Perspective article, Nature Reviews Chemistry

Quadruplex nucleic acids as targets in cancer drug discovery

Stephen Neidle

UCL School of Pharmacy

University College London

London WC1N 1AX, UK

s.neidle@ucl.ac.uk

0044 207 7535969

Abstract

Quadruplex nucleic acids were once just laboratory curiosities. These helical four-stranded structures are known to form from nucleic acid sequences rich in guanine through Hoogsteen-type hydrogen bonding. They are, however, now emerging from this status to become significant targets for small-molecule drugs, which can stabilise the structure and thereby promote selective down-regulation of gene expression, telomerase inhibition and DNA damage at telomeres. The ability of most quadruplex-binding small molecules to stabilise a number of different cellular quadruplexes may well be an advantage. This can generate a poly-gene response and thus is able to simultaneously affect a number of key driver genes in a human cancer, to potential therapeutic benefit.

Introduction

Nucleic acids are the direct targets for the majority of cytotoxic drugs, which are still extensively used in the treatment of a wide range of human cancers in spite of their toxicity, propensity to induce drug resistance and very limited cellular selectivity. There have been attempts at moving beyond this “blunderbuss” approach, by for example targeting particular duplex DNA sequences in order to selectively down-regulate gene expression using designed compounds, notably polyamides. These in particular have been used to demonstrate cellular and *in vivo* activity related to their predicted on-target effects but these large molecules have not reached the stage of clinical assessment, in spite of the elegance of the chemistry involved[1]. By contrast, conjugates of highly cytotoxic compounds (such as DNA minor-groove covalent binders) with antibodies are showing considerable promise. A recent example is a conjugate of a pyrrolobenzodiazepine derivative with cell-surface antigens that are over-expressed in some haematological cancers[2]. Clinical trials of this conjugate in B-cell non-Hodgkin’s lymphoma and B-cell acute lymphoblastic leukaemia are currently underway. However antibody-containing approaches are still of unproven applicability for a number of major cancers such as pancreatic cancer, as well as being very costly, an increasingly important issue in many health systems.

This article discusses an alternative chemistry-based approach, involving the targeting of higher-order, non-duplex four-stranded DNA sequences. It has the potential ability to overcome the non-selectivity of traditional cytotoxic agents, and also has possible efficacy in a wide range of genetically complex solid tumour types as well as with haematological cancers. This approach is one that has been largely developed in academia, resulting in a large body of chemical and biological data in the public domain. If this approach leads to successful clinical candidate compounds, it might also lead to very considerable future cost advantages over the traditional industry route of discovery and development. This article will focus on the current status of the field in relation to its potential for anticancer therapy

Basics of quadruplex architecture

The study of four-stranded nucleic acids has come a long way since their original characterization, by a seminal fibre-diffraction study over 50 years ago, of the molecular structure formed from a gel of the aggregated mononucleotide

guanosine monophosphate[3], using the same methodology that was employed a decade earlier in the elucidation of the structure of the double helix itself. The remarkable insight of Gellert, Lipsett and Davies in this work was to suggest that the unusual stability of the guanosine monophosphate aggregate was due to the strong Hoogsteen hydrogen-bonding arrangement of four in-plane guanine bases, to form a guanine (G) quartet as the core repeating unit. This ensures the formation of a right-handed parallel-stranded four-fold helix, with the G-quartets replacing the base pairs in conventional duplex nucleic acids, but analogously stacking on one another. Gellert, Lipsett and Davies also suggested, in a second insight, that the cavity at the centre of each G-quartet was of sufficient size to accommodate a water molecule. Subsequent biophysical and structural studies have shown that a metal ion normally occupies this cavity (Figure 1), with a strong preference for potassium or sodium ions[4].

The potential biological significance of the four-fold helical arrangement was realised over 20 years later by the demonstration that analogous arrangements are also possible with short-length guanine-rich oligonucleotides[5] and in particular with the repetitive G-rich sequences found in telomeric DNA at the ends of eukaryotic chromosomes, in some immunoglobulin switch regions, and in certain gene promoter regions[6-8]. These non-duplex structures were termed quadruplexes and can be formed from parallel[9], anti-parallel[10] and mixed backbone orientations[11] – this suggested a wide diversity of possible quadruplex types, which has been confirmed and greatly extended by a large number of subsequent biophysical, chemical protection, NMR and X-ray crystallographic studies (see for example, Figure 2 and refs. 12-20 for a representative selection of structural studies in this area). Quadruplexes are potentially highly stable structures, especially in the high potassium ion environment (ca 100 mM) of cell nuclei.

A typical quadruplex comprises 3-4 stacked G-quartets (the core), held together by four nucleotide strands, which can be continuous, to form an intramolecular quadruplex. Quadruplexes can also be formed from two strands (bimolecular) or four (tetramolecular). Quadruplex arrangements of all types and from a variety of biological contexts, can be further stabilised by the binding of small molecules[21-23], which inhibits quadruplex unwinding to other nucleic acid species. This small-molecule intervention and stabilisation forms the basis of a wide range of

quadruplex-mediated biological effects, most of which have been studied for their potential as therapeutic approaches, principally in human cancer. A small number of quadruplex studies have been focussed in other therapeutic areas such as HIV[17].

Quadruplexes and human telomeric DNA

The quadruplexes formed by human telomeric DNA sequences, which have been especially well-studied, are formed from repeats of the telomeric repeating unit d(GGGTTA). Four consecutive units can form stable intramolecular quadruplexes, with a variety of quadruplex topologies observed in a number of X-ray[13] and NMR[10,12,14-20] structures (Figure 2). These are dependent on factors such as the nature of the metal ion and the 5' and 3' flanking sequences. A further critical factor in defining topology appears to be quadruplex sequence concentration, with NMR and circular dichroism studies consistently indicating the presence of mixed parallel/anti-parallel-type backbone orientation (hybrid) species in potassium ion-containing dilute solution[14,18]. The nature of the dominant species in concentrated solution (more relevant to cellular conditions) is still controversial, although there is increasing evidence that the all-parallel backbone form[13,20,24] is dominant, and may also be present in a cellular environment[25,26].

The terminal 100-200 nucleotides at the 3′ end of human telomeric DNA is single-stranded[27] so can in principle form quadruplexes more readily than duplex DNA, at least in the absence of associated proteins (primarily several copies of the Protection of Telomere protein hPOT1[28]) since this overhang DNA is unconstrained by a double-helical arrangement. Several consecutive quadruplexes can be formed along the length of the single-stranded end[29,30], although to date there is no detailed atomic-level experimental structural data on such an array of tandem quadruplex repeats, or indeed on the telomere itself. Various tandem quadruplex models are available, with individual quadruplex units having parallel[31] or hybrid-type topology[32]. These models may be useful in the absence of experimental structural data since the tandem telomeric quadruplex concept has implications for selective small-molecule binding to telomeres, which are discussed further in Box 1.

