



G4-quadruplex-binding proteins: review and insights into selectivity

Vanessa Meier-Stephenson^{1,2,3}

Received: 17 November 2021 / Accepted: 4 April 2022 / Published online: 20 April 2022
© The Author(s) 2022

Abstract

There are over 700,000 putative G4-quadruplexes (G4Qs) in the human genome, found largely in promoter regions, telomeres, and other regions of high regulation. Growing evidence links their presence to functionality in various cellular processes, where cellular proteins interact with them, either stabilizing and/or anchoring upon them, or unwinding them to allow a process to proceed. Interest in understanding and manipulating the plethora of processes regulated by these G4Qs has spawned a new area of small-molecule binder development, with attempts to mimic and block the associated G4-binding protein (G4BP). Despite the growing interest and focus on these G4Qs, there is limited data (in particular, high-resolution structural information), on the nature of these G4Q-G4BP interactions and what makes a G4BP selective to certain G4Qs, if in fact they are at all. This review summarizes the current literature on G4BPs with regards to their interactions with G4Qs, providing groupings for binding mode, drawing conclusions around commonalities and highlighting information on specific interactions where available.

Keywords G4-quadruplexes · Quadruplex-binding proteins · DNA–protein interactions · RNA–protein interactions

Introduction

G4-quadruplexes (G4Q) are guanine-rich secondary nucleic acid structures that form from stacks of the planar orientation of four guanosine residues held together by Hoogsteen bonds and stabilized by metal ions, typically K⁺ ions (Fig. 1). The nucleic acid sequence that has the potential to form these structures is typically described in the formula $G_xN_yG_xN_yG_xN_yG_x$, where $x \geq 2$ guanosine residues and y is 1–7 nucleotides (N). They can occur in both RNA and DNA, and intramolecularly from a single strand of nucleic acid, or with multiple, typically 2 or 4 strands. Arrangements can vary in the nature of their stacking, forming parallel, antiparallel, and various hybrid formations, and while DNA G4Qs in nature can display a variety of these forms, RNA structures tend to be largely parallel (Zhang et al.

2010). Throughout the human genome, there are estimates of > 700,000 potential G4Q sequences (Hansel-Hertsch et al. 2017). Interestingly, these sequences show a high predominance in promoter regions, telomeres, and untranslated regions in mRNA (Hansel-Hertsch et al. 2017; Huppert and Balasubramanian 2005; Rhodes and Lipps 2015; Rigo et al. 2017). They are also found in the promoter regions and untranslated terminal regions (UTRs) of viruses (Fleming et al. 2016; Frasson et al. 2020; Lavezzo et al. 2018; Meier-Stephenson et al. 2021; Metifiot et al. 2014; Perrone et al. 2017). Their cross-species presence at these key regulation areas suggests a likely functional role in the DNA/RNA processing. Indeed, a growing body of evidence clearly links key cellular functions with these G4Q structures, in including transcription, translation, immunoglobulin class switch, and genome stabilization (Da Ros et al. 2021; Dalloul et al. 2018; Lerner et al. 2020; Siddiqui-Jain et al. 2002b; Wolfe et al. 2014). G4Qs have therefore become of great interest to the drug development realm for being able to target the respective downstream processes directed by these G4Qs (Asamitsu et al. 2019; Balasubramanian et al. 2011; McLuckie et al. 2013; Ohnmacht et al. 2015; Tauchi et al. 2006; Xu et al. 2017).

Current G4Q small-molecule binders and dyes appear to either sit atop the G4Q barrel or intercalate with the

✉ Vanessa Meier-Stephenson
meierste@ualberta.ca

¹ Department of Medicine, Division of Infectious Diseases, University of Alberta, Edmonton, AB, Canada

² Department of Medical Microbiology and Immunology, University of Alberta, Edmonton, AB, Canada

³ Li Ka Shing Institute of Virology, University of Alberta, Edmonton, AB, Canada

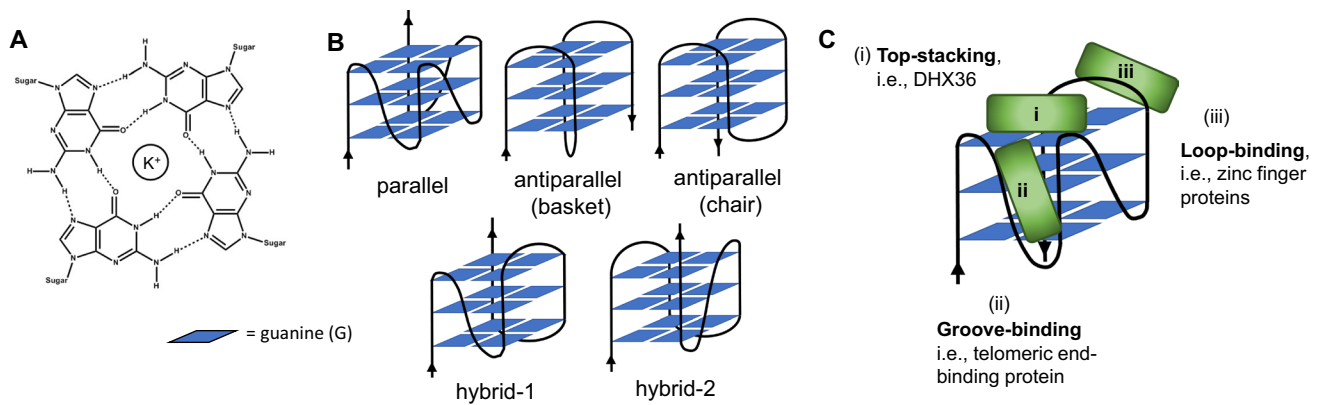


Fig. 1 Schematic of G4-quadruplexes (G4Qs), showing the G-tetrad's planar orientation (A), formed by Hoogsteen bonds and stabilized by a metal ion, typically potassium (K^+), which can stack upon one another in various orientations (B). These structures interact with var-

ious cellular proteins, which may bind in a number of different manners (C), including top-stacking (i), groove-binding (ii), and loop-binding (iii)

common G4Q core (Amato et al. 2018, 2017; Giancola and Pagano 2013; Phan et al. 2005; Sun et al. 2019). While varying affinities have been found for such molecules, there remains the challenge of improving selectivity to minimize off-target effects in the host cell. The very nature and selectivity of cellular proteins that interact with G4Qs imply that selectivity must be attainable. Determining the features of G4Q-binding proteins will be helpful in eliciting the details necessary to rationally design selective binders.

This review aims to assess the binding interactions of known G4-binding proteins (G4BPs) with their respective G4-quadruplexes, highlighting the commonalities in their interaction and subsequently features that would appear to make them selective.

G4-quadruplexes (G4Qs): form and stability

G4Q formation kinetics and stability remain a topic of much research (Grün and Schwalbe 2022; Laouer et al. 2021; Lemkul 2019; Nguyen et al. 2020; Robinson et al. 2021; Rocca et al. 2020; Spiegel et al. 2020; Wu et al. 2020). While readily formed *in vitro* with temperature manipulations, the conditions under which G4Qs form *in vivo* are incompletely understood. The conformational energy landscapes are presumed to have many potential folding intermediates with favorable minima en route to the quadruplex form (Grün and Schwalbe 2022; Rocca et al. 2020). Once formed, the thermodynamic stability of the structure (of an intramolecular three-tetrad G4Q, for this example) is approximately as stable as the helical duplex DNA of similar length (Lane et al. 2008). The conformational forms a G4Q can assume (Fig. 1B) have been shown to influence thermodynamic stability, and while predominant forms are likely for certain G-rich sequences, there are also examples of mixed forms

occurring from the same sequence (Chen and Yang 2012; Hatzakis et al. 2010; Lim et al. 2010). Even a shift along a longer G-rich stretch that changes only loop length can also influence G4Q stability (Hatzakis et al. 2010). Collectively, this suggests on ongoing flux and equilibrium among forms, rather than a fixed state. In comparing RNA and DNA G4Qs, there appears to be slightly greater thermodynamic stability of the RNA G4Q versus its DNA counterpart, likely owing to the additional hydrogen-bonding network enabled by the ribose hydroxyl groups, but always within the same order of magnitude (Arora and Maiti 2009; Joachimi et al. 2009; Zhang et al. 2010).

Formation and degradation of G4Qs have been shown *in vitro* to be dependent on a number of intrinsic factors including loop length, leading nucleic acids and incorporated ions (Bhattacharyya et al. 2016; Chen et al. 2021; Guédin et al. 2010; Hazel et al. 2004; Piazza et al. 2015). While it is possible that *in vivo* G4Qs are influenced by the same factors, there is of course the additional interplay with many cellular proteins, namely, G4-binding proteins whose roles and influences are still being defined.

G4-binding proteins (G4BPs)

There are several ways in which to categorize G4BP: one of which is by their influence on the G4s with which they interact, i.e., stabilizing versus destabilizing, but while there are a plethora of studies describing G4Q-G4BP interactions, the functional verdict remains to be determined for many pairings (Sun et al. 2019). Thus, this review has been organized by sites of binding, namely, G4BPs targeting telomeric, promoter, or RNA G4Qs. While the action of a G4BP may again be varied among groupings, the nature of the initial

physical interactions is presumed to have commonalities and is so described where able.

A preliminary note on G4BP-G4Q-binding modes

There are limited high-resolution structures of G4BPs interacting with G4Qs available, but some generalizations appear to be evolving, as will hopefully become apparent in this review. To highlight these binding modes up front for purposes of providing examples below, binding can occur at the following sites: (i) top-stacking with the upper G-quartets (i.e., atop the G4Q “barrel”), (ii) groove-binding (i.e., between the spaces of the loops), (iii) loop-binding (i.e., binding with the protruding loop nucleotides alone), or a combination of those modes (Fig. 1C). While only one to two concrete examples exist for each, others may become apparent as further biophysical and structural data become available supporting a particular mode. It is also possible that similar protein functionalities may support a particular binding mode (i.e., helicases may act through a stabilizing top-stacking interaction).

G4BPs targeting telomeres

One of the first G4Q-forming sequences identified was in the human telomere (Wang and Patel 1993). Telomeres are the terminal segments of chromosomal DNA and have a highly G-rich 3'-overhang region containing hexameric repeats of TTAGGG (Brázda et al. 2014; O'Sullivan and Karlseder 2010). Functionally, telomeres play a critical role in genome stability, degrading with each cycle of replication. They are maintained by a number of proteins to ensure the process occurs in a well-regulated manner. Many of these proteins have been identified as G4Q-binding proteins, acting on the telomeric G4Qs in various ways, anchoring, degrading, stabilizing (Brázda et al. 2014). Some of these proteins known to bind the G4Qs in this region are Protection of telomeres 1 (POT1) protein, Replication protein A (RPA), human CTC1–STN1–TEN1 (CST), Breast cancer type 1 susceptibility protein (BRCA1), heterogeneous nuclear ribonucleoprotein A1 (hnRNP A1), Bloom and Werner's syndrome proteins, and Preimplantation factor-1 (Pif1).

Protection of telomeres 1 (POT1)

POT1 protein is one of the six members of the telomere shelterin complex, which collectively prevent the telomeric overhang from being recognized as chromosomal damage (O'Sullivan and Karlseder 2010). Studies using single-molecule fluorescence resonance energy transfer (smFRET) were able to show that a monomer of POT1 binds the G4Q specifically, approaching it from the 3'-5' direction, initiating unfolding of the G4Q (Hwang et al. 2012).

It is subsequently followed by binding of a second monomer to complete the process in four sequential steps with the two monomers. Upon binding of the enhancer protein, tripeptidyl-peptidase 1 (TPP1), the POT1-TPP1 complex is thought to promote a sliding activity that unfolds then refolds the G-quadruplex (Hwang et al. 2012). Considered analogous to the POT1-TPP1 complex is the telomeric end binding protein in *Oxytricha nova* (Xin et al. 2007), and its NMR-derived structure is highlighted in Fig. 2 (Horvath and Schultz 2001). The crystal structure contains both an unwound strand of telomeric DNA and an intact G4Q at the border of the crystals' unit cell. The G4Q, as the authors also highlight, appears to be interacting only minimally and may actually be an artifact of crystal packing, but several residues do appear to have the potential, including Lys105 and Asn139 interacting with the phosphate backbone of G4Q while Tyr142 reaches into the groove to interact with several of the guanosines (Fig. 2B) (Horvath and Schultz 2001). Interestingly, the POT1 G4Q binding has been shown to be selective to antiparallel G4Qs, while parallel G4Qs are unaffected (Ray et al. 2014). If in fact groove-binding is its main mode of interaction with a G4Q, this could be supported by components of steric hindrance among the various G4Q folding forms (examples in Fig. 1).

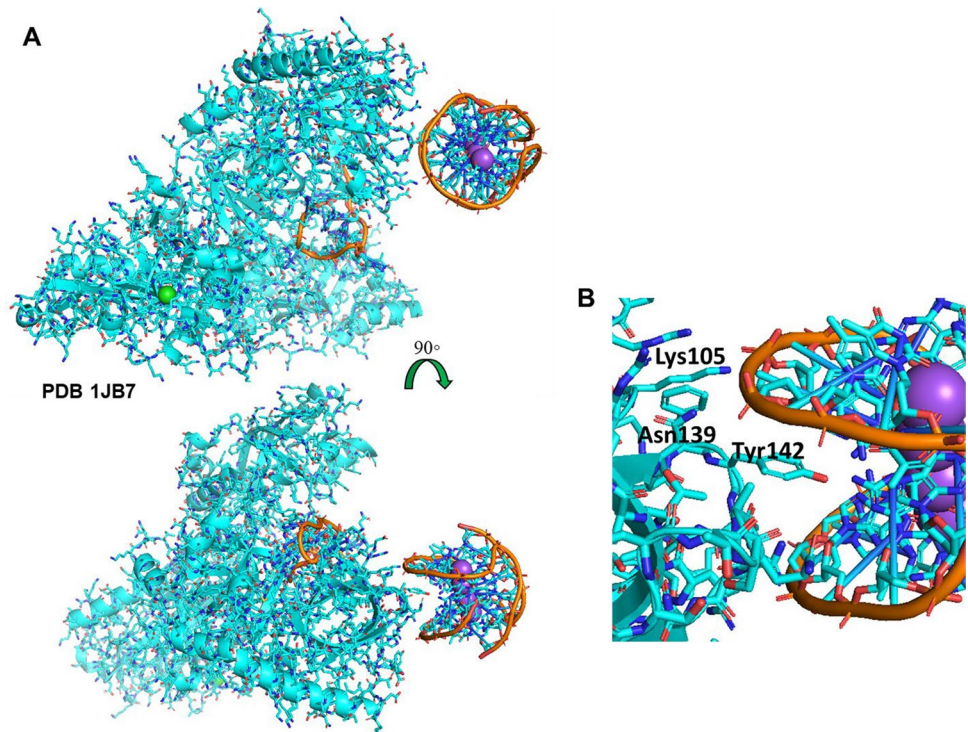
Replication protein A (RPA)

RPA protein is the most abundant single-stranded DNA (ssDNA)-binding protein in human cells (Prakash and Borgstahl 2012). It is a trimeric protein complex with multiple DNA-binding domains, mainly in its first subunit, and is heavily involved in DNA replication, repair, and recombination (Oakley and Patrick 2010; Prakash and Borgstahl 2012; Qureshi et al. 2012). Similar to POT1, RPA acts to unfold telomeric G4Qs. Unlike POT1, it approaches from the 5'-3' direction and is also able to unfold both parallel and antiparallel G4Qs (Ray et al. 2014). Crystal structures exist for the ssDNA complex, but not yet with the G4Q (Bochkarev et al. 1997; Bochkareva et al. 2001, 2002).

