Review article

G-Quadruplexes: Emerging Targets for the Structure-Based Design of Potential Anti-Cancer and Antiviral Therapies

Petar M. Mitrasinovic*

Center for Biophysical and Chemical Research, Belgrade Institute of Science and Technology, 11060 Belgrade, Serbia

** Corresponding author: E-mail: pmitrasinovic.ist-belgrade.edu.rs@tech-center.com*

Received: 01-10-2020

Abstract

G-quadruplexes (G4s) are noncanonical secondary structures that fold within guanine (G) rich strands of regulatory genomic regions. Recent evidences suggest their intimate involvement in important biological processes such as telomere maintenance, end-capping and protection, chromosome stability, gene expression, viral integration, and recombination. Mechanistic details of how and why G4 structures influence biological function indicate a rationale for treating G4s as emerging molecular targets for future therapeutics. In other words, the structural heterogeneity with well-defined binding sites, thermal stability and abundance of G4s in telomeres, oncogene promoter regions, and viral genomes make G4s attractive targets for small molecules, aimed to selectively recognize them over all other nucleic acids structures, particularly duplex forms that are most abundant in the genome. Herein, a critical survey of well-characterized G4-interactive ligands as potential tools in anti-cancer and antiviral therapies is presented. Effects that these ligands selectively exert *in vitro* and *in vivo* models are summarized. Unique ligands involved in specific G4 recognition are put forward. A key question, how to design and develop new G4 specific ligands that conform to the structural and physicochemical requirements for optimal biological activity, is discussed by considering both remarkable advances over the last few years and our recent contributions.

Keywords: Anti-cancer and antiviral therapies, gene expression, G-quadruplex, ligand, structure-based drug design, target

1. Introduction

Even though nucleic acids structures are usually imagined as a double-helical DNA that is most abundant in the genome and plays a crucial role in genetic information storage,[1] only 3% of the human genome is expressed in proteins.[2] Nucleic acids are essentially dynamic structures that influence important biological functions.[3–5] Besides folding into canonical duplex structures, single-stranded DNA may form various noncanonical structures, such as hairpin, triplex, G-quadruplex (or G-tetraplex), and i-motif structures. G-tetrad structure, defined by four Hoogsteen G-G base pairs (Figure 1a), was firstly noticed in 1910[6] and identified about fifty years later.[7] G4s fold within G-rich tracts and consist of two or more stacked G-tetrads, being selectively stabilized by centrally coordinated potassium ions to O6 of the guanines at concentrations (10–50 mM) that are substantially below the 120 mM of KCl observed in most cells types.[8–11] Stabilizing preference for monovalent cations follows the order $K^+ > Na^+ > Li^+$.[12] The intracellular monovalent cation concentration and the localized ion concentrations determine the formation of G4s and can potentially dictate their regulatory roles.[13,14] G4s can be assembled in an intramolecular (backfolded) way or from two-, three-, or four DNA strands in intermolecular structures (Figure 1b) able to adopt a large diversity of conformations and folding energies.[8] Most intramolecular G4 structures that are deposited in the public domain have been determined by nuclear magnetic resonance spectroscopy in solution.[1,15] G4s are more compact structures than duplex DNAs and contain well-defined binding sites for selective recognition by small molecules.

The presence and function of G4s in vivo are not quite clear.[8] While consensus sequence for G4 folding is not experimentally established, approximately 370,000 G-rich sequences that contain putative G4-forming motif (PQS) are present in the human genome,[8,16,17] dispersed throughout regulatory genomic regions (human telomeres, oncogene promoter regions, immunoglobulin

switch regions, ribosomal DNA)[18–21] and some regions of RNA.[22,23] Because of the self-complementary nature of duplex DNA, approximately the same number of cytosine (C)-rich motifs is present in the human genome and capable of folding into i-motif tetraplexes under slightly acidic conditions (pH=6).[8,16,17] The biological relevance of i-mo-

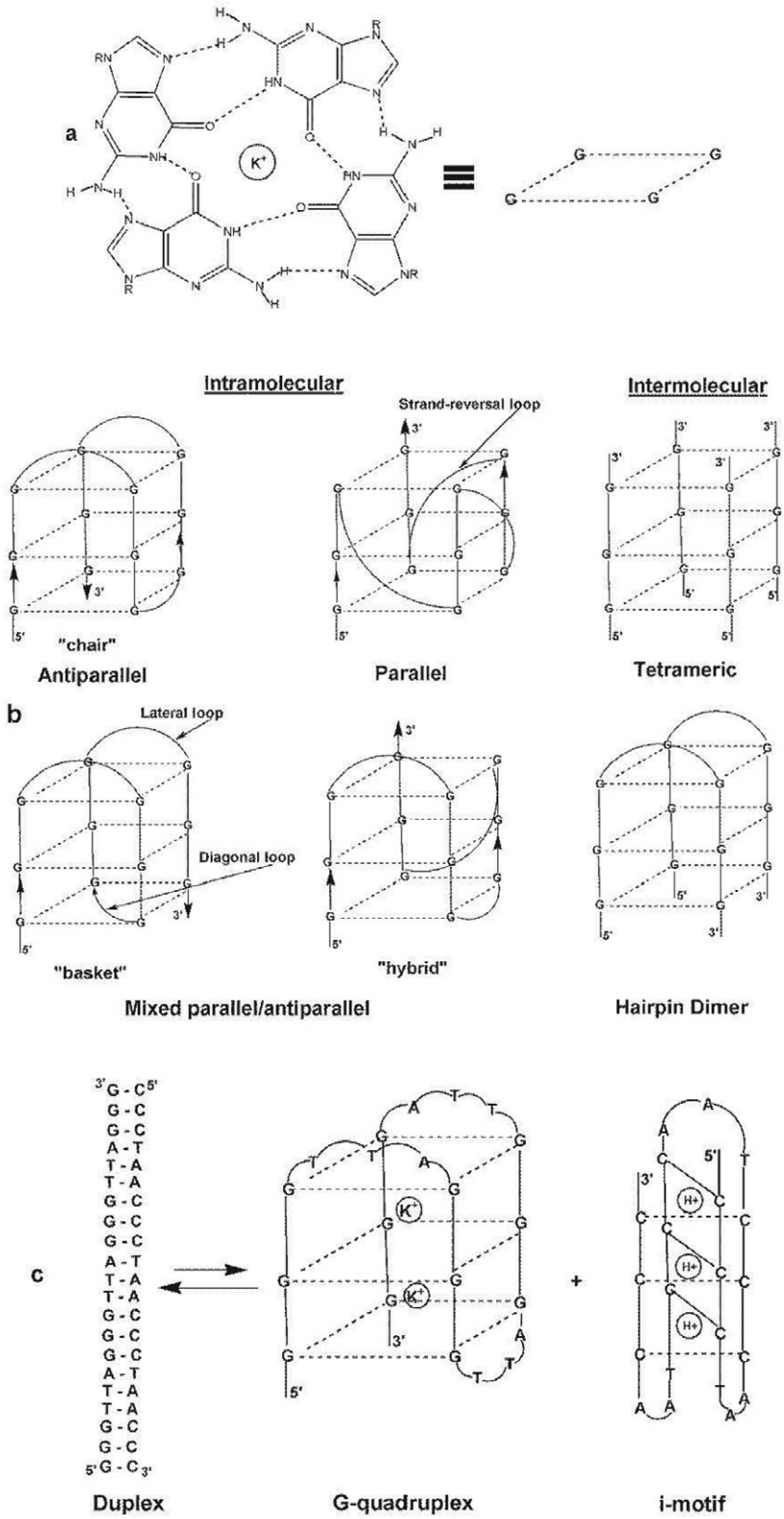

Figure 1. (**a**) G-tetrad structure. (**b**) Various G4 folding topologies. (**c**) One of several ways[24,26–34] to affect the structural equilibrium between duplex and G4/i-motif is by small-molecule binding.[8,27–29]

tif DNA *in vivo* is mainly unknown, but the possibility of having i-motfs formed under physiological conditions due to molecular binding and/or crowding interactions has been highlighted.[8,24,25] When G-C rich sequences exist as a mixture of G4/i-motif and canonical duplex DNA *in vitro*, the structural equilibrium (Figure 1c) can be affected in

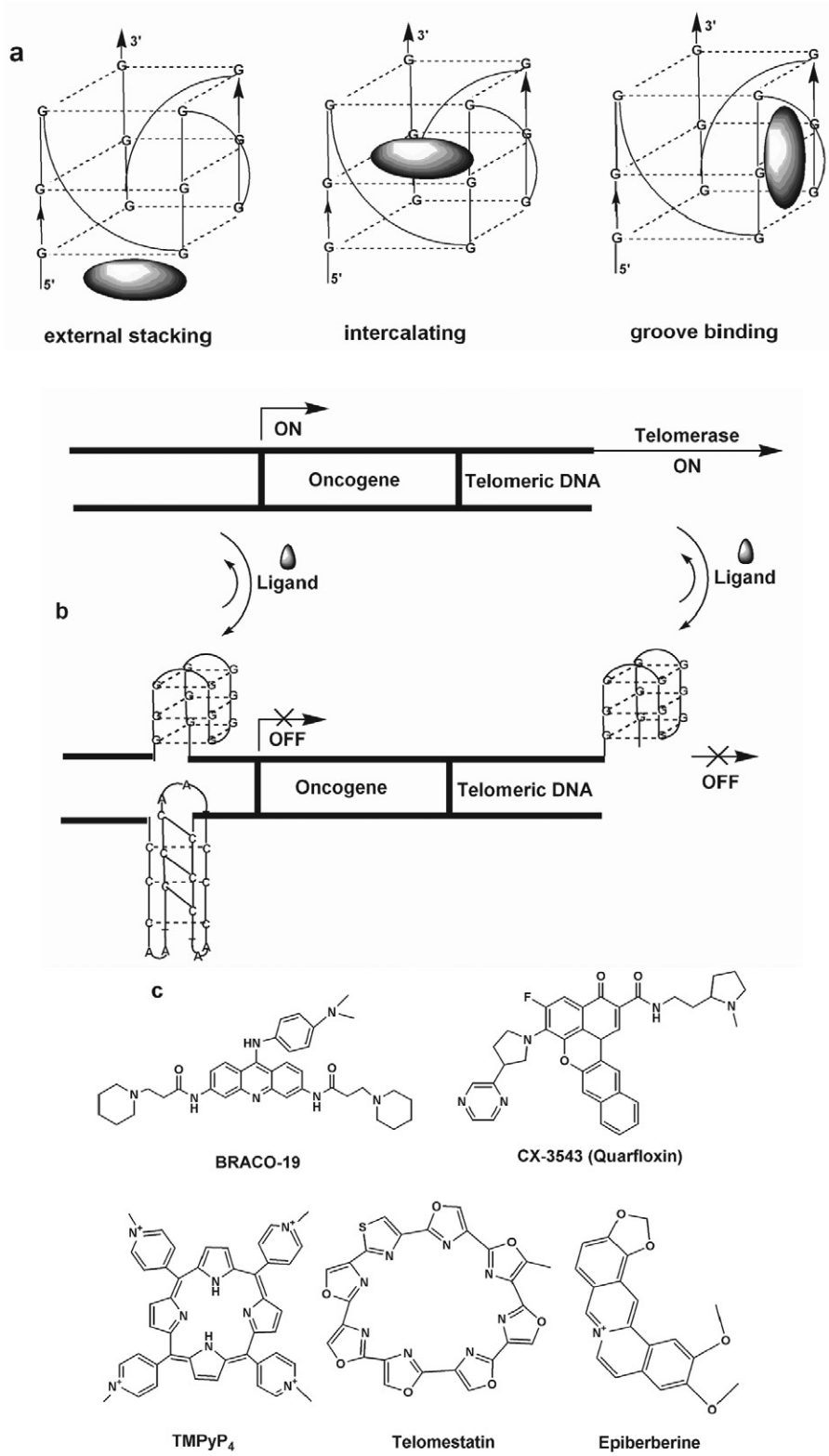

Figure 2. (**a**) Different modes of noncovalent interaction between small molecules and G4. (**b**) Regulation of gene expression and/or inhibition of telomerase activity by G4 stabilization upon ligand binding.[8] (**c**) Well-characterized G4 ligand structures having fused aromatic rings that are capable of stacking with the terminal G-tetrad.

