NIH Public Access
Author Manuscript
Future Med Chem. Author manuscript; available in PMC 2011 February 1.

Published in final edited form as:
Future Med Chem. 2010 April ; 2(4): 619–646.

Structural insights into G-quadruplexes: towards new anticancer drugs

Danzhou Yang[1,2,3,4,†] and **Keika Okamoto**[1]

[1] College of Pharmacy, The University of Arizona, 1703 E. Mabel St., Tucson, AZ 85721, USA

[2] BIO5 Institute, The University of Arizona, 1657 E. Helen St., Tucson, AZ 85719, USA

[3] Arizona Cancer Center, 1515 N. Campbell Ave., Tucson, AZ 85724, USA

[4] Department of Chemistry, The University of Arizona, Tucson, AZ 85721, USA

Abstract

DNA G-quadruplexes are DNA secondary structures formed in specific G-rich sequences. DNA sequences that can form G-quadruplexes have been found in regions with biological significance, such as human telomeres and oncogene-promoter regions. DNA G-quadruplexes have recently emerged as a new class of novel molecular targets for anticancer drugs. Recent progress on structural studies of the biologically relevant G-quadruplexes formed in human telomeres and in the promoter regions of human oncogenes will be discussed, as well as recent advances in the design and development of G-quadruplex-interactive drugs. DNA G-quadruplexes can readily form in solution under physiological conditions and are globularly folded nucleic acid structures. The molecular structures of intramolecular G-quadruplexes appear to differ from one another and, therefore, in principle may be differentially regulated and targeted by different proteins and drugs.

DNA has now been recognized to not only play a passive role in genetic-information storage but also an active role in biological processes. In addition to the Watson–Crick duplex, DNA is able to transiently form alternative DNA secondary structures within specific sequences and as a consequence of dynamic molecular events. In particular, DNA G-quadruplexes, DNA secondary structures formed in specific G-rich sequences (Figure 1), have been shown to potentially form in regions of biological significance, such as human telomeres and oncogene promoter regions. Consequently, DNA G-quadruplexes have recently emerged as a new class of novel molecular targets for anticancer drugs. In this review, we will discuss the recent progress on structural studies into the biologically relevant G-quadruplexes formed in human telomeres and in the promoter regions of human oncogenes. The well-defined targets and structural information of DNA G-quadruplexes are essential for understanding their biological functions, as well as for the rational design of small molecules targeting these structures. We will also discuss the recent progress in the design and development of G-quadruplex-interactive drugs.

[†]Author for correspondence, Tel.: +1 520 626 5969, Fax: +1 520 626 6988, yang@pharmacy.arizona.edu.

For reprint orders, please contact reprints@future-science.com

Financial & competing interests disclosure

This research was supported by the National Institutes of Health (1S10RR16659 and CA122952). The authors have no other relevant affiliations or financial involvement with any organization or entity with a financial interest in or financial conflict with the subject matter or materials discussed in the manuscript. This includes employment, consultancies, honoraria, stock ownership or options, expert testimony, grants or patents received or pending, or royalties.

No writing assistance was utilized in the production of this manuscript.

NIH-PA Author Manuscript
NIH-PA Author Manuscript
NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

G-quadruplex structure

G-quadruplexes are four-stranded DNA secondary structures that deviate from the normal duplex form of DNA. They consist of stacked G-tetrads, square-planar platforms of four guanines connected by cyclic Hoogsteen hydrogen bonding (Figure 1A), instead of the Watson–Crick hydrogen bonding in a B-DNA duplex. The guanine residues in a G-tetrad can adopt either *syn* or *anti* glycosidic conformation, in a G-strand-directional manner, with the tetrad guanines from parallel G-strands adopting the same glycosidic conformation and those from antiparallel G-strands adopting the opposite (Figure 1B). G-quadruplexes can be formed with one, two, or four-G-rich strands (Figure 1). Tetramolecular G-quadruplexes are, in general, parallel-stranded with tetrad guanines adopting *anti* glycosidic conformation.

Although the unusual ability of guanine-rich DNA solutions to form gelatinous aggregates was first noted in 1910, their exact nature was not discovered until 1962, when Gellert and co-workers proposed that these gels form planar guanine tetramers that stack into cyclic arrangements [1]. G-quadruplex formation and stabilization requires monovalent cations, in particular K^+ and Na^+, which coordinate with the eight electronegative O6 atoms of the adjacent stacked G-tetrads (Figure 1) [2,3]. Since K^+ and Na^+ are the main cations *in vivo*, G-quadruplex formation is favored under physiological conditions. A G-rich sequence may adopt different structures in the presence of different cations. The K^+ form is considered to be more biologically relevant due to its higher intracellular concentration (~140 mM) than that of Na^+ (5–15 mM). The precise location of cations between tetrads depends on the nature of the ions: Na^+ ions are observed in a range of geometries, whereas K^+ ions are always equidistant between each tetrad plane. K^+ in general is preferred over Na^+ by G-quadruplexes as K^+ has a better coordination with eight guanine O6s and a lower dehydration energy [3].

Intramolecular G-quadruplexes formed by single-stranded DNA are of intensive current research interest due to their potential formation in telomeres [4–6] and oncogene promoter sequences [7]. In contrast to tetramolecular quadruplexes, intramolecular structures form quickly and are more complex, exhibiting great conformational diversity, such as in folding topologies, loop conformations and capping structures (Figure 1C). Not only do different sequences adopt distinct topologies, but, in addition, a given sequence can fold into a variety of different conformations, as in the case of the human telomeric DNA sequence. With the recent progress in the structural studies of G-quadruplexes, numerous rules governing G-quadruplex folding have been recognized (see later discussion). However, complete prediction of a G-quadruplex conformation appears to be difficult, thus structural characterization of G-quadruplexes seems be necessary to understand each G-quadruplex structure and its ligand interactions.

Formation of DNA G-quadruplexes in biological systems

The first biologically relevant G-quadruplex formation was observed in telomeric DNA. This was based on the observed association of guanine–guanine base-paired hairpin structures in single-stranded telomeric terminal sequences of several organisms [8], which was further characterized in the telomeric DNA of *Tetrahymena* [9]. The importance of monovalent cations in the induction and stabilization of these structures was revealed by Williamson, who proposed that cations bind in the cavity formed by the guanines in each tetrad [10]. The most direct evidence of the *in vivo* existence of G-quadruplexes was established using specific antibodies against parallel and antiparallel G-quadruplexes formed in telomeric DNA of the ciliate *Stylonychia lemnae* [11]. The antibody specific for the antiparallel G-quadruplex was detected in macronuclei but not in the corresponding micronuclei of ciliates, and the signal was cell cycle-dependent and absent during replication [11]. More recently, using the same antibody in ciliates, it was shown *in vivo* that the telomere end-binding proteins TEBPα and TEBPβ are required to control G-quadruplex formation and that TEBPβ phosphorylation is needed to

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

resolve G-quadruplex structures during replication [12]. Additional evidence for G-quadruplex formation *in vivo* was provided by detection of G-quadruplex formation at human chromosomal ends by using the radiolabeled G-quadruplex ligand 360A (2,6-*N*,*N*′-methyl-quinolinio-3-yl-pyridine dicarboxamide) [13] or the fluorescent G-quadruplex ligand BMVC (3,6-*bis*(1-methyl-4-vinylpyridinium) carbazole diiodide [14].

While most of the evidence for the *in vivo* existence of G-quadruplexes is in telomeric DNA, there is a growing body of compelling evidence for the biological relevance of G-quadruplexes in noncoding regions of the genome, especially gene promoters. A significant enrichment of quadruplex-forming motifs has recently been shown in the promoter regions of human genes near transcription start sites (TSS), with a strong correlation between nuclease-hypersensitive elements (indicative of unusual DNA structures) and the G-quadruplex-forming motifs [15, 16]. It was suggested that quadruplex formation most likely occurs during processes such as transcription, replication or recombination, when the DNA duplex is actively denatured [17]. The biological relevance for the G-quadruplexes formed in gene promoters will be discussed in more details later.

In addition, G-quadruplexes have been found in other biologically relevant regions of the genome, such as immunoglobulin switch regions [17], ribosomal DNA [18] and certain regions of RNA [19–21]. For example, the quadruplex formed in the extended $(CGG)_n$ repeat of the Fragile X syndrome was found to inhibit the translation of FMR1 protein. Expression of proteins such as CBF-A or hnRNPA2, which destabilize the G-quadruplex *in vitro*, was found to increase translational efficiency [22,23].

G-quadruplex-interactive proteins

A growing list of proteins has been identified as interacting with DNA G-quadruplex structures, strongly pointing to the existence of qua-druplex DNA *in vivo* [17,24,25]. The formation of G-quadruplex structures *in vivo* within plasmid genomes transcribed in *Escherichia coli* cells was detected by electron microscopy and verified by interactions with G-quadruplex-specific nuclease GQN1 and the high-affinity G-quadruplex-binding protein nucleolin [26]. Both the β subunit of the *Oxytricha nova* telomere-binding protein TBP [27] and the *Saccharomyces cerevisiae* repressor-activator protein 1 (RAP1) [28], which are essential for telomere maintenance, promote G-quadruplex formation. Human topoisomerase I (Topo I) also promotes the formation of G-quadruplexes and binds to preformed G-quadruplexes [29,30]. The mammalian nucleolar protein nucleolin binds with high affinity to G-quadruplex structures [31]. MutSα protein, which is involved in mismatch repair, targets G-quadruplex in G loop segments and promotes synapsis of transcriptionally activated immunoglobulin switch regions [32]. Activation-induced cytosine deaminase (AID) also binds to G-quadruplex and plays a role in immunoglobulin class switch recombination [33]. On the other hand, a group of G-quadruplex destabilizing proteins has been identified. Human Pot1 is a highly conserved telomeric protein that binds to the 3′ end of single-stranded telomeric DNA and plays an important role in telomere end-capping and protection [34,35]. hPot1 has been shown to disrupt telomeric G-quadruplex structures to restore the activity of telomerase [36]. Pot1 also stimulates the activity of WRN/BLM helicases to unwind telomeric DNA [37]. Replication protein A (RPA), another highly conserved eukaryotic protein, is essential for telomere maintenance [38,39], and also unfolds telomeric G-quadruplexes [40]. Also, several single-strand binding proteins of the heterogeneous nuclear ribonucleoprotein (hnRNP) family, including hnRNP A1 [41], its derivative UP1 [42] and hnRNP D [43], resolve telomeric G-quadruplexes, presumably by trapping the single-stranded form of telomeric DNA. In addition to noncatalytic G-quadruplex-resolving proteins, the RecQ family of helicases, such as Sgs1 and Cdc13p in *S. cerevisiae*, Werner's syndrome protein (WRN) and Bloom's syndrome protein (BLM) helicases in human cells, preferentially bind and disrupt DNA G-quadruplex

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

structures and are suggested to participate in telomere maintenance by facilitating replication or recombination at telomeric ends [44–49]. The lack of RecQ helicases is associated with genomic instability, premature senescence, and accelerated telomere erosion. RecQ helicases catalytically unwind DNA G-quadruplex structures with a single-stranded 3′ tail, which requires energy from ATP hydrolysis and the presence of Mg^{2+}. It is important to note that G-quadruplex-interactive agents inhibit telomerase activity and the unwinding of G-quadruplex structures by Sgs1 and WRN/BLM helicases [44,50]. Moreover, G-quadruplex processing is most likely facilitated by G-quadruplex-specific nucleases such as yeast KEM1/SEP1 [51], human GQN1 [52] and yeast Mre11p [53], which can all cleave single-stranded DNA that neighbors G quartets. Interestingly, neither KEM1/SEP1 nor GQN1 can cleave duplex DNA, and Mre11p has much higher affinity for G-quadruplex DNA than for single- or double-stranded DNA. It is noteworthy that most proteins appear to bind to the loop regions of G-quadruplexes.

G-quadruplexes formed in human telomeres

Introduction

Telomeres are specialized DNA–protein complexes that cap the ends of linear chromosomes and provide protection against gene erosion at cell divisions, chromosomal nonhomologous end-joinings and nuclease attacks [54–56]. Human telomeres consist of tandem repeats of the hexanucleotide d $(TTAGGG)_n$ 5–10 kb in length, which terminate in a single-stranded 3′-overhang of 35–600 bases [57–61]. Telomeric DNA is extensively associated with various proteins, such as Pot1, TRF1 and TRF2, as well as TIN2, Rap1 and TPP1 [57,62–64]. The structure and stability of telomeres are closely related to cancer [4,65], aging [66], and genetic stability [56,67]. In normal cells, each cell replication results in a 50- to 200-base loss of the telomere [68–70]. After a critical shortening of the telomeric DNA is reached, the cell undergoes apoptosis [68]. However, telomeres of cancer cells do not shorten on replication, due to the activation of a reverse transcriptase, telomerase, that extends the telomeric sequence at the chromosome ends [71]. Telomerase is activated in 80–85% of human cancer cells [72] and has been suggested to play a key role in maintaining the malignant phenotype by stabilizing telomere length and integrity [73]. The G-rich sequence of human telomeric DNA has a strong propensity to form the DNA G-quadruplex secondary structure, which can inhibit the activity of telomerase [74]. In addition to the formation at the telomere end, which most likely involves intramolecular G-quadruplex structures, intermolecular G-quadruplex formation may also be involved in the T-loop invasion complex [75,76] and in meiotic chromosome synapsis and homologous recombination [9,77–79].

G-quadruplex-targeting compounds have been shown to inhibit telomerase activity, thereby making the intramolecular human telomeric DNA G-quadruplex an attractive target for cancer therapeutic intervention (Figure 2A) [4,6,65,80,81]. More recently, G-quadruplex-targeting compounds have been shown to disrupt telomere capping and induce rapid apoptosis (Figure 2B) [4,82]. As telomeric DNA is extensively associated with various proteins, disruption of telomere capping and maintenance, such as disruption of TRF2 and Pot1, can be sensed as DNA damage, which in turn rapidly activates the apoptotic pathways in the absence of telomere shortening [34,83,84]. Furthermore, G-quadruplex-targeting compounds have also been shown to inhibit the alternative lengthening of telomeres (ALT) pathway [85–87], which maintains telomere stability in a telomerase-independent manner in cancer cells (~15%) wherein telomerase is not activated.

Human telomeric G-quadruplex structures & polymorphism

Understanding of the human telomeric G-quadruplex structure under physiologically relevant conditions has been the subject of intense investigation. A 22-nt human telomeric sequence 5′-

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

AGGG(TTAGGG)$_3$-3′ (wtTel22) (Figure 3A) has been shown by NMR to form a basket-type G-quadruplex in Na^+ solution [88], which is a mixed antiparallel–parallel stranded intramolecular structure with three G-tetrads connected by one diagonal and two lateral (edgewise) TTA loops (Figure 4Ai). The same 22-nt human telomeric sequence has more recently been shown in the crystalline state in the presence of K^+ to form a parallel-stranded intramolecular G-quadruplex consisting of three G-tetrads connected with three symmetrical propeller (double-chain-reversal) TTA loops (Figure 4Aii) [89]. However, when examined in K^+ solution by NMR, this wtTel22 sequence does not form a single G-quadruplex structure (Figure 3Bi) [90]. Recent structural studies have shown that the hybrid-type intramolecular G-quadruplex structures (Figure 4Aiii) appear to be the major conformations formed in the human telomeric sequences in K^+ solution (Figure 3A) [90–95], even in the co-presence of high concentrations of Na^+ [90]. The telomeric G-quadruplexes are always in dynamic equilibrium between two conformations, the hybrid-1 and hybrid-2 structures [91]. Each hybrid-type G-quadruplex structure has distinct capping structures that may provide specific drug-binding sites. The hybrid-type conformations suggest a straightforward means for multimer formation with effective packing in the human telomeric sequence. It is important to note that the structure polymorphism appears to be intrinsic to the highly conserved human telomeric sequence and may be exploited in biological systems [5]. In this review we will focus on the structure studies from our group.

Hybrid-1 & -2 human telomeric

G-quadruplexes are closely related yet distinct in their folding structures: Large numbers of variant four-G-tract sequences containing the core wtTel22 with different flanking segments have been screened by ^{1}H NMR [90]. A few sequences, including the 26-nt sequence Tel26 (Figure 3A), form one stable major G-quadruplex in K^+ solution (Figure 3Bii). The 3′-flanking segment has been found to be critical for the formation of a stable G-quadruplex. The folding topology and the molecular structure of the G-quadruplex formed by Tel26 in K^+ solution, which is a hybrid-1 type intramolecular G-quadruplex with three G-tetrads, have been determined by NMR (PDB ID 2HY9) (Figure 4Aiii left & Figure 5B) [90,92].

Tel26 was found to quickly adopt two conformations when freshly dissolved in K^+ solution and to become stabilized in the hybrid-1 conformation [92]. Careful examination of the wild-type 26-nt human telomeric sequence wtTel26 (TTAGGG)$_4$TT (Figure 3A) showed a major intramolecular G-quadruplex, which accounts for approximately 70% of the total population in K^+ solution (Figure 3B) [91]. Albeit quite challenging, the folding topology and molecular structure of the major G-quadruplex formed by wtTel26 in K^+ solution, which is a hybrid-2 type intramolecular G-quadruplex with three G-tetrads, have also been determined by NMR (PDB ID 2JPZ) (Figure 4Aiii right & Figure 5A) [91].

While both hybrid-type telomeric structures contain three G-tetrads linked with mixed parallel/antiparallel G-strands, they differ in loop arrangements, strand orientations, G-tetrad arrangements, and capping structures (Figure 4Aiii & Figure 5A–D). The hybrid-1 structure has sequential side-lateral-lateral loops with the first TTA loop adopting the double-chain-reversal conformation, whereas the hybrid-2 structure has lateral-lateral-side loops with the last TTA loop adopting the double-chain-reversal conformation. Both hybrid-type structures contain three parallel G-strands and one antiparallel G-strand, five *syn* guanines and asymmetrical G-arrangements. The first G-tetrad (from the 5′ end) has a reversed G-arrangement from the other two G-tetrads; for the hybrid-1 structure the first G-tetrad is (*syn:syn:anti:syn*) and the bottom two are (*anti:anti:syn:anti*), whereas for the hybrid-2 structure the first G-tetrad is (*syn:anti:syn:syn*) and the bottom two are (*anti:syn:anti:anti*).