Suggestions of a biological role for telomeric quadruplexes have focussed on their possible formation at telomeres *in vivo* and include tethering chromatids together at meiosis and alignment of strand ends during recombination events.

However their potential for therapeutic use, which was first proposed in 1997[33], does not directly depend on these roles. Instead it exploits the function of the reverse-transcriptase enzyme telomerase, which is selectively expressed in the majority of human cancers, and not in normal somatic cells. Telomerase plays a central role in the initiation of tumorigenesis by catalysing the synthesis of telomeric DNA repeats onto the 3' end of chromosomes, thus maintaining them in an immortal state and hence directly contributes to cellular immortalisation.

Small molecules (initially demonstrated with a disubstituted amidoanthraquinone derivative[33]) can induce the 3' end single-stranded telomeric DNA sequence to fold into one or more consecutive quadruplexes and so stabilise them. This has the result of indirectly inhibiting telomerase activity[33], since the folded telomeric DNA 3′ end is required to be single-stranded in order to hybridise to the RNA sub-unit of telomerase at its active site. Accessibility to the catalytic domain of telomerase is necessary in order for nucleotide triphosphate substrate addition to the 3' end to build up further TTAGG repeats for telomere length maintenance, a key hallmark of cancer progression. Telomerase inhibition can thus reverse cellular immortalisation and result in anti-cancer activity, although more recent evidence has shown that the consequences of telomeric DNA quadruplex targeting can be more complex and result in DNA damage to susceptible cancer cells..

A wide range of small-molecule compounds have been shown to bind with high affinity to telomeric quadruplexes and to affect telomere function[21-23]. The trisubstituted acridine compound BRACO-19[34] (Figure 3a) for example, appears to inhibit telomerase function in cells and in tumour xenografts according to the classic telomeric quadruplex model[33]. It has anti-tumour activity *in vivo*[35], which may also be related to its ability to induce DNA damage at telomeres, for example in human glioblastoma cells[36]. This has led to the plausible hypothesis that the consequence of the displacement of telomere-associated proteins by quadruplex formation at telomere ends is to expose the ends and induce a cascade of DNA damage. This damage can be selectively non-repaired in cancer cells, which is ultimately a lethal event in these cells since they frequently have DNA repair deficiencies[37]. The DNA damage effect has been thoroughly explored with the structurally simple acridinium derivative RHPS4[38-40] and the bis(quinolinyl)pyridine-2,6-dicarboxamide derivative pyridostatin[41-43] (Figures 3b,c). RHPS4 shows *in vivo* anticancer activity in a number

of animal models; however it also shows significant off-target toxicity effects[44], mostly on the β2-adrenergic and hERG receptors, which are associated with adverse cardiovascular events. Subsequent analogue development[45] has resulted in an improved derivative with a -C≡C-CH_2NHAc substituent. This produces enhanced DNA damage together with reduced hERG inhibition. In spite of this apparent improvement, no RHPS4 analogue has progressed beyond this point to pre-clinical assessment and no plans for clinical trials have been announced.

Quadruplex-binding small molecules generally have three features in common:

- an extended heteroaromatic chromophore, which can π-π stack onto a planar G-quartet
- extended side-chains, which can extend into quadruplex grooves and loops, enhancing binding by van der Waals and electrostatic interactions
- cationic charge, which can be either at the terminus of the side-chains, or on the chromophore itself, enhancing its electron deficiency and hence strength of π-π contacts with a G-quartet

Quadruplexes in gene promoter regions

It has long been known that some gene promoter sequences contain potential quadruplex-forming sequences, especially at and close to sites of nuclease hypersensitivity (ie of enhanced ability for the two strands to dissociate under the influence of a suitable driving force). This has led to the fruitful concept that quadruplex induction and stabilisation at such sites could result in gene-specific repression of transcription[46]. The concept was originally demonstrated with the tetra-substituted porphyrin derivative TMPyP4, which binds strongly to a wide range of quadruplex nucleic acids, as well as to duplex and other nucleic acid forms. There has been much subsequent activity aimed at devising synthetic compounds and screening compound libraries as well as natural products capable of selectivity not just for quadruplexes in general, but for a particular quadruplex within the promoter of a given target gene[47,48].

Quadruplex prevalence in the human genome

This quest for selectivity has raised the wider question of quadruplex prevalence in genes, in the human genome, and ultimately in cells and tissue. The knowledge since 2003 of the complete human genome sequence has made it possible, at least in principle, to address this issue. Initial bioinformatics approaches[49,50] used the assumption that quadruplex sequences are of the type G_{3-5} X_n G_{3-5} X_o G_{3-5} X_p G_{3-5}, with loops X_n, X_o and X_p of general sequence and having 1-7 nucleotides. The sizes of the G-tracts were varied between three and five guanines in length, but were not necessarily confined to be of equal length. Both these studies independently resulted in the finding of ca 350,000 putative occurrences, with a significant number occurring within multiple G-tracts ("quadruplex islands"), where there is likely to be multiple overlapping quadruplex species. More recent predictions have suggested an even greater prevalence, in accordance with the view that quadruplexes with long (> 7 nucleotides) loops can also exist since such loops may have inherent secondary structure stability[51]. A high-throughput quadruplex sequencing approach[52] has provided an experimental view of quadruplex prevalence, with the unexpected finding of over 700,000 distinct structures, including their presence in some large genes where the bioinformatics methods had not predicted any occurrences. Both predictive and experimental approaches have found that quadruplex sequences are over-represented in promoter[53] and 5'untranslated[54] (UTR) regions of many genes.

However the presence in cells of such an overwhelming number of quadruplexes raises profound questions about the viability of cells given the known ability of quadruplexes to hinder or even halt the progression of transcription, translation or replication[55-58], depending on where quadruplexes are located within a gene or its transcribed RNA.

This apparent conundrum has been resolved, in the case of DNA quadruplexes, by a study[59] that has examined their actual prevalence in human chromatin (in an immortalised though non-oncogenic cell line), using an immunoprecipitation technique with a quadruplex-specific antibody, combined with RNA-seq sequencing. This has revealed a substantially lower number of sites (ca 10,000), than found by the earlier informatics studies. The sites are mostly in chromatin-depleted regulatory regions, especially in the promoters and 5'-UTR regions of cancer-associated genes.

