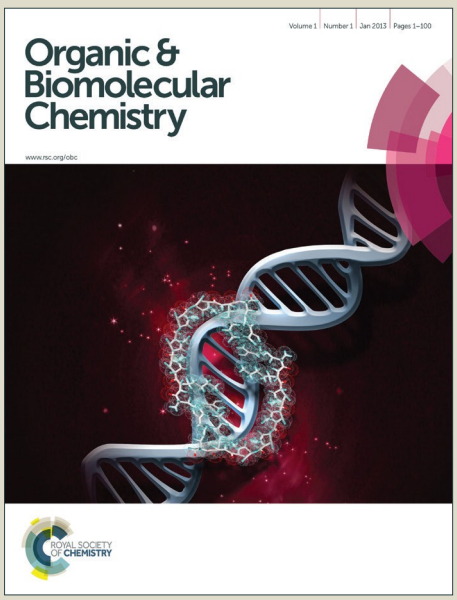
Organic & Biomolecular Chemistry

Accepted Manuscript



ROYAL SOCIETY OF CHEMISTRY

www.rsc.org/obc

Journal Name

ARTICLE

Unimolecular antiparallel G-quadruplex folding topology of 2'-5'-isoTBA sequences remains unaltered by loop composition

Received 00th January 20xx,
Accepted 00th January 20xx

DOI: 10.1039/x0xx00000x

www.rsc.org/

Manisha N. Aher, Namrata D. Erande, Moneesha Fernandes and Vaijayanti A. Kumar*

2'-5'-linked isoTBA 15mer sequence with (232) loop composition formed stable antiparallel quadruplex structures similar to the SELEX derived 15mer TBA sequence with (232) loop composition. Parallel versus antiparallel topology of 3'-5'-G-quadruplexes is largely dictated by the loop length, and it is known that the truncated loops favour parallel quadruplexes. In contrast to TBA, systematic reduction of loop length in isoTBA from (232) to (222), (131) or even (111) did not alter the antiparallel topology of the resulting 14mer, 13mer and 11mer G-rich modified isoTBA-like sequences.

Introduction

G-quadruplexes have been gaining increasing attention in the recent past because of their abundance *in vivo* and their role in meiosis, telomere maintenance, gene regulation[1] and also as targets for cancer treatments. [2] The G-quadruplex structures, their folding mechanism, protein binding to the quadruplex structures and ligands targeting the quadruplex over duplex were extensively studied by various experimental and theoretical studies.[3] Many therapeutically important G-quadruplex aptamers have been derived by SELEX and show high affinity towards biologically important targets. Examples include the thrombin-binding aptamer (TBA), anti-cancer and anti-HIV aptamers.[4,5,6] The high affinity with the target is derived from the fitting of the three dimensional shape of the quadruplex to the binding site and an array of electrostatic/hydrogen bonding interaction possibilities created in the loop region, as a consequence of the folded structures.

The thrombin-binding aptamer (TBA; d(GGTTGGTGTGGTTGG)) binds to thrombin, a large protein involved in the blood coagulation cascade.[7] The unimolecular structure of the G-quadruplex formed by TBA is well studied and is composed of two stacked G-quartets and two external TT loops and one central TGT loop, resulting in a chair-like conformation, stabilized by monovalent cations such as Na^+ and K^+.[8] The TT loops were found to be sufficient to span the narrow grooves and the TGT loop sufficed to span the wider groove of the quadruplex. The two TT loops of the quadruplex are currently believed to be interacting with thrombin anion exosite I in 1:1 stoichiometry.[9] Several chemical modifications of TBA are reported in the literature in the view of further optimizing its binding to thrombin and also to improve its stability[10] against cellular nucleases. The nucleotides in the loop region as well as the length of the loop control the topology and stability of the quadruplex structures in general and also in the case of TBA.[11,12] Shorter loop lengths (less than total 6 nucleotides in the loop) are known to destabilize the intramolecular, antiparallel folded G-quadruplex of TBA and allow only the parallel intermolecular four-stranded G-quadruplex structures. As a consequence, the thrombin binding activity of TBA sequences in which the loops are short is adversely affected.[12] In contrast, releasing the rigidity in the loop region by introducing unlocked nucleic acid monomer at selective positions in the loop region of TBA, stabilized of the intramolecular folded TBA quadruplex structure.[13]

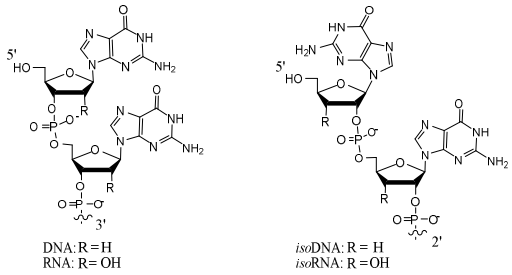


Fig. 1. Genetic 3'-5' DNA/RNA with all bases in '*anti*' conformation, and Non-genetic 2'-5' isoDNA/isoRNA showing 5' terminal guanine '*syn*' conformation.

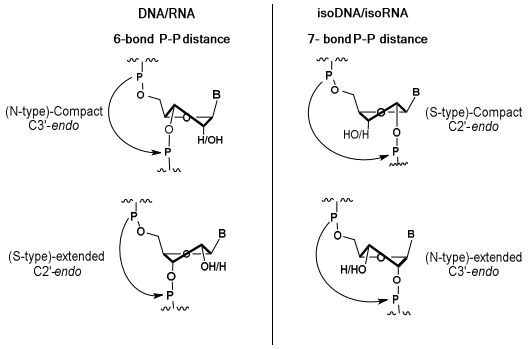


Fig. 2 Extended and compact backbone geometry in DNA/RNA and isoDNA/isoRNA.

