#### 8-amino guanine accelerates tetramolecular quadruplex formation

Julien Gros, Anna Aviñó, Jaime Lopez de la Osa, Carlos González, Laurent Lacroix, Alberto Pérez, Modesto Orozco, Ramon Eritja & Jean-Louis Mergny

## **Supplementary Information**

#### **METHODS**

#### Circular Dichroism:

CD spectra were recorded with a Jasco-810 spectropolarimeter using 1 cm pathlength quartz cuvettes in a volume of 350  $\mu$ L. Scans were performed at 4°C as previously described <sup>[1]</sup>. The CD spectra represented an average of three scans without any zero correction. Samples were preincubated at high strand concentration to allow quadruplex formation (300  $\mu$ M). Oligonucleotides were briefly heated at 90°C then immediately added to 10 mM lithium cacodylate pH 7.2, 110 mM Na<sup>+</sup>, K<sup>+</sup> or NH<sub>4</sub><sup>+</sup> buffer, and finally incubated at 4°C during 48 hours.

#### UV-isothermal association and UV-melting:

As previously described <sup>[1]</sup>, we started from isolated strands and compared association of both modified and unmodifed sequences by UV-isothermal experiments in K<sup>+</sup>, Na<sup>+</sup> or NH<sub>4</sub><sup>+</sup> conditions. To follow quadruplex formation or dissociation, one may simply record the absorbance at convenient wavelengths such as 240 <sup>[1]</sup> or 295 nm <sup>[2]</sup>. Unfortunately, 295 nm is not as convenient for sequences bearing an 8-amino guanine (**R**) for which the variation of absorbance is smaller. To study the impact of (**R**) at a given position, one could fit data using a previously published model developped in <sup>[1]</sup> and calculate association rate constants (*k*<sub>on</sub>). The corresponding order of the association reaction has been fixed at *n* = 4 as in previous studies <sup>[3,4]</sup>.

UV-melting experiments of preformed quadruplexes were recorded as previously described <sup>[1]</sup>. Association and dissociation of G4-DNA are both slow processes, leading a strong hysteresis, meaning that melting profiles do not correspond to a thermodynamic equilibrium <sup>[5]</sup>. A major consequence is that temperature of half-dissociation measured on dissociation curve (" $T_{1/2}$ ") increase when increasing the temperature gradient <sup>[1]</sup>. All values listed in *Table 1* have been obtained at 0.5°C/min.

#### Isothermal Differencial Spectra (IDS):

To construct an IDS spectrum, one simply records absorbance of the sample just before (single-stranded form) and just after (G4-DNA form) association. Each type of nucleic acid structure corresponds to a different TDS shape, meaning that this simple analysis provides a useful indication on the nature of the folded structure <sup>[6]</sup>. For G4-DNA, Isothermal Differencial Spectra (IDS) should be more relevant: in some cases, the thermal stability of the structure is so high that denaturation at high temperature is impossible.

#### Nuclear Magnetic Resonance (NMR):

Samples were suspended in either  $D_2O$  or 9:1  $H_2O/D_2O$  (5mM potassium phosphate buffer pH 7). The oligonucleotide concentration for the 2D experiments was 300  $\mu$ M. <sup>1</sup>H NMR spectra were acquired in Bruker spectrometers operating either at 600 or at 800 MHz, and equipped with

cryoprobes. 1D experiments were acquired at different temperatures, ranging from 5 to 70 °C. NOESY <sup>[7]</sup> spectra were acquired in D<sub>2</sub>O and in H<sub>2</sub>O with mixing times of 100, and 250 ms. TOCSY <sup>[8]</sup> spectra were recorded with the standard MLEV-17 spin-lock sequence and a mixing time of 80ms. In H<sub>2</sub>O experiments, water suppression was achieved by including a WATERGATE <sup>[9]</sup> module in the pulse sequence prior to acquisition.

