

NMR Structure of a Triangulenium based Long-lived Fluorescence Probe

Bound to G-quadruplex

Anita Kotar^[a], Baifan Wang^[a], Arun Shivalingham^[b], Jorge Gonzalez-Garcia^[b], Ramon Vilar^[b] and Janez Plavec^{*[a]}

^[a]Slovenian NMR Center, National Institute of Chemistry, Hajdrihova 19, 1000 Ljubljana (Slovenia); EN-FIST Center of Excellence, Trg OF 13, 1000 Ljubljana (Slovenia); Faculty of Chemistry and Chemical Technology, University of Ljubljana, Večna pot 113, Ljubljana (Slovenia)

^[b]Department of Chemistry, Imperial College London, South Kensington, London SW7 2AZ (UK)

Corresponding author: Prof Janez Plavec; E-mail: Janez.plavec@ki.si

Supplementary Information

Figures

- Figure S1. Assignment of G8, G12, G14 and G18 of 1:2 CMA-ligand complex.
- Figure S2. Imino-imino region of NOESY spectrum of 1:2 CMA-ligand complex.
- Figure S3. Anomeric-aromatic region of NOESY spectrum of 1:2 CMA-ligand complex.
- Figure S4. Sequential walk in the anomeric–aromatic region of NOESY spectrum of 1:2 CMA-ligand complex.
- Figure S5. Imino-aromatic region of NOESY spectrum of 1:2 CMA-ligand complex.
- Figure S6. ^1H - ^{13}C HSQC spectrum of 1:2 CMA-ligand complex.
- Figure S7. CD spectra of **DAOTA-M2**, free CMA G-quadruplex and 1:0.5, 1:1, 1:2 and 1:4 CMA-ligand complexes.
- Figure S8. Imino region of ^1H NMR spectra of CMA G-quadruplex and its modified analogs.
- Figure S9. Changes in chemical shifts of methyl protons of T1, T10 and T19 residues in the presence of **DAOTA-M2**.
- Figure S10. NOESY spectrum of **DAOTA-M2** at the 1:25 CMA-ligand ratio.
- Figure S11. ^1H NMR spectrum of ligand **DAOTA-M2** in the presence of folded CMA G-quadruplex at 1:10 CMA-ligand ratio.
- Figure S12. Imino-imino region of NOESY spectrum of 1:0.5 CMA-ligand.
- Figure S13. NMR variable temperature experiments.
- Figure S14. Schematic diagram of NOE connectivities.
- Figure S15. NOE cross-peaks resulting from interactions between **DAOTA-M2** and CMA G-quadruplex.
- Figure S16. ^1H NMR spectra of methyl region of modified CMA G-quadruplex CMA-A2.
- Figure S17. RMSD values of G-quadruplex backbone, G-quartet and **DAOTA-M2** molecules at the 5'- and 3'-end.
- Figure S18. Probability distribution of torsion angles of ethylene linkers between morpholino groups and triangulenium core of **DAOTA-M2**.
- Figure S19. Electrostatic surface of CMA G-quadruplex.

Tables

- Table S1. ^1H NMR chemical shifts of 1:2 CMA-ligand complex.
- Table S2. ^1H NMR chemical shifts changes induced by interactions between CMA G-quadruplex and **DAOTA-M2**.
- Table S3. NMR restraints and structural statistics.
- Table S4. Clustering analysis of binding modes of **DAOTA-M2** at the 5'-end.
- Table S5. Clustering analysis of binding modes of **DAOTA-M2** at the 3'-end.

Material and methods

- NMR spectroscopy.
- Circular Dichroism (CD) spectroscopy.
- UV melting experiments.
- NMR structure calculation.
- Data deposition.
- Molecular dynamics simulation.
- References

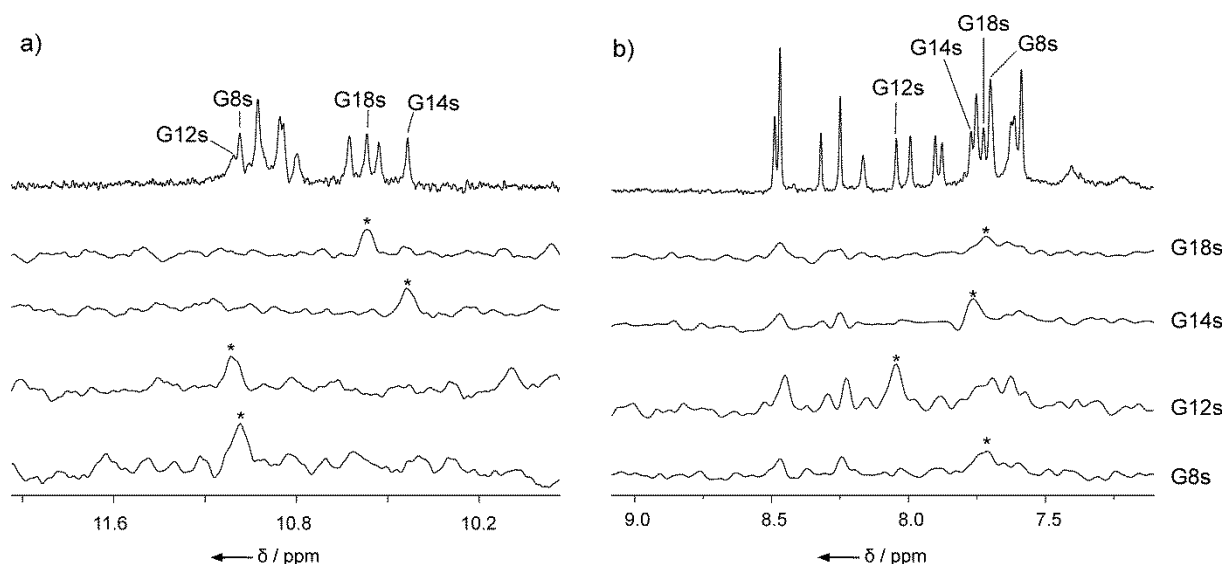


Figure S1. Assignment of G8, G12, G14 and G18 residues of 1:2 CMA-ligand complex. a) Imino region of ^1H NMR spectrum of 1:2 CMA-ligand complex (top). 1D ^{15}N -edited HSQC spectra acquired on partially (8%) residue-specific ^{15}N , ^{13}C -labelled oligonucleotides. b) Aromatic region of ^1H NMR spectrum of 1:2 CMA-ligand complex (top). 1D ^{13}C -edited HSQC spectra acquired on partially (8%) residue-specific ^{15}N , ^{13}C -labelled oligonucleotides. NMR spectra were recorded on 600 MHz spectrometer at 0.1 mM concentration of CMA, 100 mM KCl, 20 mM phosphate buffer (pH 7.0) in 90%/10% $\text{H}_2\text{O}/^2\text{H}_2\text{O}$ at 25 $^\circ\text{C}$.

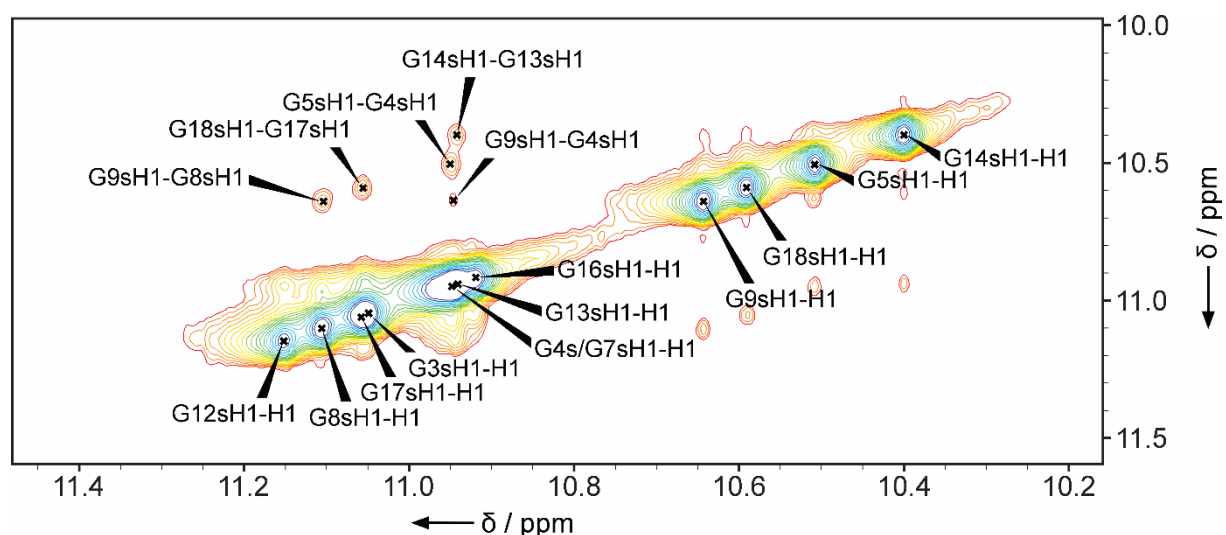


Figure S2. Imino-imino region of NOESY spectrum of 1:2 CMA-ligand complex. NMR spectrum was recorded on 600 MHz spectrometer at 0.2 mM concentration of CMA, 100 mM KCl, 10 mM lithium cacodylate buffer (pH 7.0), mixing time of 300 ms and 25 $^\circ\text{C}$ in 90%/10% $\text{H}_2\text{O}/^2\text{H}_2\text{O}$.