The discrepancy between experiment and prediction may, it is suggested, be due to the challenge of maintaining quadruplex structures within highly packaged hetereochromatin. A further study[60], this time on the prevalence of RNA quadruplexes and using reverse transcriptase stop assays and chemical foot-printing, has concluded that the number of these quadruplexes found to be folded *in vivo* in mammalian cells is rather small, by contrast with the large numbers predicted and found *in vitro*, for example, in the 5'-UTR regions of many genes. The unwinding of RNA quadruplexes to single-stranded RNAs has been ascribed to the efficiency of cellular factors such as RNA helicases.

These findings, that only a relatively small number of DNA and RNA quadruplex sites are stable (i.e. are folded) *in vivo*, and are concentrated within cancer genes, thus provides a firm basis for the concept of quadruplexes being plausible cancer targets. Their actual prevalence *in vivo* in other genomes such as viral or parasitic ones, has yet to be revealed but may well also provide a basis for selective therapeutic targeting in these systems.

But do quadruplexes really exist in cells?

The presence of quadruplexes in the nuclei of the ciliate *Stylonychia lemnae* was first demonstrated[61] in 2001 using a specific high-affinity antibody to the telomeric repeat of this eukaryote [$G_4(T_4G_4)_4$], which forms parallel quadruplexes. More recently antibodies been raised that are useable in mammalian cells, and that have been validated for a range of both inter- and intramolecular quadruplex types (positive) and other categories of nucleic acid structure (negative). These have shown foci using immunofluorescent staining that have been interpreted in terms of the presence of DNA quadruplexes in the nuclei[62-64] and RNA quadruplexes in the cytoplasm[65] of mammalian cell types, albeit in fixed cells due to the inability of these antibodies to penetrate into cells. Elevated levels of quadruplex staining have also been found in human stomach and liver cancer tissues[66], further supporting a relationship between human cancers and quadruplexes. Quadruplex formation is increased[64] in cells treated with quadruplex-binding ligands, and also in cells lacking the FANCJ helicase. This supports the concept that normal cells can use integral helicase activity to effectively resolve (ie unwind) quadruplex structures[67], whereas mutated helicases are unable to perform this function.

The challenge of directly visualizing quadruplexes in live cells is being addressed by the use of quadruplex-binding small molecules with rapid cellular and nuclear uptake capability as well as appropriate fluorescent properties[68-71]. A further important question is whether antibody or small molecule can differentiate between different quadruplex types. There is limited evidence to date that this can be achieved, although the report[72] that a quadruplex antibody can be generated which is specific for parallel quadruplexes, suggests that such a goal is feasible.

Quadruplex selectivity: is it achievable and is it necessary?

The number of active genes containing at any one time folded quadruplexes depends on the cell type and on the state of the cell cycle. Evidence on this is as yet indirect, but suggests that the number is likely to be less than the 10,000 found in cellular chromatin[59]. Is it possible though to focus small-molecule binding to a single critical quadruplex-containing gene, and if so, what cancer-related gene should this be? A hallmark of many cancer cells is the central role played by a small number of driver genes, especially in the early stages of tumour progression. Driver gene products are frequently mutated, enabling particular pathways, notably those of growth regulation, to be dis-regulated. A prominent and well-studied example is the k-RAS gene, which is frequently found to be mutated, for example, in human non-small-cell lung, pancreatic and colorectal cancers. The k-RAS gene, which together with h-RAS and n-RAS, each contains a number of putative quadruplex promoter sites, and thus can provide suitable test systems[73,74]. Promoter quadruplexes have also been mapped in a number of other well-established cancer driver genes, including c-MYC (see Boxes 2, 3), c-KIT[75-77], BCL-2[78,79] and VEGF[80]. Molecular structures have also been established for these – by high-field NMR for c-KIT, BCL-2 and VEGF, and by X-ray crystallography for c-KIT[81,82], thus providing potential starting-points for structure-based ligand design (Box 3).

Quadruplex-small-molecule selectivity can be considered on several levels:

1. Between quadruplex and other types of nucleic acid structure, particularly duplex DNA. Quadruplex:duplex selectivity is of major importance in cells since significant duplex affinity suggests strong binding to a large number of genes and can be a prelude to general cellular toxicity. Compounds such as the tetrapyridyl porphyrin TMPyP4 have high quadruplex affinity (ca 2 x 10^6 M^{-1}) that

is close to its duplex affinity (8 x 10^5 M^{-1})[83]. It turns out to be facile to screen for quadruplex selectivity and a large number of compounds have been reported that have little or no duplex affinity, in some circumstances combined with enhanced quadruplex affinity (it is common practice to employ a human telomeric quadruplex for this comparison, using a high-throughput melting temperature assay). The ability of the N,N'-bis(quinolinyl)pyridine-2,6-dicarboxamide derivative pyridostatin to stabilise this quadruplex using such a melting assay have shown an increase in melting temperature (ΔT_m) of 35° C whereas no significant ΔT_m was observed with a duplex DNA sequence. Such experiments are especially meaningful when performed with a large excess of polymeric duplex DNA, to better simulate cellular conditions[84].

The overwhelming majority of quadruplex-binding small molecules bind to quadruplexes via quasi-external stacking, in which the almost universally-present heteroaromatic chromophore is π-π stacked onto an external G-quartet face and the side-chains positioned in the quadruplex grooves. Quadruplex selectivity can be explained, at least in part, by the difference between the large, highly accessible surface area of a terminal quartet compared to the much smaller, less accessible A:T/G:C base pair surface area of a typical duplex DNA intercalation site.

2. Selectivity between different quadruplexes is a goal of many studies in this field. However since only a very small proportion (< 1%) of the total number available in the human genome have been characterised at all, quadruplex selectivity appears to be currently achievable only in a very limited sense. Typically a study will screen a small-molecule library with a small (4-6 member) panel of the best-studied quadruplexes, and in some instances selectivity between several of them can be found. However it is rare for specificity for one particular quadruplex to be found. A typical example is the alkaloid chelerythrine, which has been found to bind to three promoter quadruplexes, from the BCL-2, k-RAS and VEGF genes with Kd values of 1.92, 0.72 and 0.51 x10^6 M^{-1} respectively[85].

The difficulty of selecting a singular quadruplex from even a small pool, let alone from a whole genome, can be understood in terms of the known architectures of many quadruplexes, as seen from NMR and crystallographic studies. All share the common feature of a core of stacked G-quartets, held

together by the four phosphodiester backbones, which create four grooves. Groove dimensions depend on the length and nature of the inter-G-tract loops, which themselves provide some potential anchoring and discrimination for different ligand substituents, but these differences are really secondary to the overall global similarity of most quadruplexes. Thus the overwhelming majority conform to the pattern of structural features found in the human telomeric quadruplexes.

