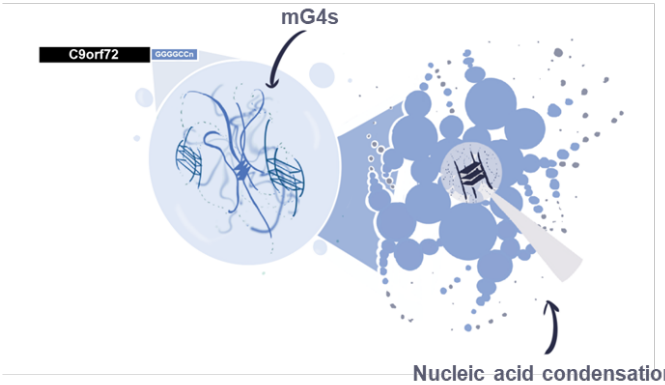
Long-range G-G base pairing regulates chromatin architecture and promotes RNA condensation in neurodegenerative diseases.

Marco Di Antonio (m.di-antonio@imperial.ac.uk)

Imperial College London, Chemistry Department, Molecular Sciences Research Hub, 82 Wood Lane, London W12 0BZ, UK
The Francis Crick Institute, 1 Midland Road, London NW1 1AT, UK

Abstract

It is well known that the Guanine base can base pair with itself by means of Hoogsteen hydrogen bonding, which can lead to the formation of DNA and RNA secondary structures known as G-quadruplex (G4).[1] Whilst G4-formation has been thoroughly investigated within short genomic sequences, G-G base pairing between distal genomic regions have been often overlooked and deemed unlikely to happen in vivo. My group has recently discovered the first human protein (CSB), a chromatin remodelling protein, that can selectively bind to multimolecular G4-structures (mG4s), suggesting that long range G-G base pairing can be potentially leveraged to orchestrate chromatin architecture.[2] More recently, my group has demonstrated that the repeat expansion of the hexanucleotide $(GGGGCC)_n$, which has been linked to the neurodegenerative diseases ALS and FTD, can form solid aggregates in a protein independent fashion by means of G-G base pairing. We showed a correlation between the emergence of multimolecular G4s (mG4s) formed by the DNA $(GGGGCC)_n$ repeats and the formation of protein free insoluble aggregates. Aggregation is dependent on K^+ concentration and repeat-length, indicating that G4-formation is essential to observe aggregates. G4-structures were detected in the aggregates by staining with the G4-specific fluorescent dye NMM. To reinforce the physiological relevance of our observations, we characterised the aggregation of RNA $(GGGGCC)_n$, which is thought to contribute to pathological aggregation in ALS/FTD. We observed that RNA repeats can aggregate at significant lower concentrations compared to DNA, suggesting that under physiological conditions RNA repeats can aggregate in the absence of any protein. Using patient-derived ALS cell lines, we validated our model by observing the same G4-based RNA aggregates in the pathological RNA *foci* that are characteristic of this disease, suggesting that nucleic acids targeting could be the key to treat neurodegenerative diseases in the future. Our findings constitute the first evidence supporting the formation of multimolecular G4-structures to drive protein-free aggregation in neurodegenerative diseases, challenging the current dogmas on the mechanisms responsible of neurodegeneration and associate protein led aggregate formation.[3]



1. Robinson, J.; Raguseo, F.; Nuccio, S. P.; Liano, D.; Di Antonio, M. *Nucleic Acids Res.* **2021**, *49* (15), 8419-8431.
2. Liano, D.; Chowdhury, S.; Di Antonio, M. *J. Am. Chem. Soc.* **2021**, 143 (49), 20988-21002.
3. Raguseo, F., Tanase, D., Malouf, L., Rubio Sanchez, R., Elani, Y., Di Michele, L., Di Antonio, M. *bioRxiv,* **2023** Pages 2023.01.31.526399 DOI: 10.1101/2023.01.31.526399