“One ring to bind them all” – Part I: The efficiency of the macrocyclic scaffold for G-quadruplex DNA recognition

David Monchaud,†,§ Anton Granzhan,† Nicolas Saettel,‡ Aurore Guédin,† Jean-Louis Mergny,§ and Marie-Paule Teulade-Fichou†,*

†Institut Curie, Section Recherche, CNRS UMR176, Centre Universitaire Paris XI, Bat. 110, 91405 Orsay, France; §Current address: Institut de Chimie Moléculaire, CNRS UMR5260, Université de Bourgogne, 21000 Dijon, France; ‡INSERM, U565, Acides Nucléiques : Dynamique, Ciblage et Fonctions Biologiques & CNRS, UMR5153, Laboratoire des Régulations et Dynamique du Génome, Muséum National d’histoire Naturelle USM 503, 43, Rue Cuvier, 75005 Paris, France; †INSERM U869, Université de Bordeaux, Institut Européen de Chimie et Biologie, 33607 Pessac cedex, France.

Introduction

G-rich DNA strands are naturally involved in duplex-DNA architecture through their association with their complementary C-rich DNA strands by the canonical Watson-Crick pairing. However, a growing body of evidence currently testifies that this canonical association is not the unique mode of stabilization of G-rich DNA in cells. Indeed, given that four guanine residues can self-associate in a planar arrangement through a Hoogsteen type hydrogen-bonding network to form a structure called G-quartet, G-rich strands can adopt a peculiar three-dimensional arrangement called G-quadruplex-DNA resulting from the stacking of several contiguous G-quartets (Fig. 1).

The formation of G-quadruplex-DNA is easily conceivable in DNA sequences that are permanently present as single-strand in cells, such as the telomeric overhang. The structural and functional integrity of this overhang is based on its association with a constellation of specific proteins, some of them belonging to the shelterin complex. Altogether, this nucleoproteic assembly caps the chromosomes, protects their integrity and is also deeply involved in the telomeric replication process. Numerous studies currently suggest that quadruplex formation in this overhang alters the structure and function of telomeres, inducing a damage response and rapid apoptosis in particular in cancer cells. Thus, over the past decade telomeric G-quadruplex DNA has been thoroughly studied with an initial focus on the possible interferences with telomerase activity.

The existence of G-quadruplex-DNA is also heavily implied in the promoter region of genes and oncogenes, and is thus assumed to play an important regulatory role in their transcription. It has been indeed demonstrated that the involvement of G-rich sequences in duplex architecture is compatible with their folding into quadruplex-structures, thanks to the breathing of duplex DNA. Nevertheless, we have to keep in mind that these G-quadruplex
forming sequences although transiently single-stranded are a priori not as easily accessible as the telomeric sequences, due to the presence of the transcription machinery.\textsuperscript{10}

Stabilization of quadruplex architecture by small molecules is thus emerging as a potential anti-cancer approach since it is thought to interfere with oncogenic expression and telomeric maintenance in cancer cells.\textsuperscript{5} Interestingly, several classes of small molecules have been developed that efficiently target G-quadruplex-DNA.\textsuperscript{11} Among them, macrocycles rapidly became popular for recognition of quadruplexes motivated by the fact that telomestatin is one of the most active known G-quadruplex ligands.\textsuperscript{12} This natural 8-ring 24-membered macroheterocycle (Fig. 2) displays exceptional affinity for quadruplex and most importantly has no of affinity for duplex DNA (\textit{vide infra}). Nevertheless, apart from several reports on telomestatin-like molecules, true polyheteroaryl analogues are still scarce, essentially because of the difficulty to synthesise such molecules. Another well-known G-quadruplex ligand – and arguably one of the most studied – is TMPyP4,\textsuperscript{13} a tetramethylpyridinium porphyrin (Fig. 2). This molecule has been widely used, essentially due to its great affinity for several quadruplex targets, as well as its commercial availability. However, the interest in employing TMPyP4 is somewhat counterbalanced by its lack of quadruplex-selectivity (\textit{vide infra}).

These two examples illustrate the main reasons that made the macrocyclic scaffold particularly interesting for targeting G-quadruplex DNA: \textit{i} - a broad aromatic surface that favours the stacking interactions with the external G-quartets of the quadruplex; \textit{ii} - a rigid structure that maximizes the quartet overlap and, combined with their large size, impedes intercalation into duplex-DNA and \textit{iii} - in the case of TMPyP4, a cationic charge that promotes the electrostatic interactions with the negatively charged biopolymer. Unfortunately, the high cationic charge of TMPyP4 represents both an advantage and a drawback due to the non specificity of electrostatic interactions which promote association with any form of DNA, but in particular duplex-DNA, thereby decreasing the binding selectivity (\textit{vide infra}).

A great deal of research effort around the macrocyclic scaffold was thus motivated by these two examples. The macrocyclic systems described up to now as G-quadruplex ligands can be divided in three different categories: \textit{i} - telomestatin-like (\textit{i.e.} neutral and rigid macrocycles), \textit{ii} - porphyrin (TMPyP4)-like (\textit{i.e.} cationic and rigid macrocycles) and \textit{iii} - the less studied family of flexible polyammonium macrocycles (\textit{i.e.} cationic and non planar), which will be described herein but also in the companion paper, \textit{“One Ring to bind them all”} – Part II (A. Granzhan \textit{et al.}) in the present issue.

\textbf{Telomestatin-like macrocycles.}

Telomestatin is the benchmark compound in terms of G-quadruplex recognition (Fig. 2). This natural compound was isolated from \textit{Streptomyces anulatus} in 2001 by Shin-ya’s
group, and has been extensively studied due to its outstanding selectivity for G-quadruplex binding and highly promising biological properties.

