Supporting Information

Intercalation of a Heterocyclic Ligand between Quartets in a G-Rich Tetrahelical Structure

Anita Kotar\textsuperscript{[a]} Vojč Kocman\textsuperscript{[a]} and Janez Plavec\textsuperscript{[a, b, c]}

cem_201904923_sm_misellaneous_information.pdf
Table of Contents
1. Materials and methods ................................................................................................................... 2
   NMR sample preparation .................................................................................................................. 2
   NMR spectroscopic experiments ..................................................................................................... 2
   $K_d$ determination from the NMR titration data .............................................................................. 2
   CD spectroscopy ............................................................................................................................. 3
   Restraints and structure calculations ............................................................................................... 3
   Data deposition ............................................................................................................................... 4
2. Supporting Information Figures ...................................................................................................... 5
3. Tables ............................................................................................................................................ 15
4. References .................................................................................................................................... 16
1. Materials and methods

NMR sample preparation
The isotopically unlabeled and residue-specific partially $^{13}$C,$^{15}$N-labelled (10% guanine residues) oligonucleotides were synthesized on K&A Laborgeraete GbR DNA/RNA Synthesizer H-8. In all cases, standard phosphoramidite chemistry was used. Deprotection was done with use of aqueous ammonia at 55 °C for 14 h. Samples were purified and desalted with the use of Amicon Ultra-15 Centrifugal Filter Devices to give NMR samples with concentration between 0.01 and 0.5 mM per strand. NMR samples of VK2 alone and in the complexes were prepared by dissolving desalted oligonucleotides in 300 or 500 μL of 90%/10% mixture of H$_2$O/2H$_2$O and in the presence of LiCl with 100 mM concentration as well as in 100 mM concentrations of KCl. pH value of the samples was set to 6.0 with the use of LiOH and HCl. Extra care was taken to ensure that only specified cations were present in NMR samples. For NMR titration experiments, 360A was first dissolved in DMSO-d6 to 10mM stock concentration. For PDS and Phen-DC3 ligands 20 mM stock solutions were prepared in water and DMSO-d6, respectively. Ligands were added stepwise into solution of oligonucleotide as indicated on the left side of NMR spectra for each ligand. The concentrations of DMSO-d6 used in this study did not cause any spectral differences, which could account for structural changes of DNA influenced by DMSO-d6. The initial batches of 360A and Phen-DC3 ligands were obtained from the group of Prof. Marie-Paule Teulade-Fichou. Further batches of ligands 360A were bought from MedChemExpress, Phen-DC3 and Piridostatin from Sigma-Aldrich.

NMR spectroscopic experiments
NMR experiments were performed on Agilent-Varian NMR System 600 and 800 MHz spectrometers equipped with triple-resonance HCN cryogenic probes. Spectra were recorded at 0 °C in 90% H$_2$O and 10% 2H$_2$O. 1D $^1$H NMR and 2D $^1$H-$^1$H NOESY spectra were recorded with the use of DPFGSE solvent suppression method. NOESY spectra were acquired with mixing times 80-200 ms. 1D $^{15}$N-edited HSQC NMR spectra were measured on ten samples which contained 10% guanine residue-specific $^{13}$C,$^{15}$N-labelled oligonucleotides. NMR spectra were referenced to external trimethylsilylpropanoic acid reference at δ 0.0 ppm in 90%/10% H$_2$O/2H$_2$O.

$K_d$ determination from the NMR titration data
By following changes in intensity of a well-resolved signal of A8H2” of VK2-360A complex, the dissociation constants ($K_d$) were determined by fitting the data to a noncooperative 1:1 binding mode using non-linear regression analysis:

$$\Delta I_{[HG]} = \frac{I_{[HG]} - I_{[HG]}}{I_{[HG]} - I_{[HG]}} [HG]$$

(1)

where $\Delta I_{[HG]}$ represents the change in integral value of A8H2” signal after each addition of the ligand, $I_{[HG]}$ is maximal change in integral values, $[H_0]$ is the concentration of VK2, $[HG]$ is the concentration of the complex present in the solution determined with the help of Equation (2)$^{11}$:

$$[HG] = \frac{1}{2} \left( [G_0] + [H_0] + \frac{1}{K_d} \right) - \frac{1}{4} \left( [G_0] + [H_0] + \frac{1}{K_d} \right)^2 - [G_0][H_0]$$

(2)
where $G_0$ and $H_0$ are total concentrations of ligand and VK2, respectively. $K_a$ is an association constant. Noteworthy, dissociation constant ($K_d$) is the inverse of the association constant ($K_d = \frac{1}{K_a}$).

