

SURVEY AND SUMMARY

DNA G-quadruplex structures: more than simple roadblocks to transcription?

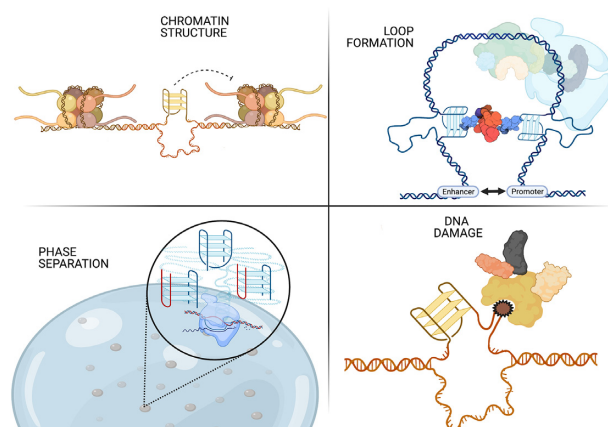
Jenna Robinson^{1,2}, Federica Raguseo^{1,2}, Sabrina Pia Nuccio^{1,2}, Denise Liano^{1,2} and Marco Di Antonio^{1,2,3,*}¹Imperial College London, Chemistry Department, Molecular Sciences Research Hub, 82 Wood Lane, London W12 0BZ, UK, ²Institute of Chemical Biology, Molecular Sciences Research Hub, 82 Wood Lane, London W12 0BZ, UK and ³The Francis Crick Institute, 1 Midland Road, London NW1 1AT, UK

Received May 26, 2021; Revised July 01, 2021; Editorial Decision July 01, 2021; Accepted July 05, 2021

ABSTRACT

It has been >20 years since the formation of G-quadruplex (G4) secondary structures in gene promoters was first linked to the regulation of gene expression. Since then, the development of small molecules to selectively target G4s and their cellular application have contributed to an improved understanding of how G4s regulate transcription. One model that arose from this work placed these non-canonical DNA structures as repressors of transcription by preventing polymerase processivity. Although a considerable number of studies have recently provided sufficient evidence to reconsider this simplistic model, there is still a misrepresentation of G4s as transcriptional roadblocks. In this review, we will challenge this model depicting G4s as simple 'off switches' for gene expression by articulating how their formation has the potential to alter gene expression at many different levels, acting as a key regulatory element perturbing the nature of epigenetic marks and chromatin architecture.

GRAPHICAL ABSTRACT



INTRODUCTION

G-quadruplexes (G4s) are DNA structures that can arise under physiological conditions from guanine-rich DNA sequences (Figure 1) (1–4). Although the ability of guanine analogues to assemble into tetrameric structures has been known since 1962 (5), the idea that G4 structures could form in the context of genomic DNA was only seriously considered 40 years later when the crystal structure of a G4 formed by the human telomeric DNA sequence was reported for the first time (6). The very same year a different G4-forming sequence was described in the promoter region of the oncogene *MYC* by Hurley *et al.* (7). Importantly, this study was not limited to the structural characterization of the quadruplex, but also provided evidence that such a G4 was biologically active by showing *MYC*-downregulation as a direct consequence of G4 stabilization by the porphyrin-based ligand TMPyP4 (7). This observation sparked interest in non-

*To whom correspondence should be addressed. Tel: +44 20 7594 5866; Email: m.di-antonio@imperial.ac.uk

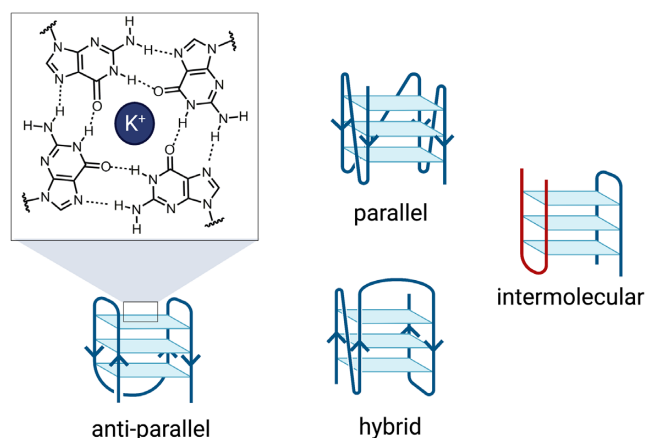


Figure 1. Schematic representation of G-quadruplex structures. G4s are constituted of four guanine bases arranged in a square planar conformation (G-tetrad) held together by Hoogsteen hydrogen bonding and further stabilized by alkali cation such as K^+ . Specific G4 topologies can be formed and include anti-parallel, parallel and hybrid structures, depending on the relative orientation of the DNA strand within the structure. Intermolecular G4 structures can also be formed when more than one DNA strand is used to generate the final structure (red and blue strands).

telomeric G4s and suggested that these structures could act as transcriptional repressors, a concept supported in several following studies on G4s formed in other oncogene promoters, such as c-KIT, BCL-2, KRAS and VEGF (8). This general notion was further consolidated by multiple bioinformatic studies revealing that G4s are enriched at gene promoters, particularly in oncogenes (9–11), supporting a role of these structures in regulating gene-expression and their potential to be targeted by small-molecules for therapeutic intervention.

However, most of the initial studies that portrayed G4s as antagonists to gene expression relied on the use of synthetic molecules as G4-stabilizers or *in vitro* models of transcription (12), typically plasmid-based, which do not take into account the complexity of endogenous gene expression within chromatin. Indeed, direct evidence to support the idea that endogenous G4 formation at gene-promoters results in transcriptional repression is still lacking. In 2012 Rodriguez *et al.* challenged the notion of G4s as direct transcriptional repressors, by revealing that treatment with the potent G4 ligand pyridostatin (PDS) caused DNA-damage at G4 sites in a transcription and replication dependent fashion (13). In this key study, which generated the first genome-wide G4 map, it was demonstrated that the DNA damage elicited by PDS at G4 sites caused gene down-regulation, rather than the formation of the G4 structure itself. The same year, Hurley and co-workers challenged their own observations made with TMPyP4 10 years earlier, reporting how treatment with G4 ligands could lead to MYC downregulation without direct targeting and stabilization of the MYC G4, but rather as an indirect consequence of global G4 stabilization elicited by the ligand (14).

More recently, the development of G4-selective antibodies has enabled genome-wide mapping of G4s by immunoprecipitation experiments followed by sequencing (ChIP-Seq). These studies have allowed for a far more direct

investigation of the association between endogenous G4-formation and gene expression, revealing an almost inverted picture to the earlier investigations. Indeed, G4 ChIP-Seq performed in keratinocytes (15) and human xenografts (16) using the G4-selective antibody BG4, has demonstrated that G4s are prevalently found at gene promoters of transcriptionally active genes, acting as transcriptional enhancers rather than repressors. The link between G4 formation and active transcription has been confirmed beyond keratinocytes, in liposarcoma cells still using BG4-based ChIP-Seq (17), as well as in a range of mammalian cell lines using a G4-selective peptide probe (18), and even with single-molecule detection of G4s within live cells under non-perturbative conditions (19). Furthermore, G4 formation seems to contribute to activated gene expression by means of many different mechanisms that span from transcription factor binding (20) to guanine oxidation (21). Despite this, there is still a misrepresentation in the scientific community of G4s acting as transcriptional roadblocks and repressors, which is reminiscent of the initial hypothesis postulated 20 years ago. In this review, we aim to provide a critical assessment of the latest research relating G4s to transcriptional regulation, discussing the genome-wide effects of G4 formation on chromatin architecture, long-range interactions, phase separation and DNA oxidation.

G4 formation in the chromatin context

To fully appreciate and investigate the role G4 formation may play in gene regulation, we need to consider the genomic context in which G4s exist. The 3D organization of genetic material in cells presents a high level of complexity, where the canonical double helix of DNA is wrapped around proteins in a macromolecular structure known as chromatin. Chromatin is made of nucleosomes: a repetitive unit consisting of histone proteins which wrap around a stretch of DNA (ca. 146 bp) contributing to the folding of DNA inside the nucleus (Figure 2A) (22,23). However, genome sequencing and microarray hybridization technologies have revealed that nucleosome occupancy is not homogenous across DNA and that there are certain genomic regions depleted of nucleosomes (24). Nucleosome depleted regions (NDRs) are stretches of more accessible DNA which act as hubs for protein binding, including those that regulate transcription (24). Interestingly, the location of NDRs within the genome is highly dynamic and dependent on chemical modifications to histones, such as methylation or acetylation, which change the interaction of histones with key proteins involved in regulating gene expression (25). Therefore, nucleosome positioning across the genome is key to define the epigenetic status of a cell and this organization appears to also be associated with the formation of certain DNA secondary structures, including G4s.