One general method to quantify the DNA affinity of a given molecule is to perform thermal denaturation experiments, either in a UV-monitored melting assay or in the so-called “FRET melting” assay. Here the thermal unfolding of the quadruplex-forming oligonucleotide F21T (FAM-G5[T2AG5]3-Tamra), doubly-labeled with a complementary pair of FRET donor and acceptor (FAM: 6-carboxyfluorescein and Tamra: 6-carboxytetramethylrhodamine) is monitored via FRET (fluorescence resonance energy transfer). Semi-quantitative evaluation of ligand binding affinity is obtained by measuring the increase in melting temperature induced by the ligand (\(\Delta T_{1/2}\)). The quadruplex-over-duplex-DNA selectivity may then be established via competitive FRET-melting which is carried out in presence of competitive duplex-DNA (ds26). In the presence of telomestatin, the increase in melting temperature is very large (\(\Delta T_{1/2} = 24^\circ\text{C}\)), and since this value is unaffected by the presence of up to 50 equiv. of competitive duplex-DNA (herein ds26, the self-complementary sequence 5’-CA2TCG2ATCGA2T2CGATC2GAT2G-3’), telomestatin stands among the most selective of G-quadruplex ligands. This property has been subsequently confirmed by independent studies using various evaluation techniques.

The enthusiasm for telomestatin was also justified by its ability to inhibit telomerase, with the reported IC\(50\) value lying, impressively, in the nanomolar range (IC\(50\) TRAP = 5 nM). However, the relationship between TRAP assay and telomerase inhibition was recently challenged on the grounds that this assay does not actually reflect the influence of the ligand on the activity of the enzyme but rather the ability of the ligand to inhibit the PCR amplification step. The direct assay, a more constraining method based on the telomerase elongation of a telomeric primer with incorporation of [\(\alpha-^{32}\text{P}\)]dGTP, provides a more reliable estimation of the telomerase inhibition ability. Interestingly, telomestatin is still highly active, staying among the best reported inhibitors, with an IC\(50\) value of 58 nM.

The structure of telomestatin is quite unusual since it is a neutral polycyclic compound, composed of five oxazole, two methyloxazole and one thiazoline rings, whose overall flatness is impeded by the presence of an \(sp^3\) carbon within the thiazoline ring (Fig. 2). Whether this molecule is aggregated in water is yet not known, however, it is likely that the strong hydrophobic character of the molecule should reinforce the hydrophobic forces that contribute to stacking interactions, and so explain its exceptional effectiveness in binding G-quadruplex. This efficiency is also assumed to rely on a perfect shape adaptation between the macrocycle and a G-quartet (Fig. 3).
Fig 3. Qualitative in silico superposition of telomestatin and a G-tetrad (extracted from PDB entry: 2A5R); guanine residues appear in gold, the carbon, nitrogen, oxygen, sulfur and hydrogen atoms of telomestatin in grey, blue, red, yellow and white respectively.

Abundant biophysical and biological investigations have been performed with telomestatin. For example it exhibits an antiproliferative activity against a wide variety of cancer cell lines with IC$_{50}$ values between 0.1 and 5 µM, including human pancreatic carcinoma, myeloid leukaemia, breast cancer, cervical, tumoral HT1080, telomerase-transformed SW39, ALT-transformed SW26, immortalized EcR293, and a panel of myeloma and neuroblastoma lines. Remarkably, telomestatin does not affect normal cell lines, with no effect observed at 5 µM on fibroblast MRC-5 cells. This observation has recently been substantiated by the finding that telomestatin does not interfere with telomere replication in normal cells. Telomestatin was also a tool of choice to investigate the biological role of G-quadruplex ligands in cells by suggesting that effects of G-quadruplex ligands might originate essentially in the disruption of the telomeric structure with subsequent displacement of protective proteins (shelterin complex) rather than in telomerase inhibition. Telomestatin was used to investigate the role of putative quadruplex formation within the promoter region of genes like VEGF, and hTERT, or oncogenes like c-Myc, Bcl-2, and RET, whose corresponding proteins are over-expressed in some cancers. Recently, telomestatin was shown to interfere with the ability of helicases to unwind G-quadruplex structures, and to affect the growth of telomerase-negative ALT (Alternative Lengthening of Telomeres) cell lines via an indirect interaction with a proteinic complex comprised of the shelterin component TRF2, the helicase BLM and the enzyme Topoisomerase IIIα. The reason why telomestatin – or more generally G-quadruplex ligands – selectively affect tumor cell lines is yet not fully understood. However, differences in plasmic membrane permeability between normal and cancer cells or in the accessibility of the telomere and in particular variations in the composition of the shelterin complex have been proposed; even if the accessibility is identical, one can imagine that the cellular responses are different.

Unexpectedly, only two studies have addressed in detail the actual binding mode of telomestatin to G-quadruplex using in silico investigations. This is particularly surprising, given that there are no structural characterizations of this interaction (NMR, X-ray crystallography) available to date. These studies support concomitant double end-stacking on quartets, with a preference for the parallel conformation of the human telomeric quadruplex. This conclusion is somewhat in contradiction with other studies based on CD spectroscopy or 125I-radioprobing, underlining that further efforts are required to clarify the interaction mode of telomestatin with quadruplexes.
Nevertheless, one major concern is that telomestatin is difficult to synthesise. Despite considerable synthetic efforts to obtain either fragments of telomestatin or telomestatin-like arrays of polyoxazoles, its total synthesis has been reported only recently, and the complexity of the proposed pathway is incompatible with large-scale preparation. This can explain why, whilst the excellence of telomestatin towards quadruplex-recognition can hardly be exaggerated, relatively few reports on telomestatin-like compounds have appeared in the literature. Among them, two hexa-oxazole macrocyclic ligands have been independently reported by Rice et al. and by Shin-ya, Nagasawa et al. The structure of these macrocycles was based on the dimeric association of two symmetrical tri-oxazole moieties through an amino-acid linker, either a valine (for HXDV, Fig. 4) or a protected serine (for S2A2-6OTD, Fig. 4) respectively.

