Quadruplex Nucleic Acids as Novel Therapeutic Targets

Journal:	*Journal of Medicinal Chemistry*
Manuscript ID	jm-2015-01835j.R1
Manuscript Type:	Perspective
Date Submitted by the Author:	n/a
Complete List of Authors:	Neidle, Stephen; University College London, UCL School of Pharmacy

Quadruplex Nucleic Acids as Novel Therapeutic Targets

Stephen Neidle*

UCL School of Pharmacy, University College London
29-39 Brunswick Square, London WC1N 1AX, UK

Keywords: quadruplex, telomere, promoters, cancer, small molecules, structure-based design, *in silico* screening

ABSTRACT

Quadruplex-forming sequences are widely prevalent in human and other genomes, including bacterial ones. These sequences are over-represented in eukaryotic telomeres, promoters and 5' untranslated regions. They can form quadruplex structures, which may be transient in many situations in normal cells since they can be effectively resolved by helicase action. Mutated helicases in cancer cells are unable to unwind quadruplexes, which are impediments to transcription, translation or replication, depending on their location within a particular gene. Small molecules that can stabilise quadruplex structures augment these effects and produce cell and proliferation growth inhibition. This article surveys the chemical biology of quadruplexes. It critically examines the major classes of quadruplex-binding small molecules that have been developed to date and the various approaches to discovering selective agents. The challenges of requiring (and achieving) small-molecule targeted selectivity for a particular quadruplex are discussed in relation to the potential of these small molecules as potentially clinically-useful therapeutic agents.

INTRODUCTION

The concept that certain guanine-rich nucleic acid sequences can form four-stranded structures is now over fifty years old, although the knowledge that such sequences form aggregates is much older.[1] Initially fibre-diffraction studies[2-4] were used to establish that polymeric runs of guanylic acid (and also guanosine monomers) form right-handed four-fold helices, and proposed that the strands are tightly held together by guanine-guanine Hoogsteen hydrogen-bonding to form guanine base quartets (termed G-quartets or G-tetrads: Figure 1a). the quartets stack together in the manner of base pairs in duplex nucleic acids, with the striking differences of having four strands, and sodium or potassium ions held centrally between quartets and coordinated to the O6 substituent of a guanine base.[5]

It was subsequently found that such guanine-rich sequences form the underlying repeat motif of telomeric DNA at the ends of all eukaryotic chromosomes[6-8] and also that short-length oligonucleotides containing such sequences can form discrete structures, termed quadruplexes[9] (Figure 1b). These structures can be formed[10-12] from

(i) a single strand, folded back three times (unimolecular quadruplexes), having the general sequence $G_a\ X_n\ G_b\ X_o\ G_c\ X_p\ G_d$ where G_{a-d} represent short guanine (G) tracts and X_{n-p} represent intervening "loop" regions of more general length and sequence

(ii) two strands, which are normally identical (bimolecular quadruplexes), each folded back once

(iii) three[13] or four strands (tetramolecular quadruplexes).

Quadruplexes can adopt a wide diversity of structures and topologies (outlined below), in striking contrast to the relative uniformity of duplex nucleic acids. A number of quadruplexes have been characterised by biophysical methods, especially by circular dichroism and

fluorescence spectroscopy.[14,15] All have the common feature of a core of stacked G-quartets, a central ion channel and four grooves, of varying dimensions (Figure 2a,b). The diversity arises in particular from (i) the variability of loop length and sequence, (ii) sequences in which the length of G-tracts is non-equivalent, and (iii) the influence of monovalent ions in the channel. Quadruplexes remained of little more than academic interest, until the early 1990s, when several seminal NMR[16-18] and X-ray crystallographic[19] studies confirmed the earlier supposition of the nature of the G-quartet and its role as the underlying structural motif in quadruplexes.

This article will discuss the background to, and current status of quadruplexes, emphasizing their targeting with small molecules for therapeutic ends. It will also examine some future directions for this rapidly-developing field, including structure-based approaches and their relevance to the development of new therapeutic agents. The interested reader is referred to the large number of existing reviews on various other aspects of quadruplex chemistry, biophysics and biology for further background reading.[10-12,14-15,20-24]

QUADRUPLEX PREVALENCE AND STABILITY

The determination of the sequence of the human genome (comprising ca three billion nucleotides) in 2002 led to the discovery of the wide prevalence of putative quadruplex sequence motifs,[25-29] in addition to their inherent occurrence in human telomeric DNA (see below). The initial search algorithms employed in two independent informatics studies[25,26] both used a basic search sequence of G_{3-5} X_n G_{3-5} X_o G_{3-5} X_p G_{3-5}, albeit with distinct algorithms. Both assumed that loops X_n, X_o and X_p would have 1-7 nucleotides, and that the size of the G-tracts varied between three and five guanines in length, but that they were not necessarily of equal length. The loop length limits were set on the assumption that shorter loops would be of greatest stability, in accordance with experimental and computational

studies.[30-32] Both these studies independently resulted in the finding of ca 350,000 putative occurrences, of which a significant number occur in longer stretches of multiple G-tracts ("quadruplex islands"), where the definition of an individual quadruplex may be ambiguous and thus there is likely to be multiple and overlapping quadruplex species. The existence of a stable quadruplex can be validated by the combined use of several spectroscopic and biophysical techniques, notably UV fluorescence[33] and circular dichroism,[34] optimally together with X-ray crystallographic[35] and NMR analyses.[36]

The concept that individual loops within a stable genomic quadruplex cannot contain more than seven nucleotides has been more recently challenged by experimental findings of quadruplex-forming sequences with long loops in a number of quadruplexes. Examples include:

(i) a nine-nucleotide propeller loop determined in the NMR structure of a quadruplex within the human CEB25 mini-satellite locus, which comprises almost perfect 52-nucleotide tandem repeats.[37]

(ii) A 26-nucleotide loop, stabilised as a hairpin and with C:G and G:G base pairing, has been reported to be formed within a quadruplex encoded in the promoter sequence of the *hTERT* gene.[38]

(iii) Very long central loops have been identified using a combined bioinformatics and experimental approach in the 5'-UTRs (untranslated regions) of human mRNAs, 10-90 nucleotides in length and with each flanking loop comprising just a single nucleotide.[39]

(iv) An NMR study has shown that the promoter of the human *BCL-2* gene contains a 13-nucleotide central loop,[40] of sequence d(CGCGGGAGGAAGG).

These results strongly suggests that the original loop length definitions[25,26] are too restrictive, and that the total number of putative quadruplexes is much greater than the original tally of ca 350,000. This is also in accord with a systematic experimental study of 80 different sequences,[41] which concluded that provided the flanking loops are short, then a very long central loop can be tolerated. The ability of this loop to form stable secondary interactions is clearly a major factor in determining stability. Shorter loops can also do this: the five-nucleotide lateral loop of sequence d(AGGAG) in the promoter of the *c-KIT* gene forms a highly stable secondary structure with two G:A base pairs, as observed in 2D-NMR[42] and crystal structures,[43,44] and persists in molecular dynamics simulations.[44,45] Although detailed structural data on the 13-nucleotide loop in the *BCL-2* promoter quadruplex[40] is not yet available, its sequence is strongly suggestive of it also containing a number of stabilising G:A base pairs.

The effects of longer loop length are however not straightforward and ultimately depend on the sequence of the loop(s) and possibly on the biological function and sequence context of the quadruplex. This has recently been illustrated in the case of the quadruplex sequence in the human mini-satellite 39-nucleotide CEB25 repeating motif, with a 9-nucleotide central loop flanked by two single-nucleotide T loops.[37] The overall quadruplex sequence is d(GGGTGGG**TGTAAGTGT**GGGTGGG). Although this central loop (shown in bold) appears to be much less structured than the long loops in other quadruplexes in for example the *BCL-2* or *c-KIT* genes, its presence still imparts some stability to this quadruplex. A subsequent detailed analysis[46] of the effects of varying loop length both *in vitro* and *in vivo* on the CEB25 mini-satellite in *Saccharomyces cerevisiae* cells has revealed that the length of the central loop is an important factor in determining genomic instability at such sequences. Short (≤ 4 nucleotide) central loop CEB25 quadruplexes have enhanced stability compared to ones with a longer loop and can act to block replication and enhance genomic instability.

The overwhelming majority of quadruplex studies, especially those related to small-molecule therapeutics, have focussed on the human genome. Putative quadruplex-forming sequences have been located using bioinformatics approaches in a number of organisms other than *homo sapiens* (there is frequently strong evolutionary conservation of quadruplex motifs in vertebrates,[47] for example between canine and humans in the *c-KIT* gene[48]). Examples of other organisms with quadruplex-containing genomes include *Saccharomyces cerevisiae*,[49] *Escherichia coli*[50,51] and a large number of bacterial species.[52] In broad terms the effects of a quadruplex sequence on transcription and translation depend on its position within an *E. coli* gene.[50] A number of quadruplex-forming sequences have been identified within the HIV genome, including within the long terminal repeat (LTR) promoter[53,54] and in the coding region of the *NEF* gene.[55] Therapeutic implications of targeting these HIV quadruplexes are discussed below. Potential quadruplex-forming sequences have also been located in the small (23 million nucleotide) malaria genome,[56] with 63 non-telomeric quadruplexes being identified in this genome using a search sequence containing up to 11 nucleotides in the loops. Of these, 16 are clustered upstream of *var* genes, and there is some evidence that several of these quadruplexes can be selectively targeted *in vitro*,[56] as demonstrated with a synthetic macrocyclic dibenzophenanthroline derivative.

The development of methodology to directly locate genomic quadruplex occurrences using a combination of next-generation sequencing and the polymerase stop assay has recently been applied to the human genome.[57] Results were consistent using either K^+ ions or stabilising ligands to induce quadruplex formation, and a large number of previously known quadruplex-forming regions have been identified and validated. In addition, and unexpectedly, quadruplexes were found in genes which had previously not been identified by computational search methods:[25,26] two such notable genes are the breast cancer susceptibility genes *BRCA1* and *BRCA2*. Overall, a total of 716,310 quadruplex occurrences were

identified, of which 451,646 had not been previously predicted. Many of these have non-canonical quadruplex structures (ie they do not conform to the $G_{3-5}X_nG_{3-5}X_oG_{3-5}X_pG_{3-5}$ pattern of quadruplex sequence norm) and/or long loops. Of particular relevance to potential drug targeting was the prediction of regions of high quadruplex density within oncogenes such as c-*MYC* and *BCL-2*.

Direct experimental observations of quadruplexes in cells have relied on the generation of quadruplex-specific antibodies, which were first successfully applied[58] to demonstrate the presence of quadruplexes in the transcriptionally active macronuclei of the ciliate organism *Stylonychia lemnae*. Interestingly, no evidence of antibody staining was found in the replication region of this organism; this prescient finding was suggested as indicative of quadruplexes being resolved (ie unwound) during replication. Another antibody, also with the ability to specifically pull-down quadruplex nucleic acid structures has been developed more recently[59] for use with human cells, initially with MCF-7 breast carcinoma cells, in tandem with deep sequencing, to map the occurrence of quadruplexes in the genome for this cell line. Quadruplex signals were found to be enriched in particular, in sub-telomeric and regulatory regions. Furthermore a representative set of enriched genes are sensitive to transcriptional down-regulation when treated with the quadruplex-specific ligand pyridostatin (**1**). Direct visualisation of quadruplex occurrence within (fixed) cancer cells has been demonstrated using the antibody approach coupled with fluorescent probes.[60,61] Interestingly the images show finite numbers of quadruplex foci, strongly suggesting that the number of stable quadruplexes within these cell lines is not large. One of these antibodies[60] (BG4) has been used to examine quadruplex occurrence in human tissues and elevated levels have been found in patient-derived stomach and liver cancers, contrasting with lower responses in normal tissues.[62] Possible causes of these elevated quadruplex levels are discussed further below, but

this data overall is strongly suggestive of underlying differences in actual quadruplex occurrence between normal and cancer cells.

Antibody-based quadruplex visualisation is restricted to fixed or permeable cells and tissues, and thus may potentially be prone to artefacts arising from the fixing process. Several groups are developing small-molecule quadruplex-specific probes that are useable in live cells. These show significant differences in fluorescent emission/excitation maxima and especially in fluorescent decay lifetimes when bound to quadruplex versus other types of nucleic acids. Examples include 3,6-bis(1-methyl-2-vinylpyridinium) carbazole diiodide[63] (o-BMVC: **2**), naphthoTASQ (a G-quartet mimetic tetra-substituted naphthalene derivative,[64] **3**) and a planar triarylmethyl carbocation[65] (**4**). These promising approaches will need cross-validation with quadruplex-binding antibodies before they become generally useable. At present neither the antibody nor the small-molecule visualisation methodology is able to identify an individual quadruplex within a genome, although this may be more readily feasible with a small genome such as that of HIV, which has only a small number of quadruplex targets.

TELOMERIC QUADRUPLEXES

The ability of eukaryotic telomeric sequences DNA, and in particular human telomeric DNA based on the hexanucleotide repeat sequence d(TTAGGG),[8,66] to form discrete quadruplex arrangements,[6,7,9,66,67] requires a transient or stable single-stranded stretch of DNA. Typically telomeric DNA can range in length from 3-7 kilobases, almost all of which except the extreme 100-200 nucleotides at the 3' end,[68] is double-stranded. This single-stranded overhang, free from the constraints of Watson-Crick base pairing, is thus in principle available for quadruplex formation and indeed such structures will spontaneously form in the absence of associated proteins. Telomeric duplex DNA may be folded into a higher-order

structure (the t-loop), and is associated with an array of telomere-binding proteins (TRF1, TRF2, RAP1, TIN2, and TPP1), collectively termed the shelterin complex.[69] The single-stranded overhang is also not normally free to form higher-order structures since it is associated with a number of copies of the single-strand telomere binding protein POT1 (hPOT1 in humans).[70] This protein thus effectively destabilises quadruplex formation at telomeric DNA ends.[71] The shelterin complex is involved in telomere length regulation, and POT1 is involved in telomere end-protection via feedback to other telomeric proteins within the shelterin complex, notably TIN2 and TRF2.[72]

The discovery of the telomere length maintenance enzyme telomerase[73] and its identification as a specialised reverse transcriptase, was followed in the mid-1990s by the finding of a profound link between the up-regulated expression of telomerase and oncogenesis.[74,75] This has been the starting-point for the surge of interest in telomeric quadruplexes, and ultimately in quadruplexes generally as therapeutic targets. The detailed chemistry at the telomerase catalytic sub-unit (hTERT in humans) active site involves assembly of nascent telomeric DNA onto the 3′ end telomeric DNA end. The process occurs on a complementary RNA template (hTR) and uses a pool of precursor mononucleotide triphosphates, with its substrate, the 3′ end of telomeric DNA, hybridising onto the template, and so the telomeric DNA must be single-stranded at this point. The enzymatic activity of hTERT can be inhibited by sequestration of this 3′ single-stranded end into a quadruplex arrangement[76], which is augmented by the stabilising influence of a quadruplex-binding small molecule.[77] Quadruplex formation in the telomeric DNA substrate effectively hinders this hybridisation from occurring. This approach to telomerase inhibition has been subsequently developed into an anti-cancer strategy[78-81] following proof-of-principle experiments demonstrating in general that inhibition of telomerase function in tumour cells (where its expression is up-regulated in ca 80-85% of all human cancers) leads to senescence and

apoptosis. These key experiments used antisense oligonucleotides,[82] dominant-negative mutants of the catalytic domain of telomerase[83] or hTERT catalytic-site small-molecule inhibitors.[84] The onset of replicative senescence normally requires telomeric DNA to shorten to a critical length and is accompanied by up-regulation of the RB pathway[85,86] via the cyclin-dependent kinase inhibitor p21, and increased expression of the tumour suppressor protein $p16^{INK4a}$. This leads to senescence induction, which is a precursor to selective cancer cell apoptosis and cell death. These cellular and molecular events have also been observed *in cellulo* with a number of quadruplex-binding small molecules, as well as demonstrating telomerase inhibitory activity and telomere shortening. Examples of telomeric quadruplex-binding agents (Figure 3) include quinoline-based triazine compounds (**5**),[87] BRACO-19 (**6**: an acridine derivative[88,89]), the perylene derivatives PM2 and PIPER (**7**)[90] and more recently, a ruthenium complex with chiral 4-(2,3-dihydroxypropyl)-formamide oxoaporphine.[91] The consequences of what is now termed telomere targeting with be discussed further in a subsequent section.

The unexpected findings[92-94] that telomeric DNA can be transcribed into discrete (non-coding) telomeric RNA molecules, termed TERRA, may indicate further potential targets for small-molecule quadruplex-mediated intervention at the telomere since TERRAs readily form stable RNA quadruplexes.[95-98] TERRA molecules are involved in regulating telomere length, via telomerase activity in telomerase-positive cancer cells, and also form TERRA RNA-telomeric DNA hybrids, which may be involved in recombination-mediated telomere length homeostatis in telomerase-negative ALT cell types.[99,100] TERRA molecules also play a role in the DNA damage response at telomeres.[101] It is not clear at present what, if any role, natural TERRA quadruplexes might play in these processes, so stabilising them with small molecules may or may not have therapeutic benefit in proliferating cells.

Quadruplex-mediated targeting of telomerase and telomere maintenance. The finding that a disubstituted amidoanthraquinone compound (**8**) could inhibit the catalytic action of telomerase from elongating the 3' single-stranded end of telomeric DNA, was correlated with its ability to fold the end into a quadruplex arrangement.[77] Subsequent synthesis and evaluation of the regioisomeric series of *bis*-amidoanthraquinones enabled structure-activity relationships to be established, which strengthened the evidence for a causal link between binding and activity.[102-104] The overall structural features needed for effective small molecule binding to a quadruplex and optimal telomerase activity within a series are closely similar and have been defined as (i) the possession of an extended heteroaromatic chromophore which can π-π stack onto a G-quartet surface of aquadruplex, and (ii) normally (at least) one flexible side-chain containing terminal cationic-charged moiety such as a pyrrolidine or piperidine group. The chemical space for quadruplex-binding telomerase-inhibiting compounds has been subsequently extended to a large number of chemotypes. These are mostly based on polyheteroaromatic compounds, for example the tetra-N-methylpyridyl-porphyrin compound (TMPyP4: **9**, well-studied but non-selective),[105] a series based on the pentacyclic dibenzophenanthroline core,[106] the N,N'-*bis*[2-(1-piperidino)ethyl]-3,4,9,10-perylenetetracarboxylic diimide PIPER[107] and a pentacyclic acridinium compound (RHPS4, 8,13-diethyl-6-methylquino[4,3,2-*kl*]acridinium iodide: **10**) and dervivatives.[108]

The concept of using substituted polycyclic and heteroaromatic compounds as probes (and potential anticancer leads) has been the dominant theme in the subsequent development of the majority of quadruplex-binding ligand libraries, whether targeted to telomeric or other categories of quadruplex.[20,22,109,110] Some of these libraries are derived from natural products, such as the core chromophores of the alkaloids berberine (**11**) and quindoline (**12**), but the majority are purely synthetic in origin. In addition, a number of categories of transition metal-

containing complexes have been developed to target telomeric quadruplexes,[111] some of which are also potent telomerase inhibitors (see for example.[112, 113])

Notable exceptions to the polycyclic + cationic substituted side-chain class are compounds based on, or derived from the macrocyclic natural product telomestatin (**13**), which comprises five unsubstituted oxazole rings, two methylated oxazole rings and a thiazoline ring with one asymmetric centre.[114,115] Even though (**13**) has no formal positive charge and the molecule is slightly non-planar, its extensive π-π overlap with a terminal G-quartet in quadruplexes[116] confers high affinity and selectivity, low duplex affinity, and is combined with potent telomerase inhibitory ability. However (**13**) is highly insoluble in aqueous media and thus challenging to formulate for biological studies or potentially as a therapeutic agent. A number of (**13**)-mimetic macrocyclic scaffolds have been reported[117], often comprising a number of oxazole rings interspersed with for example, pyridyl and/or phenyl rings or amide groups (for example, **14**).[118] These linked rings may form a fully cyclic system, as in (**13**) itself, or may be open, forming a three-sided or acute crescent-like shape, with an angle of ca 90° between each arm. The ubiquitous alkylamino type of cationic side-chain has been added to several of these ring systems, notably in derivative of the compound HXDV,[119] which has six linked oxazole rings and two valine units together with a dimethylaminoalkyl side-chain. A number of these compounds show high potency (<1 μM) in cell proliferation assays together with effective quadruplex stabilising properties. One compound (**15**) in this series has been evaluated in the MDA-MB-435 breast tumour xenograft model with evidence of diminished tumour growth relative to controls, although the relationship of the *in vivo* activity to quadruplex binding has yet to be elucidated.

The classic model of senescence and apoptosis induction via telomerase inhibition requires that telomere shortening occurs in a timely manner, progressing through to the limit of telomere length attrition (for example in MCF-7 human breast carcinoma cells, which have

a mean telomere length of ~6 kilonucleotides). Since ~100 nucleotides are lost per round of replication, critical senescence events should only be activated after a large number of rounds of replication.[120,121] Although several early studies on quadruplex-binding small molecules did observe progressive telomeric DNA shortening,[87,90] senescence and growth arrest occurred much more rapidly than predicted on the basis of the classic senescence model of critical telomere shortening being required. Rapid inhibition of cell growth and proliferation was also accompanied by changes at the telomere itself such as anaphase bridge formation and end-to-end chromosomal fusions. These changes are consistent with telomere end uncapping from associated proteins[88,90,122-4], including hPOT1, which binds to the single-stranded overhang[70] and is known to be displaced by quadruplex formation.[71] It has been suggested that the telomerase enzyme complex is physically associated with the extreme 3' end of the telomeric DNA overhang in cancer cells, with the inference that this association is displaced on quadruplex formation,[125] and thus telomere attrition is no longer the rate-determining event. These rapid senescence- and apoptosis-inducing effects have subsequently found to be induced by almost telomeric quadruplex-binding small molecules. Ligand-induced quadruplex stabilisation and telomere uncapping also results in DNA damage to telomeres and consequently initiates a cascade of responses with altered regulation of, in particular, p^{21}/$p16^{INK4a}$ kinases, p53, PARP and ATM/ATR pathways. The DNA damage responses are characterised by senescence[126] and the rapid appearance of phosphorylated histone γ-H2AX foci, often being apparent 24 hrs after compound administration.[127] These responses have been studied in detail with several quadruplex-binding small molecules, notably with the pentacyclic acridine derivative (**10**),[128-130] the macrocycle (**13**)[131] and the bis(quinolinyl)pyridine-dicarboxamide derivative (**1**).[127] There is also evidence that the observed antitumour activity of (**10**) in xenograft models is a direct consequence of the critically-important DNA damage response.[128] Intriguingly, a study of the effects of (**1**) on

the genome of several cancer cell lines using a chromatin immunoprecipitation sequencing approach[132] has revealed that this compound (which has high quadruplex-specificity and very low duplex DNA affinity[133]), produces DNA damage at both telomeres and a small number of non-telomeric regions of the genome, notably at the *SRC* gene locus.