**CD spectroscopy**

CD experiments were carried out on an Applied Photophysics Chirascan CD spectrometer over 210-340 nm wavelength range. All measurements were done in 1 mm path-length quartz cells at 0 °C. The experiments were done on samples at 0.025 mM oligonucleotide concentration per strand, 100 mM LiCl (pH 6.0), and at 0-0.1 mM concentration of ligands. It is noteworthy that CD spectra of AGCGA-quadruplexes have a characteristic profile with minimum at 245 nm, maximum at 266 nm and two shoulders at 280 and 290 nm.

Thermal stability measurements were done on samples at 0.025 mM DNA concentration and 100 mM LiCl (pH 6.0). Folding/unfolding processes were followed between 0 and 95 °C by measuring changes of CD signal at 266 or 267 nm using scanning rate of 1 °C min⁻¹. Previous studies showed that VK2 melting is not monophasic, but involves at least two transition steps (2). Thus, we analyzed the melting profile of VK2 by using a two-transitions model based on van’t Hoff formalism (3, 4):

$$\theta_T = \frac{u \cdot e^{-\frac{\Delta H_1}{RT} + \frac{\Delta H_2}{RT}} + a \cdot e^{-\frac{\Delta H_1}{RT} + \frac{\Delta H_2}{RT}}}{e^{-\frac{\Delta H_1}{RT} + \frac{\Delta H_2}{RT}} + e^{-\frac{\Delta H_1}{RT} + \frac{\Delta H_2}{RT}}} + i$$

where $\theta_T$ is the CD signal at wavelength $\lambda = 266/267$ nm at temperature $T$, $u$, $a$ and $i$ are fitting parameters, $\Delta H_1$ and $\Delta H_2$ are the enthalpies associated to each melting step, $R$ is the ideal gas constant, $T_{m1}$ and $T_{m2}$ are the melting temperatures.

**Restraints and structure calculations**

The structures of 1:1 VK2-360A complex were calculated by the simulated annealing simulations. Restraints used in the calculation were NOE-derived distance restraints (force constant 20 kcal mol⁻¹ Å⁻²), hydrogen bond (force constant 20-40 kcal mol⁻¹ Å⁻²), and $\nu_2$ restraints (force constant 200 kcal mol⁻¹ rad⁻²). Distance restraints were obtained from the NOESY spectra recorded at 200 ms mixing time. Cytosines H5-H6 cross peaks are strongly overlapped with the water signal and thus could not be used as reference. We rely on averaging the volumes of NOE of correlations of H8 and H1’ protons where anti orientation is clearly observed and we could reference the averaged volumes to the value of 3.9 Å. By this reference we classified the remaining signals as strong (1.8 - 3.6 Å), medium (2.6 – 5.0 Å) and weak (3.5 – 6.5 Å). Torsion angle restraints along glycosidic bonds (torsion angle $\chi$) for all residues were set 200°-280° for guanine and adenine residues and 170°-310° for cytosine residues, values typical for residues in anti-orientation. The force field parameters were adopted from the Generalized Amber force field using AMBER 18 program suites and basic force field leaprc.ff99SB (5). A total of 100 structures were calculated in 200 ps of NMR restrained simulated annealing simulations using the generalized Born implicit model (6). Simulated annealing calculations were initiated with random velocities. The simulated annealing simulation was as follows: in 0-5 ps, the temperature was raised from 300 K to 1000 K and held constant at 1000 K for 95 ps. Temperature was scaled down to 500 K in the next 60 ps and reduced to 100 K in the next 20 ps and was further reduced to 0 K in the last 20 ps. A family of 10
structures with the smallest restraints violations and lowest energy was subjected to a maximum of 10,000 steps of steepest descent energy minimization. Structures were visualized with UCSF Chimera.

