doi:10.2533/chimia.2009.134 Chimia 63 (2009) 134–139

Targeting G-Quadruplex DNA with Small Molecules

Nathan W. Luedtke*

Abstract: Despite a recent explosion of interest in G-quadruplex DNA, most fundamental questions regarding the possible presence and function of these intriguing structures *in vivo* remain unanswered. Cell-permeable G-quadruplex specific probes should provide researchers with new tools for studying these alternately folded DNA structures and their potential involvement in gene expression, chromosome stability, viral integration, and recombination. In this review, a survey of the binding affinity, specificity, and biological/pharmacological activities of some well-characterized G-quadruplex ligands is presented along with the potential importance of G-quadruplex DNA *in vivo*.

Keywords: DNA · Fluorescence · G-quadruplex · Nucleic acids · Tetraplex

In 1996 Nathan Luedtke received his bachelor's degree from the University of Washington where he investigated DNA alkylation with Paul Hopkins and DNA replication with Walt Fangman. Nathan received his doctoral degree in 2003 from the University of California, San Diego where he studied RNA–small molecule binding interactions with Yitzhak Tor. From 2003–2006 Nathan worked as an NIH Postdoctoral Fellow at Yale University with Alanna Schepartz where he used cell-permeable fluorophores to study protein folding in living cells. Since October 2006, Nathan has been an Assistant Professor of Organic Chemistry at the University of Zürich where he is investigating the chemistry and biology of unusual DNA structures.

1. Introduction

First discovered in 1869 by the Swiss biochemist Friedrich Miescher,[1] DNA continues to amaze and intrigue those devoted to its study. The Watson and Crick model for B-form duplex DNA,[2] which was based on over 80 years of accumulated experimental data, dramatically improved our mechanistic understanding of the replication and flow of genetic information termed "the central dogma of molecular biology" by Francis Crick in 1958.[3] Since then, the B-form double helix has usually been regarded as the biologically relevant structure of DNA, however in the context of nucleosome core particles and other nucleoprotein complexes DNA can adopt a wide variety of conformations including highly distorted A-, B-, Z-form double helices.[4]

By necessity, DNA structures are highly dynamic and their associated functions are potentially diverse. In addition to various duplex structures, single-stranded DNAs can fold into a wide variety of hairpin, triplex, G-quadruplex, and i-motif structures containing non-canonical base pairs.[5] It is estimated that approximately 200 such structures are deposited in the protein data bank as of January 2009. While the casual reader might question the biological relevance of single-stranded DNA, it is important to remember that the metabolically active form(s) of DNA (*e.g.* during replication, transcription, repair, and recombination) are single-stranded, allowing kinetic access to alternately folded DNA structures *in vivo*. One might even take the extreme point of view that the double helix is actually the inactive form of DNA and that single stranded DNAs and their protein complexes are more biologically relevant.

While single-stranded DNA can exhibit structural diversity comparable in some ways to that of RNA, DNA's potential for forming large complex structures is limited by the energetic driving forces of duplex formation. Single-stranded DNA viral genomes and telomeric DNA circles are known exceptions to this rule.[6] The thermodynamic and kinetic barriers of duplex dissociation, a prerequisite for single-stranded DNA folding, might be off-set in part or in whole by single-stranded DNA binding proteins,[7] chaperones,[8,9] small-molecule binding,[10–12] and/or negative supercoiling.[12–16]

Polypurine–polypyrimidine tracts and other repetitive sequences can form non-duplex and/or higher-order chromatin structures possibly related to a wide variety of biological activities. One family of structures originating from guanosine-rich tracts is referred to as G-quadruplex (or G-tetraplex). These structures, defined as containing two or more stacked G-tetrads (or G-quartets, Fig. 1), are selectively stabilized by potassium ions at concentrations (10–50 mM) well below the 120 mM of KCl found in most cells types.[17–19] G-tetrads can be assembled in an intramolecular (backfolded) fashion or from two-, three-, or four- strands in intermolecular structures capable of adopting a wide variety of conformations and folding energies (Fig. 2).[20–26] Two of the 26 (or more) examples of intramolecular G-quadruplex topologies are shown in Fig 2.[21,27] While G-quadruplexes can exhibit thermodynamic stabilities comparable to those of the corresponding duplex structures,[19,28–30] the presence and function(s) of these structures *in vivo* remains an open question. Intermolecular G-quadruplex structures have been proposed as intermediates or precursors of recombination and/or viral integration,[23,31,32] while intramolecular G-quadruplexes have been implicated in the

*Correspondence: Prof. Dr. N. W. Luedtke
University of Zurich
Institute of Organic Chemistry
Winterthurerstrasse 190
CH-8057 Zurich
Tel.: +41 44 635 4244
Fax: +41 44 635 6891
E-mail: luedtke@oci.uzh.ch

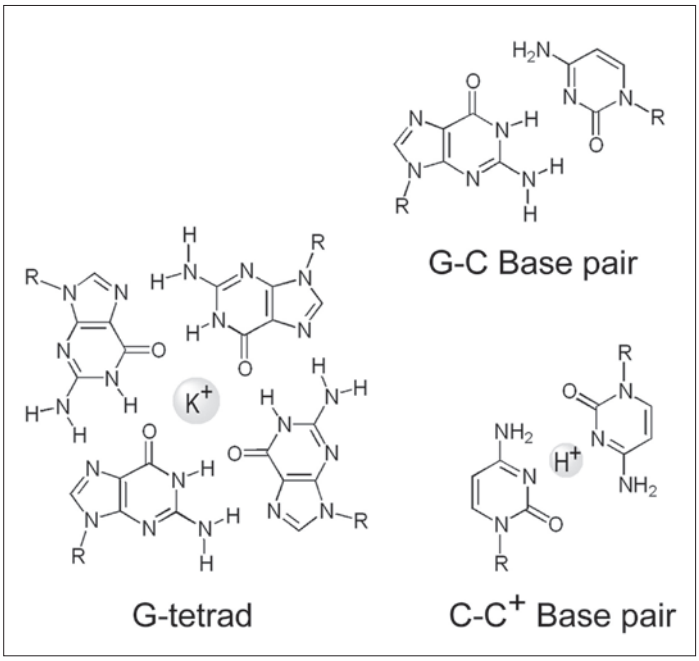

Fig. 1. Duplex DNA contains canonical G–C base pairs, while i-motif and G-quadruplex DNAs contain hemi-protonated C–C$^+$ base pairs and G-tetrads, respectively.

