SUPPLEMENTARY DATA

Self-association of short DNA loops through Minor Groove C:G:G:C Tetrads.

Júlia Viladoms, Núria Escaja, Miriam Frieden, Irene Gómez-Pinto, Enrique Pedroso & Carlos González

Table S1- ^{1}H assignments for the dimeric form of d<pCGCTCCGT> (100 mM NaCl, phosphate buffer pH7, T=35ºC).

	NH	**NH_2 (2)**	**NH_2 (1)**	**H6/H8**	**H2/H5/Me**	**H1'**	**H2'**	**H2"**	**H3'**	**H4'**	**H5'H5"**	
C1	--	n.o.	n.o.	8.02	6.11	6.48	2.26	2.59	4.57	4.18	n.a.	n.a.
G2	13.45	9.18	5.91	8.14	--	6.24	2.77	2.97	5.02	4.58	4.03	4.23
C3	--	8.79	6.53	7.43	5.38	6.28	2.33	2.44	4.98	4.31	n.a.	n.a.
T4	10.5	--	--	7.53	1.66	5.98	1.88	2.39	4.71	3.78	3.78	3.91
C5	--	n.o.	n.o.	8.02	6.11	6.44	2.26	2.60	4.63	n.a.	n.a.	n.a.
C6	--	9.14	6.99	7.74	5.98	6.31	2.22	2.73	4.93	4.37	4.05	4.18
G7	13.59	9.12	7.02	7.92	--	6.29	2.80	2.53	5.13	4.80	4.22	4.33
T8	10.80	--	--	7.42	1.86	5.93	2.40	2.40	4.64	3.46	3.23	3.63

n.o – not observed. n.a – not assigned.

Table S2- ^{1}H assignments for the dimeric form of d(TGCTTCGT) (100 mM NaCl, 10 mM $MgCl_2$, phosphate buffer pH 7, T=5ºC).

	NH	**NH_2 (2)**	**NH_2 (1)**	**H6/H8**	**H2/H5/Me**	**H1'**	**H2'**	**H2"**	**H3'**	**H4'**	**H5'H5"**	
T1	n.o.	-	-	7.72	1.93	6.41	2.48	2.62	4.90	4.30	3.89	n.a.
G2	13.52	9.11	6.04	8.11	-	6.25	2.77	2.99	5.04	4.62	4.06	4.19
C3	-	8.77	6.55	7.45	5.38	6.31	2.35	2.44	5.00	4.34	4.08	n.a.
T4	**	-	-	7.52	1.66	5.99	1.90	2.41	4.71	3.76	n.a.	n.a.
T5	n.o.	-	-	7.86	1.94	6.48	2.32	2.54	4.64	n.a.	3.91	4.21
C6	-	9.13	7.12	7.78	6.03	6.32	2.26	2.72	4.96	4.36	n.a.	n.a.
G7	13.63	9.16	7.08	8.00	-	6.22	2.83	2.54	5.16	4.81	4.23	4.35
T8	**	-	-	7.43	1.86	5.83	1.92	2.27	4.33	3.61	3.41	3.75

n.o – not observed. n.a – not assigned. ** Imino protons of T4 and T8 resonate at 10.47 and 10.83 ppm.

Table S3- ^{1}H assignments for the dimeric form of d(TCGTTGCT) (100 mM NaCl, 10 mM $MgCl_2$, phosphate buffer pH 7, T=5ºC).

	NH	**NH_2 (2)**	**NH_2 (1)**	**H6/H8**	**H2/H5/Me**	**H1'**	**H2'**	**H2"**	**H3'**	**H4'**	**H5'H5"**	
T1	n.o.	-	-	7.71	1.91	6.38	2.47	2.58	4.87	4.27	3.88	n.a.
C2	-	9.01	6.88	7.74	5.99	6.32	2.25	2.72	5.01	4.33	4.15	n.a.
G3	13.63	8.99	7.15	8.04	-	6.31	2.84	2.53	5.12	5.00	n.a.	n.a.
T4	**	-	-	7.39	1.81	5.90	1.93	2.39	4.66	3.36	3.18	3.63
T5	n.o.	-	-	7.85	1.93	6.49	2.30	2.51	4.58	4.02	n.a.	n.a.
G6	13.51	9.13	5.88	8.19	-	6.19	2.82	2.95	5.03	4.62	4.23	4.06
C7	-	8.59	6.53	7.55	5.46	6.18	2.43	2.58	4.98	4.29	4.43	n.a.
T8	**	-	-	7.56	1.62	5.80	2.33	1.98	4.40	3.81	n.a.	n.a.

n.o – not observed. n.a – not assigned. ** Imino protons of T4 and T8 resonate at 10.42 and 10.90 ppm

Table S4- Average values of the torsion angles of the dimeric structure of d<pCGCTCCGT>.

