Controlling G-quadruplex formation via lipid modification of oligonucleotide sequences

Brune Vialet, Arnaud Gissot, Romain Delzor, Philippe Barthélémy

Chem Comm

COMMUNICATION

Controlling G-quadruplex formation *via* lipid modification of oligonucleotide sequences[†]

Brune Vialet,[‡] Arnaud Gissot,[‡] Romain Delzor and Philippe Barthélémy*

G-quadruplexes (G4) represent attractive supramolecular scaffolds. In this communication, we show that the lipid modification of a G4 prone oligonucleotide sequence drastically increases the probability of forming tetramolecular parallel G4s with unprecedented conformational control over other unspecific oligomers or folds.

Guanine-rich oligonucleotides can self-assemble into four-stranded G4 (see SI for the list of acronyms) structures stabilized by π-π stacking between G-quartets and *via* Hoogsteen hydrogen bonding.[1] They are found for instance in telomeric and promoter regions, where they participate in a diversity of biological processes. Due to their biological and biophysical properties, G4s have found applications in e.g. supramolecular chemistry,[2,3] nanotechnology[4,5] or medicinal chemistry.[6,7] G4s can fold with different strand stoichiometries (1-4 strands), orientations (parallel, antiparallel, hybrid), and different number of G-quartets.[8] This conformational diversity also depends on strand concentrations, temperature, ligands and other factors. Cations also have a strong impact on both G4 folding and stability since they are coordinated to the internal carbonyl oxygen atoms of the G-quartet bases.[9]

Although naturally occurring G4s are mostly unimolecular, supramolecular chemistry has a greater interest in assemblies composed of two or more interacting species. Tetramolecular G4 appear as interesting templates since they allow for the assembly of four strands in a predictable and controlled manner. Of note, all the tetramolecular G4 structures reported so far are parallel: the four strands of the G4 proceed in the same direction. Nevertheless, tetramolecular parallel G4-prone sequences may still adopt competing flipped or mismatched geometries as well as higher order aggregates, thus limiting their utility as supramolecular scaffolds.[10,11] These undesired competing structures increase with the oligonucleotide length.[12] In addition, due to the unfavourable entropy associated with the formation of quaternary complexes,[10,13,14]

tetramolecular parallel G4s exhibit extremely slow kinetics of formation. Overall, this limits the attractiveness of tetramolecular parallel G4s as supramolecular scaffolds. Lipid-modification has been used for the hydrophobic effect-driven stabilization of tetramolecular parallel G4 assemblies.[15–17] The isolated amphiphilic tetramolecular parallel G4s are yet not observed in solution as they self-assemble into stable micellar aggregates, with no control over the correct tetramolecular parallel G4 folding. Hybrid Lipid-OligoNucleotides (LONs)[18–20] have also found applications as therapeutics[21–24] and others.[25,26]

In this communication, we show that provided the correct lipid modification is used, lipidic tetramolecular parallel G4s can be formed and manipulated to control both their folding and stability (Fig. 1). As a result, it is now possible to obtain tetramolecular parallel G4s of unprecedented length with conformational control, improved speed of formation and tuneable stability of the micellar assemblies. The oligonucleotide sequence used was chosen to contain a G-tract of 4 consecutive guanines in the middle of a 19-mer DNA sequence (Fig. S1 (ESI†)). The DNA molecule was modified at the 5'-end with different lipid phosphoramidites (Fig. 2). The phosphoramidites 1[27] and 2[28] were synthesized according to literature procedures and coupled to the 5'-end of the DNA. The randomized LON oligonucleotide sequences ('scramble') were also synthesized as controls wherein the guanines were evenly distributed along the sequence to minimize the chances of forming undesired secondary structures. When the lipid is present at the 5'-end of the oligonucleotide, the capping step during the oligonucleotide synthesis prevents abortive sequences from reacting with the lipid phosphoramidite. Provided the LON do not form stable aggregates (vide infra), their purification is thus straightforward as only the desired full-length oligonucleotide is coupled to the lipid and somewhat interacts with the reverse stationary phase. k-LON[G4] that forms stable micelles required in contrast an original purification protocol (Fig. S2 and S3 (ESI†)).