The first evidence that G4s may have a role in shaping chromatin was highlighted in a study focused on REV1, a protein belonging to the Y family of DNA polymerases, that has an essential role in ensuring replication proceeds when DNA damage occurs (26). Cells containing a mutant form of REV1 were characterized by delayed or fully compromised DNA replication, particularly at G4-forming sequences (26). The unresolved G4 structures in REV1 mu-

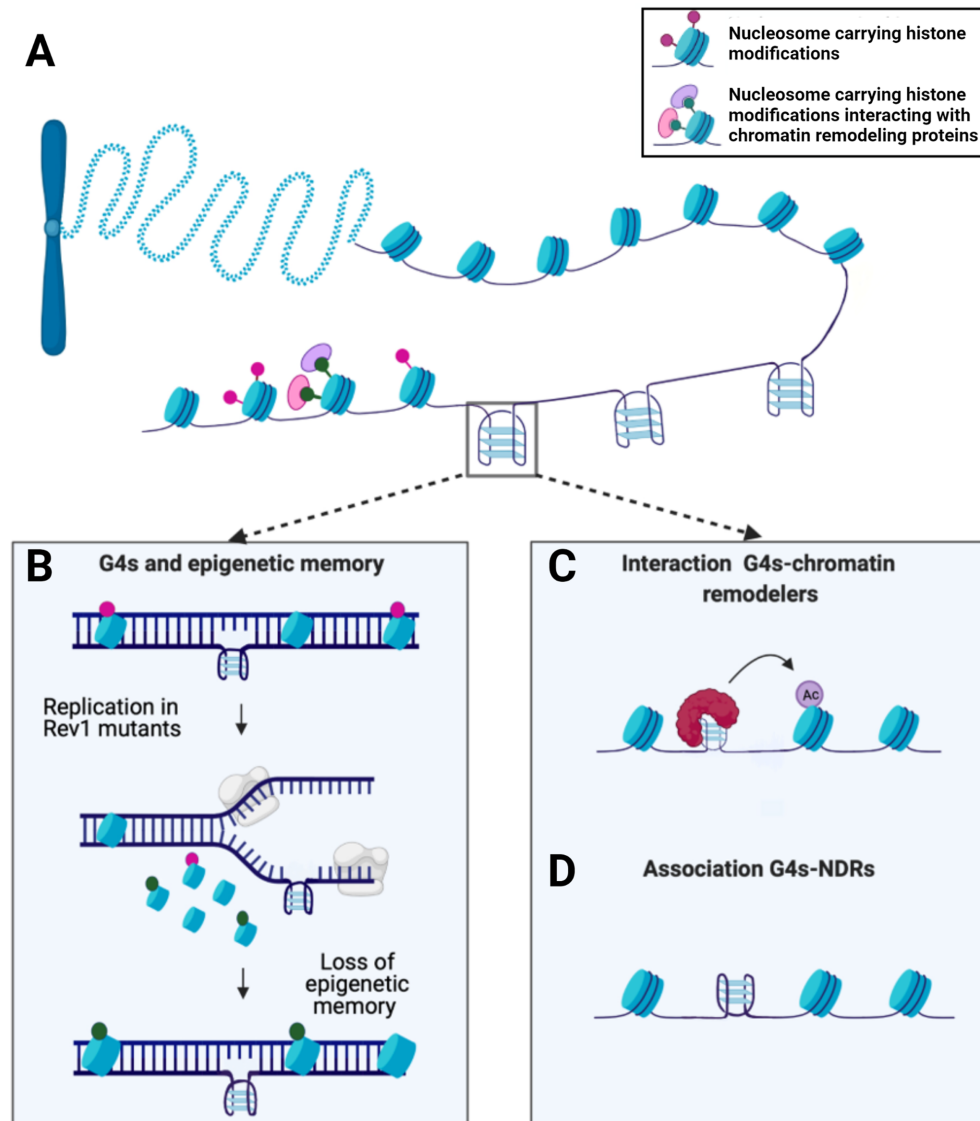


Figure 2. G4s and chromatin structure. (A) Structure of chromatin comprised of DNA wrapped around nucleosome complexes. (B) REV1 mutants unable to resolve G4s during DNA replication exhibit altered histones modifications in the newly synthesized strand and a consequent loss of epigenetic memory. (C) G4s can interact with a wide panel of proteins, including chromatin remodellers such as BRD3. (D) G4s are strongly associated with nucleosome depleted regions (NDRs) as confirmed by BG4-ChIP.

tants led to replication fork stalling, but surprisingly also had a significant impact on the epigenetic status of the cell (Figure 2B). In particular, an increase in the expression of the p-globin locus was observed. This upregulation was due to the loss of a histone modification negatively associated with transcription: H3K9 dimethylation, as well as incorporation of newly acetylated histones around the G4-forming site, promoting active transcription.

However, a later study of a separate locus revealed that unresolved G4 structures in *REV1* mutants could also lead to the loss of distinct histone marks that promote transcription, suggesting that the role of G4s in modifying the histone code is dependent on their genomic context (27). Although, these findings were limited to chicken DT40 cells, they suggested for the first time a relationship between G4s and the installation of epigenetic marks, which was more

recently confirmed by immunofluorescence studies (28). In this latest work, colocalization between G4 structures and histone modifications found in nucleosome-depleted euchromatin regions was observed, underlining the potential of G4s in moulding the epigenetic landscape.

From these findings, some open questions arise: are G4s a passive element in the deposition of histones marks or do they have a mechanistic role in shaping chromatin that is not limited to replication stalling? To unravel any mechanistic link between G4s and chromatin structure, it may be key to investigate G4-interacting proteins especially those that are known for having an active role in chromatin remodelling. An example is the protein BRD3 that was found to be one of the top hits in an unbiased microarray screening of the G4-interactome (29). BRD3 contains a bromo-domain capable of binding acetylated histones, allowing for the re-

cruitment of RNA polymerase and initiation of transcription (30). The idea that G4s interact with BRD3 within cells has been substantiated by BG4 ChIP-seq and BRD3 ChIP-seq analysis, which revealed a significant colocalization of G4 motifs within BRD3 occupancy sites (31). Additionally, the interaction between G4s and BRD3 does not prevent its binding to acetylated histones or other chromatin remodelers, suggesting that BRD3-G4 interactions might instead guide the recruitment of chromatin remodelling complexes to favour transcription at G4-sites (Figure 2C).

These findings highlighted the relevance that G4s may have in engaging key regulators of chromatin architecture, leading to the hypothesis that global changes in chromatin structure could be directly caused by G4 formation in the genome. This hypothesis is supported by bioinformatic associations made on a genome-wide scale, which demonstrated that G4-forming sequences are highly enriched at nucleosome depleted regions (Figure 2D) (10,32). Although these first studies were based on computational predictions of G4s, such results have recently been validated directly in the chromatin context by using the G4-selective antibody BG4 which confirmed the enrichment of G4s at nucleosome free sites (15). Specifically, 98% of G4s identified by BG4-ChIP coincided with NDRs, an observation that has now been recapitulated in independent studies (17). This exemplifies the interconnectedness between G4 formation and chromatin structure and strongly suggests that these structures might mark transcriptionally active regions and possibly influence nucleosome positioning.

Despite the fact that there is now significant evidence suggesting that G4s may act as an active factor in shaping the structure of chromatin, for example by altering histone modifications and nucleosome placement, it is also conceivable that G4s form at open chromatin sites as a simple consequence of DNA accessibility and negative supercoiling that is also required for active transcription (33). Future experiments must focus on assessing a direct causation between chromatin remodelling and G4 formation, underpinning whether these structures can act as recruiters of epigenetic enzymes to actively shape chromatin structure. Previous reports on BG4 ChIP-Seq have suggested that accessible chromatin is necessary but not sufficient for G4s to form (15), which indicates that these structures do not form as a simple consequence of chromatin accessibility. Identifying the regulators of their formation within open chromatin sites will be key to underpin the relevance of G4s in moulding the epigenetic landscape.

G4 formation stabilizes R-loops

In addition to nucleosome free DNA, a further structural requirement of transcription is the ability of duplex DNA to be unwound into single strands, one of which acts as the template strand for RNA polymerase, whilst the other is denoted the non-template strand. As polymerase transcribes the template strand, a three-stranded intermediate is formed comprising the two original DNA strands and the newly transcribed mRNA sequence, which make up a so-called R-loop (Figure 3A) (34). Interestingly, the requirements for R-loop generation are also compatible with G4-formation, as RNA:DNA hybrids form most stably with

C-rich regions on the template strand of DNA (35). Thus, the creation of R-loops often displaces a single-stranded G-rich sequence which may be primed for folding into a G-quadruplex (36–39). This can occur from the formation of an intramolecular G4 on the displaced strand (Figure 3A) or a DNA:RNA hybrid intermolecular G4 (Figure 3B) (40,41).

Such R-loops form significantly in actively transcribed genes and interestingly, genome-wide studies have also confirmed formation of G4-structures in transcriptionally active genes (15,17,18). Furthermore, it appears that G-quadruplexes may have a functional role in mediating transcription by actively stabilizing the R-loop (42), allowing for enhanced transcription. A recent study demonstrated this by showing that the placement of G4s on the non-template strand enhances R-loop formation resulting in significantly increased transcript output, RNA polymerase initiation and elongation (43). This study also revealed that such an effect was strand dependent, as no such increases in transcription are seen when the G4 motif was found on the template strand. Despite these interesting results, this work was limited to *in vitro* measures of transcription which may not encapsulate the complex interactions of G4s within cells. Thus, it would be worth further investigating if the presence of G4s in R-loops is similarly correlated with increased transcriptional output within a chromatin DNA context that is more representative of transcriptional processes in living cells.

In contrast to this work suggesting G4s have a positive effect on transcription by stabilizing R-loop formation, it has been noted that over-stabilization of G4s can have deleterious effects on transcription. For instance, artificially enriching G4 content either by exposing cells to G4-ligands (44) or down-regulating G4 helicases (42), increases the number of R-loops due to the positive relationship between G4 and R-loop formation (Figure 3C). These unscheduled R-loops can collide and stimulate DNA-damage in the form of double strand breaks as demonstrated by 53BP1 and γ H2AX marks that are enriched after G4 ligand incubation. In turn, such DNA damage and transcription-replication fork collisions can significantly hinder successful transcription, particularly in cells deficient in DNA damage repair proteins such as BRCA2 which are also known to be highly sensitive to G4 ligands (44,45). This is a useful example of how disrupting the natural homeostasis of G4 formation within cells can have antagonistic effects on gene expression that are not otherwise observed, highlighting the need to distinguish between such studies and those considering the epigenetic effects of endogenously formed G4s.

