry*, **55**, 2291–2299.
26. Yang,Y., Zhou,C., Zhang,T., Cheng,E., Yang,Z. and Liu,D. (2012) DNA pillars constructed from an i-motif stem and duplex branches. *Small*, **8**, 552–556.
27. Winfree,E., Liu,F., Wenzler,L.A. and Seeman,N.C. (1998) Design and self-assembly of two-dimensional DNA crystals. *Nature*, **394**, 539–544.
28. Pinheiro,A.V., Han,D., Shih,W.M. and Yan,H. (2011) Challenges and opportunities for structural DNA nanotechnology. *Nat. Nanotechnol.*, **6**, 763–772.
29. Chen,J. and Seeman,N.C. (1991) Synthesis from DNA of a molecule with the connectivity of a cube. *Nature*, **350**, 631–633.
30. Yang,H., McLaughlin,C.K., Aldaye,F.A., Hamblin,G.D., Rys,A.Z., Rouiller,I. and Sleiman,H.F. (2009) Metal-nucleic acid cages. *Nat. Chem.*, **1**, 390–396.
31. Rothemund,P.W., Papadakis,N. and Winfree,E. (2004) Algorithmic self-assembly of DNA Sierpinski triangles. *PLoS Biol.*, **2**, 2041–2053.
32. Sun,X., Hyeon Ko,S., Zhang,C., Ribbe,A.E. and Mao,C. (2009) Surface-mediated DNA self-assembly. *J. Am. Chem. Soc.*, **131**, 13248–13249.
33. Dong,Y., Yang,Z. and Liu,D. (2014) DNA nanotechnology based on i-motif structures. *Acc. Chem. Res.*, **47**, 1853–1860.
34. Mei,H., Budow,S. and Seela,F. (2012) Construction and assembly of chimeric DNA: oligonucleotide hybrid molecules composed of parallel or antiparallel duplexes and tetrameric i-motifs. *Biomacromolecules*, **13**, 4196–4204.
35. Laisné,A., Pompon,D. and Leroy,J.-L. (2010) $[C_7GC_4]_4$ association into supra molecular i-motif structures. *Nucleic Acids Res.*, **38**, 3817–3826.
36. Guittet,E., Renciuk,D. and Leroy,J.-L. (2012) Junctions between i-motif tetramers in supramolecular structures. *Nucleic Acids Res.*, **40**, 5162–5170.
37. Cao,Y., Qin,Y., Bruist,M., Gao,S., Wang,B., Wang,H. and Guo,X. (2015) Formation and dissociation of the interstrand i-motif by the sequences $d(X_nC_4Y_m)$ monitored with electrospray ionization mass spectrometry. *J. Am. Soc. Mass Spectrom.*, **26**, 994–1003.
38. Canalia,M. and Leroy,J.-L. (2009) $[5mCCTCTCTCC]_4$: an i-motif tetramer with intercalated T•T pairs. *J. Am. Chem. Soc.*, **131**, 12870–12871.
39. Webba da Silva,M., Trajkovski,M., Sannohe,Y., Ma'ani Hessari,N., Sugiyama,H. and Plavec,J. (2009) Design of a G-quadruplex topology through glycosidic bond angles. *Angew. Chem. Int. Ed.*, **48**, 9167–9170.
40. Balthasart,F., Plavec,J. and Gabelica,V. (2013) Ammonium ion binding to DNA G-quadruplexes: do electrospray mass spectra faithfully reflect the solution-phase species? *J. Am. Soc. Mass Spectrom.*, **24**, 1–8.
41. Gabelica,V., Pauw,E.D. and Rosu,F. (1999) Special feature: tutorial-meeting report-interaction between antitumor drugs and a double-stranded oligonucleotide studied by electrospray ionization mass spectrometry. *J. Mass Spectrom.*, **34**, 1328–1337.
42. Mergny,J.-L., Phan,A.-T. and Lacroix,L. (1998) Following G-quartet formation by UV-spectroscopy. *FEBS Lett.*, **435**, 74–78.
43. Mergny,J.-L. and Lacroix,L. (2003) Analysis of thermal melting curves. *Oligonucleotides*, **13**, 515–537.
44. Changenet-Barret,P., Emanuele,E., Gustavsson,T., Improta,R., Kotlyar,A.B., Markovitsi,D., Vaya,I., Zakrzewska,K. and Zikich,D. (2010) Optical properties of guanine nanowires: experimental and theoretical study. *J. Phys. Chem. C*, **114**, 14339–14346.
45. Sen,D. and Gilbert,W. (1992) Guanine quartet structures. *Methods Enzymol.*, **211**, 191–199.
46. Tran,P.L., De Cian,A., Gros,J., Moriyama,R. and Mergny,J.L. (2013) Tetramolecular quadruplex stability and assembly. *Top. Curr. Chem.*, **330**, 243–273.
47. Kypr,J., Kejnovská,I., Renčiuk,D. and Vorlíčková,M. (2009) Circular dichroism and conformational polymorphism of DNA. *Nucleic Acids Res.*, **37**, 1713–1725.
48. Dhakal,S., Lafontaine,J.L., Yu,Z., Koirala,D. and Mao,H. (2012) Intramolecular folding in human ILPR fragment with three C-rich repeats. *PLoS One*, **7**, e39271.
49. Kaushik,M., Kukreti,R., Grover,D., Brahmachari,S.K. and Kukreti,S. (2003) Hairpin–duplex equilibrium reflected in the A→B transition in an undecamer quasi-palindrome present in the locus control region of the human β-globin gene cluster. *Nucleic Acids Res.*, **31**, 6904–6915.