**Data deposition.**

The coordinates for structure of 1:1 VK2-360A complex have been deposited in the Protein Data Bank (accession number: 6SX3) and Biological Magnetic Resonance Data Bank (accession number: 34435).
2. Supporting Information Figures

**Figure S1.** Oligonucleotide sequence of VK2 and schematic presentation of topology of free VK2. Guanines are shown in blue, adenines in red, and cytosines in yellow.

**Figure S2.** Selected regions of $^1$H NMR spectra of VK2 in the presence of K$^+$ ions as 360A is titrated into the solution. The molar ratio between VK2 and 360A is indicated on the left side of spectra. The vertical scale of imino and sugar regions of all spectra were increased 5-fold compared to aromatic region. The NMR spectra were recorded at 0.2 mM VK2 per strand, 100 mM KCl, pH 6.0, 0 °C on a 600 MHz NMR spectrometer.
Figure S3. Selected regions of $^1$H NMR spectra of VK2 as 360A is titrated into solution. The molar ratio between VK2 and 360A is indicated on the left side of spectra. Assignments are shown above the individual signals and marked in black for free VK2, in magenta for VK2 in the complex, in green for 360A in the complex. The vertical scale of imino and sugar regions of all spectra was increased 5-fold compared to aromatic region. The NMR spectra were recorded at 0.2 mM VK2 per strand, 100 mM LiCl, pH 6.0, 0 °C on a 600 MHz NMR spectrometer.
Figure S4. Intensity of A8H2" signal in VK2-360A complex plotted against concentration of 360A added into solution. An overall $K_d$ value of $7 \pm 1 \mu M$ was obtained by fitting the data to a non-cooperative binding mode as described in Materials and methods section.

Figure S5. $^1$H NMR spectra of 360A alone and in the presence of VK2. The molar ratio between VK2 and 360A is indicated on the left side of spectra. Signals of unknown impurities are labeled with #. The NMR spectra were recorded at 0.01 mM VK2 per strand, 100 mM LiCl, pH 6.0, 0 °C on a 600 MHz NMR spectrometer.
Figure S6. Imino regions of $^1$H and 1D $^{15}$N-edited HSQC NMR spectra of a) free VK2 and b) VK2 in VK2-360A complex. HSQC spectra were acquired on partially (10%) guanine-specifically $^{13}$C,$^{15}$N-labelled oligonucleotides. Assignment of H1 proton resonances is indicated on the left side of each 1D HSQC spectrum. NMR spectra were recorded at 0.3-0.4 mM VK2 concentration per strand, 100 mM LiCl, pH 6, at 0 °C, on 800 MHz and 600 MHz NMR spectrometers. All NMR spectra of VK2-360A complex were acquired at 1:1 VK2:360A ratio.
Figure S7. Aromatic-anomeric regions of $^1$H-$^1$H NOESY spectrum ($\tau_m$ 200 ms) of a) free VK2 and b) VK2-360A complex. NMR spectra were recorded at 0.3 mM VK2 concentration per strand, 1.2 mM 360A concentration (b), 100 mM LiCl, pH 6, 0 °C, on an 800 MHz spectrometer. Assignments are shown above the individual signals; the signals of 360A are denoted by the letter L. The G1-A8 and G9-G15 sequential walks are depicted in blue and cyan, respectively. Sequential assignment of new VK2 cross-peaks and connections between VK2 and 360A are shown in magenta.

Figure S8. Expansions of $^1$H-$^1$H NOESY spectrum ($\tau_m$ 200 ms) of VK2-360A complex showing NOE cross-peaks between VK2 and 360A. NMR spectrum was recorded at 0.3 mM VK2 concentration per strand, 1.2 mM 360A concentration, 100 mM LiCl, pH 6, 0 °C, on an 800 MHz spectrometer. Assignments are shown above the individual signals; the signals of 360A are denoted by the letter L. Very weak signals are marked only with crosses.
**Figure S9.** Schematic presentation of position of 360A in VK2 structure and observed NOE contacts (marked with black lines) between protons of 360A and C6, G7 and A8 residues of VK2. 360A is shown in green, G7 in blue, A8 in red, and C6 in gold.

