Quadruplex Turncoats: Cation-Dependent Folding and Stability of Quadruplex-DNA Double Switches

Eric Largy, Adrien Marchand, Samir Amrane, Valérie Gabelica, Jean-Louis Mergny

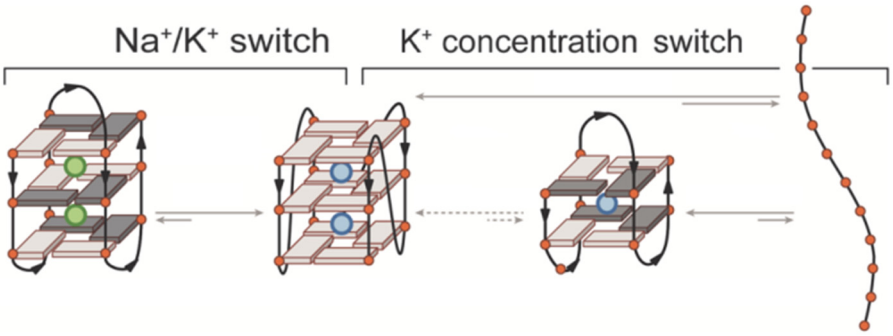
Quadruplex turncoats: cation-dependent folding and stability of quadruplex-DNA double switches

Eric Largy,[1,2,3] Adrien Marchand,[1,2,3] Samir Amrane,[1,2,3] Valérie Gabelica,[1,2,3] Jean-Louis Mergny*[1,2,3]

[1] Inserm, U1212, ARNA Laboratory, F-33000 Bordeaux, France.

[2] Univ. Bordeaux, IECB, ARNA laboratory, F-33600 Pessac, France.

[3] CNRS, UMR 5320, ARNA Laboratory, F-33600 Pessac, France.



ABSTRACT: Quadruplex (G4) nucleic acids, a family of secondary structures formed by guanine-rich sequences, exhibit an important structural polymorphism. We demonstrate here that G-rich DNA sequences may function as a double switch based on different triggers provided that their quadruplex structures and stability display a high dependence on cation nature and concentration. A first switch is based on a remarkable antiparallel to parallel conversion, taking place in a few seconds at room temperature by addition of low KCl amounts to a sodium-rich sample. The second switch involves the conversion of alternative antiparallel quadruplex structures binding only one cation, formed in the presence of submillimolar potassium or strontium concentrations, to parallel structures by increasing the cation concentration. Incidentally, extremely low K^+ or Sr^{2+} concentrations (≤ 5 equivalents) are sufficient to induce G4 formation in a buffer devoid of other G4-promoting cations, and we suggest that the alternative structures observed contain only two tetrads. Such DNA systems are biological relevant targets, can be used in nanotechnology applications, and are valuable methodological tool for the understanding of DNA quadruplex folding, notably at low cation concentrations. We demonstrate that this behavior is not restricted to a narrow set of sequences but can also be found for other G-quadruplex forming motifs, arguing for widespread applications.

1. INRODUCTION

DNA is a versatile biopolymer, which secondary structure is not limited to the classical right-handed double helix. Quadruplex (G_4) nucleic acids are a family of secondary structures formed by guanine-rich sequences. G-quadruplexes are involved in a number of key biological processes, both at the DNA and RNA level, but are also of interest for the bottom-up design of DNA-based nanostructures, because of their unique properties and orthogonal pairing rules. G-quadruplexes share the following common features: *(i)* the π-π stacking of tetrads of guanines (also called quartets) arranged in a quasi-co-planar arrangement held together by a network of 8 Hoogsteen hydrogen bonds, and *(ii)* the specific octa-coordination of monovalent (typically K^+ or Na^+) or divalent cations by the carbonyl oxygen of the quartet guanines. However, beyond these common features, G_4 nucleic acids display a variety of secondary structures, and this polymorphism is influenced by numerous variables (sequence, strand concentration, cation nature and concentration, temperature, molecular crowding agents, ligands).[1-6] The polymorphism can involve different strand stoichiometries (1 to 4 strands), strand orientations (parallel, antiparallel or hybrid), number of tetrads (at least two), and length and sequences of loops or bulges, to cite the most common. While this polymorphism makes it difficult to assess the precise structure(s) formed by a given sequence, it also provides an opportunity for specific gene targeting[3,7], DNA functionalization and nanostructure building.[8-11] Unfortunately, this polymorphism is not well understood, especially for sequences that can adopt multiple structures. This difficulty culminates with the human telomeric motif, which is one of the most – if not the most - polymorphic G_4 motif known to date.

The cation influence on G-quadruplex folding and stability is particularly prevalent,[12,13] and this makes cation-mediated switches particularly attractive. In order to be specifically bound by tetrads of guanine, a cation must be small enough to fit in the electron-rich cavity but also large enough to bridge the guanine carbonyl oxygens. Hence, Li^+ is too small to stabilize quadruplexes, Na^+ (1.18 Å) can fit in a tetrad plane and K^+ (1.51 Å) may lie in-between two tetrads (see Table S1 for ionic radii).[14,15] Various rankings of G_4 stabilization by monovalent and divalent cations have been proposed, and the general trend is: $Sr^{2+} > Ba^{2+} > K^+ > Ca^{2+} > Na^+$, NH_4^+, $Rb^+ > Mg^{2+} > Li^+ \geq Cs^+$.[16-19] Note that, in some cases (e.g. low monovalent cation concentration), some divalent cations such as Mg^{2+} can dissociate G_4 structures.[20] Besides ionic radii, factors such as the ion hydration free energy also explain the stability difference observed between Na^+ and K^+ solutions, the dehydration of the latter inducing a lesser energetic cost:[21,22] unlike diffusely bound cations, tetrad-bound cations shall lose their hydration sphere.[23] Additionally, the strontium-induced stabilization of G_4 has been explained by a strong coordination to the guanine oxygens, and a better screening of the phosphate-phosphate repulsion.[24] However, besides influencing the stability, the nature of the cations also influences the whole folding topology.[25]

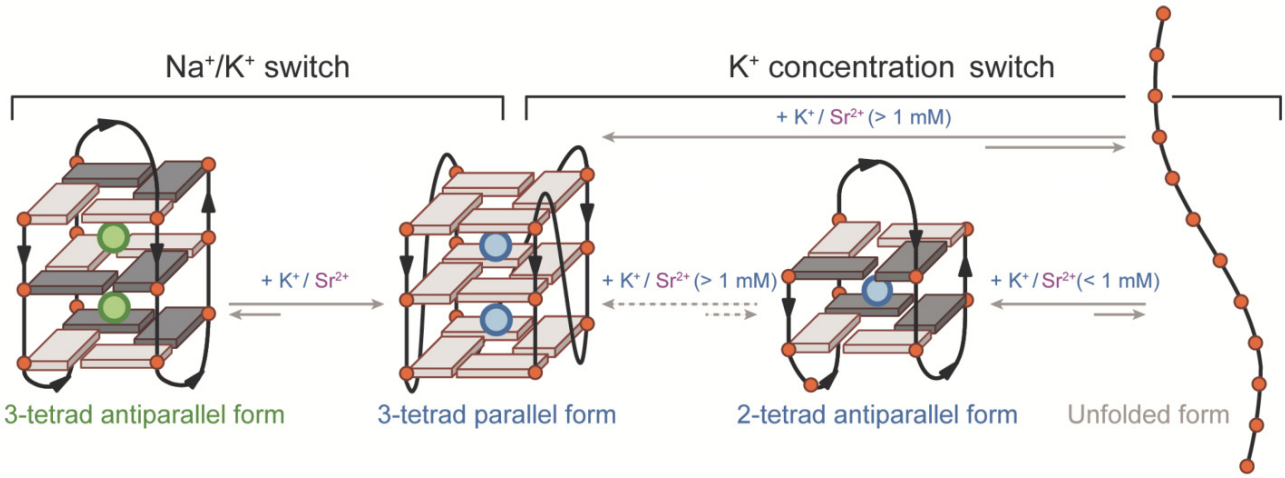


Figure 1: 222/222T can act as two different switch systems. Three different quadruplex topologies can be reached (in addition to the unfolded form), differing by the strand orientation and the number of tetrads, via tuning of the cationic conditions, flanking bases, and strand concentration (not pictured is the partial dimerization of 222). Guanines are displayed as light gray (*anti* conformation) or dark gray (*syn* conformation) rectangles as well as orange circles, and potassium and sodium cations as blue and green circles, respectively.

Cation-dependent topological changes in G4s have been observed as early as 1990 by Sen and Gilbert [26], and may find practical applications in nanotechnology or as cation probes.[12,27] More recently, a number of studies have shed light on thermodynamics and kinetics of folding of, most notably, the human telomeric sequence[25,28-36]

In this study, we wanted to better understand the cation-induced polymorphism using a simplier system that the human telomeric motif. To that aim, we launched a program aimed at discovering DNA sequences whose structures and stability display a very high cation nature and concentration dependence, enabling them to function as cation-based allosteric switches. Ideally, we were looking for sequences that would exhibit a well-defined behavior, *i.e.*, single but easily distinguishable folds in sodium and potassium. While somewhat more complex that our initial expectations, the quadruplexes studied here allowed a better understanding of cation-dependent structural changes.

Hereafter, we describe the properties of two related sequences that operate as a double switch, namely *222* (d[$(G_3T_2)_3G_3$]) and *222T* (d[$T(G_3T_2)_3G_3T$]), that both contain four triplets of guanines linked by two-thymine loops, and differ by the absence or presence, respectively, of flanking thymines. We first describe how these sequences act as an antiparallel-to-parallel G4 switch, triggered by the addition of small amounts of potassium (or strontium) to a sodium-rich sample

(Figure 1, left). A second antiparallel-to-parallel switch is then presented, triggered by an increase in KCl (or $SrCl_2$) concentration, where the starting antiparallel quadruplex likely contains only two tetrads (Figure 1, right). Variant sequences, as well as alternative group 1 and 2 cations, were investigated and provide additional insight into sequence and cation-dependency of G4 folding. Thanks to these multiple-state switch properties, such DNA systems present a functional interest (e.g. as nanomachines or nanoarchitectures),[8-11] but their structural versatility also makes them valuable methodological tool for the understanding of cation-dependent DNA G4 folding thermodynamics, kinetics, and excellent case studies for the development of associated analytical methods. Finally, we demonstrate that a number of unrelated G-rich sequences adopt the same cation switch behavior. Interestingly, mirror-image sequences exhibit somewhat different behaviors, highlighting the importance of strand polarity in G4 folding. The 222/ 222T system may provide invaluable information on the folding and interconversion pathways of G-quadruplexes. Results are presented in two distinct parts, in which we either study the conversion from one fold to the other (part I) or the cation-induced folding of an unfolded strand (part II).

2. MATERIALS AND METHODS

Oligonucleotides were purchased from Eurogentec (Seraing, Belgium) in reverse-phase purified lyophilized form. All salts and solutions were obtained at the molecular biology grade from Sigma-Aldrich (Saint-Quentin Fallavier, France). Buffer and saline solutions were prepared from ultrapure water (18.2 Ω cm^{-1} resistivity). Mass spectrometry samples were prepared in nuclease-free water (Ambion, Life technologies SAS, Saint-Aubin, France). Data analysis was performed with OriginPro 9.1 (OriginLab, Northampton, MA, USA) and Mathematica 10 (Wolfram Research, Champaign, IL, USA).