When a loop sequence itself contains one or more guanosines then added topological complexity can occur, when the guanine base becomes inserted into a G-quartet creating additional grooves. Such an arrangement has been observed, for example, in one of the two c-KIT promoter quadruplexes, which has a characteristic deep cleft as a consequence of G insertion, suitable for small molecule binding[75,81,82].

3. Cellular selectivity is a desirable goal, as measured by significant differences in cell growth inhibition, cell viability or cell death following treatment of a cancer cell line with a quadruplex-binding ligand, compared to effects on a normal non-transformed cell line. It is at this point that attributing these effects, and therefore selectivity, to the consequences of promoter quadruplex binding, become challenging. Thus, although there are a large number of reports in the literature describing small molecules that bind with high affinity to promoter quadruplex in vitro and have high growth-inhibitory potency in cancer cells, remarkably few studies to date have unequivocally validated a relationship between cause and effect. (Differences in cellular uptake between normal and cancer cell lines may also be a significant factor and needs to be taken into account). Telomeric quadruplexes are exceptions to this since it is generally accepted that certain quadruplex-binding small-molecules can result in cellular telomerase inhibition as well as DNA double-strand breaks, although this depends on cell type, as well as on the nature of the small molecule. Several studies have shown that these are the major factors contributing to the selective lethality of such agents as RHSP4 in cancer cells and in experimental models of human cancer.

It is possible to assess whether a particular gene, and its promoter quadruplex, is indeed a target, by using appropriate assays, for example using isogenic pairs of cell lines in one of which there is a specific mutation in the

promoter quadruplex sequence. An important example[86] is the study of c-MYC targeting using a 11-piperazinylquindoline derivative, (2-(4-(10H-indolo[3,2-b]quinolin-11-yl)piperazin-1-yl)-N,N-dimethylethanamine, which binds effectively to the c-MYC quadruplex *in vitro* and induces down-regulation of c-MYC expression in cells, and also results in cell growth inhibition. An exon-specific assay exploiting particular translocation features of the CA46 Burkitt's lymphoma cell line was used to demonstrate that c-MYC quadruplex targeting is not directly involved in the observed c-MYC down-regulation by this compound[86]. Whether such indirect effects on "quadruplex target" genes occur, appears to depend on the nature of the small molecule involved, and on the cell type. So the same exon-specific assay has been used to show that the ellipticine derivative with a single dimethylethanamine side-chain (GQC-05: NSC338258), by contrast, has does directly down-regulate c-MYC expression[87].

4. The demonstration that *in vivo* effects of quadruplex-binding ligands are due to changes in expression of target quadruplexes, and are not just consequences of off-target effects leading to cell death, is even more challenging. However data is emerging that not only supports the multi-targeting concept but also provides some evidence of quadruplex involvement. The tetra-substituted perylene derivative EMICORON (Figure) produces telomeric DNA damage associated with displacement of the protein hPOT1 from the single-stranded end of telomeric DNA following quadruplex formation[88]. EMICORON also strongly stabilises c-MYC and BCL-2 promoter quadruplexes *in vitro* and treated cells show down-regulation of the expression of these genes[89]. The compound has antitumour activity against several *in vivo* models of colon cancer[90], suggesting that all of these quadruplex targets (and perhaps others) contribute to this.

Promoter quadruplex effects are highly unlikely to be restricted to a single gene

Rather one can expect to observe down-regulation of the expression of a number of critical genes. Thus has been hinted at by studies of effects in a cancer cell line using gene array libraries comprising small (ca 100) numbers of genes usually in a particular pathway or linked by common function. Coincidentally two such studies[91,92] have used substituted naphthalene diimide compounds; the tetra-substituted one (MM41) was examined in the MIA-PACA2 pancreatic cancer cell line and was subsequently shown to have anti-cancer activity in a xenograft model derived from

this cell line[93]. The trisubstituted naphthalene diimide produced down-regulation of several genes in human melanoma and lung cancer cell lines, including a telomerase gene (hTERT) and the BCL-2 gene[92]. MM41 down-regulated the expression of several well-studied cancer genes such as BCL-2[91,93], although other lesser-known ones were also affected. Pyridostatin produces DNA damage both at telomeres and at several gene promoter sites, notably at a large quadruplex island in the SRC gene[42].

Quadruplex-binding small molecules as anticancer agents: progress to date

Several small molecule compounds have been shown to produce anticancer effects in animal models of human cancers. Are these effects a direct consequence of quadruplex binding? There is a strong case for telomeric DNA-damaging effects, such as those caused by, for example, the acridines RHPS4 and BRACO19, and the polycyclic perylene derivative EMICORON (Figure 3d) although as has been outlined above, effects of promoter quadruplex binding have also been observed, as they have been for several substituted naphthalene diimides.

However such studies only provide a small window on the totality of possible expression changes in all genes encoded in the human genome, comparing normal and cancer cell lines. Since most quadruplex-selective small molecules are likely to bind to multiple quadruplex promoter sites, it will be unsurprising to see the significant down-regulation of considerable numbers of genes. Since it is now established that quadruplex DNA occurrence is a hallmark of the cancer genome[59], the simultaneous poly-targeting of multiple quadruplex sites looks to be a viable, and even advantageous therapeutic strategy. It potentially offers some advantage over the traditional targeting approach to a single gene/gene product leading to down-regulation of a single pathway, especially since it is now apparent that the genetic heterogeneity and complexity of the majority of solid tumours differ not only between individuals but also varies within a single tumour[94,95]. Tumour sensitivity to quadruplex-targeting small molecules that induce DNA damage can be enhanced by combination with DNA damage repair inhibiting agents[40,96]. Related to this has been the observation that cancer cells deficient in the DNA repair proteins BRCA1 and BRCA2 are especially sensitive to quadruplex-binding small molecules – BRCA1/2 mutations occur in a small proportion of hard-to-treat breast, ovarian and pancreatic cancers[97].

One quadruplex-binding small molecule, quarfloxin[98] (CX3543: Figure 4a), designed as such by Hurley and colleagues, and originally believed to target RNA quadruplexes, has been evaluated in phase 1 and phase 2 clinical trials in human cancer. No adverse side-effects associated with quadruplex binding have been reported. A second quadruplex ligand, CX5461 (Figure 4b), has recently entered early-phase clinical trials in BRCA1/2 deficient breast cancer patients. Its behaviour is analogous to CX3543 in that both compounds induce DNA damage in repair-deficient susceptible cells and CX5461 also produces profound anticancer activity in BRCA-deficient tumours[99]. Again, the lack of toxicity that could be ascribed to quadruplex target engagement is a positive indicator for the future clinical development of these and other agents of this general class.