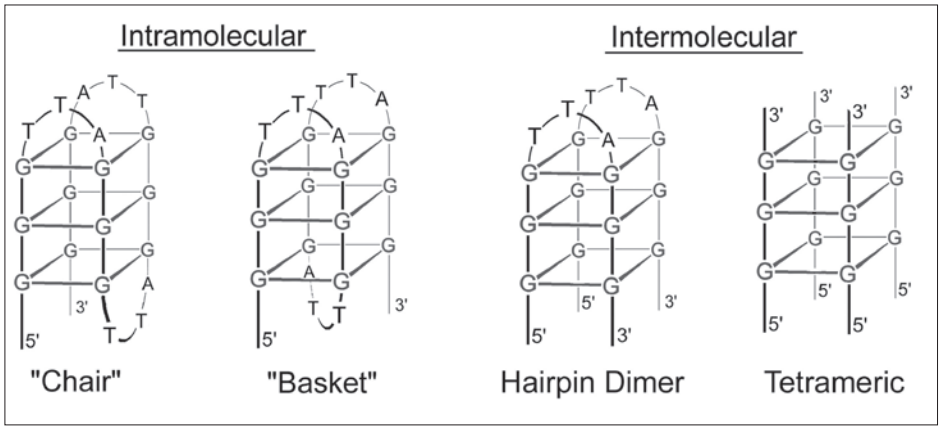

Fig. 2. G-tetrads can be assembled from a single strand of DNA in an intramolecular (backfolded) fashion where two of the 26 (or more) examples of intramolecular G-quadruplex topologies called 'chair' and 'basket' are shown.[20–27] Alternatively, two-, three-, or four DNA strands can associate in intermolecular G-quadruplex structures.

regulation of gene expression and chromosome stability.[33–35]

No consensus sequence for G-quadruplex folding has been experimentally determined, but approximately 370'000 sequences with the putative G-quadruplex-forming motif $(G_{3+}N_{1-7})_{4+}$ (where G = guanosine and N = any base) are dispersed throughout the human genome.[36,37] These sequences are concentrated in promoter regions,[14,38,39] introns,[36] 5' and 3' UTRs,[40] and at the ends of eukaryotic chromosomes.[41,42] G-rich sequences having this motif can fold into highly stable intramolecular G-quadruplex structures ($T_m \geq 90$ °C)[19] that usually fold and unfold very slowly at room temperature *in vitro* ($k_f \sim 0.001$ s^{-1}).[22] But much faster rates of folding have been observed when G-quadruplex DNA was imbedded in a molecular matrix at temperatures slightly below the T_m ($k_f \sim 1.0$ s^{-1}).[43] Molecular crowding,[44] synthetic and endogenous chaperones,[8–12] and dehydrating conditions inside the nucleus might also accelerate the rates of G-quadruplex formation *in vivo*.[22]

Due to the self-complementary nature of duplex DNA, approximately 370'000 C-rich motifs $(C_{3+}N_{1-7})_{4+}$ (where C = cytosine and N = any base) are present in the human genome.[36,37] These C-rich sequences can, under slightly acidic conditions (pH = 6), fold into a type of tetraplex called 'i-motif' that contains hemi-protonated, intercalated C–C$^+$ base pairs (Fig. 1 and Fig. 3). Almost nothing is known about the existence or biological relevance of i-motif DNA *in vivo*, but carboxylated carbon nanotubes have recently been shown to bind and stabilize i-motif DNA at pH 7 and above.[45] These results highlight the possibility that i-motifs might be formed under physiological conditions *in vivo* through molecular binding and/or crowding interactions.[30]

While direct and indirect evidence supports the presence and biological relevance of G-quadruplex structures located in the 3' single-stranded overhangs of ciliate telomeres *in vivo* (presented in Section 4), much less information is available regarding the potential relevance of the equilibrium depicted in Fig. 3, where certain G–C rich DNA sequences can exist as a mixture of G-quadruplex/i-motif and canonical duplex DNA *in vitro*.[29,30] Cell-permeable ligands that selectively bind to one or more of these structures might provide a means for probing the existence and/or controlling the function(s) associated with these structures. These same interactions might also provide a new source of therapeutic agents and targets.

2. G-Quadruplex DNA as a Potential Drug Target

Many currently used chemotherapeutic agents bind to DNA non-specifically (*e.g.* cisplatin, mitomycin C, chloroethyl nitrosoureas, daunomycin, *etc.*). The development of small molecules that specifically bind to a particular DNA secondary structure may improve cancer-specific targeting and decrease the side effects associated with chemotherapeutic treatments. G-quadruplex DNA structures are highly attractive targets given the abundance of detailed information available regarding their structures,[20,21] thermodynamic stabilities,[22] and potential biological activities – some of which might be considered cancer-specific.[21,31,33–35,46]

A growing number of groups are targeting G-quadruplex DNA with small molecules hoping to inhibit cancer growth according to two distinct mechanisms (Fig. 4). First, the over-expression of oncogenes like c-Myc, c-Kit, and KRAS might be inhibited by promoter deactivation (Section 3).[33] The second, more extensively studied mechanism, is the inhibition of telomerase, a ribonucleoprotein complex that catalyzes the 3' extension of telomeric DNA (Sections 4 and 5).[20,35,41,42,46,47]

3. Evidence for G-Quadruplex DNA as a Modulator of Gene Expression

The expansion of G-containing trinucleotide repeat sequences like (CGG), (GAA), and (GGA) is correlated with an increased probability of developing inheritable diseases such as Friedreich's ataxia, myotonic dystrophy, and Fragile X syndrome.[48] These sequences are also highly susceptible to S1 nuclease cleavage,[49] DNA repair,[50] and can fold into stable non-duplex structures *in vitro*.[48,51,52] In the case of a 14-nucleotide repeated sequence