Human CST (CTC1–STN1–TEN1)

Human CST is a trimeric complex that plays a number of roles that aid in replication, largely at the replication fork, but also has critical actions at the telomeres, ensuring their stability (Bhattacharjee et al. 2017). CST binds telomeric G4Qs, unwinds them, and facilitates the production of the complementary C-rich strand in a process known as C-strand fill-in (Bhattacharjee et al. 2017). More specifically, the CTC1 is the subunit containing the DNA-binding domain, whereby the TAGG of the repeats is found to be tightly associated with the protein in its unwound form, per cryo-EM studies (Lim et al. 2020).

Fig. 2 Example of groove-binding mode—telomeric end-binding protein of *Oxytricha nova* (OnTEBP), a protozoan analogue of human POT1 protein (PDB 1JB7) broad-view (A) and close-up (B), showing the Tyr142 in proximity to several of the G4Q guanosines and residues Lys105 and Asn139 nearer to the phosphate backbone facilitating H-bonding opportunities



Breast cancer type 1 susceptibility protein (BRCA1)

BRCA1 is a tumor suppressor protein, whereby functional dysregulation from mutations confer a high risk of breast and ovarian cancers (Ford et al. 1994). The exact mechanism of tumor suppression is not fully known, but is heavily involved in DNA repair, cell cycle regulation, and, more recently, regulation of telomeres (Ballal et al. 2009). The latter function has since been supplemented with in vitro data showing binding to G4Qs (Brázda et al. 2016). Thus, there is also a good case for BRCA1 as a direct G4BP; however, further structural data describing this interaction is needed.

Heterogeneous nuclear ribonucleoprotein A1 (hnRNP A1) and Unwinding Protein 1 (UP1)

hnRNP A1 is a ribonucleoprotein involved in RNA transport, alternative splicing, microRNA biogenesis, and transcriptional control (Hudson et al. 2014). The protein has two nucleic acid-binding sites capable of interacting with RNA or DNA. UP1 is the proteolytic cleavage product of hnRNP A1 that retains those two binding domains. These two proteins are able to unfold G4Qs to facilitate telomerase binding at the telomeres to enable telomere lengthening (Hudson et al. 2014). With focus on the UP1, mutational studies on the hexameric repeat region shows a consensus binding sequence of d(nYAGn), where Y is either a thymine or cytosine residue (Ding et al. 1999; Myers et al. 2003).

Crystal structures have been solved of UP1 bound to RNA (Morgan et al. 2015) and ssDNA (Ding et al. 1999), both highlighting the dimerization of the UP1 molecules creating linear channels for the nucleic acids in an antiparallel fashion. How the protein initially binds the telomeric G4Q is not yet clear, but in vitro studies show the UP1 is able to degrade G4Qs and its affinity for single stranded linear nucleic acid far outweighs the structured form, likely adding an energy-favorable driver to its unfolding.

The hnRNP is also able to bind and unfold G4Qs in promoter regions, including the KRAS promoter G4Q and the TRA2 β promoter. The KRAS gene is central to cell-growth signaling pathways, and the presence of the G4Q in its promoter acts as a repressor (Xodo et al. 2008). The hnRNP binding and unfolding enable the transcription to resume, which in the case of some cancers, like pancreatic adenocarcinoma, has detrimental effects (Xodo et al. 2008). Interestingly, researchers have developed a potential therapeutic G4Q decoys to bind this and other G4BPs that may act on this G4Q (Cogoi et al. 2009; Podbevšek and Plavec 2016). Another G4Q promoter target of hnRNP A1 is the TRA2 β exon 2, regulating alternative splicing of exon 2, whereby G4Q binding and subsequent unfolding facilitates inclusion of the exon 2 (Nishikawa et al. 2019). Binding and functional studies are available on these, but there are no detailed interaction studies as of yet.

Regarding hnRNP's role in binding RNA G4Qs, hnRNP A1 was shown to be involved in mRNA transport,

chaperoning the RNA into the cytoplasm and present near translation (von Hacht et al. 2014). This information was determined from pull-down studies using the matrix metalloproteinase MT3-MMP (also known as MMP16) and the actin-related protein 2 (ARPC2) and confirmed with functional and mutational studies (Serikawa et al. 2018; von Hacht et al. 2014).

Bloom protein (BLM)

BLM is a highly conserved RecQ-like helicase that has functions in DNA repair and telomeric stability. The helicase acts in a 3'-5' direction, functioning on the leading strand of replication (Sauer and Paeschke 2017). Unfolding is mediated by two domains, the RecQ C-terminal domain (RQC) and HRDC, which bind and unfold in sequential steps. The process is not necessarily direct, and there can be refolding and unfolding transitions throughout the process (Chatterjee et al. 2014). Unfolding of the G4Q appears to be made easier with a longer trailing segment, at least 6–8-nt long on the 3' end, likely implying binding, or alternatively favorable positioning of the BLM helicase (Budhathoki et al. 2014). It has also been shown that unlike many other helicases, the BLM does not appear to need adenosine triphosphate (ATP) to facilitate its unfolding (Budhathoki et al. 2014).

Werner's syndrome protein (WRN)

WRN is another RecQ-like helicase and a counterpart of BLM. It exerts its helicase function on the lagging strand of replication (Sauer and Paeschke 2017). Binding and unwinding are similarly facilitated by the RQC and HRDC domains, which are conserved throughout this family of helicases (Chatterjee et al. 2014; Lerner and Sale 2019).

Preimplantation factor-1 (Pif1)

Pif1 is a highly conserved helicase across many domains of life and has roles in maintaining genome stability (Paeschke et al. 2013; Zhou et al. 2014). It acts to unwind G4Qs, including telomeric G4Qs, moving in a 5'-3' direction. There has been some conflicting evidence in the literature regarding the nature of this enzyme's helicase functioning, including its ability to rapidly unfold the G4Q but not the dsDNA beyond or that unwinding is slow, but that it does both (Byrd and Raney 2015; Hou et al. 2015; Zhou et al. 2014). A fairly recent study highlights the importance of experimental conditions and shows that while the enzyme is ATP-dependent, some G4Q unwinding can occur in the absence of ATP, perhaps due to transient folding and unfolding of a less stable G4Q (Byrd et al. 2018). This allows the Pif1 to bind the ssDNA can be trapped by the helicase, creating a longer ssDNA overhang that allows more Pif1

enzymes to bind. After a multistep G4Q unwinding, Pif1 is able to proceed with downstream dsDNA unwinding, the rate of which is dependent on the rate of the G4Q unfolding (Byrd et al. 2018; Zhou et al. 2014). There are crystal structures available for Pif1 bound to short DNA oligomers and to dsDNA (Su et al. 2019), but none yet incorporating a G4Q.

G4BPs targeting promoters

Gene promoter regions are another site of G4Q formation and subsequent site of G4Q-binding protein recruitment. The most well-studied are those preceding oncogenes, in particular the c-MYC promoter. The c-MYC promoter and related oncogene is considered one of the master regulators in cancer biogenesis where overexpression upregulates many aspects of cellular growth and metabolism (Chen et al. 2018a; Dang 2012; Miller et al. 2012). Unsurprisingly, there are a number of proteins that bind and regulate this promoter, described in various reviews (Dang 2012; Miller et al. 2012; Wang et al. 2021). Those known specifically to bind G4Qs will be highlighted here, recognizing that this list may not be exhaustive. In addition to the c-MYC promoter, many other promoters containing G4Qs have been identified, some acted upon by the same G4BPs highlighted above. Evidence for the functional roles of G4Qs in these sites is evolving alongside the structural data, as our understanding increases regarding their roles in human diseases. The G4BPs described here are the DEAH box protein 36 (DHX36), non-metastatic factor (NM23-H2), nucleolin, MYC-associated zinc finger (MAZ), Specificity protein 1 (Sp1), Yin Yang 1 (YY1), poly (ADP-ribose) polymerase 1 (PARP1), and the transcriptional helicases, and xeroderma pigmentosum type B and D (XPB and XPD).

D-E-A-H box protein 36 (DHX36; also known as G4 resolvase 1 (G4R1), MLE-like protein 1 (MLEL1), and RNA helicase associated with AU-rich elements (RHAU))

DHX36 is a helicase belonging to the DEAH box family of enzymes and has been shown to bind and unwind both DNA- and RNA-G4Qs (Creacy et al. 2008; Lattmann et al. 2010). Its most well-characterized roles relating to G4Qs are in resolving the G4Qs in RNA, hence one of its aliases; however, there are also studies describing their role in DNA including in the telomeres (Booy et al. 2012; Lattmann et al. 2011; Sexton and Collins 2011) and in promoter regions (Huang et al. 2012).

A co-crystal structure of bovine DHX36 with the c-MYC promoter region G4Q provides insight into this interaction, supporting prior structural data and revealing new information about its binding mode and action (Fig. 3A; PDB

5VHE) (Chen et al. 2018b). First, the helicase is comprised of two RecA-like domains that together create a small positively charged channel big enough to accommodate a single strand of DNA. A further C-terminal domain consists of the oligosaccharide-binding domain (OB) and the G4Q-binding helix, or DHX36-specific motif (DSM). The DSM helix sits atop the G4Q's 5' face, while the OB domain contacts both the G4Q side and trailing single-stranded 3'-terminal DNA. Independent crystallography and NMR studies show that the quadruplex forms a stable, 3-tetrad structure in solution (Fig. 3D) (Ambrus et al. 2005; Heddi et al. 2020; Phan et al. 2004). In the co-crystal structure, however, the 3'-most guanosine trio has been “tugged” downward, resulting in the upper tetrad being formed with neighboring thymine and adenosine as the DNA strand is pulled into the cavity—a description and schematic best viewed in the original paper (Chen et al. 2018b). This allows one to envision the further collapse of the G4Q as more of the nucleotides are pulled through, linearizing the structure.

Focusing on the DSM helix and G4Q interaction in more detail (Fig. 3B), the G4Q-facing residues of the DSM create

a hydrophobic surface with its Ile65, Trp68, Tyr69, and Ala70, enabling stacking with the bases of the upper G4Q tetrad. This binding mode is akin to many small-molecule G4Q binders that have been studied (Collie et al. 2012; Medeiros-Silva et al. 2017; Nielsen et al. 2014; Ohnmacht et al. 2015; Read et al. 2001). An arginine (Arg63) of the helix reaches over to bind one of the phosphate-backbone oxygen. Studies looking at the dissociation constants, K_d , of full-length versus truncated DHX36 in combination with the G4Q report K_d values of < 10 pM and 310 nM, respectively (Booy et al. 2012; Giri et al. 2011; Lattmann et al. 2010). This suggests that at the very least the DSM alone is not the sole contributor to the binding. Indeed, as the co-crystal structure highlights, the OB domain is also involved in this interaction. It contains two loops—one interacts with the sugar phosphate backbone of the G4Q, forming extensive hydrogen bonds (largely with residues Lys860, Asn851, Gly853 and Lys 855; Fig. 3C), and the second loop interacts with the trailing 3'-strand (Chen et al. 2018b). One question that arises is whether the helicase “processes” all single-stranded DNA in this manner or whether it uses the G4Q as a

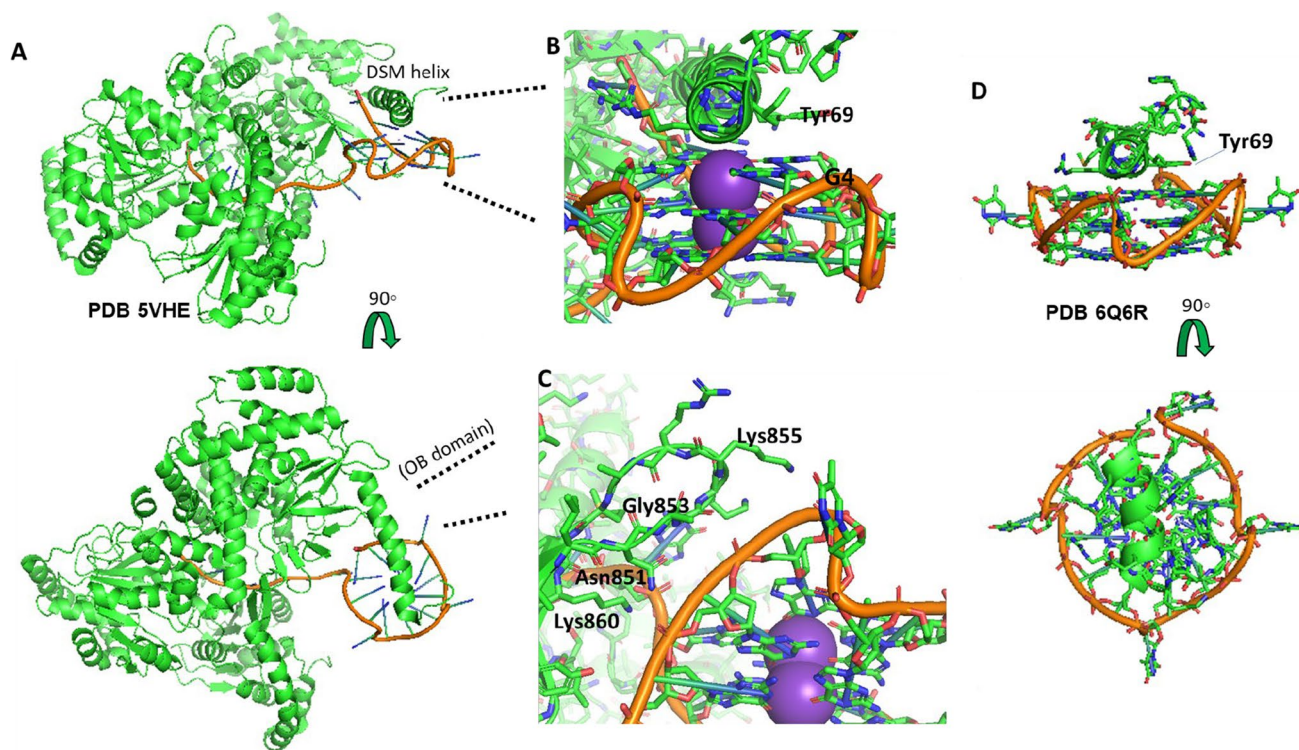


Fig. 3 Example of top-stacking—DHX36 with c-MYC promoter region G4Q. **A** High-level orientation of the structural arrangement showing the DSM helix sitting atop the G4Q, the lateral OB domain loop contacting the G4Q from the side, while the G4Q is pulled through the RecA-like domains (see text; PDB 5VHE); **B** DSM helix showing the Tyr69 oriented parallel with an upper guanosine from the tetrad facilitating π - π stacking. Other hydrophobic residues make up

the remainder of the downward facing helical residues (Ile65, Trp68, and Ala70); **C** OB domain showing the proximity for the extensive hydrogen-bonding network between the phosphate backbone of the G4Q and Lys860, Asn851, Gly853, and Lys 855. **D** Independent study of the DHX36 DSM domain with c-MYC showing similar top-stacking binding mode (PDB 6Q6R)

recognition signal to then attach to the strand and unwind it. Certainly, the K_d of the interaction where a G4Q that is being pulled through a full-length DHX36 helicase will likely be much tighter than that of the helix and G4Q in solution, but it does not give us the initial binding drive, but rather an averaged of what is likely a favorable overall action.

