Electronic Supplementary Material (ESI) for Organic & Biomolecular Chemistry

ELECTRONIC SUPPLEMENTARY INFORMATION

More than one non-canonical phosphodiester bond into the G-tract: formation of unusual parallel G-quadruplex structures

Antonella Virgilio,‡ Veronica Esposito,‡ Luciano Mayol and Aldo Galeone*

Department of Pharmacy, University of Naples "Federico II", Via D. Montesano, 49, 80131 Naples (Italy)

‡These authors contributed equally.

*Corresponding Author: galeone@unina.it

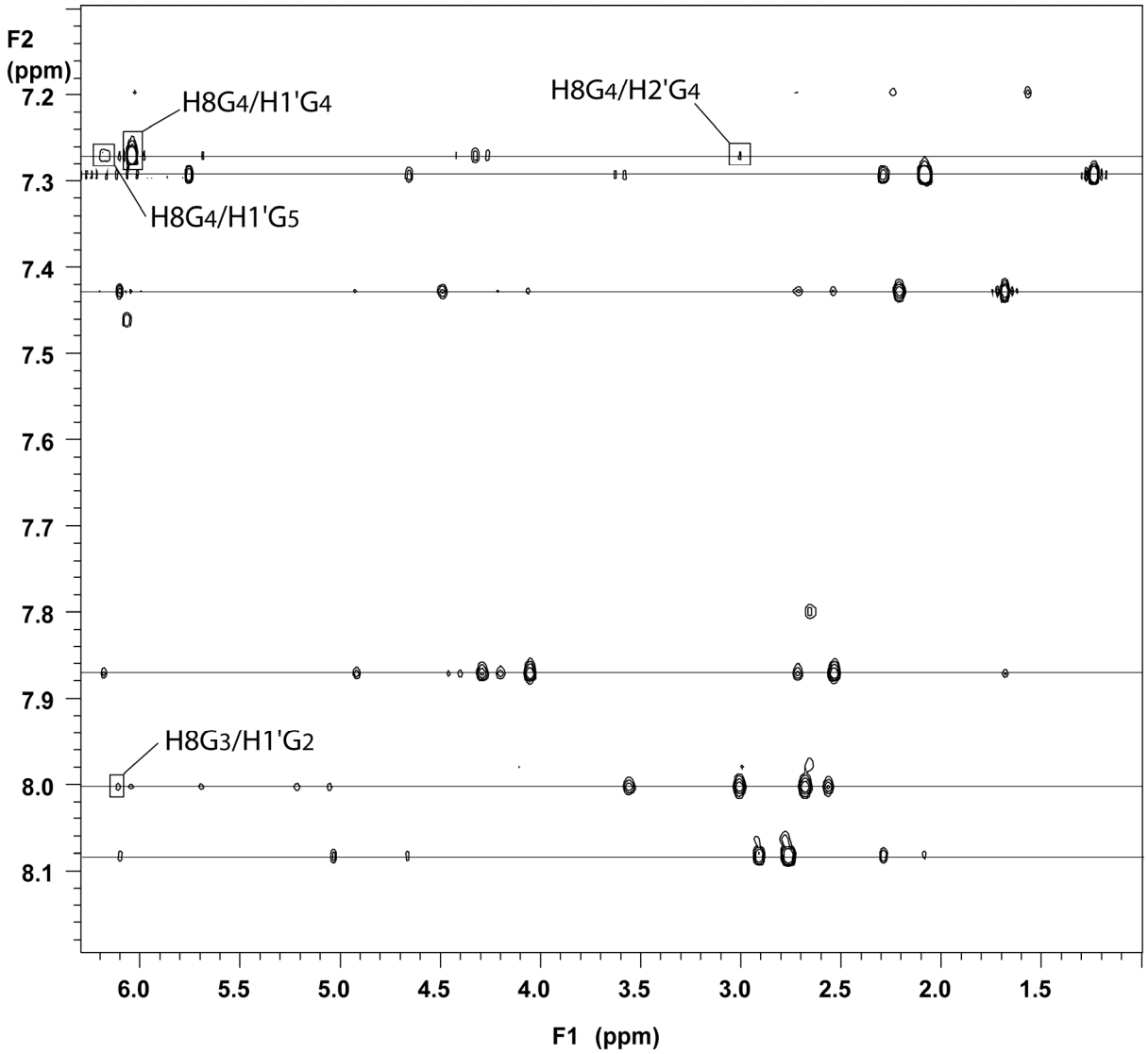

Figure S1: Expanded 2D NOESY spectrum of **Q305** (5'$T_1G_2$3'-3'$G_3G_4$5'-5'$G_5T_6$3') at 700 MHz, 25°C, 3 mM ODN, 0.6 mL (H_2O/D_2O 9:1, v/v), buffer solution: 10 mM KH_2PO_4/K_2HPO_4 (pH 7.0), 70 mM KCl, and 0.2 mM EDTA. The aromatic/sugar proton NOE connectivities are indicated by continuous black lines. The NOE cross-peaks indicating the *syn* dG and the NOE contacts between the dG nucleotides flanking the 5'-5' and the 3'-3' inversion of polarity sites are highlighted by black boxes.