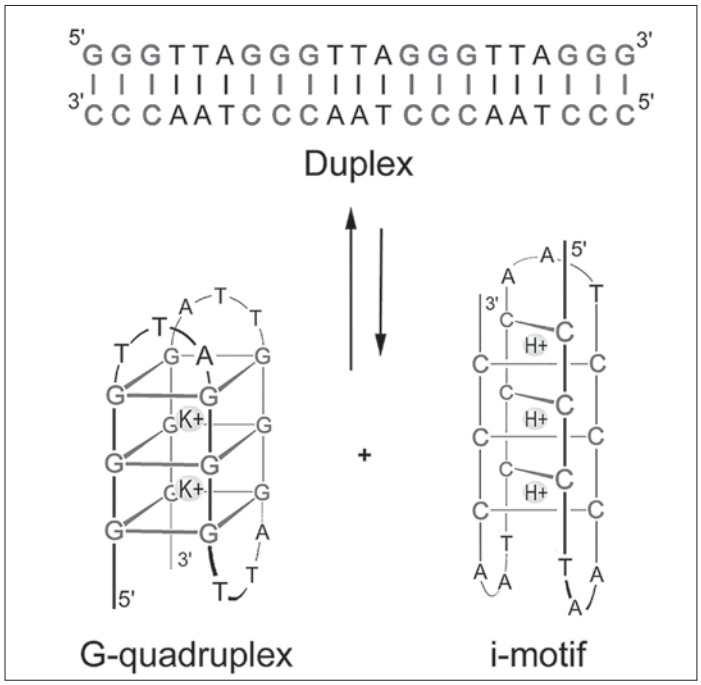

Fig. 3. The equilibrium between duplex and G-quadruplex/i-motif structures can be influenced by DNA binding proteins,[7] small-molecule binding,[10–12] negative supercoiling,[12–16] changes in pH and temperature,[29] and molecular crowding.[30]

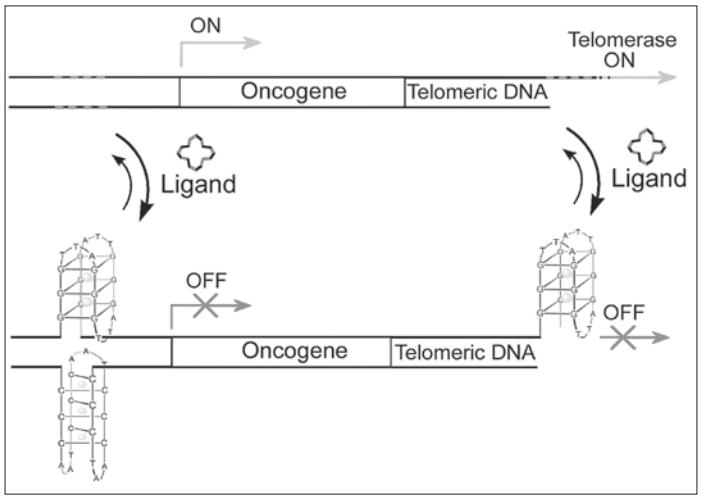

Fig. 4. Ligand-mediated stabilization of G-quadruplex DNA might facilitate the regulation of gene expression and/or the inhibition of telomerase activity.

located 363 bp upstream of the human insulin gene, a smaller number of repeats is correlated with a higher frequency of type-2 diabetes mellitus.[53] Mutations capable of disrupting the formation of G-quadruplex structures in this mini-satellite also disrupted expression of the insulin gene.[54–56] These studies were the first to suggest a direct relationship between the formation of G-quadruplexes in promoter regions and the suppression of gene transcription. Recently, a trinucleotide repeat (GGA) capable of forming G-quadruplex structure(s) was shown to have both transcriptional activation and repressor activities in the context of the c-MYB promoter.[52]

Another early example of a putative G-quadruplex-containing promoter is the non-repetitive nuclease hypersensitive element 'NHEIII$_1$' located 142 bp upstream of the P1 promoter of the c-MYC gene. By using a combination of DNA point mutations and small molecule binding interactions, a correlation between G-quadruplex stabilization and the suppression of promoter activity was established by the Hurley Lab.[57] When 0.1 mM of the cationic porphyrin TMPyP4 (presented in detail in Section 6) was incubated with cells, c-MYC promoter activity decreased by 50% or more. When a cell line with the NHEIII$_1$ deleted from the c-MYC promoter was used, little if any promoter suppression was observed after addition of TMPyP4.[57] Since then a number of other ligands that exhibit a wide range of G-quadruplex affinity and specificity have been shown to bind promoter-derived G-quadruplex *in vitro* and suppress the transcription of various genes in cell cultures.[47,58–60] Some, but not all, of the genes suppressed by TMPyP4 contain a putative G-quadruplex forming sequence in or near the promoter.[60]

Approximately 30–40% of human promoters contain a putative G-quadruplex motif.[39,61] These sequences consistently exhibit nuclease hypersensitivity and the ability to form G-quadruplex structures *in vitro*.[33,39] Since many of the genes containing potential quadruplex structures are overexpressed in cancer tissues, the deactivation of certain promoters by cell-permeable G-quadruplex ligands may provide a new route to anti-cancer therapies.[33] These results highlight the importance of developing highly selective G-quadruplex ligands that can differentiate between closely related G-quadruplex structures, as the deactivation of a large number of genes will result in non-specific toxicity like in the case of TMPyP4.[60,62] It remains an open question, however, if G-quadruplex structures in or near promoters are 'normal' regulatory elements utilized by transcriptional control machinery *in vivo*.

4. Evidence for Telomeric G-Quadruplex DNA

Pioneering work by Muller and McClintock in the 1930s revealed chromosomal ends were naturally endowed with the ability to prevent fusions with broken chromosomes and other telomeric ends.[63] Decades of subsequent studies have revealed this telomeric 'capping' function is related to a wide variety of telomere associated proteins,[64,65] attachment to the nuclear matrix,[66,67] and higher-order chromatin structures.[68–70] It should be noted that large differences in telomere length,[71] elongation mechanism(s),[6] and different telomeric end structures have been observed,[6,69] even in closely related species and cell types.

The telomeric ends of all known eukaryotic chromosomes contain a 3' single-stranded overhang.[41,63,72] These overhangs contain repetitive G-rich sequences such as $(G_{2-3}(TG)_{1-6}T)_n$ for *Saccharomyces*, $(TGGGGT)_n$ for *Tetrahymena*, $(TTTTGGGG)_n$ for *Stylonychia*, $(TTAGGG)_n$ for *Arabidopsis,* and $(TTAGGG)_n$ for vertebrates. In the late 1980s G-quadruplex structures derived from 3' telomeric overhangs were first reported.[23–26] Since then, many groups have speculated about potential relationships between telomeric G-quadruplex structures and chromosomal end capping activities.[34,35] Interestingly, nearly all reported telomeric sequences can fold into G-quadruplex structures *in vitro*.[20–26] While it may be an incredible coincidence that diverse telomeric sequences can fold into a family of closely related structures, there is growing evidence that G-quadruplex DNA plays a direct role in telomere structure and stability in certain organisms.[73–75] Direct evidence for the existence of G-quadruplex DNA at the telomeric ends of macronuclei in *Stylonychia lemnae* has been revealed *in vivo* using G-quadru-

plex-specific antibodies.[75] Subsequent studies using a combination of RNAi and antibodies have provided a direct link between telomeric end-binding proteins and the formation of G-quadruplex structures in ciliates.[73] Direct evidence supporting or refuting the formation of G-quadruplex structures in vertebrates has so far remained elusive. Indirect evidence for their existence has been inferred from competitive binding interactions between G-quadruplex-selective small molecules and endogenous binders of the 3' overhang like POT1 and telomerase.[46,74] Some studies have also claimed co-localization of G-quadruplex ligands with telomeric ends of metaphase spreads.[76,77] Further studies are needed to definitively prove (or disprove) the existence and biological function of telomeric G-quadruplex DNA in vertebrates.