Dr. V. A. Kumar, Dr. N. D. Erande, Dr. M. Fernandes and M. N. Aher
Organic Chemistry Division
CSIR-National Chemical Laboratory
Dr. Homi Bhabha Road, Pune 411008, India
E-mail: va.kumar@ncl.res.in

†Electronic supplementary information (ESI) available: Gel electrophoresis, representative CD-melting curves, CD-hysteresis curves, Thrombin binding CD-spectras and MALDI-TOF spectras

The 3'-deoxy-2'-5'-linked non-genetic isoDNA (Figure 1) exhibits high stability against cellular enzymes[14] and therefore it could be a promising alternative nucleic acid for developing DNA therapeutics. Formation of duplex and triplex structures comprising isoDNA and complementary RNA are well-documented in the literature.[15,23] In our previous study, we discovered that employing these nuclease resistant, homogeneous 2'-5' linkers in the sequence of TBA (isoTBA) could also form unimolecular folded G-quadruplex structures and could exhibit thrombin binding and anti-coagulant properties.[16] In 3'-deoxy-isoDNA, the phosphodiester linkages are C2'-C5' instead of C3'-C5' linkages in native DNA (Figure 1). This leads to an increase in the number of bonds between O5' and 3'-O-phosphorus from six in DNA to seven in isoDNA. Also, the 2'-5' linkages maintain an extended backbone geometry, as the anomeric effect and the O4'-C1'-C2'-O2' gauche effect on the substituted sugar favours N-type sugar conformation (Figure 2).[17] This extended backbone with increased number of bonds in isoTBA compared to TBA might be responsible for the relatively less stable G-quadruplex structure compared to TBA. In this particular case therefore, the flexibility in the loop region probably needs to be addressed.

We reported earlier that replacement of the central TGT loop by UGU loop, stabilized the G-quadruplex structure of isoTBA.[16] The 3'-hydroxy group in uridine would disfavour the extended N-type sugar conformation (Figure 2) compared to the 3'-deoxyuridine, due to the additional O4'-C4'-C3'-O3' gauche effect and exert rigidity in the TGT loop region due to favoured compact S-type geometry of the sugar (Figure 2).[17] We therefore envisaged that unlike in TBA, isoTBA may benefit from decreasing the loop length and we thought of exploring the effect of systematically studying the effect of decreasing the length in the loop region of isoTBA. We also synthesized the earlier reported TBA sequences with decreasing loop length, for comparison.[12] This would enable us to study the effect of the reduced loop length on the topology of the G-quadruplexes formed by different isoTBA-like sequences and compare it with isosequential TBA. We report here the unique antiparallel topology and thermal stability of all the isoTBA sequences using ultraviolet (UV) absorption, Circular dichroism (CD) spectroscopy, and Gel electrophoresis studies. The chaperone effect of thrombin to induce unomolecular antiparallel fold in these oligomers is also investigated.

Results and Discussion

Guanine-rich DNA sequences tend to form four-stranded helical structures called G-quadruplexes. Depending on the direction of the guanine-containing strands and presence of *syn* or *anti* glycosidic bond angle (GBA) of guanine, the topology of G-quadruplexes are classified in three groups: parallel (group I), mixed or hybrid (group II) and antiparallel (group III).[18] In parallel (group I) quadruplexes, all the guanosine residues adopt the same GBA (all either *anti* or *syn*) and all the four strands have the same orientation. The mixed or hybrid (group II) quadruplexes contain both, sequences of guanosines with the same type of GBA (such as *anti-anti* and *syn-syn*), as well as different types (such as *syn-anti* and *anti-syn* steps). In contrast, antiparallel G-quadruplexes (group-III) exhibit both *anti* and *syn* guanines with *syn-anti* or *anti-syn* steps. As both group-II and group-III quadruplexes possess at least one of the four strands oriented antiparallel to the others, both are considered as types of antiparallel quadruplexes. Various factors affect the topology of G-quadruplexes, such as loop length, type of nucleobase in the loop, glycosidic bond angle of guanine, etc.[19,20] Among all these, loops play a key role in determining the nature of the folding of G-quadruplexes. Balasubramanian *et al.* studied various G-quadruplex-forming sequences with varying loop lengths using UV and CD spectroscopy, molecular modelling and simulations to confirm that parallel is the only possible conformation for G-quadruplexes with shorter loops (i.e., total number of nucleotides in the loops less than 6).[19]

We synthesized oligomers with decreasing loop lengths (having total number of loop residues 7, 6, 5 and 3) containing both 3'-5'-(TBA) and 2'-5'-phosphodiester-linked (isoTBA) backbone. The synthesized oligomeric sequences and their MALDI-TOF characterization data are listed in Table 1 (ESI Figure S9).

Table 1. TBA and isoTBA loop-modified sequences, MALDI-TOF mass analysis

Entry No.	Code	Sequence	MALDI-TOFMass Calc./Obs.
1	TBA232(7)	5' GG**TT**GG**T**G**T**GG**TT**GG 3'	4726/4727
2	TBA222(6)	5' GG**TT**GG**TT**GG**TT**GG 3'	4396/4396
3	TBA131(5)	5' GG**T**GG**T**G**T**GG**T**GG 3'	4117/4116
4	TBA111(3)	5' GG**T**GG**T**GG**T**GG 3'	3484/3481
5	isoTBA232(7)	5' GG**TT**GG**T**G**T**GG**TT**GG 2'	4726/4731
6	isoTBA222(6)	5' GG**TT**GG**TT**GG**TT**GG 2'	4396/4392
7	isoTBA131(5)	5' GG**T**GG**T**G**T**GG**T**GG 2'	4117/4121
8	isoTBA111(3)	5' GG**T**GG**T**GG**T**GG 2'	3484/3483

The three-digit number in the sequence code is an indication of the number of residues in each of the three loops, from the 5' to the 3'/2' end and the number in parentheses indicates the total number of loop residues. For example, TBA232(7) bears two residues each in the TT loops and 3 residues in the central TGT loop, making the total number of loop residues 7. Loop residues are indicated in bold letters.

G-quadruplex formation of all the loop-modified sequences was evaluated by CD spectroscopy. CD spectroscopy was also used to explore the parallel, mixed or hybrid and antiparallel G-quadruplex folding topology based on well-characterized CD patterns corresponding to the topology of the G-quadruplexes.[21] The thermal stability of the loop-modified TBA and isoTBA sequences was evaluated by CD melting experiments at different strand concentration as well as with different salt concentrations. Thrombin-binding efficiency of TBA232(7) and all the isoTBA sequences along with their thermal stability experiments was also done using CD spectroscopy. The difference in the UV absorption before and after melting (thermal difference spectrum, TDS) is known to be a "fingerprint" spectrum of quadruplexes.[22] We recorded the TDS for all the sequences to verify their folding topology. These results were further supplemented by denaturing and non-denaturing gel electrophoresis.