The classic model of quadruplex-induced telomerase inhibition and telomere length attrition by small molecules with affinity for telomeric quadruplexes has evolved into a more direct mechanism in cancer cells, leading directly to telomere-induced senescence, growth inhibition and apoptosis. The telomere maintenance effects of these compounds involve telomere uncapping and induction of telomere damage, and telomerase inhibition, though it may well still play a role, does not appear to be the major factor. Selectivity for cancer cells may be a consequence of deficiencies of the DNA damage response apparatus in many cancers.[134] Initial events are the dissociation of hPOT1 from the single-stranded overhang and possibly the uncapping of telomerase from telomere ends to enable quadruplex formation, which then acts as a damage-response signal. Once formed, quadruplexes need to be rapidly unwound otherwise they pose an impediment to the normal functioning of a cell. The helicase RTEL1 provides an additional defence to telomere targeting by quadruplex stabilisation since it is able to effectively unwind telomeric quadruplexes and thus maintain telomere integrity.[135] In those instances where telomere dysfunction via small-molecule quadruplex binding has been unequivocally validated, it is a reasonable assumption that RTEL1 itself has also been disabled. The recent finding of a non-uniform distribution of quadruplexes within cell nuclei and preferential localisation in heterochromatin[136] has been suggested to be a consequence of the molecular crowding within heterochromatin, and also implies that epigenetic factors may play a role in selectivity.

It should not be assumed that strong ligand binding to human telomeric quadruplexes *in vitro* means that the cellular (and *in vivo*) mechanism of action involves telomere

maintenance dysfunction – other quadruplexes may be more significant targets in particular cell types. The available evidence points to telomere targeting being the critical mechanism of action for at least one well-characterised example, the acridines-based compound (**10**). However the increasing application of high-definition genomic probes, (with (**1**)[132] being a prominent example) is demonstrating an altogether more complex scenario, in view of the multiplicity of non-telomeric quadruplex sites in the human genome, as discussed in the next section.

PROMOTER QUADRUPLEXES

In parallel with the developing interest in telomeric quadruplexes in the 1990s, potential quadruplex arrangements began to be identified in a variety of genomic contexts and genes, initially in retinoblastoma susceptibility genes[137] and notably, in the nuclease hypersensitive element within the promoter of the *c-MYC* gene.[138-140] The occurrence of a quadruplex sequence in the promoter region of a particular gene (with *c-MYC* being the paradigm for many subsequent studies) has been developed into a second potential therapeutic quadruplex-targeting strategy in human cancers.[140,141] The concept was initially articulated as a simple inhibition of function following the induction of a stabilising quadruplex-small molecule complex (using the porphyrin compound (**9**)).[140] It has received much attention and a number of genes with putative promoter quadruplex sequences have been subsequently identified, initially by a bioinformatics approach.[142-144] A quadruplex-small molecule complex would be an effective impediment to RNA polymerase transcription. Subsequent validation of a particular sequence as being capable of forming a stable quadruplex has in large part come from biophysical studies. Typically this has involved the application of circular dichroism, fluorescence and NMR spectroscopy to isolated quadruplex sequences. Promoter

quadruplexes have also been identified in bacterial genomes,[145] although here as yet the concept has not been significantly exploited to date.

The majority of stable promoter quadruplexes reported are from human oncogenes and cancer-associated genes, and all appear to be intramolecular. In a number of instances the sequences are within established nuclease hypersensitive/transcriptional activation sites within or upstream of promoter regions. Well-studied examples of genes containing such quadruplex sites for which quadruplex formation *in vitro* has been demonstrated include c-*MYC*,[140] *BCL*-2,[40,146] h-*RAS*[147], k-*RAS*,[148] c-*KIT*,[149,150] *HIF*,[151,152] b-*RAF*,[153,154] the androgen receptor,[155] *RET*,[156,157] *HSP90*,[158] *MET*[159] and VEGF,[160] whereas promoter quadruplexes appear to under-represented in genes associated with normal cellular processes. Many validated and putative promoter quadruplexes tend to be within 1kb, more commonly immediately upstream and close to the transcription start site in these genes. Perhaps unsurprisingly they can also include or are overlapping with the G-rich SP1 transcription factor binding sites[161] d(GGGCGG), which would serve to augment their ability to repress transcription. Some of these genes, which are over-expressed in particular cancers, have been targeted with small molecules at the protein level and in several instances, compounds are in the clinic (for b-RAF, c-KIT) or in clinical trial (for example for MET, BCL-2). However several of the proteins encoded by genes in this list have long been considered to be undruggable – c-MYC and the RAS proteins are prominent examples. Quadruplex targeting at the gene level, at least in principle, offers the possibility of circumventing such roadblocks.

In striking contrast to the relative simplicity of human telomeric quadruplexes, where the quadruplex unit is four d(TTAGGG) repeats, many promoter quadruplexes have complex sequences. There high diversity in overall sequence, as would be expected, which translates into diversity in the nature of the G-tract size and in the size and sequence of the loops. Some quadruplexes unequivocally comprise just four G-tract repeats, notably the two quadruplexes

in the c-*KIT* gene,[149,150] which are 22- and 21-mers respectively. These are between -87 and -109 bp and between -140 and -160 bp upstream of the transcription initiation site. The latter is adjacent to the SP1 transcription factor binding site. By contrast, the well-studied 27-mer NHE III_1 element of the c-*MYC* promoter contains five short G-tracts, and that both four-tract quadruplexes (1234- and 2345-G-tract) can be formed,[162,163] and indeed a quadruplex can be formed with all five tracts.[164] The major quadruplex-forming sequence in the *BCL*-2 promoter is a 39-mer with six G-tracts, so that in principle up to 15 distinct four-tract intramolecular quadruplexes are possible. Unsurprisingly, quadruplexes formed from such clusters can be conformationally complex, although when several loops are short (each with ≤ two nucleotides), then individual quadruplexes tend to have parallel folds. This is the case even when one loop is long, as with a *BCL*-2 quadruplex.[40] Cell-based experiments have shown that the promoter of the *SRC* proto-oncogene is a major binding site for the well-studied quadruplex-selective compound (**1**).[132] This site has six G-tracts, suggesting that these enhanced quadruplex clustering are possible hot-spots for quadruplex-binding small molecules, although in this study not all such sites were targeted, as judged by DNA damage responses.

The promoter of the h*TERT* gene (the catalytic domain of the telomerase complex) is mutated in a number of human cancers,[165] notably in melanomas and gliomas, leading to enhanced expression of telomerase and maintenance of the malignant phenotype in these cancers. The mutations occur in a highly distinctive putative quadruplex-forming region[38] with 12 G-tracts. The topology of a major quadruplex formed in this region is currently controversial, with one model based on foot-printing analysis having a 26-nucleotide loop, contrasting with NMR studies on one segment of the sequence having a parallel and a (3+1) hybrid quadruplex in equilibrium.[166] Biophysical and simulation studies[167] on the complete

hTERT G-tract region suggest the possibility of a complex arrangement involving three stacked parallel quadruplexes, but with some cross-interaction between them.

Promoter quadruplex targeting with small molecules. The early demonstration of c-*MYC* down-regulation by a small molecule has prompted much effort to target this and a number of other quadruplexes, with a view to eventual therapeutic utility. A few examples are given here. Substituent groups are usually flexible acyclic chains terminating in cationic groups. A persistent challenge is that even if strong binding to a particular promoter quadruplex is found *in vitro* and effects consistent with down-regulation of the target gene in cells are observed, this does not necessarily constitute robust validation of the cellular effects being direct on-target ones. However evidence of down-regulation at both mRNA and protein levels is usually taken to be highly suggestive of on-target effects, and such evidence has been documented in a number of instances

BCL-2 expression has been shown to be down-regulated, for example, by members of a small library of mono-substituted quindoline (**12**) derivatives,[168] using a luciferase reporter assay to select compounds with optimal promoter quadruplex stabilising ability. *BCL*-2 expression was down-regulated at the mRNA level and reduced levels of BCL-2 protein were also observed following cell treatment with the most effective *BCL*-2 quadruplex binder in the series. *BCL*-2 expression was also shown to be down-regulated in MIA-PACA2 pancreatic cancer cells[169] and tumour xenografts[170] treated with the tetra-substituted naphthalene diimide compound MM41 (**16**). This compound has exceptional affinity for several quadruplexes, especially human telomeric and *BCL*-2 promoter ones. Selectivity at the cellular level was indicated on the basis of *BCL*-2 mRNA levels being selectively down-regulated in several pathway-focussed PCR gene arrays (for oncogenes, DNA damage response and for telomere maintenance-related genes). However *BCL*-2 down-regulation was not the sole significant gene change observed, suggesting that with this compound in this

particular cell line there are likely to be several major quadruplex targets. Thus protein levels of both BCL-2 and k-RAS were reduced in treated tumours compared to controls – k-*RAS* mutation and dysregulation is a key driver of most pancreatic cancers and cell lines. Comparisons of changes in gene expression in short-term cell culture conditions with those in long-term treated tumours need to be made with caution since in the latter such effects may be masked by or mistaken for global necrotic and apoptotic effects. A study of gene expression changes induced by a trisubstituted naphthalene diimide derivative (**17**), also using a focussed gene array approach, has similarly observed *BCL*-2 as well as *hTERT* mRNA down-regulation in several cancer cell lines including a melanoma one.[171] In accord with other observations,[169,170] no reduction in c-*MYC* expression was observed, which is notable in view of the ubiquitous role of c-*MYC* in oncogenesis and the ability of this and many other such small molecules, to bind with high affinity to c-*MYC* quadruplexes. The array data derived from this trisubstituted naphthalene diimide indicates that in some cell lines at least the pattern of modulation of gene expression correlates with putative quadruplex potential, although the small number of genes surveyed probably preclude a more definitive conclusion.

The two quadruplexes found in the promoter of the c-*KIT* oncogene can been stabilised by a variety of polycyclic-based small molecules, including substituted indenoisoquinolines, tri-substituted isoalloxazines and mono-substituted benzo[a]phenoxazines.[172-174] A target validation study has been reported[174] employing a 173-member compound library including mono-substituted benzo[a]phenoxazines (**18**), using a luciferase reporter assay approach, established in the human HGC-27 gastric cancer cell line with and without c-*KIT* promoter quadruplex formation. This enabled the assignment of quadruplex-related effects to be made with confidence. The two hit compounds from this assay showed dose-dependent reduction of c-*KIT* expression in HGC-27 cells.

The reality of small molecule-driven promoter quadruplex transcriptional down-regulation is likely to be considerably more complex than the straightforward picture outlined above and may well also involve small-molecule interactions with the opposite i-motif strand in a quadruplex-containing duplex sequence, as well as proteins such as nucleolin that would stabilise a quadruplex complex once formed. Discussion of these topics is beyond the scope of this review and the reader is referred to, for example, recent mechanistic studies on small molecule binding to the *BCL*-2 and c-*MYC* promoter loci.[175-177]

A number of studies have reported high affinity of a small molecule (often a quindoline derivative) to a particular promoter quadruplex, together with observations of down-regulation of the gene, sometimes at both mRNA and protein level. The dangers of assuming a direct cause and effect relationship between *in vitro* and cellular observations have been well illustrated by a structure-activity study on the effects of 11-piperazinylquindoline derivatives, such as (2-(4-(10H-indolo[3,2-b]quinolin-11-yl)piperazin-1-yl)-N,N-dimethylethanamine: **19**) on c-*MYC* expression.[178] Several compounds in this series are effective stabilisers of a c-*MYC* quadruplex. The lead compound inhibits cell growth and produces c-*MYC* down-regulation in cells; several other derivatives also have effects on other quadruplex-containing genes, as well as their quadruplexes. An exon-specific assay,[179] which is an elegant alternative to using a pair of isogenic cell lines, exploits particular translocation features of the CA46 Burkitt's lymphoma cell line and was used to demonstrate[178] that c-*MYC* quadruplex targeting is not directly involved in c-*MYC* down-regulation by this compound. However indirect effects on "quadruplex target" genes are not by any means universal and appear to depend on the nature of the small molecule involved, and on the cell type. For example, and in striking contrast to the study outlined above,[178] an ellipticine derivative with a single dimethylaminoethoxy side-chain (GQC-05: NSC338258), was found by a combined screening and molecular modelling procedure.[179] This compound directly

down-regulate c-*MYC* expression, using the same exon-specific assay.[179] Use of a reporter assay with quadruplex-containing and quadruplex-negative constructs in the promoter can also provide evidence of promoter quadruplex involvement,[180] although the approach by itself is less definitive than a genetic one. It has been used, for example, to study effects produced by disubstituted indolo[3,2-c]quinolines on k-*RAS* expression in colon cancer cells and their relationship to binding to the k-*RAS* promoter quadruplex,[181] as well as with compounds targeting *BCL*-2 and c-*KIT* quadruplexes.[168,174]

The fluoroquinolone-based quadruplex-binding compound quarfloxin (**21**) (CX-3543: 5-fluoro-N-(2-((S)-1-methylpyrrolidin-2-yl)ethyl)-3-oxo-6-((R)-3-(pyrazin-2-yl) pyrrolidin-1-yl)-3H-benzo[b]pyrido[3,2,1-kl] phenoxazine-2-carboxamide)[182,183] is an example of evolution in target considerations. It was originally developed as a c-*MYC* quadruplex stabiliser. Subsequent studies have indicated that it acts on ribosomal DNA (rDNA), which has a large number of putative quadruplex sequences, and then inhibits rDNA-nucleolin interactions, which in turn inhibits RNA polymerase I transcription.[183] This compound associates with nucleolin in cancer cell nuclei and causes rRNA synthesis inhibition. It produced significant reductions in tumour volume in the MDA-MB-231 breast cancer and MIA-PACA-2 pancreatic cancer xenograft models. (**21**) is distinguished by being the first-in-class, (and to date, only) quadruplex-binding compound to have entered clinical trials for human cancer. Phase 1 trials indicated good patient tolerance and a phase 2 trial was in carcinoid/neuroendocrine tumours, a tumour type for which the phase 1 trial indicated some responses. Compound (**21**) has more recently been licenced to Tetragene Inc (www.tetragene.com) in order for clinical development to continue. It will be interesting to see the application of next-generation sequencing technologies to determine the actual targets in the genome for this drug.

RNA QUADRUPLEXES

By contrast with promoter quadruplexes, the concept of quadruplex formation in mRNA sequences has the obvious attraction unwinding of duplex nucleic acid is not required.[184] Thus a given RNA G-tract sequence can more readily fold into a quadruplex form than a DNA duplex one. In addition, *in vitro* studies have indicated that RNA quadruplexes are more thermodynamically stable than their DNA counterparts[185-187] and are less liable to conformational and topological heterogeneity, probably as a consequence of the additional hydrogen-bonding possibilities provided by the 2' sugar hydroxyl group in RNAs. The finding that quadruplex sequences are over-represented in the 5'-untranslated regions (5'-UTR) of many human genes[188,189] has catalysed interest in the category of sequences have been identified in a number of genes, for example, in *BCL*-2,[190] n-RAS,[191] the human estrogen receptor alpha,[192] the MT3 matrix metalloproteinase,[193] transforming growth factor β2[194] and several oncogenes targeted by the cap-binding helicase eIF4A,[195] as well as more generally in introns.[196] There is good evidence that 5'-UTR quadruplexes are involved in post-translational gene regulation,[197,198] although the exact role of a particular quadruplex depends on its locus within the 5'-UTR. The quadruplex antibody approach has been used to visualise the occurrence of RNA quadruplexes in human cells,[199] which was enhanced when a RNA quadruplex-specific small molecule (a carboxy derivative of compound (**1**)) was used.[134]

Small-molecule targeting of 5'-UTR quadruplexes has been shown to inhibit translation, for example, with the *N-RAS*[200] and TRF2[198] genes, with observations of decreased levels of translation efficiency using *in vitro* reporter assays, using respectively pyridine-2,6-bis-quinolino-dicarboxamide and bis-quinolinium derivatives. It has been shown[201] that RNA 5'-UTR quadruplexes can be targeted in cells, in a study with the *K-RAS* gene using the photo-

activated porphyrin compound tri-meso(N-methyl-4-pyridyl), meso(N-tetradecyl-4-pyridyl) porphine. This compound has cytoplasmic rather than nuclear localisation. It down-regulates k-RAS protein levels in pancreatic cancer PANC-1 cells (which over-express this protein), by up to 80% in a dose-dependent manner, and also produces cell-growth arrest. These changes were found to be correlated with changes in mRNA levels.

There is much current interest in RNA quadruplexes formed from expansion of simple quadruplex motifs such as the hexanucleotide repeat r(G_4C_2). Such repeat expansions are found in a number of neurodegenerative diseases.[202] The r(G_4C_2) motif occurs in the non-coding region of the *C9orf72* gene and is associated with the majority of cases of the neurodegenerative diseases amyotropic lateral sclerosis (ALS) and frontotemporal dementia (FDT).[203] The repeats can form RNA quadruplexes[204,205], which have been directly related to disease progression,[204] although the exact mechanisms relating to *C9orf72* quadruplexes are not known at present. It is not clear for example, whether stabilising the quadruplexes will result in beneficial changes in their association with RNA-binding proteins, or what effects this will have on the level of further truncated *C9orf72* transcripts. A start has been made on developing small-molecule targeting of the *C9orf72* quadruplex; initial experiments[206] with the non-selective porphyrin compound (**9**) show that this ligand does bind to the *C9orf72* RNA quadruplex, which then inhibits binding of two established RNA binding proteins, hnRNPA1 and ASF/SF2 to the quadruplex. The *C9orf72* repeat expansion RNA can also in principle form a hairpin arrangement in equilibrium with a quadruplex, with the latter favoured in K^+ but not Na^+-containing solution. It is not known what the effect on this would be in the crowded environment of the cytoplasm. A focussed library of 132 RNA-binding small molecules has been used to screen the *C9orf72* RNA repeat *in vitro*.[207] The three hit compounds were found to bind to the repeats in cells using a pull-down technique and had an effect not only on repeat translation but also on numbers of RNA foci in repeat-expressing

neurons. Two of the compounds have a curved shape resembling classic DNA minor groove binding compounds; the third, most biologically active compound is an ellipticine derivative with a single piperidine (cationic) side-chain (**22**). Remarkably, very few ellipticine derivatives have been evaluated for quadruplex binding and cellular activity, a notable exception being GQC-05 (NSC338258: **20**), also with a single, though different side-chain, which down-regulates c-*MYC* expression.[180]

QUADRUPLEXES IN REPLICATION

Quadruplexes can occur at many loci[208-211] throughout the human and other genomes in addition to their well-documented presence in promoters, untranslated regions and telomeres. They are present in immunoglobulin switch regions and in breakpoint regions in many cancer genomes.[212] They may well play a more general role in gene regulation[213], consist with their well-documented non-random occurrence throughout the genome. Replication presents opportunities for DNA quadruplexes to be formed since DNA becomes unwound and transiently single-stranded. In principle their occurrence presents a stall to replication, but in practice such impediments are effectively resolved by, in particular helicases.[214,215] However when helicases involved in replication and repair, such as Pif1[216,217], FANCJ[218,219] and the Bloom's helicase are mutated,[60,220,] as is often the case in cancer cells, then replication is stalled. Quadruplex replication arrest (and the appearance of increased numbers of quadruplex foci in cells), is augmented by quadruplex-binding small molecules such as the polyoxazole compound (**13**),[60] the resulting quadruplex complexes are more resistant to helicase unwinding than the native quadruplexes.[221] The helicases XPB and XPD, which are associated with promoter activity and transcriptional regulation, also unwind quadruplexes.[215] The RTEL1 helicase unwinds and resolves quadruplexes at telomeres;[222] it is mutated in a

number of human cancers[223] although it is not known as yet whether these mutations result in impaired quadruplex unwinding.

Helicase malfunction then provides a further basis for the concept of quadruplex-targeted selectivity in many cancers, and a working hypothesis to explain the observations of selective cancer cell growth inhibition induced by quadruplex-binding small molecules *in vitro* and *in vivo*. Few systematic compound library screens have been reported to date based on the concept of their ability to stabilise quadruplexes against helicase unwinding, but these would be a useful addition to existing approaches, possibly using the recently-described rapid fluorescence-based helicase assay[224], which has been validated with a number of well-established quadruplex-binding compounds.

Structure-based approaches to finding new ligands may also be fruitful. The DEAH (Asp-Glu-Ala-His) box RNA helicase binds to and unwinds quadruplexes with parallel topology. The 2D-NMR solution structure has been determined[225] of the quadruplex-binding DEAH peptide bound to a parallel quadruplex. This shows the α-helical core of the peptide sitting on a G-quartet end-face of the quadruplex (Figure 3) in a manner reminiscent of the chromophores of typical quadruplex-binding ligands. Several basic peptide side-chains are associated with the quadruplex phosphate groups. Thus the structure reinforces the view that ligand planarity is not an essential requirement for quadruplex binding and suggests ways in which novel peptomimetic ligands could be engineered to interfere with helicase unwinding and bind with high specificity to a particular quadruplex.

QUADRUPLEX STRUCTURES

Crystallographic and 2D-NMR structural studies have determined the detailed molecular structures of DNA and RNA quadruplexes from a variety of sources (reviewed in [35,36,226, 227]). Intramolecular and intermolecular quadruplexes formed from human-derived sequences have

been studied in most detail, with the latter formed from two (bimolecular) strands. Quadruplexes with four (tetramolecular) or three separate strands are also possible. All show the common features of at least two G-quartets, stacked together with a right-handed twist and connected by loop sequences. Loops can be of several distinct types; diagonal, lateral (or edgewise), and propeller (or chain-reversal). The spaces between the four strands are termed grooves, analogous to those in duplex DNA and RNA, although by contrast with the two groove sin duplex structures, quadruplexes have four grooves. Groove dimensions are sensitive to the size and nature of the loops and to topology – the pattern of glycosidic angles plays an important role in defining overall groove width and depth. The differences in groove dimensions are import factors in enabling small molecules to differentiate between different quadruplexes. A given loop type is associated with a given set of strand directions for the two strands that it connects. These various possibilities give rise to a large number of possible folds (topologies), of which only a small number have been observed to date. The guanines forming each quartet arise from the short G-tracts; however when the tracts are of unequal length or loops themselves contain guanines, then more complex arrangements may be possible and can even dominate the overall topology.

Telomeric quadruplexes. There has been considerable emphasis on structural studies of quadruplexes formed by human telomeric DNA sequences,[228,229] assembled from the repeat d(TTAGGG), in particular on intramolecular quadruplexes containing four such repeats.[230-235] All comprise a core of stacked G-quartets assembled from the G-tracts in successive repeats, held together by d(TTA loops). The structures themselves display a variety of folds.[228,230-235] 2D-NMR studies in dilute solution in K^+ ion environments have revealed several types of fold with anti-parallel strands, including (3+1) hybrid topologies (Figure 4) in which three of the four backbone strands are in one direction (ie parallel) and the fourth is in the opposite antiparallel direction.[230-232] The precise nature of the flanking sequences appears to play a

role in stabilising one particular type over another. A single fold dominates in Na^+ solution, with two lateral and one diagonal loop, so that two backbone orientations are in one direction and the other two are anti-parallel.[228] By contrast, the crystal structures[235] of a human telomeric intramolecular and a bimolecular quadruplex containing K^+ ions shows all strands parallel in both structures and with all the loops necessarily being of the propeller type in order to achieve this strand orientation (Figure 1). There is continuing controversy as to which of these folds best represents the human intramolecular quadruplex in cellular environments; under crowding and high concentration conditions the parallel form is predominant[236,237], consistent with it also being present in cells, whereas in more dilute K^+-containing solution the (3+1) forms may predominate.