![](image1)

HVDV has been widely investigated and found to profoundly stabilise the telomeric G-quadruplex structure by UV-melting assay ($\Delta T_m = 24 ^\circ C$), without any significant binding to duplex DNA. Its association with a quadruplex-forming sequence mimicking the human telomeric sequence has recently been confirmed to be based on external stacking, with occupancy of both external G-quartets. Interestingly, the high quadruplex-selectivity of HXDV is assumed to originate from its particular telomestatin-like concave shape since the two planes defined by the trioxazole moieties present an angle of $\sim 30 ^\circ$, along with the disposition of the two amino-acid residues on the same face of the cycle thereby optimizing stacking interaction with quartets (Fig. 5). The biological effects of HXDV has been investigated on human lymphohlastoma (RPM1) and murine leukaemia (P388) cell lines, with IC$_{50}$ values for inhibition of the population doubling in the low micromolar range (0.4 and 0.5 $\mu$M, respectively). Very interestingly, it has also been shown recently that HXDV elicits anti-proliferative effect on both telomerase-positive (HeLa, A875, PC3-1 with IC$_{50}$ between 0.2 and 0.6 $\mu$M) and telomerase-negative cell lines (like SAOS2 or GM847 with IC$_{50}$ between 0.4 and 0.5 $\mu$M) mainly through an M-phase inhibitory effect, thereby contributing to the elucidation of the actual biological effects of G-quadruplex ligands.
Interestingly, to circumvent the solubility problems, various HXDV derivatives have been prepared based on the modification of the nature of the side-chain(s), from two valine residues (for HXDV), to only one valine,42 one valine and one protected lysine (HXLV-AC),43 and two lysine residues (HXDL, Fig. 4).44 Of all these derivatives, HXDL appeared as the most promising. Whilst it is not neutral at physiological pH (due to protonation of both terminal amine groups, whose pKa values are estimated at 9.9 and 10.5)45 and the characteristics of its association with quadruplex are impressive both in terms of affinity (ΔT_m = 49 °C) and selectivity since only residual stabilization of salmon testes duplex-DNA is observed. This gain in efficiency is assumed to originate in a higher solubility and in enhanced interactions with quadruplex-DNA, mixing both tetrad-stacking and electrostatic interactions.

The success obtained with HXDL was concomitantly – and independently – confirmed by Nagasawa et al., who developed a series of G-quadruplex binding macrocyclic hexaoxazoles (6OTD for 6-Oxazole Telomestatin Derivatives). The very first example of this series, S2A2-6OTD (Fig. 4),40 was rapidly surpassed by more active compounds, in which the serine residues of the side-chain were replaced by lysine (L2H2-6OTD, also called HXDL by Rice et al.) or arginine residues (L2G2-6OTD).46 Both compounds were found to be active in vitro in a range similar to that of telomestatin, notably via PCR stop assay (IC_{50} down to 0.64 µM), TRAP assay (IC_{50} down to 20 nM) and inhibition of the growth of HeLa cells (IC_{50} down to 0.5 µM). Current strategies for improvement include the dimerisation of the 6OTD scaffold (6OTD-dimer, Fig. 4),47 using a bis-amide linker between the two 6OTD moieties which is long enough span the thickness of the quadruplex architecture (~15 Å). Unexpectedly the affinity of the dimer so obtained is not significantly increased as compared with the corresponding monomer, as judged by FRET-melting data (ΔT_m = 25.1 vs 25.0 °C, respectively) and PCR stop assays (IC_{50} = 3.0 vs 2.9 µM, respectively). On the other hand, the dimer is far more selective than the monomer since a ~10-fold improvement is obtained when comparing the inhibitory IC_{50} values measured by PCR stop assay allowing the authors to claim a 800-fold selectivity. The fact that the dimer is able to form a unique 1:1 complex with G-quadruplex, as shown by ESI-MS analysis, added to a proven inability to interact with duplex-DNA, implying that this dimeric scaffold is optimized for quadruplex-recognition. Further improvement was found by replacing the hexaoxazole scaffold by the more planar heptaoxazole analogue (L1H1-7OTD, Fig. 4).48 This compound was shown to interact
strongly with quadruplex-DNA (as judged by PCR stop assay, IC$_{50} = 0.67$ $\mu$M), with a fair quadruplex- over duplex-DNA selectivity (~8-fold). Furthermore, it displays a selective cytotoxicity toward telomerase-positive cells (HeLa, IC$_{50} = 2.2$ $\mu$M) as compared to telomerase-negative cells (Saos-2, IC$_{50} > 30\mu$M), thus suggesting a telomerase-based mechanism of action. It is thus highly probable that a dimeric version of the 7OTD scaffold will appear in a near future and that it will help this scaffold to improve its quadruplex- over duplex-DNA selectivity.

