doi 10.18699/bgrs2024-11.1-16

G-quadruplexes at super-enhancers and CTCF clusters help shape topologically associating domains/replication domains

Varizhuk A.*, Pavlova Iu., Iudin M., Barinov N., Sultanov R., Bogomazova A., Klinov D., Tsvetkov V.

Lopukhin Federal Research and Clinical Center of Physical-Chemical Medicine, Moscow, Russia
** annavarizhuk@gmail.com*

Key words: G-quadruplex; topologically associating domain; replication domain; CTCF; nucleosome density; super-enhancer; BRD4

Motivation and Aim: Topologically associating domains (TADs), which are established in the G1 phase of the eukaryotic cell cycle and restrict the enhancer-promoter communication, coincide with replication domains, which determine replication timing in the S-phase. Several models of the spatiotemporal regulation of transcription and replication have been proposed to explain this observation. These include the Rosette model and the CRISI (chromatin reorganization induced selective initiation) model [1]. The former suggests the intra-TAD co-firing of high-efficiency replication origins, whereas the latter implies the relocation of the high-efficiency origins towards the TAD periphery due to TAD dynamics upon the enhancer-promoter communication. Both models provide new insights into the regulatory role of G-quadruplex DNA structures (G4s) in replication. So far, they have only been considered as possible sites of pre-replication complex assembly or replication stress triggers in helicase-deficient cells. However, a growing body of data suggests that G4s may also be involved in replication-associated chromatin remodeling, namely TAD demarcation and dynamics. The former process requires local accumulation of the CCCT-binding factor (CTCF), and the latter can be driven by liquid-liqiud phase separation (LLPS) at super-enhancers (SEs). In light of the reported colocalization of G4s with strong TAD boundaries and intra-TAD interaction hotspots in the human genome [2], we sought to investigate whether G4s promote the recruitment of CTCF or LLPS drivers.

Methods and Results: Recently, we investigated the impact of G4s on CTCF positioning based on CTCF-ChIP-seq experiments using human myeloid leukemia cells (K562) pretreated with a G4-stabilizing ligand, pyridostatin (PDS). The CTCF profiles in control or PDS-treated cells were compared to the reported profiles of G4s predicted by G4-seq or confirmed by G4-ChIP-seq. The G4s with confirmed intrinsic stability in the K562 context demonstrated a high correlation with CTCF in the control samples, with minimal response to the stabilizing ligand. Conversely, the predicted (semi-stable) G4s exhibited a weak correlation with CTCF in the control samples, but this correlation increased significantly in response to PDS [3]. The ChIP-seq results were corroborated by ChIP-PCR for several representative G4 sites, and some of the G4s were shown to bind CTCF with Kd values comparable to that obtained for the consensus CTCF binding motif in a microscale thermophoresis (MST)-based binding assay.

These findings agree with the results obtained by other groups [4] and support direct CTCF recruitment by G4s. Importantly, they do not rule out indirect G4-related mechanisms of CTCF accumulation. In particular, G4s could facilitate the positioning of CTCF at nearby cognate binding sites by modulating the local CpG methylation rates

and/or nucleosome density. The molecular basis for G4-dependent methylation control has been explored elsewhere. We aimed to elucidate the relationship between G4s and nucleosome-depleted regions (NDR).

We obtained a minimal chromatin model, which included a single histone octamer (HO) assembled on a dsDNA comprising a core Widom 601 nucleosome positioning sequence and a G4-harboring or non-G4 tail. The folding of the G4 within the dsDNA tail was confirmed by several methods, including thioflavin light-up assays and ^{1}H-NMR. HO assembly on dsDNA with G4 versus non-G4 tails was analyzed by polyacrylamide gel electrophoresis. The relative intensities of the nucleosome bands indicated a reduced efficiency of HO assembly near the G4 tail. This observation was additionally verified by atomic force microscopy (AFM)-based analysis of nucleosome morphologies in samples with G4 versus non-G4 DNA [5]. Taken together, our data confirm nucleosome exclusion around G4s, which may be favorable for CTCF clustering at TAD boundaries but would also affect the intra-TAD connectivity. Given that transcriptionally active TADs are typically hyperacetylated (H3K27ac) at G4-rich promoter and enhancer regions, we additionally verified the exclusion of HOs with acetylated H3 around G4s and obtained analogous results.

