Helical foldamers as selective G-quadruplex ligands

Alexander König,[1] Vincent Laffilé,[2] Stéphane Thore,[1] Cameron D. Mackereth,[1] Liliya Yatsunyk,[3] Yann Ferrand,[2] Eric Largy,[1*] Valérie Gabelica[1,4]*

[1] Univ. Bordeaux, CNRS, INSERM, ARNA, UMR 5320, U1212, F-33600 Bordeaux, France (eric.largy@u-bordeaux.fr)

[2] Univ. Bordeaux, CNRS, IPB, CBMN, UMR 5248, IECB, F-33600 Bordeaux, France

[3] Department of Chemistry and Biochemistry, Swarthmore College, Swarthmore, PA 19081, United States

[4] School of Pharmaceutical Sciences, University of Geneva, Geneva, Switzerland (valerie.gabelica@unige.ch)

Abstract

We investigated the G-quadruplex binding selectivity of short aromatic oligoamide helical foldamers comprising quinoline (Q) and pyridine (P) units. We found that the foldamers prefer parallel G-quadruplex structures, especially when the external G-tetrads are sterically accessible. A crystal structure of the tetramer QQPQ with the parallel G-quadruplex formed by dTGGGTTGGGTTGGGTTGGGT shows two quinoline subunits interacting with an external G-tetrad through π-stacking, and solution NMR confirms that the foldamer targets the 3' and 5' ends of this G-quadruplex. Foldamers can also selectively target sequence variants of the telomeric sequences containing adenine-to-thymine mutation in the loops. The conformational selectivity of foldamers originates from the bulkiness of oligomers with four or more subunits, which imposes steric restrictions on G-quadruplex binding. The flexibility provided by the pyridine subunits was also key to improve affinity. Mixed quinoline-pyridine foldamers are thus a promising class of selective G-quadruplex ligands, and their unique modular scaffold offers new avenues to further improve their affinity and selectivity.

Introduction

G-quadruplexes (G4s) are non-canonical DNA structures wherein four guanines associate into a tetrad through Hoogsten-type H-bonding.[1] G4 folding topologies are diverse and some sequences can even form several energetically close G4 topologies.[2–7] The structure(s) adopted by a given DNA sequence can depend on the concentration of DNA strand, the type and concentration of cation[2,8–10] or on the ionic strength and pH of the buffer.[11,12] Over the past decades, the solution-phase structures of several G4s have been resolved.[13–15] We will use a previously assembled database of resolved G4 structures featuring circular dichroism (CD), mass spectrometry (MS), UV-melting and nuclear magnetic resonance (NMR) data in both NMR and electrospray ionization mass spectrometry (ESI-MS) compatible conditions.[16]

G4-forming sequences are abundant in promoters and at the telomeres, opening avenues to develop G4-targeting ligands as antiproliferative agents or to regulate gene expression.[17–24] G4 ligands are generally flat, condensed heteroaromatic systems that allow efficient π-stacking on G-tetrads while inhibiting dsDNA intercalation.[25] Cationic side chains are often added to improve water solubility and favor electrostatic interactions with the nucleic acid backbone.[26] However, while G4 selectivity over duplexes is often achieved, most ligands lack selectivity among different G4s.[27,28] Ligands discriminating certain topologies[29–31] or topology-subclasses[32–34] have already been reported but rational design efforts have not yet yielded small molecules with significant selectivity for a specific G4 structure.[35] Molecular scaffolds that deviate from the condensed aromatic structure paradigm may be required to produce ligands with improved topology selectivity.[32]

Inspired by biopolymers, foldamers are synthetic oligomers self-organized into a defined folded structure.[36] Here we studied quinoline oligoamide foldamers that fold into a helix driven by intramolecular forces including electrostatic repulsions, intramolecular hydrogen bonds, conjugation and extensive aromatic stacking.[37–39] In the absence of any chirality inducers, helical foldamers exist as a racemic mixture of left- and right-handed helices.[40] Foldamers possess two characteristic traits of G4 ligands: 1) an aromatic heterocyclic core structure, which is derived from quinoline and 2) extended, flexible side chains. But unlike most G4 ligands, foldamers are not flat. Our foldamers have a positively charged side chain (derived from

ornithine) to enhance water solubility (Figure 1). We also introduced pyridine monomers devoid of any side chain to increase the flexibility of the helices.

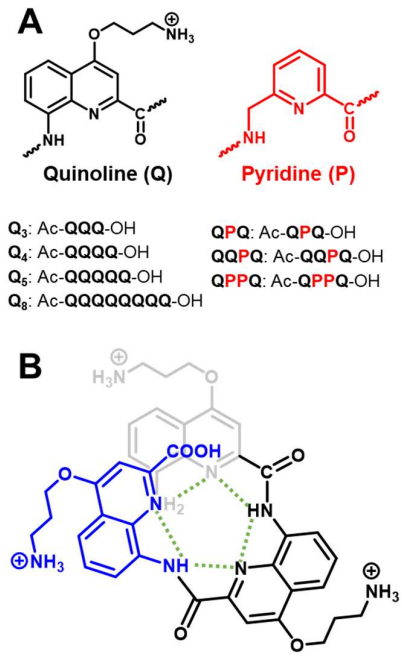

Figure 1 ***A)*** *Oligoquinoline foldamers used in this study, containing quinoline (Q) and pyridine (P) subunits. (panel B shows Q_3)* ***B)*** *Intramolecular H-bonds (green) sterically prohibit a planar structure, forcing the molecule into a (racemic) helical assembly.*

The interaction between oligo-quinoline foldamers and G4s has been studied before. They stabilize G4s, but have no effect on DNA duplexes, according to FRET-melting studies.[37,38,41] Selectivity among different G4 sequences was observed, but not explored further in terms of structure, as the main aim was to demonstrate foldamers targeting G4s over other secondary structures. The takeaway lessons from these previous studies are: 1) Negatively charged foldamers do not interact with G4s, likely due to charge repulsion.[38,42] Hence, aromatic foldamers designed as G4 ligands should carry positive charges, which are implemented through the side chain. 2) The effect of foldamer helicity on G4 binding is uncertain. It was probed by inducing the handedness quantitatively into a left (*M*) or right (*P*) configuration by covalently attaching either *(R)* or *(S)*-camphanic acid, respectively. However, introducing the camphanyl residue disrupts the foldamer-G4 interaction, indicating that the G4-foldamer interaction site is located at the *N*-terminus of the foldamer.[42] Helicenes derived from

diazaoxatriangulenium (DAOTA) were partially separated using diastereoselective precipitation. The difference in K_D value for the *(M)* and *(P)* helices was less than factor 1.5.[43] 3) Foldamers and antiparallel G4s can co-crystallize without any specific ligand-target interaction.[39]

Here, by native mass spectrometry, we serendipitously rediscovered the high binding affinity of foldamer Q_4 with the parallel-forming sequence 1XAV (from the *c-myc* promoter region), and large affinity differences with variants of the telomeric sequence. We therefore conducted a systematic investigation of a larger panel of sequences with several foldamers, and detailed biophysical and structural investigations to understand the determinants of sequence or topology preferences.

Experimental methods

Foldamer synthesis

Seven foldamer sequences (Figures S1—S2) were prepared according to the synthetic schemes in the supporting information (Figures S3—S6). NMR spectra, HPLC traces and accurate masses for each foldamer are provided in Figures S7—S20.

The aromatic oligoamides were synthesized using a microwave-assisted solid phase synthesis approach. The matrix is a low loading Cl-MPA ProTide® resin. The first monomer unit was attached to the resin using CsI/DIEA. Subsequently, each monomer was coupled to the oligomer chain iteratively through periodic cycles of deprotection. Coupling to the next monomer unit was performed using Appel conditions (PPh_3/trichloroacetonitrile) with collidine as base. Finally, trifluoroacetic acid was used to simultaneously cleave the oligomer from the resin and deprotect the Boc groups of the side chains to yield the corresponding ammonium salts. After filtration of the remaining solids, the TFA solution was evaporated then the residue was suspended in Et_2O and centrifuged. The solid was finally dissolved in water and freeze-dried.

All sequences were purified using semi-preparative HPLC on a reverse phase C18 column (mobile phase: H_2O/ACN/TFA). The purified foldamers were freeze-dried twice more to

remove TFA traces. The final product was obtained as a yellow solid with a cotton-like texture, consisting of the foldamer with one trifluoroacetate counterion per ammonium groups. Foldamer stock solutions were prepared by weighing in the purified product on a microbalance (Sartorius ME5) and dissolving them in UPLC/MS grade pure water.

Oligonucleotides and sample preparation

All oligonucleotide sequences were purchased from Eurogentec (Belgium) and dissolved in UPLC/MS grade pure water (Biosolve Chimie, France) to 1 mM DNA. Stock solutions were annealed for 3 minutes at 85°C, then any residual Na^+ ions were exchanged with 500 mM NH_4OAc, which we then flushed out with water using cellulose-matrix 3K centrifugal filter units (MerckMillipore, Ireland). We validated the desalting method by analyzing a DNA stock solution before and after desalting (Figure S21). The concentration of each DNA stock solution was determined with a UV spectrophotometer (Uvikon XL Secomam), utilizing extinction coefficients at 260 nm that were calculated from the nucleotide sequence via nearest-neighbor method.[44] An overview of the sequences studied is provided in
Table 1.
To form G4, DNA solutions were doped with either potassium chloride from Sigma Aldrich and Trimethylammonium acetate solution (TMAA) from ChemCruz or ammonium acetate solution (NH_4OAc) from Sigma Aldrich (Saint-Quentin-Fallavier, France).

Table 1 DNA sequences used for the foldamer screening experiments.

Name	**PDB**	**Sequence**	**Morphology[1]**	**θ (20°C)[2]**
1XAV	1XAV	dT**GA**G**GGT**G**GG**TA**GGGT**G**GG**TAA	Parallel G4	100%
222T		dT**GGG**TT**GGG**TT**GGG**TT**GGG**T	Polymorphic G4	98%
222T_mA		dT**GGG**TT**GGG**AA**GGG**TT**GGG**T	Polymorphic G4	92%
222T_mC		dT**GGG**TT**GGG**CC**GGG**TT**GGG**T	Polymorphic G4	96%
T30177TT	2M4P	dTT**GT**G**GT**G**GGT**G**GGT**G**GG**T	Parallel G4	100%

26CEB	2LPW	dAA**GGG**T**GGG**T**G**TAA**G**T**G**T**GGG**T**GGG**T	Parallel G4	100%
26CEB-mT		dAA**GGG**T**GGG**TTTTTTT**G**T**GGG**T**GGG**T	Parallel G4	100%
2KYP	2KYP	dC**GGG**C**GGG**C**G**CTA**GGG**A**GGG**T	Parallel G4	93%
2O3M	2O3M	dA**GGG**A**GGG**C**G**CT**GGG**A**GG**A**GGG**	Parallel G4	84%
TG4T	2O4F	(dT**GGGG**T)$_4$	Parallel G4	N/A
21G		d**GGG**TTA**GGG**TTA**GGG**TTA**GGG**	Polymorphic G4	100%
5YEY	5YEY	d**GGG**TTA**GGG**TTA**GGG**TTT**GGG**	Antiparallel G4	100%
22GT	2KF8	d**GGG**TTA**GGG**TTA**GGG**TTA**GGG**T	Antiparallel G4	98%
22GT_18T		d**GGG**TTA**GGG**TTA**GGG**TTT**GGG**T	Antiparallel G4	100%
22CTA	2KM3	dA**GGG**CTA**GGG**CTA**GGG**CTA**GGG**	Antiparallel G4	88%
TBA	148D	d**GG**TT**GG**T**G**T**GG**TT**GG**	Antiparallel G4	90%
G4T4G4	4R47	(d**GGGG**TTTT**GGGG**)$_2$	Antiparallel G4	N/A
26TTA	2JPZ	dTTA**GGG**TTA**GGG**TTA**GGG**TTA**GGG**TT	Hybrid G4	66%
Bcl2	2F8U	d**GGG**C**G**C**GGG**A**GG**AATT**GGG**C**GGG**	Hybrid G4	97%
24TTG	2GKU	dTT**GGG**TTA**GGG**TTA**GGG**TTA**GGG**A	Hybrid G4	99%
24TTG_20T		dTT**GGG**TTA**GGG**TTA**GGG**TTT**GGG**A	Hybrid G4	98%
23TAG	2JSM	dTA**GGG**TTA**GGG**TTA**GGG**TTA**GGG**	Hybrid G4	94%
2KPR	2KPR	d**GGG**T**GGGG**AA**GGGG**T**GGG**T	Hybrid G4	100%
21CCC		dCCCTAACCCTAACCCTAACCC	i-motif (pH 5.5) or single strand (pH 7)	100%
ds26		(dCAATCGGATCGAATTCGATCCGATTG)$_2$	Duplex/Hairpin	100%
DK-33		(dCGTAAATTTACG)$_2$	Duplex	100%
DK-66	1FQ2	(dCGCGAATTCGCG)$_2$	Duplex	100%
DK-100		(dCGCGGGCCCGCG)$_2$	Duplex	100%
ss24		dTGCCATGCTACTGAGATGACGCTA	Single strand	
24nonG4		dTGGGATGCGACAGAGAGGACGGGA	Single strand	
T24		dTTTTTTTTTTTTTTTTTTTTTTTT	Single strand	
A24		dAAAAAAAAAAAAAAAAAAAAAAAA	Single strand	
T6		dTTTTTT	Single strand	

[1] Most abundant secondary structure in MS sample conditions

[2] Fraction of folded DNA at 20°C in 1 mM KCl, based on UV melting data

Circular Dichroism spectroscopy

CD measurements were performed on a JASCO J-815 spectrophotometer equipped with a Lauda RE 305 temperature control system with the following parameters: 220-500 nm scan range, 50 nm/min scanning speed, 0.2 nm data pitch, 2 nm bandwidth, 2 s data integration time, 22°C temperature in the sample holder, 3 accumulations. The samples, containing 10 µM of DNA, were placed in Suprasil quartz cuvettes (Hellma) with 2 or 10 mm pathlength.