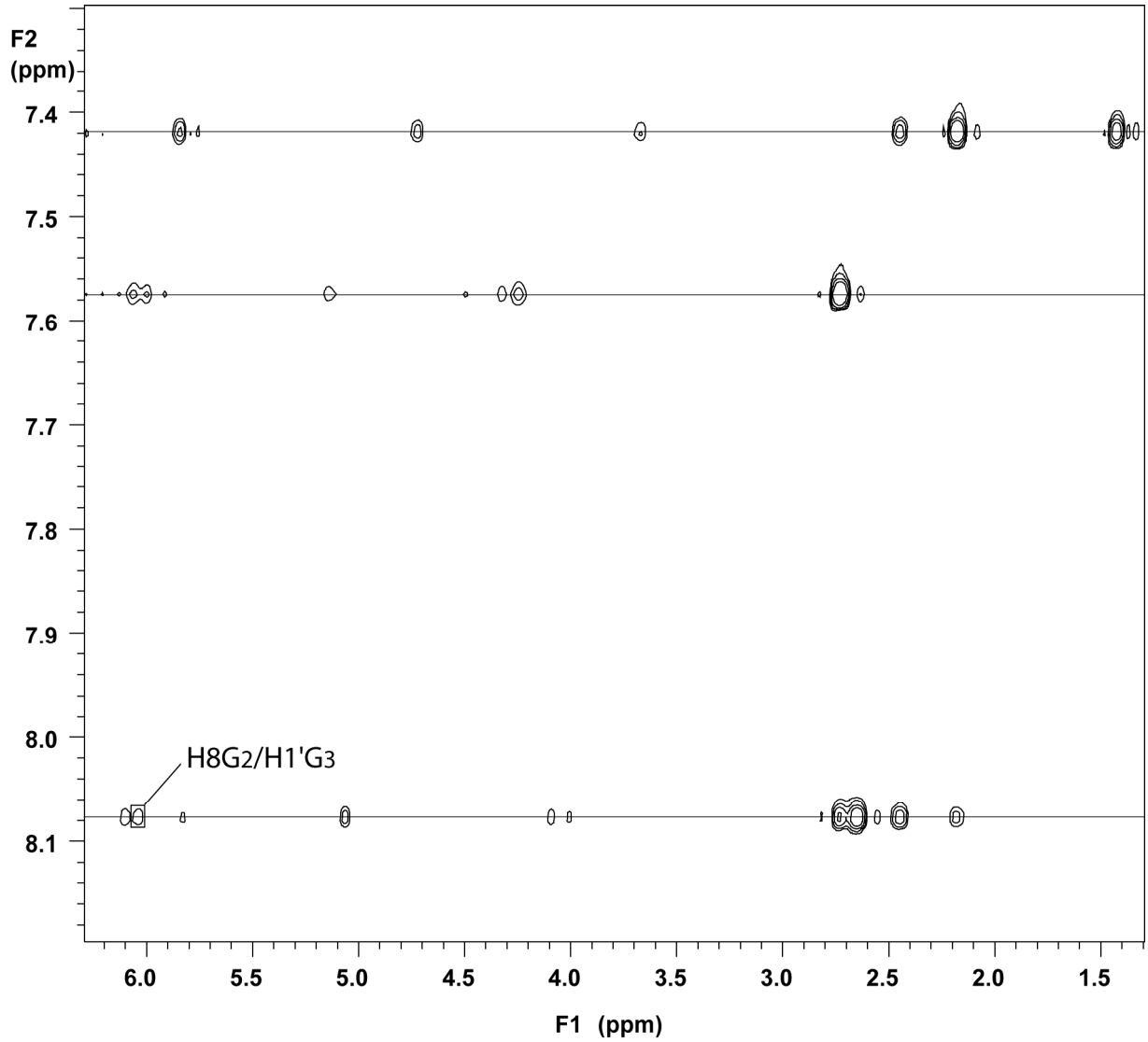

Figure S2: Expanded 2D NOESY spectrum of **Q353** (5'$T_1G_2$3'-3'G_3 5'-5'G_3 3'-3'$G_2T_1$5') at 700 MHz, 25°C, 3 mM ODN, 0.6 mL (H_2O/D_2O 9:1, v/v), buffer solution: 10 mM KH_2PO_4/K_2HPO_4 (pH 7.0), 70 mM KCl, and 0.2 mM EDTA. The aromatic/sugar proton NOE connectivities are indicated by continuous black lines. The NOE contact between the dG nucleotides flanking the 3'-3' inversion of polarity sites is highlighted by a black boxes.

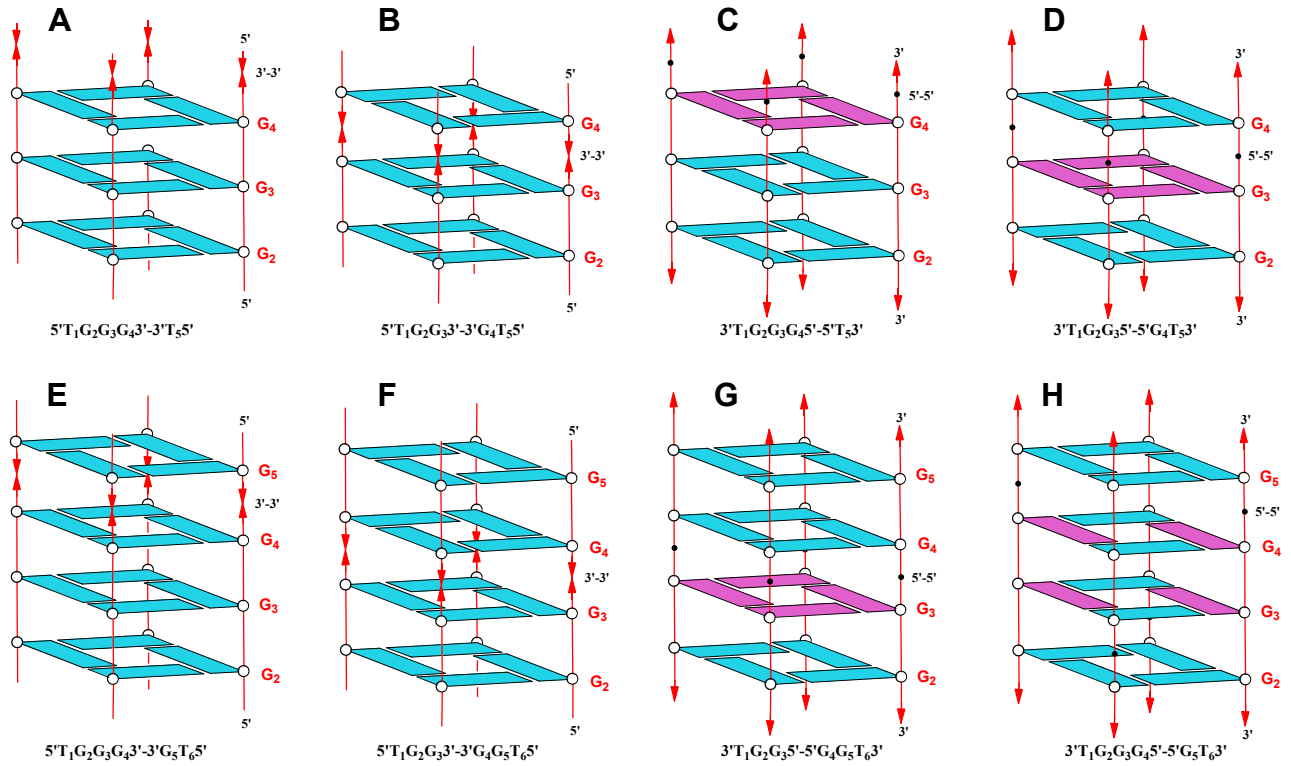

Figure S3: Schematic representation of the quadruplexes formed by previously investigated ODNs containing 5'-5' and 3'-3' inversion of polarity sites: **A-D** (see ref. 7), **E** and **H** (see ref. 10), **F** and **G** (see ref. 8). *Anti* and *syn* guanosines are in light blue and purple, respectively. The inversion of polarity sites 5'-5' and 3'-3' are indicated by black dots and red double-arrows, respectively. T residues have been omitted for clarity.