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

Loop conformations & capping structures determine the selective formation of hybrid-1 or hybrid-2 telomeric

G-quadruplex structures: The wild-type telomeric sequence wtTel26 and the modified Tel26 sequence contain the same four-G-tract core segment wtTel22, with Tel26 containing the modified 5′- and 3′-flanking AA instead of the native TT (Figure 3A). However, in K^+ solution, wtTel26 forms a major hybrid-2 structure, whereas Tel26 predominantly forms a hybrid-1 structure (Figure 4Aiii) [91,92]. The NMR structures of the two hybrid-type human telomeric G-quadruplexes are both well defined, including the two flanking sequences and the three TTA loops (Figure 5A–D). The molecular structures indicate that specific capping structures selectively stabilize the specific hybrid-type telomeric G-quadruplex. The molecular structure of the hybrid-2 G-quadruplex formed by wtTel26 in K^+ shows a T:A:T triple capping structure specific to the hybrid-2-type human telomeric G-quadruplex (Figure 5A). The well-defined T:A:T triple platform forms with T8 and A9 of the first TTA lateral loop and T25 of the 3′-flanking segment, capping the bottom end of the hybrid-2 human telomeric G-quadruplex (Figure 5C) [91]. The importance of the T:A:T capping structure in stabilizing the hybrid-2 human telomeric G-quadruplex structure has been demonstrated by mutational analysis of wtTel26 with single A-to-T and T-to-U substitutions [91]. This T:A:T triple capping structure is specific to the hybrid-2 folding and is not possible in the hybrid-1 folding (Figure 4Aiii). In the Tel26 sequence (Figure 3A), T25 is mutated to A25 and the T8:A9:T25 triple can no longer form (Figure 4Aiii); therefore, the hybrid-2 structure is no longer favored. Rather, Tel26 forms a hybrid-1 structure in K^+. Two capping platforms, specific to the hybrid-1 structure, are observed in the Tel26 structure (Figure 5B): an adenine triple structure of three naturally occurring adenines, A21, A3 and A9, at the top end (Figure 5D) and an A:T base pair structure at the bottom end of the G-quadruplex [92]. The A:T base pair-capping structure involves the mutated A25 and explains the specific stabilization effect of the hybrid-1 type structure by T25A mutation [90].

The two hybrid-type telomeric G-quadruplexes both contain a 3-nt TTA double-chain-reversal loop, which are the first and third loops in the hybrid-1 and hybrid-2 structures, respectively (Figure 4Aiii) [91,92]. As will be discussed in the promoter G-quadruplex structures, a double-chain-reversal loop conformation usually favors short loop sizes of 1 and 2 nt, likely due to the lack of stacking interactions of the groove-positioned loop residues. This may explain why the parallel-stranded telomeric G-quadruplex with three double-chain-reversal 3-nt TTA loops is not favored in K^+ solution. Actually, in the molecular structures of the two hybrid-type telomeric G-quadruplexes, the adenines of the 3-nt TTA double-chain-reversal loops (i.e., A9 in hybrid-1 and A21 in hybrid-2) both adopt a conformation which is partially stacked with the top G-tetrad (Figure 5A & B). Interestingly, the presence of a TTA loop in double-chain-reversal conformation in the hybrid-type structures allows the 5′ and 3′ ends to point in opposite directions. Thus the 3-nt double-chain-reversal loop in the hybrid-type telomeric structures may be selected to avoid the steric interference of the two ends. It is noteworthy that the presence of a 3-nt double-chain-reversal loop in the hybrid-type structures may contribute to the structural flexibility and polymorphism of human telomeric G-quadruplexes (see following discussion).

Hybrid-type structures are the predominant conformations for human telomeric sequences in K^+ solution, even in the copresence of a high concentration of Na^+

—Hybrid-type structures (Figure 4Aiii) appear to be the predominant conformation for wtTel26 and Tel26 in K^+ solution, even in the co-presence of a high concentration of Na^+. Addition of K^+ to the Na^+ solution readily converts the preformed basket-type Na^+ structure to the hybrid-type K^+ structure (Figure 3C) [90]. In contrast, addition of Na^+ to a 50-mM K^+ solution, even to a concentration of 200 mM, does not change the hybrid-type K^+ structure to the basket-type Na^+ structure (Figure 3C). The same results are also observed for the truncated

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

22-nt wtTel22 sequence. Full conversion of the Na^+ structure to its K^+ structure of Tel26 takes several hours to an overnight incubation, depending on the K^+ concentration, whereas conversion of the Na^+ structure of wtTel22 to its K^+ structure is faster than can be detected by either CD or NMR [90]. In contrast to the wtTel22 sequence that forms a single well-defined basket-type G-quadruplex in Na^+ solution [88], the extended sequences wtTel26 and Tel26 do not form a single G-quadruplex conformation in Na^+ solution [90,91].

Potential of the hybrid-type human telomeric G-quadruplexes to form higher-order multimers—Significantly, both the hybrid-type human telomeric G-quadruplexes structures provide an efficient scaffold for a compact-stacking structure of multimers in human telomeric DNA (Figure 5e). Human telomeric DNA terminates with a 3′ single-stranded overhang of 35–600 nt. The 5′ and 3′ ends of the hybrid-type structures point in opposite directions, allowing the hybrid-type G-quadruplex to be readily folded and stacked end-to-end in the elongated linear telomeric DNA strand. Intriguingly, the capping structures (the adenine triple in the hybrid-1 structure and the T:A:T triple in hybrid-2 structure) can provide not only stacking interactions between the two adjacent telomeric G-quadruplexes but also specific binding sites for small-molecule ligands to target G-quadruplexes in human telomeres (Figure 5e). Indeed, a recent study indicated that the most plausible structure of a 44-nt single-stranded human telomeric DNA is a dimer structure with each quadruplex in a different hybrid form [96]. However, different results have been shown by some other studies [97–99].

Structure polymorphism of human telomeric G-quadruplexes: is it related to the highly conserved human telomeric sequence?

For extended four-G-tract human telomeric sequences, both hybrid-1 and hybrid-2 forms appear to coexist in K^+ solution; the hybrid-2 form appears to be the major conformation (Figure 3A) [91]. For example, the hybrid-2 form accounts for approximately 70% in wtTel26 and less than 65% in wtTel27, but the hybrid-1 form can also be detected in both sequences. By contrast, for human telomeric sequences with no 3′-flanking segment (wtTel23 and wtTel24), the hybrid-1 structure appears to be the major conformation. This is consistent with the T:A:T capping structure in the hybrid-2 structure, which requires the 3′-flanking T (Figure 5C). One specific four-G-tract telomeric sequence, wtTel21-T, has very recently been found to form an interesting basket-type G-quadruplex with only two G-tetrads [100,101]. However, the two-tetrad conformation is unlikely to be a stable form of the extended human telomeric sequence, as it can hardly be detected in any telomeric sequence with a 5′-flanking segment.

While the energy difference of the two hybrid forms is quite small, so that the two forms can coexist in the extended telomeric sequences, the kinetics of the interconversion between the two forms appears to be rather slow, as very few exchange peaks were observed in Nuclear Overhauser Effect correlation spectroscopy experiments [91]. The slow kinetics of the interconversion indicates a high-energy barrier intermediate(s) for transition. We have proposed a model for the interconversion between the hybrid-type K^+ form and the basket-type Na^+ form telomeric G-quadruplex structures through a strand-reorientation mechanism (Figure 4B) [90,101]. The 5′ G-strand of the basket-1 type G-quadruplex may dissociate from the structure and swing back to the other side of the second G-stand to form a parallel-stranded structural motif with a double-chain-reversal loop. A similar strand-reorientation could occur at the 3′-end for the interconversion between the hybrid-2 and basket forms. Interestingly, the two-G-tetrad basket-type structure appears to fit the intermediate state in the proposed transition model and is likely to be a transition intermediate of the interconversion between different telomeric G-quadruplex forms [101]. The presence of a metastable intermediate state is consistent with recent reports in which the unfolding processes of the telomeric G-quadruplexes have been shown to have multiple intermediate states [102,103].

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

Unlike the G-rich sequences in gene-promoter regions that will be discussed later, the telomeric DNA sequence contains the same tandem repeats. Both hybrid-type G-quadruplexes are asymmetric structures (Figure 4Aiii), which are related to the intrinsic asymmetric residue distribution of the tandem TTA loops in the human telomeric DNA sequence. It is this intrinsic asymmetry that determines the possibility of the formation of the two very closely related but distinct hybrid-type structures. The linker segment in the telomeric sequences of human (and vertebrate) is TTA, whereas those of lower organisms contain only thymines. The presence of adenine in the TTA loop adds an asymmetry in human telomeric sequences, thereby providing a more selective basis for different capping structures, such as the T:A:T triple in the hybrid-2 structure and the adenine triple in the hybrid-1 structure, which appear to play an important role in stabilizing different forms (Figure 5A–C). The human telomeric sequence is highly conserved. Nature may have chosen this specific sequence with its asymmetry and the low energy difference between various forms. The structural polymorphism of the human telomeric sequence may provide a means of controlling the biology of human telomeres, such as imparting different protein recognition. The small energy difference between various forms may present an attractive means for drug targeting, which may readily surpass the small energy barrier and, subsequently, change the dynamics and protein recognition of human telomeres. On the other hand, this structural polymorphism may also present a challenge for specific drug targeting.

G-quadruplexes formed in human oncogene promoters

Introduction

The potential occurrence of DNA G-quadruplex has been discovered in the promoter regions of genes involved in growth and proliferation [104,105]. These genes are mostly TATA-less and all contain G/C-rich tracts in the proximal regions of promoters. It has been shown that the potential for quadruplex formation is higher within oncogenes compared with tumor suppressor genes [106]. The first experiments that suggested the existence of unusual forms of DNA in gene promoters were reported in 1982, based upon the nuclease hypersensitivity of promoter elements in the chicken β-globulin gene [107]. This was also independently demonstrated by another group [108]. These unusual DNA structures were subsequently shown to be associated with runs of guanines. Since then, the occurrence of these elements in the promoters of other genes have been reported, such as those of insulin [109], human *c-MYC* [110,111], human *VEGF* [112], human *HIF-1*α [113], human *BCL-2* [114–116], human *sMtCK* [117], human and mouse *KRAS* [118], human *c-KIT* [119,120], human *RET* [121] and human *PDGF-A* [122], in addition to chicken β-globulin [123] and mouse α7 integrin [124].

In contrast to the repeating tandems in the human telomeric sequence, the G-rich sequences of gene promoters are often composed of G-tracts with unequal numbers of guanines and various numbers of intervening bases. The promoter sequences may contain more than four-G-tracts and each of these sequences appears to be unique in its number and length of G-tracts and intervening bases [7]. Thus the G-rich promoter sequences in general are capable of forming multiple possible quadruplex structures using different combinations of G-tracts or Gs within a tract and are often a mixture of multiple conformations and loop isomers. A notable feature of the promoter quadruplexes is the prevalent occurrence of the G_3NG_3 sequence motif, which forms a robust parallel-stranded structure motif with a 1-nt double-chain-reversal loop (Figure 6). Indeed, the G_3NG_3 motif is so widespread that it has been proposed to be evolutionarily selected to serve as a stable core in the promoter intramolecular G-quadruplex structures.

The *c-MYC* gene promoter is the most extensively studied system for the G-quadruplex formation. c-MYC is a potent oncoprotein and transcription factor that plays an essential role in the control of cell growth as well as in cell fate determinations related to the induction of apoptosis [125,126]. *c-MYC* overexpression is associated with a large number of human

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

malignancies, including colon, breast, prostate, cervical, and small-cell lung carcinomas, osteosarcomas, glioblastomas, B-cell and T-cell lymphomas and myeloid leukemias [127–131]. The transcriptional regulation of *c-MYC* expression is complex, with P1 and P2 being the predominant promoters [132]. A highly conserved 27-base-pair nuclease hypersensitivity element III_1 (NHE III_1) in the proximal region of the *c-MYC* promoter controls 80–90% of the transcriptional activity, regardless of whether the P1 or P2 promoter is used (Figure 7A) [133–135]. *In vivo* footprinting using $KMnO_4$ showed that a silenced form of the *c-MYC* promoter NHE III_1 is different from the transcriptionally active state and is consistent with G-quadruplex formation [136]. More recently, the NHE III_1 has been shown to be a silencer element and the formation of a DNA G-quadruplex structure is critical for c-MYC transcriptional silencing (Figure 7B) [111]. Compounds that bind to and stabilize the G-quadruplex conformation have been shown to reduce c-MYC expression and are anti-tumorigenic [137–139].

A special requirement for promoters is that the G-quadruplex element must be generated in a region of duplex rather than single-stranded DNA, as is the case for telomeres. A report from the Levens laboratory at the National Cancer Institute demonstrated *in vivo* the effect of transcriptionally induced supercoiling in the *c-MYC* promoter that controls the rate of promoter firing [140,141]. The transcriptionally induced supercoiling can melt and convert DNA to a non-B-DNA form in the far upstream element (FUSE) in the *c-MYC* promoter, which is 1.7 Kb upstream of the P1 promoter, even in the presence of topoisomerases. The NHE III_1element, which is the G-quadruplex forming region in the *c-MYC* promoter, is much closer to the source of induced negative superhelicity, the replication machinery, and is thus likely to be subjected to greater torsional stress than the FUSE. Indeed, a report from the Hurley laboratory showed that the negative superhelicity facilitates secondary DNA structure formation under physiological conditions in a supercoiled plasmid system [142]. These results provide compelling evidence for a dynamic equilibrium between duplex, single-stranded and secondary DNA structures that determine the transcriptional state of the *c-MYC* promoter.

Importantly, two protein players involved in gene transcription have been shown to interact with the *c-MYC* promoter G-quadruplex. Nucleolin, which specifically binds and stabilizes the *c-MYC* promoter G-quadruplex, functions as a transcription repressor [143], while NM23-H2 destabilizes the *c-MYC* promoter G-quadruplex by binding to the single-stranded form and functions as a transcription activator [144,145]. In addition to the *c-MYC* promoter, it has been demonstrated *in vivo* that the G-quadruplexes form in supercoiled duplex DNA in the *VEGF* promoter [112], and *in vitro,* using single-molecule fluorescent resonance energy transfer, that a *c-KIT* promoter quadruplex was able to form within the context of double-stranded DNA [146].

Parallel-stranded G-quadruplex structure formed in the human c-MYC promoter

G-quadruplexes formed in the c-MYC promoter—The 27-mer NHE III_1 G-rich sequence MycPu27 is comprised of five consecutive runs of guanines (with three runs composed of four guanines each and two runs composed of three guanines each) (Figure 8A) and can form multiple G-quadruplex structures [110].

Native gel EMSA and DMS footprinting data showed that the 5′-run of guanines (G-tract I) is not involved in the formation of the major G-quadruplex structure of Pu27 [111]. The major G-quadruplex formed by the four consecutive 3′ runs of guanines (Myc2345) is a mixture of four loop isomers. The predominant loop isomer has been shown to be the one with G14 and G23 in the loop region, using DMS cleavage and EMSA data [147]. The folding topology of the major c-MYC G-quadruplex Myc2345 has been shown to be parallel-stranded [147,148]. The minor G-quadruplex formed in Pu27 using the 1,2,4,5 G-tracts (Myc1245) [111] has also been shown to adopt a parallel-stranded folding (Figure 8B) [148].

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

Molecular structure of the major G-quadruplexes formed in the c-MYC promoter—We have determined the solution structure of the major G-quadruplex in the c-MYC silencer element by NMR (Figure 8B), the first molecular structure of a G-quadruplex found to form in the promoter region of an oncogene [149]. The selected sequence MYC22 (Figure 8A) contains two G-to-T mutations at residue positions 14 and 23 of the wild-type c-MYC sequence, which restrict the mixture of loop isomers to the single predominant conformation of the c-MYC G-quadruplex in K^+ solution. The line width and resolution of MYC22 are much improved compared with that of the wild-type sequence (Figure 8C). Many inter-residue NOE crosspeaks are observed in 2D-NOESY (Figure 8D), which define the overall conformation of the major c-MYC G-quadruplex, Myc2345. The NOE connectivities were used for NOE-restrained molecular dynamics (RMD) calculation, which produced a family of 20 lowest-energy refined structures, with the average rms deviation being 0.88 Å including all atoms [149].

A representative structure of the NMR-refined solution structures of the major c-MYC G-quadruplex formed by MYC22 is shown in Figure 9. This is the first molecular structure showing the parallel-stranded motif with a single-nucleotide double-strand-reversal loop (G_3NG_3), a robust motif that appears to be prevalent in the G-quadruplex-forming promoter sequences. The intramolecular Myc2345 G-quadruplex consists of three G-tetrads with four parallel DNA strands that are linked by three double-strand-reversal side loops, including a double-nucleotide loop and two single-nucleotide loops. The four grooves of this G-quadruplex are of medium width. Both of the single-nucleotide loops adopt a similar conformation, with an extended sugar backbone strand while the thymine base sticks out into the solvent. The sugar backbone of the double nucleotide T14A15 loop is not as extended as the single thymine loop, with the thymine T14 base sticking out into the solvent and the adenine A15 base located toward the G-tetrad groove with the H2 end pointing to the top of the quadruplex.

The stability of the single-nucleotide parallel-stranded motif is nicely explained by the molecular structure. The right-handed twist of the two adjacent guanine strands makes the 3′-end of one G-strand very close in space to the 5′-end of the next G-strand, so that the single-nucleotide double-strand-reversal loop is very favored (Figure 9A). In fact, the single-nucleotide loop appears to be more stable than the double-nucleotide loop based on the variable temperature study, which indicates that the two ends (G16 and G13) of the TA-loop are the first melting points of the Myc2345 G-quadruplex [149].

The Myc2345 G-quadruplex structure contains very well-defined capping structures at both ends. The 5′-TGA flanking strand adopts a more or less left-handed twist backbone conformation. The A6 residue stacks very well with the G7 of the bottom tetrad, while the G5 residue lays on the side of G7 and G20 (Figure 9A). The 3′-flanking TAA adopts a well-defined fold-back structure, capping the 3′-end of the c-MYC G-quadruplex structure (Figure 9B). The terminal A25 folds back to stack with the G9 of the top G-tetrad, whereas the T23 base stacks on top of the G22 residue and forms Hoogsteen hydrogen bonding with A25. The A24 residue is stacked on top of the T23:A25 base pair. T23 is a mutation from the wild-type G23. Our molecular dynamics modeling of this G-quadruplex structure with the wild-type 3′-flanking GAA sequence demonstrated a similar stable fold-back conformation, with G23 nicely base-paired with A25 using Hoogsteen hydrogen bonding [149] (Figure 9C). The melting temperature of MYC22 is over 85°C, comparable to that of the wild-type mycPu27. The wild-type *c-MYC* promoter DNA mycPu27 has extended flanking DNA sequences and can adopt the necessary capping conformations to stabilize the major G-quadruplex conformation. It is noteworthy that the unique capping structures of the Myc2345 G-quadruplex may provide potential binding sites for drug targeting.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

Hybrid-type G-quadruplex structure formed in the human *BCL-2* promoter

The *BCL-2* gene product is a mitochondrial membrane protein that exists in delicate balance with other related proteins and functions as an inhibitor of cell apoptosis [150,151]. BCL-2 has been found to be aberrantly overexpressed in a wide range of human tumors, including B-cell and T-cell lymphomas and breast, prostate, cervical, colorectal and nonsmall cell lung carcinomas [152–157]. The major promoter P1 of the human *BCL-2* gene is a TATA-less, GC-rich promoter containing multiple transcriptional start sites and is located within a nuclease hypersensitive site (Figure 10A) [158]. A 39-bp GC-rich region adjacent to the 5′-end of the P1 promoter has been implicated as playing a major role in the regulation of BCL-2 transcription [159,160] and has been shown to bind to multiple transcription factors [158, 160–162].