**Figure S10.** Aromatic-aromatic region of $^1$H-$^1$H NOESY spectrum ($\tau_m$ 200 ms) of VK2-360A complex. Very close proximity of L-H4, G13H8 and A12H2 as well as L-H3 and G11H8 signals prevents identification of cross-peaks between A12, G13 protons and ligand protons. NMR spectrum was recorded at 0.3 mM VK2 concentration per strand, 1.2 mM 360A concentration, 100 mM LiCl, pH 6, 0 °C, on an 800 MHz spectrometer. Assignments are shown above the individual signals; the signals of 360A are denoted by the letter L.
Figure S11. a) A structure ensemble of the 10 lowest-energy structures of VK2–360A complex with the smallest restraints violations (PDB ID: 6SX3). b) The core of VK2 consisting of GAGA- and GCGC-quartets in the presence of 360A. c) Change of the ligand position in structure ensemble of the 10 lowest-energy structures of the complex.

Figure S12. The structure of the 1:1 VK2–360A complex (PDB ID: 6SX3). a) A wide groove view of the whole VK2-360A complex. b) Top view of the G-G base pairs and A4-G15 base pairs. c) The two A8 residues surrounded by G-G base pairs in the bottom fold-back arrangement. 360A is shown in green, guanines in blue, adenines in red, and cytosines in yellow.
Figure S13. The chemical structures of Phen-DC3, 360A and Piridostatin ligands from bis-quinolinium family.

Figure S14. Imino and aromatic regions of $^1$H NMR spectra of VK2 as a) PDS and b) Phen-DC3 are titrated into solution. The molar ratios between VK2 and the ligands are indicated on the left side of the spectra. The vertical scales of imino regions of all spectra were increased 5-fold compared to aromatic region. The NMR spectra were recorded at 0.2 mM of VK2, 100 mM LiCl, pH 6.0 and 0 °C on an 800 MHz spectrometer.
Figure S15. The signals of A8H2" in 1H NMR spectra of VK2 as a) Piridostatin and b) Phen-DC3 were titrated into solution. The molar ratios between VK2 and the ligands are indicated on the right side of the spectra. Assignments are shown above the corresponding spectra and marked in black for free VK2, in blue for VK2-Piridostatin complex and in purple for VK2-Phen-DC3 complex. Higher amounts of ligands are needed for formation of complexes with VK2 compared to 360A. The NMR spectra were recorded at 0.2 mM of VK2, 100 mM LiCl, pH 6.0 and 0 °C on an 800 MHz spectrometer. Signals of unknown impurities are labeled with asterisks. Complexes are also less stable compared to VK2-360A complex as it is indicated by observation of precipitation in the sample during NMR titration especially at higher amount of both ligands. Determination of $K_d$ value for VK2-Piridostatin complex was prevented due to severe precipitation observed in the sample.

Figure S16. Intensity of A8H2" signal in VK2-Phen-DC3 complex plotted against concentration of Phen-DC3 added into solution. An overall $K_d$ value of 111 ± 5 μM was obtained by fitting the data to a non-cooperative binding mode as described in Materials and methods section.
**Figure S17.** CD melting profile (0-95 °C) of free VK2 (black), VK2-360A complex (magenta), VK2-Phen-DC complex (purple) and VK2-Piridostatin complex (blue) monitored at 266 nm or 267 nm (for VK2-Phen-DC complex). Molar ellipticities at 0 °C were normalized to 1.
3. Tables