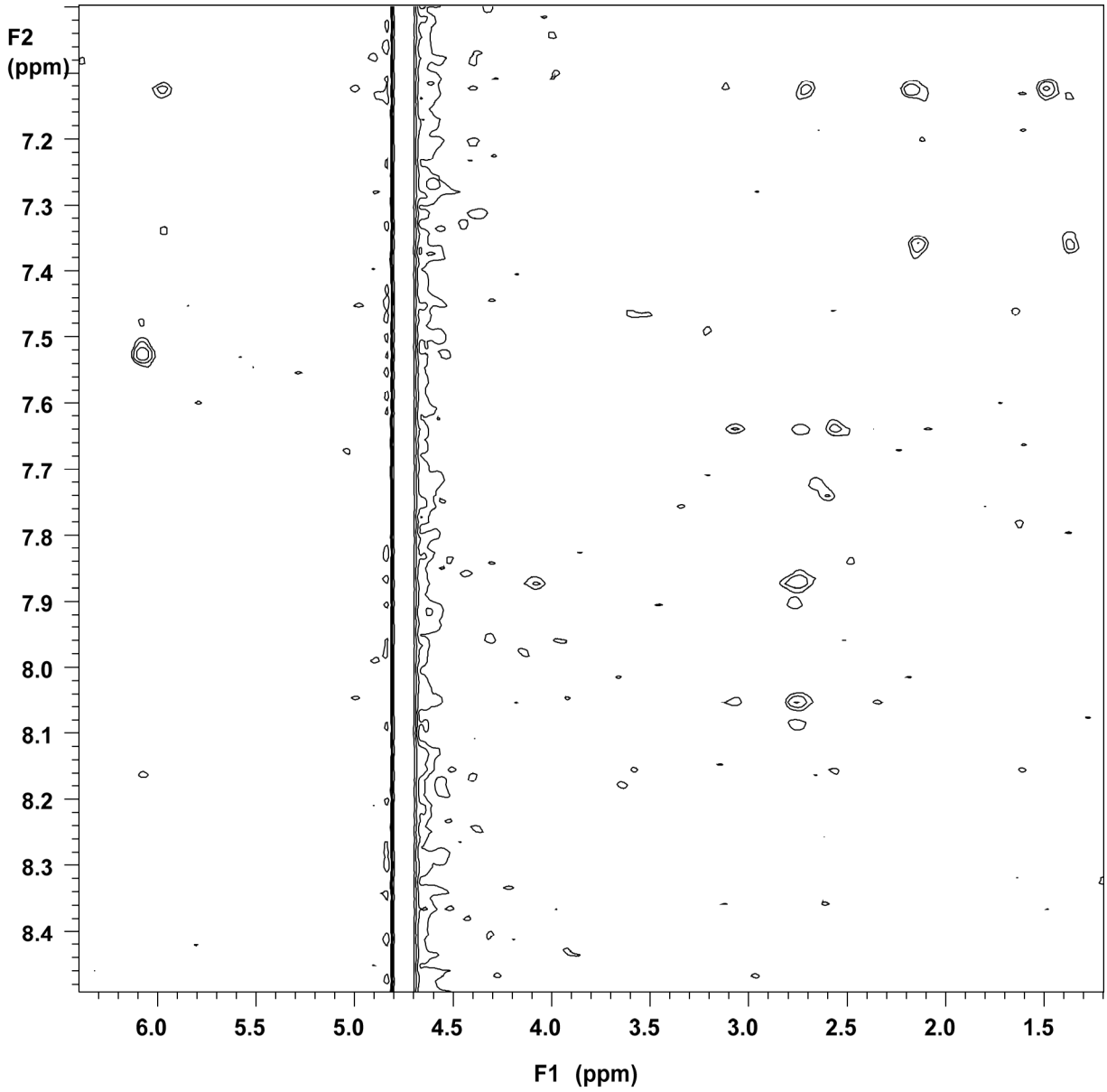

Figure S4: Expanded 2D NOESY spectrum of **Q035** (5'$T_1G_2G_3$3'-3'$G_4$5'-5'$G_5T_6$5') at 500 MHz, 25°C, 3 mM ODN, 0.6 mL (H_2O/D_2O 9:1, v/v), buffer solution: 10 mM KH_2PO_4/K_2HPO_4 (pH 7.0), 70 mM KCl, and 0.2 mM EDTA.