different ways, using DNA binding proteins,[26] small-molecule binding,[27–29] negative supercoiling,[29–33] changes in pH and temperature,[34] and molecular crowding.[24] Thus, cell-permeable and selective ligands may be viewed as potential tools for exploring the biological relevance and/or controlling the function(s) of one or more of these structures.[8]

It is widely accepted that predicting or controlling quadruplex folding is a mainly intractable problem.[35] G-rich DNA sequences are often intrinsically polymorphic *in vitro* and sensitive to pH variations, cation concentrations, or crowding conditions. An interest in the resolution of this issue is dictated by potential implementation of targeted design of quadruplexes in material, biotechnological, and therapeutic applications.[35–37] Methodological advances at a much higher resolution and throughput in the identification and characterization of G4s *in vivo* as well as *in vitro* have well expanded the knowledge of G4 structure and function.[38] Recent evidences have suggested involvement of G4s in key genome functions such as transcription, replication, genome stability, and epigenetic regulation, with many links to cancer biology.[39,40] As far as folding topology (Figure 1b) is concerned, intramolecular G4 structures have been suggested to implicate in the regulation of gene expression and chromosome stability, while intermolecular G4s have been primarily seen as intermediates or precursors of recombination and/or viral integration.[8] The mechanistic insights into G4 biology and protein interaction partners[38,39] have helped to design and develop an arsenal of molecular and chemical tools for biomedical applications,[38] with highlighting new opportunities for drug discovery.[41]

A growing number of predicted (either intramolecular or intermolecular) DNA/RNA quadruplex structures, being deposited into the public domain, enable the structure-based design of G4-interactive ligands on a continuous basis.[42–46] Ligands with specificity toward certain G4s relative to others are useful for exploring the features and functions of individual G4s in the genome.[44] Knowing that some small molecules directly bind to G4 and some others interfere with the binding between G4 structure and related binding proteins, tells that the insights into interaction with nucleic acids and into nucleic acid-protein interaction are very important.[47] Most G4 studies consider only intramolecular G4 folding, but the potential prevalence of intermolecular DNA–RNA G4s in humans has been found by bioinformatics searches,[48] indicating an urgent need for innovative research in order to be able to detect and characterize intermolecular G4 motifs *in vivo*.[38] In other words, great experimental effort and robust analysis platforms are needed to reveal their structural conformational exchange with intramolecular G4s or other structural motifs, and their potential functions in cells,[38] such as in transcription.[48] A wide variety of experimental and computational methods are used to study biomolecular interactions. Experimental techniques include isothermal titration calo-

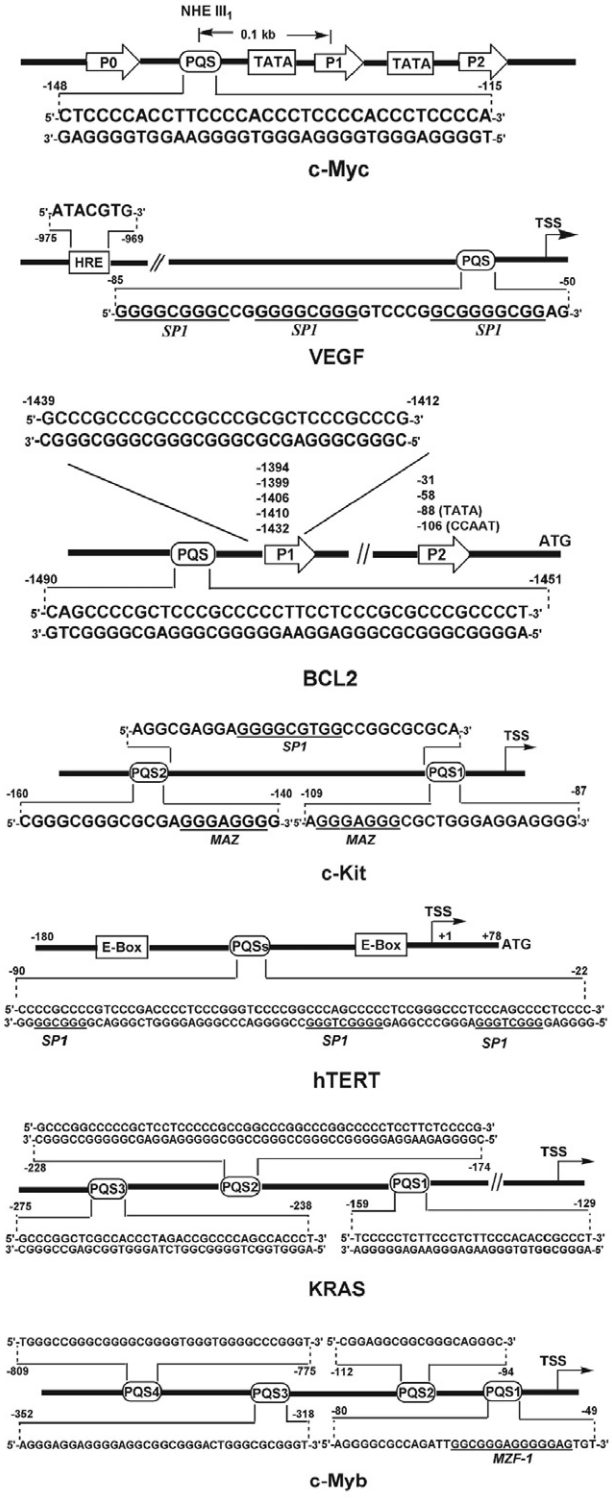

Figure 3. Gene promoters (Ps) with G4 folds.[43] c-Myc has one putative G4-forming sequence (PQS) and one nuclease hypersensitive element (NHE III_1). VEGF has one PQS that is close to the transcription start site (TSS) and one hormone response element (HRE) for regulating the transcription (SP1 – specific protein). BCL2 has two G4-forming elements that attenuate the BCL2 promoter activity. c-Kit has two PQSs that interact with transcription factors (MAZ, SP1). hTERT has a few PQSs, where the presence of two tandemly positioned G4s is proposed. KRAS has three PQSs, of which PQS1 acts as a stronger transcriptional suppressor. c-Myb has several PQSs (MZF1 – protein).

rimetry (ITC), electrospray ionization-mass spectrometry (ESI-MS), X-ray crystallography, nuclear magnetic resonance (NMR), circular dichroism (CD), ultraviolet (UV) and fluorescence spectroscopies.[49–55] Relevant computational approaches comprise virtual screening (VS), molecular docking, molecular mechanics (MM), molecular dynamics (MD), statistical thermodynamics, and bioinformatics.[56–59] Taking into account that various conformations in static structures may be due to differences in experimental conditions or procedures, the applications of MD simulations have become particularly attractive,[60–64] enabling to correlate substantial G4 domain motions and binding site rearrangements with complex formation.[65–68]

The structural heterogeneity, thermal stability and abundance of G4s in telomeres, oncogene promoter regions, and viral genomes make them appealing targets for future therapeutics. In this paper, an up-to-date survey of well-characterized G4s and G4-preferred ligands as potential targets and tools in anti-cancer and antiviral therapies is given. Effects that these ligands selectively exert *in vitro* and *in vivo* models without appreciably affecting normal cells are summarized. Unique ligands involved in specific G4 recognition through different modes of noncovalent interaction (Figure 2a) are put forward. Future perspective in conjunction with a daunting challenge – how to design small molecules that can bind selectively to each of the many possible G4 structures is, to some extent, addressed too.

2. G4-Preferred Ligands: Potential Tools in Anti-Cancer Therapy

The visualization of G4s by immunofluorescence has been enabled by the development of specific antibodies with extremely high affinity to G4s, and has shown the presence of G4s not only in human single-stranded telomeres, but also in duplex regions.[44] G4-forming sequences are observed in the promoter regions of cancer-related genes such as c-Myc,[69–72] VEGF,[73] BCL2,[74] c-Kit,[75] hTERT,[76] KRAS,[77,78] and c-Myb[79] (Figure 3).

The ends of linear chromosomes are protected by telomeres from unwanted DNA processing events that influence genome stability. Human telomeric DNA consists of 5–30 kb tandem repeats of (TTAGGG)n, which end up by a single-stranded, from-35-to-600 bases long 3′ overhang.[1,80–82] Telomere binding proteins protect G-rich telomere repeat sequences that are prone to fold into G4 structures. Small energy differences between telomeric G4 structures have been generally observed.[83,84] An intrinsic polymorphism of telomere DNA is particularly reflected through conformations adopted by the TTA loop segments.[85] The highly conserved telomeric sequence in higher eukaryotes means that potential formations of multiple G4s may be implicated in specific recognition of different structures by different proteins with the aim to control biology.[1] In the majority of human cancers cells as highly proliferative cells, telomeres are maintained by telomerase. Telomerase and its telomere substrates are potential targets for developing novel anti-cancer drugs (Figure 2b). The discovery of the naturally occurring macrocyclic compound Telomestatin (Figure 2c) with telomerase-inhibiting activity due to binding to telomeric G4 structures indicated the existence of G4s *in vivo*.[44] Telomere function, besides the inhibition of telomerase, might be influenced by targeting telomeric G4-DNA. The importance of telomere structure and its position during telomerase function, as well as its associated binding proteins has been experimentally dissected.[20] Many proteins that bind to double-stranded and/or single-stranded regions of the telomeric DNA make a nucleoprotein complex that maintains the structural integrity of telomeres *in vivo*. The effect of telomere destabilization due to small-molecule binding to DNA and consequent displacement of proteins from the complex is known as possible genotoxicity, being in relation to many G4 ligands. How this effect might be cancer-specific is unclear.[8]

Contrary to the formation of telomeric G4s in the single-stranded 3′ overhang of telomeres, promoter G4s fold in the regions of double-stranded DNA. G4 formation in promoter regions is related to genes that are responsible for cell growth and proliferation (Figure 3). The clustering of putative G4-forming sequences (PQSs) is within 1 kb upstream of the transcriptional start site (TSS). The oncogene promoters are typically TATA-less with G-rich regions in vicinal promoters. Unlike telomeric DNA, the PQSs are substantially more diverse and frequently have more than four G-tracts. A sequence is capable of forming multiple G4s through a wide variety of combinations of G-tracts or different loop isomers. The presence of the G3NG3 motif is a conspicuous feature that might have been naturally selected as a basis for G4 formation. Since the determination of c-Myc G4, parallel G4 folds have been commonly detected; most of them contain three G-tetrads and three loops (the first and the third are 1 nucleotide long, the middle loop is of variable length). In other words, each parallel G4 structure is likely to adopt unique capping and loop structures by way of its specific variable middle loop and flanking segments. The propensity of promoter sequences to form multiple and stable G4s at equilibrium is quite intriguing. For example, in the overlapping region of BCL2, the presence of two distinct conformationally interchangeable G4s suggests a mechanism for the regulation of gene transcription through specific recognition of different G4 structures by different proteins.[1] Many proteins with binding affinity to G4s have been identified.[86] Modulation of gene expression may also be influenced using different small molecules in order to recognize distinct G4s. Thus, the targeting of G4s by small molecules, aimed to disrupt the interactions between G4s and their binding proteins, emerges as a potential anti-cancer strategy.[61]

The different modes of noncovalent G4-ligand interaction include external stacking, intercalation, and groove/loop binding (Figure 2a). Experimental[87,88] and computational[61] reports have identified the π-π stacking of ligand at the end of G4 as the most stable mode. Grooves/loops have been suggested to be viable binding sites of particular importance for blocking the interaction between G4 and its binding proteins in aqueous solution.[61] The challenge of designing specific groove/loop binders stems from the groove/loop interaction mode dependence on the particular topology of groove/loop residues. However, grooves/loops offer distinct environments to gain specificity among many types of G4s by way of subtle variations of G4 topologies, groove widths, and loop sequences without affecting binding affinity.[44,61]

There are two distinct mechanisms to inhibit cancer growth through the selective stabilization of G4s by ligand molecules (Figure 2b). The first refers to the inhibition of the over-expression of oncogenes by promoter deactivation,[8,89] while the second refers to the inhibition of telomerase, a ribonucleoprotein complex that catalyzes the 3' extension of telomeric DNA.[8,90–95] The second mechanism has been more extensively studied.[8]