d<pCGCTCCGT>														
	α		β		γ		δ		ε		ζ		χ	
	Average	Order Par.	Average	Order Par	Average	Order Par.	Average	Order Par.	Average	Order Par.	Average	Order Par.	Average	Order Par.
1C			167.2	0.79	-17.8	0.41	145.2	0.98	-123.8	0.59			-126.0	1.00
2G	53.7	0.67	-177.1	0.89	-169.1	0.76	134.9	1.00	-169.4	1.00	94.2	0.67	-113.5	1.00
3C	-77.4	0.42	177.9	0.97	76.6	0.51	108.1	0.98	-104.1	0.84	-99.4	0.99	-134.1	1.00
4T	-68.0	0.99	170.2	0.93	43.7	1.00	145.1	1.00	-119.6	0.92	164.7	0.86	-147.2	1.00
5C	1.2	0.17	-173.7	0.93	-74.7	0.14	147.2	0.99	-93.1	0.86	62.1	0.91	-124.3	1.00
6C	144.1	0.21	177.1	0.76	175.5	0.84	129.1	0.98	168.6	0.53	-23.2	0.20	-115.4	0.99
7G	-85.0	0.61	174	0.98	58.3	0.86	116.3	1.00	-118.2	0.92	-72.6	0.44	-116.9	0.99
8T	-61.2	0.66	171.5	1.00	51.1	0.65	148.2	1.00	52.4	0.99	145	0.94	-154.2	1.00
9C	30.8	0.46	161.4	0.94	20.6	0.25	145.6	0.99	-116.0	0.81	51.8	0.98	-128.2	1.00
10G	59.1	0.38	176.6	0.99	-178.1	0.66	136.8	1.00	-169.6	1.00	77.8	0.41	-110.7	1.00
11C	-77.6	0.42	178.8	0.95	77.9	0.50	110.2	0.99	-98.4	0.97	-98.9	0.99	-134.6	1.00
12T	-70.4	0.99	166.9	1.00	41.0	1.00	145.7	1.00	-101.1	0.86	162.3	0.98	-146.0	1.00
13C	-50.2	0.23	-169.5	0.92	-172.2	0.10	143.8	0.99	-116.9	0.77	45.6	0.67	-125.0	1.00
14C	73.2	0.27	-174.6	0.85	177.5	0.68	128.5	0.99	168.2	0.53	-61.9	0.22	-116.5	1.00
15G	-82.3	0.62	171.7	0.99	59.3	0.87	116.9	1.00	-120.3	0.98	-76.4	0.43	-118.4	0.98
16T	-70.4	1.00	171.7	1.00	49.3	1.00	143.4	1.00			152.2	0.98	-153.6	1.00

Table S5- Average values of the torsion angles of the dimeric structure of d(TGCTTCGT).

d(TGCTTCGT)														
	α		β		γ		δ		ε		ζ		χ	
	Average	Order Par.	Average	Order Par	Average	Order Par.	Average	Order Par.	Average	Order Par.	Average	Order Par	Average	Order Par.
1T					138.2	0.35	128.2	0.90	-91.7	0.99			-42.2	1.00
2G	-161.0	0.13	-178.1	0.97	82.3	0.25	133.6	0.95	139.5	0.47	-83.6	0.97	-106.1	1.00
3C	-105.0	0.75	-177.9	1.00	47.5	1.00	99.5	0.97	-81.6	0.59	-52.2	0.30	-126.6	1.00
4T	-31.4	0.58	125.4	0.74	23.7	0.87	84.2	0.99	-108.0	0.29	122.7	0.54	-167.1	0.99
5T	7.2	0.26	143.0	0.73	-149.0	0.52	114.8	0.92	-129.4	0.48	122.5	0.42	-99.4	0.96
6C	91.0	0.54	-167.0	0.87	179.4	0.82	88.2	0.99	66.7	0.67	125.0	0.16	-123.3	1.00
7G	-143.6	0.68	-172.6	0.86	42.8	0.99	93.1	0.96	-27.6	0.62	51.6	0.64	-114.7	0.99
8T	-139.3	0.67	-162.0	0.71	41.1	0.83	100.5	1.00			76.3	0.45	-153.8	1.00
9T					159.2	0.43	129.5	0.71	-87.3	0.99			-42.2	1.00
10G	56.5	0.71	-173.0	0.90	-160.4	0.84	145.0	0.99	169.1	0.71	-89.0	0.96	-104.3	1.00
11C	-88.4	0.82	178.3	1.00	48.1	1.00	106.6	1.00	-88.1	0.99	-82.0	0.61	-126.5	1.00
12T	-50.5	0.69	131.1	0.78	36.0	1.00	85.0	1.00	-108.7	0.50	126.2	0.93	-171.2	1.00
13T	44.6	0.55	-175.7	0.56	-141.3	0.33	106.8	0.86	-149.4	0.82	100.7	0.56	-104.2	0.97
14C	67.0	0.63	-175.5	0.93	173.9	0.65	92.1	0.99	119.6	0.35	110.9	0.54	-122.7	1.00
15G	-110.7	0.61	-174.0	0.88	43.5	0.99	96.0	0.97	-45.9	0.65	-11.4	0.37	-118.6	0.99
16T	-113.0	0.76	-156.1	0.85	47.8	0.87	98.9	1.00			82.4	0.64	-155.6	1.00

Table S6- Average values of the torsion angles of the dimeric structure of d(TCGTTGCT).

