HHS Public Access

Author manuscript

Biochemistry. Author manuscript; available in PMC 2017 August 14.

Published in final edited form as:

Biochemistry. 2016 March 29; 55(12): 1749–1757. doi:10.1021/acs.biochem.6b00026.

Homologous PNA Hybridization to Noncanonical DNA G-Quadruplexes

Karen A. Kormuth[‡,§], **John L. Woolford Jr.**[‡,§], and **Bruce A. Armitage**[*,†,§]

[†]Department of Chemistry, 4400 Fifth Avenue, Pittsburgh, Pennsylvania 15213-3890, United States

[‡]Department of Biological Sciences, 4400 Fifth Avenue, Pittsburgh, Pennsylvania 15213-3890, United States

[§]Center for Nucleic Acids Science and Technology Carnegie Mellon University, 4400 Fifth Avenue, Pittsburgh, Pennsylvania 15213-3890, United States

Abstract

Potential guanine (G) quadruplex-forming sequences (QFSs) found throughout the genomes and transcriptomes of organisms have emerged as biologically relevant structures. These G-quadruplexes represent novel opportunities for gene regulation at the DNA and RNA levels. Recently, the definition of functional QFSs has been expanding to include a variety of unconventional motifs, including relatively long loop sequences (i.e., >7 nucleotides) separating adjacent G-tracts. We have identified a QFS within the 25S rDNA gene from *Saccharomyces cerevisae* that features a long loop separating the two 3′-most G-tracts. An oligonucleotide based on this sequence, QFS3, folds into a stable G-quadruplex in vitro. We have studied the interaction between QFS3 and several loop mutants with a small, homologous (G-rich) peptide nucleic acid (PNA) oligomer that is designed to form a DNA/PNA heteroquadruplex. The PNA successfully invades the DNA quadruplex target to form a stable heteroquadruplex, but with surprisingly high PNA:DNA ratios based on surface plasmon resonance and mass spectrometric results. A model for high stoichiometry PNA-DNA heteroquadruplexes is proposed, and the implications for quadruplex targeting by G-rich PNA are discussed.

Graphical abstract

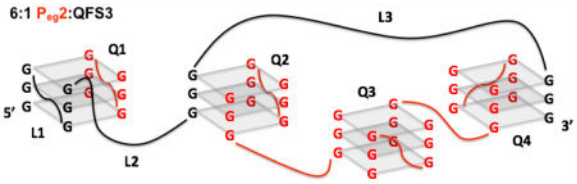
[*]**Corresponding Author**: army@cmu.edu.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.bio-chem.6b00026. Optical spectroscopy, UV melting, SPR and ESI MS data for characterization of DNA homoquadruplexes and PNA-DNA heteroquadruplexes (PDF)

The authors declare the following competing financial interest(s): One of the authors (Bruce Armitage) has a potential competing financial interest associated with this work. He holds equity in PNA Innovations, Inc, which is commercializing PNA technology invented at Carnegie Mellon University.



Certain sequence motifs within nucleic acids have the ability to adopt distinct secondary structural topologies, such as guanine (G) quadruplexes.[1] In vivo, G-quadruplexes can act as structured regulatory elements that have the ability to regulate biological processes at the DNA and RNA levels through interaction with intracellular factors.[2–4] The quad-ruplex-forming capability of regions across the human genome has been experimentally demonstrated by immunofluorescence studies relying on a G-quadruplex binding antibody, supporting the biological relevance of these structures.[5] The prevailing, sequencing-informed estimate places the number of G-quadruplexes within the human genome at >700,000.[6] The breadth of potential quadruplex-forming sequences (QFS) across DNA alone highlights the monumental challenge of identifying the functions of individual G-quadruplexes within the cell.

G-quadruplexes are formed from the mutual interaction of four sets of G-tracts via Hoogsteen H-bonding to form a series of stacked, planar G-tetrads. In an intramolecular G-quadruplex, the nucleotides separating the G-tracts become loops, which are external to the stacked G-tetrads that form the core of the structure. The canonical intramolecular G-quadruplex motif was originally defined as $G_{2–4}N_{1–7}G_{2–4}N_{1–7}G_{2–4}N_{1–7}G_{2–4}$.[7,8] This formula was designed to identify motifs having features that are most likely to support in vivo quadruplex folding. For example, short loops confer the quadruplex with greater structural stability than longer loops.[9,10] However, in recent years, the quadruplex motif definition has been expanded to include a variety of noncanonical sequences. For example, certain motifs having the backbone G-tracts broken up by one or more bulge nucleotides are able to support G-quadruplex folding. It has also become clear that the loops separating the G-tracts need not be limited to only seven nucleotides in length. A striking example of a long-looped quadruplex was identified within the human telomerase reverse transcriptase (hTERT) core promoter. This region of the gene promoter contains 12 individual G-tracts that can fold into a pair of tandem G-quadruplexes, including one with a 26 nt loop that forms its own internal hairpin, conferring enhanced stability to the structure.[12] More recent work has shown that long-looped G-quadruplexes may represent a significant portion of the intracellular quadruplex population. Specifically, potential QFSs with very long (up to 90 nt) central loops and single nucleotide flanking loops have been identified within the 5′ UTRs of human mRNAs.[13] Further, the most updated high-throughput search for potential QFSs in the human genome, which accounts for structures having bulges and long loops, suggests that unconventional G-quadruplexes are much more prevalent than previously thought.[6] This broader diversity in G-quadruplex sequence and structure enhances the possibility for developing synthetic ligands that can target and regulate specific quadruplexes within the cell.

Synthetic nucleic acid analogues, such as peptide nucleic acid (PNA),[14] are useful for targeting specific DNA and RNA sequences. G-quadruplexes are unusual targets in that they can interact with such probes not only via Watson-Crick pairing to form heteroduplexes, but also via Hoogsteen pairing to form heteroquadruplexes, which are stabilized by hetero-G-tetrads. Although it is not the focus of this report, quadruplex-targeting by heteroduplex-forming complementary probes is well established for PNA,[15–22] as well as other classes of oligonucleotides.[23–26] Interestingly, the resulting heteroduplexes appear to be less inhibitory than the native G-quadruplex to a processing ribosome[25,26] or DNA polymerase enzyme.[22]

The alternative option for sequence-based targeting of G-quadruplexes relies on G-rich probes that disrupt the native quadruplex in order to form a heteroquadruplex. We refer to G-rich oligomers that bind in this manner as *homologous* probes, even if the linker joining the G-tracts differs from the loops within the target because, to a first approximation, interaction between the two strands is limited to the guanines. We first reported this hybridization mode for targeting G-quadruplexes in 2001,[27] and there have been numerous additional examples provided by our lab[19–21,28–30] and others.[31–35] Although detailed comparisons of the two binding modes have not been made in most cases, we recently demonstrated that targeting a G-quadruplex with a heteroquadruplex-forming G-rich PNA inhibited DNA polymerase η, whereas targeting the same site with a heteroduplex-forming C-rich PNA enhanced polymerase activity.[22] The observation of opposing effects on quadruplex stability and function through the use of different targeting mechanisms motivates further investigation.

This study focuses on recognition of long-loop-containing DNA quadruplex by a G-rich PNA. We have previously found homologous PNAs having two G-tracts invade DNA quadruplexes to form 2:1 heteroquadruplexes (Scheme 1). Although we have demonstrated this binding mode for several different G-quadruplexes, the rate of hybridization can vary by more than an order of magnitude, depending on the particular fold of the quadruplex target.[30] Given the expanding definition of what constitutes a quadruplex motif, studying the effects of long loops on heteroquadruplex formation by homologous PNA probes became important to ultimately develop the most selective and potent quadruplex-targeting PNAs.

With this work we have identified a new, 3-tetrad quadruplex motif, containing a moderately long central loop in addition to a very long 3′ loop, that folds into a stable, intramolecular G-quadruplex in vitro. Further, we have found that a short, homologous PNA can spontaneously invade this long-looped quadruplex via intermolecular heteroquadruplex formation. However, in contrast to previous results with short-looped quadruplexes, the long loops allow for hybridization of additional PNA strands to the DNA, leading to formation of higher order (5:1 and 6:1 PNA:DNA) complexes. On the basis of testing of additional DNA quadruplexes, this new binding mode appears to be applicable to long-looped quadruplexes in general.