Crystal structures of small-molecule human telomeric quadruplex complexes representing several diverse chemotypes are available[238-244], with structural data available in the PDB. Structures for a number of naphthalene diimide intramolecular complexes[239-241] have been determined (Figure 5), as well as with berberine,[243] two mesoporphyrin complexes[244] and the crystal structure of a bimolecular complexes with the trisubstituted acridine compound (**6**).[238] All of these structures have parallel-stranded quadruplexes, with the ligand chromophore bound on an external G-quartet face and side-chain substituents located in quadruplex grooves. In all instances the core structure, of three stacked G-quartets, is constant and essentially unchanged. Also to a first approximation the quadruplexes in these complexes are closely similar overall to the native structures. However there are significant variations in conformation of the propeller loops and groove dimensions, with evidence of ligand-related effects.[245] Analysis of the structural data on several naphthalene diimide derivative complexes shows that loop conformations are most constrained within a series of closely-related compounds, giving confidence to conservative *in silico* design studies based on such crystal structures. These crystal structures have also highlighted the role of groove-bound

water molecules in generating a network of water-mediated hydrogen bonds to the cationic side-chain substituents of the naphthalene diimides.

Only one NMR-determined structure in this category has been reported to date[246], involving an analogue of the cyclic polyoxazole compound (**13**) with six oxazole rings together with two alkylamino side-chains (**23**), which impart aqueous solubility to the otherwise insoluble macrocycle core. A human telomeric DNA sequence was used, which forms a (3+1) hybrid fold in the native state.[230,232] This topology is retained in the complex (Figure 6), with the macrocycle stacked on one end of the quadruplex. The macrocycle has a significant out-of-plane twist, which is reflected in the non-planarity of the G-quartet on which it is stacked, indicating that G-quartets can accommodate chemotypes with substantial non-planarity without compromising overall quadruplex stability. Overall stabilisation of the six oxazole rings arises from a number of close contacts to backbone and two sideways-on loop bases, which contribute to the preference for this particular topology. The two flexible side-chains are in close hydrogen-bond contact with backbone phosphate atoms. The overall structure provides some clear pointers to ways in which affinity and selectivity can be enhanced by rational modification of, in particular, the side-chains.

Small-molecule binding tends to favour a particular quadruplex topology, dependent on the precise nature of the ligand, which can be reliably assessed, at least for telomeric quadruplexes, by CD spectroscopy.[247] Thus compound (**13**) and the pentacyclic acridines derivative (**10**), which has minimal substituents, both have a preference for anti-parallel type quadruplexes whereas the porphyrin (**9**) and a number of porphyrin derivatives as well as tetrasubstituted naphthalene diimide compounds, prefer the parallel fold. This behaviour has been rationalised on the basis of the available surface area at the chromophore binding site, the 3' or 5' terminal G-quartet.[248]

Human telomeric RNA sequences can also readily form quadruplexes. However unlike their DNA counterparts, these are altogether more conformationally rigid and fold into only a single topology in K^+ solution. This is the all-parallel-stranded form, observed both in the crystal[249] and in dilute solution by 2D-NMR;[250] both structures are closely similar to the crystal structures of the DNA telomeric quadruplex. The crystal structure has highlighted the role of the O2' ribose hydroxyl group in stabilising r(UUA) propeller loop conformation via hydrogen bonding. Only one structure of an RNA telomeric quadruplex ligand complex is currently available, involving a disubstituted acridine chromophore having two amidoalkylamino side-chains.[251] The r(UUA) propeller loops have altered conformations compared to the native structure, with the O2' ribose hydroxyl groups hydrogen bonding to loop adenine bases so that these bases swing round to become in-plane with the terminal G-quartet of the quadruplex. The resulting extended r(AGGGGA) surface is able to accommodate two side-by-side bound ligand molecules.

Promoter and other quadruplexes. 2D-NMR studies in particular have validated the existence of stable quadruplexes for a number of promoter quadruplex sequences. These are notable for their sequence diversity, and include structures for c-*MYC*,[252,253] *BCL*-2,[254,255] *VEGF*[256], h*TERT*[166] and *RET*[257] quadruplexes, all of which have fully-assigned spectra and structures deposited in the PDB. Remarkably, the majority of the major species in solution tend to have parallel topology, which is a consequence of most sequences have at least one short (n<3 nucleotide) propeller loop, whose stereochemistry requires strands to be parallel; the *VEGF* 22-mer sequence d($CG_4CG_3CCG_5CG_4$) is typical (Figure 7), with two single-nucleotide dC loops and a longer tetranucleotide loop, which plays a role in overall stabilisation of this quadruplex. This parallel propensity is also shown, for example, by the *RET* oncogene quadruplex, with one single-nucleotide dC loop and two d(GCG) propeller loops. Examination of the prevalence of single nucleotides in the quadruplex surveys of the

human genome[25,26] indicate that they probably comprise the majority of loop types in quadruplexes.

However the existence of propeller loops and of parallel strands does not necessarily lead to a simple all-parallel topology, if for example guanine bases occur within loop regions and also actively participate in quartet formation. This can result in an altogether more complex fold. An example is one of the two sequences in the c-*KIT* promoter, which has been extensively characterised by 2D-NMR[42] and by X-ray crystallography.[43,44] A non-G-tract guanosine situated between two cytosine loops, is folded back and forms part of the G-quartet core. This overall fold has not been observed in other quadruplexes, which is unsurprising given that the primary 22-mer sequence has a unique occurrence in the human genome.[256] The c-*KIT* fold (Figure 8) is conserved between solution and crystal, providing further support for the notion that the topology of quadruplexes with parallel strands is likely to be conformationally stable.

No crystal structures of small molecule complexes with non-telomeric DNA quadruplexes have been reported to date. Three 2D-NMR structures of c-*MYC* promoter quadruplexes have been determined, with the porphyrin compound (**9**)[259], the (bisquinolinium) phenanthroline compound[260] Phen-DC_3 (**24**) and a crescent-shaped mono-substituted quindoline compound with an aminoalkylamino side-chain.[261] The chromophores in all three structures are stacked over a terminal G-quartet of the quadruplex core. The complex with compound (**9**) has the substituent N-methylpyridyl groups positioned close to the edges of the grooves but not actually bound in them. Unsurprisingly this compound has modest quadruplex selectivity, being able to also bind effectively to duplex nucleic acids. The quindoline substituents of compound (**24**) have extensive overlap with the guanines of the terminal G-quartet (Figure 9) whereas the overall crescent-like shape of the molecule ensures that it would stack only poorly in a duplex intercalation site (Figure 10), in accord with its high quadruplex

selectivity. The NMR structure rationalises earlier data, on for example, analogues with pyridinium rings replacing the quinolones, which have reduced quadruplex-stabilising ability. It also suggests sites of functionalisation to enhance affinity and selectivity. The mono-substituted quindoline 2: 1 complex[261] is unusual in that a ligand molecule is bound to each terminal G-quartet face, with additional stabilisation from 5' and 3' flanking nucleotides. The short protonated aminoalkylamino side-chain attached to the quindoline skeleton is barely able to reach the nearest groove (Figure 11), suggesting ways in which to improve affinity, as well as selectivity for this particular quadruplex.

There are remarkably few detailed molecular structures available for RNA quadruplexes and none to date of 5'-UTR ones. Those determined include several in more complex flanking sequence environments than those in the structures of simple isolated DNA quadruplexes. An RNA quadruplex is the target of the Fragile X Mental Retardation Protein (FMRP) and a co-crystal structure[262] (Figure 12) shows recognition of this quadruplex by the arginine-glycine-rich (RGG) motif of the protein. This quadruplex is at the end of a helical RNA stem, where the RGG β-turn is bound. There is an extensive pattern amino-acid contacts to the RNA, stabilising the interface and its mixed-base quartet. As with the DEAH-box helicase, one sees a protein motif bound at one end of the quadruplex, analogous to small-molecule binding sites, and suggesting where small-molecule inhibitors might best act. Two co-crystal structures[263,264] of small-molecule fluorophores bound to a novel RNA quadruplex aptamer have revealed an unprecedented quadruplex fold with, in both structures, just two G-quartets. This parallel-stranded quadruplex is formed at the junction between two RNA duplex stems, with the strands folding back and each donating guanines to the quartets. The sequences involved do not obviously conform to the quadruplex rules and there are no G-tracts of more than two nucleotides in length. The two fluorophores are closely-related - in one structure[264] it is 3,5-difluoro-4-hydroxybenzylidene imidazolinone (**25**) - both are bound in a stacking

environment between one G-quartet and the adjacent base triplet (Figure 13), which are both part of an extended co-planar run of duplex and quadruplex base pairs, triplets and quartets. There are a number of contacts involving 2'-ribose hydroxyl atoms and fluorophore. Altogether these structures show that quadruplexes, perhaps RNA ones in particular, can form highly specific small-molecule binding environments, especially when part of more extensive nucleic acid structures.

QUADRUPLEX SELECTIVITY

Can small molecules have (i) selectivity for quadruplexes over duplex and other nucleic acids forms, and (ii) selectivity for a particular quadruplex? The first question is straightforward to answer, in the affirmative. There are many examples in the literature of individual compounds, and of broad chemotypes that have low or minimal affinity for duplex nucleic acids as well as high quadruplex affinity (this is typically assessed on a small number of quadruplex types). Heteroaromatic polycyclic compounds such as substituted acridines, quindolines, ellipticines and naphthalene-diimides, tend to show promiscuous binding to various types of nucleic acid structure; selectivity for the larger surface area of a terminal G-quartet increases with ligand size and especially with the number of side-chain substituents. This however is at the cost of reduced drug-like character - higher molecular weight and increased cationic charge, even though several of these compounds are showing anticancer effects *in vivo*. Many macrocyclic and crescent-shape molecules on the other hand, have built-in low duplex-binding ability by virtue of their inability to effectively bind into a duplex intercalation site (see Figure 10). These crescent-shaped molecules bear some resemblance to classic DNA duplex groove binders but crucially they do not have the shallow curvature characteristic of this class, with the three sides of the crescent being much more acutely angled in order to achieve selectivity. Examples include quindoline compounds[261] and a

number of oligo-oxazole-based compounds and several low-molecular weight meta-substituted diphenyl furans.[266]

Selectivity for a particular quadruplex has generally been approached empirically rather than by exploiting the available structural data. Even so, screening approaches have resulted in several remarkable findings. For example the crescent-shaped heptacyclic compound TOxaPy, comprising four oxazole and three pyridyl rings (**26**) is highly selective[267] for the Na^+ form of the human telomeric quadruplex and binds only very weakly to any of the K^+ forms,[229-235] indicating a marked topological preference for the antiparallel Na^+ form.[228] Compound (**26**) also shows some affinity for several promoter quadruplexes. Selectivity for human telomeric quadruplex DNA can in principle be enhanced by exploiting the potential of the 3' single-stranded telomeric DNA overhang to form tandem quadruplexes, analogous to 'beads on a string.'[268] Such selectivity has been found[269] with a hybrid oxazole-triazole compound, which can have a crescent-shaped conformation. The goal of designing small molecules selective for particular promoter quadruplexes is altogether more challenging and no studies have been reported to date utilising the available structural data. Screening methods continue to be employed, with some successes. For example, several members of a series of crescent-shaped bis(benzimidazole)-phenanthroline compounds, with additional flexible alkylamino side-chains have been found to be selective for the c-*MYC* and c-*KIT* promoter quadruplexes over a human telomeric one and duplex DNA.[270] These findings have been rationalised by molecular dynamics simulations, which emphasise the role of the side-chains in propeller loop recognition. An innovative application of small-molecule microarrays, screening 20,000 compounds, has been used to find a selective c-*MYC* promoter quadruplex inhibitor.[271] The lead compound from this approach, a disubstituted benzofuran derivative (**27**), does not have the characteristics of a conventional quadruplex binder and produces only a modest increase in c-*MYC* quadruplex melting temperature (ΔT_m), of 2.1° C.

On the other hand, measurement of binding affinity by surface plasmon resonance (SPR) gave a low μM value for K_d, sufficient to elicit a biological response. Significantly, SPR measurements on five other promoter quadruplexes did not record binding, although weak binding was found with two others (from the *BCL-2* and *RB1* genes). In accord with this selectivity, an exon-specific assay[179] using the Burkitt's lymphoma cell line CA46, demonstrated that effects were ascribable to binding to the c-*MYC* promoter quadruplex locus. Studies on various panels of cancer cell lines has confirmed and extended these finding of high selectivity, so that overall this represents the first detailed proof of concept study in this field. The lead compound (**27**), while not being an ideal drug candidate, is an appropriate starting-point for further optimisation.

LIGAND DESIGN

The overwhelming majority of current quadruplex-binding small molecule scaffolds[80, 109,265,272] have been selected on the basis of the simple premise of possessing[102-104] (i) a heteropolyaromatic planar chromophore, which binds via π-π stacking to a terminal G-quartet, and (ii) one or more flexible substituents with a cationic charge, to bind in quadruplex grooves and to loops. Notable exceptions to these broad guidelines are (i) cationic porphyrin and related scaffolds, of which the best-studied is the tetra-N-methyl-pyridyl porphyrin, compound (**9**), and (ii) a number of ligand families, all characterised by having macrocyclic or crescent-shaped scaffolds. These scaffolds are mostly based on or derived from the telomestatin and related oxazole-linked macrocyclic chemotypes. Appropriate crescent shapes can be achieved using pyridyl or phenanthroline rings disubstituted with heteroaromatic groups such as quinolines. Members of both of these broad classes of compounds bind to the terminal G-quartets in telomeric and other quadruplexes, with sufficient affinity from π-π aromatic G-quartet stacking so to not require the involvement of

additional cationic groove-binding substituents. These quadruplex binders tend to have greater quadruplex selectivity over duplex DNA or RNA than those based on polyheteroaromatic chemotypes.

New quadruplex-binding chemotypes have mostly been found as outcomes of screening chemical libraries against a measurement of quadruplex affinity or stabilisation. We are now seeing the increasing use of much larger compound libraries, which is starting to move the field away from the established concepts of what constitutes a good quadruplex-binding ligand, and promises to ensure the development in time of altogether more drug-like chemotypes. Also *in silico* and structure-based approaches are being used with increasingly broader ranges of scaffolds. The extensive structural knowledge base for the various topological folds of the human telomeric monomeric quadruplex has prompted a number of *in silico* studies[273-277] using, for the most part, well-established docking software with the crystallographic or NMR-derived structures. As well as the well-documented pitfalls of the approach in general, telomeric quadruplexes present several particular challenges as *in silico* targets:

(i) the flexibility of the loops in these structures has rarely been taken into account,[278] so "hits" from screening of small-molecule libraries can only be taken as indicative of plausible high-affinity hits and are generally not reliably predictive of ranking order. A survey[245] of known crystal structures of telomeric quadruplex complexes with small molecules has found that a number of distinct loop types are represented in this data set, although one type does predominate. There appears to be a dependence of loop type on the nature of the small molecule, although at present it is not possible to predict this. A combined spectroscopic and molecular dynamics simulation study has shown that natural mutations in loop sequences may affect loop conformations and even overall quadruplex topology.[279] This may be of especial relevance to the design of ligands specific to such mutated quadruplexes as further

sequence data on mutations in promoter regions becomes available, following the findings of hTERT promoter mutations in melanoma and other cancers.[165]

(ii) In general *in silico* studies have taken a particular telomeric quadruplex topology as a starting-point. It cannot be assumed that the chosen topology is preferentially stabilised by the resulting hits, especially since some ligands are known to alter topology, for example from anti-parallel to parallel.

(iii) *in silico* methods work best with well-defined binding pockets for example in nucleotide binding sites in kinases, which are not present in most structurally simple quadruplexes. No studies to date have been reported on *in silico* screening of more complex quadruplexes, where fragment-based approaches could also be used.

In spite of these caveats, it is apparent that the approach can successfully generate novel hits, and has the advantage of targeting quadruplex grooves as well as the more conventional terminal G-quartets. This has resulted in the discovery of a number of non-planar quadruplex-binding compounds, some of which have the potential to be developed into drug-like leads.[274-277] Docking studies have used a range of large *in silico* libraries, with the ZINC and ChemBridge databases employed, for example, to find hits against the c-*MYC* quadruplexes.[280,281] Fragment-based design has not yet been used to full develop a novel lead compound, although a study using a fragment library derived from RNA-targeting has identified the potential of this approach.[282]

Data from X-ray crystallography[35] and 2D-NMR studies[36] on molecular structures of small-molecule quadruplex complexes, notably of telomeric quadruplexes has provided the starting-points for some of these *in silico* studies. The perennial question of the relevance of quasi-solid-state crystal structures to solution or even cellular conditions, is of especial significance for quadruplexes, since some do have folds that are environment-sensitive.

These crystal structures are densely packed and crowded solvent-containing environments akin to the cellular conditions surrounding nuclear DNA – it is notable that quadruplexes have been recently reported to be localised in the highly organised DNA of heterochromatin.[135] It is therefore unsurprising that in the case of human telomeric quadruplexes - the most contentious category – studies of topology in crowded solution have mostly concurred with the crystal data. Water molecules can also play a more detailed, active and intimate role in ligand binding. The crystal structures, for example of several naphthalenediimide-telomeric quadruplex complexes,[240,241] have revealed the consistent role of bridging water molecules in the grooves via hydrogen bonds between cationic groups at the termini of ligand side-chains, and phosphate groups lining the walls of the grooves.

Although those crystal structures have not been used directly for the design of new leads, they have been useful starting-points for structure-based optimisation of existing compounds.[239] Recent 2D-NMR structures of quadruplex complexes also provide opportunities for ligand refinement and improvement.[283] An important future application of this structural information, especially for polycyclic compounds, will be to enhance their drug-like features.

Limitations of current quadruplex models. There are several potential challenges with the current approaches to measuring quadruplex-small molecule binding, which is normally assessed *in vitro* by one or more biophysical methods, using an isolated quadruplex as target. In particular there has been relatively little attention paid to the relationship of the behaviour and ligand binding of an isolated quadruplex, to its behaviour in its (DNA or RNA plus associated proteins) biological or cellular context. The higher-order telomeric quadruplexes arrangements possible in the single-stranded 3' overhang at the extreme end of eukaryotic telomeres are obvious and accessible ones to focus on, as discussed earlier in this review. The junctions of telomeric quadruplexes with duplex DNA may also be plausible ligand targets,

with ligand binding onto a d(TAT) base triplet platform formed at the junction.[284] Analogous junctions may be formed around promoter DNA quadruplexes and have been observed in two extended RNA quadruplex-small molecule complexes.[263,264] The sequences of the junctions and flanking regions may affect quadruplex topology. mRNA quadruplexes are also in principle susceptible to flanking and bystander effects, although the additional ribose 2'-OH groups may reduce flexibility and hence hinder topological change in RNA quadruplexes – this reduced flexibility has been noted in telomeric RNA quadruplexes which do not have the variability observed in dilute solution with their DNA counterparts.

Quadruplex selectivity: a real or apparent problem? This review has highlighted those studies where individual small molecule leads and starting-points have been exemplified into chemical libraries, and from these structure-activity relationships have been evaluated, for example correlating quadruplex stabilisation measured by thermal shift (ΔT_m) with telomerase inhibitory activity or changes in expression of a target gene. In reality, selectivity for individual quadruplexes has until recently been almost entirely the result of serendipitous discovery, following screening with either an individual pre-chosen quadruplex, or from a small library of (normally) promoter quadruplexes. Once a particular quadruplex has been chosen, affinity can be optimised on the basis of structure-activity relationships. Remarkably few studies have surveyed a complete genome, for example by whole-genome expression profiling[285], to verify that the chosen quadruplex is in reality a significant biological target. We have discussed in an earlier section the advisability of being able to distinguish between direct effects arising from quadruplex-small molecule binding and consequential changes in the target gene's expression, from effects that are in reality just correlated together and are not direct cause-and-effect. This latter circumstance can occur when other upstream components of a pathway, or of an intersecting pathway, are the direct targets and these may or may not be quadruplex-related.

The task of specifically targeting an individual quadruplex in say the human genome appears at first sight to be an impossibly daunting one, unless that quadruplex has features so distinctive for structure-based design or library screening. This may be more realistic with very small genomes such as the HIV one, where all the potential quadruplexes can be identified. At present this approach is not obviously feasible for the human genome. Bioinformatics surveys of both DNA and RNA human quadruplexes have shown that the majority probably have at least one short loop, suggestive of a "simple" parallel topology. Many genomic quadruplexes are therefore likely to have the same general features, ie a core of stacked quartets with the loops forming four grooves. Structural and molecular dynamics data has shown that the groove dimensions and loop conformations are not fixed, but have significant flexibility, which will change on ligand binding. Even so, it is possible to categorise a large number of these simple quadruplexes into a few structural families. More complexes quadruplexes are formed, for example when the loops themselves contain guanine residues that can actively participate in core quartet formation via insertion, when a loop is of sufficient length to form its own stable secondary structure, or when the quadruplex is an RNA one. One of the c-*KIT* quadruplexes illustrates this, with both an insertion loop and a secondary structure loop, which results in a topology that has not been observed to date in any other quadruplex. Thus such instances may provide individual quadruplex starting-points for highly selective ligand design. It might be thought then that attempting to target a simple quadruplex with more generic structural features, is unlikely to be successful. This is belied by the recent success with a microarray compound library approach to targeting the c-*MYC* quadruplex.[271] Even though the lead compound was only evaluated for selectivity with a small number of other quadruplexes (and therefore its genomic quadruplex selectivity is unknown), the fact that it was able to down-regulate c-MYC expression in a cancer cell with

some potency, demonstrates that the cell is in effect able to sort out selectivity, provided that the target plays a major role in maintaining the cell's malignant phenotype.

Although targeting a single defined quadruplex within a genome may not be readily feasible at present, it is important to achieve high selectivity over binding to duplex DNA and RNA. Competition studies that examine (for example thermal stability (ΔT_m)) effects of a ligand on a quadruplex in the presence of large excess of duplex DNA are valuable in that they have relevance to the nucleic acid biological context. Duplex nucleic acid affinity leads to non-specific cell toxicity; the fact that many quadruplex-binding small molecules have features that at first sight resemble, in particular, classic intercalative compounds, with implications of such off-target toxicity, should not deter further studies on them. It is notable that most of the quadruplex-binding compounds with reported anti-cancer activity in xenograft models, are based on classic polycyclic heteroaromatic cores, notably compounds (**6**),[89] (**10**)[128,286,], (**16**),[168] (**21**)[181] and EMICORON (N,N-bis[2-(1-piperidino)-ethyl]-1-(1-piperidinyl)-6-[2-(1-piperidino)-ethyl]-benzo[ghi] perylene-3,4:9,10-tetracarboxylic diimide.[287] This suggests that even if their genomic quadruplex selectivity is modest, it is sufficient to down-regulate key oncogenic targets.

Is it therefore necessary in practice to be confident that a single gene is being targeted at the promoter or UTR level by a small molecule? The experience of cancer therapeutics over the past 25 years suggests that in reality this does not matter as long as one can be confident of among the major quadruplex targets is a driver gene or genes that play a key role in the maintenance and progression of a particular cancer cell type. The oncogene addiction hypothesis[288,289] states that some cancers rely on the dominant action of an individual oncogene for growth and viability, and that these are appropriate targets. In principle, then targeting aberrant telomere maintenance, promoter or 5'-UTR quadruplexes may be a fruitful way of demonstrating the importance of this hypothesis, and arriving at novel drugs for the

treatment of human cancers. The quadruplex approach holds out the promise of circumventing the challenges of targeting "undruggable" targets such as c-MYC and RAS at the protein level. The next few years will determine whether this promise becomes a reality as chemical-based approaches are increasingly combined with new biological tools in order to interfere with quadruplex function.