Finally we return to the theme outlined at the outset of this article, the challenge of the ever-increasing cost of new cancer medicines, which is threatening to put promising ones out of the reach of the health budgets of an increasing number of health systems world-wide, especially as the incidence of cancer is globally increasing in line with increased longevity. The quadruplex-small-molecule story is one that to date has largely been developed in academia and this trend is set to continue into clinical development of promising compounds, provided there is continuity in significant support from funding agencies. The relatively low cost of this approach to date compared to traditional industry-led ones, will hopefully lead to a reduced cost for any licensed medicines[100].

Box 1

Small molecule selectivity for telomeric quadruplexes – targeting tandem quadruplexes

The length of the single-stranded overhang at the 3' end of human telomeric DNA (100-200 nucleotides) is sufficient for the formation of several quadruplexes in tandem[101]. Cartoon representations of two structural models for these tandem quadruplexes are shown here, based on crystallographic and NMR studies of the isolated quadruplex monomers. a. derived from the parallel-stranded quadruplex crystal structure[13]. b based on the hybrid loop NMR structures[14,18]. Both models have potential small molecule binding sites indicated by red arrows.

Several small molecule chemotypes have been reported to bind with high selectivity for tandem over isolated telomeric quadruplexes. The hybrid oxazole-triazole ligand DR4-47 shows preferential thermal stabilisation of a telomeric quadruplex dimer, and even more for the trimer, compared to its monomer binding[102]. This is possibly a consequence of effective ligand stacking to the 3' quartet of one quadruplex and to the 5' quartet of the next. Another approach, typified by the high affinity shown by a di-nickel salphen complex, with the compound shown here having 30-fold greater affinity for a quadruplex dimer compared to the monomer[103], whereas the mono-nickel salphen[104] has 5-fold selectivity. It is suggested that the dimer is able to stack on two adjacent quadruplexes, with the length of spacer being critical.

Box 2

The MYC gene as a paradigm for promoter quadruplex targeting

c-MYC is a proto-oncogene with transcription factor activity whose unregulated expression, especially when mutated, is centrally involved in the progression of a large number of human cancers. It has long been considered to be a major target for therapeutic intervention although its flexibility and lack of appropriate surface binding pockets have led to the traditional view that the MYC protein is "undruggable."

The concept that a c-MYC promoter quadruplex can circumvent this problem is derived from the region upstream of the transcription start site in the c-MYC gene. This includes two promoter sequences and one of them, the P1 promoter, contains a nuclease hypersensitive element (NHE) III1 with a 27-nucleotide sequence comprising five short G-tracts:

5´-TGGGGAGGGTGGGGAGGGTGGGGAAGG-3´

The realisation that this is a quadruplex-forming sequence has led to a large number of studies aimed at inhibiting MYC function at the gene rather than the protein level, involving the quadruplex concept. The classic study[46] on c-MYC quadruplex targeting used the porphyrin derivative TMPyP4 and demonstrated that c-MYC expression was down-regulated by this ligand, which binds tightly (albeit non-selectively) to this quadruplex region. The dominant quadruplex in potassium ion-containing solution has been subsequently identified as that formed from G- tracts 2-5. The solution structures of both this and the alterative quadruplex formed from G-tracts 1-4 have been determined by NMR methods, and several NMR structures are available of complexes with quadruplex-selective small molecules (Box 3). A large number of such compounds have also been evaluated for their effects on the stability of c-MYC (and subsequently other) promoter quadruplexes. Typically small libraries of related compounds are screened against (small) panels of quadruplex sequences under conditions in which they are either already fully folded or where binding can induce quadruplex formation. Changes in thermal melting behaviour are most commonly used as an initial assessment, measuring the ability of a compound to stabilise a quadruplex. This can be undertaken in a high-throughput mode in contrast to more quantitative studies of binding, for example using surface plasmon resonance.

c-MYC selectivity over human telomeric quadruplexes has been observed, for example at a ten-fold level for members of a series of indoloquinolines having a single aminoalkylamino side-chain, although some duplex affinity was detectable[105].

An innovative and fruitful approach using a library of 20,000 compounds with a fluorescent-tagged c-MYC quadruplex has resulted in the identification of the compound below[106]. Its effects on selectively down-regulating c-MYC expression have been validated using a cell line with a MYC translocation and by examination of several other genes, which it does not affect.

Box 3

The design and optimisation of small molecules to target a given quadruplex can be facilitated by data from NMR or crystallographic studies on complexes.

Various features can be systematically optimised:

- interactions with the terminal G-quartet, by changing the nature and size of the chromophore
- interactions with grooves and loops, by changing interacting groups and their ability to, for example, fill the available space and enhance electrostatic/van der Waals interactions

a shows the NMR-derived structure[108] of a quindoline compound (b) bound to a c-MYC quadruplex, one to each terminal G-quartet. The ligand is shown in space-filling representation, which indicates that there is room in the appropriate quadruplex groove for the amino-alkylamino side-chain could be extended in size and steric bulk, albeit at the cost of greater molecular weight and possibly reduced cellular uptake.

c, d, e shows the NMR structure of a complex between a synthetic analogue of the macrocyclic natural product telomestatin, and a (3+1) hybrid human telomeric quadruplex[109]. The macrocycle effectively stacks on to a terminal G-quartet, as shown by the extensive overlap between guanines and oxazole rings. The two alkylamino side-chains impart aqueous solubility to the macrocycle, but also effectively fill two quadruplex grooves

f, g,h show a contrasting situation from an X-ray crystallographic study[91] with the tetra-substituted naphthalene diimide compound MM41 bound to a parallel human telomeric quadruplex (shown in solvent-accessible surface representation). Water molecules are shown as mauve spheres. The figures show one quadruplex groove with a MM41 side-chain which only partially fills the groove. Instead, water molecules fill the remaining space and serve to mediate contacts between side-chain and groove walls/floor.