Univ. Bordeaux, ARNA laboratory, F-33000 Bordeaux, France.
INSERM, U1212, ARNA laboratory, F-33000 Bordeaux, France.
Fax: (+33) 05 57 57 10 15
E-mail: philippe.barthelemy@inserm.fr.
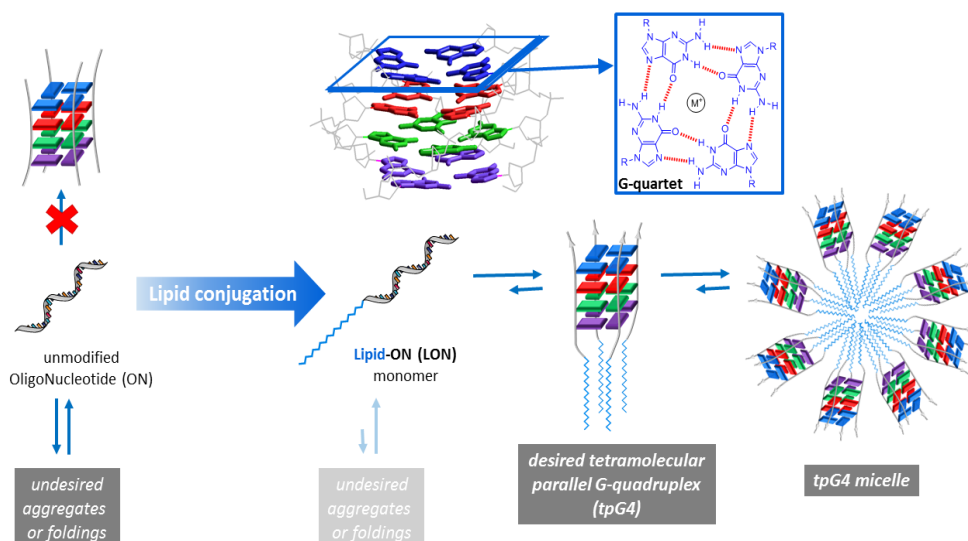‡ These two authors contributed equally to the work

COMMUNICATION

Fig 1. Lipid-driven assembly of large tetramolecular parallel G4s (the different colors for the 4 G-quartets is an aid to the eye, each quartet being structurally identical).

The nature of the supramolecular assemblies formed from these different LONs was first investigated by non-denaturing PAGE. Interestingly, both the nucleotide sequences and the nature of the lipid were shown to impact the self-assemblies (Fig. 3). In line with the other lipidic tetramolecular parallel G4s,[15,16] the di-alkylated **k-LON^{G4}** formed micelles large and more importantly stable enough to survive the sieving conditions of the PAGE matrix during electrophoresis (lane 7). The less lipophilic **1-LON^{G4}** instead seemed to partition between monomers and a more retarded tetramolecular parallel G4 band (lane 5) as confirmed by i) the CD, ii) the high K⁺-dependence, iii) the retardation in the gel and iv) ¹H NMR of **1-LON^{G4}** solutions (Fig. S4 (ESI†)). No G4 bands was observed with the unmodified **ON^{G4}** unless the salt concentration was increased (lane 3). As expected, no folds were observed for the different scramble (L)ONs (lanes 1, 4, 6). A kinetic control over an alleged equilibrium between the G4s and the monomers was first ruled out as the ratio monomer/G4 for **1-LON^{G4}** or **ON^{G4}** are clearly time independent (up to 4 weeks, PAGE or CD).

