A two-quartet G-quadruplex topology of human KIT2 is conformationally selected by a perylene derivative

Silvia Ceschi, Eric Largy, Valérie Gabelica, Claudia Sissi

Version of Record: https://www.sciencedirect.com/science/article/pii/S0300908420302212
Manuscript_5972eb87550ac512f87a3d60f2fe5601

A two-quartet G-quadruplex topology of human KIT2 is conformationally selected by a perylene derivative

Silvia Ceschi[1], Eric Largy[2], Valerie Gabelica[2*], Claudia Sissi[1*]

[1] Department of Pharmaceutical and Pharmacological Sciences, University of Padova, v. Marzolo 5, 35131, Padova, Italy
[2] Université de Bordeaux, Inserm & CNRS, Laboratoire Acides Nucléiques : Régulations Naturelle et Artificielle, (ARNA, U1212, UMR5320), IECB, 2 rue Robert Escarpit, 33607 Pessac, France

*e-mail: claudia.sissi@unipd.it; v.gabelica@iecb.u-bordeaux.fr

ABSTRACT

G-quadruplexes are promising targets for innovative anticancer therapy. Hence, many efforts are being made to find selective ligands. Drug design is often based on the available high-resolution structures, obtained for the thermodynamically stable forms. However, the complexity of the G-quadruplex folding landscape has clearly emerged in recent years, with the discovery of intermediate conformations that persist on the second to the minute time scale. In the case of the KIT2 G-quadruplex forming sequence, found within human *c-KIT* promoter, we recently identified a long-lived folding intermediate, characterized by guanine stacking in alternating orientation (as determined by circular dichroism). Given the rate of transcriptional processes, a physiological role of this arrangement should not be excluded. In the present study, we applied circular dichroism (CD) spectroscopy, native electrospray ionization mass spectrometry (ESI-MS) and electrophoretic mobility shift assays (EMSA) to show that a perylene derivative (K20) selects this topology. Interestingly, ESI-MS spectra revealed the presence of a single specifically coordinated K^+ ion in the structure, which is thus presumably composed of only two consecutive G-quartets. The parent ligand PIPER failed to promote the same conformational selection, which is therefore a process strictly dependent on the perylene side chains composition. The greater affinity of K20 for the two-quartet antiparallel topology, compared to PIPER, was finally corroborated by evaluating their binding to the KIT* G-quadruplex, which is also found within the human promoter of *c-KIT*.

Keywords: *G-quadruplex folding, c-KIT, perylene, native ESI-MS, circular dichroism*

1. INTRODUCTION

Guanine-rich DNA/RNA sequences may arrange into tetra-helical secondary structures called G-quadruplexes (G4s) [1]. They are composed of multiple stacked G-quartets (planar arrangements of four guanines interacting through Hoogsteen hydrogen bonds) and can adopt different topologies depending on inter-related features: the number of the involved strands, the relative orientation of the strands in the stem (parallel, antiparallel), the shape of the loops connecting the four parts of the stem and the stacking mode of the bases in the stem.[2] Additionally, stabilization of the structure is ensured by monovalent cations such as K^+, which intercalate between consecutive G-quartets [3],[4].

G-quadruplexes are promising targets for innovative anticancer therapy[5],[6] . Indeed, they are enriched in the promoters of several oncogenes and for many of them a role in the downregulation of the gene expression has already been demonstrated[7],[8]. However, the polymorphism of G-quadruplexes makes the rational design of G4 ligands challenging. Indeed, a given G-rich sequence can adopt different conformations 'in vitro', some of which being the most thermodynamically stable, others being kinetically favoured [9],[10],[11],[12],[13]. This also applies to KIT2, a sequence located within the promoter of the human *c-KIT* protooncogene, between -140 and -160 bp upstream of the transcription start site (TSS) [14]. In potassium-containing solutions, KIT2 arranges into two thermodynamically stable G4 structures that have been solved by NMR: a parallel three-quartet monomer and a parallel six-quartet interlocked dimer [15],[16]. Furthermore, along its folding, it assumes at least one long-lived intermediate conformation, characterized by guanine stacking in alternating orientation (as determined by circular dichroism) [17].

G4 ligands structure-based virtual screenings are usually run on the thermodynamically stable conformations, for which high-resolution structures are available. Similarly, in-solution assays performed under equilibrium conditions are typically interpreted in terms of ligand preference for the starting structure of the nucleic acid. However, the G4 structural features 'in vivo' are still a matter of concern. Interestingly, a recent bioinformatic study revealed that motifs giving rise to the 'in vitro' most thermodynamically stable G4 structures are underrepresented in the genome of a large number of species [18]. The authors suggest that evolution has selected less stable G4 conformations, which can perform a regulatory function without causing genomic instability. Furthermore, considering the rate of biological processes, the most relevant G4 conformations within a cellular context are likely to be those kinetically favoured. Given the transient nature of the latter, their selective stabilization with ligands is pivotal to gain information about their structure and function.

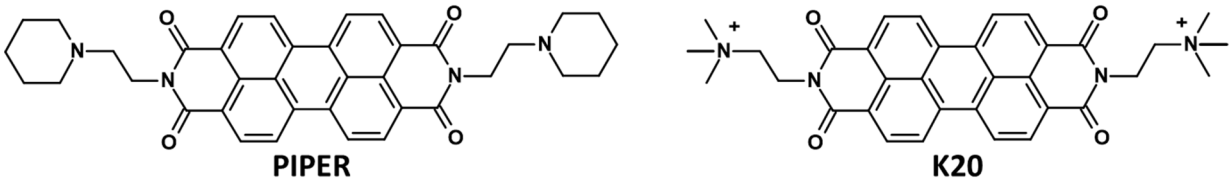
Perylene derivatives were first identified as telomeric G-quadruplex binding agents and telomerase inhibitors. Besides their binding to the intramolecular G-quadruplex, they can promote formation of di- and tetrameric G-quadruplexes from both single stranded and duplex guanine-rich DNA sequences [19],[20],[21]. Interestingly, ligand side chains containing basic amino groups modulate their preference to some intramolecular G-quadruplex topologies, owing to their binding to the DNA grooves [19],[22],[23]. In particular, a previous screening pointed out the propensity of one PIPER derivative, named K20, to preferentially drive intramolecular antiparallel G-quadruplex folding [23].

Here, we show that K20 can be successfully exploited to select an antiparallel folding intermediate of KIT2. Native ESI-MS experiments allowed us to determine both ligand and K^+ binding stoichiometries to the selected arrangement [24], revealing its two-quartet scaffold. Conversely, the parent ligand PIPER, which differs from K20 only in the terminal cyclic amino groups, does not show the same preference for the antiparallel topology.

The ability of K20 to select antiparallel G-quadruplexes in contrast to PIPER, was further confirmed by evaluating the binding of the two ligands to KIT*, which is another G-rich sequence also found within the human *c-KIT* promoter (between -138 and -117 bp upstream of the TSS) for which we recently obtained the high-resolution structure by NMR [25],[26]. This sequence represents a good control system. Indeed, in contrast to KIT2, KIT* stably folds into a single two-quartet G-quadruplex with an antiparallel topology.

