Recognition of Guanine-quadruplexes by PNA and gammaPNA Oligomers

Stanley Oyaghire

Submitted in Partial Fulfillment of the Requirements for the Degree of Doctor of Philosophy

The Department of Chemistry

Carnegie Mellon University

Pittsburgh, PA

April 21, 2016

For my parents, Bernard and Sarah Oyaghire, with whom this started, and without whose love and prayers, it would not have ended nearly as well as it did

ACKNOWLEDGEMENTS

Nothing has demanded more from me than completing the work summarized in this dissertation. For many in a less supportive context, the opportunity cost of earning the doctorate can greatly exceed the purported prestige of adding those letters to their names. One of the most influential scientists in human history attributed his prowess to being "perched on the shoulders of giants." Here are mine:

Professor Bruce Armitage has been my rock. Any university professor persuasive enough to procure funds from an expert review panel should have little trouble convincing a naïve, eager 21-year old student about the merits of their work, but it takes a special person to nurture that student through the various levels of intellectual maturity requisite for a career in scientific research. Bruce agreed to be my research advisor at a time when the word *quadruplex* registered, in my mind, little more than a building with four units; now that same word makes its way into my dissertation title, with a completely different meaning this time. I cannot end this paragraph without relating an anecdote about the man: As an undergraduate researcher coming to Pittsburgh for the first time to work in Bruce's lab, he picked me up from the bus station and drove me to get grocery supplies, before taking me to my apartment building. This act of kindness, probably out of his memory now, is the sole reason why I chose, out of other options, to come to Pittsburgh and study at Carnegie Mellon.

Other professors in the Chemistry department have aided this journey. Members of my advisory committee, Profs. Subha Das and Danith Ly, have been there from the beginning, helping me navigate all the departmental check-points *en route* to the final defense. I hope they now see the contradistinction between the naiveté of my early years and the researcher that I have become. Prof. Marcel Bruchez opened the doors of his lab to me when I began developing experiments in which my lab lacked expertise. For this I have him, and Dr. Cheryl Telmer, who directly developed crucial elements of the materials used in this work—to thank

Past members of the Armitage lab—the greatest 'Army' in the world—have been invaluable. Dr. Anisha Gupta, Dr. Ha Pham, Dr. Junriz Delos Santos, Dr. Nathaniel Shank, and Dr. Halimatu Mohammed are alumni of the group that provided training opportunities for me at different times. Most notably, Dr. Gloria Silva, with whom I share a long history of work in the lab, opened the doors of her home to me when I first arrived Pittsburgh, and continues to do so even now.

Among the present members of the group, I now have friends from whom I have learned lessons that will stay with me for a long time. I thank: Dr. Karen Kormuth and Munira Fouz, for providing an important example that resilience during difficult circumstances can yield positive outcomes; Taylor Canady, for persistently pushing me out of my intellectual comfort zone to think about broader applications of my research findings; Dr. Xiaohong Tan, for illustrating that the path to one's long-term goals can require tremendous dedication and fortitude; April Berlyoung and Dmytro Kolodieznyi, for teaching me that humility and patience are necessary virtues for adjusting to a new environment.

I also thank the high school and undergraduate researchers that worked with me. Collin Cherubim and Joe Martinez—both co-authors on my premier journal publication as a first author—are very dedicated researchers. Steve Jang, and Adam Barsouk were also fantastic to work with and brought fresh excitement and a different perspective to some long-running projects. Indeed, the act of teaching all of these students imbued me with a different level of appreciation for those who have committed their lives to delivering this service.

Although the entirety of the Chemistry department has been tremendously supportive, a few examples are worth mentioning: Rea Freeland, the calm, silver-haired Associate Dean who never lets the stresses of her job, of which there are many—percolate to her interactions with students, has been a supportive voice throughout this journey. My gratitude also goes to Prof. Catalina Achim, whose peculiar sense of humor was always effective in invigorating me during long days. The support staff including, Valerie

Bridge, Sara Wainer, Brenda Chambers, Gayathri Withers, and Patsey Haddock, have also been very helpful. Students and postdocs mostly come and go, but it is faculty and staff like these that sustain the department.

Pittsburgh has also blessed me with friends outside the context of the academy who have been very supportive. The entire family of the Redeemed Christian Church of God, Pittsburgh, has been tremendously helpful, offering me very relevant recalibration whenever I got frustrated with the scientific research community. I want to thank Pastors Gboyega and Kemi Esan, whose commitment to excellence and insightful sermons constituted a powerful driving force sustaining me through the doctoral program. Also, Emmanuel, Uyi, Yebom, Seun, Lulu, and Tilda, are friends that will stay with me for the rest of my life.

If I make it past the defense without relieving the stress of this process in harmful ways, it would be due in large part to the physical and emotional support offered by Ebun. Her love in the months leading up to the end of this process has been a buffer to neutralize bouts of doubt and panic that have occurred along the way.

My family has sacrificed so much to get me to this point. As I write this part, a drop of the emotions I feel begins to make its way down my face. My dad, Bernard Oyaghire, has been my mentor and the single most influential person in my life. His resilience in the face of challenges and commitment to his family confer on me a tremendous respect for parenthood, because of all the sacrifices it entails. My mom, Sarah Oyaghire, has been my spiritual anchor, and has bestowed on me a magnitude of love that will be hard to find anywhere else. I must thank Brenda, Martha, and Philip, who have sacrificed quality time with their older sibling that few others would relinquish without resentment.

Stanley Oyaghire

2016

Table of Contents

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

The broad and varying regulatory or causative roles of guanine (G)-quadruplexes in both biological processes and some human diseases incentivize the development of ligands able to recognize these structures and modulate their effects. These reagents might find use either as therapeutics or, more broadly, as discovery probes to elucidate both the existence and functional implications of quadruplex structures in the genome. Towards this end, this dissertation explores quadruplex recognition by peptide nucleic acid (PNA) or gamma-peptide nucleic acid (γPNA) oligomers possessing sequences that are either complementary, homologous, or both to selected quadruplex-forming sequences (QFS).

An overview of G-quadruplex structures in DNA and RNA, with particular emphasis on their structural characteristics and polymorphisms, established and/or speculated biological regulatory functions, and several strategies for recognition is presented in Chapter 1.

Chapter 2 explores quadruplex invasion by complementary γPNA oligomers directed to different regions of a QFS. We observe that while all probes bind with similar affinities to their respective sequences in the context of the QFS, there is a strong positional bias in the potency of inhibition when the γPNAs are directed against an mRNA transcript bearing the QFS in its 5'-untranslated region (UTR). This bias is all but eliminated when the probes and transcript are pre-incubated at an elevated temperature prior to the start of the translation reaction, demonstrating that kinetics exert a significant influence on translation inhibition by these probes in this context. Additionally, we find γPNAs to be functionally superior to other molecules (PNA or 2'-OMe RNA) previously explored as ligands for quadruplex invasion in other reports.

We extend this mode of recognition to additional G-rich target sequences bearing four G_2 or G_4 tracts in Chapter 3. The data presented show that the complementary γPNAs can form stable hybrids with their respective targets, overcoming, where present, the thermodynamic barrier to hybridization posed by the

quadruplex fold in the target. Importantly, these probes also elicit potent and specific inhibition in the context of an *in vitro* translation assay.

Quadruplex invasion by homologous PNA/γPNA oligomers is explored in Chapter 4. We demonstrate that all probe molecules tested are effective ligands for the QFS, since the binding reactions yield hybrids that are more stable than the starting quadruplex structure. Translation inhibition by this class of probes has also been demonstrated, but these effects were observed to be non-specific. Modifications intended to impede probe binding at sites suspected to be culpable for the non-specific effects did not improve specificity. Importantly, however, specificity is improved by the presence of relatively short DNA oligomers that compete with the mRNA transcript for probe binding, suggesting that modifications to tune specificity, while currently lacking, would enable de-conflation of specific from nonspecific effects induced by these probes.

'Chimeric' PNA oligomers possessing two domains for recognition; one complementary, and the other homologous, have been explored as QFS-interactive molecules in Chapter 5. We demonstrate that these probes are able to discriminate against an RNA target possessing only one (out of two) 'binding sites' required to engage both domains of the probe. These results provide impetus for designing PNA/γPNA oligomers that might achieve the holy grail of QFS recognition: selectivity for one quadruplex, out of many such structures reported to exist in the genome.

CHAPTER 1

GUANINE-QUADRUPLEXES: STRUCTURES, FUNCTIONS, AND STRATEGIES FOR RECOGNITION

1.1 STRUCTURAL DEVIATIONS FROM B-DNA

Although the coalition of efforts to unravel the secondary structure of DNA converged on B-DNA—a right-handed helical arrangement of antiparallel strands[1]—as the major structural conformation adopted by the information carrier, many sequence-dependent deviations exist[2]. For example, *Z-DNA* motifs[3]—characterized by alternating purines and pyrimidines, can form left-handed helices (Figure 1.1) under physiological salt conditions *in vitro*. These structures are further stabilized by negative supercoiling[4] and are predicted to relieve transcriptionally-induced torsional stress[5]. Likewise, *cruciform* structures[6] (Figure 1.1) result from negative supercoiling in regions of DNA characterized by 6-nucleotide inverted repeats, and are predicted to play a role in stabilizing certain human chromosomes[7]. Also, *triplex* structures[8]—formed by Hoogsteen hydrogen (H)-bonding interactions[9] between a third DNA strand and purine-rich duplex DNA (Figure 1.1), have been observed in buffer solutions, while motifs capable of forming these structures are enriched in the introns of genes involved in cell signaling[10].

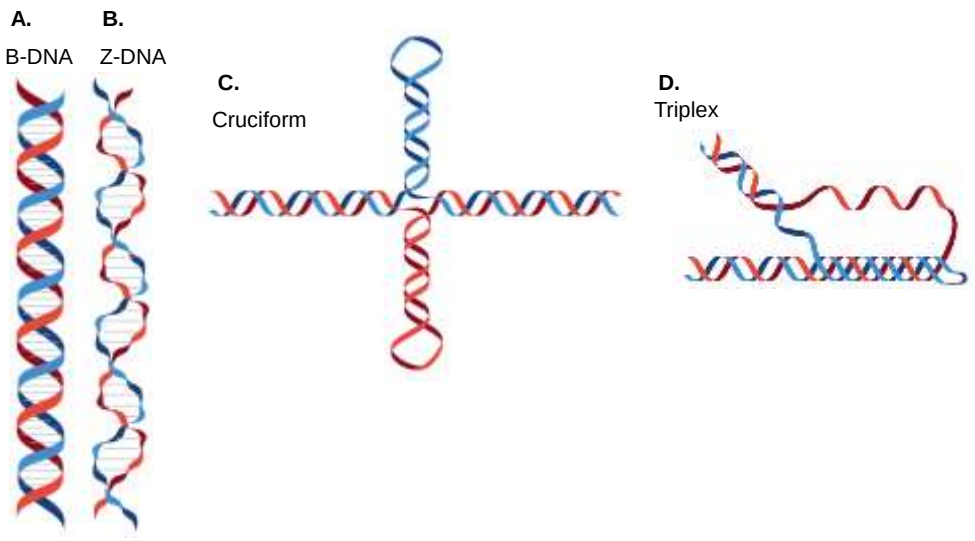


Figure 1.1: Possible structural conformations of DNA. B-DNA (**A**). Z-DNA (**B**). Cruciform (**C**). Triplex (**D**). Reviewed in REF. 2, adapted from REF. 17.

1.2 G-QUADRUPLEX STRUCTURES

While the *in vivo* existence and physiological relevance of the aforementioned structures remain under investigation, another class of non-B-form secondary structures, called guanine (G)-quadruplexes[11, 12], are receiving increasing attention due to evidence of their existence *in vivo*[13-16] and numerous reports demonstrating their regulatory and/or causative roles in biological processes[17-21] and human diseases[22], respectively.

Guanine quadruplexes are formed from G-rich oligonucleotides containing at least four G-tracts of 2 or more G residues, wherein each tract is separated by at least one nucleotide[23]. The core guanine residues are held in a square planar arrangement, wherein each guanine is both the donor and acceptor of two hydrogen bonds[11] (Figure 1.2). Co-axial stacking of these H-bonded macrocycles generates van der Waals' interactions that stabilize the quadruplex structure[23] (Figure 1.2). Further stabilizing interactions result from coordinate covalent bonding between the carbonyl groups of each guanine residue and alkali ions—with a significant bias in the stabilizing effect occurring for K^+ and Na^+.[24]

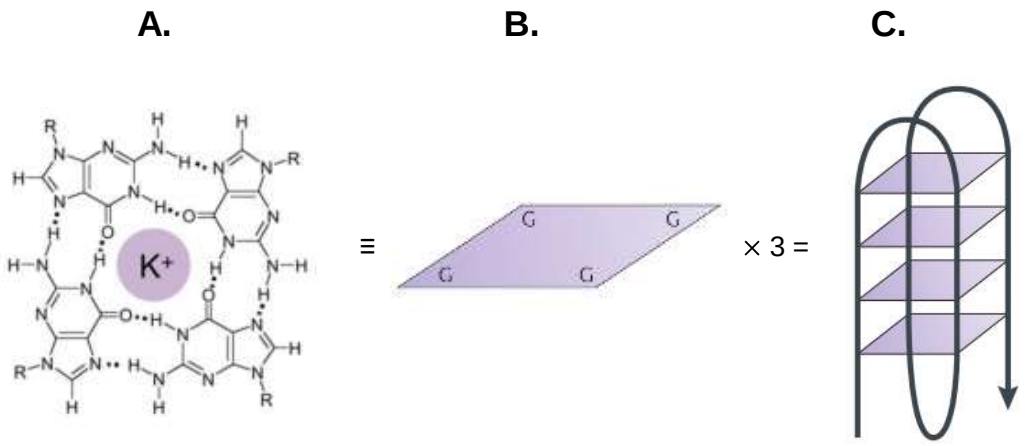


Figure 1.2: Stabilizing interactions within a G-quadruplex. Hoogsteen H-bonds and coordinate covalent interactions within the tetrad (**A**). Square planar arrangement of guanine nucleobases within G-tetrad (**B**). Tetrad stacking to form the G-quadruplex (**C**). Adapted from REF. 17.

The molecularity of the structure depends on the number of strands donating the G-tracts. Self-sufficient G-rich strands possessing all the G-tracts requisite for quadruplex formation fold into intramolecular structures[25, 26], wherein the residues involved in H-bonding interactions within the tetrad are present on the same strand (Figure 1.3). Conversely, intermolecular structures result when the G-tracts are supplied by different strands[27] (Figure 1.3). In all of these structural configurations, the intervening, non-tetrad-forming residues are extruded to the loops, directed away from quadruplex core.

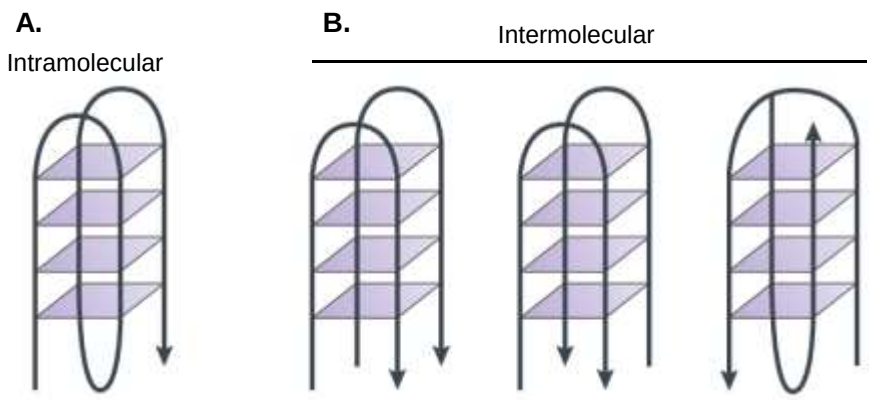


Figure 1.3: Possible structural conformations of the G-quadruplex. Structures can be intramolecular (**A**) or intermolecular (**B**). Adapted from REF. 17.

1.2.1 G-quadruplex Structures in DNA

While all the structural features enumerated above hold true for all (DNA and RNA) quadruplexes, considerable structural polymorphism exists for DNA-derived structures. The morphology of G-quadruplex structures in DNA is affected by molecular crowding, strand orientations, loop and overhang length/composition, and cation concentration/identity[28-30]. Further, although the majority of quadruplex-forming sequences (QFS) derived from functionally relevant regions of the genome were characterized by short (< 7) loops separating contiguous Gx tracts[17]—an observation that

informed early algorithms[31, 32] designed to identify QFS motifs within entire genomes—recent evidence suggests that the quadruplex structure is sufficiently robust to tolerate longer loop sequences and interruptions in the G-tracts.

For example, Palumbo *et al* have demonstrated that a G-rich DNA sequence derived from the hTERT promoter adopts a tertiary structure consisting of two end-to-end stacked G-quadruplexes—one of which features a 26-nt loop[33] (Figure 1.4A). The authors also provided evidence suggesting that the loop residues are involved in Watson-Crick H-bonding interactions to generate a duplex secondary structure within the global quadruplex tertiary structure[33] (Figure 1.4A). Further, Guédin *et al* have provided biophysical evidence to show that G-rich DNA sequences with long (up to 15 nts) intervening residues between the G_3-tracts can fold into stable intramolecular quadruplexes in physiologically relevant buffers[34]. We have recently reported stable intramolecular quadruplex folding by a G-rich DNA oligomer possessing a 17-nt intervening sequence between the last two G_3-tracts[35].

More strikingly, Mukundan and Phan utilized a rigorous repertoire of biophysical characterization methods to show that G-rich DNA sequences possessing non-G residues within the G_3-tracts could still fold into stable intramolecular quadruplex structures under physiological conditions[36, 37] (Figure 1.4B). Evidently, quadruplex folding would extrude the intervening, non-G residues as bulges within the G_3-tracts (Figure 1.4B). While the physiological relevance of these structures remains unknown, a recent report has demonstrated that non-canonical quadruplexes—of the types described above—constitute the majority (~ 70 %) of DNA G-quadruplexes in the human genome[38]. Further, the same report demonstrated that these non-anonical G-quadruplexes are the predominant quadruplex structures in genomic regions associated with regulation[38], suggesting a potential functional role for these structures.

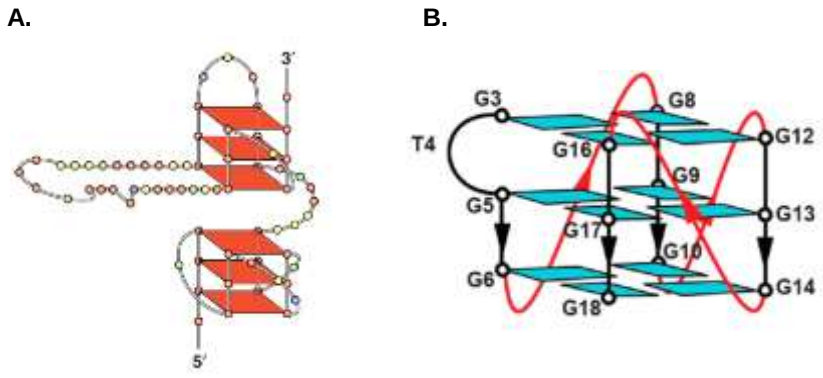


Figure 1.4: Examples of non-canonical Structures. G-quadruplexes with a long loops (**A**) or bulges have been reported (**B**). Figures adapted from REF. 32 and REF. 36, respectively.

1.2.2 G-quadruplex Structures in RNA

Significantly less polymorphism has been reported for RNA G-quadruplexes, which predominantly adopt parallel folds irrespective of sequence context and experimental conditions[26, 39, 40]. (RNA G-quadruplexes with long central loops have been reported, but these structures still conserve the parallel arrangement of strands[41].) These structures are more stable than their DNA homologs[40, 42, 43] due, in part, to an extended network of intramolecular hydrogen bonding interactions between the C2' hydroxyl groups and adjacent H-bond acceptors on neighboring backbone phosphate groups or ribose O4' atoms[44] (Figure 1.5). These interactions contribute to the favorable enthalpy of folding, while decreasing the entropic cost associated with ordering water molecules within the G-quadruplex grooves[44].

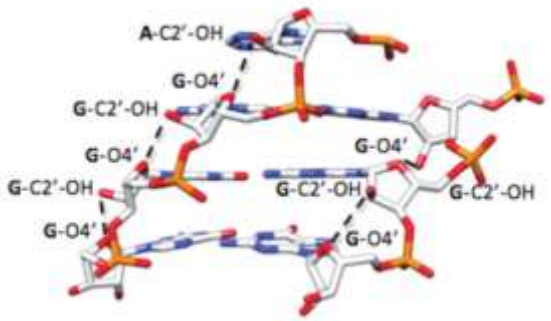


Figure 1.5: Intramolecular H-bonding interactions present in RNA G-quadruplexes. The 2'-hydroxyl group mediates interactions with O4' atoms of the adjacent sugars. Adapted from REF. 43.

1.3 (SELECTED) FUNCTIONS OF DNA G-QUADRUPLEXES

1.3.1 Regulation of Telomerase

Telomerases are evolutionarily conserved ribonucleoproteins that specifically extend the ends of linear chromosomes[45-47]. These genomic regions are characterized by a high density of G-rich repeats[48, 49] that have been directly observed to form G-quadruplexes in ciliates[13] and humans[14], and are predicted to adopt a similar structural conformation in diverse other organisms[50]. Zahler *et al* first reported that intramolecular quadruplex formation by G-rich sequences derived from the *Oxytricha* telomeric repeats deactivated these strands as substrates for recognition and polymerization by the *Oxytricha* telomerase *in vitro*[51]. These authors further proposed that quadruplex formation by the G-rich primers directly impeded recognition by the enzyme and accelerated the dissociation rates of bound substrates from its RNA subunit[51]. This finding—coupled with the subsequent discovery that telomerase is quiescent in somatic cells, but activated in the majority of cancer cells[52]—has motivated a plethora of efforts[53] to identify quadruplex-specific ligands that induce and/or stabilize telomeric quadruplex structures, with potential inhibitory effects on telomerase function (Figure 1.6).

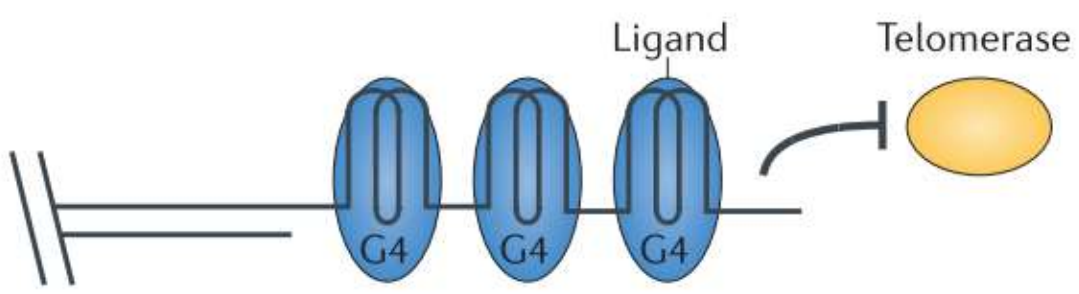


Figure 1.6: Molecules that induce and/or stabilize intramolecular quadruplex formation at telomeres are being explored as inhibitors of telomerase function. Figure adapted from REF. 17.

1.3.2 Regulation of Transcription

Although the high density of QFS motifs in the regions proximal to the promoters of protein-coding genes in humans[31, 38] and diverse other organisms (such as yeast[54], plants[55], and bacteria[56]) suggests some regulatory roles for G-quadruplexes on transcription, these effects are most clearly elucidated for the nuclease hypersensitive element (NHE) III_1 of human c-MYC[57]. Siddiqui-Jain *et al* provided the first direct evidence that a G-rich sequence within the NHE III_1 of human c-MYC folded into an intramolecular quadruplex structure that repressed transcription when placed in control of a luciferase reporter[58], thus confirming earlier speculation on the functional relevance of this motif on c-MYC transcription[59]. Further, a quadruplex-interactive molecule (TMPyP4), but not its positional isomer incapable of quadruplex binding (TMPyP2), was able to inhibit c-MYC transcription in relevant cell lines[58].

The molecular mechanism underpinning this regulation is proposed to involve the sequestration of binding sites for transcriptional activators of c-MYC, such as Sp1, by the folded conformation of the NHE III_1 (Figure 1.7).[57] (Additional silencing is speculated to occur by the sequestration of binding sites for hnRNP K by an i-motif structure formed on the C-rich strand, Figure 1.7.) It is now known that ancillary transcriptional activators such as NM23-H2 can bind and unfold the quadruplex structure, revealing the binding sites for Sp1 and other factors, such as CNBP[60]. Conversely, nucleolin—a non-specific quadruplex-binding protein that induces and/or stabilizes quadruplex structures within the NHE III_1—can potentiate inhibition of Sp1-mediated c-MYC expression (Figure 1.7).[61] These results have incentivized the development of ligands able to recognize the c-MYC G-quadruplex and perturb its interactions with NM23-H2[62, 63].

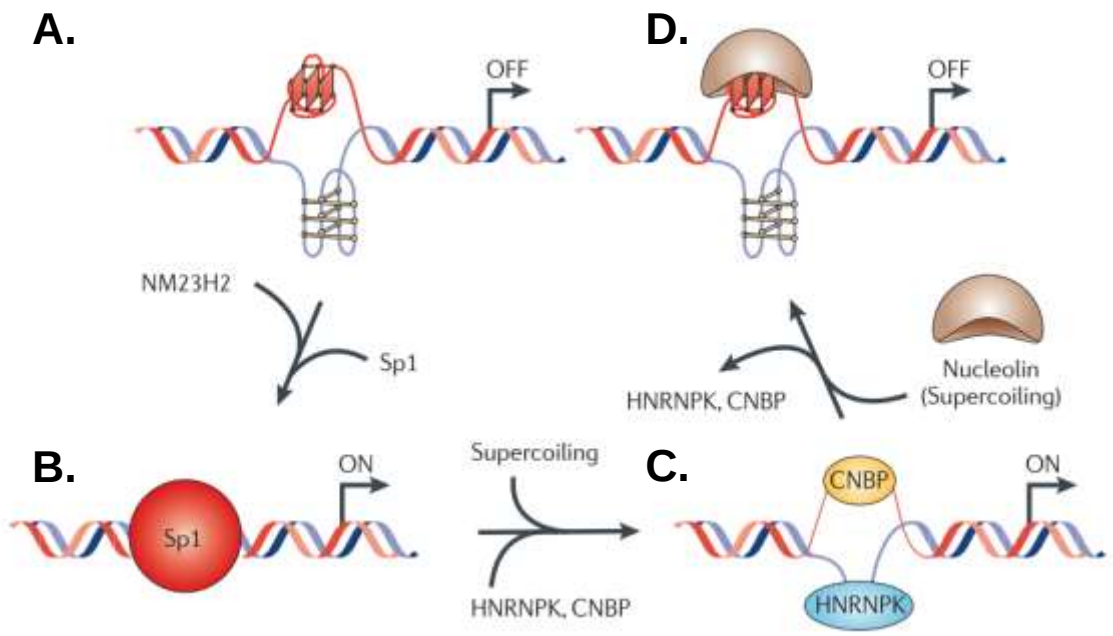


Figure 1.7: Transcriptional regulation of c-MYC by DNA G-quadruplex in NHE III_1. G-quadruplex represses transcription by occluding Sp1 binding site (**A**). NM23H2 resolves the quadruplex and is displaced by Sp1 (**B**). Supercoiling in DNA induces binding of CNBP and hnRNP K, both of which are activators of c-MYC transcription (**C**). Nucleolin displaces both transcription factors and stabilizes the quadruplex, leading to inhibition of c-MYC (**D**). Reviewed in REF. 56.

In a different example of quadruplex-mediated transcriptional regulation, Haeusler *et al* have reported that intronic G_4C_2 repeat expansions in *C9orf72*—the most frequently reported genetic cause of amyotrophic lateral sclerosis and frontotemporal dementia[64]—form intramolecular DNA G-quadruplexes that lead to an accumulation of abortive transcripts *in vitro*[22]. Quadruplex formation on the DNA template was reported to decrease the processivity of the RNA polymerase, with the severity of the impact correlating positively with the degree of repeat expansion[22]. The authors proposed a model wherein quadruplex formation downstream of the polymerase triggered pausing and eventual displacement of the polymerase, leading to eventual release of the aborted transcript[22] (Figure 1.8).

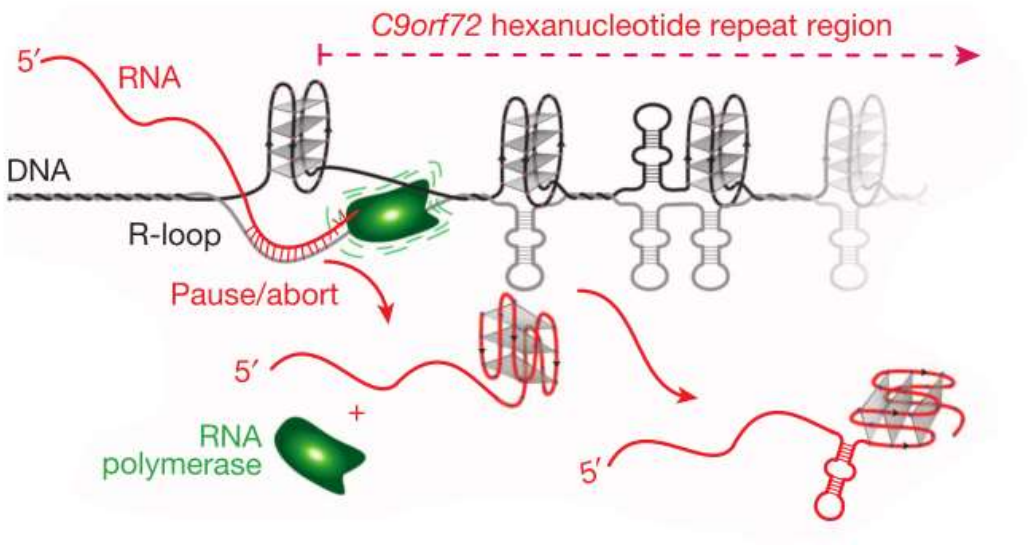


Figure 1.8: Transcriptional regulation of *C9orf72* by G_4C_2 repeat expansions. G-quadruplex structures in the DNA impair processivity of RNA polymerase. Figure adapted from REF. 22.

1.3.3 Regulation of DNA Synthesis

The transient strand separation requisite for DNA replication is predicted to facilitate quadruplex formation in G-rich regions, since this event sequesters the QFS from its complementary strand, while quadruplex folding alleviates the supercoiling introduced ahead of the replication fork. These putative quadruplexes are speculated to stall translocation of the replication machinery (Figure 1.9), thus requiring the unwinding activities of specialized DNA helicases—the absence of which can result in entire deletions of the G-rich motif. For example, London *et al* have observed that patient cells expressing a defective variant of the FANCJ quadruplex-specific DNA helicase are characterized by genomic deletions in regions predicted to adopt quadruplex folds[65]. Interestingly, loss-of-function mutations in quadruplex-specific DNA helicases is a predominant characteristic of human disorders associated with genomic instability[66, 67]—suggesting that G-quadruplex structures, and their unfolding, are also important modulators of genome integrity.

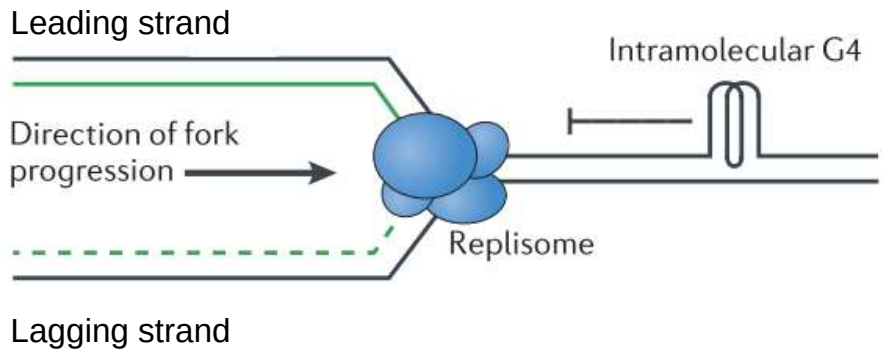


Figure 1.9: Intramolecular quadruplex formation in the leading strand might stall the replisome. Figure adapted from REF. 17.

1.3.4 Pathogen Antigenic Variation

Cahoon and Steifert have reported that a 16-base pair G-rich sequence within the pilin locus of *Neisseria gonorrhoeae*—the causative pathogen of gonorrhea—regulates homologous recombination events that control the structure of the antigen (pilin protein) displayed on the surface of the bacterium[68]. Although many pilin genes are encoded by the bacterial genome, only the gene in the *pilE* locus is expressed. High-frequency recombination reactions between the multiple silent pilin loci and the *pilE* gene are one of three antigenic variation (Av) strategies adopted by this pathogen to evade the host surveillance system[69]. Quadruplex formation by this sequence was demonstrated *in vitro*, and DNA sequence analysis of Av-deficient colonies revealed that they all bore mutations predicted to destabilize the quadruplex fold. Conversely, mutations to the sequence that did not disrupt quadruplex formation were innocuous to the Av capacity of the bacterium, demonstrating that the quadruplex structure, and not the sequence itself, was the critical modulator of recombination-mediated antigenic variation. While the molecular basis for this regulation is currently lacking, a similar QFS motif has also been reported to modulate antigenic variation in a different pathogen class[70].

1.4 (SELECTED) FUNCTIONS OF RNA G-QUADRUPLEXES

1.4.1 Regulation of Telomerase

Two QFS motifs have been proposed to regulate the function of telomerase: (1) the G-rich element present on the 5'-end of the RNA subunit (TERC) of the enzyme; (2) telomeric RNA repeats (TERRA) formed by RNA pol II-mediated transcription of the telomeric DNA C-rich strand. Gros *et al* showed that a quadruplex structure formed on the 5'-end of TERC existed in dynamic equilibrium with the canonical P1a helix established as the requisite element for boundary definition to establish the limits of reverse transcription on the 3'-template sequence[71] (Figure 1.10). The authors also observed that a quadruplex-interactive small molecule shifted the equilibrium towards the quadruplex conformation, leading to speculation that such ligands could be used to modulate the structural conformation of TERC, and thus the functional state of telomerase[71]. Further, Lattmann *et al* utilized RNA immunoprecipitation assays to show that RHAU[72]—an RNA helicase with specific quadruplex-unfolding activity, was associated with telomerase through a stable quadruplex structure on the 5'-end of TERC[73]—suggesting that such mechanisms for modulating telomerase function through TERC structural conformation might exist in natural systems.

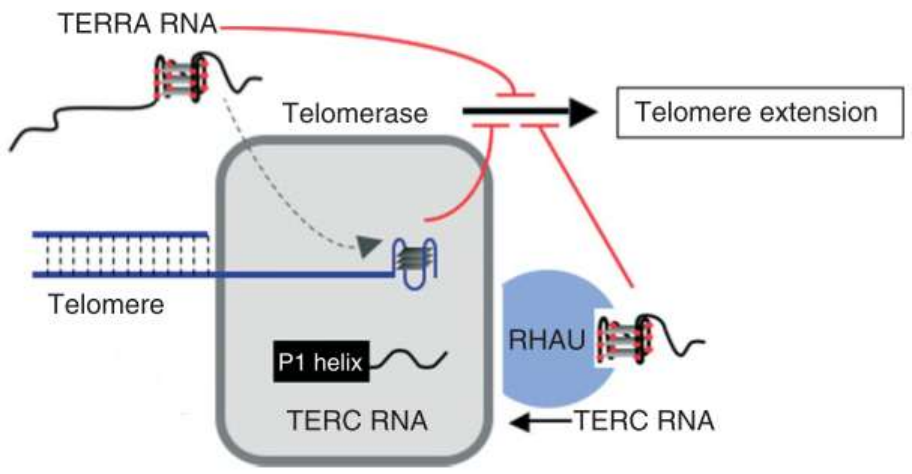


Figure 1.10: Proposed models for regulation of telomerase activity by RNA G-quadruplexes. Quadruplex formation on the 5'-end of TERC impedes formation of the functionally competent P1 helix. RHAU is proposed to unfold the quadruplex and reactivate telomerase. TERRA has been shown to inhibit telomerase activity by either direct interaction with the enzyme or formation of an intermolecular RNA-DNA quadruplex with the telomeres. Figure adapted from REF. 20.

Azzalin *et al* provided the first evidence that the C-rich strand of telomeric DNA is transcribed by the DNA-dependent RNA pol II[74]. Quadruplex formation by this sequence *in vitro* has been confirmed by various reports[44, 75, 76], while evidence for its *in vivo* existence is beginning to emerge[77]. TERRA has been shown to interact directly, and separately, with both telomerase and TERC[78]. Importantly, the interaction with the former results in potent inhibition of its telomere extension activity *in vitro*[78]. Although not yet completely elucidated, the mechanism for inhibition has been proposed to involve the sequestration of active telomerase from the DNA substrate[78]. (The affinity of TERRA for telomerase is higher than that of the DNA substrate[78].) It is currently unclear whether *intramolecular* quadruplex formation by TERRA is requisite for telomerase binding and sequestration, and some have suggested that the *intermolecular* quadruplex formed between TERRA and telomeric DNA might contribute to the inhibitory effect[27].

1.4.2 Regulation of pre-mRNA Processing (Polyadenylation)

Decorsière *et al* have reported that an RNA guanine quadruplex located downstream of the p53 polyadenylation signal was required to rescue p53 3'-end processing in cells under genotoxic stress[79]. Interestingly, the quadruplex motif activated p53 processing and expression under conditions (UV irradiation and doxorubicin treatment) in which global mRNA expression is otherwise suppressed due, in part, to the sequestration of critical processing factors in incompetent complexes. Quadruplex-mediated activation was dependent on direct interactions between the quadruplex and heterogeneous nuclear ribonucleoprotein (hnRNP) H/F[79]. The authors further proposed a model where the quadruplex-protein complex formed on the 3'-end of the mRNA was able to impede the sequestration of CsfF—a crucial component of the processing machinery—by inhibitory complexes (Figure 1.11).