Evaluation of G-quadruplex formation using CD spectroscopy

In de-ionized water alone, the CD spectrum of only the 15mer TBA232(7) exhibited a weak positive signal at ~292 nm, showing its propensity to form a unimolecular antiparallel G-quadruplex even in the absence of cations. All the other loop-modified TBA and isoTBA sequences failed to show the positive CD band ~290 nm, and

Table 2. CD analysis of TBA and isoTBA loop-modified sequences

Code	5 μM in H_2O at 5° C		5 μM in 100 mM K^+ at 5° C		20 μM in 100 mM K^+ at 5° C		5 μM in 500 mM K^+ at 5° C		Topology
	Maxima (nm)	Minima (nm)	Maxima (nm)	Minima (nm)	Maxima (nm)	Minima (nm)	Maxima (nm)	Minima (nm)	
TBA232(7)	292	255	295, 245	262	295, 245	262	292, 245	262	Antiparallel
TBA222(6)	280, 250	265	292 ,252	270	292, 252	275	292, 252	275	Hybrid or mixture of antiparallel and parallel
TBA131(5)	280, 265	235	295, 275, 260	235	295, 275, 260	235	295, 275, 260	235	Hybrid or mixture of antiparallel and parallel
TBA111(3)	255	230	260	235	260	235	255	230	Parallel
*iso*TBA232(7)	260	295	295, 250	268	295, 250	268	295 , 240	260	Antiparallel
*iso*TBA222(6)	260	294	291, 243	265	291, 243	265	291 , 243	265	Antiparallel
*iso*TBA131(5)	280-258	298 , 230	292, 245	265	292-245	265	292, 245	265	Antiparallel
*iso*TBA111(3)	250	295	290, 250	265,	295, 245	265	290, 245	270,	Antiparallel

were unable to adopt a quadruplex structure in the absence of cations (Figure 3).

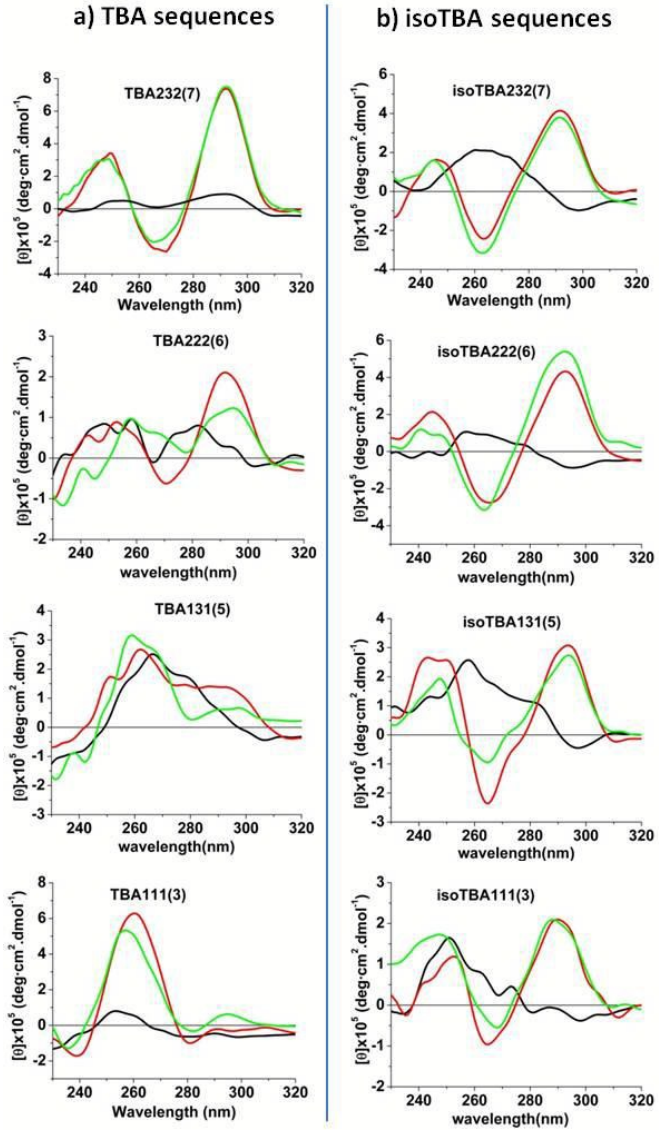


Fig. 3 CD spectra of the loop-modified sequences (**Black**-5 μM in H_2O, **Red**-5 μM in 100 mM K^+, and **Green**-5 μM in 500 mM K^+)

In the presence of K^+ ions (5 μM strand concentration, 100 mM K^+, Figure 3, Table 2), the unmodified 3'-5' TBA232(7) showed a maximum at 295 nm and a minimum at 265 nm, indicative of a unimolecular antiparallel G-quadruplex.[18] Decreasing the loop length in TBA is known to disturb the unimolecular TBA quadruplex structure and our results were similar to those reported earlier for TBA222(6) and TBA131(5).[12] These two sequences with intermediate loop lengths of six and five residues respectively displayed CD spectra with two positive bands near 292 nm and 262 nm and a negative peak at 235 nm indicating either a hybrid-type (group II) quadruplex conformation or the co-existence of multimolecular-parallel and unimolecular-antiparallel conformations. The 3'-5'-linked sequence with the shortest loops, TBA111(3), having a total of three thymine loop residues, exhibited maxima and minima at 260 nm and 235 nm respectively, indicating the group-I multimolecular parallel G-quadruplex topology.[12] Thus, in concurrence with the previous reports and loop-length phenomenon studied by Balasubramanian,[19] we observed that in case of TBA, as the loop length decreased from 232(7) to 111(3), the quadruplex topology changed from intramolecular-antiparallel to multimolecular-parallel structures. The shorter loop lengths could not support the unimolecular antiparallel quadruplex topology. We further studied the CD spectra of the isoTBA sequences with shorter loops in comparison to TBA (Figure 3). The results are tabulated in Table 2. Contrary to the results obtained for the loop-modified TBA, we observed that irrespective of loop length, all the modified *iso*TBA sequences containing 222(6), 131(5) or even 111(3) loop nucleotides, exhibited a strong positive CD signal at 295 nm and minima at ~260 nm in the presence of K^+ ions, characteristic of the unimolecular antiparallel G-quadruplex conformation. In the case of 2'-5'-linked isoTBA, extended N-type conformation of nucleotides and presence of larger P-P distance[17] might decrease the strain caused by shorter loop lengths and allow the *iso*TBA to assume an antiparallel G-quadruplex structure. Another reason for the preferred antiparallel conformation for the isoTBA sequences could be the preferred *syn* and *anti* conformations for the 5'-end- and penultimate nucleobase respectively, of 2'-5'-linked oligomers, as evident from NMR[17,23] and X-ray crystal structural studies.[24] This conformational restriction of 2'-5'-linked isoTBA oligomers may result in them not conforming to