In studying intra-TAD interactions, we focused on the formation biomolecular condensates at SEs, which typically communicate with multiple promoters. Such biocondensates are appear to be scaffolded by the SE marker BRD4, which recognizes acetylated chromatin through its tandem bromodomains and co-separates with the Mediator complex subunits due to its disordered C-terminal region [6]. The analysis of BRD4 distribution in K562 SEs using ChIP-seq data revealed high occupancy of this protein at/near NDRs, which are twice as frequent in G4-rich SEs as in G4-free SEs. Since acetylated nucleosomes are absent around G4s, the mechanism of BRD4 maintenance is currently uncertain. LLPS explains the nearly constant concentration of BRD4 along the SE. However, LLPS necessitates that DNA engages in multiple transient (weak) interactions with BRD4. Nucleosome-free dsDNA does not bind to BRD4 or facilitate LLPS at biologically relevant concentrations. Therefore, we tested representative G4s from SEs for BRD4 binding and LLPS promotion.

Representative G4s from K562 SEs were identified using MEME and FIMO algorithms. Sequences matching the highly enriched and moderately enriched motifs were selected for binding assays. Negative controls included dsDNA and iM DNA. Both G4s showed weak affinity for BRD4 with med-micromolar Kd values in both MST-based and surface plasmon resonance (SPR)-based assays, while dsDNA and iM showed negligible binding. The highly enriched motif (hereafter referred to as SE-G4) bound BRD4 more efficiently than the moderately enriched motif in both MST and SPR assays.

To assess the impact of SE-G4 on BRD4-driven LLPS, fluorescently labeled BRD4 was mixed with isolated SE-G4 or SE-G4 within dsDNA (nucleosome-free), DNA-free acetylated histones (negative control), nucleosome-free non-G4 dsDNA (negative control), or acetylated nucleosomes assembled on non-G4 dsDNA (positive control). LLPS was monitored by fluorescence microscopy. Biocondensates, visualized as micrometer-scale droplets, were observed only in the positive control and G4-containing samples. Their morphology was confirmed by AFM. To summarize, G4s from SEs facilitated LLPS *in vitro* and thus may contribute to the scaffolding of biocondensates at SEs.

Conclusion: We have shown that G4 structures, which are enriched at TAD boundaries or intra-TAD interaction hotspots, can directly recruit CTCF or increase the accessibility

of its binding sites within dsDNA by excluding nearby nucleosomes. At G4-rich SEs within transcriptionally active TADs, the exclusion of acetylated nucleosomes critical for BRD4 loading is compensated by the increased LLPS-promoting potency of G4s compared to duplex DNA. Our results add to the current understanding of G4 involvement in TAD demarcation and enhancer-promoter communication. However, it should be noted that they were obtained using simplified chromatin models and require further verification.

Funding: The study is supported by Russian Science Foundation (No. 22-15-00129).

References

1. Li Y., Xue B., Zhang M. et al. Transcription-coupled structural dynamics of topologically associating domains regulate replication origin efficiency. *Genome Biol*. 2021;22(1):206. doi 10.1186/s13059-021-02424-w
2. Hou Y., Li F., Zhang R., Li S., Liu H., Qin Z.S., Sun X. Integrative characterization of G-Quadruplexes in the three-dimensional chromatin structure. *Epigenetics*. 2019;14(9):894-911. doi 10.1080/15592294.2019.1621140
3. Tikhonova P., Pavlova Iu., Isaakova E. et al. DNA G-Quadruplexes Contribute to CTCF Recruitment. *Int J Mol Sci*. 2021;22(13):7090. doi 10.3390/ijms22137090
4. Roy S.S., Bagri S., Sengupta A., Then C.R., Kumar R., Sridharan S., Chowdhury S. Artificially inserted G-quadruplex DNA secondary structures induce long-distance chromatin activation. *eLife*. 2024;13:RP96216. doi 10.7554/eLife.96216.1
5. Pavlova I., Barinov N., Novikov R. et al. Modeling G4s in chromatin context confirms partial nucleosome exclusion and reveals nucleosome-disrupting effects of the least selective G4 ligands. *Biochimie*. 2023;204:8-21. doi 10.1016/j.biochi.2022.08.016
6. Sabari B.R., Dall'Agnese A., Boija A. et al. Coactivator condensation at super-enhancers links phase separation and gene control. *Science*. 2018;361(6400):eaar3958. doi 10.1126/science.aar3958