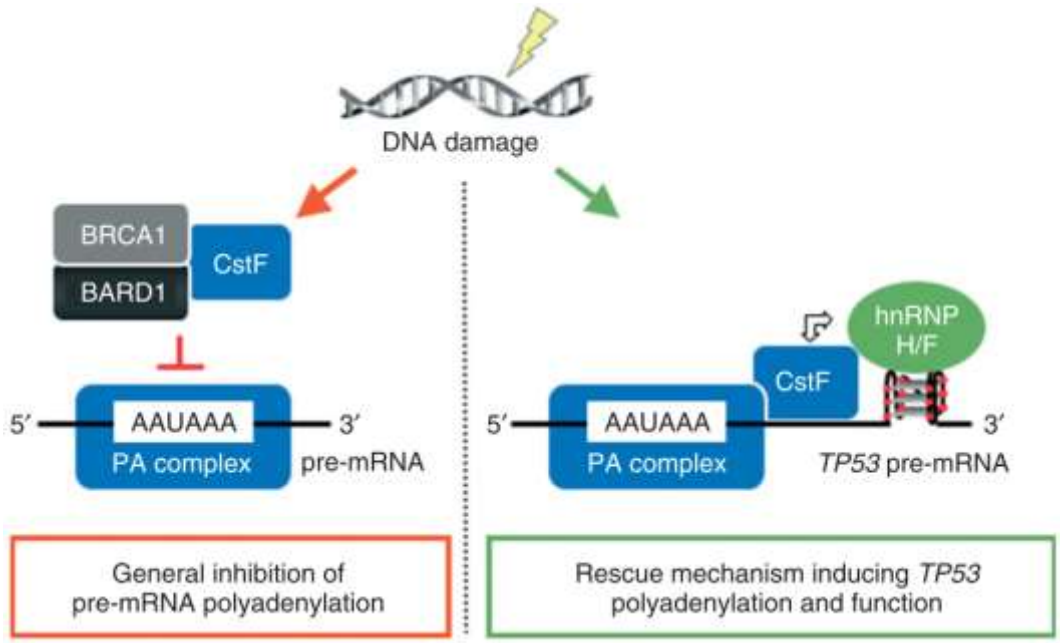


Figure 1.11: Effect of G-quadruplexes on poly-adenylation. DNA damage normally leads to sequestration of CstF by BRCA1 and BARD1 (Left). hnRNP H/F binds an RNA G-quadruplex in the 3'-UTR of TP53 and recruits CstF to the poly(A) signal, thus preventing its inactivation by BRCA1/BARD1. Figure adapted from REF. 20.

1.4.3 Regulation of pre-mRNA Processing (Alternative Splicing)

G-quadruplexes have also been shown to act as *cis* regulators of alternative splicing reactions *in vivo* (Figure 1.12). For example, a G-rich motif in intron 6 of the mRNA encoding human telomerase has been proposed to act as a silencing element by promoting alternative splicing reactions that lead to an accumulation of the inactive transcripts[80]. Quadruplex formation by this sequence has been demonstrated *in vitro*, and a quadruplex-binding small molecule with moderate selectivity for this RNA quadruplex was shown to inhibit telomerase *in vivo*[80]. The authors further demonstrated that cells treated with this ligand showed an accumulation of the inactive transcripts, while retaining the same global transcription level of the mRNA relative to control cells[80].

Didiot *et al* also demonstrated that a G-quadruplex proximal to an alternative splice site in exon 15 of the pre-mRNA for fragile x mental retardation protein (FMRP) modulated splicing events that controlled the expression levels of the different protein isoforms[81]. Interestingly, this quadruplex structure is simultaneously the site of binding for FMRP to its encoding mRNA[82] (FMR1), thus suggesting a mechanism where FMRP regulates its own splicing events by binding to the G-quadruplex.

G-quadruplexes in intron 3 of the TP53 pre-mRNA have also been postulated to act as enhancers of alternative splicing events that control the excision or inclusion of intron 2 in the mature mRNA[83]. Disrupting mutations in the quadruplex structures were reported to decrease intron 2 excision by 30 % in a reporter assay. Conversely, a small molecule ligand that stabilized the quadruplex structures amplified intron 2 excision *in vivo*. Interestingly, a bioinformatics study has identified putative QFS motifs at the 5'-end of the first intron in ~ 50 % of all human protein-coding genes, suggesting that these elements might be broadly applied as *cis*-regulators of mRNA splicing[84].

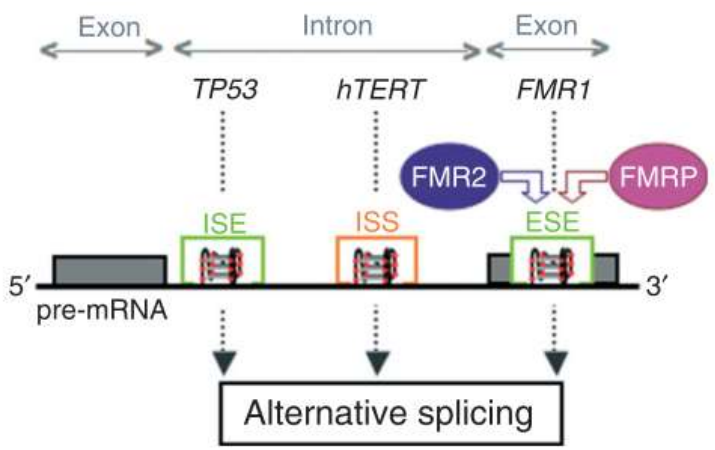


Figure 1.12: Effects of G-quadruplexes on alternative splicing. An RNA G-quadruplex in intron 6 of the human telomerase pre-mRNA silences telomerase expression by activating alternative splicing to generate inactive transcripts. An RNA G-quadruplex in exon 15 of FMR1 enhances splicing events for different isoforms. G-quadruplex in intron 3 of TP53 regulates splicing reactions on intron 2. ISE: intronic splicing enhancer; ISS: intronic splicing silencer; ESE: exonic splicing enhancer. Figure adapted from REF. 20.

1.4.4 Regulation of Translation

Perhaps the most well characterized functional roles of RNA G-quadruplexes are their regulatory effects on translation. Indeed bi-directional translation modulation (stimulation or inhibition) has been reported for RNA quadruplexes in a manner that is broadly dependent on the mechanism of translation initiation adopted by the ribosomes (Figure 1.13). *Cap-dependent translation*[85]—whereby the 40S ribosomal subunit recognizes a methylated guanosine residue on the 5'-end of the mRNA before 'scanning' to the initiation codon—has been reported, in most[86-96] but not all[95, 97] cases, to be inhibited by the presence of quadruplex structures in the 5'-untranslated region (UTR) of the message (Figure 1.13). Conversely, *cap-independent translation*[85]—where structural elements within the 5'-UTR suffice to recruit the 40S ribosomal subunit to the mRNA—has been shown (in few examples[98, 99]) to be activated by G-quadruplex structures.

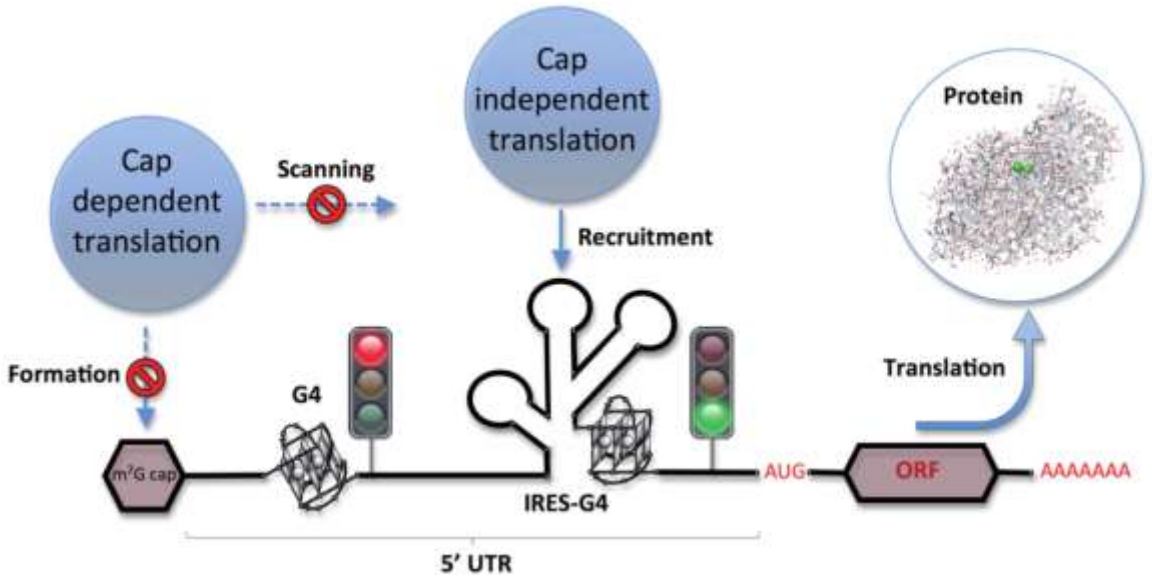


Figure 1.13: Effect of 5'-UTR RNA G-quadruplexes on translation. G-quadruplexes, in most examples, inhibit cap-dependent translation. In two reported examples, G-quadruplexes in IRESs enhance cap-independent translation. Figure adapted from REF. 18.

Schaeffer *et al* provided the first demonstration of this effect by showing that a quadruplex structure originally discovered in the coding sequence of FMR1 repressed translation when placed in the 5'-UTR of a luciferase reporter[82]. Balasubramanian and coworkers subsequently elaborated on this effect by showing that a G-rich motif present in the 5'-UTR of the human *NRAS* mRNA formed a G-quadruplex *in vitro* and repressed translation in the context of a similar reporter assay[86]. Subsequent reports from the same group showed that this inhibitory effect is modulated by the position of the G-rich motif in the context of the *NRAS* UTR,[100] and amplified by pharmacological agents known to bind and stabilize the G-quadruplex *in vitro*[101]. Importantly, this G-rich motif is conserved in both its sequence and position in orthologs of *NRAS* present in different vertebrate species[86], suggesting that this element might be an evolutionarily conserved modulator of *NRAS* expression.

Quadruplex formation and translation repression by 5'-UTR G-rich motifs have been reported for genes involved in diverse cellular processes (*ZIC-1*[93], *MT3-MMP*[92], *TRF2*[94], *ERS1*[89], *BCl2*[88], *EBAG9*[87], *FZD2*[87], *BARHL1*[87], *NCAM2*[87], *THRA*[87], *AASDHPPT*[87], *ADAM10*[96], *AKTIP*[95], *CTSB*[95])—suggesting that this mode of regulating gene expression might be broadly applied *in vivo*. The prevailing mechanistic models underpinning this regulation depend on the relative position of the quadruplex in the 5'-UTR: quadruplex structures proximal to the 5'-cap structure of the mRNAs are predicted to impede assembly of the pre-initiation complex (43S) on the message[18] (Figure 1.13). Additionally, quadruplex structures distal from the 5'-cap are proposed to stall the scanning process required for recognition of the initiation codon[18] (Figure 1.13). Interestingly, both of these models were previously proposed to explain translation inhibition by other structural elements in the 5'-UTR[102, 103], suggesting that this regulatory effect is not unique to quadruplex structures and might be a general feature of thermodynamically stable structures in the 5'-UTR.

These quadruplex-mediated repressive effects on translation are not limited to those in the 5'-UTR. Endoh *et al* previously reported translation repression by quadruplex structures inserted into the coding regions of reporter mRNAs[104]. Evidently, the quadruplexes were impervious to the intrinsic helicase activities of the ribosomes, with the degree of resistance dependent on the thermodynamic stabilities of the structures. The authors further proposed a model where quadruplex formation in the mRNA impeded entry into the A-site of the ribosome, since the dimensions of the quadruplex structure exceeded those required for facile ribosomal access[104]. In a different study, clusters of G-quadruplex structures in the coding region of the mRNA for EBNA1—an Epstein-Barr virus maintenance protein—were observed to downregulate translation[105]. The authors proposed that these structures act as *cis* regulatory elements to curtail EBNA1 expression in order to limit antigen

presentation to the host's T-cells[105]. Additionally, Arora and Suess have reported that a quadruplex in the 3'-UTR of the *PIM1* proto-oncogene repressed translation in the context of a reporter assay *in vivo*[106]. Together, these reports suggest that the regulatory effects of quadruplex structures could occur at the three different steps of translation (initiation, elongation, and termination).

Despite the prevalence of quadruplex-mediated inhibitory effects on cap-dependent translation, a few groups have reported translation stimulation by quadruplexes located in the 5'-UTR. Maiti and coworkers have reported that G-quadruplex structures in the 5'-UTRs of *FOXE3*[95] and *TGFβ2*[97] enhanced translation of reporter genes *in vivo*. Interestingly, the quadruplex structure in the latter activated translation only when it was placed in the context of the natural *TGFβ2* 5'-UTR, whereas inhibitory effects were observed when the isolated QFS was inserted into the artificial UTR of a reporter gene[97]. This result suggests that future efforts to evaluate the biological roles of quadruplex structures should refrain from conclusive assertions if the structures are studied outside their natural sequence context.

Internal entry of ribosomes to facilitate cap-independent translation[85] has also been reported to be regulated by G-quadruplex structures. Two examples exist to date to show that, when present within internal ribosome entry sites (IRES), G-quadruplexes are necessary to preserve the functionally competent structural configuration of the IRES: Bonnal *et al* reported that deletion of a QFS in the IRES of the mRNA encoding FGF-2 led to abrogation of translation initiation at downstream AUG codons[98]. The authors further showed that the G-quadruplex is part of a cluster of structures (including multiple step-loops) that are required for proper folding of the IRES. Basu and coworkers have also reported that a 17-nt G-rich sequence within IRES A of the human *VEGF* mRNA is necessary for cap-independent translation[99]. The authors provided evidence that this sequence, because of its redundancy of G-tracts, was able to adopt multiple quadruplex folds—

each of which displayed differential activating effects on translation. However, a recent report by Cammas *et al* presents evidence that this G-rich motif is functionally silent, and thus dispensable for cap-independent translation of *VEGF*[107], suggesting that additional investigations to ascertain an unambiguous nexus between VEGF expression and quadruplex formation by this sequence are necessary.

1.5 STRATEGIES FOR RECOGNITION

1.5.1 Antibodies and Peptides

The binding reactions between selected antibody fragments and QFS motifs provide some of the best evidence for the existence of G-quadruplex structures *in vivo*[13-16]. Schaffitzel *et al* provided the first demonstration of antibody-based recognition of quadruplex structures by showing that single chain variable fragments (scFvs) derived from a human combinatorial antibody library[108], and screened for binding to the parallel G-quadruplex conformer of the *Stylonychia* telomeric repeats, could react selectively with the macronucleus of the ciliate[13]. Selection by ribosome display[109] led to the identification of two candidate scFvs: one (Sty3) with 1000-fold difference in affinity between the parallel and antiparallel quadruplexes; and the other (Sty49) with similar affinities for both structures. Remarkably, the scFvs enabled determination of the conformational preference of the ciliate telomeric repeats *in vivo*, since no staining was observed with Sty3 (selective for parallel structures), while nuclear foci were detected with Sty49 (with pan affinity for parallel/antiparallel structures).[13]

In later work, Balasubramanian and coworkers used phage display[110] of a human scFv library to screen for antibody fragments selective for an intramolecular G-quadruplex derived from the promoter region of human *c-kit*[111]. The authors employed five rounds of selection to identify HF2,

an scFv with 3000-fold selectivity for the quadruplex target over GC-rich duplex structures[111]. Importantly, HF2 also showed modest preference (300-fold) for its target structure over an intramolecular quadruplex derived from a different region within the same *c-kit* promoter. When assayed for binding to other well characterized quadruplex structures derived from different gene promoter regions (*MYC* and *bcl2*), but not utilized in the negative selections—the scFv variant retained preferential, if diminished (25-fold), selectivity for its target[111]. (This level of discrimination is evidently not useful for recognition of a single quadruplex structure, since HF2 was subsequently utilized in a pull-down assay designed to enrich for quadruplex structures present in human cell lysates[112].) This result highlights an important limitation of this strategy for quadruplex recognition: it is likely that the isolation of a selective binder might be predicated on the anticipation of many possible off-target quadruplexes for use in the negative selections, thus making any unanticipated quadruplex a potential site of non-specific binding.

Another report from the same group utilized the aforementioned technique to identify a pan-quadruplex binder (HF1), using only duplex DNA in the negative selections[113]. Expression of HF1 in mammalian cells, via a plasmid bearing the coding sequence, resulted in differential expression in ~ 10% of all known protein-coding genes, presumably due to the binding reactions between the translated scFv and quadruplex structures *in vivo*[113]. Importantly, the authors observed that differentially expressed genes were enriched in QFS motifs in regulatory regions (promoters and UTRs), suggesting a potential strategy for regulating quadruplex-laden genes using plasmids expressing the protein binder.

Phage display also yielded BG4[14], a single-chain antibody variant with pan affinity for intra/intermolecular DNA quadruplexes, as well as intramolecular RNA quadruplexes. Detection of BG4 nuclear[14] and cytoplasmic[15] foci—for DNA[14] and RNA[15] G-quadruplexes, respectively,

provided the first direct evidence that both types of structures can exist in human cells. In a parallel strategy to achieve the same objective, Henderson *et al* raised monoclonal antibodies against vertebrate telomeric G-quadruplex structures by inoculating mice with the requisite (pre-folded) G-rich DNA sequence[16]. The authors reported strong nuclear staining by the antibody in a manner that was modulated by a quadruplex-binding small molecule or depletion of a quadruplex-specific DNA helicase.

Zinc-finger peptides provide an alternative recognition module for G-quadruplex structures. The first demonstration of this strategy was provided by Isalan *et al*, wherein randomly mutagenized variants of Zif268—a murine transcription factor, were selected by phage display for binding to the G-quadruplex conformer of human telomeric repeats[114]. Three rounds of selection led to the isolation of Gq1, a three-finger peptide with low nM affinity for the quadruplex target, and no detectable affinity for duplex DNA. Modeling studies on the Gq1-quadruplex complex revealed that the second finger was indispensable to the binding affinity and selectivity, and could be combined with other domains from the parent peptide to create modular two-finger peptides that retained recognition, albeit with lower affinity, for the quadruplex target[114]. Importantly, a subsequent report showed that Gq1 could stall primer extension of a quadruplex-laden target by an E. *coli* DNA polymerase, presumably by binding and stabilizing the quadruplex target[115]. Further, the authors reported that Gq1 inhibited telomerase-mediated extension of a G-rich template—an effect that was attributed to the binding of the peptide to the intramolecular quadruplex formed after addition of four or more telomeric repeats.

1.5.2 Small Molecules

The most widely explored strategy for quadruplex targeting involves the application of small, drug-like molecules able to recognize structural features unique to the quadruplex fold[116-120]. Initial interest in these molecules was predicated on the hypothesis that ligands able to induce or stabilize quadruplex formation in the telomeric 3'-overhang could preclude its recognition and extension by telomerase, ultimately inducing telomere dysfunction and/or apoptosis in cancer cells. This section provides a few representative examples of these molecules, while the interested reader is directed to REFs. 116-120 for a more comprehensive summary.

A prime example of this class of molecules is **telomestatin**[121] (Chart 1.1)—a macrocyclic polyoxazole identified by screening the secondary metabolites of *Streptomyces anulatus* for inhibitors of telomerase activity. Telomestatin was found to potently inhibit telomerase-mediated telomeric primer extension in lysates of human B lymphoma cells[121] (IC_{50} = 5 nM). Inhibitory effects have also been reported in both human cell lines[122] and animal models[123] representative of leukemia, where the ligand potentiated telomere dysfunction[122] and decreased tumor volumes[123] in the former and latter, respectively.

Also, Kim *et al* provided evidence for a direct, specific interaction between telomestatin and the G-quadruplex conformer of the human telomeric repeats—with the ligand either inducing or stabilizing, or both, quadruplex formation[124], thus leading to speculation that telomestatin induces its effect by binding to the telomeric quadruplexes. Molecular dynamics simulations revealed that the lowest energy conformation for the complex was one where two ligand molecules were stacked on the terminal tetrads of the quadruplex. More recent efforts have thus sought to tether two equivalents of a telomestatin derivative (L2H2-6OTD, Chart 1.1) using a flexible linker that allows simultaneous interactions on both terminal tetrads by one equivalent of the ligand[125]. Although

this covalent dimer (Chart 1.1) was observed to be two-fold more potent for telomerase inhibition *in vitro*[125], the additive, rather than synergistic, effects observed are unlikely for a molecule interacting simultaneously with both tetrads.

Porphyrins represent another example of quadruplex-interactive agents widely explored for their anti-telomerase activity. Hurley and coworkers first reported that 5,10,15,20-tetra-(N-methyl-4-pyridyl)porphine (**TMPyP4**, Chart 1.1) bound G-quadruplex structures formed by the human telomeric repeats[126]. Optical spectroscopic studies revealed that stacking interactions were the predominant mode of binding, and titration experiments showed that two equivalents of the ligand were associated with the quadruplex target[126]. Further, ligand binding selectively sensitized the G residues of the terminal tetrads to photocleavage, suggesting that TMPyP4 bound—as did telomestatin—by stacking to the terminal tetrads of the quadruplex. Importantly, TMPyP4 induced dose-dependent inhibition of telomerase activity in both HeLa cell extracts[126] and whole tumor cells[117, 127].

To improve the ability of the porphyrins to discriminate between quadruplex and duplex targets—an important criterion for any quadruplex-recognition strategy, Dixon *et al* introduced a coordinating metal ion in the center of the ring, and appended flexible cationic groups directed away from the porphyrin core[128]. The resulting **Mn(III) porphyrin** (Chart 1.1) showed a 10^4-fold higher affinity for the quadruplex target over GC- or AT-rich duplexes. The authors speculated that the bulky cationic groups exacerbated the steric clashes that occur when porphyrins intercalate into duplex DNA[129], thus increasing preference for the quadruplex target. While the authors do not elaborate on the binding mode of this modified ligand to the quadruplex target, an optimal fit for the binding isotherm was obtained only with a two-site model[128], suggesting that the ligand

preserves the 2:1 stoichiometry previously observed for TMPyP4. Importantly, however, the modified ligand inhibited telomerase activity, with IC_{50} = 580 nM[128].

Another molecule finding increasing utility for intervening in quadruplex function is **pyridostatin** (Chart 1.1). Originally reported by Balasubramanian and coworkers, the ligand was shown to induce significantly greater stabilization of a quadruplex target relative to duplex DNA[130]. Further, the authors showed that pyridostatin displaced POT1[131]—a component of the telomere-protecting shelterin complex[132]—from telomeric DNA *in vitro* and *in cellulo*, resulting in nuclear localization of DNA damage-response markers in the latter context[130]. Subsequent reports confirmed that ligand binding was mediated by stacking interactions with the terminal tetrad[101]. Importantly, pyridostatin, or its RNA-selective derivative[133], was subsequently utilized as a reagent to trap DNA- and RNA G-quadruplexes in human cell nuclei[14] or cytoplasm[15], respectively.

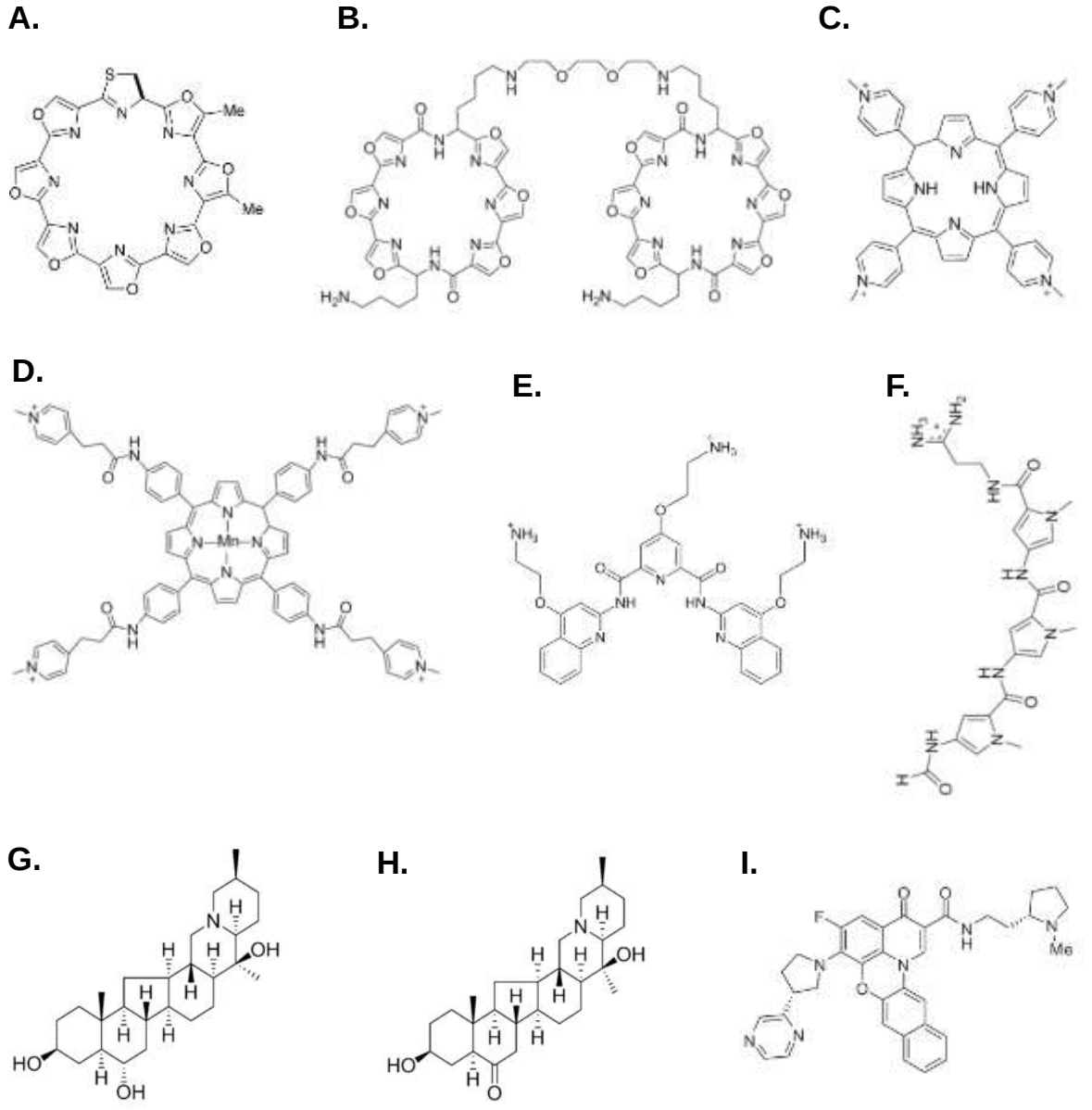


Chart 1.1: Some examples of small-molecule ligands used in quadruplex recognition. telomestatin (**A**). L2H2-6OTD dimer (**B**). TMPyP4 (**C**). Mn(III) porphyrin (**D**). pyridostatin (**E**). distamycin-A (**F**). peimine (**G**). peiminine (**H**). quarfloxin (**I**).

Although the terminal tetrads present the most obvious binding surfaces on G-quadruplexes—and are the basis for designing ligands such as anthracenes[134], quindolines[135], acridines[136], anthraquinones[137], naphthalenes[138], and helicenes[139]—alternative surfaces of the quadruplex have

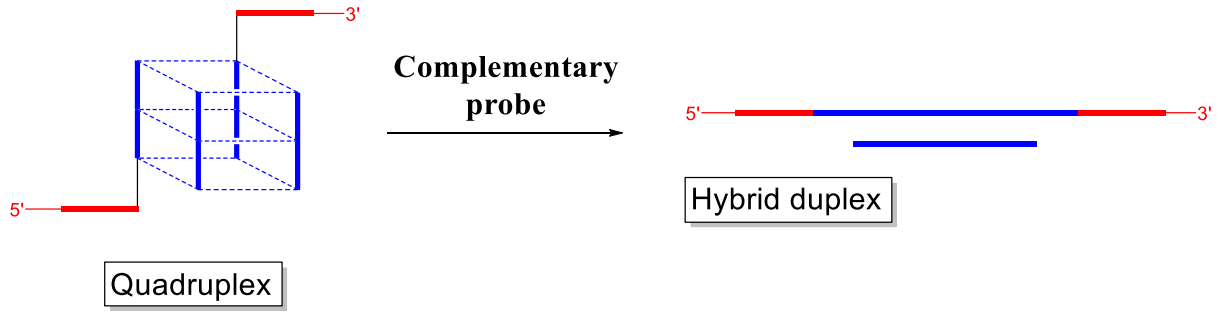
been explored. For example, Martino *et al* reported that four distamycin-A molecules (Chart 1.1) bound in the grooves of an intramolecular DNA G-quadruplex[140]. The solved structure for the complex shows that two antiparallel distamycin dimers are bound in two opposite grooves of the quadruplex target. Importantly, the binding reaction is exergonic ($\Delta G \sim -37$ kJ mol^{-1}), suggesting that tetrad stacking interactions are not the only high-affinity sites available on the quadruplex structure[140]. A similar binding model has been reported for two alkaloids, peimine and peiminine (Chart 1.1)—both of which induce significant thermal stabilization of an intramolecular quadruplex target[141].

The first demonstration of the potential clinical utility of quadruplex-interactive small molecules was provided by Drygin *et al* (Cylene Pharmaceuticals, now defunct). The authors reported that quarfloxin—a fluoroquinolone derivative (Chart 1.1), was able to displace nucleolin from ribosomal DNA G-quadruplexes in tumor cells[142]. The displacement reaction triggered mislocalization of nucleolin, thus impairing ribosome biogenesis. Quarfloxin also showed anti-proliferative activities against a broad range of cancer cell lines, and anti-tumor activity in murine models of several human cancers[142]. Despite these promising results *in vivo*, clinical trials did not progress beyond phase II. Quarfloxin has since been licensed to TetraGene (www.tetragene.com) for further clinical development.

1.5.3 Complementary Oligomeric Molecules

Recognition by oligomeric molecules possessing Watson-Crick complementary nucleobases presents the most systematic method for G-quadruplex targeting. In this case, knowledge of the primary structure of the quadruplex is usually sufficient to inform ligand design, and the binding reaction results in a structural conversion of the G-rich target strand from a quadruplex to a hybrid

duplex (Scheme 1.1), provided the latter is more thermodynamically stable than the starting quadruplex.



Scheme 1.1: QFS recognition by complementary oligomeric molecules.

A range of backbone structures have been employed to present the complementary nucleobases to the quadruplex target (Chart 1.2), including natural DNA[105, 143-150], locked nucleic acids (LNA)[151, 152], 2'-O-methly RNA (2'-OMe)[153], and peptide nucleic acids (PNA)[154-162].

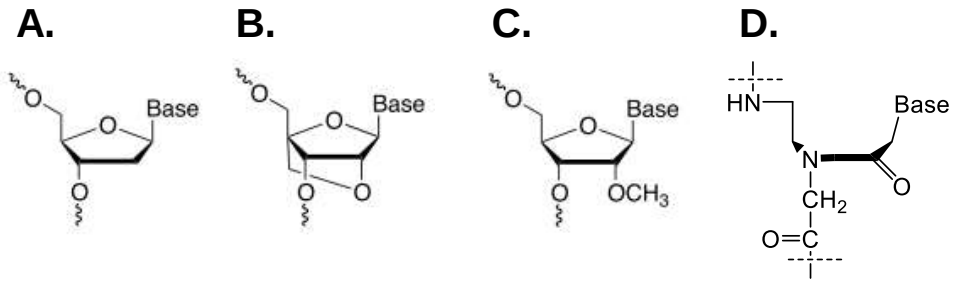


Chart 1.2: Backbone structures of oligomeric complementary molecules used for QFS recognition. DNA (**A**). LNA (**B**). 2'-OMe RNA (**C**). PNA (**D**).

In an early demonstration, Raghuraman and Cech utilized a DNA oligomer complementary to the *Oxytricha* telomeric repeats as a probe to elucidate the mechanistic details of binding between the quadruplex structure formed by the latter and a telomere-binding protein[143]. The rate of accumulation of the QFS-probe duplex was independent of the concentration of the latter, but accelerated by quadruplex-destabilizing conditions, such as elevated temperatures and low [K^+],

thus suggesting that quadruplex unfolding was the rate-limiting step[143]. The authors then used duplex accumulation to track quadruplex opening, and compared the rate of protein binding to the inferred rate of quadruplex opening. Expectedly, the opening step, and thus protein binding, was accelerated by quadruplex-destabilizing conditions, such as elevated temperatures and decreased [K^+][143]. Other demonstrations involving complementary DNA oligomers have sought to use these molecules as probes to indirectly establish quadruplex formation in the context of GC-rich double-stranded DNA[144, 147-150], or as fuel strands in nanomechanical devices[145, 146].

Murat *et al* also showed that a DNA oligomer complementary to the QFS motifs clustered in the coding region of EBNA1, an Epstein-Barr virus maintenance protein—was able to invade the RNA quadruplex *in vitro*[105]. Further, the probe was able to hybridize to its QFS targets in the context of the EBNA1 mRNA transcript and relieve the repressive effect on protein translation induced by the G-quadruplex structures[105].

Another demonstration incorporated LNA modifications into a complementary DNA strand to improve its binding affinity and hybridization kinetics to a stable DNA quadruplex[152]. Increasing LNA modifications improved the equilibrium association constant relative to the native DNA, with as much as a 25-fold difference in affinity between the unmodified probe and one probe bearing 10 LNA modifications[152]. When evaluated for their abilities to silence luciferase expression *in vivo* by invasion of a QFS motif placed upstream of the luciferase coding sequence, the most potent repressive effects were recorded for the most extensively modified probe. The authors attributed the improved silencing to the enhanced affinity and biological stability induced by LNA modification[152].

A recent report showed that 2'-OMe oligomers complementary to the central region of an RNA QFS could preclude quadruplex formation *in vitro*[153]. However, the G-rich target—once prefolded

into the quadruplex conformer, was impervious to the hybridization by the probe, presumably because the energy investment to unfold the quadruplex target was not compensated for by the ΔG for hybrid formation[153]. Although these probes were reported to stimulate translation *in cellulo*, presumably by inhibiting quadruplex formation by nascent transcripts[153]—the aforementioned limitation accentuates an important design criterion for effective complementary ligands: the chemistry of the probe must be such that the binding reaction is sufficiently exergonic to overcome the thermodynamic barrier posed by the quadruplex.

The unnatural, uncharged character[163, 164] of the PNA backbone provides a means to improve the binding affinity of complementary ligands by alleviating the repulsive interactions that might impede hybridization of the aforementioned probe types. Balasubramanian and coworkers demonstrated that a complementary PNA oligomer could trap the open conformer of a G-quadruplex formed by the human telomeric repeats[154]. Hybridization was zero-order with respect to the PNA probe, synonymous with earlier reports on QFS recognition by alternative probe types.

Our lab has extensively studied the interactions between complementary PNA oligomers and QFS targets. We first showed that a short (7mer) probe complementary to a region of a DNA QFS requisite for stable quadruplex folding was able to invade the target at moderate $[K^+]$[155]. In this example, the free energy of the hybrid was determined by both H-bonding interactions within the duplex region and stacking interactions by overhanging bases extending beyond the hybrid. Quadruplex invasion was subsequently extended to an RNA QFS target, whereby we showed that the binding reaction resulted in a relatively tight hybrid duplex (K_D = 1 nM)[156, 157]. We also observed that quadruplex invasion by the complementary PNAs was more effective for a target with longer, more accessible loop residues, relative to one with shorter loops[158], presumably because the loops of the former target facilitate initial nucleation of the probe prior to invasion.

Importantly, we have begun exploring the functional implications of this recognition mode in the context of biochemical system. In this context, we reported that a complementary PNA oligomer directed to a quadruplex-laden target was able to enhance primer extension by DNA pol η, relative to the starting quadruplex structure[159]. The results summarized in this dissertation extend our knowledge on the functional effects induced by targeting RNA G-quadruplexes with complementary PNA/γPNA probes.

1.5.4 Homologous Oligomeric Molecules

Quadruplex-interactive oligomeric molecules possessing G nucleobases contravene Watson-Crick rules for recognition—instead hybridizing to the QFS by Hoogsteen base pairing, the same interactions that mediate G-tetrad formation in the targets. By presenting the guanine nucleobases at the right intervals and in the appropriate orientations to participate in tetrad formation, these molecules effectively replace the G residues in the DNA/RNA tetrads with those contributed by the probe (Figure 1.14).