MATERIALS AND METHODS

Materials

DNA oligonucleotides were purchased from Integrated DNA Technologies (www.idtdna.com) and Fisher Scientific (www.fishersci.com). Boc/Z-protected PNA

monomers were purchased from Poly Org (www.polyorginc.com). PNA oligomers were synthesized manually on a lysine-substituted MBHA resin by standard solid-phase peptide synthesis techniques.[53,54] Monomers were activated by HBTU for 1 min in the presence of DIEA and DMF. The PNA oligomer was cleaved from the solid support using TFMSA/TFA/thioanisol/*m*-cresol (1:1:2:6), purified by reverse-phase high-performance liquid chromatography (HPLC) and characterized by matrix-assisted laser desorption ionization time-of-flight (MALDI-TOF) mass spectrometry (calculated mass: 2184, observed mass: 2183).

Equipment

Absorbance data were collected using a Varian Cary 3 spectrophotometer with a temperature controlled multicell holder. Circular dichroism experiments were completed in a Jasco J-715 spectropolarimeter having a single cell holder with a water circulating temperature controller.

UV Melting

Samples (varied concentrations of DNA and/ or PNA) were buffered in 100 mM KCl, 10 mM Tris pH 7, and 0.1 mM Na_2EDTA. The experiments were completed in two phases, first where the samples were heated to 95 °C and then annealed by slowly cooling to 15 °C (1 °C/min). The samples were then slowly heated to 95 °C. Absorbance was measured at 295 nm. Melting temperatures were determined by averaging the minimum of the first derivative of each melting curve, in triplicate (OriginPro 8). To allow for comparison with ESI-MS conditions, similar experiments were completed in buffer solution containing 150 mM NH_4OAc (pH 6.7), 10% MeOH, and 0.1% HCOOH.

Thermal Difference Spectroscopy

Thermal difference (TD) data were collected in triplicate from samples (DNA or DNA + PNA) previously annealed in UV melting experiments. UV cuvettes containing the samples were scanned individually (200 nm/min) at 20 °C. Samples were left in the heat block while the temperature was adjusted to 90 °C, and individual scans were repeated. TD spectra were produced by subtracting the absorbance spectrum at 20 °C from the absorbance spectrum at 90 °C for each sample.

Circular Dichroism

CD data were collected at 30 °C using samples (DNA or DNA + PNA) previously annealed in UV melting experiments. Each spectrum was obtained by averaging six scans collected at 100 nm/min.

Surface Plasmon Resonance (SPR)

SPR analyses were performed on a Biacore T100 instrument (GE Healthcare) using four flow cell sensor chips. CM5 sensor chips were purchased from GE Healthcare and functionalized with streptavidin (approximately 5000 RU's) using NHS-EDC coupling. 5′ biotinylated DNA ligands were individually immobilized (approximately 150 RU's) onto the streptavidin surface of separate flow cells to complete preparation of the sensor chips. Chips

were prepared for PNA injection by priming 5x in the appropriate buffer. Direct binding experiments were completed in buffer containing 10 mM HEPES pH 7.4, 100 mM KCl or LiCl, 3 mM EDTA, and 0.005% v/v polysorbate 20. PNA samples ([P_{eg2}] = 5, 10, 25, 50, 75, and 100 nM) were injected randomly and in triplicate with a total surface contact time of 420 s. In order to determine complex stoichiometry, similar direct binding experiments were completed with higher concentrations of PNA ([P_{eg2}] = 250, 500, 750, and 1000 nM) in the 100 mM KCl SPR buffer to saturate the response signal. Approximate PNA:DNA stoichiometries were then calculated using the following equation, where RU DNAimmob corresponds to the response units of immobilized DNA, RUmax PNA corresponds to the saturating response units obtained during PNA titration, MW is molecular weight and x is the number of PNA strands hybridized to each DNA strand:

$$\frac{\mathrm{RU\ \ DNAimmob}}{\mathrm{MW\ \ DNA\ \ immob}}=\frac{\mathrm{RUmax\ \ PNA}}{x(\mathrm{MW\ \ PNA})}$$

Electrospray Ionization Mass Spectrometry (ESI-MS)

A buffer solution of 150 mM NH_4OAc (pH 6.7) was used for the ESI-MS analyses and prepared with doubly distilled H_2O and filtered with a 0.22 μM filter. PNA and DNA sequences were dissolved to concentrations of 100 mM each in 150 mM NH_4OAc buffer. The QFS3 DNA sequence was ordered through Integrated DNA Technologies (Coralville, IA) and HPLC-purified. Both QFS3 DNA and P_{eg2} PNA were heated to 65 °C, slowly cooled to 25 °C, and stored at 4 °C until use. PNA and DNA complexes were mixed at their respective concentration ratios and allowed to set for up to 24 h at 4 °C. Up to 10% MeOH (v/v) was added per sample to promote solution-to-gas phase transition. Within the MeOH, a fixed percentage of the appropriate acid/base was also included to help improve signal with 0.1% formic acid (v/v) for positive ion mode analyses and 0.5% NH_3 for negative ion mode analyses.

Samples were run using a Waters Micromass ESI-Q-TOF mass spectrometer (Milford, MA). DNA samples were scanned within a range of m/z 300–3000 while PNA samples were scanned between m/z 100–2000. PNA was initially tested by ESI-MS positive ion mode at a concentration of 20 μM and 8% MeOH (0.1% HCOOH). QFS3 DNA was tested by both ESI-MS (−) and (+) ion modes at a concentration of 5 μM with 10% MeOH (0.5% NH_3) and 10% MeOH (0.1% HCOOH), respectively. PNA-QFS3 was prepared at a ratio of [6:1] where concentration of PNA was 30 μM and QFS3 concentration equaled 5 μM in 10% MeOH (0.1% HCOOH). Optimal parameters used: capillary voltage, 2400 V; sample cone voltage, 32 V; extraction voltage, 1.5 V; source temperature, 80 °C; desolvation temperature, 110 °C; collision energy, 2.0 V; cone voltage, 30 L/h; desolvation gas, 450 L/h.

RESULTS

Our lab has previously shown that short PNAs consisting of two G_3 tracts and various linkers hybridize to conventional G-quadruplexes forming 2:1 heteroquadruplexes. [19,20,28] As the definition of a quadruplex-forming sequence expands, conventional, short-looped motifs are coming to be understood as the minority within biological systems, at least based

on bioinformatic analysis.[6] Motivated in part by recent reports of quadruplexes having long (>7 nucleotides) loops between the core G-tracts, we have become interested in learning how PNA heteroquadruplex formation may change with G-quadruplexes having very different topologies.[6,13]

Identification and Design of Long Loop-Containing G-Quadruplex Motifs

This study was conducted using a series of G-quadruplex-forming DNA oligomers to assess the effects of very long loops on in vitro G-quadruplex formation and hybridization to a small, homologous (G-rich) PNA, P_{eg2} (Table 1). P_{eg2} is designed to invade and hybridize to DNA G-quadruplexes forming intermolecular, DNA/PNA heteroquadruplexes. Specifically, the two G_3 tracts allow for formation of a 3-tetrad heteroquadruplex in combination with two additional G_3 tracts donated by the target DNA The G-tracts are separated by two abasic diethylene glycol “miniPEG” (mP) units; the abasic linker was originally designed to suppress hybridization to complementary sequences, but does not impact the affinity or hybridization kinetics of heteroquadruplex formation.[29] Meanwhile, the target QFS3 was originally identified in the 25S rDNA of the yeast *Saccharomyces cerevisiae* using the algorithm $G_3X_nG_3X_nG_3X_nG_3$, with a maximum total sequence length of 50 nucleotides, i.e., the three X_n loops in total could comprise up to 38 nucleotides.[36] QFS3 is particularly interesting because it contains both a moderately long (6 nt) central loop and a very long (17 nt) 3′ loop. Myc19 is derived from a quadruplex-forming sequence found within the nuclease hypersensitivity element III_1 of the *c-MYC* promoter. [37,38] In contrast to QFS3, Myc19 has three very short (1, 2, and 1 nt) loops. Our lab has utilized Myc19 as a model G-quadruplex target for PNA probes. [28] In order to study the effects of varied loop length and placement on G-quadruplex formation, two artificial potential G-quadruplex-forming sequences were designed. $QFS3_{L2}$ and $QFS3_{L3}$ are both derived from QFS3 but each has a loop replaced with the corresponding short loop of Myc19. hTERT was included as an example G-quadruplex having an especially long central loop (26 nt) which folds into its own hairpin, greatly stabilizing the overall structure.[12]

G-rich DNA Oligomers Having Long 3′ Loops Fold into Stable, Intramolecular G-Quadruplexes in Vitro

G-quadruplex formation and stability in vitro can be detected using biophysical and spectroscopic techniques. We have confirmed that QFS3, $QFS3_{L2}$, and $QFS3_{L3}$ DNA oligomers fold into G-quadruplexes using three optical spectroscopic methods. In vitro G-quadruplex formation was detected by UV melting at 295 nm, with each oligomer producing a quadruplex-specific hypochromic transition at this wavelength. From this data we extracted a melting temperature (T_m) for each quadruplex The T_m values determined in 100 mM K^+ are significantly lower than what would be observed for a quadruplex having only short loops, consistent with the destabilizing effect that long loops are reported to have on quadruplex structures.[9,41] (Note that the T_m value for the hTERT quadruplex was determined in the presence of 10 mM KCl due to its higher thermal stability relative to the other quadruplexes.) In addition, minimal effect is observed on the T_m values recorded at 5-fold higher DNA concentration. This result confirms that each QFS forms intramolecular G-quadruplexes in vitro. Finally, replacing KCl with LiCl in the buffer significantly

destabilized the quadruplexes, as shown in Supplemental Figure S3 for QFS3, consistent with the known quadruplex-destabilizing effect of Li^{+}[42].