Our results point to the quaternary ammonium salt in the side chains of K20 as a driving element for the observed conformational selection.

Chart 1. Structures of ligands used in this study



2. MATERIALS AND METHODS

2.1 Oligonucleotides

Oligonucleotides were purchased lyophilized and RP-HPLC purified from Eurogentec (Seraing, Belgium) and used without further purification. The DNA sequences are listed in Table 1.

Table 1. DNA sequences used in this work

KIT2	5'-CGGGCGGGCGCGAGGGAGGGG-3'
KIT*	5'-GGCGAGGAGGGGCGTGGCCGGC-3'
KIT2noG4	5'-GGGACGCGACGAGAGGCCGGG-3'
KIT*noG4	5'-GGCGAGGAGTTGCGTGGCCGGC-3'

Oligonucleotides were resuspended in nuclease-free water from Thermo Fisher Scientific (Waltham, MA, USA) to obtain 1 mM stock solutions, which were then diluted to 200 µM in 10 mM Tris-HCl (>99.999%, Merck, Darmstadt, Germany) (pH 7.4) or 100 mM trimethylammonium acetate (TMAA, Ultra for HPLC, Merck, Darmstadt, Germany) (pH 7.0). Analyzed solutions contained 10 µM DNA. The concentrations of the initial stock solutions were measured by UV absorbance at 260 nm on a Uvikon XS, using molar absorption coefficients calculated with a nearest neighbour model [27]. The working solutions were doped with KCl (>99.999%, Merck, Darmstadt, Germany), annealed with 85 °C heating for 5 min and then led to equilibrate overnight at room temperature before further analysis.

2.2 Ligands

The ligand K20 was synthesized from Prof. A. Paul Krapcho (University of Vermount) as described elsewhere [22]. The ligand PIPER was purchased from Merck (Darmstadt, Germany). Ligands were carefully weighted and solubilized in DMSO to obtain stock solutions at 4.25 mM for K20 and 1.32 mM for PIPER, which were diluted in aqueous buffer just before the analysis. Ligands were added on equilibrated solutions and led to equilibrate overnight.

2.3. Circular Dichroism Spectroscopy

Circular dichroism (CD) spectra were acquired on a JASCO J-810 spectropolarimeter equipped with a Peltier temperature controller. CD spectra were recorded from 235 to 330 nm with the following parameters: scanning speed of 100 nm/min, band width of 2 nm, data interval of 0.5 nm and response of 2 s. Measurements were performed using a 1 cm path length quartz cuvette at oligonucleotide concentration of 10 µM in 10 mM Tris-HCl (pH 7.4) or in 100 mM TMAA, with the proper amount of KCl. The samples were annealed at 85 °C for 5 min and led to equilibrate before spectra acquisition. For CD titrations, increased amounts of ligands were added to distinct equilibrated samples and spectra were recorded after overnight equilibration. Observed ellipticities were converted to Molar Ellipticity which is equal to $deg \cdot cm^2 \cdot dmol^{-1}$,

calculated using the DNA residue concentration in solution. CD melting studies were performed by recording the CD signal at 265 nm from 10°C to 95°C with a heating rate of 40°C/h and 2°C data pitch.

2.4 Electrophoretic Mobility Shift Assays

Electrophoretic mobility shift assays (EMSA) were performed with ^{32}P 5′-labelled oligonucleotides. Incubation with T4 polynucleotide kinase (T4 PNK 500U 10U/µL, Thermo Fisher Scientific) and γ-^{32}P ATP (Perkin Elmer, Waltham, MA, USA) was carried out for 1h at 37 °C. After enzyme inactivation at 50 °C, DNA was extracted with phenol:chlorophorm:isoamyl alcohol (25:24:1) (Merck) and precipitated with ethanol overnight at -20 °C. Dried DNA was then resuspended in 10 mM Tris-HCl (pH 7.4). Analysed samples contained 10 µM mixture of labelled and non-labelled DNA, 50 mM KCl and proper amount of ligand. They were annealed at 85 °C for 5 min and then led to equilibrate overnight at room temperature before addition of the ligand. After equilibration, gel loading buffer (50% glycerol, 50% water, xylene cyanole, bromophenol blue) was added to the samples immediately before loading them into a 15% polyacrylamide (acrylamide/bis-acrylamide 19:1, Biosolve BV, The Netherlands) gel. Electrophoresis proceeded for 2 h in TBE buffer (0.89 M Tris-Borate, EDTA 20 mM, pH 8.0) on ice. Resolved bands on dried gels were visualized on a PhosphorImager (Amersham).

2.5 Native Electrospray Ionization Mass Spectrometry

Electrospray ionization mass spectrometry (ESI-MS) experiments were performed on a Thermo-Exactive Orbitrap mass spectrometer in the negative ion mode, equipped with a standard ESI source. Samples were injected at 5 µl/min by a syringe pump. The full scan mass range was 400–3000 m/z. The Exactive was tuned to soft conditions using the bimolecular quadruplex d$[G_4T_4G_4]_2$ in 100 mM ammonium acetate.[28] G4 solutions were prepared in 100 mM trimethyl ammonium acetate (TMAA), pH 7.0, up to 4 mM KCl. The nucleic acids were analysed at a concentration of 10 µM, adding the relevant amounts of ligands. The single strand dT_6 was also added at 2.5 µM and used as an internal standard to validate that the relative intensities reflect the relative concentrations in solution [29]. Data were analysed using Xcalibur 2.2.0 software (Thermo Fisher Scientific).

2.6 Subtraction of the non-specific K^+ adducts

The KIT2noG4 and KIT*noG4 sequences were used as references to obtain the non-specific K^+ adducts distribution. Then, the method described by Marchand et al. allowed discrimination between specific and non-specific adducts on the raw distribution obtained for the G4 forming sequences [10]. The following procedure has been used:
1- The integrals of the K^+ distribution for the non-G4 forming sequences ($I_{0K,ref}, I_{1K,ref}$, etc.), and for the sequence of interest (I_{0K}, I_{1K}, etc.) were recorded in the same electrolyte conditions.
2- For the reference, we calculated the *x*,*y*,z coefficients, which reflect the relative amount of $1K^+$, $2K^+$, and $3K^+$ non-specific adducts compared to the $0K^+$ peak.

$$x = \frac{I_{1K,ref}}{I_{0K,ref}}; y = \frac{I_{2K,ref}}{I_{0K,ref}}; z = \frac{I_{3K,ref}}{I_{0K,ref}}$$

3- The values are then used to reconstruct the specific K^+ adducts distribution thanks to the following formula:

$$0K = I_{0K} + (I_{0K} * x) + (I_{0K} * y) + (I_{0K} * z)$$

$$1K = (I_{1K} - I_{0K} * x) + (I_{1K} - I_{0K} * x) * x + (I_{1K} - I_{0K} * x) * y + (I_{1K} - I_{0K} * x) * z$$

$$2K = (I_{2K} - I_{0K} * y - (I_{1K} - I_{0K} * x)) + (I_{2K} - I_{0K} * y - (I_{1K} - I_{0K} * x)) * x + (I_{2K} - I_{0K} * y - (I_{1K} - I_{0K} * x)) * y + (I_{2K} - I_{0K} * y - (I_{1K} - I_{0K} * x)) * z$$

where 0K, 1K and 2K are the derived integrals of the specific K^+ adducts. The amount of each specific K^+ adduct has been expressed in percentage of the total amount of specific adducts for a given charge state. The derived integrals of the specific K^+ adducts were also used to derive the relative concentrations of the species in solution. After verifying that all complexes have the same response factors, we determined single point K_D at one equivalent of ligand by applying the following Equation:

$$K_D = \frac{[L]_{free} \times [\mathrm{DNA}]_{free}}{[DNA + L]}$$

The values are reported along with the standard deviation of three replicates.