the 'all *syn*' or 'all *anti*' nucleobase conformation, a prerequisite for the formation of parallel quadruplex structures.

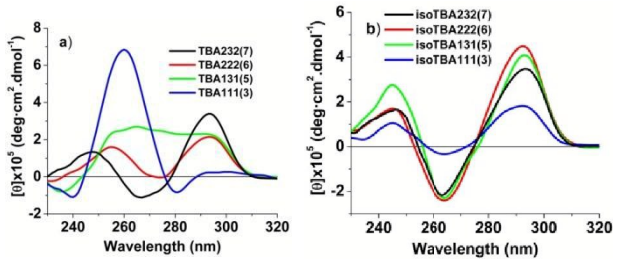


Fig. 4 CD spectra of all the loop-modified sequences at 20 µM concentration in 100 mM K^+ a) TBA sequences b) isoTBA sequences

The CD patterns for all the TBA and isoTBA sequences were studied at higher salt concentration (Figure 3) and at increased strand concentrations[25] (Figure 4) to see if this affected the unique antiparallel nature of the G-quadruplexes formed by isoTBA sequences in comparison with those by TBA sequences. In the case of TBA232(7), only increase in the intensity of bands at 295 nm and 265 nm was seen, with the overall CD pattern remaining unchanged. In the loop-restricted TBA sequences TBA222(6), TBA131(5) and TBA111(3), the increased intensity of the CD signal at 260 nm was observed in the CD spectra at higher strand concentration, with decreasing loop length, indicating higher inclination to form multimolecular parallel structures (Figure 4a). In contrast, the CD maximum at 260 nm corresponding to multimolecular parallel structure was not observed even at the higher salt or strand concentrations studied herein for 2'-5'-linked isoTBA sequences with reduced loop lengths (Figures 3 and 4b), confirming that these G-quadruplexes were indeed of unimolecular antiparallel topology. In the case of isoTBA sequences with shorter loops isoTBA222(6) and isoTBA131(5), the intensity of the CD signals at 295 nm remained comparable to that of isoTBA232(7). For the sequence isoTBA111(3) however, the intensity of the CD signal at 295nm was found to be comparatively less.

Gel Experiments

To confirm the structural and topological behaviour of the synthesized sequences, we performed denaturing and non-denaturing polyacrylamide gel electrophoresis experiments. In the denaturing gel (ESI, Figure S1), the mobility of the oligonucleotides displayed molecular weight dependency and marginally higher mobility was seen for TBA111(3) and isoTBA111(3). The non-denaturing gel (Figure 5) clearly displayed the differences caused by the differing loop length on the quadruplex topology.

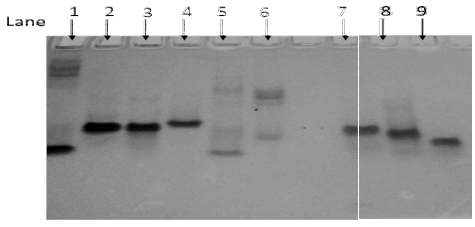


Fig. 5 Non-denaturing polyacrylamide gel mobility assay of all the TBA and isoTBA sequences in phosphate buffer (10mM, pH7.2) containing 500mM KCl. Lane1-Scrambled TBA sequence 5'-GGTGGTGGTTGTGGT-3' in water, without any added cations, Lane2-TBA232(7), Lane3-isoTBA232(7), Lane4-TBA222(6), Lane5-TBA131(5), Lane6-TBA111(3), Lane7-isoTBA222(6), Lane8-isoTBA131(5), Lane9-isoTBA111(3). The bands were visualized by UV-shadowing.

In Figure 5, Lane 1 contains the scrambled TBA sequence that lacks the G-stretches required to form the typical unimolecular antiparallel TBA quadruplex structure, but has the same nucleotide composition. Lanes 2 and 3 show the similar mobility of the 15mer unmodified TBA232(7) and isoTBA232(7), corresponding to the unimolecular antiparallel quadruplex in the presence of K^+. Lane 4 also shows the formation of unimolecular quadruplex by TBA222(6), although the CD spectrum indicated a hybrid structure for this sequence. The multiple bands in lane 5 correspond to TBA131(5). Lane 6 shows highly retarded major bands of multimolecular parallel quadruplex for TBA111(3) as seen also in CD spectral studies. Lanes 7, 8 and 9 correspond to isoTBA222(6), isoTBA131(5) and isoTBA111(3). None of the isoTBA sequences show any major retarded band in the gel, indicating the absence of multimolecular parallel quadruplexes and confirming the formation of unimolecular antiparallel quadruplexes in these cases, as also seen from their CD spectral studies.