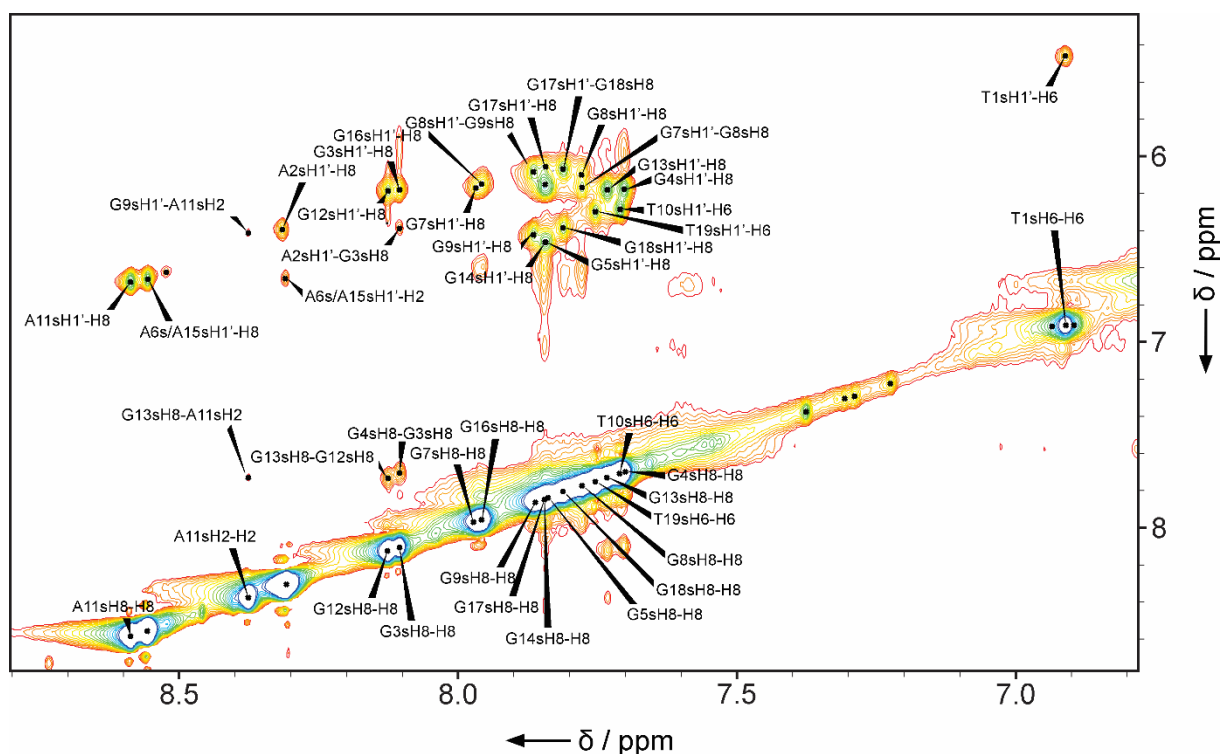


Figure S3. Anomeric-aromatic region of NOESY spectrum of 1:2 CMA-ligand complex. NMR spectrum was recorded on 600 MHz spectrometer at 0.2 mM concentration of CMA, 100 mM KCl, 10 mM lithium cacodylate buffer (pH 7.0), mixing time of 300 ms and 25 °C in 90%/10% $\text{H}_2\text{O}/^2\text{H}_2\text{O}$.

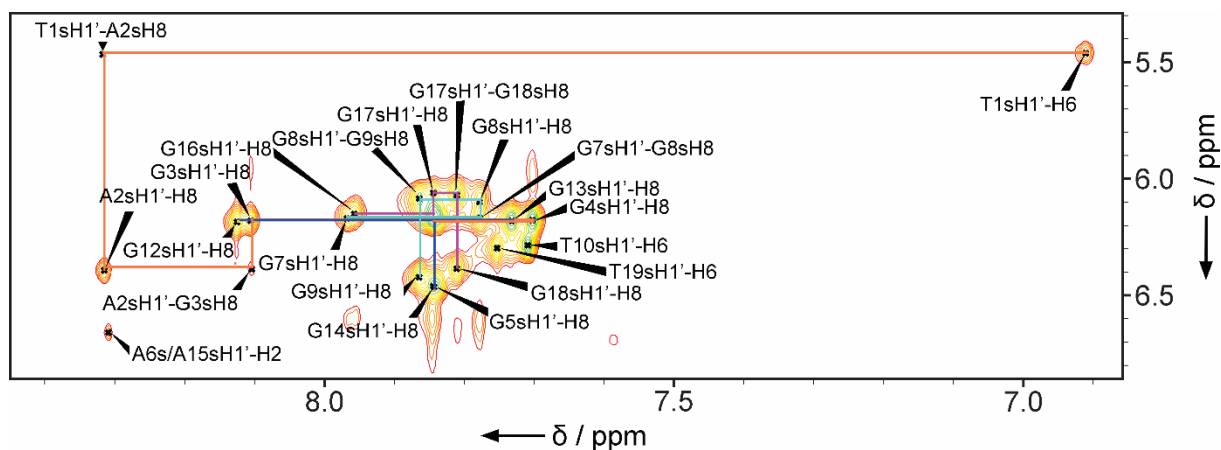


Figure S4. Sequential walk in the anomeric-aromatic region of NOESY spectrum of 1:2 CMA-ligand complex. Residue numbers denote intra-residual H1'-H6/H8 NOE interactions and colored lines denotes sequential correlations within 5'-TAGGGAGGGTAGGGAGGGT-3'. NMR spectrum was recorded on 600 MHz spectrometer at 0.2 mM concentration of CMA, 100 mM KCl, 10 mM lithium cacodylate buffer (pH 7.0), mixing time of 300 ms and 25 °C in 90%/10% $\text{H}_2\text{O}/^2\text{H}_2\text{O}$.

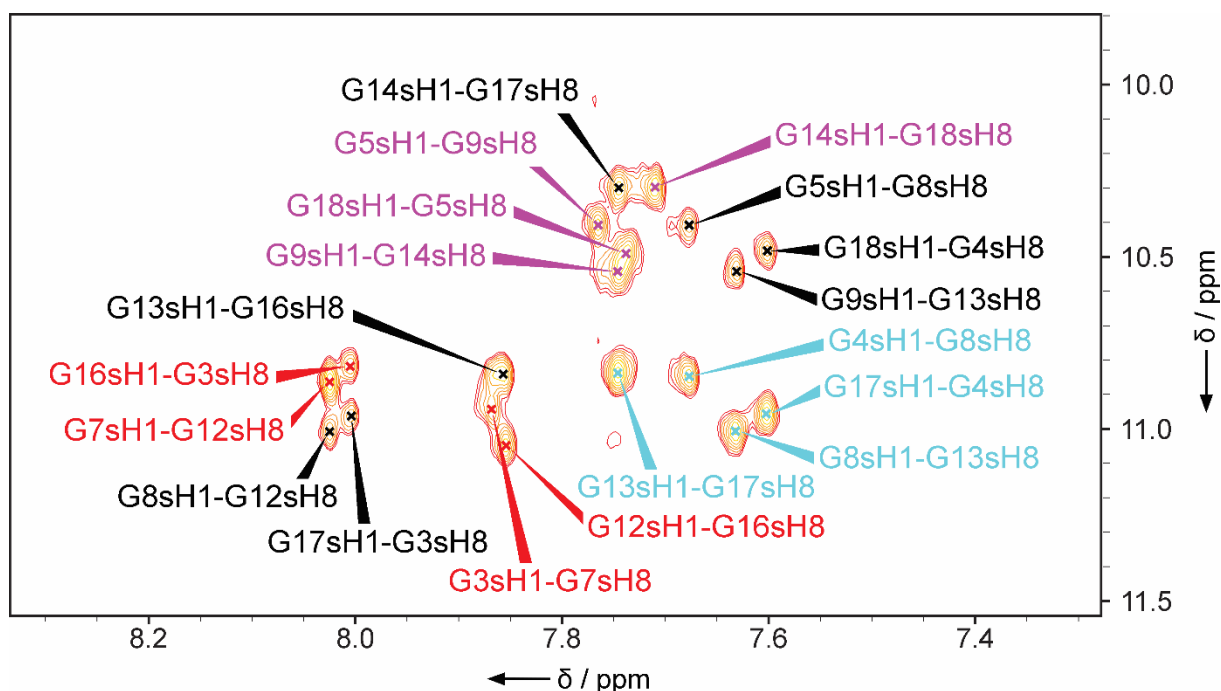


Figure S5. Imino-aromatic region of NOESY spectrum of 1:2 CMA-ligand complex. The H1-H8 cross-peaks within G3-G7-G12-G16, G4-G8-G13-G17 and G5-G9-G14-G18 quartets are designated in red, blue and magenta, respectively. Cross-peaks colored in black correspond to dipole-dipole interactions between protons in neighboring G-quartets. NMR spectrum was recorded on 600 MHz spectrometer at 0.2 mM concentration of CMA, 100 mM KCl, 10 mM lithium cacodylate buffer (pH 7.0), mixing time of 300 ms and 25 °C in 90%/10% H₂O/²H₂O.

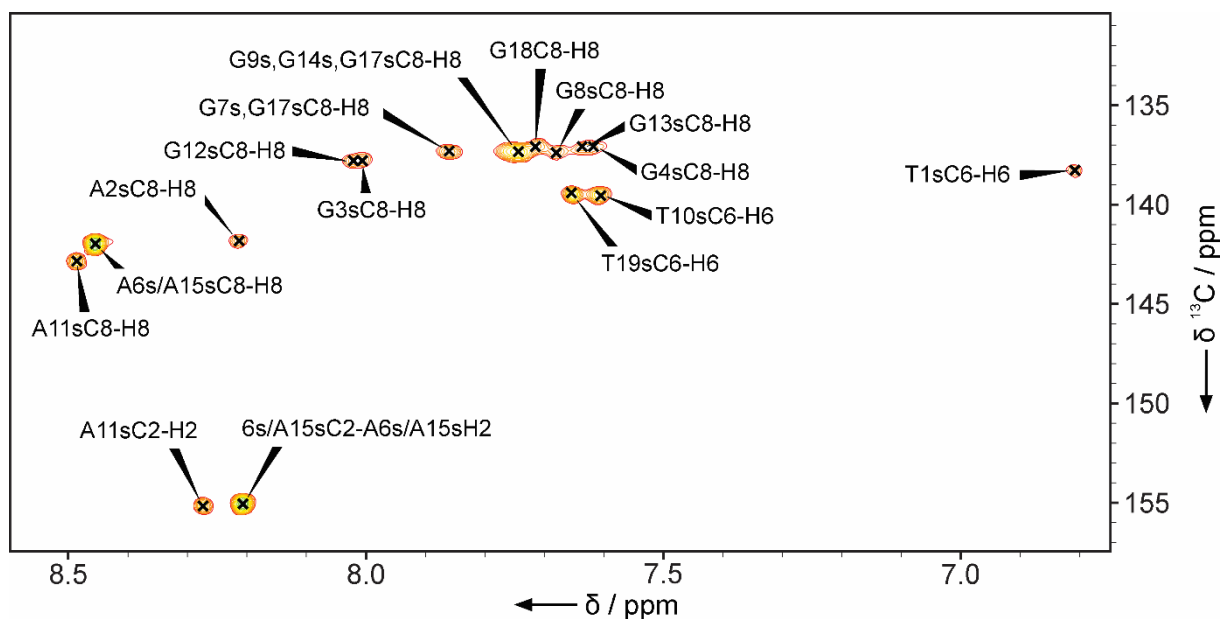


Figure S6. ¹H-¹³C HSQC spectrum of 1:2 CMA-ligand complex showing C8-H8/C6-H6 and C2-H2 correlations. NMR spectrum was recorded on 600 MHz spectrometer at 0.2 mM concentration of CMA, 100 mM KCl, 10 mM lithium cacodylate buffer (pH 7.0) and 25 °C in 90%/10% H₂O/²H₂O.