Other series of oxazole-based macrocycles have been investigated as G-quadruplex binders and notably macrocycles whose structure is based on either a tris-oxazole,\textsuperscript{49} tris-furan or tetra-furan scaffold\textsuperscript{50} (Fig. 4). Amino-terminated chains have been introduced to improve watersolubility and favour electrostatic interactions with DNA (the pKa values of the amino side-chains depicted in Fig. 4 standing in the 8.8 – 10.7 range),\textsuperscript{45} with the chain length fine-tuned to reach the quadruplex grooves. The intrinsic qualities of these compounds have been evaluated through FRET-melting assay and whilst modest results were obtained with the oxazole-based tripeptide ($\Delta T_{1/2} = 6.4$ $^\circ$C), moderate ($\Delta T_{1/2} = 10$ $^\circ$C) to good ($\Delta T_{1/2} = 17$ $^\circ$C) results were obtained with furan-based tri- and tetra-peptides respectively. Interestingly, when performed with duplex-DNA, no stabilization occurred in any case, thus inferring a high quadruplex-selectivity of this series of ligands. But a real breakthrough was achieved in increasing the aromatic surface of the ligands, since stabilization higher than 30$^\circ$C were obtained with the aminoquinoline-based cyclic oligoamide depicted in Fig. 4.\textsuperscript{51} Once more, this compound retains a high selectivity for quadruplex-DNA, since no stabilization was detected when a FRET-melting assay was performed with duplex hairpin-DNA. It is worth noting that this compound is also able to discriminate various types of quadruplex-DNA (human telomeric vs oncogenic c-kit), thus opening perspectives for discovery of a novel generation of selective and specific G-quadruplex ligands. Additionally, the very attractive in vitro biophysical results were also complemented by the low general cytotoxicity of the cyclic oligopeptides.

Despite the real viability of the neutral macrocyclic scaffold for G-quadruplex recognition, the equilibrium between hydrosolubility and activity is difficult to achieve. One improvement has been the introduction of amino side-chains that increase water-solubility and electrostatic attraction with the DNA target without altering the duplex- vs quadruplex selectivity of the ligand, which relies on the shape and rigidity of the polyheteroaryl cyclic scaffold itself. Such molecular designs have thus to be further investigated, in order to gain more insights in terms of bioavailability and selective cytotoxicity.

Porphyrin (TMPyP4)-like macrocycles.

Along with neutral and cationic cyclic polyheteroaryles, cationic porphyrins are probably the most widely used macrocyclic G-quadruplex ligands. TMPyP4, (for 5,10,15,20-tetra-(N-methyl-4-pyridyl)porphyrin),\textsuperscript{13} is a representative example of this family of ligands (Fig. 2 and 6).\textsuperscript{17} The ability of TMPyP4 to interact with DNA was first mentioned 30 years ago\textsuperscript{52} whilst its use as a G-quadruplex ligand was reported for the first time twenty years later.\textsuperscript{13,53} Since these seminal reports, this tetracationic porphyrin has been extensively studied: it has demonstrated a high affinity for G-quadruplex DNA (as monitored by FRET-melting assay, $\Delta T_{1/2} = 17$ $^\circ$C),\textsuperscript{54} but with low selectivity.\textsuperscript{54,55} TMPyP4 has been particularly used as a tool to investigate the possibility of down-regulating the expression of genes due to quadruplex formation or induction (like c-myc,\textsuperscript{56} RET,\textsuperscript{31} HIF-1$\alpha$,\textsuperscript{27} VEGF,\textsuperscript{27,58} Bel-2,\textsuperscript{30,59} KRAS,\textsuperscript{60} PDGF-A,\textsuperscript{61} c-kit,\textsuperscript{62} and even hTERT genes).\textsuperscript{28} The biological behaviour of TMPyP4 has also
been investigated in vitro against a broad variety of tumor cell lines (with IC$_{50}$ between 23 and 310 µM), including human pancreatic, breast, and prostate carcinomas, lymphoma, retinoblastoma, leukaemia and human gastrointestinal stromal tumor cell lines, as well as telomerase-transformed SW39, ALT-transformed SW26, transformed human breast and HeLa cell lines. TMPyP4 has also been tested on normal cell lines (human fibroblast and breast cell lines), in which it demonstrated moderate to acute cytotoxicity (IC$_{50}$ down to 13 µM). It has also been used as an in vitro tool for more prospective and fundamental studies, sometimes exhibiting some unexpected behaviour, including quadruplex-unfolding activity on both quadruplex-forming DNA (d(CGG) repeats, or anti-thrombin aptamer) and RNA sequences (r(CGG) repeats).

Fig 6. Interaction of TMPyP4 and a G-tetrad (adapted from PDB entry: 2A5R); guanine residues appear in gold, the carbon, nitrogen, oxygen and hydrogen atoms of TMPyP4 in grey, blue, red and white respectively.

Interestingly, despite the fact that it was quickly acknowledged that TMPyP4 has poor selectivity for quadruplex structures, this ligand is still the focus of intense research interest. Notably, a lot of structural studies are currently undertaken to investigate its binding mode with quadruplex-DNA; interestingly, TMPyP4 displays various quadruplex-interactions, ranging from intercalation between adjacent G-quartets (as suggested via Raman spectroscopy studies) to the more expected stacking on the external G-quartet (demonstrated by NMR studies), passing by a mode totally devoid of direct contacts with G-quartets (shown by X-ray analysis of obtained crystal structures), without excluding the possibility of combining several binding modes during a single recognition process (vide infra). These efforts are also currently completed by investigations dedicated to the understanding of the nature of the interaction between TMPyP4 and G-quadruplex, either in silico, or via established biophysical methods, as well as ITC, ESMS, DSC, SPR, or even HPLC methods, on a broad variety of quadruplex-architectures e.g. telomeric, c-myc, c-kit, Bel-2, TBA, the bimolecular (G$_4$T$_4$G$_4$)$_2$, or the tetramolecular (TG$_4$T) quadruplexes and even G-wires. Nevertheless, it is worth pointing out that a vast majority of these studies have been carried out in in vitro conditions that have to be considered as ‘dilute’ as compared to the crowded environment within a living cell (due to the presence of naturally occurring proteins, nucleic acids, sugars, etc). Interestingly, when the conditions are artificially crowded (using molecular crowding agents such as
ethylene glycol\textsuperscript{75c} or poly(ethylene glycol)\textsuperscript{74b,76b} (PEG), \textbf{TMPyP4} displays a higher affinity (for quadruplex-DNA)\textsuperscript{74b,75c} and selectivity (with regard to duplex-DNA)\textsuperscript{76b} than in dilute conditions, on the basis of a different recognition process (the crowded conditions influencing both the quadruplex-DNA structures and the \textbf{TMPyP4} stacking interactions (being multiple but stepwise))\textsuperscript{74b,76b}. Given that crowded conditions are more biologically relevant, these results indicate that \textbf{TMPyP4} may be more selective than anticipated in cells, and these observations will certainly revitalize \textbf{TMPyP4}-related researches.