Evaluation of G-quadruplex thermal stability using CD spectroscopy

The stability of the G-quadruplexes was followed by temperature-dependent change in the amplitude of the CD signal at 295 nm for antiparallel quadruplexes formed by isoTBA sequences in comparison with TBA232(7). The results were also compared with change in the amplitude of the CD signal at 260nm, for the parallel quadruplex formed by TBA111(3) (Table 3). At 5 µM strand concentration in the presence of 100 mM K^+, the *iso*TBA232(7) structure was found to be less stable than the control TBA232(7) (Table 3, ΔT_m = -13 °C, ESI Figure S2-a).[16] All other shorter loop-containing antiparallel G-quadruplexes of isoTBA showed a decrease in stability with decreasing loop length. isoTBA222(6) and isoTBA131(5) were equally stable, with T_m 5 °C less than that of isoTBA232(7) (ESI Figure S3-a). TBA111(3) the sequence with the shortest loop, existed in the parallel multimolecular quadruplex conformation, consequently showed the highest thermal stability compared to all other sequences (ΔT_m = +5 °C compared to TBA232(7)). In contrast, isoTBA111(3), containing the same nucleobases in the loop showed the lowest stability among all the sequences synthesized. Thus, although the extended 2'-5' linkages help to reduce the strain of the shortened loops compared to 3'-5' linkages and allow the formation of unimolecular folded structures, the formed structures were less stable compared to the antiparallel quadruplexes formed by other isoTBA and TBA sequences.

The inter- Vs intramolecular nature of the quadruplexes was confirmed by carrying out CD melting experiments at higher strand concentration (20 µM). The melting temperatures of all the isoTBA sequences were independent of concentration, similar to the TBA232(7) (at 5 µM and 20 µM; Table 3, ESI Figure S2-b & S3-b), confirming that these sequences indeed formed intramolecular (monomeric) folded structures. For all the CD melting experiments of antiparallel sequences, the hysteresis between the heating and the cooling curves was found to be negligible, again indicating unimolecular G-quadruplexes (Table 3 and ESI, Figure S4, S5).[26] The sequence TBA111(3), however, being a parallel multimolecular sequence melted at higher temperature at 20µM strand concentration. In the case of this sequence, appreciable hysteresis was also observed between the heating and cooling curves.

We further studied the CD melting of the isoTBA sequences at a higher K^+ ion concentration of 500 mM. The T_m values of all the sequences were found to be higher under these conditions, except in the case of isoTBA111(3), where marginal destabilization was observed (ESI Figure S2-c & S3-c).

Table 3. TBA and isoTBA loop-modified sequences and their CD-T_m at 295 nm at varying salt and strand concentrations

Sequence code	T_m (°C) at 5 µM, in 100 mM KCl		T_m (°C) at 20 µM, in 100 mM KCl		T_m (°C) at 5 µM, in 500 mM KCl	
	Heat	Cool	Heat	Cool	Heat	Cool
TBA232(7)	50	50	48	48	54.2	52.3
TBA111(3)	55[a]	45[a]	59.5[a]	52.5[a]	63.8[a]	53[a]
isoTBA232(7)	37	37	34	34	40	41
isoTBA222(6)	32	32	34.2	32.5	40	39.5
isoTBA131(5)	32	32	33.6	32.2	36.1	37.8
isoTBA111(3)	22	20	20	20	21.5	19.1

[a] indicate the CD-T_m values at 260 nm while all other values are at 295 nm.

UV-Thermal Difference Spectra

To distinguish the difference in the topology of the G-quadruplexes of the synthesized loop-modified sequences, we recorded their UV absorbance at different temperatures to generate the thermal difference spectra (TDS). The TDS is known to provide a fingerprint of G-quadruplex topologies.[22] Thermal difference spectra (TDS) were obtained by subtracting the spectral scan of the sample at temperatures below (i.e., 10 °C) and above (i.e., 90 °C) the melting temperatures (T_ms, Figure 6). The TDS factor is the ratio of ΔA240 nm/ΔA295 nm, where ΔA is the difference, at a given λ, between the absorbance above (i.e., 90 °C) and at a given temperature, T, below the melting temperature (where T = 5 °C, 10 °C,..., etc. upto the melting temperature, T_m).[18]

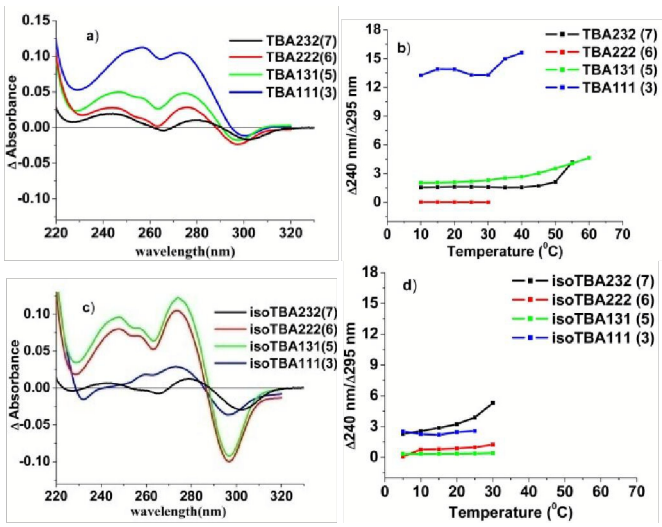


Fig. 6. a) TDS of TBA sequences, b) TDS factor of TBA sequences c) TDS of isoTBA sequences, d) TDS factor of isoTBA sequences

As the CD spectra for TBA232(7) and TBA111(3) showed maxima at 295 nm and 260 nm respectively, indicating two distinctly different G-quadruplex topologies, we first checked the TDS and TDS factor for these two sequences (Figure 6a,b). For the sequence TBA111(3), the values of TDS factor (ΔA240 nm/ΔA295 nm) appeared above 4, indicating a group I parallel quadruplex. The TDS factor of unmodified TBA232(7) was below 2, characteristic of antiparallel quadruplexes (group III). Our observation that the TDS factor for both TBA222(6) and TBA131(5) was also below 2 was in concurrence with literature,[18] where it is reported that the TDS factor cannot differentiate group-II and group-III antiparallel quadruplexes. We conducted similar studies for the isoTBA sequences. The TDS factor of all the isoTBA sequences of the study were characteristic of antiparallel G-quadruplex topology (Figure 6c,d). These results confirm the correct assignment of parallel/antiparallel quadruplex structures in concurrence with the CD signatures.