Telomeric end structures have recently been resolved at medium resolution using electron microscopy and two-dimensional gel electrophoresis.[68–70] These studies showed that telomeric ends can invade upstream duplexes to form a 'T-loop' – even in protein-free samples.[69] At first glance, the T-loop and G-quadruplex telomeric ends appear to be mutually exclusive structures, but preliminary evidence suggests that a G-quadruplex might exist at the junction between the single-stranded and double-stranded regions of the T-loop.[78,79] The exact relationship, if any, between T-loop and G-quadruplex DNA structures *in vivo* has not yet been determined.[35]

5. G-Quadruplex Ligands as Telomerase Inhibitors

Human telomeres are nucleoprotein complexes containing the repeated DNA sequence $(^{5'}GGGTTA^{3'})_n$ (n = 100–4,000) as a long duplex ending with a short (24–400 base) single-stranded 3' overhang.[80,81] Telomeric DNA gets shortened by each successive round of cellular replication, possibly resulting in a molecular clock that limits the total number of replicative cycles a normal cell can undergo.[82,83] To overcome this 'end replication problem', most cancer tissues (~85%) maintain genomic stability by telomerase-mediated extension of telomeric DNA.[84] Since telomerase is over expressed in the majority of malignant tumor cells and in relatively few somatic cells,[84] it is widely recognized as a potential cancer-specific target.[20,35,41,42,46,47] With the knowledge that folded G-quadruplex structures are inactive substrates for telomerase,[85] many researchers have developed small-molecule inhibitors of telomerase that stabilize G-quadruplex-folded structures in the 3' overhangs of telomeric DNA. Indeed, small molecule-mediated stabilization of the G-quadruplex structure(s) can effectively inhibit telomerase activity, and when applied to cells, G-quadruplex ligands can initiate apoptosis or replicative senescence. Many of the G-quadruplex ligands used in prior studies, like TMPyP4, can bind to many other nucleic acid structures,[86] making alternate mechanisms for their reported biological effects possible.[60,62,87]

Targeting telomeric G-quadruplex DNA might influence telomere function beyond the inhibition of telomerase.[46] Chromosomal ends are associated with a wide variety of proteins that bind to double-stranded and/or single-stranded regions of the telomeric DNA.[7,35] This nucleoprotein or 'shelterin' complex maintains the structural integrity of telomeres *in vivo*.[64] Small molecules that bind DNA and displace proteins from the shelterin complex can cause telomere destabilization, a possible genotoxicity associated with many G-quadruplex ligands.[74] How this type of activity might be cancer-specific remains an open question.[46]

6. Some Known G-quadruplex Ligands and Their Activities

Detailed structural analyses of G-quadruplex-ligand complexes by NMR and X-ray crystallography have demonstrated at least two types of binding sites for G-quadruplex ligands.[20,88,89] The most common is co-facial end-stacking or 'hemi intercalation' of the ligand onto one or both of the terminal G-tetrads. Other binding sites are defined by the surface features of the grooves and/or loop regions. In both cases, subtle variations of G-quadruplex topologies, groove widths, and loop sequences can facilitate selective binding interactions even between closely related G-quadruplex structures.[58,90] It is important to note, however, that G-quadruplexes are highly dynamic structures and losses in DNA conformational entropy upon ligand binding might be an important factor in determining specificity. G-quadruplexes that exist as a dynamic mixture of conformations in the unbound state, like the human telomeric sequence,[91] should generally exhibit lower ligand affinities as compared to G-quadruplexes that adopt a single conformation.[58,90]

While a large number of G-quadruplex ligands have been reported in the literature,[41,42,46,89,92–94] the cationic porphyrin TMPyP4 (Fig. 5) is the most extensively studied to date. TMPyP4 inhibits both telomerase ($IC_{50} \cong 0.7–10$ μM)[95,96] and *Taq* DNA polymerase ($IC_{50} \cong 2$ μM).[57] Telomerase inhibition by TMPyP4 is relatively insensitive to metal coordination by the porphyrin, but is highly dependent on the groups at the *meso* positions.[97] Extensive substitution of the *meso* positions with groups other than pyridinium failed to generate compounds with improved activities, but the resulting structure–activity relationships demonstrated that base stacking and charge-charge interactions are important for porphyrin-DNA binding.[97] Subsequent X-ray crystallography and NMR studies have shown that TMPyP4 can bind to G-quadruplex DNA at many different positions, including the terminal G-tetrads,[98] and the loops, grooves, and phosphodiester backbone.[99] TMPyP4 has exhibited some promising anticancer activities *in vivo*,[62,100] but it has very poor DNA specificity,[86] causes anaphase bridges in sea urchin embryos,[87] and is highly toxic *in vivo*.[62,101] When administered at 0.1 mM (a dose near its toxic limit in cell cultures) TMPyP4 has been shown to have a dramatic effect on gene expression,[60] and it down-regulates transcription of both c-Myc and the catalytic subunit of telomerase.[62] TMPyP4 binds to duplex, triplex, G-quadruplex, single-stranded, and bulk genomic DNA with similar affinities ($K_d \cong 200$ nM),[86,102] and therefore it cannot be considered a structure-selective ligand.