d(TCGTTGCT)														
	α		β		γ		δ		ε		ζ		χ	
	Average	Order Par.	Average	Order Par	Average	Order Par.	Average	Order Par.	Average	Order Par.	Average	Order Par	Average	Order Par.
1T			-131.2	0.12	-159.7	0.73	102.5	0.98	-109.9	0.73			-123.6	0.99
2C	94.7	0.37	146.9	0.36	-116.9	0.44	108.8	1.00	-179.5	0.99	101.8	0.41	-112.1	1.00
3G	-78.6	0.99	171.1	0.99	54.6	0.99	100.6	0.99	-98.3	0.72	-95.3	1.00	-143.8	1.00
4T	10.5	0.33	-176.2	0.97	118.5	0.18	154.0	0.99	-120.4	0.90	144.0	0.70	-152.1	0.99
5T	-32.3	0.44	165.6	0.54	160.5	0.64	122.2	0.98	-59.2	0.57	63.5	0.81	-84.5	0.97
6G	-59.5	0.39	-118.9	0.47	-58.3	0.70	134.1	0.99	168.1	0.84	34.4	0.26	-105.4	1.00
7C	-80.8	0.22	177.0	0.95	94.5	0.48	77.0	0.99	171.9	0.99	-84.6	0.83	-143.6	1.00
8T	-51.1	1.00	-167.7	0.99	47.3	0.99	85.4	0.99			-105.6	0.99	-159.8	1.00
9T			92.7	0.34	-177.2	0.54	101.2	0.98	-114.5	0.79			-123.6	1.00
10C	86.3	0.37	168.0	0.44	-133.5	0.32	108.7	0.99	-178.3	0.81	-34.0	0.22	-113.3	0.99
11G	-83.6	0.90	171.0	0.97	55.5	0.98	97.9	0.99	-109.3	0.66	-92.7	0.81	-141.4	1.00
12T	-17.8	0.38	-175.4	0.97	78.4	0.28	150.3	0.99	-116.6	0.73	157.4	0.60	-151.7	0.99
13T	-4.0	0.30	134.0	0.61	144.7	0.46	119.5	0.94	-83.4	0.44	48.7	0.63	-89.1	0.96
14G	-75.2	0.34	-137.8	0.58	-43.6	0.70	137.0	1.00	173.7	1.00	38.3	0.30	-108.2	1.00
15C	-65.4	0.99	165.4	1.00	51.2	0.99	76.4	1.00	173.9	0.99	-86.7	1.00	-139.9	1.00
16T	-46.4	0.86	-168.0	0.99	37.5	0.86	89.8	0.96			-106.0	0.98	-155.7	0.99

Table S7: Local helical parameters for the dimeric structure of d<pCGCTCCGT>, d(TGCTTCGT), and d(TCGTTGCT)

d<pCGCTCCGT>												
Step	Shift Dx (Å)		Slide Dy (Å)		Rise Dz (Å)		Tilt τ (º)		Roll ρ (º)		Twist Ω (º)	
	AVERAGE	SD	AVERAGE	SD	AVERAGE	SD	AVERAGE	SD	AVERAGE	SD	AVERAGE	SD
G2/C3	-0.1	0.2	-1.1	0.2	3.2	0.1	3.9	1.6	3.3	3.2	25.4	1.5
C6/G7	0.2	0.2	-0.6	0.1	3.4	0.1	-1.8	1.3	-9.4	1.5	32.3	1.5
C11/G10	0.1	0.2	-1.1	0.2	3.2	0.1	-3.3	2.2	3.2	2.2	24.9	1.7
G15/C14	-0.6	0.3	-0.7	0.2	3.3	0.3	-3.3	3.4	6.7	3.4	31.9	3.0

d(TGCTTCGT)												
Step	Shift Dx (Å)		Slide Dy (Å)		Rise Dz (Å)		Tilt τ (º)		Roll ρ (º)		Twist Ω (º)	
	AVERAGE	SD	AVERAGE	SD	AVERAGE	SD	AVERAGE	SD	AVERAGE	SD	AVERAGE	SD
G2/C3	-0.4	0.2	-1.0	0.2	4.2	0.1	-10.6	1.5	11.6	2.6	28.5	2.2
C6/G7	-0.2	0.1	-0.8	0.2	3.1	0.2	5.0	2.6	7.4	4.2	24.7	1.6
C11/G10	0.4	0.2	-1.0	0.1	4.1	0.2	9.8	1.8	12.2	3.0	28.0	1.7
G15/C14	0.2	0.2	-0.8	0.2	3.1	0.2	-5.2	2.3	9.1	4.9	24.1	1.9

d(TCGTTGCT)												
Step	Shift Dx (Å)		Slide Dy (Å)		Rise Dz (Å)		Tilt τ (º)		Roll ρ (º)		Twist Ω (º)	
	AVERAGE	SD	AVERAGE	SD	AVERAGE	SD	AVERAGE	SD	AVERAGE	SD	AVERAGE	SD
C2/G3	0.1	0.1	-0.8	0.1	3.3	0.1	-1.2	2.5	9.4	5.5	37.5	1.3
G6/C7	0.2	0.2	-0.6	0.1	3.4	0.1	-1.8	1.3	-9.4	1.5	32.3	1.5
G11/C10	-0.4	0.2	-0.7	0.2	3.4	0.3	1.5	4.0	4.7	8.2	38.0	1.1
C15/G14	-0.1	0.1	-0.6	0.1	3.5	0.1	3.2	1.4	-10.8	2.3	32.7	1.7

