

d(TG_nT) DNA sequences do not necessarily form tetramolecular G-quadruplexes

Laure Joly,^a Frédéric Rosu^a and Valérie Gabelica^{*a}

^a Physical Chemistry and Mass Spectrometry Laboratory, Department of Chemistry, University of Liège, B-4000 Liège, Belgium. Fax: +32 4 3663413; Tel: +32 4 3663432; E-mail: v.gabelica@ulg.ac.be

Electronic Supporting Information.

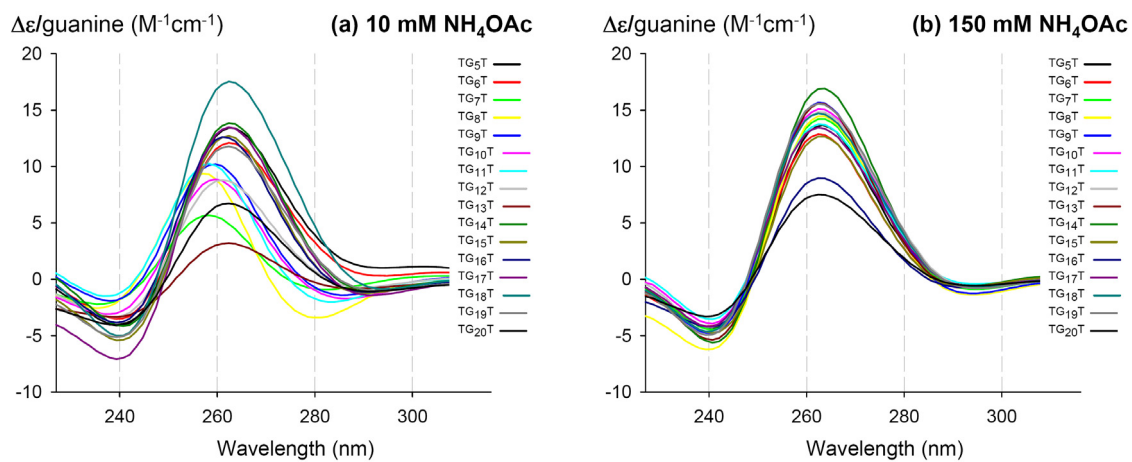


Figure S1: Complete set of CD spectra.