The 39-mer G-rich strand of the DNA in this region contains six runs of guanines separated by one or more bases, with one run of five guanines, two runs of four guanines each and three runs of three guanines each, capable of forming multiple intramolecular G-quadruplex structures (Figure 10B). By using the polymerase stop assay and DMS footprinting, it has been demonstrated that the G-quadruplex structure formed from the central four runs of guanines was the most stable G-quadruplex (bcl2Mid, Figure 10B) [115]. However, the 23mer *BCL-2* promoter DNA (bcl2Mid) of this middle region contains a run of five guanines, thus it is possible to form three loop isomers in which the three consecutive guanines involved in the G-tetrads are either at the 3′- or the 5′-end, or in the middle, straddled by one guanine on each side (Figure 10B). Using mutational analysis and NMR spectroscopy, we tested a set of 23-mer *BCL-2* promoter sequences with dual G-to-T mutations to isolate each loop isomer; we found the 15,16-G-to-T dual mutant (bcl2Mid-15,16T) forms the most stable intramolecular G-quadruplex in this *BCL-2* promoter DNA (Figure 10C) [116]. Additionally, the two mutated guanines G15 and G16 give rise to the clear DMS cleavage pattern found in the wild-type sequence, indicating that these two guanines are not involved in the G-tetrad formation of the major G-quadruplex under physiological conditions [115].

Folding topology and molecular structure of the major G-quadruplex formed in the *BCL-2* promoter—Using NMR spectroscopy, the folding topology of this major G-quadruplex formed in the *BCL-2* promoter region has been determined [116]. Bcl2Mid-15,16T forms a hybrid-2 type intramolecular G-quadruplex structure with mixed parallel/antiparallel G-strands that are connected by a single-nucleotide double-chain-reversal side loop and two lateral loops (Figure 11A). The first, third and fourth G-strands are parallel with each other and the second G-strand runs antiparallel to the other G-strands. The first three-nucleotide CGC loop in the *BCL-2* promoter sequence forms a lateral loop and appears to largely determine the overall folding of the bcl2Mid-15,16T G-quadruplex.

The molecular structure of this major G-quadruplex formed in the *BCL-2* promoter has also been determined using a NOE-restrained distance geometry and RMD approach [114]. A total of 476 NOE distance restraints, of which 168 are from inter-residue NOE interactions, were incorporated into the structure calculation. The bcl2Mid-15,16T G-quadruplex structure is very well defined, as the average rms deviation is 1.04 Å for the core G-tetrads and the first CGC loop of the 10 lowest energy structures. A representative model of this well-defined G-quadruplex structure is shown in Figure 11B. The G-quadruplex consists of three G-tetrads with mixed G-arrangements (the top two with *anti:syn:anti:anti* and the bottom with the reversed *syn:anti:syn:syn* G-arrangements) and four accessible grooves of different widths (one wide, one narrow and two medium). The three loops interact with the core G-tetrads in specific ways that define and stabilize the unique bcl2Mid-15,16T G-quadruplex structure. The first C4-G5-C6 loop adopts a lateral loop conformation with G5 and C6 stacking over the top G-tetrad, while the C4 residue falls into the wide groove I (Figure 11C). The unique

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

conformation of C4 explains the unusual NOE interactions observed between C4 and G3/G5. A10 and T15 of the second lateral loop A10-G11-G12-A13-A14-T15-T16 are uniquely paired using reversed Watson–Crick hydrogen bonding, and are well stacked with the bottom G-tetrad (Figure 11D). The remaining four residues of this long lateral loop are not very well defined and are mostly exposed to the solvent.

We have carried out systematic analysis of the bcl2Mid *BCL-2* promoter sequence with mutated loop residues to determine their functional role in G-quadruplex formation and stability [114]. The mutational analysis agrees very well with the NMR structure. For example, the two mutated residues, T15 and T16, display much higher tolerance, indicating a greater conformational flexibility of these two positions. C4 is positioned in the groove and cannot tolerate any substitution, while A10, involved in a well-defined H-bonded conformation, stacks extensively with G9 of the bottom tetrad and also cannot tolerate a substitution. The rest of the loop residues display much higher flexibility and can tolerate most substitutions. Indeed, G11 and G12, which adopt an extended loop conformation, show enhanced cleavage in the DMS footprinting study. Finally, the C20 single-nucleotide double-chain-reversal loop appears to be able to tolerate various substitutions.

Importantly, the major *BCL-2* promoter G-quadruplex (Figure 11B) is distinct from the one formed in the *c-MYC* promoter (Figure 9A) with different folding patterns, loop conformations and capping structures. It is noteworthy that the major BCL-2 G-quadruplex also contains a parallel-stranded structure motif with a single-nt double-chain-reversal loop, the C20 loop, whose conformation is similar to those observed in the c-MYC G-quadruplex.

Snapback G-quadruplex structure formed in the human *c-KIT* promoter

The c-KIT protein is a receptor tyrosine kinase that is overexpressed in many human cancers. Two G-quadruplex-forming sequences (c-KIT21 and c-KIT87up) within a nuclease hypersensitive region of the *c-KIT* promoter have been discovered [119,120]. An intramolecular parallel-stranded G-quadruplex has been suggested to form in c-KIT21 in the presence of K^+ [120]

The c-KIT87up sequence has four runs of three contiguous guanines separated by 1, 4 and 4 nt (Figure 12A). The G-quadruplex formation was first shown by NMR and UV thermal melting analysis on c-KIT87up in K^+ solution [119]. Interestingly, NMR studies showed that this c-KIT promoter c-KIT87up sequence forms a unique intramolecular G-quadruplex structure, namely the snapback parallel-stranded G-quadruplex (Figure 12B) [163]. The NMR structure of this snapback G-quadruplex showed that, despite the presence of four three-guanine runs in this sequence, the backbone of one G-strand corner of the G-tetrad core was interrupted, with a single guanine (G10) from the linker and two consecutive guanines (G21 and G22) from the fourth three-guanine run (Figure 12A). The G-quadruplex has three G-tetrads and the guanines all have *anti* glycosidic conformations, both features that are consistent with parallel-stranded G-quadruplexes. However, in addition to three double-chain-reversal loops that are observed in the normal intramolecular parallel-stranded G-quadruplex, the snapback parallel-stranded G-quadruplex also contains a lateral loop. Notably, this G-quadruplex structure also contains one G_3NG_3 sequence, which forms a parallel-stranded structure motif with a single-nt double-chain-reversal loop, which may serve as the starting point of this snapback structure (Figure 12A).

G-quadruplexes formed in other human promoters

The proximal promoters of the other human oncogenes, such as *VEGF*, *HIF-1α*, *RET*, and *KRAS*, *PDGF-A* and *c-MYB* also contain polyG/C potential regulatory elements and have been shown to form intramolecular G-quadruplex structures [112,113,118,121,122,164]. The G-

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

quadruplexes formed in the proximal promoter regions of *VEGF*, *HIF-1α*, *RET*, and *KRAS* have been proposed to be of parallel-stranded folding with three G-tetrads [112,113,118, 121]. The proximal promoter region of *PDGF-A* (Figure 12C) has been proposed to form a mixture of parallel-stranded G-quadruplexes with four-G-tetrads [122]. The GC-rich proximal promoter of human *c-MYB* is unique in its sequence, which contains three imperfect copies of $(GGA)_4$ (Figure 12D). The DNA secondary structure formed by $(GGA)_4$ in K^+ solution has been determined by NMR, consisting of a stacked heptad and tetrad (Figure 1Cv) [165]. Two $d(GGA)_4$ sequences form an intermolecular dimer consisting of a tetrad:heptad:heptad:tetrad (T:H:H:T), which is stabilized through the stacking interaction between the two heptads, and a $d(GGA)_8$ sequence forms an intramolecular T:H:H:T structure [166]. The CD and DMS footprinting studies showed that any two of the three GGAGGAGGAGG repeats of the *c-MYB* promoter can form the intramolecular T:H:H:T dimer structure [164]. The mutational analysis indicated the most preferred intramolecular T:H:H:T G-quadruplex structure is formed with two consecutive 5′ end GGAGGAGGAGG repeats. The luciferase reporter analysis showed that deletion of one or two $(GGA)_4$ repeats increased *c-MYB* promoter activity, while deletion of all three $(GGA)_4$repeats resulted in complete abrogation of *c-MYB* activity [164].

Insights into possible rules for the folding patterns of intramolecular G-quadruplexes

The *BCL-2* promoter sequence exhibits some degree of sequence similarity to one of the mutant *c-MYC* promoter sequence (Myc1245, Figure 6). Both sequences contain an extended second loop (7 nt in bcl2Mid and 6 nt in Myc1245) and a 1-nt third loop; however, the first loop in the bcl2Mid is 3 nt in length whereas the first loop in Myc1245 is 1 nt. The bcl2Mid forms a hybrid-type G-quadruplex in K^+ solution [114,116], while Myc1245 forms a parallel-stranded G-quadruplex [148]. The 3-nt first loop of bcl2Mid adopts a lateral loop conformation connecting two antiparallel G-strands, while the 1-nt first loop of Myc1245 adopts a double-chain-reversal loop conformation connecting two parallel G-strands. Consequently, the extended second loop of bcl2Mid is a lateral loop connecting two antiparallel G-strands, while the extended second loop of Myc1245 is a double-chain-reversal loop connecting two parallel G-strands (Figure 6). The 3-nt first loop of bcl2Mid appears to largely determine the overall folding of the hybrid-type, mixed parallel/anti-parallel-stranded BCL-2 G-quadruplex. It is thus clearly indicated that the parallel-stranded structure motif with double-chain-reversal loop favors short-sized loops such as 1- or 2-nt loops, with the 1-nt loop being the most favored, as observed in the G3NG3 sequence motif.

Both the *BCL-2* and *c-MYC* promoter sequences contain the same G_3NG_3 sequence motif, which forms a stable parallel-stranded structural motif with the 1-nt double-chain-reversal loop in K^+ solution. The robust parallel-stranded structural motif formed by G_3NG_3 may serve as a stable foundation for the formation of each G-quadruplex structure. It is interesting to note that the G_3NG_3 motif is commonly observed in the G-rich proximal promoter regions of human oncogenes (Figure 6). Similar to the *c-MYC* promoter sequence, the major G-quadruplex-forming sequences of the promoter regions of *c-KIT* (c-KIT21), *VEGF*, *HIF-1α* and *REl* contain the G_3NG_3 sequence motif at both ends and are suggested to form parallel-stranded intramolecular G-quadruplexes as well.

G-quadruplexes as potential anticancer drug targets

Introduction

Recognition of the biological significance of DNA G-quadruplexes has intensified research and development into G-quadruplex interactive compounds. Drug targeting of secondary DNA structures represents an entirely novel approach to anticancer drug design and development. The therapeutic possibilities of targeting G-quadruplexes were first reported by Sun *et al.* in 1997 [167], namely, that quadruplex ligands can inhibit telomerase, an enzyme that is activated

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

in most cancer cells. The subsequent discovery of the perylene derivative PIPER, which drove the formation of G-quadruplexes [168] and inhibited Sgs1-mediated G-quadruplex unwinding [44], laid the foundation for the existence of a broader mechanism for G-quadruplex ligands. Since then, diverse families of small-molecule compounds with improved specificity and affinity have been identified and/or developed, and the search for better agents is ongoing. Approaches include *in silico* and conventional screening methods, as well as rational structure-based drug design [4]. Structural data have been playing an important role in the design and development of G-quadruplex-interactive compounds. A common feature among the G-quadruplex-targeting ligands is the presence of a fused-ring system that is capable of stacking interactions with the terminal G-tetrads. Many of these ligands also contain side-chain substituents with cationic charges that have the propensity to interact with G-quadruplex grooves.

The compounds discussed in this review have all been reported to selectively bind the G-quadruplex structure over other forms of DNA. Different G-quadruplex-interactive compounds appear to selectively bind different types of G-quadruplexes, and this selectivity is likely to be related to their biological activity. For example, telomestatin has been shown to be a highly potent inhibitor of telomerase by facilitating and stabilizing the formation of the basket-type telomeric G-quadruplex [169,170]. A G-quadruplex-targeting drug, quarfloxacin® (CX-3543, Cylene Pharmaceuticals in San Diego, CA, USA), shows excellent *in vivo* activity in various solid tumors and is currently in Phase II clinical trials. Some other G-quadruplex-interactive compounds have become prospective anticancer agents that display relatively low cytotoxicity. In fact, G-quadruplex-interactive compounds themselves have contributed immensely to understanding G-quadruplexes as a therapeutic target.

G-quadruplex-interactive compounds

Quarfloxin—Quarfloxin (Figure 13A) is a first-in-class drug currently in Phase II clinical trials as a single-agent therapy for neuroendorine tumors. Quarfloxin is well tolerated in patients, and is anticipated to be in additional Phase II trials for other indications owing to its potent *in vivo* efficacy in a broad range of tumors and its considerable therapeutic window. Quarfloxin originated from the design of G-quadruplex-interactive fluoroquinolones by Laurence Hurley and his laboratory [171]. Using fluoroquinolones, a class of drugs that target duplex DNA–gyrase or DNA–topoisomerase II complexes, as a scaffold from which to design a G-quadruplex-interactive compound, the Hurley lab first developed QQ98 from norfloxacin via QQ58. QQ98 is a dual topoisomerase II/G-quadruplex-interactive compound. Based on QQ98, Cylene Pharmaceuticals developed the fluoroquinolone CX-3543 (quarfloxin) as a first-in-class G-quadruplex-interactive drug. No major toxicities have been observed with quarfloxin, and its selectivity appears to arise from both its selective uptake into the nucleolus in cancer cells and its ability to target a G-quadruplex found in rDNA versus duplex DNA (400-fold selectivity) [18]. Quarfloxin was originally designed to target the G-quadruplex found in the *c-MYC* promoter, and was shown to be highly selective for the G-quadruplex versus duplex or single-stranded DNA, and selective for the c-MYC G-quadruplex versus other G-quadruplexes. Quarfloxin lacks topoisomerase-II and -I poisoning activities and has shown potent and tumor-selective activity *in vitro* and *in vivo*. A subsequent study by Cyclene showed that quarfloxin targets a G-quadruplex found in ribosomal DNA (rDNA), which is of the same type as the *c-MYC* promoter G-quadruplex. rDNA contains vast sequences that can form G-quadruplexes, some of which have high affinity for the nucleolin protein [17]. Quarfloxin has been found to accumulate in the cell nucleolus and disrupt the binding of rDNA quadruplexes and nucleolin, redistribute nucleolin, inhibit ribosome biogenesis and induce apoptosis quite specifically in cancer cells, but not in normal cells.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

BRACO19—BRACO19, 3, 6, 9-*tri*-substituted acridine 9-[4-(*N*,*N*-dimethylamino) phenylamino]-3,6-*bis*(3-pyrrolodinopropionamido) acridine (Figure 13B), is a rationally designed trisubsti-tuted acridine derivative that appears to directly target telomeres. BRACO19 was one of the series of the compounds designed by computer modeling to exploit the unique structural features of G-quadruplex DNA, namely, to interact with three grooves of G-quadruplexes in addition to stacking on the terminal G-quartet [172,173]. A disubstituted acridine molecule that was known to bind with quadruplexes was used as a starting model and different substitutions were carried out at the third position. BRACO19 has very low cytotoxicity but much higher G-quadruplex-binding and telomerase-inhibitory activity than its parent compound. Subsequently, BRACO19 was shown to inhibit the catalytic function of telomerase in human cancer cells and destabilize the telomere capping complex [87]. BRACO19 also induced extensive end-to-end chromosomal fusions that are consistent with telomere uncapping. A distinct difference between BRACO19 and direct hTERT inhibitors that result in gradual telomere loss, is that BRACO19 produces anti-tumor effects soon after treatment. BRACO19 induced significant tumor regression within 7–10 days of the initiation of treatment in a DU-145 prostate cancer xenograft [174]. The exposure of 21NT human breast cancer cells to nonacute cytotoxic concentrations of BRACO19 resulted in a marked reduction in cell growth after only 15 days [175]. BRACO19 shows high *in vivo* activity against different types of cancers. In a xenograft model established from UXF1138L uterus carcinoma cells, BRACO-19 inhibited growth by 96% compared with controls, paralleled by loss of nuclear hTERT protein expression and an increase in atypical mitosis indicative of telomere dysfunction [176]. Despite all its favorable characteristics, the major limitations of BRACO19 are its lack of membrane permeability and small therapeutic window, which must be addressed before it can be developed into an effective clinical agent [177]. Importantly, studies of BRACO19 have contributed greatly to understanding telomeric G-quadruplex as a therapeutic target.

Telomestatin—A natural product isolated from *Streptomyces anulatus* 3533-SV4, telomestatin (Figure 13C) is the most potent small-molecule inhibitor of telomerase identified so far [169,178]. Based on the structural similarity between telomestatin and a G-quartet, its propensity to be a G-quadruplex ligand was recognized and subsequently confirmed by experimental data and molecular modeling studies [170]. Telomestatin was found to interact preferentially with intramolecular versus intermolecular G-quadruplex structures and has a 70-fold selectivity for intramolecular G-quadruplex structures over duplex DNA [171]. Telomestatin displays promising anticancer activity in human cancer cells. Treatment of different multiple myeloma cell lines with telomestatin for 3–5 weeks at the minimal effective concentration inhibited telomerase activity, reduced telomere length and caused apoptotic cell death [179]. Telomestatin treatment of primary and established acute leukemia cell lines resulted in telomere shortening and apoptosis [180,181] and increased chemosensitivity of these cells [181]. In xenograft studies, tumor tissue from telomestatin-treated animals exhibited marked apoptosis without noticeable toxicity [182]. Telomestatin has also been actively studied for its mechanism of action. Telomestatin can induce telomere dysfunction and impair chromosome integrity. Specifically, it induces the activation of ATM and Chk2 and, subsequently, increases the expression of p21 (CIP1) and p27 (KIP1), probably by inducing the ATM-dependent DNA damage response [183]. This response is probably at least in part due to the inhibition of DNA binding of POT1 and TRF2 and to the erosion of G-tails, as a consequence of G-quadruplex induction and stabilization by telomestatin [184,185]. In addition to being a promising drug candidate, the exceptional telomerase-inhibiting potency and G-quadruplex selectivity make telomestatin one of the most useful molecules in G-quadruplex research.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

TMPyP4—TMPyP4, 5, 10, 15, 20-*tetra*(*N*-methyl-4-pyridyl) porphine chloride (Figure 13D), is the product of a structure-based drug-design attempt, selected based on its physical properties, such as the presence of a fused planar ring system, positive charge and appropriate size to stack with the G-tetrads [186]. As expected, it stabilized G-quadruplexes and effectively inhibited human telomerase in HeLa cell extracts [186]. TMPyP4 exhibited significant selectivity for quadruplex DNA over duplex DNA, stacking externally with the G-tetrad. The selectivity of TMPyP4 was further substantiated in comparison with its close analogue TMPyP2, which is a poor G-quadruplex-interactive compound and has much-reduced telomerase-inhibitory activity [168]. *In vivo* data consistent with these *in vitro* data have also been obtained, in which, unlike TMPyP2, TMPyP4 decreased the rate of proliferation of sea urchin embryos and trapped the cells in mitosis. It generated anaphase bridges and displayed a chromosome destabilization-mediated antiproliferative effect [187]. Recent elucidation of the mechanism of TMPyP4 has revealed that it interacts with the G-quadruplex formed in the promoter region of *c-MYC* gene [111]. Consequently, TMPyP4 downregulates c-MYC, and this may contribute to the observed effects on telomerase by lowering hTERT, a downstream target of c-MYC [137]. However, a major hurdle in its development as a G-quadruplex target agent is its ability to bind to duplex DNA [188], RNA, RNA–DNA hybrids [189] and triplex DNA [190]. Thus, attempts have been made to generate second-generation cationic porphyrins with high selectivity for G-quadruplexes [191].