#### Theoretical calculations:

Molecular Dynamics simulations: Starting model of parallel tetraplex [d(GGGG)]<sub>4</sub> was created using crystal structure <sup>[10]</sup> and then solvated by a cubic box containing 2665 water molecules. Sodium ions inside the G4-DNA channel were placed in crystallographic positions, while extra ions needed to achieve electroneutrality were placed at optimal positions according to Poisson-Boltzman calculations. Neutral hydrated system was then subject our standard heating and equilibration procedure <sup>[11]</sup> and to 15 ns of unrestrained MD simulation at constant pressure (1 atm) and temperature (300 K) using periodic boundary conditions and Particle Mesh Ewald<sup>[12]</sup>. The snapshot obtained at nanosecond 15<sup>th</sup> was used to create models of the tetraplex containing one 8-amino guanine substitution placed at every step of the tetraplex. After equilibration, these additional tetraplexes were analyzed for 2 ns MD simulation (at the same pressure and temperature). Structures obtained at the 1st to 2nd nanosecond of each trajectory were used to generate models of tetraplexes where all four G in one step have been substituted to  $\underline{\mathbf{R}}$ . The resulting tetraplexes were re-equilibrated as noted above and subject to 2 additional ns of MD simulation (at the same pressure and temperature). PARMBSC0<sup>[13]</sup> and TIP3P<sup>[14]</sup> were used to describe force-field parameters of DNA and water, while previously developed AMBER-like parameters were used to describe  $\mathbf{R}^{[15]}$ . Newton's equations of motion were integrated every 2 fs, using SHAKE <sup>[16]</sup> to keep constant all bonds involving hydrogens. Calculations were performed using AMBER suite of programs<sup>[17]</sup>.

**Free energy calculations:** Thermodynamic integration coupled to molecular dynamics (MD/TI) <sup>[17]</sup> was used to determine the free energy change associated to the change of one or four G in the 2<sup>nd</sup> step of the [d(GGGG)]<sub>4</sub> tetraplex. This was determined as the difference in the reversible work needed to change one (or four) G by one (or four) <u>R</u> in a tetraplex and in a single stranded oligonucleotide (previously equilibrated for 2 ns). Calculations were performed in both "forward" (G to <u>R</u>) and reverse (<u>R</u> to G) directions. Mutations were done along 1, 2 and 4 ns trajectories using 21 or 41 windows divided in two halves. This complex procedure allowed us to have 12 independent estimates of every free energy change, which were then averaged to obtain values shown in the paper. Additional details of the MD/TI procedure can be found in our previous publication <sup>[18]</sup>.

**Quantum mechanical calculations:** The difference in the *syn/anti* equilibrium between guanosine and 8-amino guanosine was explored using quantum mechanical calculations simulating an aqueous solutions. Models of the two nucleosides in standard *syn* and *anti* conformations were relaxed by geometry optimization at the B3LYP/6-31G(d) level <sup>[19-20]</sup>. The effect of water on the equilibrium was simulated by using the latest version of our water-parametrized MST method <sup>[21]</sup>. The effect of the change from G to <u>R</u> in the *syn/anti* equilibrium was computed as (syn – anti) free energy difference. All QM calculations were performed using a local version of Gaussian 2003 <sup>[22]</sup> which includes the MST-routines. Temperature of 298 K was considered for entropic and thermal corrections to the energy as well as for solvation calculations.

#### Native Polyacrylamide gel electrophoresis (PAGE).

Non-denaturing gel electrophoresis allows separation of single-stranded oligonucleotides from tetramolecular G-quadruplex structures. Samples were loaded on a 20% polyacrylamide (acrylamide/bis-acrylamide 19:1) gel containing TBE 1X and the corresponding salt at 20 mM. Just before loading, samples "-" were briefly heated at 90°C and immediately after put on ice. Electrophoresis was performed at 3W in a cold room. Bands were revealed by UV-shadowing using a UV light source (254 nm) and a digital camera. Rather than using poly-dT as size markers<sup>[23]</sup>, we compared migrations of both unmodified and modified sequences with or without cations. One may notice that G4-DNA formation is not fully achieved in all cases, suggesting that association in Na<sup>+</sup> 110mM of [d(T<u>R</u>GGGT)]<sub>4</sub> would be at least as fast as [d(TG<sub>4</sub>T)]<sub>4</sub>, and faster than [d(TGGG<u>R</u>T)]<sub>4</sub>.