In another study, DHX36 is shown to interact with the Yin Yang 1 (YY1) gene, a gene that encodes for a multifunctional protein involved in tumorigenesis (Gordon et al. 2006). A G4Q exists in both the promoter of the YY1 gene and in the 5'-UTR of the mRNA it produces (Huang et al. 2012). Interestingly in their study, DHX36 was found to bind only to the DNA G4Q and not the RNA, despite both being parallel quadruplexes. This could imply a selectivity that is based on the deoxyribose chain. An alternative interpretation could be the need for trailing 3'-tail to facilitate binding of which was not included in the comparator RNA in this study (Huang et al. 2012). Other studies have also shown DHX36 to selectively bind parallel DNA G4Qs (Heddi et al. 2020; Tippana et al. 2019, 2016), but also that the single-stranded tail plays an important role in that binding (Yangyuru et al. 2018). This, coupled with the co-crystal structure above showing OB domain binding to the ssDNA, further supports this (Fig. 3C).

A more recent and important study involving DHX36 and synthetic G4Qs nicely employs small-molecule Förster resonance energy transfer (smFRET) and computational modeling to describe the process of G4Q unwinding in both DNA and RNA, which has distinct differences (Tippana et al. 2019). Again, while the DHX36 binding to G4Q DNA facilitates unfolding independent of ATP, in a presumed “tugging” and G4Q rearrangement, DHX36 binding to RNA G4Qs enable a stably unfolded state, which is followed by repetitive cycling of refolding in an ATP-dependent manner. The work was also complemented by mutational studies that show rapid dissociation of the complex with ATP hydrolysis when the DSM helix was mutated. This was in contrast to mutations in the sites responsible for binding the 3'-trailing strand (the RecA2 and OB domains), which resulted in erratic DHX36 motions that was unable to facilitate the G4Q refolding, but that maintained binding (Tippana et al. 2019). Another interesting distinction the authors also note with one of their mutations, Y69A in the DSM helix, is that the DNA G4Q readily washed off in one of its processing steps, while the same mutant RNA G4Q held fast (Tippana et al. 2019). This could imply additional stabilizing bonds with the hydroxyl groups in backbone of the RNA, or at the very least, an altered binding mode.

Non-metastatic factor (NM23-H2)

Non-metastatic factor (NM23-H2; nucleoside diphosphate kinase) is a transcriptional factor that binds the nuclease

hypersensitive element (NHE) in the c-MYC promoter that acts as a transcriptional silencer (Sengupta et al. 2019). Using chromatin immunoprecipitation, reporter assays, and FRET studies, NM23-H2 has been shown to bind and resolve this G4Q enabling transcription to proceed (Thakur et al. 2009). Importantly, NM23-H2 is also one of several G4BPs that contains an arginine-glycine-glycine (RGG) motif that is presumed to play a role in G4Q recognition and/or binding (see “Commonalities amongst G4BPs linked to binding modes”).

Nucleolin

Nucleolin is a multifunctional phosphoprotein found throughout the cell, but in high abundance in the nucleolus, for which it is named. Its many roles include ribosomal biogenesis, chromatin remodeling, transcription, and apoptosis (Ginisty et al. 1999; Mongelard and Bouvet 2007; Tajrishi et al. 2011). It is also well-known to bind G4Qs, including the c-MYC promoter G4Q, where binding, unlike with NM23-H2, stabilizes the G4Q leading to reduced transcription (González et al. 2009). Nucleolin binding has been shown in vitro with the G4Qs of the bcl-2, hTERT, VEGF, RET, PDGF-A, and c-kit promoters but the functional significance of these interactions has not yet been elucidated (González et al. 2009; Lago et al. 2017). It has also been shown to bind the hexanucleotide repeat expansion (HRE) of the *C9orf72* gene that is linked to the neurodegenerative diseases, amyotrophic lateral sclerosis (ALS), and frontotemporal dementia (FTD) (Haeusler et al. 2014). Where normally the nucleolin binds these repeats (typically 2–8 repeats of GGGGCC), in those affected, repeats can number in the thousands resulting in nucleolar dysfunction (Haeusler et al. 2014).

Regarding the nature of its binding, nucleolin has been shown to have a preference for parallel G4Qs over anti-parallel ones (González et al. 2009). Further insights into its binding with the c-MYC G4Q are gained from the studies looking at loop length mutational analysis (González et al. 2009; Lago et al. 2017). Here, a series of G4Qs with varying loop lengths reveal that while the shorter loop lengths create a more stable G4Q, nucleolin prefers to bind G4Qs with somewhat lower stability, and in particular those with at least one long loop (3–7 nt) (Lago et al. 2017).

Nucleolin, like NM23-H2, is also an RGG-motif-containing G4BP. Nucleolin was actually the first protein to have this identified whereby a repeating RGG motif (9 repeats) is found in the protein's C-terminal region (Hanakahi et al. 1999). There is also an additional RNA recognition motif (RRM) upstream of this repeating unit that was known to bind single-stranded DNA, but also more recently shown to independently bind G4Q DNA (Hanakahi et al. 1997, 1999). Together, the two regions within nucleolin are hypothesized

to be able to help fold and stabilize the G4Q. The existence of this common RGG motif in other enzymes known to bind G4Qs, namely, hnRNP A1 (Ding et al. 1999) and NM23-H2 (Dexheimer et al. 2009), has laid highlight to these motifs as recognition domains; however, the specifics of their interaction has yet to be fully elucidated (see “Commonalities amongst G4BPs linked to binding modes”).

MYC-associated zinc finger (MAZ)

MAZ is a six zinc finger protein essential to transcription and chromosomal organization (Xiao et al. 2021). For each of these functions, MAZ binds G-rich, G4Q forming regions. In the C-MYC promoter, a G4Q represses transcription and can be unwound by the MAZ protein to derepress it (Siddiqui-Jain et al. 2002a). Similarly, in the HRAS promoter, there are two upstream quadruplexes that maintain transcriptional repression, which when bound and unwound by MAZ are able to proceed with their downstream processing (Cogoi et al. 2014; Membrino et al. 2011). MAZ can also bind and unwind the Pur-1 G4Q of the IDDM2 locus to enable transcription (Lew et al. 2000). Interestingly, this G4Q-unwinding action of MAZ is unique thus far to this zinc finger protein, as other zinc finger proteins do not yet claim this action against G4Qs.

Furthermore, MAZ has also been shown to bind the upstream G4Q of the KRAS promoter enabling transcription. This insight has led to unique approach of developing G4Q-decoys for the zinc finger in pancreatic cancer cells (Cogoi et al. 2010, 2013). Despite its many roles, a structural representation of MAZ with one of its G4Qs has yet to be documented.

Specificity protein 1 (Sp1)

Specificity protein 1 (Sp1) is a ubiquitous transcription factor and member of the Krüppel-like family, which are a group of transcriptional regulators containing a triple zinc finger motif (Black et al. 2001; Vizcaino et al. 2015). Binding sites are variable including GC-boxes and G-rich quadruplex forming regions (Raiber et al. 2012). In vitro evidence exists suggesting that all three of the zinc fingers contribute to the DNA interaction, which has been a notable feature of zinc finger proteins that has also been useful in the design of synthetic zinc fingers (Al-Naama et al. 2020; Eom et al. 2016; Jamieson et al. 2003; Razin et al. 2012; Yokono et al. 1998). These manipulable features are for its interaction with *linear* DNA, however, as little structural data exists with Sp1 in combination with the G4Q form as of yet. It is beyond the scope of this review, but it may be interesting to review the details of prior such zinc finger interaction studies to determine whether the experimental conditions of each would have supported the formation of G4Qs and

thus have the potential to infer results on these structures. Of the promoter regions Sp1 is known to interact and regulate, those containing G4Qs at the site of Sp1 binding are c-KIT (Da Ros et al. 2021), HRAS (Cogoi et al. 2014; Membrino et al. 2011), and VEGF (Yokono et al. 1998). In the case of HRAS, Sp1 binding to its G4Q promoter acts as a transcriptional repressor and seems to also require MAZ (another G4Q-binding zinc finger protein, noted above) for its action (Membrino et al. 2011).

Yin Yang 1 (YY1)

Yin Yang 1 is yet another zinc finger protein recently found to target G4Qs (Li et al. 2021). Interestingly, it itself is under the control of a G4Q (see section on DHX36) (Huang et al. 2012). This zinc finger protein has integral roles in transcription, acting as a regulator of promoter-enhancer loops (Gordon et al. 2006; Li et al. 2021; Weintraub et al. 2017). A recent study nicely employs a mix of proximity ligation assays and chromatin immunoprecipitation assays to show how YY1 can bind two sites of DNA G4Qs and through dimerization bring the two pieces of DNA together, supporting its role in the positioning of promoter and enhancer sites (Li et al. 2021). Regarding its mode of binding, it was noted that using the G4Q stabilizers, TMPyP4 and pyridostatin, G4Q-binding and promoter-enhancer loop formation was greatly reduced, supporting the likely possibility that there are at least some shared points of contact in their respective interactions (Li et al. 2021).

Poly (ADP-ribose) polymerase 1 (PARP1)

PARP1 is a ubiquitous nuclear zinc protein that has a high affinity for damaged DNA and has key roles in chromatin remodeling and gene expression. It has been shown in vitro to bind the G4Qs in the promoter regions of oncogenes c-KIT, c-MYC, and KRAS (Cogoi et al. 2010; Edwards et al. 2020; Fekete et al. 2012; Soldatenkov et al. 2008). Functionally, binding at c-KIT results in catalytic activation of PARP1 (Soldatenkov et al. 2008). Evidence of gene activation upon PARP1 binding also exists for the G4Q upstream from KRAS (Cogoi et al. 2010). Furthermore, studies examining PARP1's specificity for certain G4Qs revealed the preference for parallel G4Qs, as evidenced by lack of binding to the hybrid G4Q-forming hTEL promoter region (Edwards et al. 2020). The same group considered the loop specificity as part of the incorporation, revealing that while a shorter loop enabled the G4Q bind with greater affinity to PARP1, the pentanucleotide loop in its original form (CGAGC) was required for activation of PARP1, indicating a role and incorporation of this lateral G4Q loop into the overall functional binding epitope (Edwards et al. 2020). There are X-ray- and NMR-based models of the PARP1

protein in combination with helical DNA and strand breaks (Ali et al. 2012; Bilokapic et al. 2020; Eustermann et al. 2015; Langelier et al. 2011, 2012; Patel et al. 2014), but none yet exists of PARP1 in complex with a quadruplex. Speculation would lead one to believe the same region would be responsible, but how the protein conforms to do so is still unknown.

Transcriptional helicases xeroderma pigmentosum type B and D (XPB and XPD)

XPB and XPD are both helicase components of the transcription factor II human (TFIIH) complex, an eleven-subunit transcriptional complex that plays a central role in transcription and nucleotide excision repair (NER) (Drapkin et al. 1994; Gray et al. 2014). The XPD helicase has been shown to unwind G4Qs in a 5'-3' direction in an ATP-dependent manner; the XPB helicase acts in a 3'-5' direction has been shown only to bind, but not unwind the G4Q (Gray et al. 2014). While there is a crystal structure available of the XPD Arch domain with the protein MAT1 (Peisert et al. 2020), there is no complex with DNA.

G4BPs involved in DNA replication

Fanconi Anemia Complementation group J (FANCI, also known as BRIP1 or BACH1)

FANCI belongs to the XPD-like group of helicases and facilitates a number of processes including DNA replication, homologous recombination (HR), and interstrand DNA crosslink (ICL) repair (Wu and Spies 2016). Its actions on DNA replication are facilitated through the binding and unwinding of G4Q DNA in the 5'-3' direction. The helicase appears to be critical to the process, such that in its absence, DNA replication persistently stalls at the G4Q site (Castillo Bosch et al. 2014). FANCI has been shown to unwind both intramolecular and intermolecular G-quadruplexes in vitro (Bharti et al. 2013). While it is unclear how broad the scope of G4Q binding may be for this helicase, binding studies have been performed using the telomeric G4Q ((TTAGGG)_n) and a panel of uni-, bi-, and tetramolecular G4Qs (Bharti et al. 2013; Lowran et al. 2020). The key residues of FANCI involved in the interaction are the alanine-alanine-lysine-glutamine (AAKQ) motif. Furthermore, recent results suggest the AAKQ interaction may be with the TTA loop region of the G4Q (Lowran et al. 2020).

There is a homologous G4BP found in *Caenorhabditis elegans* known as Dog-1 that supports the FANCI helicase results, but also noted here because it was also one of the first helicases shown to interact with G4Qs in vivo (Cheung et al. 2002; Kruisselbrink et al. 2008). It similarly has

critical roles in DNA repair, namely, ICL, and like FANCI, its actions on G4Qs are in the 5'-3' direction. From the above-noted mutational studies, there also does not appear to be a specific sequence or G4Q form this enzyme preferentially binds, fitting with a non-specific DNA G4Q helicase.