3. RESULTS

3.1 K20 preferentially binds to an antiparallel two-quartet conformation of KIT2.

Circular dichroism (CD) spectroscopy allows to monitor the G-quadruplex topology and to detect ligand-induced conformational changes. Parallel G-quadruplexes present the so-called type-I spectra (positive peak at 265 nm, negative peak at 240 nm). Type-II spectra (positive peaks at 265 and 295 nm, negative peak at 240 nm) are characteristic of hybrid G-quadruplexes, and type-III spectra (positive peaks at 295 and 240 nm, negative peak at 260 nm) belong to antiparallel G-quadruplexes [30]. If mixtures are present, the CD spectra will be the weighted average of all conformers present in solution and may thus resemble type-II spectra.

We performed CD titrations by recording CD spectra at increasing K20/KIT2 molar ratios (Figure 1A). Since KCl concentration higher than 100 mM strongly promotes the KIT2 folding into the dimeric G-quadruplex [16], in our assay we first equilibrated the oligonucleotide in 50 mM KCl. Under this condition, the DNA CD signature suggests the presence of parallel conformers, that contribute to the signal at 265 nm. An antiparallel component is still detectable as evidenced by the positive band at 290 nm. This means that at this KCl concentration, the equilibrium is not completely shifted towards the parallel monomeric and dimeric forms. Addition of K20 led to a decrease of the CD signal at 265 nm and an increase at 290 nm, suggesting a ligand-induced conformational shift towards an antiparallel/hybrid topology.

To check for possible ligand-induced multimerization processes, we performed native electrophoretic mobility shift assays (EMSA) with samples prepared under the same experimental conditions applied for CD titrations. Addition of K20 to KIT2 G-quadruplex, up to 2:1 perylene/DNA molar ratio, did not trigger multimerization of the oligonucleotide (Figure 1B): only two electrophoretic bands were detected, that can be attributed to the monomeric and dimeric forms among which the oligonucleotide distributes under these experimental conditions [17]. The detected amount of dimer was small, and we could observe ligand binding to this arrangement, evidenced by its decreased electrophoretic mobility. The electrophoretic mobility of the monomer increased upon K20 binding: since the ligand is positively charged, this variation reflects a ligand-induced conformational change of the nucleic acid. It has been previously reported that the antiparallel G-quadruplex of the telomeric sequence has a higher electrophoretic mobility in comparison to the parallel one; thus, we can assume that also for KIT2 this fast-running species corresponds to an antiparallel G-quadruplex [31]. At 3:1 perylene/DNA molar ratio, a very slow migrating species was detected, which likely corresponds to a ligand-induced multimer. Therefore, we limited the amount of K20 to 2:1 ligand/DNA molar ratio in all subsequent analyses.

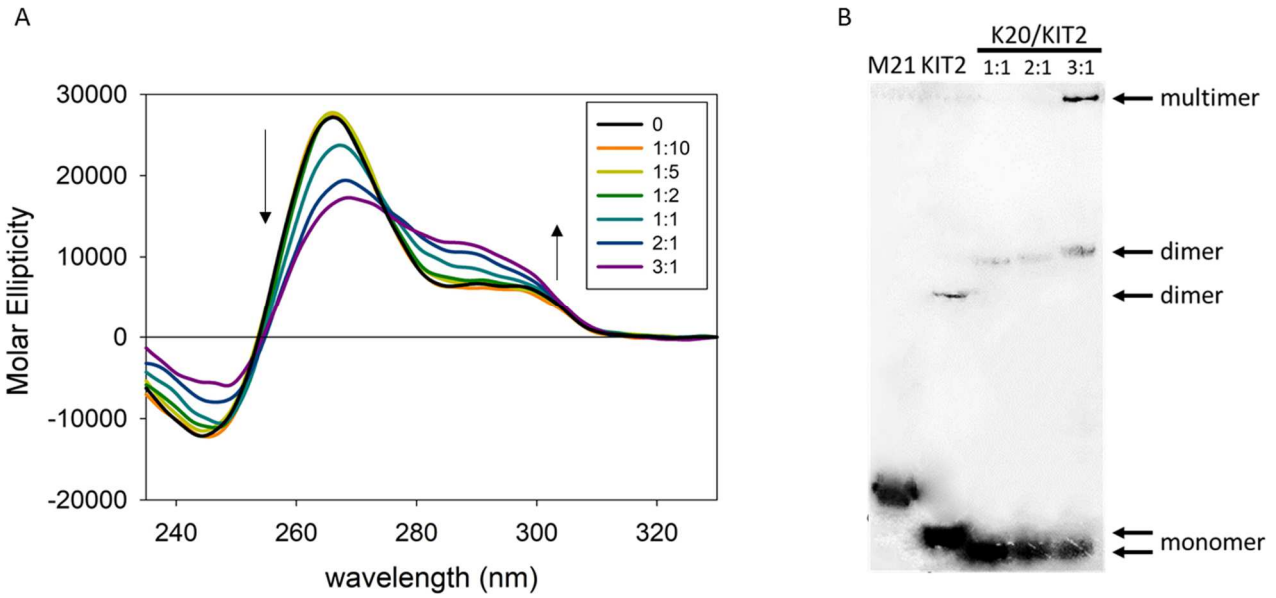


Fig. 1. *A) CD spectra of 10 µM KIT2 in 10 mM Tris-HCl (pH 7.4), 50 mM KCl at increasing K20/KIT2 molar ratios. B) EMSA of 10 µM KIT2 in 10 mM Tris-HCl (pH 7.4), 50 mM KCl at increasing K20/KIT2 molar ratios.*

Native electrospray mass spectrometry (ESI-MS) separates the different G-quadruplex scaffolds according to the number of the DNA strands that are involved, and the number of K^+ ions that are selectively coordinated in the structure, which reflects the number of consecutive G-quartets.[32],[33] Thus, it can be exploited to investigate ligand binding to polymorphic G4 forming oligonucleotides, such as KIT2.