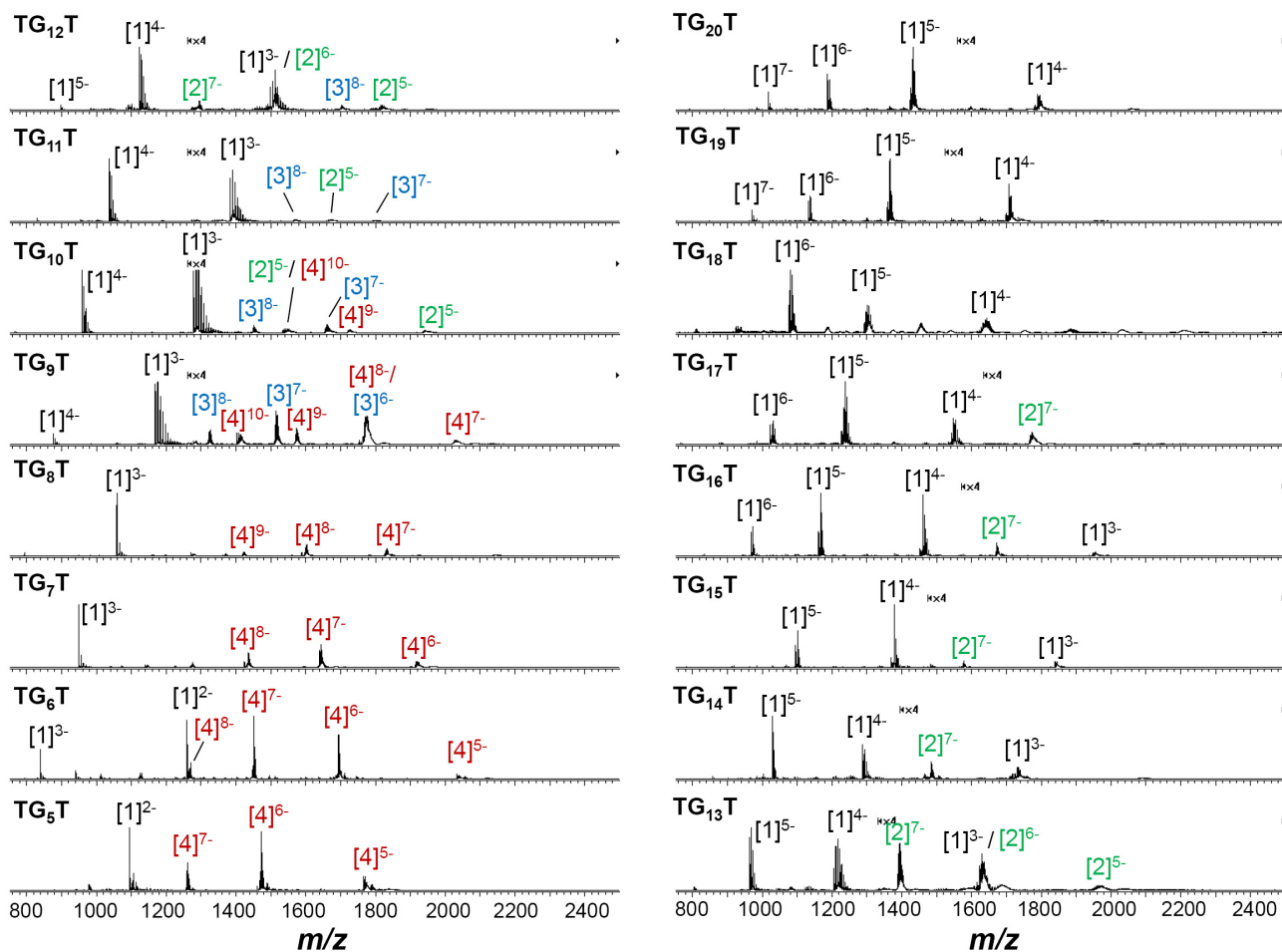


Figure S2: Complete set of ESI-MS spectra in 10 mM NH₄OAc. Note that spectra from n=9 to n=20 contain zooms (signal magnified 4 times) in the multimer region. The monomeric species is the most abundant. The spectra were smoothed (mean function, 2*20 channels) and background-subtracted for clarity.