Thus, the current status of \textbf{TMPyP4} is the opposite of 	extit{telomestatin}; it is both commercially and synthetically accessible and its interaction with quadruplexes, although multiple, is fully characterized. Finally, the simple geometric comparison allowed by the overlaps shown in Fig. 3 and 6 strongly suggests that the tetrpyrrolic macrocycle is much less suitable than the octacyclic ring for optimizing aromatic-aromatic interactions with a G-quartet. In conclusion, although \textbf{TMPyP4} remains an interesting tool that has contributed to the understanding of the ligand/quadruplex interactions and to their complexity, its biological use for probing the involvement of quadruplex in biological assays is questionable.

With an inner cavity particularly suited to coordination of a metal, \textbf{TMPyP4} (and other porphyrins, \textit{vide infra}) has also widely exploited to form metal complexes. Given the well-known modulation of DNA association on metallation of this modification was thus performed in order to improve the DNA recognition.\textsuperscript{52} \textbf{TMPyP4} has now been used to prepare Pt(\textit{II})- Cu(\textit{II})- In(\textit{III})- Zn(\textit{II})- Co(\textit{II})- Fe(\textit{III})- Ni(\textit{II})- Mn(\textit{III})- Mg(\textit{II})- and Pd(\textit{II})- complexes.\textsuperscript{13a,80} Among them, \textbf{Mn-TMPyP4} (Fig. 7) deserves particular attention since it showed a strong improvement in terms of quadruplex-vs-duplex selectivity.

![Fig 7. Chemical formulae of porphyrin-based G-quadruplex ligands](image_url)

The porphyrin core, as clearly demonstrated with \textbf{TMPyP4}, has been extensively explored as a macrocyclic scaffold for the design of G-quadruplex ligands. Among the numerous examples that have been reported, either as metal-free or in a complexed-state, the most interesting ligands are presented in Fig. 7. Even if the methods for analyzing G-quadruplex-affinity and selectivity are not yet standardised, the nature of the side-arms that surround the porphyrin core is unmistakably crucial. Indeed, while the above-mentioned \textbf{Mn-TMPyP4} was able to interact with G-quadruplex 10 times faster than with duplex-DNA, a 10,000-fold difference was obtained by moving the \textit{N}-methylpyridinium moiety from the close vicinity of the porphyrin core to the termini of flexible arms.\textsuperscript{81} This modification is assumed to greatly improve the electrostatic interaction with quadruplex grooves, while impeding intercalation into duplex DNA. This kinetic difference was evaluated through biosensor-surface plasmon resonance (SPR) measurements. The SPR technique, based on the use of surface-immobilized
oligonucleotides, enables a quantitative analysis of the binding parameters, i.e. thermodynamic (equilibrium constant, Gibbs energy of binding, stoichiometry) and kinetic \( k_{on}/k_{off} \) parameters of the interaction of small molecules with the target DNA.\(^{55,62}\) In the present case, the affinity constant determined with duplex DNA lies in the \( 10^4 \) M\(^{-1} \) range, reaching up to \( 10^8 \) M\(^{-1} \) with quadruplex-DNA. Even if more modest, similar observations were made with TQMP, a porphyrin in which \( N \)-methylpyridinium substituents were replaced by (trimethylammonium)methyl-phenol arms (Fig. 7).\(^{55,62}\) In this case, an \(~100\)-fold difference in the kinetics of the association with duplex- (\( 10^4 \) M\(^{-1} \) range) and quadruplex-DNA (\( 10^6 \) M\(^{-1} \) range) was obtained. Finally, knowing that the nature and number of the surrounding side-arms can be fine tuned in the case of TMPyP4,\(^{83}\) interesting perspectives exist to modify porphyrin-based G-quadruplex ligands in this way. It is however worth noting that the equilibrium between affinity and selectivity is difficult to access with the TMPyP4 scaffold, as demonstrated by a recent study in which a TMPyP4-related compound was included in a high-order supramolecular architecture (ruthenium coordination cubes),\(^{84}\) on the basis of the elegant idea to use the metal itself to build the macrocyclic scaffold.\(^{85}\) These cubic complexes were indeed shown to strongly bind quadruplex-DNA, but also duplex-DNA, despite their peculiar molecular volume. This lack of selectivity may originate in their highly cationic nature (eight positive charges).

A large number of structural analogues of TMPyP4 have been recently described in the literature, differing notably by the nature of the side-arms (\( N \)-methylpyridinium groups have, for example, been replaced by \( N \)-propylpyridinium,\(^{74,62}\) \( N \)-methylquinolinium,\(^{86}\) phenyl groups with short,\(^{57}\) long aminoalkyl arms,\(^{88}\) and even endowed with fluorescent reporters).\(^{89}\) The extensive family of ring-size analogues of the 16-membered porphyrins are exemplified in Fig. 8, including the 15-membered corrole, 16-membered phthalocyanine or porphyrazine and the 19-membered sapphyrin, which have all been employed for the construction of G-quadruplex ligands.