First, we acquired ESI-MS spectra to check the ability of KIT2 to properly fold into G4 under the required experimental conditions (100 mM TMAA, up to 3 mM KCl). Figure 2A shows the ESI-MS spectra of KIT2 at increasing KCl concentration, zoomed on the 5- charge state. ESI-MS spectra of a reference non-G4 forming sequence (Figure 2B), KIT2noG4 (Figure S1 Supplementary Material), with the same guanine content and length as KIT2, allowed us to detect and quantify the non-specific K^+ adducts. These latter were subtracted from the raw distribution of KIT2 K^+ adducts (See Experimental Section) to reconstruct the specific K^+ adducts distribution in the form of bar graphs (Figure 2C). This enabled us to monitor G-quadruplex formation at different KCl concentrations.

Previous ESI-MS studies were carried out at maximum 1 mM KCl. However, at 1 mM KCl, the KIT2 oligonucleotide was mainly unfolded; only small amounts of specific K^+ adducts were detected. We thus increased the KCl concentration further, while monitoring both the folding by CD, and the quality of the ESI-MS signals. When the concentration of KCl is increased, the CD spectra showed an increased signal at 265 nm (Figure S2A Supplementary Material). At 3 mM KCl, the CD spectra resembled the one recorded at 50 mM KCl, suggesting the same folding of the oligonucleotide. Due to the limited amount of KCl, the G-quadruplex structures are less stable, as revealed by the melting profile (Figure S2B Supplementary Material), but the apparent melting temperature is still 45°C. The ESI-MS spectra had a reasonable signal-to-noise ratio, and negligible amounts of dimer were detected under these experimental conditions (for full-scale ESI-MS spectra see Figure S3 Supplementary Material). The ESI-MS signals belonging to the dimeric form were too low and too noisy to be properly integrated. For this reason, we avoided quantitative data analyses.

Interestingly, for the monomer, a $1K^+$ ensemble was detected at all the tested KCl concentrations. It reveals a complex composed of two consecutive G-quartets, akin to the long-lived off-pathway intermediates described previously for telomeric and c-myc sequences [10]. Indeed, at low KCl concentration, the

equilibrium is not completely shifted towards the parallel monomeric and dimeric forms, as revealed by the residual CD contribution at 290 nm, not observed at 150 mM KCl (Figure S4 Supplementary Material).

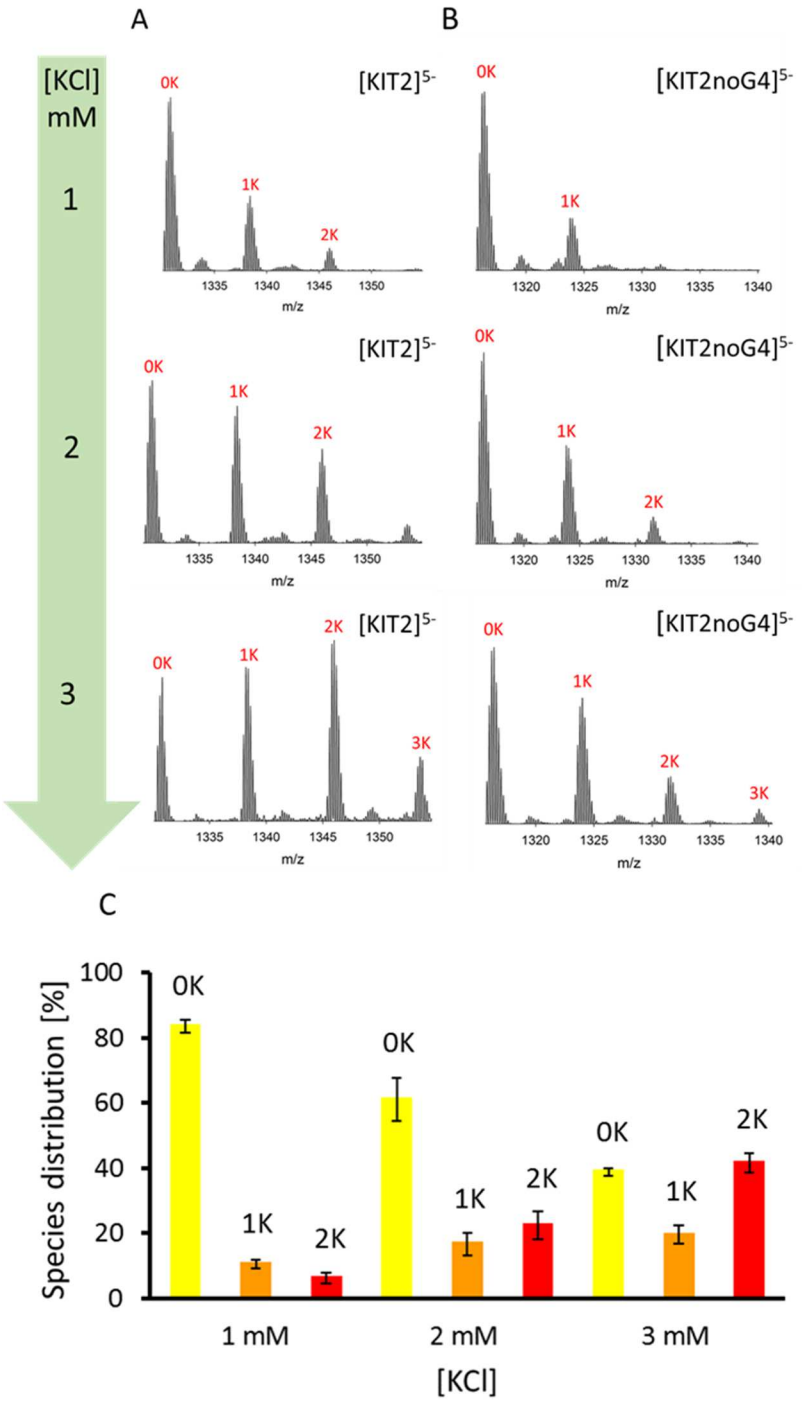


Fig. 2. *A) ESI-MS spectra of 10 µM KIT2 (A) and KIT2noG4 (B) in 100 mM TMAA at increasing KCl concentration, zoomed on the 5- charge state. C) Reconstructed relative distribution of specific K^+ adducts. Error bars are obtained from the standard deviation of three replicates.*

Figure 3 shows ESI-MS spectra of KIT2 in 3 mM KCl after the addition of one equivalent of K20 (Fig. 3A), zoomed on the 5-charge state. The ligand binding stoichiometry was exclusively 1:1. In the reconstructed specific K^+ adducts distribution (Figure 3B), the absence of the $0K^+1L$ complex suggests the incapacity of the ligand to bind to the unfolded oligonucleotide. The most represented DNA-ligand complex was the one

having a single selectively coordinated K^+ ion, while only a little amount of $2K^+1L$ complex was detected. CD spectra recorded on the same sample showed a shift toward the type-III (Figure 3E). Thus, K20 drives the conformational selection of a two-quartet topology, characterized by guanine stacking in alternating orientation.