We also used circular dichroism (CD) spectropolarimetry to confirm the presence of G-quadruplexes in each sample and gain insight into the overall topology of each structure. G-quadruplexes are often categorized based on the orientation of the G-tracts relative to each other. For example, in a parallel quadruplex, all four G-tracts run in the same direction, while an antiparallel quadruplex has two pairs of G-tracts running in opposite directions; these structural differences give rise to unique CD signatures.[43] CD spectra of QFS3, $QFS3_{L2}$, and $QFS3_{L3}$ contain characteristic G-quadruplex spectral features. The broad peak from 260 to 300 nm produced by QFS3, $QFS3_{L2}$, and $QFS3_{L3}$ suggests a more complex topology (e.g., hybrid structure) than a standard parallel (min at 240 nm, max. at 260 nm) structure (Figure 1A). Additionally, the unpaired bases in the loop should contribute to the shape of the spectrum and this contribution would be greater for longer-looped quadruplexes. The stronger signal from hTERT likely arises from the internal hairpin, present within loop 2, since B-form DNA also exhibits a positive peak at 260 nm.[12] The long loops of the other three quadruplexes are not predicted to form a base-paired secondary structure.

Thermal difference spectroscopy (TDS) provides additional insight into nucleic acid secondary structures based on features of the difference spectrum that is produced by subtracting the UV absorption spectrum of the folded nucleic acid from that of the unfolded (i.e., thermally denatured) form.[44] Each DNA oligomer in Table 1 produces a TDS containing spectral features characteristic of G-quadruplexes, in particular, the distinctive inversion near 295 nm (Figure 1B, see inset). Although this feature is weak for QFS3 and $QFS3_{L2}$, the intensity is similar to that of the known long-loop quadruplex hTERT. The spectrum of $QFS3_{L3}$ differs considerably from those of the other quadruplexes in Figure 1B, but is similar to that of Myc19 (Supporting Information, Figure S1). While the origin of the differences among the spectra is unclear, variation in thermal difference spectra produced by different DNA G-quadruplexes has been reported for other systems.[44]

PNA Invasion of Long Loop-Containing G-Quadruplexes

The G-rich PNA P_{eg2} interacts favorably with each long-looped G-quadruplex, evidenced by the increase in T_m upon the addition of two equivalents of PNA to each DNA (Table 3; Supplemental Figure S4).

CD and TDS experiments are consistent with hetero-G-quadruplex formation in preannealed samples of 2:1 PNA/ DNA (Figure 2). In the presence of PNA, the CD signal for each QFS is intensified and appears more parallel (i.e., increased positive intensity at 260 nm) than that of the DNA alone (Figure 2A; overlaid CD spectra for each quadruplex with and without PNA are shown in Supplemental Figure S5). Similar results were observed for heteroquadruplex formation between G-rich PNA and a DNA QFS based on the human telomeric repeat.[30] Analogous results are obtained from TDS, where the spectral shapes for each QFS become much more similar in the presence of PNA than for the DNA alone (compare Figures 1B and 2B).

To confirm that the PNA interacts with these DNAs via intermolecular G-quadruplex formation, we carried out SPR experiments where PNA was flowed past immobilized QFS3 (Figure 3). An obvious increase in response units was detected upon injection of increasing concentrations of PNA in buffer containing 100 mM K^+. When the experiment was repeated in buffer containing 100 mM Li^+, nearly all the response signal was lost, consistent with PNA-DNA heteroquadruplex formation. We obtained similar results for additional SPR experiments conducted with $QFS3_{L2}$, $QFS3_{L3}$, and hTERT (Supplemental Figure S6).

Sensorgrams obtained for P_{eg2} hybridization to five different G-quadruplexes at 50 nM PNA concentration are shown in Figure 4, with initial association and dissociation rates shown in Table 4. Heteroquadruplex formation for the long-looped G-quadruplexes is noticeably slower than to the short-looped quadruplex, Myc19. This result is consistent with our prior findings that heteroquadruplex formation is slower for DNA quadruplexes that adopt hybrid structures than for parallel morphologies.[30]

Detection of Higher Order Complex Formation between P_{eg2} and Long-Looped G-Quadruplexes

A curious result from the SPR experiments was that considerably more PNA was required to saturate the signal than expected for a 2:1 stoichiometry. For example, Figure 5 shows sensorgrams for binding of P_{eg2} to QFS3. The red dashed lines indicate the calculated saturating response units (RU_{max}) for 2:1, 5:1, and 6:1 stoichiometries, based on the amount of DNA immobilized on the SPR chip (determined as described in the Experimental section). On the basis of these calculations, the PNA/DNA stoichiometry appears to be either 5:1 or 6:1. (Uncertainty arises from the fact that a small amount of DNA is released from the SPR chip during the regeneration phase after each injection of PNA, i.e., the RU_{max} value for a given stoichiometry decreases over the duration of the experiment.) ESI-MS experiments also supported formation of both 5:1 and 6:1 heteroquadruplexes, as described in (Supporting Information Figures S7 and S8). High-stoichiometry (e.g., 4:1 and 5:1) complexes were also indicated by SPR for P_{eg2} binding to the other long-loop quadruplexes (Supplemental Figure S9) indicating that this phenomenon is not specific to QFS3. However, higher order complex formation does not appear to require exceptionally long loops, as the longest loop of $QFS3_{L3}$ has only six nucleotides. In contrast, binding of P_{eg2} to the short-looped quadruplex Myc19 saturates at 2:1 stoichiometry (Supplemental Figure S10), consistent with our previous results for a slightly different quadruplex-forming PNA.[28]

DISCUSSION

G-Rich Motifs with Varied Long Loop Placement Support Stable G-Quadruplex Formation

Early bioinformatics studies determined that >300,000 G-quadruplex forming sequences are present in the human genome.[7,8] While this is a large number, recent research indicates that it substantially underestimates the actual number of QFSs because quadruplex formation is supported by a considerably broader range of sequences than defined by the original search algorithm. The ability of QFSs to tolerate extra bases (bulges) within G-tracts, to adopt noncanonical folds wherein corners of the quadruplex include Gs from noncontiguous tracts, and the successful folding of quadruplexes even when long (>7 nt) loops are present,[6,12,13]

all indicate that the definition of a QFS is much broader than was originally anticipated. The results presented above represent our first effort to investigate the impact of long loops on G quadruplex recognition by a heteroquadruplex-forming PNA.

To the best of our knowledge, QFS3 represents a new quadruplex motif having a moderately long central loop in addition to an especially long 3′ loop (in contrast to a long central loop with single nucleotide flanking loops). We have shown that QFS3 and similar DNA oligomers can fold into intramolecular G-quadruplexes in vitro with thermal stabilities well above physiological temperature (30 °C for yeast), suggesting that these structures have the potential to form in vivo. Furthermore, a G-quadruplex of moderate stability may be useful to the cell as a regulatory switch especially in the presence of other molecular factors or changing intracellular conditions. QFS3 is also a potentially interesting biological quadruplex, due to its location within yeast 25S rDNA. The presence of sequences with quadruplex-forming potential in eukaryotic rDNA has been noted previously, although the biological role of such quadruplexes is yet to be determined.[36,45] In addition, the fact that QFS3 is on the nontemplate strand means that it will also be present in the 25S rRNA. While there is no evidence that it folds into a quadruplex within the mature ribosome,[46] it is possible that a quadruplex forms transiently during rRNA processing or ribosome assembly. Alternatively, a DNA-RNA heteroquadruplex could form during transcription in order to regulate ribosome levels in the cell.[47]

Long Loops Allow for Hybridization of >2 Homologous PNAs to a DNA Quadruplex

Our group has previously established a homologous targeting strategy using G-rich PNA oligomers to invade G-quadruplexes forming stable, intermolecular heteroquadruplexes.[19–21,27–30] We have now shown that long-looped G-quadruplexes are also amenable to homologous PNA hybridization. However, the long loops provide these quadruplexes with remarkable plasticity in regards to the number of PNA strands in complex with each DNA.