Raw data was blank subtracted then converted to molar ellipticity with Equation (1), where θ is the ellipticity (mdeg), c the molar concentration of oligonucleotide (mol/L), and l the path length (cm).

$$\Delta\varepsilon\ [M^{-1}cm^{-1}] = \frac{\theta}{3298.2\ c\ l} \qquad (1)$$

We provide CD spectra in MS conditions for all intramolecular G4 sequences listed in Table 1. They have either been published previously,[16] or are featured in the supporting information (Figures S22—S29).

Thermal denaturation

Thermal denaturation was monitored by UV absorption spectroscopy on a SAFAS UVmc2 double-beam spectrophotometer equipped with a high-performance Peltier temperature control unit. The melting ramps were 20→90°C (10 °C/min), 90→4°C (0.2 °C/min), 4→90°C (0.2 °C/min), 90 s data reading interval, 0.5 s averaging time. The samples containing 10 µM of DNA were placed in 10 mm path length Suprasil quartz cuvettes (Hellma). The UV absorption was taken at 260, 295 and 335 nm. The 260 nm wavelength corresponds to the DNA duplex, the 295 nm one to the G4 and the absorption at 335 nm serves as an internal reference to correct for instrumental artifacts. In addition, the buffer alone (100 mM TMAA and 0.5/1 mM KCl) was measured for blank correction.

Raw data, at 260 nm for dsDNA and 295 nm for G4s, was blank subtracted, and corrected with the 335 nm data, then converted to folded fraction θ using Equation (2), where $L0(T)$ and $L1(T)$ are the baseline values of the unfolded and folded species, respectively.[45,46]

$$\theta(T) = \frac{L0(T) - A(T)}{L0(T) - L1(T)} \qquad (2)$$

The melting temperature T_M is the temperature for which θ = 0.5. We provide $\theta_{20°C}$ for all intramolecular G4, i-motifs and duplex forming sequences, in MS conditions in Table 1. Melting curves have either been published previously,[16] or are featured in the supporting information (Figures S22—S27).

Native mass spectrometry: Instrumental setup

Native MS experiments were conducted on an Agilent 6560 IMS-QTOF. The instrumental settings were optimized to preserve non-covalent DNA structures, albeit making a compromise between signal intensity and softness of the ion transfer (Table S1).[47] Samples were infused at 3 µL/min for 11 minutes, working foldamer-by-foldamer to avoid cross-contaminations. The instrument is equipped with a drift tube for ion mobility spectrometry. The ions pass through a tube, filled with helium, and we measure the time between ions being released into the drift tube until arriving at the detector. This time is the arrival time t_A, whose correlation to the collisional cross section (CCS) is described in the Mason-Schamp equation (Equation (3)).

$$CCS = \frac{3}{16}\frac{ze}{N_0 K_0}\sqrt{\frac{2\pi}{\mu k_B T}} \quad (3)$$

Where e is the charge of the electron, k_B is the Boltzmann constant, and N_0 is the gas number density (Loschmidt number) at $T_0 = 275.15$ K and $p_0 = 1$ atm. The temperature T and the reduced mass μ ($\mu = (m_{ion}m_{He})/(m_{ion}+m_{He})$) remain practically constant within the experimental conditions (details in Table S1). The reduced mobility K_0 and charge state z differ for every ion species. We obtained the reduced mobility by measuring the arrival time at several drift voltages (Figures S30—S31) and performing a linear fit according to equation (4).

$$t_A = t_0 + \frac{L^2 p T_0}{K_0 p_0 T}\frac{1}{\Delta V} \quad (4)$$

Where t_0 is the dead time, L is the length of the drift tube (78.1 cm), p and T are the pressure and temperature inside the drift tube (which are recorded for each experiment) and K_0 is the reduced mobility of the ion.

We validated the instrumental settings before each series of measurements using $[(dTG_4T)_4\bullet(NH_4)_3]^{z-}$ as an external calibrant. We accepted the instrumental settings when the experimental CCS values were within 2% of error to our previously published CCS values (787.5 Å² [5- ion], 735.7 Å² [4-] ion).[48] The supporting information contains a step-by-step

tutorial for how to obtain CCS values and reconstruct CCS distributions (Figures S30—S31 and accompanying text).

Ligand screening with native mass spectrometry

The panel (

Table 1) includes 23 G4-forming sequences featuring different strand stoichiometries (1, 2, 4), numbers of tetrads (2 to 4), topologies (parallel, antiparallel, hybrid) and origins (*e.g.*, oncogene promoter, telomeres, aptamers). The panel also includes control sequences of alternative secondary DNA structures (duplex, i-motif, single strand). We screened the DNA panel against 4 different foldamer sequences: QPQ, QQPQ, QPPQ and Q_4. Screening samples contained 10 µM DNA, 20 µM foldamer, 4 µM dT_6, 0.5 mM KCl and 100 mM TMAA (pH 6.8) in water. 150 mM ammonium acetate (pH 6.8) was used for the sequences T_6, T_{24}, ss24, DK-66 and 21CCC. The multi-stranded G4s TG_4T and $G_4T_4G_4$ were screened in both TMAA/KCl and ammonium acetate.

First, each stoichiometry of complex is directly inferred from the mass of the complex. Then, the concentration of DNA species of any stoichiometry (M*) was determined using Equation 5, where $[M]_0$ is the total concentration of DNA in the sample and I is the intensity of the different species, obtained by integrating all signal within a specified *m/z* window (monomeric DNA: M, ligand: L).[49] The mass spectra in the supporting information show the *m/z* selection windows. The concentration of unbound foldamer is determined by difference (Eq. 6), where $[L]_0$ is the total concentration of foldamer in the sample. The dissociation constants K_{D1} and K_{D2} values for the 1:1 and 2:1 (L:M) complexes were then determined from Equations 7–8. The fraction of bound DNA (Eq. 9) is a more concise parameter to present ligand screening results.

$$[M^*] = [M]_0 \frac{I(M^*)}{I(M)+I(ML)+I(ML_2)} \quad (5)$$

$$[L] = [L]_0 - [L]_{bound} = [L]_0 - [ML] - 2[ML_2] \quad (6)$$

$$ML \rightleftharpoons M + L;\ K_{D1} = \frac{[M][L]}{[ML]} \quad (7)$$

$$ML_2 \rightleftharpoons ML + L; K_{D2} = \frac{[ML][L]}{[ML_2]} \quad (8)$$

$$fraction\ of\ bound\ DNA = \frac{[ML]+[ML_2]}{[M]+[ML]+[ML_2]} \quad (9)$$

The difference in K_{D1} and K_{D2} highlights ligand cooperativity. We define a positive cooperativity when $K_{D2} < 4K_{D1}$ and negative cooperativity when $K_{D2} > 4K_{D1}$.[50]

ESI-MS titrations

The six sequences 1XAV/222T/21G/5YEY/T24/ss24 were titrated with seven foldamers: QPQ, QPPQ, QQPQ, QQQ (Q_3), QQQQ (Q_4), QQQQQ (Q_5) and QQQQQQQQ (Q_8). The MS titration experiments follow two main objectives: 1) Obtain average K_D values from a set of measurements alongside with a standard deviation and 2) Study the effect of foldamer length on G4 affinity to aid rational ligand design. We kept the DNA concentration at 10 μM and measured at seven different foldamer concentrations: 0, 5, 10, 15, 20, 30 and 40 μM.

The data processing consists of three steps:

1. Noise correction. Three times the standard deviation of the noise was subtracted from integrated DNA signals. The ligand screening data was noise corrected as well.

2. Response analysis. Response factors depend on the gas-phase transmission yield during electrospray ionization, which can be different for the M, ML and ML_2 species. We quantify changes in response by adding dT_6 as an internal calibrant.[51] The ratio of DNA to dT_6 intensity as a function of foldamer concentration was monitored and, from changes in that ratio, we derived the effect of foldamer concentration on DNA response (Figure S32). Note that the foldamers do not interact with dT_6.

3. Dynamic fitting with the DynaFit software (BioKin, MA, USA) which derives a mathematical model from the chemical equations 7 and 8 and matches it to the experimental data with minimal residual error (Figure S33). Response factors of the complexes (ML, ML_2) were allowed to vary to reflect effects observed at step 2. K_{D1} and K_{D2} values were extracted from the fitted data.

Determination of the number of consecutive G-quartets in ligand-free and ligand-bound forms

K^+ ions bind to G4 forming sequences in two ways: in-between G-quartets (specific) or elsewhere (unspecific). Unspecific K^+ ions distribution follows a discrete probability, which can be approximated as a Poisson distribution.[52] Specific K^+ ions in G4s typically have much

higher binding affinities than unspecific ones and have well-defined stoichiometries. An adduct with *n* K^+ ions is considered specific when it is significantly more populated than its adjacent *n+1* and *n-1* adducts.

For example, in Figure 2, the main adduct of the free DNA is 1 K^+, which corresponds not only to a 2-tetrad G4 (antiparallel),[2] but also some unfolded DNA stand with non-specific potassium binding. The high relative intensity of the unfolded peak (0 K^+) is an indicator for an unspecific K^+ adduct distribution blended within all adduct peaks. The main adduct of the complex is 2 K^+, which corresponds to a 3-tetrad parallel G4. The absence of 0 and 1 K^+ adducts indicates ligand selectivity towards 3-tetrad parallel G4s over 2-tetrad antiparallel G4s and unfolded DNA.

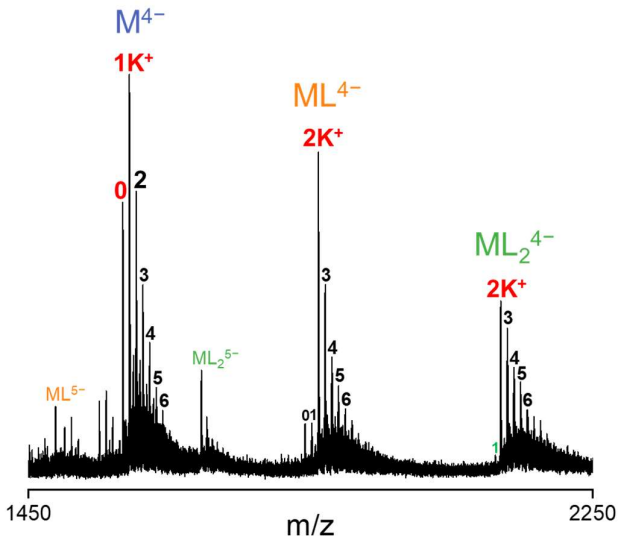

*Figure 2 Mass spectrum of 10 µM 222T_mA (dT**G**$_3$TT**G**$_3$AA**G**$_3$TT**G**$_3$T) with 20 µM Q_4 in 0.5 mM KCl and 100 mM TMAA. Labeled species are: Free DNA (M), 1:1 complex (ML) and 2:1 complex (ML_2), specific K^+ adducts (red), non-specific K^+ adducts (black).*

X-ray crystallography and structure determination

Crystallization was performed in a sitting-drop well plate (SWISSCI MRC 2 Lens Crystallisation Plate). Droplets were obtained by mixing 200 nL of a solution containing 400 µM DNA (222T), 900 µM foldamer (QQPQ) and 20 mM potassium cacodylate (pH 7.0) with an equivalent volume of crystallization screening solutions from various commercial kits. The original condition was the No. 41 of the (Hampton Research) and was composed of 100 mM KCl, 15 mM $MgCl_2$, 50 mM Tris (pH 7.5) and 10% PEG-550. We obtained clear, cuboid crystals that were 10-20 µm in size. The crystals were transferred into a cryoprotective

solution containing 100 mM KCl, 15 mM $MgCl_2$, 50 mM Tris (pH 7.5), 20% PEG-550 and 20% ethylene glycol. Finally, the crystals were flash frozen at -196°C and analyzed at the SOLEIL synchrotron (Saint-Aubin, France).