G4 formation can promote long-range DNA interactions

Although the enrichment of G4s in promoters of highly transcribed genes sprouted the notion that G4s may mediate transcriptional control (potentially by shaping local chromatin structure and R-loop formation), it is possible that this role expands beyond the local context and into additional long-range mechanisms of epigenetics. In fact, considering promoter sites directly proceeding genes constitute only a fraction of the human genome, research in this area may represent just the tip of the iceberg when it

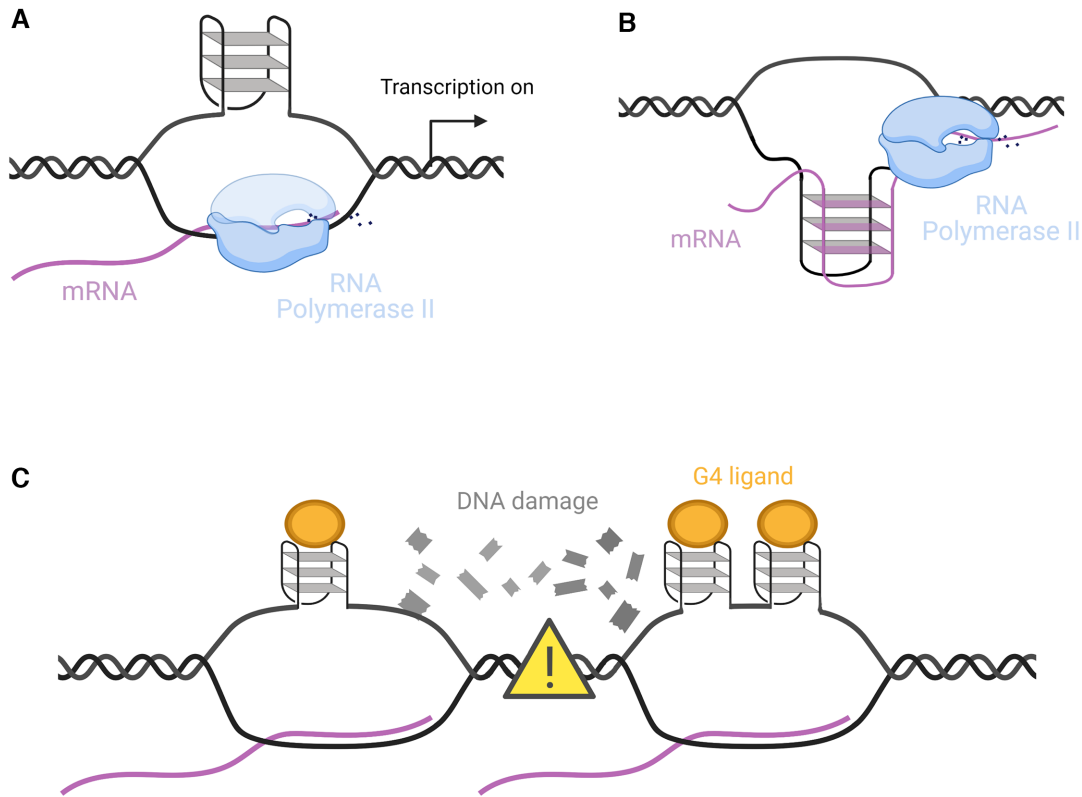


Figure 3. G4s and R-loops. (A) Intramolecular G4s formed on the non-template strand stabilize R-loops and increase transcriptional output *in vitro*. (B) Intermolecular DNA:RNA hybrid G4s may also form stabilizing R-loops on the template strand. (C) The use of G4-stabilizing small-molecules increases the prevalence of G4s and R-loops within cells. This can lead to the collision of nearby unscheduled R-loops, resulting in DNA-damage which hinders transcription.

comes to investigating how G4s control gene expression. For example, a single gene may have its transcription regulated by multiple regulatory sites littered all throughout the genome—sites which are able to exert their regulatory powers over long-distances and are often dysregulated in disease states (46–48). In recent years, compelling research has been conducted suggesting that G4s may not only be involved in proximal transcriptional control, but part of these long-range mechanisms that define an important component of the cell's epigenetic toolkit.

One of the most important long-range interactions utilized for epigenetic control is between transcriptional promoter and enhancer sites of DNA (Figure 4A) (49). While promoters mark the beginning points of transcription—the region at which RNA polymerase binds and begins its journey transcribing the code of DNA—enhancers act as auxiliary regulatory regions, recruiting additional proteins to carefully control the extent to which genes are expressed (50). This enhancer-mediated control occurs through the binding of large protein complexes containing transcription factors and cofactors, which increase the ability of RNA polymerase II to initiate and sustain transcription (50). So although promoter regions are sufficient to achieve a basal level of gene expression (51), it is through the utilization of enhancers that increases of transcription as high as 100-fold can be achieved (52). Interestingly, this control of transcription can occur over hundreds of thousands of base pairs;

(53) thus it has long been speculated that distal interactions in DNA may be facilitated by the folding of DNA into secondary structures such as loops to allow normally separated sequences to meet (Figure 4A) (49).

To test this hypothesis, countless chromatin conformation capture (3C/Hi-C) experiments have been conducted, which allow for the cross-linking and downstream sequencing of DNA regions known as 'topologically associated domains' (TADs) that are connected not in sequence, but in space (54). The results of which have shown that DNA looping occurs across the genome and can allow for distal interactions between regulatory sites. However, this work also revealed that regions where loops form in the genome are in fact not random, but instead occur in a sequence-specific fashion (55). The mechanism for this specificity is widely agreed to be achieved through the cooperation of two proteins: cohesin and CTCF (Figure 4B). Cohesin is a ring-like protein, initially discovered for its role clasping together sister chromatids during replication; however, it has a secondary function translocating down stretches of DNA to form transient loops (56). In order to prevent cohesin sliding down DNA indefinitely, the cell has utilized cohesin stop points in the form of binding to the transcription factor CTCF which is named after its interaction with CCCTC sequences in DNA. CTCF binds to and blocks the releasing function of cohesin (57), which means its presence defines the start and end points of loops and in

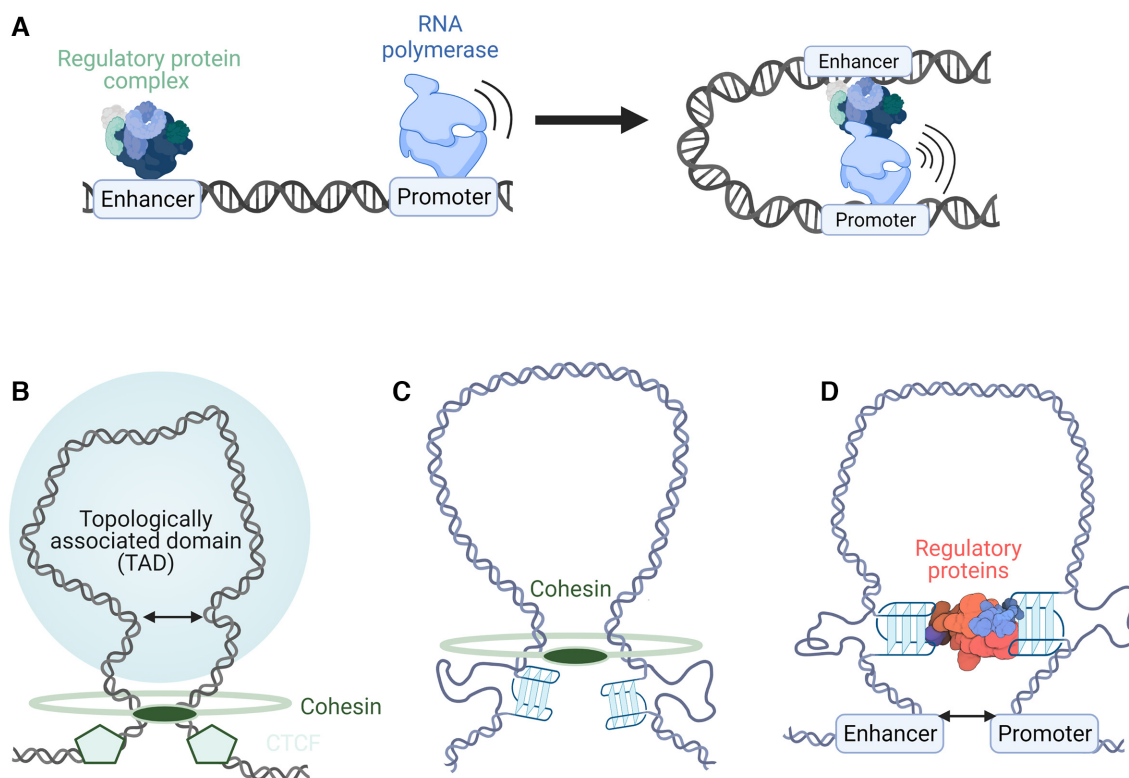


Figure 4. G4s in enhancer-promoter loop formation. (A) Schematic of how enhancers interact with their respective promoters to increase polymerase activity and transcription. (B) Cohesin and CTCF cooperate to form stable loops (topologically associated domains) in DNA. (C) G4 enrichment at loop boundaries possibly stalls cohesin. (D) G4s act as recruiters of regulatory proteins such as transcription factors that stabilize loops.

turn where the boundaries lie for topologically associated domains.

