

Review

Seven essential questions on G-quadruplexes

Sebastian L.B. König^{1,2}, Amanda C. Evans¹ and Julian L. Huppert^{1,*}

¹ Cavendish Laboratory, University of Cambridge, JJ Thomson Avenue, Cambridge, CB3 0HE, UK

² Ecole Supérieure de Biotechnologie de Strasbourg (ESBS) Boulevard Sébastien Brandt, 67412 Illkirch-Graffenstaden, France

* Corresponding author
e-mail: jlh29@cam.ac.uk

Abstract

The helical duplex architecture of DNA was discovered by Francis Crick and James Watson in 1951 and is well known and understood. However, nucleic acids can also adopt alternative structural conformations that are less familiar, although no less biologically relevant, such as the G-quadruplex. G-quadruplexes continue to be the subject of a rapidly expanding area of research, owing to their significant potential as therapeutic targets and their unique biophysical properties. This review begins by focusing on G-quadruplex structure, elucidating the intermolecular and intramolecular interactions underlying its formation and highlighting several substructural variants. A variety of methods used to characterize these structures are also outlined. The current state of G-quadruplex research is then addressed by proffering seven pertinent questions for discussion. This review concludes with an overview of possible directions for future research trajectories in this exciting and relevant field.

Keywords: DNA; G-quadruplexes; promoter; proto-oncogenes; RNA; telomeres; transcription.

Introduction

A brief background on G-quadruplex structure

Nucleic acids are composed of three building blocks: cyclic pentoses are interconnected via phosphodiester bonds and form the phosphosugar backbone. The heterocyclic bases guanine (G), thymine (T), cytosine (C), adenosine (A) and uracil (U) are attached to the sugar rings and provide the necessary chemical functionality to ensure substantial intermolecular and intramolecular forces (hydrogen bonding, π -stacking) between base pairs, resulting in stable secondary and tertiary nucleic acid architecture. The bases experience very specific hydrogen bonding interactions with each other within the well-known double helical structure, known as ‘Watson-Crick base pairing’ (1) between G and C and

between T and A. In RNA, G and C and A and U can pair up in a similar manner.

Watson-Crick hydrogen bonding, between G and C in both DNA and RNA, results from interactions between three heteroatoms on the respective nitrogen base. In the case of G, N¹, N² and O⁶ are the key atoms involved in establishing hydrogen bonding, and the face of G on which they are located is known as the Watson-Crick face (Figure 1, blue atoms). However, alternative hydrogen bonding contacts are possible for G using N² and N³ along what is known as the Hoogsteen face, which can lead to triple- or four-stranded structures (Figure 1, red atoms).

It is this Hoogsteen face, along with the Watson-Crick face, that is critical in the formation of a G-quartet (G₄), a structure involving 4 G bases hydrogen-bonded to each other (Figure 2A). Crystallographic evidence for G-quartets, or G-tetrads, was first provided by Gellert et al. in 1962 (2). G-quartets have a large π -surface, approximately twice as large as the one of a Watson-Crick base pair; hence, π -stacking is highly favored and increases the stability of these complexes. Stacks of G-quartets π -stacked on top of each other are commonly known as G-quadruplexes and are the focus of this brief review (Figure 2B).

G-quadruplex topology

The formation of G-quadruplexes is strongly stabilized by the presence of monovalent (and divalent) cations which coordinate within the quadruplex core and interact electrostatically with the O⁶ atoms (Figure 2C) (3). The stabilization of G-quadruplex formation using such monovalent cations as Na⁺ and K⁺ cations is particularly well precedented. The ability of a cation to act as a coordinatively stabilizing agent within the quadruplex core is dependent upon two factors: the amount of energy that is released upon electrostatic coordination of the cation with the O⁶ atoms of the G bases and the energy needed to dehydrate the solvated cation (4, 5). K⁺ is the strongest known coordinating monocation for G-quadruplexes (4–9).

G-quadruplex architecture is very diverse and a wide range of different topologies are known to exist. G-quadruplexes can form from either intermolecular or intramolecular interaction of G bases: unimolecular (Figure 3A) (10), bimolecular (Figure 3B) (10) and tetramolecular forms (Figure 3C–F) (10) have all been observed using crystallographic (11–15) or nuclear magnetic resonance (NMR) techniques (16–28).

It must be taken into account that nucleic acids possess directionality (5′ end to 3′ end). All four strands can be in one direction (5′ to 3′) and thus ‘parallel’ or some/all can run in different directions and hence ‘antiparallel’. The dif-

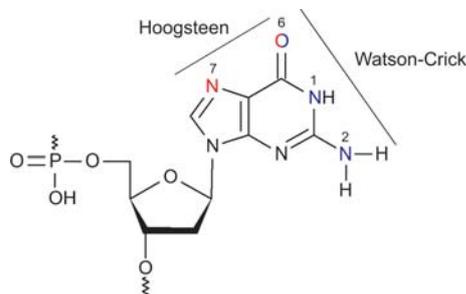


Figure 1 Molecular structure of a G nucleotide unit.

ferent relative conformations of the strands are shown in Figure 3A–E. At a molecular level, the different directionality of the strands is related to the glycosidic angle χ (O4'-C1'-N9-N4), which positions the G base and the phosphosugar backbone relative to each other (10). The different base geometries and the corresponding glycosidic angles are illustrated in Figure 4.

In intermolecular G-quadruplexes, contiguous strands are aligned next to each other without covalent connection. In intramolecular G-quadruplexes, however, loops are involved in connecting the G-bases that constitute the G-quadruplex and thus add a further level of diversity to topology. These loops can adopt several distinct conformations, as depicted in Figure 3B–E. Additionally, all quadruplex structures display grooves, which have been defined as the cavities between two adjacent G-bases bounded by the phospho-

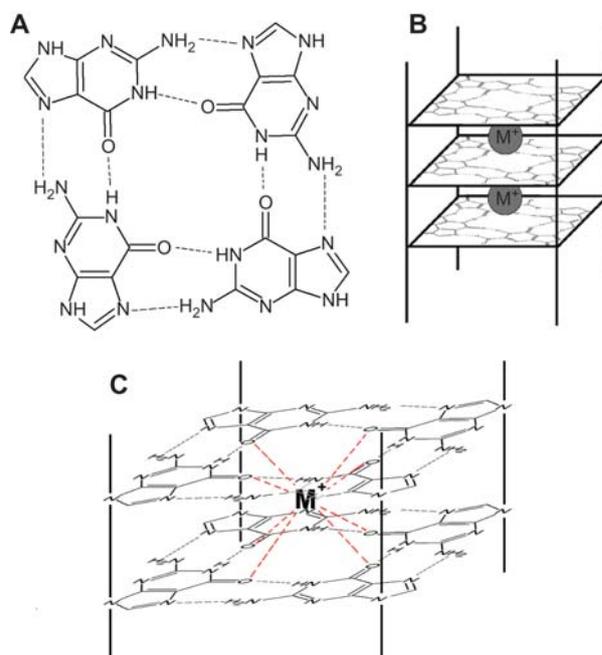


Figure 2 (A) The structure of a G-quartet. Hydrogen bonds involving the Hoogsteen face of the guanine base generate hydrogen bonding interactions. (B) Schematic structure of a G-quadruplex. (C) Electrostatic interactions between the cation and the O⁶ atoms of two G-quartets.

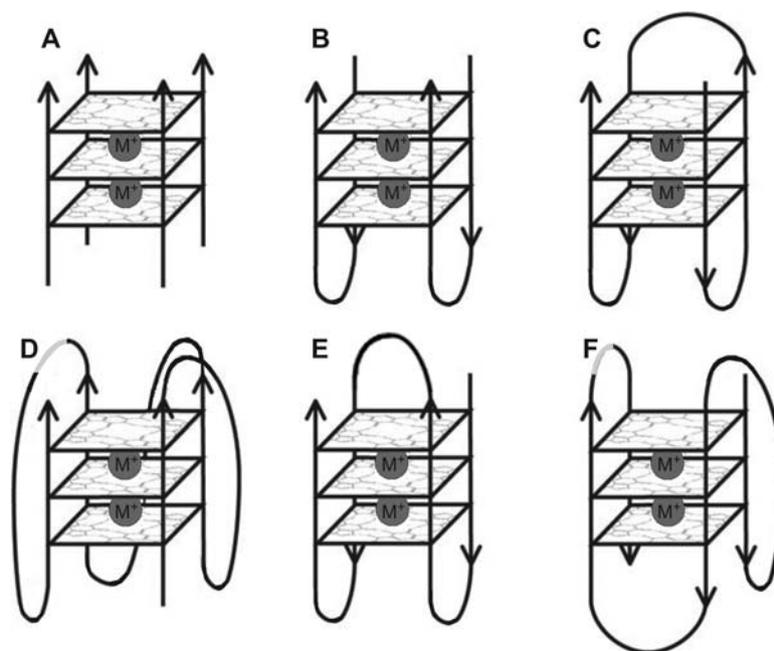


Figure 3 G-quadruplex molecularities.

(A) Tetramolecular, (B) bimolecular, (C–F) intramolecular. Different loop topologies: (B, C) lateral loops, (D) double-stranded reverse loops, (E) two lateral loops and one diagonal one. Different strand topologies: (A) parallel G-quadruplex, all four strands have the same direction; (B) antiparallel G-quadruplex, two strands for each direction, each one adjacent to a parallel strand and an antiparallel one; (C) antiparallel G-quadruplex, two strands in each direction, each one adjacent to two antiparallel strands; (F) antiparallel G-quadruplex, three strands in one direction, one strand in the other (figure modified from Huppert (10)). Permission granted.

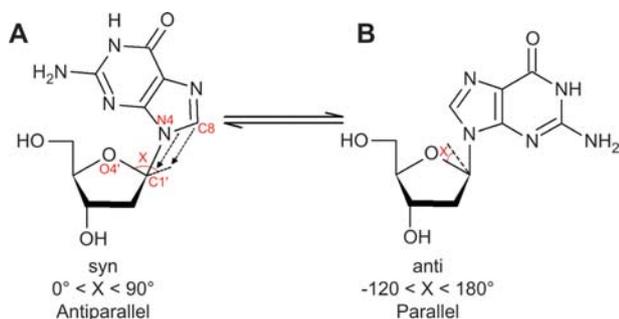


Figure 4 Preferred base geometries.

Interconversion between the anti- and syn-conformation occurs through rotation around the glycosidic bond χ , which is defined as the angle $O4'-C1', N4-C8$.

diester backbones. Groove dimensions are variable and depend on overall topology and the nature of the loops (29).

The different possible strand topologies, the flexible number of G-tetrads, and the variable composition and conformations of loops all make the G-quadruplex a very interesting nucleic acid structural motif. For more detail on G-quadruplex structure and topology, please see the review by Burge et al. (29).

Methods to study G-quadruplexes

The stability of a G-quadruplex, and hence its propensity for formation, depend upon factors such as cation type, cation concentration, solvent, presence of G-quadruplex-binding ligand and molecular crowding. There is a range of techniques that can be used to study G-quadruplex formation and stability. Complete structural elucidation on the atomic level can be achieved using either X-ray diffraction analysis or NMR spectroscopy (10). Very high resolutions below 1 Å can be achieved using X-ray diffraction (10); however, as the samples in the solid state, crystal structure determination does not necessarily reflect solution-phase nucleic acid configuration. NMR spectroscopic analysis is performed in solution and can lead to similar resolutions, but requires highly concentrated pure samples, and base substitution or isotope labeling might also be necessary to elucidate full structural detail (10). Given that the results from both methods can be in disagreement with each other, they should be regarded as complementary analytical techniques and ideally would be used in parallel (30, 31).