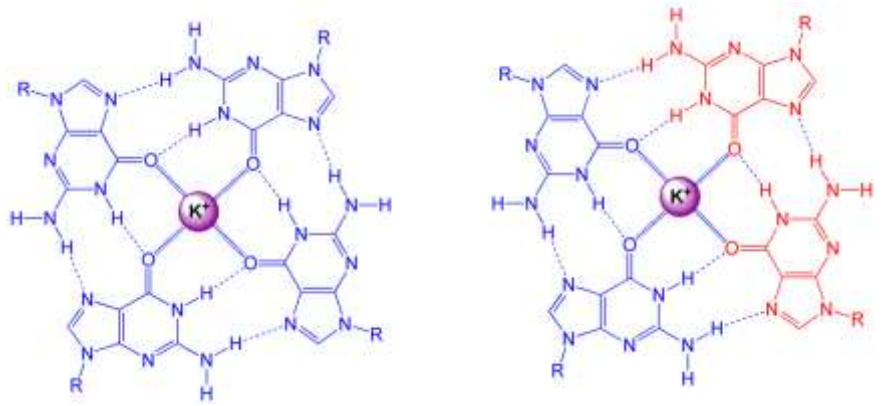
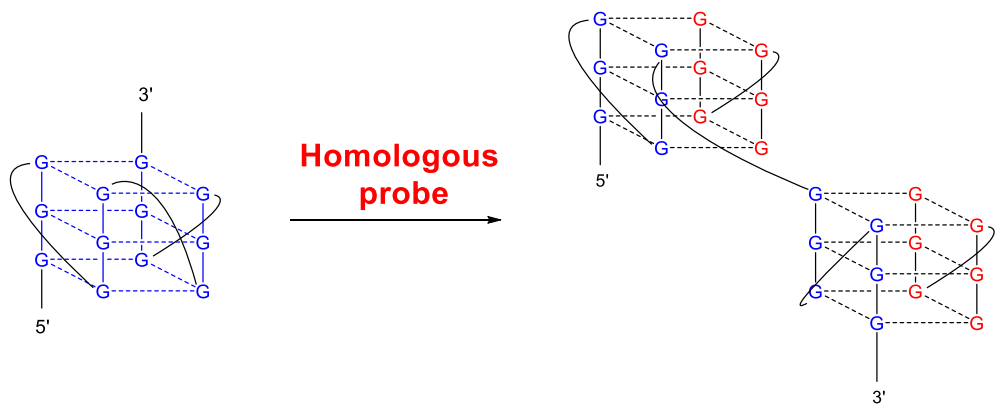


Figure 1.14: Homologous oligomeric molecules substitute guanine(s) in the DNA/RNA G-tetrad.

We provided the first demonstration of this recognition strategy by showing that a PNA oligomer homologous in sequence to the *Oxytricha nova* telomeric repeat could hybridize to its target, resulting in a stable tetramolecular complex[165]. The binding reaction—mediated by the formation of PNA-DNA hybrid tetrads—was strongly exergonic, and the resulting complex displayed cation sensitivity characteristic of quadruplex structures, albeit to a lesser extent than the starting DNA target[165]. Subsequent demonstrations extended this recognition mode to invasion of a stable intramolecular quadruplex formed by a G-rich element derived from the human *MYC* promoter[166]. Therein, we showed that the hybridization end point was a ternary complex consisting of two hybrid DNA-PNA quadruplexes (Scheme 1.2). The binding reaction was again strongly exergonic, resulting in a relatively tight complex (K_D = 5 nM)[166, 167].



Scheme 1.2: Homologous oligomeric molecules convert intramolecular G-quadruplex into hybrid quadruplexes.

More recent work has extended the utility of these probes to invasion of an intramolecular DNA G-quadruplex possessing a long (17 nt) central loop[35]. In this example, the resulting complex consisted of six PNA molecules annealed to the DNA target[35]. The observed association rate was five times slower for this complex than for the ternary structure formed with the *MYC*-derived target. Importantly, however, the PNA-DNA hybrid was more stable than the starting DNA

quadruplex target[35], thus providing a first demonstration of the potential to target these non-canonical quadruplexes with homologous probes.

Homologous recognition has also been extended to RNA G-quadruplex targets. We previously showed that a PNA oligomer homologous to *one* 7-nt sequence within an intramolecularly folded RNA G-quadruplex could hybridize to its target, resulting in *two* RNA-PNA hybrid quadruplexes[156, 157]. Evidently, *true homology*—where the nucleobases on the probe are identical to the sequence of the target—was not requisite for binding. Instead, *G-homology*—where only the G-residues requisite for binding are presented to the target—was sufficient to mediate binding of two probe molecules to the target. This degeneracy in recognition has important implications for tuning probe specificity and will be discussed further in Chapter 4.

Demonstrations of the potential to regulate biological processes by probes of this class are beginning to emerge. For example, Ito *et al* have reported that a G-rich RNA oligomer was able to inhibit EGFP expression *in vivo* by forming an intermolecular quadruplex with a G-homologous sequence inserted into the 5'-UTR or coding region, or both, of the reporter plasmid[168]. Basu and coworkers also reported that a DNA oligomer possessing domains putatively able to form hybrid quadruplex and duplex structures with requisite G-rich and flanking sites, respectively, of an mRNA could downregulate expression of both reporter and endogenous genes[169]. In recent work from our lab, we have reported that a homologous PNA oligomer directed to quadruplex-laden template could stall primer extension by DNA pol η[159]. These functional data provide incentive to explore additional effects of these probes on the biochemical and/or biological functions of G-quadruplexes, and these will be reported in Chapter 4.

1.6 CONTRIBUTIONS TO THE FIELD

The results summarized in this dissertation expand our knowledge on the strategies for G-quadruplex recognition, with particular emphasis on the functional effects induced by the sequence-directed (complementary or homologous) approaches.

In Chapter 2, we explore QFS targeting by complementary gamma (γ)-PNA oligomers. We examine the binding properties and translation inhibitory effects of three γPNA oligomers directed to different segments of an RNA QFS. Our results demonstrate that access to the QFS matter at least as much as affinity in determining the functional outcome of the binding reaction, i.e., potent inhibition by a probe is a consequence of both high affinity for the target sequence and its ability to reach the binding site in the context of a transcript. We also compare γPNA to PNA and 2'-OMe RNA oligomers—both of which have been used to invade stable quadruplexes and modulate their biochemical and/or biological effects. Our results establish the functional superiority of the γPNAs for translation inhibition in our reporter system, since the probe is more potent, more specific or both, relative to the PNA and 2'-OMe RNA.

Chapter 3 extends QFS recognition by γPNAs to targets bearing four G_2- or G_4-tracts. We observe that the respective probes form exceptionally stable hybrids with their targets, with the free energy for complex formation effectively compensating for the energy investment to unfold the quadruplex, where present. Importantly, we show that both probes produce potent and specific repressive effects on translation when directed to the 5'-ends of the respective G-rich targets in the context of a reporter transcript.

We explore quadruplex invasion and translation inhibition by *homologous* probes in Chapter 4. Our results show that the binding reactions result in hybrid quadruplexes more stable than the

starting RNA G-quadruplex. We also observe dose-dependent repressive effects on translation against both the target (quadruplex-laden) and control transcripts. The specificity is improved by introducing a DNA strand able to compete with the mRNA transcript for binding to the PNA probe. Modifications to pre-organize the probe, and thus obviate the need for the competing DNA strand, did not improve specificity, thus necessitating further investigations to elucidate the functional effects of quadruplex invasion by these probes in this context.

Finally, Chapter 5 summarizes our efforts to merge homologous and complementary domains into the same PNA probe. We show that—as isolated substructures—both domains form the expected structures with the target. However, the existence of both substructures simultaneously in the same bound probe molecule has not been proven. We show that the chimeric probe is able to discriminate between two RNA targets; one possessing binding sites to engage both domains of the probe, and the other lacking one of the two requisite binding sites. Further, we observe that the selectivity of the probe is enhanced by introducing modifications that putatively destabilize the substructure by which the deficient target is annealed to the probe.

1.7 REFERENCES

[1] Watson, J. D., and Crick, F. H. C. (1953) Molecular Structure of Nucleic Acids- A Structure for Deoxyribose Nucleic Acid, *Nature 171*, 737-738.

[2] Svozil, D., Kalina, J., Omelka, M., and Schneider, B. (2008) DNA conformations and their sequence preferences, *Nucleic Acids Res. 36*, 3690-3706.

[3] Khuu, P., Sandor, M., DeYoung, J., and Ho, P. S. (2007) Phylogenomic analysis of the emergence of GC-rich transcription elements, *Proc. Natl. Acad. Sci. U. S. A. 104*, 16528-16533.

[4] Gessner, R. V., Frederick, C. A., Quigley, G. J., Rich, A., and Wang, A. H. (1989) The molecular structure of the left-handed Z-DNA double helix at 1.0-A atomic resolution. Geometry, conformation, and ionic interactions of d(CGCGCG). *J. Biol. Chem. 264*, 7921-7935.

[5] Ha, S. C., Lowenhaupt, K., Rich, A., Kim, Y. G., and Kim, K. K. (2005) Crystal structure of a junction between B-DNA and Z-DNA reveals two extruded bases, *Nature 437*, 1183-1186.

[6] Pearson, C. E., Zorbas, H., Price, G. B., and Zannis-Hadjopoulos, M. (1996) Inverted repeats, stem-loops, and cruciforms significance for initiation of DNA replication, *J. Cell. Biochem. 63*, 1-22.

[7] Zhao, J., Bacolla, A., Wang, G., and Vasquez, K. M. (2010) Non-B DNA structure-induced genetic instability and evolution, *Cell. Mol. Life Sci. 67*, 43-62.

[8] Jain, A., Wang, G., and Vasquez, K. M. (2008) DNA triple helices: biological consequences and therapeutic potential, *Biochimie 90*, 1117-1130.

[9] Hoogsteen, K. (1963) The structure of crystals containing a hydrogen-bonded complex of 1-methylthymine and 9-methyladenine, *Acta Cryst. 16*, 907-916.

[10] Bacolla, A., Collins, J. R., Gold, B., Chuzhanova, N., Yi, M., Stephens, R. M., Stefanov, S., Olsh, A., Jakupciak, J. P., Dean, M., Lempicki, R. A., Cooper, D. N., and Wells, R. D. (2006) Long homopurine*homopyrimidine sequences are characteristic of genes expressed in brain and the pseudoautosomal region, *Nucleic Acids Res. 34*, 2663-2675.

[11] Gellert, M., Lipsett, M. N., and Davies, D. R. (1962) Helix formation by guanylic acid, *Proc. Natl. Acad. Sci. U. S. A. 48*, 2013–2018.

[12] Huppert, J. L. (2010) Structure, location and interactions of G-quadruplexes, *FEBS J. 277*, 3452-3458.

[13] Schaffitzel, C., Berger, I., Postberg, J., Hanes, J., Lipps, H. J., and Pluckthun, A. (2001) In vitro generated antibodies specific for telomeric guanine-quadruplex DNA react with Stylonychia lemnae macronuclei, *Proc. Natl. Acad. Sci. U. S. A. 98*, 8572-8577.

[14] Biffi, G., Tannahill, D., McCafferty, J., and Balasubramanian, S. (2013) Quantitative visualization of DNA G-quadruplex structures in human cells, *Nat. Chem. 5*, 182-186.

[15] Biffi, G., Di Antonio, M., Tannahill, D., and Balasubramanian, S. (2014) Visualization and selective chemical targeting of RNA G-quadruplex structures in the cytoplasm of human cells, *Nat. Chem. 6*, 75-80.

[16] Henderson, A., Wu, Y., Huang, Y. C., Chavez, E. A., Platt, J., Johnson, F. B., Brosh, R. M., Jr., Sen, D., and Lansdorp, P. M. (2014) Detection of G-quadruplex DNA in mammalian cells, *Nucleic Acids Res. 42*, 860-869.

[17] Bochman, M. L., Paeschke, K., and Zakian, V. A. (2012) DNA secondary structures: stability and function of G-quadruplex structures, *Nat. Rev. Genet. 13*, 770-780.

[18] Bugaut, A., and Balasubramanian, S. (2012) 5'-UTR RNA G-quadruplexes: translation regulation and targeting, *Nucleic Acids Res. 40*, 4727-4741.
[19] Maizels, N., and Gray, L. T. (2013) The G4 genome, *PLoS Genet. 9*, e1003468.
[20] Millevoi, S., Moine, H., and Vagner, S. (2012) G-quadruplexes in RNA biology, *Wiley Interdiscip. Rev. RNA 3*, 495-507.
[21] Rhodes, D., and Lipps, H. J. (2015) G-quadruplexes and their regulatory roles in biology, *Nucleic Acids Res. 43*, 8627-8637.
[22] Haeusler, A. R., Donnelly, C. J., Periz, G., Simko, E. A., Shaw, P. G., Kim, M. S., Maragakis, N. J., Troncoso, J. C., Pandey, A., Sattler, R., Rothstein, J. D., and Wang, J. (2014) C9orf72 nucleotide repeat structures initiate molecular cascades of disease, *Nature 507*, 195-200.
[23] Burge, S., Parkinson, G. N., Hazel, P., Todd, A. K., and Neidle, S. (2006) Quadruplex DNA: sequence, topology and structure, *Nucleic Acids Res. 34*, 5402-5415.
[24] Williamson, J. R., Raghuraman, M. K., and Cech, T. R. (1989) Monovalent cation-induced structure of telomeric DNA The G-quartet model, *Cell 59*, 871-880.
[25] Qin, Y., and Hurley, L. H. (2008) Structures, folding patterns, and functions of intramolecular DNA G-quadruplexes found in eukaryotic promoter regions, *Biochimie 90*, 1149-1171.
[26] Zhang, A. Y., Bugaut, A., and Balasubramanian, S. (2011) A sequence-independent analysis of the loop length dependence of intramolecular RNA G-quadruplex stability and topology, *Biochemistry 50*, 7251-7258.
[27] Xu, Y., Kimura, T., and Komiyama, M. (2008) Human telomere RNA and DNA form an intermolecular G-quadruplex, *Nucleic Acids Symp. Ser. (Oxf.) 52*, 169-170.
[28] Williamson, J. R. (1993) Guanine quartets, *Curr. Opin. Struct. Biol. 3*, 357-362.
[29] Williamson, J. R. (1994) G-quartet structures in telomeric DNA, *Annu. Rev. Biophys. Biomol. Struct. 23*, 703-730.
[30] Nakano, S., Miyoshi, D., and Sugimoto, N. (2014) Effects of molecular crowding on the structures, interactions, and functions of nucleic acids, *Chem. Rev. 114*, 2733-2758.
[31] Huppert, J. L., and Balasubramanian, S. (2005) Prevalence of quadruplexes in the human genome, *Nucleic Acids Res. 33*, 2908-2916.
[32] Eddy, J., and Maizels, N. (2006) Gene function correlates with potential for G4 DNA formation in the human genome, *Nucleic Acids Res. 34*, 3887-3896.
[33] Palumbo, S. L., Ebbinghaus, S. W., and Hurley, L. H. (2009) 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. 131*, 10878-10891.
[34] Guedin, A., Gros, J., Alberti, P., and Mergny, J. L. (2010) How long is too long? Effects of loop size on G-quadruplex stability, *Nucleic Acids Res. 38*, 7858-7868.
[35] Kormuth, K. A., Woolford, J. L., Jr., and Armitage, B. A. (2016) Homologous PNA Hybridization to Noncanonical DNA G-Quadruplexes, *Biochemistry 55*, 1749-1757.
[36] Mukundan, V. T., Do, N. Q., and Phan, A. T. (2011) HIV-1 integrase inhibitor T30177 forms a stacked dimeric G-quadruplex structure containing bulges, *Nucleic Acids Res. 39*, 8984-8991.
[37] Mukundan, V. T., and Phan, A. T. (2013) Bulges in G-quadruplexes: broadening the definition of G-quadruplex-forming sequences, *J. Am. Chem. Soc. 135*, 5017-5028.

[38] Chambers, V. S., Marsico, G., Boutell, J. M., Di Antonio, M., Smith, G. P., and Balasubramanian, S. (2015) High-throughput sequencing of DNA G-quadruplex structures in the human genome, *Nat. Biotechnol. 33*, 877-881.
[39] Zhang, D. H., and Zhi, G. Y. (2010) Structure monomorphism of RNA G-quadruplex that is independent of surrounding condition, *J. Biotechnol. 150*, 6-10.
[40] Zhang, D. H., Fujimoto, T., Saxena, S., Yu, H. Q., Miyoshi, D., and Sugimoto, N. (2010) Monomorphic RNA G-quadruplex and polymorphic DNA G-quadruplex structures responding to cellular environmental factors, *Biochemistry 49*, 4554-4563.
[41] Jodoin, R., Bauer, L., Garant, J. M., Mahdi Laaref, A., Phaneuf, F., and Perreault, J. P. (2014) The folding of 5'-UTR human G-quadruplexes possessing a long central loop, *RNA 20*, 1129-1141.
[42] Arora, A., and Maiti, S. (2009) Differential biophysical behavior of human telomeric RNA and DNA quadruplex, *J. Phys. Chem. B 113*, 10515-10520.
[43] Joachimi, A., Benz, A., and Hartig, J. S. (2009) A comparison of DNA and RNA quadruplex structures and stabilities, *Bioorg. Med. Chem. 17*, 6811-6815.
[44] Collie, G. W., Haider, S. M., Neidle, S., and Parkinson, G. N. (2010) A crystallographic and modelling study of a human telomeric RNA (TERRA) quadruplex, *Nucleic Acids Res. 38*, 5569-5580.
[45] Greider, C. W., and Blackburn, E. H. (1985) Identification of a specific telomere terminal transferase activity in Tetrahymena extracts, *Cell 43*, 405-413.
[46] Greider, C. W., and Blackburn, E. H. (1989) A telomeric sequence in the RNA of Tetrahymena telomerase required for telomere repeat synthesis, *Nature 337*, 331-337.
[47] Yu, G. L., Bradley, J. D., Attardi, L. D., and Blackburn, E. H. (1990) In vivo alteration of telomere sequences and senescence caused by mutated Tetrahymena telomerase RNAs, *Nature 344*, 126-132.
[48] Blackburn, E. H. (1991) Structure and function of telomeres, *Nature 350*, 569-573.
[49] Moyzis, R. K., Buckingham , J. M., Cram, L. S., Dani, M., Deaven, L. L., Jones, M. D., Meyne, J., Ratliff, R. L., and Wu, J. R. (1988) A highly conserved repetitive DNA sequence, (TTAGGG)n, present at the telomeres of human chromosomes, *Proc. Natl. Acad. Sci. U. S. A. 85*, 6622-6626.
[50] Henderson, E., Hardin, C. C., Walk, S. K., Tinoco, I. J., and E.H., B. (1987) Telomeric DNA oligonucleotides form novel intramolecular structures containing guanine·guanine base pairs, *Cell 51*, 899-908.
[51] Zahler, A. M., Williamson, J. R., Cech, T. R., and Prescott, D. M. (1991) Inhibition of telomerase by G-quartet DNA structures, *Nature 350*, 718-720.
[52] Kim, N. W., Piatyszek, M. A., Prowse, K. R., Harley, C. B., West, M. D., Ho, P. L., Coviello, G. M., Wright, W. E., Weinrich, S. L., and Shay, J. W. (1994) Specific association of human telomerase activity with immortal cells and cancer, *Science 266*, 2011-2015.
[53] Neidle, S. (2010) Human telomeric G-quadruplex: the current status of telomeric G-quadruplexes as therapeutic targets in human cancer, *FEBS J. 277*, 1118-1125.
[54] Capra, J. A., Paeschke, K., Singh, M., and Zakian, V. A. (2010) G-quadruplex DNA sequences are evolutionarily conserved and associated with distinct genomic features in Saccharomyces cerevisiae, *PLoS Comput. Biol. 6*, e1000861.
[55] Mullen, M. A., Olson, K. J., Dallaire, P., Major, F., Assmann, S. M., and Bevilacqua, P. C. (2010) RNA G-Quadruplexes in the model plant species Arabidopsis thaliana: prevalence and possible functional roles, *Nucleic Acids Res. 38*, 8149-8163.

[56] Rawal, P., Kummarasetti, V. B., Ravindran, J., Kumar, N., Halder, K., Sharma, R., Mukerji, M., Das, S. K., and Chowdhury, S. (2006) Genome-wide prediction of G4 DNA as regulatory motifs: role in Escherichia coli global regulation, *Genome Res. 16*, 644-655.
[57] Brooks, T. A., and Hurley, L. H. (2010) Targeting MYC Expression through G-Quadruplexes, *Genes & Cancer 1*, 641-649.
[58] Siddiqui-Jain, A., Grand, C. L., Bearss, D. J., and Hurley, L. H. (2002) 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. 99*, 11593-11598.
[59] Simonsson, T., Pecinka, P., and Kubista, M. (1998) DNA tetraplex formation in the control region of c-myc, *Nucleic Acids Res. 26*, 1167-1172.
[60] Dexheimer, T. S., Carey, S. S., Zuohe, S., Gokhale, V. M., Hu, X., Murata, L. B., Maes, E. M., Weichsel, A., Sun, D., Meuillet, E. J., Montfort, W. R., and Hurley, L. H. (2009) NM23-H2 may play an indirect role in transcriptional activation of c-myc gene expression but does not cleave the nuclease hypersensitive element III(1), *Mol. Cancer. Ther. 8*, 1363-1377.
[61] Gonzalez, V., Guo, K., Hurley, L., and Sun, D. (2009) Identification and characterization of nucleolin as a c-myc G-quadruplex-binding protein, *J. Biol. Chem. 284*, 23622-23635.
[62] Kang, H. J., and Park, H. J. (2009) Novel molecular mechanism for actinomycin D activity as an oncogenic promoter G-quadruplex binder, *Biochemistry 48*, 7392-7398.
[63] Ou, T. M., Lu, Y. J., Zhang, C., Huang, Z. S., Wang, X. D., Tan, J. H., Chen, Y., Ma, D. L., Wong, K. Y., Tang, J. C., Chan, A. S., and Gu, L. Q. (2007) Stabilization of G-quadruplex DNA and down-regulation of oncogene c-myc by quindoline derivatives, *J. Med. Chem. 50*, 1465-1474.
[64] DeJesus-Hernandez, M., Mackenzie, I. R., Boeve, B. F., Boxer, A. L., Baker, M., Rutherford, N. J., Nicholson, A. M., Finch, N. A., Flynn, H., Adamson, J., Kouri, N., Wojtas, A., Sengdy, P., Hsiung, G. Y., Karydas, A., Seeley, W. W., Josephs, K. A., Coppola, G., Geschwind, D. H., Wszolek, Z. K., Feldman, H., Knopman, D. S., Petersen, R. C., Miller, B. L., Dickson, D. W., Boylan, K. B., Graff-Radford, N. R., and Rademakers, R. (2011) Expanded GGGGCC hexanucleotide repeat in noncoding region of C9ORF72 causes chromosome 9p-linked FTD and ALS, *Neuron 72*, 245-256.
[65] London, T. B., Barber, L. J., Mosedale, G., Kelly, G. P., Balasubramanian, S., Hickson, I. D., Boulton, S. J., and Hiom, K. (2008) FANCJ is a structure-specific DNA helicase associated with the maintenance of genomic G/C tracts, *J. Biol. Chem. 283*, 36132-36139.
[66] Huber, M. D., Lee, D. C., and Maizels, N. (2002) G4 DNA unwinding by BLM and Sgs1p: substrate specificity and substrate-specific inhibition, *Nucleic Acids Res. 30*, 3954-3961.
[67] Mohaghegh, P., Karow, J. K., Brosh, R. M. J., Bohr, V. A., and Hickson, I. D. (2001) The Bloom's and Werner's syndrome proteins are DNA structure-specific helicases, *Nucleic Acids Res. 29*, 2843-2849.
[68] Cahoon, L. A., and Seifert, H. S. (2009) An Alternative DNA Structure Is Necessary for Pilin Antigenic Variation in Neisseria gonorrhoeae, *Science 325*, 764-767.
[69] Kline, K. A., Sechman, E. V., Skaar, E. P., and Seifert, H. S. (2003) Recombination, repair and replication in the pathogenic Neisseriae: the 3 R′s of molecular genetics of two human-specific bacterial pathogens, *Mol. Microbiol. 50*, 3-13.
[70] Walia, R., and Chaconas, G. (2013) Suggested role for G4 DNA in recombinational switching at the antigenic variation locus of the Lyme disease spirochete, *PLoS One 8*, e57792.

[71] Gros, J., Guedin, A., Mergny, J. L., and Lacroix, L. (2008) G-Quadruplex formation interferes with P1 helix formation in the RNA component of telomerase hTERC, *Chembiochem 9*, 2075-2079.
[72] Lattmann, S., Giri, B., Vaughn, J. P., Akman, S. A., and Nagamine, Y. (2010) Role of the amino terminal RHAU-specific motif in the recognition and resolution of guanine quadruplex-RNA by the DEAH-box RNA helicase RHAU, *Nucleic Acids Res 38*, 6219-6233.
[73] Lattmann, S., Stadler, M. B., Vaughn, J. P., Akman, S. A., and Nagamine, Y. (2011) The DEAH-box RNA helicase RHAU binds an intramolecular RNA G-quadruplex in TERC and associates with telomerase holoenzyme, *Nucleic Acids Res. 39*, 9390-9404.
[74] Azzalin, C. M., Reichenbach, P., Khoriauli, L., Giulotto, E., and Lingner, J. (2007) Telomeric repeat containing RNA and RNA surveillance factors at mammalian chromosome ends, *Science 318*, 798-801.
[75] Martadinata, H., Heddi, B., Lim, K. W., and Phan, A. T. (2011) Structure of long human telomeric RNA (TERRA): G-quadruplexes formed by four and eight UUAGGG repeats are stable building blocks, *Biochemistry 50*, 6455-6461.
[76] Yangyuoru, P. M., Zhang, A. Y., Shi, Z., Koirala, D., Balasubramanian, S., and Mao, H. (2013) Mechanochemical Properties of Individual Human Telomeric RNA (TERRA) G-Quadruplexes, *Chembiochem 14*, 1931-1935.
[77] Xu, Y., Suzuki, Y., Ito, K., and Komiyama, M. (2010) Telomeric repeat-containing RNA structure in living cells, *Proc. Natl. Acad. Sci. U. S. A. 107*, 14579-14584.
[78] Redon, S., Reichenbach, P., and Lingner, J. (2010) The non-coding RNA TERRA is a natural ligand and direct inhibitor of human telomerase, *Nucleic Acids Res. 38*, 5797-5806.
[79] Decorsiere, A., Cayrel, A., Vagner, S., and Millevoi, S. (2011) Essential role for the interaction between hnRNP H/F and a G quadruplex in maintaining p53 pre-mRNA 3'-end processing and function during DNA damage, *Genes & Dev. 25*, 220-225.
[80] Gomez, D., Lemarteleur, T., Lacroix, L., Mailliet, P., Mergny, J. L., and Riou, J. F. (2004) Telomerase downregulation induced by the G-quadruplex ligand 12459 in A549 cells is mediated by hTERT RNA alternative splicing, *Nucleic Acids Res. 32*, 371–379.
[81] Didiot, M. C., Tian, Z., Schaeffer, C., Subramanian, M., Mandel, J. L., and Moine, H. (2008) The G-quartet containing FMRP binding site in FMR1 mRNA is a potent exonic splicing enhancer, *Nucleic Acids Res. 36*, 4902-4912.
[82] Schaeffer, C., Bardoni, B., Mandel, J. L., Ehresmann, B., Ehresmann, C., and Moine, H. (2001) The fragile X mental retardation protein binds specifically to its mRNA via a purine quartet motif, *EMBO J. 20*, 4803-4813.
[83] Marcel, V., Tran, P. L., Sagne, C., Martel-Planche, G., Vaslin, L., Teulade-Fichou, M. P., Hall, J., Mergny, J. L., Hainaut, P., and Van Dyck, E. (2011) G-quadruplex structures in TP53 intron 3: role in alternative splicing and in production of p53 mRNA isoforms, *Carcinogenesis 32*, 271-278.
[84] Eddy, J., and Maizels, N. (2008) Conserved elements with potential to form polymorphic G-quadruplex structures in the first intron of human genes, *Nucleic Acids Res. 36*, 1321-1333.
[85] Gebauer, F., and Hentze, M. W. (2004) Molecular mechanisms of translational control, *Nat. Rev. Mol. Cell. Biol. 5*, 827-835.

[86] Kumari, S., Bugaut, A., Huppert, J. L., and Balasubramanian, S. (2007) An RNA G-quadruplex in the 5' UTR of the NRAS proto-oncogene modulates translation, *Nat. Chem. Biol. 3*, 218-221.
[87] Beaudoin, J. D., and Perreault, J. P. (2010) 5'-UTR G-quadruplex structures acting as translational repressors, *Nucleic Acids Res. 38*, 7022-7036.
[88] Shahid, R., Bugaut, A., and Balasubramanian, S. (2010) The BCL-2 5' untranslated region contains an RNA G-quadruplex-forming motif that modulates protein expression, *Biochemistry 49*, 8300-8306.
[89] Balkwill, G. D., Derecka, K., Garner, T. P., Hodgman, C., Flint, A. P., and Searle, M. S. (2009) Repression of translation of human estrogen receptor alpha by G-quadruplex formation, *Biochemistry 48*, 11487-11495.
[90] Derecka, K., Balkwill, G. D., Garner, T. P., Hodgman, C., Flint, A. P., and Searle, M. S. (2010) Occurrence of a quadruplex motif in a unique insert within exon C of the bovine estrogen receptor alpha gene (ESR1), *Biochemistry 49*, 7625-7633.
[91] Halder, K., Wieland, M., and Hartig, J. S. (2009) Predictable suppression of gene expression by 5'-UTR-based RNA quadruplexes, *Nucleic Acids Res. 37*, 6811-6817.
[92] Morris, M. J., and Basu, S. (2009) An unusually stable G-quadruplex within the 5'-UTR of the MT3 matrix metalloproteinase mRNA represses translation in eukaryotic cells, *Biochemistry 48*, 5313-5319.
[93] Arora, A., Dutkiewicz, M., Scaria, V., Hariharan, M., Maiti, S., and Kurreck, J. (2008) Inhibition of translation in living eukaryotic cells by an RNA G-quadruplex motif, *RNA 14*, 1290-1296.
[94] Gomez, D., Guedin, A., Mergny, J. L., Salles, B., Riou, J. F., Teulade-Fichou, M. P., and Calsou, P. (2010) A G-quadruplex structure within the 5'-UTR of TRF2 mRNA represses translation in human cells, *Nucleic Acids Res. 38*, 7187-7198.
[95] Agarwala, P., Pandey, S., and Maiti, S. (2014) Role of G-quadruplex located at 5' end of mRNAs, *Biochim. Biophys. Acta 1840*, 3503-3510.
[96] Lammich, S., Kamp, F., Wagner, J., Nuscher, B., Zilow, S., Ludwig, A. K., Willem, M., and Haass, C. (2011) Translational repression of the disintegrin and metalloprotease ADAM10 by a stable G-quadruplex secondary structure in its 5'-untranslated region, *J. Biol. Chem. 286*, 45063-45072.
[97] Agarwala, P., Pandey, S., Mapa, K., and Maiti, S. (2013) The G-quadruplex augments translation in the 5' untranslated region of transforming growth factor beta2, *Biochemistry 52*, 1528-1538.
[98] Bonnal, S., Schaeffer, C., Creancier, L., Clamens, S., Moine, H., Prats, A. C., and Vagner, S. (2003) A single internal ribosome entry site containing a G quartet RNA structure drives fibroblast growth factor 2 gene expression at four alternative translation initiation codons, *J. Biol. Chem. 278*, 39330-39336.
[99] Morris, M. J., Negishi, Y., Pazsint, C., Schonhoft, J. D., and Basu, S. (2010) An RNA G-Quadruplex Is Essential for Cap-Independent Translation Initiation in Human VEGF IRES, *J. Am. Chem. Soc. 132*, 17831–17839.
[100] Kumari, S., Bugaut, A., and Balasubramanian, S. (2008) Position and stability are determining factors for translation repression by an RNA G-quadruplex-forming sequence within the 5' UTR of the NRAS proto-oncogene, *Biochemistry 47*, 12664-12669.

[101] Bugaut, A., Rodriguez, R., Kumari, S., Hsu, S. T., and Balasubramanian, S. (2010) Small molecule-mediated inhibition of translation by targeting a native RNA G-quadruplex, *Org. Biomol. Chem. 8*, 2771-2776.
[102] Kozak, M. (1986) Influences of mRNA secondary structure on initiation by eukaryotic ribosomes, *Proc. Natl. Acad. Sci. U. S. A. 83*, 2850-2854.
[103] Kozak, M. (1989) Circumstances and mechanisms of inhibition of translation by secondary structure in eucaryotic mRNAs, *Mol. Cell. Biol. 9*, 5134-5142.
[104] Endoh, T., Kawasaki, Y., and Sugimoto, N. (2013) Suppression of gene expression by G-quadruplexes in open reading frames depends on G-quadruplex stability, *Angew. Chem. Int. Ed. Engl. 52*, 5522-5526.
[105] Murat, P., Zhong, J., Lekieffre, L., Cowieson, N. P., Clancy, J. L., Preiss, T., Balasubramanian, S., Khanna, R., and Tellam, J. (2014) G-quadruplexes regulate Epstein-Barr virus-encoded nuclear antigen 1 mRNA translation, *Nat. Chem. Biol. 10*, 358-364.
[106] Arora, A., and Suess, B. (2011) An RNA G-quadruplex in the 3' UTR of the proto-oncogene PIM1 represses translation, *RNA Biol. 8*, 802-805.
[107] Cammas, A., Dubrac, A., Morel, B., Lamaa, A., Touriol, C., Teulade-Fichou, M. P., Prats, H., and Millevoi, S. (2015) Stabilization of the G-quadruplex at the VEGF IRES represses cap-independent translation, *RNA Biol. 12*, 320-329.
[108] Knappik, A., Ge, L., Honegger, A., Pack, P., Fischer, M., Wellnhofer, G., Hoess, A., Wölle, J., Plückthun, A., and Virnekäs, B. (2000) Fully synthetic human combinatorial antibody libraries (HuCAL) based on modular consensus frameworks and CDRs randomized with trinucleotides, *J. Mol. Biol. 296*, 57-86.
[109] Hanes, J., Schaffitzel, C., Knappik, A., and Plückthun, A. (2000) Picomolar affinity antibodies from a fully synthetic naive library selected and evolved by ribosome display, *Nat. Biotechnol. 18*, 1287-1292.
[110] Smith, G. P. (1985) Filamentous fusion phage novel expression vectors that display cloned antigens on the virion surface, *Science 228*, 1315-1317.
[111] Fernando, H., Rodriguez, R., and Balasubramanian, S. (2008) Selective Recognition of a DNA G-Quadruplex by an Engineered Antibody, *Biochemistry 47*, 9365-9371.
[112] Lam, E. Y., Beraldi, D., Tannahill, D., and Balasubramanian, S. (2013) G-quadruplex structures are stable and detectable in human genomic DNA, *Nat. Commun. 4*, 1-8.
[113] Fernando, H., Sewitz, S., Darot, J., Tavare, S., Huppert, J. L., and Balasubramanian, S. (2009) Genome-wide analysis of a G-quadruplex-specific single-chain antibody that regulates gene expression, *Nucleic Acids Res. 37*, 6716-6722.
[114] Isalan, M., Patel, S. D., Balasubramanian, S., and Choo, Y. (2001) Selection of Zinc Fingers that Bind Single-Stranded Telomeric DNA in the G-Quadruplex Conformation, *Biochemistry 40*, 830-836.
[115] Patel, S. D., Isalan, M., Gavory, G., Ladame, S., Choo, Y., and Balasubramanian, S. (2004) Inhibition of human telomerase activity by an engineered zinc finger protein, *Biochemistry 43*, 13452-13458.
[116] Monchaud, D., and Teulade-Fichou, M. P. (2008) A hitchhiker's guide to G-quadruplex ligands, *Org Biomol Chem 6*, 627-636.
[117] Zhang, S., Wu, Y., and Zhang, W. (2014) G-quadruplex structures and their interaction diversity with ligands, *ChemMedChem 9*, 899-911.
[118] Ohnmacht, S. A., and Neidle, S. (2014) Small-molecule quadruplex-targeted drug discovery, *Bioorg Med Chem Lett 24*, 2602-2612.

[119] Neidle, S. (2016) Quadruplex Nucleic Acids as Novel Therapeutic Targets, *J. Med. Chem.* 10.1021/acs.jmedchem.5b01835.
[120] Collie, G. W., and Parkinson, G. N. (2011) The application of DNA and RNA G-quadruplexes to therapeutic medicines, *Chem. Soc. Rev. 40*, 5867-5892.
[121] Shin-ya, K., Wierzba, K., Matsuo, K., Ohtani, T., Yamada, Y., Furihata, K., Hayakawa, Y., and Seto, H. (2001) Telomestatin, a novel telomerase inhibitor from Streptomyces anulatus, *J. Am. Chem. Soc. 123*, 1262-1263.
[122] Tauchi, T., Shin-Ya, K., Sashida, G., Sumi, M., Nakajima, A., Shimamoto, T., Ohyashiki, J. H., and Ohyashiki, K. (2003) Activity of a novel G-quadruplex-interactive telomerase inhibitor, telomestatin (SOT-095), against human leukemia cells: involvement of ATM-dependent DNA damage response pathways, *Oncogene 22*, 5338-5347.
[123] Tauchi, T., Shin-ya, K., Sashida, G., Sumi, M., Okabe, S., Ohyashiki, J. H., and Ohyashiki, K. (2006) Telomerase inhibition with a novel G-quadruplex-interactive agent, telomestatin: in vitro and in vivo studies in acute leukemia, *Oncogene 25*, 5719-5725.
[124] Kim, M. Y., Vankayalapati, H., Shin-Ya, K., Wierzba, K., and Hurley, L. H. (2002) Telomestatin, a Potent Telomerase Inhibitor That Interacts Quite Specifically with the Human Telomeric Intramolecular G-Quadruplex, *J. Am. Chem. Soc. 124*, 2098-2099.
[125] Iida, K., Majima, S., Nakamura, T., Seimiya, H., and Nagasawa, K. (2013) Evaluation of the interaction between long telomeric DNA and macrocyclic hexaoxazole (6OTD) dimer of a G-quadruplex ligand, *Molecules 18*, 4328-4341.
[126] Wheelhouse, R. T., Sun, D., Han, H., Han, F. X., and Hurley, L. (1998) Cationic Porphyrins as Telomerase Inhibitors the Interaction of Tetra-(N-methyl-4-pyridyl)porphine with Quadruplex DNA, *J. Am. Chem. Soc. 120*, 3261–3262.
[127] Izbicka, E., Wheelhouse, R. T., Raymond, E., Davidson, K. K., Lawrence, R. A., Sun, D., Windle, B. E., Hurley, L. H., and Von Hoff, D. D. (1999) Effects of cationic porphyrins as G-quadruplex interactive agents in human tumor cells, *Cancer Res. 59*, 639-644.
[128] Dixon, I. M., Lopez, F., Tejera, A. M., Estève, J. P., Blasco, M. A., Pratviel, G., and Meunier, B. (2007) A G-Quadruplex Ligand with 10000-Fold Selectivity over Duplex DNA, *J. Am. Chem. Soc. 129*, 1502-1503.
[129] Lipscomb, L. A., Zhou, F. X., Presnell, S. R., Woo, R. J., Peek, M. E., Plaskon, R. R., and Williams, L. D. (1996) Structure of a DNA-porphyrin complex, *Biochemistry 35*, 2818-2823.
[130] Rodriguez, R., Müller, S., Yeoman, J. A., Trentesaux, C., Riou, J. F., and Balasubramanian, S. (2008) A Novel Small Molecule That Alters Shelterin Integrity and Triggers a DNA-Damage Response at Telomeres, *J. Am. Chem. Soc. 130*, 15758-15759.
[131] Baumann, P., and Cech, T. R. (2001) Pot1, the Putative Telomere End-Binding Protein in Fission Yeast and Humans, *Science 292*, 1171-1175.
[132] de Lange, T. (2005) Shelterin: the protein complex that shapes and safeguards human telomeres, *Genes & Dev. 19*, 2100-2110.
[133] Di Antonio, M., Biffi, G., Mariani, A., Raiber, E. A., Rodriguez, R., and Balasubramanian, S. (2012) Selective RNA versus DNA G-quadruplex targeting by in situ click chemistry, *Angew. Chem. Intl. Ed. Engl. 51*, 11073-11078.
[134] Rodriguez, R., Pantoş, G. D., Gonçalves, D. P. N., Sanders, J. K. M., and Balasubramanian, S. (2007) Ligand-Driven G-Quadruplex Conformational Switching By Using an Unusual Mode of Interaction, *Angew. Chem. 119*, 5501-5503.