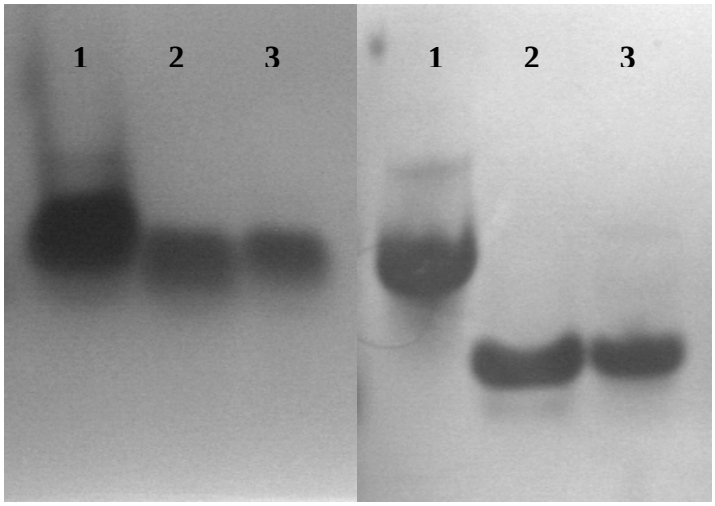


Figure S1 - Nondenaturing polyacrylamide gel electrophoresis of d(TGCTTCGT) and d(TCGTTGCT). The gel on the left is a 20% polyacrilamide containing 90 mMTris-borate (TB), run under experimental conditions where the structures are dimeric (2 mM oligonucleotide concentration, 100 mM Na^+, 50 mM Mg^{2+}, T= 4ºC). The gel on the right is a nondenaturing 20% polyacrilamide gel containing 90 mM Tris-borateEDTA (TBE) run under conditions where the two lineal oligonucleotides do not form dimeric structures (100 µM, oligonucleotide concentration, phosphate buffer 10 mM). The control oligonucleotide (**1**) is the trombin aptamer quadruplex in a 10mM KCl buffer. We have chosen this oligonucleotide as control because it is a 15mer monomeric quadruplex with a globular shape similar to the dimeric minor groove quadruplex studied in this work. As can be seen, both d(TGCTTCGT) (**2**) and d(TCGTTGCT) (**3**), migrate in the gel as monomer or dimers depending on the experimental conditions, but not as a higher order structure.

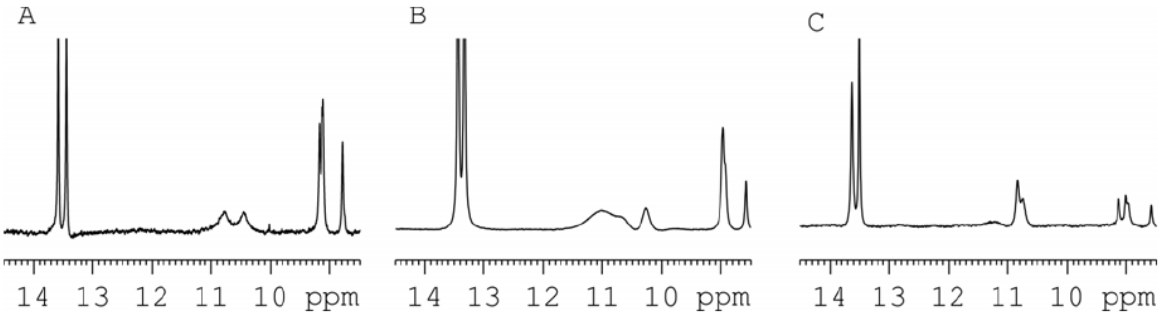


Figure S2 - Imino region of the one dimensional NMR spectra of d<pCGCTCCGT> (A), d(TGCTTCGT) (B) and d(TCGTTGCT) (C) in H_2O at 5ºC (25 mM sodium phosphate buffer, pH 7, 100 mM NaCl, 10 mM $MgCl_2$) acquired with a jump-and-return pulse sequence with a maximum excitation profile at 11.5 ppm. The broad signals observed between 10-11 ppm correspond to the imino protons of the unpaired thymines. These protons exchange rapidly with the solvent and they are not observed in spectra acquired with a WATERGATE pulse sequence for solvent suppression shown in Figure 2 of the main text.

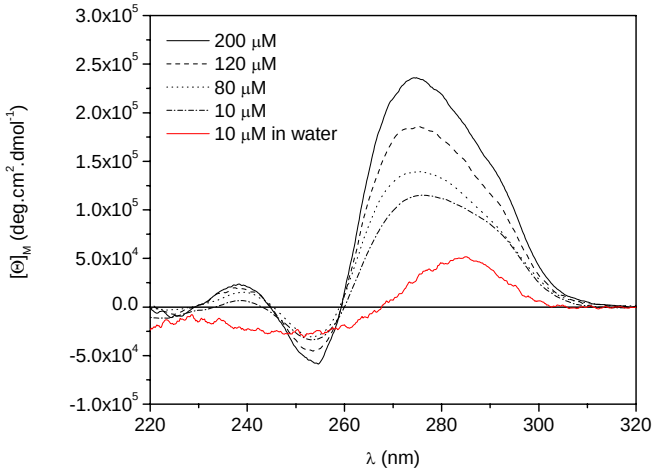


Figure S3 – Black lines: CD spectra of d<pCGCTCCGT> at different oligonucleotide concentrations recorded in 25 mM sodium phosphate buffer, pH 7, 100 mM NaCl, 10 mM $MgCl_2$. Red line. CD spectra of d<pCGCTCCGT> at 10 μM oligonucleotide concentration in water (no salt added), under these conditions the structure of d<pCGCTCCGT> is mainly monomeric.