However, CTCF binding motifs are not the only sequences enriched at loop boundaries as G-quadruplex sequences are also abundant at the edges of loops (58). Moreover, it was found that the presence of G4s at loop boundaries increases the stability of DNA loops and in turn enhances long-distance DNA interactions (58). Although the mechanism for this correlation has yet to be fully explored, the authors of this study speculated that G4s may act analogously to CTCF, stalling the progression of cohesin and thus defining the boundaries of distal interactions (Figure 4C). This is further supported by the fact that G4 motifs at loop boundaries are significantly enriched on the same strand as the CTCF binding motif, suggesting some co-operation between the action of CTCF and G4 formation (58). This work demonstrates, perhaps counter-intuitively, that the protein-stalling capabilities of G4s that have been demonstrated *in vitro* with RNA polymerase, may *in vivo* be used to enhance rather than inhibit transcription and polymerase activity by enriching loop formation. Furthermore, these findings highlight even more how the simple strategy of considering individual G4 formation at specific gene promoters might significantly underestimate the potential roles of G4s in the regulation of gene expression.

Beyond the initial extrusion of enhancer-promoter loops in DNA, G4s may further stabilize these loops by promoting the binding and recruitment of key regulatory proteins which inhibit loop collapse. It is possible that this is done

somewhat indirectly as G4s are correlated with more open, accessible chromatin and also reduced DNA methylation both of which promote protein-DNA interactions (59,60). However, various *in vitro* studies have demonstrated the high affinity of multiple regulatory proteins for G4 structures including, eukaryotic transcription factors SP1 (61), MAZ (62) and YY1 (63) and viral regulatory proteins such as the transcription factor ICP4 (64). Additionally, proteins that form part of large transcription-enhancing complexes such as the transcriptional co-activator BRD3 (29,31) and the chromatin regulating protein PARP-1 (65) have been shown to interact strongly with G4s. Within cells, enhancers and promoters containing G4-forming sequences additionally have significantly higher levels of transcription factor binding, which is accompanied by increased levels of RNA pol II occupancy and transcriptional activity (17,20,58). This enrichment is seen even when controlling for differences in chromatin accessibility (58) and the G-richness (20) of a given DNA region suggesting that it is the formation of the G-quadruplex structure itself that is responsible for enhanced transcription, as opposed to a simple correlation with open chromatin or GC content.

These findings support a model of G-quadruplex structures as scaffolds for protein binding, which may promote both local and remote interactions of gene regulation sites (Figure 4D). This hypothesis has been substantiated further with work studying the interactions of G-quadruplexes with the transcription factor ying-yang 1 (YY1), a protein well-established for its importance in creating promoter-

enhancer loops (63). In order to establish more directly the role that G4s may have on transcription factor binding and DNA looping, this study utilized three distinct methods of G4 perturbation and subsequently measured associated changes to gene expression within cells. This included: (i) the over-expression of a G4-helicase which conventionally unwinds G4 secondary structures; (ii) the disruption of G4-forming sequences through subtle CRISPR-mediated gene editing; and (iii) the use of G4-binding ligands to displace native G4-protein interactions.

In each case, perturbation of G4 formation reduced YY1 binding and looping interactions at G4-sites by an order of magnitude. This was additionally accompanied by significantly altered expression of genes not only directly associated with G4 regions through their promoters, but through distal enhancer sites that interact with said genes via loops as visualized with Hi-C experiments (63). This work is an encouraging example of how it is possible to design experiments to test, in a cellular context, the link between G4s and gene expression, the results of which highlight the currently under-studied proposition that G4s may act as long-distance regulators of transcription.

G4s may trigger liquid-liquid phase separation events

In recent years it has been discovered that the expression of genes can be mediated by phase separation events that allow for the concentration of transcriptional machinery within membrane-less organelles inside the nucleus (66,67). Liquid-liquid phase separation (LLPS) relies on a combination of weakly interacting forces between nucleic acids and the low-complexity domains of DNA/RNA-binding proteins (68). This occurs, for instance, in the nucleolus, which is a dense region within the nucleus that contain clusters of ribosomal DNA (69). However, LLPS can also occur outside of the nucleolus in ‘super-enhancer’ regions of the genome characterized by closely localized transcriptional enhancer sequences that bind transcription factors, chromatin remodelling proteins and RNA polymerase II and thus activate transcription (Figure 5A) (70).

This was first demonstrated with the transcriptional co-activator proteins BRD4 (bromodomain-containing protein 4) and MED1 (Mediator 1) which are able to link *via* intermolecular interactions at their disordered domains to create liquid-condensates specifically at super-enhancers (66). The liquid-like state of the BRD4 and MED1 aggregates presents several parallels to membrane-less organelles such as nucleoli and P-granules that are also seen in the nucleus. Additionally, disruption of liquid condensates via treatment with hexanediol, significantly reduced the recruitment of BRD4 and MED1 to super-enhancers (66). As such co-factors are essential for super-enhancer activity, it is thought that the formation of these liquid-condensates may be a unique mechanism to recruit proteins to enhancer regions and achieve transcriptional enhancement.

Although many of the studies considering LLPS in a biological context have historically focused on the protein component of the aggregate, recent work has revealed that certain nucleic acid structures can also act as nucleation sites for phase-separation, even in the absence of proteins.

For instance, protein-free aggregation has been observed *in vitro* for RNA repeat expansions, nucleobase homopolymers and some mRNAs by means of RNA–RNA interactions including base pairing, base stacking and other long-range promiscuous interactions (71). Similarly, G4s have the ability to form networks by connecting multiple nucleic acid strands in an intermolecular configuration or by π -stacking between G-tetrads of different G4s, suggesting that they have the potential to cause phase separation exploiting a similar mechanism (Figure 5B). This was exemplified in work directly associating G4s with LLPS in live cells within short root RNA (72) and the C9orf72 expansion repeat (73). In these studies, it was shown that G4-triggered phase separation is dependent on exposure of the system to G4-favouring conditions, such as increasing potassium concentration, local concentration of nucleic acids and relative guanine content. Furthermore, guanine-rich sequences have been proposed to stimulate the formation of stress granules in the cytoplasm that regulate gene translation, offering yet another piece of evidence to correlate G4 formation to LLPS (74).

The hypothesis that G4s stimulate LLPS by forming intermolecular networks has also been supported by work considering the potential formation of intermolecular G4s within regulatory regions of the genome. For instance, one computational study showed that ‘half-G4’ sequences (containing two instead of four guanine runs) are enriched in promoter and enhancer regions of DNA. It is speculated that when the two guanine runs are brought into close proximity, they are able to hydrogen bond together and thus assemble into a ‘full-G4’, that may mediate promoter-enhancer interactions (75). Similarly, it has been found that regions with a high density of guanine runs are capable of forming quadruplexes over unconventionally long distances and are significantly enriched within super-enhancer regions (76).

Thus, the formation of intermolecular or long-distance G4 structures in such regions may be a mechanism of DNA networking that stimulates LLPS and promotes enhancer activity (Figure 5C). This is particularly relevant in light of our recent findings on the nucleolar protein CSB, which we found to exhibit picomolar affinity for intermolecular G4s whilst displaying negligible binding to intramolecular G4s (77). The characterization of CSB as the first selective intermolecular G4 interactor combined with the known nucleolar localization of this protein suggests that intermolecular G4s may form within the nucleoli and contribute to its phase-separated state, however, this is yet to be explicitly demonstrated.

As such long-range and intermolecular interactions are favoured in crowded environments (78), it is possible that G4s are even more likely to form in the crowded bubbles of liquid condensates (79). Thus, this work brings forward a series of evidence suggesting that G4s might act as nucleation points for LLPS and in turn regulatory elements that may influence transcriptional control. As LLPS offers an interesting mechanism by which the cell increases the local concentration of regulatory proteins and is subsequently able to increase the transcription rate of a given gene, exploring such mechanisms in relation to G4s may yield insightful results. However, investigations into this area are still in their

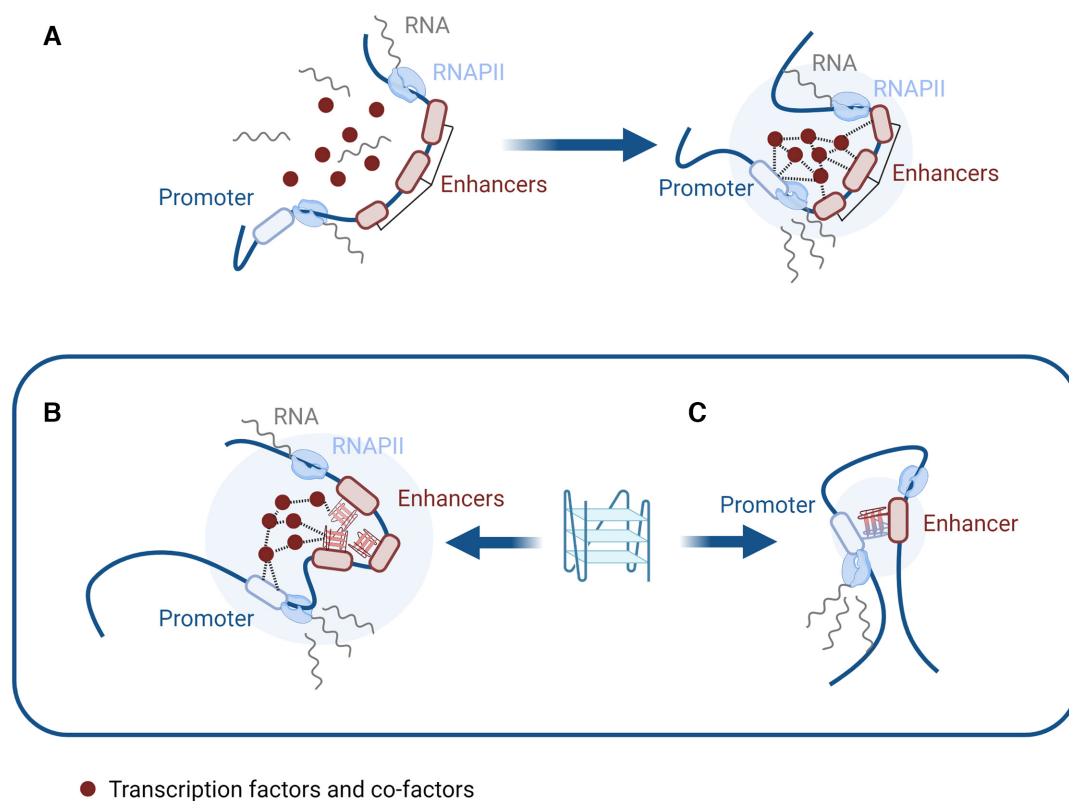


Figure 5. G4s and phase separation. (A) Super-enhancers interact with promoter regions via transcription factors, co-factors and chromatin remodelers triggering LLPS that enhances transcription. (B) G4s at super-enhancer sites may promote aggregation via intermolecular interactions such as π -stacking between quadruplexes. (C) 'half-G4' sequences can assemble intermolecularly to mediate enhancer-promoter interactions and phase separation.

infancy and demonstrate the need for additional work on the diverse roles that physical and chemical changes associated with G4 formation might play in the broader context of gene regulation.