G4BPs targeting RNA

RNA G4Qs, like their DNA counterparts, also play roles in critical cellular processes, including termination of transcription, telomerase activity, alternative splicing, and regulation of translation (Kharel et al. 2020; Lyu et al. 2021; Song et al. 2016). Structurally, RNA G4Qs are similar to those of DNA, but are much more thermodynamically stable, owing to the more extensive network of hydrogen bonding afforded from the ribose C2' hydroxyl groups (Arora and Maiti 2009; Joachimi et al. 2009; Zhang et al. 2011, 2010) (noted above). Another feature of naturally occurring RNA G4Qs is that they are essentially all parallel G4Qs (with few exceptions (Xiao et al. 2017)). That being said, the structural similarities between RNA and DNA G4Qs still hold, it is not surprising to find that many of the proteins able to bind DNA G4Qs are also able to bind RNA G4Qs. Of those G4BPs already described above, those with identified binding and/or roles in both include one of the most well-studied RNA G4BPs, DHX36, and nucleolin and the hnRNPs. Other RNA G4BPs include DHX9, the heterogeneous nuclear ribonucleoproteins (hnRNP), serine/arginine-rich splicing factors (SRSF), the AF4(ALL1-fused gene from chromosome 4)/FMR2 (fragile X mental retardation 2) (AFF) family of proteins and, of course, ribosomal proteins.

D-E-A-H box protein 9 (DHX9; also known as nuclear DNA helicase II (NDH II) and RNA helicase A (RHA))

DHX9 is an ATP-dependent helicase, similar to DHX36 described above, belonging to the SF2 superfamily of helicases (Chakraborty and Grosse 2011). It also has the ability to unwind G4Qs, with a preference for RNA substrates (Murat et al. 2018). A transcriptome-wide analysis reveals the importance of DHX9 (and DHX36) in the translational control of many transcripts, whereby the G4Qs will cause ribosomal queuing and potential alternate translations if not unwound (Murat et al. 2018). While DHX36 has the DSM helix providing the top-stacking interaction, DHX9 does not have a similar motif. It does contain an RGG box motif in its C-terminus, however, which may be able to facilitate G4Q binding as has been seen in several other G4BPs (Chakraborty and Grosse 2011). DHX36 also has an RGG-box region in its N-terminus, but this region does not appear to have interaction with the G4Q (Fig. 2), though

the N-terminal region is often truncated in many studies on the enzyme (Chen et al. 2018b).

Serine/arginine-rich splicing factors (SRSF)

SRSFs are a highly conserved family of RNA binding proteins involved in alternative and constitutive splicing, mRNA transport, translation, mRNA decay, and genome stabilization (Long and Caceres 2008). They contain an N-terminal RNA recognition motif (RRM) and C-terminal arginine-serine (RS)-domains that appears to have a strong regulatory role in the proteins' functioning, whereby phosphorylation of these domains can enhance binding to other RS-containing splicing proteins bringing them together (Lin and Fu 2007; Long and Caceres 2008). SRSF was originally identified to bind G4Q RNA from pull-down studies (noted above in hnRNP section) using MMP16 and ARPC2 (Serikawa et al. 2018; von Hacht et al. 2014). Further structural information is not yet available, nor whether G4Q-binding plays a specific role in the functional aspects of these groups of proteins.

Fragile X Mental Retardation Protein (FMRP)

Fragile X Mental Retardation Protein (FMRP) is an RNA regulatory protein belonging to the AF4 (ALL1-fused gene from chromosome 4)/FMR2 (fragile X mental retardation 2) (AFF) family of proteins that chaperones mRNA from the nucleus and has roles in transcriptional control. Its malfunction is linked to several human disorders, including Fragile X syndrome and autism (Darnell et al. 2001; Phan et al. 2011). FMRP binds G4Qs via its RGG motif and this alone appears to be necessary, at least for binding to sequence clone 1 (*sc1*) RNA, per extensive mutational studies [*sc1* was selected from a pool of random RNA oligomers (*sc*'s) to bind FMRP] (Darnell et al. 2001). Both crystal structure and NMR data are available on this motif's interaction with the G4Q in *sc1*, which shows the RGG motif forming a β -turn with 13 of the peptide's amino acids and fitting nicely into the junction at the base of the G4Q and the duplex RNA (Fig. 4—PDB 5DE5) (Phan et al. 2011; Vasilyev et al. 2015). In this structure, the RGG motif bound to two consecutive G-C base pairs of the duplex, with its Arg8, Arg10 and Gly11 and the type I β -turn (with intramolecular hydrogen bond between Gly12 and Arg15) extends upward towards the G4Q. Its furthest most reaching residue, Arg15 however, only interacts with the G7 and A17 nucleotides, neither of which are part of the specific G4Q structure (Fig. 4). Just the same, it is speculated that the stabilization afforded through this loop binding in the junction between the duplex and G4Q may be a driver of stability (Vasilyev et al. 2015).

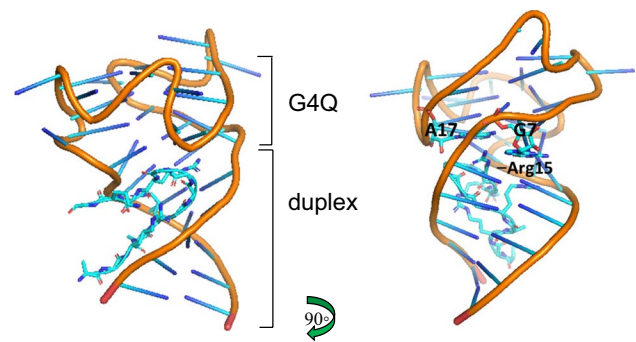


Fig. 4 FMRP's 13-amino acid β turn folding into the groove at the junction of duplex and G4Q DNA (PDB 5DE5). The uppermost amino acid, Arg15, interacts with G7 and A17 nucleotides, which are not part of the G4Q structure. Binding is thought to promote stabilization of the G4Q (see text)

Ribosomal proteins

G4Qs are found in the 5'-UTR of many mRNAs. From the MMP16 and ARPC2 pulldown studies described above, ribosomal proteins were identified from both the 60S and 40S subunits (von Hacht et al. 2014). G4Qs in the 5'-UTRs are thought to act as more of a translational regulator, slowing down protein translation while the ribosome waits for a helicase to unwind the structure (Bugaut and Balasubramanian 2012; Huppert et al. 2008). It may be possible that the recognition of the G4Qs by ribosomal proteins plays further roles, but this has yet to be determined.

G4BPs that degrade G4 DNA

Research on G4BPs with the ability to degrade G4 DNA remains an evolving field, however, the two of which we have the most information are the human nuclease, GQN1 (G quartet nuclease 1) and the *Saccharomyces cerevisiae* Mre11 protein (ScMre11p) (Ghosal and Muniyappa 2005, 2007; Sun et al. 2001). These proteins are nucleases that appear to use the G4 as an anchor for cleavage of the G4 DNA at a site nearby, degrading the single strand.

GQN1 cuts the single-stranded DNA region 2–5 nt 5' of the barrel formed by stacked G-quartets and is independent of the upstream sequence. The nuclease does not degrade duplex or single-stranded DNA or G4 RNA (Sun et al. 2001). Another nuclease is the *Saccharomyces cerevisiae* Mre11 protein (ScMre11p), which also shows high binding affinity for G4 DNA over single- or double-stranded DNA. In this case, the binding of ScMre11p to G4 DNA facilitates endonucleolytic cleavage at G residues that flank the G-quartets (Ghosal and Muniyappa 2005, 2007).

There are currently no enzymes known to be able to degrade the G4Q structure in its wound form.

Commonalities amongst G4BPs linked to binding modes

The different functions G4BPs have are likely dictated or at least heavily linked to the binding mode the protein assumes in interacting with a G4Q. G4BPs that unwind a multitude of different G4Qs, such as DHX36, cannot be too selective toward G4Qs. Thus, a top-stacking binding mode would appear to be more practical—stabilizing the stacked column of G4-tetrads while another part of protein unwinds the G4Q like a spool (Fig. 3).

A subset of G4BPs are the groove-binders, such as nucleolin or FMRP, whereby the presence of an RGG motif, shown to fold into a β -loop, is able to fit into a groove of the G4Q and act as a stabilizer (Masuzawa and Oyoshi 2020; Vasilyev et al. 2015). The presence of the RGG motif alone is not enough to define its function, however, as other G4BPs contain this region (i.e., hnRNP A1 and NM23-H2) and these proteins have differing roles from stabilizing to unwinding (Dexheimer et al. 2009; Ding et al. 1999; Huang et al. 2018; Thandapani et al. 2013). From the stabilization perspective, there at least appears to be some support in the FMRP data (Fig. 4; Phan et al. 2011; Vasilyev et al. 2015)). Given the abundant links with the repeating RGG-motif in many highly regulatory proteins, there is much still to explore with how this contributes to the interaction and functioning around the G4Q.

A comment on zinc finger protein binding

Zinc finger proteins (ZFP) are an interesting subset of G4BPs that warrant highlighting separately. There are currently no experimentally derived high-resolution structures of a ZFP bound to a G4Q. There is one very nice combination biophysics-computational modelling study of a synthetic ZFP, Gq1, a zinc finger protein originally derived from a phage display library coupled with a systematic series of biophysical binding assays (Isalan et al. 2001; Ladame et al. 2006; Patel et al. 2004). In the study, researchers compared binding of Gq1 and the yeast zinc finger homologue, Zif268 to telomeric G4Q or Zif268's (duplex) dsDNA binding sequence. In particular, they systematically swapped out the “fingers” of one for the other to determine which had the greatest effects on G4Q affinity and selectivity. Their results showed that any one finger of Gq1 could be replaced with the corresponding finger of Zif268, without losing quadruplex affinity or duplex discrimination, but when two fingers were exchanged, where one was the second finger, both tenfold reductions in binding and loss of G4Q-duplex

discrimination resulted (Ladame et al. 2006). The modelling studies go on to support the biophysical data, providing a resultant lowest energy coordinate file, whose arrangement is outlined in Fig. 5. Here, the model supports more of the loop-binding mode of interaction. The first of the “fingers” binds with two of the three protruding nucleotides (T12 and A13), while the second “finger” binds the third loop nucleotide (T11) and simultaneously holds the phosphate backbone (Fig. 5A, B). The third “finger” focusses on the neighboring loop, extensively binding the phosphate backbone of the outpouching loop (Fig. 5C).

This work provides us with some of the first supporting data for how this class of proteins may in fact interact with these structures.

G4Q loops affording the basis for selectivity?

With its many critical cellular roles, it would follow logically that G4Qs must have some degree of selectivity to bind the correct G4BPs. This concept has been demonstrated in many studies of G4Q-G4BP interactions (Heddi et al. 2020; Huang et al. 2012; Lago et al. 2017). With the barrel core is similar amongst G4Qs, it is the orientation and length of the loops that make one G4Q different from the next and likely the mechanism by which selectivity is afforded.

G4Q stability is inversely linked to loop length—the shorter the loop, the more stable the G4Q. Guedin et al. performed a nice systematic assessment of this whereby they studied the melting temperatures (T_m) of a large panel of G4Q with varying loop lengths, showing that a length greater than 9 nt had a destabilizing effect on the structure, especially if multiple longer length loops were incorporated (Guédin et al. 2010). The c-MYC promoter G4Q is an example of a highly stable, short loop-length G4Q, while the TRA2 β has a single 8-nt loop (Table 1). It is possible that only some types of helicases are able to overcome a more stably-folded G4Q and be identified as a binding protein, such as that of the DHX36 helicase and its action on the c-MYC, whereby crystal structure evidence suggests it pulls the G4Q through its RecA-like channel (Fig. 3A; PDB 5VHE) (Chen et al. 2018b).

The loops of G4Q-forming regions have also been shown to influence the type of G4Q able to form (i.e., parallel vs anti-parallel) (Bugaut and Balasubramanian 2008; Hazel et al. 2004; Tippiana et al. 2014). G4BPs appear to have preference for some arrangements over others. For example, DHX36 appears to favor G4Qs with a parallel conformation (but able to bind anti-parallel), and if in fact its binding mode is from the top of the G4Q barrel, this could be reasoned as a less sterically hindered arrangement for this form of binding. Conversely, nucleolin while again favoring parallel G4Qs,

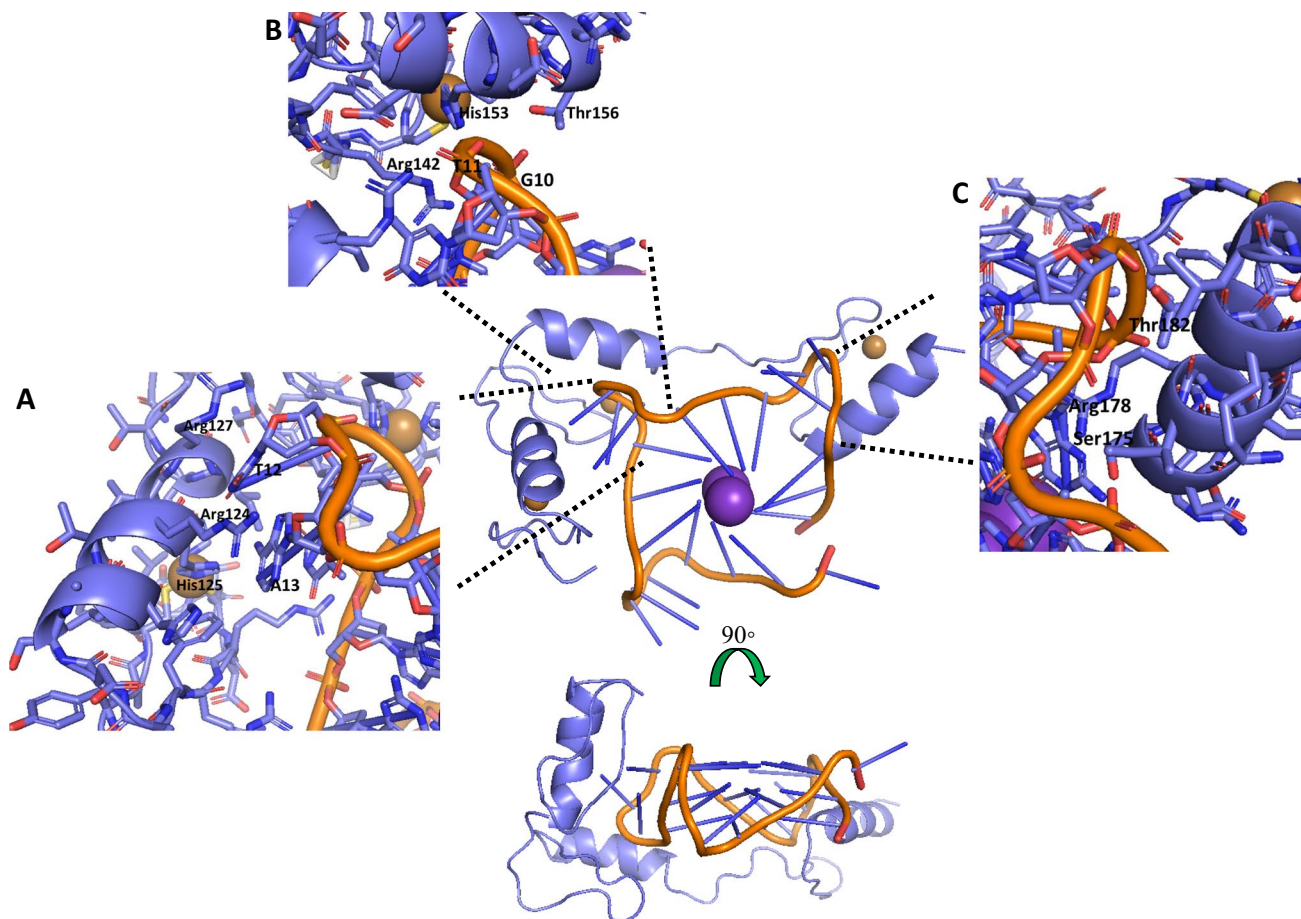


Fig. 5 Example of loop-binding mode—synthetic zinc finger, Gq1 targeting a telomeric G4Q. Computationally derived model (PDB from Ladame, et al. 2006), showing overall arrangement (center) and key residues from **A** the first “finger,” whereby His125, Arg124, and Arg127 interact with the two outward-directed T12 and A13 nucleo-

tides; **B** the second “finger,” where His153 and Thr156 bind with phosphate backbone of G10, while Arg142 wraps under to bind the other protruding nucleotide, T11; and **C** the third “finger,” where the Ser175, Arg178, and Thr182 create extensive H-bonds with the phosphate backbone of the loop

has a strong preference for long-looped G4Qs, implying a side-on interaction with the structure (Lago et al. 2017).