[135] Xu, W., Tan, J. H., Chen, S. B., Hou, J. Q., Li, D., Huang, Z. S., and Gu, L. Q. (2011) Studies on the binding of 5-N-methylated quindoline derivative to human telomeric G-quadruplex, *Biochem. Biophys. Res. Commun. 406*, 454-458.
[136] Collie, G. W., Sparapani, S., Parkinson, G. N., and Neidle, S. (2011) Structural basis of telomeric RNA quadruplex--acridine ligand recognition, *J. Am. Chem. Soc. 133*, 2721-2728.
[137] Sun, D., Thompson, B., Cathers, B. E., Salazar, M., Kerwin, S. M., Trent, J. O., Jenkins, T. C., Neidle, S., and Hurley, L. H. (1997) Inhibition of Human Telomerase by a G-Quadruplex-Interactive Compound, *J. Med. Chem. 40*, 2113-2116.
[138] Collie, G. W., Promontorio, R., Hampel, S. M., Micco, M., Neidle, S., and Parkinson, G. N. (2012) Structural basis for telomeric G-quadruplex targeting by naphthalene diimide ligands, *J. Am. Chem. Soc. 134*, 2723-2731.
[139] Shinohara, K., Sannohe, Y., Kaieda, S., Tanaka, K., Osuga, H., Tahara, H., Xu, Y., Kawase, T., Bandol, T., and Sugiyama, H. (2010) A chiral wedge molecule inhibits telomerase activity, *J. Am. Chem. Soc. 132*, 3778-3782.
[140] Martino, L., Virno, A., Pagano, B., Virgilio, A., Di Micco, S., Galeone, A., Giancola, C., Bifulco, G., Mayol, L., and Randazzo, A. (2007) Structural and thermodynamic studies of the interaction of distamycin A with the parallel quadruplex structure [d(TGGGGT)]4, *J. Am. Chem. Soc. 129*, 16048-16056.
[141] Li, Q., Xiang, J., Li, X., Chen, L., Xu, X., Tang, Y., Zhou, Q., Li, L., Zhang, H., Sun, H., Guan, A., Yang, Q., Yang, S., and Xu, G. (2009) Stabilizing parallel G-quadruplex DNA by a new class of ligands: two non-planar alkaloids through interaction in lateral grooves, *Biochimie 91*, 811-819.
[142] Drygin, D., Siddiqui-Jain, A., O'Brien, S., Schwaebe, M., Lin, A., Bliesath, J., Ho, C. B., Proffitt, C., Trent, K., Whitten, J. P., Lim, J. K., Von Hoff, D., Anderes, K., and Rice, W. G. (2009) Anticancer activity of CX-3543: a direct inhibitor of rRNA biogenesis, *Cancer Res. 69*, 7653-7661.
[143] Raghuraman, M. K., and Cech, T. R. (1990) Effect of monovalent cation-induced telomeric DNA structure on the binding of Oxytricha telomeric protein, *Nucleic Acids Res. 18*, 4543-4552.
[144] Phan, A. T., and Mergny, J. L. (2002) Human telomeric DNA G-quadruplex , i-motif and Watson ± Crick double helix, *Nucleic Acids Res. 30*, 4618-4625.
[145] Li, J. J., and Tan, W. (2002) A Single DNA Molecule Nanomotor, *Nano Lett. 2*, 315–318.
[146] Alberti, P., and Mergny, J. L. (2003) DNA duplex-quadruplex exchange as the basis for a nanomolecular machine, *Proc. Natl. Acad. Sci. U. S. A. 100*, 1569-1573.
[147] Kumar, N., and Maiti, S. (2004) Quadruplex to Watson-Crick duplex transition of the thrombin binding aptamer: a fluorescence resonance energy transfer study, *Biochem. Biophys. Res. Commun. 319*, 759-767.
[148] Kumar, N., and Maiti, S. (2005) The effect of osmolytes and small molecule on Quadruplex-WC duplex equilibrium: a fluorescence resonance energy transfer study, *Nucleic Acids Res. 33*, 6723-6732.
[149] Kumar, N., Sahoo, B., Varun, K. A., Maiti, S., and Maiti, S. (2008) Effect of loop length variation on quadruplex-Watson Crick duplex competition, *Nucleic Acids Res. 36*, 4433-4442.
[150] Arora, A., Nair, D. R., and Maiti, S. (2009) Effect of flanking bases on quadruplex stability and Watson-Crick duplex competition, *FEBS J. 276*, 3628-3640.

[151] Kumar, N., and Maiti, S. (2007) Role of Locked Nucleic Acid Modified Complementary Strand in Quadruplex Watson - Crick Duplex Equilibrium, *J. Phys. Chem. B 111*, 12328-12337.
[152] Kumar, N., Patowary, A., Sivasubbu, S., Petersen, M., and Maiti, S. (2008) Silencing c-MYC Expression by Targeting Quadruplex in P1 Promoter Using Locked Nucleic Acid Trap, *Biochemistry 47*, 13179-13188.
[153] Rouleau, S. G., Beaudoin, J. D., Bisaillon, M., and Perreault, J. P. (2015) Small antisense oligonucleotides against G-quadruplexes: specific mRNA translational switches, *Nucleic Acids Res. 43*, 595-606.
[154] Green, J. J., Ying, L., Klenerman, D., and Balasubramanian, S. (2003) Kinetics of unfolding the human telomeric DNA quadruplex using a PNA trap, *J. Am. Chem. Soc. 125*, 3763–3767.
[155] Datta, B., and Armitage, B. A. (2001) Hybridization of PNA to Structured DNA Targets: Quadruplex Invasion and the Overhang Effect, *J. Am. Chem. Soc. 123*, 9612–9619.
[156] Marin, V. L., and Armitage, B. A. (2005) RNA Guanine Quadruplex Invasion by Complementary and Homologous PNA Probes, *J. Am. Chem. Soc. 127*, 8032–8033.
[157] Marin, V. L., and Armitage, B. A. (2006) Hybridization of Complementary and Homologous Peptide Nucleic Acid Oligomers to a Guanine Quadruplex-Forming RNA, *Biochemistry 45*, 1745–1754.
[158] Gupta, A., Lee, L., Roy, S., Tanious, F. A., Wilson, W. D., Ly, D. H., and Armitage, B. A. (2013) Strand invasion of DNA quadruplexes by PNA: comparison of homologous and complementary hybridization, *ChemBioChem 14*, 1476–1484.
[159] Murphy, C. T., Gupta, A., Armitage, B. A., and Opresko, P. L. (2014) Hybridization of G-quadruplex-forming peptide nucleic acids to guanine-rich DNA templates inhibits DNA polymerase eta extension, *Biochemistry 53*, 5315-5322.
[160] Amato, J., Gabelica, V., Borbone, N., Rosu, F., De Pauw, E., Oliviero, G., Piccialli, G., and Mayol, L. (2008) A short C-rich PNA fragment capable to form novel G-quadruplex-PNA complexes, *Nucleic Acids Symp. Ser. (Oxf) 52*, 167-168.
[161] Amato, J., Oliviero, G., De Pauw, E., and Gabelica, V. (2009) Hybridization of short complementary PNAs to G-quadruplex forming oligonucleotides: An electrospray mass spectrometry study, *Biopolymers 91*, 244-255.
[162] Amato, J., Pagano, B., Borbone, N., Oliviero, G., Gabelica, V., Pauw, E. D., D'Errico, S., Piccialli, V., Varra, M., Giancola, C., Piccialli, G., and Mayol, L. (2011) Targeting G-quadruplex structure in the human c-Kit promoter with short PNA sequences, *Bioconjug. Chem. 22*, 654-663.
[163] Nielsen, P. E., Egholm, M., Berg, R. H., and Buchardt, O. (1991) Sequence-selective recognition of DNA by strand displacement with a thymine-substituted polyamide, *Science 254*, 1497-1500.
[164] Egholm, M., Buchardt, O., Christensen, L., Behrens, C., Freier, S. M., Driver, D. A., Berg, R. H., Kim, S. K., Norden, B., and Nielsen, P. E. (1993) PNA hybridizes to complementary oligonucleotides obeying the Watson–Crick hydrogen-bonding rules, *Nature 365*, 566-568.
[165] Datta, B., Schmitt, C., and Armitage, B. A. (2003) Formation of a PNA 2 - DNA 2 Hybrid Quadruplex, *J. Am. Chem. Soc. 125*, 4111-4118.
[166] Roy, S., Tanious, F. A., Wilson, W. D., Ly, D. H., and Armitage, B. A. (2007) High-affinity homologous peptide nucleic acid probes for targeting a quadruplex-forming sequence from a MYC promoter element, *Biochemistry 46*, 10433–10443.

[167] Lusvarghi, S., Murphy, C. T., Roy, S., Tanious, F. A., Sacui, I., Wilson, W. D., Ly, D. H., and Armitage, B. A. (2009) Loop and backbone modifications of peptide nucleic acid improve g-quadruplex binding selectivity, *J. Am. Chem. Soc. 131*, 18415-18424.
[168] Ito, K., Go, S., Komiyama, M., and Xu, Y. (2011) Inhibition of translation by small RNA-stabilized mRNA structures in human cells, *J. Am. Chem. Soc. 133*, 19153-19159.
[169] Bhattacharyya, D., Nguyen, K., and Basu, S. (2014) Rationally induced RNA:DNA G-quadruplex structures elicit an anticancer effect by inhibiting endogenous eIF-4E expression, *Biochemistry 53*, 5461-5470.

CHAPTER 2

G-QUADRUPLEX INVASION AND TRANSLATION INHIBITION BY COMPLEMENTARY γPNA OLIGOMERS

2.1 INTRODUCTION

Many reports establish aberrant gene expression as crucial to the pathogenesis of several diseases[1-4]—making synthetic molecules able to regulate gene expression valuable to the biomedical research community as either therapeutics or, more broadly, as exploratory tools to delineate pathways that contribute to the diseased state. While the targets for early synthetic efforts were limited to proteins with anomalous expression in relevant disease tissues, interest has slowly accrued in recognizing additional elements annotated with regulatory roles in gene expression. Towards this end, guanine (G)-quadruplexes—nucleic acid structural elements formed by specific arrangements of guanine nucleotides—have become important as targets for pharmacological intervention[5-9] (reviewed in Chapter 1).

Although several parallel strategies for quadruplex recognition are being developed (reviewed in Chapter 1), the most intuitive, systematic methods rely on oligomeric molecules presenting complementary nucleobases at appropriate intervals to form Watson-Crick base pairs[10] with the quadruplex-forming sequence (QFS). An early demonstration utilizing this strategy showed that a DNA oligomer complementary to the *Oxytricha* telomeric repeat was useful as a probe in elucidating the mechanistic details of binding between the G-rich motif and a cognate protein[11]. Hybridization of the probe with the DNA target was used to extrapolate the kinetics for unfolding of the quadruplex structure inherent in the latter, with the results being compared to the rate of formation of the protein-DNA repeat complex[11]. The authors further demonstrated that QFS binding—whether by the protein or complementary probe, was preceded by a rate-limiting quadruplex 'opening' step, a conclusion that was supported in later studies of QFS hybridization to additional complementary ligands[11].

Subsequent demonstrations used complementary DNA oligomers either as probes to indirectly establish the existence of quadruplex structures under physiologically relevant conditions[12-14], or as 'fuel' strands[15-17] to propel molecular motions in so-called nanomechanical devices. In all of these examples, yields of the resulting duplex hybrids were mitigated by quadruplex-stabilizing conditions, such as low temperatures[16], high $[K^+]$[11, 16], and elevated amounts of osmolytes/cosolvents[18].

While the biochemical and/or biological relevance of these demonstrations may differ, the lessons gleaned on quadruplex recognition by complementary probes remain consistent: Hybridization to the QFS is beset by both thermodynamic and kinetic barriers—both of which are exacerbated by quadruplex-stabilizing conditions[18], and alleviated by converse conditions that induce unfolding[16]. The thermodynamic barrier exists because the binding site for the complementary probe is sequestered within a stable intrinsic structure, requiring that the ΔG for complex formation be sufficient to compensate for the energy investment to unfold the quadruplex. Likewise, the requirement for quadruplex unfolding to reveal the binding site for the complementary probe introduces a kinetic barrier[19] that is overcome in a manner dependent on the probe's ability to hybridize the open conformation of the QFS, provided quadruplex opening is the rate-limiting step.

The aforementioned criteria therefore enable predictions of desirable properties of QFS-binding complementary ligands: The applied probe must possess high affinity for the target, so that the binding reaction is sufficiently exergonic to compensate for quadruplex unfolding. Also, hybridization is likely to be accelerated if the probe is devoid of negative charges that decelerate binding through electrostatic repulsions. Both of these criteria have been satisfied (separately) in previous demonstrations by a combination of strategies, including application of long DNA oligomers complementary to the entire QFS[16], non-systematic introduction of modified

nucleotides at random positions to improve binding affinities[20, 21], and backbone modifications that obviate negative charges on the complementary probe[19].

We previously showed that a short (7mer) PNA probe complementary to crucial residues of a stable DNA QFS could invade the quadruplex target, resulting in a PNA-DNA hybrid duplex[22]. Our results suggested that the binding energy of the PNA to its complementary site, and the enthalpic contribution from stacking of overhanging, unhybridized nucleotides—both conflated into the ΔG of complex formation—were sufficient to compensate for the energetic cost of unfolding the DNA target[22]. This recognition strategy was also extended to an RNA quadruplex target, wherein we showed that another short PNA probe was able to overcome the target structure, yielding a relatively tight (K_D = 1 nM) hybrid[23, 24].

Although PNA oligomers expanded the repertoire of QFS-recognizing complementary ligands, important limitations have impeded their broader applications. For example, quadruplex invasion by the short PNA probe described above was most facile below 10 mM K^+, whereas elevated K^+ concentrations—conditions likely to exist *in vivo*—precluded hybridization, due to the increased thermodynamic stability of the quadruplex target[22]. We also showed that hybridization of a short PNA oligomer directed to a surface-immobilized QFS was impeded by the kinetic barrier posed by the stable quadruplex[25]. Further, the combined relevance of these two limitations in the context of a biochemical assay was provided by Murphy *et al*, wherein experiments were performed at 10 mM K^+ to demonstrate the impact of PNA-mediated quadruplex invasion on primer extension by a processing DNA polymerase[26].

γPNA[27, 28] oligomers present attractive alternative reagents that might overcome the limitations of PNA probes in the context of quadruplex invasion at high salt concentrations (or *in vivo*), where the quadruplex target is likely to be thermodynamically and kinetically stable. By featuring a

stereogenic center at the γ position of the PNA molecule, these modified probes are pre-organized into helical conformers, with the helical outcome dependent on the stereochemistry at the γ carbon[27, 28]. γPNA oligomers also possess superior affinity for complementary DNA/RNA targets, and—where diethylene glycol units are installed at the γ position—feature improved aqueous solubility, relative to their PNA homologs[28].

This chapter presents evidence of quadruplex invasion and translation inhibition by complementary γPNA oligomers designed to hybridize distinct regions of an RNA QFS. The data presented here show that all γPNA probes studied readily overcome both the thermodynamic and kinetic barriers to hybridization (at 100 mM K^+), resulting in γPNA-RNA hybrids that are more stable than the intrinsic quadruplex fold. Although these probes all display repressive effects on translation, there is a strong, kinetically influenced positional bias in the potency of inhibition—with the probe directed to the 5'-end of the QFS displaying a 5 – 6 fold higher potency than the other γPNAs. Finally, evidence is also presented here to support the functional superiority of γPNAs over other recently-utilized complementary oligomers for translation inhibition mediated by RNA quadruplex invasion in this context.

2.2 MATERIALS AND METHODS

2.2.1 γPNA/RNA/DNA Oligomers

The γPNA/PNA oligomers (Chart 2.1) used for all experiments were purchased from PNA Innovations Inc. (www.pnainnovations.com) and included C-terminal L-lysine residues. The γPNA oligomer used as the scrambled control for γ5' in the *in vitro* translation assay is presented below. RNA/2'-OMe RNA and DNA oligomers used were obtained from Integrated DNA Technologies (www.idtdna.com). Sequences of all biotinylated DNA oligonucleotides used as capture strands in SPR competition experiments and the RNA oligomer designed for SPR direct-binding experiments are also given below. (RNA/DNA sequences written 5'-3' and γPNA sequence written C-N. Underlined Gs are predicted to participate in G-tetrads.)

	Sequence
4G3 RNA (for SPR)	AGACCCAAGCACUAUAAGCUAGCGGGAGGGCGGGUCUGGGCGAUCC
γ5' (scrambled overhang)	TCGAACCCTCCC (γPNA)
γ5' capture strand	biotin-ATTACTAGCGGGAGGGATTA (DNA)
γ3' capture strand	biotin-ATTAGTCTGGGCGATCATTA (DNA)
γCen capture strand	biotin-ATTAGAGGGCGGGTCTATTA (DNA)
PNA5'	GATCGCCCTCCC
OMe5'	GAUCGCCCUCCC

2.2.2 UV Melting Experiments

Thermal melting experiments were performed on a Varian Cary 300 spectrophotometer equipped with a thermoelectrically-controlled multicell holder. Samples were prepared in a buffer containing 10 mM Tris-HCl (pH 7.4), 0.1 mM Na/Li_2EDTA, and different concentrations of KCl/LiCl. The solutions were incubated at 95 °C for 5 min and cooled to 15 °C at a rate of 1 °C /min. Subsequently, the annealed samples were incubated at 15 °C for 5 min, and a heating ramp

was applied at 1 °C/min up to 95 °C. Melting curves were generated by monitoring absorbance values at 275 nm (for hetero-duplexes) and 295 nm (for quadruplexes) every 0.5 °C. Where possible, the melting temperature (T_m) was determined from a first-derivative plot of the respective melting curve. Each reported melting temperature is the average of three independent experiments.

2.2.3 Circular Dichroism (CD) Spectropolarimetry

CD spectra were obtained on a Jasco J-715 circular dichroism spectropolarimeter equipped with a water-circulating temperature controller. Samples were prepared in a buffer containing 10 mM Tris-HCl (pH 7.4), 0.1 mM Na/Li_2EDTA, and different concentrations of KCl/LiCl. The samples were incubated at 95 °C for 5 min and cooled slowly to room temperature. All spectra were collected at 37 °C. Each spectrum was collected after 6 different scans (200 – 360 nm) at a scan rate of 100 nm/min and baseline corrected. The CD spectra for 3 different samples were collected and averaged.

2.2.4 Surface Plasmon Resonance (SPR) Experiments

All SPR experiments were performed on a Biacore T100 instrument equipped with a four-channel CM5 sensor chip (GE Health Care). The sensor chip was coated with a carboxylmethyl dextran matrix that allows further functionalization with streptavidin via a standard NHS-EDC coupling procedure. Covalent immobilization of streptavidin was continued until 7000 response units (RU) of the protein were captured on each of the four channels (flow cells). The final step of the sensor design involved non-covalent immobilization of a biotinylated capture strand on the streptavidin-modified surface.

The sensor surface was primed by a buffer injection [100 mM KCl/LiCl, 10 mM Tris-HCl (pH 7.4), 3 mM Na/Li_2EDTA, and 0.005 % v/v P-20 surfactant] prior to all experiments. The capture

strand utilized for direct-binding studies with 4G3 RNA was a biotinylated 15mer γPNA oligomer that was designed to hybridize to a complementary sequence appended to the 5'-end of the RNA QFS. The active flow cell was modified with a high density (1000 RU) of the biotinylated γPNA capture strand. Immobilization of the target RNA was then performed by injecting a 100 nM solution (pre-annealed for 2 h) over the active flow cell for 400 s (flow rate = 50 µL/min) until a medium density (300 RU) of the RNA accumulated on the surface. A short dissociation phase (20 s) was introduced to remove any unhybridized or loosely-bound RNA from the surface.

To study γPNA hybridization to the immobilized RNA, 50 nM of each probe was injected for 400 s (flow rate = 30 µL/min) over both the active and reference flow cells, after which the buffer injection was restarted and continued for 600 s to monitor dissociation of the hybrids. The sensor surface was regenerated with a solution of 10 mM NaOH and 1 M NaCl (2x for 30 s), and another buffer injection (120 s) was introduced to wash off residual regeneration solution. The sensor surface was replenished with fresh RNA, and subsequent binding studies were continued by simple iterations of the aforementioned steps. The initial association rates were calculated based on the times required to reach 0.3 normalized response units of bound γPNA, similar to a procedure utilized by us in an earlier report[29].

For the competition assay[30], 1000 RU of a DNA capture strand was immobilized on the surface of the chip. The flow cell containing the capture strand was subsequently calibrated for the free γPNA concentration by injecting a series of solutions containing 5 – 50 nM of the free probe for 200 s. The slopes for the sensorgrams were obtained over a 10 s window (beginning 30 s post-injection) and plotted against the concentration of free γPNA to obtain a linear regression curve.

A fixed concentration of each γPNA probe (10 or 20 nM) was then incubated with increasing concentrations of the pre-annealed RNA target at 25 °C for 2 hours prior to injection over the

sensor surface. The slopes for the resulting sensorgrams were then used to obtain the amounts of free γPNA in equilibrium with the hybrid using the previously-generated regression curve (equation 1). Further, the concentration of bound γPNA was calculated by assuming that the system obeys the mass conservation condition (equation 2).

$$[\gamma PNA]_{free} = \frac{Slope_{40-50\,s}}{M} \quad \boldsymbol{Eqn\ 1}$$

$$M = slope\ of\ calibration\ curve$$

$$[\gamma PNA]_{Total} = [\gamma PNA]_{free} + [\gamma PNA]_{bound} \quad \boldsymbol{Eqn\ 2}$$

The bound-probe concentration, in turn, corresponds to the amount of hybrid (duplex) formed for a system involving two interacting components (equation 3). Therefore, the accumulation of the duplex after each addition of the RNA target could be determined and fit to a Langmuir isotherm that accounts for ligand depletion (equation 4), where $[P]_T$ = total γPNA concentration; $[R]_T$ = total RNA concentration; K_D = equilibrium dissociation constant.

$$[\gamma PNA]_{bound} \equiv [Duplex] = [\gamma PNA]_{Total} - [\gamma PNA]_{free} \quad \boldsymbol{Eqn\ 3}$$

$$[Duplex] = \frac{b - \sqrt{b^2 - 4 \bullet [P]_T \bullet [R]_T}}{2 \bullet [P]_T} \quad \boldsymbol{Eqn\ 4}$$

$$b = [R]_T + [P]_T + K_D$$

2.2.5 Template Production and *In Vitro* Transcription

The DNA template containing the target sequence upstream of the firefly luciferase gene (**4G3** reporter) was amplified by Polymerase Chain Reaction (PCR) using Taq DNA polymerase (New England Biolabs) and two sequence specific primers: forward primer (5'-AATACGCAAACCGCCTCTC-3') and reverse primer (5'-GGTGATGTCGGCGATATAGG-3'). A similar procedure was also used to obtain a control reporter without the target sequence

upstream of the luciferase coding sequence (**No G3** reporter). The resulting DNA fragments were purified using a GeneJET PCR purification kit (ThermoFisher) according to the manufacturer's instructions. The integrity and length of each PCR product were verified using a 1 % agarose gel and electrophoresis.

The purified DNA templates were transcribed *in vitro* using a cocktail consisting of T7 RNA polymerase (100 units) and a mixture of the ribonucleotide triphosphates (500 µM of each NTP) in a total reaction volume of 100 µL at 37 ^{o}C for 2 h. The resulting mRNA transcripts were purified using the GeneJET RNA clean-up Micro Kit (Thermo Scientific) according to the manufacturer's instructions. The final mRNA concentrations were estimated by UV spectroscopy. The extinction coefficient of each mRNA transcript was assumed to be the sum of the molar absorptivities of each nucleobase, and the recorded absorbance (at 260 nm) was assumed to be unaffected by any (undetermined) secondary folding of the transcripts. The nucleotide sequences for the UTRs in both the experimental (4G3) and control (No G3) transcripts are presented below, with the quadruplex-forming fragment, where present, highlighted in bold and upper-case.

4G3: gagacccaagcuuucagauccgcuagc**GGGAGGGCGGGUCUGGG**cgauccagccaccaug

No G3: gaucuaauaucuacuuaagaacacaaaacucgagaaccaug

2.2.6 Cell-Free Translation

A translation reaction was typically performed by incubating the purified transcripts in a mixture containing nuclease-treated rabbit reticulocyte lysate (70 % v/v) (Promega), 10 µM amino acid mixtures minus leucine, 10 µM amino acid mixtures minus methionine and 20 units RNasin ribonuclease inhibitor (50 µL total reaction volume) at 30 ^{o}C for 1.5 h. Experiments to examine the effect of each γPNA on luciferase production were performed by pre-incubating the mRNA

with requisite amino acids and increasing concentrations of the γPNA in a buffer containing 79 mM KCl and 7.9 mM Tris-HCl (pH = 7.4) to a total volume of 19.5 µL at 37 ^{o}C for 1 h. Each subsequent translation reaction was started by adding 30.5 µL of the rabbit reticulocyte lysate to the pre-incubated mixture, and the final mixture was incubated at 30 ^{o}C for 1.5 h. Finally, we performed a luciferase assay by incubating 10 µL of the translation products with 50 µL of a reagent cocktail (D-luciferin, Mg^{2+}, and ATP) (Promega) to estimate the relative amounts of enzyme produced from the mRNA reporter. Luciferase activity was estimated as relative light units (RLU) on a Tecan Infinite M1000 Plate Spectrometer.

2.3 RESULTS

2.3.1 Target Selection and Probe Design

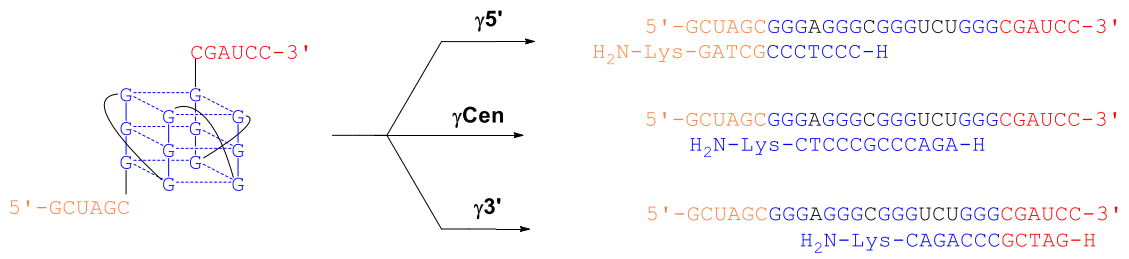
We selected as our target, a guanine (G)-rich RNA sequence derived from the 5'-untranslated region (UTR) of the human *NRAS* mRNA. The original sequence was previously shown to form a stable, three-tetrad, intramolecular G-quadruplex that represses translation in the context of a luciferase reporter model[31]—an effect that is reinforced by pharmacological[32] and/or sequence perturbations[33] that stabilize the structure, and attenuated by converse destabilizing mutations[33]. Our target sequence (called **4G3 RNA**, Chart 2.1) bears few deviations from the original sequence[31] reported by Balasubramanian. (1) We deleted one guanine residue from the second G-tract of the original NRAS QFS to simplify characterization of hybrids formed with homologous PNA/γPNA probes (reported in Chapter 4). (2) Additional mixed-sequence bases were also introduced on both termini of the target to examine any contributions of overhang recognition by complementary probes in driving quadruplex invasion. Importantly, as demonstrated below, both changes did not perturb the biophysical and biochemical characteristics of the quadruplex target.

Chart 2.1: Sequences of γPNA and RNA oligomers used in this study. (RNA sequences written 5'-3' and γPNA sequences written C-N. Underlined Gs are predicted to participate in G tetrads.)

Oligomer	Sequence
4G3 RNA	GCUAGCGGGAGGGCGGGUCUGGGCGAUCC
γ5'	GATCGCCCTCCC
γCen	CTCCCGCCCAGA
γ3'	CAGACCCGCTAG
5' rComp	CUAGCGGGAGGG
3' rComp	GAGGGCGGGUCU
Cen rComp	GUCUGGGCGAUC

Three complementary γPNA oligomers (Chart 2.1) were designed and tested for their abilities to invade the **4G3 RNA** quadruplex target and perturb its biochemical function. **γ5'** was designed to recognize the first two G-tracts of the QFS, in addition to the adjacent 5' overhanging nucleotides. The same recognition pattern was expected by **γ3'** on the opposite end of the QFS, while the binding site for **γCen** was embedded completely within the QFS (Scheme 2.1). Two additional probes—representing molecule classes recently reported[26, 34] to invade stable G-quadruplexes, with consequent biological/biochemical effects—were also tested in our system: **PNA5'** and **OMe5'**—PNA and 2'-OMe homologues of **γ5'**, respectively—were both tested for their abilities to invade our quadruplex target.

Scheme 2.1



Invasion of folded 4G3 RNA quadruplex by complementary γPNA oligomers.

2.3.2 Characterization of 4G3 RNA

The RNA target was predicted to fold into a three-tetrad intramolecular quadruplex, with loops of 1-, 1-, and 3- nucleotides, respectively, as previously reported for a very similar G-rich sequence. Circular dichroism (CD) experiments on samples containing 2 μM of the target in 1 mM KCl elicited a minimum and a maximum at 245 and 265 nm, respectively (Figure 1A, closed circles). Repeating the experiments in a 1 mM LiCl buffer produced the same peaks, but with ~ 30 % weaker intensities (Figure 2.1A, open circles). The peak positions observed herein are consistent with a parallel fold within the quadruplex structure[35, 36], consistent with the known monomorphic character of RNA G-quadruplexes[37]. Additionally, the attenuation in signal intensities upon cation substitution is well known for G-quadruplex structures[35] and has been attributed both to the higher desolvation energy[38] and small size of Li^+ relative to K^+, both of which converge to decrease the ΔG of quadruplex folding.

Thermal denaturation analyses on the same samples produced a hypochromic transition in 1 mM KCl that was eliminated upon cation substitution to Li^+ (Figure 2.1B). This result in consistent with quadruplex formation in the former buffer, and destabilization in the latter[36]. Interestingly, no residual quadruplex folding was observed by UV-melting in the LiCl buffer (Figure 2.1B, open circles), whereas weak quadruplex signatures were observed by CD experiments, under the same conditions (Figure 2.1A, open circles). While it is plausible that these divergent results represent the differential sensitives of these techniques—as the proportion of folded structures might be below the threshold detectable by UV-melting, it is also possible that the relatively weak CD peaks in 1 mM LiCl might be due to a non-quadruplex, yet helical fold present in this structure under these conditions.

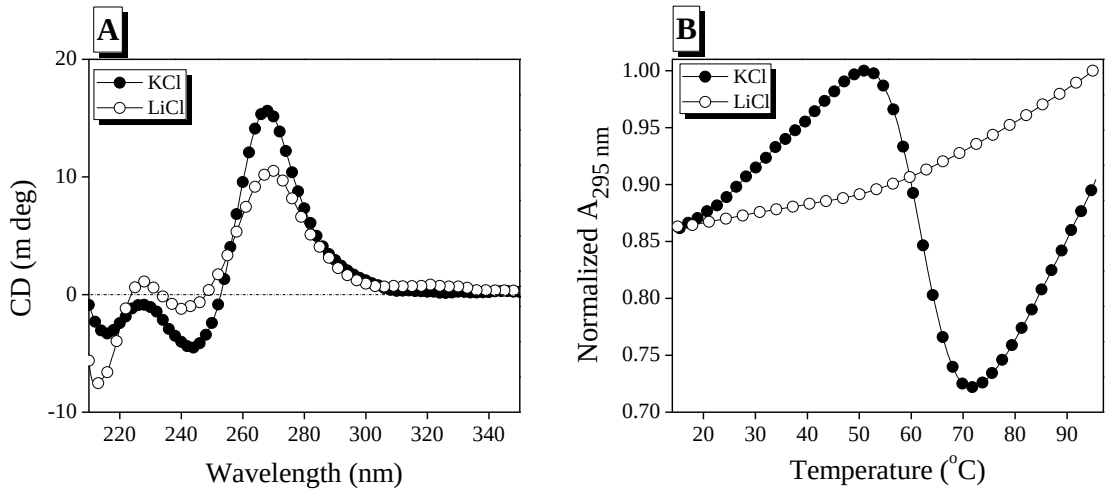


Figure 2.1: Biophysical characterization of 4G3 RNA quadruplex. CD spectra (**A**) and UV-melting curves (**B**) of 4G3 RNA target. All samples contained 2 µM RNA buffered in 1 mM KCl (closed circles) or 1 mM LiCl (open circles).

Importantly, the quadruplex structure is very stable in 1 mM KCl (T_m = 62 ± 1 °C)—a property that is further amplified in 100 mM KCl (Figure 2.2). Additionally, quadruplex folding is unimolecular, as a 5-fold increase in strand concentration (to 10 µM) did not perturb the thermal stability of the structure (Figure A1).

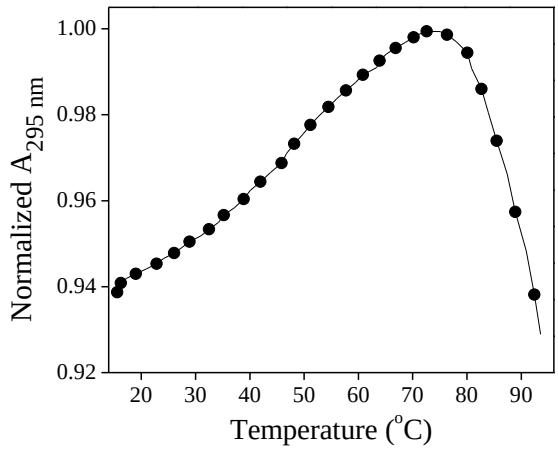


Figure 2.2: Quadruplex stability is enhanced by higher [K^+]. UV-melting curves for 2 µM 4G3 RNA in 100 mM KCl.

Biochemical characterization of the RNA target was performed by inserting the QFS into position 28 within the 5'-UTR of a firefly luciferase transcript (**4G3** reporter). A deletion control (**No G3** reporter), devoid of the QFS within the same position, was examined in parallel to ascertain any quadruplex-dependent effects on translation. (Quadruplex formation within the **4G3** reporter was assumed under the conditions of the assay.) The **4G3** reporter was translated 50 % less efficiently than the control transcript (Figure 2.3), indicating a repressive role for this quadruplex structure, consistent with many other reports describing inhibitory RNA quadruplexes[31, 33, 39-43]—including one for a very similar sequence[31]. Taken together, our data suggest that **4G3 RNA** forms a stable G-quadruplex structure with a repressive effect on protein synthesis, when placed in the 5'-UTR of a reporter transcript.

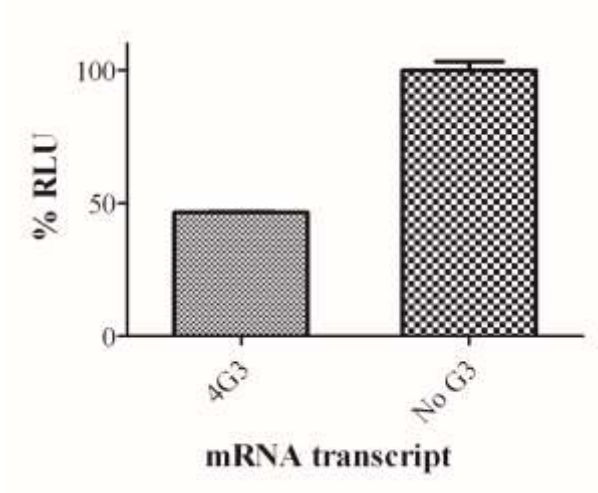


Figure 2.3: Biochemical Characterization of 4G3 RNA. Luminescence intensity (expressed as % RLU) for 4G3 reporter and No G3 control transcripts.

2.3.3 Characterization of 4G3 RNA-γPNA Hybrids

We next sought to characterize the hybrids formed between **4G3 RNA** and the complementary γPNA oligomers. Samples containing the target and each probe were prepared in 100 mM KCl buffer to simulate a realistic thermodynamic barrier to hybridization by the γPNA oligomers. (The high salt concentration was shown above to stabilize the quadruplex target.) Melting curves generated at 295 nm[44] were characterized by a hyperchromic transition (Figure 2.4A), in contrast to the hypochromic trend previously observed for the target alone. This inversion in chromicity indicates that the quadruplex fold of **4G3 RNA** is abolished in the presence of the complementary probes.