AUTHOR INFORMATION

Corresponding author

Email s.neidle@ucl.ac.uk

Phone +44-2077535969

Notes

The author declares no competing financial interests.

Biography

Stephen Neidle is an Imperial College London graduate (BSc chemistry, PhD, DSc). He started his independent academic career at the Department of Biophysics at King' College London, becoming one of the first Cancer Research Campaign Career Development Fellows, and spent some while at the Fox Chase Cancer Center, Philadelphia working in collaboration with Helen Berman. He moved in 1985 to the Institute of Cancer Research UK as a Cancer Research Campaign Life Fellow and Professor of Biophysics. In 2002 he was appointed to the Chair of Chemical Biology at the School of Pharmacy, University of London (now part of University College London). Stephen Neidle's major research interests are in drug design and discovery involving nucleic acid-small molecule interactions applied to cancer, anti-infective and neuro-degenerative diseases.

ACKNOWLEDGEMENTS

I am grateful for support of recent quadruplex studies in my laboratory from Johnson and Johnson (UK) and the Pancreatic Cancer Research Fund (UK), as well as earlier support from Cancer Research UK, the Medical Research Council (UK) and the Wellcome Trust.

ABBREVIATIONS USED

hTERT, human Telomerase Reverse Transcriptase; BCL-2, B-cell Lymphoma 2; HIV, Human Immunodeficiency Virus; NEF, Negative Regulatory Factor; BRCA1, 2, Breast Cancer tumour suppressor genes; MYC, avian Myelocytomatosis viral oncogene homolog; TRF1, 2, Telomeric Repeat Factor 1, 2; hPOT1, human Protection of Telomeres; RB, RetinoBlastoma; TERRA, Telomeric Repeat-containing RNA; ALT, Alternative Lengthening of Telomeres; PARP, Poly (ADP-Ribose) Polymerase; ATM, Ataxia Telangiectasia Mutated; ATR, Ataxia Telangiectasia and Rad3-related; SRC: Rous Sarcoma; RTEL1, Regulator of Telomere Elongation Helicase 1; NHE: Nuclease Hyper-responsive Element; h, k, n-RAS, Harvey, Kirsten sarcoma virus, neuroblastoma Rat Sarcoma ;FANCJ, Fanconi Anemia group J protein; XPB, D: Xeroderma Pigmentosum types B, D; VEGF, Vascular Endothelial Growth Factor; PDB, Protein Data Bank; RET, Rearranged during Transfection

REFERENCES

(1) Bang, I. Untersuchungen über die guanylsäre. *Biochem. Zeitschrift* **1910**, *26*, 293-311.

(2) Gellert, M.; Lipsett, M. N.; Davies, D. R. Helix formation by guanylic acid. *Proc. Natl. Acad. Sci. USA* **1962**, *48*, 2013-2018.

(3) Arnott, S.; Chandrasekaran, R.; Marttila, C. M. Structures for polyinosinic acid and polyguanylic acid. *Biochem. J.* **1974**, *141*, 537-543.

(4) Zimmerman, S. B.; Cohen, G. H.; Davies, D. R. X-ray fiber diffraction and model-building study of polyguanylic acid and polyinosinic acid. *J. Mol. Biol*. **1975**, *92*, 181-192.

(5) Howard, F. B.; Miles, H. T. Poly(inosinic acid) helices: essential chelation of alkali metal ions in the axial channel. *Biochemistry* **1982**, *21*, 6736-6745.

(6) Sundquist, W. I.; Klug, A. Telomeric DNA dimerizes by formation of guanine tetrads between hairpin loops. *Nature* **1989**, *342*, 825-829.

(7) Sen, D.; Gilbert, W. A sodium-potassium switch in the formation of four-stranded G4-DNA. *Nature* **1990**, *344*, 410-414.

(8) Meyne, J.; Ratliff, R. L.; Moyzis, R. K. Conservation of the human telomere sequence $(TTAGGG)_n$ among vertebrates. *Proc. Natl. Acad. Sci. USA* **1989**, *86,* 7049-7053.

(9) Henderson, E.; Hardin, C. C.; Walk, S. K.; Tinoco, I. Jr.; Blackburn, E. H. Telomeric DNA oligonucleotides form novel intramolecular structures containing guanine-guanine base pairs. *Biochemistry* **1987**, *51*, 899-908.

(10) Davis J. T. G-quartets 40 years later: from 5′-GMP to molecular biology and supramolecular chemistry. *Angew. Chem. Int. Ed. Engl.* **2004**, *43*, 668-698.

(11) Burge, S.; Parkinson, G. N.; Hazel, P.; Todd, A. K.; Neidle, S. Quadruplex DNA: sequence, topology and structure. *Nucleic Acids Res*. **2006**, *34*, 5402-5415.

(12) Patel, D. J.; Phan, A. T.; Kuryavyi, V. Human telomere, oncogenic promoter and 5'-UTR G-quadruplexes: diverse higher order DNA and RNA targets for cancer therapeutics. *Nucleic Acids Res.* **2007,** *35*, 7429-7455.

(13) Zhou, J.; Bourdoncle, A.; Rosu, F.; Gabelica, V.; Mergny, J.-L. Tri-G-quadruplex: controlled assembly of a G-quadruplex structure from three G-rich strands. *Angew. Chem. Int. Ed. Engl.* **2012**, *51*, 11002-11005.

(14) Neidle, S.; Balasubramanian, S. (eds.) *Quadruplex Nucleic Acids*. **2006**, RSC Publishing, UK.

(15) Chaires, J. B.; Graves, D. E (eds.) *Quadruplex Nucleic Acids*. **2013**, Springer-Verlag, Berlin.

(16) Smith, F. W.; Feigon, J. Quadruplex structure of *Oxytricha* telomeric DNA oligonucleotides. *Nature* **1992**, *356*, 164-168.

(17) Aboul-ela, F.; Murchie, A. I.; Lilley, D. M. NMR study of parallel-stranded tetraplex formation by the hexadeoxynucleotide d(TG_4T). *Nature* **1992**, *360*, 280-282.

(18) Wang, Y.; Patel D. J. Solution structure of a parallel-stranded G-quadruplex DNA. *J. Mol. Biol.* **1993**, *234***,** 1171-1183.

(19) Laughlin, G.; Murchie, A. I. H.; Norman, D. G.; Moore, M. H.; Moody, P. C. E.; Lilley, D. M. J.; Luisi, B. The high-resolution crystal structure of a parallel-stranded guanine tetraplex. *Science* **1994**, *265*, 520-524.

(20) Neidle, S. *Therapeutic Applications of Quadruplex Nucleic Acids*, **2012**, Academic Press, San Diego.

(21) Bidzinska, J.; Cimino-Reale, G.; Zaffaroni, N.; Folini, M. G-quadruplex structures in the human genome as novel therapeutic targets. *Molecules* **2013**, *18*, 12368-91235.

(22) Müller, S.; Rodriguez, R. G-quadruplex interacting small molecules and drugs: from bench towards bedside. *Expert. Rev. Clin. Pharmacol*. **2014**, *7*, 663-679.

(23) Cree, S. L.; Kennedy, M. A. Relevance of G-quadruplex structures to pharmacogenetics. *Frontiers Pharmacol*. **2014**, *5*, 160.

(24) Parrotta, L.; Ortuso, F.; Moraca, F.; Rocca, R.; Costa, G.; Alcaro, S.; Artese, A. Targeting unimolecular G-quadruplex nucleic acids: a new paradigm for the drug discovery? *Expert Opin. Drug Discov*. **2014**, *9*, 1167-1187.

(25) Todd, A. K.; Johnston, M.; Neidle, S. Highly prevalent putative quadruplex sequence motifs in human DNA. *Nucleic Acids Res.* **2005**, *33*, 2901-2907.

(26) Huppert, J. L., Balasubramanian, S. Prevalence of quadruplexes in the human genome. *Nucleic Acids Res.* **2005**, *33*, 2908–2916.

(27) Yadav, V. K.; Abraham, J. K.; Mani, P.; Kulshrestha, R.; Chowdhury. S. QuadBase: genome-wide database of G4 DNA - occurrence and conservation in human, chimpanzee, mouse and rat promoters and 146 microbes. *Nucleic Acids Res*. **2008**, *36*, D381-385.

(28) Ryvkin, P.; Hershman, S. G. Wang, L. S.; Johnson, F. B. Computational approaches to the detection and analysis of sequences with intramolecular G-quadruplex forming potential. *Methods Mol. Biol*. **2010**, *608*, 39-50.

(29) Cao, K.; Ryvkin, P.; Johnson, F. B. Computational detection and analysis of sequences with duplex-derived interstrand G-quadruplex forming potential. *Methods* **2012**, *57*, 3-10.

(30) Risitano, A.; Fox, K. R. Influence of loop size on the stability of intramolecular DNA quadruplexes. *Nucleic Acids Res*. **2004**, *32*, 2598-2606.

(31) Hazel, P.; Huppert, J. H.; Balasubramanian, S.; Neidle, S. Loop-length dependent folding of G-quadruplexes. *J. Amer. Chem. Soc*. **2004**, *125*, 16405-16415.

(32) Bugaut, A.; Balasubramanian, S. A sequence-independent study of the influence of short loop lengths on the stability and topology of intramolecular DNA G-quadruplexes. *Biochemistry* **2008**, *47*, 689-697.

(33) Mergny, J.-L.; Phan, A. T.; Lacroix, L. Following G-quartet formation by UV-spectroscopy. *FEBS Lett*. **1998**, *435*, 74-78.

(34) Vorlíčková, M.; Kejnovská, I.; Sagi, J.; Renčiuk, D.; Bednářová, K,.;Motlová, J.; Kypr, J. Circular dichroism and guanine quadruplexes. *Methods* **2012**, *57*, 64-75.

(35) Neidle, S. The structures of quadruplex nucleic acids and their drug complexes. *Curr. Opin. Struct. Biol.* **2009***, 19*, 239-250.

(36) Adrian, M.; Heddi, B.; Phan, A. T. NMR spectroscopy of G-quadruplexes. *Methods* **2012**, *57*, 11-24.

(37) Amrane, S.; Adrian, M.; Heddi, B.; Serero, A.; Nicolas, A.; Mergny, J.-L.; Phan, A. T. Formation of pearl-necklace monomorphic G-quadruplexes in the human CEB25 minisatellite. *J. Amer. Chem. Soc*. **2012**, *134*, 5807-5816.

(38) Palumbo, S. L.; Ebbinghaus, S. W.; Hurley, L. H. Formation of a unique end-to-end stacked pair of G-quadruplexes in the hTERT core promoter with implications for inhibition of telomerase by G-quadruplex-interactive ligands. *J. Amer. Chem. Soc*. **2012**, *131*, 10678-10691.

(39) Jodoin, R.; Bauer, L.; Garant, J.-M.; Mahdi Laaref, A.; Phaneuf, F.; Perreault, J.-P. The folding of 5'-UTR human G-quadruplexes possessing a long central loop. *RNA* **2014**, *20*, 1129-1141.

(40) Agrawal, P.; Lin, C.; Mathad, R. I., Carver, M.; Yang, D. Z. The major G-quadruplex formed in the human BCL-2 proximal promoter adopts a parallel structure with a 13-nt loop in K+ solution. *J. Amer. Chem. Soc*. **2014**, *136*, 1750-1753.

(41) Guédin, A.; Gros, J.; Alberti, P.; Mergny, J.-L. How long is too long? Effects of loop size on G-quadruplex stability. *Nucleic Acids Res*. **2010**, *38*, 7858-7868.

(42) Phan, A. T.; Kuryavyi, V.; Burge, S.; Neidle, S.; Patel, D. J. Structure of an unprecedented G-quadruplex scaffold in the human *c-kit* promoter. *J. Amer. Chem. Soc*. **2007**, *129*, 4386-4392.

(43) Wei, D.; Parkinson, G. N.; Reszka, A. P.; Neidle, S. Crystal structure of a c-kit promoter quadruplex reveals the structural role of metal ions and water molecules in maintaining loop conformation. *Nucleic Acids Res*. **2012**, *40*, 4691-4700.

(44) Wei, D., Husby, J.; Neidle, S. Flexibility and structural conservation in a c-KIT G-quadruplex. *Nucleic Acids Res*. **2015**, *43*, 629-644.

(45) Islam, B.; Stadlbauer, P.; Krepl, M.; Koca, J.; Neidle, S.; Haider, S.; Sponer, J. Extended molecular dynamics of a c-kit promoter quadruplex. *Nucleic Acids Res*. **2015**, doi: 10.1093/nar/gkv785.

(46) Piazza, A.; Adrian, M.; Samazan, F.; Heddi, B.; Hamon, F.; Serero, A.; Lopes, J.; Teulade-Fichou, M.-P.; Phan, A. T.; Nicolas, A. Short loop length and high thermal stability determine genomic instability induced by G-quadruplex-forming minisatellites. *EMBO J*. **2015**, *34*, 1718-1734.

(47) Frees, S.; Menendez, C.; Crum, M.; Bagga, P. S. QGRS-Conserve: a computational method for discovering evolutionarily conserved G-quadruplex motifs. *Human Genomics* **2014**, *8*, 8.

(48) Da Rosn S.; Zorzan, E.; Giantin, M.; Zorro Shahidian, L.; Palumbo, M.; Dacasto, M.; Sissi, C. Sequencing and G-quadruplex folding of the canine proto-oncogene KIT promoter region: might dog be used as a model for human disease? *PLoS One* **2014**, *9*, e103876.

(49) Capra, J. A.; Paeschke, K.; Singh, M.; Zakian, V. A. G-quadruplex DNA sequences are evolutionarily conserved and associated with distinct genomic features in *Saccharomyces cerevisiae*. *PLoS Comput. Biol*. **2010**, *6*, e1000861.

(50) Rawal, P.; Kummarasetti, V. B.; Ravindran, J.; Kumar, N.; Halder, K.; Sharma, R.; Mukerji, M.; Das, S. K.; Chowdhury, S. Genome-wide prediction of G4 DNA as regulatory motifs: role in *Escherichia coli* global regulation. *Genome Res*. **2006**, *16*, 644-655.

(51) Holder, I. T.; Hartig, J. S. A matter of location: influence of G-quadruplexes on *Escherichia coli* gene expression. *Chem. Biol*. **2014**, *21*, 1511-1521.

(52) Du, X.; Wojtowicz, D.; Bowers, A. A.; Levens, D.; Benham, C. J.; Przytycka, T. M. The genome-wide distribution of non-B DNA motifs is shaped by operon structure and suggests the transcriptional importance of non-B DNA structures in *Escherichia coli*. *Nucleic Acids Res*. **2014**, *41*, 5965-5977.

(53) Perrone, R.; Nadai, M.; Frasson, I.; Poe, J. A.; Butovskaya, E.; Smithgall, T. E.; Palumbo, M.; Palù, G.; Richter, S. N. A dynamic G-quadruplex region regulates the HIV-1 long terminal repeat promoter. *J. Med. Chem*. **2013**, *56*, 6521-6530.

(54) Amrane, S.; Kerkour, A.; Bedrat, A.; Vialet, B.; Andreola, M.-L.; Mergny, J.-L. Topology of a DNA G-quadruplex structure formed in the HIV-1 promoter: a potential target for anti-HIV drug development. *J. Amer. Chem. Soc*. **2014**, *136*, 5249-5252.

(55) Perrone, R.; Nadai, M.; Poe, J. A.; Frasson, I.; Palumbo, M.; Palù, G.; Smithgall, T. E.; Richter, S. N. Formation of a unique cluster of G-quadruplex structures in the HIV-1 Nef coding region: implications for antiviral activity. *PLoS One* **2013**, *8*, e73121.

(56) Smargiasso, N.; Gabelica, V.; Damblon, C.; Rosu, F.; De Pauw, E.; Teulade-Fichou, M.-P.; Rowe, J. A.; Claessens, A. Putative DNA G-quadruplex formation within the promoters of *Plasmodium falciparum var* genes. *BMC Genomics* **2009**, *10*, 362.

(57) Chambers, V. S.; Marsico, G.; Boutell, J. M.; Di Antonio, M.; Smith, G. P.; Balasubramanian, S. High-throughput sequencing of DNA G-quadruplex structures in the human genome. *Nature Biotechnol*. **2015**, *33*. 877-881.

(58) Schaffitzel, C.; Berger, I.; Postberg, J.; Hanes, J.; Lipps, H. J.; Pluckthun, A. *In vitro* generated antibodies specific for telomeric guanine-quadruplex DNA react with *Stylonychia lemnae* macronuclei. *Proc. Natl. Acad. Sci. USA* **2001**, *98*, 8572-8577.

(59) Lam E.Y.; Beraldi D.; Tannahill D.; Balasubramanian S. G-quadruplex structures are stable and detectable in human genomic DNA. *Nature Commun*. **2013**, *4*, 1796-1780.

(60) Biffi, G.; Tannahill, D.; McCafferty, J.; Balasubramanian, S. Quantitative visualization of DNA G-quadruplex structures in human cells. *Nature Chem*. **2013**, *5*, 182-186.

(61) Henderson, A.; Wu, Y.; Huang, Y. C.; Chavez, E. A.; Platt, J.; Johnson, F. B.; Brosh, R. M. Jr.; Sen D.; Lansdorp, P. M. Detection of G-quadruplex DNA in mammalian cells. *Nucleic Acids Res*. **2013**, *42*, 860-869.

(62) Biffi, G.; Tannahill, D.; Miller, J.; Howat, W. J.; Balasubramanian, S. Elevated levels of G-quadruplex formation in human stomach and liver cancer tissues. *PLoS One* **2014**, *9*, e102711.

(63) Tseng, T.-Y.; Chien, C.-H.; Chu, J.-F.; Huang, W.-C.; Lin, M.-Y.; Chang, C.-C.; Chang. T.-C. Fluorescent probe for visualizing guanine-quadruplex DNA by fluorescence lifetime imaging microscopy. *J. Biomed. Opt*. **2013**, *18*, 101309.

(64) Laguerre, A.; Hukezalie, K.; Winckler, P.; Katranji, F.; Chanteloup, G.; Pirrotta, M.; Perrier-Cornet, J.-M.; Wong, J. M. Y.; Monchaud, D. Visualization of RNA-quadruplexes in live cells. *J. Amer. Chem. Soc*. **2015**, *137*, 8521-8525.

(65) Shivalingam, A.; Angeles Izquierdo, M.; Le Marois, A.; Vysˇniauskas, A.; Suhling, K.; Kuimova, M. K.; Vilar, R. The interactions between a small molecule and G-quadruplexes are visualized by fluorescence lifetime imaging microscopy. *Nature Commun*. **2015**, *6*, 8178.

(66) Bochman, M. L.; Paeschke, K.; Zakian, V. A. DNA secondary structures: stability and function of G-quadruplex structures. *Nature Rev. Genetics* **2012**, *13*, 770-780.

(67) McElligott, R.; Wellinger, R. J. The terminal DNA structure of mammalian chromosomes. *EMBO J.* **1997**.*16*, 3705-3714.

(68) Wright, W. E.; Tesmer, V. M.; Huffman, K. E.; Levene, S. D.; Shay, J. W. Normal human chromosomes have long G-rich telomeric overhangs at one end. *Genes Dev*. **1997**, *11*, 2801-2809.

(69) De Lange, T. How shelterin solves the telomere end-protection problem. *Cold Spring Harb. Symp. Quant. Biol*. **2010**, *75*, 167-177.

(70) Baumann, P.; Cech, T. R. Pot1, the putative telomere end-binding protein in fission yeast and humans. *Science* **2001**, *292*, 1171-1175.

(71) Zaug, A. J.; Podell, E. R.; Cech, T. R. Human POT1 disrupts telomeric G-quadruplexes allowing telomerase extension in vitro. *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 10864-10869.

(72) Frescas, D.; de Lange, T. TRF2-tethered TIN2 can mediate telomere protection by TPP1/POT1. *Mol. Cell Biol*. **2014**, *34*, 1349-1362.

(73) Greider, C.W.; Blackburn, E. H. Identification of a specific telomere terminal transferase activity in *Tetrahymena* extracts. *Cell* **1985**, *43*, 405-413.

(74) Kim, N. W.; Piatyszek, M. A.; Prowse, K. R.; Harley, C. B.; West, M. D.; Ho, P. L.; Corviello, G. M.; Wright, W. E.; Weinrich, S. L.; Shay, J. W. Specific association of human telomerase activity with immortal cells and cancer. *Science* **1994**, *266*, 2011-2015.

(75) Counter, C. M.; Hirte, H. W.; Bacchetti, S., Harley, C. B. Telomerase activity in human ovarian carcinoma. *Proc. Natl. Acad. Sci. U.S.A* **1994**, *91*, 2900-2904.

(76) Zahler, A. M.; Williamson, J. R.; Cech, T. R.; Prescott, D. M. Inhibition of telomerase by G-quartet DNA structures. *Nature* **1991**, *350*, 718-750.

(77) Sun, D.; Thompson, B.; Cathers, B. E.; Salazar, M.; Kerwin, S. M.; Trent, J. O.; Jenkins, T. C.; Neidle, S. Hurley, L. H. Inhibition of human telomerase by a G-quadruplex-interactive compound. *J. Med. Chem*. **1997**, *40*, 2113-2116.

(78) Oganesian, L.; Bryan, T. M. Physiological relevance of telomeric G-quadruplex formation: a potential drug target. *BioEssays* **2007**, *29*, 155-165.

(79) De Cian, A.; Lacroix, L.; Douarre, C.; Temime-Smaali, N.; Trentesaux, C.; Riou, J.-F.; Mergny, J.-L. Targeting telomeres and telomerase. *Biochimie* **2008**, *90*, 131-155.

(80) Sissi, C.; Palumbo, M. Telomeric G-quadruplex architecture and interactions with potential drugs. *Curr. Pharm. Design* **2014**, *20*, 6489-6509.

(81) Sekaran, V.; Soares, J.; Jarstfer, M. B. Telomere maintenance as a target for drug discovery. *J. Med. Chem.* **2014,** *57*, 521-538.

(82) Chen, Z.; Monia, B. P.; Corey, D. R. Telomerase inhibition, telomere shortening, and decreased cell proliferation by cell permeable 2'-O-methoxyethyl oligonucleotides. *J. Med. Chem*. **2002**, *45*, 5423-5425.

(83) Hahn, W. C.; Stewart, S. A.; Brooks, M. W.; York, S. G.; Eaton, E.; Kurachi, A.; Beijersbergen, R. L.; Knoll, J. H. M.; Meyerson, M.; Weinberg, R. A. Inhibition of telomerase limits the growth of human cancer cells. *Nature Medicine* **1999**, *10*, 1164-1170.

(84) Damm, K.; Hemmann, U.; Garin-Chesa, P.; Hauel, N.; Kauffmann, I.; Priepke, H.; Niestroj, C.; Daiber, C.; Enenkel, B.; Guilliard, B.; Lauritsch, I.; Muller, E.; Pascolo, E.; Sauter, G.; Pantic, M.; Martens, U. M.; Wenz, C.; Lingner, J.; Kraut, N.; Rettig, W. J. and Schnapp, A. A highly selective telomerase inhibitor limiting human cancer cell proliferation. *EMBO J.* **2001**, *20*, 6958-6968.