G4s stimulate expression by promoting DNA oxidation and repair

An additional area in which epigenetic control appears at the G4 level can surprisingly be seen when considering how cells respond to environmental stresses. An example of this stress is demonstrated by guanine oxidation induced by reactive oxygen species (ROS) that arise from multiple exogenous and endogenous factors such as exposure to radiation and metabolism (80,81). Among the four DNA bases, guanine has the lowest redox potential, therefore, guanines are one of the most frequent sites to be oxidized by ROS generating 8-oxo-7,8-dihydroguanine (OG) DNA (Figure 6A) (82,83). Consequently G4s, being composed of several runs of guanines, are easy targets for guanine oxidation (84). As oxidizing species can potentially affect many cellular processes, resulting in ageing or cancer (85,86), exposure to ROS is a constant challenge for cells which have developed efficient repair mechanisms to cope with such mutagenic stress (81). However, the stimulation of these DNA damage repair pathways is not always a negative event and may in fact have important implications for transcriptional control particularly in relation to G4s.

This has been explored in recent studies which demonstrated that OG formation in gene promoters containing G4-forming sequences can stimulate DNA repair mechanisms and promote gene activation (87–89). In one study, a luciferase reporter assay was employed to investigate the effects on gene expression when inserting OG bases into the *VEGF* and *NTHL1* promoters (87). Surprisingly, a 300% increase in luciferase expression was measured with the OG plasmid compared to the plasmid without the oxidized guanine and this activation was related to the formation of G4s in these promoters. Under normal circumstances, the presence of an OG would trigger a DNA damage repair pathway known as base excision repair (BER) in which the oxidised base is removed leaving an apurinic (AP) site in the DNA sequence (90). Subsequently, another BER-related protein known as APE1 would cleave the apurinic region so it can be repaired by downstream polymerase action (Figure 6B). However, in this study it was shown that when this oxidation occurs in a G4-forming region that contains a fifth extra run ('spare tyre') of guanines, an alternative G4 structure can form which causes the AP site to be extruded into a loop.

This looped G4 structure is then recognized and bound by APE1, which reduces the cleaving ability of the protein and instead unmasks a separate function of the protein as a recruiter of transcription factors (Figure 6C) (87,91). Indeed, previous data reported APE1 as a multifunctional enzyme which can form part of a large transcriptional complex which includes HIF-1, STAT3 and CBP/p300 leading

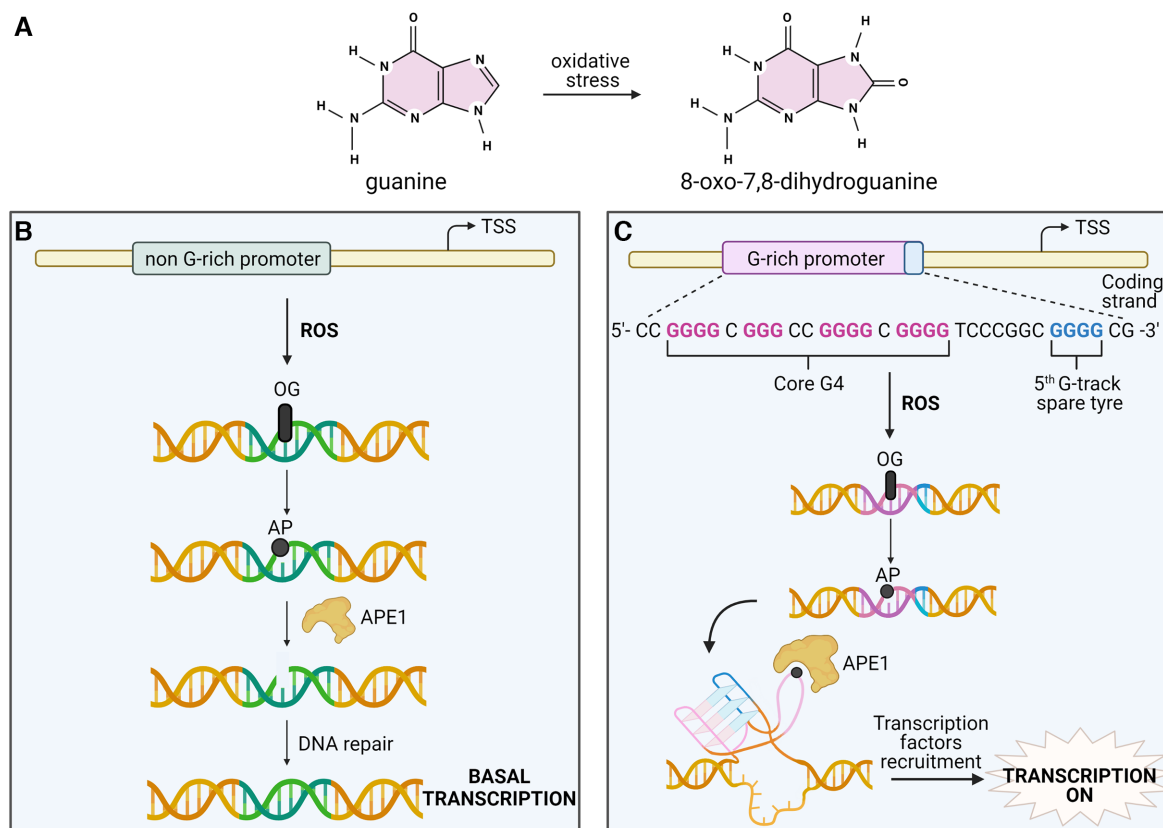


Figure 6. G4s and DNA damage. (A) Guanine is frequently oxidized by oxidative stress generating 8-oxo-7,8-dihydroguanine (OG). (B) OG formation upon ROS damage in a gene promoter without a G4 causes the formation of an apurinic (AP) site that is recognized and cleaved by APE1. This cleaved site is subsequently repaired without increasing the gene transcription level. (C) OG formation in a gene promoter containing a core G4 sequence (in pink) and a fifth G-track (spare tyre in blue) causes the formation of an alternative G4-structure and the extrusion of the AP site into a loop. The structure is recognized and bound by APE1 which reduces the cleaving ability of APE1 and stimulates the recruitment of transcription factors with consequent transcriptional activation.

to gene activation (92,93). They hypothesize that the accumulation of transcription factors through the ref-1 domain of APE1 explains the transcriptional enhancement that is seen only when this alternative G4 forms in vicinity to OG sites and hinders the usual nuclease function of APE1 (87). In support of this, they did not observe increased luciferase expression when the OG damage was inserted in *VEGF* promoter with only four G-tracks, suggesting that the new G4 conformation cannot be formed once the AP is extruded into a loop. Similar results were also observed in subsequent studies on other gene promoters such as *PCNA*, as well as in studies considering the location and strand dependency of this phenomenon, providing additional examples of the role of G4s as regulators of gene expression upon G oxidation (88,89,94).

Whilst mechanistically informative, a limitation of these studies is the reliance on plasmid systems with synthetically inserted G4-rich promoters which may not be representative of the role of G4s in a cellular and chromatin context. Despite this, similar results were observed when considering plasmids transfected into human glioblastoma cells where guanines were oxidized using the cytokine TNF α which naturally stimulates oxidative stress within cells, suggesting that these findings are consistent even when tested under more physiologically relevant conditions (95). Further-

more, a separate analysis considered endogenously formed cellular G4s using G4 ChIP-Seq and showed a significant enrichment of APE1 at G4 sites within human cells (96). In this study they additionally verified that the binding of APE1 to G4s not only stabilizes the G4 structure but increases the residence time of the protein at sites of DNA damage (96), further consolidating a mechanism linking G4 oxidation with transcriptional enhancement within cells.

In addition to this, APE1 is not the only protein that appears to be recruited to OG DNA by G4s: a recent report showed that oxidized quadruplexes can be bound by and enhance Poly ADP-ribose polymerase-1 (PARP-1) activity (97). Similarly to APE1, PARP-1 is an important effector of the BER pathway which recognizes various DNA lesions (98) and seems to respond to the formation of the looped G4 structure that arises during DNA damage. Additionally, ChIP-qPCR studies within pancreatic cancer cells suggested that OG formation promotes the recruitment of transcription factor proteins MAZ and hnRNP A1 to the G4-rich *KRAS* promoter and may promote the transcription process (99).

Altogether this work demonstrates that an apparently mutagenic event such as the oxidation of guanine bases by ROS, in a G4 context, could actually be considered as a way to epigenetically control gene expression. Such DNA dam-

age coincides with relaxed helical tension and is particularly likely to occur in guanine-rich regions (87), thus promoting the formation of G4s which in turn appear to tune the activity and recruitment of transcriptional regulators. In this way, DNA damage in the form of guanine oxidation at G4-sites might represent a novel mechanism by which G4s naturally contribute towards transcriptional control within cells.