Coupling this discussion are the examples of G4BPs that are selective of RNA over DNA and vice versa. For example, DHX36’s selective binding of YY1’s promoter G4Q, but not its mRNA equivalent, suggesting a G4Q backbone interaction is also involved (Huang et al. 2012). It is likely that both loops and core G4Q structural features play some role in providing the appropriate selectivity and specificity needed for the G4BP interaction and subsequent cellular pathway outcome.

Binding versus functioning

Binding of a G4BP to a G4Q does not equate to functioning. This is demonstrated importantly with the effects of loop modifications on the c-KIT promoter G4Q in complex with

PARP1 (Edwards et al. 2020). PARP1 affinity for the G4Q increased as the loop features were removed, however, the complex was no longer able to activate PARP (Edwards et al. 2020). In this study, there was a clear link with the pentanucleotide loop as playing a role in this interaction, as noted above. This also highlights the fact that protein functioning in nature and the energetics of an interaction do not always hold a linear correlation. This is of course an important distinction as we try to design functional assays to support our biophysical claims in *in vivo* systems.

Concluding remarks

G4BPs play a central role in many cellular processes, from transcription to translation to genomic stabilization. With the abundance of G4Q sequences present in the cell, it follows that a G4BP must have some selective manner of

Table 1 G4-quadruplex binding proteins (G4BPs)

	G4-sequence	Function, if known	Comments
Telomeric G4BPs			
POT1	(TTAGGG) _n	Unfolding of G4 (and refolding with POT1-TPP1 complex)	Selective to anti-parallel G4Qs
RPA	(TTAGGG) _n	Unfolds G4Qs	Unfolds both parallel and anti-parallel G4Qs
Human CST	TTAGGG.AATCCC	Unfolds G4Qs	Complex of 3 proteins; CTC1 contains the DNA-binding site
hnRNP A1 and UPI	(TTAGGG) _n	Unfolds G4Qs	Specifically, nYAGn seq Binds RNA and DNA
BLM	(TTAGGG) _n	Unfolds G4Qs (leading strand) in 3'-5'	Needs a ssDNA spacer between G4Q and the replication fork to function Similar to WRN
WRN	(TTAGGG) _n	Unfolds G4Qs (lagging strand) in 3'-5'	Similar to BLM
BRCA1	(TTAGGG) _n		Binds G4Qs in vitro
Pif1	(TTAGGG) _n	Unfolds 5'-3' direction	
Replication G4BPs			
FANCI		Unfolds G4Qs in 5'-3' direction	unwinds both intramolecular and inter-molecular G4Qs
Promoter G4BPs			
Sp1	G4 (C) G3 (CC) G5 (C) G4 (TCCC C) G4 (CGG) (VEGF) CCCGGGCGGGCGCGAGGGAGG GGAGG (c-KIT) CGGGGCGGGGCGGGGGCGGGG GCG (HRAS)		Binds parallel and antiparallel G4Qs
DHX36	Many sequences	Unfolds G4Qs in 3'-5'	Strong preference for binding parallel DNA and RNA G4Qs over antiparallel; requires trailing 3' end
Nucleolin	TGGGGAGGGTGGGGAGGGTGG GGAAGG (c-MYC) (GGGGCC) _n (HRE)	Stabilizes G4Q to suppress transcription	Preferentially binds parallel G4Qs but can bind both
NM23-H2	TGGGGAGGGTGGGGAGGGTGG GGAAGG (c-MYC)	Unfolds G4Q to enable transcription	
MAZ	TGGGGAGGGTGGGGAGGGTGG GGAAGG (c-MYC) ACAGGGGTGTGGGG (Pur-1) TCGGGTTGCGGGCGCAGGGCA CGGGCG and CGGGGCGGGGCG GGGGCGGGGCG (HRAS) GGGAGGGAGGGAAGGAGGGAG GGAGGGA (KRAS)	Unfolds G4Qs to enable transcription	
PARP1	C3G3CG3CGCGAG3AG4AG2 (c-KIT) TGGGGAGGGTGGGGAGGGTGG GGAAGG (c-MYC) GGGAGGGAGGGAAGGAGGGAG GGAGGGA (KRAS)		Recognizes parallel G4Qs; binding with c-KIT activates PARP1
XPD/XPB		XPD: 5'-3' direction XPB: 3'-5' direction	
hnRNP	GGGAGGGAGGGAAGGAGGGAG GGAGGGA (KRAS) GGGGTGGGGCCCTGCGAGGGC GGG (TRA2β)	Unfolds G4Q	

Table 1 (continued)

	G4-sequence	Function, if known	Comments
RNA G4BPs			
hnRNP A1	AACGAGGGAGGGAGGGAGAGG GAGAGA- (MMP16) AGCCGGGGCUGGGCGGG GACCG GGCUUGU (ARPC2)	Unfolds G4Q	
DHX36	Many sequences		
DHX9	Many sequences	Unfolds G4Qs in 3'-5'	Preference for RNA substrates; requires a 3' single-stranded tail for initial binding
FMRP	GUGUGGAAGGAGUGGCUG GGUUG(sc1)	Stabilizes the G4Q	

interacting with its G4Q. Such, binding modes discussed here that may afford selectivity are the top-, groove- and loop-binding interactions. While only a few high-resolution structural examples exist of each, combinations of these modes in these interacting species are apparent. As the field of G4Q-G4BPs continues to evolve—new binding proteins discovered and structures solved, further insights into their functions and cellular pathogenesis will be gained.

Author contribution The author is the sole contributor to the design, preparation and revision of this work and approves it for submission.

Declarations

Ethics approval As this is a review, no ethics approvals were needed or sought, nor were there any human subjects involved, and therefore no need for participation or publication consents.

Conflict of interest The author declares no competing interests.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Ali AAE, Timinszky G, Arribas-Bosacoma R, Kozłowski M, Hassa PO, Hassler M, Ladurner AG, Pearl LH, Oliver AW (2012) The zinc-finger domains of PARP1 cooperate to recognize DNA strand breaks. *Nat Struct Mol Biol* 19:685–692. <https://doi.org/10.1038/nsmb.2335>
- Amato J, Pagano A, Capasso D, Di Gaetano S, Giustiniano M, Novellino E, Randazzo A, Pagano B (2018) Targeting the BCL2 gene promoter G-quadruplex with a new class of furopyridazinone-based molecules. *ChemMedChem* 13:406–410. <https://doi.org/10.1002/cmdc.201700749>
- Amato J, Pagano A, Cosconati S, Amendola G, Fotticchia I, Iaccarino N, Marinello J, De Magis A, Capranico G, Novellino E, Pagano B, Randazzo A (2017) Discovery of the first dual G-triplex/G quadruplex stabilizing compound: a new opportunity in the targeting of G-rich DNA structures? *Biochimica et biophysica acta* 1861:1271–1280. <https://doi.org/10.1016/j.bbagen.2016.11.008>
- Ambrus A, Chen D, Dai J, Jones RA, Yang D (2005) Solution structure of the biologically relevant G quadruplex element in the human c-MYC promoter. Implications for G-quadruplex stabilization. *Biochemistry* 44:2048–2058. <https://doi.org/10.1021/bi048242p>
- Arora A, Maiti S (2009) Differential biophysical behavior of human telomeric RNA and DNA quadruplex. *J Phys Chem B* 113:10515–10520. <https://doi.org/10.1021/jp810638n>
- Asamitsu S, Obata S, Yu Z, Bando T, Sugiyama H (2019) Recent progress of targeted G-quadruplex-preferred ligands toward cancer therapy. *Molecules (Basel, Switzerland)* 24:429. <https://doi.org/10.3390/molecules24030429>
- Balasubramanian S, Hurley LH, Neidle S (2011) Targeting G-quadruplexes in gene promoters: a novel anticancer strategy? *Nat Rev Drug Discov* 10:261–275. <https://doi.org/10.1038/nrd3428>
- Ballal RD, Saha T, Fan S, Haddad BR, Rosen EM (2009) BRCA1 localization to the telomere and its loss from the telomere in response to DNA damage. *J Biol Chem* 284:36083–36098. <https://doi.org/10.1074/jbc.M109.025825>
- Bharti SK, Sommers JA, George F, Kuper J, Hamon F, Shin-ya K, Teulade-Fichou MP, Kisker C, Brosh RM Jr (2013) Specialization among iron-sulfur cluster helicases to resolve G-quadruplex DNA structures that threaten genomic stability. *J Biol Chem* 288:28217–28229. <https://doi.org/10.1074/jbc.M113.496463>
- Bhattacharjee A, Wang Y, Diao J, Price CM (2017) Dynamic DNA binding, junction recognition and G4 melting activity underlie the telomeric and genome-wide roles of human CST. *Nucleic Acids Res* 45:12311–12324. <https://doi.org/10.1093/nar/gkx878>
- Al-Naama N, Mackeh R, Kino T (2020) C2H2-type zinc finger proteins in brain development, neurodevelopmental, and other neuropsychiatric disorders: systematic literature-based analysis. *Front Neurol* 11:1–13. <https://doi.org/10.3389/fneur.2020.00032>