Scheme 2A illustrates possible structures for 5:1 and 6:1 PNA:DNA heteroquadruplexes. After opening of the quad-ruplex, the first PNA molecule hybridizes analogous to the proposed binding mode for short-loop quadruplexes[28] to form heteroquadruplex Q1. However, the long third loop hinders binding of another PNA molecule to both the third and fourth G-tracts. (The extended loop likely provides the molecule with sufficient flexibility to accommodate binding of additional PNA strands.) Rather, two PNA molecules combine with the third G-tract to form a termolecular heteroquadruplex (Q2). Two PNAs also combine with the fourth G-tract in this manner (Q4), giving rise to the 5:1 complex. This model leaves a dangling G-tract for each of the two PNAs that are bound within Q2 and Q4, which can be cross-linked into an all-PNA quadruplex using the sixth PNA molecule (Q3). Further, these tandem quadruplexes can potentially stack end-to-end (Scheme 2B), enhancing the stability of the overall structure, which could explain the drive toward higher-stoichiometry complexes.[12,48] While this model stands in contrast to the “beads-on-a-string” conformation where the individual quadruplexes do not stack, as has been proposed for long telomeric DNAs,[49–51] the differences in linker length and composition for the PNA-DNA heteroquadruplexes reported here could give rise to different preferred structures.[50]

Regardless of the structural details, the model for a 6:1 complex is appealing because the total number of G-tracts is 16, with four contributed by QFS3 and two contributed by each of the 6 PNA molecules. If the total number of G-tracts is a multiple of 4, a structure can form in which each of the G-tracts participates in a homo- or heteroquadruplex.

SPR also indicated the formation of higher-order hetero-quadruplexes between P_{eg2} and DNA quadruplexes $QFS3_{L2}$, $QFS3_{L3}$, and hTERT. The results for $QFS3_{L3}$, which has a maximum loop length of only 6 nt, suggests that the quadruplex target need not have an exceptionally long loop before higher order PNA hybridization stoichiometries become possible. Nevertheless, the fact that Myc19 only allows formation of 2:1 heteroquadruplexes demonstrates that there is a lower limit on loop lengths that permit formation of high stoichiometry complexes.

Evidence for formation of higher stoichiometry PNA-DNA heteroquadruplexes was recently reported by Panyutin and co-workers, through investigation of PNA ($G_4T_4G_4$) hybridization to the homologous target sequence in a DNA oligonucleotide, as well as in the context of a DNA hairpin loop and a plasmid. Observation of multiple gel-shifted bands indicated formation of at least two different heteroquadruplexes of differing mobility, while optical spectroscopy detected at least one of these heteroquadruplexes with a 3:1 stoichiometry (PNA/ DNA). Interestingly, the target only contained two G-tracts, precluding formation of its own intramolecular G-quadruplex In our QFS3 system, which has four G-tracts, we propose a stoichiometry of 5:1 or 6:1, with the latter being double of what Panyutin et al. found for a truncated target with half as many G-tracts.

Implications for Biological G-Quadruplex Targeting

We have shown that long loop-containing G-quadruplexes can be invaded by homologous, G-rich PNA However, the tendency of these long loops to accommodate higher order complex formation in vitro is not ideal for biological applications of this targeting method, since the higher PNA concentrations required to achieve stable binding would increase the likelihood of unwanted, off-target interactions. Therefore, future probe designs for targeting long-loop quadruplexes should focus on decreasing the stoichiometry of these complexes. A potential solution is the combination of homologous and complementary domains, allowing for bimodal quadruplex hybridization, i.e. heteroduplex and— quadruplex formation by the same probe, as reported recently by Basu and co-workers. The complementary domain could be targeted either to an external flanking region or within the long loop, while the homologous domain then associates with two nearby G-tracts to form the heteroquadruplex part of the final structure. Optimization of the relative affinities of the two domains should then allow selective targeting of one quadruplex (or quadruplex-forming sequence) among the millions of QFSs present in the genome and transcriptome.

This work provides new insight into potential targeting of a specific class of G-quadruplexes. Additional studies will be required to develop a better understanding of how loop length, structure, and sequence affect PNA hybridization. Nevertheless, we expect that the prevalence of unusual quadruplex motifs in biology will require the development of new targeting strategies to allow identification and modulation of G-quadruplex functions in the future.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

The authors thank Prof. W. David Wilson and Sarah Laughlin-Toth of Georgia State University for acquiring mass spectrometry data and for helpful discussions and the David Scaife Family Charitable Foundation for financial support of this research (Award 141RA01).

ABBREVIATIONS AND ACRONYMS

G	guanine
QFS	quadruplex-forming sequence
PNA	peptide nucleic acid
hTERT	human telomerase reverse transcriptase
Myc19	quadruplex-forming element in human *C-MYC* promoter
miniPEG	diethylene glycol
TDS	thermal difference spectroscopy
CD	circular dichroism
SPR	surface plasmon resonance
ESI-MS	electrospray ionization mass spectrometry

References

1. Burge S, Parkinson GN, Hazel P, Todd AK, Neidle S. Quadruplex DNA: Sequence, Topology and Structure. Nucleic Acids Res. 2006; 34:5402–5415. [PubMed: 17012276]
2. Patel DJ, Phan AT, Kuryavyi V. Human Telomere, Oncogenic Promoter and 5'-UTR G-Quadruplexes: Diverse Higher Order DNA and RNA Targets for Cancer Therapeutics. Nucleic Acids Res. 2007; 35:7429–7455. [PubMed: 17913750]
3. Balasubramanian S, Neidle S. G-Quadruplex Nucleic Acids as Therapeutic Targets. Curr. Opin. Chem. Biol. 2009; 13:345–353. [PubMed: 19515602]
4. Lipps HJ, Rhodes D. G-Quadruplex Structures: *In Vivo* Evidence and Function. Trends Cell Biol. 2009; 19:414–422. [PubMed: 19589679]
5. Biffi G, Tannahill D, McCafferty J, Balasubramanian S. Quantitative Visualization of DNA G-Quadruplex Structures in Human Cells. Nat. Chem. 2013; 5:182–186. [PubMed: 23422559]
6. Chambers VS, Marsico G, Boutell JM, Di Antonio M, Smith GP, Balasubramanian S. High-Throughput Sequencing of DNA G-Quadruplex Structures in the Human Genome. Nat. Biotechnol. 2015; 33:877–881. [PubMed: 26192317]
7. Huppert JL, Balasubramanian S. Prevalence of Quadruplexes in the Human Genome. Nucleic Acids Res. 2005; 33:2908–2916. [PubMed: 15914667]
8. Todd AK, Johnston M, Neidle S. Highly Prevalent Putative Quadruplex Sequence Motifs in Human DNA. Nucleic Acids Res. 2005; 33:2901–2907. [PubMed: 15914666]
9. Guedin A, Gros J, Alberti P, Mergny J-L. How Long is Too Long? Effects of Loop Size on G-Quadruplex Stability. Nucleic Acids Res. 2010; 38:7858–7868. [PubMed: 20660477]