Several data sets were collected at the French synchrotron Soleil on the beamline Proxima I to a maximum resolution of 2.51 Å. The raw images were processed with XDS.[53] The structure was solved using the Phaser-Molecular Replacement procedure from the software package Phenix.[54] The structure was solved using the G4 core from the previously solved structure of the 222T sequence (PDB: 6P45).[55] We tested various models with and without the connecting nucleotide to help remove model bias. We improved the G4 model by several rounds of manual rebuilding and added the QQPQ foldamer once the electron density from the G4 was filled. QQPQ foldamer restrains were generated using the webserver from GlobalPhasing[56] and added to the refinement procedure in Phenix. The final model included two potassium ions located between the G-tetrads of the same G4 and one potassium ion located between the G-tetrads of two dimerizing G4s coordinated by the guanine bases, several magnesium ions and 18 nucleotides of the G4 sequence (missing only the first and the last thymine nucleotide). The summary of data collection and refinement statistics is presented in Table S5. Figures were prepared using ChimeraX.[57]

NMR spectroscopy

NMR spectra were recorded on a 700 MHzBruker Avence Neo spectrometer equipped with a triple resonance gradient standard probe, or a 800 MHz Bruker Avance Neo spectrometer equipped with a triple resonance cryoprobe. Topspin version 4.1 (Bruker BioSpin) was used for data collection. Processing and analysis for 1D spectra also used Topspin, whereas the 2D spectra were processed using NMRPipe[58] followed by analysis with Sparky (T.D. Goddard and D.G. Kneller, University of California). The parallel and monomorphic G4 sequence T95-2T ($dT_2G_3TG_3TG_3TG_3T$, PDB: 2LK7) was used as a substitute for 222T which displayed poor quality NMR spectra.[59] Identification of key foldamer-binding thymine bases was aided by using three mutated sequences, each of which replaced one thymine with (deoxy)uracil: 2LK7_1U (d**U**$TG_3TG_3TG_3TG_3T$), 2LK7_2U (dT**U**$G_3TG_3TG_3TG_3T$) and 2LK7_18U ($dT_2G_3TG_3TG_3TG_3$**U**). The 2LK7 and 5YEY NMR samples contained 100 µM DNA in 10 mM potassium phosphate buffer (pH 7, contains 16 mM K^+) in 90%/10% H_2O/D_2O without or with

added QQPQ. The 2D $^{1}H,^{1}H$-NOESY used a mixing time of 200 ms, and the $^{1}H,^{1}H$-TOCSY used a mixing time of 60 ms, with both measured at 278 K.

Every DNA sequence was measured with and without QQPQ. 1D-NMR spectra of QQPQ alone were measured in d6-DMSO instead of H_2O/D_2O and potassium phosphate.

Results and Discussion

Foldamer specificity for different nucleic acid secondary structures

To have a first ranking of foldamer binding affinities, we recorded native ESI-MS spectra at one DNA:foldamer concentration ratio. For example, in Figure 3 we can visually compare the mass spectra of foldamers QPPQ and QQPQ with vastly different binding affinities for the same DNA strand 5YEY. The relative abundance of complex (ML) compared to free DNA (M) readily shows that QQPQ binds to the sequence 5YEY with much higher affinity than QPPQ. The calculated K_D values differ by more than one order of magnitude.

The low intensity of 2:1 complex compared to the 1:1 complex indicates that the 5YEY G4 only has one accessible binding site. In contrast, foldamer Q_4 binds to the parallel G4 222T at two binding sites (see Figure 2).

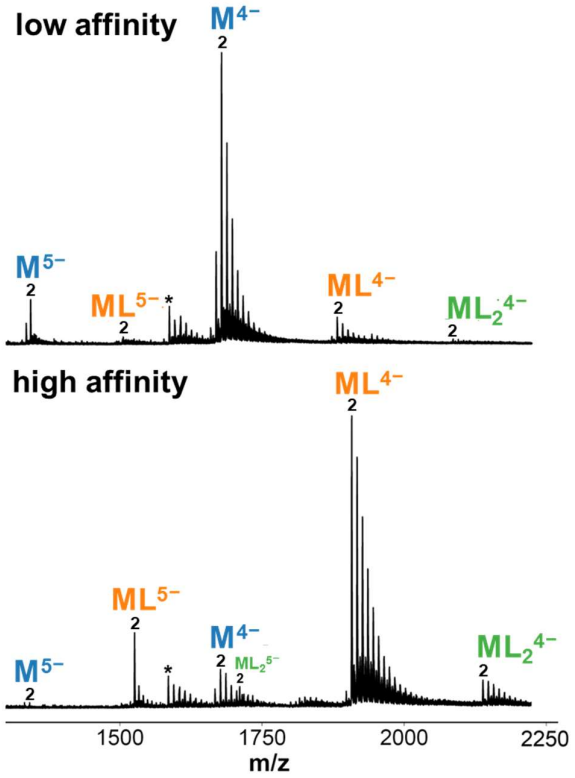

*Figure 3 Mass spectra of 10 µM 5YEY (d**G**$_3$TTA**G**$_3$TTA**G**$_3$TTT**G**$_3$) with **Top**: 20 µM QPPQ. (K_{D1} = 75 µM, K_{D2} = 40 µM) **Bottom**: 20 µM QQPQ. (K_{D1} = 1.5 µM, K_{D2} = 22 µM). Sample matrix contains: 0.5 mM KCl and 100 mM TMAA (pH 6.8). Labeled species are: Unbound DNA (M), 1:1 complex (ML), 2:1 complex (ML_2), 2K^+ adduct (2), depurinated DNA (*).*

We carried out this general screening with four foldamers (Q_4, QQPQ, QPPQ and QPQ), and 34 DNA sequences that include well-characterized G4 structures and other DNA secondary structures as controls (i-motif, duplexes, and single-stranded DNA). Figure 4A presents the ligand affinity results as a heatmap. The detailed data, including K_D values, are in **Error! Reference source not found.** and the mass spectra (showing the foldamer and K^+ stoichiometries) are shown in Figures S34—S69. Overall, the data indicate that foldamers have a strong preference for parallel G4s and bind very little to DNA controls. Among different foldamers, QQPQ has the most varied affinities and selectivities.

All four studied foldamers bind parallel G4s with K_D values in the low μM range (Table S2). But not all parallel G4s are equivalent. First, 222T_mA and 2KYP are partially unfolded in 0.5 mM K^+. Therefore, there is a mixture of targetable structures (folded G4) and non-targetable structures (unfolded species). This explains that the fraction of bound DNA is lower compared to the other (fully folded) sequences. Next, the G-tetrads of 2O3M and TG4T are less exposed than those of common parallel G4s with three propeller loops. 2O3M has a lateral and a snapback loop can obstruct the two terminal G-tetrads. TG4T has eight flanking thymines that can form a sparsely populated T4 tetramer, which also protects the G4 from 5'-5'-dimerization.[60] We therefore hypothesize that foldamers bind to the external G-tetrads, but that loops/bases covering the terminal tetrads will impede foldamer binding for steric reasons.

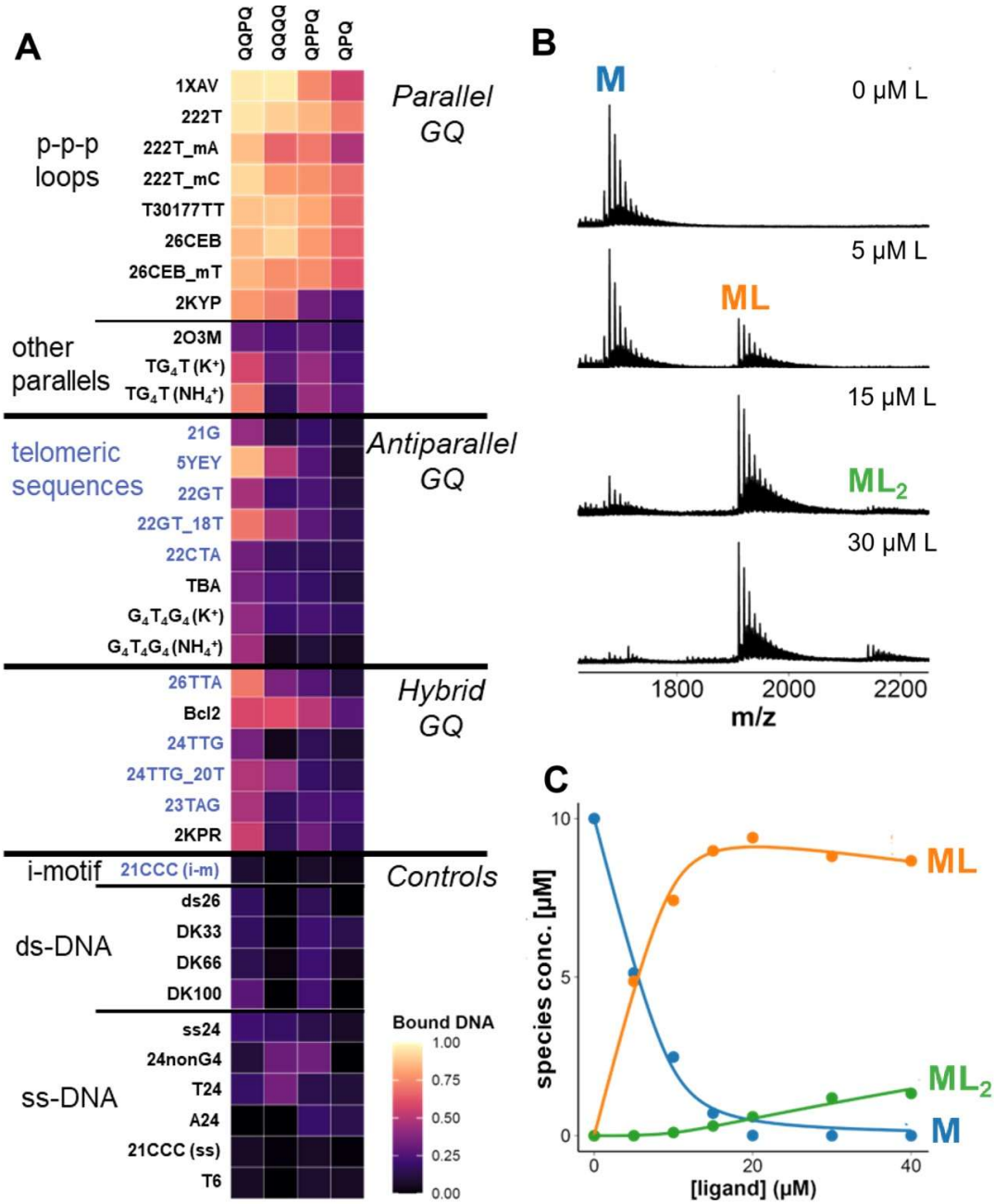

Figure 4 ***A)*** *Heatmap showing the fraction of DNA bound for all DNA/foldamer combinations, calculated from mass spectra. Sample conditions are: 10 μM DNA, 20 μM foldamer, 0.5 mM KCl, 100 mM TMAA (pH 6.8). DNA sequences are listed in*

Table 1. Telomeric-related sequences are highlighted in blue. ***B)*** *Titration data for 5YEY and QQPQ, showing an extract of the raw mass spectra.* ***C)*** *Processed data. The dynamic fitting curves match the experimental datapoints and return* $K_{D1} = 0.54 \pm 0.1$ *μM,* $K_{D2} = 200 \pm 70$ *μM.*

QPQ and QPPQ are highly selective to parallel G4s with three propeller loops but have moderate affinities. These foldamers are more likely crescent-shaped than helical, because the quinoline-oligomer chain is not long enough (it takes 5 Q units for 2 helical turns).[36] QQPQ and Q_4 are more stable helices; as indicated by the network of intramolecular H-bonds for QQPQ by NOESY (see NMR section). Their affinity to parallel G4s is 3 to 4 times higher.