Circular Dichroism (CD) experiments were performed with a JASCO J-815 spectropolarimeter equipped with a JASCO CDF-426S Peltier temperature controller, using quartz cells of 2 mm path length. The scans were recorded at 20 °C or 4 °C, from 210 to 350 nm with the following parameters: 0.2 nm data pitch, 2 nm bandwidth, 0.5 s response, 50 nm min^{-1} scanning speed, and are the result of three accumulations (except for time-dependent experiments where single scans were acquired at 200 nm min^{-1}). Solutions were prepared by diluting thermally-renatured oligonucleotide samples at 10 µM in lithium cacodylate (20 mM, pH 7.2) or trimethylammonium acetate (TMAA; *qs* for a 120 mM salt concentration), supplemented with various concentrations of group 1 or 2 cation chlorides. The CD data were blank-subtracted and normalized to molar dichroic absorption ($\Delta\varepsilon$) based on nucleoside concentration using Equation 1, with θ the ellipticity in millidegrees, c the nucleoside concentration in mol L^{-1}, and l the path length in cm.

$$\Delta\varepsilon = \theta / (32980 \times c \times l) \quad (1)$$

In time-dependent experiments, plots at 3 characteristic wavelengths (typically 240, 260, and 295 nm) were simultaneously fitted with single (j = 1) and double exponential (j = 2) models (Equation 2), where $\Delta\varepsilon_0$ is the offset, A_i and k_i are respectively the span and the rate constant of the process i, and are shared parameters during the fitting process.

$$\Delta\varepsilon = \Delta\varepsilon_0 + \sum_{i=1}^{j}[A_i \times \exp(-k_i \times t)] \quad (2)$$

The relative quality of both models was assessed for all data sets using the corrected Akaike information criterion (AICc) and the extra sum-of-squares F tests. The double exponential model was selected in each case as the most relevant.

Half-life $t_{1/2}$ and time constant τ were calculated following their classical definition:

$$t_{1/2}^{i} = \ln(2)/k_i \quad (3)$$

$$\tau_i = 1/k_i \quad (4)$$

In CD screening experiments, spectra were acquired in lithium cacodylate supplemented with 100 mM of either KCl or NaCl to assess whether the sequences display distinct topologies in sodium- and potassium-rich conditions. Then, KCl (100 mM) was added to the Na^+-containing sample at constant temperature, and a CD spectrum was acquired 3—4 hours later. In titration experiments of Na^+ samples in lithium cacodylate, spectra were recorded after a 3-min incubation period following KCl or $SrCl_2$ addition. Titrations of unfolded samples were performed with independent solutions prepared in TMAA, and incubated overnight.

Dual wavelength parametric test of the two state model (referred herein as "Wallimann plots"[37]) were constructed by plotting the CD intensity at a given wavelength versus the CD intensity at another wavelength, then performing a linear regression on the data set. Three wavelengths were used (~ 240, 260 and 295 nm). Linearity of the plot suggests a 2-state model, while deviation from the linearity indicates the presence of at least 3 species.

Singular Value Decomposition (SVD) was adapted in Wolfram Mathematica 10 from the protocol of Gray and Chaires.[38] Raw data of potassium concentration-dependent or time-dependent experiments (matrix D) were decomposed into three matrices U (basis spectra), S (singular values) and V (amplitude vectors as a function of potassium concentration or time, respectively).

$$D = USV \quad (5)$$

The scree plots (a chart associating the components to their singular value) were generated for the ten highest singular values, and the corresponding percentages of total variance were obtained from the relative variance (RV) of singular values s_i:

$$RV = \frac{s_i^2}{\sum_i s_i^2} \quad (6)$$

In a first approach, the significant components were selected so that the sum of their contribution reaches at least 99.75% of the total variance.

1H NMR experiments were performed on a 700 MHz NMR Bruker spectrometer, at 20 °C, on 300-µM oligonucleotide samples. The jump-and-return method (1D ^{1}H JR) was used for water suppression in 1D spectrum acquisition.

Size Exclusion-HPLC was conducted as described previously, on a UltiMate 3000 UHPLC system (Thermo Scientific Dionex, Sunnyvale, CA, USA), equipped with an autosampler, a diode array detector and a ThermoAcclaim SEC-300 column (4.6×300 mm; 5-µm hydrophilic polymethacrylate resin spherical particles, 300 Å pore size) set at 20 °C.[6] Elution was performed at 0.150 ml/min in Tris•HCl 50 mM, pH 7.5, supplemented with either 100 mM NaCl, 100 mM KCl, or 100 mM NaCl + 5 mM KCl.

UV absorption spectroscopy. Melting temperatures were examined by measuring the changes in absorbance at 295 nm as a function of the temperature, using a UVmc2 double-beam spectrophotometer (SAFAS, Monte Carlo, Monaco) equipped with a high-performance Peltier temperature controller and a thermostatable ten-cell holder. The samples containing the oligonucleotide in cacodylate buffer (20 mM $LiAsMe_2O_2$, pH 7.2), supplemented by various concentrations of chloride salts of a group 1 or 2 cation and lithium chloride (*qs* for a 120 mM salt concentration), were first heated to 95 °C, then the absorbance was monitored at 260, 295 and 350 nm on a cycle composed of a cooling down to 4 °C at a rate of 0.2 °C min^{-1}, then heating back up at 95 °C at the same rate. Raw data were blank subtracted, and melting temperatures were subsequently obtained by determining the maximum of the first derivative data, assuming only two thermodynamic states. Thermodynamic parameters were determined from the data restricted to a 10 °C temperature window centered on the melting temperature. The association constant K_A of the folding reaction (equation 7) is calculated for each data point assuming the association of two cations with one DNA molecule, unless otherwise specified.

$$unfolded + 2K^+ \rightleftharpoons quadruplex \cdot 2K^+ \quad (7a)$$

$$K_a = \frac{[quadruplex \cdot 2K^+]}{[unfolded] \times [K^+]^2} \quad (7b)$$

Subsequently, $\Delta G°_{folding}$ is calculated for each point using equation 8 [39]. The enthalpy $\Delta H°$ and entropy $\Delta S°$ of folding are extracted from the intercept and slope, respectively, of the $\Delta G°_{fold}$ versus T plot.

$$\Delta G°_{fold} = -RTln(K_A) = \Delta H°_{fold} - T.\Delta S°_{fold} \quad (8)$$

Thermal Difference Spectra (TDS) were obtained by subtraction of the UV-spectra (λ = 210—350 nm) of a given sample in lithium cacodylate (20 mM, pH 7.2) supplemented with chloride salts of a group 1 or 2 cation (100 mM), acquired at 95 °C and 4 °C.[40]

Isothermal difference spectra (IDS) were calculated by subtraction of the UV-spectra of a given sample acquired in absence or presence (100 or 0 mM for steady state measurements, and 5 mM added for time-dependent experiments) of monovalent or divalent cations, at 20 °C. IDS are not strictly identical to TDS because the absorbance spectra of the unfolded and folded species are temperature-dependent. IDS therefore better reflect the absorbance properties of the folded species. Time-dependent IDS were performed by addition of 5 mM KCl to a DNA sample monitored by a scan every 25 seconds (λ = 220—330 nm, 1.0 nm step, and 0.1 s integration time). Rate constants k were determined by non-linear fitting of the time evolution of the absorbance at 295 nm, with an exponential model, discarding the first point. The Arrhenius equation 9 where R is the gas constant and A is the pre-exponential factor, linearized to equation 10, was used to determine the activation energy E_a by plotting $\ln(k)$ versus $1/T$, where the slope is $-E_a/R$.

$$k = Ae^{-\frac{E_a}{RT}} \quad (9)$$

$$\ln(k) = \frac{-E_a}{R} \cdot \frac{1}{T} + \ln(A) \quad (10)$$

Similarly, the Gibbs free energy of activation is determined with the Eyring–Polanyi equation:

$$k = \frac{k_B T}{h} e^{-\frac{\Delta G^{\ddagger}}{RT}} \quad (11)$$

Where $\Delta G^{\ddagger}$ is the Gibbs energy of activation, k_B is Boltzmann's constant, and h is Planck's constant; linearized as follows (equation 12):

$$ln\frac{k}{T} = \frac{-\Delta H^{\ddagger}}{R} \cdot \frac{1}{T} + ln\frac{k_B}{h} + \frac{\Delta S^{\ddagger}}{R} \quad (12)$$

Where $\Delta H^{\ddagger}$ is the enthalpy of activation and $\Delta S^{\ddagger}$ is the entropy of activation. Plotting $\ln(k/T)$ versus $1/T$ gives access to $\Delta H^{\ddagger}$ and $\Delta S^{\ddagger}$ via the slope ($-\Delta H^{\ddagger}/R$) and intercept ($\Delta S^{\ddagger}/R + \ln(k_B/h)$).

Native electrospray mass spectrometry (ESI-MS). Samples were prepared by mixing oligonucleotides at a 10-µM strand concentration with 0—1 mM cation and TMAA (*qs* for a 120 mM salt concentration). Native ESI-MS spectra were obtained using an LCT Premier mass spectrometer (Waters, Manchester, UK). The ESI source voltage was set to 2.2 kV with a desolvation temperature of 60°C, and a sample cone voltage set at 200 V. The source pressure was adjusted at 45 mbar and monitored with a Center Two probe (Oerlikon Leybold Vacuum, Cologne, Germany). Injection was performed at 200 µl h^{-1}. Mathematical subtraction of the non-specific potassium adducts was performed as described by Gabelica *et al.*,[41] based on a method first described by Klassen *et al.*[42] This allows distinguishing potassium cation bound by the G-tetrads from those that are diffusely bound. Briefly, the integrals of the potassium cation distribution for a non-G4 forming sequence d[T($GCGT_2$)$_3$GCG]) are measured, then normalized to the zero-potassium peak, and finally used to reconstruct the specific potassium stoichiometry (see Figure S1 and detailed explanations in Supplementary section). The method was validated with 2-tetrad (22GT: d[G_3($TTAG_3$)$_3$T]) and 3-

tetrad (24TTG: d[$TTG_3(TTAG_3)_3A$]; c-myc Pu24: d[$TGAG_3TG_4AG_3TG_4AAG_2$]) G4-forming sequences.[41] Binding constants were determined with Dynafit 4.1 (Biokin, Watertown, MA, USA) *via* fitting by iterations the folding equation 7, decomposed in two distinct elementary equations 13 and 14 (where global $K_D = K_{D1} \times K_{D2}$).

$$unfolded + K^+ \rightleftharpoons quadruplex \cdot K^+ \quad (13a)$$

$$K_{D1} = \frac{[quadruplex \cdot K^+]}{[unfolded] \times [K^+]} \quad (13b)$$

$$quadruplex \cdot K^+ + K^+ \rightleftharpoons quadruplex \cdot 2K^+ \quad (14a)$$

$$K_{D2} = \frac{[quadruplex \cdot 2K^+]}{[quadruplex \cdot K^+] \times [K^+]} \quad (14b)$$