50. Bardin,C. and Leroy,J.L. (2008) The formation pathway of tetramolecular G-quadruplexes. *Nucleic Acids Res.*, **36**, 477–488.
51. Rosu,F., Gabelica,V., Poncelet,H. and De Pauw,E. (2010) Tetramolecular G-quadruplex formation pathways studied by electrospray mass spectrometry. *Nucleic Acids Res.*, **38**, 5217–5225.
52. Zhu,H. and Lewis,F.D. (2007) Pyrene excimer fluorescence as a probe for parallel G-quadruplex formation. *Bioconjugate Chem.*, **18**, 1213–1217.

53. Senior,M.M., Jones,R.A. and Breslauer,K.J. (1988) Influence of loop residues on the relative stabilities of DNA hairpin structures. *Proc. Natl. Acad. Sci. U.S.A.*, **85**, 6242–6246.
54. Lu,M., Guo,Q. and Kallenbach,N.R. (1992) Structure and stability of sodium and potassium complexes of dT_4G_4 and dT_4G_4T. *Biochemistry*, **31**, 2455–2459.
55. Leroy,J.L., Gehring,K., Kettani,A. and Gueron,M. (1993) Acid multimers of oligodeoxycytidine strands: stoichiometry, base-pair characterization, and proton exchange properties. *Biochemistry*, **32**, 6019–6031.
56. Watson,J.D. and Crick,F.H. (1953) Molecular structure of nucleic acids. *Nature*, **171**, 737–738.
57. Mergny,J.L., De Cian,A., Ghelab,A., Sacca,B. and Lacroix,L. (2005) Kinetics of tetramolecular quadruplexes. *Nucleic Acids Res.*, **33**, 81–94.
58. Frank-Kamenetskii,M.D. and Mirkin,S.M. (1995) Triplex DNA structures. *Annu. Rev. Biochem.*, **64**, 65–95.
59. Virgilio,A., Esposito,V., Citarella,G., Pepe,A., Mayol,L. and Galeone,A. (2012) The insertion of two 8-methyl-2′-deoxyguanosine residues in tetramolecular quadruplex structures: trying to orientate the strands. *Nucleic Acids Res.*, **40**, 461–475.
60. Uesugi,S. and Ikehara,M. (1977) Carbon-13 magnetic resonance spectra of 8-substituted purine nucleosides. Characteristic shifts for the *syn* conformation. *J. Am. Chem. Soc.*, **99**, 3250–3253.
61. Esposito,V., Randazzo,A., Piccialli,G., Petraccone,L., Giancola,C. and Mayol,L. (2004) Effects of an 8-bromodeoxyguanosine incorporation on the parallel quadruplex structure $[d(TGGGT)]_4$. *Org. Biomol. Chem.*, **2**, 313–318.
62. Virgilio,A., Esposito,V., Randazzo,A., Mayol,L. and Galeone,A. (2005) 8-methyl-2′-deoxyguanosine incorporation into parallel DNA quadruplex structures. *Nucleic Acids Res.*, **33**, 6188–6195.
63. Dutta,K., Fujimoto,T., Inoue,M., Miyoshi,D. and Sugimoto,N. (2010) Development of new functional nanostructures consisting of both DNA duplex and quadruplex. *Chem. Commun. (Camb.)*, **46**, 7772–7774.
64. Ghodke,H.B., Krishnan,R., Vignesh,K., Kumar,G., Narayana,C. and Krishnan,Y. (2007) The I-tetraplex building block: rational design and controlled fabrication of robust 1D DNA scaffolds through non-Watson–Crick interactions. *Angew. Chem. Int. Ed.*, **119**, 2700–2703.
65. Chiorcea-Paquim,A.M., Santos,P.V., Eritja,R. and Oliveira-Brett,A.M. (2013) Self-assembled G-quadruplex nanostructures: AFM and voltammetric characterization. *Phys. Chem. Chem. Phys.*, **15**, 9117–9124.
66. Wong,A. and Wu,G. (2003) Selective binding of monovalent cations to the stacking G-quartet structure formed by guanosine 5′-monophosphate: a solid-state NMR study. *J. Am. Chem. Soc.*, **125**, 13895–13905.
67. Kunstelj,K., Federiconi,F., Spindler,L. and Drevensek-Olenik,I. (2007) Self-organization of guanosine 5′-monophosphate on mica. *Colloids Surf. B. Biointerfaces*, **59**, 120–127.
68. Hessari,N.M., Spindler,L., Troha,T., Lam,W.C., Drevensek-Olenik,I. and da Silva,M.W. (2014) Programmed self-assembly of a quadruplex DNA nanowire. *Chemistry*, **20**, 3626–3630.
69. Wang,H., Nora,G.J., Ghodke,H. and Opresko,P.L. (2011) Single molecule studies of physiologically relevant telomeric tails reveal POT1 mechanism for promoting G-quadruplex unfolding. *J. Biol. Chem.*, **286**, 7479–7489.
70. Petraccone,L., Spink,C., Trent,J.O., Garbett,N.C., Mekmaysy,C.S., Giancola,C. and Chaires,J.B. (2011) Structure and stability of higher-order human telomeric quadruplexes. *J. Am. Chem. Soc.*, **133**, 20951–20961.
71. Kankia,B. (2014) Tetrahelical monomolecular architecture of DNA: a new building block for nanotechnology. *J. Phys. chem. B*, **118**, 6134–6140.