#### SUPPLEMENTARY REFERENCES:

- [1] J. L. Mergny, A. De Cian, A. Ghelab, B. Saccà, L. Lacroix, *Nucleic Acids Res.* 2005, 33, 81.
- [2] J. L. Mergny, A. T. Phan, L. Lacroix, *FEBS Lett.* **1998**, *435*, 74.
- [3] J. R. Wyatt, P. W. Davis, S. M. Freier, *Biochemistry* **1996**, *35*, 8002.
- [4] R. Stefl, T. E. Cheatham, N. Spackova, E. Fadrna, I. Berger, J. Koca, J. Sponer, *Biophys. J.* 2003, *85*, 1787.
- [5] J. L. Mergny, L. Lacroix, *Oligonucleotides* **2003**, *13*, 515.
- [6] J. L. Mergny, J. Li, L. Lacroix, S. Amrane, J. B. Chaires, *Nucleic Acids Res.* 2005, 33, e138.
- [7] A. Kumar, R. R. Ernst, K. Wuthrich, *Biochem. Biophys. Res. Commun.* **1980**, *95*, 1.
- [8] A. Bax, D. J. Davies, J. Magn. Reson. 1985, 65, 355.
- [9] M. Piotto, V. Saudek, V. Sklenar, *Journal of Biomolecular NMR* 1992, 2, 661.
- [10] G. Laughlan, A. I. H. Murchie, D. G. Norman, M. H. Moore, P. C. E. Moody, D. M. J. Lilley, B. Luisi, *Science* 1994, 265, 520.
- [11] G. C. Shields, C. A. Laughton, M. Orozco, Journal of the American Chemical Society 1997, 119, 7463.
- [12] T. Darden, D. York, L. Pedersen, Journal of Chemical Physics 1993, 98, 10089.
- [13] A. Perez, I. Marchan, D. Svozil, J. Sponer, T. E. Cheatham, 3rd, C. A. Laughton, M. Orozco, *Biophys J* 2007, 92, 3817.
- [14] W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey, M. L. Klein, Journal of Chemical Physics 1983, 79, 926
- [15] A. Avino, J. C. Morales, M. Frieden, B. G. de la Torre, R. G. Garcia, E. Cubero, F. J. Luque, M. Orozco, F. Azorin, R. Eritja, *Bioorganic & Medicinal Chemistry Letters* **2001**, *11*, 1761.
- [16] J. P. Ryckaert, G. Ciccotti, H. J. C. Berendsen, Journal of Computational Physics 1977, 23, 327
- [17] D. A. Case, T. E. Cheatham, T. Darden, H. Gohlke, R. Luo, K. M. Merz, A. Onufriev, C. Simmerling, B. Wang, R. J. Woods, *Journal of Computational Chemistry* 2005, 26, 1668.
- [18] J. P. de la Osa, C. Gonzalez, R. Gargallo, M. Rueda, E. Cubero, M. Orozco, A. Avino, R. Eritja, *Chembiochem* **2006**, 7, 46
- [19] A. D. Becke, *Journal of Chemical Physics* **1993**, *98*, 5648.
- [20] C. T. Lee, W. T. Yang, R. G. Parr, *Physical Review B* 1988, 37, 785.
- [21] I. Soteras, C. Curutchet, A. Bidon-Chanal, M. Orozco, F. J. Luque, *Journal of Molecular Structure-Theochem* 2005, 727, 29.

[22] Gaussian 03, Revision C.02, M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, J. A. Montgomery, Jr., T. Vreven, K. N. Kudin, J. C. Burant, J. M. Millam, S. S. Iyengar, J. Tomasi, V. Barone, B. Mennucci, M. Cossi, G. Scalmani, N. Rega, G. A. Petersson, H. Nakatsuji, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, M. Klene, X. Li, J. E. Knox, H. P. Hratchian, 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, P. Y. Ayala, K. Morokuma, G. A. Voth, P. Salvador, J. J. Dannenberg, V. G. Zakrzewski, S. Dapprich, A. D. Daniels, M. C. Strain, O. Farkas, D. K. Malick, A. D. Rabuck, K. Raghavachari, J. B. Foresman, J. V. Ortiz, Q. Cui, A. G. Baboul, S. Clifford, J. Cioslowski, B. B. Stefanov, G. Liu, A. Liashenko, P. Piskorz, I. Komaromi, R. L. Martin, D. J. Fox, T. Keith, M. A. Al-Laham, C. Y. Peng, A. Nanayakkara, M. Challacombe, P. M. W. Gill, B. Johnson, W. Chen, M. W. Wong, C. Gonzalez, and J. A. Pople, Gaussian, Inc., Wallingford CT, 2004.

[23] I. Kejnovska, J. Kypr, M. Vorlickova, Biochem. Biophys. Res. Commun. 2007, 353, 776.

Tetraplex	Ion-bases	Stacking	H-bond	Total
1R	-8.6	+1.3	+1.7	-5.6
4R	-27.6	+5.1	+6.7	-15.8

**Sup. Table S1**. Difference in different interactions energies between  $2^{nd}$  step of the  $d(GGGG)_4$  tetraplex and those containing 1 (R) and 4 (4R) 8 amino guanines and that step.

Values include: ion-nucleobase interaction energy, stacking (both intra and inter-strand) and hydrogen-bond. Values are in kcal/mol and a negative sign means that the substitution stabilizes the structure.