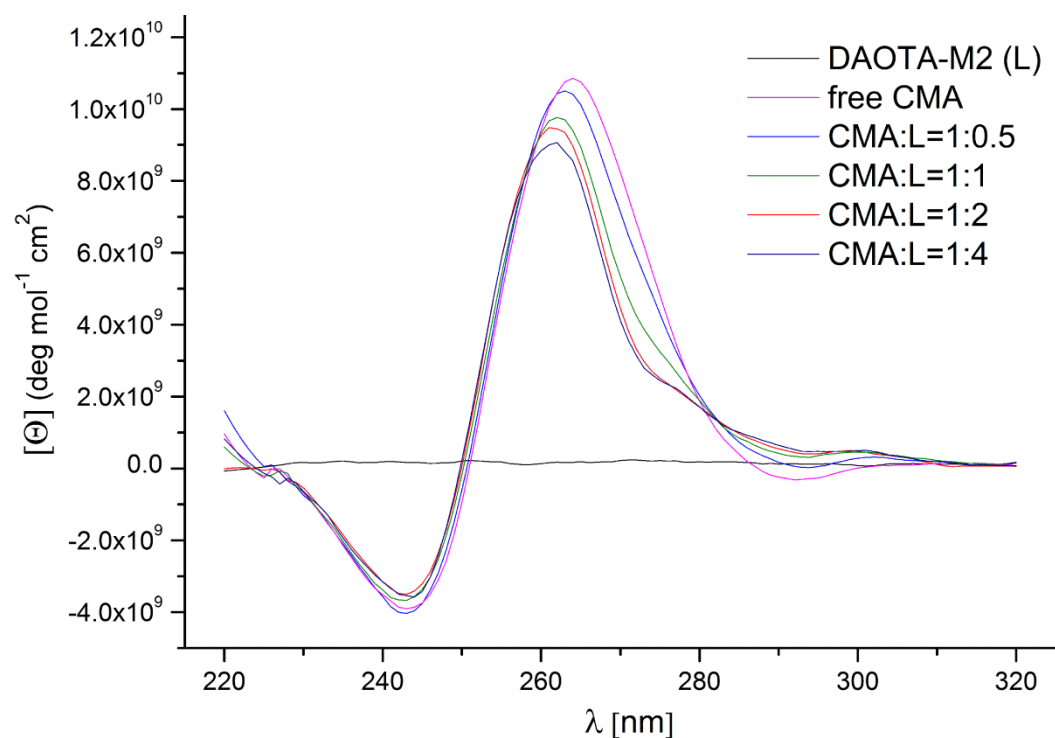


Figure S7. CD spectra of **DAOTA-M2**, free CMA G-quadruplex and 1:0.5, 1:1, 1:2 and 1:4 CMA-ligand complexes. The spectra were recorded at 40 μM concentration of CMA, 100 mM KCl, 20 mM phosphate buffer (pH 7.0) at 25 $^{\circ}\text{C}$.

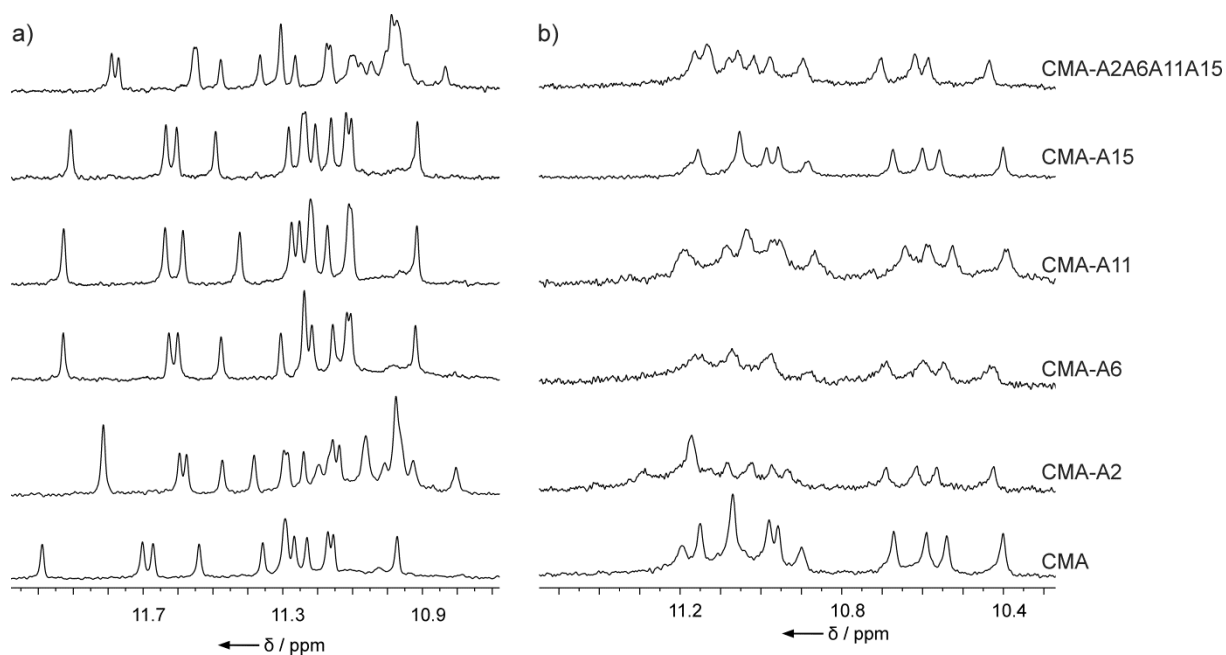


Figure S8. Imino region of ^1H NMR spectra of CMA G-quadruplex and its modified analogs where A2 (CMA-A2), A6 (CMA-A6), A11 (CMA-A11), A15 (CMA-A15) and all adenines (CMA-A2A6A11A15) were replaced with propyl spacer C3. a) G-quadruplexes alone. b) In the presence of two equivalents of **DAOTA-M2**. NMR spectra were recorded on 600 MHz spectrometer at 100 mM KCl, 20 mM phosphate buffer (pH 7.0) and 25 $^{\circ}\text{C}$ in 90%/10% $\text{H}_2\text{O}/^2\text{H}_2\text{O}$.

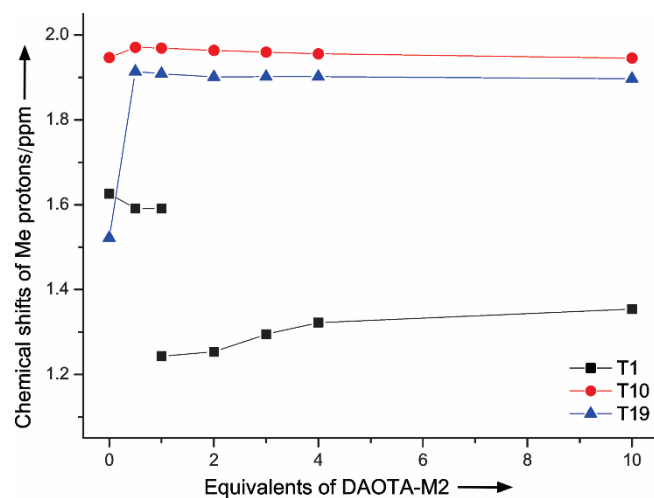


Figure S9. Changes in chemical shifts of methyl protons of T1 (black), T10 (red) and T19 (blue) residues in the presence of various equivalents of **DAOTA-M2**. The concentration of CMA G-quadruplex was 0.2 mM.

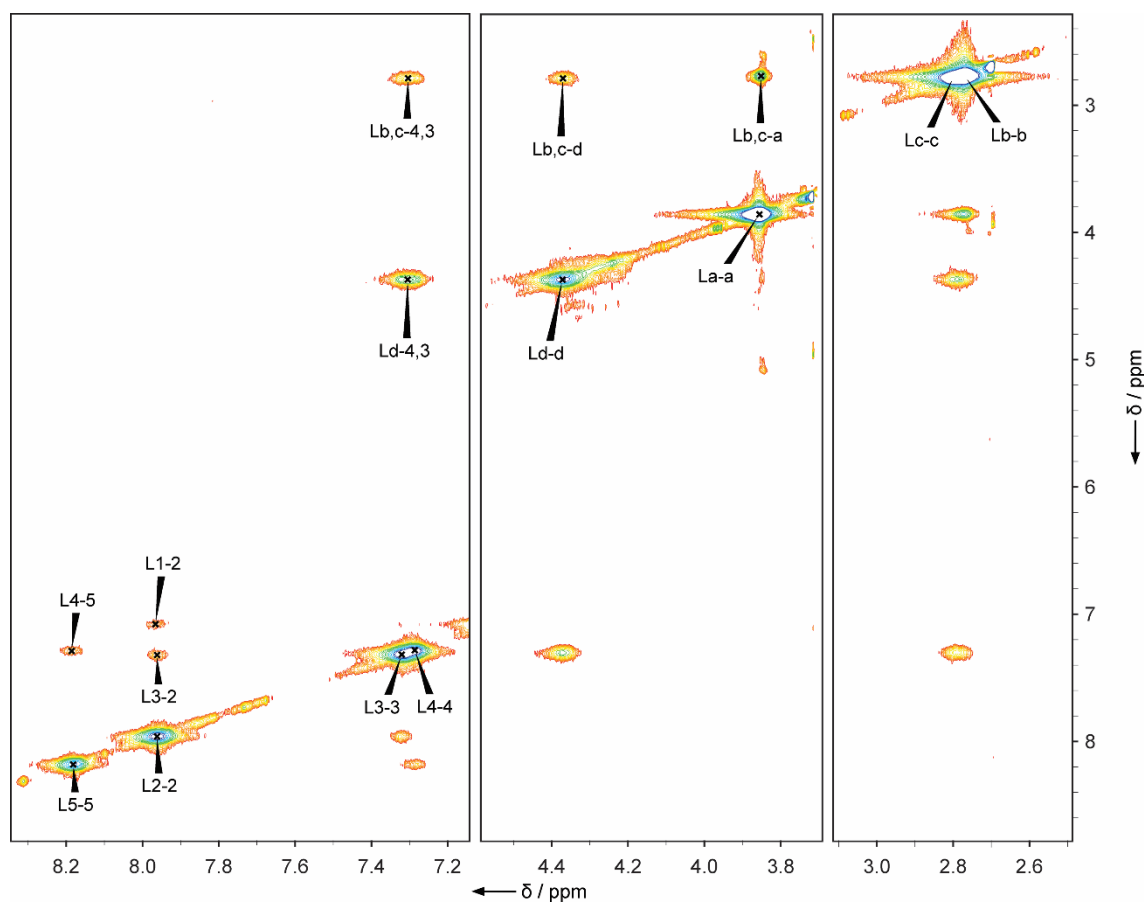


Figure S10. NOESY spectrum of **DAOTA-M2** at the 1:25 CMA-ligand ratio. Diagonal cross-peaks corresponding to the ligand were marked in according to the numbering showed in Figure 1a. NMR spectrum was recorded on 600 MHz spectrometer at 0.5 mM concentration of **DAOTA-M2**, 100 mM KCl, 20 mM phosphate buffer (pH 7.0), mixing time of 300 ms and 25 °C in 90%/10% H₂O/²H₂O.