The anionic porphyrin N-methyl mesoporphyrin (NMM), which is commercially available as a mixture of four regioisomers, has excellent G-quadruplex specificity,[86] but relatively poor G-quadruplex affinity ($K_d \cong 2–10$ μM).[103,104] Changes in the fluorescence properties of NMM can be used to detect G-quadruplex DNA *in vitro*.[103] Recently, NMM was found to up-regulate genes having promoters with a high potential for G-quadruplex formation, and it also suppressed rDNA activity in yeast.[38] The binding of rDNA (known to have a high potential for G-quadruplex formation) and the subsequent suppression of ribosome biogenesis is the reported mode of action of CX-3543 (quarfloxin), a small molecule drug candidate currently in phase II clinical trials.[105]

Telomestatin (SOT-095), a natural product isolated from *Streptomyces anulatus* 3533-SV4,[106] is one of the strongest and most specific inhibitors of telomerase reported to date ($IC_{50} \cong 1$ μM).[95,96] Telomestatin has molecular dimensions similar to those of G-tetrad DNA and can bind to various G-quadruplexes with modest affinity ($K_d \cong 30$ nM).[107] Telomestatin exhibits good selectivity for intramolecular versus intermolecular G-quadruplex structures,[108] and it has a 70-fold lower affinity for duplex DNA.[108] Telomestatin induces telomere shortening in treated cells more rapidly than is expected for a single mechanism involving telomerase inhibition.[109,110] Recent studies have shown that telomere uncapping and the

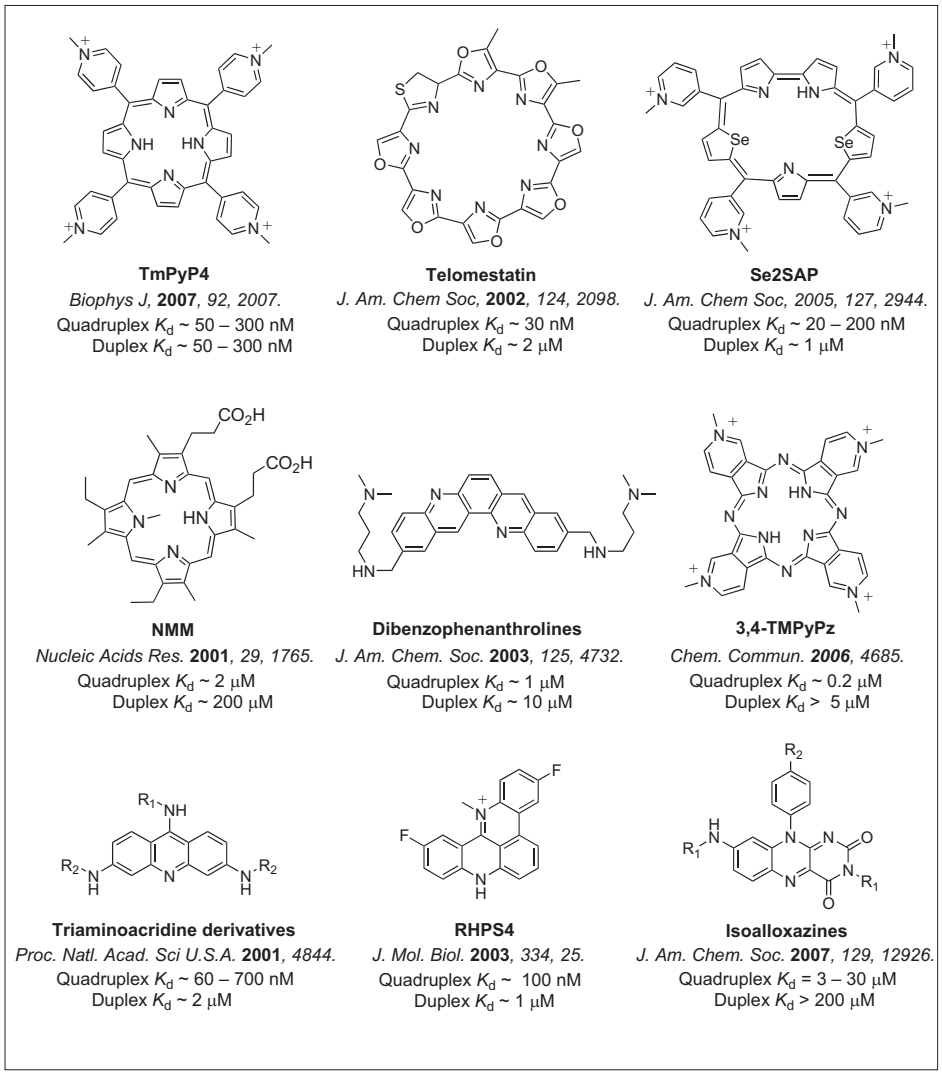

Fig. 5. Some of the known G-quadruplex ligands and their estimated dissociation constants for duplex and G-quadruplex DNA. Comprehensive surveys of G-quadruplex ligands have appeared in recent review articles.[89,94]

loss of telomeric DNA is related to the competition between telomestatin and POT1 – a shelterin protein that binds to the 3' single-stranded overhang.[74] While it is unknown if this type of activity might be cancer-selective, telomestatin induces senescence and apoptosis in a number of different tumor cell types and exhibits less toxicity towards normal progenitor cells.[108–111] The administration of telomestatin to mice containing U937 xenografts reduced tumor volumes without causing signs of toxicity.[112] A total synthesis of telomestatin has recently been reported,[113] but little is known about the physicochemical and structural features important for its DNA binding and pharmacological activities. No SARs have been reported for telomestatin, due, in part, to the extreme difficulties associated with the synthesis of its highly constrained macrocyclic core.[114] Synthetic analogs of telomestatin containing six or seven oxazole units have recently been prepared and show good G-quadruplex affinity and specificity.[115,116]

7. Concluding Remarks

In recent decades, great strides have been made towards the *sequence*-specific targeting of nucleic acids with antisense oligonucleotides and polyamide minor groove binders.[117,118] These molecules have facilitated highly selective gene silencing, and have found numerous applications in basic research and medicine. The pharmacological potential of these compounds, however, is severely limited by the size of molecule needed (typically 2,000–5,000 Da) to achieve sequence specificity in the context of the human genome. Compared to duplex DNA, G-quadruplex and i-motif structures have much more compact structures that contain well-defined binding sites for small molecules.

Small molecules capable of structure-selective DNA binding may provide an exciting new avenue for the development of anti-cancer agents and molecular probes. As novel therapeutics, they have the potential to extend and improve the lives of those suffering from one of the most devastating and common causes of premature death. A major impediment to the validation of G-quadruplex DNA as a new anti-cancer drug target is the lack of small molecules that bind to G-quadruplex DNA with high affinity ($K_d \leq 1$ nM) and high specificity (10'000-fold or lower affinity to all other nucleic acids). G-quadruplex-specific antibodies generated by *in vitro* evolution may provide key tools for target validation,[75,119] while the design and synthesis of new high affinity G-quadruplex ligands will provide new drug candidates and molecular probes. These molecules will provide researchers with new tools for studying the potential relationships between DNA folding and gene expression, chromosome stability, viral integration, and recombination.