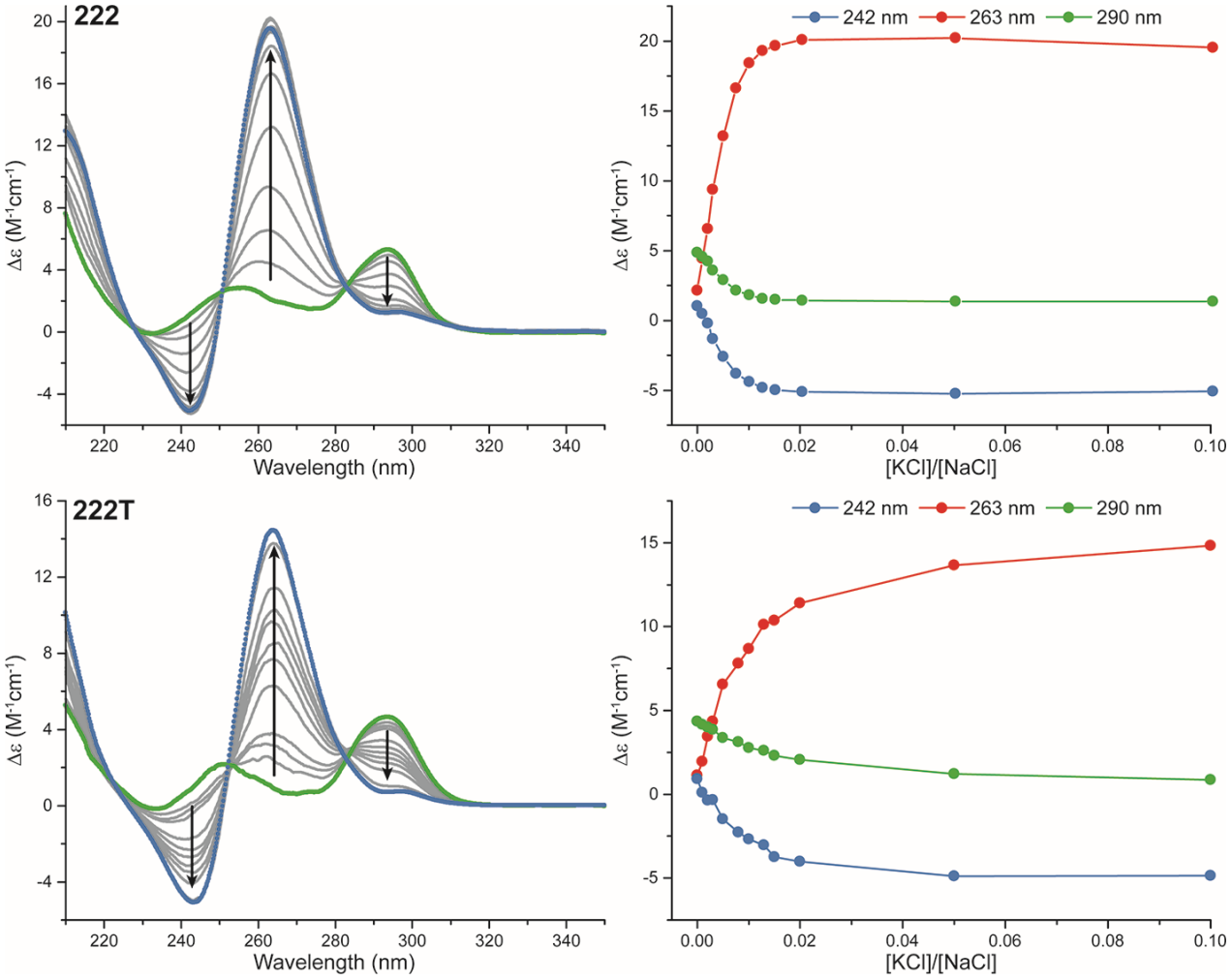


Figure 2: Left: CD titration of 10 µM 222 (top) and 222T (bottom) solutions containing NaCl (100 mM) by KCl (0—10 mM, from green to blue). Arrows indicate the potassium-dependent evolution of the characteristic bands. Right: Plots of the potassium-dependent evolution of the three characteristic bands.

3. RESULTS

3.1. A SODIUM-TO-POTASSIUM SWITCH

3.1.1. Steady-state characterization of sequences *222* and *222T*

Two sequences with promising cation-dependent topological switch properties, *222* (d[$(G_3T_2)_3G_3$]) and *222T* (d[$T(G_3T_2)_3G_3T$]), were identified by a circular dichroism-based screening, where we looked for large differences between the K^+ and Na^+ CD spectra (the 25 other sequences analyzed during this initial screen are presented at the end of part I). At a 10 μM strand concentration, both *222* and *222T* display an antiparallel signature in 100-mM sodium containing solutions (spectra characterized by positive bands around 295 and 250 nm, and a local minimum around 275 nm, arising from *anti*/*syn* or *syn*/*anti* guanine stacks), and a parallel signature in presence of 100 mM KCl (positive band at 260 nm, negative band at 240 nm, observed for *anti*/*anti* guanine stacks). Upon addition of increasing KCl concentrations (0—100 mM) to the 100-mM Na^+ containing samples, a clear conversion from the antiparallel to the parallel signature is observed, and is complete at a relatively low $[K^+]/[Na^+]$ ratio (5—10 % KCl; Figure 2), suggesting an efficient structural conversion and a preference for K^+ ions. In both cases, the presence of isodichroic points (250 and 282 nm) suggest a 2-state model, and hence a "switch" behavior. Singular value decomposition (SVD; see the experimental section for details) corroborates the presence of only two significant species at equilibrium (Figures S2 and S3). Experiments performed at 2.5 and 25 μM strand concentrations indicated that the conversion is DNA concentration-independent (data not shown).

^{1}H NMR experiments confirm that the sodium and potassium samples adopt distinct structures, and that the sodium-sample can be converted to the potassium fold by addition of 10 mM KCl (Figure S4). Interestingly, this topological switching is observed in less than ten minutes at room temperature. The time-dependence of this phenomenon is presented in more details in the following section.

Size-exclusion HPLC of 10 μM samples, eluted in 100 mM KCl or NaCl, or a mixture of NaCl and KCl (100/5 mM), indicates that the two oligonucleotides adopt an intramolecular structure in all the investigated conditions (Figure S5).[6] However, in order to account for the dilution of the samples during the course of the elution, and more generally investigate the oligonucleotides' behavior at high strand concentration, 250-μM samples are analyzed. A minor peak (11% relative intensity), corresponding to a dimer species, is detected for *222* only (Figure S6). The presence of flanking bases on G-rich sequences often inhibits the formation of higher-order structures, which may explain the absence of dimer for *222T*.[43]

The topology of the two sequences is assessed with a set of additional mono and divalent cations (NH_4^+, Rb^+, Li^+, Sr^{2+}, Ca^{2+}, Mg^{2+}), all from chloride salt sources. The CD signatures of solutions containing 100 mM of either of these cations are unique (for a given sequence), except for Li^+ and

Mg^{2+} that lead to similar spectra, probably because these cations do not promote the formation of G4s at 20 °C (*vide infra*) (Table S2 and Figure S7). A given cation does not necessarily give the exact same signature for both sequences although a similar trend is observed. Strontium seems to promote the formation of parallel structures as already observed for human telomeric sequences.[44] However, the presence of a small but significant positive band centered around 300 nm for *222T* indicates the presence of a small amount of antiparallel G4. This band is shifted as compared to what is observed with usual cations (290—295 nm), maybe because the larger Sr^{2+} cations distort partially the G4 (longer inter-tetrad distance), as proposed previously for TBA.[45,46] This very valuable feature allowing one to readily monitor the displacement of sodium by strontium cations will be further used below.

The formation of G4 is further confirmed for all these cations (except lithium and magnesium) by thermal difference spectra (TDS; Figures S8, S9 and S10 and Table S2). Karsisiotis and colleagues have suggested that the topology of a G4 can be inferred from the TDS,[47] using the $\Delta A_{240nm}/\Delta A_{295nm}$ ratio to distinguish parallel topologies from hybrid and antiparallel topologies. The use of TDS signature is however not optimal because (i) high melting temperatures (e.g. potassium or strontium) lead to decreased TDS intensities, and (ii) temperature-dependent changes in molar absorption coefficient contaminate the signature. We thus analyze IDS fingerprints instead, which essentially display the same characteristic bands than TDS but is performed at constant temperature (herein, 20 °C) thus avoiding temperature-dependent artefacts. For both oligonucleotides, potassium gives a three-fold higher $\Delta A_{240nm}/\Delta A_{295nm}$ ratio than sodium (Table S2 and Figure S11), corroborating the CD results and in agreement with previous reports.[47] Other cations inducing a parallel fold give similar ratios, such as Ca^{2+} with *222* or Rb^{+} with *222T*, whereas antiparallel or mixed topologies do not deviate significantly from the sodium values.

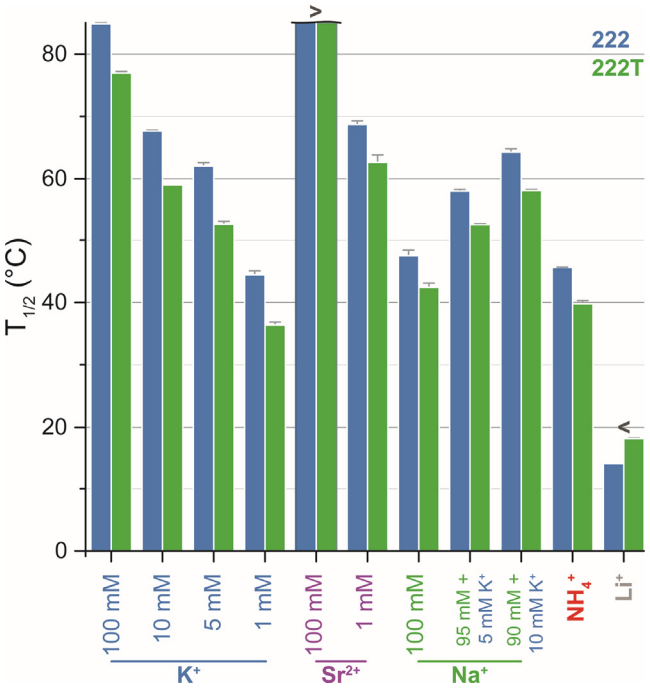


Figure 3: Cation-dependent $T_{1/2}$ values (°C) of 222 and 222T in lithium cacodylate (20 mM, pH 7.2), inferred from the melting curves presented in SI (Figures S12 and S13). Where necessary, the salt concentration is adjusted with lithium chloride (total salt concentration: 120 mM). Symbols > and < indicate that the precise T_m cannot be inferred from the melting curve because the transition is not entirely visible, and is thus respectively greater or smaller than indicated. $T_{1/2}$ values have been calculated as average of two experiments. Full results are shown in Table S3.