Well-characterized G4-interactive ligands, BRACO-19, CX-3543 (Quarfloxin), TMPyP4, and Telomestatin that have modest binding affinities to duplex DNAs[44] are given in Figure 2c. The chemical structure of BRACO-19 is composed of fused aromatic rings that are capable of stacking with the terminal G-tetrad and of three side chains that branch out of its heteroaromatic core. Many similar ligands (like CX-3543) have one or more cationic side chains that are inclined to interact with G4 grooves/loops. An early idea that an optimal G4-preferred ligand structure contains large, planar, symmetric and cyclic rings, such as those of TMPyP4 and Telomestatin (Figure 2c) in order to maximize stacking interactions with the external G-tetrad has been closely associated with low specificity among intramolecular G4s.[1] BRACO-19 is one of the most studied G4 ligands so far. The studies of BRACO-19 have greatly contributed to the treatment of telomeric G4s as potential therapeutic targets. BRACO-19 has shown high anti-cancer activities *in vivo*, such as in a UXF1138L uterus carcinoma xenograft and in a DU-145 prostate cancer xenograft. Despite all these favorable functional features, the lack of membrane permeability and small therapeutic window have been identified as the major limitations of BRACO-19, which must be resolved before any attempt to develop an effective clinical agent.[42] Quarfloxin was the first-in-class ligand with considerable therapeutic window that had completed Phase II trials as a drug candidate, well-tolerated in patients against neuroendocrine tumors, carcinoid tumors, and lymphoma. Quarfloxin targets a G4 from the c-Myc promoter region to disrupt the G4-nucleolin complexes, and its G4-binding was reported as inhibiting RNA biogenesis. The Phase III of human cancer clinical trials is not currently proceeding due to high albumin binding.[43]

G4s and G4-interactive ligands as potential targets and tools in anti-cancer therapy are given in Table 1. Besides small molecules, G4-binding metal complexes have been recognized as promising anti-cancer drugs.[96] In general, ligand-mediated stabilization of the G4 structure(s) effectively inhibits telomerase activity or oncogene over-expression and, when applied to cells, most G4 ligands initiate antiproliferative effect (apoptosis) and/or replicative senescence.[8] Ni-P, Quercetin, TH3, IZCZ-3, Benzofuran derivative, and Furopyridazinone derivative cause negligible cytotoxicity to normal somatic cells *in vivo*. By avoiding many of the problems underlying the therapeutic use of oligonucleotides, the G-rich VEGFq oligonucleotide has contributed to a novel approach to specific inhibition of gene expression *in vivo*, which can be applied to the wide array of genes whose promoters contain quadruplex-forming sequences.[97] The chemical structure of each ligand, underlined with its respective target topology is displayed in Figure 4. Among these fifteen ligands, eleven prefer parallel, three prefer hybrid, and one prefers dimeric G4 binding. It is known that the induction of a quadruplex or change of a quadruplex conformation upon binding may be one of the most powerful methods to exert a desired biological effect.[51] If a ligand selectively interacts with different G4 topologies, the particular ligand is expected to easily regulate the conformational switch by surpassing the energy barriers between distinct G4 structures in Na^+ or K^+ solution. An NMR structural analysis has revealed that a berberine derivative, epiberberine (Figure 2c), discriminates a hybrid type 2 telomere G4 from the other adoptable topologies and promoter G4s (c-Myc, BCL2, and PDGFR).[98] Also the ability of epiberberine to convert the other conformations, such as telomere G4 hybrid type 1 and antiparallel (basket type) G4s, into the type 2 hybrid topology has been reported.[43,98] It has been recently concluded that specific targeting of G4s by small molecules represents a promising strategy to study the function of targets inside a living cell without influencing their intact states.[44]

The way in which CM03 (Figure 4) has been designed to target multiple effector pathways in pancreatic ductal adenocarcinoma (PDAC) deserves more attention.[99] The co-crystal structure of MM41 (Figure 4) with an intramolecular human telomeric parallel G4 has been the starting point for CM03 design. Even though the nature and structures of target G4s are unknown, NMR and crystal structures illustrate that some features are common to all G4s, particularly a core of stacked G-tetrads with small-molecule binding at the end of the core. The chromophore of MM41 has been somewhat asymmetrically stacked to the terminal G-tetrad. Indeed, one of the four substituent chains has not been positioned as the other three side chains with respect to G4 due to its orientation away from the G4 surface. The particular side chain has not been capable of making effective contacts with a G4 groove, so that its contribution to overall bind-

Table 1. G4 ligands with anti-cancer activities. Updated data reported previously.[43]

Target	G4 topology	Cell Line	Cancer Type	Ligand[Ref.]	Effect
telomere	hybrid	MDA-MB-231	breast cancer (adenocarcinoma)	Ni-P[107,108]	Cancer stem cell-specific apoptosis, bulk cancer-specific apoptosis and senescence, negligible cytotoxicity to normal somatic cells
		MCF-7	breast cancer (adenocarcinoma)		
telomere	dimeric G4s	SiHa	squamous cell carcinoma	IZNP1[109]	Apoptosis, senescence
telomere	parallel	A549	lung adenocarcinoma	MM41[99] CM03[99]	Antiproliferative activity (apoptosis), BCL2 and KRAS as secondary targets
		MCF-7	breast cancer (adenocarcinoma)		
		MIA PaCa-2 PANC-1	pancreatic ductal adenocarcinoma		
c-Myc	parallel	HeLa	cervical cancer	Quercetin[110]	Apoptosis, mild cytotoxicity to normal cell line
c-Myc	parallel	A549	lung cancer	TH3[111]	Antiproliferative effect (apoptosis), negligible cytotoxicity to normal somatic cells
		HeLa	cervical cancer		
c-Myc	parallel	SiHa	squamous cell carcinoma	IZCZ-3[112]	Antiproliferative effect (apoptosis), negligible cytotoxicity to normal somatic cells
		HeLa	cervical cancer		
		Huh7	liver cancer		
		A375	malignant melaoma		
c-Myc	parallel	L363, MM1S, MM1R etc.	myeloma	Benzofuran derivative[113]	Antiproliferative effect (apoptosis), negligible cytotoxicity to normal cells
c-Myc	parallel	HCT116	colorectal carcinoma	Tz 1[114]	Apoptosis
VEGF	parallel	A549	lung cancer	VEGFq[97]	Autophagic apoptosis
BCL2	hybrid	Jurkat	human acute T cell leukemia	Furopyridazinone derivative[115]	Antiproliferative effect (apoptosis), negligible cytotoxicity to normal cells
c-Kit	parallel	MCF-7	breast adenocarcinoma	AQ1[116]	Antiproliferative effect (apoptosis)
		HGC-27	gastric carcinoma		
hTERT	hybrid with stem loop	MCF7	breast adenocarcinoma	GTC365[117]	Apoptosis, senescence
KRAS	parallel	HCT16 SW620	colorectal carcinoma	Indoloquinoline derivatives[118]	Apoptosis
c-Myb	parallel	MCF7	breast adenocarcinoma	Topotecan[119]	Repressed expression, uncertain specificity

ing has been minimal. Relative to MM41, an optimal compound has been hypothesized to contain three substituents and to bind with similar affinity, as well as to have the advantage of lower molecular weight and reduced overall cationic charge. Thus, MM41 has been a suboptimal drug candidate due to its higher molecular weight and four positive charges, while CM03 has been an improved rationally designed derivative of MM41 and a novel lead candidate compound for potential therapy against human PDAC. Particular promoter G-quadruplexes have not been assumed as targets. Global genome transcriptome profiling has been employed to determine which genes are affected by the rationally designed G4-interactive small molecule. Consequently, potential targets at the whole genome level in two pancreatic cancer cell lines have been determined. With *in vitro* cell assays and *in vivo* models for human PDAC, CM03 has been identified as a highly selective and potent G4-binding ligand.[99]

The dynamics of noncovalent interaction between a structurally representative set of small molecules (BRACO-19, $TMPyP_4$, CX-3543, 10074-G5, Telomestatin, Tet-

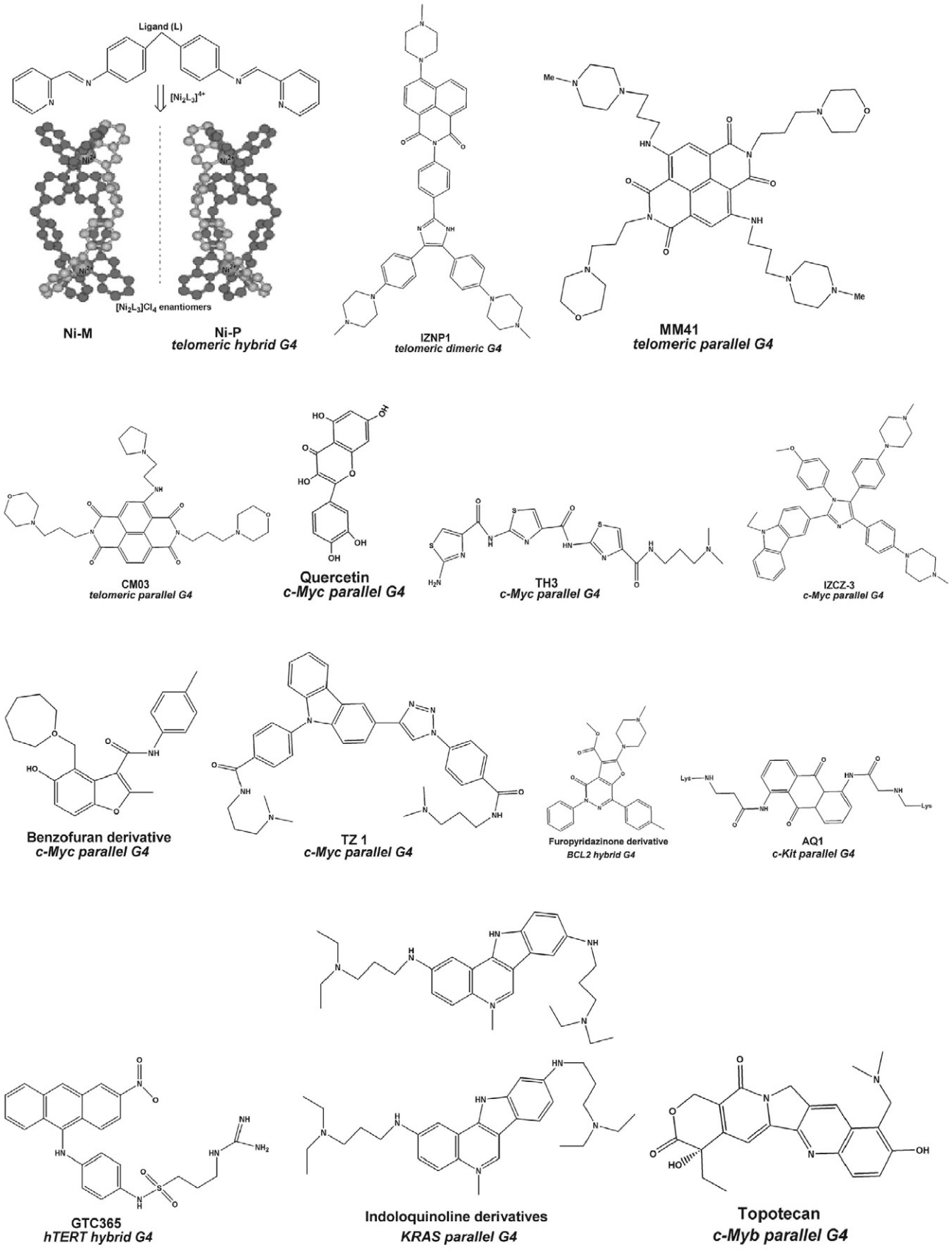

Figure 4. Chemical structures of G4 ligands (with denoted target topologies by italic) that exhibit anti-cancer activities (Table 1).