CONCLUSIONS

In this review, we have explored a number of pathways by which quadruplex DNA may regulate gene transcription with an emphasis on how G4 formation may actually associate with transcriptional activation. This is in direct contrast to the initial and perhaps outdated perspective where G4s are labelled as direct inhibitors of polymerase activity and transcription. Despite this, the goal of this review is not to argue that G4s act solely as transcriptional activators, as there may be individual cases where G4s are repressive in nature (100), but to critically highlight the substantial evidence showing that G4s are globally correlated with transcriptional enhancement rather than repression. This forces us to assess the way in which we have been investigating G4s in terms of gene regulation.

Early studies considered how individual G4s behave within plasmid systems or relied on ligands that artificially stabilize dynamic G4 structures and may additionally displace natural G4-binding partners. This work contributed to a narrative that depicts G4s as transcriptional roadblocks and might not reflect at all the endogenous role of these structures. Such a narrative is still often re-iterated in the current literature and reflects a simplistic view of G4 formation and transcriptional regulation that we strongly feel needs to be updated in light of recent studies. Going forward, it is important to consider the G4 not as an isolated entity within a specific genomic location, but as a structure that exists as part of an interconnected network of other biomolecules in living cells.

Recent research conducted in this spirit has revealed multiple exciting mechanisms that may explain how G4s act as transcriptional regulators. This includes the role of G4s in mediating the placement of histone marks and in interacting with chromatin remodelling proteins, thus shaping the higher order structure of chromatin. Additionally, the notion that G4s can influence the formation and stability of transcriptional loops including R-loops and long-range enhancer-promoter loops has been demonstrated. Even relatively new mechanisms considering intermolecular G4s as triggers of liquid-liquid phase separation in combination with selective interactors (such as CSB) and DNA-damage induced transcriptional activation are being explored. Altogether these novel mechanisms strongly challenge the simple ‘on/off’ switch role that it is still often associated to G4 formation.

By considering G4s as more than simple knot-like impediments for RNA polymerase to overcome, it will be possible to develop these new avenues of research which may be particularly important for investigations of transcriptional dysfunction in pathologies such as cancer and ageing. It follows that further research in this area may unravel a distinct layer of epigenetic regulation in which G4s are implicated as

multi-faceted regulatory elements in a complex cellular environment.

ACKNOWLEDGEMENTS

All figures were created on Biorender.com
The authors are grateful to Professor Ramon Vilar for helpful discussions on the manuscript.

FUNDING

M.D.A. is grateful to BBSRC that fully supports his research group with a BBSRC David Phillips Fellowship [BB/R011605/1]; J.R. is supported by an EPSRC and NIHR Imperial Biomedical Research Centre scholarship [EP/S023518/1]; F.R. is supported by a Leverhulme Trust, Leverhulme Cellular Bionics scholarship [EP/S023518/1]; S.N. is supported by an EPSRC scholarship [EP/S023518/1]; D.L. is supported by an Imperial College DTP scholarship funded by the Chemistry Department. Funding for open access charge: Imperial College London.

Conflict of interest statement. None declared.

REFERENCES

1. Raguseo, F., Chowdhury, S., Minard, A. and Di Antonio, M. (2020) Chemical-biology approaches to probe DNA and RNA G-quadruplex structures in the genome. *Chem. Commun.*, **56**, 1317–1324.
2. Spiegel, J., Adhikari, S. and Balasubramanian, S. (2020) The structure and function of DNA G-Quadruplexes. *Trends Chem.*, **2**, 123–136.
3. Kwok, C.K. and Merrick, C.J. (2017) G-Quadruplexes: prediction, characterization, and biological application. *Trends Biotechnol.*, **35**, 997–1013.
4. Mukherjee, A.K., Sharma, S. and Chowdhury, S. (2019) Non-duplex G-Quadruplex structures emerge as mediators of epigenetic modifications. *Trends Genet.*, **35**, 129–144.
5. Gellert, M., Lipsett, M.N. and Davies, D.R. (1962) Helix formation by guanylic acid. *Proc. Natl. Acad. Sci. U.S.A.*, **48**, 2013–2018.
6. Parkinson, G.N., Lee, M.P.H. and Neidle, S. (2002) Crystal structure of parallel quadruplexes from human telomeric DNA. *Nature*, **417**, 876–880.
7. Siddiqui-Jain, A., Grand, C.L., Bearss, D.J. and Hurley, L.H. (2002) Direct evidence for a G-quadruplex in a promoter region and its targeting with a small molecule to repress c-MYC transcription. *Proc. Natl. Acad. Sci. U.S.A.*, **99**, 11593–11598.
8. Balasubramanian, S., Hurley, L.H. and Neidle, S. (2011) Targeting G-quadruplexes in gene promoters: A novel anticancer strategy? *Nat. Rev. Drug Discov.*, **10**, 261–275.
9. Huppert, J.L. and Balasubramanian, S. (2007) G-quadruplexes in promoters throughout the human genome. *Nucleic Acids Res.*, **35**, 406–413.
10. Halder, K., Halder, R. and Chowdhury, S. (2009) Genome-wide analysis predicts DNA structural motifs as nucleosome exclusion signals. *Mol. Biosyst.*, **5**, 1703–1712.
11. Eddy, J. and Maizels, N. (2008) Conserved elements with potential to form polymorphic G-quadruplex structures in the first intron of human genes. *Nucleic Acids Res.*, **36**, 1321–1333.
12. Kim, N. (2017) The interplay between G-quadruplex and transcription. *Curr. Med. Chem.*, **26**, 2898–2917.
13. Rodriguez, R., Miller, K.M., Forment, J.V., Bradshaw, C.R., Nikan, M., Britton, S., Oelschlaegel, T., Xhemalce, B., Balasubramanian, S. and Jackson, S.P. (2012) Small-molecule-induced DNA damage identifies alternative DNA structures in human genes. *Nat. Chem. Biol.*, **8**, 301–310.
14. Boddupally, P.V.L., Hahn, S., Beman, C., De, B., Brooks, T.A., Gokhale, V. and Hurley, L.H. (2012) Anticancer activity and cellular repression of c-MYC by the G-quadruplex-stabilizing