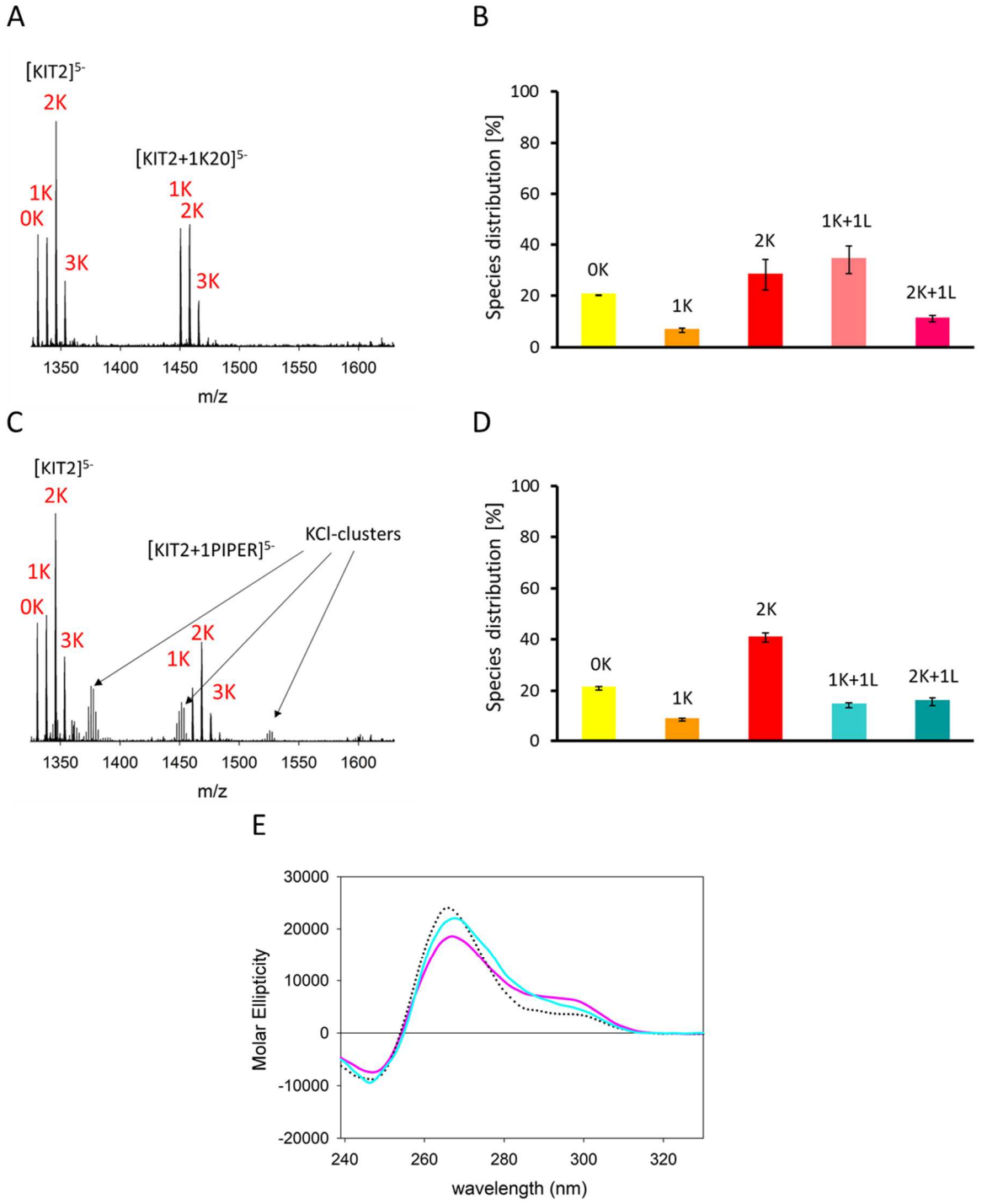


Fig. 3. *A) ESI-MS spectra of 10 µM KIT2 in 100 mM TMAA, 3 mM KCl, after incubation with one equivalent of K20, zoomed on the 5-charge state. B) Reconstructed relative distribution of specific K^+ adducts. Error bars are obtained from the standard deviation of three replicates. C) ESI-MS spectra of 10 µM KIT2 in 100 mM TMAA, 3 mM KCl, after incubation with one equivalent of PIPER, zoomed on the 5-charge state. Arrows point to KCl-clusters that are sometimes detected when electrospraying alkali halides. [34] D) Reconstructed relative distribution of specific K^+ adducts. Error bars are obtained from the standard deviation of three replicates. E) CD spectra of 10 µM KIT2 in 100 mM TMAA,3 mM KCl, before (dotted line) and after (solid line) incubation with one equivalent of K20 (pink line) or PIPER (cyan line).*

3.2 PIPER is unable to drive the same conformational selection of the KIT2 two-quartet antiparallel topology.

As we did for K20, we first checked the propensity of PIPER to induce multimerization of the oligonucleotide, by performing EMSA at increasing ligand/DNA molar ratios (Figure S5 Supplementary Material). We did not observe multimerization up to 3:1 ligand/DNA molar ratio. Interestingly, addition of PIPER did not cause an increment of the electrophoretic mobility of the monomeric G-quadruplex but only a smear of the corresponding band. This suggests a different and less efficient binding mode of PIPER vs K20.

We then applied native ESI-MS to simultaneously determine K^+ and ligand binding stoichiometries. Figure 3C shows ESI-MS spectra of KIT2 in 3 mM KCl after incubation with one equivalent of PIPER, zoomed on the 5- charge state. As for K20, the observed ligand binding stoichiometry was exclusively 1:1. Selectivity for the G-quadruplex structure is proved by the absence of the $0K^+1L$ ensemble. After reconstruction of the specific K^+ adducts distribution (Figure 3D), we observed that the $1K^+1L$ and the $2K^+1L$ complexes were equally represented, suggesting a reduced ability of PIPER to drive the conformational selection of the two-quartet topology. Consistently, CD spectra of the same samples showed only minor changes at 265 nm and 290 nm (Figure 3E).

3.3 K20, compared to PIPER, binds with greater affinity to the two-quartet antiparallel KIT* G-quadruplex.

The preferential binding of K20 to the antiparallel G4 topology, not displayed by its parent ligand PIPER, was further demonstrated by evaluating their binding to the two-quartet antiparallel KIT* G-quadruplex. We acquired ESI-MS spectra of KIT* in 4 mM KCl (Figure S6 Supplementary Material), since at lower [KCl] the oligonucleotide was not folded. ESI-MS spectra of a reference non-G4 forming sequence, KIT*noG4, acquired in the same experimental conditions, allowed reconstructing the specific K^+ adducts distribution. A single specifically coordinated K^+ ion was detected, and this is in line with the reported two-quartet structure. However, the oligonucleotide was only partially folded, hence, CD spectra of the same sample did not fully resemble the one of an antiparallel G-quadruplex (Figure S6D Supplementary Material).