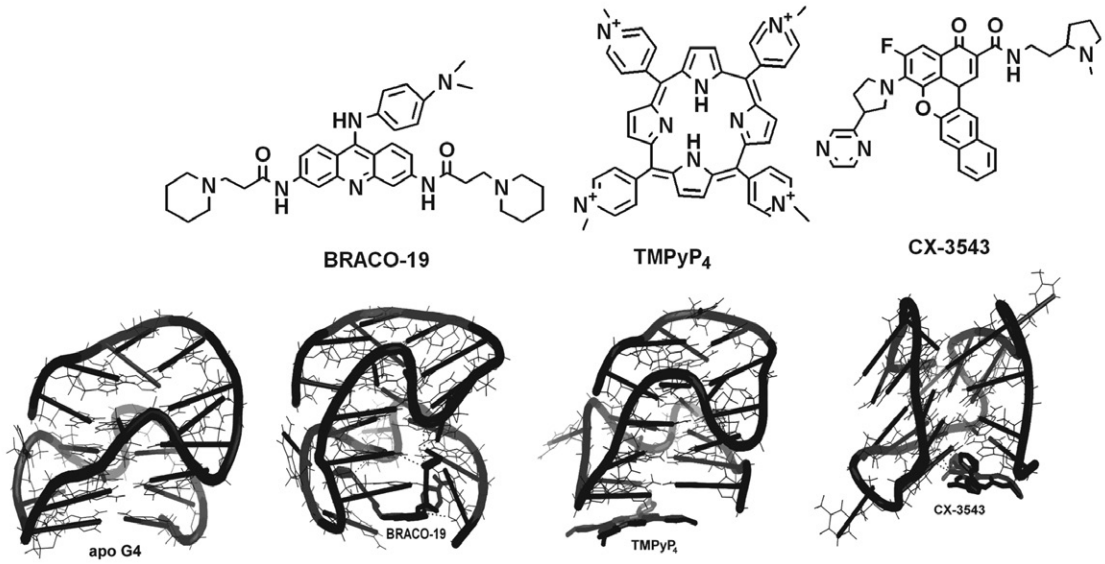

Figure 5. External stacking of BRACO-19, $TMPyP_4$ and CX-3543 to apo (ligand-free) G4 from the c-Myc promoter in the stable regime of molecular dynamics (MD) simulation. In a representative set of structurally diversified ligands, these three molecules were established to have the highest affinity for G4. Only BRACO-19 was shown to be a thermodynamically favorable binder by increasing the conformational flexibility of G4 in the asymptotic ($t \rightarrow \infty$) regime of MD simulation.[68]

rahydropalmatine, Sanguinarine, Hoechst 33258, Benzophenanthridine derivative, Nitidine Chloride, Piperine, 12459, Quercetin, Quindoline, Berberine, and Flavopiridol) and a G-quadruplex formed in the c-Myc oncogene promoter region was recently explored in a systematic fashion from a rigorous biophysical point of view.[68] In fact,

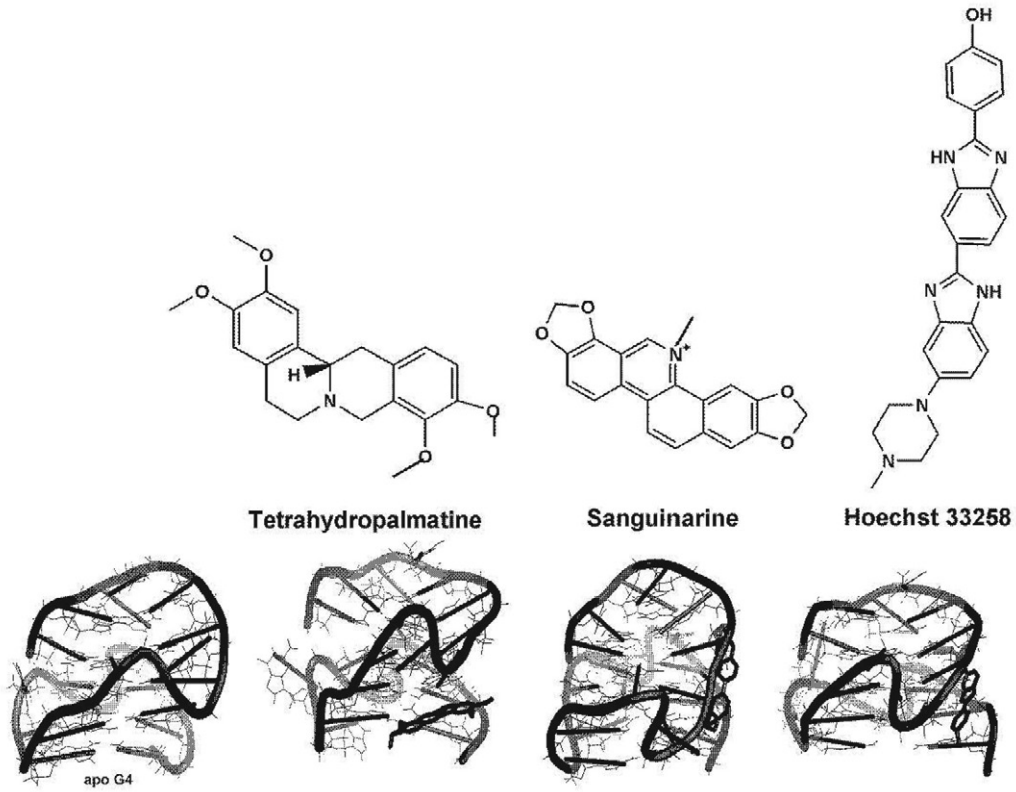

Figure 6. Groove binding of Tetrahydropalmatine, Sanguinarine and Hoechst 33258 to apo (ligand-free) G4 from the c-Myc promoter in the stable regime of molecular dynamics (MD) simulation. In a representative set of structurally diversified ligands, these three molecules were established to have the highest affinity for G4. Only Tetrahydropalmatine was shown to be a thermodynamically favorable binder by increasing the conformational flexibility of G4 in the asymptotic ($t \rightarrow \infty$) regime of MD simulation.[68]

the thermodynamic consequences of apo (ligand-free) G4 conformational flexibility change upon ligand binding have been investigated in the asymptotic regime ($t \rightarrow \infty$) of MD simulation, obtained by extrapolating the stable regime to infinitely long MD simulation. BRACO-19, TMPyP$_4$ and CX-3543 have shown the highest affinity to the G4 by stacking to the bottom G-tetrad of G4 (Figure 5). However, only BRACO-19 has been found to be a thermodynamically preferable binder by increasing the conformational flexibility of G4 (Figure 5), with a somewhat larger (by about 3 kcal mol^{-1}) contribution to the additional flexibility of G4 from the sugar-phosphate backbone than from the complete system of nucleobases. In addition, Tetrahydropalmatine, Sanguinarine and Hoechst 33258 have exhibited the highest affinity to the target by groove binding (Figure 6). However, only Tetrahydropalmatine has been found to be a thermodynamically favorable binder by increasing the conformational flexibility of G4 (Figure 6), mainly through the complete system of nucleobases. Therefore, two distinct mechanisms by way of which small molecules interact with G4 are associated with increased conformational flexibility and increased conformational rigidity of apo G4 upon ligand binding respectively.[68]

Even though pure tetrad-binding mode is more stable than groove/loop binding mode, grooves/loops are viable binding sites that are of interest for the structure-based drug design. Grooves/loops with distinct environments help in tuning ligand specificity among many types of G4s without affecting binding affinity.[44] Thus, multiple binding modes, which include external stacking and/or intercalation and/or groove/loop binding of two or more ligands simultaneously, have attracted certain attention.[61] This type of binding is less stable than external stacking, likely due to the ability of groove/loop-binding ligands to induce loop rearrangements and destabilize the overall binding by displacing the interaction of the side chains of G-tetrad-binding ligands with the grooves/loops of G4. There are indications that a combined – G-tetrad and groove binding of ligands enhances G4 conformational rigidity, reflected through the decreased conformational flexibility of both G-tetrads and the backbone.[61] For rationalizing this aspect in the case of G4 from the c-Myc promoter region, a relevant structural basis was proposed to include two unique – thermodynamically preferred small molecules: the external stacking of BRACO-19 and the groove binding of Tetrahydropalmatine simultaneously.[68]

Binding sites defined by the surface features of the groove/loop regions can be used to stimulate selective binding interactions, even between closely related G4 structures.[8] Subtle variations of G4 topologies, groove widths, and loop sequences are associated with a highly dynamic nature of G4 structures, which have propensity to lose conformational entropy upon ligand binding.[101] This factor in determining specificity is important in order to distinguish G4s with lower ligand affinities that exist as a dynamic mixture of conformations in the unbound state (human telomere) from G4s that adopt a single conformation.[8] A small molecule, with binding affinity to increase the conformational entropy of G4 by stacking at the end of

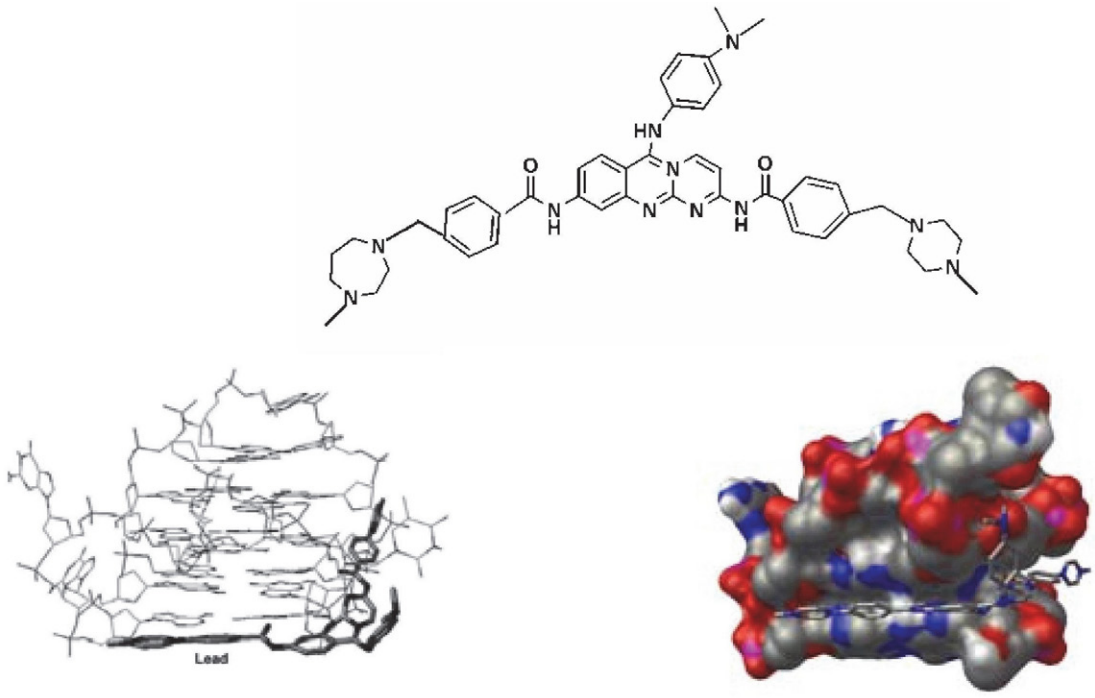

Figure 7. Proposal of lead candidate structure to interact with the c-Myc promoter G4 through external stacking and groove binding simultaneously. This proposal was based on HTVS experiments employing the key pharmacophore features of BRACO-19 for the search of the KEGG databases.[100]

Table 2. G4s that are identified in viral genomes and their active ligands. Updated data reported previously.[46]

Virus	Name	Size (nm)	Genome	No. of G4s[Ref.]	Active Ligands[Ref.]
Human Immunodeficiency Virus 1	HIV-1	ø 120	(+)ssRNA 9.75 kb	12[120]	BRACO-19[120–122] TMPyP4[122,123] PIPER[123] c-exNDI[124] Nitidine Chloride[105] Benzophenanthridine derivative[105]
Herpes Simplex Virus 1	HSV-1	ø 125	dsDNA 152 kb	316[125,126]	BRACO-19[126] c-exNDI[127]
Epstein-Barr Virus	EBV	ø 120–180	dsDNA 172 kb	13[125]	BRACO-19[128] PDS[129] PhenDC3[130]
Kaposi's Sarcoma associated Herpes Virus	KSHV	ø 125	dsDNA 170 kb	52[125,131]	PhenDC3[131]
Human Herpes Virus 6	HHV-6	ø 200	dsDNA 162 kb	43[125]	BRACO-19[132]
Hepatitis C Virus	HCV	ø 60	(+)ssRNA 9.6 kb	2[133]	TMPyP4[133] PDP[133]
Human Papilloma Virus	HPV	ø 60	circular dsDNA 8 kb	8[134]	
Zika Virus	ZIKV	ø 50	(+)ssRNA 11 kb	8[135]	
Severe Acute Respiratory Syndrome Corona Virus	SARS CoV	ø 200	(+)ssRNA 30 kb		
Hepatitis B Virus	HBV	ø 42	partially circular dsDNA 3.2 kb	1[136]	TMPyP4[136] PDS[136]
Ebola Virus	EBOV cylindrical	ø 80	(-)ssRNA 18.9 kb	1[137]	TMPyP4[137]