- 11-piperazinylquinoline is not dependent on direct targeting of the G-quadruplex in the c-MYC promoter. *J. Med. Chem.*, **55**, 6076–6086.
15. Hänsel-Hertsch, R., Beraldi, D., Lensing, S.V., Marsico, G., Zyner, K., Parry, A., Di Antonio, M., Pike, J., Kimura, H., Narita, M. *et al.* (2016) G-quadruplex structures mark human regulatory chromatin. *Nat. Genet.*, **48**, 1267–1272.
 16. Hänsel-Hertsch, R., Simeone, A., Shea, A., Hui, W.W.I., Zyner, K.G., Marsico, G., Rueda, O.M., Bruna, A., Martin, A., Zhang, X. *et al.* (2020) Landscape of G-quadruplex DNA structural regions in breast cancer. *Nat. Genet.*, **52**, 878–883.
 17. Lago, S., Nadai, M., Cernilogar, F.M., Kazerani, M., Domínguez Moreno, H., Schotta, G. and Richter, S.N. (2021) Promoter G-quadruplexes and transcription factors cooperate to shape the cell type-specific transcriptome. *Nat. Commun.*, **12**, 3885.
 18. Zheng, K.W., Zhang, J.Y., He, Y.De, Gong, J.Y., Wen, C.J., Chen, J.N., Hao, Y.H., Zhao, Y. and Tan, Z. (2020) Detection of genomic G-quadruplexes in living cells using a small artificial protein. *Nucleic Acids Res.*, **48**, 11706–11720.
 19. Di Antonio, M., Ponjavic, A., Radzevičius, A., Ranasinghe, R.T., Catalano, M., Zhang, X., Shen, J., Needham, L.M., Lee, S.F., Klenerman, D. *et al.* (2020) Single-molecule visualization of DNA G-quadruplex formation in live cells. *Nat. Chem.*, **12**, 832–837.
 20. Spiegel, J., Cuesta, S.M., Adhikari, S., Hänsel-Hertsch, R., Tannahill, D. and Balasubramanian, S. (2021) G-quadruplexes are transcription factor binding hubs in human chromatin. *Genome Biol.*, **22**, 117.
 21. Fleming, A.M. and Burrows, C.J. (2020) Interplay of guanine oxidation and G-Quadruplex folding in gene promoters. *J. Am. Chem. Soc.*, **142**, 1115–1136.
 22. Kornberg, R.D. (1974) Chromatin structure: a repeating unit of histones and DNA. *Science*, **184**, 868–871.
 23. Noll, M. (1974) Subunit structure of chromatin. *Nature*, **251**, 249–251.
 24. Jiang, C. and Pugh, B.F. (2009) Nucleosome positioning and gene regulation: advances through genomics. *Nat. Rev. Genet.*, **10**, 161–172.
 25. Bannister, A.J. and Kouzarides, T. (2011) Regulation of chromatin by histone modifications. *Cell Res.*, **21**, 381–395.
 26. Sarkies, P., Reams, C., Simpson, L.J. and Sale, J.E. (2010) Epigenetic instability due to defective replication of structured DNA. *Mol. Cell*, **40**, 703–713.
 27. Schiavone, D., Guilbaud, G., Murat, P., Papadopoulou, C., Sarkies, P., Prioleau, M., Balasubramanian, S. and Sale, J.E. (2014) Determinants of G quadruplex-induced epigenetic instability in REV1-deficient cells. *EMBO J.*, **33**, 2507–2520.
 28. Komůrková, D., Svobodová Kovaříková, A. and Bártová, E. (2021) G-Quadruplex structures colocalize with transcription factories and nuclear speckles surrounded by acetylated and dimethylated histones H3. *Int. J. Mol. Sci.*, **22**, 1995.
 29. Vlasenok, M., Levchenko, O., Basmanov, D., Klinov, D., Varizhuk, A. and Pozmogova, G. (2018) Data set on G4 DNA interactions with human proteins. *Data Br.*, **18**, 348–359.
 30. Lamonica, J.M., Deng, W., Kadauke, S., Campbell, A.E., Gamsjaeger, R., Wang, H., Cheng, Y., Billin, A.N., Hardison, R.C., Mackay, J.P. *et al.* (2011) Bromodomain protein Brd3 associates with acetylated GATA1 to promote its chromatin occupancy at erythroid target genes. *Proc. Natl. Acad. Sci. U.S.A.*, **108**, E159–E168.
 31. Pavlova, I.I., Tsvetkov, V.B., Isaakova, E.A., Severov, V.V., Khomyakova, E.A., Laciš, I.A., Lazarev, V.N., Lagarkova, M.A., Pozmogova, G.E. and Varizhuk, A.M. (2020) Transcription-facilitating histone chaperons interact with genomic and synthetic G4 structures. *Int. J. Biol. Macromol.*, **160**, 1144–1157.
 32. Du, Z., Zhao, Y. and Li, N. (2009) Genome-wide colonization of gene regulatory elements by G4 DNA motifs. *Nucleic Acids Res.*, **37**, 6784–6798.
 33. Selvam, S., Koirala, D., Yu, Z. and Mao, H. (2014) Quantification of topological coupling between DNA superhelicity and G-quadruplex formation. *J. Am. Chem. Soc.*, **136**, 13967–13970.
 34. Belotserkovskii, B.P., Tornaletti, S., D'Souza, A.D. and Hanawalt, P.C. (2018) R-loop generation during transcription: formation, processing and cellular outcomes. *DNA Repair (Amst.)*, **71**, 69–81.
 35. Huppert, J.L. (2008) Thermodynamic prediction of RNA-DNA duplex-forming regions in the human genome. *Mol. Biosyst.*, **4**, 686–691.
 36. Miglietta, G., Russo, M. and Capranico, G. (2020) G-quadruplex-R-loop interactions and the mechanism of anticancer G-quadruplex binders. *Nucleic Acids Res.*, **48**, 11942–11957.
 37. Duquette, M.L., Handa, P., Vincent, J.A., Taylor, A.F. and Maizels, N. (2004) Intracellular transcription of G-rich DNAs induces formation of G-loops, novel structures containing G4 DNA. *Genes Dev.*, **18**, 1618–1629.
 38. Zhao, Y., Zhang, J.Y., Zhang, Z.Y., Tong, T.J., Hao, Y.H. and Tan, Z. (2017) Real-Time detection reveals responsive cotranscriptional formation of persistent intramolecular DNA and intermolecular DNA:RNA Hybrid G-quadruplexes stabilized by R-Loop. *Anal. Chem.*, **89**, 6036–6042.
 39. Zheng, K.W., He, Y.De, Liu, H.H., Li, X.M., Hao, Y.H. and Tan, Z. (2017) Superhelicity constrains a localized and R-loop-dependent formation of G-Quadruplexes at the upstream region of transcription. *ACS Chem. Biol.*, **12**, 2609–2618.
 40. Wanrooij, P.H., Uhler, J.P., Shi, Y., Westerlund, F., Falkenberg, M. and Gustafsson, C.M. (2012) A hybrid G-quadruplex structure formed between RNA and DNA explains the extraordinary stability of the mitochondrial R-loop. *Nucleic Acids Res.*, **40**, 10334–10344.
 41. Zheng, K.W., Xiao, S., Liu, J.Q., Zhang, J.Y., Hao, Y.H. and Tan, Z. (2013) Co-transcriptional formation of DNA: RNA hybrid G-quadruplex and potential function as constitutional cis element for transcription control. *Nucleic Acids Res.*, **41**, 5533–5541.
 42. Tan, J., Wang, X., Phoon, L., Yang, H. and Lan, L. (2020) Resolution of ROS-induced G-quadruplexes and R-loops at transcriptionally active sites is dependent on BLM helicase. *FEBS Lett.*, **594**, 1359–1367.
 43. Lee, C.Y., McNerney, C., Ma, K., Zhao, W., Wang, A. and Myong, S. (2020) R-loop induced G-quadruplex in non-template promotes transcription by successive R-loop formation. *Nat. Commun.*, **11**, 3392.
 44. De Magis, A., Manzo, S.G., Russo, M., Marinello, J., Morigi, R., Sordet, O. and Capranico, G. (2019) DNA damage and genome instability by G-quadruplex ligands are mediated by R loops in human cancer cells. *Proc. Natl. Acad. Sci. U.S.A.*, **116**, 816–825.
 45. Xu, H., Di Antonio, M., McKinney, S., Mathew, V., Ho, B., O'Neil, N.J., Santos, N.D., Silvester, J., Wei, V., Garcia, J. *et al.* (2017) CX-5461 is a DNA G-quadruplex stabilizer with selective lethality in BRCA1/2 deficient tumours. *Nat. Commun.*, **8**, 14432.
 46. Field, A. and Adelman, K. (2020) Evaluating enhancer function and transcription. *Annu. Rev. Biochem.*, **89**, 213–234.
 47. Herz, H.M. (2016) Enhancer deregulation in cancer and other diseases. *Bioessays*, **38**, 1003–1015.
 48. Bradner, J.E., Hnisz, D. and Young, R.A. (2017) Transcriptional addiction in cancer. *Cell*, **168**, 629–643.
 49. Schoenfelder, S. and Fraser, P. (2019) Long-range enhancer–promoter contacts in gene expression control. *Nat. Rev. Genet.*, **20**, 437–455.
 50. Stadhouders, R., van den Heuvel, A., Kolovos, P., Jorna, R., Leslie, K., Grosveld, F. and Soler, E. (2012) Transcription regulation by distal enhancers: Who's in the loop? *Transcription*, **3**, 181.
 51. Haberle, V. and Stark, A. (2018) Eukaryotic core promoters and the functional basis of transcription initiation. *Nat. Rev. Mol. Cell Biol.*, **19**, 621–637.
 52. Beagan, J.A., Pastuzyn, E.D., Fernandez, L.R., Guo, M.H., Feng, K., Titus, K.R., Chandrashekar, H., Shepherd, J.D. and Phillips-Cremins, J.E. (2020) Three-dimensional genome restructuring across timescales of activity-induced neuronal gene expression. *Nat. Neurosci.*, **23**, 707–717.
 53. Higgs, D.R. (2020) Enhancer–promoter interactions and transcription. *Nat. Genet.*, **52**, 470–471.
 54. Dixon, J.R., Selvaraj, S., Yue, F., Kim, A., Li, Y., Shen, Y., Hu, M., Liu, J.S. and Ren, B. (2012) Topological domains in mammalian genomes identified by analysis of chromatin interactions. *Nature*, **485**, 376–380.
 55. Ren, G., Jin, W., Cui, K., Rodriguez, J., Hu, G., Zhang, Z., Larson, D.R. and Zhao, K. (2017) CTCF-mediated enhancer-promoter interaction is a critical regulator of cell-to-cell variation of gene expression. *Mol. Cell*, **67**, 1049–1058.