Melting curves generated at 275 nm[24]—to directly examine hybrid duplex melting, were characterized by a general hyperchromic trend (Figure 2.4B). However, the presence of multiple transitions and absence of clear inflection points made further interpretation of this data set difficult (Figure 2.4B). Additional support for the existence of hybrid duplexes was obtained from CD spectra, wherein a minimum at 295 nm—characteristic of PNA-RNA hybrid duplexes[24, 45], was observed for all γPNA-RNA complexes (Appendix, Figure A2). These spectra were not equivalent to the mathematical aggregates of those for 4G3 RNA and the γPNAs, suggesting that the CD signatures observed here are characteristic of the hybrid duplexes (Figure A3). Importantly, the minima and maxima for all hybrid duplexes were similar to those observed for the complexes formed with truncated RNA targets (rComp) possessing only the bases complementary to each probe (Appendix, Figure A4). Taken together, the composite of thermal denaturation and CD analyses suggest that the γPNA oligomers can overcome the thermodynamic barrier to hybridization posed by the 4G3 RNA target, with the binding reaction resulting in complexes that are more stable than the starting quadruplex structure.

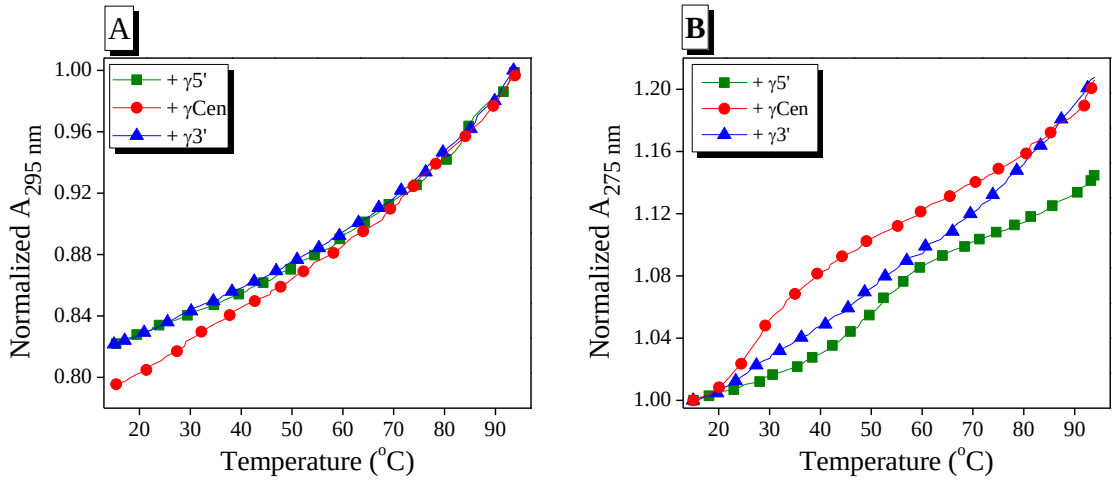


Figure 2.4: Biophysical characterization of 4G3-γPNA Hybrids. UV melting curves at 295 nm (**A**) and 275 nm (**B**). All samples contained 2 µM 4G3 RNA and 2 µM γ5' (green), γ3' (blue), or γCen (red) in 100 mM KCl.

Measurements of the dissociation constants (K_D) were next performed for the respective γPNA-4G3 hybrids. Although direct extrapolation of thermodynamic parameters for hybridization is possible from UV-melting curves[46], such analyses were precluded herein by the complex melting profiles recorded for these hybrid complexes (*vide supra*). Therefore, K_D values for the hybrids were determined by a surface plasmon resonance (SPR) competition method[30]. Originally reported by Roy *et al*[30], this method facilitates homogenous, label-free analyses of biomolecular interactions. A comprehensive description of the experimental protocol is provided in the methods section. In brief, an SPR chip—laden with a high density of a DNA capture strand, is calibrated with increasing concentrations of the relevant complementary γPNA. The sensor responses from each injection are used to construct a regression curve relating instrument signal to free γPNA concentration. A fixed amount of the γPNA is then titrated with increasing concentrations of **4G3 RNA**, and the sensor responses are used to compute the bound PNA (duplex) concentrations. The

binding isotherm, resulting from a plot of [duplex] against [**4G3 RNA**], is fit to a Langmuir model that accounts for ligand depletion at high RNA concentrations.

The sensorgrams from calibration and competition, along with the linear regression curve used for each γPNA are presented in the Appendix (Figures A5 – A7). The binding isotherms for the three hybrid duplexes are presented in Figure 5, and K_D values predicted by the Langmuir model are summarized in Table 1. We observe that all γPNAs bind with similar affinities to the **4G3 RNA** target, with K_D values in the low (2 – 4) nM range (Table 2.1). This result is not surprising, as all γPNA probes tested here are challenged by the same intrinsic quadruplex structure of the RNA target and have similar G-C contents (Chart 2.1). Additionally, the fact that γPNA oligomers directed to different regions of the quadruplex target exhibit similar K_D values suggests that the intrinsic affinities of the probes for their respective target sequences within the QFS were similar.

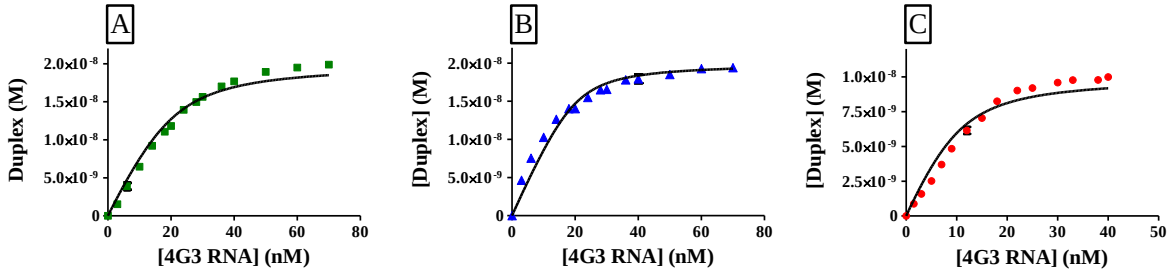


Figure 2.5: Binding Isotherms for 4G3-γPNA Hybrids. γ5' + 4G3 (**A**). γ3' + 4G3 (**B**). γCen + 4G3 (**C**). Data fit to Langmuir model presented in Methods (section 2.2.4).

Table 2.1: Equilibrium dissociation constants (K_D, nM) for γPNA-RNA hybrid duplexes formed with 4G3 RNA quadruplex and truncated complementary RNA targets.

γPNA	4G3 RNA	rComp
γ5'	4.2 ± 0.3	3.8 ± 0.2
γ3'	1.9 ± 0.2	0.0015 ± 0.0004
γCen	2.8 ± 0.3	33 ± 1.7

2.3.4 Characterization of rComp-γPNA Hybrids

To test the above hypothesis, we performed identical analyses on the truncated RNA targets bearing only the complementary sequences for the respective γPNAs oligomers (Chart 2.1). Hybrid duplex formation between each respective rComp and the corresponding γPNA was verified by UV melting and CD experiments (Figure 2.6). However the melting data collected here—as with the **4G3 RNA** samples, precluded direct K_D determination due to the incomplete transitions recorded at even the highest temperatures.

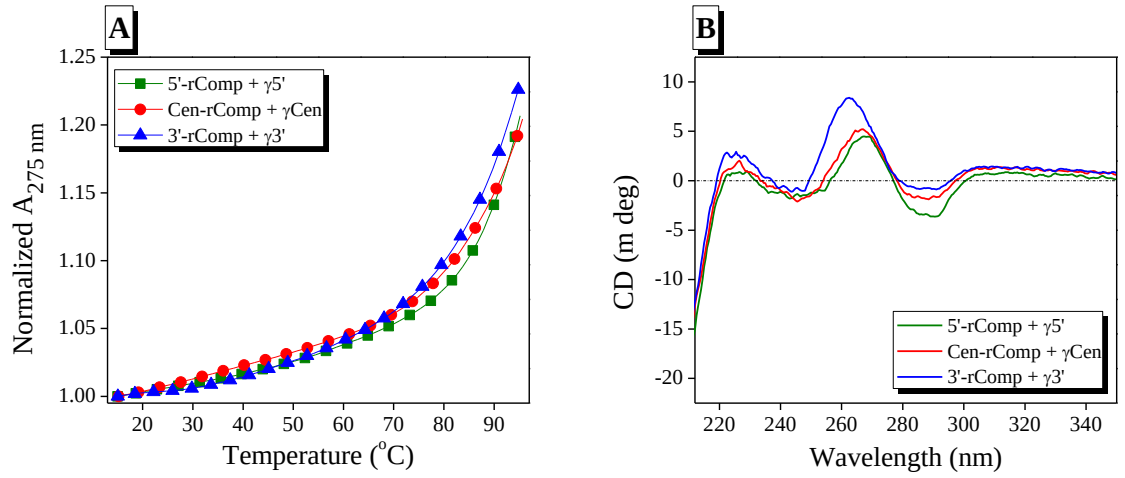


Figure 2.6: Biophysical characterization of rComp-γPNA Hybrids. UV melting curves at 275 nm (**A**) and CD Spectra (**B**). All samples contained 2 μM of the respective rComp and were buffered in 100 mM KCl.

SPR competition experiments, performed as described previously, were therefore used to determine K_D values for the resulting hybrid duplexes. Sensorgrams from the competition experiments are presented in the Appendix (Figure A8). The binding isotherms are presented in Figure 2.7, and K_D values (Langmuir model) are summarized in Table 1. Interestingly, the affinities for the rComp-γPNA hybrids differ by four orders of magnitude—with **γ3'** hybridizing to its complementary sequence (**3'-rComp**) with 22,000-fold higher affinity than **γCen**. [We currently do not have an explanation for the fact that the affinity of the probes for their rComp targets is only marginally improved (for **γ5'**) and even less (for **γCen**), relative to the **4G3** target.] This disparity in binding affinities is greater than that recorded for the **4G3 RNA** hybrids (Table 2.1). One possible explanation for this result is that the structures adopted by the rComp targets in solution are sufficiently disparate to impose significantly different levels of resistance to hybridization of the respective probes. For example, intra- or intermolecular structures formed by one rComp oligomer will depress its affinity for the complementary γPNA probe, provided the targets for the other probes are devoid of such structures.

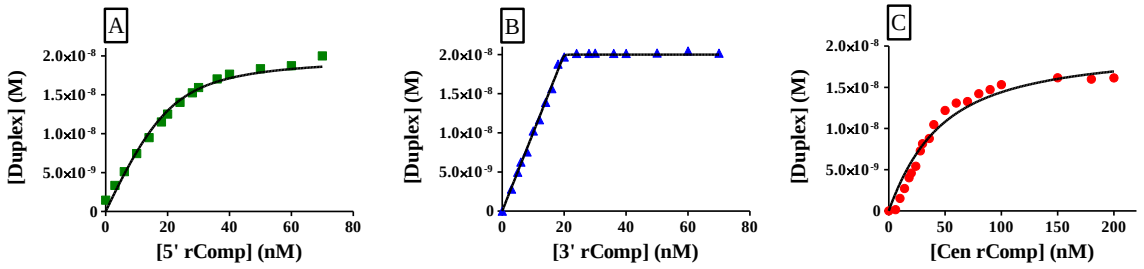


Figure 2.7: Binding Isotherms for rComp-γPNA Hybrids. γ5' + 5'-rComp (**A**). γ3' + 3'-rComp (**B**). γCen + Cen rComp (**C**). Data fit to Langmuir model presented in Methods (section 2.2.4).

To test for the presence of intra- or intermolecular structures formed by the rComp targets, we examined melting and CD profiles of each rComp oligomer in 100 mM KCl—the same conditions under which affinities were measured. **5'-rComp** and **Cen-rComp** elicit hypochromic melting

transitions characteristic of stable quadruplexes, while **3'-rComp** did not possess quadruplex melting signatures (Figure 2.8A). CD experiments performed on the same samples also showed characteristic quadruplex peaks for the former sequences, which were attenuated in the latter (Figure 2.8B). Closer examination of the respective sequences reveals that **5'-rComp** and **Cen-rComp** both possess two G_3 tracts separated by single-nucleotide loops—sequence requirements for intermolecular quadruplex formation involving two or four strands. Only one G_3 tract is present in **3'-rComp**, implying that quadruplex formation by this sequence can be accomplished only by the assembly of four strands.

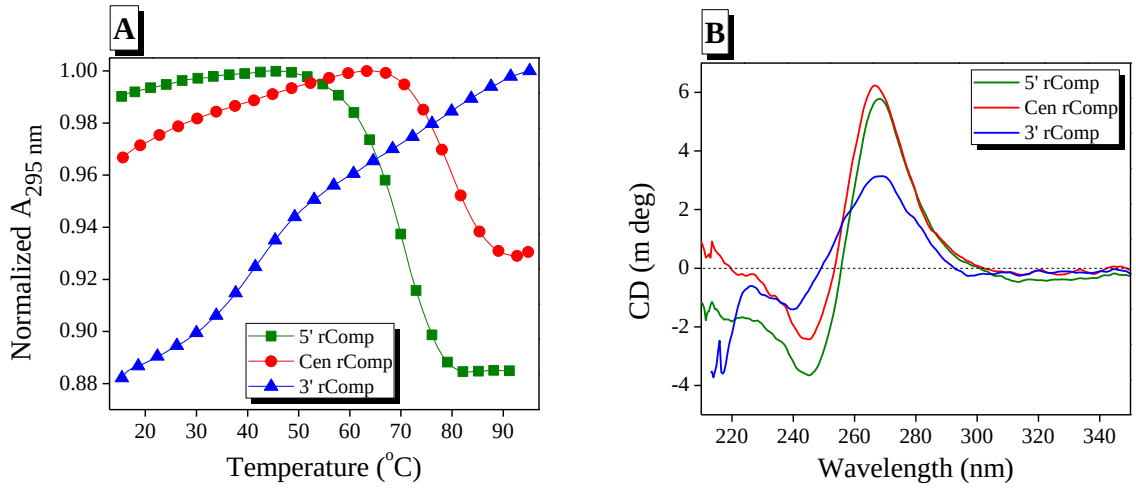


Figure 2.8: Biophysical characterization of rComp targets. UV melting curves at 295 nm (**A**) and CD Spectra (**B**) for 2 μM 5'-rComp (green), 3'-rComp (blue), and Cen-rComp (red). All samples were buffered in 100 mM KCl.

Although the quadruplex structures—where they exist—are stable at the concentrations (μM) used for the melting experiments, it is not clear whether they would be stable at the lower concentrations (nM) used for the SPR competition analyses. Additionally, K_D determination for the **γCen/Cen-rComp** hybrid in 100 mM LiCl, conditions expected to disfavor quadruplex folding—showed only a 25-fold recovery in binding affinity (33 nM vs 1.3 nM, Figure A9). This result suggests that the

depression in binding affinity for the truncated, structured target is not sufficiently explained by intermolecular folding of **Cen-rComp**. It will therefore be interesting to measure the binding affinities of **γ3'** and **γ5'** for their respective targets under identical conditions (100 mM LiCl) in order to further elucidate the extent to which intermolecular structures, or their absence, affect binding thermodynamics. Taken together, however, our analyses suggest that the γPNA oligomers can form stable hybrids with the truncated RNA targets, regardless of the any intermolecular secondary folding.

2.3.5 Effect of γPNAs on mRNA Translation

We next assessed the effect of each γPNA on translation of the luciferase reporter containing the RNA QFS (**4G3**) and the control reporter (**No G3**) lacking this element. Dose response curves are shown in Figure 2.9 for each γPNA, and IC_{50} values are summarized in Table 2.2. All γPNA oligomers tested here are potent inhibitors of translation, with IC_{50} values in the low – mid nM range (Table 2.2). We recorded minute to moderate activities of the probes against the **No G3** reporter, even at the highest probe concentrations (Figure 2.9)—demonstrating that these probes inhibit translation via a specific recognition event. Fidelity in recognition was further demonstrated by pre-incubating the probes with the cell lysates prior to introduction of a control reporter transcript to begin the translation reaction. (The control transcript in this case is supplied by Promega and codon-optimized for expression in this lysate kit.) All γPNA probes remained inert to the control reporter, even at concentrations corresponding to two times the respective IC_{50} values (Figure A10)—demonstrating that these probes have no detectable effect on ribosomal translation efficiency in this system.

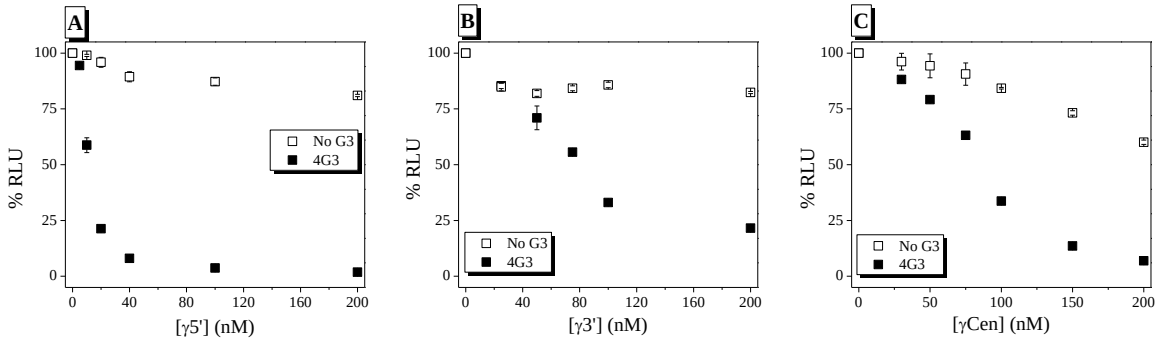


Figure 2.9: Effect of γPNAs on luciferase mRNA translation. Dose-response curves for γ5' (**A**), γ3' (**B**), and γCen (**C**) against 4G3 target (filled squares) and No G3 control (open squares) mRNAs. [mRNA] = 10 nM and γPNA-RNA samples were incubated for one hr at 37 °C prior to adding to lysate. Data are normalized to 100% for the samples that lacked any γPNA. Error bars represent standard deviation of three independent trials.

Table 2.2. IC_{50} values (nM) for γPNA antisense oligomers targeted to 4G3 RNA at two different temperatures.

γPNA	**IC_{50} (nM) 37 ºC**	**IC_{50} (nM) 60 ºC**
γ5'	15	15
γ3'	75	20
γCen	90	20

We also tested a variation of **γ5'** in which the first 5 nucleotides were scrambled to eliminate complementarity to the 5'-overhang of the quadruplex, and found that the partially scrambled γPNA did not effectively inhibit translation (Figure 2.10). This attenuation in potency upon eliminating overhang complementarity is further evidence of the sequence selectivity of these probes, and highlights overhang recognition as a salient determinant of the invasion and inhibitory effects of the complementary γPNAs in this context (discussed further below).

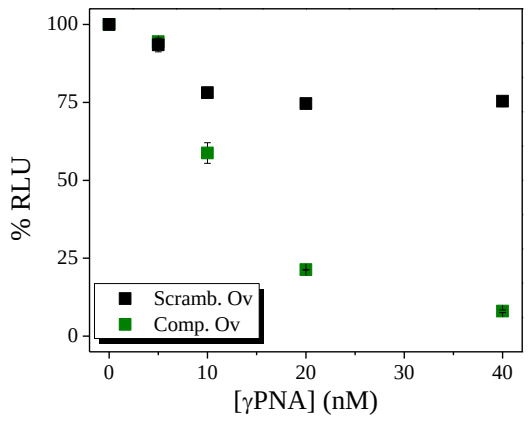


Figure 2.10: Effect of overhang recognition on translation inhibition by γ5'. Dose-response curves for γ5' bearing complementary (green) or scrambled (black) overhang recognition element.

The dose-response curves for the three γPNAs against the **4G3** reporter (summarized in the Appendix, Figure A11) were next compared to elucidate any position-dependent effects on quadruplex invasion and translation inhibition. We observed a strong positional bias in the potency of inhibition, with the probe directed to the 5'-end of the QFS inhibiting translation at 5 – 6 fold higher potency than probes directed to other positions (Figure A11 and Table 2.2). This variation in potencies exceeds the difference in K_D values recorded for the respective hybrids (Table 2.1), suggesting that binding affinities alone are not sufficient to explain the biochemical activities observed for these probes. (In any case, the most potent probe, **γ5'**, is simultaneously the weakest binder of the **4G3 RNA** target (K_D = 4.2 nM))

2.3.6 Effect of Kinetics on Translation Inhibition by γPNAs

We therefore considered the possibility that kinetic effects might be contributing to the variation in potency. In the initial cell-free translation experiments, the γPNA and mRNA were incubated at 37 °C for 60 min prior to addition to the rabbit reticulocyte lysate to begin translation. We repeated these experiments but increased the incubation temperature to 60 °C to enhance hybridization rates prior to performing translation at 37 °C. IC_{50} values determined under these conditions showed nearly identical results for the three γPNAs (Figure 2.11 and Table 2.2), all but eliminating the variation observed after incubation at lower temperature, consistent with a kinetic contribution to the discrimination observed at 37 °C.

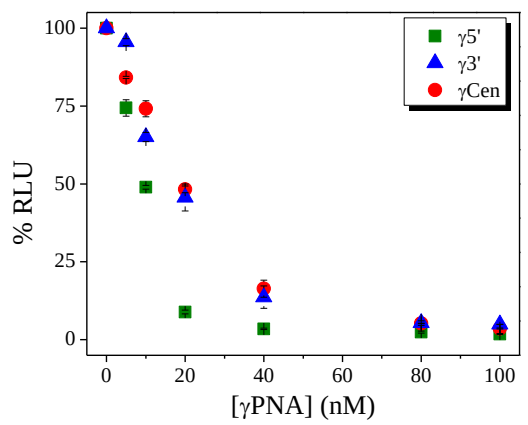


Figure 2.11: Comparing antisense inhibition by γ5' (green squares), γ3' (blue triangles), and γCen (red circles) following incubation at 60 °C. Data are normalized to 100% for the samples that lacked any γPNA.

Although kinetic effects largely explain the disparate potencies, it is unlikely that the K_D values are entirely irrelevant. The divergence between affinity and activity might reflect our inability to

accurately simulate the conditions of the translation assay during measurement of the former. Binding affinities were obtained by an SPR competition method (described above), wherein a fixed concentration of the γPNA probes was titrated with increasing concentrations of a *synthetic, truncated* RNA strand (~ 40 nt). However, the translation assay challenges the γPNA probes with an *in vitro transcribed* RNA fragment (~ 2000 nt). It is therefore possible that K_D measurements with the truncated strand fail to account for adjacent or long-range interactions that could perturb the stability of the quadruplex and/or its interactions with the complementary probes. 'Realistic' affinity measurements are limited here by the unavailability of mRNA transcript in sufficient amounts and, even where present, the need to unambiguously validate the binding stoichiometry of the mRNA-γPNA hybrids prior to curve fitting.

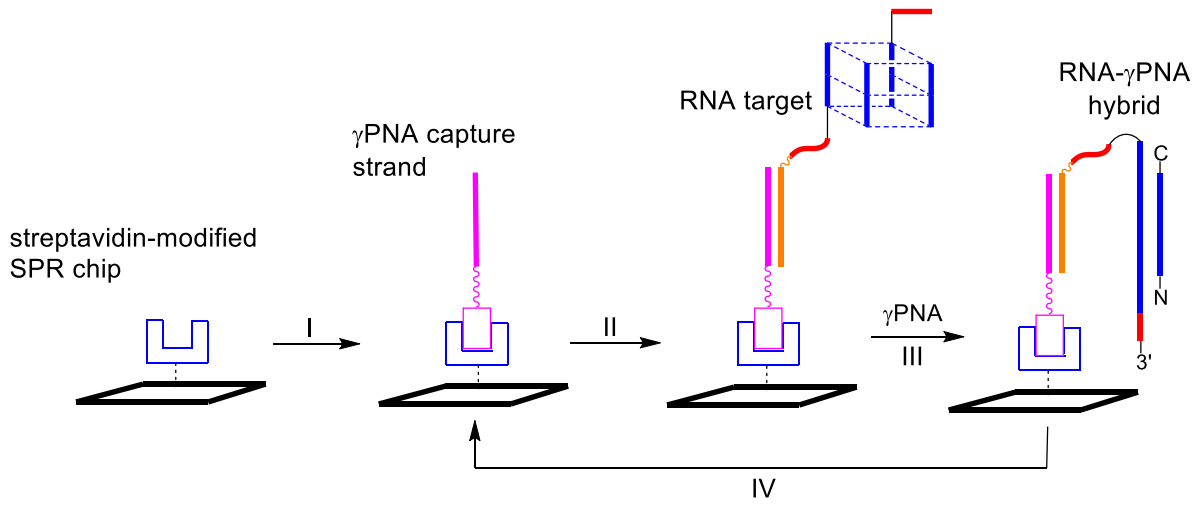
2.3.7 Monitoring Kinetics of γPNA Hybridization to 4G3 RNA

In order to further interrogate the interactions between **4G3 RNA** target and the respective γPNA oligomers, we utilized SPR direct-binding experiments in which the individual probes were flowed over the surface of a target (RNA)-laden SPR chip. Direct-binding analyses involving immobilized RNA targets are traditionally difficult to execute, especially when the reactions yield high-affinity, slowly-dissociating complexes—as is the case in our system.

To circumvent this technical limitation, we developed a method that allowed us to replenish the sensor surface with fresh RNA after each injection cycle (Scheme 2.2). Specifically, **4G3 RNA** was extended by adding nucleotides to the 5'-end which were complementary to a biotinylated γPNA capture strand that was deposited on the SPR chip. Flowing the RNA over the γPNA-functionalized chip resulted in immobilization of the RNA by Watson-Crick pairing. (A similar method was recently reported by Smolke and coworkers for DNA/RNA aptamer immobilization

on an SPR chip[47].) Hybridization to the γPNA capture strand did not affect the thermal stability of the RNA quadruplex (Figure A12). The strength of the γPNA-RNA duplex was sufficiently high to minimize dissociation of the RNA QFS during the subsequent introduction of the antisense molecules (Figure A13). After each hybridization experiment, the antisense molecule and RNA QFS were removed by washing with NaOH, followed by reconstitution of the chip by addition of fresh RNA QFS.

Scheme 2.2



I: Immobilization of biotinylated γPNA on streptavidin-modified SPR chip.
II: Hybridization of 4G3 RNA to capture strand via 5'-tag.
III: γPNA probes bind to non-covalently immobilized RNA quadruplex target.
IV: Regeneration removes resulting complex, releasing the γPNA capture strand for subsequent reactions.

As shown in Figure 2.12, we observed a slight kinetic advantage (1.3 – 1.6 fold) for hybridization to the 5'-end of **4G3 RNA** (by **γ5'**), relative to other positions of the QFS (Figure 12A, see legend). However, the accessibility of this position is counter-balanced by *faster* release of the bound probe (Figure 2.12A). While these results, by themselves, do not provide insight into the association/dissociation kinetics of the γPNA oligomers with their cognate RNA targets outside of

the quadruplex context—the faster off-rates for the two terminally-directed probes (**γ5'** and **γ3'**) are consistent with our previous report describing stabilization of PNA-DNA hybrid duplexes by overhanging, unhybridized nucleotides[22]. It is therefore possible that the centrally-directed probe (**γCen**)—although deprived of an obvious nucleation site, forms a kinetically stable complex due to stacking of overhanging bases on both ends of the duplex.

Interestingly, the trends in relative initial association and dissociation rates for the respective probes did not change when identical experiments were repeated in 100 mM LiCl (Figure 2.12B)—demonstrating that the differences observed herein are not due to quadruplex folding, and might be attributable to some intrinsic properties of the probes (*vide infra*). Importantly, the kinetic differences recorded here, while statistically significant ($p = 0.0363$), are insufficient to explain the strong positional bias observed from the *in vitro* translation experiments. As with the affinity measurements, it is possible that the information derivable from these SPR experiments is limited by the use of a truncated target, instead of the full mRNA transcript. However, our method for immobilizing RNA molecules (summarized above) should alleviate this challenge in future studies.

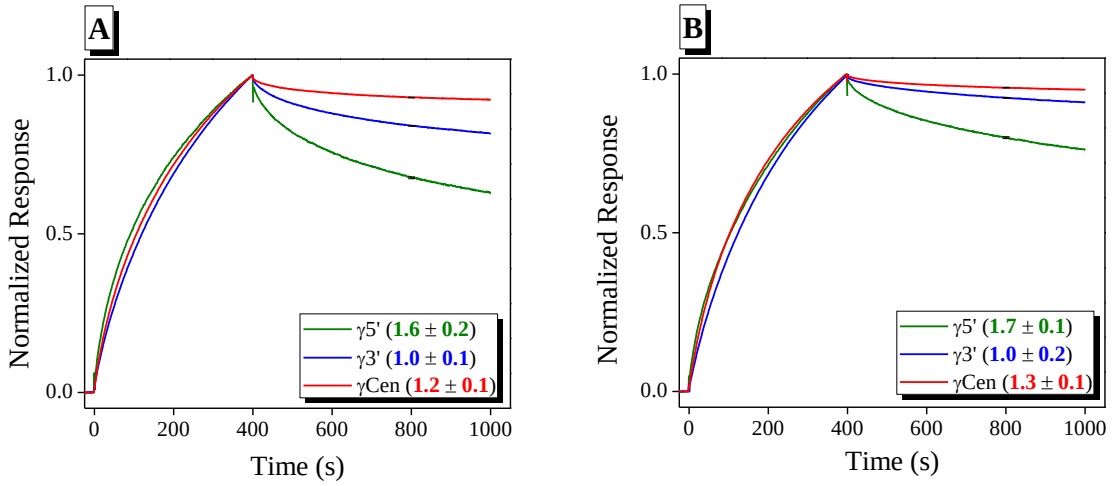


Figure 2.12: Sensorgrams for hybridization of 50 nM γ5' (green curve), γ3' (blue curve), or γCen (red curve) to immobilized 4G3 RNA. Running buffer contained 10 mM Tris-HCl (pH 7.4), 3 mM Na_2EDTA, and 100 mM KCl (A) or 100 mM LiCl (B). All binding experiments were performed at 25 ^{o}C. Error bars at t = 800 s.

2.3.8 Evidence of Intramolecular Folding in γPNA Probes

The advantage potentiated by overhang recognition on the 5'-end of the QFS (for **γ5'**) was expected to be available on the 3'-end (to **γ3'**), to the extent the 3'-overhanging nucleotides are themselves not sequestered by intra/intermolecular folds, or occluded by adjacent secondary structures downstream of the quadruplex. We therefore sought to examine the three γPNA probes for evidence of existing structures that might preclude access to their respective binding sites, and thereby manifest as weaker invasion or potency of some probes relative to others. As shown in Figure 2.13A, UV-melting analyses of the γPNA oligomers, in the absence of the RNA targets, revealed hyperchromic melting transitions, with the highest T_m (51 °C) recorded for **γ3'** (Figure 2.13A). (For comparison, **γ5'** and **γCen** elicit melting transitions at 35 °C and 31 °C, respectively.)

One possible explanation for this result is that these probes all adopt structures in solution—the most stable being present in **γ3'**—which have to be disrupted for the probes to access their complementary sites on the target. However, extensive interpretation of this result is precluded by the fact that helical preorganization of the probes enforces a stacking arrangement of the nucleobases—the unfolding of which would also produce the hyperchromic transitions recorded here. The latter possibility is supported by the fact that the melting transitions are concentration-independent, as a five-fold increase in strand concentration did not perturb the ostensible T_m values (Figure 2.13B). Regardless of the physical interpretation for these results, the fact that **γCen** inhibits translation with six-fold lower potency than **γ5'**, while forming a structure comparable—in thermal stability, at least—to the latter, suggests that intramolecular secondary folding (or nucleobase stacking) is not a crucial determinant of the inhibitory effects induced by these probes.

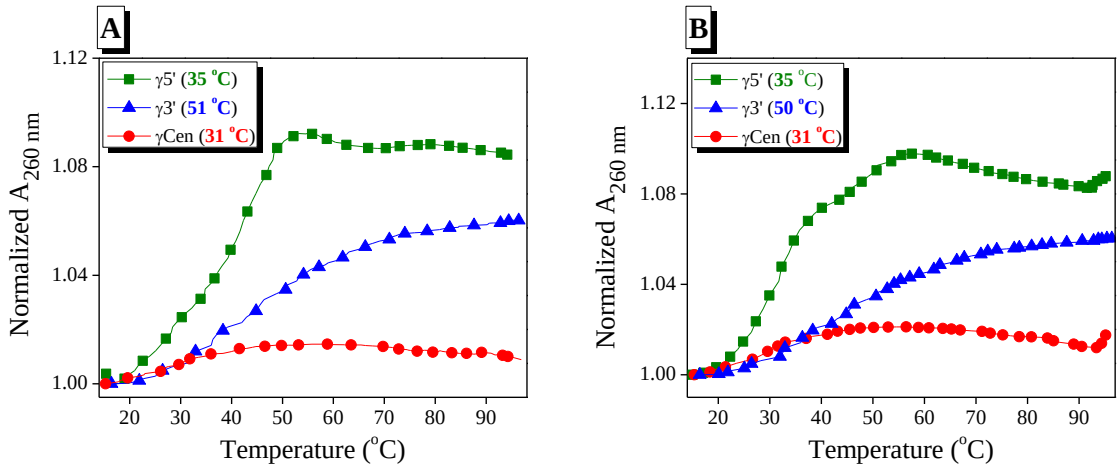


Figure 2.13: Characterization of γPNA oligomers. UV-melting curves at 260 nm for γ5' (green curve), γ3' (blue curve), or γCen (red curve) at 2 μM (**A**) or 10 μM (**B**) strand concentrations. All samples contained 100 mM KCl. T_m values are presented in parenthesis in the legend.

2.3.9 Comparing γPNA with PNA and 2'-OMe Oligomers: Translation Inhibition

The results presented above show that γPNA oligomers can invade a stable RNA G-quadruplex and exert potent inhibition of translation at low nM concentrations. We next assessed the antisense effects of PNA and 2'-OMe oligomers directed to the same position (5'-end) of the quadruplex as the most potent γPNA in the context of the luciferase assay. These probes were selected based on recent reports demonstrating their abilities to invade stable G-quadruplexes and modulate polymerase extension (for PNA[26]) and translation (for 2'-OMe[34]). Ancillary melting experiments also demonstrated that the hypochromic transition of **4G3 RNA** was abolished in the presence of either of the aforementioned probes, demonstrating their capacity to overcome the thermodynamic barrier to hybridization posed by the quadruplex structure (Figure A14).

Figure 2.14 presents dose-response curves for both oligomers against the target and control reporters. (IC_{50} values are summarized in Table 2.3.) The PNA probe (**PNA5'**) is as potent as the γPNA probe, but also retains significant activity against the control transcript (Figure 2.14A). The 2'-OMe oligomer (**OMe5'**) is ~ 15-fold less potent than **γ5'** (250 nM vs 15 nM, respectively), although this probe, unlike the PNA oligomer, displayed only modest non-specific effects against the control transcript (Figure 2.14B). While it is likely that the reduced potency of **OMe5'** is attributable, at least in part, to the lower affinity of 2'-OMe oligomers compared with γPNA, we also considered the possibility that kinetic effects could again contribute to the differences.

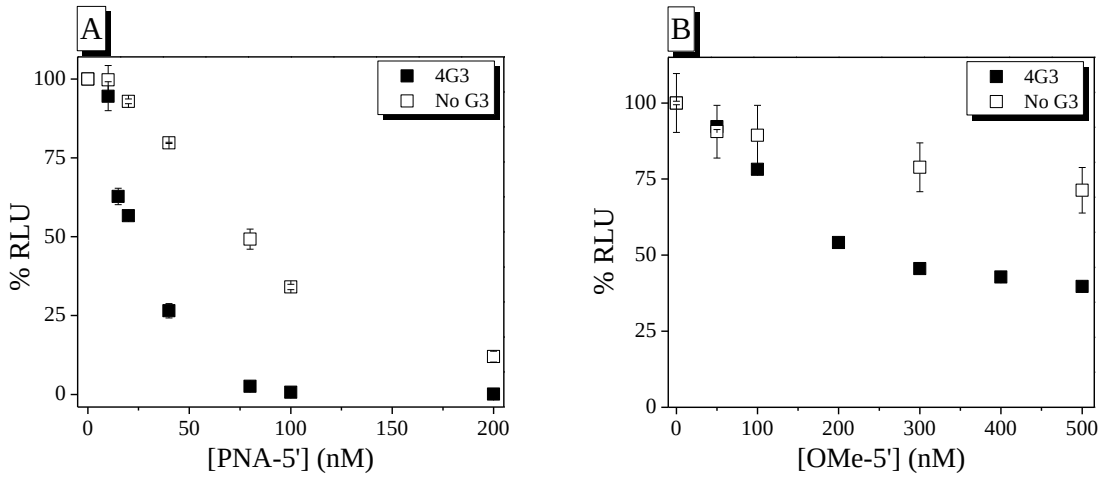


Figure 2.14: Effects of alternative antisense molecules on mRNA translation. Dose response curves for PNA5' (**A**) or OMe5' (**B**) oligomers against 4G3 target (filled squares) or No G3 control (open squares) reporters. [mRNA] = 10 nM, and RNA-PNA/2'OMe samples were incubated for one hr at 37 °C prior to adding to lysate. Data are normalized to 100% for the samples that lacked any antisense probe.

Table 2.3. IC_{50} values (nM) for different antisense oligomers targeted to 4G3 RNA.