![Fig 8. Poly-heterocyclic macrocyclic cores used in the structures of G-quadruplex ligands. The number in the center corresponds to the number of atoms in the central ring.](image)

Corroles are tetrapyrdolic macrocyclic structures, lacking one bridging \( sp^2 \) carbon with respect to porphyrins. Interestingly, the resulting TMPyP4 analogue, i.e. substituted with \( N \)-methylpyridinium groups (Fig. 9),\(^{90}\) is only trisubstituted, with a peculiar spatial disposition of pyridinium moieties relative to each other (Fig. 10).\(^{17}\) This corrole efficiently stabilizes quadruplex-DNA, as judged from CD-melting experiments (\( T_m \) increase of 36 °C), with a fair selectivity, by SPR measurements (affinity constants in the range of \( 10^3 \) and \( 10^4 \) M\(^{-1} \) for duplex- and quadruplex-DNA, respectively). Copper-corroles have also been studied and proved to be good stabilizing agents for quadruplexes formed from sequences of both the human telomere and the c-myc promoter.\(^{91}\) Porphyrazine, the poly-aza tetra(pyrolopyridine) analogue of porphyrin (Fig. 8 and 9) has also been widely used in the design of G-quadruplex ligands, because of its steric properties (close to that of the porphyrin), but also because its synthesis opens novel perspectives for structural diversity. It enables, for example, to fuse the...
pyridinium moieties with the cyclic scaffold, thereby leading to a flat, electron-deficient and structurally frozen macrocycle, with an excellent overlap of a G-tetrad (Fig. 10).$^{17}$ **3,4-TMPyPz**, used either as metal-free or as Zn-complex, shows a high affinity and selectivity for quadruplex-DNA, as demonstrated by UV-vis and SPR experiments (affinity constants in the range of $10^6$ M$^{-1}$ for quadruplex-DNA and $<10^5$ for duplex-DNA), with a refined binding mode, since a unique binding site was determined using both of these techniques.$^{92}$ It is worth noting that **3,4-TMPyPz** has also been used as a tool to investigate the role of G-quadruplex ligands in a totally telomere-independent mechanism, i.e. an innovative gene-silencing strategy based on the control by a G-quadruplex ligand of the dicing of quadruplex-forming short hairpin RNAs that actively participate to the RNA interference (RNAi) pathway.$^{93}$

![Fig 9. Chemical structures of corrole-, porphyrine-, phthalocyanine- and sapphyrin-based G-quadruplex ligands](image)

Phthalocyanines are structurally close to porphyrines, with the pyridine rings that compose the porphyrine scaffold replaced by benzene rings. The zinc phthalocyanine complex **ZnPc** (Fig. 9)$^{94}$ carries the classical amino side-chains in the close vicinity of the macrocyclic scaffold. The results are interesting, since they demonstrate that a highly charged compound (up to 8 positive charges), whilst presenting an expected high degree of quadruplex affinity, can concomitantly retain a satisfactory level of selectivity (on the basis of SPR and CD-melting data) - the difference in kinetic parameters with duplex- vs quadruplex-DNA being in the range of one order of magnitude. When the aminoalkyl side-arms of **ZnPc** are replaced by guanidinium moieties, the corresponding **Zn-DIGP** (for tetrakis(diisopropylguanidino)zinc phthalocyanine, Fig. 9) elicits quite exceptional properties, both in terms of quadruplex affinity ($K_d < 2$ nM for c-myc quadruplex) and quadruplex- vs duplex-DNA selectivity (**Zn-DIGP** displays a 5000-fold higher affinity for c-myc quadruplex than for calf thymus DNA). In addition, **Zn-DIGP** exhibits “turn-on” fluorescent properties, making it a particularly interesting quadruplex-selective dye.$^{95}$ Very recent studies have substantiated the interest in this series of molecules, confirming their good cellular uptake and inhibition of promoter-quadruplex-mediated genes (KRAS).$^{96}$
The expanded diselenosapphyrin Se2SAP (Fig. 9) represents a novel class of chemically challenging compounds but with a very interesting ability to interact only with its preferential target, the oncogenic c-myc derived quadruplex. Indeed, Se2SAP discerns, firstly, among the DNAs of various nature (duplex- and quadruplex-DNA, with a selectivity factor of ~600 as judged by a competitive Taq polymerase assay), and secondly, among quadruplex-DNA of various structures (Se2SAP binds specifically c-myc, as compared to human telomeric quadruplex, but displays also a higher affinity for the quadruplex found in the promoter region of VEGF). The substitution of nitrogen by selenium atoms in the backbone of the ligand was motivated by the decrease of the known photocytotoxicity of the porphyrins, as confirmed by the low general cytotoxicity of Se2SAP in HeLa cells (no effect at 200 µM dose). One major drawback to the use of Se2SAP in cells stands in its limited bioavailability, which remains to be improved for further biological investigations.