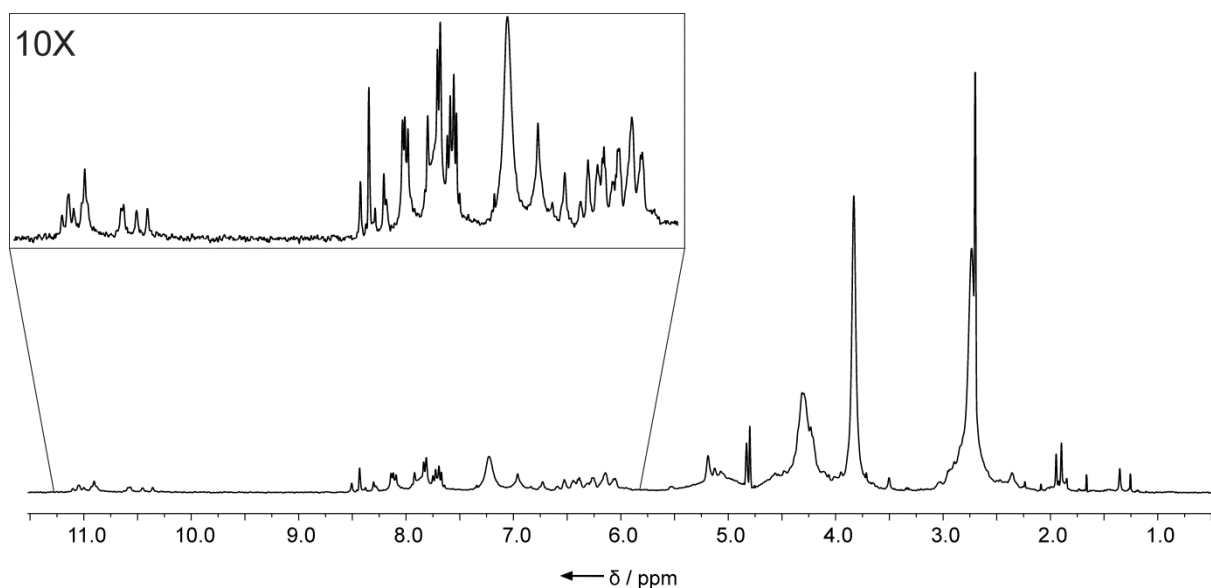


Figure S11. ^1H NMR spectrum of ligand **DAOTA-M2** in the presence of folded CMA G-quadruplex at 1:10 CMA-ligand ratio. NMR spectrum was recorded on 600 MHz spectrometer at 1.0 mM concentration of **DAOTA-M2**, 100 mM KCl, 20 mM phosphate buffer (pH 7.0) and 25 °C in 90%/10% $\text{H}_2\text{O}/^2\text{H}_2\text{O}$.

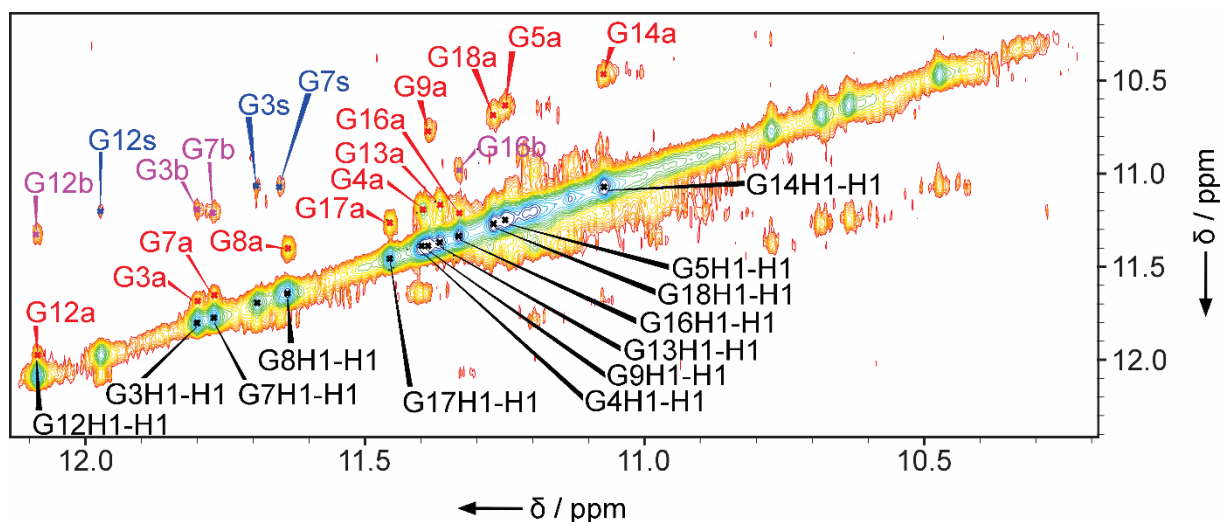


Figure S12. Imino-imino region of NOESY spectrum of CMA G-quadruplex at the 1:0.5 CMA-ligand ratio. Diagonal cross-peaks corresponding to the free CMA G-quadruplex are labelled in black. Cross-peaks arising from chemical exchange between free CMA and CMA3, CMA5, CMA6 are labelled in red, magenta and blue, and denoted by final letters 'a', 'b' and 's', respectively. NMR spectrum was recorded on 600 MHz spectrometer at 0.2 mM concentration of CMA, 100 mM KCl, 20 mM phosphate buffer (pH 7.0), mixing time of 40 ms and 25 °C in 90%/10% $\text{H}_2\text{O}/^2\text{H}_2\text{O}$.

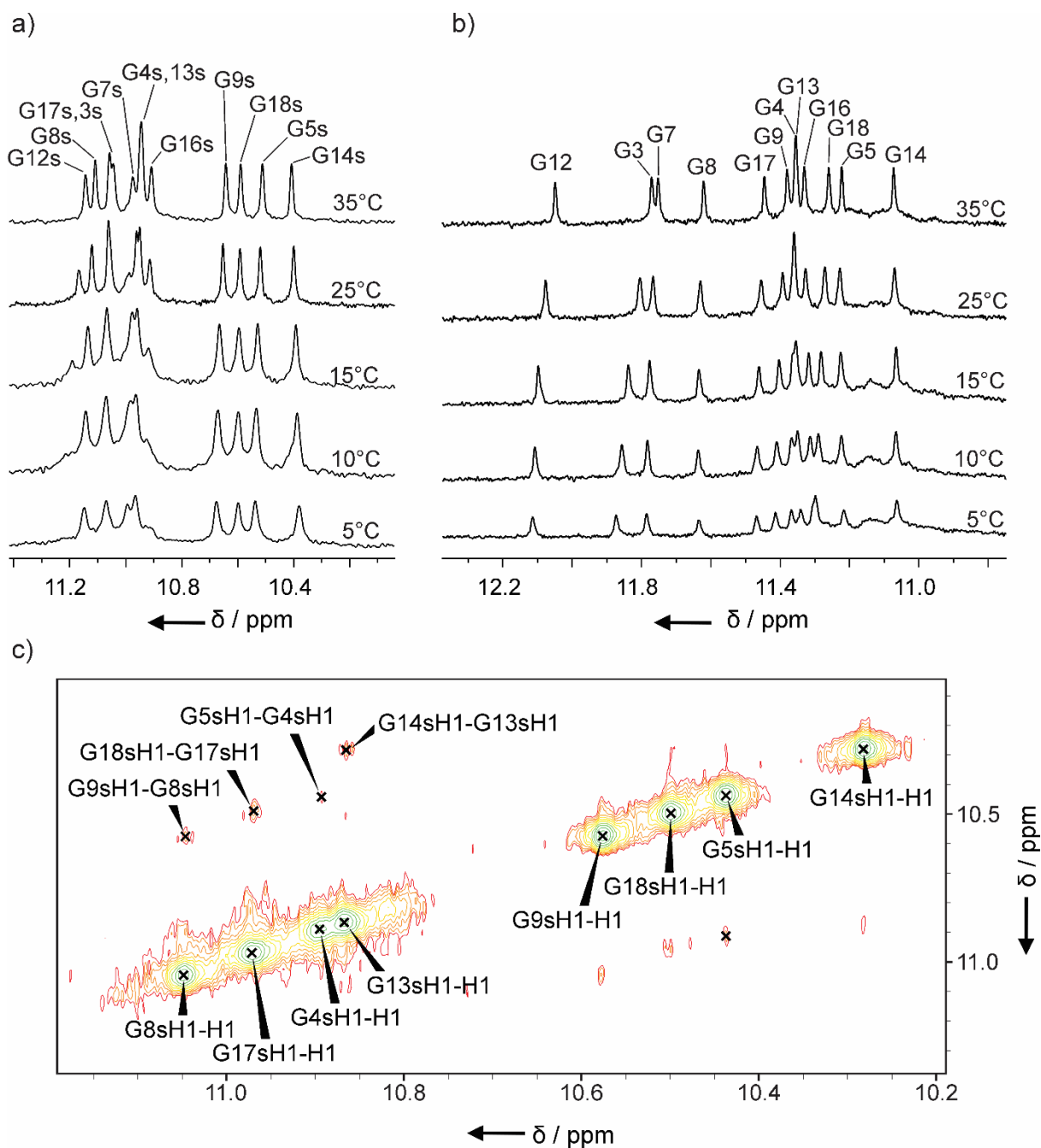


Figure S13. NMR variable temperature experiments. Imino region of ^1H NMR spectra of a) 1:2 CMA-ligand complex and b) free CMA G-quadruplex at temperatures between 5 and 35 °C indicated on the right. c) Imino-imino region of NOESY spectrum at mixing time of 300 ms at 5 °C. NMR spectra were recorded on 800 MHz spectrometer at 0.2 mM concentration of CMA, 100 mM KCl, 20 mM phosphate buffer (pH 7.0) in 90%/10% $\text{H}_2\text{O}/^2\text{H}_2\text{O}$.