Chaperone Effect of Thrombin

It has been earlier demonstrated that thrombin can act as a molecular chaperone for the folding of TBA232(7).[27] In our earlier studies we observed that the isomeric 2'-5'-linked isoTBA sequence also could be induced to attain a folded G-quadruplex by the chaperone activity of thrombin.[16] We performed CD experiments with all the sequences of the study, in the presence of increasing concentrations of thrombin at low temperature (Figure 7). In the case of native TBA232(7), a CD maximum at 295 nm was observed even in the absence of either thrombin or K^+, revealing the preference of G-quadruplex structure as discussed above. We observed an increase in the CD signal amplitude with a concomitant shift towards 300 nm upon incremental addition of thrombin. All the other 3'-5'-linked TBA sequences with decreasing loop lengths did not show any change in CD pattern upon addition of thrombin indicating inability of thrombin to assist the folding of these sequences into a quadruplex structure (ESI, Figure S6). In the case of all the 2'-5'-linked isoTBA sequences, the CD maximum at 295 nm and thus, the G-quadruplex structure was not evident in the absence of either thrombin or K^+ ions. Upon incremental addition of thrombin, the CD maximum at 295nm emerged for isoTBA232(7), isoTBA222(6) and isoTBA131(5). Isoelliptic points were observed at ~280 nm and ~255 nm in the case of isoTBA232(7) and isoTBA222(6), indicating the two-state nature of the structural transition.[27] However, in the case of isoTBA111(3), thrombin was unable to induce the sequence to adopt the antiparallel G-quadruplex structure. The temperature-dependent stability of these induced G-quadruplex structures was determined by recording the change in the CD amplitude at 295 nm with increasing temperature. The strength of the TBA232(7) G-quadruplex was the highest with T_m = 22 °C, followed by isoTBA232(7) (T_m = 13 °C), isoTBA222(6) and isoTBA131(5) ($T_m \leq 10$ °C for both sequences). The CD signal at 295 nm was restored upon addition of K^+ ions and the quadruplexes were found to be as stable as with K^+ ions alone (ESI, Figure S7, S8)

Table 4. CD data showing the chaperone effect of thrombin on the TBA and isoTBA loop-modified sequences and T_m values of the resulting quadruplexes

Sequence code	+ Thrombin		Binding to Thrombin	T_m at 5 µM with Thrombin
	Maxima (nm)	Minima (nm)		Heat/cool (°C)
TBA232 (7)	298, 248	278-268	Yes	22/22
isoTBA232(7)	290, 245	265	Yes	13/13
isoTBA222(6)	291, 250	265	Yes	<10/<10
isoTBA131(5)	296, 248	272	Yes	<10/<10
isoTBA111(3)	260	295	No	--

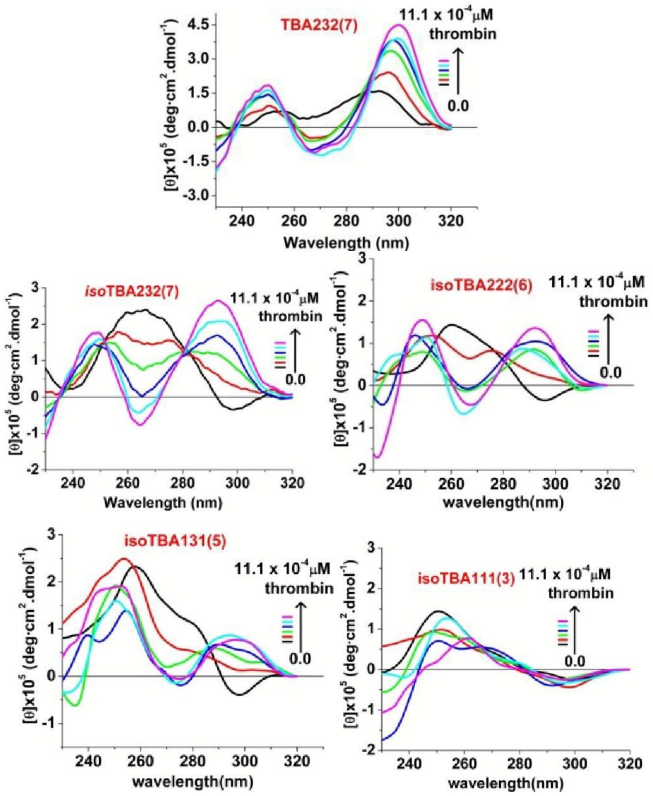


Fig. 7 Changes in CD signal upon addition of thrombin to a) TBA232(7), b) isoTBA232(7), c) isoTBA222(6), d) isoTBA131(5) and e) isoTBA111(3). (**Black**-0.0, **Red**-2.22 X 10^{-4}µM, **Green**-4.44 X 10^{-4}µM, **Blue**-6.66 X 10^{-4}µM, **Cyan**-8.88 X 10^{-4}µM, **Magenta**-11.1 X 10^{-4}µM)

Conclusions

The 2'-5'-linked isoTBA sequences of the present study were found to form antiparallel G-quadruplex structures, independent of loop length in presence of K^+ ions. This result is in contrast to previously reported studies of 3'-5'- phosphodiester-linked G-quadruplexes where reducing the loop length resulted in a change from antiparallel to parallel quadruplex structural topology. The decreasing loop-length in the isoTBA sequences however, adversely affected the thermal stability of these regioisomeric antiparallel quadruplexes and also the thrombin chaperon effect to fold them into stable quadruplex structures was reduced for oligomers with shorter loops.

MATERIAL AND METHODS

Oligonucleotide synthesis and purification

3'-5'- and 2'-5'-linked oligonucleotides were synthesized in-house on a Bioautomation Mermade-4 DNA synthesizer employing β-cyanoethyl phosphoramidite chemistry. The 2'-deoxy-3'-phosphoramidites were obtained from ChemGenes and 3'-deoxy-2'-phosphoramidites from Glen Research. Universal columns procured from Bioautomation were used for 2'-5'-linked oligomer synthesis. Oligonucleotides were cleaved from the solid support by treating with aqueous ammonia at 60° C for 6 h and then concentrated. Oligonucleotides were purified by RP-HPLC on a C18 column using a Waters system (Waters Delta 600e quaternary solvent delivery system, 2998 photodiode array detector and Empower2 chromatography software). An increasing gradient of acetonitrile in 0.1 M triethylammonium acetate (pH 7.0) was used.