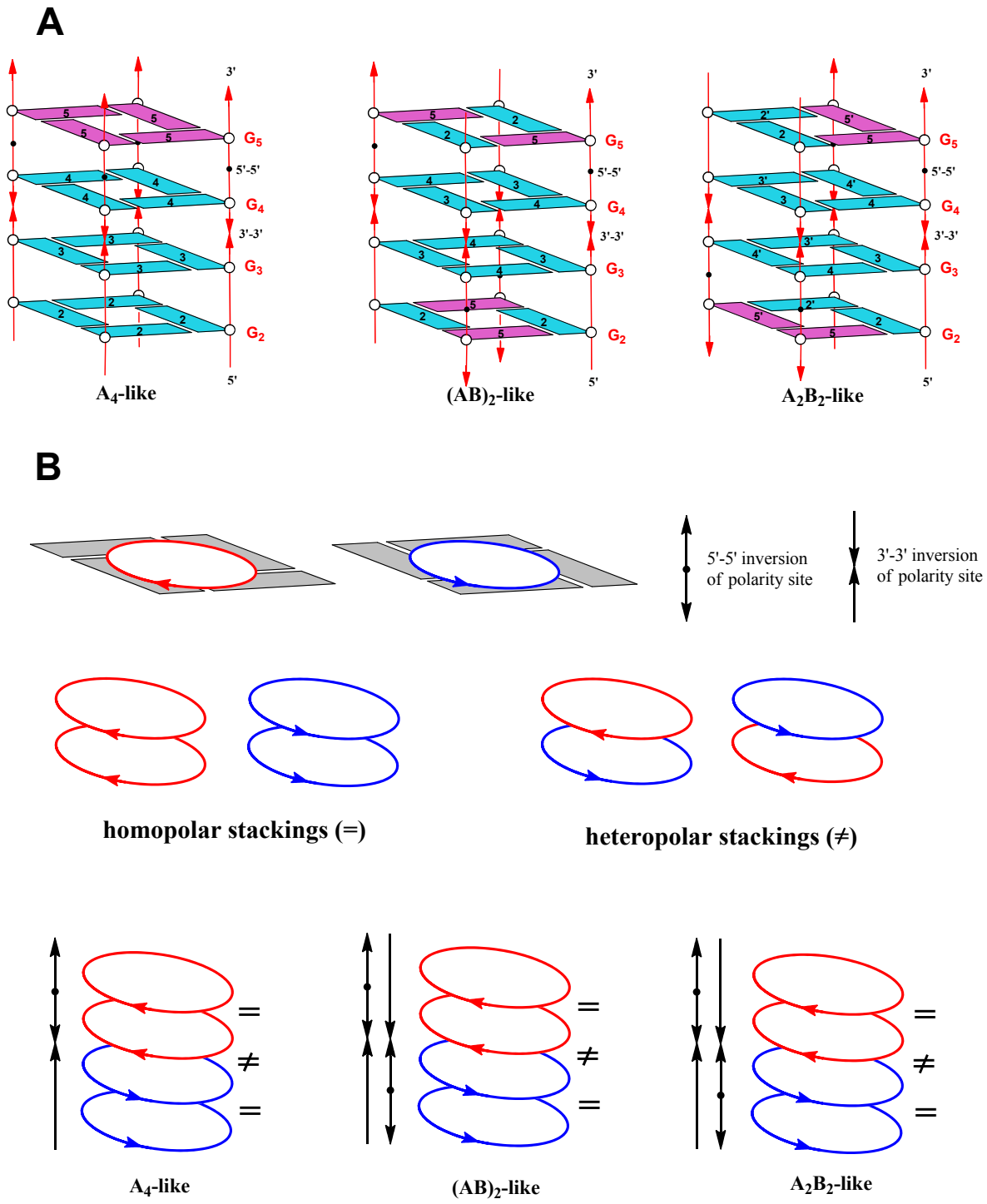

Figure S5: (A) Schematic representation of the quadruplexes hypothesized for **Q035**. *Anti* and *syn* guanosines are in light blue and purple, respectively. The inversion of polarity sites 5'-5' and 3'-3' are indicated by black dots and red double-arrows, respectively. T residues have been omitted for clarity. (B) Sketches of the stacking arrangements for the quadruplex structures in B.

Comparison of CD profiles of Q350, Q305, Q035, Q535 and Q353 with other quadruplex structures

For convenience, we will discuss separately CD data concerning ODNs **Q350** and **Q305** containing two inversion of polarity sites, from those containing three inversion of polarity sites, namely **Q535** and **Q353**. Since, according to the NMR data, **Q035** seems to form a complex mixture of G-quadruplex structures, its CD profile will be discussed apart.

ODNs **Q350** and **Q305** show very similar CD profiles (Fig. S5A), both of them being characterized by a positive band at 246 and a negative one at 265 nm. In either case, a further positive band is present in their CD spectra (around 291 and 298 nm for **Q350** and **Q305**, respectively). Concerning **Q350**, NMR data (see the main article) have shown the coexistence in solution of a major species (about 70%, containing an all-*syn* G-tetrad) and a minor species (containing only *anti* guanines) (Fig. 4 in the main article). However it should be noted that both species are characterized by a G-tetrads stacking arrangement similar to that of the G-quadruplex structure formed by **Q305** (Fig. S5A), namely two hetero- and one homopolar stacks, thus accounting for the similarity of their CD profiles. At the best of our knowledge, only a G-quadruplex structure is known to adopt a G-tetrad arrangement involving two hetero- and one homopolar stacking, namely the structure formed by ODN TMGGMT (**F14**) in which M = 8-methyl-2'-deoxyguanosine (ref. 14). In the region 220-280 nm, its CD spectrum (Fig. S6A) is quite similar to those of **Q350** and **Q305**, apart from a slight shift towards higher wavelengths, while the positive band around 295 nm is more intense. Since **F14** contains only canonical 3'-5' phosphodiester bonds, the observed differences could be ascribed to the influence of the sites of inversion on the twist angles between two G-tetrads stacked through a non-canonical phosphodiester bond.

The G-quadruplex structure formed by ODN **Q535** shows a CD profile characterized by a negative (266 nm) and two positive bands (245 and 294 nm) (Fig. S6B). According to the NMR data (see the main article), it adopts a G-quadruplex structure in which two homo- and one heteropolar stacks occur (Fig. S6B). An analogous G-tetrad arrangement has been observed also in **Q33** (ref. 8) whose CD profile, although showing positive and negative bands roughly in the same spectral regions, differs from that of **Q535**, due to a shift towards higher wavelengths and a stronger band around 265 nm.