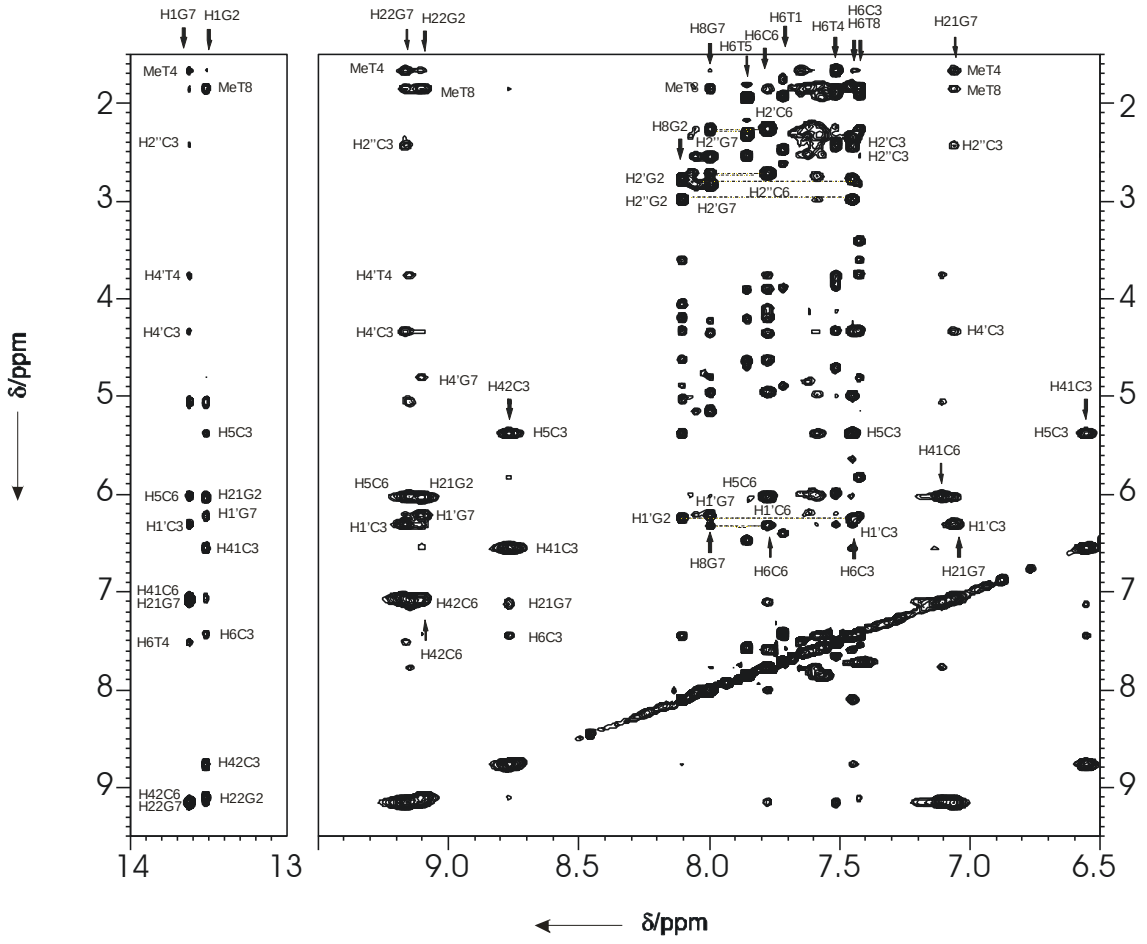


Figure S4 - Regions of the NOESY spectrum (t_m=150 ms) of d(TGCTTCGT) in H_2O (4 mM oligonucleotide concentration, 100 mM NaCl, 10 mM $MgCl_2$, T=5ºC, pH 7).