**Table S1.** $^1$H NMR chemical shifts changes induced by interactions between VK2 and 360A.$^a$

| Residue / Proton | H1  | H8/H6 | H1’ | H2  | H2’ | H3  | H4  |
|------------------|-----|-------|-----|-----|-----|-----|-----|
| G1               | 0.03| 0.01  | 0.03| 0.04| 0.06| 0.03| 0.02|
| G2               | 0.04| 0.02  | 0.04| 0.00| 0.00| 0.04| 0.04|
| G3               | 0.01| 0.06  | 0.01| 0.05| 0.03| nd  |
| A4               | /   | 0.00  | 0.03| 0.01| 0.10| 0.55| 0.36|
| G5               | 0.04| 0.07  | -0.52| 0.04| 0.06| 0.09| 0.15|
| C6               | /   | 0.07  | 0.06| 0.18| 0.14| 0.10| 0.10|
| G6               | /   | 0.12  | 0.06| 0.18| 0.14| 0.10| 0.10|
| G7               | 0.06| 0.14  | 0   | -0.09| -0.16| -0.15| nd  |
| G7               | 0.02| -0.04 | 0   | -0.09| -0.16| -0.15| nd  |
| A8               | /   | 0.40  | -0.02| -0.50| -0.65| 0.34| 0.06|
| A8               | /   | 0.17  | -0.02| -0.50| -0.65| 0.34| 0.06|
| G9               | 0.33| 0.01  | -0.17| 0.38| 1.52| 0.25| 0.22|
| G10              | 0.02| -0.63 | -0.12| 0.01| 0.53| 0.43| 0.23|
| G11              | 0.29| 0.10  | 0.05| 0.09| 0.07| 0.00| 0.13|
| A12              | /   | -0.09 | 0.11| 0.09| 0.10| 0.04| 0.05|
| G13              | 0.04| 0.01  | 0.03| 0.10| -0.01| 0.03| 0.06|
| G14              | /   | 0.04  | 0.05| 0.03| 0.05| 0.10| 0.07|
| G15              | nd  | -0.05 | 0.07| 0.36| -0.19| 0.059| -0.01|
| A16              | /   | 0.02  | 0.18| -0.03| 0.03| 0.13| nd  |

$^a$ $^1$H NMR chemical shifts are given in ppm are measured in 90% $\text{H}_2\text{O} / 10\% ^1\text{H}_2\text{O}, 0 \degree \text{C}, 0.2 \text{mM concentration of VK2 oligonucleotide per strand}, 0.2 \text{mM concentration of 360A}, 100 \text{mM HCl (pH 6.0)}$ and referenced to TMSpa (δ 0.0 ppm). nd - not determined.

**Table S2.** $^1$H NMR chemical shifts of 1:1 VK2-360A complex.$^a$

| Residue / Proton | H1  | H8/H6 | H2/H5 | H1’ | H2  | H2’ | H3  | H4  |
|------------------|-----|-------|-------|-----|-----|-----|-----|-----|
| G1               | 11.09| 7.66  |       | 5.84| 2.09| 2.22| 4.66| 4.05|
| G2               | 10.74| 7.66  |       | 5.59| 2.11| 2.22| 4.78| 4.08|
| G3               | /    | 7.68  |       | 5.71| 2.43| 2.50| 4.84| nd  |
| A4               | /    | 7.94  | nd    | 6.18| 2.42| 2.69| 4.33| 4.01|
| G5               | 12.93| 7.95  |       | 5.87| 2.52| 2.42| 4.82| 4.21|
| C6               | /    | 6.89  | 4.90  | 6.02| 2.09| 1.37| 4.74| 4.17|
| C6               | /    | 6.83  |       | 6.02| 2.09| 1.37| 4.74| 4.17|
| G7               | 10.28| 8.11  |       | 6.02| 2.89| 3.05| 4.69| nd  |
| G7               | 10.32| 8.28  |       | 6.02| 2.89| 3.05| 4.69| nd  |
| A8               | /    | 7.07  | 7.76  | 5.42| 1.64| -0.13| 4.43| 4.27|
| A8               | /    | 7.31  | 7.76  | 5.42| 1.64| -0.13| 4.43| 4.27|
| G9               | 10.17| 6.92  |       | 5.83| 2.09| 1.43| 4.49| 4.11|
| G10              | 10.92| 8.03  |       | 5.80| 2.10| 1.44| 4.43| 4.08|
| G11              | 10.14| 8.04  |       | 5.87| 2.82| 2.58| 4.49| 4.08|
| A12              | /    | 7.40  | 7.92  | 5.45| 2.26| 0.97| 4.74| 4.26|
| G13              | 12.78| 7.85  |       | 5.36| 2.42| 2.49| 4.74| 4.26|
| C14              | /    | 6.95  | 4.99  | 5.87| 1.52| 2.09| 4.67| 4.08|
| G15              | *    | 8.04  |       | 5.45| 2.25| 2.49| 4.56| 4.24|
| A16              | /    | 7.61  | 7.75  | 5.46| 2.64| 2.34| 4.61| nd  |