Some other techniques used to analyze G-quadruplex structure are rather specific for one particular feature of the G-quadruplex structure and cannot be used to resolve the whole structure. Yet they are not less useful, because they often do not require extensive sample preparation. Among a number of techniques, UV-spectroscopy, circular dichroism and fluorometric analysis (fluorescence resonance energy transfer techniques) are the most widely used. The formation and stability of a G-quadruplex structure at certain temperatures can be assessed by monitoring the absorption at 295 nm and is characterized by a strong hypochromic shift (Figure 5) (32–34). Such UV analysis can provide melting tem-

perature determination and, following detailed van't Hoff analysis, further details about the thermodynamics of quadruplex formation such as ΔG , ΔH and ΔS (10). Circular dichroism (CD) can be used to distinguish all-parallel from antiparallel structures (Figure 6) (35); however, prediction of quadruplex structure from CD spectral data is highly complicated, as only limited theory exists to date (36). Fluorescence resonance energy transfer (FRET) can be used to report the separation of any two points of a G-quadruplex. This technique is as frequently used as the UV melting experiments described earlier. However, it is costly to add two fluorescent probes to a G-quadruplex and they might alter the structure and its biophysical properties (10). For a complete overview on these techniques, in addition to further information about other analytical techniques for G-quadruplex structural elucidation, please see the review published by Ou et al. (37).

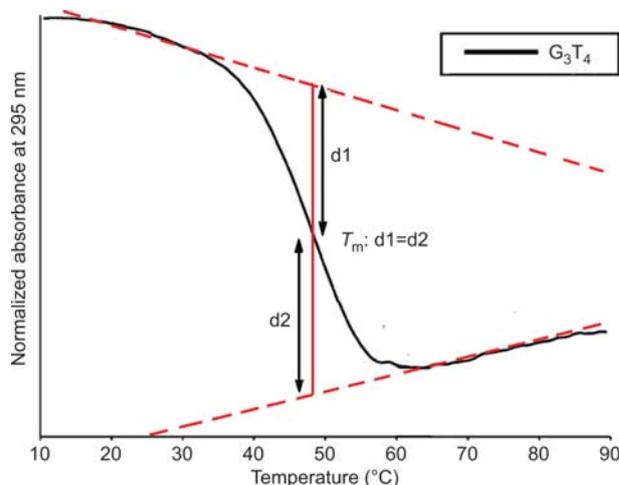


Figure 5 G-quadruplexes undergo a hypochromic shift at 295 nm upon melting.

The midpoint ($d1=d2$) is called melting temperature T_m . The dotted lines correspond to baselines. Figure modified from Hazel et al. (32). Permission granted.

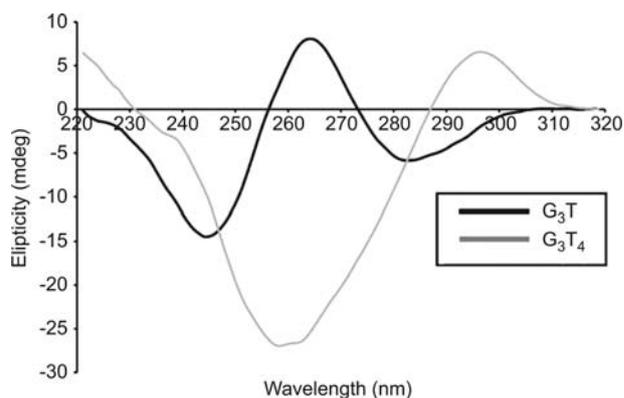


Figure 6 Sample CD spectra depicting G-quadruplex formation. Parallel quadruplex: maximum at 260 nm, minimum at 240 nm (G_3T). Antiparallel quadruplex: maximum at 295 nm, minimum at 260 nm (G_3T_4) (32).

A significant body of research has been published exploring G-quadruplex formation and topology. But why is it so important to study and understand G-quadruplex structure and stability? What might the biological role of G-quadruplexes be and how might an understanding of structure and stability assist in elucidating and optimizing function? To what extent are the structure and the stability of a G-quadruplex-forming sequence predictable? What is the evidence for the *in vivo* formation of this exotic secondary structure and what role(s) might it play in a biological organism? What are the potential applications of this motif, biological and biotechnological? This review intends to address each of these essential questions in turn and thus outline the current state of G-quadruplex research.

Is it possible to predict G-quadruplex structure and stability?

Based on the data obtained using the biophysical analytical methods described above, it has been demonstrated that it is possible to approximately predict whether a sequence might be likely to form a quadruplex *in vitro* using ‘the folding rule’ (38). This rule predicts that any sequence $G_n X_{1-7} G_n X_{1-7} G_n X_{1-7} G_n$ (G =guanine, $3 < n < 5$, X =any base) will form a G-quadruplex and has been implemented algorithmically in a tool called quadparser, which is freely available to download from the internet (www.quadruplex.org) (39). This algorithm can be used to scan a sequence of putative G-quadruplex sequences (PQSs); indeed, a scan of the human genome predicted the existence of 376 000 G-quadruplex forming sequences (38). However, these predictive data should be interpreted carefully: quadparser can only scan nucleic acid sequences, which does not actually entail proof of formation either *in vitro* or *in vivo* formation. A powerful prediction tool would then take into account all factors affecting G-quadruplex stability. For further details on computational prediction of G-quadruplex formation, please see the reviews published by Todd (40), Ryvkin et al. (41) and Huppert (42).

The stability of any given particular G-quadruplex structure can be affected by numerous factors. The stabilizing effect of cations has already been described; however, there remains some confusion regarding possible quadruplex destabilization upon exposure to lithium ions. Risitano et al. has reported that lithium ions exert a strong inhibitory effect of Li^+ on G-quadruplex formation but results from the Lacroix group do not confirm this effect (43–45). Thermodynamic G-quadruplex stability is known to increase with the number of G-quartets involved due to extended π - π stacking (23). Additionally, G-quadruplex topology also affects its stability: thermodynamic and NMR data suggest that the composition and length of the loops can also have an impact on G-quadruplex stability. The parallel strand conformation has been found to be thermodynamically more favorable than the antiparallel conformation for short loops, whereas long loops give rise to antiparallel G-quadruplexes (Figure 7) (23, 45–51).

This phenomenon leads to structural polymorphism for intermediate loop lengths (Figure 8) (52). In fact, a multitude of different G-quadruplex topologies have been found to exist for one G-quadruplex forming sequence, depending upon environmental conditions (53). For example, the human telomeric motif can adopt an antiparallel configuration in the presence of sodium ions, but has been shown to adopt an all-parallel conformation in the presence of potassium ions (14, 54, 55). Furthermore, sequences exist that form very unusual and unpredictable structures: c-kit87up [d(AGGGAGGGCGCTGGGAGGAGGG)] folds in such a way that a singular non-G-tract G participates in the formation of a G-tetrad core, despite the presence of four G-tracts in the sequence that each contain three consecutive G units (56, 57).

Within a double-stranded context, such as the genome, G-quadruplexes exist in constant competition with duplex DNA: indeed, the interconversion between duplex and quadruplex has been studied for a range of sequences (6, 58–62). Thus, why would G-quadruplexes form at all in the presence of the complementary strand, given the high stability of the duplex form? Results suggest that under near-physiological conditions, double-stranded DNA predominantly does not form a G-quadruplex. However, there are additional factors within a cell that can shift the equilibrium to favor quadruplex formation, such as chromatin structure, duplex supercoiling, molecular crowding and the presence of specific proteins and small molecule ligands capable of interacting with a G-quadruplexes (37, 63–66). Sun and Hurley have recently shown that negative superhelicity favors the formation of G-quadruplexes in double-stranded DNA under near-physiological conditions (67).

Predicting G-quadruplex stability is therefore rather difficult, even for simple examples. Several factors affecting G-quadruplex stability have yet to be studied and the factors

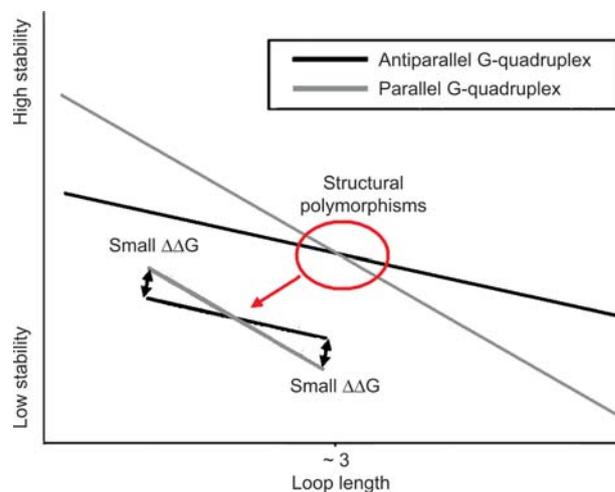


Figure 7 G-quadruplex stability as a function of loop length. G-quadruplexes with short loops preferably adopt parallel conformation; for long loops, the antiparallel conformation is energetically favored. Structural polymorphisms exist mainly for quadruplexes with intermediate loop length.

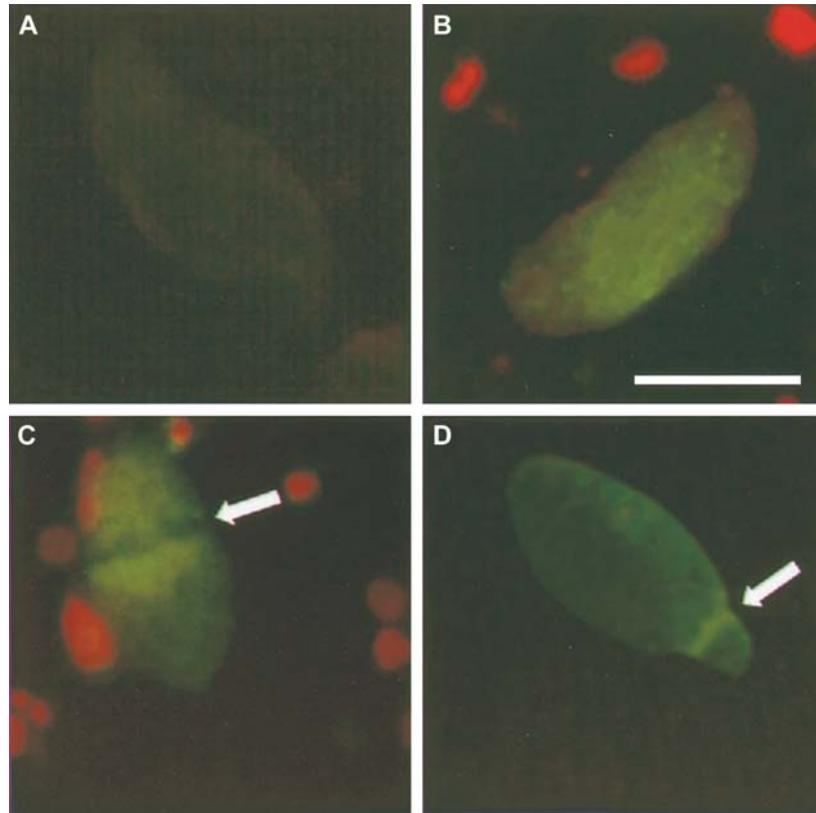


Figure 8 Antibody immunofluorescence staining and EM of *Stylotricha lemnae* nuclei. *Stylotricha* macronuclei and micronuclei were probed with the selected telomeric G-quadruplex DNA-specific scFv antibody probes Sty3 (binds to parallel G-quadruplexes) and Sty49 (no topology specificity). Incubation of the macronuclei with Sty3 did not produce a signal above the background level where no scFv was used (A). Sty49, by contrast, produced a clear signal when applied to vegetative macronuclei (B). Replication and telomere elongation occurs in the macronucleus in a morphologically distinct region, the replication band; however, no binding of Sty49 to this nuclear region was observed (C), even though the DNA concentration, and thus the concentration of telomere sequences, is higher there than elsewhere in the macronucleus, as evidenced by *in situ* hybridization with an FITC-labeled telomeric probe (D). As an internal positive control, in some experiments a few algae were added to the nuclei (red fluorescence in B and C due to the presence of chlorophyll within the algae). Figure modified from Schaffitzel et al. (52). Permission granted.

that have been studied can often not be quantified energetically. However, the accurate prediction of G-quadruplex stability from its sequence is necessary to scan genomes for PQSs with a low error rate and therefore of fundamental interest. Stegle et al. proposed a new machine-learning approach to predict the stability of an unknown G-quadruplex from an existing dataset of experimental values (68). This approach could be the key to an improved G-quadruplex prediction tool.