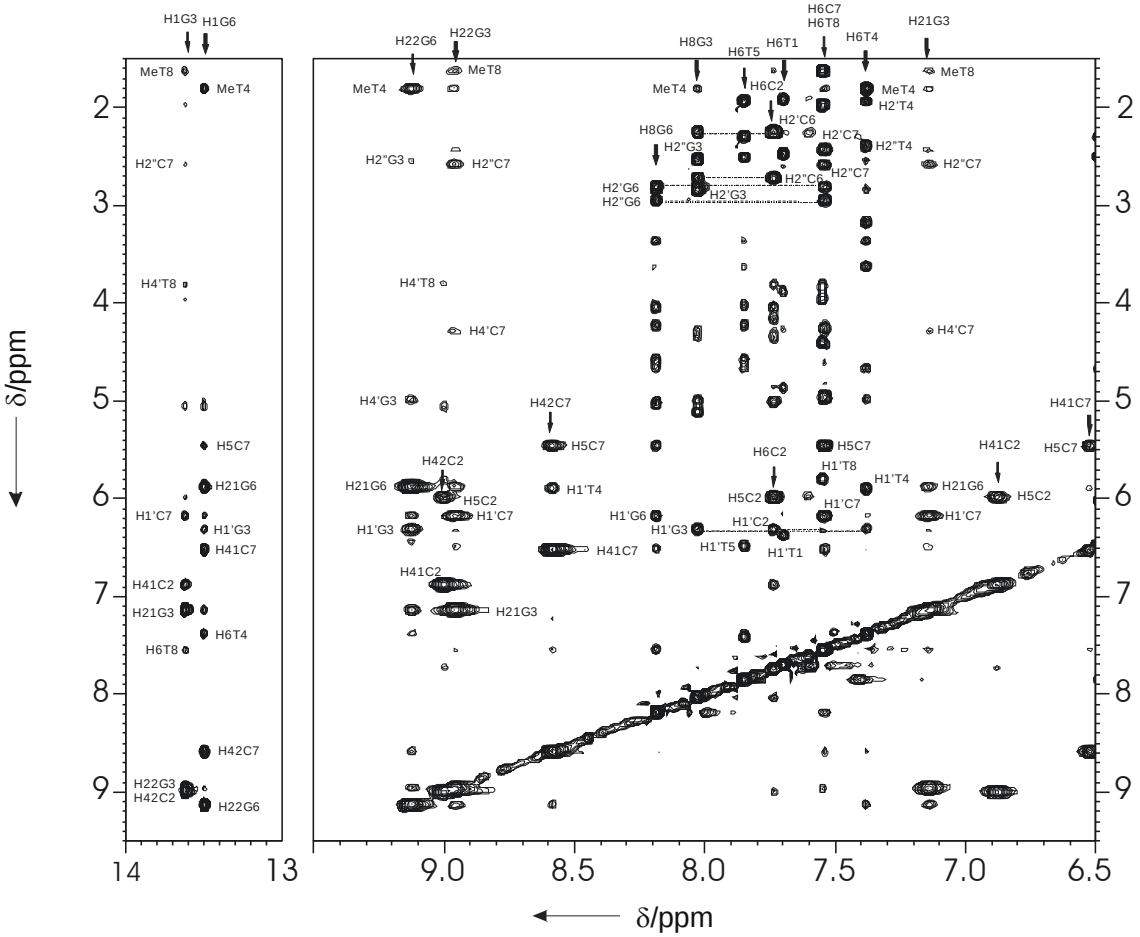


Figure S5 - Regions of the NOESY spectrum (t_m=150 ms) of d(TCGTTGCT) in H_2O (2 mM oligonucleotide concentration, 100 mM NaCl, 10 mM $MgCl_2$, T=5ºC, pH 7).

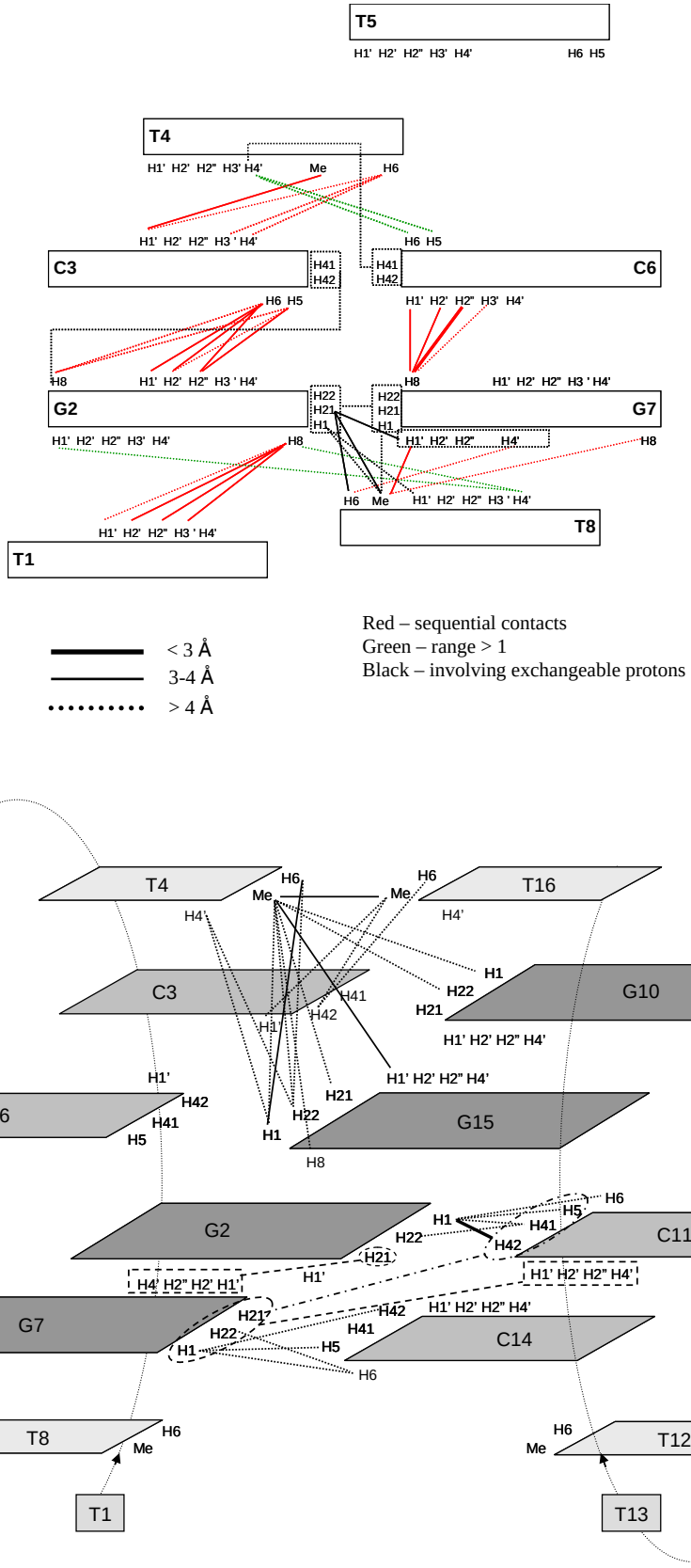
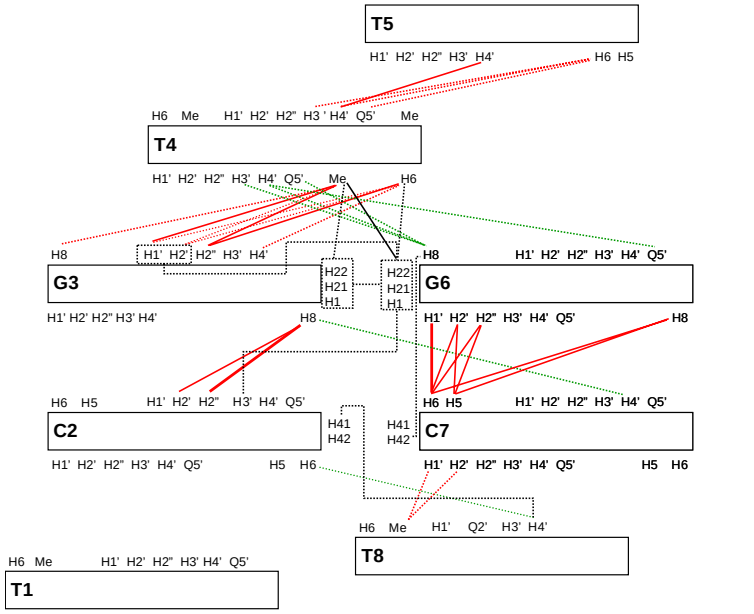