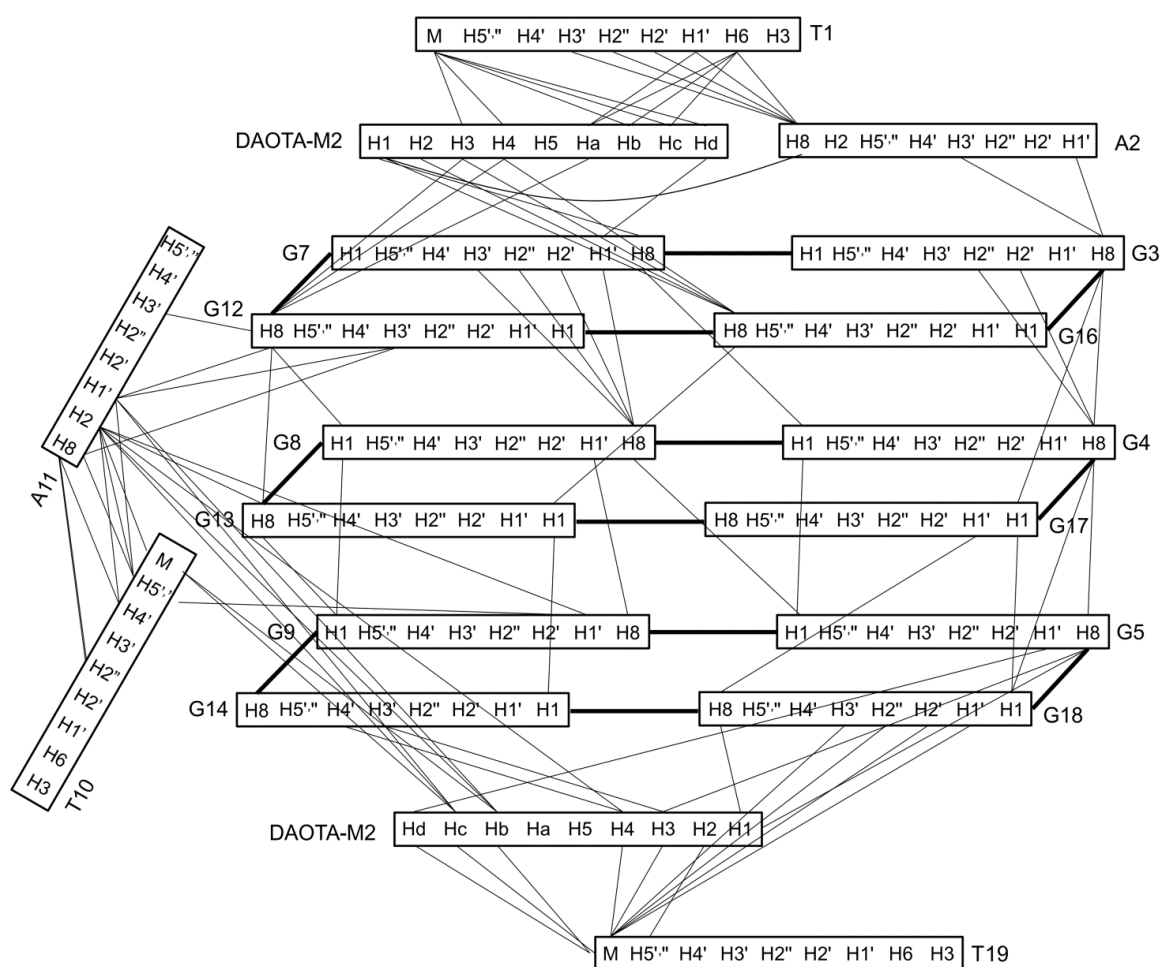


Figure S14. Schematic diagram of NOE connectivities. The black lines represent intermolecular NOE cross-peaks between CMA G-quartet and **DAOTA-M2**, inter-residual NOE cross-peaks of CMA G-quadruplex and of **DAOTA-M2**.

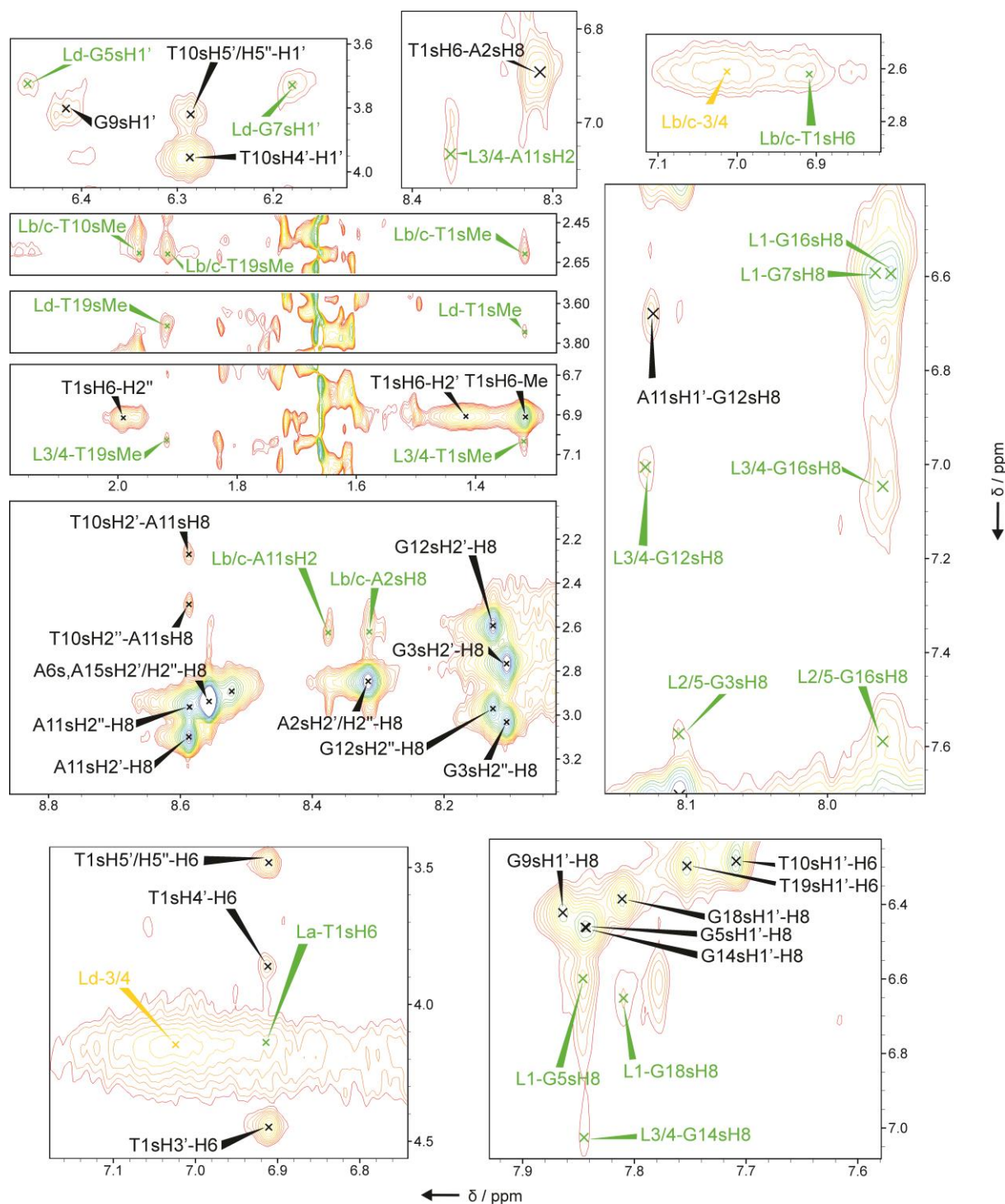


Figure S15. NOE cross-peaks resulting from interactions between **DAOTA-M2** and CMA G-quadruplex. Different regions of NOESY spectrum show intermolecular NOE cross-peaks between **DAOTA-M2** and CMA G-quadruplex (green), inter-residual NOE cross-peaks of CMA G-quadruplex (black) and of **DAOTA-M2** (yellow). NMR spectrum was recorded on 600 MHz spectrometer at 0.2 mM concentration of CMA, 100 mM KCl, 10 mM lithium cacodylate buffer (pH 7.0), mixing time of 300 ms and 25 °C in 90%/10% H₂O/²H₂O.

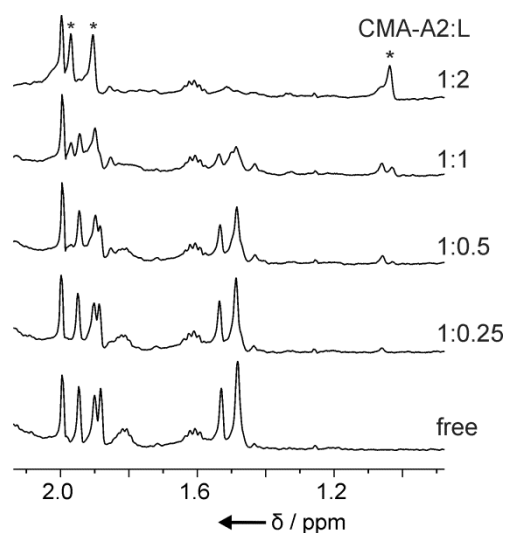


Figure S16. ^1H NMR spectra of methyl region of modified CMA G-quadruplex where A2 was replaced with propyl spacer C3 (CMA-A2) as **DAOTA-M2** (L) is titrated into solution. The molar ratio between CMA-A2 and the ligand is indicated on the right. With the star are marked the signals with similar fingerprint as was observed in methyl region of NMR spectrum of 1:2 CMA-ligand complex. NMR spectrum was recorded on 600 MHz spectrometer at 0.2 mM concentration of CMA, 100 mM KCl, 20 mM phosphate buffer (pH 7.0) and 25 °C in 90%/10% $\text{H}_2\text{O}/^2\text{H}_2\text{O}$.