- Bhattacharyya D, Mirihana Arachchilage G, Basu S (2016) Metal cations in G-quadruplex folding and stability. *Front Chem* 4:38. <https://doi.org/10.3389/fchem.2016.00038>
- Bilokapic S, Suskiewicz MJ, Ahel I, Halic M (2020) Bridging of DNA breaks activates PARP2-HPF1 to modify chromatin. *Nature* 585:609–613. <https://doi.org/10.1038/s41586-020-2725-7>
- Black AR, Black JD, Azizkhan-Clifford J (2001) Sp1 and kruppel-like factor family of transcription factors in cell growth regulation and cancer. *J Cell Physiol* 188:143–160. <https://doi.org/10.1002/jcp.1111>
- Bochkarev A, Pfuetzner RA, Edwards AM, Frappier L (1997) Structure of the single-stranded-DNA binding domain of replication protein A bound to DNA. *Nature* 385:176–181. <https://doi.org/10.1038/385176a0>
- Bochkareva E, Belegu V, Korolev S, Bochkarev A (2001) Structure of the major single-stranded DNA binding domain of replication protein A suggests a dynamic mechanism for DNA binding. *The EMBO journal* 20:612–618. <https://doi.org/10.1093/emboj/20.3.612>
- Bochkareva E, Korolev S, Lees-Miller SP, Bochkarev A (2002) Structure of the RPA trimerization core and its role in the multistep DNA-binding mechanism of RPA. *EMBO J* 21:1855–1863. <https://doi.org/10.1093/emboj/21.7.1855>
- Booy EP, Meier M, Okun N, Novakowski SK, Xiong S, Stetefeld J, McKenna SA (2012) The RNA helicase RHAU (DHX36) unwinds a G4-quadruplex in human telomerase RNA and promotes the formation of the P1 helix template boundary. *Nucleic Acids Res* 40:4110–4124. <https://doi.org/10.1093/nar/gkr1306>
- Brázda V, Hároníková L, Liao JCC, Fojta M (2014) DNA and RNA quadruplex-binding proteins. *Int J Mol Sci* 15:17493–17517. <https://doi.org/10.3390/ijms151017493>
- Brázda V, Hároníková L, Liao JCC, Fridrichová H, Jagelská EB (2016) Strong preference of BRCA1 protein to topologically constrained non-B DNA structures. *BMC Mol Biol* 17:14. <https://doi.org/10.1186/s12867-016-0068-6>
- Budhathoki JB, Ray S, Urban V, Janscak P, Yodh JG, Balci H (2014) RecQ-core of BLM unfolds telomeric G-quadruplex in the absence of ATP. *Nucleic Acids Res* 42:11528–11545. <https://doi.org/10.1093/nar/gku856>
- Bugaut A, Balasubramanian S (2008) A sequence-independent study of the influence of short loop lengths on the stability and topology of intramolecular DNA G-quadruplexes. *Biochemistry* 47:689–697. <https://doi.org/10.1021/bi701873c>
- Bugaut A, Balasubramanian S (2010) 5'-UTR RNA G-quadruplexes: translation regulation and targeting. *Nucleic Acids Res* 40:4727–4741. <https://doi.org/10.1093/nar/gks068>
- Byrd AK, Bell MR, Raney KD (2018) Pif1 helicase unfolding of G-quadruplex DNA is highly dependent on sequence and reaction conditions. *J Biol Chem* 293:17792–17802. <https://doi.org/10.1074/jbc.RA118.004499>
- Byrd AK, Raney KD (2015) A parallel quadruplex DNA is bound tightly but unfolded slowly by pif1 helicase. *J Biol Chem* 290:6482–6494. <https://doi.org/10.1074/jbc.M114.630749>
- Castillo Bosch P, Segura-Bayona S, Koole W, van Heteren JT, Dewar JM, Tijsterman M, Knipscheer P (2014) FANCD1 promotes DNA synthesis through G-quadruplex structures. *EMBO J* 33:2521–2533. <https://doi.org/10.15252/embj.201488663-2533>
- Chakraborty P, Grosse F (2011) Human DHX9 helicase preferentially unwinds RNA-containing displacement loops (R-loops) and G-quadruplexes. *DNA Repair (Amst)* 10:654–665. <https://doi.org/10.1016/j.dnarep.2011.04.013>
- Chatterjee S, Zagelbaum J, Savitsky P, Sturzenegger A, Huttner D, Janscak P, Hickson ID, Gileadi O, Rothenberg E (2014) Mechanistic insight into the interaction of BLM helicase with intra-strand G-quadruplex structures. *Nat Commun* 5:5556. <https://doi.org/10.1038/ncomms6556>
- Chen H, Liu H, Qing G (2018a) Targeting oncogenic Myc as a strategy for cancer treatment. *Signal Transduct Target Ther* 3:5. <https://doi.org/10.1038/s41392-018-0008-7>
- Chen J, Cheng M, Salgado Gilmar F, Stadlbauer P, Zhang X, Amrane S, Guédin A, He F, Šponer J, Ju H, Mergny J-L, Zhou J (2021) The beginning and the end: flanking nucleotides induce a parallel G-quadruplex topology. *Nucleic Acids Res* 49:9548–9955. <https://doi.org/10.1093/nar/gkab681>
- Chen MC, Tippana R, Demeshkina NA, Murat P, Balasubramanian S, Myong S, Ferré-D'Amaré AR (2018b) Structural basis of G-quadruplex unfolding by the DEAH/RHA helicase DHX36. *Nature* 558:465–469. <https://doi.org/10.1038/s41586-018-0209-9>
- Chen Y, Yang D (2012) Sequence, stability, and structure of G-quadruplexes and their interactions with drugs. *Curr Protoc Nucleic Acid Chem Chapter: Unit17.5*. <https://doi.org/10.1002/0471142700.nc1705s50>
- Cheung I, Schertzer M, Rose A, Lansdorp PM (2002) Disruption of dog-1 in *Caenorhabditis elegans* triggers deletions upstream of guanine-rich DNA. *Nat Genet* 31:405–409. <https://doi.org/10.1038/ng928>
- Cogoi S, Paramasivam M, Filichev V, Géci I, Pedersen EB, Xodo LE (2009) Identification of a new quadruplex motif in the KRAS promoter and design of pyrene-modified G4-decoys with antiproliferative activity in pancreatic cancer cells. *J Med Chem* 52:564–568. <https://doi.org/10.1021/jm800874t>
- Cogoi S, Paramasivam M, Membrino A, Yokoyama KK, Xodo LE (2010) The KRAS promoter responds to Myc-associated zinc finger and poly(ADP-ribose) polymerase 1 proteins, which recognize a critical quadruplex-forming G4-element. *J Biol Chem* 285:22003–22016. <https://doi.org/10.1074/jbc.M110.101923>
- Cogoi S, Shchekotikhin AE, Xodo LE (2014) HRAS is silenced by two neighboring G-quadruplexes and activated by MAZ, a zinc-finger transcription factor with DNA unfolding property. *Nucleic Acids Res* 42:8379–8388. <https://doi.org/10.1093/nar/gku574>
- Cogoi S, Zorzet S, Rapozzi V, Géci I, Pedersen EB, Xodo LE (2013) MAZ-binding G4-decoy with locked nucleic acid and twisted intercalating nucleic acid modifications suppresses KRAS in pancreatic cancer cells and delays tumor growth in mice. *Nucleic Acids Res* 41:4049–4064. <https://doi.org/10.1093/nar/gkt127>
- Collie GW, Promontorio R, Hampel SM, Micco M, Neidle S, Parkinson GN (2012) Structural Basis for Telomeric G-Quadruplex Targeting by Naphthalene Diimide Ligands. *J Am Chem Soc* 134:2723–2731. <https://doi.org/10.1021/ja2102423>
- Creacy SD, Routh ED, Iwamoto F, Nagamine Y, Akman SA, Vaughn JP (2008) G4 resolvase 1 binds both DNA and RNA tetramolecular quadruplex with high affinity and is the major source of tetramolecular quadruplex G4-DNA and G4-RNA resolving activity in HeLa cell lysates. *J Biol Chem* 283:34626–34634. <https://doi.org/10.1074/jbc.M806277200>
- Da Ros S, Nicoletto G, Rigo R, Ceschi S, Zorzan E, Dacasto M, Giannini M, Sissi C (2021) G-quadruplex modulation of SP1 functional binding sites at the KIT proximal promoter. *Int J Mol Sci* 22:329–341. <https://doi.org/10.3390/ijms22010329>
- Dalloul Z, Chenuet P, Dalloul I, Boyer F, Aldigier JC, Laffleur B, El Makhour Y, Ryffel B, Quesniaux VFJ, Togbé D, Mergny JL, Cook-Moreau J, Cogné M (2018) G-quadruplex DNA targeting alters class-switch recombination in B cells and attenuates allergic inflammation. *J Allergy Clin Immunol* 142:1352–1355. <https://doi.org/10.1016/j.jaci.2018.06.011>
- Dang CV (2012) MYC on the path to cancer. *Cell* 149:22–35. <https://doi.org/10.1016/j.cell.2012.03.003>

- Darnell JC, Jensen KB, Jin P, Brown V, Warren ST, Darnell RB (2001) Fragile X mental retardation protein targets G quartet mRNAs important for neuronal function. *Cell* 107:489–449. [https://doi.org/10.1016/s0092-8674\(01\)00566-9](https://doi.org/10.1016/s0092-8674(01)00566-9)
- Dexheimer TS, Carey SS, Zuohe S, Gokhale VM, Hu X, Murata LB, Maes EM, Weichsel A, Sun D, Meuillet EJ, Montfort WR, Hurlley LH (2009) NM23-H2 may play an indirect role in transcriptional activation of c-myc gene expression but does not cleave the nuclease hypersensitive element III(1). *Mol Cancer Ther* 8:1363–1377. <https://doi.org/10.1158/1535-7163.Mct-08-1093>
- Ding J, Hayashi MK, Zhang Y, Manche L, Krainer AR, Xu RM (1999) Crystal structure of the two-RRM domain of hnRNP A1 (UP1) complexed with single-stranded telomeric DNA. *Genes Dev* 13:1102–1115. <https://doi.org/10.1101/gad.13.9.1102>
- Drapkin R, Reardon JT, Ansari A, Huang JC, Zawel L, Ahn K, Sancar A, Reinberg D (1994) Dual role of TFIIF in DNA excision repair and in transcription by RNA polymerase II. *Nature* 368:769–772. <https://doi.org/10.1038/368769a0>
- Edwards AD, Marecki JC, Byrd AK, Gao J, Raney Kevin D (2020) G-Quadruplex loops regulate PARP-1 enzymatic activation. *Nucleic Acids Res* 49:416–431. <https://doi.org/10.1093/nar/gkaa1172>
- Eom KS, Cheong JS, Lee SJ (2016) Structural analyses of zinc finger domains for specific interactions with DNA. *J Microbiol Biotechnol* 26:2019–2029. <https://doi.org/10.4014/jmb.1609.09021>
- Eustermann S, Wu WF, Langelier MF, Yang JC, Easton LE, Riccio AA, Pascal JM, Neuhaus D (2015) Structural basis of detection and signaling of DNA single-strand breaks by human PARP-1. *Mol Cell* 60:742–754. <https://doi.org/10.1016/j.molcel.2015.10.032>
- Fekete A, Kenesi E, Hunyadi-Gulyas E, Durgó H, Berko B, Dunai ZA, Bauer PI (2012) The guanine-quadruplex structure in the human c-myc gene's promoter is converted into B-DNA form by the human poly(ADP-Ribose)polymerase-1. *PLoS one* 7:e42690. <https://doi.org/10.1371/journal.pone.0042690>
- Fleming AM, Ding Y, Alenko A, Burrows CJ (2016) Zika virus genomic rna possesses conserved quadruplexes characteristic of the flaviviridae family. *ACS Infectious Diseases* 2:674–681. <https://doi.org/10.1021/acsinfecdis.6b00109>
- Ford D, Easton DF, Bishop DT, Narod SA, Goldgar DE (1994) Risks of cancer in BRCA1-mutation carriers. *Breast Cancer Linkage Consortium. Lancet* 343:692–695. [https://doi.org/10.1016/s0140-6736\(94\)91578-4](https://doi.org/10.1016/s0140-6736(94)91578-4)
- Frasson I, Soldà P, Nadai M, Lago S, Richter SN (2020) Transcription factor recruitment by parallel quadruplexes to promote transcription: the case of herpes simplex virus-1 ICP4. *bioRxiv:2020.2007.2014.202747*. <https://doi.org/10.1101/2020.07.14.202747>
- Ghosal G, Muniyappa K (2005) *Saccharomyces cerevisiae* Mre11 is a high-affinity G4 DNA-binding protein and a G-rich DNA-specific endonuclease: implications for replication of telomeric DNA. *Nucleic Acids Res* 33:4692–4703. <https://doi.org/10.1093/nar/gki777>
- Ghosal G, Muniyappa K (2007) The characterization of *Saccharomyces cerevisiae* Mre11/Rad50/Xrs2 complex reveals that Rad50 negatively regulates Mre11 endonucleolytic but not the exonucleolytic activity. *J Mol Biol* 372:864–882. <https://doi.org/10.1016/j.jmb.2007.07.013>
- Giancola C, Pagano B (2013) Energetics of ligand binding to G-quadruplexes. *Top Curr Chem* 330:211–242. https://doi.org/10.1007/128_2012_347
- Ginisty H, Sicard H, Roger B, Bouvet P (1999) Structure and functions of nucleolin. *J Cell Sci* 112(Pt 6):761–772
- Giri B, Smaldino PJ, Thys RG, Creacy SD, Routh ED, Hantgan RR, Lattmann S, Nagamine Y, Akman SA, Vaughn JP (2011) G4 resolvase 1 tightly binds and unwinds unimolecular G4-DNA. *Nucleic Acids Res* 39:7161–7178. <https://doi.org/10.1093/nar/gkr234>
- González V, Guo K, Hurley L, Sun D (2009) Identification and characterization of nucleolin as a c-mycG-quadruplex-binding protein. *J Biol Chem* 284:23622–23635. <https://doi.org/10.1074/jbc.M109.018028>
- Gordon S, Akopyan G, Garban H, Bonavida B (2006) Transcription factor YY1: structure, function, and therapeutic implications in cancer biology. *Oncogene* 25:1125–1142. <https://doi.org/10.1038/sj.onc.1209080>
- Gray LT, Vallur AC, Eddy J, Maizels N (2014) G quadruplexes are genome-wide targets of transcriptional helicases XPB and XPD. *Nature chemical biology* 10:313–318. <https://doi.org/10.1038/nchembio.1475>
- Grün JT, Schwalbe H (2022) Folding dynamics of polymorphic G-quadruplex structures. *Biopolymers* 113:e23477. <https://doi.org/10.1002/bip.23477>
- Guédin A, Gros J, Alberti P, Mergny J-L (2010) How long is too long? Effects of loop size on G-quadruplex stability. *Nucleic Acids Res* 38:7858–7868. <https://doi.org/10.1093/nar/gkq639>
- Haeusler AR, Donnelly CJ, Periz G, Simko EAJ, Shaw PG, Kim M-S, Maragakis NJ, Troncoso JC, Pandey A, Sattler R, Rothstein JD, Wang J (2014) C9orf72 nucleotide repeat structures initiate molecular cascades of disease. *Nature* 507:195–200. <https://doi.org/10.1038/nature13124>
- Hanakahi LA, Dempsey LA, Li M-J, Maizels N (1997) Nucleolin is one component of the B cell-specific transcription factor and switch region binding protein, LR1. *Proceedings of the National Academy of Sciences* 94:3605–3610. <https://doi.org/10.1073/pnas.94.8.3605>
- Hanakahi LA, Sun H, Maizels N (1999) High affinity interactions of nucleolin with G-G-paired rDNA. *J Biol Chem* 274:15908–15912. <https://doi.org/10.1074/jbc.274.22.15908>
- Hansel-Hertsch R, Di Antonio M, Balasubramanian S (2017) DNA G-quadruplexes in the human genome: detection, functions and therapeutic potential. *Nat Rev Mol Cell Biol* 18:279–284. <https://doi.org/10.1038/nrm.2017.3>
- Hatzakis E, Okamoto K, Yang D (2010) Thermodynamic stability and folding kinetics of the major G-quadruplex and its loop isomers formed in the nuclease hypersensitive element in the human c-Myc promoter: effect of loops and flanking segments on the stability of parallel-stranded intramolecular G-quadruplexes. *Biochemistry* 49:9152–9160. <https://doi.org/10.1021/bi100946g>
- Hazel P, Huppert J, Balasubramanian S, Neidle S (2004) Loop-Length-Dependent Folding of G-Quadruplexes. *J Am Chem Soc* 126:16405–16415. <https://doi.org/10.1021/ja045154j>
- Heddi B, Cheong VV, Schmitt E, Mechulam Y, Phan AT (2020) Recognition of different base tetrads by RHAU (DHX36): X-ray crystal structure of the G4 recognition motif bound to the 3'-end tetrad of a DNA G-quadruplex. *J Struct Biol* 209:107399. <https://doi.org/10.1016/j.jsb.2019.10.001>
- Horvath MP, Schultz SC (2001) DNA G-quartets in a 1.86 Å resolution structure of an *Oxytricha nova* telomeric protein-DNA complex. *J Mol Biol* 310:367–377. <https://doi.org/10.1006/jmbi.2001.4766>
- Hou X-M, Wu W-Q, Duan X-L, Liu N-N, Li H-H, Fu J, Dou S-X, Li M, Xi X-G (2015) Molecular mechanism of G-quadruplex unwinding helicase: sequential and repetitive unfolding of G-quadruplex by Pif1 helicase. *Biochemical Journal* 466:189–199. <https://doi.org/10.1042/bj20140997>
- Huang W, Smaldino PJ, Zhang Q, Miller LD, Cao P, Stadelman K, Wan M, Giri B, Lei M, Nagamine Y, Vaughn JP, Akman SA, Sui G (2012) Yin Yang 1 contains G-quadruplex structures in its promoter and 5'-UTR and its expression is modulated by G4 resolvase 1. *Nucleic Acids Res* 40:1033–1049. <https://doi.org/10.1093/nar/gkr849>