Apart from parallel G4s, the QQPQ foldamer tightly binds to 5YEY (K_{D1} = 1.5 μM), but also to TG4T, 22GT_18T and 26TTA ($K_{D1} \approx 5$ μM). 5YEY and 22GT_18T are telomeric mutants with a TTA→TTT mutation in the third loop, which induces a switch from a 2-tetrad to a 3-tetrad antiparallel G4.[61] QQPQ and Q_4 are sensitive to this structural switch: The affinities to the mutants is 5-10 times higher than to the 'wild-type' sequences (compare: 21G/5YEY, 22GT/22GT_18T, 24TTG/24TTG_20T). 26TTA is a hybrid-2 G4 that is targeted by QQPQ but no other foldamer in this panel. We cannot assess whether QQPQ is selective towards hybrid-2 over hybrid-1 topology.

None of the foldamers interact well with most antiparallel G4s, where the G-tetrads are usually covered by base triads and/or diagonal loops on both sides (21G, 22GT, 22CTA, TBA, G4T4G4). This confirms that foldamers selectively target G4s whose configuration allows exposed G-tetrads. The low affinity towards alternative DNA structures (i-motif, duplex, single strand) underlines the G4 specificity of foldamers, except for very specific cases described below. The three-dimensional shape of foldamers prevents duplex intercalation, as they are too bulky to lodge in between base pairs.[43] Among the control DNA sequences, the foldamers were irresponsive to changes in G-richness, molecularity or strand length.

Characterization of affinities with ESI-MS titrations

To further characterize binding affinities and stoichiometries, we selected six DNA sequences (two single strands, two parallel G4s, and two derivatives of the telomeric sequences that differ by a single base), and seven foldamers: QPQ, QQPQ, QPPQ, Q_3, Q_4, Q_5 and Q_8. Figure 4B shows titration data for 5YEY and QQPQ. All K_D values are listed in Table S3, and estimated response factors of each DNA:Foldamer complex are in Table S4. The full titration datasets are in Figures S70—S110.

The control sequence ss24 is a single-stranded 24-mer with 25% of each nucleobase. None of the foldamers target this sequence (K_{D1} > 100 μM) save for Q_3, which has significant

populations of 1:1 complex, 2:1 complex and 3:1 complex (Figure S98). Q_3 is thus the least G4-specific ligand in the panel. T24 is another single-stranded control sequence that does not interact with QPQ, QQPQ or QPPQ ($K_D > 100$ μM), but it attracted our attention during the screening with Q_4. Intriguingly, it forms high-affinity 2:1 complex with Q_3, Q_4 and Q_5 ($K_{D2} < 1$ μM), including a 4:1 complex with Q_3 ($K_{D4} = 0.01 \pm 4$ μM). Although our foldamers are racemic and display no CD signal on their own, the induced CD in the presence of T24 matches the CD signature of a right-handed foldamer helix (Figure S111). Our hypothesis is that T24 and Q_n foldamers associate into a right-handed double helix, wherein the foldamer is dimerized in order to accommodate the full length of the T24 strand. We also tried the RNA counterpart U24 (Figure S112), because U-rich motifs play a role in RNA expression,[62,63] but there was no induced CD in U24 upon adding Q_n foldamers.

The parallel G4s 1XAV and 222T show the highest affinity. All foldamers form high-affinity 1:1 complex with 222T ($K_{D1} < 10$ μM). The strongest binding partner for 1XAV is Q_4 ($K_{D1} = 75 \pm 59$ nM), the weakest one is Q_8 ($K_{D1} = 28 \pm 6.7$ μM). All foldamers form 2:1 complex ($K_{D2} \approx 1$ μM), except for Q_8 and QPQ. Q_3 is the only foldamer to form a 3:1 complex, but based on our results with a control sequence ss24, the interaction might not be G4 specific.

Finally, the two telomeric sequences, 21G and 5YEY, are identical, save one nucleotide in the third loop (A in 21G and T in 5YEY). This mutation, however, changes the polymorphic 21G to the antiparallel 5YEY. Q_5, Q_4 and QQPQ form a stable 1:1 complex with 5YEY ($K_{D1} < 10$ μM), but not 21G. Q_4 is the most selective foldamer, with a 20-fold change in K_{D1} between 21G and 5YEY. We also compared the sequences 22GT (antiparallel with 2 quartets) and 24TTG (very stable hybrid-1 G4) and the corresponding sequences with the same A→T mutation (22GT_18T and 24TTG_20T), and observed the same increase of affinity for the sequence variant containing the thymine (Figure 4A).

Overall, the ESI-MS titrations indicate that foldamers have the highest affinity (in the low/sub-μM range) towards parallel G4s (1XAV and 222T), alongside a few notable exceptions (5YEY, T24). Foldamer length is important and there seems to be a 'sweet spot' at about 4-5 monomers. Having only 3 monomers comes with a decline in affinity (QPQ) or G4 specificity (Q_3). Having over 5 monomers leads to no affinity increase, but loss of DNA signal that scales with increasing number of Q monomers (Figure S32). The reason is that each Q monomer adds a positive charge to the foldamer, which reduces charge repulsion among DNA polyanions, promoting DNA aggregation. We confirm this hypothesis from 1) seeing precipitation in concentrated DNA

solution when an excess of foldamer is added 2) loss of CD signature for a stable (i.e. resistant to disruption) parallel G4 when Q_8 is added (Figure S113).

Preferred topologies and foldamer-induced topology switching

For our structural investigation we selected QQPQ because (1) it was the foldamer with the highest G4 affinity and (2) we were intrigued by its ability to selectively target some telomeric repeat sequences, among which we picked 5YEY and 26TTA. We picked T30177TT as a stable parallel-stranded G4. 222T and 222T_mA are polymorphic in 0.5 mM K^+ with a mix of 3-tetrad parallel and 2-tetrad antiparallel G4.[2] The equilibrium of 222T in our experimental conditions is shifted towards the parallel form, while 222T_mA is shifted towards the antiparallel form. We characterized all complexes by mass spectrometry (MS), ion mobility spectrometry (IMS), CD, and thermal stability, and present selected results in Figure 5.

MS provides a quick readout on ligand stoichiometry, potassium binding stoichiometry and the quantity of each species. The number of specific K^+ ions correlates to the number of G-tetrads. IMS lets us track changes in structure even when there is no change in mass. The ions' collisional cross section (CCS) depends on the ion size, shape and charge state (within constant experimental settings). Adding ligands is increasing the size of the complex and is expected to lead to a CCS increment. In addition, significant conformational differences for a given size are expected to lead to significant differences in ion mobility, resulting in different CCS distributions. Thus, the shape of the CCS distribution for each ligand stoichiometry indicates how many topologies could be present in solution. CD signatures of DNA alone and with foldamer informs on how the foldamer affects the G4 topology. Finally, the difference in melting temperature with and without foldamer (ΔT_m) provides a quick snapshot of how potent a ligand is at stabilizing the G4 structure. When there are sufficiently populated species with significantly different melting temperatures, multiple melting transitions are visible.

For T30177TT, 2 K^+ is the dominant adduct in the mass spectra. It has a parallel CD signature and is 100% folded at 25°C. We are thus dealing with a stable, 3-tetrad parallel G4. Adding foldamer does not change the K^+ adduct distribution, the shape of the CCS distribution or the CD signature, meaning that foldamer binding does not significantly change the G4 structure. The slight drop in CD intensity is likely due to foldamer-induced DNA aggregation (Figure S113). The melting temperature increases by 17 °C, but two distinct melting transitions appear.

We know from previous temperature-controlled MS studies on G4 ligands that the 1:1 complex usually melts before the 2:1 complex.[64] This situation could apply here as well.

222T and 222T_mA are polymorphic in 0.5 mM K^+. The CCS distributions of free DNA are therefore the sum of multiple Gaussian functions, each representing a different set of conformers. Upon foldamer binding, the CCS distributions are less convoluted (i.e., they look more like a single Gaussian distribution), highlighting that the foldamer selectively targets certain conformers within the conformational ensemble of polymorphic sequences. Figure 5 shows that the complexes (ML/ML_2) have 2 specific K^+ ions, therefore the complex must be a 3-tetrad G4. Adding foldamer shifts the CD spectra towards a parallel signature (222T_mA has a remarkable signature switch. The foldamer targets parallel G4 and stabilizes it ($\Delta T_m \approx +25$ °C). Similar to T30177TT, the DNA/foldamer system has two separated, broadened melting transitions which likely correspond to the 1:1 and 2:1 complex.

5YEY has only one accessible binding site, as the intensity of 1:1 complex is significantly higher than that of 2:1 complex in the mass spectrum (see also the titration in Figure S87). The foldamer affects G4 folding, causing a CD band at 260 nm to increase. The CD signature indicates that the 1:1 complex could be a hybrid G4, but the CCS distributions give no conclusive evidence whether there is only one hybrid species or multiple species with different topologies. The affinity of the 1:1 complex between 5YEY and QQPQ ($K_{D1} = 0.54 \pm 0.1$ μM) is comparable to that observed for parallel G4s. The melting curve is steeper and shifts less ($\Delta T_m = +8.2$ °C) compared to those for parallel G4s. That implies that the complex formation with 5YEY is more enthalpy-driven.[46]

26TTA mirrors the behavior of 222T/222T_mA: The K^+ adduct distribution shifts towards $2K^+$, a parallel CD signature emerges and the UV melting transition broadens in the presence of QQPQ. The foldamer promotes a structural rearrangement of 26TTA, probably from hybrid to mostly parallel topology. As a result, the complex has two binding sites.

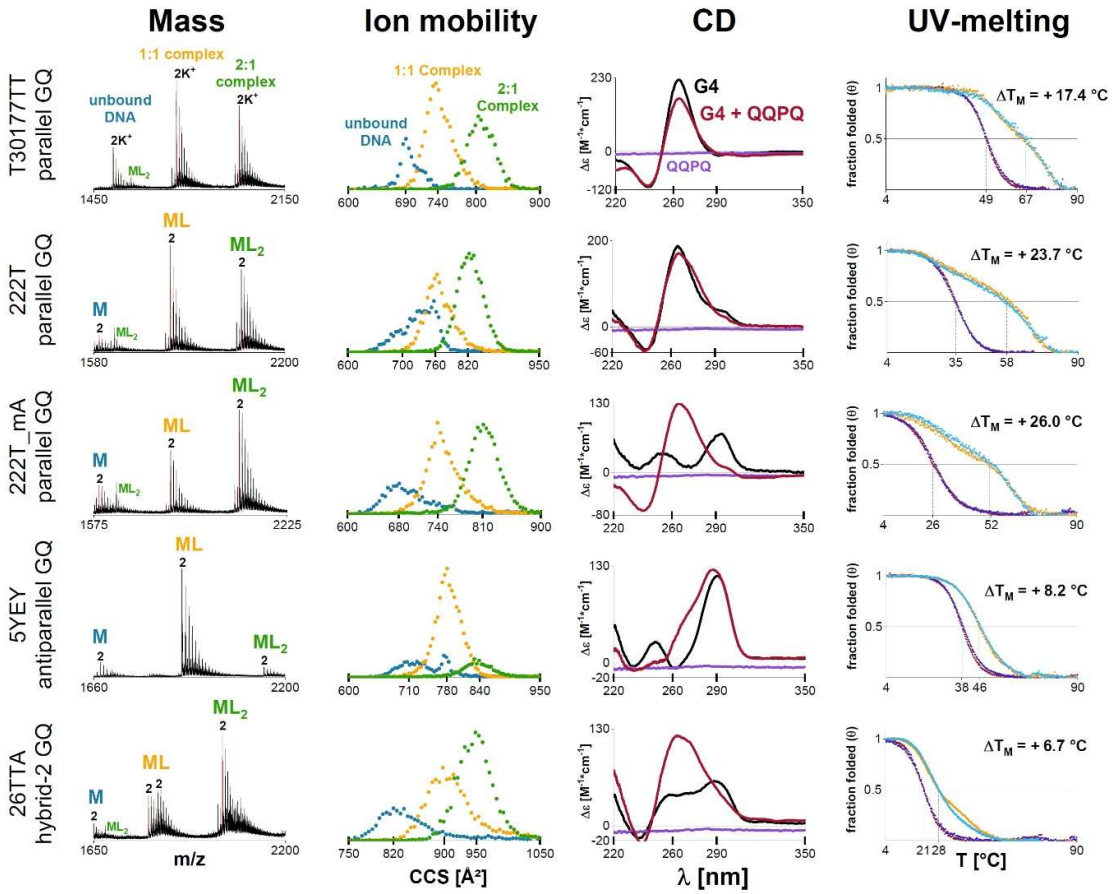

Figure 5 The effect of foldamer QQPQ on five different G4s. Samples contain 10 µM DNA, 20 µM QQPQ, 0.5 mM KCl, 100 mM TMAA (pH 6.8). ***1st column****: Mass spectra showing the 4-charge state (5- for 26TTA).****2nd column****: CCS distributions extracted from the 2K⁺ adduct of each species, highlighted with a red line in the corresponding mass spectrum (in column 1).* ***3rd column****: CD spectra of the complex (red) and control spectra with free DNA (black) and free foldamer (purple).* ***4th column****: UV-melting curves with foldamer (heating: orange; cooling: cyan) and without foldamer (heating: crimson, cooling: navy).*

X-ray crystallography

We used X-ray crystallography to test the end-stacking hypothesis based on the CD and native MS results. Molecular crowding induces a bias for the parallel topology,[65,66] so we selected the parallel G4 sequences 1XAV and 222T to maintain the solution-phase topology and promote crystallization with QQPQ or Q_4. We also studied the 5YEY/QQPQ and T24/Q_4 complexes due to their specific binding properties that could provide additional mechanistic insights. We only obtained crystals for the 222T/QQPQ complex, whose structure is shown in Figure 6. Atomic and diffraction data have been deposited in the PDB under the accession code: 8QN2. Crystal and structure refinement information is deposited in Table S5. Electron density maps are shown in Figures S114—S116.