Figure S6 - Schematic representation of intra- (top) and inter-molecular (bottom) distance constraints for d(TGCTTCGT). Constraints are classified in three categories according to their upper distance limit.



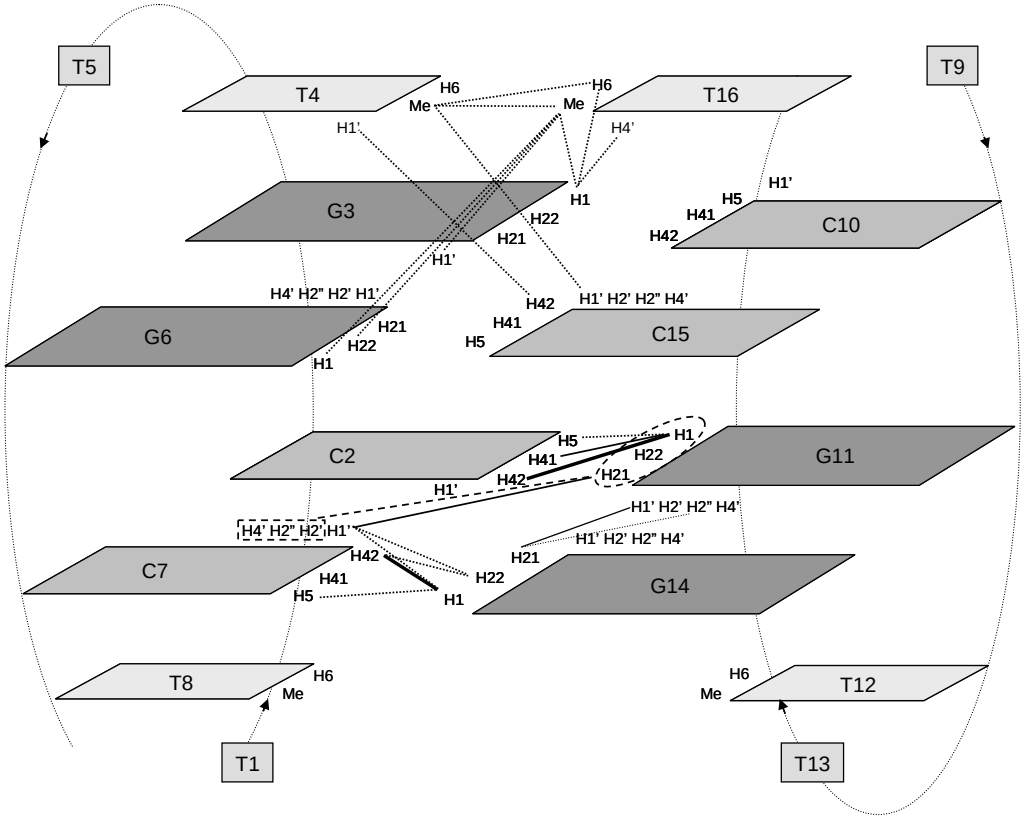
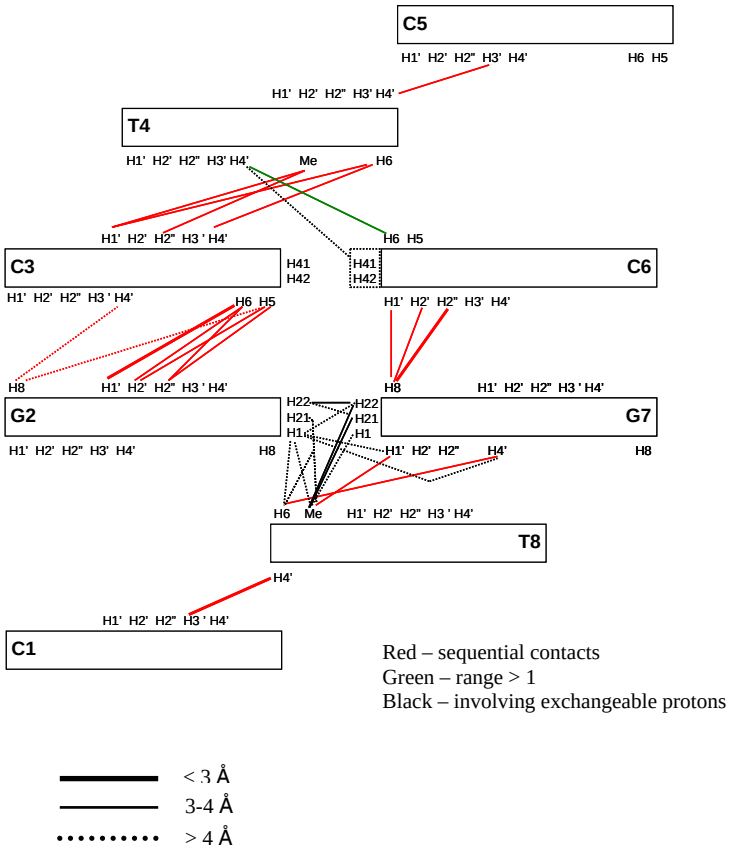