307A—307A is a 2, 6-pyridine-dicarboxamide derivative (Figure 13e) that was identified based on its high selectivity of G-quadruplex-forming oligo-nucleotides compared with mutated ones that cannot form G-quadruplexes [192]. It was further shown to be equipotent against c-MYC and telomeric G-quadruplex-forming sequences, and to be potentially useful in inhibiting *c-MYC* gene transcription in tumor cells. 307A and other members of this series inhibited cell proliferation at very low concentrations and induced apoptosis within a few days in a dose-dependent manner in telomerase-positive glioma cells. They also had antiproliferative effects in SAOS-2, a cell line in which telomere maintenance involves the ALT mechanism. The apoptosis induced by these compounds was preceded by multiple alterations of the cell cycle, and these effects were not associated with telomere shortening but directly related to telomere instability involving telomere-end fusion and anaphase-bridge formation [193]. In a more recent study, the role of 307A and its close analog 360A in G-quadruplex stabilization were further substantiated, wherein incubation of oligomers with these compounds resulted in extensive formation of G-quadruplexes *in vitro*. 360A was also shown not only to bind to and lock into preformed quadruplexes, but also to act as a chaperone for tetramolecular complexes [194].

RHPS4—The five-ring acridine, 3, 11-difluoro-6,8,13-trimethyl-8H-quino[4,3,2-kl] acridinium methosulfate (Figure 13F), RHPS4 was initially thought to inhibit telomerase owing to structural similarities to previously identified G-quadruplex interactive telomerase inhibitors [86]. Long-term exposure of 21NT breast cancer or A431 vulval cancer cells to nonacute cytotoxic concentrations of RHPS4 resulted in a decrease in cell growth after 15 days, as well as a marked reduction in cellular telomerase activity and cell cycle changes. An interesting observation is the lack of reduction in telomere length during this period. Further studies showed that the short-term biological activity of RHPS4 was not caused by telomere shortening but by telomere dysfunction, based on the presence of telomeric fusions, polynucleated cells and typical images of telophase bridges in treated cells [195]. RHPS4 was also shown to interact with quadruplexes by stacking at the end. The efficacy of RHPS4 as a part of a combination therapy has been studied; the combination of RHPS4 with paclitaxol (or taxol), doxorubicin (or adriamycin) and the experimental therapeutic agent 17-(allylamino)-17-demethoxygeldanamycin conferred increased sensitivity to RHPS4-treated MCF-7 cells [196]. In another study, the combination of RHPS4 with taxol caused tumor remissions and

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

further enhancement of telomere dysfunction [197]. RHPS4 is a highly promising G-quadruplex ligand in preclinical development and it is anticipated to enter the clinic in the near future [198].

Triazines—Triazines are a new series of G-quadruplex agents that exhibit potent and specific antitelomerase activity with an IC_{50} in the nanomolar range [85]. 12459 (Figure 13g) and another triazine known as 115404 were initially chosen based on the delayed growth arrest in tumor cells treated with subapoptotic doses in a long-term study [85]. Initial studies on the activity of 115404 in telomerase-negative ALT cells showed a mechanism different from telomerase inhibition. Subsequent studies with 12459 showed that telomerase activity is not the main target for 12459 but hTERT functions are likely to contribute to a resistance phenotype [199]. In addition to senescence, 12459 also induces apoptosis on short-term treatment of the human lung adenocarcinoma A549 cells [85], with a parallel rapid degradation of the telomeric G-overhang [200]. The apoptosis is mediated, at least in part, by BCL-2, which is not a critical determinant of the long-term senescence induced by 12459 [200]. Interestingly, 12459 also works by a novel mechanism for a G-quadruplex ligand, impairing the splicing machinery of hTERT by stabilizing quadruplexes located in the hTERT intron 6 [201].

Other potential G-quadruplex-targeting drugs

Numerous other G-quadruplex-interacting pharmacophores and agents, including diamidoanthraquinones such as GSU1051 [167,202], perylenes such as PIPER [203] and quindoline derivatives [204], have shown various levels of selectivity and potency in binding to G-quadruplexes. Many such agents are currently in various stages of preclinical testing and some of them will likely enter the clinic in the near future. In addition to being potential drug targets, DNA G-quadruplexes have also been shown to be potential cancer therapeutics themselves. AS1411 (Antisoma, London, UK) is an unmodified 26-base G-quadruplex-forming oligonucleotide that can inhibit the growth of malignant cells by inducing apoptosis. AS1411 is currently in Phase II trials for the treatment of renal cancer and acute myeloid leukemia. AS1411 has been shown to have cancer-selective antiproliferative activity against a wide range of malignant cell types [205]. The mode of action of AS1411 is suggested to be related to its binding to cell surface nucleolin, whose density is much higher in cancer cells, to induce internalization of the oligonucleotide into the cells where it affects other nucleolin functions. Mechanistic studies indicated that stable G-quadruplex formation and binding to nucleolin protein were necessary for the nuclease resistance, cellular uptake and activity of AS1411. Finally, antibodies that can specifically bind G-quadruplexes are being developed [11,206].

Conclusion

The DNA G-quadruplex secondary structures have been demonstrated to be potential regulatory elements in regions of biological significance, such as in human telomeres and in the promoter regions of a number of important growth-related genes. Significantly, DNA G-quadruplexes can readily form in solution under physiological conditions. One important feature for the intra-molecular DNA G-quadruplexes is that they are globularly folded nucleic acid structures, which are uniquely determined by the primary nucleotide sequences up to 20–30 bases in length, in a manner analogous to protein folding from a primary amino acid sequence. Thus, unlike duplex B-DNA, NMR solution structures of globularly-folded intramolecular G-quadruplexes can be calculated from an arbitrary extended single-stranded DNA as a starting model. The presence of G-quadruplexes in different regions of the genome may provide a design challenge for targeted drug development; however, the molecular structures of intramolecular G-quadruplexes appear to differ from one another and, therefore, in principle, can be differentially regulated and targeted by different proteins and drugs. As

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

discussed in this review, the great conformational diversity of the G-quadruplex DNA secondary structures comes not only from different folding patterns, but also from specific loop conformations and capping structures. These specific folding and loop structures provide unique drug-binding pocket(s) for each G-quadruplex.

Future perspective

More studies are expected to emerge on promoter DNA G-quadruplexes and their structures, as more gene promoters are found to be able to form G-quadruplexes and multiple G-quadruplexes may form in a single gene promoter. Structural studies of promoter quadruplexes are expected to provide new insights into G-quadruplex topology and DNA sequence-structure relationships. While more quadruplex-specific ligands are being identified, understanding of their binding is currently limited as few structures are available. More structural studies on drug-G-quadruplex complexes are anticipated to address this need. In addition, a better understanding of the biological roles of G-quadruplexes and G-quadruplex-interactive proteins should emerge from research targeting these issues. We also expect to see more reports on quadruplexes formed in RNAs.

From a drug-development perspective, research on developing quadruplex-interactive compounds is expected to be highly active. Small-molecule compounds will still remain the focus for quadruplex-interactive drug design, with a particular emphasis on the selectivity toward G-quadruplex structure over duplex DNA. New quadruplex ligands are also expected to cover a wide range of chemical entities. While the interest in telomere-targeting agents is likely to continue, more interest will be seen in promoter G-quadruplex-targeting agents. Additional quadruplex-targeting compounds are expected to enter preclinical and clinical studies in the near future. More studies are also expected to come out on the mechanism and clinical potential of quadruplex ligands, especially in view of the frequency of potential quadruplexes in the human genome.

Executive summary

- DNA G-quadruplexes have recently emerged as a new class of novel molecular targets for anticancer drugs.
- A growing list of proteins has been identified to interact with DNA G-quadruplex structures, strongly pointing to the existence of quadruplex DNA *in vivo*.
- Intramolecular G-quadruplexes are of intensive current research interest due to their potential formation in telomeres and oncogene promoter regions. Intramolecular structures form quickly in solution under physiological conditions and exhibit great conformational diversity, such as in folding topologies, loop conformations and capping structures.
- The first biologically relevant G-quadruplex formation was observed in telomeric DNA. G-quadruplex-interactive compounds can inhibit telomerase activity, thereby making the intramolecular human telomeric DNA G-quadruplex an attractive target for cancer therapeutic intervention.
- Intramolecular G-quadruplexes formed in extended human telomeric sequences in K^+ solution appear to be two equilibrating hybrid-type structures determined by loop conformations and capping structures, with the hybrid-2 form being the major conformation; both forms provide an efficient scaffold for a compact-stacking multimer structure in human telomeric DNA.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

- The highly conserved human telomeric sequence with the same tandem repeats can adopt various G-quadruplex forms with low energy differences; structural polymorphism appears to be intrinsic to the human telomeric sequence.
- The potential occurrence of DNA G-quadruplex has been found to be concentrated in the proximal promoter regions of genes involved in growth and proliferation. The formation of G-quadruplex secondary structures in a gene promoter appears to be closely related to transcription-induced supercoiling.
- The G-rich sequences of gene promoters are often more complex, with each promoter sequence being unique and capable of forming multiple quadruplex structures.
- A highly conserved NHE III_1 in the proximal region of the *c-MYC* promoter controls 80–90% of the *c-MYC* transcriptional activity and forms parallel-stranded G-quadruplexes, which have been shown to be a silencer element. The molecular structure of the major *c-MYC* promoter G-quadruplex provides insights into the robust formation of the single-nucleotide parallel-stranded motifs (G_3NG_3) in K^+ solution.
- The major G-quadruplex formed in a G/C-rich region in the proximal promoter of the *BCL-2* gene adopts a hybrid-type G-quadruplex structure, distinct from that formed in the *c-MYC* promoter.
- A unique snapback G-quadruplex structure is formed in the human *c-KIT* promoter sequence.
- The c-*MYC*, *BCL-2*, and *c-KIT* promoter regions, as well as the promoter regions of *VEGF*, *HIF-1α*, *RET*, and *KRAS*, all contain the G_3NG_3 sequence motif. The robust 1-nt loop parallel-stranded structural motif formed by G_3NG_3 may serve as a stable foundation for G-quadruplex formation.
- Recognition of the biological significance of DNA G-quadruplexes has intensified research and development of G-quadruplex-interactive compounds, some of which are prospective anticancer agents that display relatively low cytotoxicity. A G-quadruplex-targeting drug has entered Phase II clinical trials.

Acknowledgments

We thank Dr Megan Carver for proofreading the paper.

Key Terms

DNA G-quadruplexes	DNA secondary structures formed in specific G-rich sequences. DNA G-quadruplexes are four-stranded DNA secondary structures, consisting of stacked G-tetrads, square planar platforms of four guanines connected by cyclic Hoogsteen hydrogen bonding
Oncogene promoters	Numerous oncogenes have been found to be TATA-less and contain G-rich/C-rich tracts in the proximal regions of promoters. The potential occurrence of DNA G-quadruplex has been suggested for the G-rich/C-rich oncogene promoters
syn or *anti*	Sugar glycosidic conformation of a nucleotide
Intramolecular G-quadruplex	G-quadruplex formed by one DNA molecule

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

Telomeric DNA	Chromosome ends, consisting of tandem repeats of the hexanucleotide $(TTAGGG)_n$ for humans, 5–10 kb in length, terminating in a single-stranded 3′-overhang of 35–600 bases. The G-rich human telomeric DNA has a strong propensity to form the DNA G-quadruplex secondary structure, which can inhibit the activity of telomerase
Quadruplex structure diversity	Conformational variations of quadruplexes, including different folding topologies, different loop and capping conformations

Bibliography

1. Gellert M, Lipsett MN, Davies DR. Helix formation by guanylic acid. Proc Natl Acad Sci USA 1962;48:2013–2018. [PubMed: 13947099]
2. Sen D, Gilbert W. A sodium-potassium switch in the formation of four-stranded G4-DNA. 1990;344 (6265):410–414.
3. Hud, NV.; Plavec, J. The role of cations in determining quadruplex structure and stability. In: Neidle, S., editor. Quadruplex Nucleic Acids. Royal Society of Chemistry Publishing; Cambridge UK: 2006. p. 100-130.
4. Neidle S, Parkinson G. Telomere maintenance as a target for anticancer drug discovery. Nat Rev Drug Discov 2002;1(5):383–393. [PubMed: 12120414]
5. Dai JX, Carver M, Yang DZ. Polymorphism of human telomeric quadruplex structures. Biochimie 2008;90(8):1172–1183. [PubMed: 18373984]
6. Punchihewa, C.; Yang, D. Therapeutic targets and drugs II: G-quadruplex and G-quadruplex inhibitors. In: Hiyama, K., editor. Cancer Drug Discovery and Development:Telomeres and Telomerase in Cancer. Humana Press; Totowa, NJ, USA: 2009. p. 251-280.
7. Qin Y, Hurley LH. Structures folding patterns and functions of intramolecular DNA G-quadruplexes found in eukaryotic promoter regions. Biochimie 2008;90(8):1149–1171. [PubMed: 18355457]
8. Henderson E, Hardin CC, Walk SK, Tinoco I Jr, Blackburn EH. Telomeric DNA oligonucleotides form novel intramolecular structures containing guanine-guanine base pairs. Cell 1987;51(6):899–908. [PubMed: 3690664]
9. Sundquist WI, Klug A. Telomeric DNA dimerizes by formation of guanine tetrads between hairpin loops. Nature 1989;342(6251):825–829. [PubMed: 2601741]
10. Williamson JR, Raghuraman MK, Cech TR. Monovalent cation-induced structure of telomeric DNA: the G-quartet model. Cell 1989;59(5):871–880. [PubMed: 2590943]
11. Schaffitzel C, Berger I, Postberg J, Hanes J, Lipps HJ, Pluckthun A. *In vitro* generated antibodies specific for telomeric guanine-quadruplex DNA react with Stylonychia lemnae macronuclei. Proc Natl Acad Sci USA 2001;98(15):8572–8577. [PubMed: 11438689]
12. 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(10):847–854. [PubMed: 16142245]
13. Granotier C, Pennarun G, Riou L, et al. Preferential binding of a G-quadruplex ligand to human chromosome ends. Nucl Acids Res 2005;33(13):4182–4190. [PubMed: 16052031]
14. Chang CC, Kuo IC, Lin JJ, et al. A novel carbazole derivative BMVC: a potential antitumor agent and fluorescence marker of cancer cells. Chem Biodivers 2004;1(9):1377–1384. [PubMed: 17191915]
15. Huppert JL, Balasubramanian S. G-quadruplexes in promoters throughout the human genome. Nucl Acids Res 2007;35(2):406–413. [PubMed: 17169996]
16. Verma A, Yadav VK, Basundra R, Kumar A, Chowdhury S. Evidence of genome-wide G4 DNA-mediated gene expression in human cancer cells. Nucl Acids Res 2009;37(13):4194–4204. [PubMed: 19211664]
17. Maizels N. Dynamic roles for G4 DNA in the biology of eukaryotic cells. Nat Struct Mol Biol 2006;13 (12):1055–1059. [PubMed: 17146462]