References

1. See for example: Hargrove, A. E. *et al*. Tumor repression of VCaP xenografts by a pyrrole-imidazole polyamide. *PLoS One* **10**, e0143161 (2015).
2. Flynn, M. J. *et al*. ADCT-301, a pyrrolobenzodiazepine (PBD) dimer-containing antibody-drug conjugate (ADC) targeting CD25-expressing hematological malignancies. *Mol. Cancer Ther*. **15**, 2709-2721 (2016).
3. Gellert, M., Lipsett, M. N. & Davies, D. R. Helix formation by guanylic acid. *Proc. Natl. Acad. Sci. USA* **48**, 2013-2018 (1962).
4. Bhattacharyya, D., Mirihana Arachchilage, G. & Basu, S. Metal cations in G-quadruplex folding and stability. *Front. Chem*. **4**, 38 (2016).
5. 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. *Cell* **51**, 899-908 (1987).
6. Sen, D. & Gilbert, W. Formation of parallel four-stranded complexes by guanine-rich motifs in DNA and its implications for meiosis. *Nature* **334**, 364-366 (1988).
7. Sundquist, W. I. & Klug, A. Telomeric DNA dimerizes by formation of guanine tetrads between hairpin loops. *Nature* **342**, 825-829 (1989).
8. Simonsson, T., Pecinka, P. & Kubista, M. DNA tetraplex formation in the control region of c-myc. *Nucleic Acids Res*. **26**, 1167-1172 (1998).
9. Laughlan, G. *et al*. The high-resolution crystal structure of a parallel-stranded guanine tetraplex. *Science* **265**, 520-524 (1994).
10. Wang, Y. & Patel, D. J. Solution structure of the human telomeric repeat d[$AG_3(T_2AG_3)_3$] G-tetraplex. *Structure* **1**, 263-282 (1993).
11. Burge, S., Parkinson, G. N., Hazel, P., Todd, A. K. & Neidle, S. Quadruplex DNA: sequence, topology and structure. *Nucleic Acids Res*. **34**, 5402-5415 (2006).
12. Adrian, M., Heddi, B. & Phan, A. T. NMR spectroscopy of G-quadruplexes. *Methods* **57**, 11-24 (2012).
13. Parkinson, G. N., Lee, M. P. & Neidle, S. Crystal structure of parallel quadruplexes from human telomeric DNA. *Nature* **417**, 876-880 (2002).

14. Luu, K. N. *et al.* Structure of the human telomere in K^+ solution: an intramolecular (3 + 1) G-quadruplex scaffold. *J. Am. Chem. Soc.* **128**, 9963-9970 (2006).
15. Brčić, J. & Plavec, J. Solution structure of a DNA quadruplex containing ALS and FTD related GGGGCC repeat stabilized by 8-bromodeoxyguanosine substitution. *Nucleic Acids Res*. **43**, 8590-8600 (2015).
16. Vasilyev, N. *et al*. 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* **112**, 5391-5400 (2015).
17. De Nicola, B. *et al*. Structure and possible function of a G-quadruplex in the long terminal repeat of the proviral HIV-1 genome. *Nucleic Acids Res*. **44**, 6442-6451 (2016).
18. 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.* **35**, 4927-4740 (2007).
19. 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. Am. Chem. Soc*. **131**, 4301-4309 (2009).
20. Heddi, B. & Phan, A. T. Structure of human telomeric DNA in crowded solution. *J. Am. Chem. Soc.* **133**, 9824-9833 (2013).
21. Monchaud, D. & Teulade-Fichou, M. P. A hitchhiker's guide to G-quadruplex ligands. *Org. Biomol. Chem.* **6**, 627-636 (2008).
22. Müller, S. & Rodriguez, R. G-quadruplex interacting small molecules and drugs: from bench towards bedside. *Expert. Rev. Clin. Pharmacol*. **7**, 663-679 (2014).
23. Islam, M. K., Jackson, P. J., Rahman, K. M. & Thurston, D. E. Recent advances in targeting the telomeric G-quadruplex DNA sequence with small molecules as a strategy for anticancer therapies. *Future Med. Chem*. **8**, 1259-1290 (2016).
24. Kejnovská, I., Vorlíčková, M., Brázdová, M. & Sagi, J. Stability of human telomere quadruplexes at high DNA concentrations. *Biopolymers* **101**, 428-438 (2014).

25. Liu, H. Y. *et al*. Conformation selective antibody enables genome profiling and leads to discovery of parallel G-quadruplex in human telomeres. *Cell Chem. Biol*. **23**, 1261-1270 (2016).
26. Moye, A. L. et al. Telomeric G-quadruplexes are a substrate and site of localization for human telomerase. *Nature Commun*. **6**, 7643 (2015).
27. 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*. **11**, 2801-2809 (1997).
28. 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* **102**, 10864-10869 (2005).
29. Yu, H. Q., Miyoshi, D. & Sugimoto, N. Characterization of structure and stability of long telomeric DNA G-quadruplexes. *J. Am. Chem. Soc*. **128**, 15461-15468 (2006).
30. Bugaut, A. & Alberti, P. Understanding the stability of DNA G-quadruplex units in long human telomeric strands. *Biochimie* **113**, 125-133 (2015).
31. Haider, S. M. & Neidle, S. A molecular model for drug binding to tandem repeats of telomeric G-quadruplexes. *Biochem. Soc. Trans*. **37**, 583-588 (2009).
32. Petraccone, L. *et al*. Structure and stability of higher-order human telomeric quadruplexes. *J. Am. Chem. Soc.* **133**, 20951-20961 (2011).
33. Sun, D. *et al.* Inhibition of human telomerase by a G-quadruplex-interactive compound. *J. Med. Chem.* **40**, 2113-2116 (1997).
34. Campbell, N. H., Parkinson, G. N., Reszka, A. P. & Neidle, S. Structural basis of DNA quadruplex recognition by an acridine drug. *J. Am. Chem. Soc.* **130**, 6722-6724 (2008).
35. Burger, A. M. *et al.* The G-quadruplex-interactive molecule BRACO-19 inhibits tumor growth, consistent with telomere targeting and interference with telomerase function. *Cancer Res.* **65**, 1489-1496 (2005).
36. Zhou, G. *et al*. Telomere targeting with a novel G-quadruplex-interactive ligand BRACO-19 induces T-loop disassembly and telomerase displacement in human glioblastoma cells. *Oncotarget* **7**, 14925-14939 (2016).
37. Brown, J. S., O'Carrigan, B., Jackson, S. P. & Yap, T. A. Targeting DNA repair in cancer: beyond PARP inhibitors. *Cancer Discov*. **7**, 20-37 (2017).