2 mins (CD)), with the remaining **1-LON^{G4}** molecules trapped in other folding pathways (vide infra). Nevertheless, **1-LON^{G4}** and **k-LON^{SC}** do form micelles as judged by agarose gel, DLS and viscosizing (Fig. 3B, lanes 3,5 and 6, Fig. S4D (ESI†)). These micelles do not however survive the sieving conditions of PAGE. This instability is increased in the presence of EDTA due to efficient chelation of Mg²⁺. Mg²⁺ was found to stabilize all the LON micellar assemblies irrespective of their nucleotide sequence (G4 or scramble), probably as a result of an entropic gain associated with the Mg²⁺-mediated neutralization of the numerous micellar anionic charges. For instance, **k-LON^{SC}** no longer migrates during PAGE with increased amounts of Mg²⁺ just as what is observed in the absence of salts with the stronger **k-LON^{G4}** micelles (Fig. 3, lane 7). The kinetic stability of **k-LON^{G4}** micelles over those formed from **k-LON^{SC}** was further evidenced by the partitioning or the lack thereof of respectively **k-LON^{SC}** and **k-LON^{G4}** monomers into micelles of the neutral triton (Fig. S5 (ESI†)).
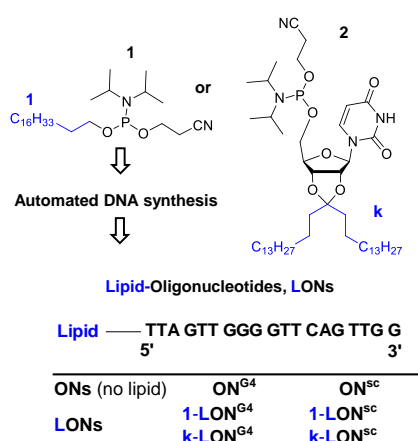

Fig 2. Chemical structures of lipid phosphoramidites. (1) octadecyl, (2) ketal "k" and DNA sequences of the different (L)ONs studied.

We further characterized the kinetics of assembly. The **1-LON^{G4}** molecules that form tetramolecular parallel G4 fold quickly (t <
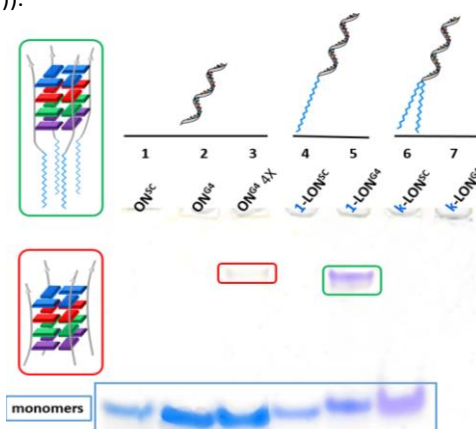
Fig 3. Native PAGE of the different (L)ONs in the presence of 1X salts (except lane 3). The absence of band for k-LON^{G4} (lane 7) results from the formation of stable micellar aggregates

Tetramolecular parallel G4 formation occurs at the surface of **k-LON^{G4}** micelle as judged by the CD signature and K⁺-dependence observed in agarose gel for **k-LON^{G4}** solutions (Fig. S6 (ESI†)). In line with the findings by Wilner *et al* on short G4–

forming LONs,[25] our results suggest that the G4-prone sequence of **k-LON^{G4}** stabilizes the micelles, with a lesser effect by **1-LON^{G4}** (Fig. 4B, lanes 3 and 4) in which individual **1-LON^{G4}** G4s are visible (Fig. 3, lane 5).
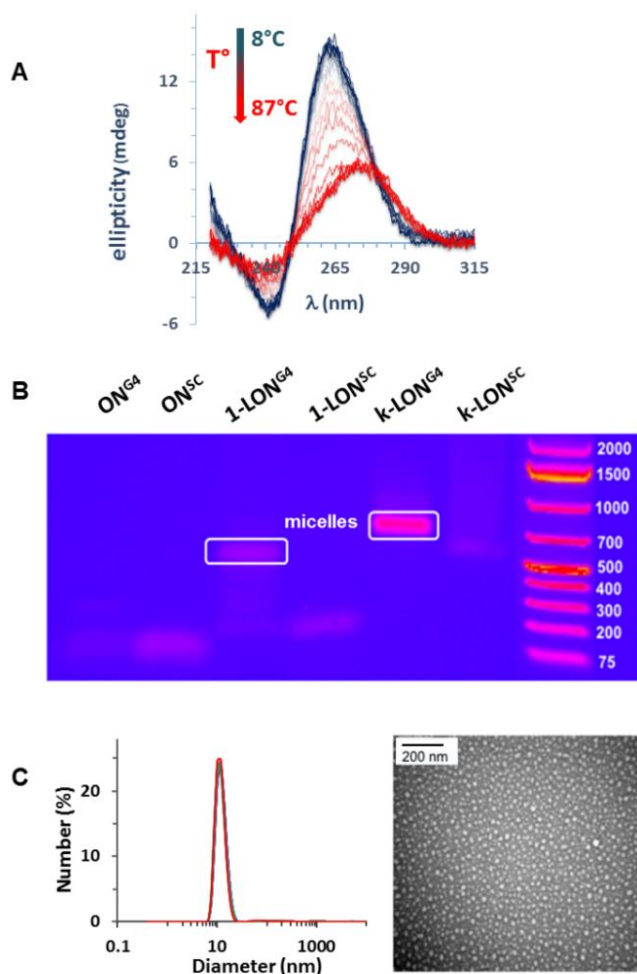