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

18. Drygin D, Siddiqui-Jain A, O'Brien S, et al. Anticancer activity of CX-3543: a direct inhibitor of rRNA biogenesis. Cancer Res 2009;69(19):7653–7661. [PubMed: 19738048]
19. Wieland M, Hartig JS. RNA quadruplex-based modulation of gene expression. Chem Biol 2007;14 (7):757–763. [PubMed: 17656312]
20. Kumari S, Bugaut A, Huppert JL, Balasubramanian S. An RNA G-quadruplex in the 5′ UTR of the NRAS proto-oncogene modulates translation. Nat Chem Biol 2007;3(4):218–221. [PubMed: 17322877]
21. Kostadinov R, Malhotra N, Viotti M, Shine R, D'Antonio L, Bagga P. GRSDB: a database of quadruplex forming G-rich sequences in alternatively processed mammalian pre-mRNA sequences. Nucl Acids Res 2006;34:D119–D124. [PubMed: 16381828]
22. Weisman-Shomer P, Cohen E, Hershco I, et al. The cationic porphyrin TMPyP4 destabilizes the tetraplex form of the fragile X syndrome expanded sequence d(CGG)(n). Nucl Acids Res 2003;31 (14):3963–3970. [PubMed: 12853612]
23. Khateb S, Weisman-Shomer P, Hershco I, Loeb LA, Fry M. Destabilization of tetraplex structures of the fragile X repeat sequence (CGG)n is mediated by homolog-conserved domains in three members of the hnRNP family. Nucl Acids Res 2004;32(14):4145–4154. [PubMed: 15302914]
24. Oganesian L, Bryan TM. Physiological relevance of telomeric G-quadruplex formation: a potential drug target. Bioessays 2007;29(2):155–165. [PubMed: 17226803]
25. Fry M. Tetraplex DNA and its interacting proteins. Front Biosci 2007;12:4336–4351. [PubMed: 17485378]
26. 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(13):1618–1629. [PubMed: 15231739]
27. Fang G, Cech TR. Characterization of a G-quartet formation reaction promoted by the β-subunit of the *Oxytricha* telomere-binding protein. Biochemistry 1993;32:11646–11657. [PubMed: 8218232]
28. Giraldo R, Rhodes D. The yeast telomere-binding protein RAP1 binds to and promotes the formation of DNA quadruplexes in telomeric DNA. EMBO J 1994;13(10):2411–2420. [PubMed: 8194531]
29. Arimondo PB, Riou JF, Mergny JL, et al. Interaction of human DNA topoisomerase I with G-quartet structures. Nucl Acids Res 2000;28(24):4832–4838. [PubMed: 11121473]
30. Marchand C, Pourquier P, Laco GS, Jing N, Pommier Y. Interaction of human nuclear topoisomerase I with guanosine quartet-forming and guanosine-rich single-stranded DNA and RNA oligonucleotides. J Biol Chem 2002;277(11):8906–8911. [PubMed: 11756434]
31. Hanakahi LA, Sun H, Maizels N. High affinity interactions of nucleolin with G-G-paired rDNA. J Biol Chem 1999;274(22):15908–15912. [PubMed: 10336496]
32. Larson ED, Duquette ML, Cummings WJ, Streiff RJ, Maizels N. MutS[α] binds to and promotes synapsis of transcriptionally activated immunoglobulin switch regions. Curr Biol 2005;15(5):470–474. [PubMed: 15753043]
33. Duquette ML, Pham P, Goodman MF, Maizels N. AID binds to transcription-induced structures in c-MYC that map to regions associated with translocation and hypermutation. 2005:5791.
34. Baumann P, Cech TR. Pot1, the putative telomere end-binding protein in fission yeast and humans. Science 2001;292(5519):1171–1175. [PubMed: 11349150]
35. Lei M, Podell ER, Cech TR. Structure of human POT1 bound to telomeric single-stranded DNA provides a model for chromosome end-protection. Nat Struct Mol Biol 2004;11(12):1223–1229. [PubMed: 15558049]
36. Lei M, Podell ER, Baumann P, Cech TR. DNA self-recognition in the structure of Pot1 bound to telomeric single-stranded DNA. Nature 2003;426(6963):198–203. [PubMed: 14614509]
37. Opresko PL, Mason PA, Podell ER, et al. POT1 stimulates RecQ helicases WRN and BLM to unwind telomeric DNA substrates. J Biol Chem 2005;280(37):32069–32080. [PubMed: 16030011]
38. Schramke V, Luciano P, Brevet V, et al. RPA regulates telomerase action by providing Est1p access to chromosome ends. Nat Genet 2004;36(1):46–54. [PubMed: 14702040]
39. Cohen S, Jacob E, Manor H. Effects of single-stranded DNA binding proteins on primer extension by telomerase. Biochim Biophys Acta 2004;1679(2):129–140. [PubMed: 15297146]

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

40. Salas TR, Petruseva I, Lavrik O, et al. Human replication protein A unfolds telomeric G-quadruplexes. Nucl Acid Res 2006;34(17):4857–4865.
41. Zhang QS, Manche L, Xu RM, Krainer AR. hnRNP A1 associates with telomere ends and stimulates telomerase activity. RNA 2006;12(6):1116–1128. [PubMed: 16603717]
42. Fukuda H, Katahira M, Tsuchiya N, et al. Unfolding of quadruplex structure in the G-rich strand of the minisatellite repeat by the binding protein UP1. Proc Natl Acad Sci USA 2002;99(20):12685–12690. [PubMed: 12235355]
43. Enokizono Y, Konishi Y, Nagata K, et al. Structure of hnRNP D complexed with single-stranded telomere DNA and unfolding of the quadruplex by heterogeneous nuclear ribonucleoprotein D. J Biol Chem 2005;280(19):18862–18870. [PubMed: 15734733]
44. Han HY, Bennett RJ, Hurley LH. Inhibition of unwinding of G-quadruplex structures by Sgs1 helicase in the presence of *N,N′-bis*[2-(1-piperidino)ethyl]-3,4,9,10-perylenetetracarboxylic diimide a G-quadruplex-interactive ligand. Biochemistry 2000;39(31):9311–9316. [PubMed: 10924124]
45. Cohen H, Sinclair DA. Recombination-mediated lengthening of terminal telomeric repeats requires the Sgs1 DNA helicase. Proc Natl Acad Sci USA 2001;98(6):3174–3179. [PubMed: 11248051]
46. Huang PH, Pryde FE, Lester D, et al. SGS1 is required for telomere elongation in the absence of telomerase. Curr Biol 2001;11(2):125–129. [PubMed: 11231130]
47. Lin Y-C, Shih J-W, Hsu C-L, Lin J-J. Binding and partial denaturing of G-quartet DNA by Cdc13p of *Saccharomyces cerevisiae*. J Biol Chem 2002;276:47671–47674. [PubMed: 11585819]
48. Orren DK, Theodore S, Machwe A. The Werner syndrome helicase/exonuclease (WRN) disrupts and degrades D-loops *in vitro*. Biochemistry 2002;41(46):13483–13488. [PubMed: 12427008]
49. Lee JY, Kozak M, Martin JD, Pennock E, Johnson FB. Evidence that a RecQ helicase slows senescence by resolving recombining telomeres. PLoS Biol 2007;5(6):e160. [PubMed: 17550308]
50. Li JL, Harrison RJ, Reszka AP, et al. Inhibition of the Bloom's and Werner's syndrome helicases by G-quadruplex interacting ligands. Biochemistry 2001;40(50):15194–15202. [PubMed: 11735402]
51. Liu Z, Frantz JD, Gilbert W, Tye BK. Identification and characterization of a nuclease activity specific for G4 tetrastranded DNA. Proc Natl Acad Sci USA 1993;90(8):3157–3161. [PubMed: 8475054]
52. Sun H, Yabuki A, Maizels N. A human nuclease specific for G4 DNA. Proc Natl Acad Sci USA 2001;98(22):12444–12449. [PubMed: 11675489]
53. Ghosal G, Muniyappa K. Saccharomyces cerevisiae Mre11 is a high-affinity G4 DNA-binding protein and a G-rich DNA-specific endonuclease: implications for replication of telomeric DNA. Nucl Acids Res 2005;33(15):4692–4703. [PubMed: 16116037]
54. Blackburn EH. Telomere states and cell fates. Nature 2000;408(6808):53–56. [PubMed: 11081503]
55. van Steensel B, Smogorzewska A, de Lange T. TRF2 protects human telomeres from end-to-end fusions. Cell 1998;92(3):401–413. [PubMed: 9476899]
56. Hackett JA, Feldser DM, Greider CW. Telomere dysfunction increases mutation rate and genomic instability. Cell 2001;106(3):275–286. [PubMed: 11509177]
57. de Lange T. Shelterin: the protein complex that shapes and safeguards human telomeres. Genes Dev 2005;19(18):2100–2110. [PubMed: 16166375]
58. Moyzis RK, Buckingham JM, Cram LS, et al. A highly conserved repetitive DNA sequence (TTAGGG)n present at the telomeres of human chromosomes. Proc Natl Acad Sci USA 1988;85(18):6622–6626. [PubMed: 3413114]
59. Wright WE, Tesmer VM, Huffman KE, Levene SD, Shay JW. Normal human chromosomes have long G-rich telomeric overhangs at one end. Genes Dev 1997;11(21):2801–2809. [PubMed: 9353250]
60. Colgin LM, Reddel RR. Telomere maintenance mechanisms and cellular immortalization. Curr Opin Genet Dev 1999;9(1):97–103. [PubMed: 10072358]
61. Sfeir AJ, Chai WH, Shay JW, Wright WE. Telomere-end processing: the terminal nucleotides of human chromosomes. Mol Cell 2005;18(1):131–138. [PubMed: 15808515]
62. Court R, Chapman L, Fairall L, Rhodes D. How the human telomeric proteins TRF1 and TRF2 recognize telomeric DNA: a view from high-resolution crystal structures. EMBO Rep 2005;6(1):39–45. [PubMed: 15608617]

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

63. Xin H, Liu D, Wan M, et al. TPP1 is a homologue of ciliate TEBP-β and interacts with POT1 to recruit telomerase. Nature 2007;445(7127):559–562. [PubMed: 17237767]
64. Wang F, Podell ER, Zaug AJ, et al. The POT1-TPP1 telomere complex is a telomerase processivity factor. Nature 2007;445(7127):506–510. [PubMed: 17237768]
65. Hurley LH, Wheelhouse RT, Sun D, et al. G-quadruplexes as targets for drug design. Pharmacol Ther 2000;85(3):141–158. [PubMed: 10739869]
66. Bodnar AG, Ouellette M, Frolkis M, et al. Extension of life-span by introduction of telomerase into normal human cells. Science 1998;279(5349):349–352. [PubMed: 9454332]
67. Sun D, Lopez-Guajardo CC, Quada J, Hurley LH, Von Hoff DD. Regulation of catalytic activity and processivity of human telomerase. Biochemistry 1999;38(13):4037–4044. [PubMed: 10194316]
68. Harley CB, Futcher AB, Greider CW. Telomeres shorten during ageing of human fibroblasts. Nature 1990;345(6274):458–460. [PubMed: 2342578]
69. Hastie ND, Dempster M, Dunlap MG, Thompsson AM, Green DK, Allshire RC. Telomere reduction in human colorectal carcinoma with aging. Nature 1990;346:866–868. [PubMed: 2392154]
70. Mehle C, Ljungberg B, Roos G. Telomere shortening in renal cell carcinoma. Cancer Res 1994;54:236–241. [PubMed: 8261445]
71. Greider CW, Blackburn EH. Identification of a specific telomere terminal transferase activity in Tetrahymena extracts. Cell 1985;43(2):405–413. [PubMed: 3907856]
72. Kim NW, Piatyszek MA, Prowse KR, et al. Specific association of human telomerase activity with immortal cells and cancer. Science 1994;266(5193):2011–2015. [PubMed: 7605428]
73. Hanahan D, Weinberg RA. The hallmarks of cancer. Cell 2000;100(1):57–70. [PubMed: 10647931]
74. Zahler AM, Williamson JR, Cech TR, Prescott DM. Inhibition of telomerase by G-quartet DNA structures. Nature 1991;350(6320):718–720. [PubMed: 2023635]
75. Griffith J, Bianchi A, de Lange T. TRF1 promotes parallel pairing of telomeric tracts *in vitro*. J Mol Biol 1998;278:79–88. [PubMed: 9571035]
76. Stansel RM, de Lange T, Griffith JD. T-loop assembly *in vitro* involves binding of TRF2 near the 38 telomeric overhang. EMBO J 2001;20(19):5532–5540. [PubMed: 11574485]
77. Sen D, Gilbert W. Formation of parallel four-stranded complexes by guanine-rich motifs in DNA and its implications for meiosis. Nature 1988;334(6180):364–366. [PubMed: 3393228]
78. Anuradha S, Muniyappa K. Molecular aspects of meiotic chromosome synapsis and recombination. Prog Nucl Acid Res Mol Biol 2005;79:49–132.
79. Shukla AK, Roy KB. Rec A-independent homologous recombination induced by a putative fold-back tetraplex DNA. Biol Chem 2006;387(3):251–256. [PubMed: 16542145]
80. Hurley LH. DNA and its associated processes as targets for cancer therapy. Nat Rev Cancer 2002;2 (3):188–200. [PubMed: 11990855]
81. Mergny JL, Helene C. G-quadruplex DNA: a target for drug design. Nat Med 1998;4(12):1366–1367. [PubMed: 9846570]
82. De Cian A, Lacroix L, Douarre C, et al. Targeting telomeres and telomerase. Biochimie 2007;90(1): 131–155. [PubMed: 17822826]
83. Karlseder J, Smogorzewska A, de Lange T. Senescence induced by altered telomere state not telomere loss. Science 2002;295(5564):2446–2449. [PubMed: 11923537]
84. Blackburn EH. Switching and signaling at the telomere. Cell 2001;106(6):661–673. [PubMed: 11572773]
85. Riou JF, Guittat L, Mailliet P, et al. Cell senescence and telomere shortening induced by a new series of specific G-quadruplex DNA ligands. Proc Natl Acad Sci USA 2002;99(5):2672–2677. [PubMed: 11854467]
86. Gowan SM, Heald R, Stevens MF, Kelland LR. Potent inhibition of telomerase by small-molecule pentacyclic acridines capable of interacting with G-quadruplexes. Mol Pharmacol 2001;60(5):981–988. [PubMed: 11641426]
87. Incles CM, Schultes CM, Kempski H, Koehler H, Kelland LR, Neidle S. A G-quadruplex telomere-targeting agent produces p16-associated senescence and chromosomal fusions in human prostate cancer cells. Mol Cancer Therap 2004;3(10):1201–1206. [PubMed: 15486186]

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

88. Wang Y, Patel DJ. Solution structure of the human telomeric repeat d[AG3(T2AG3)3] G-tetraplex. Structure 1993;1(4):263–282. [PubMed: 8081740]

89. Parkinson GN, Lee MP, Neidle S. Crystal structure of parallel quadruplexes from human telomeric DNA. Nature 2002;417(6891):876–880. [PubMed: 12050675]

90. Ambrus A, Chen D, Dai J, Bialis T, Jones RA, Yang DZ. Human telomeric sequence forms a hybrid-type intramolecular G-quadruplex structure with mixed parallel/antiparallel strands in potassium solution. Nucl Acids Res 2006;34(9):2723–2735. [PubMed: 16714449]

91. Dai JX, Carver M, Punchihewa C, Jones RA, Yang DZ. Structure of the hybrid-2 type intramolecular human telomeric G-quadruplex in K^+ solution: insights into structure polymorphism of the human telomeric sequence. Nucl Acids Res 2007;35(15):4927–4940. [PubMed: 17626043]

92. Dai JX, Punchihewa C, Ambrus A, Chen D, Jones RA, Yang DZ. Structure of the intramolecular human telomeric G-quadruplex in potassium solution: a novel adenine triple formation. Nucl Acids Res 2007;35(7):2440–2450. [PubMed: 17395643]

93. Xu Y, Noguchi Y, Sugiyama H. The new models of the human telomere d[AGGG(TTAGGG)(3)] in K^+ solution. Bioorg Med Chem 2006;14(16):5584–5591. [PubMed: 16682210]

94. Luu KN, Phan AT, Kuryavyi V, Lacroix L, Patel DJ. Structure of the human telomere in K^+ solution: an intramolecular (3+1) G-quadruplex scaffold. J Am Chem Soc 2006;128(30):9963–9970. [PubMed: 16866556]

95. Phan AT, Kuryavyi V, Luu KN, Patel DJ. Structure of two intramolecular G-quadruplexes formed by natural human telomere sequences in K^+ solution. Nucl Acids Res 2007;35(19):6517–6525. [PubMed: 17895279]

96. Petraccone L, Trent JO, Chaires JB. The tail of the telomere. J Am Chem Soc 2008;130(49):16530–16532. [PubMed: 19049455]

97. Chang CC, Chien CW, Lin YH, Kang CC, Chang TC. Investigation of spectral conversion of d(TTAGGG)(4) and d(TTAGGG)(13) upon potassium titration by a G-quadruplex recognizer BMVC molecule. Nucl Acids Res 2007;35(9):2846–2860. [PubMed: 17430965]

98. Haider S, Parkinson GN, Neidle S. Molecular dynamics and principal components analysis of human telomeric quadruplex multimers. Biophys J 2008;95(1):296–311. [PubMed: 18375510]

99. Renciuk D, Kejnovska I, Skolakova P, Bednarova K, Motlova J, Vorlickova M. Arrangements of human telomere DNA quadruplex in physiologically relevant K^+ solutions. Nucl Acids Res 2009;37(19):6625–6634. [PubMed: 19717545]

100. Lim KW, Amrane S, Bouaziz S, et al. Structure of the human telomere in K^+ solution: a stable basket-type G-quadruplex with only two G-tetrad layers. J Am Chem Soc 2009;131(12):4301–4309. [PubMed: 19271707]

101. Zhang Z, Dai J, Veliath E, Jones RA, Yang D. Structure of a two-G-tetrad intramolecular G-quadruplex formed by a variant human telomeric sequence in K^+ solution: insights into the interconversion of human telomeric G-quadruplex structures. Nucl Acid Res 2010;38(3):1009–1021.

102. Antonacci C, Chaires JB, Sheardy RD. Biophysical characterization of the human telomeric $(TTAGGG)_{(4)}$repeat in a potassium solution. Biochemistry 2007;46(15):4654–4660. [PubMed: 17381076]

103. Gray RD, Chaires JB. Kinetics and mechanism of K^+- and Na^+-induced folding of models of human telomeric DNA into G-quadruplex structures. Nucl Acids Res 2008;36(12):4191–4203. [PubMed: 18567908]

104. Rustighi A, Tessari MA, Vascotto F, Sgarra R, Giancotti V, Manfioletti G. A polypyrimidine/polypurine tract within the Hmga2 minimal promoter: a common feature of many growth-related genes. Biochemistry 2002;41(4):1229–1240. [PubMed: 11802722]

105. Hurley LH, Siddiqui-Jain A. Developing therapeutics to target oncogenes. Genet Eng News 2005;25(1):26–26.

106. Eddy J, Maizels N. Gene function correlates with potential for G4 DNA formation in the human genome. Nucl Acids Res 2006;34(14):3887–3896. [PubMed: 16914419]

107. Larsen A, Weintraub H. An altered DNA conformation detected by s1 nuclease occurs at specific regions in active chick globin chromatin. Cell 1982;29(2):609–622. [PubMed: 6288265]