$^a$ $^1$H NMR chemical shifts are given in ppm are measured in 90% $\text{H}_2\text{O} / 10\% ^1\text{H}_2\text{O}, 0 \degree \text{C}, 0.2 \text{mM concentration of VK2 oligonucleotide per strand}, 0.2 \text{mM concentration of 360A}, 100 \text{mM HCl (pH 6.0)}$ and referenced to TMSpa (δ 0.0 ppm). nd - not determined. *Many signals were observed between 10 and 11 ppm for G15 residue as shown on Figure S6b.
Table S3. NMR restraints and structural statistics for calculation of high-resolution structure of 1:1 VK2-360A complex through simulated annealing calculation.

| NOE-derived distance restraints                        |       |
|--------------------------------------------------------|-------|
| Total                                                  | 314   |
| Intra-residue NOEs                                      | 159   |
| Inter-residue                                          | 155   |
| Sequential                                             | 89    |
| Long-range                                             | 30    |
| Ligand restraints                                       | 36    |
| Torsion angle restraints                                | 31    |
| Hydrogen bond restraints                                | 37    |

| Deviation from idealized covalent geometry              |       |
|--------------------------------------------------------|-------|
| Average distance violation (Å)                          | 0.11 ± 0.04 |
| Maximum distance violation (Å)                          | 0.3   |
| Average torsion angle violation (deg)                   | 0.89 ± 0.08 |
| Maximum torsion angle violation (deg)                   | 1.104 |

| Pairwise heavy-atom r.m.s.d. (Å)                        |       |
|--------------------------------------------------------|-------|
| AGCGA-quadruplex with ligand                            | 2.6 ± 0.3 |
| AGCGA-quadruplex                                       | 2.6 ± 0.6 |
| Core of AGCGA-quadruplex (residues G5-G7 and A12-C14)  | 1.2 ± 0.2 |
| Core of AGCGA-quadruplex with ligand                    | 1.1 ± 0.2 |
| Hydrophobic pocket of AGCGA-quadruplex (residues G7-A12)| 1.3 ± 0.3 |
| Core and GG and GA base pairs (residues G1-A4)          | 2.6 ± 0.6 |
| Core and GG base pairs (residues G9-G11)                | 1.3 ± 0.3 |
| GG and GA base pairs (residues G1-A4)                   | 3.0 ± 0.7 |
| GG base pairs (residues G9-G11)                         | 1.3 ± 0.4 |
| Core and GG base pairs (residues G9-G11) and ligand     | 1.2 ± 0.3 |
| Without G1-A4 and G15 residues                          | 1.2 ± 0.2 |

r.m.s.d. – root mean squared deviation

4. References

1. L. Fielding: NMR methods for the determination of protein–ligand dissociation constants. *Prog. Nucl. Magn. Reson. Spectrosc.* 2007; 51: 219-242.
2. S. Hadži, V. Kocman, D. Oblak, J. Plavec, J. Lah: Energetic Basis of AGCGA-Rich DNA Folding into a Tetrahelical Structure. *Angew. Chem. Int. Ed. Engl.* 2019; 58 (8): 2387-2391.
3. R. D. Gray, R. Buscaglia, J. B. Chaires: Populated intermediates in the thermal unfolding of the human telomeric quadruplex. *J. Am. Chem. Soc.* 2012; 134 (40): 16834-16844.
4. N. J. Greenfield: Using circular dichroism collected as a function of temperature to determine the thermodynamics of protein unfolding and binding interactions. *Nat. Protoc.* 2006; 1 (6): 2527-2535.
5. D. A. Case, I. Y. Ben-Shalom, S. R. Brozell, D. S. Cerutti, I. Cheatham, T. E., V. W. D. Cruzeiro, et al.: AMBER 2018, University of California, San Francisco. 2018.
6. A. Onufriev, D. Bashford, D. A. Case: Modification of the generalized Born model suitable for macromolecules. *J. Phys. Chem. B* 2000; 104 (15): 3712-3720.