The CD spectrum of ODN **Q353** shows a positive band at 245 nm and a negative one at 263 nm (Fig. S6B). A further positive, very weak band at 286 nm is also detectable, contrary to the CD profiles of the other modified ODNs investigated in which a marked positive band is present in this spectral region. In the case of ODN **Q353**, NMR data point to a G-quadruplex structure characterized by three heteropolar stacks. A similar strand arrangement could be observed in ODN TGMGMT (M = 8-methyl-2'-deoxyguanosine) (ref. 14) containing only canonical 3'-5' phosphodiester linkages, whose CD spectrum partially resembles that of **Q353**, except for two features already observed in the previous comparison between CD profiles of **Q305** and **F14**, namely, a slight shift towards higher wavelengths and a higher band at about 290 nm (Fig. S6C). As far as the CD spectrum of ODN **Q035** is concerned, it displays two positive bands at 246 and 294 nm and a negative band at 266 nm. It should be noted that both the quadruplex structures previously hypothesized for the main species in solution (Fig. S5A) show both an identical tetrad arrangement, characterized by two homopolar- and one heteropolar stacking (Fig. S6D). Notably, the same tetrad arrangement has been observed also for **Q535**, whose normalized CD spectrum is almost super-imposable to that of **Q035** (Fig. S6D). Then, although the CD profile of **Q035** has not helped us to discriminate between a A_4-like (parallel-like) and a $(AB)_2$-like quadruplex structure (Fig. S5A), it corroborates our initial supposition based both on the molecular symmetry inferred by NMR data and the generalizations deduced by the structural preferences of other ODNs containing inversion of polarity sites.

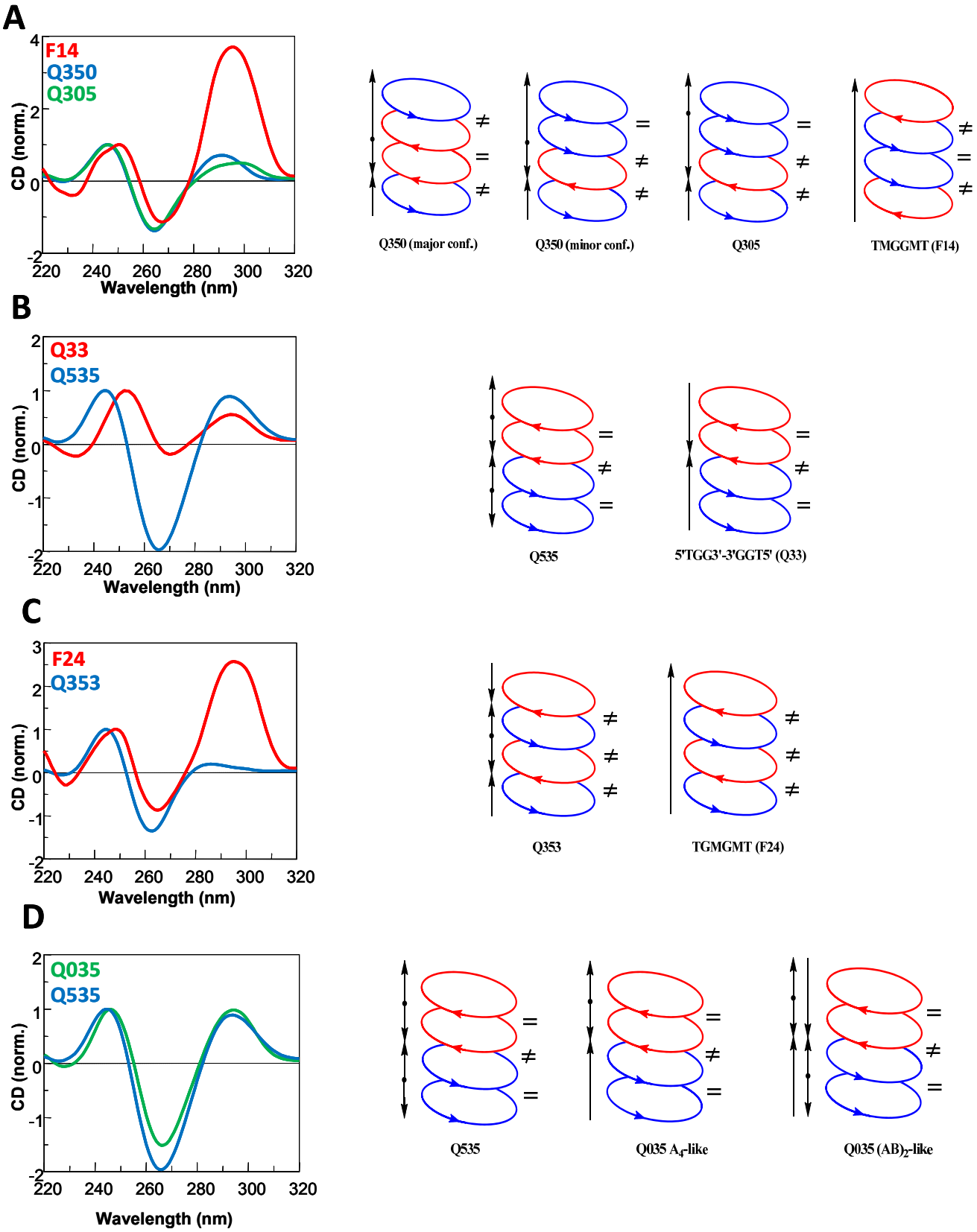

Figure S6: Comparisons between CD profiles (left) of quadruplex structures sharing the same stacking arrangement (sketches on the right; see text and Fig. S5 and for details). **A**: **F14**, **Q350** and **Q305**; **B**: **Q33** and **Q535**; **C**: **F24** and **Q353**; **D**: **Q035** and **Q535**. CD spectra were normalized at +1 for the band occurring in the 245-255 nm range.