Acknowledgements

I would like to thank Dr. Jawad Alzeer, Heidi W. Luedtke, Balayeshwanth R. Vummidi, and Dr. Oliver Scholz for their critical reading of this manuscript. Generous financial support provided by the Swiss National Science Foundation (grant #116868) and the University of Zürich is gratefully acknowledged.

Received: February 2, 2009

[1] R. Dahm, *Hum. Genet*. **2008**, *122*, 565.
[2] J. D. Watson, F. H. C. Crick, *Nature* **1953**, *171*, 737.
[3] F. Crick, *Nature* **1970**, *227*, 561.
[4] D. Svozil, J. Kalina, M. Omelka, B. Schneider, *Nucleic Acids Res*. **2008**, *36*, 3690.
[5] J. Suhnel, *Biopolymers* **2001**, *61*, 32.
[6] C. Groff-Vindman, A. J. Cesare, S. Natarajan, J. D. Griffith, M. J. McEachern, *Mol. Cell. Biol.* **2005**, *25*, 4406.
[7] M. Fry, *Front. Biosci.* **2007**, *12*, 4336.
[8] P. Armas, S. Nasif, N. B. Calcaterra, *J. Cell. Biochem.* **2008**, *103*, 1013.
[9] E. Baldrich, C. K. O'Sullivan, *Ana. Biochem.* **2005**, *341*, 194.
[10] H. Y. Han, C. L. Cliff, L. H. Hurley, *Biochemistry* **1999**, *38*, 6981.
[11] A. De Cian, J. L. Mergny, *Nucleic Acids Res*. **2007**, *35*, 2483.
[12] D. Y. Sun, K. X. Guo, J. J. Rusche, L. H. Hurley, *Nucleic Acids Res*. **2005**, *33*, 6070.
[13] F. Kouzine, D. Levens, *Front. Biosci.* **2007**, *12*, 4409.
[14] P. Rawal, V. B. R. Kummarasetti, J. Ravindran, N. Kumar, K. Halder, R. Sharma, M. Mukerji, S. K. Das, S. Chowdhury, *Genome Res.* **2006**, *16*, 644.
[15] F. Kouzine, S. Sanford, Z. Elisha-Feil, D. Levens, *Nat. Struct. Mol. Biol.* **2008**, *15*, 146.
[16] A. K. Shukla, K. B. Roy, *J. Biochem.* **2006**, *139*, 35.
[17] N. V. Hud, F. W. Smith, F. A. L. Anet, J. Feigon, *Biochemistry* **1996**, *35*, 15383.
[18] T. G. Voss, C. D. Fermin, J. A. Levy, S. Vigh, B. Choi, R. F. Garry, *J. Virol.* **1996**, *70*, 5447.
[19] A. Risitano, K. R. Fox, *Biochemistry* **2003**, *42*, 6507.
[20] D. J. Patel, A. T. Phan, V. Kuryavyi, *Nucleic Acids Res.* **2007**, *35*, 7429.
[21] S. Burge, G. N. Parkinson, P. Hazel, A. K. Todd, S. Neidle, *Nucleic Acids Res.* **2006**, *34*, 5402.
[22] A. N. Lane, J. B. Chaires, R. D. Gray, J. O. Trent, *Nucleic Acids Res*. **2008**, *36*, 5482.
[23] D. Sen, W. Gilbert, *Nature* **1988**, *334*, 364.