10. Hazel P, Huppert J, Balasubramanian S, Neidle S. Loop-Length-Dependent Folding of G-Quadruplexes. J. Am. Chem. Soc. 2004; 126:16405–16415. [PubMed: 15600342]
11. Mukundan VT, Phan AT. Bulges in G-Quadruplexes: Broadening the Definition of G-Quadruplex-Forming Sequences. J. Am. Chem. Soc. 2013; 135:5017–5028. [PubMed: 23521617]
12. Palumbo SL, Ebbinghaus SW, Hurley LH. Formation of a Unique End-to-End Stacked Pair of G-Quadruplexes in the hTERT Core Promoter with Implications for Inhibition of Telomerase by G-Quadruplex-Interactive Ligands. J. Am. Chem. Soc. 2009; 131:10878–10891. [PubMed: 19601575]
13. Jodoin R, Bauer L, Garant J-M, Mahdi Laaref A, Phaneuf F, Perreault J-P. The Folding of 5′-UTR Human G-Quadruplexes Possessing a Long Central Loop. RNA. 2014; 20:1129–1141. [PubMed: 24865610]
14. Egholm M, Buchardt O, Christensen L, Behrens C, Freier SM, Driver DA, Berg RH, Kim SK, Norden B, Nielsen PE. PNA Hybridizes to Complementary Oligonucleotides Obeying the Watson-Crick Hydrogen-Bonding Rules. Nature. 1993; 365:566–568. [PubMed: 7692304]
15. Datta B, Armitage BA. Hybridization of PNA to Structured DNA Targets: Quadruplex Invasion and the Overhang Effect. J. Am. Chem. Soc. 2001; 123:9612–9619. [PubMed: 11572682]
16. Green JJ, Ying L, Klenerman D, Balasubramanian S. Kinetics of Unfolding the Human Telomeric DNA Quadruplex Using a PNA Trap. J. Am. Chem. Soc. 2003; 125:3763–3767. [PubMed: 12656607]
17. Amato J, Oliviero G, De Pauw E, Gabelica V. Hybridization of Short Complementary PNAs to G-Quadruplex Forming Oligonucleotides: an Electrospray Mass Spectrometry Study. Biopolymers. 2009; 91:244–255. [PubMed: 19065573]
18. Amato J, Pagano B, Borbone N, Oliviero G, Gabelica V, De Pauw E, D'Errico S, Piccialli V, Varra M, Giancola C, Piccialli G, Mayol L. Targeting G-Quadruplex Structure in the Human c-Kit Promoter with Short PNA Sequences. Bioconjugate Chem. 2011; 22:654–663.
19. Marin VL, Armitage BA. RNA Guanine Quadruplex Invasion by Complementary and Homologous PNA Probes. J. Am. Chem. Soc. 2005; 127:8032–8033. [PubMed: 15926825]
20. Marin VL, Armitage BA. Hybridization of Complementary and Homologous Peptide Nucleic Acid Oligomers to a Guanine Quadruplex-Forming RNA. Biochemistry. 2006; 45:1745–1754. [PubMed: 16460021]
21. Gupta A, Lee L-L, Tanious F, Wilson WD, Ly DH, Armitage BA, Roy S. Strand Invasion of DNA Quadruplexes by PNA: Comparison of Homologous and Complementary Hybridization. ChemBioChem. 2013; 14:1476–1484. [PubMed: 23868291]
22. Murphy CT, Gupta A, Armitage BA, Opresko PL. Hybridization of G-quadruplex-forming peptide nucleic acids to guanine-rich DNA templates inhibits DNA polymerase eta extension. Biochemistry. 2014; 53:5315–5322. [PubMed: 25068499]
23. Kumar N, Maiti S. Role of Locked Nucleic Acid Modified Complementary Strand in Quadruplex/Watson-Crick Duplex Equilibrium. J. Phys. Chem. B. 2007; 111:12328–12337. [PubMed: 17914789]
24. Kumar N, Patowary A, Sivasubbu S, Petersen M, Maiti S. Silencing c-MYC Expression by Targeting Quadruplex in P1 Promoter Using Locked Nucleic Acid Trap. Biochemistry. 2008; 47:13179–13188. [PubMed: 19053274]
25. Murat P, Zhong J, Lekieffre L, Cowieson NP, Clancy JL, Preiss T, Balasubramanian S, Khanna R, Tellam J. G-quadruplexes regulate Epstein-Barr virus-encoded nuclear antigen 1 mRNA translation. Nat. Chem. Biol. 2014; 10:358–364. [PubMed: 24633353]
26. Rouleau SG, Beaudoin J-D, Bisaillon M, Perreault J-P. Small Antisense Oligonucleotides Against G-Quadruplexes: Specific mRNA Translational Switches. Nucleic Acids Res. 2015; 43:595–606. [PubMed: 25510493]
27. Datta B, Schmitt C, Armitage BA. Formation of a PNA2-DNA2 Hybrid Quadruplex. J. Am. Chem. Soc. 2003; 125:4111–4118. [PubMed: 12670232]
28. Roy S, Tanious FA, Wilson WD, Ly DH, Armitage BA. High-affinity homologous peptide nucleic acid probes for targeting a quadruplex-forming sequence from a MYC promoter element. Biochemistry. 2007; 46:10433–10443. [PubMed: 17718513]

29. Lusvarghi S, Murphy CT, Roy S, Tanious FA, Sacui I, Wilson WD, Ly DH, Armitage BA. Loop and Backbone Modifications of PNA Improve G Quadruplex Binding Selectivity. J. Am. Chem. Soc. 2009; 131:18415–18424. [PubMed: 19947597]
30. Roy S, Zanotti KJ, Murphy CT, Tanious FA, Wilson WD, Ly DH, Armitage BA. Kinetic Discrimination in Recognition of DNA Quadruplex Targets by Guanine-Rich Hetero-quadruplex-Forming PNA Probes. Chem. Commun. 2011; 47:8524–8526.
31. Englund EA, Xu Q, Witschi MA, Appella DH. PNA-DNA Duplexes, Triplexes, and Quadruplexes Are Stabilized with trans-Cyclopentane Units. J. Am. Chem. Soc. 2006; 128:16456–16457. [PubMed: 17177367]
32. Ishizuka T, Yang J, Komiyama M, Xu Y. G-Rich Sequence-Specific Recognition and Scission of Human Genome by PNA/DNA Hybrid G-Quadruplex Formation. Angew. Chem., Int. Ed. 2012; 51:7198–7202.
33. Ito K, Go S, Komiyama M, Xu Y. Inhibition of Translation by Small RNA-Stabilized mRNA Structures in Human Cells. J. Am. Chem. Soc. 2011; 133:19153–19159. [PubMed: 22007660]
34. Usui K, Okada A, Kobayashi K, Sugimoto N. Control of Guanine-Rich DNA Secondary Structures Depending on the Protease Activity Using a Designed PNA Peptide. Org. Biomol. Chem. 2015; 13:2022–2025. [PubMed: 25519192]
35. Gaynutdinov TI, Englund EA, Appella DH, Onyshchenko MI, Neumann RD, Panyutin IG. G-Quadruplex Formation Between G-Rich PNA and Homologous Sequences in Oligonucleotides and Supercoiled Plasmid DNA. Nucleic Acid Ther. 2015; 25:78–84. [PubMed: 25650982]
36. Hershman SG, Chen Q, Lee JY, Kozak ML, Yue P, Wang L-S, Johnson FB. Genomic Distribution and Functional Analysis of Potential G-Quadruplex-Forming Sequences in Saccharomyces cerevisae. Nucleic Acids Res. 2007; 36:144–156. [PubMed: 17999996]
37. Siddiqui-Jain A, Grand CL, Bearss DJ, Hurley LH. Direct Evidence for a G-Quadruplex in a Promoter Region and its Targeting with a Small Molecule to Repress c-MYC Transcription. Proc. Natl. Acad. Sci. U. S. A. 2002; 99:11593–11598. [PubMed: 12195017]
38. Simonsson T, Pecinka P, Kubista M. DNA Tetraplex Formation in the Control Region of c-myc. Nucleic Acids Res. 1998; 26:1167–1172. [PubMed: 9469822]
39. Phan AT, Modi YS, Patel DJ. Propeller-Type Parallel-Stranded G-Quadruplexes in the Human c-myc Promoter. J. Am. Chem. Soc. 2004; 126:8710–8716. [PubMed: 15250723]
40. Mergny J-L, Phan A-T, Lacroix L. Following G-Quartet Formation by UV-Spectroscopy. FEBS Lett. 1998; 435:74–78. [PubMed: 9755862]
41. Kwok CK, Sherlock ME, Bevilacqua PC. Effect of Loop Sequence and Loop Length on the Intrinsic Fluorescence of G-Quadruplexes. Biochemistry. 2013; 52:3019–3021. [PubMed: 23621657]
42. Williamson JR, Raghuraman MK, Cech TR. Monovalent Cation-Induced Structure of Telomeric DNA: The G-Quartet Model. Cell. 1989; 59:871–880. [PubMed: 2590943]
43. Vorlíckova M, Kejnovska I, Sagi J, Renciuk D, Bednarova K, Motlova J, Kypr J. Circular Dichroism and Guanine Quadruplexes. Methods. 2012; 57:64–75. [PubMed: 22450044]
44. Mergny J-L, Li J, Lacroix L, Amrane S, Chaires JB. Thermal Difference Spectra: A Specific Signature for Nucleic Acid Structures. Nucleic Acids Res. 2005; 33:e138. [PubMed: 16157860]
45. Hanakahi LA, Sun H, Maizels N. High Affinity Interactions of Nucleolin with G-G-Paired rDNA. J. Biol. Chem. 1999; 274:15908–15912. [PubMed: 10336496]
46. Ben-Shem A, Jenner L, Yusupova G, Yusupov M. Crystal structure of the eukaryotic ribosome. Science. 2010; 330:1203–1209. [PubMed: 21109664]
47. Wanrooij PH, Uhler JP, Simonsson T, Falkenberg M, Gustafsson CM. G-Quadruplex Structures in RNA Stimulate Mitochondrial Transcription Termination and Primer Formation. Proc. Natl. Acad. Sci. U. S. A. 2010; 107:16072–16077. [PubMed: 20798345]
48. Tothova P, Krafcikova P, Vigalsky V. Formation of Highly Ordered Multimers in G-Quadruplexes. Biochemistry. 2014; 53:7013–7027. [PubMed: 25347520]
49. Yu H-Q, Miyoshi D, Sugimoto N. Characterization of Structure and Stability of Long Telomeric DNA G-Quadruplexes. J. Am. Chem. Soc. 2006; 128:15461–15468. [PubMed: 17132013]

50. Yu H, Gu X, Nakano S, Miyoshi D, Sugimoto N. Beads-on-a-String Structure of Long Telomeric DNAs Under Molecular Crowding Conditions. J. Am. Chem. Soc. 2012; 134 2006010.1021/ja305384c.