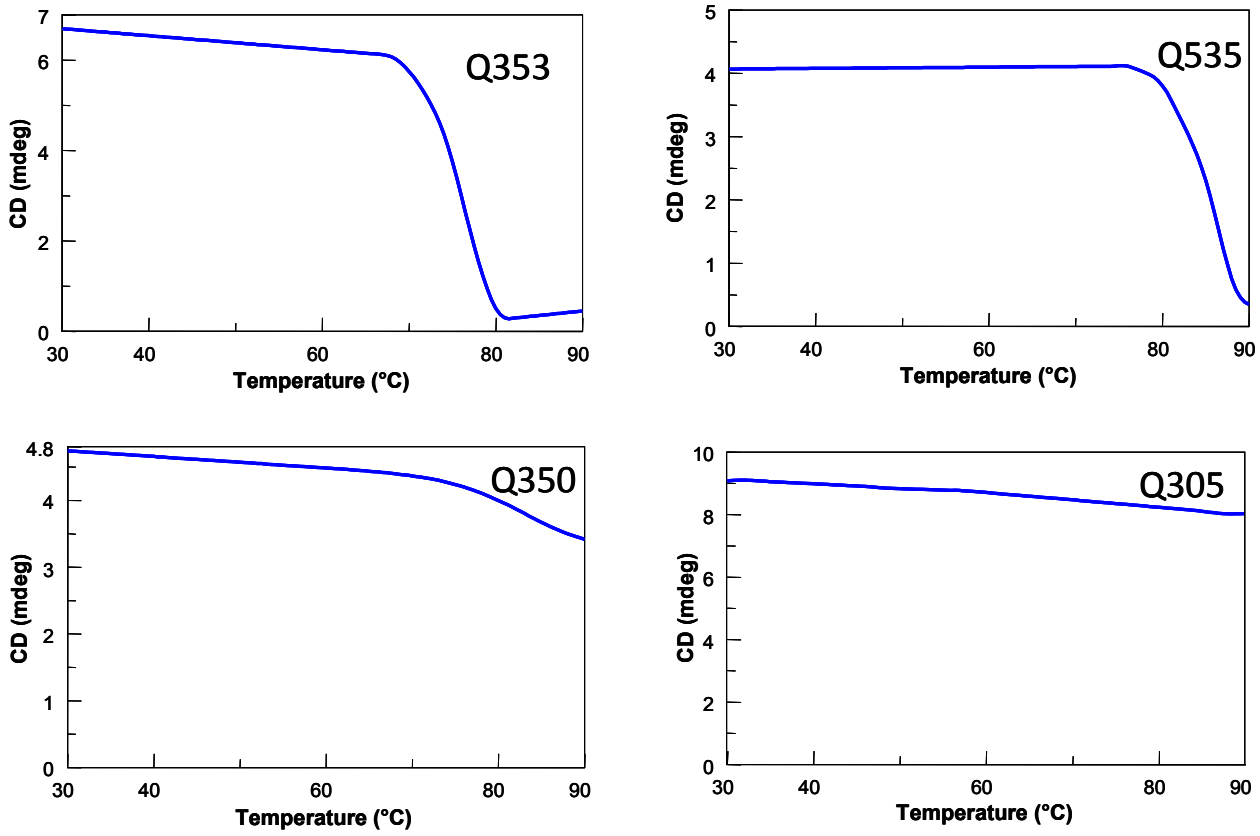

Figure S7: CD melting profiles of quadruplex structures formed by ODNs **Q353**, **Q535**, **Q350** and **Q305**. See experimental section for details.

Discussion on the gel electrophoresis (PAGE) results

It is well known that the electrophoretic migration for quadruplex structures depends on the interplay between several factors including molecular shape, size and charge. As for ODNs with three inversions (**Q353** and **Q535**), their electrophoretic profile shows a behavior rather different with the band of **Q353** which migrates faster. The different behaviors of **Q353** and **Q535** could be tentatively explained mainly by two features: 1) the major compactness of **Q353** compared to **Q535**, due to the presence of two 3'-3' inversion of polarity sites shorter than the 5'-5' ones; this consideration can be invoked to account for the similar faster migration of denatured **Q353** than denatured **Q535**, as well; 2) the minor negative charge of **Q535** compared to **Q353**, which may be ascribed to the better ability of the 5'-5' inversion of polarity sites than the 3'-3' ones to function as cation occupancy sites (ref. 11) These observations are also in accordance with the behaviors of **Q350** and **Q305**, possessing the same number and type of sites and showing quite similar electrophoretic motilities. A further interesting datum comes out from the electrophoresis experiment, concerning the presence of several minor bands for many of the modified quadruplexes. As far **Q350** is concerned, NMR measurement pointed out to the presence in solution of two different quadruplex structures that is confirmed by the electrophoretic profile, as well. The number and the intensity of the bands in the other cases could be explained through the hierarchic rules and generalizations described in the main article. According to these, in a parallel-like quadruplex, a G3'-3'G step only can assume the form G_{anti}3'-3'G_{anti}, thus accounting for presence of a unique band in the electrophoretic profile of **Q353**. As for **Q535**, taking into account that a G5'-5'G step can assume both the form G_{anti}5'-5'G_{anti} and G_{anti}5'-5'G_{syn} (more probable), three parallel-like quadruplex structures could form adopting three strands types, namely 3'TG_{syn}5'-5'G_{anti}3'-3'G_{anti}5'-5'G_{syn}T3' (the major conformation according to the NMR data), 3'TG_{anti}5'-5'G_{anti}3'-3'G_{anti}5'-5'G_{syn}T3' and 3'TG_{anti}5'-5'G_{anti}3'-3'G_{anti}5'-5'G_{anti}T3', thus explaining the presence of a major and two minor bands in the gel. Similar considerations could account for the presence of a further, slower migrating weak band in the case of **Q305** that could be due to a minor parallel-like quadruplex structure adopting the strand type 5'TG_{anti}3'-3'$G_{anti}G_{anti}$5'-5'G_{anti}T3', apart from the main parallel-like quadruplex structure that has proven to adopt a strand type 5'TG_{anti}3'-3'$G_{anti}G_{syn}$5'-5'G_{anti}T3'. These considerations would suggest that the glycosidic conformation adopted by the residues can affect the electrophoretic motility of a quadruplex structure, thus confirming similar results obtained in investigation concerning ODNs containing 8-methyl-2'-deoxyguanosines (ref. 14 and 15).
Concerning **Q035**, it appears as a single large band, apparently in contrast with the NMR data that suggest more than one conformation in solution. These data could be explained by the comparable electrophoretic motilities of the different conformations adopted by **Q035** (thus originating a large band), particularly taking into consideration that all the three conformations hypothesized for this ODN (Fig. S5A) are characterized by the same number of *syn* (one residue per strand) and *anti* guanosines (three residues per strand) .