Fig 4. CD melting profile of 1-LON^{G4} (A) Native agarose gel of the (L)ONs (2X salt) (B) DLS and TEM of 1-LON^{G4} (C).

Interestingly, **k-LON^{G4}** exhibited very well-defined micellar bands compared to its scramble analog (Fig. 4B, lanes 5 and 6 respectively). No voluminous aggregates being observed by DLS or viscosizing (Fig. S7 + S8 (ESI†)), the prominent trailing shoulders observed with **k-LON^{SC}** more likely resulted from unspecific adsorptive interactions between the highly concentrated free scramble DNA sequences at the micelle surface and the agarose gel matrix during electrophoresis. These unspecific interactions increase with DNA sizes and decrease with increasing salt concentrations,[29] in line with our observations (Fig. S7 (ESI†)). Conversely, the favorable intra-micellar G4 formation within the **k-LON^{G4}** hydrophilic micellar corona may prevent unspecific adsorption of the LON with the agarose matrix.

Overall, our results suggest that the stability of the LON micellar assemblies is efficiently tuned playing around 1) the nature of the lipid, 2) the sequence of the oligonucleotide and 3) the nature of the salts present. Hence, the combination of a dialkyl lipid together with a G4-prone oligonucleotide efficiently "freezes" micelles. In contrast, the monoalkyl, non-G4 **1-LON^{SC}** does not give any micelles in PAGE or

agarose gels unless high strand concentrations (> 50μM in DLS) and Mg^{2+} are used**. k-LON^{SC}** and **1-LON^{G4}** constitute interesting intermediate cases as potential drug delivery nano-cargoes. Conventional lipid-based micelles are appealing low-nanometric drug carriers but suffer from high kinetic instability for in vivo purposes. On the other hand, very stable micelles may not release their drug payload. Our system may therefore constitute an interesting alternative in this regard.

Tetramolecular G4s reported in the literature are parallel and usually formed from short oligonucleotide sequences (6-8 nt in length in general). Indeed, dimers, trimers, mismatched quadruplexes and pentamers) and/or other undesired folds are found kinetically trapped in solution during the folding of the perfect tetramolecular parallel G4.[10,11] Given their molecularity of four, their kinetics of formation are extremely slow at micromolar concentrations.[27] We indeed find that no or very little assembled tetramolecular parallel G4s are observed with 5 μM unmodified **ON^{G4}** (only a faint retarded band being visible (Fig. 4B, first lane)) even after 48h. In contrast, we find that the presence of the alkyl chain in **1-LON^{G4}** drastically accelerated the kinetics of tetramolecular parallel G4 formation under the same conditions, such that we were unable to follow its rate of formation by CD and thus could not further pursue kinetic aspects of the assembly process. Directed oligomerization of **1-LON^{G4}** in the correct parallel orientation from micelle formation is likely responsible for this impressive rate acceleration. Despite this increased rate of folding in the presence of the lipid, perfect folding of the tetramolecular parallel G4 is still not achieved as seen by the amount of remaining 'monomers' (Fig. 3, lane 5). We observe that the monomer and folded species are not free to equilibrate in solution as supported by the high thermal stability of **1-LON^{G4}** tetramolecular parallel G4 (Fig. S9 (ESI†)) and consistent ratio of **1-LON^{G4}** monomer and G4 over time. Instead, the **1-LON^{G4}** tetramolecular parallel G4 corresponds to an energetic minimum, with the other energetic pathways leading to kinetically trapped species stable enough to prevent equilibration with monomers, but not stable enough to survive the sieving conditions of the PAGE. This hypothesis was confirmed by heating **1-LON^{G4}** samples above the melting temperature of these undesired species but below the Tm of the perfectly matched tetramolecular parallel G4 (the likely thermodynamically favored complex). The now unfolded monomers were then available to form matched tetramolecular parallel G4s upon cooling, although several temperature cycles were required for near complete conversion (Fig. S10 (ESI†)).