The thermal stability of the two sequences is investigated by UV-melting experiments, monitored at 295 nm, at a 10-µM strand concentration (Figure 3 and Figure S12, and Table S3).[48] In all cationic conditions described hereafter except Rb^+ and Mg^{2+}, *222T* is less stable than *222* by 5 to 10 °C, which is similar to the decrease in stability observed for some other sequences that have flanking bases.[49,50] Very high melting temperatures (T_m) are observed in presence of 100 mM KCl (85.0 and 77.0 °C, for *222* and *222T*), lowering with decreasing potassium concentrations (Figure S13). On the other hand, the T_m obtained in 100 mM of sodium only reaches 46.6 and 41.8 °C, respectively. This remarkable T_m difference (> 35 °C at 100 mM) certainly accounts for the topological switch properties of these two oligonucleotides. Furthermore, a mixture of 95 mM NaCl and 5 mM KCl stabilizes *222* and *222T* by 10 °C as compared to 100 mM NaCl alone. As no significant difference of T_m or $\Delta G°$ is observed when 95 mM LiCl substitute the sodium in the 95 mM NaCl/5 mM KCl sample, it is likely that the potassium cations displace entirely the sodium ones, rather than giving a mixed di-cation G4 [51]. Interestingly, the melting curve of *222* in 100 mM KCl is characterized by a slight inflection at lower temperatures (~60 °C; Figure S12), likely due to the presence of the small amount of dimer that was evidenced by SE-HPLC. Accordingly, the T_m of

222 decreases slightly at lower strand concentrations (5 and 1 µM, in 100 mM KCl), and the inflection in the curve disappears (Figure S14). Conversely, *222T* does not show any strand-concentration dependence, also in accordance with the size-exclusion chromatography results.

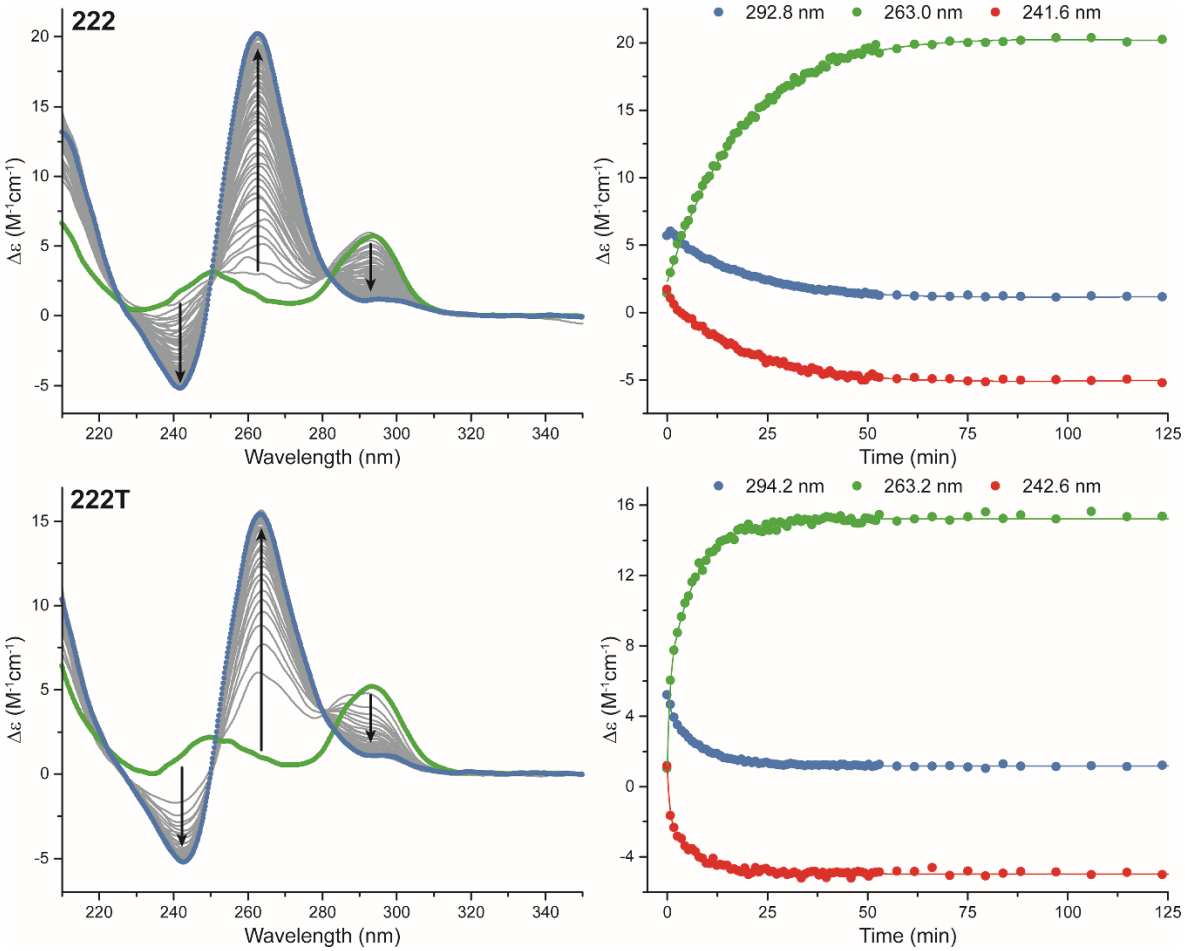


Figure 4: Left panels: time-dependent CD of 222 (top) and 222T (bottom) in 100-mM Na^+ (green) after addition of 5 mM KCl (final spectra: blue), at 4 °C. Arrows show the time evolution of the CD signal for characteristic bands. Right panels: plot of the time-dependence of three characteristic bands. Lines are obtained by global fitting with shared parameters using a double-exponential model.

The divalent Sr^{2+} cation induces the highest melting temperature (ΔT_m > 37 and 45 °C vs. Na^+). Lowering its concentration does affect the T_m, but it remains high at low concentrations (> 60 °C at 1 mM). In fact, very low strontium concentrations are sufficient to entirely fold the system (*vide infra*). Calcium is the only other tested cation that stabilizes *222* and *222T* significantly more than sodium ($\Delta\Delta G°$ ≈ 6 kJ/mol). As expected, lithium is the least efficient cation of our set. One cannot exclude that part or the totality of the G4 stabilization is provided by cation traces from the

oligonucleotide sample, rather than, or additionally to, lithium. It also shows that Li^+ should not be considered as a G4-destabilizing agent but rather as a G4 "indifferent" cation.

3.1.2. Cation-induced conformational switching kinetics

NMR experiments suggest that the structural switch is very fast at room temperature for both sequences. Indeed, at 20 °C, the addition of 5 mM of potassium induces an interconversion of 10-µM oligonucleotide solutions on the second timescale, *i.e.* faster than the CD measurement allows us to monitor. The following experiments are therefore carried out at 4 °C in order to slow down the topological conversion enough to be appreciable (Figure 4 and Figure S15). Plots at three characteristic wavelengths are fitted with a double exponential model (Tables S4), where the first phenomenon likely accounts for the quick exchange of sodium by potassium cations ($t_{1/2} \leq 1$ min, at 4 °C), as suggested for the human telomeric sequence [27]. After this fast initial exchange (green to gray curves), the set of spectra displays clear isodichroic points (250 and 280 nm) and it this thus not surprising that a global fitting with shared parameters at three different wavelengths can be successfully achieved. The interconversion seems faster for *222T* than *222* ($t_{1/2}$ = 314 vs. 771 s, respectively). A similar difference is observed when the conversion is triggered by only 1 mM KCl (Figures S15 and S16; relaxation times are ten-fold shorter for *222T* vs. *222*), although in this case it is not complete for *222T* (the 290 nm band intensity only lowers from roughly 50%), in accordance with the titration performed previously (Figure 2). Because of the fairly important dead time of the CD experiments, the cation exchange is largely imperceptible by SVD analysis, which consequently yields only two significant species (Figures S17 and S18). “Wallimann plots" (see experimental section) are dual wavelength parametric test of the two state model that were generated here by plotting the CD intensity at one wavelength vs. the intensity at another wavelength (Figure S19). Linearity of the plot typically indicates that a two state model is observed. All built plots are linear, with the notable exception of the first point, and this suggests that after the fast initial cation exchange, a two-state model is observed (Table S5).[37]

The time-dependence of the switch is further monitored by IDS, since, as mentioned earlier, small but appreciable and reproducible differences are observed between the potassium and sodium spectra (Figures S20 and S21, left, translated into IDS differences of 10% (for *222*), to 14.5% (for *222T*) Figures S20 & S21, right panels). Hence, the conversion is followed by UV-vis spectroscopy, at temperatures ranging from 1 to 10 °C. This is to our knowledge the first report of a structural G4 interconversion monitored by UV-absorbance, illustrating the use of IDS for the analysis of subtle structural transitions.

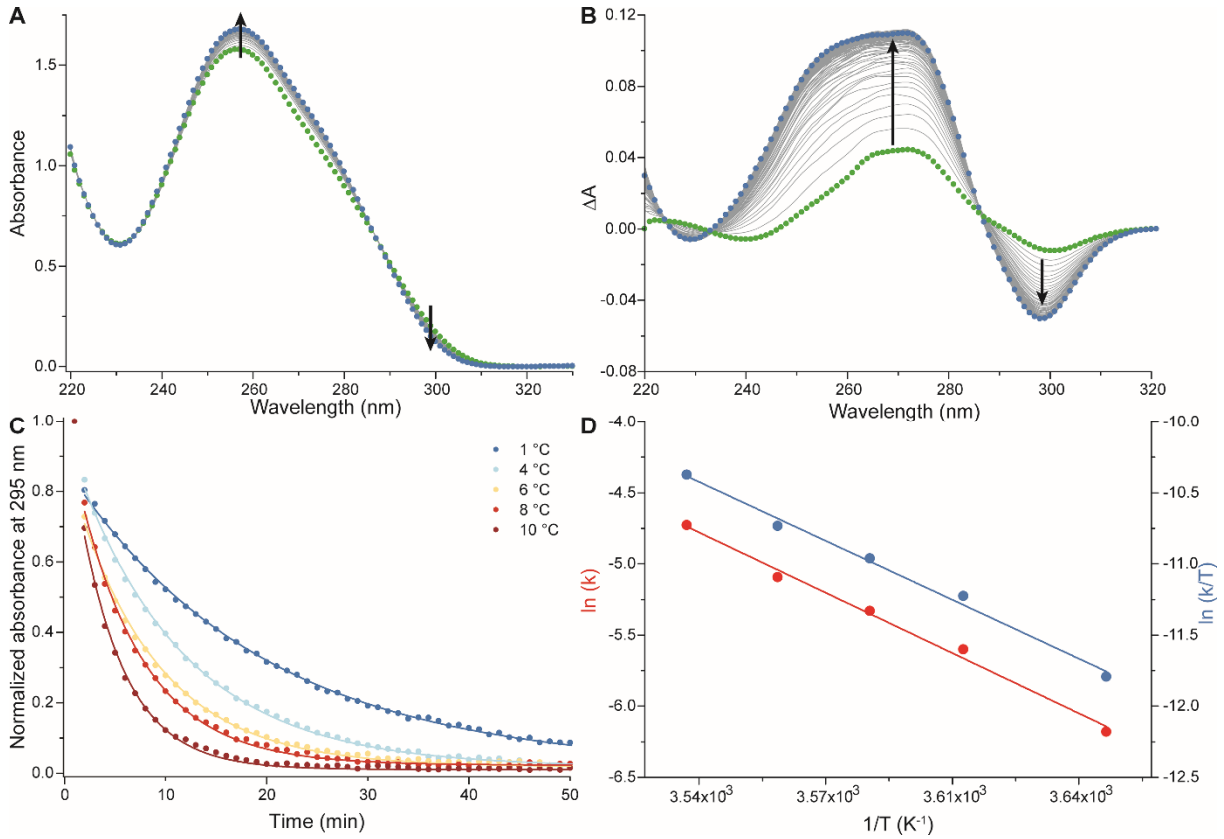


Figure 5: Time-dependent UV-vis spectra of 22*2*T in 100-mM Na^+ (green) after addition of 5 mM KCl (final spectra: blue), at 4 °C (A), the corresponding IDS signatures (the first IDS shown, in green, is obtained after addition of potassium and 15 seconds experimental dead time) (B), the time-evolution of the absorbance at 295 nm at various temperatures (C), and plots verifying the Arrhenius (red; R^2 = 0.99048) and Eyring-Polanyi equations (blue; R^2 = 0.99002) in this temperature range (D).