(85) Jacobs, J J.; de Lange, T. p16INK4a as a second effector of the telomere damage pathway. *Cell Cycle* **2005**, *4*, 1364-1368.

(86) Salama, R.; Sadaie, M.; Hoare, M.; Narita, M. Cellular senescence and its effector programs. *Genes Dev*. **2014**, *28*, 99-114.

(87) Riou, J.-F.; Guittat, L.; Mailliet, P.; Laoui, A.; Renou, E.; Petitgenet, O.; Mégnin-Chanet, F.; Hélène, C.; Mergny, J.-L.. Cell senescence and telomere shortening induced

by a new series of specific G-quadruplex DNA ligands. *Proc. Natl. Acad. Sci. USA* **2002**. *99*, 2672-2677.

(88) Incles, C. M.; Schultes, C. M.; Kempski, H.; Koehler, H.; Kelland, L. R.; Neidle, S. A G-quadruplex telomere targeting agent produces p16-associated senescence and chromosomal fusions in human prostate cancer cells. *Mol. Cancer Ther*. **2004**, *3*, 1201-1206.

(89) Taka, T.; Huang, L.; Wongnoppavich, A.; Tam-Chang, S. W.; Lee, T. R.; Tuntiwechapikul, W. Telomere shortening and cell senescence induced by perylene derivatives in A549 human lung cancer cells. *Biorg. Med. Chem*. **2013**, *21*, 883-890.

(90) Burger, A. M.; Dai, F.; Schultes, C. M.; Reszka, A. P.; Moore, M. J.; Double, J. A.; Neidle, S. The G-quadruplex-interactive molecule BRACO-19 inhibits tumor growth, consistent with telomere targeting and interference with telomerase function. *Cancer Res*. **2005**, *65*, 1489-1496.

(91) Chen, Z.-F.; Qin, Q.-P.; Qin, J.-L.; Zhou, J.; Li, Y.-L.; Li, N.; Liu, Y.-C.; Liang, H. Water-soluble ruthenium(II) complexes with chiral 4-(2,3-dihydroxypropyl)-formamide oxoaporphine (FOA): in vitro and in vivo anticancer activity by stabilization of G-quadruplex DNA, inhibition of telomerase activity, and induction of tumor cell apoptosis. *J. Med. Chem*. **2015**, *58*, 4771-4789.

(92) Schoeftner, S.; Blasco, M. A. Developmentally regulated transcription of mammalian telomeres by DNA-dependent RNA polymerase II. *Nature Cell Biol*. **2008**, *10*, 228-236.

(93) Azzalin, C. M.; Reichenbach, P.; Khoriauli, L.; Giulotto, E.; Lingner, J. Telomeric repeat containing RNA and RNA surveillance factors at mammalian chromosome ends. *Science* **2007**, *318*, 798-801.

(94) López de Silanes, I.; d'Alcontres, M. S.; Blasco, M. A. TERRA transcripts are bound by a complex array of RNA-binding proteins. *Nature Commun*. **2010**, *1*, 1-9.

(95) Randall, A.; Griffith, J. D. Structure of long telomeric RNA transcripts: the G-rich RNA forms a compact repeating structure containing G-quartets. *J. Biol. Chem*. **2009**, *284*, 13980-13986.

(96) Martadinata, H.; Phan, A. T. Structure of propeller-type parallel-stranded RNA G-quadruplexes, formed by human telomeric RNA sequences in K^+ solution. *J. Amer. Chem. Soc*. **2009**, *131*, 2570-2578.

(97) Collie, G. W.; Haider, S. M.; Neidle, S.; Parkinson, G. N. A crystallographic and modelling study of a human telomeric RNA (TERRA) quadruplex. *Nucleic Acids Res.* **2010**, *38*, 5569-5580.

(98) Martadinata, H.; Phan, A. T. Structure of human telomeric RNA (TERRA): stacking of two G-quadruplex blocks in K^+ solution. *Biochemistry* **2013**, *52*, 2176-2183.

(99) Cusanelli, E.; Chartrand, P. Telomeric repeat-containing RNA TERRA: a noncoding RNA connecting telomere biology to genome integrity.*Front. Genet*. **2015**, *6*, 143.

(100) Arora, R.; Azzalin, C. M. Telomere elongation chooses TERRA ALTernatives. *RNA Biol*. **2015**, *12*, 938-941.

(101) Porro A, Feuerhahn S, Delafontaine J, Riethman H, Rougemont J, Lingner J. Functional characterization of the TERRA transcriptome at damaged telomeres. *Nature Commun*. **2014**, *5*, 5379.

(102) Perry, P. J.; Gowan, S. M.; Reszka, A. P.; Polucci, P.; Jenkins, T. C.; Kelland, L. R.; Neidle, S. 1,4- and 2,6-disubstituted amidoanthracene-9,10-dione derivatives as inhibitors of human telomerase. *J. Med. Chem*. **1998**, *41*, 3253-3260.

(103) Perry, P. J.; Reszka, A. P.; Wood, A. A.; Read, M. A.; Gowan, S. M.; Dosanjh, H. S.; Trent, J. O.; Jenkins, T. C.; Kelland, L. R.; Neidle, S. Human telomerase inhibition by regioisomeric disubstituted amidoanthracene-9,10-diones. *J. Med. Chem*. **1998**, *41*, 4873-4884.

(104) Read, M. A.; Wood, A. A.; Harrison, J. R.; Gowan, S. M.; Kelland, L. R.; Dosanjh, H. S.; Neidle, S. Molecular modeling studies on G-quadruplex complexes of telomerase inhibitors: structure-activity relationships. *J. Med. Chem*. **1999**, *42*, 4538-4546.

(105) Wheelhouse, R. T.; Sun, D. K.; Han, H. Y.; Han, F. X. G.; Hurley, L. H. Cationic porphyrins as telomerase inhibitors: the interaction of tetra-(*N*-methyl-4-pyridyl)porphine with quadruplex DNA. *J. Amer. Chem. Soc.* **1998,** *120*, 3261-3262.

(106) Mergny, J.-L. Lacroix, L.; Teulade-Fichou, T.; Hounsou, C.; Guittat, M.; Hourau, M.; Arimondo, P. B.; Vigneron, J.P.; Lehn, J.-M.; Riou, J.-F.; Garestier, T.; Hélène, C.; Telomerase inhibitors based on quadruplex ligands selected by a fluorescence assay. *Proc. Natl. Acad. Sci. USA* **2001**. *98*, 3062-3067.

(107) Han, H.; Bennett, R. J.; Hurley, L. H. 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*, 9311-9316.

(108) Gowan, S. M.; Heald, R.; Stevens, M. F.; Kelland, L. R. Potent inhibition of telomerase by small-molecule pentacyclic acridines capable of interacting with G-quadruplexes. *Mol. Pharm.* **2001,** *60*, 981-988.

(109) Monchaud, D.; Teulade-Fichou, M.-P. A hitchhiker's guide to G-quadruplex ligands. *Org. Biomol. Chem.* **2008,** *6*, 627-636.

(110) Franceschin, M. (2009) G-quadruplex DNA structures and organic chemistry: more than one connection. *Eur. J. Org. Chem*. **2009**, 2215-2220.

(111) Ralph, S. F. Quadruplex DNA: a promising drug target for the medicinal inorganic chemist. *Curr. Topics Med. Chem*. **2011**, *11*, 572-590.

(112) Reed, J. E.; Arnal, A. A.; Neidle, S.; Vilar, R. Stabilization of G-quadruplex DNA and inhibition of telomerase activity by square-planar nickel(II) complexes. *J. Amer. Chem. Soc.* **2006,** *128*, 5992-5993.

(113) Chen, Z.-F.; Qin, Q.-P.; Qin, J.-L.; Zhou, J.; Li, Y.-L.; Li, N.; Liu, Y.-C.; Liang, H. Water-soluble ruthenium(II) complexes with chiral 4-(2,3-dihydroxypropyl)-formamide oxoaporphine (FOA): in vitro and in vivo anticancer activity by stabilization of G-Quadruplex DNA, inhibition of telomerase activity, and induction of tumor cell apoptosis. *J. Med. Chem*. **2015**, *58*, 4771-4789.

(114) Kim, M. Y.; Vankayalapati, H.; Shin-Ya, K.; Wierzba, K.; Hurley, L. H. Telomestatin, a potent telomerase inhibitor that interacts quite specifically with the human telomeric intramolecular G-quadruplex. *J. Amer. Chem. Soc.* **2002,** *124*, 2098-2099.

(115) Linder, J.; Garner, T. P.; Williams, H. E. L.; Searle, M. S.; Moody, C. J. Telomestatin: formal total synthesis and cation-mediated interaction of its *seco*-derivatives with G-quadruplexes. *J. Amer. Chem. Soc*. **2011**, *133*, 1044-1051.

(116) Chung, W. J.; Heddi, B.; Tera, M.; Iida, K.; Nagasawa, K.; Phan, A. T. Solution structure of an intramolecular (3 + 1) human telomeric G-quadruplex bound to a telomestatin derivative. *J. Amer. Chem. Soc*. **2013**, *135*, 13495-13501.

(117) Nielsen, M. C.; Ulven, T. Macrocyclic G-quadruplex ligands. *Curr. Med. Chem*. **2010**, *17*, 3438-3448.

(118) Rzuczek, S. G.; Pilch, D. S.; Liu, A.; Liu, L., LaVoie, E. J.; Rice, J. E. Macrocyclic pyridyl polyoxazoles: selective RNA and DNA G-quadruplex ligands as antitumor agents. *J. Med. Chem*. **2010**, *53*, 3632-3644.

(119) Satyanarayana, M.; Kim, Y. A.; Rzuczek, S. G.; Pilch, D. S.; Liu, A. A.; Liu, L. F.; Rice, J. E.; LaVoie, E. J. Macrocyclic hexaoxazoles: Influence of aminoalkyl substituents

on RNA and DNA G-quadruplex stabilization and cytotoxicity. *Bioorg. Med. Chem. Lett.* **2010**, *20*, 3150-3154.

(120) Nandakumar, J.; Cech, T. R. Finding the end: recruitment of telomerase to telomeres. *Nature Rev. Mol. Cell Biol*. **2013**, *14*, 69-82.

(121) Blackburn, E. H. Telomeres and telomerase: the means to the end. *Angew. Chem. Int. Ed. Engl*. **2010**, *49*, 7405-7421.

(122) Gomez, D.; O'Donohue, M. F.; Wenner, T.; Douarre, C.; Macadré, J.; Koebel, P.; Giraud-Panis, M.-J.; Kaplan, H.; Kolkes, A.; Shin-Ya, K.; Riou, J.-F. 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*, 6908-6912.

(123) Gunaratnam, M.; Greciano, O.; Martins, C.; Reszka, A. P.; Schultes, C. M.; Morjani, H.; Riou, J.-F.; Neidle, S. Mechanism of acridine-based telomerase inhibition and telomere shortening. *Biochem. Pharmacol*. **2007**, *74*, 679-689.

(124) Sabater, L.; Nicolau-Travers, M.-L.; De Rache, A.; Prado, E.; Dejeu, J.; Bombarde, O.; Lacroix, J.; Calsou, P.; Defrancq, E.; Mergny, J.-L.; Gomez, D.; Pratviel, G. The nickel(II) complex of guanidinium phenyl porphyrin, a specific G-quadruplex ligand, targets telomeres and leads to POT1 mislocalization in culture cells. *J. Biol. Inorg. Chem*. **2015**, *20*, 729-738.

(125) Leonetti, C.; Amodei, S.; D'Angelo, C.; Rizzo, A.; Benassi, B.; Antonelli, A.; Elli, R.; Stevens, M. F. G.; D'Incalci, M.; Zupi, G.; Biroccio, A. 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*, 1138-1146.

(126) d'Adda di Fagagna, F. Living on a break: cellular senescence as a DNA-damage response. *Nat. Rev. Cancer* **2008**, *8*, 512-522.

(127) Rodriguez, R.; Müller, S.; Yeoman, J. A.; Trentesaux, C.; Riou, J.-F.; Balasubramanian, S. A novel small molecule that alters shelterin integrity and triggers a DNA-damage response at telomeres. *J. Amer. Chem. Soc*. **2008**, *130*, 15758-15759.

(128) Salvati, E.; Leonetti, C.; Rizzo, A.; Scarsella, M.; Mottolese, M.; Galati, R.; Sperduti, I.; Stevens, M. F. G.; D'Incalci, M.; Biasco, M.; Chiorino, G.; Bauwens, S.; Horard, B.; Gilson, E.; Stoppacciaro, A.; Zupi, G.; Biroccio, A. Telomere damage induced by the G-quadruplex ligand RHPS4 has an antitumor effect. *J. Clin. Investig*. **2007**, *117*, 3236-3247.

(129) Rizzo, A.; Salvati, E.; Porru, M.; D'Angelo, C.; Stevens, M. F.; D'Incalci, M.; Leonetti, C.; Gilson, E.; Zupi, G.; Biroccio, A. Stabilization of quadruplex DNA perturbs telomere replication leading to the activation of an ATR-dependent ATM signaling pathway. *Nucleic Acids Res*. **2009**, *37*, 5353-5364.

(130) Salvati, E.; Scarsella, M.; Porru, M.; Rizzo, A.; Iachettini, S.; Tentori, L.; Graziani, G.; D'Incalci, M.; Stevens, M. F. G.; Orlandi, A.; Passeri, D.; Gilson, E.; Zupi, G.; Leonetti, C.; Biroccio, A. PARP1 is activated at telomeres upon G4 stabilization: possible target for telomere-based therapy. *Oncogene* **2010**, *29*, 6280-6293.

(131) Tauchi, T.; Shin-ya, K.; Sashido, G.; Sumi, M.; Nakajima, A.; Shimamoto, T.; Ohyashiki, J. H.; Ohyashiki, K. Activity of a novel G-quadruplex-interactive telomerase inhibitor, telomestatin (SOT-095), against human leukemia cells: involvement of ATM-dependent DNA damage response pathways. *Oncogene* **2003**, ***22***, 5338-5347.

(132) Rodriguez, R; Miller, K. M.; Forment, J. V.; Bradshaw, C. R.; Nikan, M.; Britton, S.; Oelschlaegel, T.; Xhemalce, B.; Balasubramanian, S.; Jackson, S. P. Small-molecule-

induced DNA damage identifies alternative DNA structures in human genes. *Nature Chem. Biol*. **2012**, *8*, 301-310.

(133) Müller, S.; Sanders, D. A.; Di Antonio, M.; Matsis, S.; Riou, J.-F.; Rodriguez, R.; Balasubramanian, S. Pyridostatin analogues promote telomere dysfunction and long-term growth inhibition in human cancer cells. *Org. Biomol. Chem*. **2012**, *10*, 6537-6546.

(134) Dietlein, F.; Thelen, L.; Reinhardt, H. C. Cancer-specific defects in DNA repair pathways as targets for personalized therapeutic approaches.*Trends Genet*. **2014**, *30*, 326-339.

(135) Vannier, J.-B.; Pavicic-Kaltenbrunner, V.; Petalcorin, M. I.; Ding H.; Boulton, S. J. RTEL1 dismantles T loops and counteracts telomeric G4-DNA to maintain telomere integrity. *Cell* **2012**, *149*, 795-806.

(136) Hoffmann, R. F.; Moshkin, Y. M.; Mouton, S.; Grzeschik, N. A.; Kalicharan, R. D.; Kuipers, J,; Wolters, A. H.; Nishida, K.; Romashchenko, A. V.; Postberg, J.; Lipps, H.; Berezikov, E.; Sibon, O. C.; Giepmans, B. N.; Lansdorp, P. M. Guanine quadruplex structures localize to heterochromatin. *Nucleic Acids Res*. **2015**, doi: 10.1093/nar/gkv900.

(137) Murchie, A. I.; Lilley, D. M. Retinoblastoma susceptibility genes contain 5′ sequences with a high propensity to form guanine-tetrad structures. *Nucleic Acids Res*. **1992**, *20*, 49-53.

(138) Basu, S.; Wickstrom, E. Temperature and salt dependence of higher order structure formation by antisense *c-myc* and *c-myb* phosphorothioate oligodeoxyribonucleotides containing tetraguanylate tracts. *Nucleic Acids Res*. **1998**, *25*, 1327-1332.

(139) Simonsson, T.; Pecinka, P.; Kubista, M. DNA tetraplex formation in the control region of c-myc. *Nucleic Acids Res*. **1998**, *26*, 1167-1172.

(140) Siddiqui-Jain, A.; Grand, C. L.; Bearss, D. J.; Hurley, L. H. Direct evidence for a G-quadruplex in a promoter region and its targeting with a small molecule to repress c-MYC transcription. *Proc. Natl. Acad. Sci. USA* **2002**, 99, 11593-11598.

(141) Balasubramanian, S.; Hurley, L. H.; Neidle, S. Targeting G-quadruplexes in gene promoters: a novel anticancer strategy? *Nat. Rev. Drug. Discovery* **2011**, *10*, 261-275.

(142) Huppert, J. L.; Balasubramanian, S. G-quadruplexes in promoters throughout the human genome. *Nucleic Acids Res*. **2007**, *35*, 406–413.

(143) Eddy, J.; Maizels, N. Conserved elements with potential to form polymorphic G-quadruplex structures in the first intron of human genes. *Nucleic Acids Res*. **2008,** *36*, 1321-1333.

(144) Du, Z.; Zhao, Y.; Li, N. Genome-wide colonization of gene regulatory elements by G4 DNA motifs. *Nucleic Acids Res*. **2009**, *37*, 6784-6798.

(145) Beaume, N.; Pathak, R.; Yadav, V. K.; Kota, S.; Misra, H. S.; Gautam, H. K.; Chowdhury, S. Genome-wide study predicts promoter-G4 DNA motifs regulate selective functions in bacteria: radioresistance of *D. radiodurans* involves G4 DNA-mediated regulation. *Nucleic Acids Res*. **2013**, *41*, 76-89.

(146) Dai, J.; Chen, D.; Jones, R. A; Hurley, L. H.; Yang, D. NMR solution structure of the major G-quadruplex structure formed in the human BCL2 promoter region. *Nucleic Acids Res*. **2006**, *34*, 5133-5144.

(147) Membrino, A.; Cogoi, S.; Pedersen, E. B.; Xodo, L. E. G4-DNA formation in the HRAS promoter and rational design of decoy oligonucleotides for cancer therapy. *PLoS One* **2011**, *6*, e24421.

(148) Cogoi, S., Paramasivam, M., Spolaore, B., Xodo, L. E. Structural polymorphism within a regulatory element of the human KRAS promoter: formation of G4-DNA recognized by nuclear proteins. *Nucleic Acids Res*. **2008**, *36*, 3765-3780.

(149) Rankin, S.; Reszka, A. P.; Huppert, J.; Zloh, M.; Parkinson, G. N.; Todd, A. K.; Ladame, S.; Balasubramanian, S.; Neidle, S. Putative DNA quadruplex formation within the human *c-kit* oncogene. *J. Amer. Chem. Soc*. **2005**, *127*, 10584-10589.

(150) Fernando, H.; Reszka, A. P.; Huppert, J.; Ladame, S.; Rankin, S.; Venkitaraman, A. R.; Neidle, S.; Balasubramanian, S. A conserved quadruplex motif located in a transcription activation site of the human c-kit oncogene. *Biochemistry* **2006**, *45*, 7854–7860.

(151) De Armond, R.; Wood, S.; Sun, D.; Hurley, L. H.; Ebbinghaus, S. W. Evidence for the presence of a guanine quadruplex forming region within a polypurine tract of the hypoxia inducible factor 1α promoter. *Biochemistry* **2005**, *44*, 16341-16350.

(152) Lombardo, C. M.; Welsh, S. J.; Strauss, S. J.; Dale, A. G.; Todd, A. K.; Nanjunda, R.; Wilson, W. D.; Neidle, S. A novel series of G-quadruplex ligands with selectivity for HIF-expressing osteosarcoma and renal cancer cell lines. *Bioorg. Med. Chem. Lett*. **2012**, *22*, 5984-5988.

(153) Wei, D.; Todd, A. K.; Zloh, M.; Gunaratnam, M.; Parkinson, G. N.; Neidle. S. Crystal structure of a promoter sequence in the B-raf gene reveals an intertwined dimer quadruplex. *J. Amer. Chem. Soc*., **2013**, *135*, 19319-19329.

(154) Greco, M. L.; Folini, M.; Sissi, C. Double stranded promoter region of BRAF undergoes to structural rearrangement in nearly physiological conditions. *FEBS Lett*. **2015**, *589*, 2117-2123.

(155) Mitchell, T.; Ramos-Montoya, A.; Di Antonio, M.; Murat, P.; Ohnmacht, S.; Micco, M.; Jurmeister, S.; Fryer, L.; Balasubramanian, S.; Neidle, S.; Neal, D. E. Downregulation of androgen receptor transcription by promoter G-quadruplex stabilization as a potential alternative treatment for castrate-resistant prostate cancer. *Biochemistry*, **2013**, *52*, 1429-1436.

(156) Guo, K.; Pourpak, A.; Beetz-Rogers, K.; Gokhale, V.; Sun, D.; Hurley, L. H. Formation of pseudosymmetrical G-quadruplex and i-motif structures in the proximal promoter region of the RET oncogene. *J. Amer. Chem. Soc*., **2007**, *129*, 10220-10228.

(157) Tong, X.; Lan, W.; Zhang, X.; Wu, H.; Liu, M.; Cao, C. Solution structure of all parallel G-quadruplex formed by the oncogene RET promoter sequence. *Nucleic Acids Res*. **2011**, *39*, 6753-6763.

(158) Ohnmacht, S. A.; Micco, M.; Petrucci, V.; Todd, A. K.; Reszka, A. P.; Gunaratnam, M.; Carvalho, M. A.; Zloh, M.; Neidle, S. A sequence in the HSP90 promoter forms G-quadruplex structures with selectivity for phenyl bis-oxazole derivatives. *Bioorg. Med. Chem. Lett*. **2012**, *22*, 5930-5935.

(159) Yan, J.; Zhao, X.; Liu, B.; Yuan, Y.; Guan, Y. An intramolecular G-quadruplex structure formed in the human MET promoter region and its biological relevance. *Molec. Carcinogenesis* **2015**, DOI: 10.1002/mc.22330.

(160) Agrawal, P.; Hatzakis, E.; Guo, K.; Carver, M.; Yang, D. Solution structure of the major G-quadruplex formed in the human VEGF promoter in K^+: insights into loop interactions of the parallel G-quadruplexes. *Nucleic Acids Res*. **2013**, *41*, 10584-10592

(161) Todd, A. K.; Neidle, S. The relationship of potential G-quadruplex sequences in cis-upstream regions of the human genome to SP1-binding elements. *Nucleic Acids Res*. **2008**, *36*, 2700-2704.

(162) Yang, D.; Hurley, L. H. Structure of the biologically relevant G-quadruplex in the c-MYC promoter. *Nucleosides Nucleotides Nucleic Acids* **2006**, *25*, 951-968.

(163) Mathad, R. I.; Hatzakis, E.; Dai, J.; Yang, D. c-MYC promoter G-quadruplex formed at the 5'-end of NHE III1 element: insights into biological relevance and parallel-stranded G-quadruplex stability. *Nucleic Acids Res*. **2011**, *39*, 9023-9033.