38. Cookson, J. C. *et al.* Pharmacodynamics of the G-quadruplex-stabilizing telomerase inhibitor 3,11-difluoro-6,8,13-trimethyl-8*H*-quino[4,3,2-*kl*] acridinium methosulfate (RHPS4) *in vitro*: activity in human tumor cells correlates with telomere length and can be enhanced, or antagonized, with cytotoxic agents. *Mol. Pharmacol.* **68**, 1551-1558 (2005).
39. Salvati, E. *et al.* Telomere damage induced by the G-quadruplex ligand RHPS4 has an antitumor effect. *J. Clin. Invest.* **117**, 3236-3247 (2007).
40. Salvati, E. *et al*. PARP1 is activated at telomeres upon G4 stabilization: possible target for telomere-based therapy. *Oncogene* **29**, 6280-6293 (2010).
41. Rodriguez, R. *et al.* A novel small molecule that alters shelterin integrity and triggers a DNA-damage response at telomeres. *J. Am. Chem. Soc.* **130**, 15758-15759 (2008).
42. Rodriguez, R. *et al*. Small-molecule-induced DNA damage identifies alternative DNA structures in human genes. *Nature Chem. Biol*. **8**, 301-310 (2012).
43. Müller, S. *et al*. Pyridostatin analogues promote telomere dysfunction and long-term growth inhibition in human cancer cells. *Org. Biomol. Chem*. **10**, 6537-6546 (2012).
44. Iachettini, S. *et al*. On and off-target effects of telomere uncapping G-quadruplex selective ligands based on pentacyclic acridinium salts. *J. Exp. Clin. Cancer Res*. **32**, 68 (2013).
45. Rizzo, A. *et al*. Identification of novel RHPS4-derivative ligands with improved toxicological profiles and telomere-targeting activities. *J. Exp. Clin. Cancer Res*. **33**, 81 (2014).
46. 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* **99**, 11593-11598 (2002).
47. Balasubramanian, S., Hurley, L. H. & Neidle, S. Targeting G-quadruplexes in gene promoters: a novel anticancer strategy? *Nature Rev. Drug. Discovery* **10**, 261-275 (2011).
48. Rigo, R., Palumbo, M. & Sissi, C. G-quadruplexes in human promoters: A challenge for therapeutic applications. *Biochim. Biophys. Acta* **4165**, 30518-30519 (2016).

49. Todd, A. K., Johnston, M. & Neidle, S. Highly prevalent putative quadruplex sequence motifs in human DNA. *Nucleic Acids Res.* **33**, 2901-2907 (2005).
50. Huppert, J. L. & Balasubramanian, S. Prevalence of quadruplexes in the human genome. *Nucleic Acids Res.* **33**, 2908-2916 (2005).
51. Bedrat, A., Lacroix, L. & Mergny, J.-L. Re-evaluation of G-quadruplex propensity with G4Hunter. *Nucleic Acids Res*. **44**, 1746-1759 (2016).
52. 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.
53. Huppert, J. L. & Balasubramanian, S. G-quadruplexes in promoters throughout the human genome. *Nucleic Acids Res.* **35**, 406-413 (2007).
54. Bugaut, A. & Balasubramanian, S. 5'-UTR RNA G-quadruplexes: translation regulation and targeting. *Nucleic Acids Res*. **40**, 4727-4741 (2012).
55. Maizels, N. & Gray, L. T. The G4 genome. *PLoS Genet*. **9**, e1003468 (2013).
56. Murat, P. & Balasubramanian, S. Existence and consequences of G-quadruplex structures in DNA. *Curr. Opin. Genet. Dev*. **25**, 22-29 (2014).
57. Rhodes, D. & Lipps, H. J. G-quadruplexes and their regulatory roles in cancer. *Nucleic Acids Res*. **43**, 8627-8637 (2015).
58. Bochman, M. L., Paeschke, K. & Zakian, V. A. DNA secondary structures: stability and function of G-quadruplex structures. *Nature Rev. Genetics* **13**, 770-780 (2012).
59. Hänsel-Hertsch, R. *et al*. G-quadruplex structures mark human regulatory chromatin. *Nature Genet*. **48**, 1267-1272 (2016).
60. Guo, J. U. & Bartel, D. P. RNA G-quadruplexes are globally unfolded in eukaryotic cells and depleted in bacteria. *Science* **353**, (2016).
61. Schaffitzel, C., Berger, I., Postberg, J., Hanes, J., Lipps, H. J. & Plückthun, A. *In vitro* generated antibodies specific for telomeric guanine-quadruplex DNA react with *Stylonychia lemnae* macronuclei. *Proc. Natl. Acad. Sci. USA* **98**, 8572-8577 (2001).
62. Lam E.Y., Beraldi D., Tannahill D. & Balasubramanian S. G-quadruplex structures are stable and detectable in human genomic DNA. *Nature Commun*. **4**, 1796-1780 (2013).

63. Biffi, G., Tannahill, D., McCafferty, J. & Balasubramanian, S. Quantitative visualization of DNA G-quadruplex structures in human cells. *Nature Chem*. **5**, 182-186 (2013).
64. Henderson, A. *et al*. Detection of G-quadruplex DNA in mammalian cells. *Nucleic Acids Res*. **42**, 860-869 (2013).
65. 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*. **6**, 75-80 (2014).
66. 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* **9**, e102711 (2014).
67. Mendoza, O., Bourdoncle, A., Boulé, J.-B., Brosh, R. M. Jr. & Mergny, J.-L. G-quadruplexes and helicases. *Nucleic Acids Res*. **44**, 1989-2006 (2016).
68. Tseng, T.-Y. *et al*. Fluorescent probe for visualizing guanine-quadruplex DNA by fluorescence lifetime imaging microscopy. *J. Biomed. Opt*. **18**, 101309 (2013).
69. Laguerre, A. *et al*. Visualization of RNA-quadruplexes in live cells. *J. Am. Chem. Soc*. **137**, 8521-8525 (2015).
70. Shivalingam, A. *et al*. The interactions between a small molecule and G-quadruplexes are visualized by fluorescence lifetime imaging microscopy. *Nature Commun*. **6**, 8178 (2015).
71. Doria, F., Nadai, M., Zuffo, M., Perrone, R., Freccero. M. & Richter, S. N. A red-NIR fluorescent dye detecting nuclear DNA G-quadruplexes: in vitro analysis and cell imaging. *Chem. Commun*. **53**, 2268-2271 (2017).
72. Liu, H. Y. *et al*. conformation selective antibody enables genome profiling and leads to discovery of parallel G-quadruplex in human telomeres. *Cell Chem. Biol*. **23**, 1261-1270 (2016).
73. Cogoi, S. & Xodo, L. E. G4 DNA in ras genes and its potential in cancer therapy. *Biochim. Biophys. Acta* **1859**, 663-674 (2016).
74. Paulo, A. & Francisco, A. P. Oncogene expression modulation in cancer cell lines by DNA G-quadruplex-interactive small molecules. *Curr. Med. Chem*. **23**, (2016), in press. DOI: 10.2174/0929867323666160829145055.