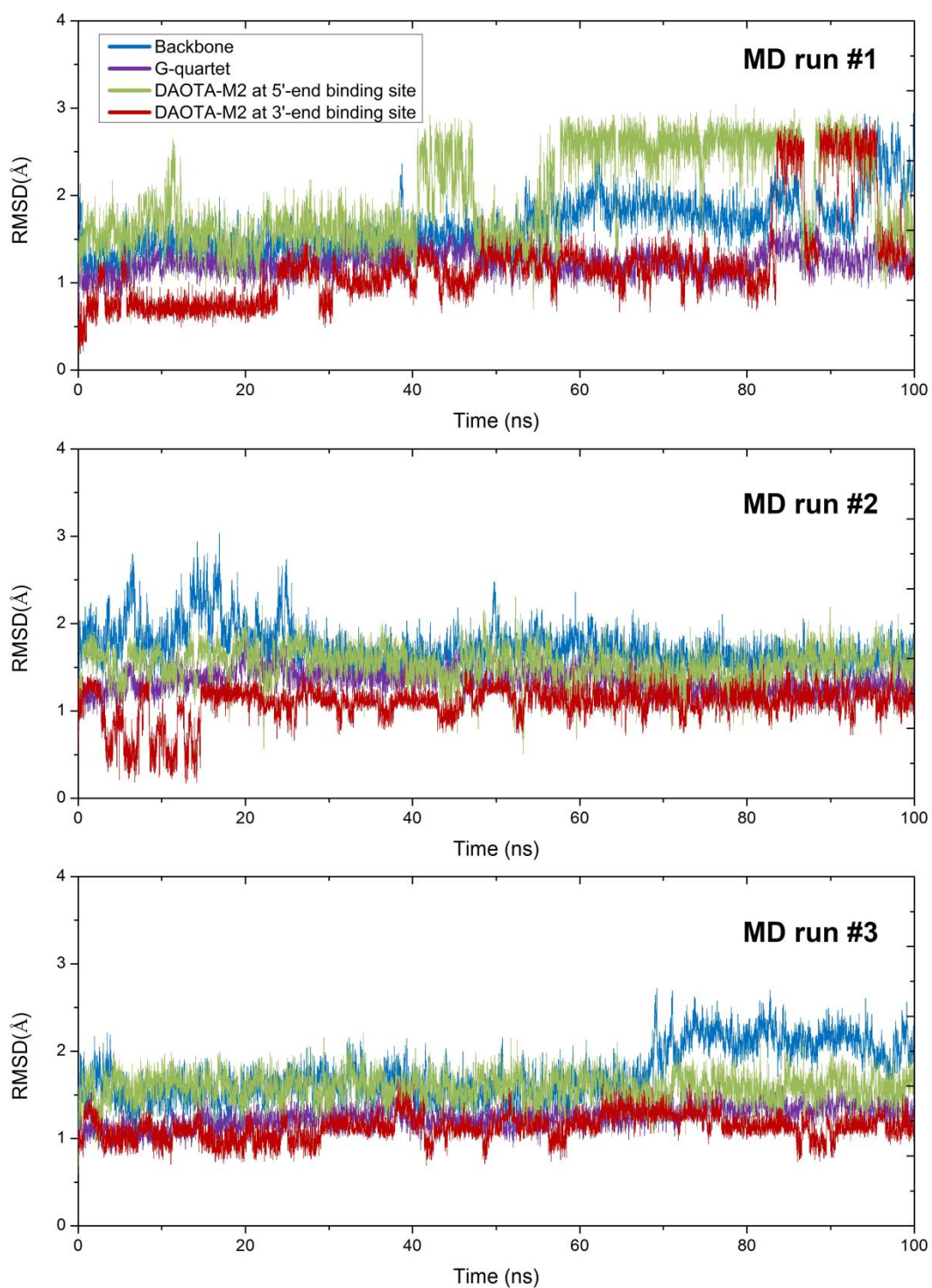


Figure S17. RMSD values of G-quadruplex backbone, G-quartet and **DAOTA-M2** molecules at the 5'- and 3'-end binding site vs time during three independent 100 ns MD simulations.

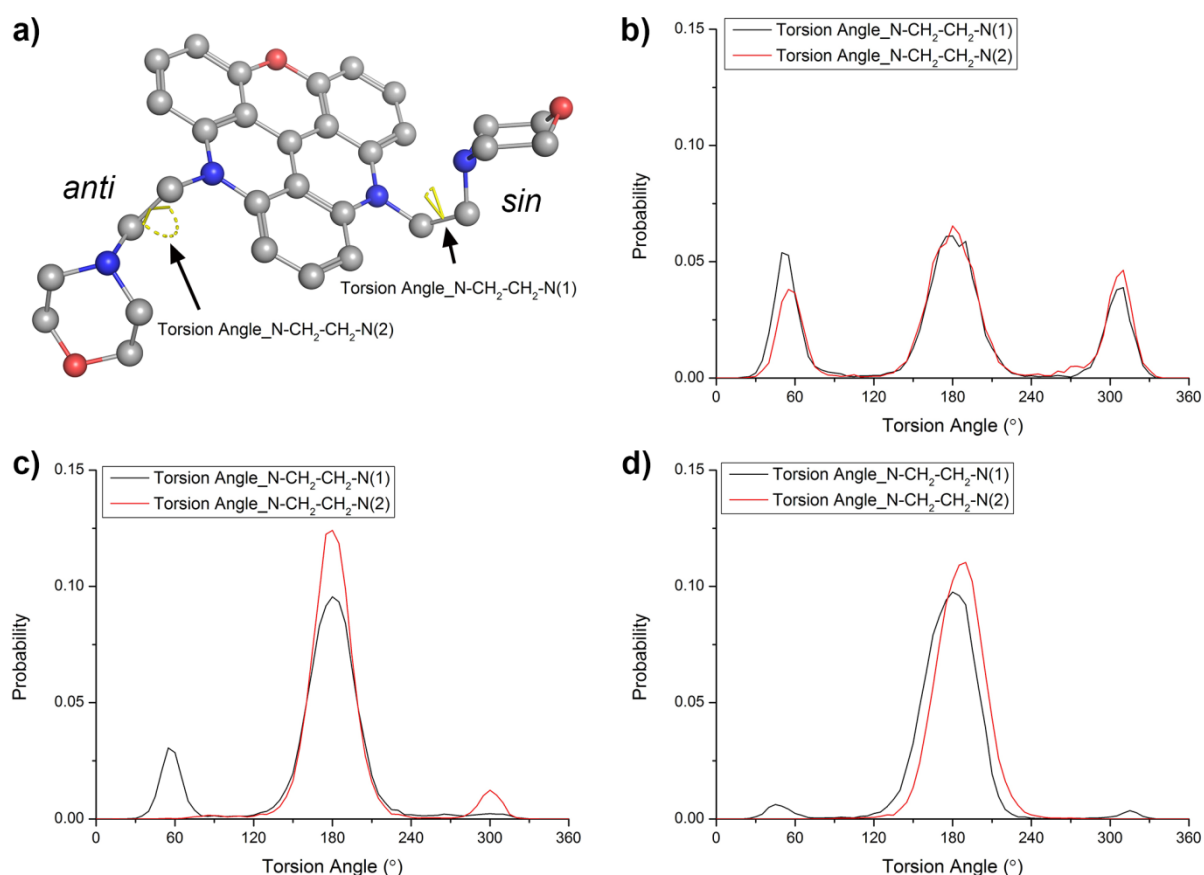


Figure S18. Probability distribution of torsion angles of ethylene linkers between morpholino groups and triangulenium core of **DAOTA-M2**. Data were calculated from MD simulation. a) Illustration of the torsion angles of ethylene linkers and its *syn* and *anti* conformations. Hydrogen atoms are removed for clarity. The probability distribution of the torsion angles of ethylene linkers from the MD simulation of b) free ligand, c) ligand bound to the 5'-G-quartet and d) ligand bound to the 3'-G-quartet.

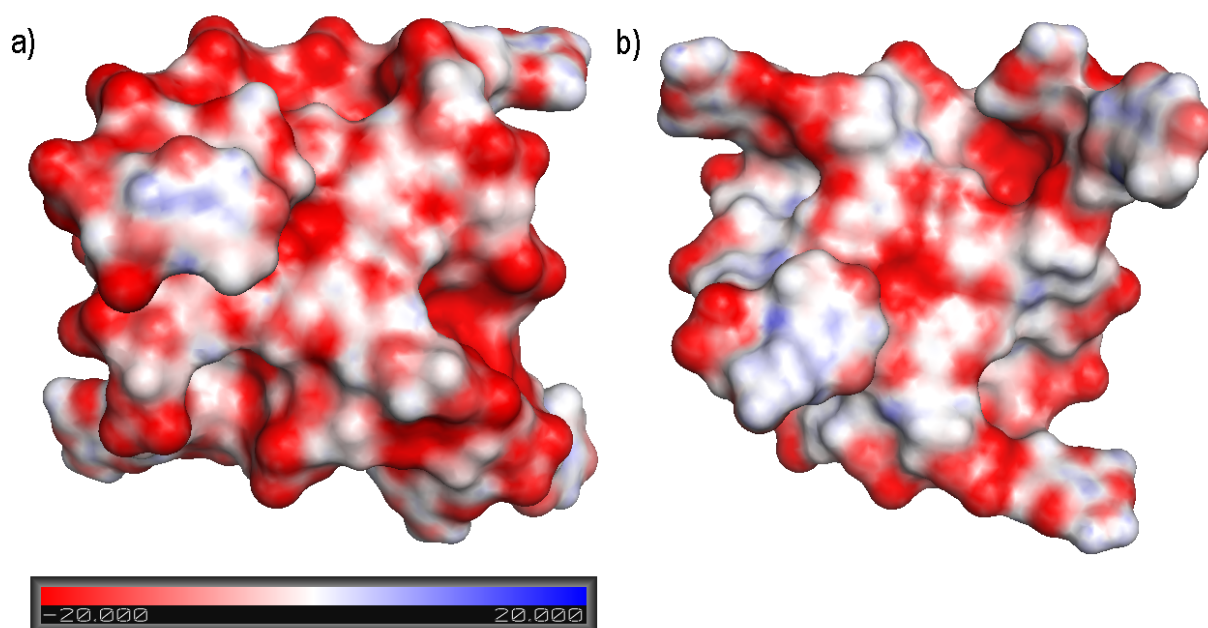


Figure S19. Electrostatic surface of CMA G-quadruplex calculated using Adaptive Poisson-Boltzmann Solver (APBS). The map was calculated using multiple Debye-Huckel boundary conditions, nonlinear Poisson-Boltzman equation with 40 points as surface density at 298 K. Scale is reported in dimensionless units. a) The 5'-end and b) the 3'-end of CMA G-quadruplex.