- Huang Z-L, Dai J, Luo W-H, Wang X-G, Tan J-H, Chen S-B, Huang Z-S (2018) Identification of G-Quadruplex-Binding Protein from the Exploration of RGG Motif/G-Quadruplex Interactions. *J Am Chem Soc* 140:17945–17955. <https://doi.org/10.1021/jacs.8b09329>
- Hudson JS, Ding L, Le V, Lewis E, Graves D (2014) Recognition and binding of human telomeric G-quadruplex DNA by unfolding protein 1. *Biochemistry* 53:3347–3356. <https://doi.org/10.1021/bi500351u>
- Huppert JL, Balasubramanian S (2005) Prevalence of quadruplexes in the human genome. *Nucleic Acids Res* 33:2908–2916. <https://doi.org/10.1093/nar/gki609>
- Huppert JL, Bugaut A, Kumari S, Balasubramanian S (2008) G-quadruplexes: the beginning and end of UTRs. *Nucleic Acids Res* 36:6260–6268. <https://doi.org/10.1093/nar/gkn511>
- Hwang H, Buncher N, Opresko PL, Myong S (2012) POT1-TTP1 regulates telomeric overhang structural dynamics. *Structure* 20:1872–1880. <https://doi.org/10.1016/j.str.2012.08.018>
- Isalan M, Patel SD, Balasubramanian S, Choo Y (2001) Selection of zinc fingers that bind single-stranded telomeric DNA in the G-quadruplex conformation. *Biochemistry* 40:830–836. <https://doi.org/10.1021/bi001728v>
- Jamieson AC, Miller JC, Pabo CO (2003) Drug discovery with engineered zinc-finger proteins. *Nat Rev Drug Discov* 2:361–368. <https://doi.org/10.1038/nrd1087>
- Joachimi A, Benz A, Hartig JS (2009) A comparison of DNA and RNA quadruplex structures and stabilities. *Bioorg Med Chem* 17:6811–6815. <https://doi.org/10.1016/j.bmc.2009.08.043>
- Kharel P, Becker G, Tsvetkov V, Ivanov P (2020) Properties and biological impact of RNA G-quadruplexes: from order to turmoil and back. *Nucleic Acids Res* 48:12534–12555. <https://doi.org/10.1093/nar/gkaa1126>
- Kruisselbrink E, Guryev V, Brouwer K, Pontier DB, Cuppen E, Tijsterman M (2008) Mutagenic capacity of endogenous G4 DNA underlies genome instability in FANCD1-defective *C. elegans*. *Curr Biol* 18:900–905. <https://doi.org/10.1016/j.cub.2008.05.013>
- Ladame S, Schouten JA, Roldan J, Redman JE, Neidle S, Balasubramanian S (2006) Exploring the recognition of quadruplex DNA by an engineered Cys2-His2 zinc finger protein. *Biochemistry* 45:1393–1399. <https://doi.org/10.1021/bi050229x>
- Lago S, Tosoni E, Nadai M, Palumbo M, Richter SN (2017) The cellular protein nucleolin preferentially binds long-looped G-quadruplex nucleic acids. *Biochimica et Biophysica Acta (BBA) – Genera Subjects* 1861(1371–1381). <https://doi.org/10.1016/j.bbagen.2016.11.036>
- Lane AN, Chaires JB, Gray RD, Trent JO (2008) Stability and kinetics of G-quadruplex structures. *Nucleic Acids Research* 36:5482–5515. <https://doi.org/10.1093/nar/gkn517>
- Langelier MF, Planck JL, Roy S, Pascal JM (2011) Crystal structures of poly(ADP-ribose) polymerase-1 (PARP-1) zinc fingers bound to DNA: structural and functional insights into DNA-dependent PARP-1 activity. *J Biol Chem* 286:10690–10701. <https://doi.org/10.1074/jbc.M110.202507>
- Langelier MF, Planck JL, Roy S, Pascal JM (2012) Structural basis for DNA damage-dependent poly(ADP-ribosylation) by human PARP-1. *Science* 336:728–732. <https://doi.org/10.1126/science.1216338>
- Laouer K, Schmid M, Wien F, Changenet P, Hache F (2021) Folding dynamics of DNA G-quadruplexes probed by millisecond temperature jump circular dichroism. *J Phys Chem B* 125:8088–8098. <https://doi.org/10.1021/acs.jpcc.1c01993>
- Lattmann S, Giri B, Vaughn JP, Akman SA, Nagamine Y (2010) Role of the amino terminal RHAU-specific motif in the recognition and resolution of guanine quadruplex-RNA by the DEAH-box RNA helicase RHAU. *Nucleic Acids Res* 38:6219–6233. <https://doi.org/10.1093/nar/gkq372>
- Lattmann S, Stadler MB, Vaughn JP, Akman SA, Nagamine Y (2011) The DEAH-box RNA helicase RHAU binds an intramolecular RNA G-quadruplex in TERC and associates with telomerase holoenzyme. *Nucleic Acids Res* 39:9390–9404. <https://doi.org/10.1093/nar/gkr630>
- Lavezzo E, Berselli M, Frasson I, Perrone R, Palu G, Brazzale AR, Richter SN, Toppo S (2018) G-quadruplex forming sequences in the genome of all known human viruses: A comprehensive guide. *PLoS Comput Biol* 14:e1006675. <https://doi.org/10.1371/journal.pcbi.1006675>
- Lemkul JA (2019) Same fold, different properties: polarizable molecular dynamics simulations of telomeric and TERRA G-quadruplexes. *Nucleic Acids Research* 48:561–575. <https://doi.org/10.1093/nar/gkz1154>
- Lerner LK, Holzer S, Kilkenny ML, Šviković S, Murat P, Schiavone D, Eldridge CB, Bittleston A, Maman JD, Branzei D, Stott K, Pellegrini L, Sale JE (2020) Timeless couples G-quadruplex detection with processing by DDX11 helicase during DNA replication. *The EMBO Journal* 39(e104185). <https://doi.org/10.15252/embj.2019104185>
- Lerner LK, Sale JE (2019) Replication of G Quadruplex DNA. *Genes* 10:95
- Lew A, Rutter WJ, Kennedy GC (2000) Unusual DNA structure of the diabetes susceptibility locus IDDM2 and its effect on transcription by the insulin promoter factor Pur-1/MAZ. *Proceedings of the National Academy of Sciences* 97:12508–12512. <https://doi.org/10.1073/pnas.97.23.12508>
- Li L, Williams P, Ren W, Wang MY, Gao Z, Miao W, Huang M, Song J, Wang Y (2021) YY1 interacts with guanine quadruplexes to regulate DNA looping and gene expression. *Nat Chem Biol* 17:161–168. <https://doi.org/10.1038/s41589-020-00695-1>
- Lim CJ, Barbour AT, Zaug AJ, Goodrich KJ, McKay AE, Wuttke DS, Cech TR (2020) The structure of human CST reveals a decameric assembly bound to telomeric DNA. *Science* 368:1081–1085. <https://doi.org/10.1126/science.aaz9649>
- Lim KW, Lacroix L, Yue DJ, Lim JK, Lim JM, Phan AT (2010) Coexistence of two distinct G-quadruplex conformations in the hTERT promoter. *J Am Chem Soc* 132:12331–12342. <https://doi.org/10.1021/ja101252n>
- Lin S, Fu X-D (2007) SR proteins and related factors in alternative splicing. *Adv Exp Med Biol* 623:107–122. https://doi.org/10.1007/978-0-387-77374-2_7
- Long Jennifer C, Caceres Javier F (2008) The SR protein family of splicing factors: master regulators of gene expression. *Biochem J* 417:15–27. <https://doi.org/10.1042/bj20081501>
- Lowran K, Campbell L, Popp P, Wu CG (2020) Assembly of a G-quadruplex repair complex by the FANCD1 DNA helicase and the REV1 polymerase. *Genes* 11(5). <https://doi.org/10.3390/genes11010005>
- Lyu K, Chow EY-C, Mou X, Chan T-F, Kwok Chun K (2021) RNA G-quadruplexes (rG4s): genomics and biological functions. *Nucleic Acids Res* 49:5426–5450. <https://doi.org/10.1093/nar/gkab187>
- Masuzawa T, Oyoshi T (2020) Roles of the RGG domain and RNA recognition motif of nucleolin in G-quadruplex stabilization. *ACS Omega* 5:5202–5208. <https://doi.org/10.1021/acsomega.9b04221>
- McLuckie KIE, Di Antonio M, Zecchini H, Xian J, Caldas C, Krippeendorff B-F, Tannahill D, Lowe C, Balasubramanian S (2013) G-Quadruplex DNA as a molecular target for induced synthetic lethality in cancer cells. *J Am Chem Soc* 135:9640–9643. <https://doi.org/10.1021/ja404868t>
- Medeiros-Silva J, Guédin A, Salgado GF, Mergny J-L, Queiroz JA, Cabrita EJ, Cruz C (2017) Phenanthroline-bis-oxazole ligands

- for binding and stabilization of G-quadruplexes. *Biochimica et Biophysica Acta (BBA) - General Subjects* 1861:1281–1292. <https://doi.org/10.1016/j.bbagen.2016.11.024>
- Meier-Stephenson V, Badmalia MD, Mrozowich T, Lau KCK, Schultz SK, Gemmill DL, Osiowy C, van Marle G, Coffin CS, Patel TR (2021) Identification and characterization of a G-quadruplex structure in the pre-core promoter region of hepatitis B virus covalently closed circular DNA. *J Biol Chem* 296:100589. <https://doi.org/10.1016/j.jbc.2021.100589>
- Membrino A, Cogoi S, Pedersen EB, Xodo L (2011) G4-DNA Formation in the HRAS Promoter and Rational Design of Decoy Oligonucleotides for Cancer Therapy. *PLoS one* 6:e24421. <https://doi.org/10.1371/journal.pone.0024421>
- Metifiot M, Amrane S, Litvak S, Andreola ML (2014) G-quadruplexes in viruses: function and potential therapeutic applications. *Nucleic Acids Res* 42:12352–12366. <https://doi.org/10.1093/nar/gku999>
- Miller DM, Thomas SD, Islam A, Muench D, Sedoris K (2012) c-Myc and cancer metabolism. *Clinical cancer research: an official journal of the American Association for Cancer Research* 18:5546–5553. <https://doi.org/10.1158/1078-0432.CCR-12-0977>
- Mongelard F, Bouvet P (2007) Nucleolin: a multiFACeTed protein. *Trends in Cell Biology* 17:80–86. <https://doi.org/10.1016/j.tcb.2006.11.010>
- Morgan CE, Meagher JL, Levensgood JD, Delproposto J, Rollins C, Stuckey JA, Tolbert BS (2015) The first crystal structure of the UPI domain of hnRNP A1 bound to RNA reveals a new look for an old RNA binding protein. *Journal of molecular biology* 427:3241–3257. <https://doi.org/10.1016/j.jmb.2015.05.009>
- Murat P, Marsico G, Herdy B, Ghanbarian A, Portella G, Balasubramanian S (2018) RNA G-quadruplexes at upstream open reading frames cause DHX36- and DHX9-dependent translation of human mRNAs. *Genome Biol* 19:229. <https://doi.org/10.1186/s13059-018-1602-2>
- Myers JC, Moore SA, Shamoo Y (2003) Structure-based incorporation of 6-methyl-8-(2-deoxy-beta-ribofuranosyl)isoxanthopterin into the human telomeric repeat DNA as a probe for UPI binding and destabilization of G-tetrad structures. *J Biol Chem* 278:42300–42306. <https://doi.org/10.1074/jbc.M306147200>
- Nguyen TQN, Lim KW, Phan AT (2020) Folding kinetics of G-quadruplexes: duplex stem loops drive and accelerate G-quadruplex folding. *J Phys Chem B* 124:5122–5130. <https://doi.org/10.1021/acs.jpcc.0c02548>
- Nielsen MC, Larsen AF, Abdikadir FH, Ulven T (2014) Phenanthroline-2,9-bis-triazoles as selective G-quadruplex ligands. *Eur J Med Chem* 72:119–126. <https://doi.org/10.1016/j.ejmech.2013.11.027>
- Nishikawa T, Kuwano Y, Takahara Y, Nishida K, Rokutan K (2019) HnRNPA1 interacts with G-quadruplex in the TRA2B promoter and stimulates its transcription in human colon cancer cells. *Sci Rep* 9:10276. <https://doi.org/10.1038/s41598-019-46659-x>
- O'Sullivan RJ, Karlseder J (2010) Telomeres: protecting chromosomes against genome instability. *Nature reviews Molecular cell biology* 11:171–181. <https://doi.org/10.1038/nrm2848>
- Oakley GG, Patrick SM (2010) Replication protein A: directing traffic at the intersection of replication and repair. *Front Biosci (Landmark Ed)* 15:883–900. <https://doi.org/10.2741/3652>
- Ohnmacht SA, Marchetti C, Gunaratnam M, Besser RJ, Haider SM, Di Vita G, Lowe HL, Mellinas-Gomez M, Diocou S, Robson M, Šponer J, Islam B, Barbara Pedley R, Hartley JA, Neidle S (2015) A G-quadruplex-binding compound showing anti-tumour activity in an in vivo model for pancreatic cancer. *Sci Rep* 5:11385. <https://doi.org/10.1038/srep11385>
- Paeschke K, Bochman ML, Garcia PD, Cejka P, Friedman KL, Kowalczykowski SC, Zakian VA (2013) Pif1 family helicases suppress genome instability at G-quadruplex motifs. *Nature* 497:458–462. <https://doi.org/10.1038/nature12149>
- Patel MR, Bhatt A, Steffen JD, Chergui A, Murai J, Pommier Y, Pascal JM, Trombetta LD, Fronczek FR, Talele TT (2014) Discovery and structure-activity relationship of novel 2,3-dihydrobenzofuran-7-carboxamide and 2,3-dihydrobenzofuran-3(2H)-one-7-carboxamide derivatives as poly(ADP-ribose)polymerase-1 inhibitors. *J Med Chem* 57:5579–5601. <https://doi.org/10.1021/jm5002502>
- Patel SD, Isalan M, Gavory G, Ladame S, Choo Y, Balasubramanian S (2004) Inhibition of human telomerase activity by an engineered zinc finger protein that binds G-quadruplexes. *Biochemistry* 43:13452–13458. <https://doi.org/10.1021/bi048892t>
- Peissert S, Sauer F, Grabarczyk DB, Braun C, Sander G, Poterszman A, Egly J-M, Kuper J, Kisker C (2020) In TFIID the Arch domain of XPD is mechanistically essential for transcription and DNA repair. *Nat Commun* 11:1667. <https://doi.org/10.1038/s41467-020-15241-9>
- Perrone R, Lavezzo E, Palù G, Richter SN (2017) Conserved presence of G-quadruplex forming sequences in the Long Terminal Repeat Promoter of Lentiviruses. *Sci Rep* 7:2018. <https://doi.org/10.1038/s41598-017-02291-1>
- Phan AT, Kuryavyi V, Darnell JC, Serganov A, Majumdar A, Ilin S, Raslin T, Polonskaia A, Chen C, Clain D (2011) Structure-Function Studies of FMRP RGG Peptide Recognition of an RNA Duplex-Quadruplex Junction. *Nat Struct Mol Biol* 18:796–804. <https://doi.org/10.1038/nsmb.2064>
- Phan AT, Kuryavyi V, Gaw HY, Patel DJ (2005) Small-Molecule Interaction with a Five-Guanine-Tract G-Quadruplex Structure from the Human MYC Promoter. *Nat Chem Biol* 1:167–173. <https://doi.org/10.1038/nchembio723>
- Phan AT, Modi YS, Patel DJ (2004) Propeller-Type Parallel-Stranded G-Quadruplexes in the Human c-myc Promoter. *J Am Chem Soc* 126:8710–8716. <https://doi.org/10.1021/ja048805k>
- Piazza A, Adrian M, Samazan F, Heddi B, Hamon F, Serero A, Lopes J, Teulade-Fichou M-P, Phan AT, Nicolas A (2015) Short loop length and high thermal stability determine genomic instability induced by G-quadruplex-forming minisatellites. *The EMBO journal* 34:1718–1734. <https://doi.org/10.15252/embj.201490702>
- Podbevšek P, Plavec J (2016) KRAS promoter oligonucleotide with decoy activity dimerizes into a unique topology consisting of two G-quadruplex units. *Nucleic acids Res* 44:917–925. <https://doi.org/10.1093/nar/gkv1359>
- Prakash A, Borgstahl GEO (2012) The Structure and Function of Replication Protein A in DNA Replication. In: MacNeill S (ed) *The eukaryotic replisome: a guide to protein structure and function*. Springer Netherlands, Dordrecht, pp 171–196
- Qureshi MH, Ray S, Sewell AL, Basu S, Balci H (2012) Replication protein A unfolds G-quadruplex structures with varying degrees of efficiency. *J Phys Chem B* 116:5588–5594. <https://doi.org/10.1021/jp300546u>
- Raiber EA, Kranaster R, Lam E, Nikan M, 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. <https://doi.org/10.1093/nar/gkr882>
- Ray S, Bandaria JN, Qureshi MH, Yildiz A, Balci H (2014) G-quadruplex formation in telomeres enhances POT1/TPP1 protection against RPA binding. *Proceedings of the National Academy of Sciences* 111:2990–2995. <https://doi.org/10.1073/pnas.1321436111>
- Razin SV, Borunova VV, Maksimenko OG, Kantidze OL (2012) Cys2His2 zinc finger protein family: classification, functions, and major members. *Biochemistry Biokhimiia* 77:217–226. <https://doi.org/10.1134/s0006297912030017>
- Read M, Harrison RJ, Romagnoli B, Tanious FA, Gowan SH, Reszka AP, Wilson WD, Kelland LR, Neidle S (2001) Structure-based