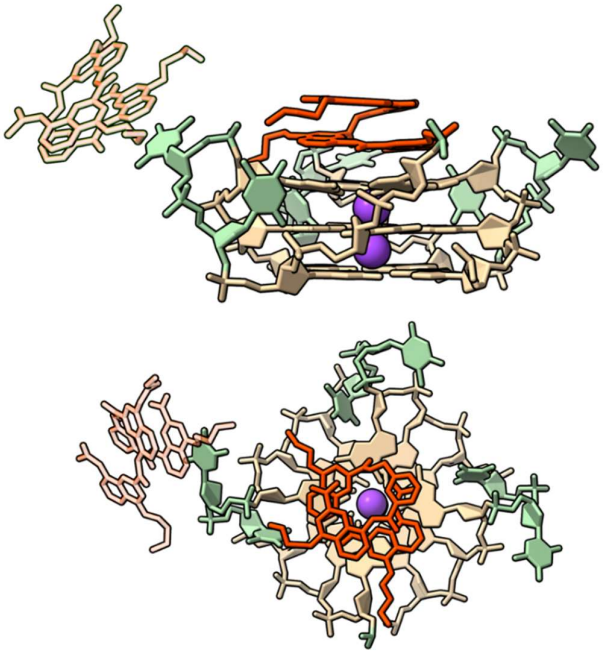

*Figure 6 X-ray crystal structure of G4 222T (dT**G**$_3$TT**G**$_3$TT**G**$_3$TT**G**$_3$T) and foldamer QQPQ (PDB entry code: 8QN2) from the side (**top**) and from above (**bottom**). Color coding: two QQPQ molecules in red (the binding site not existing in solution phase is set to 50% transparency), guanine in tan, thymine in green, potassium in purple. The crystallization matrix contains 100 mM KCl, 15 mM $MgCl_2$, 50 mM Tris (pH 7.5), 10% PEG-550 and a 9:4 ratio of foldamer:DNA.*

222T adopts a parallel-stranded topology and forms a 5'-5' crystallographic dimer. Two QQPQ foldamers interact with one 222T molecule. The first QQPQ has two quinolines π-stacking onto the 3' G-tetrad. The QQ unit is flattened at the binding site to allow more contact area for π-stacking. We suspect that the P unit adds flexibility through its methylene group, which can freely rotate. The flexibility is needed to reduce molecular tension in the foldamer helix caused by the QQ pair flattening. Such molecular architecture explains why QQPQ has the highest G4 affinity: It has two Qs for π-stacking, followed by a P which acts as a step, after which the molecule can resume its helical shape. The top view illustrates how the diameter of the foldamer helix matches with the size of a G-tetrad. The propeller loops are not limiting the foldamer's access to the G-tetrads and therefore do not interfere with foldamer binding.

According to our ESI-MS titrations, the 222T/QQPQ complex is saturated in a 3-fold excess of foldamer, where it forms a 2:1 complex (Figure S80). We expected to have one ligand on each end of the G4. However, the 5' interface is utilized for G4 dimer formation and the second foldamer instead stacks onto a thymine base of a TT loop (Figure 6). Parallel G4s tend to dimerize in concentrated solution because of their exposed G-tetrads. The preferred interface for dimerization is 5' to 5' end.[67,68] The foldamer sitting atop the thymine establishes a contact with a neighboring unit cell (Figure S117). Such crystal contacts are needed to promote crystal formation and growth, but they can only persist in the solid phase.

The electron density map of the G4 is well-resolved (Figure S114), but the electron density of the foldamer is less defined (high B-factors, see Figure S116), suggesting high dynamics. Several factors can contribute to these dynamics: (1) rotational freedom along the helical axis, (2) exchanges between left- and right-handed helix, since the foldamer is racemic, (3) the foldamer latching onto the G-tetrad from its C-terminus or N-terminus, and (4) conformational freedom of the foldamer side chains. We tested models where either the C- or N-terminus of QQPQ interacts with the 3' G-tetrad (Figure S118) and found no significant change in refinement statistics that would suggest any C/N-terminal arrangement being more favorable than the others. So even though the crystal structure depicts a static image of the foldamer-DNA-complex, dynamic exchanges are expected in solution. Alternatively, the foldamer binds to different G4s with different termini and the observed structure is a statistical summary of all arrangements.

In summary, we have validated one binding site being the 3' G-tetrad and the binding mode being π-stacking but need supporting evidence for the second binding site and the foldamer orientation.

NMR spectroscopy

NMR allows us to investigate foldamer binding to G4 DNA in solution and thus validate (or not) our X-ray structural studies. We could not obtain well-resolved NMR spectra with 222T, but instead used 2LK7 (dTT**G**$_3$T**G**$_3$T**G**$_3$T**G**$_3$T), whose structure is very similar to 222T but is more dynamically constrained (1 nt loops) and protected from dimerization (TT motif at 5').[59] QQPQ targets 2LK7 with sub-μM affinity (Figure S119, Table S3). All 1D and 2D NMR spectra of 2LK7/QQPQ are gathered in the supporting information (Figures S120—S134).

Our key findings are summarized in Figure 7. Figure 7A shows a ^{1}H-^{1}H NOESY spectrum of the 2LK7/QQPQ complex. We observe two networks of NOE contacts that correspond to the amide protons of the foldamer (designated *NH1* to *NH4* in Figure 7D). The amide protons are in close contact through the H-bond network that enables the helical foldamer arrangement. The detection of (exchangeable) amide protons in the 1D spectrum (Figure 7C) further indicates that these protons are protected from solvent exchange. These factors support the helical folding of QQPQ in solution phase.

Figure 7A shows two distinct NOE networks of the same amide protons (lime green and blue). The two significantly different chemical environments for the foldamer must be the first and second binding site, in agreement with the 2:1 complex observed in the mass spectra and our crystal structure. While we do see shifts for the 3' G-tetrad, we do not observe any perturbation of the thymine loop region, indicating that this interaction with loop thymines likely does not take place in solution. A more suitable candidate for the second binding site is the 5' G-tetrad, which is inaccessible in the solid phase because of G4 dimerization.

We then labeled the 5' and 3' end of 2LK7 by switching the flanking dT nucleotides with dU (which lacks the characteristic methyl group of thymine). These chemical changes, albeit small, cause visible chemical shifts near each respective thymine base. Figure 7B shows how we assigned a binding site to each NOE network by illustrating the NOE between the amide group (*NH1*) and the methyl group (*HAc*) at the N-terminus of the foldamer. Mutating the 3' end to uracil (U18) causes a chemical shift in the blue NOE network, while the green NOE network remains unaffected. Likewise, mutating the 5' end (U2) causes a chemical shift in the green

NOE network, while the blue NOE network remains unaffected. Therefore, one foldamer must be close to the 5' end and another one close to the 3' end. While the foldamer that interacts with the 3' G-tetrad was clearly seen in the crystal structure, we can now safely deduce that in solution the second foldamer is π-stacking onto the 5' G-tetrad.

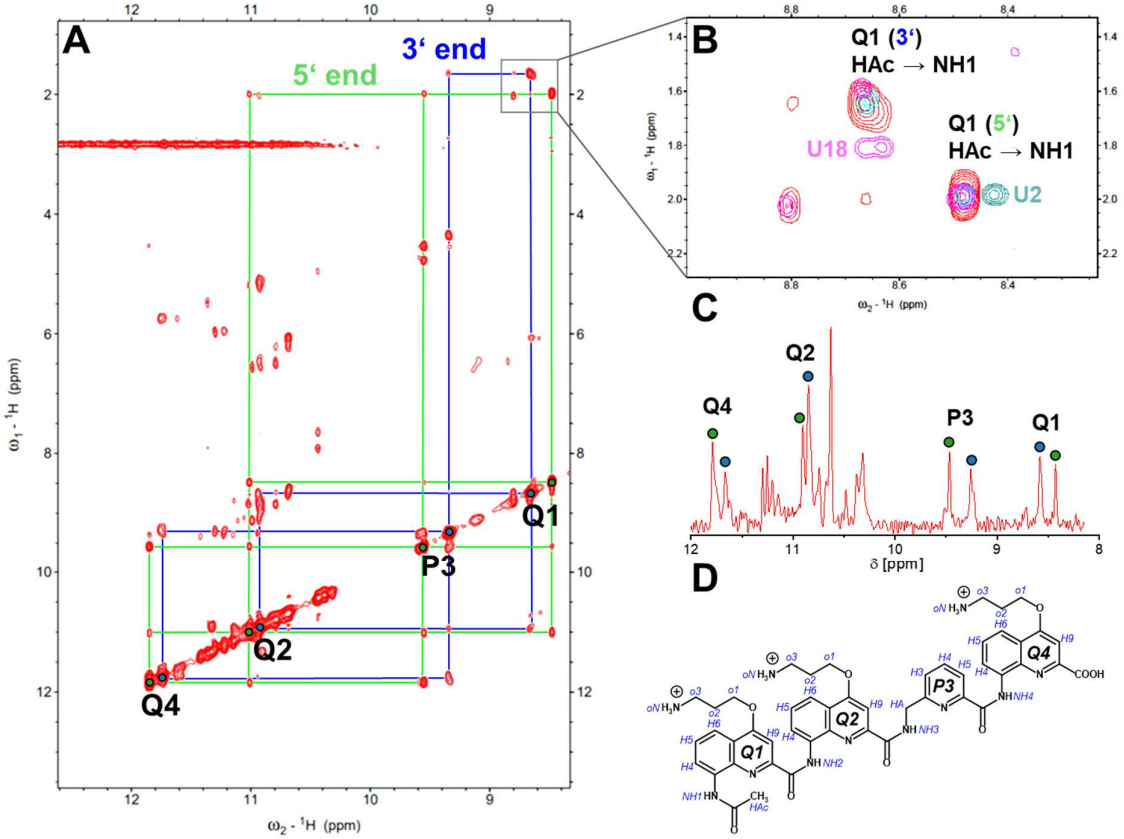

*Figure 7 A) ^{1}H-NOESY of 2LK7 (dTT**G**$_3$T**G**$_3$T**G**$_3$T**G**$_3$T) in excess of QQPQ, showing two NOE networks connecting the foldamer amide protons (NH), indicating that the foldamer is located in two distinct chemical environments. B) Marking the flanking nucleotides through dT→dU mutations induces chemical shifts in the foldamer on the 5' end (pos. 1,2) and 3' end (pos. 18) of the G4. C) 1D-NMR of 2LK7 in 3-fold excess of QQPQ, showing the guanine H1-proton region. Foldamer amide protons are highlighted. D) Scheme of QQPQ with proton labels shown in blue. Samples contain 100 µM 2LK7, 300 µM QQPQ, 10 mM KP buffer (pH 7.0) in 90/10 H_2O/D_2O.*

Specificity for variants of the telomeric sequence

In the above characterization of the foldamer-bound G4s based on the telomeric sequence, we observed intriguing specificities for 5YEY,for sequences wherein the third loop is TTT, and also for 26TTA. This latter G4 is not very stable in MS-compatible K^+ solutions, and therefore the quadruplex can switch its conformation towards a parallel-stranded topology. For 5YEY, our data indicate that QQPQ induces significant base stacking changes in 5YEY (Figure 5). We were thus curious about the structure formed by the 5YEY-QQPQ complex. Unfortunately, we could not crystalize the complex, but we were able to probe the solution states of the free and bound 5YEY by NMR spectroscopy (Figure 8A).