The kinetically trapped species likely correspond to mismatched parallel quadruplexes and perhaps other oligomers; as already observed with unmodified tetramolecular parallel G4s.[10,11] Indeed, although the solution is clearly enriched in perfectly matched **1-LON^{G4}** tetramolecular parallel G4 after the temperature cycles (Fig. S10 (ESI†)), the intensity of the tetramolecular parallel G4 CD signal for the enriched solution does not increase (Fig. S11 (ESI†)). Of note, these undesired folds are likely also present in **k-LON^{G4}** micelles (as well as in the other lipidic tetramolecular parallel G4 reported so far) but not observable because of the stability of the micelles. Importantly, the presence of Na$^+$ ions is required for the temperature cycles to work. In fact, the averaged Tm of the **1-LON^{G4}** tetramolecular parallel G4s is >95°C in the presence of K$^+$ or Mg^{2+} and only about 40°C with Na$^+$ (Fig. S12 (ESI†)). Accordingly, the kinetically

trapped imperfect tetramolecular parallel G4 species may remain stabilized in the presence of K^+ and Mg^{2+}. This effect would explain why no melting of these undesired species is observed. This, together with the absence of decrease in the tetramolecular parallel G4 CD intensity in the presence of Na^+ point at imperfect tetramolecular parallel G4s as the likely main contributors to the undesired folds.

We capitalized on this impressive preference of K^+ over Na^+ in the stabilization of the G4 from **1-LON^{G4}** to develop a biologically-relevant switch. We mimicked intra- and extracellular salt concentrations and demonstrated that **1-LON^{G4}** forms tetramolecular parallel G4s only in intracellular salt conditions at 37°C (Tm>90°C), while no G4 signature is observed at this temperature in the presence of extracellular salt concentrations (Tm<20°C, Fig. S13 (ESI†)).

Finally, the increase in salt concentration is also likely responsible for the stabilization of these imperfect tetramolecular parallel G4s (lipid-modified or not, see for instance lanes 2 and 3, Fig. 2). As of Mg^{2+}, divalent cations do not necessarily stabilize G4s (they may lead to electrostatic repulsion in the central core of the quadruplexes when present in excess).[13,30] In our case, Mg^{2+} may well help bridging and neutralizing the phosphate groups of the long flanking regions around the G4 core. The thermodynamics behind these remarkable Na^+ and Mg^{2+} salt effects are currently under investigation.

In conclusion, lipid modification of a long G4-prone oligonucleotide sequence is critical to favor the thermodynamics and kinetics of tetramolecular parallel G4 formation over other folds or oligomeric states. The nature of the lipid in the LON structure is important to observe, manipulate the individual tetramolecular G4s and tune the stability of the micellar assemblies. Our data highlight the potential of LON for the construction of parallel G4s of unprecedented length. This provides a means to construct promising and potentially switchable supramolecular architectures for nanotechnology and nucleic acid-based therapeutics like the LON switchable G4 demonstrated in this communication that forms only in intracellular salt conditions.