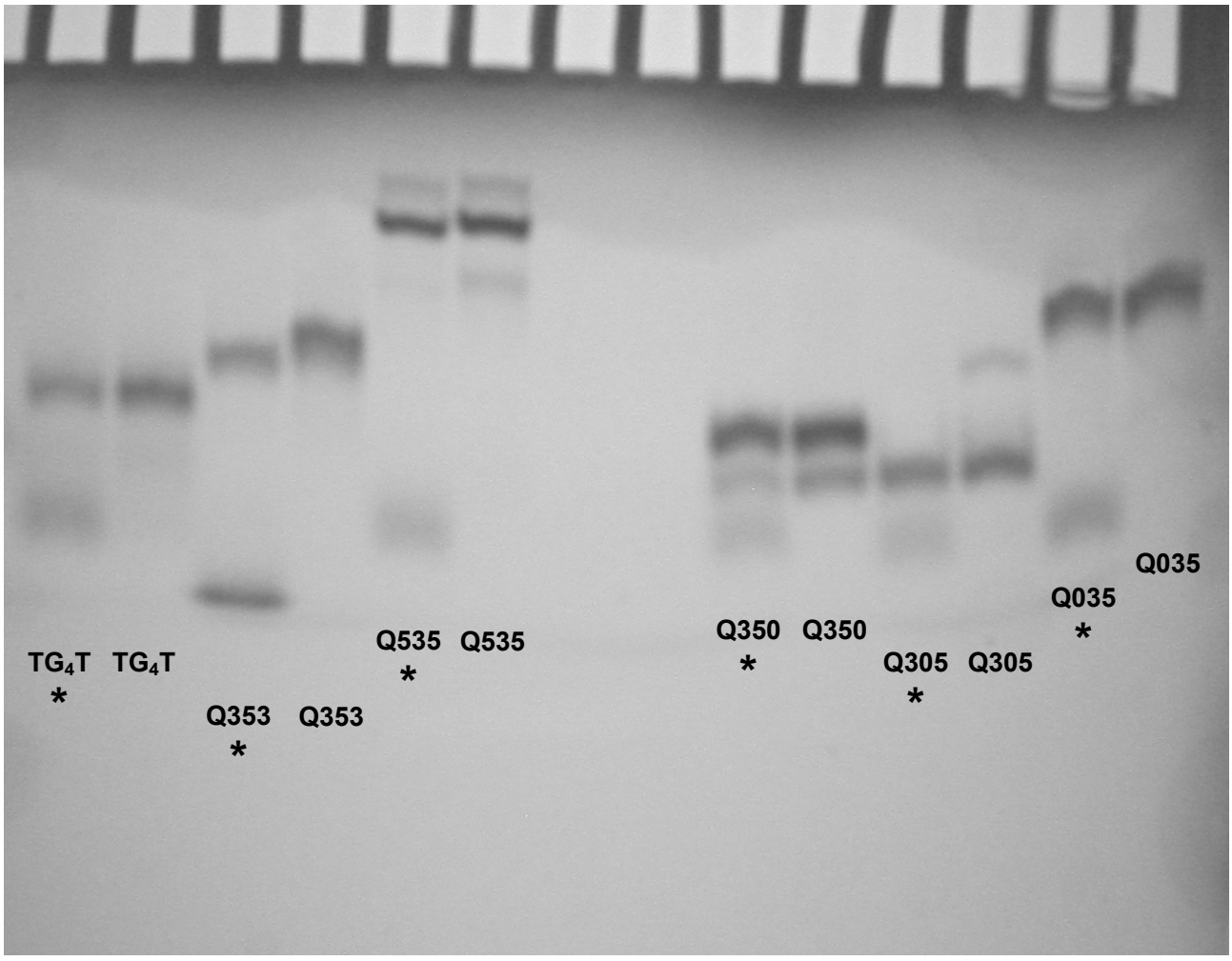

Figure S8: Non-denaturing PAGE. Samples were loaded on a non-denaturing gel, run at 5°C. Oligonucleotides were revealed by UV shadowing. Lanes marked by asterisks correspond to the denatured sequences pre-treated with LiOH (0.1 M, 5' at 80°C) followed by neutralization by HCl 0.1 M and immediate loading on the gel. The native samples have been prepared in the NMR buffer (10 mM KH_2PO_4, 70 mM KCl and 0.2 mM EDTA, pH=7).