NOESY data indicate that only one foldamer NOE network is present (Figure 8B), which reaffirms the 1:1 complex seen in our ESI-MS results (Figure 5). The (lime-colored) NOE network and NMR shifts match well with the lime-colored network in Figure 7, where the foldamer was bound at the 5' end. Although this observation does not prove that in 5YEY the binding site is also located on the 5' end, it is consistent with this interpretation.

The 5YEY structure contains an A:T base pair formed between A6 in the first loop and the mutated T18 in the third loop.[69] A clear indicator of the A:T base pair in the NMR spectrum is the observation of the H-bonded H3 amide proton of T18 within the chemical shift range typical of Watson-Crick base-paired thymine (Figure 8A, top spectrum). In the presence of QQPQ, this characteristic peak disappears from this characteristic region (Figure 8A, bottom spectrum), indicating that the 5YEY/QQPQ complex no longer contains the A:T base pair. Our MS data show 2 specific K^+ ions for both free 5YEY and the 5YEY/QQPQ complex, therefore both free and bound 5YEY have three G-tetrads, but both the circular dichroism spectrum and the ion mobility data suggest that a topology switch could have occurred. Finally, we also recall that the only non-G4 forming sequence to which foldamers significantly bind is the polythymine dT_{24}.

We can imagine two scenarios of QQPQ binding to 5YEY. The first scenario is that the G4 topology of 5YEY remains unchanged. The A:T base pair dissociates to make space for the foldamer helix, allowing it to π-stack onto the 5'-opposing G-tetrad. However, foldamer binding by stacking on the quartet on the same side of two lateral loops is difficult to reconcile with the observation that, in parallel G4s, steric hindrance on the terminal quartets diminishes the foldamer affinity. We thus favor a second scenario in which the G4 undergoes structural rearrangement. The CD results indicate G-quartet stacking typical of hybrid structures. Hybrid

structures typically have a lateral loop on each side of the G-stack core, but recalling the non-negligible affinity of foldamers for polythymines, we suggest that the G-quartet on the side of thymine loops (and T or TT overhangs in 22GT_18T and 24TTG_20T) could be the foldamer binding site.

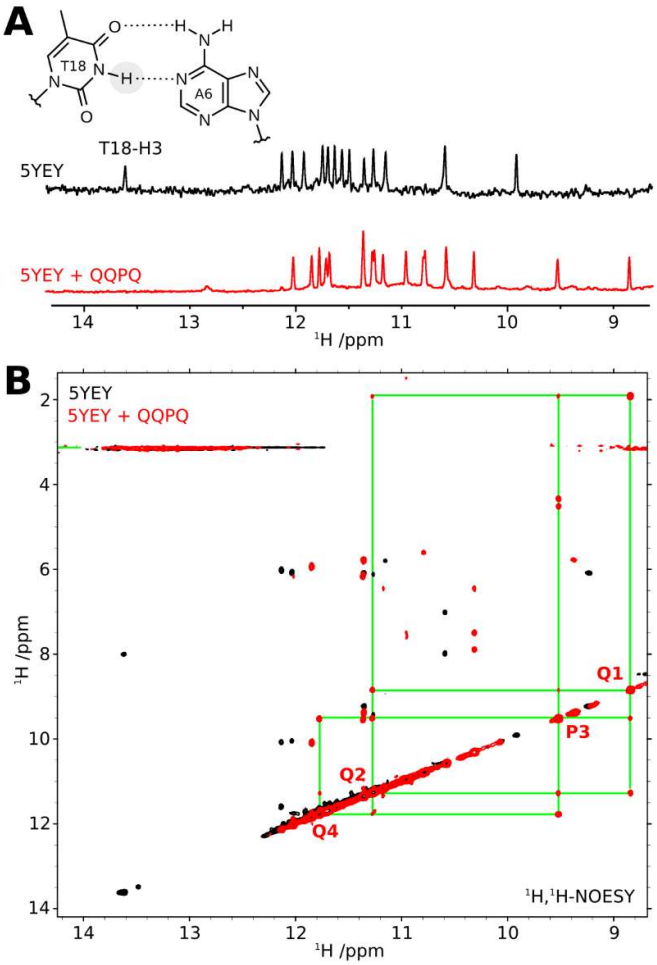

*Figure 8 A) In the ligand-free form of 5YEY (black spectrum), the T18-H3 proton is observable and in the region typical of a Watson-Crick basepair (to A6). In the presence of one equivaent of QQPQ (red spectrum), the Watson-Crick peak is no longer present. B) $^{1}H,^{1}H$-NOESY of 5YEY (d**G**$_3$TTA**G**$_3$TTA**G**$_3$TT<u>T</u>**G**$_3$) with an excess of QQPQ, showing one NOE network connecting the foldamer amide protons (NH). The single NOE network suggests that the foldamer is located in one distinct chemical environment. 1D spectra were collected at 700 MHz, and 2D spectra at 800 MHz. Samples were measured at 278 K and contained 100 µM 5YEY in 10 mM KP buffer (pH 7.0) in 90/10 H_2O/D_2O, without and with 100 µM QQPQ.*

Molecular dynamics of the complexes

Crystallography experiments indicate that the bound QQPQ remains quite dynamic, showing no clear preference for a specific enantiomer and binding through either its N- or C-terminus. Additionally, these experiments did not allow observation of binding at the 5'-end of the 222T due to its dimerization, nor the peculiar binding to 5YEY. To gain further insights into these aspects, we conducted GPU-enabled one-microsecond molecular dynamics simulations using the OL21 Amber force field (see methods in supporting information). The 1:1 complex from the crystal structure served as a reference, with 222T as a monomer and without the second foldamer interacting with a neighboring unit cell. Based on this reference complex, three other 1:1 complexes were prepared wherein QQPQ binds through its C-terminus (3CM) or as the (P) enantiomer (3NP and 3CP). Additionally, four 2:1 complexes were prepared to study the 5'-end binding: 3NP/5NP, 3NP/5CP, 3NM/5NM, and 3NM/5CM (Figure S135).

MD simulations show that 222T can accommodate QQPQ on both end tetrads without causing steric hindrance or structural modification to either partner (Figures S135—S139). In particular, QQPQ retains its helicity. Both enantiomers bind through their N- and C-termini without clear preference, which aligns with the extensive binding and absence of induced CD. The side chains can bind to the backbone and, to a lower extent sugars (Figures S140—S141). When QQPQ is bound in 3', the side chains can establish H-bonds with the O4 and O2 acceptors of loop and terminal thymines, but this is almost never the case in 5' (Figure S142).

A notable finding is that QQPQ can rotate along its helical axis (Figures S143—S145), sometimes completing a full rotation within a microsecond of simulation (Figure S146). Additionally, the side chains exhibit significant dynamics and do not form long-lasting interactions with DNA (Figures S140—S142). These observations likely account for the poor electron density of QQPQ in crystallography. Similarly, the 5' and 3' dT are highly dynamic, which explains their absence from the crystal data. However, there were a few instances where their N3 H-bound to the amide oxygen of QQPQ, suggesting a potential particular affinity for thymines (Figure 9 and Figure S147).

Finally, we investigated the binding of QQPQ to 5YEY. A complex was assembled from the hybrid 1 structure of 2JSM, adjusted to match the sequence of 5YEY and capped with QQPQ above the 5'-tetrad. The MD trajectory was particularly stable (Figure S148), and provided two key observations supporting the conversion of 5YEY to a hybrid 1 topology with the disruption of the A6:T18 base pair. First, the 5'-tetrad of 5YEY can accommodate QQPQ, as it is only

occupied by the T16T17T18 lateral loop, which uses minimal space in the absence of the A6:T18 base pair (Figure S149A). Second, the disruption of this base pair allows for the formation of a stable hydrogen bond between the N3 atom of T18 and the amide connecting the N-terminus quinolines of QQPQ (Figure S149A, B). Again, this binding mode is similar to that of the 5' T1 with 222T described above (Figure S147), and is consistent with the proximity of QQPQ and T2, when the former stacks on the 5'-end of 2LK7 (Figure 8). Finally, T17 stacks over the C-terminus quinoline of QQPQ, further stabilizing the complex (Figure S149C).

The sequence and topology preference of oligoamide quinoline/pyridine containing foldamers can be summarized in Figure 9. The foldamers can stack on terminal G-quartets, preferably when they are sterically accessible. If the two termini are accessible, 2:1 complexes can form (Figure 9A). Exceptionally, binding to some sequences forming non-parallel stranded structures can also occur (e.g., Figure 9B, proposed topology for the 1:1 complex between the 5YEY sequence and QQPQ), provided that there are thymines in the loop, because thymines can form hydrogen bonds with the outward-pointing amide carbonyl groups of the foldamer (Figure 9C, a snapshot extracted from MD simulations).

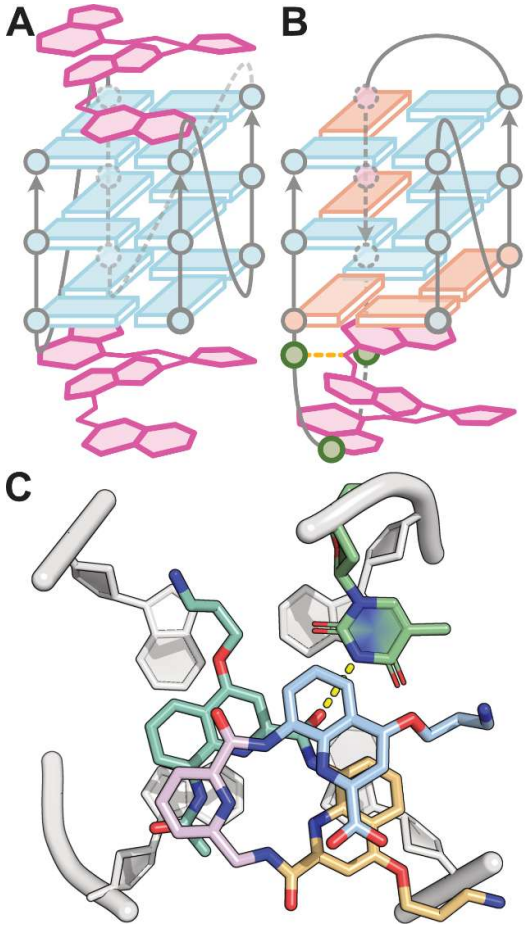

Figure 9: Scheme of the deduced foldamer ligand binding modes. A) Stacking on terminal quartets of a parallel sequence (e.g., 222T), when they are accessible. B) Stacking on a terminal quartet of a non-parallel sequence (e.g., 5YEY), with loop thymines in proximity. C) A snapshot of molecular dynamics started from the crystal structure 8QN2, showing a transient hydrogen bond between an outward pointing amide carbonyl of the foldamer and the N3—H group of a thymine.

Conclusions

Our comprehensive survey of the interaction between G4s and small oligoamide quinoline/pyridine foldamers reveals exciting opportunities to develop sequence- and structure-selective ligands. These foldamers do not target double-stranded DNA, single strands or i-motifs. They selectively target parallel G4s, forming 1:1 and 2:1 complex with sub-μM K_D values. The crystal structure reveals two quinoline units in the foldamer helix interact with an exposed G-tetrad by π-stacking. The 'bulkiness' of the foldamer helix leads makes it selective for parallel G4 topology with unobstructed 3' and 5' G-tetrads. The helix clashes with nucleobases surrounding the G-tetrad, making it sensitive to loops and flanking nucleotides. We also found that thymines in the loops favor the interaction of foldamers with some specific sequences, offering further opportunities to fine-tune the selectivity.

We believe that through rational design, the structure of the foldamer helix can be adjusted to adapt itself to the terminal G-tetrad but also to the surrounding nucleobases. For example, we showed that adding flexibility to the foldamer by incorporating pyridine-based monomers did increase the affinity while preserving G4 specificity. Another intriguing starting point is the particularly high affinity of the QQPQ foldamer to the 5YEY antiparallel variant of the telomeric sequence, which calls for more in-depth structural studies to understand how foldamers could target non-parallel G4s.