Do G-quadruplexes form *in vivo*?

Several approaches to determine the occurrence of G-quadruplexes under physiological conditions have been reported in the literature. *In silico* studies have found that both prokaryotic and eukaryotic genomes contain many PQSs (39, 69, 70). Within the human genome, these sequences (approximately 376 000) often appear to correlate with functional genomic regions (71). They are particularly abundant in both

the G-rich telomere sequences found at the ends of chromosomes and in eukaryotic gene promoters (39, 56, 70, 72, 73). However, one of the fundamental problems with using quadparser as a means to determine where G-quadruplexes form throughout the genome is that it does not take into account the fact that a complementary strand within a stretch of double-stranded DNA will influence the stability of G-quadruplex formation. Given that the equilibrium between double-stranded and G-quadruplex states mainly lies to the left, the first experimental approaches initially focused on G-quadruplexes formation in PQSs that would necessarily occur in single-stranded form, such as the telomeric 3'-overhang or RNA (71). For G-quadruplex formation to occur in the presence of duplex DNA, canonical Watson-Crick base pairing must be made less favorable (71). The two strands must be separated and the single strands stabilized and as this is an energetically unfavorable process, enhanced stability of the single-stranded structure must be promoted.

So far, research efforts have concentrated on the telomeric PQSs, in particular the telomeric 3'-overhang sequences.

Ciliates are an ideal model organism for studying the formation of G-quadruplexes in this region, as their cells contain transcriptionally active macronuclei displaying telomeres at higher concentrations than found in other genomes. Anti-telomeric G-quadruplex antibodies were used to specifically bind to antiparallel G-quadruplex conformations formed in the *Stylotricha* 3' telomeric overhang d(TTTTGGGG)₂ *in vitro*, followed by *in situ* immunostaining using antibodies to allow electron microscopy (EM) visualization (Figure 8) (52). The results showed that G-quadruplexes were absent during replication, suggesting that they must be resolved during telomere replication (Figure 9) (74). This is supportive of G-quadruplexes repressing telomerase activity and providing a capping function for telomeres, preventing them from being continually elongated as they are in tumor cells (52, 71). Indirect approaches have provided evidence for the formation of G-quadruplexes in the 3'-overhang region of ciliates: assuming that telomeric G-quadruplexes form *in vivo* and do possess biological functionality, one would expect their occurrence to be tightly regulated. Two telomere-end binding proteins, TEBP α and TEBP β , that specifically recognize the d(TTTTGGGG)₂-overhang of *Oxytricha*, have been identified (52, 73). Knocking down either of these proteins in *Stylotricha* using RNAi was observed by EM to lead to loss of G-quadruplex antibody staining (75). The hypothesis that G-quadruplexes play a regulatory role in replication, especially during the S-phase of the cell cycle, is further supported by the fact that the activity of TEBP β is subjected to cell cycle-dependent regulation (75).

However, *in vivo* evidence in the 3' telomeric overhang region in vertebrate cells has not been as thoroughly investigated. Lower telomere concentration in vertebrates does not allow for the use of immunostaining techniques for visualization – hence these methods have proven less definitive for demonstrating *in vivo* vertebrate G-quadruplex formation. Support for the presence of G-quadruplexes in human telomeres has been recently reported using a fluorescent cyanine dye that specifically binds to G-quadruplexes, allowing their visualization *in vitro* (76). Interestingly, the G-quadruplex-folding activity that occurs in TEBP β is not conserved in its vertebrate homolog TPP1 (71, 77).

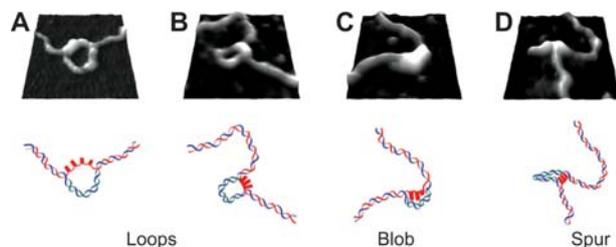


Figure 9 Zoomed atomic force microscopy (AFM) images of G-quadruplex structures with corresponding diagrammatic representations of the DNA arrangement. Areas (140–140 nm²) showing regions of transcribed plasmids containing loops (A, B), a blob (C) and a spur (D).

The blue lines in the diagrams represent the non-coding strand of the DNA, the red lines represent the G-rich coding DNA strand and the green lines represent hybridized mRNA (74).

G-quadruplexes could potentially form in other duplex regions of the telomere region, but this duplex structure would need to be separated in order to favor G-quadruplex formation. Separation events in the double-stranded telomere regions typically only occur during replication and transcription of telomeric DNA into telomeric-repeat-containing RNA (TERRA) (78). If G-quadruplexes were formed during these cellular events telomere-specific helicases would be needed to unravel the quadruplex to ensure that replication and transcription is not blocked. Knockout experiments on the ‘regulator of telomere length’ gene (RTEL) in mice and embryonic stem cells resulting in chromosome-end fusions suggest that the protein product of the RTEL gene might function by unwinding secondary structures such as G-quadruplexes (79). Moreover, the telomerase-specific helicases WRN and BLM (*Human*) and Sgs1p (*Saccharomyces cerevisiae*) have been shown to unwind G-quadruplexes *in vitro* and contain one conserved G-quadruplex-binding domain (80–83).

Is there also evidence for the occurrence of G-quadruplexes in other parts of the genome? Some G-rich genes form an unusual post-transcriptional structure in which the RNA produced remains bound to the template DNA strand; this leaves the coding strand free and causes loop formation (84). These structures are known as RNA/DNA hybrids and the loops are known as G-loops. Duquette et al. have shown that intracellular transcription of G-rich regions can lead to the formation of G-quadruplexes in these G-loops *in vitro* using electron microscopy (84). These findings were supported by Neaves et al. who have demonstrated that transcription of the plasmid pPH600 results in RNA/DNA hybrid formation and folding of the G-loop into putative G-quadruplex structures using atomic force microscopy (AFM) (Figure 9) (74). Further indirect evidence for non-telomeric G-quadruplex formation has been provided by knockdown experiments for a number of DNA helicases (DOG-1: *Caenorhabditis elegans*, FANCF: *human*) that are thought to unravel G-quadruplex structures (85–87). The downregulation of these helicases results in cells that tend to accumulate large genomic deletions in G-rich sequences predicted to include PQSs (39, 85–87). Recent support for the presence of G-quadruplexes in the promoter regions of genes has also arisen from studies using whole-genome expression microarrays. The G-quadruplex stabilizing ligand TMPyP4 has been demonstrated to affect expression of 1200 genes, 700 of which are known to contain PQSs (88–90), providing further support for the hypothesis that G-quadruplexes provide a regulatory role in transcription. These observations are supported by recent findings of Fernando et al. who developed a single-chain antibody that is selective for G-quadruplex DNA over double-stranded DNA (91). When this antibody is expressed in human cells it significantly affects the expression of a wide variety of genes, in a manner that correlates with the presence of predicted G-quadruplexes.

Messenger RNA contains untranslated regions (UTRs), known as introns, before the start codon (5'-UTR) and after the stop codon (3'-UTR) that are not translated into the protein sequence. There are approximately 3000 mRNAs for

which the formation of PQSs in 5'-UTRs has been predicted using the quadparser algorithm (92, 93). These motifs can function as translational regulators: the Balasubramanian group has reported that the presence of a G-quadruplex in the 5'-UTR of the NRAS proto-oncogene mRNA has been shown to modulate translational activity *in vitro* (92). *In vivo* studies support the existence of mRNA G-quadruplexes: Wieland and Hartig have constructed mRNA-based G-rich elements that mask the ribosome binding site by folding into four-stranded structures. The suppression of gene expression was found to correlate with the stability of inserted G-quadruplexes (94). Arora et al. demonstrated that the Zic-1 RNA G-quadruplex represses protein synthesis inside eukaryotic cells (95).

The significant enrichment of PQSs in the functional parts of the genome and the results of the *in vitro* and *in vivo* studies described above suggest that G-quadruplexes form in living organisms and could very well possess key biological regulatory function within the telomeric and promoter regions of the genome. However, given that most *in vivo* experiments have been carried out on the telomeric regions of ciliates, future investigations should now focus on other putative G-quadruplex-forming sequences of the genome and on the transcriptome. New methods must be developed to directly detect G-quadruplexes in vertebrate cells. For a complete review on *in vivo* evidence for G-quadruplexes, please read the paper by Lipps and Rhodes (71).

What biological role could G-quadruplexes play?

As described in the previous section, computational studies have shown that numerous PQSs exist throughout the functional parts of prokaryotic and eukaryotic genomes, such as promoter regions and telomeres. As these motifs are highly conserved, owing to their significant biological roles, these computational studies suggest with statistical significance that G-quadruplexes are likely to possess necessary biological functionality.

Telomeres are single-stranded repeats of TTAGGG from 12 to 200 bases in length (96). They are found at the ends of chromosomes and are thought to protect them from deterioration. Telomeres gradually shorten with each cell replication cycle; however, an enzyme called telomerase is known to add telomeric repeats of TTAGGG to the 3'-end of the chromosome and thus protect the telomeres from deterioration over time. G-quadruplex formation has been shown to inhibit telomerase action *in vitro* and could therefore act as a negative regulator of telomere regulation (Figure 10A) (97). However, Zhang et al. recently showed that the opposite might be the case: the yeast telomerase subunit Est1p, which is known to recruit telomerase, was shown *in vitro* to promote the formation of G-quadruplexes (98). Telomeric sequences in Est1p-deficient cells were gradually shortened, suggesting that G-quadruplexes might actually play a positive regulatory role in the maintenance of telomere length (98). In either case, G-quadruplexes possibly play a crucial

role in the determination of cell lifetime and they are therefore likely to be involved in cancer. Kim et al. have shown that telomerase is expressed in 80–85% of cancer cells, but such activity is absent in somatic cells (99).

Promoters regulate gene transcription and often serve as binding sites for proteins called transcription factors, which serve to further regulate transcriptional activity. Computational studies have shown that promoter-associated quadruplex motifs are often found either proximal to or overlapping with transcription factor binding sites, suggesting that G-quadruplexes could be cis-acting regulators of gene expression (Figure 10B) (71, 100, 101). The presence of G-quadruplex forming areas in promoter regions has been confirmed *in vitro* for a number of genes: the chicken β -globin, which encodes the β -subunit of the oxygen-transporting hemoglobin complex; the oncogenes c-myc, c-kit, bcl-2, VEGF, RET, Rb and k-ras; and the Hypoxia-inducible factor 1 which responds to changes in available oxygen in the cellular environment (102–109). The effect of a putative G-quadruplex on the transcriptional activity of a gene has only been extensively studied for the NHEIII-promoter region of the oncogene c-myc. This promoter region has been shown to regulate 90% of gene transcription *in vitro* (89). Furthermore, 80% downregulation of c-MYC expression in MiaPaCa-2 cells was observed *in vivo* after 48 h treatment with 100 μ M TMPyP4, a G-quadruplex-binding ligand that stabilizes quadruplex formation (90).

As mentioned in the previous section, the presence of G-quadruplexes in the UTR of mRNA is suggested by the results of several computational and experimental studies (92, 93). The incidence of PQSs in these regions is not random: the Huppert group has shown that their incidence shows significant strand asymmetry and positional bias. In 5'-UTRs, G-quadruplex motifs tend to exist towards the 5'-end of the 5'-UTR supportive of function relating to translation initiation. With regard to 3'-UTRs, G-quadruplex motifs tend to cluster immediately after the 3'-end of the mRNA, particularly in cases of genes that have a proximal gene in the 3' direction. Therefore, G-quadruplexes in these regions could serve as a translational regulator and terminator/upregulator of gene transcription (Figure 10C) (93).