[24] J. R. Williamson, M. K. Raghuraman, T. R. Cech, *Cell* **1989**, *59,* 871.
[25] E. H. Blackburn, *Nature* **1991**, *350*, 569.
[26] W. I. Sundquist, A. Klug, *Nature* **1989**, *342*, 825.
[27] M. Webba da Silva, *Chemistry* **2007**, *13*, 9738.
[28] Y. Qin, E. M. Rezler, V. Gokhale, D. Sun, L. H. Hurley, *Nucleic Acids Res.* **2007**, *35*, 7698.
[29] A. T. Phan, J. L. Mergny, *Nucleic Acids Res*. **2002**, *30*, 4618.
[30] D. Miyoshi, S. Matsumura, S. Nakano, N. Sugimoto, *J. Am. Chem. Soc*. **2004**, *126*, 165.
[31] A. K. Shukla, K. B. Roy, *Biol. Chem.* **2006**, *387*, 251.
[32] V. De Fonzo, E. Bersani, F. Aluffi-Pentini, V. Parisi, *Med. Hypotheses* **2001**, *57*, 103.
[33] Y. Qin, L. H. Hurley, *Biochimie* **2008**, *90*, 1149.
[34] N. Maizels, *Nat. Struct. Mol. Biol*. **2006**, *13*, 1055.
[35] L. Oganesian, T. M. Bryan, *Bioessays* **2007**, *29*, 155.
[36] J. L. Huppert, S. Balasubramanian, *Nucleic Acids Res*. **2005**, *33*, 2908.
[37] A. K. Todd, M. Johnston, S. Neidle, *Nucleic Acids Res.* **2005**, *33*, 2901.
[38] S. G. Hershman, Q. Chen, J. Y. Lee, M. L. Kozak, P. Yue, L. S. Wang, F. B. Johnson, *Nucleic Acids Res*. **2008**, *36*, 144.
[39] J. L. Huppert, S. Balasubramanian, *Nucleic Acids Res.* **2007**, *35*, 406.
[40] J. L. Huppert, A. Bugaut, S. Kumari, S. Balasubramanian, *Nucleic Acids Res*. **2008**, *36*, 6260.
[41] J. L. Mergny, J. F. Riou, P. Mailliet, M. P. Teulade-Fichou, E. Gilson, *Nucleic Acids Res.* **2002**, *30*, 839.
[42] J. Cuesta, M. A. Read, S. Neidle, *Mini Rev. Med. Chem*. **2003**, *3*, 11.
[43] J. W. Shim, Q. Tan, L. Q. Gu, *Nucleic Acids Res*. **2008**, doi: 10.1093/nar/gkn968.
[44] D. Miyoshi, N. Sugimoto, *Biochimie* **2008**, *90*, 1040.
[45] X. Li, Y. H. Peng, J. S. Ren, X. G. Qu, *Proc. Natl. Acad. Sci. U. S. A*. **2006**, *103*, 19658.
[46] A. De Cian, L. Lacroix, C. Douarre, N. Temime-Smaali, C. Trentesaux, J. F. Riou, J. L. Mergny, *Biochimie* **2008**, *90*, 131.
[47] M. A. Shammas, R. J. S. Reis, C. Li, H. Koley, L. H. Hurley, K. C. Anderson, N. C. Munshi, *Clin. Cancer Res*. **2004**, *10*, 770.
[48] P. Catasti, X. Chen, S. V. S. Mariappan, E. M. Bradbury, G. Gupta, *Genetica* **1999**, *106*, 15.
[49] T. Aoki, K. S. Koch, H. L. Leffert, *J. Mol. Biol.* **1997**, *267*, 229.
[50] M. M. Slean, G. B. Panigrahi, L. P. Ranum, C. E. Pearson, *DNA Repair* **2008**, *7*, 1135.
[51] A. Bacolla, J. E. Larson, J. R. Collins, J. Li, A. Milosavljevic, P. D. Stenson, D. N. Cooper, R. D. Wells, *Genome Res*. **2008**, *18*, 1545.
[52] S. L. Palumbo, R. M. Memmott, D. J. Uribe, Y. Krotova-Khan, L. H. Hurley, S. W. Ebbinghaus, *Nucleic Acids Res*. **2008**, *36*, 1755.
[53] W. Bastian, *J. Ped. Endocrin. Metab*. **2002**, *15*, 471.
[54] G. C. Kennedy, M. S. German, W. J. Rutter, *Nat. Genetics* **1995**, *9*, 293.
[55] P. Catasti, X. Chen, R. K. Moyzis, E. M. Bradbury, G. Gupta, *J. Mol. Biol.* **1996**, *264*, 534.
[56] A. Lew, W. J. Rutter, G. C. Kennedy, *Proc. Natl. Acad. Sci. U. S. A.* **2000**, *97*, 12508.
[57] A. Siddiqui-Jain, C. L. Grand, D. J. Bearss, L. H. Hurley, *Proc. Natl. Acad. Sci. U. S. A*. **2002**, *99*, 11593.
[58] M. Bejugam, S. Sewitz, P. S. Shirude, R. Rodriguez, R. Shahid, S. Balasubramanian, *J. Am. Chem. Soc*. **2007**, *129*, 12926.
[59] D. Sun, W. J. Liu, K. X. Guo, J. J. Rusche, S. Ebbinghaus, V. Gokhale, L. H. Hurley, *Mol. Cancer Therapeut.* **2008**, *7*, 880.
[60] Y. Mikami-Terao, M. Akiyama, Y. Yuza, T. Yanagisawa, O. Yamada, H. Yamada, *Cancer Lett.* **2008**, *261*, 226.
[61] R. Zhang, Y. Lin, C. T. Zhang, *Nucleic Acids Res*. **2008**, *36*, D372.
[62] C. L. Grand, H. Y. Han, R. M. Munoz, S. Weitman, D. D. Von Hoff, L. H. Hurley, D. J. Bearss, *Mol. Cancer Therapeut.* **2002**, *1*, 565.
[63] A. D. Bolzan, M. S. Bianchi, *Rev. Mutat. Res*. **2006**, *612*, 189.
[64] T. de Lange, *Genes Dev.* **2005**, *19*, 2100.
[65] E. L. Denchi, T. de Lange, *Nature* **2007**, *448*, 1068.
[66] M. E. Luderus, B. vanSteensel, L. Chong, O. C. M. Sibon, F. F. M. Cremers, T. deLange, *J. Cell Biol.* **1996**, *135*, 867.
[67] F. Jonsson, J. Postberg, C. Schaffitzel, H. J. Lipps, *Chromosome Res.* **2002**, *10*, 445.
[68] J. D. Griffith, L. Comeau, S. Rosenfield, R. M. Stansel, A. Bianchi, H. Moss, T. de Lange, *Cell* **1999**, *97*, 503.
[69] K. G. Murti, D. M. Prescott, *Proc. Natl. Acad. Sci. U. S. A.* **1999**, *96*, 14436.
[70] J. L. Munoz-Jordan, G. A. M. Cross, T. de Lange, J. D. Griffith, *EMBO J.* **2001**, *20*, 579.
[71] E. Gilson, A. Londono-Vallejo, *Cell Cycle* **2007**, *6*, 2486.
[72] J. H. Kim, W. T. Kim, I. K. Chung, *Plant Mol. Biol*. **1998**, *36*, 661.
[73] K. Paeschke, T. Simonsson, J. Postberg, D. Rhodes, H. J. Lipps, *Nat. Struct. Mol. Biol.* **2005**, *12*, 847.
[74] D. Gomez, T. Wenner, B. Brassart, C. Douarre, M. F. O'Donohue, V. El Khoury, K. Shin-Ya, H. Morjani, C. Trentesaux, J. F. Riou, *J. Biol. Chem.* **2006**, *281*, 38721.