51. Petraccone L, Spink C, Trent JO, Garbett NC, Mekmaysy CS, Giancola C, Chaires JB. Structure and Stability of Higher-Order Human Telomeric Quadruplexes. J. Am. Chem. Soc. 2011; 133:20951–20961. [PubMed: 22082001]

52. Bhattacharyya D, Nguyen K, Basu S. Rationally Induced RNA:DNA G-Quadruplex Structures Elicit an Anticancer Effect by Inhibiting Endogenous eIF-4E Expression. Biochemistry. 2014; 53:5461–5470. [PubMed: 25072107]

53. Christensen L, Fitzpatrick R, Gildea B, Petersen KH, Hansen HF, Koch T, Egholm M, Buchardt O, Nielsen PE, Coull J, Berg RH. Solid-Phase Synthesis of Peptide Nucleic Acids. J. Pept. Sci. 1995; 1:175–183. [PubMed: 9222994]

54. Koch, T. PNA Synthesis by Boc Chemistry. In: Nielsen, PE., editor. Peptide Nucleic Acids: Protocols and Applications. 2. Horizon Bioscience; Norfolk: 2004. p. 37-60.

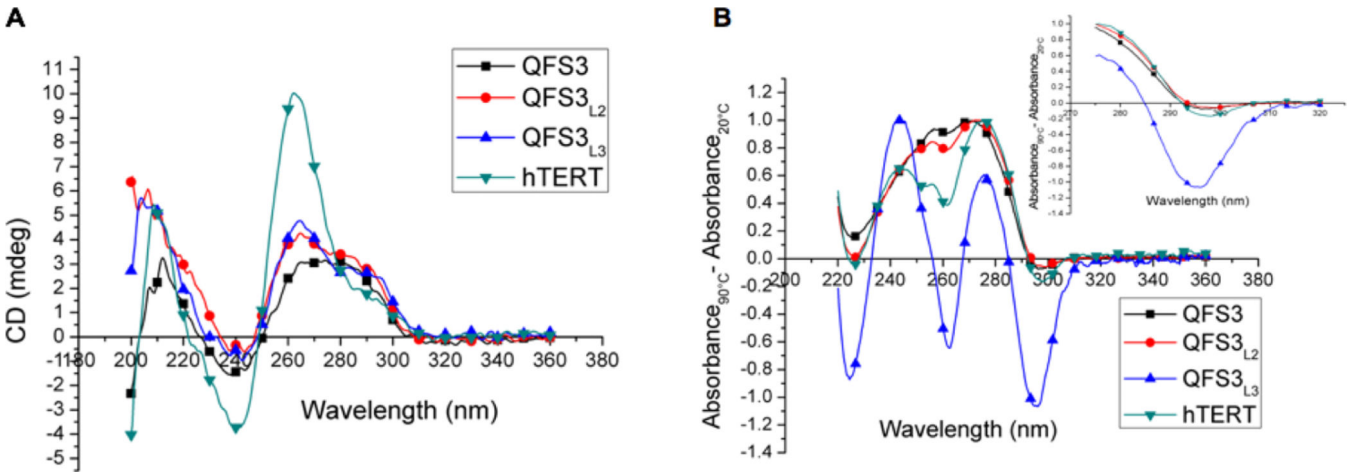


Figure 1.

Long loop-containing DNAs fold into G-quadruplexes in vitro. (A) CD spectra of QFS3, $QFS3_{L2}$, and $QFS3_{L3}$ in 100 mM K^+. (B) TDS of QFS3, $QFS3_{L2}$, $QFS3_{L3}$, and hTERT in 100 mM K^+. The inset graph highlights the region around 295 nm, which includes a quadruplex-specific inversion for each oligomer.

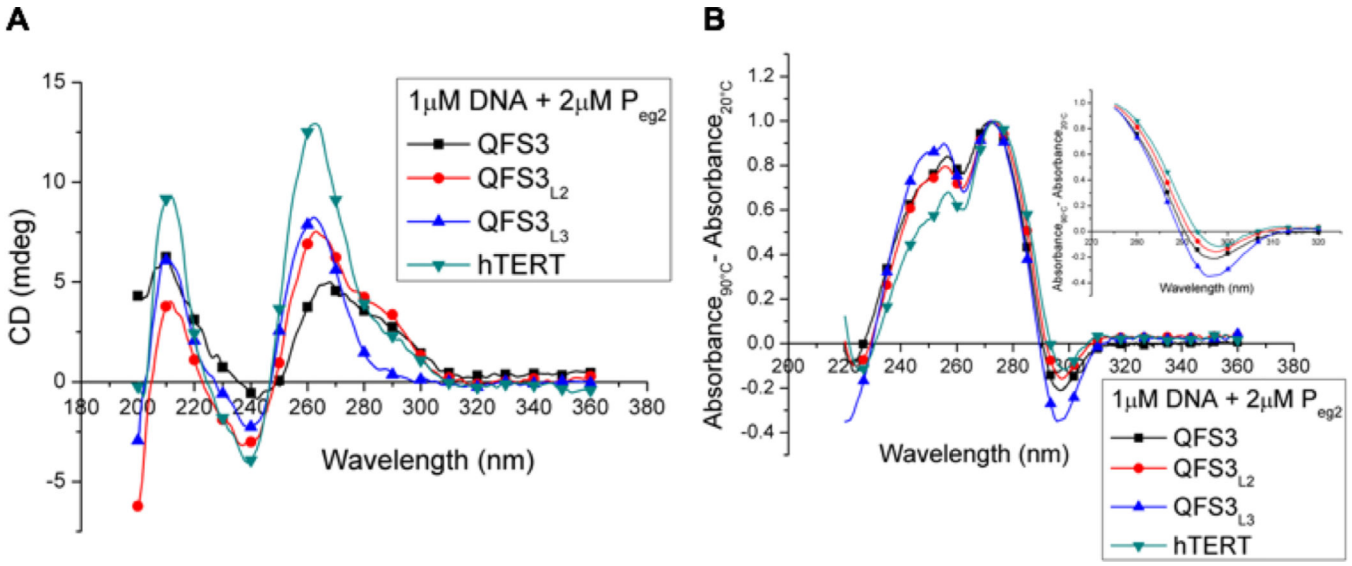


Figure 2.

P_{eg2} forms DNA/PNA heteroquadruplexes with long loop-containing DNA G-quadruplexes. (A) CD spectra of each P_{eg2}/DNA complex in 100 mM K^+ suggests overall more parallel topology (min at 240 nm, max at 260 nm) than those of the DNA alone (Figure 1A). (B) TD spectroscopy in 100 mM K^+ detects G-quadruplex formation for each complex as evidenced by the inversion near 295 nm. The inset graph highlights the region around 295 nm.