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

108. Wood WI, Felsenfeld G. Chromatin structure of the chicken B-globin gene region – sensitivity to Dnase-I, micrococcal nuclease and Dnase-II. J Biol Chem 1982;257(13):7730–7736. [PubMed: 6282852]
109. Hammond-Kosack MC, Dobrinski B, Lurz R, Docherty K, Kilpatrick MW. The human insulin gene linked polymorphic region exhibits an altered DNA structure. Nucl Acids Res 1992;20(2):231–236. [PubMed: 1741248]
110. Simonsson T, Pecinka P, Kubista M. DNA tetraplex formation in the control region of c-MYC. Nucl Acids Res 1998;26(5):1167–1172. [PubMed: 9469822]
111. 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. PNAS 2002;99(18): 11593–11598. [PubMed: 12195017]
112. Sun DY, Guo KX, 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. Nucl Acids Res 2005;33(18):6070–6080. [PubMed: 16239639]
113. De Armond R, Wood S, Sun DY, Hurley LH, Ebbinghaus SW. Evidence for the presence of a guanine quadruplex forming region within a polypurine tract of the hypoxia inducible factor 1α promoter. Biochemistry 2005;44(49):16341–16350. [PubMed: 16331995]
114. Dai JX, Chen D, Jones RA, Hurley LH, Yang DZ. NMR solution structure of the major G-quadruplex structure formed in the human BCL2 promoter region. Nucl Acids Res 2006;34(18):5133–5144. [PubMed: 16998187]
115. Dexheimer TS, Sun D, Hurley LH. Deconvoluting the structural and drug-recognition complexity of the G-quadruplex-forming region upstream of the bcl-2 P1 promoter. J Am Chem Soc 2006;128 (16):5404–5415. [PubMed: 16620112]
116. Dai J, Dexheimer TS, Chen D, et al. 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(4):1096–1098. [PubMed: 16433524]
117. Yafe A, Etzioni S, Weisman-Shomer P, Fry M. Formation and properties of hairpin and tetraplex structures of guanine-rich regulatory sequences of muscle-specific genes. Nucl Acids Res 2005;33 (9):2887–2900. [PubMed: 15908587]
118. Cogoi S, Quadrifoglio F, Xodo LE. G-rich oligonucleotide inhibits the binding of a nuclear protein to the Ki-ras promoter and strongly reduces cell growth in human carcinoma pancreatic cells. Biochemistry 2004;43(9):2512–2523. [PubMed: 14992588]
119. Rankin S, Reszka AP, Huppert J, et al. Putative DNA quadruplex formation within the human c-KIT oncogene. J Am Chem Soc 2005;127(30):10584–10589. [PubMed: 16045346]
120. Fernando H, Reszka AP, Huppert J, et al. A conserved quadruplex motif located in a transcription activation site of the human c-kit oncogene. Biochemistry 2006;45(25):7854–7860. [PubMed: 16784237]
121. Guo K, Pourpak A, Beetz-Rogers K, Gokhale V, Sun D, Hurley LH. Formation of pseudosymmetrical G-quadruplex and i-motif structures in the proximal promoter region of the RET oncogene. J Am Chem Soc 2007;129(33):10220–10228. [PubMed: 17672459]
122. Qin Y, Rezler EM, Gokhale V, Sun D, Hurley LH. Characterization of the G-quadruplexes in the duplex nuclease hypersensitive element of the PDGF-A promoter and modulation of PDGF-A promoter activity by TMPyP4. Nucl Acids Res 2007;35(22):7698–7713. [PubMed: 17984069]
123. Woodford KJ, Howell RM, Usdin K. A novel K^+-dependent DNA-synthesis arrest site in a commonly occurring sequence motif in eukaryotes. J Biol Chem 1994;269(43):27029–27035. [PubMed: 7929444]
124. Etzioni S, Yafe A, Khateb S, Weisman-Shomer P, Bengal E, Fry M. Homodimeric MyoD preferentially binds tetraplex structures of regulatory sequences of muscle-specific genes. J Biol Chem 2005;280(29):26805–26812. [PubMed: 15923190]
125. Marcu KB, Bossone SA, Patel AJ. MYC function and regulation. Ann Rev Biochem 1992;61(1): 809–858. [PubMed: 1497324]
126. Pelengaris S, Khan M. The many faces of c-MYC. Archives of Biochemistry and Biophysics 2003;416(2):129–136. [PubMed: 12893289]

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

127. Kinzler KW, Vogelstein B. Lessons from hereditary colorectal cancer. Cell 1996;87(2):159–170. [PubMed: 8861899]
128. Nupponen NN, Kakkola L, Koivisto P, Visakorpi T. Genetic alterations in hormone-refractory recurrent prostate carcinomas. Am J Pathol 1998;153(1):141–148. [PubMed: 9665474]
129. Dang CV. c-Myc target genes involved in cell growth apoptosis and metabolism. Mol Cell Biol 1999;19(1):1–11. [PubMed: 9858526]
130. Nesbit CE, Tersak JM, Prochownik EV. MYC oncogenes and human neoplastic disease. Oncogene 1999;18(19):3004–3016. [PubMed: 10378696]
131. Schlagbauer-Wadl H, Griffioen M, van Elsas A, et al. Influence of increased c-Myc expression on the growth characteristics of human melanoma. J Investigative Dermatology 1999;112(3):332–336.
132. Marcu, KB.; Patel, AJ.; Yang, Y. C-Myc in B-Cell Neoplasia. Elsevier; Oxford, UK: 1997. Differential regulation of the c-MYC P1 and P2 promoters in the absence of functional tumor suppressors: implications for mechanisms of deregulated MYC transcription; p. 47-56.
133. Sakatsume O, Tsutsui H, Wang Y, et al. Binding of THZif-1, a MAZ-like zinc finger protein to the nuclease-hypersensitive element in the promoter region of the c-MYC protooncogene. J Biol Chem 1996;271(49):31322–31333. [PubMed: 8940139]
134. Tomonaga T, Levens D. Activating transcription from single stranded DNA. PNAS 1996;93(12): 5830–5835. [PubMed: 8650178]
135. Berberich SJ, Postel EH. Puf/Nm23-H2/Ndpk-B transactivates a human C-Myc promoter-cat gene via a functional nuclease hypersensitive element. Oncogene 1995;10(12):2343–2347. [PubMed: 7784082]
136. Michelotti GA, Michelotti EF, Pullner A, Duncan RC, Eick D, Levens D. Multiple single-stranded cis elements are associated with activated chromatin of the human c-myc gene *in vivo*. Molecular and Cellular Biology 1996;16(6):2656–2669. [PubMed: 8649373]
137. Grand CL, Han H, Munoz RM, et al. 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(8):565–573. [PubMed: 12479216]
138. Ou TM, Lu YJ, Zhang C, et al. Stabilization of G-quadruplex DNA and down-regulation of oncogene c-myc by quindoline derivatives. J Med Chem 2007;50(7):1465–1474. [PubMed: 17346034]
139. Liu J-N, Deng R, Guo J-F, et al. Inhibition of myc promoter and telomerase activity and induction of delayed apoptosis by SYUIQ-5, a novel G-quadruplex interactive agent in leukemia cells. Leukemia 2007;21(6):1300–1302. [PubMed: 17392822]
140. Lavelle C. DNA torsional stress propagates through chromatin fiber and participates in transcriptional regulation. Nat Struct Mol Biol 2008;15(2):123–125. [PubMed: 18250629]
141. Kouzine F, Sanford S, Elisha-Feil Z, Levens D. The functional response of upstream DNA to dynamic supercoiling *in vivo*. Nat Struct Mol Biol 2008;15(2):146–154. [PubMed: 18193062]
142. 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(9):2863–2874. [PubMed: 19385599]
143. Gonzalez V, Guo K, Hurley L, Sun D. identification and characterization of nucleolin as a c-myc G-quadruplex-binding protein. J Biol Chem 2009;284(35):23622–23635. [PubMed: 19581307]
144. Dexheimer TS, Carey SS, Zuohe S, et al. NM23-H2 may play an indirect role in transcriptional activation of c-myc gene expression but does not cleave the nuclease hypersensitive element III1. Mol Cancer Ther 2009;8:1363.
145. Thakur RK, Kumar P, Halder K, et al. Metastases suppressor NM23-H2 interaction with G-quadruplex DNA within c-MYC promoter nuclease hypersensitive element induces c-MYC expression. Nucl Acids Res 2009;37(1):172–183. [PubMed: 19033359]
146. Shirude PS, Okumus B, Ying L, Ha T, Balasubramanian S. Single-molecule conformational analysis of G-quadruplex formation in the promoter DNA duplex of the proto-oncogene c-kit. J Am Chem Soc 2007;129(24):7484–7485. [PubMed: 17523641]
147. Seenisamy J, Rezler EM, Powell TJ, et al. The dynamic character of the G-quadruplex element in the c-MYC promoter and modification by TMPyP4. J Am Chem Soc 2004;126(28):8702–8709. [PubMed: 15250722]

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

148. Phan AT, Modi YS, Patel DJ. Propeller-type parallel-stranded G-quadruplexes in the human c-myc promoter. J Am Chem Soc 2004;126(28):8710–8716. [PubMed: 15250723]

149. Ambrus A, Chen D, Dai J, Jones RA, Yang DZ. Solution structure of the biologically relevant G-quadruplex element in the human c-MYC promoter. Implications for G-quadruplex stabilization. Biochemistry 2005;44(6):2048–2058. [PubMed: 15697230]

150. Hockenbery D, Nunez G, Milliman C, Schreiber RD, Korsmeyer SJ. Bcl-2 is an inner mitochondrial membrane protein that blocks programmed cell death. Nature 1990;348(6299):334–336. [PubMed: 2250705]

151. Vaux DL, Cory S, Adams JM. Bcl-2 gene promotes haemopoietic cell survival and cooperates with c-myc to immortalize pre-B cells. Nature 1988;335(6189):440–442. [PubMed: 3262202]

152. Akagi T, Kondo E, Yoshino T. Expression of Bcl-2 protein and Bcl-2 mRNA in normal and neoplastic lymphoid tissues. Leuk Lymphoma 1994;13(1–2):81–87. [PubMed: 8025526]

153. Joensuu H, Pylkkanen L, Toikkanen S. Bcl-2 protein expression and long-term survival in breast cancer. Am J Pathol 1994;145(5):1191–1198. [PubMed: 7977649]

154. Tjalma W, De Cuyper E, Weyler J, et al. Expression of bcl-2 in invasive and in situ carcinoma of the uterine cervix. Am J Obstetrics Gynecology 1998;178(1):113–117.

155. Pezzella F, Turley H, Kuzu I, et al. Bcl-2 protein in non-small-cell lung carcinoma. N Engl J Med 1993;329(10):690–694. [PubMed: 8393963]

156. McDonnell TJ, Troncoso P, Brisbay SM, et al. Expression of the protooncogene bcl-2 in the prostate and its association with emergence of androgen-independent prostate cancer. Cancer Research 1992;52(24):6940–6944. [PubMed: 1458483]

157. Baretton GB, Diebold J, Christoforis G, et al. Apoptosis and immunohistochemical bcl-2 expression in colorectal adenomas and carcinomas. Aspects of carcinogenesis and prognostic significance. Cancer 1996;77(2):255–264. [PubMed: 8625232]

158. Seto M, Jaeger U, Hockett RD, et al. Alternative promoters and exons somatic mutation and deregulation of the Bcl-2-Ig fusion gene in lymphoma. Embo J 1988;7(1):123–131. [PubMed: 2834197]

159. Young RL, Korsmeyer SJ. A negative regulatory element in the bcl-2 5′-untranslated region inhibits expression from an upstream promoter. Mol Cell Biol 1993;13(6):3686–3697. [PubMed: 8388542]

160. Heckman C, Mochon E, Arcinas M, Boxer LM. The WT1 protein is a negative regulator of the normal bcl-2 allele in t(14;18) lymphomas. J Biol Chem 1997;272(31):19609–19614. [PubMed: 9235968]

161. Gomez-Manzano C, Mitlianga P, Fueyo J, et al. Transfer of E2F-1 to human glioma cells results in transcriptional up-regulation of Bcl-2. Cancer Research 2001;61(18):6693–6697. [PubMed: 11559537]

162. Liu YZ, Boxer LM, Latchman DS. Activation of the Bcl-2 promoter by nerve growth factor is mediated by the p42/p44 MAPK cascade. Nucl Acids Res 1999;27(10):2086–2090. [PubMed: 10219080]

163. 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(14):4386–4392. [PubMed: 17362008]

164. Palumbo SL, Memmott RM, Uribe DJ, Krotova-Khan Y, Hurley LH, Ebbinghaus SW. A novel G-quadruplex-forming GGA repeat region in the c-myb promoter is a critical regulator of promoter activity. Nucl Acids Res 2008;36(6):1755–1769. [PubMed: 18252774]

165. Matsugami A, Ouhashi K, Kanagawa M, et al. An intramolecular quadruplex of (GGA)4 triplet repeat DNA with a G:G:G:G tetrad and a G(:A):G(:A):G(:A):G heptad and its dimeric interaction. J Mol Biol 2001;313(2):255–269. [PubMed: 11800555]

166. Matsugami A, Okuizumi T, Uesugi S, Katahira M. Intramolecular higher order packing of parallel quadruplexes comprising a G:G:G:G tetrad and a G(:A):G(:A):G(:A):G heptad of GGA triplet repeat DNA. J Biol Chem 2003;278(30):28147–28153. [PubMed: 12748183]

167. Sun D, Thompson B, Cathers BE, et al. Inhibition of human telomerase by a G-quadruplex-interactive compound. J Med Chem 1997;40(14):2113–2116. [PubMed: 9216827]

168. Han H, Rangan A, Hurley LH. Selective interaction of cationic porphyrins with different types of G-quadruplex structures. Clin Cancer Res 1999;5:3852S–3852S.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

169. Shin-ya K, Wierzba K, Matsuo K, et al. Telomestatin a novel telomerase inhibitor from Streptomyces anulatus. J Am Chem Soc 2001;123(6):1262–1263. [PubMed: 11456694]

170. Kim MY, Vankayalapati H, Shin-Ya K, Wierzba 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(10):2098–2099. [PubMed: 11878947]

171. Kim MY, Duan WH, Gleason-Guzman M, Hurley LH. Design synthesis and biological evaluation of a series of fluoroquino-anthroxazines with contrasting dual mechanisms of action against topoisomerase II and G-quadruplexes. J Med Chem 2003;46(4):571–583. [PubMed: 12570378]

172. Read M, Cuesta J, Basra I, et al. Rational design approaches to increase the potency of G-quadruplex-mediated telomerase inhibitors. Clin Cancer Res 2001;7(11):3797S–3797S.

173. Read M, Harrison RJ, Romagnoli B, et al. Structure-based design of selective and potent G-quadruplex-mediated telomerase inhibitors. Proc Natl Acad Sci USA 2001;98(9):4844–4849. [PubMed: 11309493]

174. Kelland LR. Overcoming the immortality of tumour cells by telomere and telomerase based cancer therapeutics – current status and future prospects. Eur J Cancer 2005;41(7):971–979. [PubMed: 15862745]

175. Gowan SM, Harrison JR, Patterson L, et al. A G-quadruplex-interactive potent small-molecule inhibitor of telomerase exhibiting *in vitro* and *in vivo* antitumor activity. Mol Pharmacol 2002;61 (5):1154–1162. [PubMed: 11961134]

176. Burger AM, Dai FP, Schultes CM, et al. The G-quadruplex-interactive molecule BRACO-19 inhibits tumor growth consistent with telomere targeting and interference with telomerase function. Cancer Res 2005;65(4):1489–1496. [PubMed: 15735037]

177. Taetz S, Baldes C, Murdter TE, et al. Biopharmaceutical characterization of the telomerase inhibitor BRACO19. Pharm Res 2006;23(5):1031–1037. [PubMed: 16715394]

178. De Cian A, Cristofari G, Reichenbach P, et al. Reevaluation of telomerase inhibition by quadruplex ligands and their mechanisms of action. Proc Natl Acad Sci USA 2007;104(44):17347–17352. [PubMed: 17954919]

179. Shammas MA, Reis RJS, Li C, et al. Telomerase inhibition and cell growth arrest after telomestatin treatment in multiple myeloma. Clin Cancer Res 2004;10(2):770–776. [PubMed: 14760100]

180. Nakajima A, Tauchi T, Sashida G, et al. Telomerase inhibition enhances apoptosis in human acute leukemia cells: possibility of antitelomerase therapy. Leukemia 2003;17(3):560–567. [PubMed: 12646945]

181. Sumi M, Tauchi T, Sashida G, et al. A G-quadruplex-interactive agent telomestatin (SOT-095), induces telomere shortening with apoptosis and enhances chemosensitivity in acute myeloid leukemia. Int J Oncol 2004;24(6):1481–1487. [PubMed: 15138591]

182. Tauchi T, Shin-ya K, Sashida G, et al. Telomerase inhibition with a novel G-quadruplex-interactive agent telomestatin: *in vitro* and *in vivo* studies in acute leukemia. Oncogene 2006;25(42):5719–5725. [PubMed: 16652154]

183. Tauchi T, Shin-ya K, Sashida G, et al. 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(34):5338–5347. [PubMed: 12917635]

184. Gomez D, O'Donohue MF, Wenner T, et al. The G-quadruplex ligand telomestatin inhibits POT1 binding to telomeric sequences *in vitro* and induces GFP-POT1 dissociation from telomeres in human cells. Cancer Res 2006;66(14):6908–6912. [PubMed: 16849533]

185. Tahara H, Shin-Ya K, Seimiya H, Yamada H, Tsuruo T, Ide T. G-quadruplex stabilization by telomestatin induces TRF2 protein dissociation from telomeres and anaphase bridge formation accompanied by loss of the 3′ telomeric overhang in cancer cells. Oncogene 2006;25(13):1955–1966. [PubMed: 16302000]

186. Wheelhouse RT, Sun D, Han H, Han FX, Hurley LH. Cationic porphyrins as telomerase inhibitors: the interaction of tetra-(*N*-methyl-4-pyridyl)porphine with quadruplex DNA. J Am Chem Soc 1998;120:3261–3262.

187. Izbicka E, Nishioka D, Marcell V, et al. Telomere-interactive agents affect proliferation rates and induce chromosomal destabilization in sea urchin embryos. Anti-Cancer Drug Des 1999;14(4):355–365.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

188. Guliaev AB, Leontis NB. Cationic 5,10,15,20-tetrakis(*N*-methylpyridinium-4-yl)porphyrin fully intercalates at 5′-CG-3′ steps of duplex DNA in solution. Biochemistry 1999;38(47):15425–15437. [PubMed: 10569925]

189. Uno T, Hamasaki K, Tanigawa M, Shimabayashi S. Binding of meso-tetrakis(*N*-methylpyridinium-4-yl)porphyrin to double helical RNA and DNA–RNA hybrids. Inorg Chem 1997;36(8):1676–1683. [PubMed: 11669759]

190. Lee YA, Kim JO, Cho TS, Song R, Kim SK. Binding of meso-tetrakis(*N*-methylpyridium-4-yl) porphyrin to triplex oligonucleotides: evidence for the porphyrin stacking in the major groove. J Am Chem Soc 2003;125(27):8106–8107. [PubMed: 12837072]

191. Dixon IM, Lopez F, Tejera AM, et al. A G-quadruplex ligand with 10000-fold selectivity over duplex DNA. J Am Chem Soc 2007;129(6):1502–1503. [PubMed: 17283987]

192. Lemarteleur T, Gomez D, Paterski R, Mandine E, Mailliet P, Riou J-F. Stabilization of the c-myc gene promoter quadruplex by specific ligands' inhibitors of telomerase. Biochem Biophys Res Comm 2004;323(3):802–808. [PubMed: 15381071]

193. Pennarun G, Granotier C, Gauthier LR, Gomez D, Boussin FD. Apoptosis related to telomere instability and cell cycle alterations in human glioma cells treated by new highly selective G-quadruplex ligands. Oncogene 2005;24(18):2917–2928. [PubMed: 15735722]

194. De Cian A, Mergny JL. Quadruplex ligands may act as molecular chaperones for tetramolecular quadruplex formation. Nucl Acids Res 2007;35(8):2483–2493. [PubMed: 17395639]

195. Leonetti C, Amodei S, D'Angelo C, et al. Biological activity of the G-quadruplex ligand RHPS4 (3,11-difluoro-6,8,13-trimethyl-8H-quino[4,3,2-kl]acridinium methosulfate) is associated with telomere capping alteration. Mol Pharmacol 2004;66(5):1138–1146. [PubMed: 15304549]

196. Cookson JC, Dai F, Smith V, et al. Pharmacodynamics of the G-quadruplex-stabilizing telomerase inhibitor 3,11-difluoro-6,8,13-trimethyl-8H-quino[4,3,2-kl] acridinium methosulfate (RHPS4) *in vitro*: activity in human tumor cells correlates with telomere length and can be enhanced or antagonized with cytotoxic agents. Mol Pharmacol 2005;68(6):1551–1558. [PubMed: 16150933]

197. Phatak P, Cookson JC, Dai F, et al. 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(8):1223–1233. [PubMed: 17406367]

198. Kelland L. Targeting the limitless replicative potential of cancer: the telomerase/telomere pathway. Clin Cancer Res 2007;13(17):4960–4963. [PubMed: 17785545]

199. Gomez D, Aouali N, Londono-Vallejo A, et al. Resistance to the short term antiproliferative activity of the G-quadruplex ligand 12459 is associated with telomerase overexpression and telomere capping alteration. J Biol Chem 2003;278(50):50554–50562. [PubMed: 14525974]

200. Douarre C, Gomez D, Morjani H, et al. Overexpression of Bcl-2 is associated with apoptotic resistance to the G-quadruplex ligand 12459 but is not sufficient to confer resistance to long-term senescence. Nucl Acids Res 2005;33(7):2192–2203. [PubMed: 15831792]

201. Gomez D, Lemarteleur T, Lacroix L, Mailliet P, Mergny J-L, Riou J-F. Telomerase downregulation induced by the G-quadruplex ligand 12459 in A549 cells is mediated by hTERT RNA alternative splicing. Nucl Acids Res 2004;32(1):371–379. [PubMed: 14729921]

202. Sun D, Thompson B, Salazar M, Jenkins T, Neidle S, Hurley LH. Human telomerase inhibition by G-tetraplex interactive compounds. Proc Am Assoc Cancer Res 1997;38:637.