56. Chien, R., Zeng, W., Kawauchi, S., Bender, M.A., Santos, R., Gregson, H.C., Schmiesing, J.A., Newkirk, D.A., Kong, X., Ball, A.R. *et al.* (2011) Cohesin mediates chromatin interactions that regulate mammalian β -globin expression. *J. Biol. Chem.*, **286**, 17870–17878.
57. Li, Y., Haarhuis, J.H.I., Sedeño Cacciato, A., Oldenkamp, R., van Ruiten, M.S., Willems, L., Teunissen, H., Muir, K.W., de Wit, E., Rowland, B.D. *et al.* (2020) The structural basis for cohesin–CTCF-anchored loops. *Nature*, **578**, 472–476.
58. Hou, Y., Li, F., Zhang, R., Li, S., Liu, H., Qin, Z.S. and Sun, X. (2019) Integrative characterization of G-quadruplexes in the three-dimensional chromatin structure. *Epigenetics*, **14**, 894–911.
59. Bell, A.C. and Felsenfeld, G. (2000) Methylation of a CTCF-dependent boundary controls imprinted expression of the Igf2 gene. *Nature*, **405**, 482–485.
60. Medvedeva, Y.A., Khamis, A.M., Kulakovskiy, I.V., Ba-Alawi, W., Bhuyan, M.S.I., Kawaji, H., Lassmann, T., Harbers, M., Forrest, A.R.R. and Bajic, V.B. (2014) Effects of cytosine methylation on transcription factor binding sites. *BMC Genomics*, **15**, 119.
61. Eun-Ang, R., Ramon, K., Enid, L., Mehran, N. and Balasubramanian, S. (2012) A non-canonical DNA structure is a binding motif for the transcription factor SP1 in vitro. *Nucleic Acids Res.*, **40**, 1499–1508.
62. Cogoi, S., Paramasivam, M., Membrino, A., Yokoyama, K.K. and Xodo, L.E. (2010) The KRAS promoter responds to Myc-associated zinc finger and poly(ADP-ribose) polymerase 1 proteins, which recognize a critical quadruplex-forming GA-element. *J. Biol. Chem.*, **285**, 22003–22016.
63. Li, L., Williams, P., Ren, W., Wang, M.Y., Gao, Z., Miao, W., Huang, M., Song, J. and Wang, Y. (2020) YY1 interacts with guanine quadruplexes to regulate DNA looping and gene expression. *Nat. Chem. Biol.*, **17**, 161–168.
64. Frasson, I., Soldà, P., Nadai, M., Lago, S. and Richter, S.N. (2021) Parallel G-quadruplexes recruit the HSV-1 transcription factor ICP4 to promote viral transcription in herpes virus-infected human cells. *Commun. Biol.*, **4**, 510.
65. Cogoi, S., Paramasivam, M., Spolaore, B. and Xodo, L.E. (2008) Structural polymorphism within a regulatory element of the human KRAS promoter: Formation of G4-DNA recognized by nuclear proteins. *Nucleic Acids Res.*, **36**, 3765–3780.
66. Sabari, B.R., Dall'Agnese, A., Boija, A., Klein, I.A., Coffey, E.L., Shrinivas, K., Abraham, B.J., Hannett, N.M., Zamudio, A.V., Manteiga, J.C. *et al.* (2018) Coactivator condensation at super-enhancers links phase separation and gene control. *Science*, **361**, eaar3958.
67. Boija, A., Klein, I.A., Sabari, B.R., Dall'Agnese, A., Coffey, E.L., Zamudio, A.V., Li, C.H., Shrinivas, K., Manteiga, J.C., Hannett, N.M. *et al.* (2018) Transcription factors activate genes through the phase-separation capacity of their activation domains. *Cell*, **175**, 1842–1855.
68. Wolozin, B. and Ivanov, P. (2019) Stress granules and neurodegeneration. *Nat. Rev. Neurosci.*, **20**, 649–666.
69. McStay, B. (2016) Nucleolar organizer regions: genomic 'dark matter' requiring illumination. *Genes Dev.*, **30**, 1598–1610.
70. Wang, X., Cairns, M.J. and Yan, J. (2019) Super-enhancers in transcriptional regulation and genome organization. *Nucleic Acids Res.*, **47**, 11481–11496.
71. Van Treeck, B. and Parker, R. (2018) Emerging roles for intermolecular RNA-RNA interactions in RNP assemblies. *Cell*, **174**, 791–802.
72. Zhang, Y., Yang, M., Duncan, S., Yang, X., Abdelhamid, M.A.S., Huang, L., Zhang, H., Benfey, P.N., Waller, Z.A.E. and Ding, Y. (2019) G-quadruplex structures trigger RNA phase separation. *Nucleic Acids Res.*, **47**, 11746–11754.
73. Fay, M.M., Anderson, P.J. and Ivanov, P. (2017) ALS/FTD-associated C9ORF72 repeat RNA promotes phase transitions in vitro and in cells. *Cell Rep.*, **21**, 3573–3584.
74. Byrd, A.K., Zybailov, B.L., Maddukuri, L., Gao, J., Marecki, J.C., Jaiswal, M., Bell, M.R., Griffin, W.C., Reed, M.R., Chib, S. *et al.* (2016) Evidence that G-quadruplex DNA accumulates in the cytoplasm and participates in stress granule assembly in response to oxidative stress. *J. Biol. Chem.*, **291**, 18041–18057.
75. Hegyi, H. (2015) Enhancer-promoter interaction facilitated by transiently forming G-quadruplexes. *Sci. Rep.*, **5**, 9165.
76. Williams, J.D., Houserova, D., Johnson, B.R., Dyniewski, B., Berroyer, A., French, H., Barchie, A.A., Bilbrey, D.D., Demeis, J.D., Ghee, K.R. *et al.* (2020) Characterization of long G4-rich enhancer-associated genomic regions engaging in a novel loop:loop 'G4 Kissing' interaction. *Nucleic Acids Res.*, **48**, 5907–5925.
77. Liano, D. and Di Antonio, M. (2021) Cockayne Syndrome B protein selectively interacts and resolves intermolecular 1 DNA G-quadruplex structures. *bioRxiv* doi: <https://doi.org/10.1101/2021.03.25.436565>, 25 March 2021, preprint: not peer reviewed.
78. André, A.A.M. and Spruijt, E. (2020) Liquid–liquid phase separation in crowded environments. *Int. J. Mol. Sci.*, **21**, 5908.
79. Zheng, K., Chen, Z., Hao, Y. and Tan, Z. (2009) Molecular crowding creates an essential environment for the formation of stable G-quadruplexes in long double-stranded DNA. *Nucleic Acids Res.*, **38**, 327–338.
80. Ward, J.F. (2000) Complexity of damage produced by ionizing radiation. In: *Cold Spring Harbor Symposia on Quantitative Biology*. Cold Spring Harbor Laboratory Press, Vol. **65**, pp. 377–382.
81. Markkanen, E. (2017) Not breathing is not an option: how to deal with oxidative DNA damage. *DNA Repair (Amst.)*, **59**, 82–105.
82. Steenken, S. and Jovanovic, S.V. (1997) How easily oxidizable is DNA? One-electron reduction potentials of adenosine and guanosine radicals in aqueous solution. *J. Am. Chem. Soc.*, **119**, 617–618.
83. Cadet, J., Wagner, J.R., Shafirovich, V. and Geacintov, N.E. (2014) One-electron oxidation reactions of purine and pyrimidine bases in cellular DNA. *Int. J. Radiat. Biol.*, **90**, 423–432.
84. Clark, D.W., Phang, T., Edwards, M.G., Geraci, M.W. and Gillespie, M.N. (2012) Promoter G-quadruplex sequences are targets for base oxidation and strand cleavage during hypoxia-induced transcription. *Free Radic. Biol. Med.*, **53**, 51–59.
85. Finkel, T., Serrano, M. and Blasco, M.A. (2007) The common biology of cancer and ageing. *Nature*, **448**, 767–774.
86. Tubbs, A. and Nussenzweig, A. (2017) Endogenous DNA damage as a source of genomic instability in cancer. *Cell*, **168**, 644–656.
87. Fleming, A.M., Ding, Y. and Burrows, C.J. (2017) Oxidative DNA damage is epigenetic by regulating gene transcription via base excision repair. *Proc. Natl. Acad. Sci. U.S.A.*, **114**, 2604–2609.
88. Fleming, A.M., Zhu, J., Ding, Y. and Burrows, C.J. (2019) Location dependence of the transcriptional response of a potential G-quadruplex in gene promoters under oxidative stress. *Nucleic Acids Res.*, **47**, 5049–5060.
89. Fleming, A.M., Zhu, J., Ding, Y. and Burrows, C.J. (2017) 8-Oxo-7, 8-dihydroguanine in the context of a gene promoter G-quadruplex is an on-off switch for transcription. *ACS Chem. Biol.*, **12**, 2417–2426.
90. David, S.S., O'Shea, V.L. and Kundu, S. (2007) Base-excision repair of oxidative DNA damage. *Nature*, **447**, 941–950.
91. Broxson, C., Hayner, J.N., Beckett, J., Bloom, L.B. and Tornaletti, S. (2014) Human AP endonuclease inefficiently removes abasic sites within G4 structures compared to duplex DNA. *Nucleic Acids Res.*, **42**, 7708–7719.
92. Bhakat, K.K., Mantha, A.K. and Mitra, S. (2009) Transcriptional regulatory functions of mammalian AP-endonuclease (APE1/Ref-1), an essential multifunctional protein. *Antioxidants Redox Signal.*, **11**, 621–637.
93. Gray, M.J., Zhang, J., Ellis, L.M., Semenza, G.L., Evans, D.B., Watowich, S.S. and Gallick, G.E. (2005) HIF-1 α , STAT3, CBP/p300 and Ref-1/APE are components of a transcriptional complex that regulates Src-dependent hypoxia-induced expression of VEGF in pancreatic and prostate carcinomas. *Oncogene*, **24**, 3110–3120.
94. Redstone, S.C.J., Fleming, A.M. and Burrows, C.J. (2019) Oxidative modification of the potential G-Quadruplex sequence in the PCNA gene promoter can turn on transcription. *Chem. Res. Toxicol.*, **32**, 437–446.
95. Fleming, A.M., Zhu, J., Howpay Manage, S.A. and Burrows, C.J. (2019) Human NEIL3 gene expression regulated by epigenetic-like oxidative DNA modification. *J. Am. Chem. Soc.*, **141**, 11036–11049.
96. Roychoudhury, S., Pramanik, S., Harris, H.L., Tarpley, M., Sarkar, A., Spagnol, G., Sorgen, P.L., Chowdhury, D., Band, V., Klinkebiel, D. *et al.* (2020) Endogenous oxidized DNA bases and APE1 regulate the formation of G-quadruplex structures in the genome. *Proc. Natl. Acad. Sci. U.S.A.*, **117**, 11409–11420.

97. Edwards,A.D., Marecki,J.C., Byrd,A.K., Gao,J. and Raney,K.D. (2021) G-Quadruplex loops regulate PARP-1 enzymatic activation. *Nucleic Acids Res.*, **49**, 416–431.
98. Reynolds,P., Cooper,S., Lomax,M. and O'Neill,P. (2015) Disruption of PARP1 function inhibits base excision repair of a sub-set of DNA lesions. *Nucleic Acids Res.*, **43**, 4028–4038.
99. Cogoi,S., Ferino,A., Miglietta,G., Pedersen,E.B. and Xodo,L.E. (2018) The regulatory G4 motif of the Kirsten ras (KRAS) gene is sensitive to guanine oxidation: Implications on transcription. *Nucleic Acids Res.*, **46**, 661–676.
100. Matsumoto,K., Okamoto,K., Okabe,S., Fujii,R., Ueda,K., Ohashi,K. and Seimiya,H. (2021) G-quadruplex-forming nucleic acids interact with splicing factor 3B subunit 2 and suppress innate immune gene expression. *Genes Cells*, **26**, 65–82.