2. Tables

Table S1. ^1H NMR chemical shifts of 1:2 CMA-ligand complex.^[a]

Proton	H6/H8	H2/Me H1/Me	H1'	H2'/H2''	H3'
Residue					
T1	6.01	1.32	5.46	1.42/1.99	4.44
A2	8.32	/	6.34	2.85	5.12
G3	8.11	11.05	6.18	2.76/3.03	5.02
G4	7.70	10.95	6.18	2.64/3.00	5.05
G5	7.84	10.51	6.46	/	5.20
A6	8.56	8.31	6.66	/	5.23
G7	7.97	10.96	6.17	2.46/2.92	5.18
G8	7.78	11.11	6.08	2.63/2.86	5.07
G9	7.86	10.64	6.42	/	5.07
T10	7.71	1.97	6.23	2.29/2.49	4.77
A11	8.59	8.38	6.68	3.10/2.96	5.19
G12	8.13	11.16	6.19	2.59/2.97	5.06
G13	7.73	10.94	6.18	2.63/2.98	5.05
G14	7.85	10.40	6.46	/	5.20
A15	8.56	8.31	6.66	/	5.23
G16	7.96	10.92	6.15	/	5.20
G17	7.85	11.06	6.08	/	5.10
G18	7.81	10.59	6.39	/	5.11
T19	7.75	1.92	6.30	2.36/2.37	5.06

^[a] ^1H NMR chemical shifts are given in ppm are measured in 90% H_2O / 10% $^2\text{H}_2\text{O}$, 25 °C, 0.2 mM concentration of CMA oligonucleotide per strand, 0.4 mM concentration of **DAOTA-M2**, 100 mM KCl and 10 mM lithium cacodylate buffer (pH 7.0) and referenced to TMS.

Table S2. ^1H NMR chemical shifts changes induced by interactions between CMA G-quadruplex and **DAOTA-M2**.^[a]

Proton	$\delta(\text{Free CMA}) - \delta(1:2 \text{ CMA-ligand complex})$
T1 Me	0.38
A2 H8	-0.33
G3 H1	0.73
G4 H1	0.41
G5 H1	0.71
A6,15 H2	-0.01
A6,15 H8	-0.01
G7 H1	0.70
G8 H1	0.49
G9 H1	0.73
T10 Me	-0.01
A11 H2	-0.03
A11 H8	-0.14
G12 H1	0.90
G13 H1	0.41
G14 H1	0.67
G16 H1	0.44
G17 H1	0.39
G18 H1	0.68
T19 Me	-0.38

^[a] ^1H NMR chemical shifts are given in ppm are measured in 90% H_2O / 10% $^2\text{H}_2\text{O}$, 25 °C, 0.2 mM concentration of CMA oligonucleotide per strand, 0.4 mM concentration of **DAOTA-M2**, 100 mM KCl and 10 mM lithium cacodylate buffer (pH 7.0) and referenced to TMS.

Table S3. NMR restraints and structural statistics.

NMR distance and torsion angle restraints		
NOE-derived distance restraints	non-exchangeable	ex-changeable
Within G-quadruplex		
Intra-residue	143	0
Sequential ($i, i + 1$)	38	4
Long-range ($i, >i + 1$)	2	20
With DAOTA-M2 molecules	38	0
Hydrogen bond restraints	24	
Torsion angle restraints	57	
Structure statistics		
Violations		
Mean NOE restraint violation (Å)		0.119±0.030
Max. NOE restraint violation (Å)		0.186
Deviations from idealized geometry		
Bond length (Å)		0.011±0.000
Bond angle (°)		2.236±0.023
Pairwise heavy atom RMSD (Å)		
G-quadruplex with ligands		1.485±0.278
G-quadruplex		1.516±0.354
G-quartet core		0.847±0.145
G-quartet core with the 5'-terminal(T1 and A2)		0.953±0.185
G-quartet core with A6 and A15		1.429±0.339
G-quartet core with A10 and T11		0.979±0.187
G-quartet core with the 3'-terminal(T19)		0.834±0.139
G-quartet core with the 5'-terminal and ligands		1.097±0.211
G-quartet core with the 3'-terminal and ligands		0.878±0.139

Table S4. Clustering analysis of binding modes of **DAOTA-M2** at the 5'-end.

#Cluster	No. of structures	Percentage (%)	RMSD (Å) ^[a]	S.D. ^[b]
1	82647	34.4	2.846	0.891
2	63089	26.3	2.159	0.768
3	36501	15.2	2.564	0.814
4	22059	9.2	2.87	0.931
5	14338	6.0	1.762	0.507
6	6422	2.7	3.28	0.848
7	5575	2.3	2.837	0.862
8	3014	1.3	2.161	0.705
9	2482	1.0	2.632	0.94
10	2039	0.8	2.246	0.717
11	1834	0.8	3.042	0.986

^[a] Average distance between points in the cluster.^[b] Standard deviation of points in the cluster.**Table S5.** Clustering analysis of binding modes of **DAOTA-M2** at the 3'-end.

#Cluster	No. of structures	Percentage (%)	RMSD (Å) ^[a]	S.D. ^[b]
1	204508	85.2	2.879	1.017
2	29463	12.3	2.079	0.828
3	2192	0.9	2.334	0.658
4	1677	0.7	2.165	0.658
5	1339	0.6	1.954	0.559
6	821	0.3	2.859	0.825

^[a] Average distance between points in the cluster.^[b] Standard deviation of points in the cluster.

Material and methods

Sample preparation. The isotopically unlabeled and residue-specific low-enrichment (8% ^{15}N , ^{13}C -labelled) DNA oligonucleotide d[TAGGGAGGGTAGGGAGGGT] designed as CMA, was synthesized on K&A Laborgeraete GbR DNA/RNA Synthesizer H-8. In all cases, the standard phosphoramidite chemistry was used. Deprotection was done with the use of aqueous ammonia at 50 °C for 12 h. Samples were purified and desalted with the use of Amicon Ultra-15 Centrifugal Filter Units. The samples of CMA were prepared at 300 μL in 90%/10% $\text{H}_2\text{O}/^2\text{H}_2\text{O}$ in 10 mM lithium cacodylate or in 20 mM potassium phosphate buffer (pH 7.0) and in the presence of 100 mM KCl to give a 0.2 mM oligonucleotide concentration per strand. The samples containing 100% $^2\text{H}_2\text{O}$ were prepared as describe above, than lyophilized and subsequently dissolved in 300 μL 100% $^2\text{H}_2\text{O}$ solution. NMR samples with different ratios of oligonucleotide and **DAOTA-M2** were prepared by dissolving dry **DAOTA-M2** in 1 μL of DMSO- d_6 and then dissolve in aqueous solution of the CMA oligonucleotide or alternatively by titrating 13.6 mM **DAOTA-M2** solution in DMSO- d_6 into aqueous solution of the CMA oligonucleotide. The concentrations of DMSO- d_6 used in this study did not cause any spectral differences, which could account for structural changes of DNA influenced by DMSO- d_6 .^[10] **DAOTA-M2** was completely soluble in aqueous solutions at least up to 1.0 mM concentration. The fingerprint in NMR spectra of CMA G-quadruplex in lithium cacodylate and potassium phosphate buffers was the same indicating that buffer did not cause any changes in structure of G-quadruplex.

NMR spectroscopy. NMR data were collected on Agilent (Varian) NMR System 600 and 800 MHz spectrometers in the temperature range from 5 to 35 °C in 90%/10% $\text{H}_2\text{O}/^2\text{H}_2\text{O}$ or in 100% $^2\text{H}_2\text{O}$. ^1H and NOESY spectra were recorded with the use of the DPGSE solvent suppression method. NOESY spectra were acquired at mixing times in the rage between 40 and 300 ms. TOCSY spectra were acquired at mixing times between 20 and 40 ms in 100% $^2\text{H}_2\text{O}$ solution. ROESY spectra were recorded at mixing time of 40 ms. NMR spectra were processed and analyzed by using VNMRJ (Varian Inc.) and the Sparky (UCSF) software. Cross-peak assignment and integration was achieved using the software SPARKY (UCSF). NOE (Nuclear Overhauser Effect) distance restraints for protons were obtained from a 2D NOESY spectrum recorded with a mixing time of 300 ms. The average volume of H7-H6 cross-peaks of T1, T10 and T19 was used as the distance reference of 3.0 Å. Cross-peaks were classified as strong (1.8-3.6 Å), medium (2.6-5.0 Å) and weak (3.5-6.5 Å).

Circular Dichroism (CD) spectroscopy. CD experiments were carried out on CD spectrometer 62A DS (Aviv, ZDA) at 25 °C over the wavelength range between 220 and 320 nm. Measurements were made in the 0.1 cm path-length quartz cells. The experiments were done on samples at 40 μM oligonucleotide concentration per strand and 20 mM potassium phosphate buffer (pH 7.0) and at 0-160 μM concentration of **DAOTA-M2**.

UV melting experiments. Thermal stability measurements were performed on a Varian CARY-100 BIO UV-VIS spectrophotometer (Varian Inc.) equipped with a thermoelectric temperature controller. The experiments were done on samples at 40 μM CMA concentration and 20 mM potassium phosphate buffer (pH 7.0). Folding/unfolding processes were followed between 10 and 95 °C by measuring absorbance at 260 and 295 nm using scanning rate of 0.5 °C min^{-1} .