ESI-MS spectra of KIT* in 4 mM KCl, recorded after incubation with one equivalent of K20 (Figure 4A) revealed selective ligand binding to the two-quartet G-quadruplex arrangement, with a 1:1 (KIT*:K20) stoichiometry. Accordingly, the CD spectra acquired on the same sample showed higher contribution at 290 nm (Figure 4E), supporting the ability of K20 to drive the antiparallel folding of the oligonucleotide. PIPER, instead, showed negligible binding to KIT*, as revealed by the ESI-MS and CD spectra. (Figure 4C/E) The modest amount of detected complex corresponded to a 1:1:1 (KIT*:K^+:ligand) binding stoichiometry.

For the binding of K20 and PIPER to KIT*, we could assume equal response factors of the species (See Figure S7 Supplementary Material), hence, we could derive equilibrium dissociation constants from ESI-MS experiments. Values are reported in Table 2.

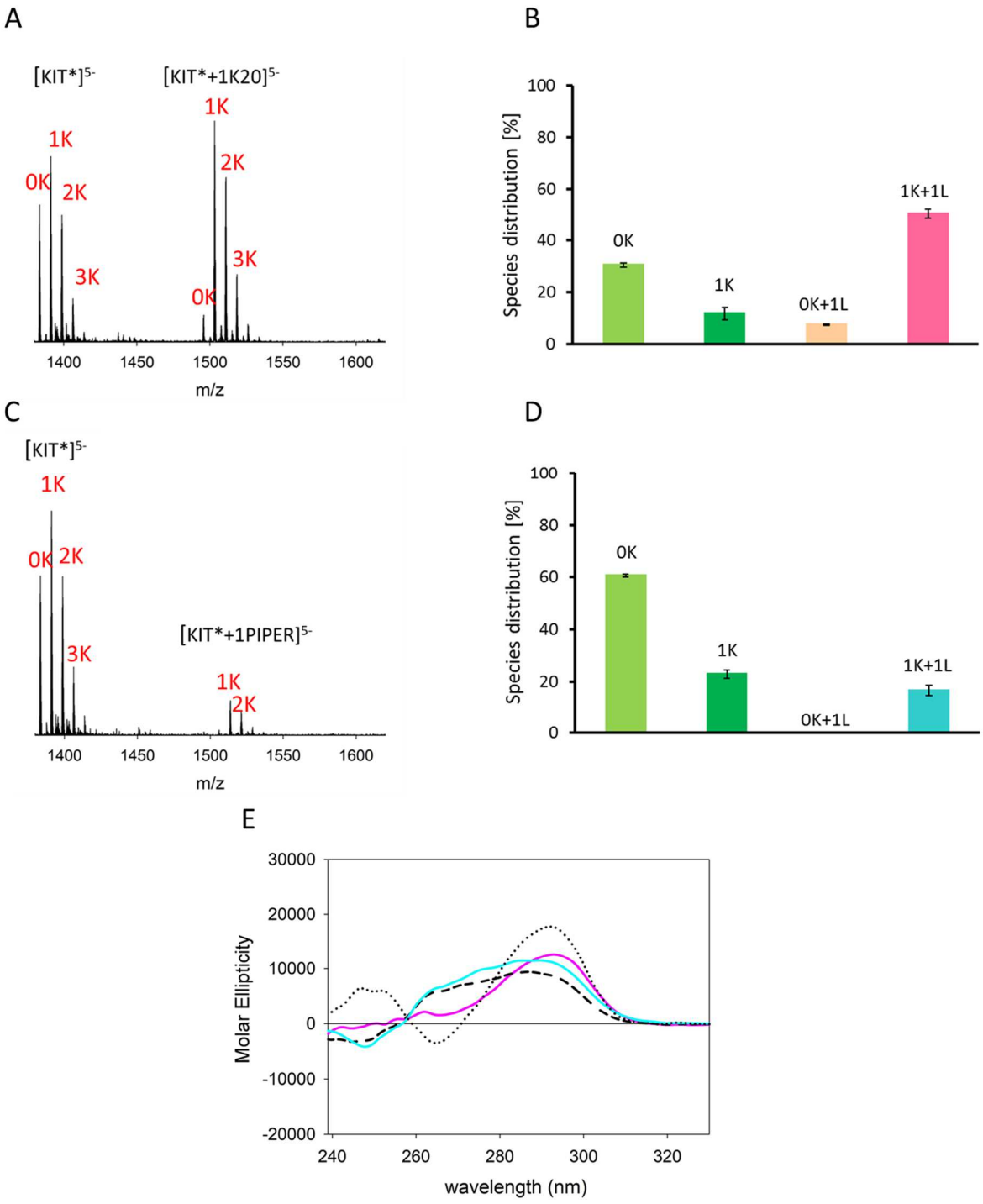


Fig. 4. *A) ESI-MS spectra of 10 µM KIT* in 100 mM TMAA, 4 mM KCl, after incubation with one equivalent of K20, zoomed on the 5-charge state. B) Reconstructed relative distribution of specific K^+ adducts. Error bars are obtained from the standard deviation of three replicates. C) ESI-MS spectra of 10 µM KIT* in 100 mM TMAA, 4 mM KCl, after incubation with one equivalent of PIPER, zoomed on the 5-charge state. D) Reconstructed relative distribution of specific K^+ adducts. Error bars are obtained from the standard deviation of three replicates. E) CD spectra of 10 µM KIT*in 10 mM Tris-HCl, 150 mM KCl (dotted line) and in 100 mM TMAA, 4 mM KCl, before (dashed line) and after (solid line) incubation with one equivalent of K20 (pink line) or PIPER (cyan line).*

Table 2. Equilibrium dissociation constants obtained from ESI-MS experiments performed in 100 mM TMAA, 4 mM KCl, at 1:1 ligand/DNA molar ratio

	ligand **K20**	**PIPER**
sequence **KIT***	$K_D x10^6$(M) 1.0±0.3*	$K_D x10^6$(M) 12±2*

**s.d. from three replicates*

4. DISCUSSION

In a previous study we dissected the G4 folding pathway of KIT2 under different ionic conditions [17]. We identified a short-lived intermediate with antiparallel content (as derived by its circular dichroism spectra). In the presence of K^+, it rapidly evolves toward parallel structures, while it is stabilized in the presence of Na^+. However, we did not obtain structural information about this arrangement. In the presence study, native ESI-MS experiments performed at low K^+ concentration, allowed us to detect and characterize this conformation by means of K^+ binding stoichiometry. Interestingly, the presence of a single specifically coordinated K^+ ion, revealed its two-quartet scaffold.
Furthermore, the perylene derivative K20 was able to drive the conformational selection of this topology, against the more thermodynamically stable parallel three-quartet one. Interestingly, the parent ligand PIPER did not show the same binding preference. The greater affinity of K20 for the two-quartet antiparallel topology, compared to PIPER, was further supported by the fact that it binds to the antiparallel two-quartet KIT* G-quadruplex with a 10-fold smaller dissociation constant. The different binding properties of these ligands can be explained by considering the role of the side chains. Indeed, basic perylene side chains are likely to interact with the phosphate groups in the DNA grooves [19], and they may determine the ligand preference for a given G-quadruplex topology. Taking KIT* as a comparative model for the antiparallel topology, it shows two wide and two narrow grooves that can accommodate the perylene side chains. We can assume that, in the case of K20, thanks to its permanently charged quaternary nitrogen with small size substituents, the energetic gain is significantly larger when the interaction occurs in the narrow vs medium/wide grooves. This would justify either its higher affinity for KIT* and its ability to trap KIT2 in the antiparallel conformation.