Antisense	IC_{50} (nM)
γ5'	15
PNA5'	20
OMe5'	250

To test this hypothesis, luciferase experiments were performed by incubating the **4G3 RNA** transcript with a constant concentration of **OMe5'** (250 nM), varying incubation time for the probe with the target RNA prior to the start of translation. As shown in Figure 2.15, we observed substantial time-dependent improvements in translation inhibition by **OMe5'**, indicative of slower hybridization kinetics for this probe relative to the homologous γPNA. Control experiments performed by pre-incubating the mRNA transcript alone for 5 h showed reduced translation (60 %) from the reporter, possibly due to mRNA degradation (Figure A15). Importantly, however, this attenuation in luciferase production was less than the inhibition observed in the presence of the

probe (Figure A15), suggesting that the enhanced antisense effects reported in Figure 2.15 are due—at least in part, to the specific activity of the **OMe5'** probe.

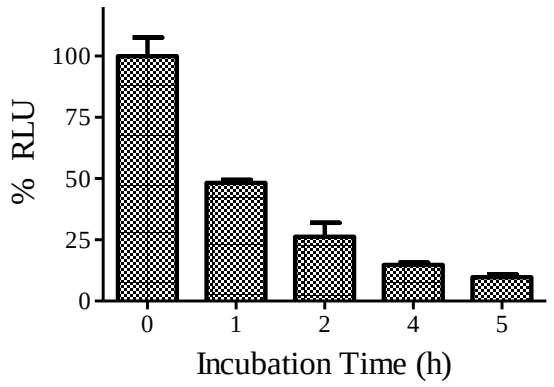


Figure 2.15: Effect of incubation time on translation inhibition by OMe5'. [RNA] = 10 nM and [OMe5'] = 250 nM. All reactions were incubated at 37 °C for increasing durations prior to addition to lysate.

2.3.10 Comparing γPNA with PNA and 2'-OMe Oligomers: Binding Kinetics

SPR direct-binding experiments were next employed to assess the contributions of any differential binding kinetics for the different probe types to the disparate potencies recorded in the *in vitro* translation assays. Again, these analyses were enabled by the noncovalent immobilization of the RNA quadruplex on the SPR chip using the method described in Scheme 2.2.

The three probes exhibit drastically different binding profiles to the immobilized target (Figure 2.16). Quadruplex invasion by **PNA5'** is more efficient than **γ5'**, with as much as a seven fold difference between the initial association rates for the two probes recorded when experiments were performed in 100 mM KCl (Figure 2.16A). These results are surprising, since a previous report directly comparing hybridization by both PNA and γPNA to an immobilized complementary DNA

showed that the association rates were not significantly perturbed by the introduction of γ modifications[28]. **PNA5'** was also released from the target at a much faster rate than **γ5'**, consistent with a previous report showing that γ-modifications suppress dissociation rates of γPNA relative to the unmodified PNA[28]. The difference between the initial association rates for the two probes decreases to 3-fold when the quadruplex is destabilized in 100 mM LiCl, suggesting that the differential kinetics are not due entirely to the quadruplex structure (Figure 2.16B), and might be attributable to some inherent structure in the probes.

Hybridization by **OMe5'** was more than 40-fold slower than **γ5'** (Figure 2.16A), a trend that was slightly improved when the experiment was repeated in 100 mM LiCl (Figure 2.16B). The fact that quadruplex destabilization did not result in significant enhancement in hybridization suggests that the binding profile observed herein is due to the intrinsic properties of the probe. 2'-OMe oligomers retain the negatively charged, phosphodiester backbone of RNA/DNA. It is therefore possible that the electrostatic repulsion between this probe and the RNA target impedes facile hybridization, even when the quadruplex is destabilized. Importantly, the slow association rate for this probe helps explain its higher IC_{50} recorded in the luciferase assay (Figure 2.14), and the progressive improvements in translation inhibition following extended incubation times reported in Figure 2.16. Taken together, these results establish the functional superiority of γPNA probes for translation inhibition mediated by quadruplex invasion *in vitro*, over other molecules (PNA and 2'-OMe) recently reported to invade stable G-quadruplexes and modulate their functions.

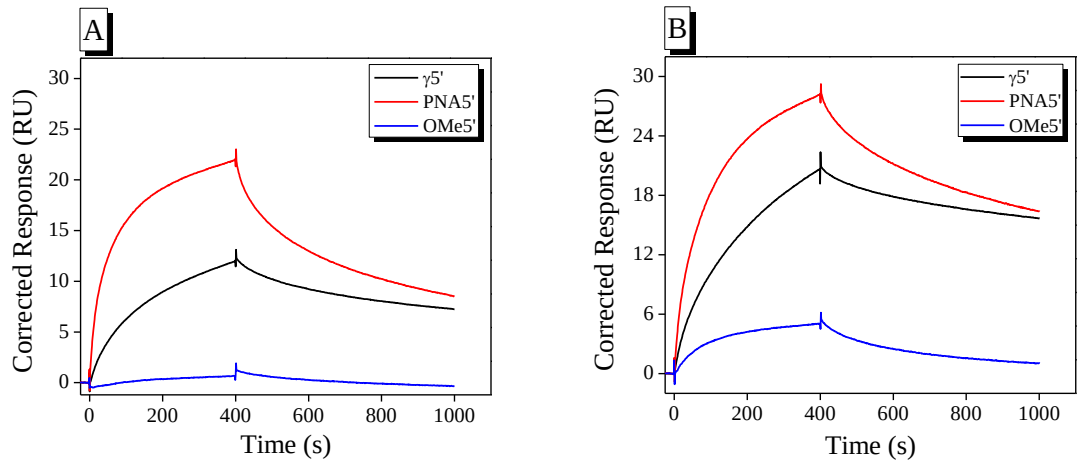


Figure 2.16: Sensorgrams for hybridization of 50 nM γ5' (black curve), PNA5' (red curve), or OMe5' (blue curve) to immobilized 4G3 RNA. Running buffer contained 10 mM Tris-HCl (pH 7.4), 3 mM Na_2EDTA, and 100 mM KCl (**A**) or 100 mM LiCl (**B**). All binding experiments were performed at 25 ºC. The same amount of RNA was immobilized prior to each injection of the probe.

2.4 DISCUSSION

The results presented above show that antisense γPNA 12mers directed to a quadruplex-forming sequence within the 5'-UTR of a firefly luciferase reporter mRNA inhibit translation with IC_{50} values in the low-mid nanomolar range. Importantly, the probes minimally inhibit a control transcript lacking the QFS and have no significant effect on overall ribosomal translation efficiency, demonstrating useful selectivity for future intracellular applications. We also observed an interesting variation in the antisense potencies depending on the site of recognition within the QFS, with hybridization to the 5'- and 3'-ends of the quadruplex and flanking regions yielding more potent inhibition than an internal site after pre-incubation at 37 ºC. Interestingly, this discrepancy in the antisense potencies was substantially diminished after pre-incubation at 60 ºC, demonstrating that kinetic effects exert significant control over the efficiency of translation

inhibition by these probes. We illustrate below, some of the factors that contribute to these potent, temperature-responsive, and position-dependent antisense effects observed for the γPNA probes.

The inherently high affinity[28] of γPNA for RNA drives quadruplex invasion and stable duplex formation at low γPNA and RNA concentrations. The intrinsic affinity of the probes is especially important in this context since the stable structure of **4G3 RNA** imposes a thermodynamic penalty to hybridization that has to be compensated by the free energy of complex formation. This energetic barrier is likely to be exacerbated in a cellular context, where high salt concentrations could further stabilize the quadruplex target.

Our results also demonstrate that access to the QFS matters at least as much as affinity in determining the functional outcome of the binding reaction. The significantly more potent antisense effect observed for **γ5'** than **γ3'** or **γCen** following incubation at 37 ºC occurs presumably because the flanking nucleotides on the 5'-end of the QFS facilitate initial nucleation of the probe in the vicinity of the quadruplex target. The resulting increased local concentration for **γ5'** might drive hybrid formation to a greater extent than for **γ3'** or **γCen**. We observe that the positional bias of antisense inhibition is significantly minimized following incubation at 60 ºC. It is likely that the elevated temperature destabilizes the quadruplex structure as well as nearby secondary structure, facilitating γPNA hybridization, suggesting that the rate-determining step is unfolding of the quadruplex secondary structure, consistent with a report from Green *et al*[19] for hybridization of a complementary PNA to a DNA quadruplex derived from the human telomere sequence.

γPNA also exhibited important advantages over other antisense molecules (PNA and 2'-OMe) recently reported[26, 34] to invade stable G-quadruplexes and modulate their downstream biochemical/biological effects. Translation inhibition by the γPNA probe was more potent, more

specific or both, relative to the homologous PNA and 2'-OMe oligomers. The PNA probe is as potent as the γPNA oligomer, but the significant off-target activity of the former makes it difficult to de-conflate its specific from non-specific effects in the context of the target reporter. It is possible that the rigid structure predicted for the γPNA oligomer makes it less tolerant than the unmodified PNA to bulges resulting from off-target binding. Additionally, the slower off-rate observed for γPNA relative to PNA in the SPR binding analyses suggests that the former might be kinetically trapped within its corresponding hybrid duplex with the target, thereby limiting its availability to potential off-target sites. The γPNA probe was also more potent and specific than the 2'-OMe oligomer. In this context, the negatively-charged character of the phosphodiester backbone likely decelerates hybridization by the 2'-OMe probe and destabilizes its hybrid with the target. Regardless of the basis for these differences, the assay results provide evidence for the functional superiority of γPNA over PNA and 2'-OMe in translation inhibition for this target.

Our results can be compared to other recent reports of antisense approaches to modulate quadruplex function. In an important example, Murat *et al* reported that complementary DNA and RNA probes could invade the guanine quadruplexes formed within the coding sequence (CDS) of the Epstein-Barr virus maintenance protein and facilitate translation by previously-stalled ribosomes both *in vitro* and *in vivo*[48]. Evidently, ribosomal translocation was more adversely affected by the cluster of quadruplex structures in the native mRNA than by the hybrid duplexes formed by the antisense agents, resulting in increased translation[48].

Rouleau *et al* have provided a more closely related example by directing complementary 2'-OMe RNA antisense oligonucleotides (ASOs) to three-tetrad RNA quadruplexes located within the 5'-UTR of both reporter and endogenous genes[34]. The quadruplex targets studied by the authors featured one long (12-13 nts) unstructured central loop and two peripheral single-nucleotide loops.

The authors demonstrated that complementary probes designed to bind the central two G-tracts and intervening loop nucleotides inhibited quadruplex formation *in vitro* and modestly enhanced translation *in cellulo*, an effect that was attributed to inhibition of quadruplex formation within the 5'-UTR[34]. As in the previous case by Murat *et al*, this indicates that the heteroduplex formed by the antisense oligonucleotide and the mRNA is *less inhibitory* than is the endogenous quadruplex structure[34], in contrast to our findings. (Interestingly, *in vitro* experiments indicated that the antisense oligonucleotide was unable to bind to a pre-folded RNA quadruplex. Nevertheless, statistically significant effects on translation were observed for both cellular reporters and endogenous mRNA targets, indicating that the ASOs were capable of hybridizing to their targets *in cellulo*.) At this point, it is unclear why the 2'-OMe ASO reported by Rouleau et al had a weak stimulatory effect on translation[34], whereas our experiments with a 2'-OMe ASO led to weak inhibition.

Eukaryotic translation is a complicated multi-step process controlled mainly by a rate-limiting initiation step[49]. Briefly, the 40S ribosomal subunit binds to specific structures in the 5'-UTR and scans this region of the transcript for the start codon[49]. Location of this signal is followed by recruitment of the 60S subunit to form the complete (80S) ribosomal machinery that proceeds with translation across the CDS[49]. In contrast to the aforementioned reports[34, 48], we find that the quadruplex-to-hybrid duplex structural conversion results in further translation inhibition, indicating that the γPNA-RNA hybrids are stronger impediments to ribosomal translocation than the starting quadruplex, even with the relatively short (12 bp) length of the heteroduplexes.

At least three important factors might explain the divergence of our results from previous efforts to target RNA quadruplexes using antisense molecules. First, the position (UTR or CDS) of the hybrid resulting from quadruplex invasion might determine the effect of an antisense molecule on

ribosomal translocation. Second, the fact that the ribosomal structure varies depending on its position on the transcript (i.e. 40S in the 5'-UTR, 80S in the CDS) might impart differential effects on processing of structures encountered along the transcript. Third, even for hybrids formed in the same position, the affinity of the probes employed (e.g. 2'-OMe RNA or γPNA) could determine their functional outcomes.

In this context, we recently reported that a complementary PNA 8mer could bind to a DNA G quadruplex and enhance the accumulation of full-length products from primer extension by DNA pol η[26]. Interestingly, higher concentrations of the PNA probe simultaneously elicited strong inhibition at the predicted binding sites and increased the amounts of full-length products, suggesting that some of the bound PNA molecules were displaced by the translocating polymerase[26]. These results are consistent with the hypothesis that PNA oligomers—carefully tuned in affinities for their targets, could be used to stimulate, rather than inhibit, nucleic acid processing. It is therefore possible that, in the current demonstration, variation of the length of the γPNA without altering the target site could very well result in a transition from translation stimulation to inhibition as the γPNA length increases. Regardless of the basis for this divergence between our results and published data, the ability to enhance and suppress translation is likely to be useful, since regulation—by the strictest definition—requires that we can drive expression in both directions.

Much remains to be learned about quadruplex invasion and antisense effects by γPNAs. We intend to examine the number of overhanging bases required to drive efficient invasion, since secondary structures up- and down-stream of the quadruplex are likely to impact the accessibility of its adjacent nucleotides to antisense molecules. Application of structure-analysis methods such as SHAPE[50] or footprinting[51] and correlation with antisense activity might lead to a better

understanding of the extent to which local secondary structure affects γPNA hybridization to quadruplexes. Enhancing gene expression with γPNAs is also of interest to us. Such stimulation might be achieved by carefully tuning oligomer length to be simultaneously sufficient for quadruplex invasion yet susceptible to ribosomal displacement.

It would also be interesting to examine the positional bias (if any) of γPNA antisense effects *in vivo*. Towards this end, additional modifications would be required on the probes for efficient intracellular delivery. Cellular membrane translocation of γPNAs might be achieved by conjugating the probes to cell-penetrating peptides originally utilized for delivery of peptide fragments[52]. Additionally, Ly and coworkers have observed efficient auto-transfection of PNA oligomers possessing guanidinium functional groups at the γ-position of the PNA backbone[53]. Delivery of γPNAs by polymeric nanoparticles has also been reported and represents a viable alternative[54].

In conclusion, our results demonstrate the efficacy of relatively short complementary γPNAs for invading a stable RNA G quadruplex structure and inhibiting translation in cell lysate. The high affinity of γPNA and the potential for discriminating among closely related quadruplex targets based on recognition of flanking nucleotides, as recently demonstrated by Basu and coworkers[55], gives confidence that sequence-based targeting is a viable strategy for regulating G-quadruplex function.

2.5 REFERENCES

[1] Malina, A., Cencic, R., and Pelletier, J. (2011) Targeting Translation Dependence in Cancer, *Oncotarget 2*, 76-88.

[2] Meric, F., and Hunt, K. K. (2002) Translation initiation in cancer a novel target for therapy, *Mol. Cancer Ther. 1*, 971-979.

[3] Clemens, M. J., and Bommer, U. A. (1999) Translational control the cancer connection, *Int. J. Biochem. Cell Biol. 31*, 1-23.

[4] Penagarikano, O., Mulle, J. G., and Warren, S. T. (2007) The pathophysiology of fragile x syndrome, *Annu. Rev. Genomics Hum. Genet. 8*, 109-129.

[5] Balasubramanian, S., Hurley, L. H., and Neidle, S. (2011) Targeting G-quadruplexes in gene promoters: a novel anticancer strategy?, *Nat. Rev. Drug Discov. 10*, 261-275.

[6] Bochman, M. L., Paeschke, K., and Zakian, V. A. (2012) DNA secondary structures: stability and function of G-quadruplex structures, *Nat. Rev. Genet. 13*, 770-780.

[7] Brooks, T. A., and Hurley, L. H. (2010) Targeting MYC Expression through G-Quadruplexes, *Genes & Cancer 1*, 641-649.

[8] Collie, G. W., and Parkinson, G. N. (2011) The application of DNA and RNA G-quadruplexes to therapeutic medicines, *Chem. Soc. Rev. 40*, 5867-5892.

[9] Zhang, S., Wu, Y., and Zhang, W. (2014) G-quadruplex structures and their interaction diversity with ligands, *ChemMedChem 9*, 899-911.

[10] Watson, J. D., and Crick, F. H. C. (1953) Molecular Structure of Nucleic Acids- A Structure for Deoxyribose Nucleic Acid, *Nature 171*, 737-738.

[11] Raghuraman, M. K., and Cech, T. R. (1990) Effect of monovalent cation-induced telomeric DNA structure on the binding of Oxytricha telomeric protein, *Nucleic Acids Res. 18*, 4543-4552.

[12] Phan, A. T., and Mergny, J. L. (2002) Human telomeric DNA G-quadruplex , i-motif and Watson ± Crick double helix, *Nucleic Acids Res. 30*, 4618-4625.

[13] Arora, A., Nair, D. R., and Maiti, S. (2009) Effect of flanking bases on quadruplex stability and Watson-Crick duplex competition, *FEBS J. 276*, 3628-3640.

[14] Kumar, N., Sahoo, B., Varun, K. A., Maiti, S., and Maiti, S. (2008) Effect of loop length variation on quadruplex-Watson Crick duplex competition, *Nucleic Acids Res. 36*, 4433-4442.

[15] Li, J. J., and Tan, W. (2002) A Single DNA Molecule Nanomotor, *Nano Lett. 2*, 315–318.

[16] Alberti, P., and Mergny, J. L. (2003) DNA duplex-quadruplex exchange as the basis for a nanomolecular machine, *Proc. Natl. Acad. Sci. U. S. A. 100*, 1569-1573.

[17] Alberti, P., Bourdoncle, A., Sacca, B., Lacroix, L., and Mergny, J. L. (2006) DNA nanomachines and nanostructures involving quadruplexes, *Org. Biomol. Chem. 4*, 3383-3391.

[18] Kumar, N., and Maiti, S. (2005) The effect of osmolytes and small molecule on Quadruplex-WC duplex equilibrium: a fluorescence resonance energy transfer study, *Nucleic Acids Res. 33*, 6723-6732.

[19] Green, J. J., Ying, L., Klenerman, D., and Balasubramanian, S. (2003) Kinetics of unfolding the human telomeric DNA quadruplex using a PNA trap, *J. Am. Chem. Soc. 125*, 3763–3767.

[20] Kumar, N., and Maiti, S. (2007) Role of Locked Nucleic Acid Modified Complementary Strand in Quadruplex Watson - Crick Duplex Equilibrium, *J. Phys. Chem. B 111*, 12328-12337.

[21] Kumar, N., Patowary, A., Sivasubbu, S., Petersen, M., and Maiti, S. (2008) Silencing c-MYC Expression by Targeting Quadruplex in P1 Promoter Using Locked Nucleic Acid Trap, *Biochemistry 47*, 13179-13188.
[22] Datta, B., and Armitage, B. A. (2001) Hybridization of PNA to Structured DNA Targets: Quadruplex Invasion and the Overhang Effect, *J. Am. Chem. Soc. 123*, 9612–9619.
[23] Marin, V. L., and Armitage, B. A. (2005) RNA Guanine Quadruplex Invasion by Complementary and Homologous PNA Probes, *J. Am. Chem. Soc. 127*, 8032–8033.
[24] Marin, V. L., and Armitage, B. A. (2006) Hybridization of Complementary and Homologous Peptide Nucleic Acid Oligomers to a Guanine Quadruplex-Forming RNA, *Biochemistry 45*, 1745–1754.
[25] Gupta, A., Lee, L., Roy, S., Tanious, F. A., Wilson, W. D., Ly, D. H., and Armitage, B. A. (2013) Strand invasion of DNA quadruplexes by PNA: comparison of homologous and complementary hybridization, *ChemBioChem 14*, 1476–1484.
[26] Murphy, C. T., Gupta, A., Armitage, B. A., and Opresko, P. L. (2014) Hybridization of G-quadruplex-forming peptide nucleic acids to guanine-rich DNA templates inhibits DNA polymerase eta extension, *Biochemistry 53*, 5315-5322.
[27] Dragulescu-Andrasi, A., Rapireddy, S., Frezza, B. M., Gayathri, C., Gil, R. R., and Ly, D. H. (2006) A Simple γ-Backbone Modification Preorganizes Peptide Nucleic Acid into a Helical Structure, *J. Am. Chem. Soc. 2006*, 10258–10267.
[28] Sahu, B., Sacui, I., Rapireddy, S., Zanotti, K. J., Bahal, R., Armitage, B. A., and Ly, D. H. (2011) Synthesis and characterization of conformationally preorganized, (R)-diethylene glycol-containing gamma-peptide nucleic acids with superior hybridization properties and water solubility, *J. Org. Chem. 76*, 5614-5627.
[29] Roy, S., Zanotti, K. J., Murphy, C. T., Tanious, F. A., Wilson, W. D., Ly, D. H., and Armitage, B. A. (2011) Kinetic discrimination in recognition of DNA quadruplex targets by guanine-rich heteroquadruplex-forming PNA probes, *Chem. Commun. 47*, 8524-8526.
[30] Roy, S., Tanious, F. A., Wilson, W. D., Ly, D. H., and Armitage, B. A. (2007) High-affinity homologous peptide nucleic acid probes for targeting a quadruplex-forming sequence from a MYC promoter element, *Biochemistry 46*, 10433–10443.
[31] Kumari, S., Bugaut, A., Huppert, J. L., and Balasubramanian, S. (2007) An RNA G-quadruplex in the 5' UTR of the NRAS proto-oncogene modulates translation, *Nat. Chem. Biol. 3*, 218-221.
[32] Bugaut, A., Rodriguez, R., Kumari, S., Hsu, S. T., and Balasubramanian, S. (2010) Small molecule-mediated inhibition of translation by targeting a native RNA G-quadruplex, *Org. Biomol. Chem. 8*, 2771-2776.
[33] Kumari, S., Bugaut, A., and Balasubramanian, S. (2008) Position and stability are determining factors for translation repression by an RNA G-quadruplex-forming sequence within the 5' UTR of the NRAS proto-oncogene, *Biochemistry 47*, 12664-12669.
[34] Rouleau, S. G., Beaudoin, J. D., Bisaillon, M., and Perreault, J. P. (2015) Small antisense oligonucleotides against G-quadruplexes: specific mRNA translational switches, *Nucleic Acids Res. 43*, 595-606.
[35] Vorlickova, M., Kejnovska, I., Sagi, J., Renciuk, D., Bednarova, K., Motlova, J., and Kypr, J. (2012) Circular dichroism and guanine quadruplexes, *Methods 57*, 64-75.
[36] Williamson, J. R. (1993) Guanine quartets, *Curr. Opin. Struct. Biol. 3*, 357-362.

[37] Zhang, A. Y., Bugaut, A., and Balasubramanian, S. (2011) A sequence-independent analysis of the loop length dependence of intramolecular RNA G-quadruplex stability and topology, *Biochemistry 50*, 7251-7258.
[38] Hud, N. V., Smith, F. W., Anet, F. A., and Feigon, J. (1996) The Selectivity for K+ versus Na+ in DNA Quadruplexes Is Dominated by Relative Free Energies of Hydration: A Thermodynamic Analysis by 1H NMR, *Biochemistry 35*, 15383-15390.
[39] Morris, M. J., and Basu, S. (2009) An unusually stable G-quadruplex within the 5'-UTR of the MT3 matrix metalloproteinase mRNA represses translation in eukaryotic cells, *Biochemistry 48*, 5313-5319.
[40] Halder, K., Wieland, M., and Hartig, J. S. (2009) Predictable suppression of gene expression by 5'-UTR-based RNA quadruplexes, *Nucleic Acids Res. 37*, 6811-6817.
[41] Gomez, D., Guedin, A., Mergny, J. L., Salles, B., Riou, J. F., Teulade-Fichou, M. P., and Calsou, P. (2010) A G-quadruplex structure within the 5'-UTR of TRF2 mRNA represses translation in human cells, *Nucleic Acids Res. 38*, 7187-7198.
[42] Agarwala, P., Pandey, S., and Maiti, S. (2014) Role of G-quadruplex located at 5' end of mRNAs, *Biochim. Biophys. Acta 1840*, 3503-3510.
[43] Jodoin, R., Bauer, L., Garant, J. M., Mahdi Laaref, A., Phaneuf, F., and Perreault, J. P. (2014) The folding of 5'-UTR human G-quadruplexes possessing a long central loop, *RNA 20*, 1129-1141.
[44] Mergny, J. L., Phan, A. T., and Lacroix, L. (1998) Following G-quartet formation by UV-spectroscopy, *FEBS Lett. 435*, 74-78.
[45] Egholm, M., Buchardt, O., Christensen, L., Behrens, C., Freier, S. M., Driver, D. A., Berg, R. H., Kim, S. K., Norden, B., and Nielsen, P. E. (1993) PNA hybridizes to complementary oligonucleotides obeying the Watson–Crick hydrogen-bonding rules, *Nature 365*, 566-568.
[46] Marky, L. A., and Breslauer, K. J. (1987) Calculating thermodynamic data for transitions of any molecularity from equilibrium melting curves, *Biopolymers 26*, 1601-1620.
[47] Chang, A. L., McKeague, M., Liang, J. C., and Smolke, C. D. (2014) Kinetic and equilibrium binding characterization of aptamers to small molecules using a label-free, sensitive, and scalable platform, *Anal. Chem. 86*, 3273-3278.
[48] Murat, P., Zhong, J., Lekieffre, L., Cowieson, N. P., Clancy, J. L., Preiss, T., Balasubramanian, S., Khanna, R., and Tellam, J. (2014) G-quadruplexes regulate Epstein-Barr virus-encoded nuclear antigen 1 mRNA translation, *Nat. Chem. Biol. 10*, 358-364.
[49] Gebauer, F., and Hentze, M. W. (2004) Molecular mechanisms of translational control, *Nat. Rev. Mol. Cell. Biol. 5*, 827-835.
[50] Wilkinson, K. A., Merino, E. J., and Weeks, K. M. (2006) Selective 2'-hydroxyl acylation analyzed by primer extension (SHAPE): quantitative RNA structure analysis at single nucleotide resolution, *Nat. Protoc. 1*, 1610-1616.
[51] Tijerina, P., Mohr, S., and Russell, R. (2007) DMS footprinting of structured RNAs and RNA-protein complexes, *Nat. Protoc. 2*, 2608-2623.
[52] Copolovici, D. M., Langel, K., Eriste, E., and Langel, Ü. (2014) Cell-Penetrating Peptides Design, Synthesis, and Applications, *Acs Nano 8*, 1972-1994.
[53] Sahu, B., Chenna, V., Lathrop, K. L., Thomas, S. M., Zon, G., Livak, K. J., and Ly, D. H. (2009) Synthesis of Conformationally Preorganized and Cell-Permeable Guanidine-Based γ-Peptide Nucleic Acids (γGPNAs), *J. Org. Chem. 74*, 1509–1516.

[54] Bahal, R., McNeer, N. A., Ly, D. H., Saltzman, W. M., and Glazer, P. M. (2014) Nanoparticle for delivery of antisense γPNA oligomers targeting CCR5, *Artificial DNA: PNA & XNA 4*, 49-57.
[55] Bhattacharyya, D., Nguyen, K., and Basu, S. (2014) Rationally induced RNA:DNA G-quadruplex structures elicit an anticancer effect by inhibiting endogenous eIF-4E expression, *Biochemistry 53*, 5461-5470.

CHAPTER 3

INVASION OF FOUR- AND TWO-TETRAD G-QUADRUPLEXES BY COMPLEMENTARY γPNA OLIGOMERS

3.1 INTRODUCTION

Guanine-rich sequences with potential to form four- or two-tetrad quadruplexes are becoming interesting targets for recognition due to recent reports assigning them causative roles in certain human diseases. Specifically, expansions of G_4-[1] and G_2-containing[2] repeats have been implicated in the pathophysiology of certain neurodegenerative[1] and neurodevelopmental[2] disorders. Concomitant folding by some of these sequences into four-[3-5] or two-tetrad[6, 7] quadruplexes is now well established, while the molecular mechanisms underlying their respective contributions to disease progression are at different levels of development.

For example, stable, four-tetrad, intramolecular quadruplexes within the G_4C_2 intronic repeats of the *C9orf72* gene—the most frequently reported genetic cause of both amyotrophic lateral sclerosis and frontotemporal dementia[8]—induce premature transcription termination[1], with the accumulation of abortive transcripts correlating positively with the degree of repeat expansion[1]. The molecular mechanism underpinning this phenomenon involves the sequestration of otherwise-active nucleolin by G-quadruplex structures in the transcribed mRNA—resulting in a dispersed nuclear distribution of the protein, rather than localized accumulation in the nucleolus[1]. Also, expansions of the G_2C repeat in the 5'-untranslated region (UTR) of the fragile X mental retardation 1 (FMR1) gene result in loss of function of the encoded protein (FMRP), leading to fragile x syndrome (FXS)[2]. Although quadruplex formation by the G-rich repeats has been established *in vitro* for synthetic, truncated DNA sequences[6]—their existence in the context of the full-length genomic DNA, and/or their precise contributions to disease pathology remain nebulous and constitute an active research area[9].

Both the established and speculated contributions of these G-quadruplex structures to disease pathology incentivize the development of specific sequence-recognition agents that could invade

four- or two-tetrad quadruplex targets and function as either therapeutics or probes for potential quadruplex function. However, while repeat expansions are prime candidates for antisense intervention[10, 11], quadruplex-laden targets pose unique challenges to antisense targeting[12], since the binding sites—although known, are not immediately accessible to the ligands. Further, four-tetrad quadruplexes, even when present as isolated structures[5], and two-tetrad quadruplexes, when present in clusters[13]—can pose a significant thermodynamic barrier to hybridization—likely exacerbated—as in *C9orf72*[1], by competing protein ligands.

We recently showed that γPNA[14, 15] oligomers complementary to different regions of an RNA quadruplex forming sequence (QFS) could hybridize to their respective target sequences (Chapter 2), even when the latter were partially or fully sequestered within a stable, three-tetrad, intramolecular G-quadruplex. The binding event resulted in exceptionally stable hybrid duplexes that were characterized by low-nM dissociation constants (K_D). We also showed that these probes elicited dose-dependent inhibition of translation when directed against a QFS motif inserted into the 5-UTR of a luciferase reporter gene. Interestingly, we observed a strong positional bias in the potency of inhibition, with greater repressive effects observed for the probe directed to the 5'-end of the QFS (Chapter 2).

The data presented in this chapter show that γPNA oligomers complementary to the 5'-ends of both G_2 and G_4-rich RNA targets can hybridize to their respective cognate sequences, leading to stable hybrid duplexes with low/sub-nM dissociation constants. In some cases, γPNA hybridization is accompanied by unfolding of a very stable existing quadruplex structure, necessitating that a second equivalent of the probe anneal to an additional site on the QFS to prevent refolding. The probes also readily overcome the kinetic barrier to hybridization posed by the stable quadruplex structure (where present)—an effect that is further enhanced by quadruplex-

destabilizing buffer conditions. Finally, the probes are potent and specific inhibitors of translation when directed against G-rich inserts in the 5'-UTR of a reporter transcript.

3.2 MATERIALS AND METHODS

3.2.1 γPNA/RNA/DNA Oligomers

The γPNA oligomers (Chart 3.1) used for all experiments were purchased from PNA Innovations Inc. (www.pnainnovations.com) and included C-terminal L-lysine residues. RNA and DNA oligomers used were obtained from Integrated DNA Technologies (www.idtdna.com). Sequences of all biotinylated DNA oligonucleotides used as capture strands in SPR competition experiments and the RNA oligomers designed for SPR direct-binding experiments are also given below. (RNA/DNA sequences written 5'-3'.)

	Sequence
4G4 RNA (for SPR)	AGACCCAAGCACUAUAAGCUAGCGGGGAGGGGCGGGGUCUGGGGCGAUCC
4G2 RNA (for SPR)	AGACCCAAGCACUAUAAGCUAGCGGAAGGCCGGUUCUGGCCGAUCC
γ4G4 capture strand	biotin- ATTACTAGCGGGGAGGATTA (DNA)
γ4G2 capture strand	biotin- ATTACTAGCGGAAGGCATTA (DNA)

3.2.2 UV Melting Experiments

Thermal melting experiments were performed on a Varian Cary 300 spectrophotometer equipped with a thermoelectrically-controlled multicell holder. Samples were prepared in a buffer containing 10 mM Tris-HCl (pH 7.4), 0.1 mM Na/Li_2EDTA, and different concentrations of KCl/LiCl. The solutions were incubated at 95 °C for 5 min and cooled to 15 °C at a rate of 1 °C /min. Subsequently, the annealed samples were incubated at 15 °C for 5 min, and a heating ramp was applied at 1 °C/min up to 95 °C. Melting curves were generated by monitoring absorbance values at 275 nm (for hetero-duplexes) and 295 nm (for quadruplexes) every 0.5 °C. Where

possible, the melting temperature (T_m) was determined from a first-derivative plot of the respective melting curve. Each reported melting temperature is the average of three independent experiments.

3.2.3 Circular Dichroism (CD) Spectropolarimetry

CD spectra were obtained on a Jasco J-715 circular dichroism spectropolarimeter equipped with a water-circulating temperature controller. Samples were prepared in a buffer containing 10 mM Tris-HCl (pH 7.4), 0.1 mM Na/Li_2EDTA, and different concentrations of KCl/LiCl. The samples were incubated at 95 °C for 5 min and cooled slowly to room temperature. All spectra were collected at 37 °C. Each spectrum was collected after 6 different scans (200 – 360 nm) at a scan rate of 100 nm/min and baseline corrected. The CD spectra for 3 different samples were collected and averaged.

3.2.4 Surface Plasmon Resonance (SPR) Experiments

All SPR experiments were performed on a Biacore T100 instrument equipped with a four-channel CM5 sensor chip (GE Health Care). The sensor chip was coated with a carboxylmethyl dextran matrix that allows further functionalization with streptavidin via a standard NHS-EDC coupling procedure. Covalent immobilization of streptavidin was continued until 7000 response units (RU) of the protein were captured on each of the four channels (flow cells). The final step of the sensor design involved non-covalent immobilization of a biotinylated capture strand on the streptavidin-modified surface.

The sensor surface was primed by a buffer injection [100 mM KCl/LiCl, 10 mM Tris-HCl (pH 7.4), 3 mM Na/Li_2EDTA, and 0.005 % v/v P-20 surfactant] prior to all experiments. The capture strand utilized for direct-binding studies with either 4Gx RNA was a biotinylated 15mer γPNA oligomer that was designed to hybridize to a complementary sequence appended to the 5’-end of

the RNA QFS. The active flow cell was modified with a high density (1000 RU) of the biotinylated γPNA capture strand. Immobilization of the target RNA was then performed by injecting a 100 nM solution (pre-annealed for 2 h) over the active flow cell for 400 s (flow rate = 50 µL/min) until a medium density (300 RU) of the RNA accumulated on the surface. A short dissociation phase (20 s) was introduced to remove any unhybridized or loosely-bound RNA from the surface.

To study γPNA hybridization to the immobilized RNA, 50 nM of each probe was injected for 400 s (flow rate = 30 µL/min) over both the active and reference flow cells, after which the buffer injection was restarted and continued for 600 s to monitor dissociation of the hybrids. The sensor surface was regenerated with a solution of 10 mM NaOH and 1 M NaCl (2x for 30 s), and another buffer injection (120 s) was introduced to wash off residual regeneration solution. The sensor surface was replenished with fresh RNA, and subsequent binding studies were continued by simple iterations of the aforementioned steps. The initial association rates were calculated based on the times required to reach 5 RUs of bound γPNA, similar to a procedure utilized by us in an earlier report[16].

For the competition assay[17], 1000 RU of a DNA capture strand was immobilized on the surface of the chip. The flow cell containing the capture strand was subsequently calibrated for the free γPNA concentration by injecting a series of solutions containing 5 – 50 nM of the free probe for 200 s. The slopes for the sensorgrams were obtained over a 10 s window (beginning 30 s post-injection) and plotted against the concentration of free γPNA to obtain a linear regression curve.

A fixed concentration of each γPNA probe (10 or 20 nM) was then incubated with increasing concentrations of the pre-annealed RNA target at 25 °C for 2 hours prior to injection over the sensor surface. The slopes for the resulting sensorgrams were then used to obtain the amounts of free γPNA in equilibrium with the hybrid using the previously-generated regression curve

(equation 1). Further, the concentration of bound γPNA was calculated by assuming that the system obeys the mass conservation condition (equation 2).

$$[\gamma PNA]_{free} = \frac{Slope_{40-50\,s}}{M} \quad \boldsymbol{Eqn\ 1}$$

$$M = slope\ of\ calibration\ curve$$

$$[\gamma PNA]_{Total} = [\gamma PNA]_{free} + [\gamma PNA]_{bound} \quad \boldsymbol{Eqn\ 2}$$

The bound-probe concentration, in turn, corresponds to the amount of hybrid (duplex) formed for a system involving two interacting components (equation 3). Therefore, the accumulation of the duplex after each addition of the RNA target could be determined and fit to a Langmuir isotherm that accounts for ligand depletion (equation 4), where $[P]_T$ = total γPNA concentration; $[R]_T$ = total RNA concentration; K_D = equilibrium dissociation constant.

$$[\gamma PNA]_{bound} \equiv [Duplex] = [\gamma PNA]_{Total} - [\gamma PNA]_{free} \quad \boldsymbol{Eqn\ 3}$$

$$[Duplex] = \frac{b - \sqrt{b^2 - 4 \bullet [P]_T \bullet [R]_T}}{2 \bullet [P]_T} \quad \boldsymbol{Eqn\ 4}$$

$$b = [R]_T + [P]_T + K_D$$

3.2.5 Template Production and *In Vitro* Transcription

The DNA template containing the target sequence upstream of the firefly luciferase gene (4Gx reporter) was amplified by Polymerase Chain Reaction (PCR) using Taq DNA polymerase (New England Biolabs) and two sequence specific primers: forward primer (5'-AATACGCAAACCGCCTCTC-3') and reverse primer (5'-GGTGATGTCGGCGATATAGG-3'). A similar procedure was also used to obtain a control reporter without the target sequence upstream of the luciferase coding sequence. The resulting DNA fragments were purified using a

GeneJET PCR purification kit (ThermoFisher) according to the manufacturer's instructions. The integrity and length of each PCR product were verified using a 1 % agarose gel and electrophoresis.