MALDI-TOF mass

MALDI-TOF mass spectra were recorded on Voyager-De-STR (Applied Biosystems). The matrix used for analysis was THAP (2, 4, 6-trihydroxyacetophenone) and ammonium citrate as matrix in 2:1 proportion.

CD spectroscopy

CD spectra were recorded on a Jasco J-815 CD spectrometer equipped with a Jasco PTC-424S/15 peltier system. 2mm path-length quartz cuvettes were used for a sample volume of 500µl and strand concentration of 5 µM or 20 µM in potassium phosphate buffer (10 mM, pH 7.2) containing 100 mM or 500 mM KCl. Oligomers in buffer were annealed by heating at 95 °C for 5 min, then slowly cooling to room temperature followed by refrigeration for 5 to 6 h before use. Spectral scans over a range of 320 nm to 200 nm were collected as accumulations of 3 scans at a scanning rate of 100 nm min^{-1}. CD melting was performed by monitoring CD intensity at 295 nm against temperature over a range of 5-90 °C at a heating rate of 2 °C per min.

UV-Thermal Difference Spectra

UV-absorbance scans of the TBA and isoTBA oligomers were recorded using 10mm pathlength quartz cells on a Cary 300 Bio UV-Visible Spectrophotometer (Varian) or an Analytik Jena SPECORD® 200 plus spectrometer equipped with peltier-controlled temperature controller and a scanning speed of 300 nm/min. The TBA oligomers (5 µM strand concentration) were annealed in potassium phosphate buffer (10 mM, pH 7.2), containing 100 mM KCl. The concentration was calculated on the basis of absorbance from molar extinction coefficients of the corresponding nucleobases of DNA. UV spectra over 200 nm-320 nm range were recorded over a temperature range of 5 or 10 to 90 °C. A 3 min equilibration period at each measurement was allowed to ensure homogeneous sample temperature. The autozero function was applied on the corresponding buffer at 10 °C. TDS factors were calculated as the absolute values of $\Delta A240$ nm/$\Delta A295$ nm, where $\Delta A\lambda$ is the difference, at the given wavelength λ, between the absorbance above (90 °C) and at a given temperature T, below the melting temperature (where T = 5 °C, 10 °C,..., etc. upto the melting temperature, T_m).[18]

Gel electrophoresis

The quadruplex-forming ability and the nature of the quadruplexes formed by the synthesized oligomers were assessed by denaturing and non-denaturing polyacrylamide gel electrophoresis using 20% gels. As a reference, the scrambled TBA sequence, viz., 5'-GGTGGTGGTTGTGGT-3' (15-mer) was used. For the non-denaturing gel, TBA and *iso*TBA oligomers (350 µM strand concentration) were annealed in 10 mM potassium phosphate buffer (pH 7.2) containing 500 mM KCl prior to loading on the gel. The gel was run in TBE buffer at a voltage of 150 V at 10 °C till the bromophenol blue dye

had migrated to 3/4th gel length. The bands were visualized by UV-shadowing. For the denaturing gel containing 7 M urea, the TBA and isoTBA oligomers (350 μM strand concentration) were annealed in de-ionized water, and the gel was run in TBE buffer at 25 °C at 150 V.

Acknowledgements

A research grant (GenCODE, BSC0123) from CSIR, New Delhi, is gratefully acknowledged.

Notes and references

1 W. E. Wright, V. M. Tesmer, K. E. Huffiman, S. D. Levene, J. W. Shay, *Genes Dev*., 1997, **11**, 2801-2809; D. Rhodes, Hans J. Lipps Nucleic Acids Res. 2015, doi: 10.1093/nar/gkv862

2 V. S. Chambers, G. Marsico, J. M. Boutell, M. Di Antonio, G. P. Smith, S. Balasubramanian *Nat. Biotech.* 2015, **33**, 877-881.; G. Biffi, D. Tannahill, J. McCafferty, S. Balasubramanian, *Nat. Chem.* 2013, **5**, 182–186; S. Balasubramanian, L. H. Hurley, S. Neidle, *Nat. Rev. Drug Discov*. 2011, **10**, 261–275; D. Sen, W. Gilbert, *Nature*, 1988, **334**, 364–366.

3 A. Rajendran, M. Endo, K. Hidaka, P. L. T. Tran, J. L. Mergny, H. Sugiyama, *Nucleic Acids Res*., 2013, 41, 8738-8747; T. Mashimo, H. Yagi, Y. Sannohe, A. Rajendran, H. Sugiyama . *J. Am. Chem. Soc*., 2010, 132, 14910-18; R. Stefl, T.E. Cheatham III, N. Spacková, E. Fadrná, I. Berger, J. Koca, J. Sponer, *Biophys. J*., 2003, 85, 1787-1804; A. Rajendran, M. Endo, K. Hidaka, H. Sugiyama, *Angew. Chem. Int. Ed*. 2014, 53, 4107-4112; S. Lyonnais, R. J. Gorelick, J. L. Mergny, E. L. Cam, G. Mirambeau, *Nucleic Acids Res*. 2003, 31, 5754-5763; A. Rajendran, M. Endo, K. Hidaka, P. L. T. Tran, J. L. Mergny, R. J. Gorelick, H. Sugiyama *J. Am. Chem. Soc*., 2013, 135, 18575-85; L. Guittat, P. Alberti, F. Rosu, S. V. Miert, E. Thetiot, L. Pieters, V. Gabelica, E. D. Pauw, A. Ottaviani, J. F. Riou, J. L. Mergny, *Biochimie*, 2003, 85, 535-547; A. Rajendran, M. Endo, K. Hidaka, P. L. T. Tran, M. P. T Fichou, J. L. Mergny, H. Sugiyama, *RSC Adv*., 2014, 4, 6346-6355; A. Rajendran, M. Endo, K. Hidaka, M. P. T. Fichou, J. L. Mergny, H. Sugiyama, *Chem. Commun*., 2015, 51, 9181-9184.

4 K. Y. Wang, S. McCurdy, R. G. Shea, S. Swaminathan, P. H. Bolton, *Biochemistry,* 1993, **32**, 1899–1904.

5 H. Y. Qi, C. P. Lin, X. Fu, L. M. Wood, A. A. Liu, Y. C. Tsai, Y. J. Chen, C. M. Barbieri, D. S. Pilch, L. F. Liu, *Cancer Res*., 2006, **66**, 11808–11816.