203. Han HY, Cliff CL, Hurley LH. Accelerated assembly of G-quadruplex structures by a small molecule. Biochemistry 1999;38(22):6981–6986. [PubMed: 10353809]

204. Zhou J-L, Lu Y-J, Ou T-M, et al. Synthesis and evaluation of quindoline derivatives as G-quadruplex inducing and stabilizing ligands and potential inhibitors of telomerase. J Med Chem 2005;48(23): 7315–7321. [PubMed: 16279791]

205. 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. Experimental Mol Pathol 2009;86(3 SI): 151–164.

206. Fernando H, Rodriguez R, Balasubramanian S. Selective recognition of a DNA G-quadruplex by an engineered antibody. Biochemistry 2008;47(36):9365–9371. [PubMed: 18702511]

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

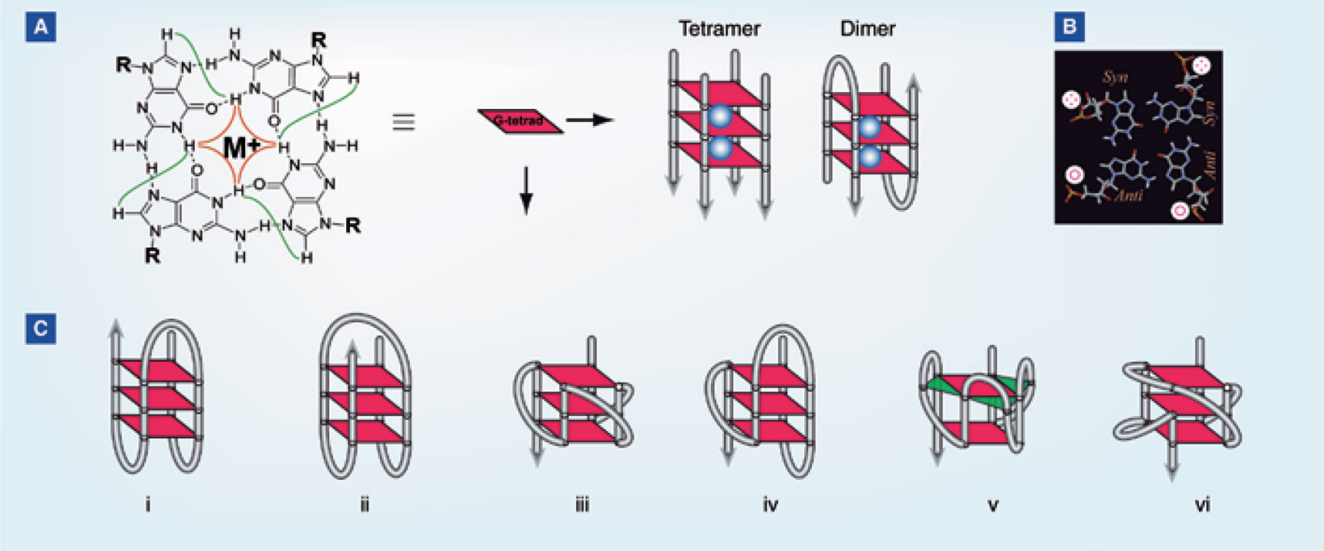

Figure 1.

(A) G-tetrad, a square-planar alignment of four guanines connected by cyclic Hoogsteen hydrogen bonding between the N1, N2 and O6, N7 of guanine bases (left). The H1–H1 and H1–H8 connectivity patterns detectable in NOESY experiments is also shown. A schematic tetrameric and dimeric G-quadruplex composed of three G-tetrads (right). Cations (K^+ or Na^+), shown as blue balls, are needed to stabilize G-quadruplexes by coordinating with the eight electronegative carbonyl oxygen O6 atoms of the adjacent G-tetrads. (B) The guanines in a G-tetrad can adopt either *syn* or *anti* glycosidic conformation. The guanines from the parallel G-strands adopt the same glycosidic conformation and the guanines from the antiparallel G-strands adopt the opposite glycosidic conformations. (C) Examples of monomeric (intramolecular) G-quadruplexes with different folding structures.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

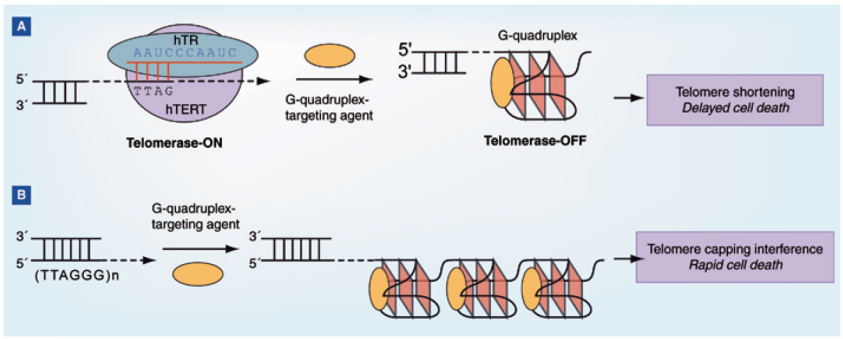

Figure 2.
(A) Mechanism of telomerase inhibition by G-quadruplex-targeting compounds. (B) Mechanism of drug-mediated interference of telomere capping by G-quadruplex-targeting compounds.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

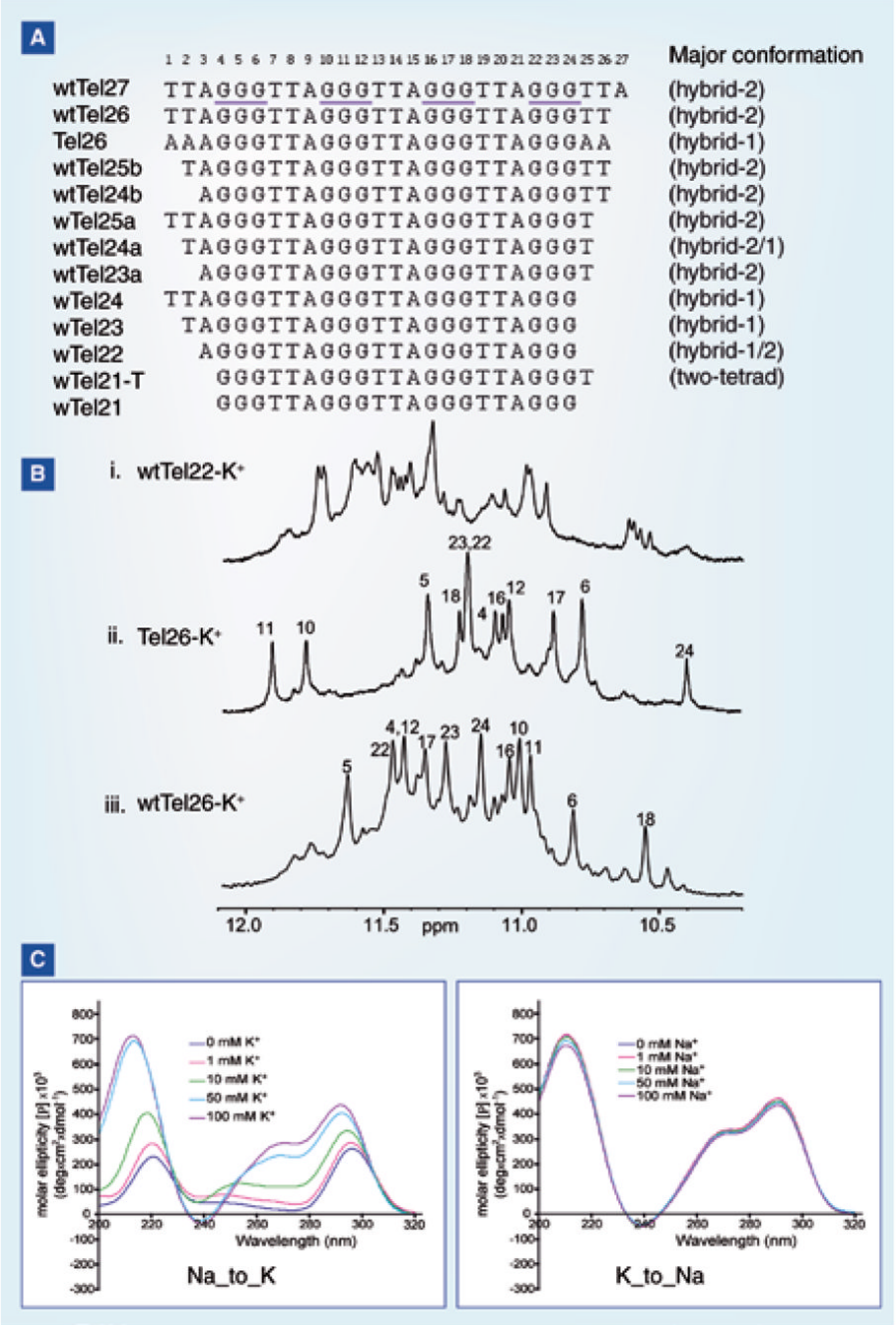

Figure 3.

(A) Four-G-tract human telomeric sequences with different flanking sequences. The numbering system is shown above wtTel27. The major conformation formed in each sequence is indicated. (B) The imino proton region of the 1D ^{1}H NMR of (i) wtTel22, (ii) Tel26 with assignment and (iii) wtTel26 with assignment in K^+ solution. (C) Titration experiments of K^+ in the presence of 150 mM Na^+ for Tel26 (left), and titration experiments of Na^+ in the presence of 100 mM K^+ for Tel26 (right), monitored by circular dichroism spectroscopy.

NIH-PA Author Manuscript
NIH-PA Author Manuscript
NIH-PA Author Manuscript

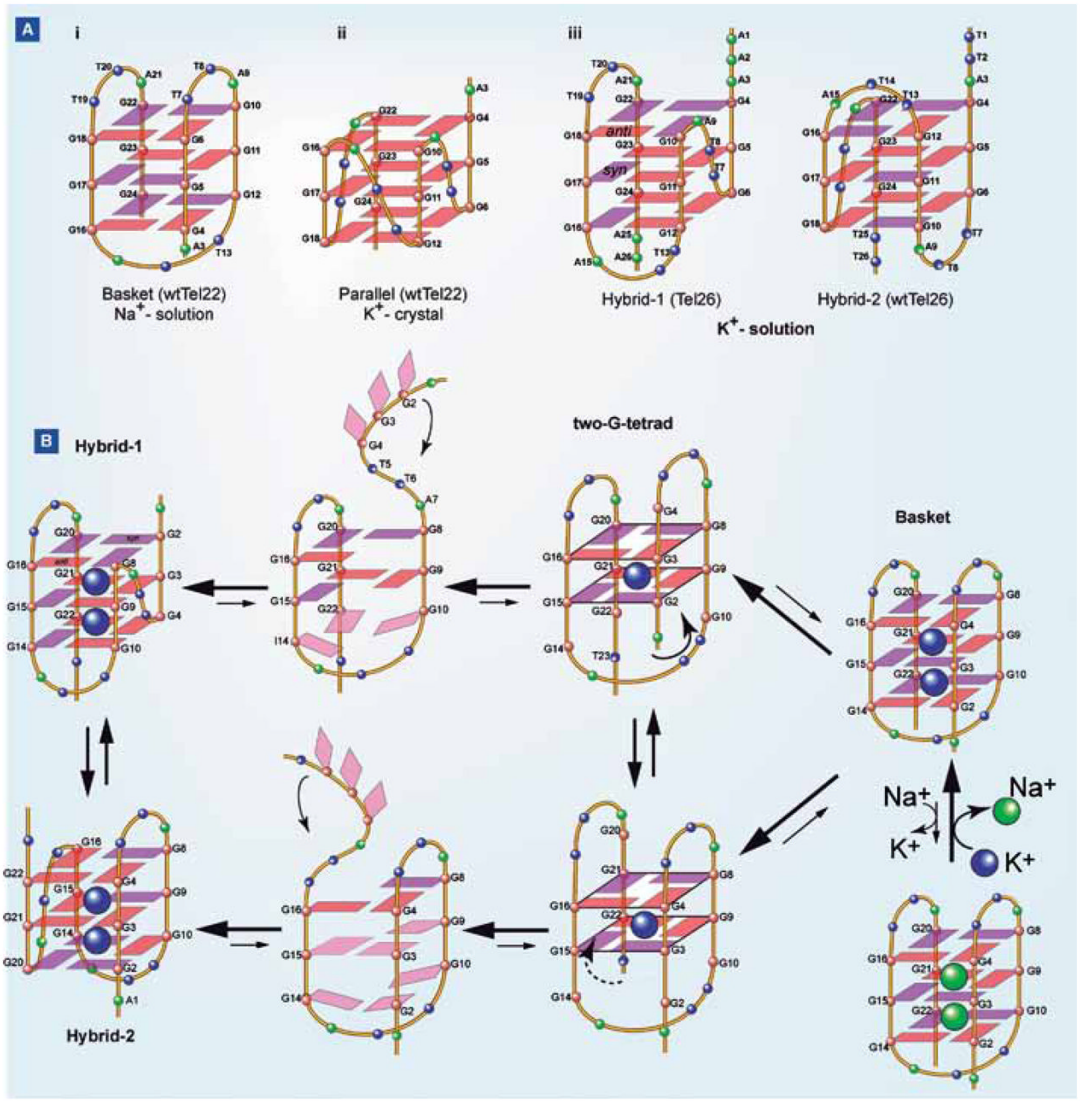

Figure 4.

(A) Folding topology of the basket-type intramolecular G-quadruplex formed by wtTel22 in Na^+ solution as determined by NMR (i). Folding topology of the propeller-type parallel-stranded intramolecular G-quadruplex formed by wtTel22 in the presence of K^+ in crystalline state (ii). Folding topologies of the hybrid-1 (major conformation in Tel26) (iii, left) and hybrid-2 (major conformation in wtTel26) (iii, right) intramolecular G-quadruplexes in K^+ solution. The numbering system is based on wtTel26. Yellow box: (*anti*) guanine; red box: (*syn*) guanine. (B) A model of the interconversion between the basket-type (Na^+) and the hybrid-type (K^+) telomeric G-quadruplexes through a strand-reorientation mechanism. A two-tetrad form is likely to be a transition intermediate of the interconversion between different telomeric structures.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

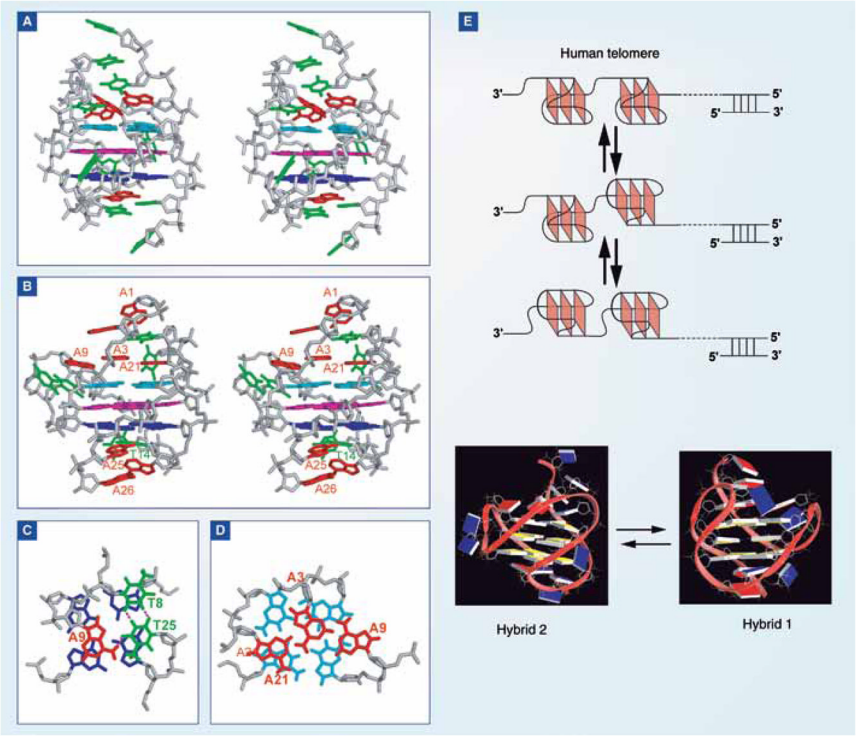

Figure 5.