As foldamers are made by solid-phase synthesis, other variants can be easily conceived and synthesized. We showed that combinations of four monomer ensured high affinity and some selectivity. Some obvious next steps would be to conceive foldamer compounds with other subunits to modulate the stacking (first two subunits), the flexibility (third subunit), and to explore a variety of sidechains. Overall, our study provides fundamental guidance to those looking to use foldamers as a highly promising ligand scaffold that could be moulded to selectively target a G4 target of interest.

Author Contributions

L.Y. first serendipitously discovered foldamer-G4 interactions and recorded preliminary data. Supervised by Y.F., V.L. synthesized, purified and characterized the foldamer. A.K. prepared and measured the ligand screening, titration, CD and UV melting experiments. A.K. and E.L. carried out MS data processing and visualization. S.T. supervised A.K. during crystallogenesis experiments. S.T. solved and refined the crystal structure. C.M. supervised A.K. for NMR experiments and data processing. A.K. wrote the manuscript with revisions and edits from all authors. V.G. conceived and supervised the entire work and acquired funding.

Conflicts of Interest

The authors declare no conflict of interest.

Acknowledgements

We thank Dr. Frédéric Rosu for precious advice on native MS of G4s and useful discussions, Benjamin Lienard and Guilhem Ribera for the technical support during their internship. The work benefited from access to NMR and MS at the "Plateforme de BioPhysico-Chimie Structurale" of the European Institute of Chemistry and Biology (IECB), and the DOREMI CALI v3 cluster of the "Mésocentre de Calcul Intensif Aquitain" (Université de Bordeaux, France). This work was funded by the Agence National de la Recherche (project ANR-18-CE29-0013 POLYnESI to VG and YF) and by the French Ministry of Higher Education and Research (MESRI PhD fellowship to AK).

References

(1) Parkinson, G. N. The Reality of Quadruplex Nucleic Acids as a Therapeutic Target. In *Guanine Quartets*; 2012; pp 237–247. https://doi.org/10.1039/9781849736954-00237.

(2) Largy, E.; Marchand, A.; Amrane, S.; Gabelica, V.; Mergny, J.-L. Quadruplex Turncoats: Cation-Dependent Folding and Stability of Quadruplex-DNA Double Switches. *Journal of the American Chemical Society*, 2016, *138*, 2780–2792. https://doi.org/10.1021/jacs.5b13130.

(3) Marchand, A.; Gabelica, V. Folding and Misfolding Pathways of G-Quadruplex DNA. *Nucleic Acids Research*, 2016, *44*, 10999–11012. https://doi.org/10.1093/nar/gkw970 %J Nucleic Acids Research.

(4) Frelih, T.; Wang, B.; Plavec, J.; Šket, P. Pre-Folded Structures Govern Folding Pathways of Human Telomeric G-Quadruplexes. *Nucleic Acids Research* **2020**, *48* (4), 2189–2197. https://doi.org/10.1093/nar/gkz1235.

(5) Gabelica, V. A Pilgrim's Guide to G-Quadruplex Nucleic Acid Folding. *Biochimie*, 2014, *105*, 1–3. https://doi.org/10.1016/j.biochi.2014.08.007.

(6) Šponer, J.; Bussi, G.; Stadlbauer, P.; Kührová, P.; Banáš, P.; Islam, B.; Haider, S.; Neidle, S.; Otyepka, M. Folding of Guanine Quadruplex Molecules–Funnel-like Mechanism or Kinetic Partitioning? An Overview from MD Simulation Studies. *Biochimica et Biophysica Acta (BBA) - General Subjects* **2017**, *1861* (5, Part B), 1246–1263. https://doi.org/10.1016/j.bbagen.2016.12.008.

(7) Largy, E.; Mergny, J.-L. Shape Matters: Size-Exclusion HPLC for the Study of Nucleic Acid Structural Polymorphism. *Nucleic Acids Research* **2014**, *42* (19), e149. https://doi.org/10.1093/nar/gku751.

(8) Hardin, C. C.; Watson, T.; Corregan, M.; Bailey, C. Cation-Dependent Transition between the Quadruplex and Watson-Crick Hairpin Forms of d(CGCG3GCG). *Biochemistry* **1992**, *31* (3), 833–841. https://doi.org/10.1021/bi00118a028.

(9) Neidle, S.; Balasubramanian, S. The Role of Cations in Determining Quadruplex Structure and Stability. In *Quadruplex Nucleic Acids*; 2006; pp 100–130. https://doi.org/10.1039/9781847555298-00100.

(10) Largy, E.; Mergny, J.-L.; Gabelica, V. Role of Alkali Metal Ions in G-Quadruplex Nucleic Acid Structure and Stability. *The Alkali Metal Ions: Their Role for Life*, 2016, 203–258. https://doi.org/10.1007/978-3-319-21756-7_7.

(11) Khristenko, N.; Amato, J.; Livet, S.; Pagano, B.; Randazzo, A.; Gabelica, V. Native Ion Mobility Mass Spectrometry: When Gas-Phase Ion Structures Depend on the Electrospray Charging Process. *J. Am. Soc. Mass Spectrom.* **2019**, *30* (6), 1069–1081. https://doi.org/10.1007/s13361-019-02152-3.

(12) Griffeycor, R. H.; Sasmor, H.; Greig, M. J. Oligonucleotide Charge States in Negative Ionization Electrospray-Mass Spectrometry Are a Function of Solution Ammonium Ion Concentration. *J Am Soc Mass Spectrom* **1997**, *8* (2), 155–160. https://doi.org/10.1016/S1044-0305(96)00200-0.

(13) Burge, S.; Parkinson, G. N.; Hazel, P.; Todd, A. K.; Neidle, S. Quadruplex DNA: Sequence, Topology and Structure. *Nucleic Acids Research* **2006**, *34* (19), 5402–5415. https://doi.org/10.1093/nar/gkl655.

(14) Huppert, J. L. Four-Stranded Nucleic Acids: Structure, Function and Targeting of G-Quadruplexes. *Chem. Soc. Rev.* **2008**, *37* (7), 1375–1384. https://doi.org/10.1039/B702491F.

(15) Phan, A. T. Human Telomeric G-Quadruplex: Structures of DNA and RNA Sequences. *The FEBS Journal* **2010**, *277* (5), 1107–1117. https://doi.org/10.1111/j.1742-4658.2009.07464.x.

(16) Ghosh, A.; Largy, E.; Gabelica, V. DNA G-Quadruplexes for Native Mass Spectrometry in Potassium: A Database of Validated Structures in Electrospray-Compatible Conditions. *Nucleic Acids Research* **2021**, *49* (4), 2333–2345. https://doi.org/10.1093/nar/gkab039.

(17) Chaudhuri, R.; Bhattacharya, S.; Dash, J.; Bhattacharya, S. Recent Update on Targeting C-MYC G-Quadruplexes by Small Molecules for Anticancer Therapeutics. *J. Med. Chem.* **2021**, *64* (1), 42–70. https://doi.org/10.1021/acs.jmedchem.0c01145.

(18) Patel, D. J.; Phan, A. T.; Kuryavyi, V. Human Telomere, Oncogenic Promoter and 5′-UTR G-Quadruplexes: Diverse Higher Order DNA and RNA Targets for Cancer Therapeutics. *Nucleic Acids Research* **2007**, *35* (22), 7429–7455. https://doi.org/10.1093/nar/gkm711.

(19) Kosiol, N.; Juranek, S.; Brossart, P.; Heine, A.; Paeschke, K. G-Quadruplexes: A Promising Target for Cancer Therapy. *Molecular Cancer* **2021**, *20* (1), 40. https://doi.org/10.1186/s12943-021-01328-4.

(20) Dai, J.; Carver, M.; Yang, D. Polymorphism of Human Telomeric Quadruplex Structures. *Biochimie* **2008**, *90* (8), 1172–1183. https://doi.org/10.1016/j.biochi.2008.02.026.

(21) Phan, A. T.; Mergny, J.-L. Human Telomeric DNA: G-Quadruplex, i-Motif and Watson-Crick Double Helix. *Nucleic Acids Res* **2002**, *30* (21), 4618–4625. https://doi.org/10.1093/nar/gkf597.

(22) Phan, A. T. Human Telomeric G-Quadruplex: Structures of DNA and RNA Sequences: Human Telomeric G-Quadruplex Structures. *FEBS Journal* **2010**, *277* (5), 1107–1117. https://doi.org/10.1111/j.1742-4658.2009.07464.x.

(23) Gao, J.; Pickett, H. A. Targeting Telomeres: Advances in Telomere Maintenance Mechanism-Specific Cancer Therapies. *Nat Rev Cancer* **2022**, *22* (9), 515–532. https://doi.org/10.1038/s41568-022-00490-1.

(24) Huppert, J. L.; Balasubramanian, S. Prevalence of Quadruplexes in the Human Genome. *Nucleic Acids Research* **2005**, *33* (9), 2908–2916. https://doi.org/10.1093/nar/gki609.

(25) Monchaud, D.; Allain, C.; Bertrand, H.; Smargiasso, N.; Rosu, F.; Gabelica, V.; De Cian, A.; Mergny, J.-L.; Teulade-Fichou, M.-P. Ligands Playing Musical Chairs with G-Quadruplex DNA: A Rapid and Simple Displacement Assay for Identifying Selective G-Quadruplex Binders. *Biochimie* **2008**, *90* (8), 1207–1223. https://doi.org/10.1016/j.biochi.2008.02.019.

(26) Micheli, E.; Lombardo, C. M.; D'Ambrosio, D.; Franceschin, M.; Neidle, S.; Savino, M. Selective G-Quadruplex Ligands: The Significant Role of Side Chain Charge Density in a Series of Perylene Derivatives. *Bioorganic & Medicinal Chemistry Letters* **2009**, *19* (14), 3903–3908. https://doi.org/10.1016/j.bmcl.2009.03.106.

(27) Ma, Y.; Iida, K.; Nagasawa, K. Topologies of G-Quadruplex: Biological Functions and Regulation by Ligands. *Biochemical and Biophysical Research Communications* **2020**, *531* (1), 3–17. https://doi.org/10.1016/j.bbrc.2019.12.103.

(28) Müller, D.; Saha, P.; Panda, D.; Dash, J.; Schwalbe, H. Insights from Binding on Quadruplex Selective Carbazole Ligands. *Chemistry – A European Journal* **2021**, *27* (50), 12726–12736. https://doi.org/10.1002/chem.202101866.

(29) Dickerhoff, J.; Dai, J.; Yang, D. Structural Recognition of the MYC Promoter G-Quadruplex by a Quinoline Derivative: Insights into Molecular Targeting of Parallel G-Quadruplexes. *Nucleic Acids Research* **2021**, *49* (10), 5905–5915. https://doi.org/10.1093/nar/gkab330.

(30) Ritson, D. J.; Moses, J. E. A Fragment Based Click Chemistry Approach towards Hybrid G-Quadruplex Ligands: Design, Synthesis and Biophysical Evaluation. *Tetrahedron* **2012**, *68* (1), 197–203. https://doi.org/10.1016/j.tet.2011.10.066.

(31) Reznichenko, O.; Cucchiarini, A.; Gabelica, V.; Granzhan, A. Quadruplex DNA-Guided Ligand Selection from Dynamic Combinatorial Libraries of Acylhydrazones. *Org. Biomol. Chem.* **2021**, *19* (2), 379–386. https://doi.org/10.1039/D0OB01908A.

(32) Hamon, F.; Largy, E.; Guédin-Beaurepaire, A.; Rouchon-Dagois, M.; Sidibe, A.; Monchaud, D.; Mergny, J.-L.; Riou, J.-F.; Nguyen, C.-H.; Teulade-Fichou, M.-P. An Acyclic Oligoheteroaryle That Discriminates Strongly between Diverse G-Quadruplex Topologies. *Angewandte Chemie International Edition* **2011**, *50* (37), 8745–8749. https://doi.org/10.1002/anie.201103422.

(33) Lin, C.; Wu, G.; Wang, K.; Onel, B.; Sakai, S.; Shao, Y.; Yang, D. Molecular Recognition of the Hybrid-2 Human Telomeric G-Quadruplex by Epiberberine: Insights into Conversion of Telomeric G-Quadruplex Structures. *Angewandte Chemie International Edition* **2018**, *57* (34), 10888–10893. https://doi.org/10.1002/anie.201804667.