[75] C. Schaffitzel, I. Berger, J. Postberg, J. Hanes, H. J. Lipps, A. Pluckthun, *Proc. Natl. Acad. Sci. U. S. A.* **2001**, *98*, 8572.
[76] C. C. Chang, I. C. Kuo, I. F. Ling, C. T. Chen, H. C. Chen, P. J. Lou, J. J. Lin, T. C. Chang, *Anal. Chem*. **2004**, *76*, 4490.
[77] C. Granotier, G. Pennarun, L. Riou, F. Hoffschir, L. R. Gauthier, A. De Cian, D. Gomez, E. Mandine, J. F. Riou, J. L. Mergny, P. Mailliet, B. Dutrillaux, F. D. Boussin, *Nucleic Acids Res*. **2005**, *33*, 4182.
[78] Y. Xu, H. Sato, Y. Sannohe, K.-i. Shinohara, H. Sugiyama, *J. Am. Chem. Soc.* **2008**, *130*, 16470.
[79] Y. Xu, H. Sato, K.-i. Shinohara, M. Komiyama, H. Sugiyama, *Nucleic Acids Symp. Ser.* **2007**, *51*, 243.
[80] G. Cimino-Reale, E. Pascale, E. Battiloro, G. Starace, R. Verna, E. D'Ambrosio, *Nucleic Acids Res.* **2001**, *29*, E35.
[81] H. Riethman, *Ann. Rev. Genomics Hum. Genetics* **2008**, *9*, 1.
[82] S. A. Stewart, R. A. Weinberg, *Oncogene* **2002**, *21*, 627.
[83] C. Bischoff, H. C. Petersen, J. Graakjaer, K. Andersen-Ranberg, J. W. Vaupel, V. A. Bohr, S. Kolvraa, K. Christensen, *Epidemiology* **2006**, *17*, 190.
[84] N. W. Kim, M. A. Piatyszek, K. R. Prowse, C. B. Harley, M. D. West, P. L. Ho, G. M. Coviello, W. E. Wright, S. L. Weinrich, J. W. Shay, *Science* **1994**, *266*, 2011.
[85] A. M. Zahler, J. R. Williamson, T. R. Cech, D. M. Prescott, *Nature* **1991**, *350*, 718.
[86] J. Ren, J. B. Chaires, *Biochemistry* **1999**, *38*, 16067.
[87] E. Izbicka, D. Nishioka, V. Marcell, E. Raymond, K. K. Davidson, R. A. Lawrence, R. T. Wheelhouse, L. H. Hurley, R. S. Wu, D. D. Von Hoff, *Anticancer Drug Des.* **1999**, *14*, 355.
[88] S. Neidle, G. N. Parkinson, *Biochimie* **2008**, *90*, 1184.
[89] D. Monchaud, M. P. Teulade-Fichou, *Org. Biomol. Chem.* **2008**, *6*, 627.
[90] A. Arora, S. Maiti, *J. Phys. Chem. B* **2008**, *112*, 8151.
[91] J. X. Dai, M. Carver, D. Z. Yang, *Biochimie* **2008**, *90*, 1172.
[92] K. M. Doherty, S. Sharma, R. Gupta, R. M. Brosh, Jr., *Recent Pat. Anticancer Drug Discov.* **2006**, *1*, 185.
[93] J. H. Tan, L. Q. Gu, J. Y. Wu, *Mini Rev. Med. Chem.* **2008**, *8*, 1163.
[94] T. M. Ou, Y. J. Lu, J. H. Tan, Z. S. Huang, K. Y. Wong, L. Q. Gu, *ChemMedChem* **2008**, *3*, 690.
[95] J. Reed, M. Gunaratnam, M. Beltran, A. P. Reszka, R. Vilar, S. Neidle, *Anal. Biochem*. **2008**, *380*, 99.
[96] A. De Cian, G. Cristofari, P. Reichenbach, E. De Lemos, D. Monchaud, M. P. Teulade-Fichou, K. Shin-Ya, L. Lacroix, J. Lingner, J. L. Mergny, *Proc. Natl. Acad. Sci. U. S. A*. **2007**, *104*, 17347.
[97] D. F. Shi, R. T. Wheelhouse, D. Sun, L. H. Hurley, *J. Med. Chem*. **2001**, *44*, 4509.
[98] A. T. Phan, V. Kuryavyi, H. Y. Gaw, D. J. Patel, *Nat. Chem. Biol.* **2005**, *1*, 167.
[99] G. N. Parkinson, R. Ghosh, S. Neidle, *Biochemistry* **2007**, *46*, 2390.
[100] A. Villanueva, G. Jori, *Cancer Lett.* **1993**, *73*, 59.
[101] C. H. Sibata, V. C. Colussi, N. L. Oleinick, T. J. Kinsella, *Expt. Opin. Pharmacother*. **2001**, *2*, 917.
[102] M. W. Freyer, R. Buscaglia, K. Kaplan, D. Cashman, L. H. Hurley, E. A. Lewis, *Biophys. J.* **2007**, *92*, 2007.
[103] H. Arthanari, S. Basu, T. L. Kawano, P. H. Bolton, *Nucleic Acids Res.* **1998**, *26*, 3724.
[104] X. Wu, N. Maizels, *Nucleic Acids Res*. **2001**, *29*, 1765.
[105] P. Bates, J. L. Mergny, D. Yang, *EMBO Repts.* **2007**, *8*, 1003.
[106] K. Shin-ya, K. Wierzba, K. Matsuo, T. Ohtani, Y. Yamada, K. Furihata, Y. Hayakawa, H. Seto, *J. Am. Chem. Soc.* **2001**, *123*, 1262.
[107] M. Y. Kim, H. Vankayalapati, K. Shin-Ya, K. Wierzba, L. H. Hurley, *J. Am. Chem. Soc.* **2002**, *124*, 2098.
[108] M. Y. Kim, M. Gleason-Guzman, E. Izbicka, D. Nishioka, L. H. Hurley, *Cancer Res*. **2003**, *63*, 3247.
[109] D. Gomez, N. Aouali, A. Renaud, C. Douarre, K. Shin-Ya, J. Tazi, S. Martinez, C. Trentesaux, H. Morjani, J. F. Riou, *Cancer Res.* **2003**, *63*, 6149.
[110] T. Tauchi, K. Shin-Ya, G. Sashida, M. Sumi, A. Nakajima, T. Shimamoto, J. H. Ohyashiki, K. Ohyashiki, *Oncogene* **2003**, *22*, 5338.
[111] 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*, 770.
[112] T. Tauchi, K. Shin-ya, G. Sashida, M. Sumi, S. Okabe, J. H. Ohyashiki, K. Ohyashiki, *Oncogene* **2006**, *25*, 5719.
[113] T. Doi, M. Yoshida, K. Shin-ya, T. Takahashi, *Org. Lett*. **2006**, *8*, 4165.
[114] J. Deeley, A. Bertram, G. Pattenden, *Org. Biomol. Chem.* **2008**, *6*, 1994.
[115] G. S. Minhas, D. S. Pilch, J. E. Kerrigan, E. J. LaVoie, J. E. Rice, *Bioorg. Med. Chem. Lett.* **2006**, *16*, 3891.
[116] D. S. Pilch, C. M. Barbieri, S. G. Rzuczek, E. J. Lavoie, J. E. Rice, *Biochimie* **2008**, *90*, 1233.
[117] I. Lebedeva, C. A. Stein, *Annu. Rev. Pharmacol. Toxicol.* **2001**, *41*, 403.
[118] P. B. Dervan, R. M. Doss, M. A. Marques, *Curr. Med. Chem. Anticancer Agents* **2005**, *5*, 373.
[119] H. Fernando, R. Rodriguez, S. Balasubramanian, *Biochemistry* **2008**, *47*, 9365.