We found that the DNA:ligand binding stoichiometry was 1:1 for both ligands, on both the two-quartet and three-quartet scaffolds. Given the reported binding mode of perylene ligands to G-quadruplexes, which is likely to involve end-stacking interactions of the aromatic core with the G-tetrads [19], the 1:1 stoichiometry suggests an asymmetric nature of the tetrads in all the tested G-quadruplex topologies. This is in line with the reported high-resolution structures of KIT2 and KIT*. Indeed, in the parallel three-quartet topology of KIT2, Kuryavyi and co-workers highlighted the presence of a reversed C1•A13 non-canonical pair stacked over the top G-quartet [16], while for KIT* we evidenced a G10•C18 base pair under the bottom G-quartet [26]. We thus expect an asymmetric nature of the G-quartets even in the KIT2 two-quartet antiparallel topology, for which high-resolution structure is not yet available.

5. CONCLUSIONS

It becomes clear that the mere oligonucleotide sequence is not sufficient to uniquely determine the G-quadruplex topology adopted in solution. Indeed, alternative conformations can be explored by a given G-

rich sequence, having different kinetic and thermodynamic profiles. Actually, some of them may not even exist 'in vivo', where the G-rich sequence is inserted into a double-stranded DNA. Thus, to validate a given G-quadruplex forming sequence as a pharmacological target and to properly drive the design of effective selective ligands, it is pivotal to understand which conformation is the most important to exert its biological function within a physiological context. K20 and PIPER share a great structure similarity and common duplex-DNA recognition properties, but we showed that they diverge on the preference for G-quadruplex topology. In addition, these derivatives show a poor short-term cytotoxicity, a property that, on the bases of data on structurally related perylene and naphthalendiimide derivatives, appears to be not simply related to lack of uptake of the protonated fraction [22],[35],[36]. Given these evidences, they can be used as selective probes for the different G-quadruplex topologies during *in vitro* transcription assays or, eventually, *in cell* dual-luciferase reporter assays. As fat it concerns c-KIT, this will eventually lead to the identification of the most relevant conformation of KIT2 for the physiological regulation of the oncogene expression, with the final aim to design selective ligands for it.

FUNDING

This work was supported by University of Padova (grant#SISS_SID19_01 and SC PhD fellowship)

AUTHOR CONTRIBUTIONS:

- conception of the work: CS, VG

- collection of data: SC, EL

- analysis of data: SC, EL, VG, CS

- writing of manuscript: SC, EL, VG, CS

REFERENCES

[1] D. Sen, W. Gilbert, Formation of parallel four-stranded complexes by guanine-rich motifs in DNA and its implications for meiosis, Nature. 334 (1988) 364–366. https://doi.org/10.1038/334364a0.

[2] M. Webba da Silva, Geometric formalism for DNA quadruplex folding, Chem. Weinh. Bergstr. Ger. 13 (2007) 9738–9745. https://doi.org/10.1002/chem.200701255.

[3] A. Wong, G. Wu, Selective binding of monovalent cations to the stacking G-quartet structure formed by guanosine 5'-monophosphate: a solid-state NMR study, J. Am. Chem. Soc. 125 (2003) 13895–13905. https://doi.org/10.1021/ja0302174.

[4] E. Largy, J.-L. Mergny, V. Gabelica, Role of Alkali Metal Ions in G-Quadruplex Nucleic Acid Structure and Stability, Met. Ions Life Sci. 16 (2016) 203–258. https://doi.org/10.1007/978-3-319-21756-7_7.

[5] S. Balasubramanian, L.H. Hurley, S. Neidle, Targeting G-quadruplexes in gene promoters: a novel anticancer strategy?, Nat. Rev. Drug Discov. 10 (2011) 261–275. https://doi.org/10.1038/nrd3428.

[6] R. Rigo, M. Palumbo, C. Sissi, G-quadruplexes in human promoters: A challenge for therapeutic applications, Biochim. Biophys. Acta. 1861 (2017) 1399–1413. https://doi.org/10.1016/j.bbagen.2016.12.024.

[7] T.A. Brooks, S. Kendrick, L. Hurley, Making sense of G-quadruplex and i-motif functions in oncogene promoters, FEBS J. 277 (2010) 3459–3469. https://doi.org/10.1111/j.1742-4658.2010.07759.x.