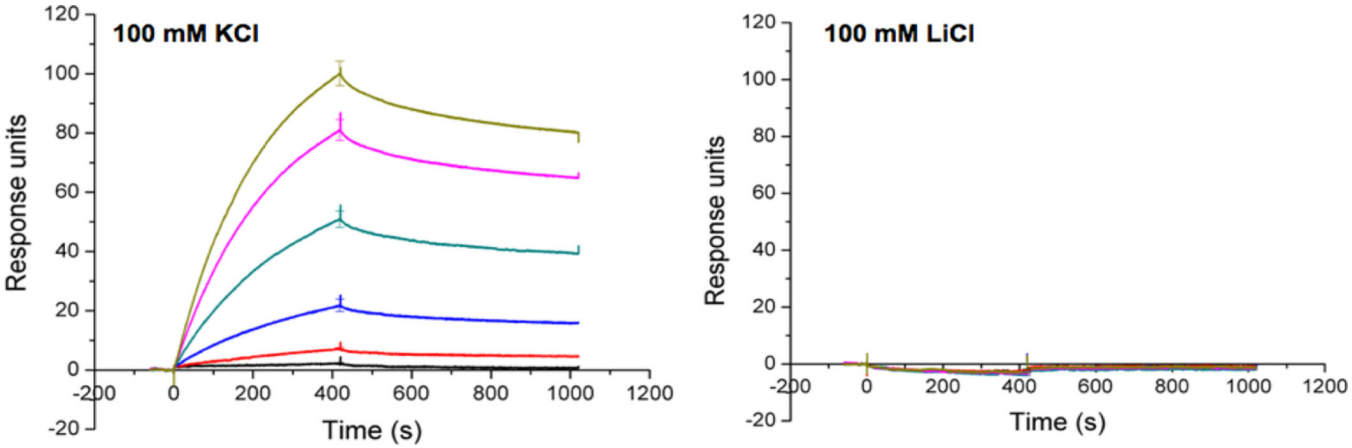


Figure 3.

P_{eg2} hybridization to QFS3 results in PNA/DNA heteroquadruplex formation. SPR sensorgrams of P_{eg2} binding to immobilized QFS3 in 100 mM KCl (left). In 100 mM LiCl, the response is lost, indicating the molecules do not interact in quadruplex-destabilizing conditions (right). [P_{eg2}] = 5, 10, 25, 50, 75, and 100 nM.

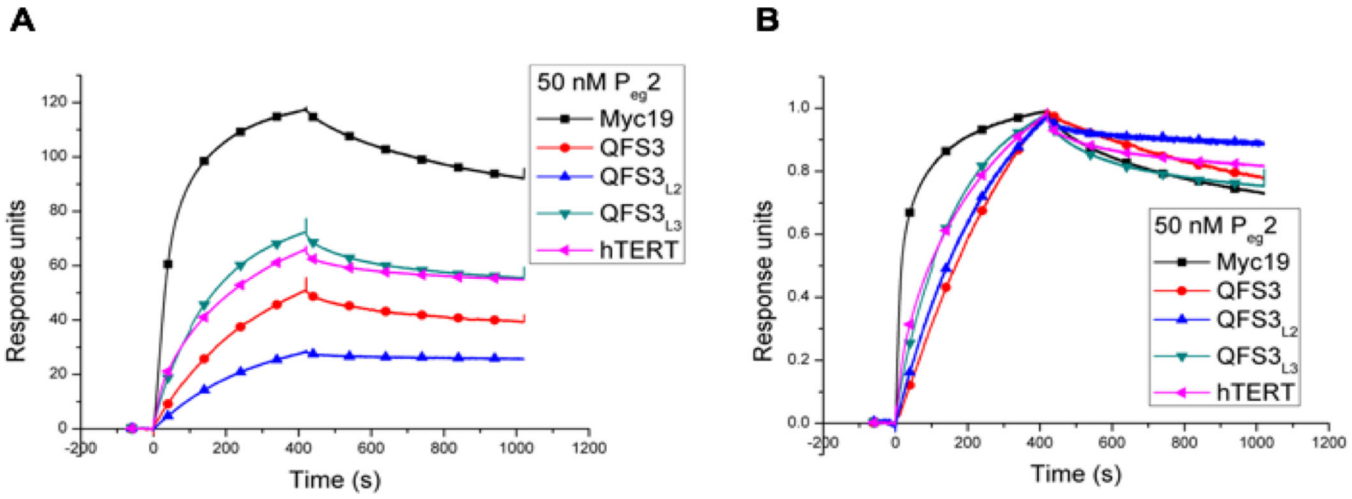


Figure 4.

P_{eg2} readily associates with long-looped G-quadruplexes. (A) SPR sensorgrams of 50 nM P_{eg2} binding to various immobilized DNA quadruplexes in 100 mM K^+. (B) Sensorgrams are normalized to the maximum RU_{max} value for each quadruplex to allow for visual comparison of on/off rates of P_{eg2} to each target.

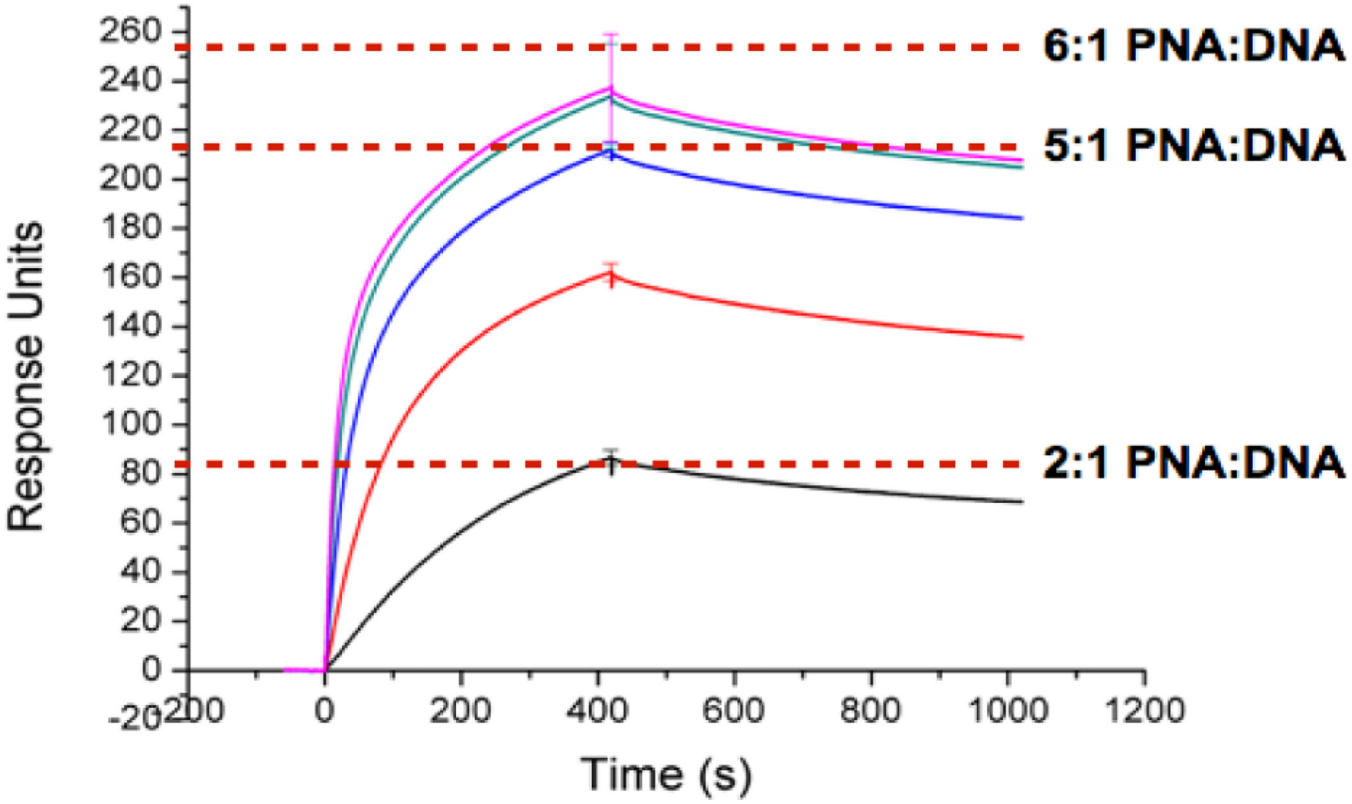
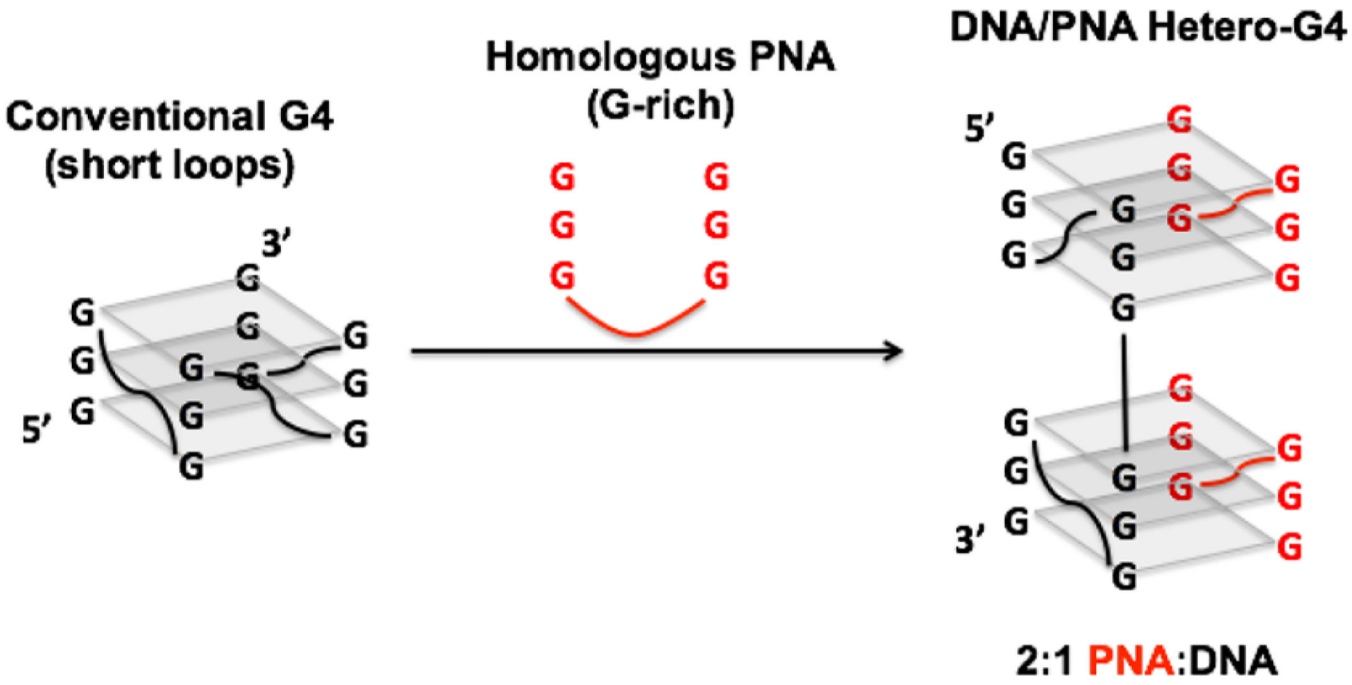