Corresponding reaction rates are extracted from the time evolution of the absorbance at 295 nm, a characteristic wavelength for the study of G4 structures (Figure 5, Figures S20, S21 and S22).[40] The energy of activation is determined from the Arrhenius equation in this temperature range (E_a = 92 ± 6 and 101 ± 5 kJ/mol, respectively for *22*2 and *22*2*T*) and is almost identical to the enthalpy of activation obtained from the Eyring-Polanyi equation ($\Delta H^{\ddagger}$ = 90 ± 6 and 98 ± 5 kJ/mol). The use of alternative wavelengths (230 nm for *22*2 and 250 nm for *22*2*T*) gives results lying within the standard deviation: E_a = 87 ± 6 and 102 ± 6 kJ/mol, $\Delta H^{\ddagger}$ = 84 ± 6 and 102 ± 6 kJ/mol (Figure S23). As expected from the previous observation that the rate of reaction increases with the temperature, the activation energy is positive. It can also be noted that the IDS profiles of *22*2 and *22*2*T* are radically different, suggesting that these two oligonucleotides adopt somewhat different structures in a given cationic environment, despite having very close sequences and CD signatures. This may account for the differences observed in PAGE migration (not shown), thermal stability, rate of

reactions, NMR spectra, SE-HPLC relative elution volumes, and potassium binding capabilities. The dimer of 222 could in principle also account for this difference, but this is less likely since only small amounts are detected by SE-HPLC.

Strontium is also particularly efficient at stabilizing the G4 structures formed by *222* and *222T*. However, addition of 5 mM $SrCl_2$ in 100 mM-Na^+ containing *222* or *222T* solutions results in a much *slower* interconversion at 20 °C (i.e. around 60 minutes vs. < 1 min with KCl) (Figures S24). Interestingly, strontium provides insights into the structure interconversion mechanism thanks to the CD band shift described previously, which allows the deconvolution of the displacement of Na^+ by Sr^{2+}, and the topological conversion *per se*. Within a minute after addition of Sr^{2+} to the Na^+-containing sample, a clear 295 to 300 nm band shifting is observed with only a little change in intensity at 260 nm. This suggests that sodium is first displaced by strontium, and the structural conversion takes place after. This does not necessarily prove that a similar mechanism applies for the sodium/potassium exchange, but is in line with the mechanism proposed so far for the human telomeric sequence.[27]

3.1.3. Sodium-to-potassium structural switching is not an isolated case

To explore the origins of the differences in stability and dimerization of *222* and *222T*, the variant sequences *222-5'T* and *222-3'T*, where a single flanking thymine is present, in 5' or 3' respectively, were investigated. For both sequences, a sodium-potassium conversion is still observed by addition of 5 mM KCl, although it is not entirely complete in the case of the 5'T-variant (Figure S25). Faster kinetics are observed in the latter case, and a 2-state model can be inferred from the SVD (again with the notable exception of the first spectrum) (Figures S26 and S27). Melting curves of these two sequences in presence of 5, 10, or 100 mM KCl, or 100 mM NaCl, reveal interesting information (Figure S28). First, *222* and *222-3'T* display similar stabilities and are more stable than *222T* and *222-5'T*, which also share comparable T_ms. The 5'-dT nucleotide therefore destabilizes the G4, while the 3'-dT does not affect the stability. Second, an inflection of the melting curves at low temperature can be observed for *222-5'T*, similarly to *222,* but not for *222-3'T* (and *222T*), suggesting that the 3'-dT base inhibits the dimer formation. This is further confirmed by SE-HPLC experiments performed at a 250 µM strand concentration, where *222-5'T* presents a minor dimer peak (Figure S29). Overall, these results suggests the formation of a stacked dimer by *222* and *222-5'T*, whereas *222T* and *222-3'T* exist solely as monomers.

Twenty-five additional sequences extracted from the literature (Table S6), either synthetic constructs or biologically relevant sequences, are screened by CD, by comparing spectra obtained in 100 mM KCl, 100 mM NaCl, and after addition of 100 mM KCl in the sodium samples (Figure S30). These sequences contain at least four repeats of a minimum of three guanines, except H-Bi-G4, which forms a bimolecular G4. All are likely to form three-tetrad G4s. All the investigated

sequences do display distinct signatures in Na^+ and K^+ conditions. Antiparallel signatures are obtained in sodium conditions only, while parallel or hybrid/mixtures spectra are observed regardless of the cation nature, suggesting that the antiparallel, 3-tetrad topology is disfavored in potassium. Most of the sequences show a preference for potassium over sodium, as the spectrum obtained under mixed conditions is often indistinguishable from the one in potassium only, in accordance with the usual higher stability observed for G4 in K^+ *vs*. Na^+. Additional sequences were then prepared in 100 mM NaCl and titrated by KCl (0—10 mM) (Figures S31 and S32). They are characterized by a clear antiparallel-to-parallel (*331*, *133*, *232*, *223*, and *322*), hybrid-to-parallel (*161*), and antiparallel-to-hybrid switch (*G5T2*), or no visible switch (*161-T* and *161-TT*). Isodichroic points can be observed in most cases, supporting a "pseudo" two-state model (Na^+ and K^+ topologies), but none of these sequences interconvert completely with potassium concentrations as low as observed with *222*/*222T*. The couple *331*/*133* and *223*/*322* have mirror-image sequences, but somewhat different behaviors, highlighting the importance of strand polarity in G4 folding. SE-HPLC demonstrates that, at low strand concentration, these oligonucleotides form intramolecular G4s in various cationic conditions, albeit for *161* and *G5T2* that also partly fold into intermolecular species (Figure S33). Variants of *161*, namely *161-T* and *161-TT*, were designed to form exclusively intramolecular G4s thanks to the additions of flanking thymines.[6,52] They however do not exhibit a clear topological transition anymore, since both the potassium and sodium forms are now parallel. The same remark applies to *222-TT*, although to a lesser extent as it co-exists as a mixture of parallel and antiparallel G4s when stabilized by sodium (Figure S31).

In conclusion, several different sequences could potentially serve as switches, and combining some with different relative stabilities could eventually expand the range of cation concentrations at which these switches could operate. Adding flanking bases to the G4 core helps inhibiting multimer formation, but it is also important to take into account that flanking bases also affect the topology of intramolecular structures.

3.2. POTASSIUM CONCENTRATION SWITCH

In part I, we have explored the structural conversion of two sequences as a function of the cation nature. Hereafter, we study the behavior of these sequences upon increasing cation concentration, starting from an unfolded strand in presence of a cation "hostile" to quadruplex formation (TMA). Hence, the topology of *222*/*222T* oligonucleotides as a function of the concentration in potassium (0—100 mM) is studied in the presence of various concentrations of TMAA buffer to ensure a constant salt concentration (120 mM). TMA is too bulky to be coordinated by the O6 of tetrad guanines and is therefore a cation of choice to fix the ionic strength and neutralize the backbone phosphates while allowing one to study the specific binding of potassium (or other cations) to the tetrads. Tetrabutylammonium phosphate was used for the same purpose by Gray and Chaires [25],

but TMAA has the advantage of being a volatile electrolyte compatible with mass spectrometry experiments [53]. This approach fundamentally differs from the data presented in Part I, in that we are starting here from an unfolded strand rather than from a different quadruplex fold (compare the CD spectra obtained in zero KCl in Figures 2 and 6).

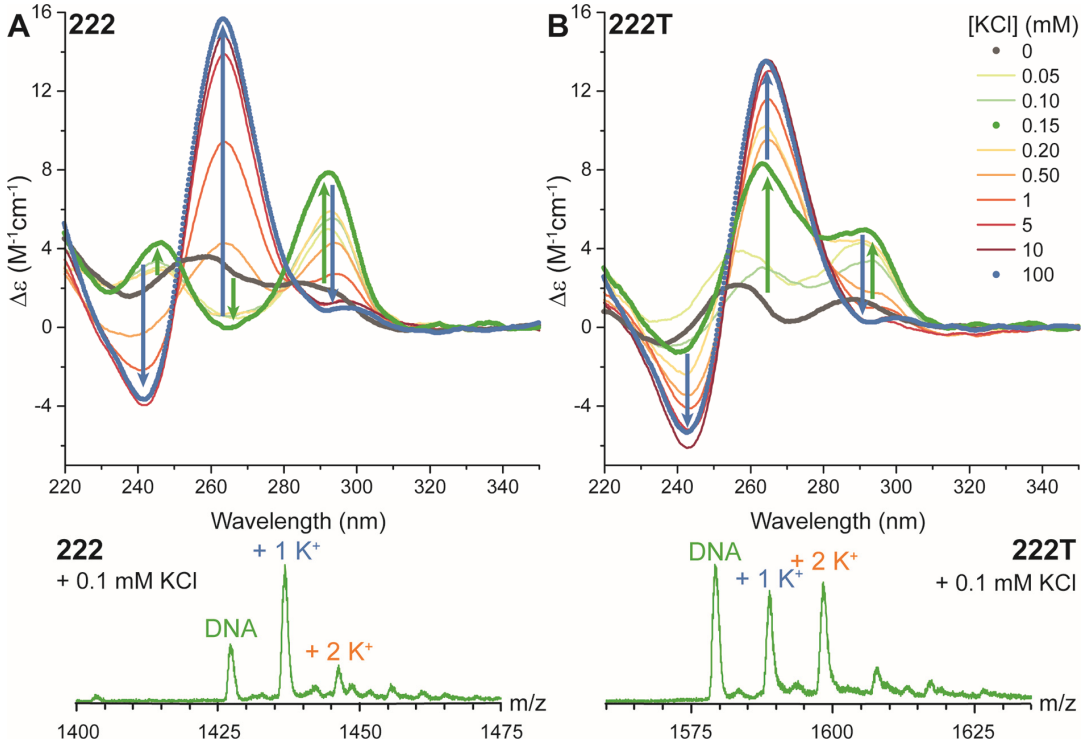


Figure 6: Top panel: Titration of A) 222 and B) 222T (10 μM) in TMAA (concentration adjusted to keep the total salt concentration constant at 120 mM) by KCl (0—100 mM). Spectra acquired in presence of 0 and 100 mM KCl are colored in dark gray and blue, respectively. The spectra with the highest molar ellipticity at 295 nm is colored in green. The green arrows indicate the appearance of an antiparallel signal at low KCl concentration (< 200 μM), while the blue arrows show the rise of the expected parallel fold at higher potassium concentrations (0.2—100 mM) with the concomitant disappearance of the 295 nm antiparallel signal. Bottom panel: Corresponding ESI-MS spectra for the 0.1-mM KCl samples, before correction for diffuse binding, focused on the 4- charge state (DNA alone and bound to one or two potassium cations). Full spectra are reported in Figure S35.