G4 from an oncogene promoter region, can be hypothesized as a unique, specific pharmacophore for the identification of new lead candidates by high-throughput virtual screening (HTVS).[68,100] A lead candidate compound (Figure 7), predicted to recognize the c-Myc promoter G4 through external stacking and groove binding at the same time, was designed by HTVS experiments in combination with analog design.[100] The key pharmacophore features of BRACO-19 have been used for the search of the Kyoto Encyclopedia of Genes and Genomes (KEGG) databases in order to generate hit-to-lead candidates. Two crucial features rationalize the visible G4-stabilizing advantages of the concomitant external stacking and groove binding of the lead candidate over the external stacking of BRACO-19. The first is a flexible aromatic core of the lead candidate relative to a rigid one of BRACO-19. The second is a more polar surface of the lead candidate (by about 51 Å^2) than that of BRACO-19. The conformational flexibility of small molecules is generally more preferable compared to their locking in a presumed bioactive G4 conformation.[100] Structure-based virtual screening and cell-based screening approaches, as well as biophysical and/or biological assays define an acceptable framework for the determination of completely new types of bioactive G4-interactive ligands.[44,102]

3. G4-Preferred Ligands: Potential Tools in Antiviral Therapy

The presence of G4s in viruses has attracted more attention during the last few years. The viruses include those involved in recent epidemics, such as the Zika and Ebola viruses. Putative G4-forming sequences are usually locat-

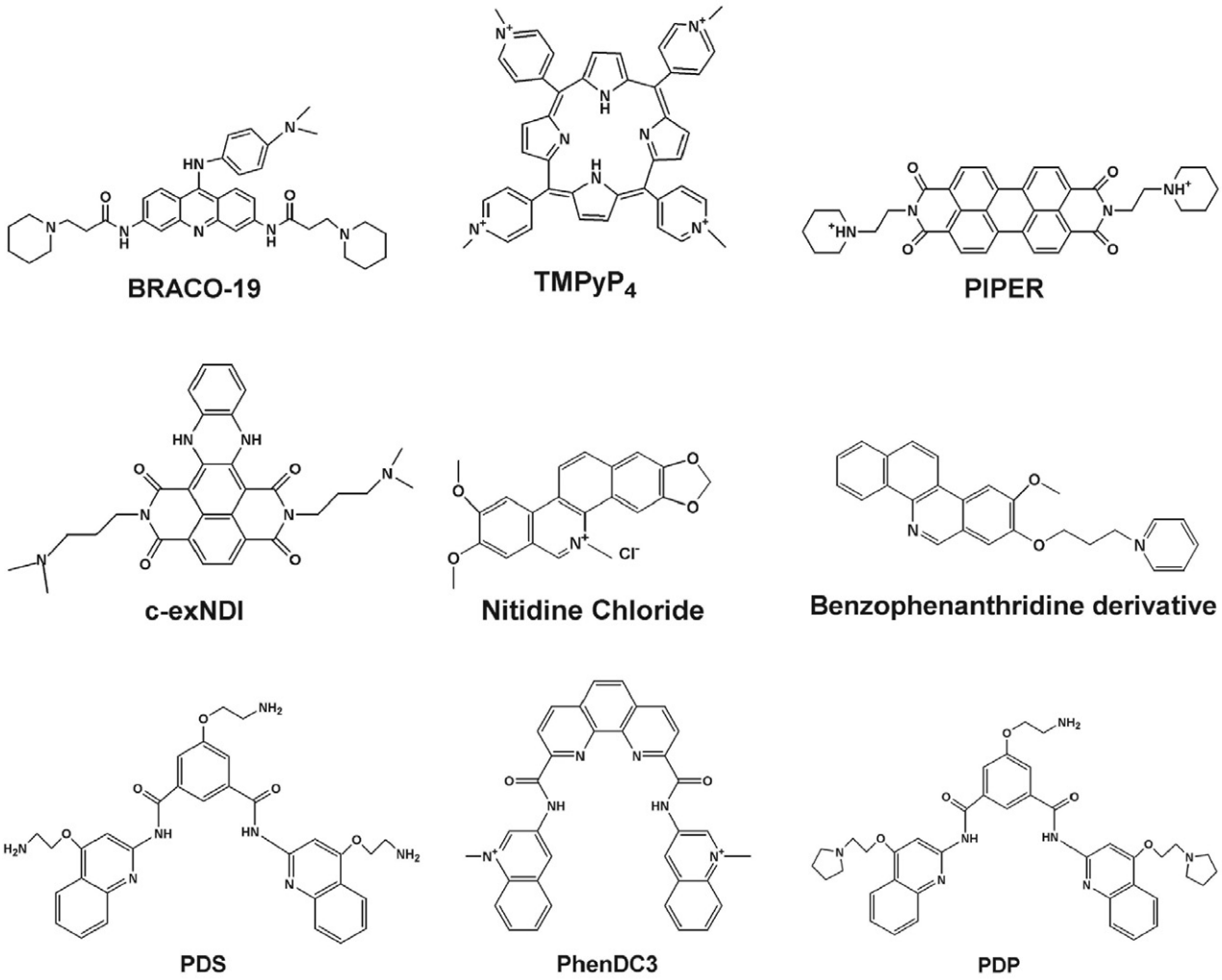

Figure 8. Chemical structures of ligands that prefer G4s in viral genomes.

ed in regulatory regions of the viral genomes and implicated in key viral processes; in some cases, their involvement in viral latency has been reported too.[46] G4 ligands are tools that have been developed and tested in order to study the complexity of G4-mediated mechanisms in the viral life cycle. They have also been viewed as potential therapeutic agents.[46] G4s that are identified in viral genomes, as well as their active ligands are summarized in Table 2. The chemical structures of the ligands are shown in Figure 8. Promising antiviral effects of G4 ligands have been generally related to G4-mediated mechanisms of action both at the genome level and at transcriptional level.[46] G4-forming oligonucleotides as potential antiviral agents have been previously reviewed in great detail,[103,104] so that they are not considered in the present review article.

Experimental research based on ESI-MS, CD spectrometry, and DMS footprinting has indicated the formation of a G4 within a G-rich sequence that is located between –76 and –57 bp in the HIV-1 promoter.[105] The CD melting experiment has also shown that, among eight natural small molecules (Nitidine Chloride – NC, Benzophenanthridine derivative – BPD, Jatrorrhizine, Tetrahydropalmatine, Toddalolactone, Coptisine, Piperine, and Astragalin), NC and BPD have the highest and nearly equal affinities to the HIV-1 promoter G4.[105] The binding modes of NC and BPD have been elaborated using sophisticated computational methods,[66] demonstrating that NC is a thermodynamically unfavorable binder by increasing the conformational rigidity of apo G4 and that BPD is a thermodynamically favorable binder by increasing the conformational flexibility of apo G4 in the asymptotic ($t \rightarrow \infty$) regime of MD simulation (Figure 9).

In addition to HTVS methods or structure-based design with ahead-presumed features, fragment-based drug discovery (FBDD) may be a valuable approach to the generation of new pharmacophores that specifically recognize G4 structures. This approach is based on the generation of molecular fragment small libraries screened against the receptor in order to further synthetically elaborate them into lead compounds. For example, one of the heterocyclic molecules (Figure 10) has been shown to specifically recognize G4 from the HIV-1 long terminal repeat (LTR) pro-

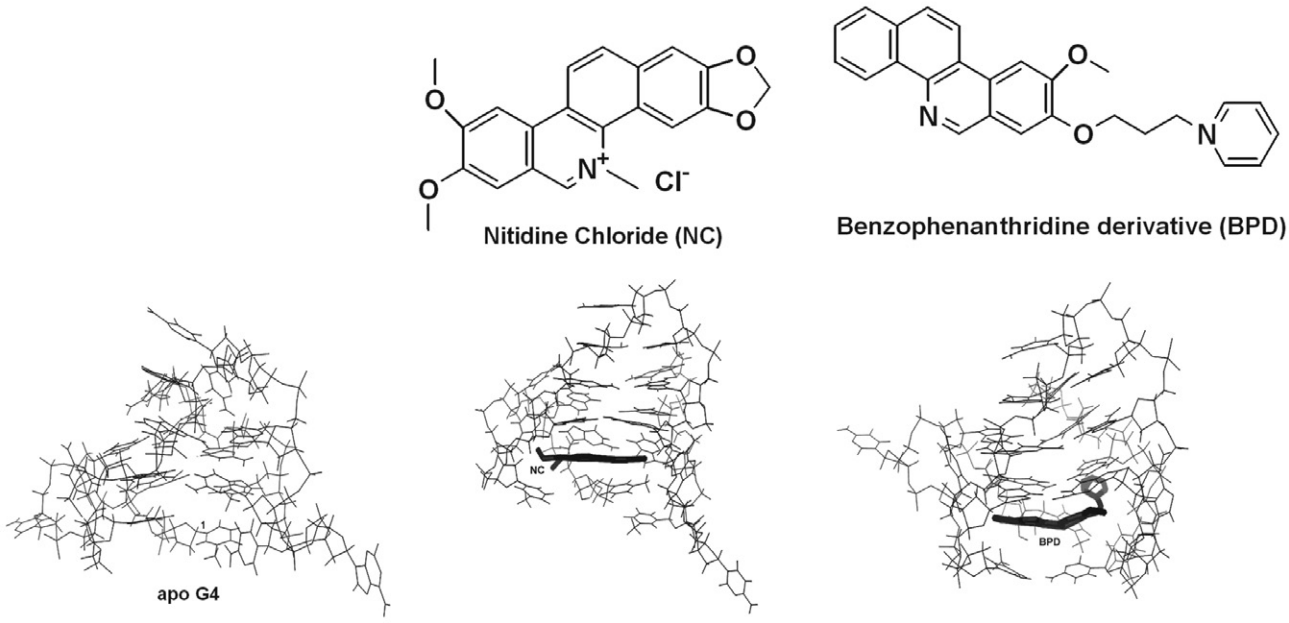

Figure 9. Among eight natural small molecules, NC and BPD were shown to have the most pronounced (and mutually comparable) affinity for G4 from the HIV-1 promoter.[105] NC is a thermodynamically unfavorable binder by increasing the conformational rigidity of apo G4, while BPD is a thermodynamically favorable binder by increasing the conformational flexibility of apo G4 in the asymptotic ($t \to \infty$) regime of MD simulation.[66]

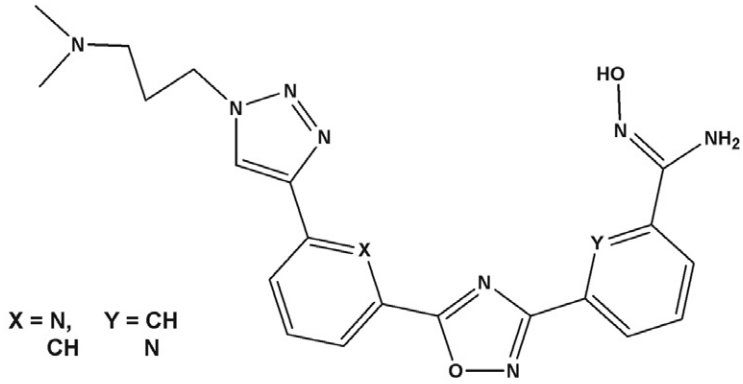

Figure 10. One of the heterocyclic molecules was shown to prefer G4 from the HIV-1 LTR promoter region and to represent a potential pharmacophore for the development of novel ligands with unexpected chemical features. These compounds were developed using a FBDD approach.[106]

moter region and to represent a potential pharmacophore for the development of novel ligands with unexpected chemical features.[106] Size and poor pharmacokinetics are the main obstacles in the development of G4-interactive ligand. FBDD can be a relevant approach to the development of compounds that have smaller sizes and more drug-like properties.