Figure S7 - Schematic representation of intra- (top) and inter-molecular (bottom) distance constraints for d(TCGTTGCT). Constraints are classified in three categories according to their upper distance limit.



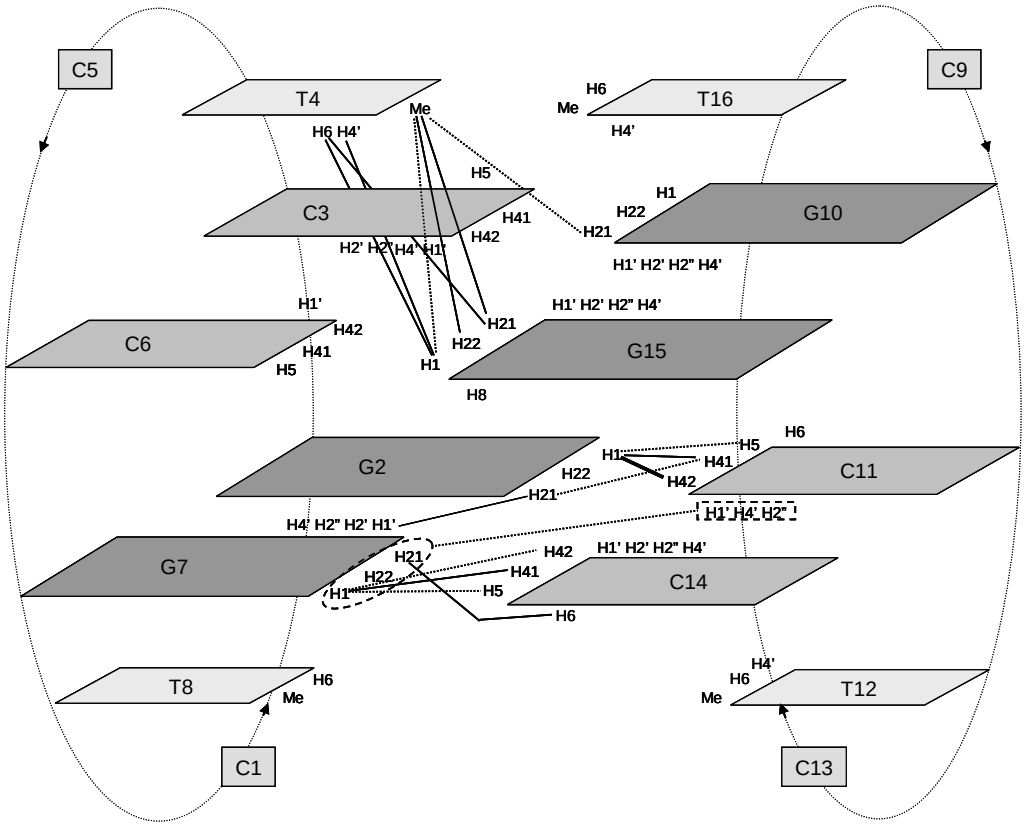


Figure S8 - Schematic representation of intra- (top) and inter-molecular (bottom) distance constraints for d<pCGCTCCGT>. Constraints are classified in three categories according to their upper distance limit.

Distinction between intra- and inter-molecular distances: The symmetry of these dimeric structures implies that for every correlation cross-peak observed between two protons, two different distance constraints are possible, the intra- and the inter-molecular distance constraints. A number of cross-peaks can be unequivocally assigned as intra- or inter-molecular distances. For the rest of the cross-peaks, a preliminary model of the structure based on similar structures is used to decide between intra- or inter-molecular contacts and, with the restrictions obtained a first round of preliminary structure calculation with DYANA is performed. Those distance constraints that are systematically violated (> 1Å) are reassigned and again introduced in the structure calculation. This process is repeated several times until a consistent distance constraints collection is obtained. The cross-peaks that remain ambiguous after this process are not included in the final structure calculations.