6 J. R. Wyatt, T. A. Vickers, J. L. Roberson, R. W. Buckheit, Jr., T. Klimkait, E. DeBaets, P. W. Davis, B. Rayner, J. L. Imbach, D. J. Ecker, *Proc. Natl. Acad. Sci. USA*, 1994, **91**, 1356–1360.

7 L. C. Bock, L. C. Griffin, J. A. Latham, E. H. Vermaas, J. J. Toole, *Nature* 1992, **355**, 564-566.; P. Schultze, R. F. Macaya, J. Feigon, *J. Mol. Biol.* 1994, 235, 1532–1547.

8 J. A. Kelly, J. Feigon, T. O. Yeates, *J. Mol. Biol*. 1996, **256**, 417-422.;

9 I. Russo Krauss, A. Merlino, A. Randazzo, E. Novellino, L. Mazzarella, F. Sica, *Nucleic Acids Res.,* 2012, **40**, 8119-8128.

10 A. Avino, C. Fabrega, M. Tintore, R. Eritja *Current Pharmaceutical Design,* 2012, **18**, 2036-2047; B. Saccà, L. Lacroix, J. L. Mergny, *Nucleic Acids Res.,* 2005, **33**, 1182-1192.

11 A. Avino, G. Portella, R. Ferreira, R. Gargallo, S. Mazzini, V. Gabelica, M. Orozco, R. Eritja, *FEBS Journal*, 2014, **281**, 1085–1099; S. Nagatoishi, N. Isono, K. Tsumoto, N. Sugimoto, *Biochimie*, 2011, **93**, 1231-1238.

12 A. Joachimi, A. Benz, J. S. Hartig, *Bioorg. Med. Chem.*, 2009, **19**, 6811-6815; I. Smirnov, R. H. Shafer, *Biochemistry,* 2000**, 39***,* 1462-1468.

13 A. Pasternak, F. J. Hernandez, L. M. Rasmussen, B. Vester, J. Wengel, *Nucleic Acids Res*., 2011, **39,** 1155-1164.

14 E. R. Kandimalla, A. Manning, Q. Zhao, D. R. Shaw, R. A. Byrn, V. Sasisekharan, S. Agrawal, *Nucleic Acids Res.* 1997, **25**, 370-378.

15 T. L. Sheppard, R. C. Breslow, *J. Am. Chem. Soc.* 1996, **40**, 9810-9811; T. P. Prakash, K. -E. Jung, C. Switzer, *Chem. Commun.* 1996, 1793-1795; K. -E. Jung, C. Switzer, *J. Am. Chem. Soc.* 1994, **116**, 6059-6061; P. A. Giannaris, M. J. Damha, *Nucleic Acids Res*. 1993, **21**, 4742-4749; N. Erande, A. D. Gunjal, M. Fernandes, V. A. Kumar, *Chem. Commun*. 2011, **47**, 4007-4009; N. Erande, A. D. Gunjal, M. Fernandes, R. Gonnade, V. A. Kumar, *Org. Biomol. Chem.* 2013, **11**, 746-757; H. Sawai, J. Seki, H. Ozaki, *J. Biomol. Struct. Dyn.* 1996, **13**, 1043-1051; M. J. Damha, A. Noronha, *Nucleic Acids Res.* 1998, **26**, 5152-5156.

16 A. D. Gunjal, M. Fernandes, N. Erande, P. R. Rahamohanan, V. A. Kumar, *Chem. Comm*. 2014, **50**, 605-607.

17 M. Polak, M. Manoharan, G. B. Inamati, J. Plavec, *Nucleic Acids Res.* 2003, **31**, 2066-2076.; B. J. Premraj, S. Raja, N. S. Bhavesh, K. Shi, R. V. Hosur, M. Sundaralingam, N. Yathindra, *Eur. J. Biochem.,* 2004, **271**, 2956-2966.

18 A. I. Karsisiotis, N. M. Hessari, E. Novellino, G. P. Spada, A. Randazzo, M. Webba da Silva, *Angew. Chem. Int. Ed. Engl.* 2011, **50**, 10645-10648.

19 P. Hazel, J. Huppert, S. Balasubramanian, S. Neidle, *J. Am. Chem. Soc*. 2004, **126**, 16405–16415.

20 P. K. Dominick, M. B. Jarstfer, *J. Am. Chem. Soc*. 2004, **126**, 5050–5051.

21 N. Berova, K. Nakanishi (Eds.: N. Berova, K. Nakanishi, R.W. Woody), Wiley-VCH, New York, 2nd edn, 2000, p. 337; N. Berova, L. Di Bari, G. Pescitelli, *Chem. Soc. Rev*. 2007, **36,** 914-931; D. M. Gray, J. -D. Wen, C. W. Gray, R. Repges, G. Raabe, J. Fleischhauer, Chirality 2008, **20**, 431–440; M. Vorlickova, I. Kejnovska, J. Sagi, D. Renciuk, K. Bednarova, J. Motlova, J. Kypr, *Methods* 2012, **57**, 64-75; A. Randazzo, G. P. Spada, M. W. da Silva, *Top. Curr. Chem.* 2013, **330**, 67-86.

22 J. L. Mergny, J. Li, L. Lacroix, S. Amrane, J. B. Chaires, *Nucleic Acids Res.* 2005, **33**, e138.

23 R. Krishan, T. P. Seshadri, *Biopolymers*, 1994, **34**, 1637-1646.

24 B. J. Premraj, N. Yathindra, *J. Biomol. Struct. Dyn*. 1998, **16**, 313–328.

25 J. L. Mergny, L. Lacroix, *Oligonucleotides*, 2003, **13**, 515–537.

26 H. Emmanuel, O. Keika, Y. Danzhou, *Biochemistry* 2010, **49**,9152–9160; P. A. Rachwal, K. R. Fox, *Methods*, 2007, **43**, 291–301.

27 E. Baldrich, C. K. O'Sullivan, *Anal Biochem.* 2005*,* **341,** 194–197; S. Nagatoishi, Y. Tanaka, K. Tsumoto, *Biochem. Biophys. Res. Commun*. 2007**, 352,** 812-817.