[8] R. Hänsel-Hertsch, M. Di Antonio, S. Balasubramanian, DNA G-quadruplexes in the human genome: detection, functions and therapeutic potential, Nat. Rev. Mol. Cell Biol. 18 (2017) 279–284. https://doi.org/10.1038/nrm.2017.3.
[9] R.D. Gray, J.O. Trent, J.B. Chaires, Folding and Unfolding Pathways of the Human Telomeric G-Quadruplex, J. Mol. Biol. 426 (2014) 1629–1650. https://doi.org/10.1016/j.jmb.2014.01.009.
[10] A. Marchand, V. Gabelica, Folding and misfolding pathways of G-quadruplex DNA, Nucleic Acids Res. 44 (2016) 10999–11012. https://doi.org/10.1093/nar/gkw970.
[11] Y. Cheng, Q. Tang, Y. Li, Y. Zhang, C. Zhao, J. Yan, H. You, Folding/unfolding kinetics of G-quadruplexes upstream of the P1 promoter of the human BCL-2 oncogene, J. Biol. Chem. 294 (2019) 5890–5895. https://doi.org/10.1074/jbc.RA119.007516.
[12] R.D. Gray, J.O. Trent, S. Arumugam, J.B. Chaires, Folding Landscape of a Parallel G-Quadruplex, J. Phys. Chem. Lett. 10 (2019) 1146–1151. https://doi.org/10.1021/acs.jpclett.9b00227.
[13] P. Stadlbauer, P. Kührová, L. Vicherek, P. Banáš, M. Otyepka, L. Trantírek, J. Šponer, Parallel G-triplexes and G-hairpins as potential transitory ensembles in the folding of parallel-stranded DNA G-Quadruplexes, Nucleic Acids Res. 47 (2019) 7276–7293. https://doi.org/10.1093/nar/gkz610.
[14] H. Fernando, A.P. Reszka, J. Huppert, S. Ladame, S. Rankin, A.R. Venkitaraman, S. Neidle, S. Balasubramanian, A Conserved Quadruplex Motif Located in a Transcription Activation Site of the Human c-kit Oncogene, Biochemistry. 45 (2006) 7854–7860. https://doi.org/10.1021/bi0601510.
[15] S.-T.D. Hsu, P. Varnai, A. Bugaut, A.P. Reszka, S. Neidle, S. Balasubramanian, A G-Rich Sequence within the c-kit Oncogene Promoter Forms a Parallel G-Quadruplex Having Asymmetric G-Tetrad Dynamics, J. Am. Chem. Soc. 131 (2009) 13399–13409. https://doi.org/10.1021/ja904007p.
[16] V. Kuryavyi, A.T. Phan, D.J. Patel, Solution structures of all parallel-stranded monomeric and dimeric G-quadruplex scaffolds of the human c-kit2 promoter, Nucleic Acids Res. 38 (2010) 6757–6773. https://doi.org/10.1093/nar/gkq558.
[17] R. Rigo, W.L. Dean, R.D. Gray, J.B. Chaires, C. Sissi, Conformational profiling of a G-rich sequence within the c-KIT promoter, Nucleic Acids Res. 45 (2017) 13056–13067. https://doi.org/10.1093/nar/gkx983.
[18] E. Puig Lombardi, A. Holmes, D. Verga, M.-P. Teulade-Fichou, A. Nicolas, A. Londoño-Vallejo, Thermodynamically stable and genetically unstable G-quadruplexes are depleted in genomes across species, Nucleic Acids Res. 47 (2019) 6098–6113. https://doi.org/10.1093/nar/gkz463.
[19] O.Yu. Fedoroff, M. Salazar, H. Han, V.V. Chemeris, S.M. Kerwin, L.H. Hurley, NMR-Based Model of a Telomerase-Inhibiting Compound Bound to G-Quadruplex DNA, Biochemistry. 37 (1998) 12367–12374. https://doi.org/10.1021/bi981330n.
[20] A. Rangan, O.Y. Fedoroff, L.H. Hurley, Induction of duplex to G-quadruplex transition in the c-myc promoter region by a small molecule, J. Biol. Chem. 276 (2001) 4640–4646. https://doi.org/10.1074/jbc.M005962200.
[21] H. Han, C.L. Cliff, L.H. Hurley, Accelerated assembly of G-quadruplex structures by a small molecule, Biochemistry. 38 (1999) 6981–6986. https://doi.org/10.1021/bi9905922.
[22] C. Sissi, L. Lucatello, A. Paul Krapcho, D.J. Maloney, M.B. Boxer, M.V. Camarasa, G. Pezzoni, E. Menta, M. Palumbo, Tri-, tetra- and heptacyclic perylene analogues as new potential antineoplastic agents based on DNA telomerase inhibition, Bioorg. Med. Chem. 15 (2007) 555–562. https://doi.org/10.1016/j.bmc.2006.09.029.
[23] C. Pivetta, L. Lucatello, A. Paul Krapcho, B. Gatto, M. Palumbo, C. Sissi, Perylene side chains modulate G-quadruplex conformation in biologically relevant DNA sequences, Bioorg. Med. Chem. 16 (2008) 9331–9339. https://doi.org/10.1016/j.bmc.2008.08.068.
[24] A. Marchand, A. Granzhan, K. Iida, Y. Tsushima, Y. Ma, K. Nagasawa, M.-P. Teulade-Fichou, V. Gabelica, Ligand-induced conformational changes with cation ejection upon binding to human telomeric DNA G-quadruplexes, J. Am. Chem. Soc. 137 (2015) 750–756. https://doi.org/10.1021/ja5099403.
[25] E.-A. Raiber, R. Kranaster, E. Lam, M. Nikan, S. Balasubramanian, A non-canonical DNA structure is a binding motif for the transcription factor SP1 in vitro, Nucleic Acids Res. 40 (2012) 1499–1508. https://doi.org/10.1093/nar/gkr882.

[26] A. Kotar, R. Rigo, C. Sissi, J. Plavec, Two-quartet kit* G-quadruplex is formed via double-stranded pre-folded structure, Nucleic Acids Res. 47 (2019) 2641–2653. https://doi.org/10.1093/nar/gky1269.
[27] M.J. Cavaluzzi, P.N. Borer, Revised UV extinction coefficients for nucleoside-5'-monophosphates and unpaired DNA and RNA, Nucleic Acids Res. 32 (2004) e13–e13. https://doi.org/10.1093/nar/gnh015.
[28] F. Balthasart, J. Plavec, V. Gabelica, Ammonium Ion Binding to DNA G-Quadruplexes: Do Electrospray Mass Spectra Faithfully Reflect the Solution-Phase Species?, J. Am. Soc. Mass Spectrom. 24 (2013) 1–8. https://doi.org/10.1007/s13361-012-0499-3.
[29] V. Gabelica, F. Rosu, E. De Pauw, A Simple Method to Determine Electrospray Response Factors of Noncovalent Complexes, Anal. Chem. 81 (2009) 6708–6715. https://doi.org/10.1021/ac900785m.
[30] A.I. Karsisiotis, N.M. Hessari, E. Novellino, G.P. Spada, A. Randazzo, M. Webba da Silva, Topological characterization of nucleic acid G-quadruplexes by UV absorption and circular dichroism, Angew. Chem. Int. Ed Engl. 50 (2011) 10645–10648. https://doi.org/10.1002/anie.201105193.
[31] I. Kejnovská, K. Bednárová, D. Renciuk, Z. Dvoráková, P. Školáková, L. Trantírek, R. Fiala, M. Vorlícková, J. Sagi, Clustered abasic lesions profoundly change the structure and stability of human telomeric G-quadruplexes, Nucleic Acids Res. 45 (2017) 4294–4305. https://doi.org/10.1093/nar/gkx191.
[32] A. Marchand, V. Gabelica, Native Electrospray Mass Spectrometry of DNA G-Quadruplexes in Potassium Solution, J. Am. Soc. Mass Spectrom. 25 (2014) 1146–1154. https://doi.org/10.1007/s13361-014-0890-3.
[33] J. Gros, F. Rosu, S. Amrane, A. De Cian, V. Gabelica, L. Lacroix, J.-L. Mergny, Guanines are a quartet's best friend: impact of base substitutions on the kinetics and stability of tetramolecular quadruplexes, Nucleic Acids Res. 35 (2007) 3064–3075. https://doi.org/10.1093/nar/gkm111.
[34] C. Hao, R.E. March, T.R. Croley, J.C. Smith, S.P. Rafferty, Electrospray ionization tandem mass spectrometric study of salt cluster ions. Part 1— Investigations of alkali metal chloride and sodium salt cluster ions, J. Mass Spectrom. 36 (2001) 79–96. https://doi.org/10.1002/jms.107.
[35] T. Taka, K. Joonlasak, L. Huang, T. Randall Lee, S.-W.T. Chang, W. Tuntiwechapikul, Down-regulation of the human VEGF gene expression by perylene monoimide derivatives, Bioorg. Med. Chem. Lett. 22 (2012) 518–522. https://doi.org/10.1016/j.bmcl.2011.10.089.
[36] A.M. Zhirov, D.A. Kovalev, D.V. Ulshina, S.V. Pisarenko, O.P. Demidov, I.V. Borovlev, Diazapyrenes: interaction with nucleic acids and biological activity, Chem. Heterocycl. Compd. (2020) 1–20. https://doi.org/10.1007/s10593-020-02717-1.

GRAPHICAL ABSTRACT