(A) Stereo view of the representative NMR structure of the hybrid-2 telomeric G-quadruplex formed by wtTel26 in K^+ solution. (B) Stereo view of the representative NMR structure of the hybrid-1 telomeric G-quadruplex formed by Tel26 in K^+ solution. (C) The bottom view of the T:A:T triple capping the bottom G-tetrad (blue) of hybrid-2 telomeric G-quadruplex, with the potential hydrogen bonds shown as dashed lines. (D) Top view of the adenine triple (red) capping the top G-tetrad (cyan) of hybrid-1 telomeric G-quadruplex. The top (from 5′-end) G-tetrad is in cyan, the middle G-tetrad is in magenta, and the bottom G-tetrad is in blue. (E) A model showing a DNA secondary structure composed of compact-stacking multimers of hybrid-type G-quadruplexes in human telomeres, with an equilibrium between hybrid-1 and hybrid-2 forms in K^+ solution. The representative NMR structures, which are distinct, of hybrid-1 and hybrid-2 telomeric G-quadruplexes are shown (guanine: yellow; adenine: red; thymine: blue).

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

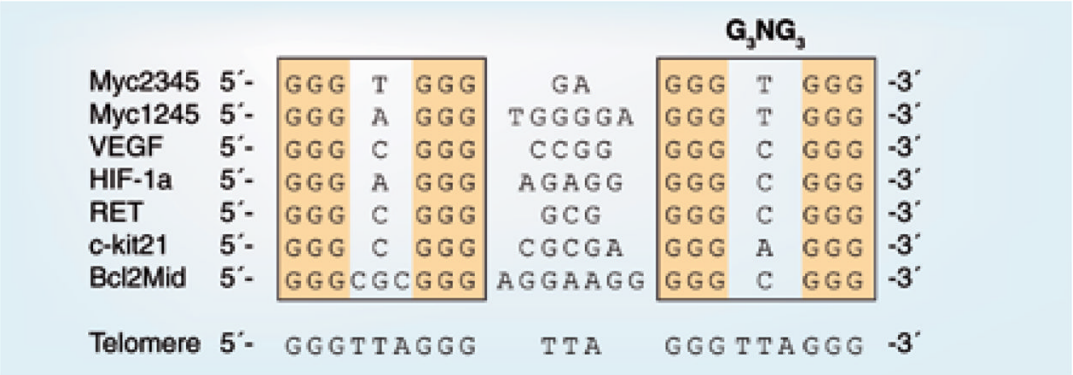

Figure 6.

Comparison of G-quadruplex-forming sequences in selected gene promoters. The telomeric sequence is also shown as a comparison. All the promoter G-rich sequences shown contain the G_3NG_3 motif; except for the BCL-2 sequence, they have all been shown to form parallel-stranded G-quadruplexes.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

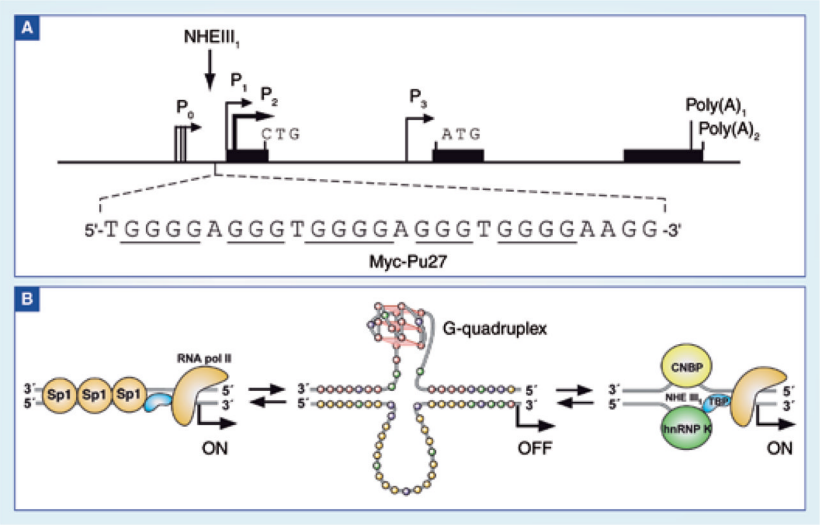

Figure 7.

(A) The promoter structure of the human *c-MYC* gene. The G-rich NHE III_1 is shown, with guanine runs underlined. (B) Alternative forms of the NHE III_1 of the *c-MYC* promoter associated with transcriptional activation or silencing.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

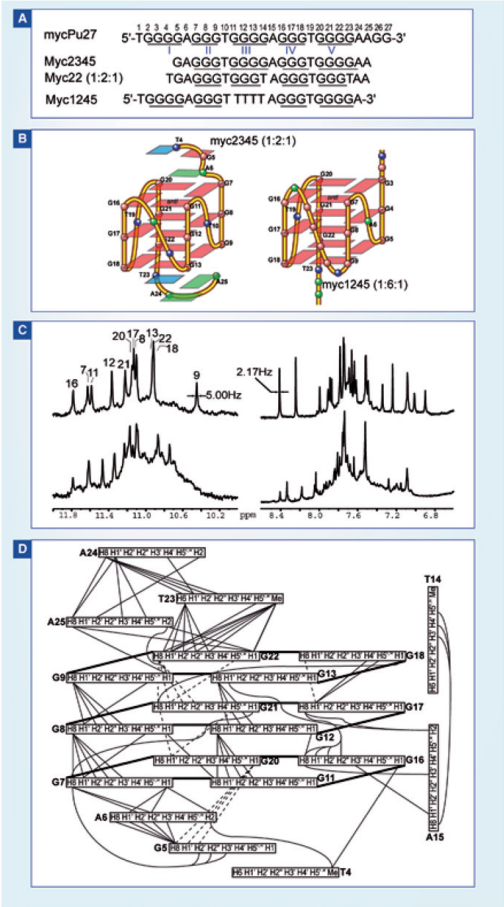

Figure 8.

(A) The *c-MYC* promoter sequence and its modifications. (B) The folding structure of the major G-quadruplex formed in the *c-MYC* promoter, Myc2345(1:2:1) (left) and the minor G-quadruplex, Myc1245 (right). (C) 1D NMR spectra of the T-Myc2345 sequence (bottom) and its major loop isomer Myc22 (top). (D) Inter-residue NOEs of MYC22, which forms the major G-quadruplex of the *c-MYC* promoter. The NOE connectivities clearly define the quadruplex conformation and provide distance restraints for structure calculation.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

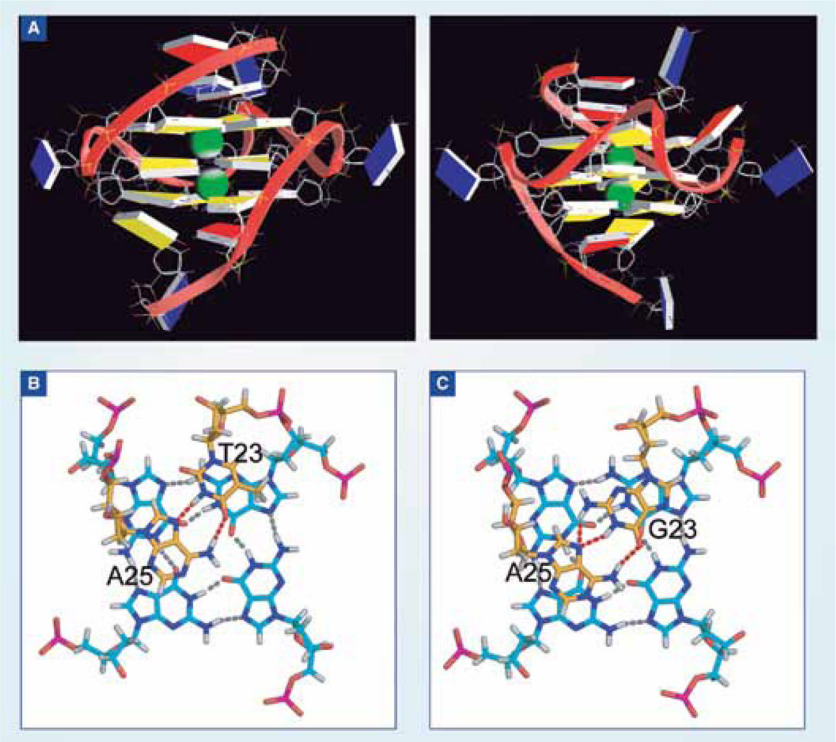

Figure 9.

(A) Representative NMR structure of the major *c-MYC* promoter G-quadruplex formed by Myc22, a parallel-stranded structure, in two opposite views. Two potassium ions coordinated between the G-tetrads are included for the calculation and are shown as green spheres. (Guanine: yellow; adenine: red; thymine: blue) The 3′-end view of the G-quadruplex with mutant T23 (B) and wild-type G23 (C). The hydrogen bonds of the top tetrad (black) and the T/G23:A25 base pair (green) are shown in dashed lines.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

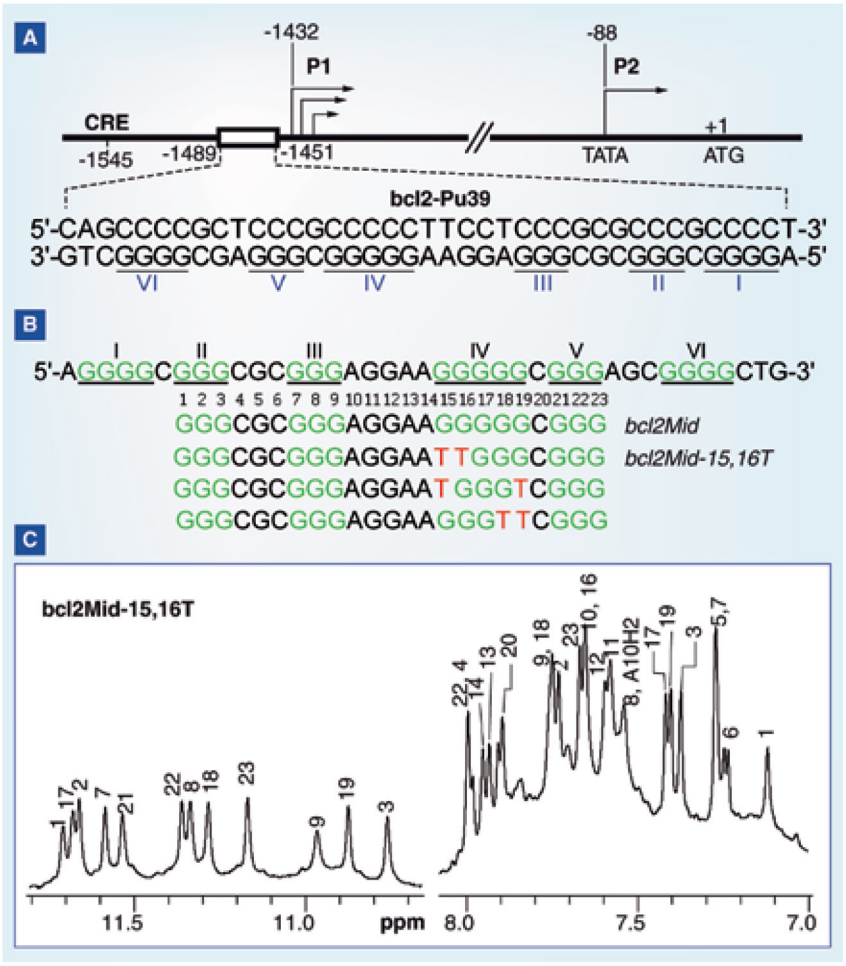

Figure 10.

(A) Promoter structure of the human *BCL-2* gene. The G/C-rich region of the promoter is shown, with guanine runs underlined. (B) The *BCL-2* promoter sequence and its modifications. (C) 1D NMR spectrum of the bcl2Mid-15,16T, with assignments, which forms the major G-quadruplex in the *BCL-2* promoter.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

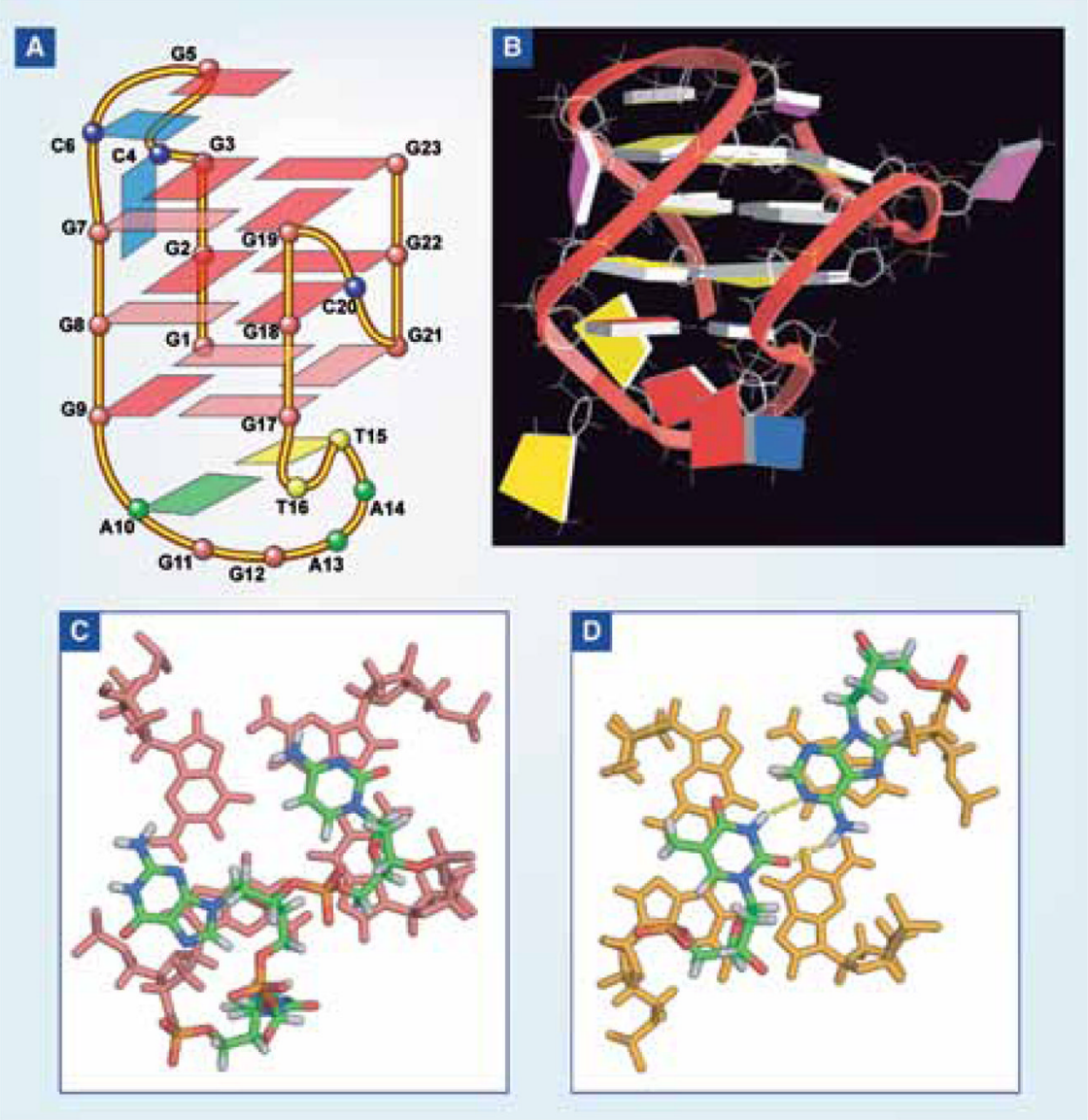

Figure 11.

(A) Folding structure of the major G-quadruplex formed in the *BCL-2* promoter. (B) A representative NMR structure of the major *BCL-2* promoter G-quadruplex. The (C) 5′-end and (D) 3′-end view of the major *BCL-2* promoter G-quadruplex. The hydrogen bonds of the A10:T15 base pair are shown in dashed lines.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

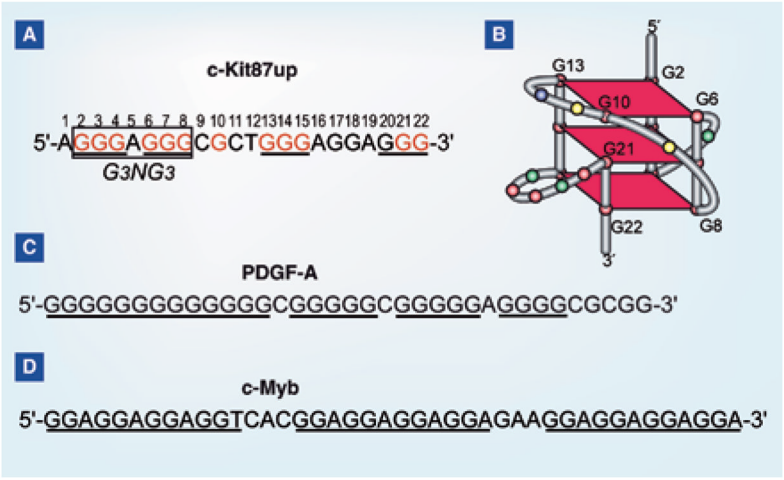

Figure 12.

(A) Sequence of one G-rich region of the human *c-KIT* gene promoter (c-KIT87up). (B) The folding topology of the NMR-determined structure formed by c-KIT87up [168]. (C) The sequence of the G-rich region of the human *PDGF-A* gene promoter. (D) The sequence of the G-rich region of the human *c-MYB* gene promoter.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

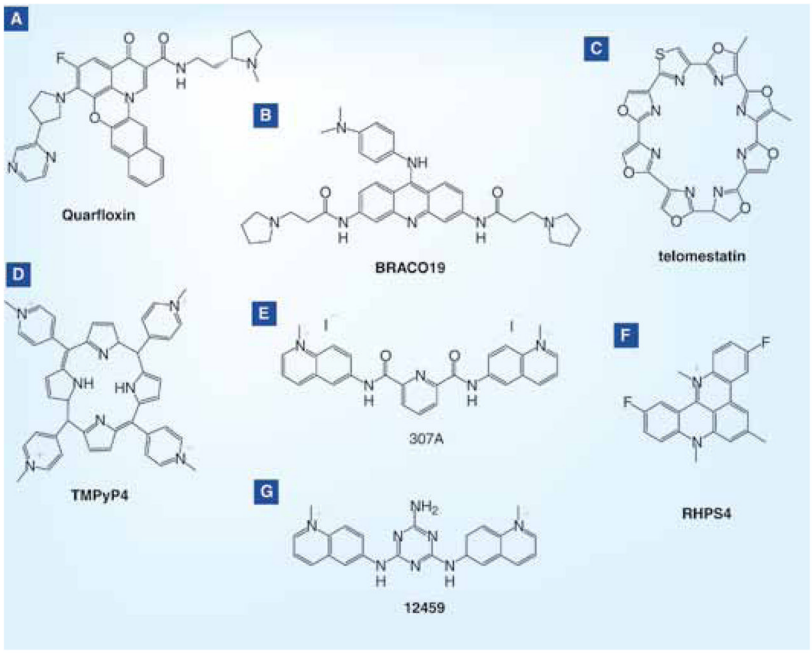

Figure 13.
G-quadruplex-interactive small-molecule compounds.