The purified DNA templates were transcribed *in vitro* using a cocktail consisting of T7 RNA polymerase (100 units) and a mixture of the ribonucleotide triphosphates (500 µM of each NTP) in a total reaction volume of 100 µL at 37 ^{o}C for 2 h. The resulting mRNA transcripts were purified using the GeneJET RNA clean-up Micro Kit (Thermo Scientific) according to the manufacturer's instructions. The final mRNA concentrations were estimated by UV spectroscopy. The extinction coefficient of each mRNA transcript was assumed to be the sum of the molar absorptivities of each nucleobase, and the recorded absorbance (at 260 nm) was assumed to be unaffected by any (undetermined) secondary folding of the transcripts. The nucleotide sequences for the UTRs in both the experimental (4Gx) and control transcripts are presented below, with the quadruplex-forming fragment, where present, highlighted in bold and upper-case, and underlined.

4G4: gagacccaagcuuucagauccgcuagc**GGGGAGGGGCGGGGUCUGGGG**cgauccagccaccaug

4G2: gagacccaagcuuucagauccgcuagc**GGAAGGCCGGUUCUGG**ccgauccagccaccaug

Control: gaucuaauaucuacuuaagaacacaaaacucgagaaccaug

3.2.6 Cell-Free Translation

A translation reaction was typically performed by incubating the purified transcripts in a mixture containing nuclease-treated rabbit reticulocyte lysate (70 % v/v) (Promega), 10 µM amino acid mixtures minus leucine, 10 µM amino acid mixtures minus methionine and 20 units RNasin ribonuclease inhibitor (50 µL total reaction volume) at 30 ^{o}C for 1.5 h. Experiments to examine the effect of each γPNA on luciferase production were performed by pre-incubating the mRNA with requisite amino acids and increasing concentrations of the γPNA in a buffer containing 79

mM KCl and 7.9 mM Tris-HCl (pH = 7.4) to a total volume of 19.5 µL at 37 ^{o}C for 1 h. Each subsequent translation reaction was started by adding 30.5 µL of the rabbit reticulocyte lysate to the pre-incubated mixture, and the final mixture was incubated at 30 ^{o}C for 1.5 h. Finally, we performed a luciferase assay by incubating 10 µL of the translation products with 50 µL of a reagent cocktail ($_{D}$-luciferin, Mg^{2+}, and ATP) (Promega) to estimate the relative amounts of enzyme produced from the mRNA reporter. Luciferase activity was estimated as relative light units (RLU) on a Tecan Infinite M1000 Plate Spectrometer.

3.3 RESULTS

3.3.1 Target Selection and Probe Design

Two G-rich targets (**4G4** and **4G2**, Chart 3.1) were selected based on the **4G3** sequence reported in Chapter 2. We previously showed that **4G3** folds into an intramolecular G-quadruplex consisting of three stacked G-tetrads, with consequent repressive effects on translation (Chapter 2). By inserting (as in **4G4**) or mutating (as in **4G2**) one guanine within each G_3 tract of **4G3**, we constructed additional reporter transcripts predicted to possess quadruplex motifs with divergent biophysical and/or biochemical properties relative to **4G3**. These constructs were expected to yield insight into any existing nexus between the stability of the quadruplex motif and its repressive effects on translation—as has been suggested in a previous report.[18] Further, two- and four-tetrad quadruplexes—as are predicted for **4G4** and **4G2**, respectively—have recently been studied as regulatory or causative elements in gene expression[19] and human disease[1]. Therefore, we sought to explore these structures as targets for recognition by complementary γPNA oligomers.

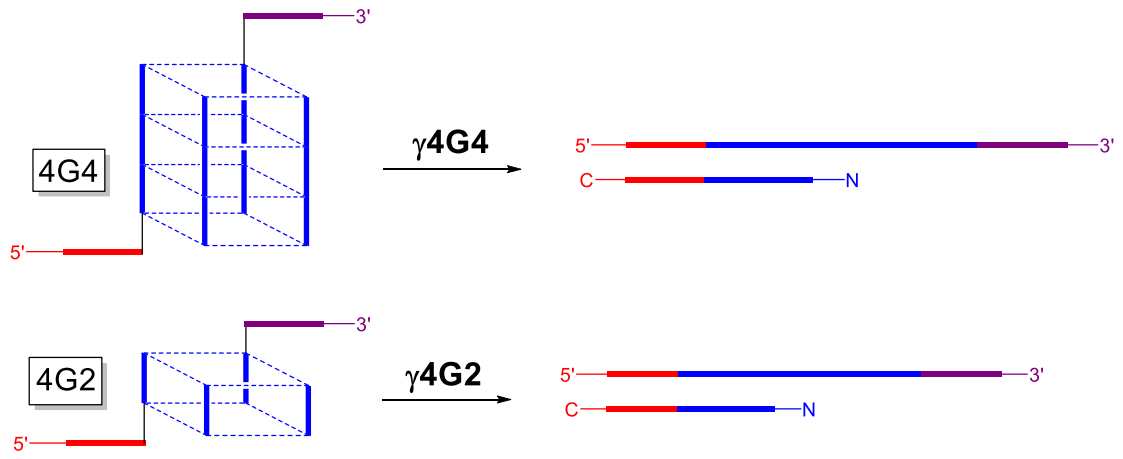
Chart 3.1: Sequences of γPNA and RNA oligomers used in this study. (RNA sequences written 5'-3' and γPNA sequences written C-N. Underlined Gs are predicted to participate in G tetrads.)

Oligomer	Sequence
4G4 RNA	GCUAGCGGGGAGGGGCGGGGUCUGGGGCGAUCC
4G2 RNA	GCUAGCGGAAGGCCGGUUCUGGCCGAUCC
γ4G4	GATCGCCCCTCC
γ4G2	GATCGCCTTCCG
4G4 rComp	CUAGCGGGGAGG
4G2 rComp	CUAGCGGAAGGC

12mer γPNAs (Chart 3.1) complementary to the 5'-end of each respective QFS were employed as probes in our study. The choice of this position (5'-end) was informed by our previous observation of a positional, kinetically influenced bias in the potency of inhibition by complementary γPNAs

targeted to different regions of a QFS motif in the context of a luciferase reporter assay (Chapter 2). Further analyses revealed that recognition of the 5'-overhanging nucleotides of the quadruplex was crucial to translation inhibition by the most potent probe (Chapter 2). We therefore designed **γ4G4** and **γ4G2** (Chart 3.1) to recognize the 5'-overhanging nucleotides of their respective quadruplex targets, in addition to seven internal nucleotides of the QFS (Scheme 3.1).

Scheme 3.1



Recognition of 4G4 or 4G2 RNA targets by γPNA oligomers complementary to the respective 5'-ends

3.3.2 Biophysical Characterization of 4G4 and 4G2 RNAs

3.3.2.1 4G4

Quadruplex folding by this sequence was predicted to result in four stacked G-tetrads, with non-bonded yet internal residues extruded to the loops (1, 1, and 3 nts, respectively). CD spectra of **4G4** (in KCl) possessed a minimum and a maximum at 245 and 265 nm, respectively—both of which were diminished when identical experiments were performed in LiCl (Figure 3.1A). Both the positions of the peaks and their attenuation upon the cation switch are consistent with parallel quadruplex[20] folding by this RNA sequence in KCl, and its unfolding in LiCl[21]. Evidence of

quadruplex stability was also obtained from UV melting analyses, where we observed a hypochromic melting transition[22] for **4G4** in 1 mM KCl (T_m = 61.9 ± 0.7 °C), which was eliminated in 1 mM LiCl (Figure 3.1B). Importantly, quadruplex stability is concentration-independent, as melting experiments performed with a five-fold higher concentration of the RNA showed a similar transition temperature (Figure B1). This result confirms that the folding reaction is unimolecular and results in an intramolecular quadruplex, as was suggested in a previous report[18] on a very similar sequence.

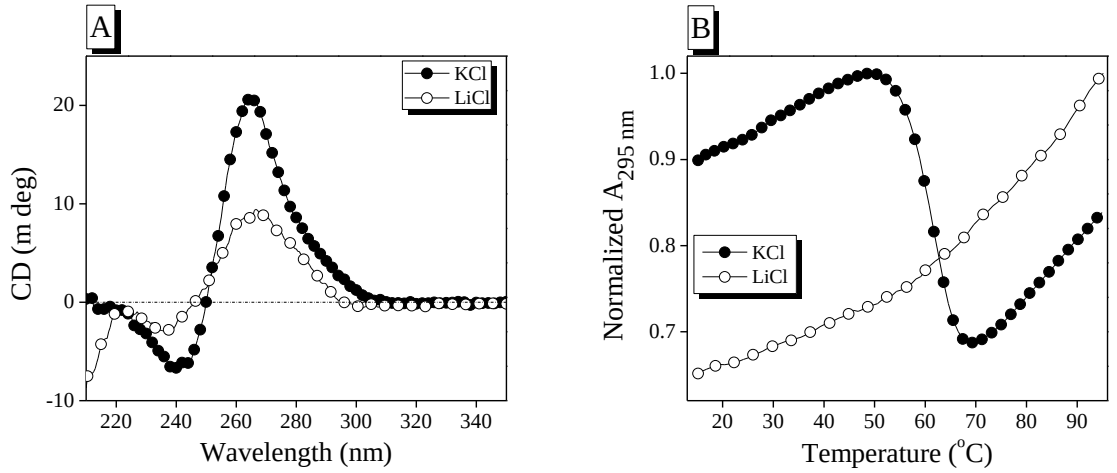


Figure 3.1. Biophysical characterization of 4G4 RNA quadruplex. CD spectra (**A**) and UV-melting curves (**B**) of 4G4 RNA target. All samples contained 2 µM RNA buffered in 1 mM KCl (closed circles) or 1 mM LiCl (open circles).

3.3.2.2 **4G2**

Although this sequence was predicted to fold into a quadruplex bearing two stacked G-tetrads, CD analyses revealed no differences, in either peak positions or intensities, between the K^+ and Li^+-bound forms of the RNA (Figure 3.2A). UV-melting analyses on **4G2** also showed identical hyperchromic transitions in KCl and LiCl buffers (Figure 3.2B). While these data suggest that **4G2** does not adopt a quadruplex fold under the conditions of our experiments, it is possible that folding

might occur at elevated salt concentrations (>100 mM K^+) and/or the crowded environment of the cell lysates—neither of which are accurately simulated here.

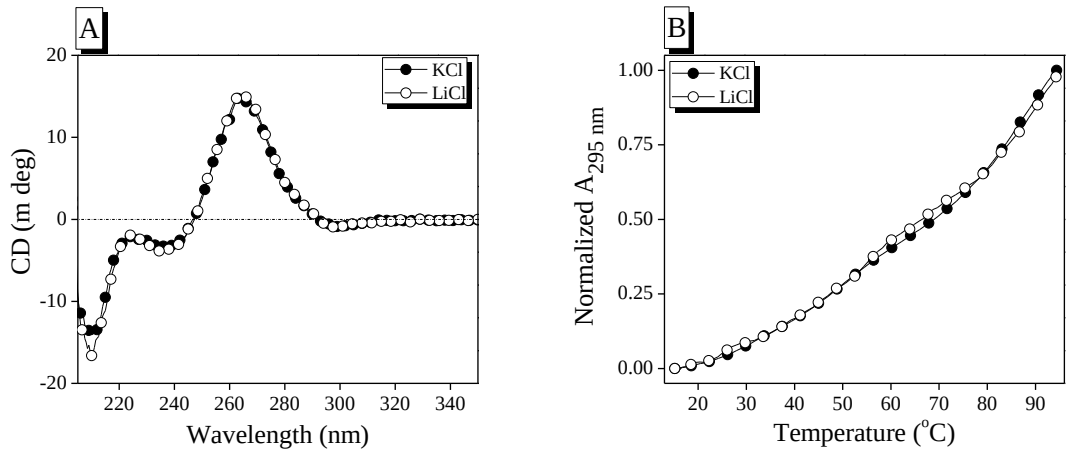


Figure 3.2. Biophysical characterization of 4G2 RNA quadruplex. CD spectra (**A**) and UV-melting curves (**B**) of 4G4 RNA target. All samples contained 2 µM RNA buffered in 100 mM KCl (closed circles) or 100 mM LiCl (open circles).

3.3.3 Biochemical Characterization of 4Gx RNAs

We next sought to examine the effects of the **4G2** and **4G4** RNA sequences on luciferase production from transcripts bearing the respective G-rich elements in their 5'-UTRs. Figure 3.3 shows translation efficiencies (presented as % RLU) for **4G2** and **4G4** reporter transcripts, relative to the control reporter devoid of a G-rich insert. (Data for **4G3** is also included for comparison.) We observe 20 % translation inhibition by **4G4**, an effect less than that observed with **4G3** (Figure 3.3). Further, translation inhibition by **4G2**—previously observed to be absent a stable quadruplex fold—was greatest among the three inserted G-rich motifs.

An explanation for this magnitude of inhibition by **4G2** RNA is unknown. It is possible that the UTR for this reporter possesses a stable alternate structure with greater repressive effects on

translation than **4G3** or **4G4**. Therefore, structure elucidation methods[23, 24] to determine the fold adopted by this UTR in the context of the mRNA transcript will prove useful in future studies. Importantly, however, all three 4Gx constructs studied here exhibit disparate translation efficiencies relative to the control, suggesting a possible—if incompletely understood—regulatory effect on translation by these G-rich elements.

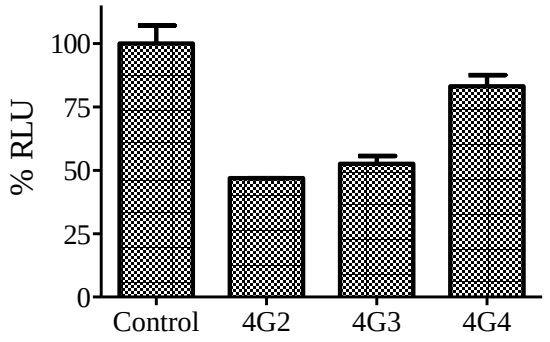


Figure 3.3. Biochemical Characterization of 4Gx RNAs. Luminescence intensity (expressed as % RLU) for 4Gx reporter and No G3 control transcripts.

3.3.4 Characterization of 4Gx-γPNA Hybrids

3.3.4.1 4G4 + γ4G4

The γPNA oligomer complementary to the 5'-end of the **4G4** QFS (**γ4G4**) was examined for its ability to form a stable hybrid with its complementary sequence, part of which would be sequestered by the stable quadruplex fold observed for **4G4**. Results from CD and UV-melting analyses on the **4G4-γ4G4** duplex are presented in Figure 3.4. While CD spectra alone are insufficient evidence of hybridization, as the peak positions for parallel RNA G-quadruplexes[25] and RNA-PNA hybrid duplexes[26] are similar, the spectrum presented in Figure 3.4A shows a minimum at 295 nm that is absent in that for **4G4** (Figure 3.1). This peak has been observed in

previous reports of RNA-PNA duplexes[26] and suggests hybridization between the two binding partners tested here.

The thermal stability for the hybrid could not be assessed, as melting curves generated at 275 nm yielded incomplete, non-cooperative melting transitions (Figure 3.4B). However, the fact that we fail to observe a complete transition at 95 °C suggests a high thermal stability for the hybrid. Importantly, the hypochromic transition recorded for **4G4** (Figure 3.1B) is abolished in the presence of an equimolar amount of **γ4G4** (Figure B2), suggesting that under these conditions, and for these concentrations of probe and target, the resulting hybrid duplex is more stable than the **4G4** quadruplex.

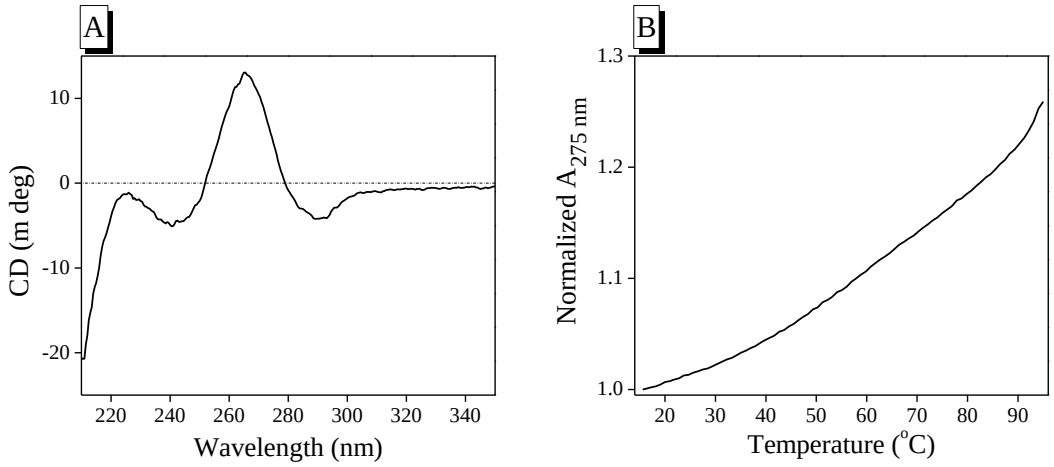


Figure 3.4. Biophysical Characterization of 4G4-γ4G4 Hybrid. CD spectrum (**A**) and UV-melting curve (**B**) of 4G4-γ4G4. All samples contained 2 µM RNA and 2 µM γPNA in 100 mM KCl.

SPR competition experiments[17] were employed to obtain the K_D value for the hybrid duplex. Figure 3.5 presents the binding isotherm constructed by titrating **γ4G4** with increasing concentrations of the **4G4** RNA. (For sensorgrams and calibration curve, see Figure B3.) We observed that the γPNA forms a relatively tight hybrid with the RNA target, with K_D = 0.94 nM.

Although the fit is non-ideal (Figure 3.5), this sub-nanomolar affinity is impressive, since the intrinsic structure of the RNA target poses a significant thermodynamic barrier to hybridization that has to be compensated by the free energy of the binding reaction. (Interestingly, the data fit better to a two-site model with two identical K_D values. However, we currently have no reasonable physical interpretation for that result.) It is likely that the high G-C content (75 %) of the probe imparts significant thermodynamic stability to the resulting duplex, thus compensating for the energy investment to unfold the quadruplex structure in the target. Additionally, quadruplex invasion by the probe results in a relatively long (20 nt) stretch of overhanging bases on the 3'-end of the RNA that could stabilize the duplex by stacking on the terminal hybrid base pair, as we have suggested in a previous report[27].

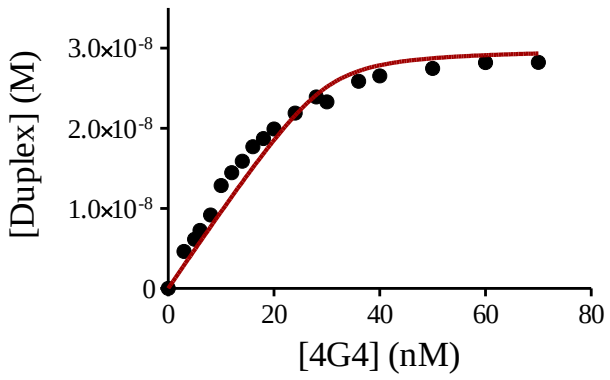


Figure 3.5. Binding Isotherm for 4G4-γ4G4 Hybrid.

Table 3.1. Equilibrium dissociation constants (K_D, nM) for γPNA-RNA hybrid duplexes formed with 4Gx RNA and truncated complementary RNA targets.

γPNA	4Gx RNA	4Gx-rComp
γ4G4	0.94 ± 0.33	3.02 ± 0.35
γ4G2	0.14 ± 0.03	0.19 ± 0.02

3.3.4.2 **4G2 + γ4G2**

Hybridization by **γ4G2** to its complementary sequence within the **4G2** QFS was expected to be unhindered, as no quadruplex folding was detected in the latter (Figure 3.2A). Indeed, the minimum at 295 nm in the CD spectrum for the hybrid confirmed duplex formation (Figure 3.6A), as was observed for the **4G4-γ4G4** reaction. Further, the resulting hybrid is very thermally stable, eliciting an incomplete transition even at 95 °C (Figure 3.6B).

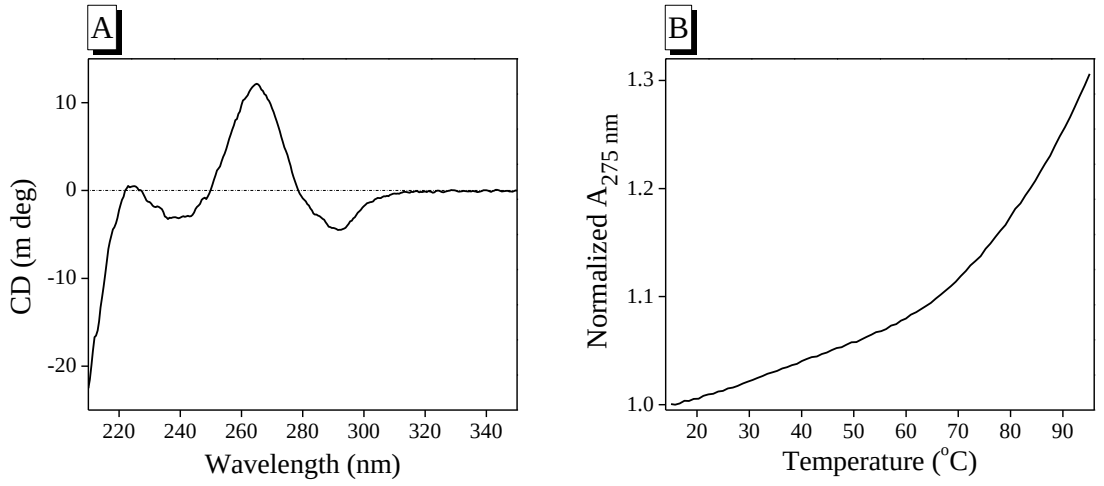


Figure 3.6. Biophysical Characterization of 4G2-γ4G2 Hybrid. CD spectrum (**A**) and UV-melting curve (**B**) of 4G2-γ4G2. All samples contained 2 μM RNA and 2 μM γPNA in 100 mM KCl.

SPR competition experiments[17] were used to construct the binding isotherm presented in Figure 3.7. (See Figure B4 for sensorgrams and calibration curve.) The data fit well to a one-site binding model, yielding a K_D = 140 pM (Figure 3.7 and Table 3.1). The higher affinity recorded for this probe, relative to **γ4G4**, is likely a consequence of the absence of a stable quadruplex fold in its RNA target. Importantly, the affinity recorded here is similar—in order of magnitude—to values reported earlier for hybridization of γPNAs to unstructured complementary RNA targets[14].

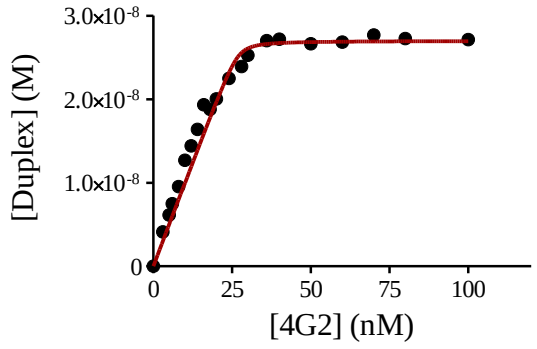


Figure 3.7. Binding Isotherm for the 4G2-γ4G2 Hybrid.

3.3.5 Characterization of rComp-γPNA Hybrids

3.3.5.1 4G4-rComp + γ4G4

To further understand hybridization of **γ4G4** to its complementary sequence within the QFS, a series of biophysical experiments were next performed on the hybrid duplex formed between the probe and a truncated RNA target (**4G4-rComp**) bearing only the nucleotides complementary to **γ4G4**. A CD spectrum of the hybrid possessed a minimum at 295 nm (Figure 3.8A), characteristic of antiparallel PNA-RNA duplexes.[26] The incomplete melting transition (Figure 3.8B)—although indicative of a high thermal stability for the hybrid, necessitated the application of SPR competition experiments to determine the corresponding K_D value.

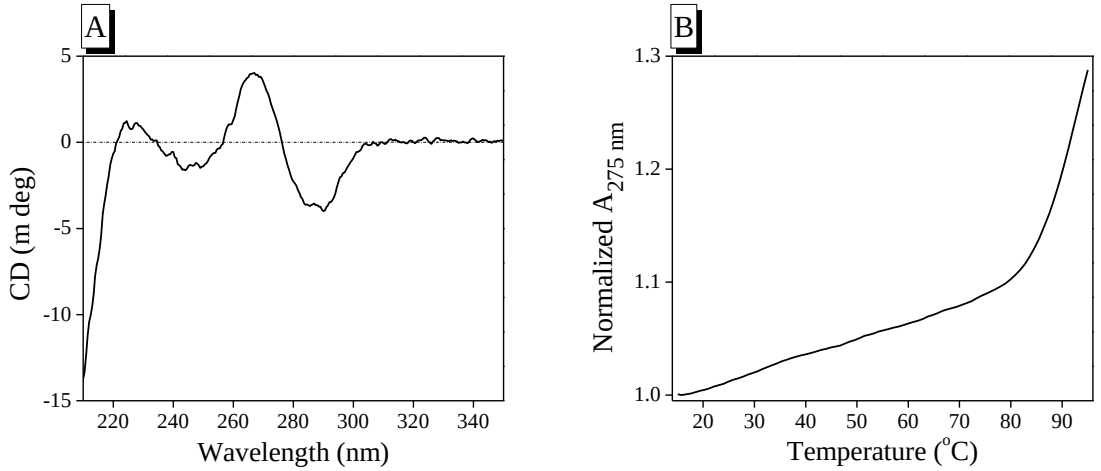


Figure 3.8. Biophysical Characterization of 4G4rComp-γ4G4 Hybrid. CD spectrum (**A**) and UV-melting curve (**B**) of 4G4rComp-γ4G4. All samples contained 2 µM RNA and 2 µM γPNA in 100 mM KCl.

The binding isotherm for the hybrid is presented in Figure 3.9. The data fit well to a one-site binding model, and the K_D (3.02 nM, Table 3.1) is indicative of a relatively tight complex. (SPR sensorgrams and calibration curve are presented in Figure B5.) Interestingly, the affinity of **γ4G4** for the truncated target (**4G4-rComp**) is ~ 3-fold lower than that for the **4G4** target (Table 3.1). We have not performed melting experiments to check for the existence of a stable structure in **4G4-rComp** that might also present a thermodynamic barrier to duplex formation with the probe.

However, such resistance—if it exists, is unlikely to be greater than that posed by the **4G4** target, since this sequence (**4G4-rComp**) possesses only one G_4 tract, and thus can only form a tetramolecular quadruplex consisting of four stacked G-tetrads (Chart 3.1). (Intermolecular quadruplex formation in the truncated target might also account for the 15-fold higher K_D recorded for the **γ4G4/4G4-rComp** hybrid, relative to the **γ4G2/4G2-rComp** hybrid, Table 3.1) Additionally, while it is likely that the overhanging bases in the longer target—and absent in **4G4-**

rComp, confer greater thermodynamic stability to the hybrid formed with the former, it is currently not clear whether such effects will account for a 200 % variation in the K_D values between the respective duplexes.

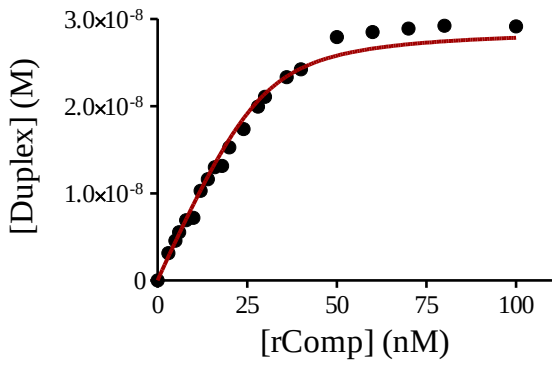


Figure 3.9. Binding Isotherm for the 4G4rComp-γ4G4 Hybrid.

3.3.5.2 4G2-rComp + γ4G2

Identical analyses were performed to understand hybridization of **γ4G2** to its complementary sequence. CD and UV-melting (Figure 3.10A and B, respectively) experiments confirmed formation of a stable hybrid between the probe and its truncated complementary sequence (**4G2-rComp**). As with **γ4G4**—the thermal stability of the hybrid, and consequent incomplete melting analysis, necessitated SPR competition experiments for affinity measurements. Figure 3.11 presents the binding isotherm constructed by titrating **γ4G2** with increasing concentrations of **4G2-rComp**. (SPR sensorgrams and calibration curve presented in Figure B6.)

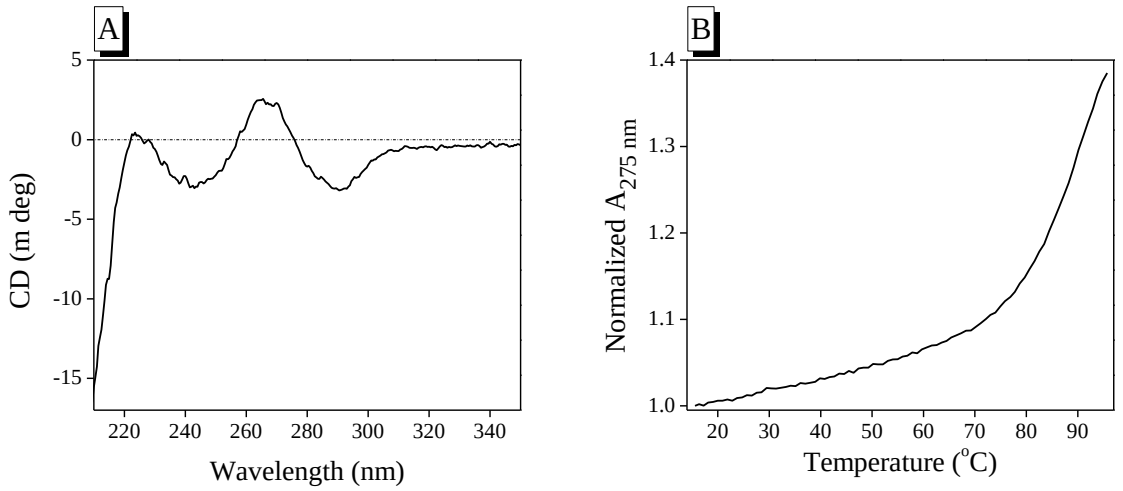


Figure 3.10. Biophysical Characterization of 4G2rComp-γ4G2 Hybrid. CD spectrum (**A**) and UV-melting curve (**B**) of 4G2rComp-γ4G2. All samples contained 2 µM RNA and 2 µM γPNA in 100 mM KCl.

We obtained a K_D value = 190 pM for the hybrid duplex (Table 3.1). Interestingly, this value is similar to that obtained with the full-length (**4G2**) target—suggesting that the thermodynamic penalty incurred by the probe is the same for hybridization to both the full-length and truncated targets. These data are consistent with results from CD and UV-melting experiments showing that **4G2** does not form a stable quadruplex.

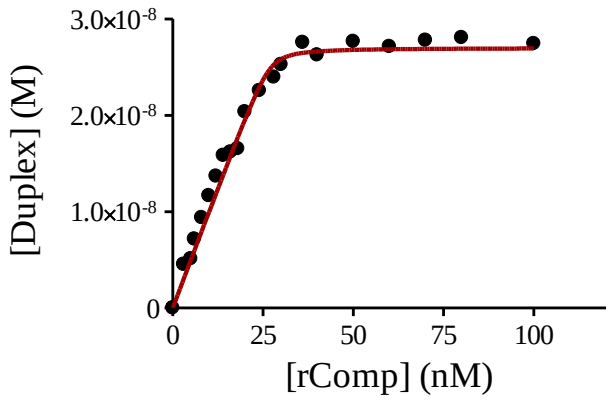


Figure 3.11. Binding Isotherm for the 4G2rComp-γ4G2 Hybrid.

3.3.6 SPR Direct-Binding Analyses on 4Gx Targets

To directly interrogate the interactions between the γPNA probes and their respective G-rich targets, we employed SPR direct-binding experiments wherein the RNA targets were non-covalently immobilized on the SPR chip (see Chapter 2). The sensorgrams are corrected for the immobilization levels of the RNA target—and, where a buffer change was made, care was taken to ensure identical immobilization levels for the RNA target under both buffer conditions.

3.3.6.1 **4G4 + γ4G4**

While the γPNA probe readily overcomes the kinetic barrier to hybridization posed by the stable **4G4** quadruplex (Figure 3.12A, black curve), destabilization of the target in Li^+ led to a strong (~ 6 fold) amplification in binding (Figure 3.12A, red curve), presumably due to the increased accessibility of the complementary sequence. We also observed faster release kinetics in the presence of Li^+ than in K^+. This result is surprising, since the latter buffer is expected to stabilize the QFS resulting from dissociation, potentially accelerating **γ4G4** release. Importantly, however, the acceleration in binding following target destabilization is an interesting demonstration of the inertia to binding introduced by a stable structure in the target strand.

3.3.6.2 **4G2 + γ4G2**

Figure 3.12B presents sensorgrams for **γ4G2** binding to- and dissociating from the immobilized **4G2** RNA. While we observe binding between the probe and target under both conditions, no cation-specific binding enhancements are observed in this interaction, unlike with **γ4G4**. This consistency in binding responses following ostensible target destabilization (Figure 3.12B) is

further evidence that **4G2** RNA does not fold into a stable quadruplex under these conditions. We currently have no explanation for why so little **γ4G2** bound to the RNA in either buffer, since all sensorgrams were corrected for the immobilization level of the RNA target.

3.3.7 4G3 + γ4G3

Results from identical analyses with **γ4G3** (**γ5'** in Chapter 2) were compared with those for the **γ4G4** and **γ4G2** probes. Quadruplex invasion was accomplished by this probe in K^+ (Figure 3.12C, black curve), an outcome that was amplified when the target was destabilized in Li^+ (Figure 3.12C, red curve). Interestingly the signal (binding) enhancement recorded upon quadruplex destabilization for this probe was less than that for **γ4G4** (6.7-fold for **γ4G4**; 3.6-fold for **γ4G3**). This result is not due to disparate binding responses to the stable quadruplexes, since both probes elicit identical binding profiles to their respective targets in KCl (Figure B7a). Rather, we observe that the binding responses in LiCl account for this difference (Figure B7b), suggesting that the aforementioned buffer has a greater 'opening' effect on the **4G4** than the **4G3** target. While the basis for this result is unknown, our results, taken together, show that the complementary γPNA probes can overcome the inherent structure of RNA quadruplexes and gain access to their binding sites.

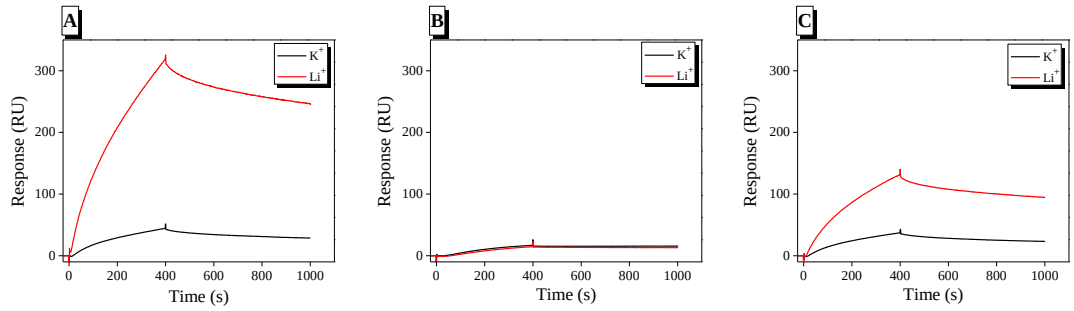


Figure 3.12. Sensorgrams for hybridization of 50 nM γPNAs to 4Gx RNA targets. γ4G4 + 4G4 (**A**) γ4G2 + 4G2 (**B**) γ4G3 + 4G3 (**C**). Running buffer contained 10 mM Tris-HCl (pH 7.4), 3 mM Na_2EDTA, and 100 mM KCl (black curve) or 100 mM LiCl (red curve). All binding experiments were performed at 25 °C. Sensorgrams are corrected for the immobilization level of each RNA.

3.3.8 Effects of γ4G4 and γ4G2 on Translation

We next examined the functional implications of γPNA binding in the context of a luciferase reporter assay. Luciferase mRNA transcripts bearing the G-rich inserts in their 5'-UTRs were titrated with increasing concentrations of the respective complementary γPNA probes. Control experiments were also performed using luciferase transcript devoid of the 5'-UTR G-rich inserts.

3.3.8.1 γ4G4

We observed dose-dependent inhibition of translation by this probe, with IC_{50} = 15 nM (Figure 3.13A and Table 1). Additionally, no activity was detected against the control reporter, even at the highest dose of the probe (Figure 3.13A, open squares), suggesting that translation inhibition results from a specific binding event. The potency recorded for this probe is impressive, since recognition of its binding site is likely accompanied by unfolding of a stable four-tetrad quadruplex. Further, the inertness of the probe to the control reporter is important, since the C4

tracts within the probe could, in principle, bind to other G-rich, non-QFS domains present in the coding sequence.

3.3.8.2 γ4G2

This probe also displayed potent inhibition of translation (IC_{50} = 20 nM, Table 3.1) in the presence of the target reporter, with only modest off-target inhibition at the highest concentration tested, when applied against the control transcript (Figure 3.13B). Although concerning, the off-target effect is de-emphasized by the near-absolute inhibition of the target reporter with 75 nM **γ4G2**—a concentration at which we see minimal effects on the control. Taken together, these results establish complementary γPNA oligomers as effective probes for invasion of stable G-quadruplexes (where present in 5'-UTRs), with resultant inhibitory effects on translation.