Figure 5.

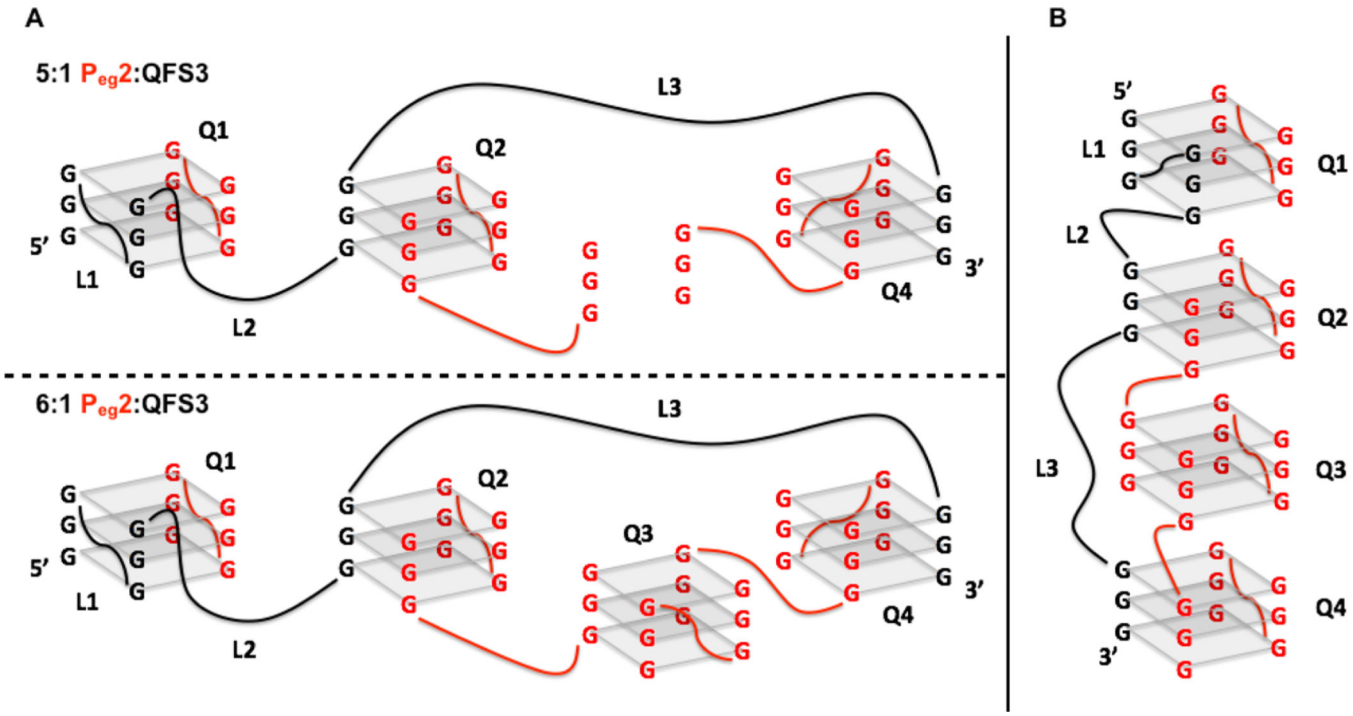
P_{eg2} forms a higher order heteroquadruplex with the long-looped DNA G-quadruplex, QFS3. SPR sensorgrams of P_{eg2} binding to immobilized QFS3 in 100 mM K^+. [P_{eg2}] = 100, 250, 500, 750, and 1000 nM. The red dashed lines indicate theoretical RU_{max} for 2:1, 5:1, and 6:1 PNA:DNA.



Scheme 1.

A Homologous Targeting Strategy for PNA Hybridization to G-Quadruplex DNA[a]

[a]A short, homologous PNA can invade a DNA G-quadruplex forming a 2PNA:1DNA heteroquadruplex.



Scheme 2.

(A) Model 5:1 and 6:1 PNA:DNA Heteroquadruplex Structures. (B) Q1-Q4 Represented in an End-to-End Stacked Conformation

Table 1

Sequences of PNA Probes and DNA G-Quadruplex Targets[a]

	PNA (N-to-C)
Peg2	H-GGG-mP-mP-GGG-Lys-NH_2
	DNA (5′-to-3′)
Myc19[39]	AGGGTGGGGAGGGTGGGGA
QFS3	GGGTCGGGTAGTGAGGGCCTTGGTCAGACGCAGCGGG
$QFS3_{L2}$	GGGTCGGGGAGGGCCTTGGTCAGACGCAGCGGG
$QFS3_{L3}$	GGGTCGGGTAGTGAGGGTGGG
hTERT[12]	GGGGCTGGGCCGGGGACCCGGGAGGGGTCGGGACGGGGCGGGG

[a]G-tracts predicted or reported to participate in tetrad formation shown in bold and underlined.

Table 2

Concentration Dependence of DNA Quadruplex Melting Temperature (T_m, °C) Determined in 100 mM K^+

DNA	1 μM DNA	5 μM DNA
QFS3	54.9 ± 3.0	57.0 ± 0.2
$QFS3_{L2}$	56.7 ± 1.3	56.7 ± 0.3
$QFS3_{L3}$	68.7 ± 0.2	68.3 ± 0.6
hTERT[a]	64.8 ± 1.6	61.8 ± 0.8

[a]Measured in 10 mM K^+.

Table 3

PNA-DNA Heteroquadruplex Melting Temperature (T_m, °C) and Stabilization Relative to DNA Homoquadruplex (ΔT_m, °C) Determined in 100 mM K^+ [a]

DNA	T_m	ΔT_m
QFS3	61.6 ± 0.4	6.7
$QFS3_{L2}$	75.6 ± 3.1	18.9
$QFS3_{L3}$	77.2 ± 3.9	8.5
hTERT [b]	>95	>30

[a] [DNA] = 1 μM, [Peg2] = 2 μM.

[b] Measured in 10 mM K^+.

Table 4

PNA Association (On) and Dissociation (Off) Rates in RU•s^{-1} and Ratio (Off/On) Determined at 50 nM P_{eg2}

DNA	On Rate	Off Rate	Ratio
Myc19	0.501 ± 0.008	0.086 ± 0.003	0.17
QFS3	0.171 ± 0.007	0.338 ± 0.001	0.18
$QFS3_{L2}$	0.115 ± 0.006	0.014 ± 0.004	0.12
$QFS3_{L3}$	0.338 ± 0.017	0.077 ± 0.004	0.20
hTERT	0.353 ± 0.004	0.048 ± 0.004	0.14