Conflicts of interest and acknowledgments

Notes and references

1   J. Choi and T. Majima, *Chem. Soc. Rev.*, 2011, **40**, 5893–5909.
2   J. T. Davis, *Angew. Chem. Int. Ed.*, 2004, **43**, 668–698.
3   E. Hasuike, A. M. Akimoto, R. Kuroda, K. Matsukawa, Y. Hiruta, H. Kanazawa and R. Yoshida, *Chem. Commun.*, 2017, **53**, 3142–3144.
4   B. Shlyahovsky, Y. Li, O. Lioubashevski, J. Elbaz and I. Willner, *ACS Nano*, 2009, **3**, 1831–1843.
5   O. Mendoza, J.-L. Mergny, J.-P. Aimé and J. Elezgaray, *Nano Lett.*, 2016, **16**, 624–628.
6   S. A. Ohnmacht, C. Marchetti, M. Gunaratnam, R. J. Besser, S. M. Haider, G. D. Vita, H. L. Lowe, M. Mellinas-Gomez, S. Diocou, M. Robson, J. Šponer, B. Islam, R. B. Pedley, J. A. Hartley and S. Neidle, *Sci. Rep.*, 2015, **5**, 11385.
7   C. K. Kwok and S. Balasubramanian, *Angew. Chem. Int. Ed.*, 2015, **54**, 6751–6754.
8   E. Largy and J.-L. Mergny, *Nucleic Acids Res.*, 2014, **42**, e149–e149.
9   A. Włodarczyk, P. Grzybowski, A. Patkowski and A. Dobek, *J. Phys. Chem. B*, 2005, **109**, 3594–3605.
10   C. Bardin and J. L. Leroy, *Nucleic Acids Res.*, 2008, **36**, 477–488.
11   F. Rosu, V. Gabelica, H. Poncelet and E. D. Pauw, *Nucleic Acids Res.*, 2010, **38**, 5217–5225.
12   S. Burge, G. N. Parkinson, P. Hazel, A. K. Todd and S. Neidle, *Nucleic Acids Res.*, 2006, **34**, 5402–5415.
13   J.-L. Mergny, A. De Cian, A. Ghelab, B. Saccà and L. Lacroix, *Nucleic Acids Res.*, 2005, **33**, 81–94.
14   J. Zhou, F. Rosu, S. Amrane, D. N. Korkut, V. Gabelica and J. L. Mergny, *Methods*, 2014, **67**, 159–168.
15   H. Liu, K. D. Moynihan, Y. Zheng, G. L. Szeto, A. V. Li, B. Huang, D. S. Van Egeren, C. Park and D. J. Irvine, *Nature*, 2014, **507**, 519–522.
16   S. E. Wilner, S. E. Sparks, D. Cowburn, M. E. Girvin and M. Levy, *J. Am. Chem. Soc.*, 2015, **137**, 2171–2174.
17   G. Koutsoudakis, A. P. de León, C. Herrera, M. Dorner, G. Pérez-Vilaró, S. Lyonnais, S. Grijalvo, R. Eritja, A. Meyerhans, G. Mirambeau and J. Díez, *Antimicrob. Agents Chemother.*, 2017, AAC.02354-16.
18   A. Gissot, M. Camplo, M. W. Grinstaff and P. Barthélémy, *Org. Biomol. Chem.*, 2008, **6**, 1324–1333.
19   A. Patwa, A. Gissot, I. Bestel and P. Barthélémy, *Chem. Soc. Rev.*, 2011, **40**, 5844–5854.
20   M. Kwak and A. Herrmann, *Chem. Soc. Rev.*, 2011, **40**, 5745.
21   M. Raouane, D. Desmaële, G. Urbinati, L. Massaad-Massade and P. Couvreur, *Bioconjug. Chem.*, 2012, **23**, 1091–1104.
22   A. Patwa, A. Gissot, K. Oumzil and P. Barthélémy, in *DNA in Supramolecular Chemistry and Nanotechnology*, E. Stulz and G. Clever, Wiley., 2015, pp. 276–293.
23   S. Cogoi, U. Jakobsen, E. B. Pedersen, S. Vogel and L. E. Xodo, *Sci. Rep.*, 2016, **6**, srep38468.
24   B. Ugarte-Uribe, S. Grijalvo, S. N. Pertíñez, J. V. Busto, C. Martín, A. Alagia, F. M. Goñi, R. Eritja and I. Alkorta, *Bioorg. Med. Chem.*, 2017, **25**, 175–186.
25   A. Gissot, K. Oumzil, A. Patwa and P. Barthélémy, *New J. Chem.*, 2014, **38**, 5129–5134.
26   U. Jakobsen and S. Vogel, *Org. Biomol. Chem.*, 2016, **14**, 6985–6995.
27   M. J. Palte and R. T. Raines, *J Am Chem Soc*, 2012, **134**, 6218–6223.
28   A. Gissot, C. Di Primo, I. Bestel, G. Giannone, H. Chapuis and P. Barthélémy, *Chem. Commun.*, 2008, 5550–5552.
29   S. S. Smith, T. E. Gilroy and F. A. Ferrari, *Anal. Biochem.*, 1983, **128**, 138–151.
30   N. V. Hud and J. Plavec, in *Quadruplex Nucleic Acids*, eds. S. Neidle and S. Balasubramanian, 2006, pp. 100–130.