(164) Phan, A. T.; Kuryavyi, V.; Gaw, H. Y.; Patel, D. J. Small-molecule interaction with a five-guanine-tract G-quadruplex structure from the human MYC promoter. *Nature Chem. Biol*. **2005**, *1*, 167-173.

(165) Vinagre, J.; Almeida, A.; Pópulo, H.; Batista, R.; Lyra, J.; Pinto, V.; Coelho, R.; Celestino, R.; Prazeres, H.; Lima, L.; Melo, M.; da Rocha, A. G.; Preto, A.; Castro, P.; Castro, L.; Pardal, F.; Lopes, J. M.; Santos, L. L.; Reis, R. M.; Cameselle-Teijeiro, J.; Sobrinho-Simões, M.; Lima, J.; Máximo, V.; Soares, P. Frequency of TERT promoter mutations in human cancers. *Nature Commun*. **2013**, *4*, 2185.

(166) Lim, K. W.; Lacroix, L.; Yue, D. J.; Lim, J. K.; Lim, J. M.; Phan, A. T. Coexistence of two distinct G-quadruplex conformations in the hTERT promoter. *J. Amer. Chem. Soc*., **2010**, *132*, 12331-12342.

(167) Chaires, J. B.; Trent, J. O.; Gray, R. D.; Dean, W. L.; Buscaglia, R.; Thomas, S. D.; Miller, D. M. An improved model for the hTERT promoter quadruplex. *PLoS One* **2014**, *9*, e115580.

(168) Wang, X.-D.; Ou, T.-M.; Lu, Y.-J.; Li, Z.; Xu, Z.; Xi, C.; Tan, J.-H.; Huang, S.-L.; An, L.-K.; Li, D.; Gu, L.-Q.; Huang, Z.-S. Turning off transcription of the bcl-2 gene by stabilizing the bcl-2 promoter quadruplex with quindoline derivatives. *J. Med. Chem*. **2010**, *53*, 4390-4398.

(169) Micco, M.; Collie, G. W.; Dale, A. G.; Ohnmacht, S. A.; Pazitna, I.; Gunaratnam, M.; Reszka, A. P.; Neidle, S. Structure-based design and evaluation of naphthalene diimide G-quadruplex ligands as telomere targeting agents in pancreatic cancer cells. *J. Med. Chem*. **2013**, *56*, 2959-2974.

(170) Ohnmacht, S. A.; Marchetti, C.; Gunaratnam, M.; Besser, R. J.; Haider, S. M.; Di Vita, G.; Lowe, H. L.; Mellinas-Gomez, M.; Diocou, S.; Robson, M.; Šponer, J.; Islam, B.; Pedley R. B.; Hartley, J. A.; Neidle, S. A G-quadruplex-binding compound showing

anti-tumour activity in an *in vivo* model for pancreatic cancer. *Scientific Rep*. **2015**, *5*, 11385.

(171) Nadai, M.; Cimino-Reale, G.; Sattin, G.; Doria, F.; Butovskaya, E.; Zaffaroni, N.; Freccero, M.; Palumbo, M.; Richter, S. N.; Folini, M. Assessment of gene promoter G-quadruplex binding and modulation by a naphthalene diimide derivative in tumor cells. *Int. J. Oncol*. **2015**, *46*, 369-380.

(172) Bejugam, M.; Gunaratnam, M.; Müller, S.; Sanders, D. A.; Sewitz, S.; Fletcher, J.; Neidle, S.; Balasubramanian, S. Targeting the c-Kit promoter quadruplexes with 6-substituted indenoisoquinolines, a novel class of G-quadruplex stabilising small molecule ligands. *ACS Med. Chem. Lett*. **2010**, *1*, 306-310.

(173) Bejugam, M.; Sewitz, S.; Shirude, P. S.; Rodriguez, R.; Shahid, R.; Balasubramanian, S. Trisubstituted isoalloxazines as a new class of G-quadruplex binding ligands: small molecule regulation of c-kit oncogene expression. *J. Amer. Chem. Soc*. **2007**, *129*, 12926-12927.

(174) McLuckie, K. I.; Waller, Z. A.; Sanders, D. A.; Alves, D.; Rodriguez, R.; Dash, J.; McKenzie, G. J.; Venkitaraman, A. R.; Balasubramanian, S. G-quadruplex-binding benzo[a]phenoxazines down-regulate c-KIT expression in human gastric carcinoma cells. *J. Amer. Chem. Soc*. **2011**, *133*, 2658-2663.

(175) Brooks, T. A.; Hurley, L. H Targeting MYC expression through G-quadruplexes. *Genes Cancer* **2010**, *1*, 641-649.

(176) Kendrick, S.; Kang, H. J.; Alam, M. P.; Madathil, M. M.; Agrawal, P.; Gokhale, V.; Yang, D.; Hecht, S. M.; Hurley, L. H. The dynamic character of the BCL2 promoter i-motif provides a mechanism for modulation of gene expression by compounds that bind selectively to the alternative DNA hairpin structure. *J. Amer. Chem. Soc*. **2014**, *136*, 4161-4171.

(177) Kang, H. J.; Kendrick, S.; Hecht, S. M.; Hurley, L. H. The transcriptional complex between the BCL2 i-motif and hnRNP LL is a molecular switch for control of gene expression that can be modulated by small molecules. *J. Amer. Chem. Soc*. **2014**, 136, 4172-4185.

(178) Boddupally, P. V.; Hahn, S.; Beman, C.; De, B.; Brooks, T. A.; Gokhale, V.; Hurley, L. H. Anticancer activity and cellular repression of c-MYC by the G-quadruplex-stabilizing 11-piperazinylquindoline is not dependent on direct targeting of the G-quadruplex in the c-MYC promoter. *J. Med. Chem*. **2012**, *55*, 6076-6086.

(179) Brown, R. V.; Danford, F. L.; Gokhale, V.; Hurley, L. H.; Brooks, T. A. Demonstration that drug-targeted down-regulation of MYC in non-Hodgkins lymphoma is directly mediated through the promoter G-quadruplex. *J. Biol. Chem*. **2011**, *286*, 41018-41027

(180) Brooks, T. A.; Hurley, L. H. The role of supercoiling in transcriptional control of MYC and its importance in molecular therapeutics. *Nature Rev. Cancer* **2010**, *1*, 641-649.

(181) Lavrado, J.; Brito, H.; Borralho, P. M.; Ohnmacht, S. A.; Kim, N. S.; Leitão, C.; Pisco, S.; Gunaratnam, M.; Rodrigues, C. M.; Moreira, R.; Neidle, S.; Paulo, A. KRAS oncogene repression in colon cancer cell lines by G-quadruplex binding indolo[3,2-c]quinolines. *Scientific Rep*. **2015**, *5*, 9696.

(182) Duan, W.; Rangan, A.; Vankayalapati, H.; Kim, M.Y.; Zeng, Q.; Sun, D.; Han, H.; Fedoroff, O.Y.; Nishioka, D.; Rha, S.Y.; Izbicka, E.; Von Hoff, D.D.; Hurley, L.H. Design and synthesis of fluoroquinophenoxazines that interact with human telomeric G-quadruplexes and their biological effects. *Mol. Cancer Ther.* **2001**, *1*, 103-120.

(183) Drygin, D.; Siddiqui-Jain, A.; O'Brien, S.; Schwaebe, M.; Lin, A.; Bliesath, J.; Ho, C. B.; Proffitt, C.; Trent, K.; Whitten, J. P.; Lim, J. K.; Von Hoff, D.; Anderes, K.; Rice, W.

G. Anticancer activity of CX-3543: a direct inhibitor of rRNA biogenesis. *Cancer Res*. **2009**, *69*, 7653-7661.

(184) Wieland, M.; Hartig, J. S. RNA quadruplex-based modulation of gene expression. *Chem.Biol*. **2007**, *14*, 757-763.

(185) Joachimi, A., Benz, A. & Hartig, J. S. (2009) A comparison of DNA and RNA quadruplex structures and stabilities. *Bioorg. Med. Chem*. **17**, 6811-6815.

(186) Chaires, J. B. Human telomeric G-quadruplex: thermodynamic and kinetic studies of telomeric quadruplex stability. *FEBS J*. **2010**, *277*, 1098-1106.

(187) Sugimoto, N. Noncanonical structures and their thermodynamics of DNA and RNA under molecular crowding: beyond the Watson-Crick double helix. *Int. Rev. Cell Mol. Biol*. **2014**, *307*, 205-273.

(188) Huppert, J. L.; Bugaut, A.; Kumari, S.; Balasubramanian, S. G-quadruplexes: the beginning and end of UTRs. *Nucleic Acids Res*. **2008**, *36*, 6260-6268.

(189) Bugaut, A.; Balasubramanian, S. 5'-UTR RNA G-quadruplexes: translation regulation and targeting. *Nucleic Acids Res*. **2012**, *40*, 4727-4741.

(190) Shahid, R.; Bugaut, A.; Balasubramanian, S. The BCL-2 5' untranslated region contains an RNA G-quadruplex-forming motif that modulates protein expression. *Biochemistry* **2010**, *49*, 8300-8306.

(191) Kumari, S.; Bugaut, A.; Huppert, J. L.; Balasubramanian, S. An RNA G-quadruplex in the 5' UTR of the NRAS proto-oncogene modulates translation. *Nature Chem. Biol*. **2007**, *3*, 218-221.

(192) Derecka, K.; Balkwill, G. D.; Garner, T. P.; Hodgman, C.; Flint, A. P.; Searle, M. S. Occurrence of a quadruplex motif in a unique insert within exon C of the bovine estrogen receptor alpha gene (ESR1). *Biochemistry* **2010**, *49*, 7625-7633.

(193) Morris, M. J.; Basu, S. An unusually stable G-quadruplex within the 5'-UTR of the MT3 matrix metalloproteinase mRNA represses translation in eukaryotic cells. *Biochemistry* **2009**, *48*, 5313-5319.

(194) Agarwala, P.; Pandey, S.; Mapa, K.; Maiti, S. The G-quadruplex augments translation in the 5' untranslated region of transforming growth factor β2. *Biochemistry* **2013**, *52*, 1528-1538.

(195) Wolfe, A. L.; Singh, K.; Zhong, Y.; Drewe, P.; Rajasekhar, V. K.; Sanghvi, V. R.; Mavrakis, K. J.; Jiang, M.; Roderick, J. E.; Van der Meulen, J.; Schatz, J. H.; Rodrigo, C. M.; Zhao, C.; Rondou, P.; de Stanchina, E.; Teruya-Feldstein, J.; Kelliher, M. A.; Speleman, F.; Porco, J. A. Jr.; Pelletier, J.; Rätsch, G.; Wendel, H. G. RNA G-quadruplexes cause eIF4A-dependent oncogene translation in cancer. *Nature* **2014**, *513*, 65-70.

(196) Todd, A. K.; Neidle, S. Mapping the sequences of potential guanine quadruplex motifs. *Nucleic Acids Res*. **2011**, *39*, 4917-4927.

(197) Arora, A.; Dutkiewicz, M.; Scaria, V.; Hariharan, M.; Maiti, S.; Kurreck, J. Inhibition of translation in living eukaryotic cells by an RNA G-quadruplex motif. *RNA* **2008**, *14*, 1290-1296.

(198) Gomez, D.; Guédin, A.; Mergny, J.-L.; Salles. B.; Riou, J.-F.; Teulade-Fichou, M.-P.; Calsou, P. A G-quadruplex structure within the 5'-UTR of TRF2 mRNA represses translation in human cells. *Nucleic Acids Res*. **2010**, *38*, 7187-7198.

(199) Biffi, G.; Di Antonio, M.; Tannahill, D.; Balasubramanian, S. Visualization and selective chemical targeting of RNA G-quadruplex structures in the cytoplasm of human cells. *Nature Chem*. **2014**, *6*, 75-80.

(200) Bugaut, A.; Rodriguez, R.; Kumari, S.; Hsu, S. T.; Balasubramanian, S. Small molecule-mediated inhibition of translation by targeting a native RNA G-quadruplex. *Org. Biomol. Chem*. **2010**, *8*, 2771-2776.

(201) Faudale, M.; Cogoi, S.; Xodo, L. E. Photoactivated cationic alkyl-substituted porphyrin binding to g4-RNA in the 5'-UTR of KRAS oncogene represses translation. *Chem. Commun*. **2012**, *48*, 874-876.

(202) Simone, R.; Fratta, P.; Neidle, S.; Parkinson, G. N.; Isaacs, A. M. G-quadruplexes: Emerging roles in neurodegenerative diseases and the non-coding transcriptome. *FEBS Lett*. **2015**, *589*, 1653-1668.

(203) Haeusler A.R.; Donnelly C.J.; Periz G.; Simko E.A.; Shaw P.G.; Kim M.S.; Maragakis N.J.; Troncoso J.C.; Pandey A.; Sattler R., Rothstein, J. D.; Wang, J. C9orf72 nucleotide repeat structures initiate molecular cascades of disease. *Nature* **2014**, *507*, 195-200.

(204) Fratta P.; Mizielinska S.; Nicoll A.J.; Zloh M.; Fisher E. M.; Parkinson G.; Isaacs A. M. C9orf72 hexanucleotide repeat associated with amyotrophic lateral sclerosis and frontotemporal dementia forms RNA G-quadruplexes. *Scientific Rep*. **2012**, *2*, 1016.

(205) Reddy K.; Zamiri B.; Stanley S.Y.; Macgregor R. B. Jr.; Pearson C. E. The disease-associated r(GGGGCC)$_n$ repeat from the C9orf72 gene forms tract length-dependent uni- and multimolecular RNA G-quadruplex structures. *J. Biol. Chem*. **2013**, *288*, 9860-9866.

(206) Zamiri, B.; Reddy K.; Macgregor, R. B. Jr,; Pearson, C. E. TMPyP4 porphyrin distorts RNA G-quadruplex structures of the disease-associated r(GGGGCC)$_n$ repeat of the C9orf72 gene and blocks interaction of RNA-binding proteins. *J. Biol. Chem*. **2014**, *289*, 4653-4659.

(207) Su, Z.; Zhang, Y.; Gendron, T. F.; Bauer, P. O.; Chew, J.; Yang, W. Y.; Fostvedt, E.; Jansen-West, K.; Belzil, V. V.; Desaro, P.; Johnston, A.; Overstreet, K.; Oh, S. Y.; Todd,

P. K.; Berry, J. D.; Cudkowicz, M. E.; Boeve, B. F.; Dickson, D.; Floeter, M. K.; Traynor, B. J.; Morelli, C.; Ratti, A.; Silani, V.; Rademakers, R.; Brown, R. H.; Rothstein, J. D.; Boylan, K. B.; Petrucelli, L.; Disney, M. D. Discovery of a biomarker and lead small molecules to target r(GGGGCC)-associated defects in c9FTD/ALS. *Neuron* **2014**, *83*, 1043-1050.

(208) De, S.; Michor, F. DNA secondary structures and epigenetic determinants of cancer genome evolution. *Nature Struct. Mol. Biol*. **2011**, *18*, 950-955.

(209) Tarsounas, M.; Tijsterman, M. Genomes and G-quadruplexes: for better or for worse. *J. Mol. Biol*. **2013**, *425*, 4782-4789.

(210) Maizels, N.; Gray, L. T. The G4 genome. *PLoS Genet*. **2013**, *9*, e1003468.

(211) Murat, P.; Balasubramanian, S. Existence and consequences of G-quadruplex structures in DNA. *Curr. Opin. Genet. Dev*. **2014**, *25*, 22-29.

(212) Nambiar, M.; Goldsmith, G.; Moorthy, B. T.; Lieber, M. R.; Joshi, M. V.; Choudhary, B.; Hosur R, V.; Raghavan, S. C. Formation of a G-quadruplex at the BCL2 major breakpoint region of the t(14;18) translocation in follicular lymphoma. *Nucleic Acids Res*. **2011**, *39*, 936-948.

(213) Rhodes, D.; Lipps, H. J. G-quadruplexes and their regulatory roles in cancer. *Nucleic Acids Res*. **2015**, *43*, 8627-8637.

(214) Bharti, S. K.; Sommers, J. A.; George, F.; Kuper, J.; Hamon, F.; Shin-ya, K.; Teulade-Fichou, M.-P.; Kisker, C.; Brosh, R. M. Jr. Specialization among iron-sulfur cluster helicases to resolve G-quadruplex DNA structures that threaten genomic stability. *J. Biol. Chem*. **2013**, *288*, 28217-28229.

(215) Gray, L. T.; Vallur, A. C.; Eddy, J.; Maizels, N. G quadruplexes are genomewide targets of transcriptional helicases XPB and XPD. *Nature Chem. Biol*. **2014**, *10*, 313-318.

(216) Ribeyre, C.; Lopes, J.; Boulé, J.-B.; Piazza, A.; Guédin, A.; Zakian, V. A.; Mergny, J.-L.; Nicolas, A. The yeast Pif1 helicase prevents genomic instability caused by G-quadruplex-forming CEB1 sequences in vivo. *PLOS Genet*. **2009**, 5:e1000475.

(217) Paeschke, K.; Bochman, M. L.; Garcia, P. D.; Cejka, P.; Friedman, K. L.; Kowalczykowski, S. C.; Zakian, V. A. Pif1 family helicases suppress genome instability at G-quadruplex motifs. *Nature* **2013**, *497*, 458-462.

(218) Castillo Bosch, P.; Segura-Bayona, S.; Koole, W.; van Heteren, J. T.; Dewar, J. M.; Tijsterman, M.; Knipscheer, P. FANCJ promotes DNA synthesis through G-quadruplex structures. *EMBO J*. **2014**, *33*, 2521-2533.

(219) Brosh, R. M. Jr.; Cantor, S. B. Molecular and cellular functions of the FANCJ DNA helicase defective in cancer and in Fanconi anemia. *Front. Genet*. **2014**, *5*, 372.

(220) Nguyen, G. H.; Tang, W.; Robles, A. I.; Beyer, R. P.; Gray, L. T.; Welsh, J. A.; Schetter, A. J.; Kumamoto, K.; Wang, X. W.; Hickson, I. D.; Maizels, N.; Monnat, R. J. Jr.; Harris, C. C. Regulation of gene expression by the BLM helicase correlates with the presence of G-quadruplex DNA motifs. *Proc. Natl. Acad. Sci. USA* **2014**, *111*, 9905-9910.

(221) Li, J.-L.; Harrison, R. J.; Reszka, A. P.; Brosh, R. M. Jr.; Bohr, V. A.; Neidle, S.; Hickson, I. D. Inhibition of the Bloom's and Werner's syndrome helicases by G-quadruplex interacting ligands. *Biochemistry* **2001**, *40*, 15194-15202.

(222) Vannier, J.-B.; Pavicic-Kaltenbrunner, V.; Petalcorin, M. I.; Ding, H.; Boulton, S. J. RTEL1 dismantles T loops and counteracts telomeric G4-DNA to maintain telomere integrity. *Cell* **2012**, *149*, 795-806.

(223) Vannier, J.-B.; Sarek, G.; Boulton, S. J. RTEL1: functions of a disease-associated helicase. *Trends Cell Biol*. **2014**, *24*, 416-425.

(224) Mendoza, O.; Gueddouda, N. M.; Boulé, J.-B.; Bourdoncle, A.; Mergny, J.-L. A fluorescence-based helicase assay: application to the screening of G-quadruplex ligands. *Nucleic Acids Res*. **2015**, *43*, e71.

(225) Heddi, B.; Cheong, V. V.; Martadinata, H.; Phan, A. T. Insights into G-quadruplex specific recognition by the DEAH-box helicase RHAU: Solution structure of a peptide-quadruplex complex. *Proc. Natl. Acad. Sci. USA* **2015**, *115*, 9608-9613.

(226) Campbell, N.; Collie, G. W.; Neidle, S. Crystallography of DNA and RNA G-quadruplex nucleic acids and their ligand complexes. *Curr. Protoc. Nucleic Acid Chem.* **2012**, Ch. 17, Unit 17.6.

(227) Karsisiotis, A. I.; Webba da Silva, M. Structural probes in quadruplex nucleic acid structure determination by NMR. *Molecules* **2012**, *57*, 13073-13086.

(228) Wang, Y.; Patel D. J. Solution structure of the human telomeric repeat d[$AG_3(T_2AG_3)_3$] G-tetraplex. *Structure* **1993**, *1***,** 263-282.

(229) Phan, A. T.; Patel, D. J. Two-repeat human telomeric d(TAGGGTTAGGGT) sequence forms interconverting parallel and antiparallel G-quadruplexes in solution: distinct topologies, thermodynamic properties, and folding/unfolding kinetics. *J. Amer. Chem. Soc.* **2003***, 125*, 15021-15027.

(230) Ambrus, A.; Chen, D.; Dai, J.; Bialis, T.; Jones, R. A.; Yang, D. Human telomeric sequence forms a hybrid-type intramolecular G-quadruplex structure with mixed parallel/antiparallel strands in potassium solution. *Nucleic Acids Res.* **2006**, *34*, 2723-2735.

(231) Dai, J.; Carver, M.; Punchihewa, C.; Jones, R. A.; Yang, D. Structure of the hybrid-2 type intramolecular human telomeric G-quadruplex in K^+ solution: insights into structure polymorphism of the human telomeric sequence. *Nucleic Acids Res.* **2007**, *35*, 4927-4740.

(232) Luu, K. N.; Phan, A. T.; Kuryavyi, V.; Lacroix, L.; Patel, D. J. Structure of the human telomere in K^+ solution: an intramolecular (3+1) G-quadruplex scaffold. *J. Amer. Chem. Soc.* **2006***, 128*, 9963-9970.

(233) Lim, K. W.; Amrane, S.; Bouaziz, S.; Xu, W.; Mu, Y.; Patel, D. J.; Luu, K. N.; Phan, A. T. Structure of the human telomere in K^+ solution: a stable basket-type G-quadruplex with only two G-tetrad layers. *J. Amer. Chem. Soc.* **2009**, *131*, 4301-4309.

(234) Phan, A. T.; Kuryavyi, V.; Luu, K. N.; Patel, D. J. Structure of two intramolecular G-quadruplexes formed by natural human telomere sequences in K^+ solution. *Nucleic Acids Res.* **2007***, 35*, 6517-6525.

(235) Parkinson, G. N.; Lee, M. P. H.; Neidle, S. Crystal structure of parallel quadruplexes from human telomeric DNA. *Nature* **2002**, *417*, 876-880.

(236) Heddi, B.; Phan, A. T. Structure of human telomeric DNA in crowded solution. *J. Amer. Chem. Soc.* **2013***, 133*, 9824-9833.

(237) Renčiuk, D.; Kejnovská, I.; Školáková, P.; Bednářová, K.; Motlová, J.; Vorlíčková, M. Arrangements of human telomere DNA quadruplex in physiologically relevant K^+ solutions. *Nucleic Acids Res.* **2009**, *37*, 6625-6634.

(238) Campbell, N. H.; Parkinson, G. N.; Reszka, A. P.; Neidle, S. Structural basis of DNA quadruplex recognition by an acridine drug. *J. Amer. Chem. Soc.* **2008***, 130*, 6722-6724.

(239) Parkinson, G. N.; Cuenca, F.; Neidle, S. Topology conservation and loop flexibility in quadruplex-drug recognition: crystal structures of inter- and intramolecular telomeric DNA quadruplex-drug complexes. *J. Mol. Biol.* **2008**, *381*, 1145-1156.