The overall data suggest that G-quadruplexes could be involved in multiple cellular pathways; however, the *in vivo* formation remains unreported. Initial investigations into *in vivo* G-quadruplex formation are supportive of their existence and, indeed, of the fundamental role they can play in gene regulation and expression, but definitive proof of *in vivo* quadruplex formation remains a goal yet to be achieved in order to confirm their biological functions.

Is there an evolutionary pressure on G-quadruplexes?

Regarding the sheer volume of PQSs throughout the human genome, it is unlikely that all PQSs have biological functionality. A study of G-quadruplex genomic evolutionary conservation is therefore of fundamental interest for identifying

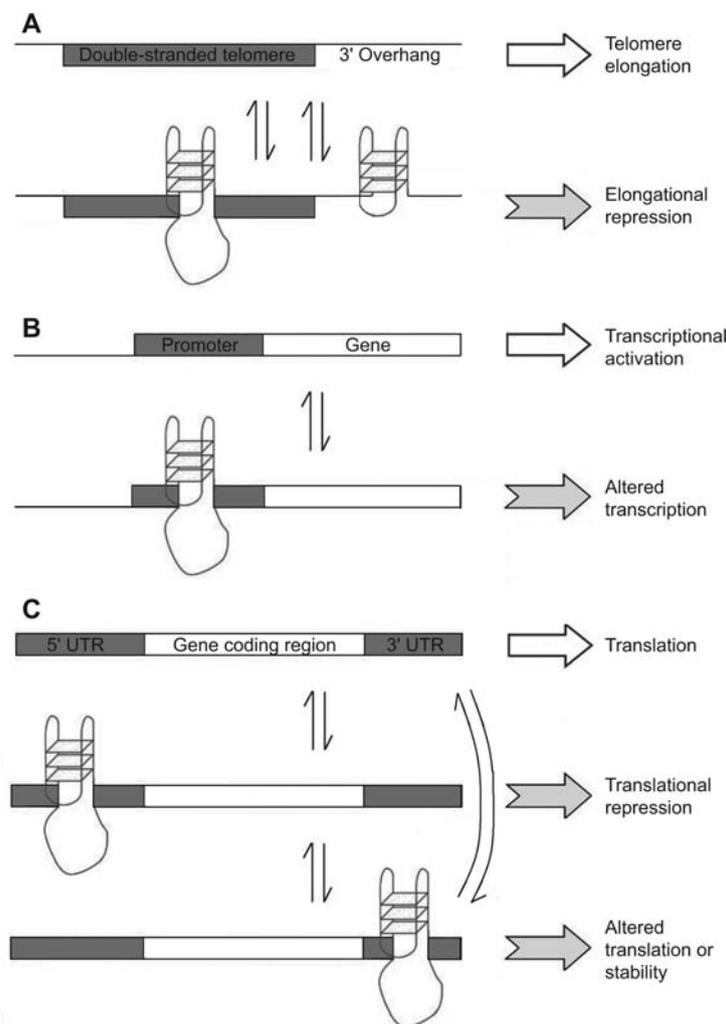


Figure 10 (A) Formation of G-quadruplexes in the telomeric region: G-quadruplexes could prevent telomerase activity. (B) Formation of a G-quadruplex in a promoter: at the simplest level, quadruplex formation can act as a steric block to gene expression transcription machinery. (C) Formation of a G-quadruplex in the UTR of a messenger RNA: G-quadruplexes in the 5'-UTR can downregulate translation, their occurrence in the 3'-UTR can alter gene translation.

functionally significant G-quadruplex sequences. Verma et al. have found more than 700 taxonomically related PQSs in gene promoter regions of human, mouse and rat genomes, suggesting that these sequences are conserved within the promoter (110). Sequence conservation has been found to be more likely in runs of G-bases which are known to form stable G-quadruplexes (111). These results would support G-quadruplex-forming sequences possessing biological function, as conservation often correlates with biological importance.

Statistically, short PQS patterns involving one-base loops should be the most predominate, as confirmed by a series of computational studies carried out by Huppert and Balasubramanian and Todd et al. (38, 112). However, the prevalence of PQSs appears to be subject to evolutionary pressure, as the number of PQSs found was fewer than expected (38). Wong and Huppert recently reported a computational study identifying potential evolutionary effects on PQS conserva-

tion and composition to identify functional relevant sequences and, out of 600 genomes tested, all but one contained fewer PQSs than expected statistically, suggesting that PQS formation is generally suppressed (39). In prokaryotic genomes, PQS loops have often been found to be identical with regard to length and composition with loops of 3, 3, 3 bases or 3, 2, 3 bases particularly abundant in *Escherichia coli* (39). Studies of human variation have shown that G-quadruplexes have in general been lost from the genome since our last common ancestor with chimpanzees, with only a few newly forming; however, these new ones have had a selective advantage and spread rapidly through the population.

These evolutionary studies suggest that G-quadruplexes are statistically likely to play a biological role in the genome. However, no such study has been carried out for the transcriptome so far and it would be of fundamental interest to find out whether G-quadruplexes are also repressed on the

mRNA level or whether there exist preferred topologies consistent with one biological role. The identification of the proteins that bind to the taxonomically related genes reported by Verma et al. would serve to further confirm this hypothesis. Further investigations could also be performed exploring to what extent the G-quadruplex binding regions are conserved in mouse, human and rat, as interspecies similarities and differences could prove to be of further predictive interest.

Do G-quadruplexes have potential as therapeutic targets?

Current tumor chemotherapeutics commonly utilize agents that are cytotoxic despite considerable disadvantages (e.g., generation of tumor resistance and high cellular toxicity) that use of these agents present (113). Such side effects are frequently the result of nonspecific and undesirable interactions of the active agent with the biological system into which it has been introduced. It would therefore be optimal to produce agents that have both specificity of action and efficient and effective reactivity. Indeed, given that replication and transcription of the genome is upregulated in most cancer cells, it would be most desirable to design and develop chemotherapeutics that could specifically address this upregulation at the nucleic acid level. Given the prevalence of G-quadruplexes in proto-oncogenic or telomeric regions of the genome, which have both been indicated to play significant roles in cancer progression, G-quadruplex binding ligands could be used to stabilize or disrupt a G-quadruplex within one of these regions in order to specifically modulate the activity of an enzyme while circumventing the drawbacks of current cancer treatment. Considerable effort has been focused on the development of such binding/stabilizing ligand such as Quarfloxin, a G-quadruplex-binding ligand that has been shown to disrupt the interaction of an rDNA/protein complex and is already in phase II clinical trials (www.clinicaltrials.gov) (114, 115).

Three binding modes are prevalent for G-quadruplex-binding ligands: they can stack upon the terminal G-quartet via π - π interactions (external stacking), intercalate between two G-quartets or bind to the grooves between two adjacent phosphosugar backbones (32). Owing to these common binding modes, most G-quadruplex binding ligands reported to date share certain common structural characteristics. Firstly, they are frequently based on a polycyclic aromatic core that can have effective π - π interactions with the G-quartet. By incorporating one or more cationic side chains (usually protonated amino groups) to the core aromatic rings for electrostatically interaction with the negatively charged phosphodiester backbone of the DNA, G-quadruplex binding can be further enhanced (37, 113, 116–118). The stability provided by these ligand-quadruplex interactions has been confirmed experimentally by several structural NMR and X-ray diffraction studies (13, 119–123). However, the role that other structural G-quadruplex features play in ligand recognition is less well understood. For example, based on the results of

diffraction analyses performed by Balasubramanian and Neidle and Campbell et al. suggest that the flexibility of loops that contain three or more nucleotides could also play a role in ligand specificity (113, 122). For a complete overview on all core ligand structures synthesized to date, please see the review by Ou et al. (37).

Most G-quadruplex-binding ligands have been targeted to bind to telomeric repeats, as these single-stranded regions are believed to be the most likely PQSs to form G-quadruplexes *in vivo* (*vide supra*). Several studies have demonstrated that telomeres could be highly relevant anticancer therapeutic targets (99). Telomeres in cancer cells are known to be gradually elongated by the action of the catalytic telomerase reverse transcriptase subunit hTERT (*human*) and knock-down experiments on hTERT using siRNA, antisense and small molecule inhibitors have been reported to result in reduced growth of cancer cells (117, 124). Currently, one of the tightest known G-quadruplex-binding ligand targeting the human telomeric repeat (TTAGGG)_n is telomestatin (Figure 11A), a naturally occurring macrocycle found in *Streptomyces anulatus* (125). Other potent molecules designed to target telomeric DNA reported to demonstrate *in vivo* anticancer activity are the polycycle RHSP4 (Figure 11B) and the acridine compound BRACO-19 (Figure 11C) (126–128). Cell death due to displacement of bound proteins from the 3'-overhang upon ligand binding, and possibly also telomerase uncapping is often induced after a few days of exposure to these molecules (129, 130).

The hypothesis that G-quadruplexes localized in the promoter regions of genes can affect the transcriptional activity of a gene has been supported by a range of *in vitro* studies for such genes as c-myc, a (transcription factor) and k-ras, (a GTPase involved in signal transduction). This has led to the synthesis of a range of G-quadruplex-binding ligands for controlling proto-oncogenic gene expression. The G-quadruplex-binding molecule TMpyP4 has been demonstrated to downregulate transcription of the c-myc gene as well as other oncogenes and cell cycle controlling genes (Figure 12) (90). This same ligand has also been observed to downregulate transcriptional activity of the k-ras gene (131). The c-kit proto-oncogene contains two G-quadruplex-forming motifs, both of which have been successfully targeted using an isoalloxazine compound (Figure 13) which leads to decreased c-kit mRNA levels in c-kit expressing cell lines (132).

What are the molecular requirements for a small molecule to be a suitable G-quadruplex binding ligand? Strictly speaking, potential drugs would optimally target a singular G-quadruplex substructure within an entire genome. The ligand would possess double specificity: it would both discriminate duplex DNA from G-quadruplex DNA and also favor stabilization of one specific G-quadruplex above all to others. In reality, there still is a long way to go. Synthesis of small molecules possessing much higher affinity for G-quadruplexes than for double-stranded DNA has been achieved (116, 117); however, the sheer number of PQSs throughout the human genome appears to make singular G-quadruplex ligand discrimination a daunting task (38). Selectivity could be increased by designing molecules that recognized the full

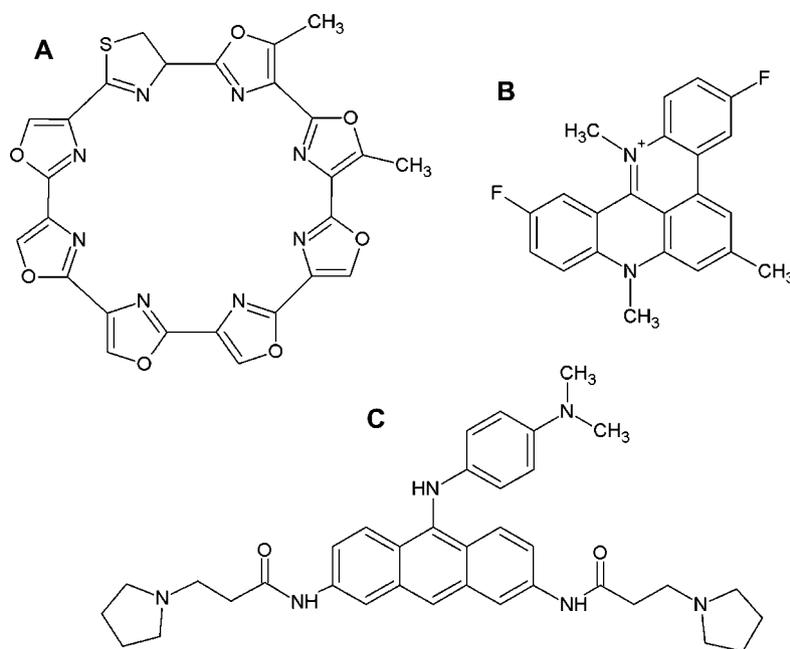


Figure 11 Aromatic molecules such as the natural product telomestatin (A), the fluorinated polycyclic methylenecrithinium salt RHSP4 (B) and the tri-substituted acridine BRACO-19 (C).

range of the topology of G-quadruplexes. Many substructural motifs, for example, the number of G-stacks or the relative directionality of the strands, are not addressed by the small molecules that have been developed to date. A detailed structural analysis of the epitopes used for the immunostaining experiments described previously (52) could lead to the development of a new class of small molecules mimicking G-quadruplex-specific antibodies. Are there other possible strategies to improve selectivity? Modular design could be the key to higher specificity. For example, a fairly selective G-quadruplex ligand could be linked to a molecule that specifically binds to a nucleic acid binding protein in the immediate vicinity of the G-quadruplex structure. However, this approach requires a deeper understanding of the biological environment proximate to G-quadruplexes than is currently available. If one were to succeed in increasing affinity and specificity of such ligands, G-quadruplex-targeting molecules would have the potential to become a very powerful new class of drugs.