NMR structure calculation. The structures of 1:2 CMA-ligand complex were calculated by the simulated annealing (SA) simulations base on NOE-derived distance restraints. The partial charges of **DAOTA-M2** were generated by geometry optimization and electrostatistical potential (ESP) calculations with Gaussian 09^[1] at the level of HF/6-31G*, followed by RESP fitting^[2] via antechamber^[3] module of AmberTool 15.^[4] Other force field parameters were adopted from the Generalized Amber force field.^[5] SA simulations were performed using the CUDA version of pmemd module of AMBER 14 program suites^[4, 6] and Cornell *et al.* force field basic version parm99^[7] with the bsc0^[8], χOL4 ^[9], $\epsilon/\zeta\text{OL1}$ ^[10] and βOL1 ^[11] refinements. The initial extended single-stranded DNA structure was obtained using the leap module of AMBER 14. A total of 100 structures were calculated in 80 ps of NMR restrained simulated annealing (SA) simulations using the generalized Born implicit model.^[12] The cut-off for non-bonded interactions was 999 Å and the SHAKE algorithm^[13] for hydrogen atoms was used with the 0.4 fs time steps. For each SA simulation, a random velocity was

used. The SA simulation was as follows: in 0-2 ps, the temperature was raised from 300 K to 1000 K and held constant at 1000 K for 38 ps. Temperature was scaled down to 500 K in the next 24 ps and reduced to 100 K in the next 8 ps and was further reduced to 0 K in the last 8 ps. Restraints used in the calculation were NOE-derived distance restraints (force constant $20 \text{ kcal mol}^{-1} \text{ \AA}^{-2}$), hydrogen bond (force constant $40 \text{ kcal mol}^{-1} \text{ \AA}^{-2}$), torsion angle χ , ϵ (region $170^\circ - 290^\circ$) and v_2 restraints (force constant $200 \text{ kcal mol}^{-1} \text{ rad}^{-2}$). The torsion angle (χ , v_2 and ϵ) restraints were applied for all of residues based on the intensity of respective H8-H1' 2D NOESY cross-peaks (for χ , anti-region, $200^\circ - 280^\circ$), observation of high $J_{\text{H1'H2'}}$ coupling constants corresponding to predominance of S-type sugar puckering (for v_2 , region $-40^\circ - -20^\circ$) and J_{HP} coupling constants related to ϵ (region $170^\circ - 290^\circ$).^[14] NOE-derived distance restraints were used after the first 3000 steps. All structures were minimized with a maximum of 10 000 steps of energy minimization and a family of 10 structures was selected based on the smallest restraints violations and lowest energy.

Data deposition. The coordinates for structure of CMA G-quadruplex-**DAOTA-M2** complex have been deposited in the Protein Data Bank (accession codes : 5LIG) and Biological Magnetic Resonance Data Bank (accession codes : 34025).

Molecular dynamics simulation. The **DAOTA-M2** molecules and G-quadruplex complex were placed in a truncated octahedral box of TIP3P water molecules with the box border at least $10 \text{ \AA}$ away from any atoms of the G-quadruplex. Extra K^+ ions were added to neutralize the negative charges of the G-quadruplex. After the models were built, the ion positions were randomized using CPPTRAJ^[15] by swapping random water and ion positions such that no ion was closer than $4 \text{ \AA}$ to another and all ions were greater than $6 \text{ \AA}$ away from the DNA to avoid any biasing created by the initial placement of the ions.

The force field for MD simulation was the same as the SA simulation. The simulations were performed with the CUDA version of pmemd module of AMBER 14^[6, 16]. Prior to MD simulation, the system were subjected to a series of minimizations and equilibration. The equilibration protocol started with 500 steps of steepest descent minimization followed by 500 steps of conjugate gradient minimization with $25 \text{ kcal mol}^{-1} \text{ \AA}^{-2}$ position restraints on DNA and ligand atoms. Then the system was heated from 0 to 300 K during 100 ps with position restraints of $25 \text{ kcal mol}^{-1} \text{ \AA}^{-2}$ on G-quadruplex. Afterwards, the system underwent minimization with $5 \text{ kcal mol}^{-1} \text{ \AA}^{-2}$ restraints on DNA atoms using 500 steps of steepest descent method followed by 500 steps of conjugate gradient. Continuing in position restraints on DNA and ligand atoms of $5 \text{ kcal mol}^{-1} \text{ \AA}^{-2}$, the system was equilibrated during 50 ps at constant temperature of 300 K and pressure of 1 atm. Then an analogous series of alternating minimizations and equilibrations were performed, consecutively using decreasing position restraints of 4, 3, 2 and 1 $\text{kcal mol}^{-1} \text{ \AA}^{-2}$. Finally, an equilibration using position restraints of $0.5 \text{ kcal mol}^{-1} \text{ \AA}^{-2}$ and starting velocities from the previous equilibration, followed by a short 50 ps molecular dynamics took place without any restraints. Pressure coupling used during equilibration was set to 0.2, coupling during the last molecular dynamics phase was set to 5.

The production simulation were carried out at constant pressure of 1 atm and constant temperature of 300 K maintained using Langevin dynamics with a collision frequency of 2.0, respectively. Periodic boundary conditions were used, and electrostatic interactions were calculated by the particle mesh Ewald method^[17] with the non-bonded cutoff set to $9 \text{ \AA}$. The SHAKE algorithm^[13] was applied to bonds involving hydrogens, and a 2 fs integration step was used. The production run was carried out for continuous 100 ns and the snapshots were written at every 1 ps. Three MD simulations were performed with different initial velocity distributions. Trajectories were analyzed using the CPPTRAJ module of AMBER. Cluster analysis was performed on the **DAOTA-M2** molecules binding at the 5'- and 3'-G-quartet with the 5'- and 3'-termini residues using the average-linkage hierarchical agglomerative method.^[18] Coordinate RMSD was used as the distance metric. The critical distance ϵ value was set to $40 \text{ \AA}$ and the sieve value was set to 25. Only the non-hydrogen atoms were used in the clustering. MD simulations of free **DAOTA-M2** were also performed with the same protocol and parameters as described above.

4. References

- [1] G. W. T. M.J. Frisch, H.B. Schlegel, G.E. Scuseria, M.A. Robb, J.R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G.A. Petersson, H. Nakatsuji, M. Caricato, X. Li, H.P. Hratchian, A.F. Izmaylov, J. Bloino, G. Zheng, J.L. Sonnenberg, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, J.A. Montgomery Jr., J.E. Peralta, F. Ogliaro, M. Bearpark, J.J. Heyd, E. Brothers, K.N. Kudin, V.N. Staroverov, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J.C. Burant, S.S. Iyengar, J. Tomasi, M. Cossi, N. Rega, J.M. Millam, M. Klene, J.E. Knox, J.B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R.E. Stratmann, O. Yazyev, A.J. Austin, R. Cammi, C. Pomelli, J.W. Ochterski, R.L. Martin, K. Morokuma, V.G. Zakrzewski, G.A. Voth, P. Salvador, J.J. Dannenberg, S. Dapprich, A.D. Daniels, Ö. Farkas, J.B. Foresman, J.V. Ortiz, J. Cioslowski, D.J. Fox., *Gaussian 09 (Revision A. 02)*. Gaussian, Inc., Wallingford, CT (2009).
- [2] C. I. Bayly, P. Cieplak, W. D. Cornell, P. A. Kollman, *J. Phys. Chem.* **1993**, 97, 10269-10280.
- [3] J. M. Wang, W. Wang, P. A. Kollman, *Abstr. Pap. Am. Chem. S.* **2001**, 222, U403-U403.
- [4] J. T. B. D.A. Case, R.M. Betz, D.S. Cerutti, T.E. Cheatham, III, T.A. Darden, R.E. Duke, T.J. Giese, H. Gohlke, A.W. Goetz, N. Homeyer, S. Izadi, P. Janowski, J. Kaus, A. Kovalenko, T.S. Lee, S. LeGrand, P. Li, T. Luchko, R. Luo, B. Madej, K.M. Merz, G. Monard, P. Needham, H. Nguyen, H.T. Nguyen, I. Omelyan, A. Onufriev, D.R. Roe, A. Roitberg, R. Salomon-Ferrer, C.L. Simmerling, W. Smith, J. Swails, R.C. Walker, J. Wang, R.M. Wolf, X. Wu, D.M. York and P.A. Kollman, *AMBER 2015, University of California, San Francisco*. **2015**.
- [5] J. M. Wang, R. M. Wolf, J. W. Caldwell, P. A. Kollman, D. A. Case, *J. Comput. Chem.* **2004**, 25, 1157-1174.
- [6] R. Salomon-Ferrer, A. W. Goetz, D. Poole, S. Le Grand, R. C. Walker, *J. Chem. Theory Comput.* **2013**, 9, 3878-3888.
- [7] W. D. Cornell, P. Cieplak, C. I. Bayly, I. R. Gould, K. M. Merz, D. M. Ferguson, D. C. Spellmeyer, T. Fox, J. W. Caldwell, P. A. Kollman, *J. Am. Chem. Soc.* **1996**, 118, 2309-2309.
- [8] A. Perez, I. Marchan, D. Svozil, J. Sponer, T. E. Cheatham, III, C. A. Loughton, M. Orozco, *Biophys. J* **2007**, 92, 3817-3829.
- [9] M. Krepl, M. Zgarbova, P. Stadlbauer, M. Otyepka, P. Banas, J. Koca, T. E. Cheatham, 3rd, P. Jurecka, J. Sponer, *J. Chem. Theory Comput.* **2012**, 8, 2506-2520.
- [10] M. Zgarbova, F. Javier Luque, J. Sponer, T. E. Cheatham, III, M. Otyepka, P. Jurecka, *J. Chem. Theory Comput.* **2013**, 9, 2339-2354.
- [11] M. Zgarbova, J. Sponer, M. Otyepka, T. E. Cheatham, III, R. Galindo-Murillo, P. Jurecka, *J. Chem. Theory Comput.* **2015**, 11, 5723-5736.
- [12] aA. Onufriev, D. Bashford, D. A. Case, *J. Phys. Chem. B* **2000**, 104, 3712-3720; bA. Onufriev, D. Bashford, D. A. Case, *Proteins* **2004**, 55, 383-394.
- [13] J. P. Ryckaert, G. Ciccotti, H. J. C. Berendsen, *J. Comput. Phys.* **1977**, 23, 327-341.
- [14] Neidle, S., *Principles of Nucleic Acid Structure*; Academic Press: London, 2010.
- [15] D. R. Roe, T. E. Cheatham, III, *J. Chem. Theory Comput.* **2013**, 9, 3084-3095.
- [16] aY. Ding, A. M. Fleming, L. He, C. J. Burrows, *J. Am. Chem. Soc.* **2015**, 137, 9053-9060; bR. Stefl, F. C. Oberstrass, J. L. Hood, M. Jourdan, M. Zimmermann, L. Skrisovska, C. Maris, L. Peng, C. Hofr, R. B. Emeson, F. H. T. Allain, *Cell* **2010**, 143, 225-237; cA. T. Phan, V. Kuryavyi, H. Y. Gaw, D. J. Patel, *Nat. Chem. Biol.* **2005**, 1, 167-173.
- [17] aT. Darden, D. York, L. Pedersen, *J. Chem. Phys.* **1993**, 98, 10089-10092; bU. Essmann, L. Perera, M. L. Berkowitz, T. Darden, H. Lee, L. G. Pedersen, *J. Chem. Phys.* **1995**, 103, 8577-8593.
- [18] J. Shao, S. W. Tanner, N. Thompson, T. E. Cheatham, III, *J. Chem. Theory Comput.* **2007**, 3, 2312-2334.