(240) Collie, G. W.; Promontorio, R.; Hampel, S. M.; Micco, M.; Neidle, S.; Parkinson, G. N. Structural basis for telomeric G-quadruplex naphthalene diimide ligand targeting *J. Amer. Chem. Soc*. **2012**, *134*, 2723-2731.

(241) Micco, M.; Collie, G. W.; Dale, A. G.; Ohnmacht, S. A.; Pazitna, I.; Gunaratnam, M.; Reszka, A. P.; Neidle, S. Structure-based design and evaluation of naphthalene diimide G-quadruplex ligands as telomere targeting agents in pancreatic cancer cells. *J. Med. Chem*. **2013**, *56*, 2959-2974.

(242) Campbell, N. H.; Karim, N. H.; Parkinson, G. N.; Gunaratnam, M.; Petrucci, V.; Todd, A. K.; Vilar, R.; Neidle, S. Molecular basis of structure-activity relationships between salphen metal complexes and human telomeric DNA quadruplexes. *J. Med. Chem*. **2012**, *55*, 209-222.

(243) Bazzicalupi, C.; Ferraroni, M.; Bilia, A. R.; Scheggi, F.; Gratteri, P. The crystal structure of human telomeric DNA complexed with berberine: an interesting case of stacked ligand to G-tetrad ratio higher than 1:1. *Nucleic Acids Res*. **2012**, *41*, 632-638.

(244) Nicoludis, J. M.; Miller, S. T.; Jeffrey, P. D.; Barrett, S. P.; Rablen, P. R.; Lawton, T. J.; Yatsunyk, L. A. Optimized end-stacking provides specificity of N-methyl mesoporphyrin IX for human telomeric G-quadruplex DNA. *J. Amer. Chem. Soc*. **2012**, *134*, 20446-20456.

(245) Collie, G. W.; Campbell, N. H.; Neidle, S. Loop flexibility in human telomeric quadruplex small-molecule complexes. *Nucleic Acids Res.* **2015,** *43*, 4785-4799.

(246) Chung, W. J.; Heddi, B.; Tera, M.; Iida, K.; Nagasawa, K.; Phan, A. T. Solution structure of an intramolecular (3 + 1) human telomeric G-quadruplex bound to a telomestatin derivative. *J. Amer. Chem. Soc*. **2013**, *135*, 13495-13501.

(247) Vorlíčková, M.; Kejnovská, I.; Sagi, J.; Renčiuk D.; Bednářová, K.; Motlová, J.; Kypr, J. Circular dichroism and guanine quadruplexes. *Methods* **2014**, *67*, 159-168.

(248) Haider, S. M.; Autiero, I.; Neidle, S. Surface area accessibility and the preferred topology of telomeric DNA quadruplex-ligand complexes. *Biochimie*. **2011**, *93*, 1275-1279.

(249) Collie, G. W.; Haider, S.; Neidle, S.; Parkinson, G. N. A crystallographic and modelling study of a human telomeric RNA (TERRA) quadruplex. *Nucleic Acids Res.* 2010**,** *38*, 5569-5580.

(250) Martadinata, H.; Phan, A. T. Structure of propeller-type parallel-stranded RNA G-quadruplexes, formed by human telomeric RNA sequences in K+ solution. *J. Amer. Chem. Soc*. **2009**, *131*, 2570-2578.

(251) Collie, G. W.; Sparapani, S.; Parkinson, G. N.; Neidle, S. Structural basis of telomeric RNA quadruplex-acridine ligand recognition *J. Amer. Chem. Soc*. **2011**, *133*, 2721-2728.

(252) Ambrus, A.; Chen, D.; Dai, J.; Jones, R. A.; Yang, D. Solution structure of the biologically relevant G-quadruplex element in the human c-MYC promoter. Implications for G-quadruplex stabilization. *Biochemistry* **2005**, *44*, 2048-2058.

(253) Mathad, R. I.; Hatzakis, E., Dai, J.; Yang, D. c-MYC promoter G-quadruplex formed at the 5'-end of NHE III1 element: insights into biological relevance and parallel-stranded G-quadruplex stability. *Nucleic Acids Res*. **2011**, *39*, 9023-9033.

(254) Dai, J.; Chen, D.; Jones, R. A.; Hurley, L. H.; Yang, D. NMR solution structure of the major G-quadruplex structure formed in the human BCL2 promoter region. *Nucleic Acids Res*. **2006**, *34*, 5133-5144.

(255) Agrawal, P.; Lin, C.; Mathad, R. I.; Carver, M.; Yang, D. The major G-quadruplex formed in the human BCL-2 proximal promoter adopts a parallel structure with a 13-nt loop in K^+ solution. *J. Amer. Chem. Soc*. **2014**, *136*, 1750-1753.

(256) Agrawal, P.; Hatzakis, E.; Guo, K.; Carver, M.; Yang, D. Solution structure of the major G-quadruplex formed in the human VEGF promoter in K+: insights into loop interactions of the parallel G-quadruplexes. *Nucleic Acids Res*. **2013**, *41*, 10584-10592.

(257) Tong, X.; Lan, W.; Zhang, X.; Wu, H.; Liu, M.; Cao, C. Solution structure of all parallel G-quadruplex formed by the oncogene RET promoter sequence. *Nucleic Acids Res*. **2011**, *39*, 6753-6763.

(258) Todd, A. K.; Haider, S. M.; Parkinson, G. N.; Neidle, S. Sequence occurrence and structural uniqueness of a G-quadruplex in the human *c-kit* promoter. *Nucleic Acids Res.* **2007**, *35*, 5799–5808.

(259) Phan, A. T.; Kuryavyi, V.; Gaw, H. Y.; Patel, D. J. Small-molecule interaction with a five-guanine-tract G-quadruplex structure from the human MYC promoter. *Nature Chem. Biol*. **2005**, *1*, 167-173.

(260) Chung, W. J.; Heddi, B.; Hamon, F.; Teulade-Fichou, M.-P.; Phan, A. T. Solution structure of a G-quadruplex bound to the bisquinolinium compound Phen-DC(3). *Angew. Chem. Int. Ed. Engl*. **2014**, *53*, 999-1002.

(261) Dai, J.; Carver, M.; Hurley, L. H.; Yang, D. Solution structure of a 2:1 quindoline-c-MYC G-quadruplex: insights into G-quadruplex-interactive small molecule drug design. *J. Amer. Chem. Soc*. **2011**, *133*, 17673-17680.

(262) Vasilyev, N.; Polonskaia, A.; Darnell, J. C.; Darnell, R. B.; Patel, D. J.; Serganov, A. Crystal structure reveals specific recognition of a G-quadruplex RNA by a β-turn in the RGG motif of FMRP. *Proc. Natl. Acad. Sci. USA* **2015**, *112*, E5391-400.

(263) Warner, K. D.; Chen, M. C.; Song, W.; Strack, R. L.; Thorn, A.; Jaffrey, S. R.; Ferré-D'Amaré, A. R. Structural basis for activity of highly efficient RNA mimics of green fluorescent protein. *Nature Struct. Mol. Biol*. **2014**, *21*, 658-663.

(264) Huang, H.; Suslov, N. B.; Li, N. S.; Shelke, S. A.; Evans, M. E.; Koldobskaya, Y.; Rice, P. A.; Piccirilli, J. A. A G-quadruplex-containing RNA activates fluorescence in a GFP-like fluorophore. *Nature Chem. Biol*. **2014**, *10*, 686-691.

(265) Ohnmacht, S. A.; Neidle, S. Small-molecule quadruplex-targeted drug discovery. *Bioorg. Med. Chem. Lett*. **2014**, *24*, 2602-2612.

(266) Ohnmacht, S. A.; Varavipour, E.; Nanjunda, R.; Pazitna, I.; Di Vita, G.; Gunaratnam, M.; Kumar, A.; Ismail, M. A.; Boykin, D. W.; Wilson, W. D.; Neidle, S. Discovery of new G-quadruplex binding chemotypes. *Chem. Commun*. **2014**, *50*, 960-963.

(267) Hamon, F.; Largy, E.; Guédin-Beaurepaire, A.; Rouchon-Dagois, M.; Sidibe, A.; Monchaud, D.; Mergny, J.-L.; Riou, J.-F.; Nguyen, C.-H.; Teulade-Fichou, M.-P. An acyclic oligoheteroaryle that discriminates strongly between diverse G-quadruplex topologies. *Angew. Chem. Int. Ed. Engl*. **2011**, *50*, 8745-8749.

(268) Yu, H.; Gu, X.; Nakano, S. I.; Miyoshi, D.; Sugimoto, N. Beads-on-a-string structure of long telomeric DNAs under molecular crowding conditions. *J. Amer. Chem. Soc*. **2012**, *134*, 20060-20069.

(269) Cousins, A. R.; Ritson, D.; Sharma, P.; Stevens, M. F.; Moses, J. E.; Searle, M. S. Ligand selectivity in stabilising tandem parallel folded G-quadruplex motifs in human telomeric DNA sequences. *Chem. Commun*. **2014**, *50*, 15202-15205.

(270) Dhamodharan, V.; Harikrishna, S.; Bhasikuttan, A. C.; Pradeepkumar, P. I. Topology specific stabilization of promoter over telomeric G-quadruplex DNAs by bisbenzimidazole carboxamide derivatives. *ACS Chem. Biol*. **2015**, *10*, 821-833.

(271) Felsenstein, K. M.; Saunders, L. B.; Simmons, J. K.; Leon, E.; Calabrese, D. R.; Zhang, S.; Michalowski, A.; Gareiss, P.; Mock, B. A.; Schneekloth, J. S. Jr. Small Molecule microarrays enable the identification of a selective, quadruplex-binding inhibitor of MYC expression. *ACS Chem. Biol*. **2016**, *11*, 139-148

(272) Li, Q.; Xiang, J. F.; Yang, Q. F.; Sun, H. X.; Guan, A.J.; Tang, Y.L. G4LDB: a database for discovering and studying G-quadruplex ligands. *Nucleic Acids Res*. **2013**, *41*, D1115-23.

(273) Ma, D. L.; Ma, V. P.; Chan, D. S.; Leung, K. H.; Zhong, H. J.; Leung, C. H. In silico screening of quadruplex-binding ligands. *Methods* **2012**, *57*, 106-114.

(274) Cosconati, S.; Marinelli, L.; Trotta, R.; Virno, A.; Mayol, L.; Novellino, E.; Olson, A. J.; Randazzo, A. Tandem application of virtual screening and NMR experiments in the discovery of brand new DNA quadruplex groove binders. *J. Amer. Chem. Soc*. **2009**, *131*, 16336-16337.

(275) Cosconati, S.; Marinelli, L.; Trotta. R.; Virno, A.; De Tito, S.; Romagnoli, R.; Pagano, B.; Limongelli, V.; Giancola, C.; Baraldi, P.G.; Mayol, L.; Novellino, E.; Randazzo A. Structural and conformational requisites in DNA quadruplex groove binding: another piece to the puzzle. *J. Amer. Chem. Soc.* **2010**, *132*, 6425-6433.

(276) Alcaro, S.; Musetti, C.; Distinto, S.; Casatti, M.; Zagotto, G.; Artese, A.; Parrotta, L.; Moraca, F.; Costa, G.; Ortuso, F.; Maccioni, E.; Sissi, C. Identification and characterization of new DNA G-quadruplex binders selected by a combination of ligand and structure-based virtual screening approaches. *J. Med. Chem*. **2013**, *56*, 843-855.

(277) Di Leva, F. S.; Zizza, P.; Cingolani, C.; D'Angelo, C.; Pagano, B.; Amato, J.; Salvati, E.; Sissi, C.; Pinato, O.; Marinelli, L.; Cavalli, A.; Cosconati, S.; Novellino, E.; Randazzo, A.; Biroccio, A. Exploring the chemical space of G-quadruplex binders: discovery of a novel chemotype targeting the human telomeric sequence. *J. Med. Chem*. **2013**, *56*, 9646-9654.

(278) Husby, J.; Todd, A. K.; Platts, J. A.; Neidle, S. Small-molecule G-quadruplex interactions: Systematic exploration of conformational space using multiple molecular dynamics. *Biopolymers* **2013**, *99*, 989-1005.

(279) Sattin, G.; Artese, A.; Nadai, M.; Costa, G.; Parrotta, L.; Alcaro, S.; Palumbo, M.; Richter, S. N. Conformation and stability of intramolecular telomeric G-quadruplexes: sequence effects in the loops. *PLoS One* **2013**, *8*, e84113.

(280) Kang, H.-J.; Park, H.-J. In silico identification of novel ligands for G-quadruplex in the c-MYC promoter. *J. Comp. Aided Mol. Des*. **2015**, *29*, 339-348.

(281) Hou, J.-Q.; Chen, S.- B.; Zan, L.-P.; Ou, T.-M.; Tan, J. H.; Luyt, L. G.; Huang, Z.-S. Identification of a selective G-quadruplex DNA binder using a multistep virtual screening approach. *Chem. Commun*. **2015**, *51*, 198-201.

(282) Nasiri, H. R.; Bell, N. M.; McLuckie, K. I.; Husby, J.; Abell, C.; Neidle, S, Balasubramanian S. Targeting a c-MYC G-quadruplex DNA with a fragment library. *Chem. Commun*. **2014**, *50*, 1704-1707.

(283) Yang, D.; Okamoto, K. Structural insights into G-quadruplexes: towards new anticancer drugs. *Future Med. Chem*. **2010**, *2*, 619-646

(284) Russo Krauss, I.; Ramaswamy, S.; Neidle, S.; Haider, S.; Parkinson, G. N. Structural insights into the quadruplex-duplex 3' interface formed from a telomeric repeat: a potential molecular target. *J. Amer. Chem. Soc*. **2016**, DOI: 10.1021/jacs.5b10492.

(285) Rodriguez, R.; Miller, K. M. Unravelling the genomic targets of small molecules using high-throughput sequencing. *Nature Rev. Genet*. **2014**, *15*, 783-796.

(286) Rizzo, A.; Iachettini, S.; Zizza, P.; Cingolani, C.; Porru, M.; Artuso, S.; Stevens, M.; Hummersone, M.; Biroccio, A.; Salvati, E.; Leonetti, C. Identification of novel RHPS4-derivative ligands with improved toxicological profiles and telomere-targeting activities. *J. Exp. Clin. Cancer Res*. **2014**, *33*, 81.

(287) Porru, M.; Artuso, S.; Salvati, E.; Bianco, A.; Franceschin, M.; Diodoro, M. G.; Passeri, D.; Orlandi, A.; Savorani, F.; D'Incalci, M.; Biroccio, A.; Leonetti, C. Targeting G-quadruplex DNA structures by EMICORON has a strong antitumor efficacy against advanced models of human colon cancer. *Mol. Cancer Ther*. **2015**, *14*, 2541-2551.

(288) Weinstein, I. B. Cancer. Addiction to oncogenes--the Achilles heal of cancer. *Science* **2002**, *297*, 63-64.

(289) Pagliarini, R.; Shao, W.; Sellers, W. R. Oncogene addiction: pathways of therapeutic response, resistance, and road maps toward a cure. *EMBO Rep*. **2015**, *16*, 280-296.

Figure legends

Figure 1

a. Stick representation of a guanine G-quartet, with hydrogen bonds shown as dashed lines. A potassium ion at the centre of the quartet is also shown. This is coordinated to O6 atoms of the quartet and also to those of the stacked adjacent quartet.

b. Representation of the assembly of an intramolecular quadruplex from a human telomeric DNA sequence (PDB id 1KF1)[235], illustrating the donation of one guanine from each of the four G-tracts to form an individual G-quartet. Three quartets stack together to form this quadruplex, linked by propeller loops

Figure 2

a. Cartoon representation of a quadruplex (PDB id 1JPQ), with various features highlighted and labelled. The stack of four G-quartets constitute the quadruplex core. Guanine bases are coloured blue. The loops contain thymine bases, coloured cyan.

b. View of the same quadruplex, now looking down onto the quartet plane.

Figure 3

A cartoon view of the DEAH helicase peptide-quadruplex complex, as determined by 2D-NMR.[225] The α-helix core of the peptide is shown sitting on an external G-quartet face (PDB id 2N21).

Figure 4

A view of a (3+1) hybrid human telomeric quadruplex, determined by 2D-NMR (PDB id 2HY9).[231]

Figure 5

View showing the crystal structure[241] of an intramolecular human telomeric DNA G-quadruplex bound by the tetra-substituted naphthalene-diimide compound (**16**), shown in stick form (PDB id 3UYH).

Figure 6

Two orthogonal views of the 2D-NMR structure[246] of a complex between a human telomeric quadruplex and a telomestatin analogue (**23**), shown in stick representation (PDB id 2MB3).

Figure 7

Cartoon view of the parallel quadruplex formed from a sequence in the human VEGF promoter (PDB id 2M27), determined by 2D-NMR.[256]

Figure 8

Cartoon view of the high-resolution crystal structure[43] for a quadruplex formed from a sequence in the c-*KIT* promoter (PDB id 3QXR).

Figure 9

View of the 2D-NMR structure[260] of a c-*MYC* promoter quadruplex, with bound (bisquinolinium) phenanthroline Phen-DC_3 (PDB id 2MGN), projected onto the plane of the G-quartets. Note the extensive π-π overlap between the four guanines of the terminal quartet and the two quinolinium and phenanthroline moieties.

Figure 10

A molecular model for Phen-DC_3 intercalated between two Watson-Crick base pairs in duplex DNA. Note that although the phenanthroline group is stacked between the base pairs, the two quinolinium groups are necessarily completely protruding into solvent and do not make any contacts with the DNA.

Figure 11

View of the 2:1 complex between a mono-substituted quindoline compound[261] (shown in stick representation) and a c-*MYC* promoter complex, determined by 2D-NMR (PDB id 2L7V).

Figure 12

Co-crystal structure[262] of the RNA quadruplex which is the target of the Fragile X Mental Retardation Protein (FMRP) bound to the arginine-glycine-rich (RGG) motif of the protein. The close-up view shows the part of the RNA stem and quadruplex, in cartoon form, together with the peptide (with the β-turn and backbone in cartoon form and side-chain atoms in stick form), bound onto a mixed quartet surface at one end of the quadruplex.

Figure 13

a. Close-up view of the crystal structure[264] (PDB id 4Q9R) RNA quadruplex region of the RNA aptamer mimicking green fluorescent protein, bound to a substituted hydroxybenzylidene imidazolinone molecule (shown in space-filling mode).
b. Close-up looking onto the plane of the G-quartets, showing the overlap between the bound hydroxybenzylidene imidazolinone molecule (in stick form) and the guanines.

Figure 1a

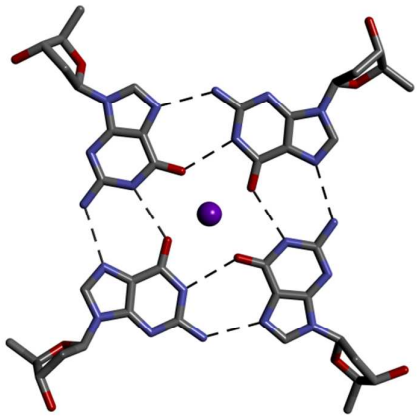


Figure 1b

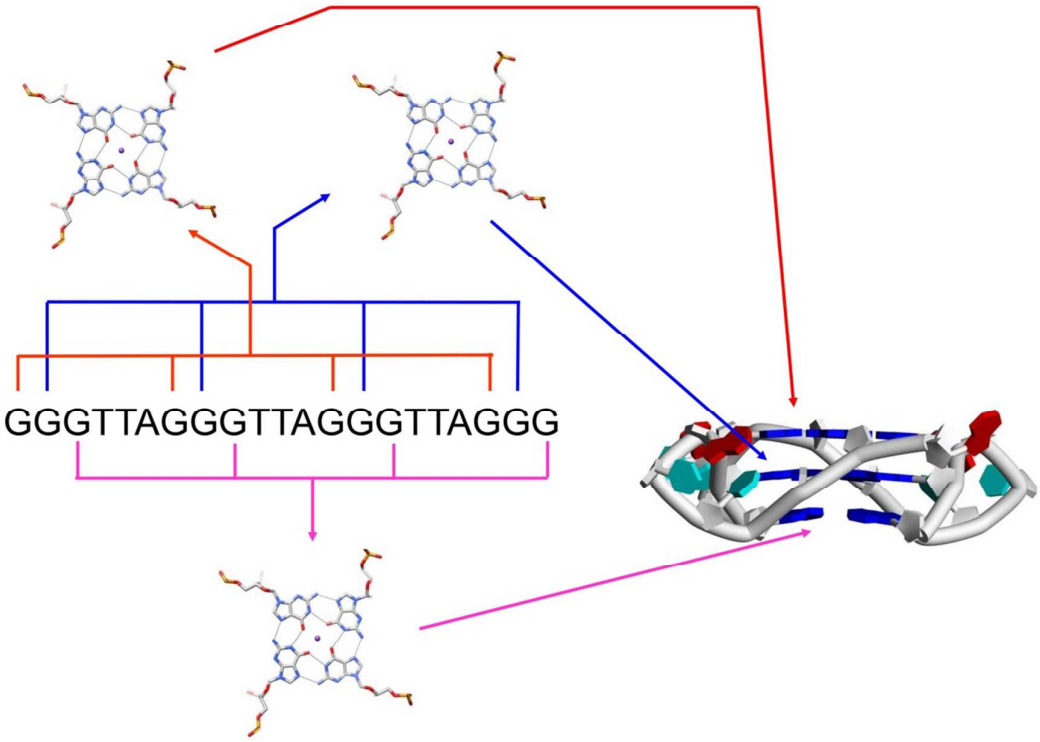


Figure 2a

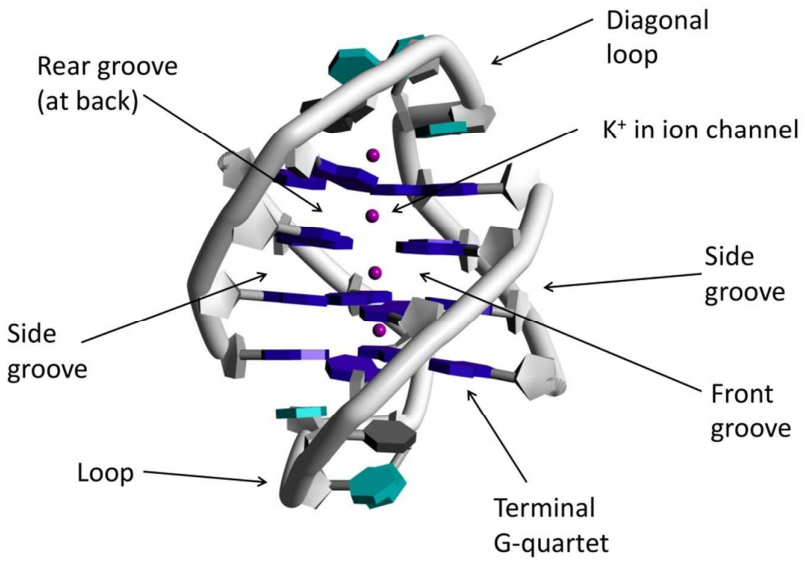


Figure 2b

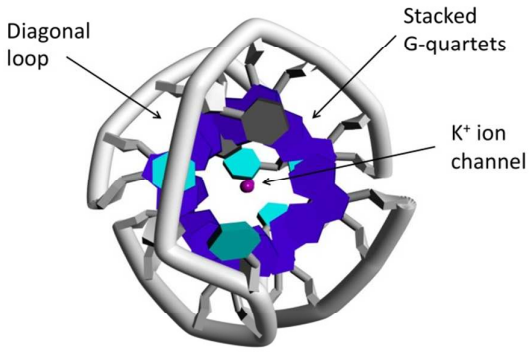


Figure 3

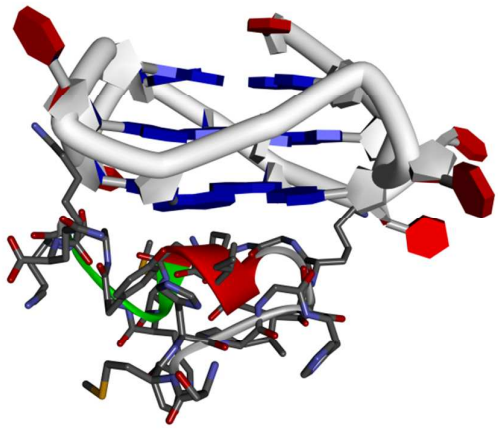


Figure 4

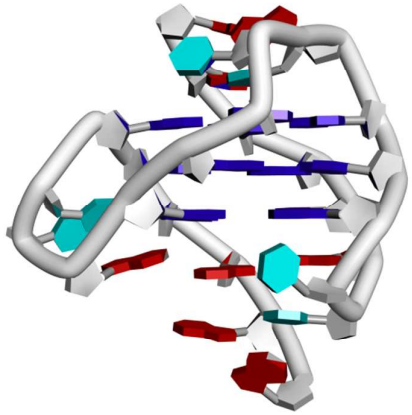


Figure 5

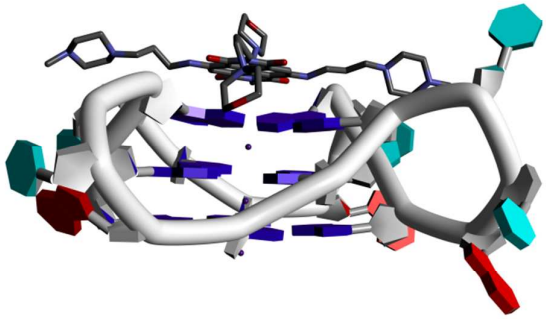
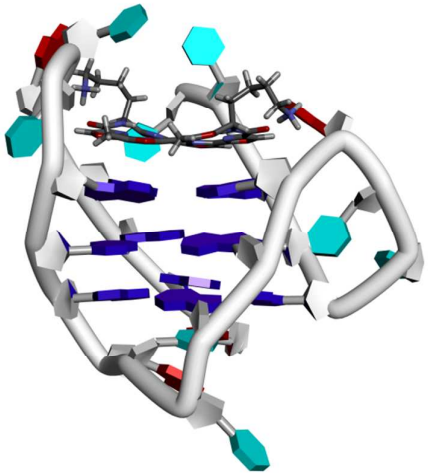