75. 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. Am. Chem. Soc*. **129**, 4386-4392 (2007).
76. Hsu, S. T., Varnai, P., Bugaut, A., Reszka, A. P., Neidle, S. & Balasubramanian, S. A G-rich sequence within the c-kit oncogene promoter forms a parallel G-quadruplex having asymmetric G-tetrad dynamics. *J. Am. Chem. Soc*. **131**, 13399-13409 (2009).
77. Kuryavyi, V, Phan, A. T. & Patel, D. J. Solution structures of all parallel-stranded monomeric and dimeric G-quadruplex scaffolds of the human c-kit2 promoter. *Nucleic Acids Res.* **38**, 6757-6773 (2010).
78. 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*. **34**, 5133-5144 (2006).
79. 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. Am. Chem. Soc*. **136**, 1750-1753 (2014).
80. 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*. **41**, 10584-10592 (2013).
81. 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*. **40**, 4691-4700 (2012).
82. Wei, D., Husby, J. & Neidle, S. Flexibility and structural conservation in a c-KIT G-quadruplex. *Nucleic Acids Res*. **43**, 629-644 (2015).
83. Martino, L., Pagano, B., Fotticchia, I., Neidle, S. & Giancola, C. Shedding light on the interaction between TMPyP4 and human telomeric quadruplexes. *J. Phys. Chem*. B. **113**, 14779-14786 (2009).
84. Guédin, A., Lacroix, L. & Mergny, J.-L. Thermal melting studies of ligand DNA interactions. *Methods Mol. Biol*. **613**, 25-35 (2010).

85. Jana, J. *et al*. Chelerythrine down regulates expression of VEGFA, BCL2 and KRAS by arresting G-quadruplex structures at their promoter regions. *Scientific Rep*. **7**, 40706 (2017).
86. Boddupally, P. V. *et al*. 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.* **55**, 6076-6086 (2012).
87. 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*. **286**, 41018-41027 (2011).
88. Franceschin, M. *et al*. Aromatic core extension in the series of N-cyclic bay-substituted perylene G-quadruplex ligands: increased telomere damage, antitumor activity, and strong selectivity for neoplastic over healthy cells. *ChemMedChem* **7**, 2144-2154 (2012).
89. Porru, M. *et al*. Targeting G-quadruplex DNA structures by EMICORON has a strong antitumor efficacy against advanced models of human colon cancer. *Mol. Cancer Ther*. **14**, 2541-2551 (2015).
90. Micheli, E. *et al*. Perylene and coronene derivatives binding to G-rich promoter oncogene sequences efficiently reduce their expression in cancer cells. *Biochimie* **125**, 223-231 (2016).
91. Micco, M. *et al*. Structure-based design and evaluation of naphthalene diimide G-quadruplex ligands as telomere targeting agents in pancreatic cancer cells. *J. Med. Chem*. **56**, 2959-2974 (2013).
92. Nadai, M. *et al*. Assessment of gene promoter G-quadruplex binding and modulation by a naphthalene diimide derivative in tumor cells. *Int. J. Oncol*. **46**, 369-380 (2015).
93. Ohnmacht, S. A. *et al*. A G-quadruplex-binding compound showing anti-tumour activity in an *in vivo* model for pancreatic cancer. *Scientific Rep*. **5**, 11385 (2015).
94. Alizadeh, A. A. *et al*. (2015). Toward understanding and exploiting tumor heterogeneity. *Nature Medicine* **21**, 846-853 (2015).

95. Jamal-Hanjani, M., Quezada, S. A., Larkin, J. & Swanton C. Translational implications of tumor heterogeneity. *Clin. Cancer Res*. **21**, 1258-1266 (2015).
96. Hänsel-Hertsch, R., Di Antonio, M. & Balasubramanian, S. DNA G-quadruplexes in the human genome: detection, functions and therapeutic potential. *Nature Rev. Mol. Cell Biol*. (2017) doi:10.1038/nrm.2017.3
97. Zimmer, J. *et al*. Targeting BRCA1 and BRCA2 Deficiencies with G-Quadruplex-Interacting Compounds. *Mol. Cell* **61**, 449-460 (2016).
98. Drygin, D. *et al*. Anticancer activity of CX-3543: a direct inhibitor of rRNA biogenesis. *Cancer Res*. **69**, 7653-7661 (2009).
99. Xu, H. *et al*. CX-5461 is a DNA G-quadruplex stabilizer with selective lethality in BRCA1/2 deficient tumours. *Nature Commun*. **8**, 14432 (2017).
100. Workman, P., Draetta, G. F., Schellens, J. H. & Bernards, R. How Much Longer Will We Put Up With $100,000 Cancer Drugs? *Cell* **168**, 579-583 (2017).
101. 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. Am. Chem. Soc*. **134**, 20060-20069 (2012).
102. Cousins, A. R. *et al*. Ligand selectivity in stabilising tandem parallel folded G-quadruplex motifs in human telomeric DNA sequences. *Chem. Commun*. **50**, 15202-15205 (2014).
103. Vilar, R. *et al*. Di-nickel-salphen complexes as binders of human telomeric dimeric G-quadruplexes. *Chemistry* (2017) doi: 10.1002/chem.201700276.
104. 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. Am. Chem. Soc*. **128**, 5992-5993 (2006).
105. Funke, A., Dickerhoff, J. & Weisz, K. Towards the development of structure-selective G-quadruplex-binding indolo[3,2-b]quinolines. *Chemistry* **22**, 3170-3181 (2016).
106. Felsenstein, K. M. *et al*. Small molecule microarrays enable the identification of a selective, quadruplex-binding inhibitor of MYC expression. *ACS Chem. Biol*. **11**, 139-48 (2016).
107. 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. Am. Chem. Soc*. **133**, 17673−17680 (2011).

108. 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. Am. Chem. Soc*. **135**, 13495−13501 (2013).

Figure legends

Figure 1 The G-quartet, showing four in-plane guanine bases linked by Hoogsteen hydrogen bonds, shown as dashed lines. Each guanine base participates in four hydrogen bonds, two involving donors from one face and two with acceptors on the adjacent face. The central metal ion (with K^+ being the most strongly bound) is normally coordinated to eight O6 atoms from two adjacent quartets, in a bipyramidal antiprismatic arrangement.

Figure 2 Cartoon representation of three distinct topologies observed for intramolecular telomeric quadruplexes. Arrows indicate strand polarity and rectangles show the guanine bases. a. the antiparallel form found in Na^+ solution by NMR[10], b. the all-parallel form observed in a crystal structure of the K^+ form[13], c. a hybrid structure found by NMR in K^+ solution[14].

Figure 3 Structures of selected quadruplex-binding ligands

Figure 4 Structures of clinical quadruplex-binding compounds