3.2.1. Steady-state characterization: at low KCl concentration, *222* forms an antiparallel, 2-quartet quadruplex

As expected, TMAA does not promote the formation of G4 in the absence of other cations (**Figure** 6A-B, dark gray spectra). Surprisingly, low concentrations of K^+ (e.g. 0.05—0.2 mM for *222*) promotes an "antiparallel" fold for 222 (Figure 6A, green arrows), rather than the expected "parallel" fold described in the first part of this manuscript. The parallel structure is eventually obtained upon further increasing the KCl concentration (blue arrows) (see also Figure S34). Hence, at a concentration of 150 µM in KCl, the CD signature of *222* is purely antiparallel. The presence of flanking bases tends to promote the parallel topology *vs*. the antiparallel one, and hence the corresponding solution of *222T* contains a mixture of antiparallel and parallel fold at 150 µM KCl. The titration was also monitored by mass spectrometry (ESI-MS) in order to detect and quantify the number of potassium ions specifically bound to the tetrads (Figures 6, bottom, S35 and S36). The potassium binding stoichiometry gives a hint regarding the number of tetrads since there are usually one more tetrad than octa-coordinated cations.[13] A control sequence was titrated to estimate the amount of diffusively bound cations (Figure S37 and Table S7), and to mathematically subtract that contribution from the experimental spectra of *222* and *222T* in order to extract the fraction of specifically bound complex (Figure 7).

At 100 µM KCl (dashed line), *222* mostly binds one potassium cation, whereas *222T* is a mixture of populations that bind either zero, one, or two potassium, in similar proportions (Figures 6-7). This is reflected in CD by a pure antiparallel signature for *222*, and a parallel/antiparallel mixture for *222T*. In both cases, a fraction of the populations do not bind potassium, which is consistent with UV-melting experiments indicating that these sequences are not fully folded at 20 °C in presence of less than 500 µM KCl (Figure S38 and S39). ESI-MS shows that as much as 80% of the population of *222* binds a single cation at 100—200 µM potassium concentrations, whereas it never exceeds 20% for *222T* (Figure 7), which is also qualitatively observed by CD. This translates into a K_D of the second binding site being lower than the first binding site one for *222T*, but not for *222*. Note that, at 1 mM in KCl, *222* partly forms a dimer, in which the binding of five potassium cations suggests a 6-tetrad structure compatible with the stacking of two monomers, as suggested by SE-HPLC and UV-melting experiments conducted in potassium-rich conditions (Figure S35). Moreover, the ^{1}H NMR spectra obtained at low potassium concentrations are significantly different than the one obtained from potassium-rich samples, confirming the presence of a distinct set of species (Figure S4).

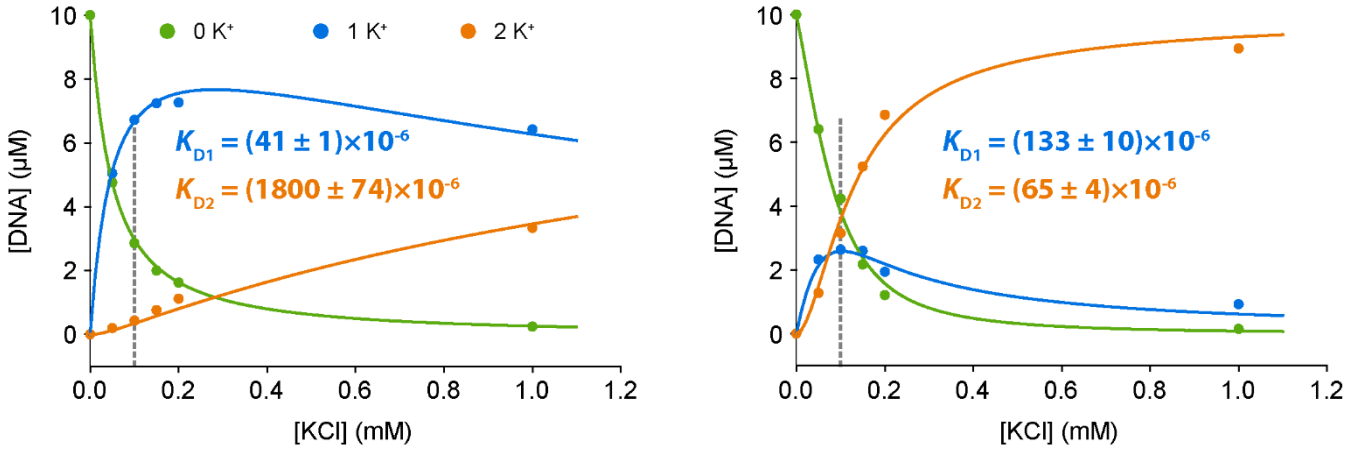


Figure 7: Concentration of the unfolded (0 K^+), 1 K^+ and 2 K^+ bound G4 versus the concentration of potassium chloride, as determined by ESI-MS. Lines are the result of a non-linear fitting procedure based on chemical equations, and resulting sequential K_D's are indicated on the respective graphs. The binding or 3 or more K^+ is negligible after mathematical subtraction of non-G4-specific adducts.

Further analysis of the melting data also hints at the presence of distinct species at low potassium concentrations, differing by the number of bound cations. For a given oligonucleotide, the Gibbs free energy of folding at 20 °C are fairly similar at 5, 10, and 100 mM (around -56 and -50 kJ/mol for *222* and *222T*, respectively), but increases significantly at 1 mM and lower ($\Delta\Delta G° \approx 7$ and 4 kJ/mol), consistent with the presence of another folding equilibrium producing a less stable quadruplex (Table S3; Figure S40). We hypothesized that the antiparallel structure observed at lower potassium concentration, coordinating to only one potassium cation, may only contain two tetrads, whereas the more stable parallel species contains three tetrads, as expected from the sequence. An alternative hypothesis for the 1-K^+ binding species would involve the folding of the oligonucleotide into a 3-tetrad G4 where one cation binding site is empty. This seems unlikely since the coordination of cations in the electronegative cavities of G4 is usually regarded as the folding driving force.[25]

3.2.2. Time-dependence of the concentration-based conformational switch

When 5 mM potassium chloride is added to a *222*/*222T* solution in TMAA only, maximum folding of the antiparallel species is observed after one minute at 20 °C, before converting slowly to the expected parallel structure (Figure 8 and Figure S41). At 4°C, the transformation is slower and the presence of the antiparallel structure is more noticeable, most notably for *222*. In both cases, the conversion is visibly faster for *222T*. Time-dependent UV-vis experiments, performed at 10 °C, also highlight the relative slow conversion rate compared to the Na^+ to K^+ switch (9 and 23 times times

slower, for *222* and *222T* respectively), and the longer lifetime of the antiparallel species for *222* (Figures S42 and S43). In absence of another cation (except TMA), the antiparallel 1-K^+ bound G4 acts either as a folding intermediate or as a kinetic trap. The latter case seems likely given the long-lived non-equilibrium folding state observed, relatively to the much faster Na^+ to K^+ topological conversion described above. Formation of 2-tetrad G4s requires only two partners (the oligonucleotide and one potassium cation), and can be achieved by a variety of loop combinations [54], whereas a canonical parallel three-tetrad G4 can only be obtained by a single loop combination (all double chain reversal), and requires an extra partner (an additional potassium cation), which is kinetically more complex. The difference of stability between the antiparallel and parallel structures is smaller for *222T* than *222*, as seen from ESI-MS and UV-melting experiments, which may account for the greater transience of this fold for the former sequence.

Strontium also possibly induces a 2-tetrad antiparallel G4, at submillimolar concentrations as observed by CD (Figures S44) and ESI-MS (Figure S45). Surprisingly low amounts of strontium are required to fold *222* and *222T* into an antiparallel G4, with full folding being readily achieved at a 50 µM concentration (50% folding is observed with a stoichiometric amount of Sr^{2+}). Therefore, formation of a structure including a second strontium ion may be delayed by this stable competitor. Alternatively, it is possible that the formation of 3-tetrad structures involve one empty coordination site, in order to minimize the electrostatic repulsion. The coordination of Sr^{2+} ions in every other coordination site has been observed in crystals of dimers of $r(UG_4U)_4$. [55] Unfortunately, the mathematical subtraction of non-specific strontium adducts is not possible due to the clear preferential outer-sphere coordination of strontium to Hoogsteen faces of guanines,[56] leading to a significant amount of binding of strontium to the non-G4 forming control sequences (where all Hoogsteen faces of guanines are available), and therefore preclude any signal deconvolution and subsequent quantification.

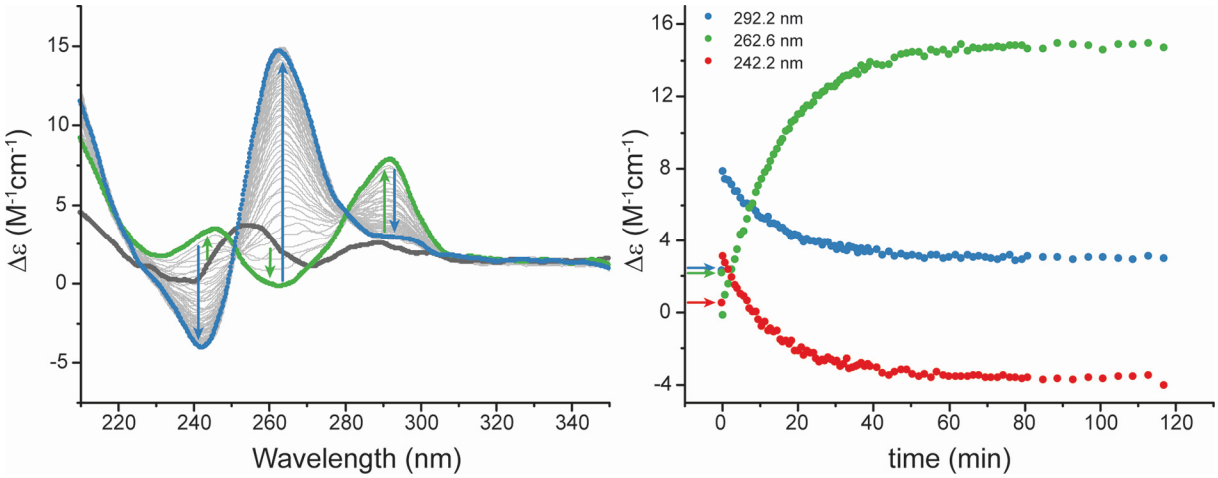


Figure 8: Time-dependent CD of *222*, prepared in TMAA with no other cation (dark grey) after addition of 5 mM KCl, at 20°C. An antiparallel intermediate (green) is observed after one minute, then converts to a parallel signal (final spectrum: blue). Right: Corresponding plots of the normalized ellipticity at three characteristic wavelengths, where arrows highlight the initial point at t=0.