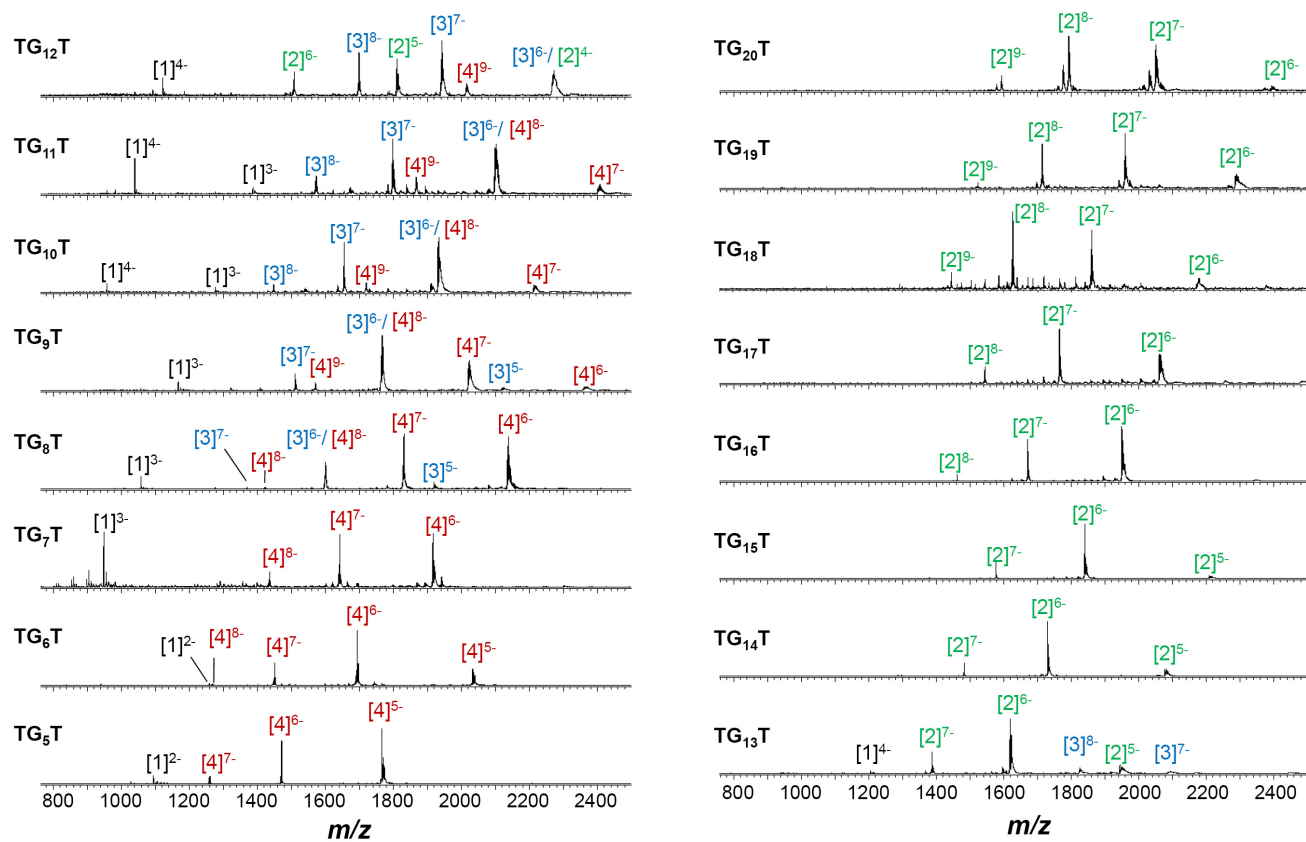


Figure S3: Complete set of ESI-MS spectra in 150 mM NH₄OAc. The spectra were smoothed (mean function, 2*20 channels) and background-subtracted for clarity.

Table S1: For sequences of different G-tract length (n), summary of the number of strands (s) detected in the assemblies, number of ammonium ions (x) retained in each species observed by ESI-MS, deduction of the number of G-quartets ($x+1$), the number of guanines involved in G-quartets ($4*(x+1)$), and the number of guanines not participating to G-quartets ($s*n - 4*(x+1)$). The number of strands of the species with the largest mass spectral peak areas is underlined.

n	150 mM NH ₄ OAc					10 mM NH ₄ OAc				
	s	x	Quartets ($x+1$)	G in quartet: $4*(x+1)$	G out of quartet: $s*n-4*(x+1)$	s	x	Quartets ($x+1$)	G in quartet: $4*(x+1)$	G out of quartet: $s*n-4*(x+1)$
5	<u>4</u>	4	5	20	0	<u>4</u>	4	5	20	0
6	<u>4</u>	5	6	24	0	<u>4</u>	5	6	24	0
7	<u>4</u>	6	7	28	0	<u>4</u>	5-6	6-7	24-28	4-0
8	<u>4</u>	5-6	6-7	24-28	8-4	<u>4</u>	6-7	7-8	28-32	4-0
	<u>3</u>	4	5	20	4	<u>1</u>	0 ^(a)	--	--	--
9	<u>4</u>	7-8	8-9	32-36	4-0	<u>1</u>	0 ^(a)	--	--	--
	<u>3</u>	4	5	20	7	<u>3</u>	0 ^(a)	--	--	--
						<u>4</u>	7-8	8-9	32-36	4-0
10	<u>4</u>	8	9	36	4	<u>1</u>	0 ^(a)	--	--	--
	<u>3</u>	5	6	24	6	<u>2</u>	0 ^(a)	--	--	--
						<u>3</u>	0 ^(a)	--	--	--
						<u>4</u>	0 ^(a)	--	--	--
11	<u>4</u>	9-10	10-11	40-44	4-0	<u>1</u>	0 ^(a)	--	--	--
	<u>3</u>	6	7	28	13	<u>2</u>	0 ^(a)	--	--	--
						<u>3</u>	0 ^(a)	--	--	--
12	<u>3</u>	7	8	32	4	<u>1</u>	0 ^(a)	--	--	--
	<u>2</u>	4	5	20	4	<u>2</u>	0 ^(a)	--	--	--
						<u>3</u>	0 ^(a)	--	--	--
13	<u>3</u>	8	9	36	3	<u>1</u>	0 ^(a)	--	--	--
	<u>2</u>	4	5	20	6	<u>2</u>	0 ^(a)	--	--	--
14	<u>2</u>	5	6	24	4	<u>1</u>	0 ^(a)	--	--	--
15	<u>2</u>	5	6	24	6	<u>1</u>	2	3	12	3
16	<u>2</u>	5	6	24	8	<u>1</u>	2	3	12	4
17	<u>2</u>	5	6	24	12	<u>1</u>	2	3	12	5
18	<u>2</u>	6	7	28	8	<u>1</u>	2	3	12	6
19	<u>2</u>	7	8	32	6	<u>1</u>	2	3	12	7
20	<u>2</u>	7	8	32	8	<u>1</u>	2	3	12	8

^(a) Only a broad distribution of sodium adducts was detected (for a typical spectrum, see inset of Figure 2, $n=14$, in the main text). As this distribution is statistical and starts with zero adducts, we estimate that there is not a specific number of G-quartets formed in these structures.