4. Conclusions and Future Perspective

G-quadruplexes are naturally forming structures under physiological conditions, stabilized by monovalent cations present in cells. Over two hundreds quadruplex structures, either intramolecular or intermolecular, are currently deposited in the public domain such as the Protein Data Bank. Most G4 studies consider only intramolecular G4 folding. However, the potential prevalence of intermolecular DNA/RNA G4s in humans has been indicated by bioinformatics searches. It means that innovative research is urgently needed with the aim to detect and characterize intermolecular G4 motifs *in vivo*, that is, their structural conformational exchange with intramolecular G4s or other structural motifs, and their potential functions in cells.[38] The structural diversity, thermal stability and abundance of G4s in telomeres, oncogene promoter regions, and viral genomes make them attractive targets for potential anti-cancer and antiviral therapies.

Although the nature and structures of target G4s are unknown, NMR and crystal structures show that some features are common to all G4s; e.g. a core of stacked

G-tetrads with small-molecule binding at one of the ends of the core. One of the features has been recently revealed by molecular dynamics simulations,[66,68] because distinct conformations that are often observed in static, experimentally determined structures may be the consequences of the differences in experimental conditions or procedures. The thermodynamic consequences of apo (ligand-free) G4 conformational flexibility change upon complex formation have been observed in the asymptotic regime ($t \rightarrow \infty$) of MD simulation. Two dissected mechanisms of G4-small molecule interaction are associated with increased conformational flexibility and increased conformational rigidity of apo target upon ligand binding, thereby being thermodynamically favorable and unfavorable respectively. A small molecule with binding affinity to increase the conformational flexibility of G4 through π-π stacking at the end of G4 can be conceivable as a unique, specific pharmacophore for designing novel lead candidate compounds by high-throughput virtual screening.[66,68] Virtual screening has been demonstrated to be effective in reducing the initial number of potential candidates.[44] In this way a lead candidate structure has been predicted to target a G4 from the c-Myc promoter region through external stacking and groove binding simultaneously.[100] This approach would have useful implications for overcoming the challenge of designing specific groove/loop binders, which stems from the groove/loop interaction mode dependences on the particular G4 topologies, groove widths, and loop sequences. Therefore, the use of grooves/loops offers distinct environments aimed to gain specificity among many types of G4s without influencing binding affinity.[44]

In contrast to HTVS methods or structure-based design with pre-set features, fragment-based drug discovery, which is based on the generation of molecular fragment small libraries screened against the receptor to further synthetically convert them into lead compounds, may be a valuable approach to the generation of new pharmacophores that specifically recognize G4 nucleic acid structures. The sizes and poor pharmacokinetic properties of G4-interactive ligands are the main glitches in their development. By adding up fragments to singly recognize the target, FBDD can be seen as a relevant approach to the development of compounds that have smaller sizes and more drug-like properties.[106]

An appropriate framework for identifying totally new types of bioactive G4-interactive ligands is currently defined by structure-based virtual screening methods and cell-based screening approaches.[44,106] Specific targeting of G4s by small molecules is and will be a promising tool for studying the behavior of targets inside a living cell without influencing their intact states.[44]

Particular promoter G4s should not be assumed as prior targets, indicating that single G4 promoter targeting strategy is not quite a suitable approach. In fact, the knowledge of potential targets at the whole genome level is needed. Global genome transcriptome profiling can be exploited for the determination of which genes are affected by a rationally designed G4-interactive small molecule. As a consequence, the selectivity and potency of a new G4-preferred compound can be evaluated using *in vitro* cell assays and *in vivo* models. A relevant example is the successful design, synthesis and identification of CM03 as a novel lead candidate for the potential therapy against human pancreatic cancer.[99]

This review article is imagined to inspire ongoing efforts of modern chemists and pharmacists to target G4 structures.

5. References

1. B. Onel, C. Lin, D. Z. Yang, *Sci. China Chem.* **2014**, *57(12)*, 1605–1614. **DOI:**10.1007/s11426-014-5235-3
2. B. E. Bernstein, E. Birney, I. Dunham, E. D. Green, C. Gunter, M. Snyder, *Nature* **2012**, *489(7414)*. 57–74. **DOI:**10.1038/nature11247
3. P. M. Mitrasinovic, *J. Chem. Inf. Model.* **2015**, *55(2)*, 421–433. **DOI:**10.1021/ci5006965
4. A. Rakic, P. M. Mitrasinovic, *J. Serb. Chem. Soc.* **2008**, *73(1)*, 41–53. **DOI:**10.2298/JSC0801041R
5. P. M. Mitrasinovic, *J. Struct. Biol.* **2006**, *153(3)*, 207–222. **DOI:**10.1016/j.jsb.2005.12.001
6. I. Bang, *Biochemische Zeitschrif* **1910**, *26*, 293–311.
7. M. Gellert, M. N. Lipsett, D. R. Davies, *Proc. Natl. Acad. Sci. U. S. A.* **1962**, *48(12)*, 2013-2018. **DOI:**10.1073/pnas.48.12.2013
8. N. W. Luedtke, *Chimia* **2009**, *63(3)*, 134–139. **DOI:**10.2533/chimia.2009.134
9. N. V. Hud, F. W. Smith, F. A. L. Anet, J. Feigon, *Biochemistry* **1996**, *35(48)*, 15383–15390. **DOI:**10.1021/bi9620565
10. T. G. Voss, C. D. Fermin, J. A. Levy, S. Vigh, B. Choi, R. F. Garry, *J. Virol.* **1996**, *70(8)*, 5447–5454. **DOI:**10.1128/JVI.70.8.5447-5454.1996
11. A. Risitano, K. R. Fox, *Biochemistry* **2003**, *42(21)*, 6507–6513. **DOI:**10.1021/bi026997v
12. D. Sen, W. Gilbert, *Nature* **1990**, *344(6265)*, 410–414. **DOI:**10.1038/344410a0
13. D. Miyoshi1, A. Nakao, N. Sugimoto, *Nucleic Acids Res.* **2003**, *31(4)*, 1156–1163. **DOI:**10.1093/nar/gkg211
14. D. Bhattacharyya, G. Mirihana Arachchilage, S. Basu, *Front. Chem.* **2016**, *4*, 38. **DOI:**10.3389/fchem.2016.00038
15. G. Marsico, V. S. Chambers, A. B. Sahakyan, P. McCauley, J. M. Boutell, M. Di Antonio, S. Balasubramanian, *Nucleic Acids Res.* **2019**, *47(8)*, 3862–3874. **DOI:**10.1093/nar/gkz179
16. J. L. Huppert, S. Balasubramanian, *Nucleic Acids Res.* **2005**, *33(9)*, 2908–2916. **DOI:**10.1093/nar/gki609
17. A. K. Todd, M. Johnston, S. Neidle, *Nucleic Acids Res.* **2005**, *33(9)*, 2901–2907. **DOI:**10.1093/nar/gki553
18. M. Kaushik, S. Kaushik, A. Bansal, S. Saxena, S. Kukreti, *Curr. Mol. Med.* **2011**, *11(9)*, 744–769. **DOI:**10.2174/156652411798062421

19. G. Wu, L. Chen, W. Liu, D. Yang, *Molecules* **2019**, *24(8)*, 1578. **DOI:**10.3390/molecules24081578
20. L. I. Jansson, J. Hentschel, J. W. Parks, T. R. Chang, C. Lu, R. Baral, C. R. Bagshaw, M. D. Stone, *Proc. Natl. Acad. Sci. U. S. A.* **2019**, *116(19)*, 9350–9359. **DOI:**10.1073/pnas.1814777116
21. A. Küçükakdağ Doğu, Ö Persil Çetinkol, *Turk. J. Chem.* **2019**, *43(4)*, 1040–1051. **DOI:**10.3906/kim-1903-6
22. S. K. Wang, Y. Wu, T. M. Ou, *Curr. Topics Med. Chem.* **2015**, *15(19)*, 1947–1956. **DOI:**10.2174/1568026615666150515145733
23. R. Tippana, M. C. Chen, N. A. Demeshkina, A. R. Ferré-D'Amaré, S. Myong, *Nature Comm.* **2019**, *10(1)*, 1855. **DOI:**10.1038/s41467-019-09802-w
24. D. Miyoshi, S. Matsumura, S. Nakano, N. Sugimoto, *J. Am. Chem. Soc.* **2004**, *126(1)*, 165–169. **DOI:**10.1021/ja036721q
25. X. Li, Y. H. Peng, J. S. Ren, X. G. Qu, *Proc. Natl. Acad. Sci. U. S. A.* **2006**, *103(52)*, 19658–19663. **DOI:**10.1073/pnas.0607245103
26. M. Fry, *Front. Biosci.* **2007**, *12*, 4336–4351. **DOI:**10.2741/2391
27. H. Y. Han, C. L. Cliff, L. H. Hurley, *Biochemistry* **1999**, *38(22)*, 6981–6986. **DOI:**10.1021/bi9905922
28. A. De Cian, J. L. Mergny, *Nucleic Acids Res.* **2007**, *35(8)*, 2483–2493. **DOI:**10.1093/nar/gkm098
29. D. Y. Sun, K. X. Guo, J. J. Rusche, L. H. Hurley, *Nucleic Acids Res.* **2005**, *33(18)*, 6070–6080. **DOI:**10.1093/nar/gki917
30. F. Kouzine, D. Levens, *Front. Biosci.* **2007**, *12*, 4409–4423. **DOI:**10.2741/2398
31. P. Rawal, V. B. Kummarasetti, J. Ravindran, N. Kumar, K. Halder, R. Sharma, M. Mukerji, S. K. Das, S. Chowdhury, *Genome Res.* **2006**, *16(5)*, 644–655. **DOI:**10.1101/gr.4508806
32. F. Kouzine, S. Sanford, Z. Elisha-Feil, D. Levens, *Nat. Struct. Mol. Biol.* **2008**, *15(2)*, 146–154. **DOI:**10.1038/nsmb.1372
33. A. K. Shukla, K. B. Roy, *J. Biochem.* **2006**, *139(1)*, 35–39. **DOI:**10.1093/jb/mvj009
34. A. T. Phan, J. L. Mergny, *Nucleic Acids Res.* **2002**, *30(21)*, 4618–4625. **DOI:**10.1093/nar/gkf597
35. S. A. Dvorkin, A. I. Karsisiotis, M. Webba da Silva, *Sci. Adv.* **2018**, *4(8)*, eaat3007. **DOI:**10.1126/sciadv.aat3007
36. S. Asamitsu, M. Takeuchi, S. Ikenoshita, Y. Imai, H. Kashiwagi, N. Shioda, *Int. J. Mol. Sci.* **2019**, *20(12)*, E2884. **DOI:**10.3390/ijms20122884
37. J. Ida, S. K. Chan, J. Glökler, Y. Y. Lim, Y. S. Choong, T. S. Lim, *Molecules* **2019**, *24(6)*, E1079. **DOI:**10.3390/molecules24061079
38. C. K. Kwok, C. J. Merrick, *Trends Biotech.* **2017**, *35(10)*, 997–1013. **DOI:**10.1016/j.tibtech.2017.06.012
39. J. Spiegel, S. Adhikari, S. Balasubramanian, *Trends Chem.* **2020**, *2(2)*, 123–136. **DOI:**10.1016/j.trechm.2019.07.002
40. K. G. Zyner, D. S. Mulhearn, S. Adhikari, S. Martinez Cuesta, M. Di Antonio, N. Erard, G. J. Hannon, D. Tannahill, S. Balasubramanian, *Elife* **2019**, *8*, e46793. **DOI:**10.7554/eLife.46793
41 J. Bidzinska, G. Cimino-Reale, N. Zaffaroni, M. Folini, *Molecules* **2013**, *18(10)*, 12368–12395. **DOI:**10.3390/molecules181012368
42. D. Yang, K. Okamoto, *Future Med. Chem.* **2010**, *2(4)*, 619–646. **DOI:**10.4155/fmc.09.172
43. S. Asamitsu, S. Obata, Z. Yu, T. Bando, H. Sugiyama, *Molecules* **2019**, *24(3)*, e429. **DOI:**10.3390/molecules24030429
44. S. Asamitsu, T. Bando, H. Sugiyama, *Chem. Eur. J.* **2019**, *25(2)*, 417–430. **DOI:**10.1002/chem.201802691
45. D. Musumeci, C. Riccardi, D. Montesarchio, *Molecules* **2015**, *20(9)*, 17511–17532. **DOI:**10.3390/molecules200917511
46. E. Ruggiero, S. N. Richter, *Nucleic Acids Res.* **2018**, *46(7)*, 3270–3283. **DOI:**10.1093/nar/gky187
47. Z. Y. Sun, X. N. Wang, S. Q. Cheng, X. X. Su, T. M. Ou, *Molecules* **2019**, *24*, 396. **DOI:**10.3390/molecules24030396
48. K. W. Zheng, S. Xiao, J. Q. Liu, J. Y. Zhang, Y. H. Hao, Z. Tan, *Nucleic Acids Res.* **2013**, *41(10)*, 5533–5541. **DOI:**10.1093/nar/gkt264
49. B. Pagano, S. Cosconati, V. Gabelica, L. Petraccone, S. De Tito, L. Marinelli, V. La Pietra, F. S. di Leva, I. Lauri, R. Trotta, E. Novellino, C. Giancola, A. Randazzo, *Curr. Pharm. Des.* **2012**, *18(14)*, 1880–1899. **DOI:**10.2174/138161212799958332
50. A. Nuthanakanti, I. Ahmed, S. Y. Khatik, K. Saikrishnan, S. G. Srivatsan, *Nucleic Acids Res.* **2019**, *47(12)*, 6059–6072. **DOI:**10.1093/nar/gkz419
51. M. Chen, G. Song, C. Wang, D. Hu, J. Ren, X. Qu, *Biophys. J.* **2009**, *97(7)*, 2014–2023. **DOI:**10.1016/j.bpj.2009.07.025
52. V. Viglasky, L. Bauer, K. Tluckova, P. Javorsky, *J. Nucleic Acids* **2010**, *2010*, Article ID 820356. **DOI:**10.4061/2010/820356
53. R. D. Gray, L. Pettracone, R. Buscaglia, J. B. Chaires, *Methods Mol. Biol.* **2010**, *608*, 121–136. **DOI:**10.1007/978-1-59745-363-9_8
54. B. DeMarco, S. Stefanovic, A. Williams, K. R. Moss, B. R. Anderson, G. J. Bassell, M. R. Mihailescu, *PLoS ONE* **2019**, *14(5)*, e0217275. **DOI:**10.1371/journal.pone.0217275
55. M. I. Umar, D. Ji, C. Y. Chan, C. K. Kwok, *Molecules* **2019**, *24*, 2416. **DOI:**10.3390/molecules24132416
56. R. C. Monsen, J. O. Trent, *Biochimie* **2018**, *152*, 134–148. **DOI:**10.1016/j.biochi.2018.06.024
57. S. Haider, *J. Ind. Institute Sci.* **2018**, *98(3)*, 325–339. **DOI:**10.1007/s41745-018-0083-3
58. V. Brazda, J. Kolomaznik, J. Lysek, M. Bartas, M. Fojta, J. Stastny, J. L. Mergny, *Bioinformatics* **2019**, *35(18)*, 3493–3495. **DOI:**10.1093/bioinformatics/btz087
59. A. Pavlov, P. M. Mitrasinovic, *Curr. Org. Chem.* **2010**, *14(2)*, 129–137. **DOI:**10.2174/138527210790069866
60. L. Petraccone, N. C. Garbett, J. B. Chaires, J. O. Trent, *Biopolymers* **2010**, *93(6)*, 533–548. **DOI:**10.1002/bip.21392
61. Q. Hou, S. B. Chen, J. H. Tan, H. B. Luo, D. Li, L. Q. Gu, Z. H. Huang, *J. Comput. Aided Mol. Des.* **2012**, *26*, 1355–1368. **DOI:**10.1007/s10822-012-9619-1
62. J. Husby, A. K. Todd, J. A. Platts, S. Neidle, *Biopolymers* **2013**, *99(12)*, 989–1005. **DOI:**10.1002/bip.22340
63. F. S. Di Leva, E. Novellino, A. Cavalli, M. Parrinello, V. Limongelli, *Nucleic Acids Res.* **2014**, *42(9)*, 5447–5455. **DOI:**10.1093/nar/gku247
64. J. K. Zhou, D. Y. Yang, S. Y. Sheu, *Phys. Chem. Chem. Phys.* **2015**, *17(19)*, 12857–12869. **DOI:**10.1039/C5CP00378D
65. F. Moraca, J. Amato, F. Ortuso, A. Artese, B. Pagano, E. Novel-