4. DISCUSSION

222T and *222* can adopt at least four distinct sets of structures: unfolded, Na^+-bound antiparallel, 1-cation/2-tetrad antiparallel, and 2-cation/3-tetrad parallel. This switching system is complex enough to perform advanced functions, yet more tractable than the human telomeric sequences because it favors a parallel conformation rather than a range of hybrid conformations. Switching between the conformations is triggered by cation nature and cation concentration (Figure 1). *222* and *222T* display an impressive Na^+ to K^+-dependent topology switch, with high conversion rates being observed with small amounts of potassium. Overall, the results support a 2-step mechanism where the sodium is first displaced by potassium, followed by the topological conversion *per se*. No other long-lived intermediate is identified. Our results highlight the significant effect of flanking residues on the topology, stability, and folding/conversion kinetics of G4s. The formation of dimers can be readily controlled by tuning the flanking sequences and the strand concentration. On the one hand, *222* requires less potassium than *222T* (1 vs 5 mM K^+) to switch, and can partly form a dimer, whereas on the other hand *222T* exhibits higher conversion reaction rates and folds uniquely into monomeric G4s. The sequence can therefore be fine-tuned depending on the exact functionalities intended for the switch.

Overall, the stabilization of these sequences by cations follows the usual trend (Sr^{2+} > K^+ > Ca^{2+} > Na^+, NH_4^+, Rb^+ > Mg^{2+}, Li^+), but the impressive T_m difference between potassium- and sodium-containing samples account probably for the interconverting properties. Generally, the G4 stability

in KCl solutions is inversely proportional to the loop length, with short loops favoring parallel structures.[43,57] The average melting temperature of libraries of G4s having at least two bases (A or T) in all of their loops is 65 ± 4 °C in 150 mM KCl, lower than what is found for *222* and *222T* in 100-mM KCl solutions (> 75 °C).[43] In 150 mM NaCl, libraries of G4 with at least 2-nucleotide loops are on average antiparallel and have an average melting temperature of 51 °C, in line with what is reported herein. Note also that all cations leading to the most stable structures (Sr^{2+}, K^{+}, Ca^{2+}) mainly or entirely promote the formation of parallel G4s, and that the addition of flanking bases seems to disfavor the antiparallel fold. Finally, although strontium leads to higher melting stabilities, the observed interconversion rate is significantly slower.

222 and *222T* can also act as a potassium/strontium-concentration dependent switch. Sub-millimolar concentrations of these cations promote an antiparallel topology, while higher concentrations lead to a parallel topology. The antiparallel fold is more favored in the case of *222* than *222T*. Furthermore, on the basis of ESI-MS and UV-melting experiments, we propose that this alternative structure contains only two tetrads. The folding in a 2-tetrad G4 is likely explained by (i) the presence of low concentration of free cation, and (ii) the relative stability of this fold that is dependent on guanine orientations. At low potassium concentration, there is little free potassium available to populate the binding sites offered by G4s, and a 2-tetrad structure may consequently be favored. Moreover, the folding into 2-tetrad structures is entropically favored since there are a higher number of accessible folds.[54] Cang *et al*. have shown by free energy analysis of molecular dynamics experiments that an all *syn*/*anti* stack-containing topology is more favorable than the *anti*/*anti* counterpart by 15 kcal/mol, for 2-tetrad structures.[58] The antiparallel structure may then be favored for the 2-tetrad *222* and *222T* structures because it enables to populate uniquely the more stable *syn*/*anti* step. At higher concentrations, enough free potassium is available to displace the equilibrium towards a 3-tetrad structure, where the parallel topology seems to be energetically preferred. *Anti*/*syn* and *syn*/*anti* guanine stacking are not geometrically – and hence energetically – identical ($\Delta E_{syn/anti} < \Delta E_{anti/syn}$).[59] Consequently, the oligonucleotide might adopt a parallel topology to avoid the unfavorable *anti*/*syn* (or *syn*/*syn*) step that would arise from an additional tetrad following an antiparallel arrangement. Interestingly, a number of other 2-tetrad G4s found in potassium solutions do actually adopt an antiparallel fold (e.g. TBA [60], *Bombyx mori* telomere [61], HIV-PRO1 [62], 21CTA [63]). On the other hand, to our knowledge no intramolecular, potassium-stabilized, 3-tetrad, 2+2 antiparallel G4 structure were solved by X-Ray diffraction or NMR so far,[59] which is in line with our observations on a number of sequences. Furthermore, libraries of short-looped G4 forming sequences have been shown to predominantly adopt parallel topologies at high K^{+} concentrations.[43]

222-derived sequences are very efficient topological Na^{+} to K^{+} switches, and they are not isolated cases in the G4–forming sequence space. We demonstrated that other sequences display similar

properties, albeit with lower potency. Overall, this suggests that a very specific couple of sequences exhibiting remarkable properties have been identified, but also that equally interesting systems could be unearthed. Using BLAST [64,65], the *222T* sequence in particular (and *a fortiori* the *222* one, albeit with a higher probability to be found owing to its shorter sequence) is found at different locations in the human genome, either in transcribed or untranscribed regions (Table S8). By screening the Gene Expression Omnibus (GEO) database,[66,67] one significant hit was uncovered for the regulation of Beta-secretase 2 (BACE2 gene; E = 0.002) that may be involved in Alzheimer's disease.[68] *222T* was also found in diverse bacterial organisms (34 hits) including pathogenic species such as *Neisseria meningitides* (*meningococcus*; 16 consecutive 5'-TG_3T repeats in the gene NMB1261 of the *Neisseria meningitidis* MC58 chromosome coding for a type III restriction-modification system EcoPI enzyme), *Neisseria gonorrhoeae* (*gonococcus*), and the Avian Pathogenic *Escherichia coli* strain APEC O78. It is a fairly highly repeated sequence in *Paenibacillus* strains such as *Paenibacillus sp*. FSL H7-0737 (thirty-five 5'-TG_3T repeats constituting most of the H70737_RS03380 gene), *Paenibacillus sp*. FSL P4-0081 (sixteen), *Paenibacillus odorifer* strain DSM 15391 (fifteen), and *Paenibacillus sp.* FSL R7-0273 (twelve repeats upstream from an amino acid permease encoding gene). The sequence *222T* is thus a potential biological target whose structural versatility could be taken advantage of for specific gene regulation, considering the relatively low number of hits for a sequence that short. As we found other synthetic sequences that also respond in a switch manner, it is extremely likely that several other biologically relevant sequences will do so as well. This relatively simple system may also provide clues on the structural heterogeneity of more complex / more polymorphic cases such as a the human telomeric quadruplex.

5. CONCLUSIONS

In summary, the salient points of this work are:

(i) We extensively characterized two DNA G4 topological double switch sequences that can act as a simple and inexpensive multiple-state switching device with putative application in nanotechnologies (nanoarchitecture, nanomachines) and biology.[8-11] The triggering rate and outcome can be readily controlled by a number of parameters including the cations' nature and concentration, the strand concentration and sequence, and the temperature. For instance, FRET-tagged oligonucleotides are fast and sensitive reporters of the presence of low potassium amounts in sodium-rich solutions (Figure S46). In the same vein, since the first report about the stabilization of quadruplex structures by Sr^{2+} cations,[69] G4-forming oligonucleotides have been used as probes for the detection of these toxic cations.[70,71] Finally, they may also be used to study the binding properties of ligands (small molecules, proteins) in terms of preferred topology,[72,73] number of tetrads,[72] oligomerization,[74] and folding induction.[75]

(ii) The study of G4-forming sequences in presence of low cation concentrations lead to the discovery of unexpected folding behavior, such as a 2-tetrad antiparallel structure formed at low concentrations of potassium. The study of other sequences in such conditions is currently ongoing and may allow to better understand the folding of G4 nucleic acids, and notably whether the 2-tetrad structure is a folding intermediate or a competitive structure.

(iii) More generally, a number of the observations reported here provide insight into sequence and cation dependent folding of quadruplex nucleic acids, including for a small library of interconverting oligonucleotides that go beyond the heavily studied human telomeric and TBA sequences. These are sequences of choice for studying the cation-dependent folding of quadruplexes, in terms of preferred topologies, loop sequences, flanking bases, and multimerization. Many are also putative biological targets within the human genome or human pathogen genomes.

(iv) The sequences presented here may also be valuable benchmarks for the development of analytical methods aimed at studying G4 structures, as exemplified herein with the use of an original method (time-dependent isothermal difference UV absorption spectra, or IDS).

ASSOCIATED CONTENT

Supporting Information. Additional experimental procedures (diffusively bound potassium cleaning in native ESI-MS experiments), data, and extra figures. This material is available free of charge via the Internet at http://pubs.acs.org.

AUTHOR INFORMATION

Corresponding Author

* jean-louis.mergny@inserm.fr.

Author Contributions

The manuscript was written through contributions of all authors. All authors have given approval to the final version of the manuscript.

Funding Sources

This research was funded by the Agence Nationale de la Recherche (OligoSwitch [ANR-12-IS07–0001] 'Quarpdiem' [ANR-12-BSV8–0008–01] & 'VIBBnano' [ANR-10-NANO-04–03]), the Conseil Régional Aquitaine (Grant 20121304005 to V.G.), the EU (FP7-PEOPLE-2012-CIG-333611 to V.G.) and INSERM (ATIP-Avenir Grant R12086GS to V.G.).

ACKNOWLEDGMENT

We gratefully acknowledge Marie Toulisse and Miho Takematsu for their kind technical assistance, Frédéric Rosu at the structural biophysico-chemistry platform of the *Institut Européen de Chimie et Biologie* for the access to mass spectrometers and his invaluable support, Dr. Amina Bedrat, Dr. Valentina D'Atri and Dr. Marleen Renders for fruitful discussions.

ABBREVIATIONS

CD: Circular Dichroism; ESI-MS: Electrospray Ionisation Mass Spectrometry; FRET: Förster Resonance Energy Transfer; G4: quadruplex; IDS: Isothermal Difference Spectra; NMR: Nuclear Magnetic Resonance; SE-HPLC: Size Exclusion High-Performance Liquid Chromatography; SVD: Singular Value Decomposition; TBA: Thrombin Binding Aptamer; TDS: Thermal Difference Spectra; TMA: Tri Methyl Ammonium; TMAA: Tri Methyl Ammonium Acetate.

REFERENCES

(1) Choi, J.; Majima, T. *Chemical Soc. Rev.* **2011**, *40*, 5893.

(2) Cer, R. Z.; Donohue, D. E.; Mudunuri, U. S.; Temiz, N. A.; Loss, M. A.; Starner, N. J.; Halusa, G. N.; Volfovsky, N.; Yi, M.; Luke, B. T.; Bacolla, A.; Collins, J. R.; Stephens, R. M. *Nucleic Acids Res.* **2013**, *41*, D94.

(3) Patel, D. J.; Phan, A. T.; Kuryavyi, V. *Nucleic Acids Res.* **2007**, *35*, 7429.

(4) Dai, J.; Carver, M.; Yang, D. *Biochimie* **2008**, *90*, 1172.

(5) Le, H. T.; Miller, M. C.; Buscaglia, R.; Dean, W. L.; Holt, P. A.; Chaires, J. B.; Trent, J. O. *Org. Biomol. Chem.* **2012**, *10*, 9393.

(6) Largy, E.; Mergny, J. L. *Nucleic Acids Res.* **2014**, *42*, e149.

(7) Parkinson, G. N. In *Guanine Quartets: Structure and Application*; Spindler, L., Fritzsche, W., Eds.; Royal Society of Chemistry: Cambridge, UK, 2012, p 237.