Figure 6



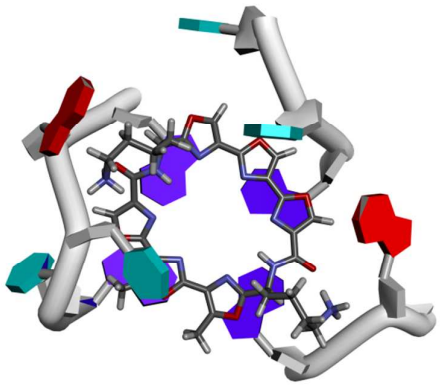


Figure 7

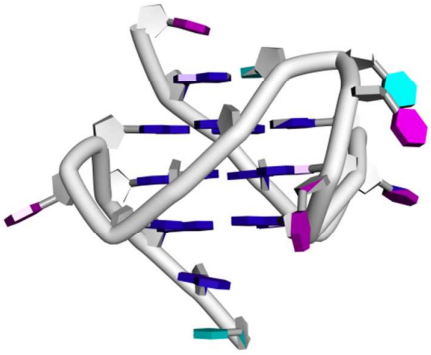


Figure 8

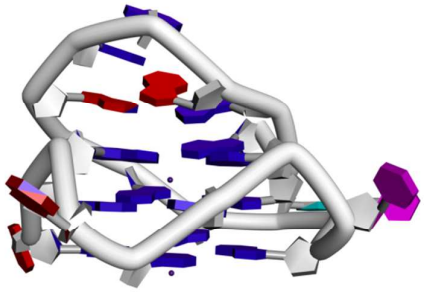


Figure 9

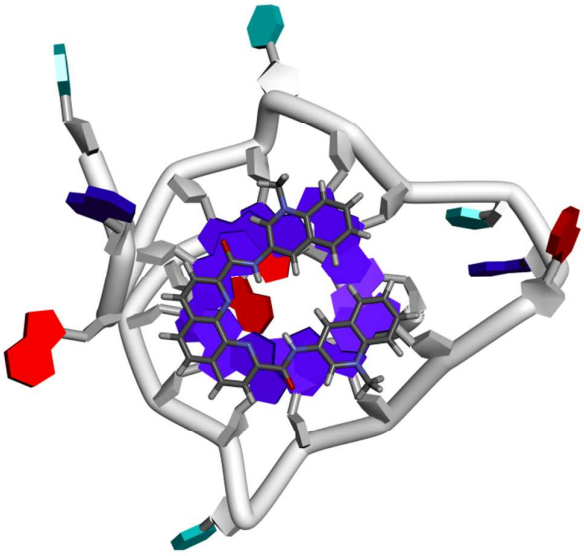


Figure 10

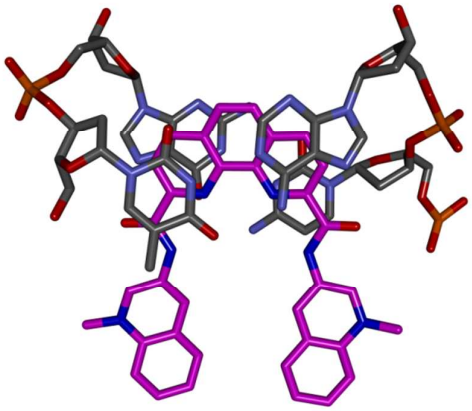


Figure 11

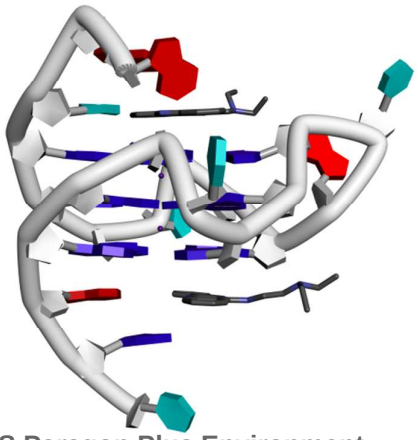


Figure 12

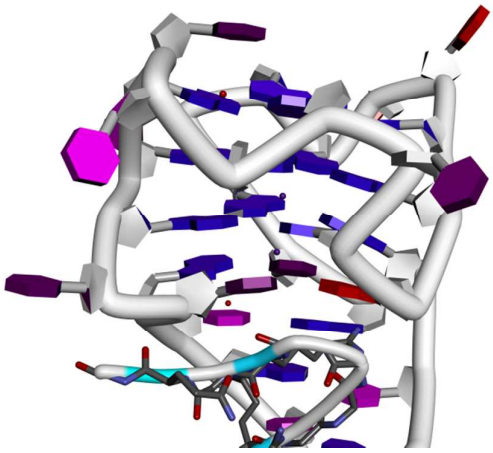
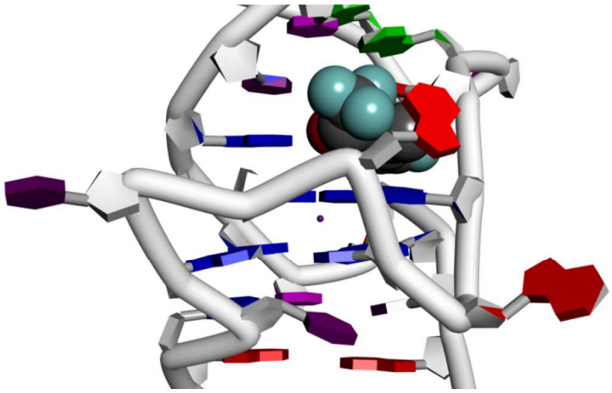
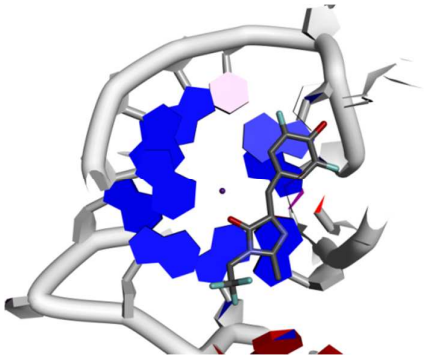


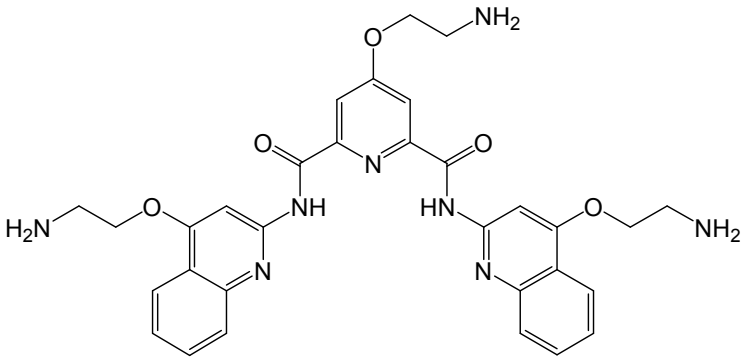
Figure 13

b **b**

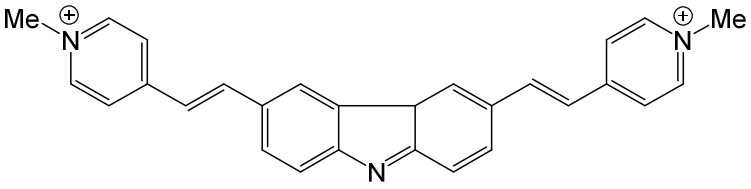




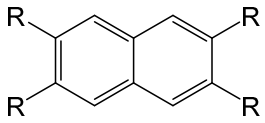
1



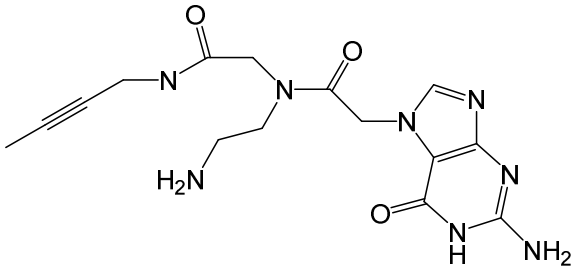
2



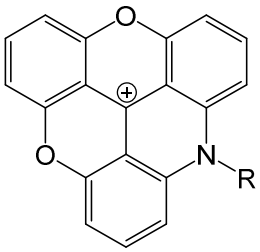
3



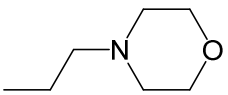
R =



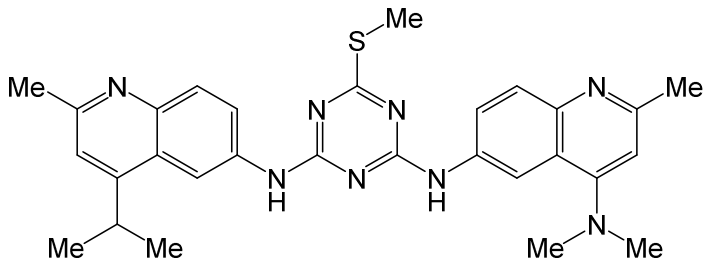
4



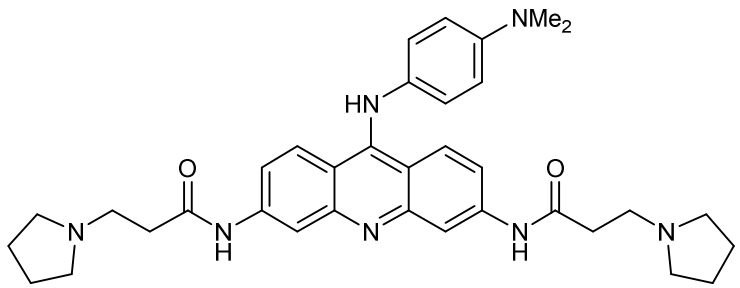
R =



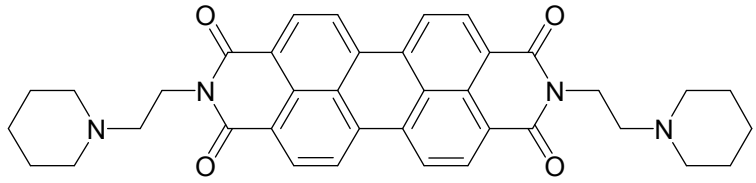
5



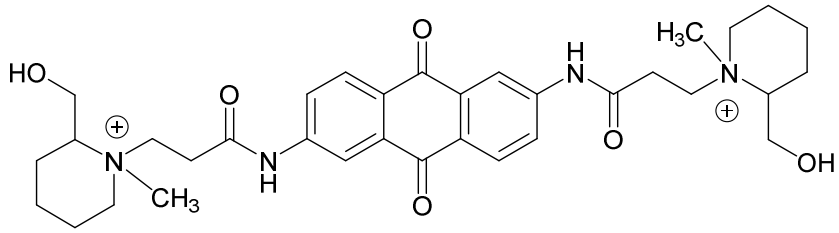
6



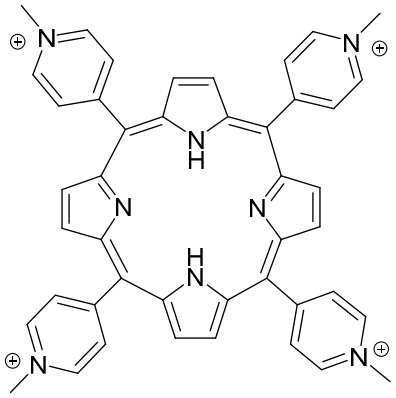
7



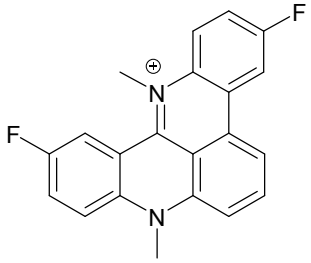
8



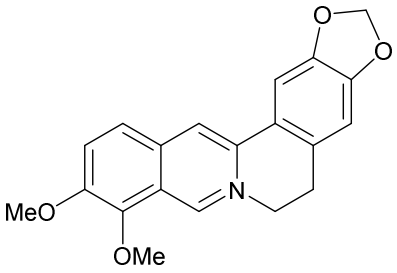
9



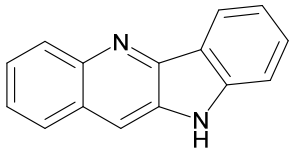
10



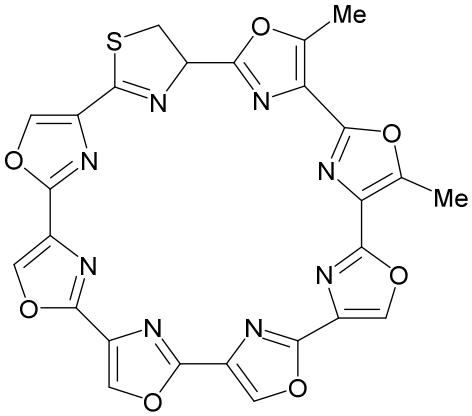
11



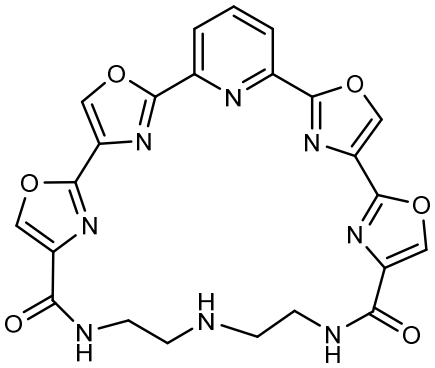
12



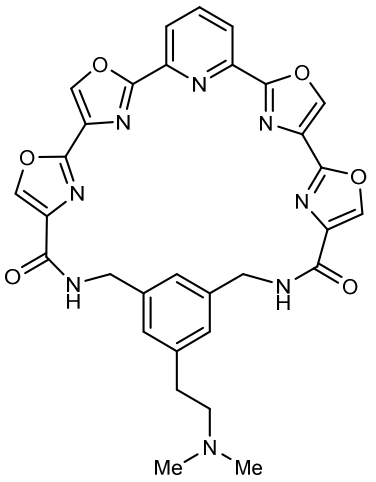
13



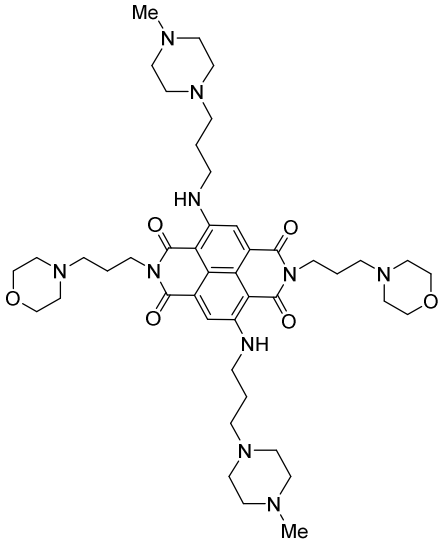
14



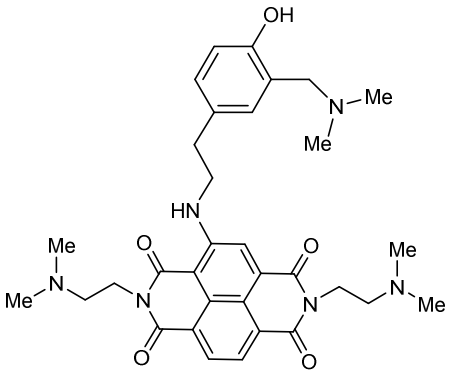
15



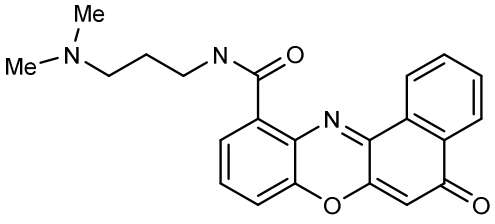
16



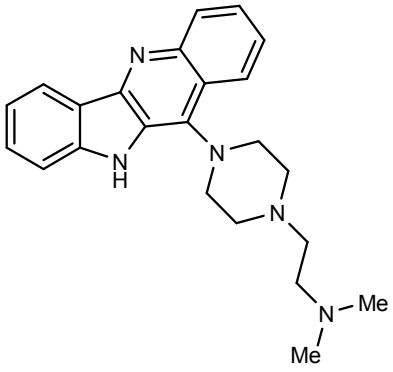
17



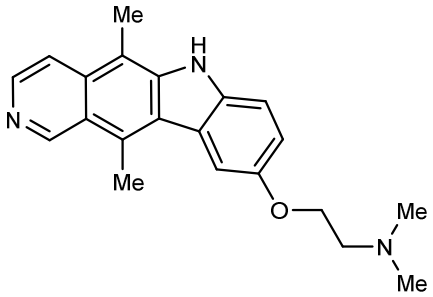
18



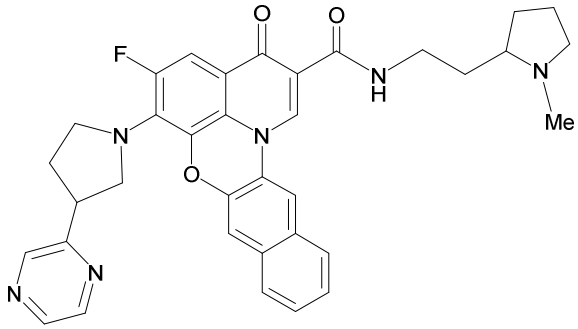
19



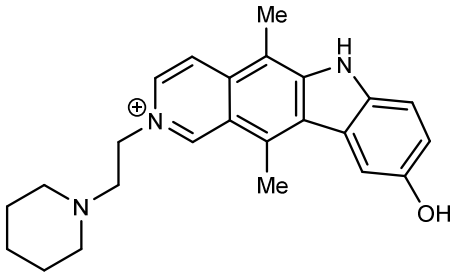
20



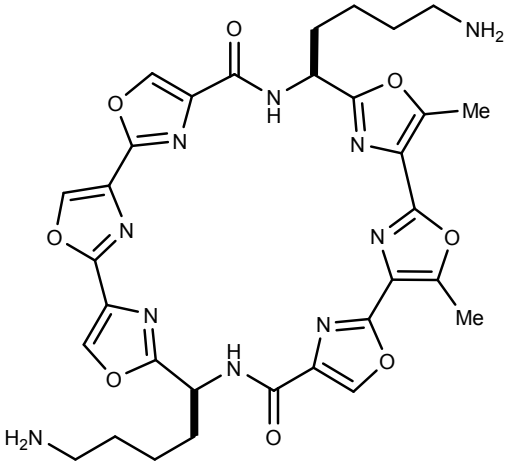
21



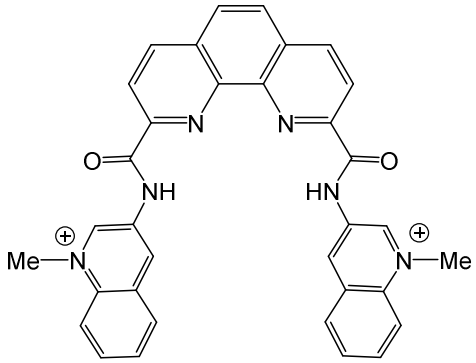
22



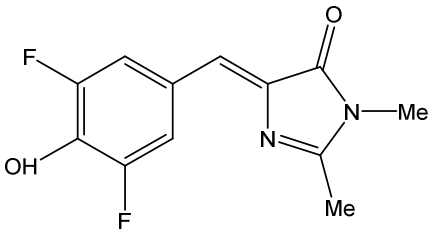
23



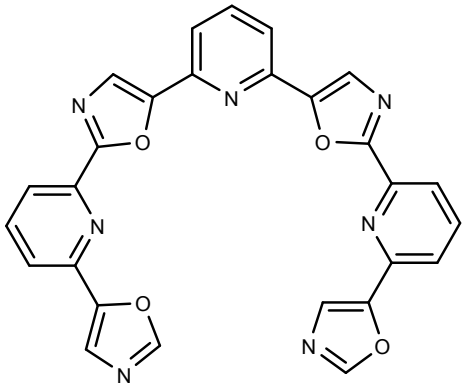
24



25



26



27

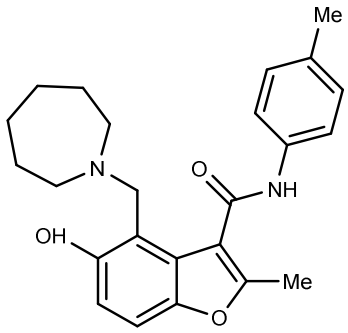


Table of Contents graphic