(34) McQuaid, K. T.; Takahashi, S.; Baumgaertner, L.; Cardin, D. J.; Paterson, N. G.; Hall, J. P.; Sugimoto, N.; Cardin, C. J. Ruthenium Polypyridyl Complex Bound to a Unimolecular Chair-Form G-Quadruplex. *J. Am. Chem. Soc.* **2022**, *144* (13), 5956–5964. https://doi.org/10.1021/jacs.2c00178.

(35) Largy, E.; Teulade-Fichou, M.-P. Screening for Quadruplex Binding Ligands: A Game of Chance? **2012**. https://doi.org/10.1039/9781849736954-00248.

(36) Guichard, G.; Huc, I. Synthetic Foldamers. *Chem. Commun.* **2011**, *47* (21), 5933–5941. https://doi.org/10.1039/C1CC11137J.

(37) Jena, P. V.; Shirude, P. S.; Okumus, B.; Laxmi-Reddy, K.; Godde, F.; Huc, I.; Balasubramanian, S.; Ha, T. G-Quadruplex DNA Bound by a Synthetic Ligand Is Highly Dynamic. *J. Am. Chem. Soc.* **2009**, *131* (35), 12522–12523. https://doi.org/10.1021/ja903408r.

(38) Müller, S.; Laxmi-Reddy, K.; Jena, P. V.; Baptiste, B.; Dong, Z.; Godde, F.; Ha, T.; Rodriguez, R.; Balasubramanian, S.; Huc, I. Targeting DNA G-Quadruplexes with Helical Small Molecules, 2014, *15*, 2563–2570. https://doi.org/10.1002/cbic.201402439.

(39) Mandal, P. K.; Baptiste, B.; Langlois d'Estaintot, B.; Kauffmann, B.; Huc, I. Multivalent Interactions between an Aromatic Helical Foldamer and a DNA G-Quadruplex in the Solid State. *ChemBioChem* **2016**, *17* (20), 1911–1914. https://doi.org/10.1002/cbic.201600281.

(40) Qi, T.; Maurizot, V.; Noguchi, H.; Charoenraks, T.; Kauffmann, B.; Takafuji, M.; Ihara, H.; Huc, I. Solvent Dependence of Helix Stability in Aromatic Oligoamide Foldamers. *Chem. Commun.* **2012**, *48* (51), 6337–6339. https://doi.org/10.1039/C2CC31533E.

(41) Shirude, P. S.; Gillies, E. R.; Ladame, S.; Godde, F.; Shin-ya, K.; Huc, I.; Balasubramanian, S. Macrocyclic and Helical Oligoamides as a New Class of G-Quadruplex Ligands. *J. Am. Chem. Soc.* **2007**, *129* (39), 11890–11891. https://doi.org/10.1021/ja073775h.

(42) Delaurière, L.; Dong, Z.; Laxmi-Reddy, K.; Godde, F.; Toulmé, J.-J.; Huc, I. Deciphering Aromatic Oligoamide Foldamer–DNA Interactions. *Angewandte Chemie International Edition* **2012**, *51* (2), 473–477. https://doi.org/10.1002/anie.201106208.

(43) Summers, P. A.; Thomas, A. P.; Kench, T.; Vannier, J.-B.; Kuimova, M. K.; Vilar, R. Cationic Helicenes as Selective G4 DNA Binders and Optical Probes for Cellular Imaging. *Chem. Sci.* **2021**, *12* (43), 14624–14634. https://doi.org/10.1039/D1SC04567A.

(44) Tataurov, A. V.; You, Y.; Owczarzy, R. Predicting Ultraviolet Spectrum of Single Stranded and Double Stranded Deoxyribonucleic Acids. *Biophysical Chemistry* **2008**, *133* (1), 66–70. https://doi.org/10.1016/j.bpc.2007.12.004.

(45) Mergny, L., Jean-Louis; Lacroix. Analysis of Thermal Melting Curves. *Oligonucleotides*, 2003, 515–537.

(46) Mergny, J. L.; Lacroix, L. UV Melting of G-Quadruplexes. *Curr Protoc Nucleic Acid Chem*, 2009, *Chapter 17*, Unit 17.1. https://doi.org/10.1002/0471142700.nc1701s37.

(47) Gabelica, V.; Livet, S.; Rosu, F. Optimizing Native Ion Mobility Q-TOF in Helium and Nitrogen for Very Fragile Noncovalent Structures. *Journal of the American Society for Mass Spectrometry*, 2018, *29*, 2189–2198. https://doi.org/10.1021/jasms.8b05701.

(48) Marchand, A.; Livet, S.; Rosu, F.; Gabelica, V. Drift Tube Ion Mobility: How to Reconstruct Collision Cross Section Distributions from Arrival Time Distributions? *Anal. Chem.* **2017**, *89* (23), 12674–12681. https://doi.org/10.1021/acs.analchem.7b01736.

(49) Largy, E.; König, A.; Ghosh, A.; Ghosh, D.; Benabou, S.; Rosu, F.; Gabelica, V. Mass Spectrometry of Nucleic Acid Noncovalent Complexes. *Chem. Rev.* **2021**. https://doi.org/10.1021/acs.chemrev.1c00386.

(50) Rosu, F.; De Pauw, E.; Gabelica, V. Electrospray Mass Spectrometry to Study Drug-Nucleic Acids Interactions. *Biochimie* **2008**, *90* (7), 1074–1087. https://doi.org/10.1016/j.biochi.2008.01.005.

(51) Gabelica, V.; Rosu, F.; De Pauw, E. A Simple Method to Determine Electrospray Response Factors of Noncovalent Complexes. *Analytical Chemistry*, 2009, *81*, 6708–6715. https://doi.org/10.1021/ac900785m.

(52) Daubenfeld, T.; Bouin, A.-P.; van der Rest, G. A Deconvolution Method for the Separation of Specific Versus Nonspecific Interactions in Noncovalent Protein-Ligand Complexes Analyzed by ESI-FT-ICR Mass Spectrometry. *Journal of the American Society for Mass Spectrometry* **2006**, *17* (9), 1239–1248. https://doi.org/10.1016/j.jasms.2006.05.005.

(53) Kabsch, W. Automatic Processing of Rotation Diffraction Data from Crystals of Initially Unknown Symmetry and Cell Constants. *Journal of Applied Crystallography* **1993**, *26* (6), 795–800. https://doi.org/10.1107/S0021889893005588.

(54) Liebschner, D.; Afonine, P. V.; Baker, M. L.; Bunkóczi, G.; Chen, V. B.; Croll, T. I.; Hintze, B.; Hung, L.-W.; Jain, S.; McCoy, A. J.; Moriarty, N. W.; Oeffner, R. D.; Poon, B. K.; Prisant, M. G.; Read, R. J.; Richardson, J. S.; Richardson, D. C.; Sammito, M. D.; Sobolev, O. V.; Stockwell, D. H.; Terwilliger, T. C.; Urzhumtsev, A. G.; Videau, L. L.; Williams, C. J.; Adams, P. D. Macromolecular Structure Determination Using X-Rays, Neutrons and Electrons: Recent Developments in Phenix. *Acta Cryst D* **2019**, *75* (10), 861–877. https://doi.org/10.1107/S2059798319011471.

(55) Lin, L. Y.; McCarthy, S.; Powell, B. M.; Manurung, Y.; Xiang, I. M.; Dean, W. L.; Chaires, B.; Yatsunyk, L. A. Biophysical and X-Ray Structural Studies of the (GGGTT)3GGG G-Quadruplex in Complex with N-Methyl Mesoporphyrin IX. *PLOS ONE* **2020**, *15* (11), e0241513. https://doi.org/10.1371/journal.pone.0241513.

(56) Smart, O. S.; Womack, T. O.; Sharff, A.; Flensburg, C.; Keller, P.; Paciorek, W.; Vonrhein, C.; Bricogne, G. Grade2 Version 1.4.1, Global Phasing Ltd.

(57) Pettersen, E. F.; Goddard, T. D.; Huang, C. C.; Meng, E. C.; Couch, G. S.; Croll, T. I.; Morris, J. H.; Ferrin, T. E. UCSF ChimeraX: Structure Visualization for Researchers, Educators, and Developers. *Protein Science* **2021**, *30* (1), 70–82. https://doi.org/10.1002/pro.3943.

(58) Delaglio, F.; Grzesiek, S.; Vuister, G. W.; Zhu, G.; Pfeifer, J.; Bax, A. NMRPipe: A Multidimensional Spectral Processing System Based on UNIX Pipes. *J Biomol NMR* **1995**, *6* (3), 277–293. https://doi.org/10.1007/BF00197809.

(59) Do, N. Q.; Phan, A. T. Monomer–Dimer Equilibrium for the 5′–5′ Stacking of Propeller-Type Parallel-Stranded G-Quadruplexes: NMR Structural Study. *Chemistry – A European Journal* **2012**, *18* (46), 14752–14759. https://doi.org/10.1002/chem.201103295.

(60) Šket, P.; Plavec, J. Tetramolecular DNA Quadruplexes in Solution: Insights into Structural Diversity and Cation Movement. *J. Am. Chem. Soc.* **2010**, *132* (36), 12724–12732. https://doi.org/10.1021/ja104889t.

(61) Lenarčič Živković, M.; Rozman, J.; Plavec, J. Adenine-Driven Structural Switch from a Two- to Three-Quartet DNA G-Quadruplex. *Angewandte Chemie International Edition* **2018**, *57* (47), 15395–15399. https://doi.org/10.1002/anie.201809328.

(62) Akashi, M.; Shaw, G.; Hachiya, M.; Elstner, E.; Suzuki, G.; Koeffler, P. Number and Location of AUUUA Motifs: Role in Regulating Transiently Expressed RNAs. *Blood* **1994**, *83* (11), 3182–3187. https://doi.org/10.1182/blood.V83.11.3182.3182.

(63) Munoz-Tello, P.; Rajappa, L.; Coquille, S.; Thore, S. Polyuridylation in Eukaryotes: A 3′-End Modification Regulating RNA Life. *Biomed Res Int* **2015**, *2015*, 968127. https://doi.org/10.1155/2015/968127.

(64) Marchand, A.; Rosu, F.; Zenobi, R.; Gabelica, V. Thermal Denaturation of DNA G-Quadruplexes and Their Complexes with Ligands: Thermodynamic Analysis of the

Multiple States Revealed by Mass Spectrometry. *J. Am. Chem. Soc.* **2018**, *140* (39), 12553–12565. https://doi.org/10.1021/jacs.8b07302.

(65) Li, J.; Correia, J. J.; Wang, L.; Trent, J. O.; Chaires, J. B. Not so Crystal Clear: The Structure of the Human Telomere G-Quadruplex in Solution Differs from That Present in a Crystal. *Nucleic Acids Research* **2005**, *33* (14), 4649–4659. https://doi.org/10.1093/nar/gki782.

(66) Zhang, D.-H.; Fujimoto, T.; Saxena, S.; Yu, H.-Q.; Miyoshi, D.; Sugimoto, N. Monomorphic RNA G-Quadruplex and Polymorphic DNA G-Quadruplex Structures Responding to Cellular Environmental Factors. *Biochemistry* **2010**, *49* (21), 4554–4563. https://doi.org/10.1021/bi1002822.

(67) Smargiasso, N.; Rosu, F.; Hsia, W.; Colson, P.; Baker, E. S.; Bowers, M. T.; De Pauw, E.; Gabelica, V. G-Quadruplex DNA Assemblies: Loop Length, Cation Identity, and Multimer Formation. *J. Am. Chem. Soc.* **2008**, *130* (31), 10208–10216. https://doi.org/10.1021/ja801535e.

(68) Kogut, M.; Kleist, C.; Czub, J. Why Do G-Quadruplexes Dimerize through the 5'-Ends? Driving Forces for G4 DNA Dimerization Examined in Atomic Detail. *PLOS Computational Biology* **2019**, *15* (9), e1007383. https://doi.org/10.1371/journal.pcbi.1007383.

(69) Liu, C.; Zhou, B.; Geng, Y.; Tam, D. Y.; Feng, R.; Miao, H.; Xu, N.; Shi, X.; You, Y.; Hong, Y.; Tang, B. Z.; Lo, P. K.; Kuryavyi, V.; Zhu, G. A Chair-Type G-Quadruplex Structure Formed by a Human Telomeric Variant DNA in K+ Solution. *Chem. Sci.* **2018**, *10* (1), 218–226. https://doi.org/10.1039/C8SC03813A.