Are there other potential applications for G-quadruplexes?

In addition to targeting of G-quadruplex-forming sequences with small molecules or proteins in a pharmacological context, G-quadruplex structures could also be deployed within a nanobiotechnological context. The formation of G-quadruplex polymorphisms is known to be cation type-dependent, and this property has already been harnessed to develop a range of cation sensors. Kong et al. have designed a highly sodium-tolerant potassium sensor using a G-quadruplex-forming oligonucleotide sequence and the fluorescent dye

crystal violet (Figure 14A) (134); this system is also able to discriminate between parallel and antiparallel G-quadruplexes by characteristic fluorescence intensities (Figure 14B) (133). Other groups have employed G-tracts immobilized onto the surface of a gold electrode to measure variations in current intensity in response to potassium-induced G-quadruplex formation (135).

Nanodevices incorporating G-quadruplexes capable of linear or rotational movements have also been investigated. Alberti and Mergny have reported a device that performs an extension-contraction movement involving a single 21-base G-quadruplex-forming oligonucleotide in a duplex-quadruplex equilibrium (136). Hou et al. have immobilized a G-quadruplex-forming DNA sequence onto a synthetic nanopore and showed that pore size varies according to K^+ concentration and that increasing potassium ion concentration triggers a conformational change in the oligonucleotide substructure (137). Such biomimetic systems could prove useful for studying *in vivo* functioning of biological pores.

Self-assembly of G-tracts has also been reported as a means towards constructing scaffolds for functional nanomaterials, with spatial arrangement being directed by synthetic point modifications (138). Wong et al. have reported a non-covalent assembly of 16 guanosine monomers that can be crosslinked to obtain a G-quadruplex-like structure capable of transporting Na^+ across membranes, with $^{23}Na^+$ studies confirming these results (139).

Other potential applications of G-quadruplex-forming sequences have been reported. Aptamers are nucleic acid molecules that specifically bind to various biologically important molecules. The so-called thrombin-binding aptamer [TBA, d(GGTTGGTGTGGTTGG)] specifically binds

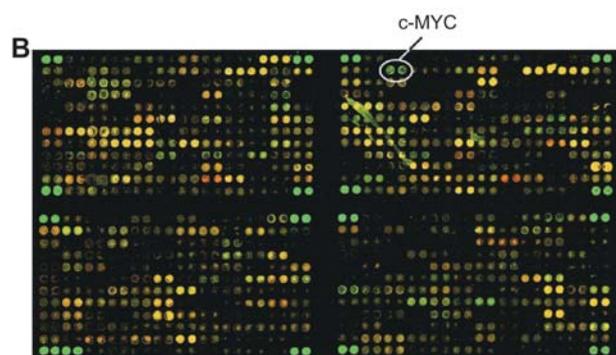
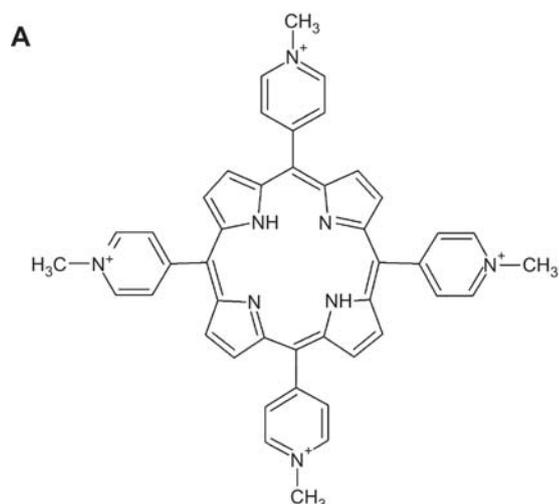


Figure 12 Molecular structure of TMPyP4 (A). cDNA microarray analysis of HeLa S3 cells treated with 100 μM TMPyP4 for 48 h. The c-MYC proto-oncogene, the most consistently downregulated gene in this study, is highlighted. A number of other genes, including c-MYC-regulated genes, proto-oncogenes and cell cycle controllers, were also specifically affected by TMPyP4. Figure modified from Grand et al. (90). Permission granted.

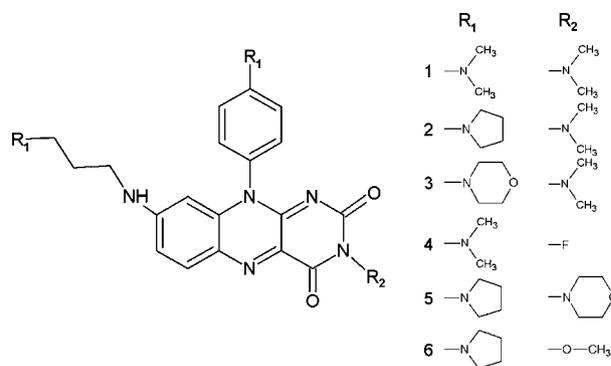


Figure 13 Structures of isoalloxazine ligands using the residues 1–6.

Figure modified from Bejugam et al. (132). Permission granted.

thrombin, a coagulation protein in the blood stream, and inhibits its activity (140–142). With a view to nanodevice application, Yoshida et al. have developed an aptameric enzyme subunit, composed of an enzyme-binding aptamer (TBA) and a target molecule-binding aptamer (143). Binding to the target protein was detected by monitoring thrombin activity (143).

A remarkable property of G-quadruplex-forming sequences might lead to the use of G-rich anticancer aptamers, as discovered by Choi et al. (144, 145). Upon the discovery of a nucleolin-targeted G-rich 26-mer known as AS1411 (144), which has now entered phase II clinical trials, they synthesized a series of G-rich oligonucleotides of various lengths by random incorporation of bases and analyzed them with regard to their CD stability, cellular uptake, protein binding and ant proliferative activity (145). Unexpectedly, those oligonucleotides whose CD spectra were supportive of G-quadruplex formation were shown to strongly inhibit cancer cell growth (145). Moreover, these oligonucleotides were observed to be more stable, with an cellular uptake rate (145). Their cancer-selective antiproliferative activity could

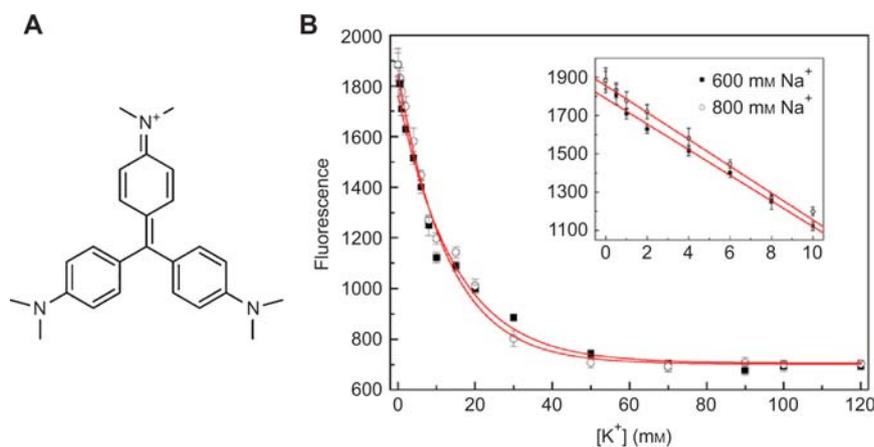


Figure 14 Triphenylmethane dye crystal violet (A). Fluorescence spectrum of the crystal violet/G-quadruplex complex as a function of K^+ ion concentration in the presence of 600 mM Na^+ ion (■) or 800 mM Na^+ ion (○). The insert shows the fluorescence change of the crystal violet/G-quadruplex complex in the K^+ concentration range of 0–10 mM. The solid lines represent linear fits to the data. In this experiment, the fluorescence was detected 40 min after K^+ was added. Figure modified from Kong et al. (133). Permission granted.

be due to these G-rich sequences binding to proteins involved in cancer growth and by G-quadruplex formation causing a modulation of their activity. Therefore, G-quadruplexes themselves could function as an anticancer drug.

What about the complementary strand? An introduction to the *i*-motif

A consideration of G-quadruplexes should also take into account another DNA substructure known to exist in nucleic acid regions which are complementary to quadruplex-forming domains. *In vitro* experimental evidence for the complementary strand to the human telomeric sequence $d(\text{GG-GTTA})_n$, known as an '*i*-motif', was reported by Gehring et al. in 1993 (146). This unusual structure is based on protonated C-C base pairs whose stability is strongly dependent on low pH (Figure 15A) (147, 148). These stack upon each other via π - π interaction in an intercalated manner, forming the *i*-motif (Figure 15B). A variety of *i*-motif structural conformations, involving different intercalations and looping topologies, have been reported (147, 149). No *in vivo* evidence has thus far been reported for the occurrence of *i*-motifs; the non-physiological low pHs at which *i*-motifs are stable could indeed prevent their physiological occurrence. However, they can form at near-physiological pH, and given the relatively small unfavorable free energy at physiological pH, it is not hard to imagine that small changes in supercoiling, bending, local pH, ligand binding and other factors could induce stable *i*-motif formation *in vivo*. Potential *i*-motif forming sequences are abundant in the biologically relevant part of the genome; owing to their complementarity to PQSs, the prevalence of their occurrence will be just as frequent as that found for all duplex PQSs. Interestingly, C-rich sequences appear in the centromere of some human chromosomes: Gallego et al. have reported that the CENP-B box, the binding site of centromere protein B, folds back on itself and forms a dimeric *i*-motif (150, 151).

The *i*-motif has potential applications in nanotechnology as it can function as a one-dimensional scaffold (152). An

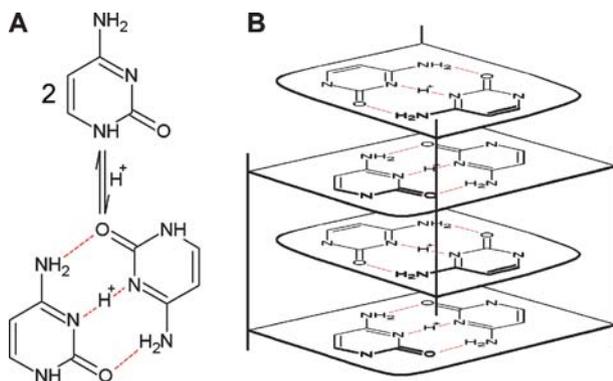


Figure 15 (A) Two C bases are capable of hydrogen bonding upon protonation. (B) Schematic structure of the *i*-motif.

The vertical lines in bold font represent the phosphosugar backbones.

ingenious nanodevice to map spatial and temporal pH changes inside living cells involving an *i*-motif was recently reported by Modi et al. (153). Their group synthesized a 'molecular hinge' of double-stranded DNA with a single-stranded C-rich overhang and fluorescent markers at both ends. At sufficiently low pH, the former rigid structure folds upon itself and the two extremities form an *i*-motif. This process was followed via FRET between the two fluorophores. This is an efficient reporter system with a relatively wide pH range 5.5 to 6.8 and could potentially be used to map endosome maturation spatially and temporally *in vivo*.