A further noteworthy family of porphyrins is hemin and related compounds. Hemin, also known as Fe(III)-protoporphyrin IX (Fig. 11), is the oxidized version of heme (Fe(II)-protoporphyrin IX), widely known as a cofactor of hemoglobin (and myoglobin) that actively participates in the reversible binding of dioxygen. Since the pioneering work of Sen et al., hemin is known to bind tightly to quadruplex architecture; however, hemin has not been investigated as a G-quadruplex-ligand per se, but has been thoroughly investigated for its DNAzyme activity when bound to quadruplex. Indeed, the hemin/quadruplex complex acquires the ability to catalyze the H₂O₂-mediated oxidation of precursors such as ABTS (2,2'-azino-bis(3-ethylbenzothiazoline-6-sulfonic acid)), luminol or TMB (3,3',5,5'-tetramethylbenzidine), therefore causing an easily detectable color change. This particular
activity has found numerous applications, including the detection of telomerase activity, single-stranded DNA, SNP (single nucleotide polymorphism), DNA and DNA analytes, proteins (like nucleolin, lysozyme, or thrombin), various cations (such as Cu$^{2+}$, Hg$^{2+}$, Pb$^{2+}$, or K$^+$) or enzyme co-factors (like L-histidine), as well as a use for constructing DNA logic gates, and it has been also implied in the development of novel in vitro assays for screening G-quadruplex ligands.

A hemin-related porphyrin, NMM (or N-methyl mesoporphyrin IX) has also been thoroughly investigated for its ability to interact with G-quadruplex. NMM (Fig. 11) is interesting as the unique representative of the negatively charged macrocyclic G-quadruplex ligands (its two carboxylic groups being deprotonated at physiological pH). NMM was initially studied since it behaves as stable transition-state analogue for ferrochelatase enzyme, which catalyzes the insertion of Fe$^{2+}$ ions into protoporphyrin in the final step of the heme biosynthesis. Whilst NMM became rapidly known as highly quadruplex-selective, it was only sparingly studied due to a moderate affinity for its DNA target, despite an excellent overlap with the G-tetrad (Fig. 12). However, in recent studies in yeast, it has shown a very interesting ability to control the regulation of genes in direct connection with the probability of quadruplex-formation in their promoters. Additionally, given that the effect of NMM is felt at both telomeres and loci throughout the genome, this study provides novel lines of evidence of the possible in vivo roles of G-quadruplexes. Finally, NMM has also been used as a novel tool for isolating G-quadruplex DNA from a mixture of nucleic acids derived from in vivo sources, via its grafting on a sepharose resin, thus offering an innovative way to determine whether or not the in silico detected putative quadruplex-forming sequences (QFS) exist in the cellular context as quadruplex DNA.

![Fig 12](image-url). Qualitative in silico superposition of NMM and a G-tetrad (extracted from PDB entry: 2A5R); guanine residues appear in gold, the carbon, nitrogen, oxygen and hydrogen atoms of NMM in grey, blue, red and white respectively.

Polyammonium cyclophane-type macrocycles.

As compared to the telomestatin, porphyrins and related compounds, polyammonium cyclophane-type macrocycles have not been extensively studied as G-quadruplex ligands. Mainly two types of cyclophane-like macrocycles have been studied for their interaction with quadruplexes: the cyclobisintercalator family (CBI, Fig. 13) and the neomycin-capped aromatic platforms (exemplified by NCQ, Fig.13).
The CBI macrocycles contain two flat aromatic units, usually derived from simple intercalators linked together by polyamine chains. This particular scaffold is highly soluble at physiological pH thanks to protonation of the four benzyl nitrogen atoms of the linkers. With regard to DNA binding the CBIs present unique features that are i- a strong association with DNA bases via aromatic π-π interactions, and ii- a very low affinity for duplex-DNA. This family of compounds has already been studied for its recognition of various unusual DNA structures mainly mismatch-containing DNA, abasic sites, and trinucleotide repeats. However, only two members of this family, namely BOQ₁ (for Bis-Ortho-Quinacridine, Fig. 13) and BisA (for Bis-Acridine, Fig. 13) have been already evaluated for their ability to interact with quadruplex-DNA.

Fig 13. Chemical formulae of non-planar macrocyclic G-quadruplex ligands CBIs: BOQ₁ (left), BisA (centre) and the neomycine capped scaffold NCQ (right).

BisA and more particularly BOQ₁ display promising G-quadruplex-binding properties. Both compounds are composed of two large polycyclic aromatic moieties namely acridine or quinacridine that enable efficient π-stacking interactions with nucleic bases in particular guanines. These dimeric macrocycles were found to bind much better to quadruplexes than their planar monomeric acyclic counterparts. This strong association has been demonstrated by numerous techniques, including FRET-melting assay, SPR and equilibrium dialysis: BOQ₁ and BisA efficiently induce thermal stabilization of the human telomeric quadruplex (ΔT1/2 = 28 and 15 °C, respectively) displaying high affinity for this target (Kₐ = 1.2×10⁷ M⁻¹ for BOQ₁ measured by SPR and 1.1×10⁵ M⁻¹ for BisA, estimated by dialysis). The difference between the two compounds clearly demonstrates the importance of the size of the aromatic unit in the binding interaction with the quadruplex structure.

Most interestingly the two compounds exhibit a high selectivity for quadruplex- over duplex-DNA attributed to their particular topology. CBIs are clearly distinct from the planar rigid macrocycles mentioned precedently since they exist in a semi-closed conformation with the two aromatic units facing each other in more or less parallel planes (Fig. 13 and 14). This particular non-planar conformation impedes binding to duplex-DNA, both for steric reasons and because the length of the linkers does not enable bis-intercalation between contiguous base pairs due to the neighbour-exclusion principle. In the case of BOQ₁, this particular topology was found both in the free state (Fig. 14A-C) and in the complex formed with the human telomeric quadruplex (Fig. 14D), meaning that there is little conformational change upon binding. This would suggest that the ligand is already preorganized to fit into the quadruplex target resulting in a gain in the binding entropy. In addition, the modelling studies performed both with the telomeric and with the c-myc quadruplex demonstrate that the semi-closed conformation of BOQ₁ enables interaction via a mixed binding mode, combining G-quartet and loop (Fig. 14D, c-myc not shown). In fact, careful examination of this complex
strongly suggests the occurrence of electrostatic and H-bonding interactions between the linkers and the loop backbone (Fig. 14D).