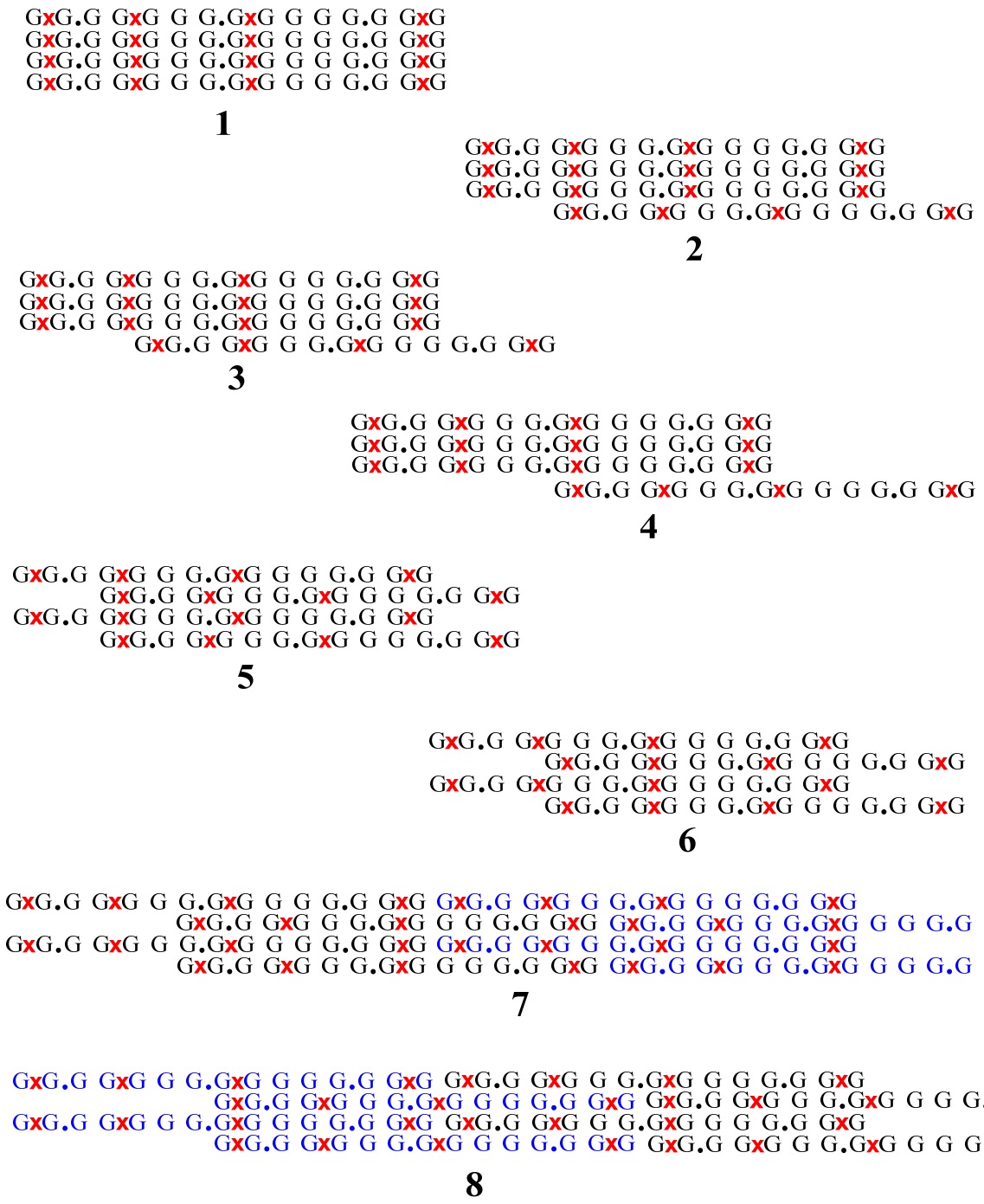

Figure S9: Possible strand arrangements of ODN 5'G3'-3'G5'-5'GG3'-3'GGG5'-5'G3'-3'GGGG5'-5'-GG3'-3'G5'. 3'-3' and 5'-5' inversion of polarity sites are indicated by red x and black dots, respectively. The sequence has been suitably designed in such a way that, in arrangement **1,** each type of inversion of polarity site (3'-3' or 5'-5') occurs next to sites of the same type along the G-quadruplex complex. This is not the case for other alternative arrangements (for example **2**-**8**).

Q350 5'$T_1G_2$3'-3'$G_3$5'-5'$\underline{G}_4G_5T_6$3' (major conf.)								
	H8/H6	H1'	H2'/H2''	H3'	H4'	H5'/H5''	CH_3	NH
T_1	7.37	5.83	2.12/2.40	4.70	4.48	N.A.	1.37	-
G_2	8.07	6.12	2.70/2.84	5.10	4.37	N.A.	-	11.53
G_3	7.96	6.07	2.77/2.84	5.18	4.52	4.21/4.15	-	N.A.
$\underline{G}_4$	7.26	5.95	2.89/3.22	5.10	4.42	4.37	-	N.A.
G_5	7.89	5.97	2.47/2.59	4.98	4.39	4.23/4.07	-	11.40
T_6	7.24	6.00	2.11	4.44	3.97	N.A.	1.69	-

Q350 5'$T_1G_2$3'-3'$G_3$5'-5'$G_4G_5T_6$3' (minor conf.)								
	H8/H6	H1'	H2'/H2''	H3'	H4'	H5'/H5''	CH_3	NH
T_1	7.25	5.82	1.88/2.26	N.A.	N.A.	N.A.	1.43	-
G_2	8.08	5.91	2.86	N.A.	N.A.	N.A.	-	N.A.
G_3	7.70	5.84	2.71/2.86	5.14	N.A.	N.A.	-	N.A.
G_4	7.69	N.A.	2.53/2.86	N.A.	N.A.	N.A.	-	N.A.
G_5	7.68	6.15	2.45/2.63	4.31	N.A.	N.A.	-	N.A.
T_6	7.32	6.03	2.14	4.92	4.43	4.21	1.65	-

Q305 5'$T_1G_2$3'-3'$G_3\underline{G}_4$5'-5'$G_5T_6$3'								
	H8/H6	H1'	H2'/H2''	H3'	H4'	H5'/H5''	CH_3	NH
T_1	7.30	5.75	2.09/2.29	4.66	4.03	3.63/3.58	1.24	-
G_2	8.09	6.10	2.76/2.91	5.03	4.05	N.A.	-	11.44
G_3	8.01	5.69	2.56/2.68	5.22	4.43	N.A.	-	11.24
$\underline{G}_4$	7.27	6.04	3.01/3.56	5.05	4.33	4.26	-	11.39
G_5	7.87	6.18	2.53/2.72	4.92	4.20	N.A.	-	11.29
T_6	7.43	6.10	2.21	4.49	4.05	4.21/4.06	1.68	-

Table S1: Proton chemical shifts for quadruplexes formed by ODNs **Q350** and **Q305** (700 MHz, T = 25°C) in 10 mM KH_2PO_4/K_2HPO_4, 70 mM KCl and 0.2 mM EDTA (pH 7.0). N.A. = not assigned. The underlined residues adopt a *syn* glycosidic conformation.

Q535 3'T_1<u>G_2</u>5'-5'G_3 3'-3'G_3 5'-5'<u>G_2</u>$T_1$3'								
	H8/H6	H1'	H2'/H2''	H3'	H4'	H5'/H5''	CH_3	NH
T_1	7.12	5.97	2.10/2.18	4.53	3.97	N.A.	1.50	-
<u>G_2</u>	7.56	6.08	2.73/3.15	4.98	4.43	4.26/4.08	-	11.24
G_3	7.86	6.22	2.66/2.80	5.16	4.48	4.22/4.09	-	11.20

Q353 5'$T_1G_2$3'-3'G_3 5'-5'G_3 3'-3'$G_2T_1$5'								
	H8/H6	H1'	H2'/H2''	H3'	H4'	H5'/H5''	CH_3	NH
T_1	7.42	5.84	2.18/2.45	4.73	3.68	N.A.	1.42	-
G_2	8.08	6.09	2.65/2.73	5.07	4.39	4.09/4.00	-	11.46
G_3	7.58	6.02	2.73	5.13	4.52	4.32/4.25	-	11.32

Table S2: Proton chemical shifts for quadruplexes formed by ODNs **Q535** and **Q353** (700 MHz, T = 25°C) in 10 mM KH_2PO_4/K_2HPO_4, 70 mM KCl and 0.2 mM EDTA (pH 7.0). N.A. = not assigned. The underlined residues adopt a *syn* glycosidic conformation.