Outlook

The G-quadruplex field has blossomed in recent years and has provided a wide range of fascinating hypotheses and proofs of principle. It has led to some very valuable detailed work in structural biology and in medicinal chemistry, and some elegant discoveries in biophysics. The real challenge over the next 5–10 years is for this field's promise to be truly realized, with clearer evidence that these structures play an important role *in vivo*, that they can be effectively and selectively targeted by drugs, and/or that they can play an important role in bionanotechnology. We believe that the potential inherent in these structures will soon be actualized, and we encourage other researchers to address the seven questions we have posed in this article in order to further our understanding of G-quadruplexes.

Highlights

- DNA and RNA can form complex secondary structures, including guanine-rich G-quadruplexes.
- These structures are fascinating in their own right and could play important roles in biology and nanotechnology.
- We are beginning to be able to predict the structure and stability of G-quadruplexes.
- There is increasing evidence that G-quadruplexes can form *in vivo*, although further work is required to demonstrate their functions.
- G-quadruplexes appear to play a wide range of biological roles, at least *in vitro*, including in telomeres regulating transcription and regulating translation.
- Evolutionary pressures act on G-quadruplex-forming motifs, reducing their number in many genomic regions, but increasing it in regulatory domains.
- Many ligands have been developed that can bind G-quadruplexes, from small molecules to proteins. These show good selectivity for G-quadruplexes over duplex DNA, but as yet are not highly specific for unique G-quadruplex sequences.
- The G-quadruplex motif can be useful in nanotechnology as an aptamer, a structural switch or as a wire.

- C-rich complements can form their own four-stranded structure known as an '*i*-motif': this can form biologically and is also useful as a pH-sensitive element in nanotechnology.

Conflict of interest statement

None declared.

Acknowledgements

J.L.H. is a Research Councils UK Academic Fellow. This work was supported by the Isaac Newton Trust. A.C.E. is supported by the Isaac Newton Trust. S.L.B.K. is supported by the German Academic Exchange Service DAAD.

References

- Watson JD, Crick FHC. Molecular structure of nucleic acids: a structure for deoxyribose nucleic acid. *Nature* 1953; 171: 737–8.
- Gellert M, Lipsett MN, Davies DR. Helix formation by guanylic acid. *Proc Natl Acad Sci USA* 1962; 48: 2013–8.
- Mergny JL, De Cian A, Ghelab A, Sacca B, Lacroix L. Kinetics of tetramolecular quadruplexes. *Nucleic Acids Res* 2005; 33: 81–96.
- Hud NV, Smith FW, Anet FAL, Feigon J. The selectivity for K⁺ versus Na⁺ in DNA quadruplexes is dominated by relative free energies of hydration: a thermodynamic analysis by 1H NMR. *Biochemistry* 1996; 35: 15383–90.
- Gu J, Leczczyński J. Origin of Na⁺/K⁺ selectivity of the guanine tetraplexes in water: the theoretical rationale. *J Phys Chem A* 2002; 106: 529–32.
- Hardin CC, Watson T, Corregan M, Bailey C. Cation-dependent transition between the quadruplex and Watson-Crick hairpin forms of d(CGCG3GCG). *Biochemistry* 1992; 31: 833–41.
- Włodarczyk A, Grzybowski P, Patkowski A, Dobek A. Effect of ions on the polymorphism effective charge, and stability of human telomeric DNA. Photon correlation spectroscopy and circular dichroism studies. *J Phys Chem B* 2005; 109: 3594–605.
- Guschlbauer W, Chantot JF, Thiele D. Four-stranded nucleic acid structures 25 years later: from guanosine gels to telomeric DNA. *Biomol Struct Dyn* 1990; 8: 2013–8.
- Venczel EA, Sen D. Parallel and antiparallel G-DNA structures from a complex telomeric sequence. *Biochemistry* 1993; 32: 6220–8.
- Huppert JL. Four-stranded nucleic acids: structure, function and targeting of G-quadruplexes. *Chem Soc Rev* 2008; 37: 1375–84.
- Horváth MP, Schultz SC. DNA G-quartets in a 1.86 Å resolution structure of an *Oxytricha nova* telomeric protein-DNA complex. *J Mol Biol* 2001; 310: 367–77.
- Phillips K, Dauter Z, Murchie AI, Lilley DM, Luisi B. The crystal structure of a parallel-stranded guanine tetraplex at 0.95 Å resolution. *J Mol Biol* 1997; 273: 171–82.
- Haider S, Parkinson GN, Neidle S. Structure of a G-quadruplex-ligand complex. *J Mol Biol* 2003; 326: 117–25.
- Parkinson GN, Lee MP, Neidle S. Crystal structure of parallel quadruplexes from human telomeric DNA. *Nature* 2002; 417: 876–80.
- Padmanabhan K, Padmanabhan KP, Ferrara JD, Sadler JE, Tulinsky A. The structure of alpha-thrombin inhibited by a 15-mer single-stranded DNA aptamer. *J Biol Chem* 1993; 268: 17651–4.
- Krishnan-Ghosh Y, Liu D, Balasubramanian S. Formation of an interlocked quadruplex dimer by d(GGGT). *J Am Chem Soc* 2004; 126: 11009–16.
- Phan AT, Patel DJ. Two-repeat human telomeric d(TAGGGTTAGGGT) sequence forms interconverting parallel and antiparallel G-quadruplexes in solution: distinct topologies, thermodynamic properties and folding/unfolding kinetics. *J Am Chem Soc* 2003; 125: 15021–7.
- Phan AT, Modi YS, Patel DJ. Two-repeat tetrahymena telomeric d(TGGGGTTGGGGT) sequence interconverts between asymmetric dimeric G-quadruplexes in solution. *J Mol Biol* 2004; 338: 93–102.
- Crnugelj M, Hud NV, Plavec J. The solution structure of d(G4T4G3)2: a bimolecular G-quadruplex with a novel fold. *J Mol Biol* 2002; 320: 911–24.
- Smith FW, Lau FW, Feigon J. d(G3T4G3) forms an asymmetric diagonally looped dimeric quadruplex with guanosine 5'-syn-syn-anti and 5'-syn-anti-anti N-glycosidic conformations. *Proc Natl Acad Sci USA* 1994; 91: 10546–50.
- Crnugelj M, Sket P, Plavec J. Small change in a G-rich sequence, a dramatic change in topology: new dimeric G-quadruplex folding motif with unique loop orientations. *J Am Chem Soc* 2003; 125: 7866–71.
- Liu H, Kugimiya A, Matsugami A, Katahira M, Uesugi S. Quadruplex structures of RNA 14-mer, r(GGAGGUUU-UGGAGG) and DNA 14-mer, d(GGAGGTTTTGAGG). *Nucleic Acids Res* 2002; 2 (Suppl): 177–8.
- Wang Y, Patel DJ. Solution structure of the human telomeric repeat d[AG3(T2AG3)3] G-tetraplex. *Structure* 1993; 1: 263–82.
- Wang Y, Patel DJ. Solution structure of the tetrahymena telomeric repeat d(T2G4)4 G-tetraplex. *Structure* 1994; 2: 1141–56.
- Kuryavyi V, Majumdar A, Shallop A, Chernichecko N, Scripkin E, Jones R, Patel DJ. A double chain reversal loop and two diagonal loops define the architecture of a unimolecular DNA quadruplex containing a pair of stacked G(syn)-G(syn)-G(anti)-G(anti) tetrads flanked by a G-(T-T) Triad and a T-T-T triple. *J Mol Biol* 2001; 310: 181–94.
- Phan AT, Modi YS, Patel DJ. Propeller-type parallel-stranded G-quadruplexes in the human c-myc promoter. *J Am Chem Soc* 2004; 126: 8710–6.
- Ambrus A, Chen D, Dai J, Jones RA, Yang D. Solution structure of the biologically relevant G-quadruplex element in the human c-MYC promoter: implications for G-quadruplex stabilization. *Biochemistry* 2005; 44: 2046–58.
- Schultze P, Macaya RF, Feigon J. Three-dimensional solution structure of the thrombin binding DNA aptamer d(GGTTGGTGTGGTTGG). *J Mol Biol* 1994; 235: 1532–47.
- Burge S, Parkinson GN, Hazel P, Todd AK, Neidle S. Quadruplex DNA: sequence, topology and structure. *Nucleic Acids Res* 2006; 34: 5402–15.
- Keniry MA. Quadruplex structures in nucleic acids. *Biopolymers* 2000; 56: 123–46.
- Li J, Correia JJ, Wang L, Trent JO, Chaires JB. Not so crystal clear: the structure of the human telomere G-quadruplex in