- design of selective and potent G quadruplex-mediated telomerase inhibitors. *Proceedings of the National Academy of Sciences* 98:4844–4849. <https://doi.org/10.1073/pnas.081560598>
- Rhodes D, Lipps HJ (2015) G-quadruplexes and their regulatory roles in biology. *Nucleic Acids Res* 43:8627–8637. <https://doi.org/10.1093/nar/gkv862>
- Rigo R, Palumbo M, Sissi C (2017) G-quadruplexes in human promoters: A challenge for therapeutic applications. *Biochimica et Biophysica Acta* 1861:1399–1413. <https://doi.org/10.1016/j.bbagen.2016.12.024>
- Robinson J, Raguseo F, Nuccio SP, Liano D, Di Antonio M (2021) DNA G-quadruplex structures: more than simple roadblocks to transcription? *Nucleic Acids Res* 49:8419–8431. <https://doi.org/10.1093/nar/gkab609>
- Rocca R, Palazzesi F, Amato J, Costa G, Ortuso F, Pagano B, Randazzo A, Novellino E, Alcaro S, Moraca F, Artese A (2020) Folding intermediate states of the parallel human telomeric G-quadruplex DNA explored using Well-Tempered Metadynamics. *Sci Rep* 10:3176. <https://doi.org/10.1038/s41598-020-59774-x>
- Sauer M, Paeschke K (2017) G-quadruplex unwinding helicases and their function in vivo. *Biochem Soc Trans* 45:1173–1182. <https://doi.org/10.1042/bst20170097>
- Sengupta P, Bhattacharya A, Sa G, Das T, Chatterjee S (2019) Truncated G-quadruplex isomers cross-talk with the transcription factors to maintain homeostatic equilibria in c-MYC transcription. *Biochemistry* 58:1975–1991. <https://doi.org/10.1021/acs.biochem.9b00030>
- Serikawa T, Spanos C, von Hacht A, Budisa N, Rappsilber J, Kurreck J (2018) Comprehensive identification of proteins binding to RNA G-quadruplex motifs in the 5' UTR of tumor-associated mRNAs. *Biochimie* 144:169–184. <https://doi.org/10.1016/j.biochi.2017.11.003>
- Sexton AN, Collins K (2011) The 5' guanosine tracts of human telomerase RNA are recognized by the G-quadruplex binding domain of the RNA helicase DHX36 and function to increase RNA accumulation. *Mol Cell Biol* 31:736–743. <https://doi.org/10.1128/MCB.01033-10>
- Siddiqui-Jain A, Grand CL, Bearss DJ, Hurley LH (2002a) Direct evidence for a G-quadruplex in a promoter region and its targeting with a small molecule to repress c-MYC transcription. *Proceedings of the National Academy of Sciences* 99:11593–11598. <https://doi.org/10.1073/pnas.182256799>
- Siddiqui-Jain A, Grand CL, Bearss DJ, Hurley LH (2002b) Direct evidence for a G-quadruplex in a promoter region and its targeting with a small molecule to repress c-MYC transcription. *Proceedings of the National Academy of Sciences of the United States of America* 99:11593–11598. <https://doi.org/10.1073/pnas.182256799>
- Soldatenkov VA, Vetcher AA, Duka T, Ladame S (2008) First evidence of a functional interaction between DNA quadruplexes and poly(ADP-ribose) polymerase-1. *ACS chemical biology* 3:214–219. <https://doi.org/10.1021/cb700234f>
- Song J, Perreault JP, Topisirovic I, Richard S (2016) RNA G-quadruplexes and their potential regulatory roles in translation. *Translation (Austin)* 4:e1244031. <https://doi.org/10.1080/21690731.2016.1244031>
- Spiegel J, Adhikari S, Balasubramanian S (2020) The structure and function of DNA G-quadruplexes. *Trends Chem* 2:123–136. <https://doi.org/10.1016/j.trechm.2019.07.002>
- Su N, Byrd AK, Bharath SR, Yang O, Jia Y, Tang X, Ha T, Raney KD, Song H (2019) Structural basis for DNA unwinding at forked dsDNA by two coordinating Pif1 helicases. *Nat Commun* 10:5375. <https://doi.org/10.1038/s41467-019-13414-9>
- Sun H, Yabuki A, Maizels N (2001) A human nuclease specific for G4 DNA. *Proceedings of the National Academy of Sciences of the United States of America* 98:12444–12449. <https://doi.org/10.1073/pnas.231479198>
- Sun ZY, Wang XN, Cheng SQ, Su XX, Ou TM (2019) Developing Novel G-Quadruplex Ligands: from Interaction with Nucleic Acids to Interfering with Nucleic Acid Protein Interaction. *Molecules* 24:396–424. <https://doi.org/10.3390/molecules24030396>
- Tajrishi MM, Tuteja R, Tuteja N (2011) Nucleolin: The most abundant multifunctional phosphoprotein of nucleolus. *Commun Integr Biol* 4:267–275. <https://doi.org/10.4161/cib.4.3.14884>
- Tauchi T, Shin-ya K, Sashida G, Sumi M, Okabe S, Ohyashiki JH, Ohyashiki K (2006) Telomerase inhibition with a novel G-quadruplex-interactive agent, telomestatin: in vitro and in vivo studies in acute leukemia. *Oncogene* 25:5719–5725. <https://doi.org/10.1038/sj.onc.1209577>
- Thakur RK, Kumar P, Halder K, Verma A, Kar A, Parent J-L, Basundra R, Kumar A, Chowdhury S (2009) Metastases suppressor NM23-H2 interaction with G-quadruplex DNA within c-MYC promoter nuclease hypersensitive element induces c-MYC expression. *Nucleic Acids Res* 37:172–183. <https://doi.org/10.1093/nar/gkn919>
- Thandapani P, O'Connor TR, Bailey TL, Richard S (2013) Defining the RGG/RG motif. *Mol Cell* 50:613–623. <https://doi.org/10.1016/j.molcel.2013.05.021>
- Tippana R, Chen MC, Demeshkina NA, Ferré-D'Amaré AR, Myong S (2019) RNA G-quadruplex is resolved by repetitive and ATP-dependent mechanism of DHX36. *Nat Commun* 10:1855. <https://doi.org/10.1038/s41467-019-09802-w>
- Tippana R, Hwang H, Opresko PL, Bohr VA, Myong S (2016) Single-molecule imaging reveals a common mechanism shared by G-quadruplex-resolving helicases. *Proceedings of the National Academy of Sciences* 113:8448–8453. <https://doi.org/10.1073/pnas.1603724113>
- Tippana R, Xiao W, Myong S (2014) G-quadruplex conformation and dynamics are determined by loop length and sequence. *Nucleic Acids Res* 42:8106–8114. <https://doi.org/10.1093/nar/gku464>
- Vasilyev N, Polonskaia A, Darnell JC, Darnell RB, Patel DJ, Sengavon A (2015) Crystal structure reveals specific recognition of a G-quadruplex RNA by a β -turn in the RGG motif of FMRP. *Proceedings of the National Academy of Sciences of the United States of America* 112:E5391–E5400. <https://doi.org/10.1073/pnas.1515737112>
- Vizcaino C, Mansilla S, Portugal J (2015) Sp1 transcription factor: A long-standing target in cancer chemotherapy. *Pharmacol Ther* 152:111–124. <https://doi.org/10.1016/j.pharmthera.2015.05.008>
- von Hacht A, Seifert O, Menger M, Schütze T, Arora A, Konthur Z, Neubauer P, Wagner A, Weise C, Kurreck J (2014) Identification and characterization of RNA guanine-quadruplex binding proteins. *Nucleic Acids Res* 42:6630–6644. <https://doi.org/10.1093/nar/gku290>
- Wang C, Zhang J, Yin J, Gan Y, Xu S, Gu Y, Huang W (2021) Alternative approaches to target Myc for cancer treatment. *Signal Transduct Target Ther* 6:117. <https://doi.org/10.1038/s41392-021-00500-y>
- Wang Y, Patel DJ (1993) Solution structure of the human telomeric repeat d[AG3(T2AG3)3] G-tetraplex. *Structure* 1:263–282. [https://doi.org/10.1016/0969-2126\(93\)90015-9](https://doi.org/10.1016/0969-2126(93)90015-9)
- Weintraub AS, Li CH, Zamudio AV, Sigova AA, Hannett NM, Day DS, Abraham BJ, Cohen MA, Nabet B, Buckley DL, Guo YE, Hnisz D, Jaenisch R, Bradner JE, Gray NS, Young RA (2017) YY1 Is a Structural Regulator of Enhancer-Promoter Loops. *Cell* 171:1573–1588.e1528. <https://doi.org/10.1016/j.cell.2017.11.008>
- Wolfe AL, Singh K, Zhong Y, Drewe P, Rajasekhar VK, Sanghvi VR, Mavrakis KJ, Jiang M, Roderick JE, Van der Meulen J (2014) RNA G-quadruplexes cause EIF4A-dependent oncogene

- translation in cancer. *Nature* 513:65–70. <https://doi.org/10.1038/nature13485>
- Wu CG, Spies M (2016) G- quadruplex recognition and remodeling by the FANCD1 helicase. *Nucleic Acids Res* 44:8742–8753. <https://doi.org/10.1093/nar/gkw574>
- Wu W-Q, Zhang M-L, Song C-P (2020) A comprehensive evaluation of a typical plant telomeric G-quadruplex (G4) DNA reveals the dynamics of G4 formation, rearrangement, and unfolding. *J Biol Chem* 295:5461–5469. <https://doi.org/10.1074/jbc.RA119.012383>
- Xiao C-D, Ishizuka T, Xu Y (2017) Antiparallel RNA G-quadruplex Formed by Human Telomere RNA Containing 8-Bromoguanosine. *Sci Rep* 7:6695. <https://doi.org/10.1038/s41598-017-07050-w>
- Xiao T, Li X, Felsenfeld G (2021) The Myc-associated zinc finger protein (MAZ) works together with CTCF to control cohesin positioning and genome organization. *Proc Nat Acad Sci* 118:e2023127118. <https://doi.org/10.1073/pnas.2023127118>
- Xin H, Liu D, Wan M, Safari A, Kim H, Sun W, O'Connor MS, Songyang Z (2007) TPP1 is a homologue of ciliate TEBP-beta and interacts with POT1 to recruit telomerase. *Nature* 445:559–562. <https://doi.org/10.1038/nature05469>
- Xodo L, Paramasivam M, Membrino A, Cogoi S (2008) Protein hnRNPA1 binds to a critical G-rich element of KRAS and unwinds G-quadruplex structures: implications in transcription. *Nucleic Acids Symp Ser* 2004:159–160. <https://doi.org/10.1093/nass/nrn081>
- Xu H, Di Antonio M, McKinney S, Mathew V, Ho B, O'Neil NJ, Santos ND, Silvester J, Wei V, Garcia J, Kabeer F, Lai D, Soriano P, Banáth J, Chiu DS, Yap D, Le DD, Ye FB, Zhang A et al (2017) CX-5461 is a DNA G-quadruplex stabilizer with selective lethality in BRCA1/2 deficient tumours. *Nat Commun* 8:14432. <https://doi.org/10.1038/ncomms14432>
- Yangyuoru PM, Bradburn DA, Liu Z, Xiao TS, Russell R (2018) The G-quadruplex (G4) resolvase DHX36 efficiently and specifically disrupts DNA G4s via a translocation-based helicase mechanism. *J Biol Chem* 293:1924–1932. <https://doi.org/10.1074/jbc.M117.815076>
- Yokono M, Saegusa N, Matsushita K, Sugiura Y (1998) Unique DNA binding mode of the N-terminal zinc finger of transcription factor Sp1. *Biochemistry* 37:6824–6832. <https://doi.org/10.1021/bi9727646>
- Zhang AYQ, Bugaut A, Balasubramanian S (2011) A sequence-independent analysis of the loop length dependence of intramolecular RNA G-quadruplex stability and topology. *Biochemistry* 50:7251–7258. <https://doi.org/10.1021/bi200805j>
- Zhang D-H, Fujimoto T, Saxena S, Yu H-Q, Miyoshi D, Sugimoto N (2010) Monomorphic RNA G-quadruplex and polymorphic DNA G-quadruplex structures responding to cellular environmental factors. *Biochemistry* 49:4554–4563. <https://doi.org/10.1021/bi1002822>
- Zhou R, Zhang J, Bochman ML, Zakian VA, Ha T (2014) Periodic DNA patrolling underlies diverse functions of Pif1 on R-loops and G-rich DNA. *eLife* 3:e02190. <https://doi.org/10.7554/eLife.02190>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.