lino, S. Alcaro, M. Parrinello, V. Limongelli, *Proc. Natl. Acad. Sci. U. S. A.* **2017**, *114(11)*, E2136–E2145. **DOI:**10.1073/pnas.1612627114
66. P. M. Mitrasinovic, *J. Biomol. Struct. Dyn.* **2018**, *36(9)*, 2292–2302. **DOI:**10.1080/07391102.2017.1358670
67. B. Machireddy, H. J. Sullivan, C. Wu, *Molecules* **2019**, *24(6)*, E1010. **DOI:**10.3390/molecules24061010
68. P. M. Mitrasinovic, *Croat. Chem. Acta* **2019**, *92(1)*, 43–57. **DOI:**10.5562/cca3456
69. C. L. Grand, T. J. Powell, R. B. Nagle, D. J. Bearss, D. Tye, M. Gleason-Guzman, L. H. Hurley, *Proc. Natl. Acad. Sci. U. S. A.* **2004**, *101(16)*, 6140–6145. **DOI:**10.1073/pnas.0400460101
70. A. Siddiqui-Jain, C. L. Grand, D. J. Bearss, L. H. Hurley, *Proc. Natl. Acad. Sci. U. S. A.* **2002**, *99(18)*, 11593–11598. **DOI:**10.1073/pnas.182256799
71. E. V. Prochownik, P. K. Vogt, *Genes & Cancer* **2010**, *1(6)*, 650–659. **DOI:**10.1177/1947601910377494
72. S. Balasubramanian, L. H. Hurley, S. Neidle, *Nat. Rev. Drug Discov.* **2011**, *10(4)*, 261–275. **DOI:**10.1038/nrd3428
73. D. Sun, K. Guo, J. J. Rusche, L. H. Hurley, *Nucleic Acids Res.* **2005**, *33(18)*, 6070–6080. **DOI:**10.1093/nar/gki917
74. T. S. Dexheimer, D. Sun, L. H. Hurley, *J. Am. Chem. Soc.* **2006**, *128(16)*, 5404–5415. **DOI:**10.1021/ja0563861
75. S. Rankin, A. P. Reszka, J. Huppert, M. Zloh, G. N. Parkinson, A. K. Todd, S. Ladame, S. Balasubramanian, S. Neidle, *J. Am. Chem. Soc.* **2005**, *127(30)*, 10584–10589. **DOI:**10.1021/ja050823u
76. S. L. Palumbo, S. W. Ebbinghaus, L. H. Hurley, *J. Am. Chem. Soc.* **2009**, *131(31)*, 10878–10891. **DOI:**10.1021/ja902281d
77. S. Cogoi, F. Quadrifoglio, L. E. Xodo, *Biochemistry* **2004**, *43(9)*, 2512–2523. **DOI:**10.1021/bi035754f
78. R. K. Morgan, H. Batra, V. C. Gaerig, J. Hockings, T. A. Brooks, *Biochim. Biophys. Acta (BBA) Gene Regul. Mech.* **2016**, *1859(2)*, 235–245. **DOI:**10.1016/j.bbagrm.2015.11.004
79. R. G. Ramsay, T. J. Gonda, *Nat. Rev. Cancer* **2008**, *8(7)*, 523–534. **DOI:**10.1038/nrc2439
80. R. K. Moyzis, J. M. Buckingham, L. S. Cram, M. Dani, L. L. Deaven, M. D. Jones, J. Meyne, R. L. Ratliff, J. R. Wu, *Proc. Natl. Acad. Sci. U. S. A.* **1988**, *85(18)*, 6622–6626. **DOI:**10.1073/pnas.85.18.6622
81. W. E. Wright, V. M. Tesmer, K. E. Huffman, S. D. Levene, J. W. Shay, *Genes Dev.* **1997**, *11(21)*, 2801–2809. **DOI:**10.1101/gad.11.21.2801
82. A. J. Sfeir, W. H. Chai, J. W. Shay, W. E. Wright, *Mol. Cell* **2005**, *18(1)*, 131–138. **DOI:**10.1016/j.molcel.2005.02.035
83. J. Dai, M. Carver, C. Punchihewa, R. A. Jones, D. Z. Yang, *Nucleic Acids Res.* **2007**, *35(15)*, 4927–4940. **DOI:**10.1093/nar/gkm522
84. L. Petraccone, C. Spink, J. O. Trent, N. C. Garbett, C. S. Mekmaysy, C. Giancola, J. B. Chaires, *J. Am. Chem. Soc.* **2011**, *133(51)*, 20951–20961. **DOI:**10.1021/ja209192a
85. J. Dai, M. Carver, D. Yang, *Biochimie* **2008**, *90*, 1172–1183. **DOI:**10.1016/j.biochi.2008.02.026
86. V. Brázda, L. Hároníková, J. C. Liao, M. Fojta, *Int. J. Mol. Sci.* **2014**, *15(10)*, 17493–17517. **DOI:**10.3390/ijms151017493
87. B. J. Chen, Y. L. Wu, Y. Tanaka, W. Zhang, *Int. J. Biol. Sci.* **2014**, *10(10)*, 1084–1096. **DOI:**10.7150/ijbs.10190
88. S. Neidle, *Curr. Opin. Struct. Biol.* **2009**, *19*, 239–250. **DOI:**10.1016/j.sbi.2009.04.001
89. Y. Qin, L. H. Hurley, *Biochimie* **2008**, *90(8)*, 1149–1171. **DOI:**10.1016/j.biochi.2008.02.020
90. D. J. Patel, A. T. Phan, V. Kuryavyi, *Nucleic Acids Res.* **2007**, *35(22)*, 7429–7455. **DOI:**10.1093/nar/gkm711
91. L. Oganesian, T. M. Bryan, *Bioessays* **2007**, *29(2)*, 155–165. **DOI:**10.1002/bies.20523
92. J. L. Mergny, J. F. Riou, P. Mailliet, M. P. Teulade-Fichou, E. Gilson, *Nucleic Acids Res.* **2002**, *30(4)*, 839–865. **DOI:**10.1093/nar/30.4.839
93. J. Cuesta, M. A. Read, S. Neidle, *Mini Rev. Med. Chem.* **2003**, *3(1)*, 11–21. **DOI:**10.2174/1389557033405502
94. A. De Cian, L. Lacroix, C. Douarre, N. Temime-Smaali, C. Trentesaux, J. F. Riou, J. L. Mergny, *Biochimie* **2008**, *90(1)*, 131–155. **DOI:**10.1016/j.biochi.2007.07.011
95. M. A. Shammas, R. J. Shmookler Reis, C. Li, H. Koley, L. H. Hurley, K. C. Anderson, N. C. Munshi, *Clin. Cancer Res.* **2004**, *10(2)*, 770–776. **DOI:**10.1158/1078-0432.CCR-0793-03
96. S. N. Georgiades, N. H. Abd Karim, K. Suntharalingam, R. Vilar, *Angew. Chem. Int. Ed.* **2010**, *49*, 4020–4034. **DOI:**10.1002/anie.200906363
97. D. Muench, F. Rezzoug, S. D. Thomas, J. Xiao, A. Islam, D. M. Miller, K. C. Sedoris, *PLoS ONE* **2019**, *14(1)*, e0211046. **DOI:**10.1371/journal.pone.0211046
98. C. Lin, G. Wu, K. Wang, B. Onel, S. Sakai, Y. Shao, D. Yang, *Angew. Chem. Int. Ed.* **2018**, *57*, 10888–10893. **DOI:**10.1002/anie.201804667
99. C. Marchetti, K. G. Zyner, S. A. Ohnmacht, M. Robson, S. M. Haider, J. P. Morton, G. Marsico, T. Vo, S. Laughlin-Toth, A. A, Ahmed, G. Di Vita, I. Pazitna, M. Gunaratnam, R. J. Besser, A. C. G. Andrade, S. Diocou, J. A. Pike, D. Tannahill, R. B. Pedley, T. R. J. Evans, W. D. Wilson, S. Balasubramanian, S. Neidle, *J. Med. Chem.* **2018**, *61(6)*, 2500–2517. **DOI:**10.1021/acs.jmedchem.7b01781
100. P. M. Mitrasinovic, *Acta Chim. Slov.* **2020**, *67*, 1–10. **DOI**:10.17344/acsi.2019.5105
101. G. Sattin, A. Artese, M. Nadai, G. Costa, L. Parrotta, S. Alcaro, M. Palumbo, S. N. Richter, *PLoS ONE* **2013**, *8(12)*, e84113. **DOI:**10.1371/journal.pone.0084113
102. R. Rocca, F. Moraca, G. Costa, M. Nadai, M. Scalabrin, C. Talarico, S. Distinto, E. Maccioni, F. Ortuso, A. Artese, S. Alcaro, S. N. Richter, *Biochim. Biophys. Acta (BBA) Gen. Subj.* **2017**, *1861(5PtB)*, 1329–1340. **DOI:**10.1016/j.bbagen.2016.12.023
103. D. Musumeci, C. Riccardi, D. Montesarchio, *Molecules* **2015**, *20(9)*, 17511–17532. **DOI:**10.3390/molecules200917511
104. C. Platella, C. Riccardi, D. Montesarchio, G. N. Roviello, D. Musumeci, *Biochim. Biophys. Acta (BBA) Gen. Subj.* **2017**, *1861(5PtB)*, 1429–1447. **DOI:**10.1016/j.bbagen.2016.11.027
105. W. Wang, Y. Sui, L. Zhang, W. Tan, X. He, X. Xie, *Can. J. Chem.* **2016**, *94(1)*, 60–65. **DOI:**10.1139/cjc-2015-0215
106. M. Tassinari, A. Lena, E. Butovskaya, V. Pirota, M. Nadai, M.