- solution differs from that present in a crystal. *Nucleic Acids Res* 2005; 33: 4649–59.
32. Hazel P, Huppert JL, Balasubramanian S, Neidle S. Loop-length dependent folding of G-quadruplexes. *J Am Chem Soc* 2004; 126: 16405–15.
 33. Rachwal PA, Fox KR. Quadruplex melting. *Methods* 2007; 43: 291–301.
 34. Mergny JL, Phan A, Lacroix L. Following G-quartet formation by UV-spectroscopy. *FEBS Lett* 1998; 435: 74–8.
 35. Paramasivan S, Rujan I, Bolton PH. Circular dichroism of quadruplex DNAs: applications to structure, cation effects and ligand binding. *Methods* 2007; 43: 324–31.
 36. Gray DM, Wen JD, Gray CW, Repges R, Repges C, Raabe G, Fleischhauer J. Measured and calculated CD spectra of G-quartets stacked with the same or opposite polarities. *Chirality* 2008; 20: 431–40.
 37. Ou TM, Lu YJ, Tan JH, Huang TS, Wong KY, Gu LQ. G-quadruplexes: targets in anticancer drug design. *Chem Med Chem* 2008; 3: 690–713.
 38. Huppert JL, Balasubramanian S. Prevalence of quadruplexes in the human genome. *Nucleic Acids Res* 2005; 33: 2908–16.
 39. Wong HM, Stegle O, Rodgers SM, Huppert JL. A toolbox for predicting G-quadruplex formation and stability. *J Nucl Ac* 2010; in press.
 40. Todd AK. Bioinformatics approaches to quadruplex sequence location. *Methods* 2007; 43: 246–51.
 41. Ryvkin P, Hershman SG, Wang LS, Johnson FB. Computational approaches to the detection and analysis of sequences with intramolecular G-quadruplex forming potential. *Methods Mol Biol* 2010; 608: 39–50.
 42. Huppert JL. Hunting G-quadruplexes. *Biochimie* 2008; 90: 1140–8.
 43. Risitano A, Fox KR. Stability of intramolecular DNA quadruplexes: comparison with DNA duplexes. *Biochemistry* 2003; 42: 6507–13.
 44. Risitano A, Fox KR. The stability of intramolecular DNA quadruplexes with extended loops forming inter- and intra-loop duplexes. *Org Biomol Chem* 2003; 1: 1852–5.
 45. Risitano A, Fox KR. Influence of loop size on the stability of intramolecular DNA quadruplexes. *Nucleic Acids Res* 2004; 32: 2598–606.
 46. Wang Y, Patel DJ. Solution structure of a parallel-stranded G-quadruplex DNA. *J Mol Biol* 1993; 234: 1171–83.
 47. Aboul-ela F, Murchie AIH, Norman DG, Lilley DMJ. Solution structure of a parallel-stranded tetraplex formed by d(TG4T) in the presence of sodium ions by nuclear magnetic resonance spectroscopy. *J Mol Biol* 1994; 243: 458–71.
 48. Vorlickova M, Chladkova J, Kejnovska I, Fialova M, Kypr J. Guanine tetraplex topology of human telomere DNA is governed by the number of (TTAGGG) repeats. *Nucleic Acids Res* 2005; 33: 5851–60.
 49. Neidle S, Parkinson GN. The structure of telomeric DNA. *Curr Opin Struct Biol* 2003; 13: 275–83.
 50. Phan AT, Kuryavyi V, Burge S, Neidle S, Patel DJ. Structure of an unprecedented G-quadruplex scaffold in the human c-kit promoter. *J Am Chem Soc* 2007; 129: 4386–92.
 51. Todd AK, Haide SM, Parkinson GN, Neidle S. Sequence occurrence and structural uniqueness of a G-quadruplex in the human c-kit promoter. *Nucleic Acids Res* 2007; 35: 5799–808.
 52. Schaffitzel C, Berger I, Postberg J, Hanes J, Lipps HJ, Plückthun A. *In vitro* generated antibodies specific for telomeric guanine-quadruplex DNA react with *Stylonychia lemnae* macronuclei. *Proc Natl Acad Sci USA* 2001; 98: 8572–7.
 53. Simonsson T. G-quadruplex DNA structures – variations on a theme. *Biol Chem* 2001; 383: 621–8.
 54. Sen D, Gilbert W. Guanine quartet structures. *Methods Enzymol* 1992; 211: 191–9.
 55. Neidle S, Balasubramanian S, editors. *Quadruplex nucleic acids*, 1st ed., Cambridge, UK: Royal Society of Chemistry, 2006.
 56. Petraccone L, Erra E, Duro I, Esposito V, Randazzo A, Mayol L, Mattia CA, Barone G, Giancola C. Relative stability of quadruplexes containing different number of G-tetrads. *Nucleosides Nucleotides Nucleic Acids* 2005; 24: 757–60.
 57. Lu M, Guo Q, Kallenbach NR. Thermodynamics of G-tetraplex formation by telomeric DNAs. *Biochemistry* 1993; 32: 598–601.
 58. Li W, Wu P, Ohmichi T, Sugimoto N. Characterization and thermodynamic properties of quadruplex/duplex competition. *FEBS Lett* 2002; 526: 77–81.
 59. Miura T, Thomas GJ. Structure and dynamics of interstrand guanine association in quadruplex telomeric DNA. *Biochemistry* 1994; 33: 7848–56.
 60. Datta B, Armitage BA. Hybridization of PNA to structured DNA targets: quadruplex invasion and the overhang effect. *J Am Chem Soc* 2001; 123: 9612–9.
 61. Deng H, Braunlin WH. Duplex to quadruplex equilibrium of the self-complementary oligonucleotide d(GGGGCCCC). *Biopolymers* 1995; 35: 677–81.
 62. Salazar M, Thompson BD, Kerwin SM, Hurley LH. Thermally induced DNA-RNA hybrid to G-quadruplex transitions: possible implications for telomere synthesis by telomerase. *Biochemistry* 1996; 35: 16110–5.
 63. Strahan GD, Keniry MA, Shafer RH. NMR structure refinement and dynamics of the K+-[d(G3T4G3)]₂ quadruplex via particle mesh Ewald molecular dynamics simulations. *Biophys J* 1998; 76: 968–81.
 64. Gottschling DE, Zakian VA. Telomere proteins: specific recognition and protection of the natural termini of *Oxytricha* macronuclear DNA. *Cell* 1986; 47: 195–205.
 65. Gray JT, Celander DW, Price CM, Cech TR. Cloning and expression of genes for the *Oxytricha* telomere-binding protein: specific subunit interactions in the telomeric complex. *Cell* 1991; 67: 807–14.
 66. Fang G, Cech TR. The b subunit of *Oxytricha* telomere binding protein promotes G-quartet formation by telomeric DNA. *Cell* 1993; 74: 875–85.
 67. Sun D, Hurley LH. The importance of negative superhelicity in inducing the formation of G-quadruplex and i-Motif structures in the c-myc promoter: implications for drug targeting and control of gene expression. *J Med Chem* 2009; 52: 2863–74.
 68. Stegle O, Payet L, Mergny J-L, MacKay DJC, Huppert JL. Predicting and understanding the stability of G-quadruplexes. *Bioinformatics* 2009; 25: i374–82.
 69. Rawal P, Kummarasetti VB, Ravindran J, Kumar N, Halder K, Sharma R, Mukerji M, Das SK, Chowdhury S. Genome-wide prediction of G4 DNA as regulatory motifs: role in *Escherichia coli* global regulation. *Genome Res* 2006; 16: 644–55.
 70. Johnson JE, Smith JS, Kozak ML, Johnson FB. *In vivo* veritas: using yeast to probe the biological functions of G-quadruplexes. *Biochimie* 2008; 90: 1250–63.
 71. Lipps HJ, Rhodes D. G-quadruplex structures: *in vivo* evidence and function. *Trends Cell Biol* 2009; 19: 414–22.
 72. Maizels N. Dynamic roles for G4 DNA in the biology of eukaryotic cells. *Nat Struct Mol Biol* 2006; 13: 1055–9.

73. Huppert JL, Balasubramanian S. G-quadruplex in the human genome. *Nucleic Acids Res* 2007; 35: 406–13.
74. Neaves KL, Huppert JL, Henderson RM, Edwardson JM. Direct visualization of G-quadruplexes in DNA using atomic force microscopy. *Nucleic Acids Res* 2009; 37: 6269–75.
75. Paeschke K, Simonsson T, Postberg J, Rhodes D, Lipps HJ. Telomere end-binding proteins control the formation of G-quadruplex DNA structures *in vivo*. *Nat Struct Mol Biol* 2005; 12: 847–54.
76. Yang Q, Xiang J, Yang S, Zhou Q, Li Q, Tang Y, Xu G. Verification of specific G-quadruplex structure by using a novel cyanine dye supramolecular assembly: I. recognizing mixed G-quadruplex in human telomeres. *Chem Commun (Camb)* 2009; 1103–5.
77. Paeschke K, Juranek S, Simonsson T, Hempel A, Rhodes D, Lipps HJ. Telomerase recruitment by the telomere end binding protein-b facilitates G-quadruplex DNA unfolding in ciliates. *Nat Struct Mol Biol* 2008; 15: 598–604.
78. Azzalin CM, Reichenbach P, Khoriatou L, Giulotto E, Lingner J. Telomeric repeat containing RNA and RNA surveillance factors at mammalian chromosome ends. *Science* 2007; 318: 798–801.
79. Ding H, Schertzer M, Wu X, Gertsenstein M, Selig S, Kam-mori M, Pourvali R, Poon S, Vulto I, Chavez E, Tam PP, Nagy A, Lansdorp PM. Regulation of murine telomere length by Rtel: an essential gene encoding a helicase-like protein. *Cell* 2004; 117: 873–86.
80. Fry M, Loeb LA. Human Werner syndrome DNA helicase unwinds tetrahelical structures of the fragile X syndrome repeat sequence d(CGG)_n. *J Biol Chem* 1999; 274: 12797–802.
81. Sun H, Karow JK, Hickson ID, Maizels N. The Bloom's syndrome helicase unwinds G4 DNA. *J Biol Chem* 1998; 273: 27587–92.
82. Huber MD, Duquette ML, Shiels JC, Maizels N. A conserved G4 DNA binding domain in RecQ family helicases. *J Mol Biol* 2006; 358: 1071–80.
83. Sun H, Bennett RJ, Maizels N. The *Saccharomyces cerevisiae* Sgs1 helicase efficiently unwinds GG paired DNAs. *Nucleic Acids Res* 1999; 27: 1978–84.
84. Duquette ML, Handa P, Vincent JA, Taylor AF, Maizels N. Intracellular transcription of G-rich DNAs induces formation of G-loops, novel structures containing G4 DNA. *Genes Dev* 2004; 18: 1618–29.
85. Cheung I, Schertzer M, Rose AM, Lansdorp PM. Disruption of dog-1 in *Caenorhabditis elegans* triggers deletions upstream of guanine-rich DNA. *Nat Genet* 2002; 31: 405–9.
86. Zhao Y, Tarailo-Graouac M, O'Neil NJ, Rose AM. Spectrum of mutational events in the absence of DOG-1/FANCI in *Caenorhabditis elegans*. *DNA Repair (Amst)* 2008; 7: 1846–54.
87. London TB, Barber LJ, Mosedale G, Kelly GP, Balasubramanian S, Hickson ID, Boulton SJ, Hiom K. FANCI is a structure-specific DNA helicase associated with the maintenance of genomic G/C tracts. *J Biol Chem* 2008; 283: 36132–9.
88. Verma A, Yadav VK, Basundra R, Kumar A, Shantanu C. Evidence of genome-wide G4 DNA-mediated gene expression in human cancer cells. *Nucleic Acids Res* 2009; 37: 4194–204.
89. Siddiqui-Jain A, Grand CL, Bearss DJ, Hurley LH. Direct evidence for a G-quadruplex in a promoter region and its targeting with a small molecule to repress c-MYC transcription. *Proc Natl Acad Sci USA* 2002; 99: 11593–8.
90. Grand CL, Han H, Munoz RM, Weitman S, Von Hoff DD, Hurley LH, Bearss DJ. The cationic porphyrin TMPyP4 down-regulates c-MYC and human telomerase reverse transcriptase expression and inhibits tumor growth *in vivo*. *Mol Cancer Ther* 2002; 1: 565–73.
91. Fernando H, Sewitz S, Darot J, Tavaré S, Huppert JL, Balasubramanian S. Genome-wide analysis of a G-quadruplex-specific single-chain antibody that regulates gene expression. *Nucleic Acids Res* 2009; 37: 6716–22.
92. Kumari S, Bugaut A, Juppert JL, Balasubramanian S. An RNA G-quadruplex in the 5' UTR of the NRAS proto-oncogene modulates translation. *Nat Chem Biol* 2007; 3: 218–21.
93. Huppert JL, Bugaut A, Kumari S, Balasubramanian S. G-quadruplexes: the beginning and end of UTRs. *Nucleic Acids Res* 2008; 36: 6260–8.
94. Wieland M, Hartig JS. RNA quadruplex-based modulation of gene expression. *Chem Biol* 2007; 14: 757–63.
95. Arora A, Dutkiewicz M, Scaria V, Hariharan M, Maiti S, Kur-reck J. Inhibition of translation in living eukaryotic cells by an RNA G-quadruplex motif. *RNA* 2008; 14: 1290–6.
96. Makarov VL, Hirose Y, Langmore JP. Long G tails at both ends of human chromosomes suggest a C strand degradation mechanism for telomere shortening. *Cell* 1997; 88: 657–66.
97. Zahler AM, Williamson JR, Cech TR, Prescott DM. Inhibition of telomerase by G-quartet DNA structures. *Nature* 1991; 350: 718–20.
98. Zhang ML, Tong XJ, Fu XH, Zhou BO, Wang J, Liao XH, Li QJ, Shen N, Ding J, Zhou JQ. Yeast telomerase subunit Est1p has guanine quadruplex-promoting activity that is required for telomere elongation. *Nat Struct Mol Biol* 2010; 17: 202–10.
99. Kim NW, Piatyszek MA, Prowse KR, Harley CB, West MD, Ho PL, Coviello GM, Wright WE, Weinrich SL, Shay JW. Specific association of human telomerase activity with immortal cells and cancer. *Science* 1994; 266: 2011–5.
100. Todd AK, Neidle S. The relationship of potential G-quadruplex sequences in cis-upstream regions of the human genome to SP1-binding elements. *Nucleic Acids Res* 2008; 36: 2700–4.
101. Eddy J, Maizels N. Conserved elements with potential to form polymorphic G-quadruplex structures in the first intron of human genes. *Nucleic Acids Res* 2008; 36: 1321–33.
102. Howell RM, Woodfors KJ, Weitzmann MN, Usdin K. The chicken β -globulin gene promoter forms a novel "cinched" tetrahelical structure. *J Biol Chem* 1996; 271: 5208–14.
103. Simonsson T, Pecinka P, Kubista M. DNA tetraplex formation in the control region of c-myc. *Nucleic Acids Res* 1998; 26: 1167–72.
104. Rankin S, Reszka AP, Huppert JL, Zloh M, Parkinson GH, Todd AK, Ladame S, Balasubramanian S, Neidle S. Putative DNA quadruplex formation within the human c-kit oncogene. *J Am Chem Soc* 2005; 127: 10584–9.
105. Fernando H, Reszka AP, Huppert JL, Ladame S, Rankin S, Venkitaraman AR, Neidle S, Balasubramanian S. A conserved quadruplex motif located in a transcription activation site of the human c-kit oncogene. *Biochemistry* 2006; 45: 7854–60.
106. Dai J, Dexheimer TS, Chen D, Carver M, Ambrus A, Jones RA, Yang D. An intramolecular G-quadruplex structure with mixed parallel/antiparallel G-strands formed in the human BCL-2 promoter region in solution. *J Am Chem Soc* 2006; 128: 1096–8.
107. Sun D, Guo K, Rusche JJ, Hurley LH. Facilitation of a structural transition in the polypurine/polypyrimidine tract within the proximal promoter region of the human VEGF gene by the presence of potassium and G-quadruplex-interactive agents. *Nucleic Acids Res* 2005; 33: 6070–80.
108. de Armond R, Wood S, Sun D, Hurley LH, Ebbinghaus SW. Evidence for the presence of a guanine quadruplex forming