(8) Davis, J. T. *Angw. Chem. Int. Ed.* **2004**, *43*, 668.

(9) Zhou, J.; Amrane, S.; Korkut, D. N.; Bourdoncle, A.; He, H. Z.; Ma, D. L.; Mergny, J. L. *Angew. Chem. Int. Ed.* **2013**, *52*, 7742.

(10) Tintoré, M.; Eritja, R.; Fábrega, C. *ChemBioChem* **2014**, *15*, 1374.

(11) Alberti, P.; Bourdoncle, A.; Saccà, B.; Lacroix, L.; Mergny, J.-L. *Org. Biomol. Chem.* **2006**, *4*, 3383.

(12) Hud, N. V.; Plavec, J. In *Quadruplex Nucleic Acids*; Neidle, S., Balasubramanian, S., Eds.; Royal Society of Chemistry: Cambridge, UK, 2006, p 301.

(13) Engelhart, A. E.; Plavec, J.; Persil, Ö.; Hud, N. V. In *Nucleic Acid-Metal Ion Interactions*; Hud, N. V., Neidle, S., Eds.; Royal Society of Chemistry: Cambridge, UK, 2009, p 433.

(14) Burge, S.; Parkinson, G. N.; Hazel, P.; Todd, A. K.; Neidle, S. *Nucleic Acids Res.* **2006**, *34*, 5402.

(15) Shannon, R. D. *Acta Crystallogr. A.* **1976**, *32*, 751.

(16) Hardin, C. C.; Watson, T.; Corregan, M.; Bailey, C. *Biochemistry* **1992**, *31*, 833.

(17) Włodarczyk, A.; Grzybowski, P.; Patkowski, A.; Dobek, A. *J. Phys. Chem. B* **2005**, *109*, 3594.

(18) Venczel, E. A.; Sen, D. *Biochemistry* **1993**, *32*, 6220.

(19) Kankia, B. I.; Marky, L. A. *J. Am. Chem. Soc.* **2001**, *123*, 10799.

(20) Hardin, C. C.; Perry, A. G.; White, K. *Biopolymers* **2000**, *56*, 147.

(21) Hud, N. V.; Smith, F. W.; Anet, F. A.; Feigon, J. *Biochemistry* **1996**, *35*, 15383.

(22) Gu, J.; Leszczynski, J. *J. Phys. Chem. A* **2002**, *106*, 529.

(23) Misra, V. K.; Draper, D. E. *Biopolymers* **1998**, *48*, 113.

(24) Martín-Hidalgo, M.; Rivera, J. M. *Chem. Commun.* **2011**, *47*, 12485.

(25) Gray, R. D.; Chaires, J. B. *Biophys. Chem.* **2011**, *159*, 205.

(26) Sen, D.; Gilbert, W. *Nature* **1990**, *344*, 410.

(27) Gray, R. D.; Li, J.; Chaires, J. B. *J. Phys. Chem. B* **2009**, *113*, 2676.

(28) Gray, R. D.; Trent, J. O.; Chaires, J. B. *J. Mol. Biol.* **2014**, *426*, 1629.

(29) Buscaglia, R.; Gray, R. D.; Chaires, J. B. *Biopolymers* **2013**, *99*, 1006.

(30) Gray, R. D.; Chaires, J. B. *Methods* **2012**, *57*, 47.

(31) Gray, R. D.; Buscaglia, R.; Chaires, J. B. *J. Am. Chem. Soc.* **2012**, *134*, 16834.

(32) Miller, M. C.; Le, H. T.; Dean, W. L.; Holt, P. A.; Chaires, J. B.; Trent, J. O. *Org. Biomol. Chem.***2011**, *9*, 7633.

(33) Stadlbauer, P.; Trantirek, L.; Cheatham, T. E., 3rd; Koca, J.; Šponer, J. *Biochimie* **2014**, *105*, 22.

(34) Shek, Y. L.; Noudeh, G. D.; Nazari, M.; Heerklotz, H.; Abu-Ghazalah, R. M.; Dubins, D. N.; Chalikian, T. V. *Biopolymers* **2014**, *101*, 216.

(35) Li, M. H.; Wang, Z. F.; Kuo, M. H.; Hsu, S. T.; Chang, T. C. *J. Phys. Chem. B* **2014**, *118*, 931.

(36) Marchand, A.; Ferreira, R.; Tateishi-Karimata, H.; Miyoshi, D.; Sugimoto, N.; Gabelica, V. *J. Phys. Chem. B* **2013**, *117*, 12391.

(37) Wallimann, P.; Kennedy, R. J.; Miller, J. S.; Shalongo, W.; Kemp, D. S. *J. Am. Chem. Soc.***2003**, *125*, 1203.

(38) Gray, R. D.; Chaires, J. B. *Curr. Protoc. Nucleic Acid Chem.* **2011**, *Chapter 17*, Unit 17.4.

(39) Mergny, J. L.; Lacroix, L. *Oligonucleotides* **2003**, *13*, 515.

(40) Mergny, J. L.; Li, J.; Lacroix, L.; Amrane, S.; Chaires, J. B. *Nucleic Acids Res.* **2005**, *33*, e138.

(41) Marchand, A.; Granzhan, A.; Iida, K.; Tsushima, Y.; Ma, Y.; Nagasawa, K.; Teulade-Fichou, M. P.; Gabelica, V. *J. Am. Chem. Soc.* **2015**, *137*, 750.

(42) Sun, J.; Kitova, E. N.; Wang, W.; Klassen, J. S. *Anal. Chem.* **2006**, *78*, 3010.

(43) Smargiasso, N.; Rosu, F.; Hsia, W.; Colson, P.; Baker, E. S.; Bowers, M. T.; De Pauw, E.; Gabelica, V. *J. Am. Chem. Soc.* **2008**, *130*, 10208.

(44) Pedroso, I. M.; Duarte, L. F.; Yanez, G.; Baker, A. M.; Fletcher, T. M. *Biochem. Biophys. Res. Commun.* **2007**, *358*, 298.

(45) De Rache, A.; Kejnovská, I.; Vorlíčková, M.; Buess-Herman, C. *Chem. Eur. J.* **2012**, *18*, 4392.

(46) Mao, X.-A.; Marky, L. A.; Gmeiner, W. H. *J. Biomol. Struct. Dyn.* **2004**, *22*, 25.

(47) Karsisiotis, A. I.; Hessari, N. M.; Novellino, E.; Spada, G. P.; Randazzo, A.; Webba da Silva, M. *Angew. Chem. Int. Ed.* **2011**, *50*, 10645.

(48) Mergny, J. L.; Lacroix, L. *Curr. Protoc. Nucleic Acid Chem.* **2009**, *Chapter 17*, Unit 17.1.

(49) Arora, A.; Maiti, S. *J. Phys. Chem. B* **2008**, *112*, 8151.

(50) Viglasky, V.; Bauer, L.; Tluckova, K.; Javorsky, P. *J. Nucleic Acids* **2010**, *2010*, Article ID 820356.

(51) Sket, P.; Črnugelj, M.; Plavec, J. *Nucleic Acids Res.* **2005**, *33*, 3691.

(52) Mukundan, V. T.; Phan, A. T. *J. Am. Chem. Soc.***2013**, *135*, 5017.

(53) Marchand, A.; Gabelica, V. *J. Am. Soc. Mass Spectrom.* **2014**, *25*, 1146.

(54) Webba da Silva, M. *Chem. Eur. J.* **2007**, *13*, 9738.

(55) Deng, J.; Xiong, Y.; Sundaralingam, M. *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98*, 13665.

(56) Ennifar, E.; Walter, P.; Dumas, P. *Nucleic Acids Res.* **2003**, *31*, 2671.

(57) Bugaut, A.; Balasubramanian, S. *Biochemistry* **2008**, *47*, 689.

(58) Cang, X.; Šponer, J.; Cheatham, T. E. *Nucleic Acids Res.* **2011**, *39*, 4499.

(59) Lech, C. J.; Heddi, B.; Phan, A. T. *Nucleic Acids Res.* **2013**, *41*, 2034.

(60) Schultze, P.; Macaya, R. F.; Feigon, J. *J. Mol. Biol.* **1994**, *235*, 1532.

(61) Amrane, S.; Ang, R. W. L.; Tan, Z. M.; Li, C.; Lim, J. K. C.; Lim, J. M. W.; Lim, K. W.; Phan, A. T. *Nucleic Acids Res.* **2009**, *37*, 931.

(62) Amrane, S.; Kerkour, A.; Bedrat, A.; Vialet, B.; Andréola, M.-L.; Mergny, J.-L. *J. Am. Chem. Soc.* **2014**, *136*, 5249.

(63) Lim, K. W.; Alberti, P.; Guédin, A.; Lacroix, L.; Riou, J.-F.; Royle, N. J.; Mergny, J.-L.; Phan, A. T. *Nucleic Acids Res.* **2009**, *37*, 6239.

(64) Madden, T. In *The NCBI Handbook*; McEntyre, J., Ostell, J., Eds.; National Center for Biotechnology Information (US): Bethesda (MD), 2002.

(65) Johnson, M.; Zaretskaya, I.; Raytselis, Y.; Merezhuk, Y.; McGinnis, S.; Madden, T. L. *Nucleic Acids Res.* **2008**, *36*, 5.

(66) Barrett, T.; Wilhite, S. E.; Ledoux, P.; Evangelista, C.; Kim, I. F.; Tomashevsky, M.; Marshall, K. A.; Phillippy, K. H.; Sherman, P. M.; Holko, M.; Yefanov, A.; Lee, H.; Zhang, N.; Robertson, C. L.; Serova, N.; Davis, S.; Soboleva, A. *Nucleic Acids Res.* **2013**, *41*, 991.

(67) Edgar, R.; Domrachev, M.; Lash, A. E. *Nucleic Acids Res.* **2002**, *30*, 207.

(68) Rochin, L.; Hurbain, I.; Serneels, L.; Fort, C.; Watt, B.; Leblanc, P.; Marks, M. S.; De Strooper, B.; Raposo, G.; van Niel, G. *Proc. Natl. Acad. Sci. U.S.A.* **2013**, *110*, 10658.

(69) Chen, M. *Biochemistry* **1992**, 31, 3769

(70) Qu, K., Zhao, C., Ren, J., Qu, X. *Mol. Biosyst*. **2012**, 8, 779

(71) K.-H. Leung, V. P.-Y. Ma, H.-Z. He, D. S.-H. Chan, H. Yang, C.-H. Leung, D.-L. Ma, *RSC Advances* **2012**, *2*, 8273.

(72) 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. *Angew. Chem. Int. Ed.* **2011**, *50*, 8745.

(73) Largy, E.; Saettel, N.; Hamon, F.; Dubruille, S.; Teulade-Fichou, M.-P. *Curr. Pharm. Des.* **2012**, *18*, 1992.

(74) Zhao, C.; Wu, L.; Ren, J.; Xu, Y.; Qu, X. *J. Am. Chem. Soc.* **2013**, *135*, 18786.

(75) Nakatani, K.; Hagihara, S.; Sando, S.; Sakamoto, S.; Yamaguchi, K.; Maesawa, C.; Saito, I. *J. Am. Chem. Soc.* **2003**, *125*, 662.