Freccero, F. Doria, S. N. Richter, *Molecules* **2018**, *23(8)*, 1874. **DOI:**10.3390/molecules23081874
107. H. Yu, X. Wang, M. Fu, J. Ren, X. Qu, *Nucleic Acids Res.* **2008**, *36(17)*, 5695–5703. **DOI:**10.1093/nar/gkn569
108. H. Qin, C. Zhao, Y. Sun, J. Ren, X. Qu, *J. Am. Chem. Soc.* **2017**, *139(45)*, 16201–16209. **DOI:**10.1021/jacs.7b07490
109. M. Hu, S. Chen, B. Wang, T. Ou, L. Gu, J. Tan, Z. Huang, *Nucleic Acids Res.* **2016**, *45(4)*, 1606–1618. **DOI:**10.1093/nar/gkw1195
110. A. Tawani, S. K. Mishra, A. Kumar, *Sci. Rep.* **2017**, *7(1)*, 3600. **DOI:**10.1038/s41598-017-03906-3
111. D. Dutta, M. Debnath, D. Müller, R. Paul, T. Das, I. Bessi, H. Schwalbe, J. Dash, *Nucleic Acids Res.* **2018**, *46(11)*, 5355–5365. **DOI:**10.1093/nar/gky385
112. M. Hu, Y. Wang, Z. Yu, L. Hu, T. Ou, S. Chen, Z. Huang, J. Tan, *J. Med. Chem.* **2018**, *61(6)*, 2447–2459. **DOI:**10.1021/acs.jmedchem.7b01697
113. K. M. Felsenstein, L. B. Saunders, J. K. Simmons, E. Leon, D. R. Calabrese, S. Zhang, A. Michalowski, P. Gareiss, B. A. Mock, J. S. Jr. Schneekloth, *ACS Chem. Biol.* **2015**, *11(1)*, 139–148. **DOI:**10.1021/acschembio.5b00577
114. D. Panda, P. Saha, T. Das, J. Dash, *Nat. Commun.* **2017**, *8*, 16103. **DOI:**10.1038/ncomms16103
115. J. Amato, A. Pagano, D. Capasso, S. Di Gaetano, M. Giustiniano, E. Novellino, A. Randazzo, B. Pagano, *ChemMedChem* **2018**, *13(5)*, 406–410. **DOI:**10.1002/cmdc.201700749
116. E. Zorzan, S. Da Ros, C. Musetti, L. Z. Shahidian, N. F. R. Coelho, F. Bonsembiante, S. Létard, M. E. Gelain, M. Palumbo, P. Dubreuil, M. Giantin, C. Sissi, M. Dacasto, *Oncotarget* **2016**, *7(16)*, 21658–21675. **DOI:**10.18632/oncotarget.7808
117. H. Kang, Y. Cui, H. Yin, A. Scheid, W. P. Hendricks, J. Schmidt, A. Sekulic, D. Kong, J. M. Trent, V. A. Gokhale, *J. Am. Chem. Soc.* **2016**, *138(41)*, 13673–13692. **DOI:**10.1021/jacs.6b07598
118. J. Lavrado, H. Brito, P. M. Borralho, S. A. Ohnmacht, N. Kim, C. Leitão, S. Pisco, M. Gunaratnam, C. M. Rodrigues, R. Moreira, S. Neidle, A. Paulo, *Sci. Rep. 2015*, *5*, 9696. **DOI:**10.1038/srep09696
119. F. Li, J. Zhou, M. Xu, G. Yuan, *Int. J. Biol. Macromol.* **2018**, *107(PtB)*, 1474–1479. **DOI:**10.1016/j.ijbiomac.2017.10.010
120. R. Perrone, M. Nadai, I. Frasson, J. A. Poe, E. Butovskaya, T. E. Smithgall, M. Palumbo, G. Palu, S. N. Richter, *J. Med. Chem.* **2013**, *56(16)*, 6521–6530. **DOI:**10.1021/jm400914r
121. R. Perrone, E. Butovskaya, D. Daelemans, G. Palu, C. Pannecouque, S. N. Richter, *J. Antimicrob. Chemother.* **2014**, *69(12)*, 3248–3258. **DOI:**10.1093/jac/dku280
122. D. Piekna-Przybylska, G. Sharma, S. B. Maggirwar, R. A. Bambara, *Cell Cycle* **2017**, *16(10)*, 968–978. **DOI:**10.1080/15384101.2017.1312225
123. R. Perrone, M. Nadai, J. A. Poe, I. Frasson, M. Palumbo, G. Palu, T. E. Smithgall, S. N, Richter, *PLoS ONE* **2013**, *8(8)*, e73121. **DOI:**10.1371/journal.pone.0073121
124. R. Perrone, F. Doria, E. Butovskaya, I. Frasson, S. Botti, M. Scalabrin, S. Lago, V. Grande, M. Nadai, M. Freccero, S. N. Richter, *J. Med. Chem.* **2015**, *58(24)*, 9639–9652. **DOI:**10.1021/acs.jmedchem.5b01283
125. B. Biswas, M. Kandpal, U. K. Jauhari, P. Vivekanandan, *BMC Genomics* **2016**, *17(1)*, 949. **DOI:**10.1186/s12864-016-3282-1
126. S. Artusi, M. Nadai, R. Perrone, M. A. Biasolo, G. Palu, L. Flamand, A. Calistri, S. N. Richter, *Antiviral Res.* **2015**, *118*, 123–131. **DOI:**10.1016/j.antiviral.2015.03.016
127. S. Callegaro, R. Perrone, M. Scalabrin, F. Doria, G. Palu, S. N. Richter, *Sci. Rep.* **2017**, *7(1)*, 2341. **DOI:**10.1038/s41598-017-02667-3
128. J. Norseen, F. B. Johnson, P. M. Lieberman, *J. Virol.* **2009**, *83(20)*, 10336–10346. **DOI:**10.1128/JVI.00747-09
129. P. Murat, J. Zhong, L. Lekieffre, N. P. Cowieson, J. L. Clancy, T. Preiss, S. Balaswwwubramanian, R. Khanna, J. Tellam, *Nat. Chem. Biol.* **2014**, *10(5)*, 358–364. **DOI:**10.1038/nchembio.1479
130. M. J. Lista, R. P. Martins, G. Angrand, A. Quillevere, C. Daskalogianni, C. Voisset, M. P. Teulade-Fichou, R. Fahraeus, M. Blondel, *Microb. Cell* **2017**, *4(9)*, 305–307. **DOI:**10.15698/mic2017.09.590
131. A. Madireddy, P. Purushothaman, C. P. Loosbroock, E. S. Robertson, C. L. Schildkraut, S. C. Verma, *Nucleic Acids Res.* **2016**, *44(8)*, 3675–3694. **DOI:**10.1093/nar/gkw038
132. S. Gilbert-Girard, A. Gravel, S. Artusi, S. N. Richter, N. Wallaschek, B. B. Kaufer, L. Flamand, *J. Virol.* **2017**, *91(14)*, e00402–17. **DOI:**10.1128/JVI.00402-17
133. S. R. Wang, Y. Q. Min, J. Q. Wang, C. X. Liu, B. S. Fu, F. Wu, L. Y. Wu, Z. X. Qiao, Y. Y. Song, G. H. Xu, Z. G. Wu, G. Huang, N. F. Peng, R. Huang, W. X. Mao, S. Peng, Y. Q. Chen, Y. Zhu, T. Tian, X. L. Zhang, X. Zhou, *Sci. Adv.* **2016**, *2(4)*, e1501535. **DOI:**10.1126/sciadv.1501535
134. K. Tluckova, M. Marusic, P. Tothova, L. Bauer, P. Sket, J. Plavec, V. Viglasky, *Biochemistry* **2013**, *52(41)*, 7207–7216. **DOI:**10.1021/bi400897g
135. A. M. Fleming, Y. Ding, A. Alenko, C. J. Burrows, *ACS Infect. Dis.* **2016**, *2(10)*, 674–681. **DOI:**10.1021/acsinfecdis.6b00109
136. B. Biswas, M. Kandpal, P. Vivekanandan, *Nucleic Acids Res.* **2017**, *45(19)*, 11268–11280. **DOI:**10.1093/nar/gkx823
137. S. R. Wang, Q. Y. Zhang, J. Q. Wang, X. Y. Ge, Y. Y. Song, Y. F. Wang, X. D. Li, B. S. Fu, G. H. Xu, B. Shu, P. Gong, B. Zhang, T. Tian, X. Zhou, *Cell Chem. Biol.* **2016**, *23(9)*, 1113–1122. **DOI:**10.1016/j.chembiol.2016.07.019

Povzetek

G-kvadrupleksi (G4) so nekanonske sekundarne strukture, ki se zvijejo znotraj vijačnic, bogatih z gvaninom (G), v regulatornih genskih regijah. Nedavni dokazi kažejo na njihovo tesno vključenost v pomembne biološke procese, kot so vzdrževanje telomer, zaščita koncev vijačnic, stabilnost kromosomov, izražanje genov, integracija virusov in rekombinacija. Mehanistične podrobnosti, kako in zakaj strukture G4 vplivajo na biološko funkcijo, kažejo na utemeljenost obravnave G4 kot potencialnih molekulskih tarč za bodoče terapevtike. Z drugimi besedami, strukturna heterogenost z natančno določenimi vezavnimi mesti, termična stabilnost in pogostnost G4 v telomerih, onkogenskih promotorskih regijah in virusnih genomih naredijo G4 za privlačne tarče za majhne molekule, katerih cilj je selektivno prepoznavanje med vsemi drugimi strukturami nukleinskih kislin, zlasti dupleksne oblike, ki so v genomu najbolj pogoste. V članku je predstavljen kritičen pregled dobro opisanih ligandov, ki interagirajo z G4, kot potencialnih orodij za zdravljenje raka in protivirusnih terapij. Učinki, ki jih ti ligandi selektivno izvajajo v *in vitro* in *in vivo* modelih, so povzeti. Predstavljeni so edinstveni ligandi, ki sodelujejo v specifičnem prepoznavanju G4. Ključno vprašanje, kako oblikovati in razviti nove G4 specifične ligande, ki ustrezajo strukturnim in fizikalno-kemijskim zahtevam za optimalno biološko aktivnost, je obravnavano ob upoštevanju izjemnega napredka v zadnjih nekaj letih in naših nedavnih prispevkov.