**Sup. Figure S1:** Normalized Isothermal Differential Spectra of  $[d(T\underline{R}GGGGT)]_4$ ,  $[d(TGGGG\underline{R}T)]_4$ ,  $[d(TGGGG\underline{R}T)]_4$ ,  $[d(TGGG\underline{R}T)]_4$ ,  $[d(TG_5T)]_4$  and  $[d(TG_4T)]_4$  in KCl.

**Sup. Figure S2:** <sup>1</sup>H-NMR spectra of  $[d(TG_5T)]_4$ ,  $[d(T\underline{R}GGGGT)]_4$ ,  $[d(TGGGG\underline{R}T)]_4$ , and  $[d(TGG\underline{R}GGT)]_4$  at different temperatures. Experiments are carried out in H<sub>2</sub>O/D<sub>2</sub>O 9:1 pH 7 buffer in 10 mM K<sup>+</sup> phosphate and 5 mM KCl. [oligo] = 300  $\mu$ M.

**Sup. Figure S3:** <sup>1</sup>H-NMR spectra of  $[d(TG_4T)]_4$ ,  $[d(T\underline{R}GGGT)]_4$ ,  $[d(TGGG\underline{R}T)]_4$ , and  $[d(TGG\underline{R}GGT)]_4$  at different temperatures. Experiments are carried out in H<sub>2</sub>O/D<sub>2</sub>O 9:1 pH 7 buffer in 10 mM K<sup>+</sup> phosphate and 5 mM KCl. [oligo] = 300  $\mu$ M.

Sup. Figure S4: Aromatic- (H2', H2") region of the NOESY of  $[d(TG_5T)]_4$ , and  $[d(TGGGG\underline{R}T)]_4$ . 250ms mixing time in H<sub>2</sub>O/D<sub>2</sub>O 9:1 pH 7, buffer 10 mM K<sup>+</sup> phosphate and 5 mM KCl .[oligo] = 300  $\mu$ M.

**Sup. Figure S5:** Superposition of tetraplex structures containing none (green), one (blue) or four (yellow)  $\underline{\mathbf{R}}$  at the second position. Details of the tetrads are given in the lower part of the slide.

Sup. Figure S6: Example of an association experiment.

 $d(TG_4\underline{R}T)$  in K<sup>+</sup> conditions: 50 mM KCl supplemented with LiCl 60 mM to keep ionic strength constant (110 mM) in all experiments. Single-strand to quadruplex transition is characterised by a time-dependent increase in absorbance at 295nm while an opposite trend is observed at 240 nm. Datas are well fited with the model used (plain lines under experimental points).

**Sup. Figure S7:** Representation of the dipole direction (located over the centre of masses) for G (bottom) and <u>R</u> (top). The better orientation along the C=O bond in the later case is evident.



 $[d(TGGGGGT)]_4$ 

T=5C 11.5 ppi 11. 10. 8.5 M T=15C 11.0 8.5 11.5 10.5 10.0 9.5 9.0 ppi Ш T=25C 11.0 8.5 11.5 10. pp T=35C 11.0 10.0 9.5 8.5 10 5 9.0 8.0 . ppi T=45C 11.5 9.0 8.5 11. 10. 9.5 ppi 8.0 T=55C 11.0 8.5 11.5 10.5 10.0 9.5 9.0 l T=65C 11.5 11.0 10.0 9.5 9.0 ppi U T=75C 11.5 11.0 8.5 8.0 7.5 10.5 10.0 9.5 9.0 ppi

T=15C -hu 12.0 11.5 11.0 10.5 9.5 T=25C J 12.0 11.5 11.0 10.5 10.0 9.5 9.0 8.1 T=35C 12.0 11.5 11.0 T=45C wM), ----Ar 12.0 11.5 11.0 10.5 10.0 9.5 9.0 8.5 8.0 7.5 M T=55C 12.0 10.5 11.5 11.0 9 5 10.0 9.0 8.0  ${\rm str}^{\rm s}$ т=65С

8.5



# $[d(TGGRGGT)]_4$



## $[d(TGGGGRT)]_4$



Sup. Figure S2

## Supplementary Material (ESI) for Chemical Communications This journal is (c) The Royal Society of Chemistry 2008 T=5C

12.0 11.5



Sup. Figure S3

Supplementary Material (ESI) for Chemical Communications This journal is (c) The Royal Society of Chemistry 2008

#### $[D(TGGGGGT)]_4$



#### $[D(TGGGGRT)]_4$