- region within a polypurine tract of the Hypoxia Inducible Factor 1R promoter. *Biochemistry* 2005; 44: 16341–50.
109. Smith TG, Robbins PA, Ratcliffe PJ. The human side of hypoxia-inducible factor. *Br J Haematol* 2008; 141: 325–34.
 110. Verma A, Halder K, Halder R, Yadav VK, Rawal P, Thakur RK, Mohd F, Sharma A, Chowdhury S. Genome-wide computational and expression analyses reveal G-quadruplex DNA motifs as conserved cis-regulatory elements in human and related species. *J Med Chem* 2008; 51: 5641–9.
 111. Nakken S, Rognes T, Hovig E. The disruptive positions in human G-quadruplex motifs are less polymorphic and more conserved than their neutral counterparts. *Nucleic Acids Res* 2009; 37: 5749–56.
 112. Todd AK, Johnston M, Neidle S. Highly prevalent putative quadruplex sequence motifs in human DNA. *Nucleic Acids Res* 2005; 33: 2901–7.
 113. Balasubramanian S, Neidle S. G-quadruplex nucleic acids as therapeutic targets. *Curr Opin Chem Biol* 2009; 13: 345–53.
 114. Marschke RF, Ricart AD, Von Hoff DD, Lim JK, Papadopoulos K. Phase I clinical trial of CX-3543, a pro-apoptotic anti-tumor agent. *J Clin Oncol* 2006; 24: 3082.
 115. Drygin D, Siddiqui-Jain A, O'Brien S, Schwaebe M, Lin A, Bliesath J, Ho CB, Proffitt C, Trent K, Whitten JP, Lim JKC, Von Hoff D, Anderes K, Rice WG. Anticancer activity of CX-3543: a direct inhibitor of rRNA biogenesis. *Cancer Res* 2009; 69: 7653–61.
 116. Monchaud D, Teulade-Fichou MP. A hitchhiker's guide to G-quadruplex ligands. *Org Biomol Chem* 2008; 6: 627–36.
 117. de Cian A, Lacroix L, Douarre C, Temime-Smaali N, Trenteaux C, Riou J-F, Mergny J-L. Targeting telomeres and telomerase. *Biochimie* 2008; 90: 131–55.
 118. Sun D, Thompson B, Cathers BE, Salazar M, Kerwin SM, Trent JO, Jenkins TC, Neidle S, Hurley LH. Inhibition of human telomerase by a G-quadruplex-interactive compound. *J Med Chem* 1997; 40: 2113–6.
 119. Gavathiotis E, Heald RA, Stevens MF, Searle MS. Drug recognition and stabilisation of the parallel stranded DNA quadruplex d(TTAGGGT)₄ containing the human telomeric repeat. *J Mol Biol* 2003; 334: 25–36.
 120. Parkinson GN, Ghosh R, Neidle S. Structural basis for binding of porphyrin to human telomeres. *Biochemistry* 2007; 46: 2390–7.
 121. Fedoroff OY, Salazar M, Han H, Chemeris VV, Kerwin SM, Hurley LH. NMR-based model of a telomerase-inhibiting compound bound to G-quadruplex DNA. *Biochemistry* 1998; 37: 12367–74.
 122. Campbell NH, Parkinson GN, Reszka AP, Neidle S. Structural basis of DNA quadruplex recognition by an acridine drug. *J Am Chem Soc* 2008; 130: 6722–4.
 123. Parkinson GN, Cuenca F, Neidle S. Topology conservation and loop flexibility in quadruplex-drug recognition: crystal structures of inter- and intramolecular telomeric DNA quadruplex-drug complexes. *J Mol Biol* 2008; 381: 1145–56.
 124. Takada T, Hayashi T, Arakawa M, Kominami R. Telomere elongation frequently observed during tumor metastasis. *Jpn J Cancer Res* 1992; 83: 1124–7.
 125. Kim MY, Vankayalapati H, Shin-ya K, Wierzbka K, Hurley LH. Telomestatin, a potent telomerase inhibitor that interacts quite specifically with the human telomeric intramolecular G-quadruplex. *J Am Chem Soc* 2002; 124: 2098–9.
 126. Tauchi T, Shin-ya K, Sashida G, Sumi M, Okabe S, Ohyashiki JH, Ohyashiki K. Telomerase inhibition with a novel G-quadruplex-interactive agent, telomestatin: in vitro and in vivo studies in acute leukemia. *Oncogene* 2006; 25: 5719–25.
 127. Burger AM, Dai F, Schultes CM, Reszka AP, Moore MJ, Double JA, Neidle S. The G-quadruplex interactive molecule BRACO-19 inhibits tumor growth, consistent with telomere targeting and interference with telomerase function. *Cancer Res* 2005; 65: 1489–96.
 128. Phatak P, Cookson JC, Dai F, Smith V, Gartenhaus RB, Stevens MF, Burger AM. Telomere uncapping by the G-quadruplex ligand RHPS4 inhibits clonogenic tumour cell growth in vitro and in vivo consistent with a cancer stem cell targeting mechanism. *Br J Cancer* 2007; 96: 1223–33.
 129. Tauchi T, Shin-ya K, Sashida G, Sumi M, Nakajima A, Shimamoto T, Ohyashiki JH, Ohyashiki K. Activity of a novel G-quadruplex-interactive telomerase inhibitor, telomestatin (SOT-095), against human leukemia cells: involvement of ATM-dependent DNA damage response pathways. *Oncogene* 2003; 22: 5338–47.
 130. Salvati E, Leonetti C, Rizzo A, Scarsella M, Mottolese M, Galati R, Sperduti I, Stevens MF, D'Incalci M, Blasco M, Chiorino G, Horard B, Gilson E, Stoppacciaro A, Zupi G, Biroccio A. Telomere damage induced by the G-quadruplex ligand RHPS4 has an antitumor effect. *J Clin Invest* 2007; 117: 3236–47.
 131. Cogo S, Xodo LE. G-quadruplex formation within the promoter of the KRAS proto-oncogene and its effect on transcription. *Nucleic Acids Res* 2006; 34: 2536–49.
 132. Bejugam M, Sewitz S, Shirude PS, Rodriguez R, Shahid R, Balasubramanian S. Trisubstituted isoalloxazines as a new class of G-quadruplex binding ligands: small molecule regulation of c-kit oncogene expression. *J Am Chem Soc* 2007; 129: 12926–7.
 133. Kong DM, Ma YE, Guo JH, Yang W, Shen HX. Fluorescent sensor for monitoring structural changes of G-quadruplexes and detection of potassium ion. *Anal Chem* 2009; 81: 2678–84.
 134. Kong DM, Guo JH, Yang W, Ma YE, Shen HX. Crystal violet-G-quadruplex complexes as fluorescent sensors for homogeneous detection of potassium ion. *Biosens Bioelectron* 2009; 25: 88–93.
 135. Wu ZS, Chen CR, Shen GL, Yo RQ. Reversible electronic nanoswitch based on DANN G-quadruplex conformation: a platform for single-step, reagentless potassium detection. *Biomaterials* 2008; 29: 2689–96.
 136. Alberti P, Mergny JL. DNA duplex quadruplex exchange as the basis for a nanomolecular machine. *Proc Natl Acad Sci USA* 2003; 100: 1569–73.
 137. Hou X, Guo W, Xia F, Nie FQ, Dong H, Tian Y, Wen L, Wang L, Cao L, Yang Y, Xue J, Song Y, Wang Y, Liu D, Jiang L. A biomimetic potassium responsive nanochannel: G-quadruplex DNA conformational switching in a synthetic nanopore. *J Am Chem Soc* 2009; 131: 7800–5.
 138. Lena S, Masiero S, Pieraccini S, Spada GP. Guanosine hydrogen-bonded scaffolds: a new way to control the bottom-up realization of well-defined nanoarchitectures. *Chem Eur J* 2009; 15: 7792–806.
 139. Wong A, Fettinger JC, Forman SL, Davis JT, Gang W. The sodium ions inside a lipophilic G-quadruplex channel as probed by solid-state ²³Na NMR. *J Am Chem Soc* 2002; 124: 742–3.
 140. Paborsky L, McCurdy S, Griffin L, Toole J, Leung LL. The single-stranded DNA aptamer-binding site of human thrombin. *J Biol Chem* 1993; 268: 20808–11.
 141. Tsiang M, Gibbs CS, Griffin L, Dunn K, Leung LL. Selection of a suppressor mutation that restores affinity of an oligonu-

- cleotide inhibitor for thrombin using in-vitro genetics. *J Biol Chem* 1995; 270: 19370–6.
142. Hall SW, Gibbs CS, Leung LL. Identification of critical residues on thrombin mediating its interaction with fibrin. *Thromb Haemost* 2001; 86: 1466–74.
 143. Yoshida W, Sode K, Ikebukuro K. Aptameric enzyme subunit for homogeneous protein sensing. *Nucleic Acids Symp Ser* 2007; 51: 99–100.
 144. Bates PJ, Laber DA, Miller DM, Thomas SD, Trent JO. Discovery and development of the G-rich oligonucleotide AS1411 as a novel treatment for cancer. *Exp Mol Pathol* 2009; 86: 151–64.
 145. Choi EW, Nayak LV, Bates PJ. Cancer-selective antiproliferative activity is a general property of some G-rich oligonucleotides. *Nucleic Acids Res* 2010; 38: 1623–35.
 146. Gehring K, Leroy JL, Guéron M. A tetrameric structure with protonated cytosine cytosine base pairs. *Nature* 1993; 363: 561–5.
 147. Leroy JL, Guéron M, Mergny JL, Hélène C. Intramolecular folding of a fragment of the cytosine-rich strand of telomeric DNA into an i-motif. *Nucleic Acids Res* 1994; 22: 1600–6.
 148. Mergny JL, Lacroix L, Han X, Leroy JL, Hélène C. Intramolecular folding of pyrimidine oligodeoxynucleotides into an i-DNA motif. *J Am Chem Soc* 1995; 117: 8887–98.
 149. Phan AT, Leroy JL. Intramolecular i-motif structures of telomeric DNA. *J Biomol Struct Dyn* 2000; Special Issue S2: 245–52.
 150. Gallego J, Chou SH, Reid BR. Centromeric pyrimidine strands fold into an intercalated motif by forming a double hairpin with a novel T.G.G.T tetrad: solution structure of the d(TCCCGTTTCCA) dimer. *J Mol Biol* 1997; 273: 840–56.
 151. Gilbert DE, Feigon J. Multistranded DNA structures. *Curr Opin Struct Biol* 1999; 9: 305–14.
 152. Ghodke HB, Krishnan R, Vignesh K, Kumar GVP, Narayana C, Krishnan Y. The I-tetraplex building block: rational design and controlled fabrication of robust 1D DNA Scaffolds via non-Watson Crick self assembly. *Angew Chem Int Ed* 2007; 46: 2646–9.
 153. Modi S, Goswami D, Gupta GD, Mayor S, Krishnan Y. A DNA nanomachine that maps spatial and temporal pH changes inside living cells. *Nat Nanotechnol* 2009; 4: 325–30.