The second category of polyammonium cyclophane-type macrocycles namely the neomycin-capped macrocycles were developed with the goal to validate the ditopic interactive design. To this aim three aromatic platforms (acridine, phenanthroline, quinacridine) able to stack on a G-quartet were combined with a well-known loop-binding motif i.e. the aminoglycoside neomycin. The most active prototype of this series is the quinacridine derivative NCQ shown in Fig. 13. This ligand that has a high degree of molecular flexibility is a potent binder of the telomeric quadruplex as shown by the FRET melting assay ($\Delta T_{1/2} = 15^\circ C$) whereas it elicits a poor affinity for duplex-DNA. Most importantly NCQ elicits a high degree of selectivity for looped-quadruplex (i.e. intramolecular telomeric quadruplex), compared with the tetramolecular quadruplex lacking the loop motifs. The preference for intramolecular over tetramolecular quadruplexes confirms the potential of the multitopic approach and altogether this work proposes a new innovative design for achieving intra-quadruplex selectivity. Subsequent studies have since then fully supported this approach using various G-quadruplex recognition scaffolds, including a macrocyclic oxazole-based tripeptide or a small-molecule ligand (acridine) equipped with loop and/or groove recognition elements like carbohydrates or peptides.

![Fig 14. Structure of BOQ1 (A), side (B) and front views (C) of the lowest-energy conformation during the molecular dynamic simulation (see Ref. 126) and modelled interaction between BOQ1 and the human telomeric quadruplex (D, see Ref. 121c); guanine residues appear in gold, the carbon, nitrogen, oxygen and hydrogen atoms in grey, blue, red and white respectively.](image_url)

In conclusion, polyammonium cyclophane-type macrocycles have been only sparingly studied as G-quadruplex ligands. They are able to establish various types of interaction with their DNA target, thereby opening the innovative perspective of a selective interaction with quadruplex-DNA as a function of the nature of the quadruplex structure itself. This intra-quadruplex selectivity, motivated by the putative presence of quadruplexes at various genomic localizations (vide supra), is one of the major issues that will have to be addressed in the near future. The results previously obtained with CBIs and presented herein, although limited to a few examples, are promising. On this solid basis, an extensive study of binding of CBIs to G-quadruplex has been undertaken which is further reported in the companion paper (“One Ring to bind them all” – Part II. by A. Granzhan et al.), a research article in the present issue.
Conclusion

Macrocyclic scaffolds are particularly attractive for designing selective G-quadruplex ligands essentially for the two following reasons: on one hand they show a poor affinity for the “standard” B-DNA conformation, due to their sterically difficult intercalation between the base pairs of the double helix; on the other hand, in contrast, they fit nicely with the external G-quartets of quadruplexes that constitute accessible planar sites of large aromatic area. Although synthetic accesses to macrocycles are often difficult, this chemical class is nevertheless a fascinating tool that can be chemically engineered to generate new biological properties. The few macrocyclic families depicted herein have only begun to mine this incredible potential and we hope that the present article will be helpful to understand what the crucial guidelines are to control their efficiencies and selectivities for recognizing quadruplexes.

Acknowledgments

The authors would like to thank sincerely their co-workers for the synthetic work and stimulating scientific discussions related to G-quadruplex DNA and ligands; a particular acknowledgment will be dedicated to A. De Cian, M. Kaiser and L. Lacroix. The authors are also grateful to all the collaborators who contributed to the works and to all their colleagues worldwide whose works are described herein. A special acknowledgment to J. R. R. Tolkien for the title inspiration and to J. Martin for careful reading of the manuscript.

References

4 a) T. de Lange, Gene Dev., 2005, 19, 2100.
8 For a recent review, see: Y. Qin, L. H. Hurley, Biochimie, 2008, 90, 1149.
17 Stable conformations of the ligands structure were generated using the Dreiding force-field of the calculator plugins from ChemAxon (Marvin 5.0.1, 2008, ChemAxon (http://www.chemaxon.com)), except for TMPyP4 for which the structure was extracted from the PDB (PDB ID 2HR1). Superposition was performed using Maestro (Maestro Academic Version 8.0.314, (c) 1999-2007 Schrödinger, LLC). Molecular graphics images were produced using the UCSF Chimera package from the Resource for Biocomputing, Visualization, and Informatics at the University of California, San Francisco (supported by NIH P41 RR-01081): E.F. Pettersen, T.D. Goddard, C.C. Huang, G.S. Couch, D.M. Greenblatt, E.C. Meng, T.E. Ferrin, *J. Comput. Chem.*, 2004, 25, 1605.
45 Calculator Plugins were used for pKa calculation, Marvin 5.0.1, 2008, ChemAxon, http://www.chemaxon.com


P. Ragazzon, J. B. Chaires, Methods, 2007, 43, 313.


A BOQx4H+ minimized structure was imported as a MOL2 file into Maestro (Maestro Academic Version 9.0.109, (c)1999-2009 Schrödinger, LLC). The system was then built within Maestro via the Desmond Application (Desmond Molecular Dynamics System, version 2.2.6.2.3, D. E. Shaw Research, New York, NY, 2009), using the predefined TIP4PEW Solvent Model, a cubic boundary box with 10 Å edges and Cl⁻ as neutralizing ions. Molecular Dynamics studies were then performed with Desmond, with a 4.8 ns simulation time and a NPT Ensemble class at 300 K and 1.01325 bar, allowing the model to relax before