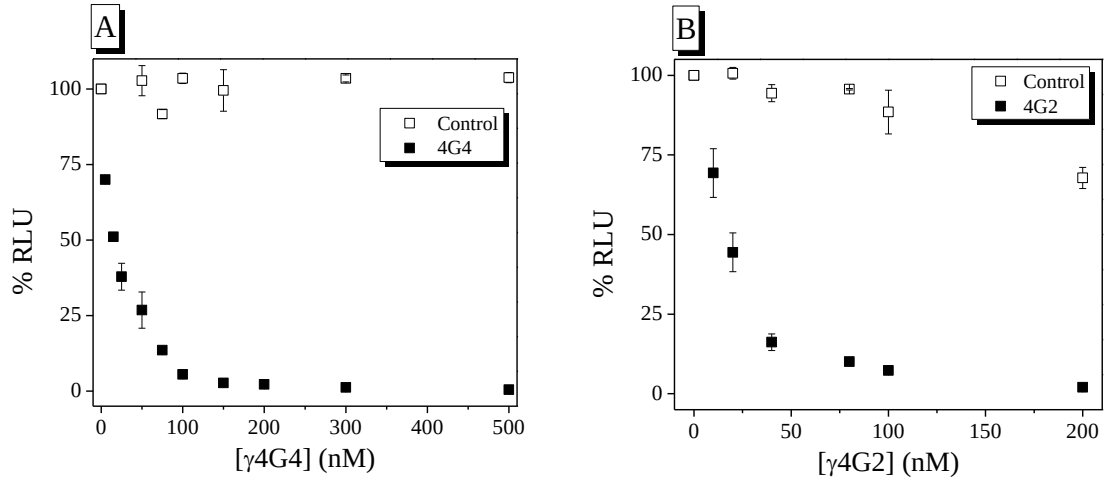


Figure 3.13. Effects of γPNAs on mRNA translation. Dose response curves for γ4G4 (**A**) or γ4G2 (**B**) oligomers against 4Gx target (filled squares) or control (open squares) reporters. [mRNA] = 10 nM, and RNA-γPNA samples were incubated for one hr at 37 °C prior to adding to lysate. Data are normalized to 100% for the samples that lacked any γPNA probe.

3.4 DISCUSSION

The results presented in this chapter demonstrate that γPNA oligomers complementary to the 5'-ends of G-rich RNA targets can hybridize with high affinities to their respective target sequences. Quadruplex folding—where present, is sufficient to decelerate and weaken, but insufficient to preclude, hybridization between the probes and targets. The resulting hybrids are exceptionally stable, with the ΔG for complexation compensating for the energy investment to unfold any existing quadruplex structure. Importantly, the γPNA oligomers elicit potent inhibition of translation mediated by a specific binding event.

Extending the recognition to **4G4** and **4G2** RNA targets is an important advance in the application of complementary γPNA oligomers for quadruplex invasion, since these targets display disparate biophysical and biochemical properties relative to the **4G3** RNA target studied in Chapter 2. **4G4** forms a stable, intramolecularly folded quadruplex that exerts a moderately inhibitory effect on translation. Conversely, we failed to observe characteristic quadruplex signatures for **4G2** RNA, even though this sequence, when inserted into the 5'-UTR of the reporter transcript—produced strong repressive effects. It is possible that quadruplex formation by **4G4** is disfavored by competing secondary structures that are less obstinate to 43S scanning in the 5'-UTR. Additionally, a putative **4G2** quadruplex could be reinforced by molecular crowding in the cell lysates or adjacent structures in the 5'-UTR which coalesce to pose a strong impediment to scanning. Together, our data show that these targets have differential effects on translation, while

also accentuating the need to apply structure elucidation methods, such as SHAPE[23] and/or foot printing[24] analyses, to understand/verify quadruplex folding in future reports.

Our results on the basal levels of translation for the 4Gx reporters contrast with a previous report from Kumari *et al* showing systematic enhancements in the translation efficiency for the quadruplex-laden transcripts with progressive destabilization of the quadruplex structures[18]. Importantly, however, the aforementioned authors presented the 4Gx sequences in the context of the *natural* (236 nt) *NRAS* 5'-UTR[18], whereas our results were obtained for 4Gx sequences inserted into a *synthetic* (~60 nt) 5'-UTR. It is possible that the different sequence context applied here induces alternate folding patterns for the inserted G-rich motifs or reinforces the existing quadruplex folds to different extents, resulting in effects on translation that cannot be explained by biophysical data obtained for the isolated sequences in buffer solution.

Indeed, the importance of sequence context in elucidating the putative functional effects of G-quadruplex structures was further demonstrated by a report from Maiti and coworkers, wherein a QFS within the 5'-UTR of TGFβ2 mRNA *repressed* translation when placed in the context of an artificial 5'-UTR, but *enhanced* translation when in the context of the natural TGFβ2 5'-UTR[28]. These results emphasize the importance of comprehensive structural characterization of the G-rich motifs within the natural sequence context prior to conclusive statements on their biochemical or biological effects.

Data presented in Chapter 2 highlighted the relevance of overhang recognition to translation inhibition by the most potent probe. The extent to which this contribution affects the activities of **γ4G4** and **γ4G2** is likely determined, in part, by the structural context of the QFS in the 5'-UTR, i.e. structures adjacent to the quadruplex might occlude overhanging bases to different extents. In general, however, future efforts should seek to correlate structural insight on the entire quadruplex-

laden UTR with the activities of complementary ligands, especially those directed to flanking bases.

In conclusion, these results demonstrate the potential for applying complementary γPNA oligomers as reagents for translation inhibition mediated by quadruplex invasion. Recent advances in the delivery of γPNA oligomers[29, 30], coupled with the rapid pace of discovery of QFS motifs and/or their functions, make us confident that this class of reagents will find broad use in the biomedical research community.

3.5 REFERENCES

[1] Haeusler, A. R., Donnelly, C. J., Periz, G., Simko, E. A., Shaw, P. G., Kim, M. S., Maragakis, N. J., Troncoso, J. C., Pandey, A., Sattler, R., Rothstein, J. D., and Wang, J. (2014) C9orf72 nucleotide repeat structures initiate molecular cascades of disease, *Nature 507*, 195-200.

[2] Penagarikano, O., Mulle, J. G., and Warren, S. T. (2007) The pathophysiology of fragile x syndrome, *Annu. Rev. Genomics Hum. Genet. 8*, 109-129.

[3] Fratta, P., Mizielinska, S., Nicoll, A. J., Zloh, M., Fisher, E. M., Parkinson, G., and Isaacs, A. M. (2012) C9orf72 hexanucleotide repeat associated with amyotrophic lateral sclerosis and frontotemporal dementia forms RNA G-quadruplexes, *Sci. Rep. 2*, 1016.

[4] Zhou, B., Liu, C., Geng, Y., and Zhu, G. (2015) Topology of a G-quadruplex DNA formed by C9orf72 hexanucleotide repeats associated with ALS and FTD, *Sci. Rep. 5*, 16673.

[5] Reddy, K., Zamiri, B., Stanley, S. Y., Macgregor, R. B., Jr., and Pearson, C. E. (2013) The disease-associated r(GGGGCC)n repeat from the C9orf72 gene forms tract length-dependent uni- and multimolecular RNA G-quadruplex structures, *J. Biol. Chem. 288*, 9860-9866.

[6] Fry, M., and Loeb, L. A. (1994) The fragile X syndrome d(CGG)n nucleotide repeats form a stable tetrahelical structure, *Proc. Natl. Acad. Sci. U. S. A. 91*, 4950-4954.

[7] Kettani, A., Kumar, R. A., and Patel, D. J. (1995) Solution structure of a DNA quadruplex containing the fragile X syndrome triplet repeat, *J. Mol. Biol. 254*, 638-656.

[8] DeJesus-Hernandez, M., Mackenzie, I. R., Boeve, B. F., Boxer, A. L., Baker, M., Rutherford, N. J., Nicholson, A. M., Finch, N. A., Flynn, H., Adamson, J., Kouri, N., Wojtas, A., Sengdy, P., Hsiung, G. Y., Karydas, A., Seeley, W. W., Josephs, K. A., Coppola, G., Geschwind, D. H., Wszolek, Z. K., Feldman, H., Knopman, D. S., Petersen, R. C., Miller, B. L., Dickson, D. W., Boylan, K. B., Graff-Radford, N. R., and Rademakers, R. (2011) Expanded GGGGCC hexanucleotide repeat in noncoding region of C9ORF72 causes chromosome 9p-linked FTD and ALS, *Neuron 72*, 245-256.

[9] Wells, R. D. (2009) Mutation spectra in fragile X syndrome induced by deletions of CGG*CCG repeats, *J. Biol. Chem. 284*, 7407-7411.

[10] Mulders, S. A., van den Broek, W. J., Wheeler, T. M., Croes, H. J., van Kuik-Romeijn, P., de Kimpe, S. J., Furling, D., Platenburg, G. J., Gourdon, G., Thornton, C. A., Wieringa, B., and Wansink, D. G. (2009) Triplet-repeat oligonucleotide-mediated reversal of RNA toxicity in myotonic dystrophy, *Proc. Natl. Acad. Sci. U. S. A. 106*, 13915-13920.

[11] Wheeler, T. M., Sobczak, K., Lueck, J. D., Osborne, R. J., Lin, X., Dirksen, R. T., and Thornton, C. A. (2009) Reversal of RNA Dominance by Displacement of Protein Sequestered on Triplet Repeat RNA, *Science 325*, 336-339.

[12] Donnelly, C. J., Zhang, P. W., Pham, J. T., Haeusler, A. R., Mistry, N. A., Vidensky, S., Daley, E. L., Poth, E. M., Hoover, B., Fines, D. M., Maragakis, N., Tienari, P. J., Petrucelli, L., Traynor, B. J., Wang, J., Rigo, F., Bennett, C. F., Blackshaw, S., Sattler, R., and Rothstein, J. D. (2013) RNA toxicity from the ALS/FTD C9ORF72 expansion is mitigated by antisense intervention, *Neuron 80*, 415-428.

[13] Murat, P., Zhong, J., Lekieffre, L., Cowieson, N. P., Clancy, J. L., Preiss, T., Balasubramanian, S., Khanna, R., and Tellam, J. (2014) G-quadruplexes regulate Epstein-Barr virus-encoded nuclear antigen 1 mRNA translation, *Nat. Chem. Biol. 10*, 358-364.

[14] Sahu, B., Sacui, I., Rapireddy, S., Zanotti, K. J., Bahal, R., Armitage, B. A., and Ly, D. H. (2011) Synthesis and characterization of conformationally preorganized, (R)-diethylene glycol-containing gamma-peptide nucleic acids with superior hybridization properties and water solubility, *J. Org. Chem. 76*, 5614-5627.
[15] Dragulescu-Andrasi, A., Rapireddy, S., Frezza, B. M., Gayathri, C., Gil, R. R., and Ly, D. H. (2006) A Simple γ-Backbone Modification Preorganizes Peptide Nucleic Acid into a Helical Structure, *J. Am. Chem. Soc. 2006*, 10258–10267.
[16] Roy, S., Zanotti, K. J., Murphy, C. T., Tanious, F. A., Wilson, W. D., Ly, D. H., and Armitage, B. A. (2011) Kinetic discrimination in recognition of DNA quadruplex targets by guanine-rich heteroquadruplex-forming PNA probes, *Chem. Commun. 47*, 8524-8526.
[17] Roy, S., Tanious, F. A., Wilson, W. D., Ly, D. H., and Armitage, B. A. (2007) High-affinity homologous peptide nucleic acid probes for targeting a quadruplex-forming sequence from a MYC promoter element, *Biochemistry 46*, 10433–10443.
[18] Kumari, S., Bugaut, A., and Balasubramanian, S. (2008) Position and stability are determining factors for translation repression by an RNA G-quadruplex-forming sequence within the 5' UTR of the NRAS proto-oncogene, *Biochemistry 47*, 12664-12669.
[19] Morris, M. J., Negishi, Y., Pazsint, C., Schonhoft, J. D., and Basu, S. (2010) An RNA G-quadruplex is essential for cap-independent translation initiation in human VEGF IRES, *J. Am. Chem. Soc. 132*, 17831-17839.
[20] Vorlickova, M., Kejnovska, I., Sagi, J., Renciuk, D., Bednarova, K., Motlova, J., and Kypr, J. (2012) Circular dichroism and guanine quadruplexes, *Methods 57*, 64-75.
[21] Williamson, J. R. (1993) Guanine quartets, *Curr. Opin. Struct. Biol. 3*, 357-362.
[22] Mergny, J. L., Phan, A. T., and Lacroix, L. (1998) Following G-quartet formation by UV-spectroscopy, *FEBS Lett. 435*, 74-78.
[23] Wilkinson, K. A., Merino, E. J., and Weeks, K. M. (2006) Selective 2'-hydroxyl acylation analyzed by primer extension (SHAPE): quantitative RNA structure analysis at single nucleotide resolution, *Nat. Protoc. 1*, 1610-1616.
[24] Tijerina, P., Mohr, S., and Russell, R. (2007) DMS footprinting of structured RNAs and RNA-protein complexes, *Nat. Protoc. 2*, 2608-2623.
[25] Zhang, A. Y., Bugaut, A., and Balasubramanian, S. (2011) A sequence-independent analysis of the loop length dependence of intramolecular RNA G-quadruplex stability and topology, *Biochemistry 50*, 7251-7258.
[26] Egholm, M., Buchardt, O., Christensen, L., Behrens, C., Freier, S. M., Driver, D. A., Berg, R. H., Kim, S. K., Norden, B., and Nielsen, P. E. (1993) PNA hybridizes to complementary oligonucleotides obeying the Watson Crick hydrogen-bonding rules, *Nature 365*, 566-568.
[27] Datta, B., and Armitage, B. A. (2001) Hybridization of PNA to Structured DNA Targets: Quadruplex Invasion and the Overhang Effect, *J. Am. Chem. Soc. 123*, 9612–9619.
[28] Agarwala, P., Pandey, S., Mapa, K., and Maiti, S. (2013) The G-quadruplex augments translation in the 5' untranslated region of transforming growth factor beta2, *Biochemistry 52*, 1528-1538.
[29] Bahal, R., McNeer, N. A., Ly, D. H., Saltzman, W. M., and Glazer, P. M. (2014) Nanoparticle for delivery of antisense γPNA oligomers targeting CCR5, *Artificial DNA: PNA & XNA 4*, 49-57.

[30] Sahu, B., Chenna, V., Lathrop, K. L., Thomas, S. M., Zon, G., Livak, K. J., and Ly, D. H. (2009) Synthesis of Conformationally Preorganized and Cell-Permeable Guanidine-Based γ-Peptide Nucleic Acids (γGPNAs), *J. Org. Chem. 74*, 1509–1516.

CHAPTER 4

G-QUADRUPLEX INVASION AND TRANSLATION INHIBITION BY HOMOLOGOUS PNA/γPNA OLIGOMERS

4.1 INTRODUCTION

The rapid pace of discovery of novel regulatory roles for guanine (G)-quadruplexes incentivizes the development of ligands capable of recognizing these non-canonical structures, with resultant effects on their biochemical and/or biological functions (reviewed in Chapter 1). Such reagents would find use either as therapeutics—where an unambiguous nexus exists between quadruplex formation and disease pathogenesis[1], or as discovery probes for elucidating the mechanistic contributions of quadruplex structures to biological processes. Towards this end, sequence-recognizing agents that anneal to the primary structure, rather than shape, of the quadruplex are being developed, in part because of the potential for specific binding, and the requirement for little upfront knowledge of the quadruplex structure (beyond its sequence) ahead of ligand design.

Although oligomeric complementary ligands[2-11] that anneal to the quadruplex-forming sequence (QFS) by Watson-Crick hydrogen bonding[12] represent the most obvious members of this reagent class—our group[9, 13-19], and others[20-23], have demonstrated the potential for QFS recognition by peptide nucleic acid (PNA) oligomers presenting homologous (G-rich) nucleobases to the QFS target. Our first demonstration of this alternative recognition mode utilized a PNA molecule homologous to the *Oxytricha nova* telomeric repeat to hybridize a synthetic DNA strand based on the repeat motif[13]. Hybridization occurred via a strong exergonic reaction, resulting in a very stable PNA-DNA hetero-quadruplex that aligned the PNA N-terminus with the DNA 5'-end[13]. Further, the resulting complex displayed cation sensitivity characteristic of quadruplex structures, even though the impacts of varying cation concentrations were less pronounced than in an analogous DNA homoquadruplex[13].

We extended this recognition mode to an RNA quadruplex-forming target, whereby a PNA oligomer homologous to a region of the RNA sequestered in the quadruplex fold was shown to invade the target—resulting in a more stable RNA-PNA hybrid quadruplex[14, 15]. The binding reaction resulted in a ternary complex consisting of two PNA molecules annealed to the same RNA strand[14, 15]. This result led to the insight that *true homology*—where the PNA nucleobases are entirely homologous to those of the target—was not requisite for binding. Rather, *'G-homology'*—where G-nucleobases are presented at the same intervals in the probe as they occur in the target—was sufficient to mediate heteroquadruplex formation[14, 15].

The same principle was demonstrated in a subsequent report from our group, wherein we showed that a homologous PNA probe formed a ternary complex with a DNA QFS, despite the presence of only one perfectly homologous sequence on the target[16]. In fact, the binding reaction resulted in a very tight complex (K_D = 5 nM)—an impressive feat, given the relatively small size (8mer) of the probe, and the extrusion of non-bonded intervening residues[16].

This degeneracy in recognition—where G-homology instead of perfect homology would suffice for binding—proved useful in subsequent attempts to tune the selectivity of homologous PNA oligomers for quadruplex-forming targets instead of potential C-rich, duplex-forming off-targets. In this context, we showed that the replacement of intervening residues in a homologous probe with abasic residues was sufficient to discourage hybridization to a C-rich strand[17]. The redesigned probe, which remained G-homologous to a DNA QFS, formed stable hybrid quadruplexes with the DNA target[17]. Importantly, the dissociation constant for the hybrid was unperturbed by the modification in the probe, demonstrating the potential for the heteroquadruplex to tolerate considerable modifications in the PNA component[17].

Further demonstration of this tolerance for probe modification was provided when we showed that left-handed preorganization of the PNA—afforded by incorporation of D-amino acid-derived G monomers—did not preclude quadruplex invasion and formation of stable hybrids[17]. Importantly, this modification significantly mitigated hybridization to a C-rich sequence[17], thus priming this probe as a tool specific for quadruplex recognition in intact cells and/or cell lysates, without collateral binding to C-rich off-targets.

The data presented in this chapter summarize our efforts to utilize homologous PNA oligomers as reagents to modulate translation of a reporter transcript bearing a QFS element in its 5'-UTR. We demonstrate that the PNA probes can form stable hybrids with the QFS target, resulting in hybrids that are more stable than the quadruplex structure inherent in the latter. The probes elicit non-specific inhibition in the context of an *in vitro* translation assay—an effect that persists despite incorporation of γ-modifications[24, 25] to induce left-handed helical pre-organization, but is remediated by DNA strands that compete with the mRNA transcript for probe binding. Although we cannot currently define translation inhibition as a direct result of a specific recognition event between the homologous probes and quadruplex target in this context, these results provide insight on possible off-target binding sites that mediate the non-specific repressive effects of the PNAs.

4.2 MATERIALS AND METHODS

4.2.1 PNA/γPNA/DNA/RNA

All t-boc-protected PNA monomers were purchased from PolyOrg, Inc. (Leominster, MA). The PNA/γPNA oligomers (Chart 1) used for all experiments were synthesized using standard solid-phase peptide synthesis[26] protocols and included C-terminal L-lysine residues. Pure oligomers were isolated after purification using reverse-phase HPLC (Waters 2695 Separations Module) and confirmed by MALDI-TOF mass spectrometry (Applied Biosystems, Voyager DE sSTR) using α-cyano-4-hydroxycinnamic acid as the matrix. The γPNA oligomer (sequence below) used as the RNA capture strand in SPR experiments was purchased from PNA Innovations Inc. (www.pnainnovations.com) and was functionalized with a biotin on the ε-amino group of the C-terminal lysine residue. RNA and DNA oligomers used were obtained from Integrated DNA Technologies (www.idtdna.com). Sequences of the DNA and RNA oligonucleotides used in SPR direct-binding experiments are given below. (RNA/DNA sequences written 5'-3' and γPNA sequence written C-N.)

	Sequence
4G3 RNA (for SPR)	AGACCCAAGCACUAUAAGCUAGCGGGAGGGCGGGUCUGGGCGAUCC
C-rich DNA (for SPR)	biotin-TTTTTCCCACCCCTCCCACCCCT
RNA capture strand	biotin-TCTGGGTTCGTGATA (γPNA)

4.2.2 UV Melting Experiments

Thermal melting experiments were performed on a Varian Cary 300 spectrophotometer equipped with a thermoelectrically-controlled multicell holder. Samples were prepared in a buffer containing 10 mM Tris-HCl (pH 7.4), 0.1 mM Na/Li_2EDTA, and different concentrations of KCl/LiCl. The solutions were incubated at 95 °C for 5 min and cooled to 15 °C at a rate of 1 °C

/min. Subsequently, the annealed samples were incubated at 15 °C for 5 min, and a heating ramp was applied at 1 °C/min up to 95 °C. Melting curves were generated by monitoring 295 nm every 0.5 °C. Where possible, the melting temperature (T_m) was determined from a first-derivative plot of the respective melting curve. Each reported melting temperature is the average of three independent experiments.

4.2.3 Circular Dichroism (CD) Spectropolarimetry

CD spectra were obtained on a Jasco J-715 circular dichroism spectropolarimeter equipped with a water-circulating temperature controller. Samples were prepared in a buffer containing 10 mM Tris-HCl (pH 7.4), 0.1 mM Na/Li_2EDTA, and different concentrations of KCl/LiCl. The samples were incubated at 95 °C for 5 min and cooled slowly to room temperature. All spectra were collected at 37 °C. Each spectrum was collected after 6 different scans (200 – 360 nm) at a scan rate of 100 nm/min and baseline corrected. The CD spectra for 3 different samples were collected and averaged.

4.2.4 Surface Plasmon Resonance (SPR) Experiments

All SPR experiments were performed on a Biacore T100 instrument equipped with a four-channel CM5 sensor chip (GE Health Care). The sensor chip was coated with a carboxylmethyl dextran matrix that allows further functionalization with streptavidin via a standard NHS-EDC coupling procedure. Covalent immobilization of streptavidin was continued until 7000 response units (RU) of the protein were captured on each of the four channels (flow cells). The final step of the sensor design involved non-covalent immobilization of a biotinylated capture strand on the streptavidin-modified surface.

For direct-binding experiments with immobilized DNA, a low-density chip (~100 RU) was prepared to minimize mass transfer effects. The sensor surface was primed with a buffer injection [100 mM KCl/LiCl, 10 mM Tris-HCl (pH 7.4), 3 mM Na/Li_2EDTA, and 0.005 % v/v P-20 surfactant] prior to introduction of a fixed concentration (25 nM) of the different PNA/γPNA probes. Sample injection was continued for 420 s (flow rate = 50 µL/min), after which the buffer injection was restarted and continued for 600 s to monitor dissociation.

To perform similar experiments on an RNA target, we immobilized a biotinylated 15mer γPNA oligomer designed to hybridize to a complementary sequence appended to the 5'-end of the RNA QFS. The active flow cell was modified with a high density (1000 RU) of the biotinylated γPNA capture strand. Immobilization of the target RNA was then performed by injecting a 100 nM solution (pre-annealed for 2 h) over the active flow cell for 400 s (flow rate = 50 µL/min) until a medium density (300 RU) of the RNA accumulated on the surface. A short dissociation phase (20 s) was then introduced to remove any unhybridized or loosely-bound RNA from the surface.

To study hybridization to the immobilized RNA, 50 nM of each probe was injected for 400 s (flow rate = 30 µL/min) over both the active and reference flow cells, after which the buffer injection was restarted and continued for 600 s to monitor dissociation of the hybrids. The sensor surface was regenerated with a solution of 10 mM NaOH and 1 M NaCl (2x for 30 s), and another buffer injection (120 s) was introduced to wash off residual regeneration solution. The sensor surface was replenished with fresh RNA, and subsequent binding studies were continued by simple iterations of the aforementioned steps. In both experiments (DNA/RNA direct-binding), the initial association rates were calculated based on the times required to reach 5 response units of bound PNA/γPNA, similar to a procedure utilized by us in an earlier report[18].

4.2.5 Template Production and *In Vitro* Transcription

The DNA template containing the target sequence upstream of the firefly luciferase gene (4G3 reporter) was amplified by Polymerase Chain Reaction (PCR) using Taq DNA polymerase (New England Biolabs) and two sequence specific primers: forward primer (5'-AATACGCAAACCGCCTCTC-3') and reverse primer (5'-GGTGATGTCGGCGATATAGG-3'). A similar procedure was also used to obtain a control reporter (No G3) without the target sequence upstream of the luciferase coding sequence. The resulting DNA fragments were purified using a GeneJET PCR purification kit (ThermoFisher) according to the manufacturer's instructions. The integrity and length of each PCR product were verified using a 1 % agarose gel and electrophoresis.

The purified DNA templates were transcribed *in vitro* using a cocktail consisting of T7 RNA polymerase (100 units) and a mixture of the ribonucleotide triphosphates (500 µM of each NTP) in a total reaction volume of 100 µL at 37 ^{o}C for 2 h. The resulting mRNA transcripts were purified using the GeneJET RNA clean-up Micro Kit (Thermo Scientific) according to the manufacturer's instructions. The final mRNA concentrations were estimated by UV spectroscopy. The extinction coefficient of each mRNA transcript was assumed to be the sum of the molar absorptivities of each nucleobase, and the recorded absorbance (at 260 nm) was assumed to be unaffected by any (undetermined) secondary folding of the transcripts. The nucleotide sequences for the UTRs in both the experimental (4G3) and control transcripts are presented below, with the quadruplex-forming fragment, where present, highlighted in bold and upper-case.

4G3: gagacccaagcuuucagauccgcuagc**GGGAGGGCGGGUCUGGG**cgauccagccaccaug

No G3: gaucuaauaucuacuuaagaacacaaaacucgagaaccaug

4.2.6 Cell-Free Translation

A translation reaction was typically performed by incubating the purified transcripts in a mixture containing nuclease-treated rabbit reticulocyte lysate (70 % v/v) (Promega), 10 µM amino acid mixtures minus leucine, 10 µM amino acid mixtures minus methionine and 20 units RNasin ribonuclease inhibitor (50 µL total reaction volume) at 30 ^{o}C for 1.5 h. Experiments to examine the effect of each γPNA on luciferase production were performed by pre-incubating the mRNA with requisite amino acids and increasing concentrations of the γPNA in a buffer containing 79 mM KCl and 7.9 mM Tris-HCl (pH = 7.4) to a total volume of 19.5 µL at 37 ^{o}C for 1 h. Each subsequent translation reaction was started by adding 30.5 µL of the rabbit reticulocyte lysate to the pre-incubated mixture, and the final mixture was incubated at 30 ^{o}C for 1.5 h. Finally, we performed a luciferase assay by incubating 10 µL of the translation products with 50 µL of a reagent cocktail ($_{D}$-luciferin, Mg^{2+}, and ATP) (Promega) to estimate the relative amounts of enzyme produced from the mRNA reporter. Luciferase activity was estimated as relative light units (RLU) on a Tecan Infinite M1000 Plate Spectrometer.

4.3 RESULTS

4.3.1 Target Selection and Probe Design

The target (called **4G3**) reported in this chapter is an RNA sequence consisting of four G_3 tracts, interspersed with 1-, 1-, and 3-nt non-G residues, respectively (Chart 4.1). We previously showed that the sequence adopts an intramolecular quadruplex fold possibly consisting of three stacked G-tetrads—with the intervening residues extruded to the loops (Chapter 2) [8]. Finally, the quadruplex structure exerts a repressive effect on translation when inserted into the 5'-untranslated region (UTR) of a luciferase reporter transcript (Chapter 2) [8].

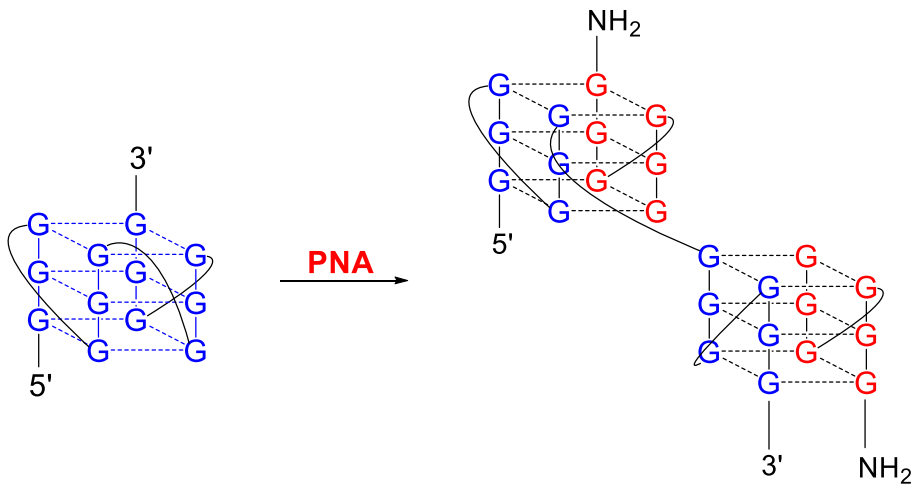
Chart 4.1: Sequences of PNA/γPNA and RNA oligomers used in this study. (RNA sequence written 5'-3', and PNA/γPNA sequences written N-C. Underlined Gs in RNA are predicted to participate in G tetrads. γ-modifications in PNA are presented in uppercase.)

Oligomer	Sequence
4G3 RNA	GCUAGCGGGAGGGCGGGUCUGGGCGAUCC
$\mathbf{P_{AG}}$	ggg(ag)ggg
$\mathbf{P_{eg2}}$	ggg(eg)$_2$ggg
^{3,6}**D-**$\mathbf{P_{eg2}}$	ggG(eg)$_2$ggG
^{4,5,6}**D-**$\mathbf{P_{eg2}}$	ggg(eg)$_2$GGG

Two homologous probes ($\mathbf{P_{AG}}$ and $\mathbf{P_{eg2}}$, Chart 4.1) were initially selected for quadruplex targeting based on previous reports[16, 17] demonstrating their capacities to invade a stable intramolecular DNA quadruplex and form hetero-quadruplexes. ($\mathbf{P_{AG}}$ was originally called $\mathbf{P_{myc}}$ in previous reports[16, 17].) Although originally designed to bind a different target, we show that these probes can invade the **4G3** target and form stable hybrid quadruplexes (Scheme 4.1). We also tested ^{3,6}**D-**$\mathbf{P_{eg2}}$ and ^{4,5,6}**D-**$\mathbf{P_{eg2}}$—two derivatives of $\mathbf{P_{eg2}}$ that bore γ-hydroxymethyl modifications[25] introduced at strategic positions (described later) to induce left-handed helical preorganization in the probe

(Chart 4.1). The effects of these modifications on the kinetics of invasion and potency/specificity of translation inhibition is also reported below.

Scheme 4.1



Invasion of folded 4G3 RNA quadruplex by homologous PNA oligomers

4.3.2 Characterization of RNA-PNA Hybrids

To investigate the capacity of $\mathbf{P_{AG}}$ and $\mathbf{P_{eg2}}$ to invade the stable quadruplex structure inherent in **4G3**, UV-melting and CD analyses (Figure 4.1) were performed on the RNA target in the presence and absence of the respective probes. Our results show that the hypochromic transition recorded for melting of the **4G3** quadruplex target is shifted towards higher temperatures in the presence of 2 equivalents of either homologous probe (Figure 4.1A). (The choice of this stoichiometry is based on previous reports from our group showing that a homologous PNA probe similar to those reported here forms a ternary complex with its quadruplex-forming RNA target[14, 15].) This shift in the position of the hypochromic transition has previously been ascribed to the melting of a more stable hybrid quadruplex formed between the probe and target, relative to the starting RNA intramolecular quadruplex[14, 15].

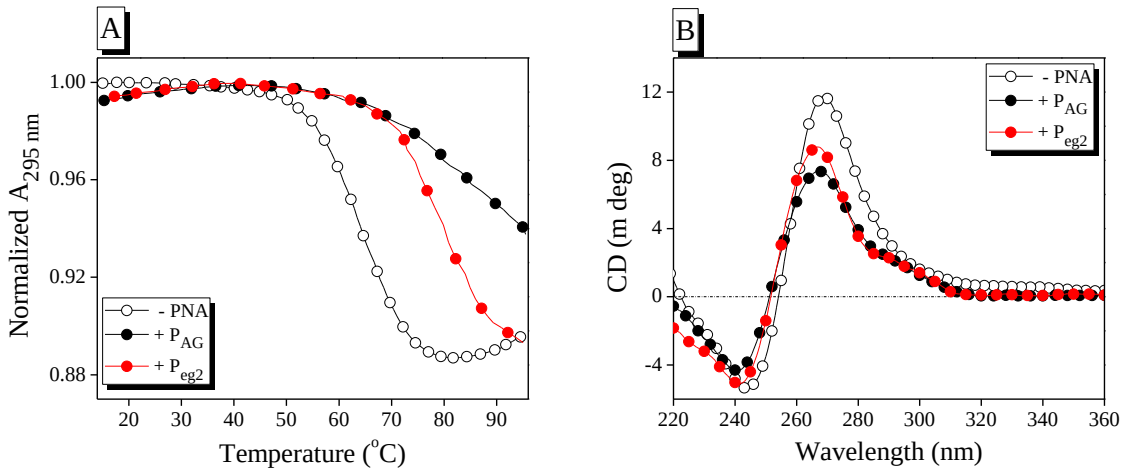


Figure 4.1: Biophysical characterization of RNA-PNA Hybrid quadruplexes. UV-melting (**A**) and CD spectra (**B**) of 2.5 µM 4G3 RNA target in the absence (open circles) or presence of 5 µM P_{AG} (black filled circles) or P_{eg2} (red filled circles). All samples contained 1 mM KCl, 10 mM Tris-HCl (pH 7.4), and 0.1 mM Na_2EDTA.

Changes in the CD spectrum for the RNA target upon introduction of the homologous probes are less dramatic than those recorded in the UV-melting experiments (Figure 4.1B): While the peak positions remain the same in both the RNA and RNA-PNA hybrid quadruplexes, we observed an attenuation in the amplitude of the CD signals recorded for the latter (Figure 4.1B). This reduction in signal, even with nearly twice as many nucleobases in the hetero-quadruplex, is probably a consequence of the partial relaxation in the helical twist of the RNA target to accommodate the PNA probes. Importantly, the peak positions are characteristic of parallel hetero-quadruplex structures—in which the N-termini of the PNA probes are aligned with the 5'-end of the RNA target[13]. Together, our data thus suggest that these homologous probes can invade the quadruplex target and form very stable parallel hybrid quadruplexes.

4.3.3 Effects of P_{AG} and P_{eg2} on Translation

To investigate the functional implications of PNA binding in the context of a biochemical assay, we employed the luciferase reporter system[8, 27] previously utilized to demonstrate the inhibitory effects on translation exerted by complementary γPNA oligomers (Chapters 2 and 3). As in previous chapters, two luciferase transcripts were tested: (1) a reporter transcript bearing the **4G3** sequence in its 5'-UTR was used to examine the effects of PNA-mediated quadruplex invasion on translation. (2) We also tested a control transcript (**No G3**) ostensibly devoid of a binding site for the G-rich homologous probes to determine the specificity of an induced effects on translation.

4.3.3.1 P_{AG}

Titration of the **4G3** reporter transcript with increasing concentrations of $\mathbf{P_{AG}}$ elicited modest dose-dependent inhibitory effects on translation (Figure 4.2A, filled squares). However, control experiments with the **No G3** transcript produced an overlapping dose-response curve, indicating that the effects recorded against **4G3** are not due exclusively to quadruplex recognition. While hybrid formation between this probe and a *synthetic, truncated* (~40 nt) RNA target was demonstrated in buffer solution, it is not clear whether the same reaction would occur with an *in vitro transcribed* (~ 2000 nt) RNA in the context of mammalian cell lysates.

Further, although intermolecular quadruples formed between G-rich targets and homologous *RNA* probes have been shown to inhibit GFP expression[28], such repressive effects on translation have not been reported for *PNA*-containing hybrid quadruplexes. It is therefore possible that the PNA-RNA quadruplex—if formed, is efficiently resolved by the translocating 43 S ribosomal subunit[29]. However, regardless of the basis for the similar results in the presence and absence of the quadruplex target, it is clear that alternative, if off-target, PNA binding sites mediate translation inhibition in both the **No G3** control and **4G3** target reporters.

4.3.3.2 $\mathbf{P_{eg2}}$

Experiments with this probe also produced identical dose-response curves against the reporter (**4G3**) and control transcripts (Figure 4.2B), as was observed for $\mathbf{P_{AG}}$. These results are surprising, since this probe bears modifications[17] introduced to disfavor recognition of potential C-rich off-target sequences (*vide infra*) speculated to mediate binding of $\mathbf{P_{AG}}$ to the control transcript. Interestingly, translation inhibition by $\mathbf{P_{eg2}}$, even though non-specific, is more potent than the effects recorded with $\mathbf{P_{AG}}$ (Figure 4.2).

Further, the non-specific repressive effect of this probe is not due to direct inhibition of the ribosomes (possibly by rRNA binding), since incubating 250 nM $\mathbf{P_{eg2}}$ with cell lysate prior to addition of a control transcript (supplied by Promega) did not perturb translation (Figure C1). The inertness of this probe to the commercially available control transcript is surprising, and may be a consequence of the fact that this transcript is codon-optimized by the manufacturer for luciferase expression in this lysate system. Therefore, any mRNA molecules left unhybridized might be sufficiently well-translated to compensate for any loss in signal induced by PNA binding.

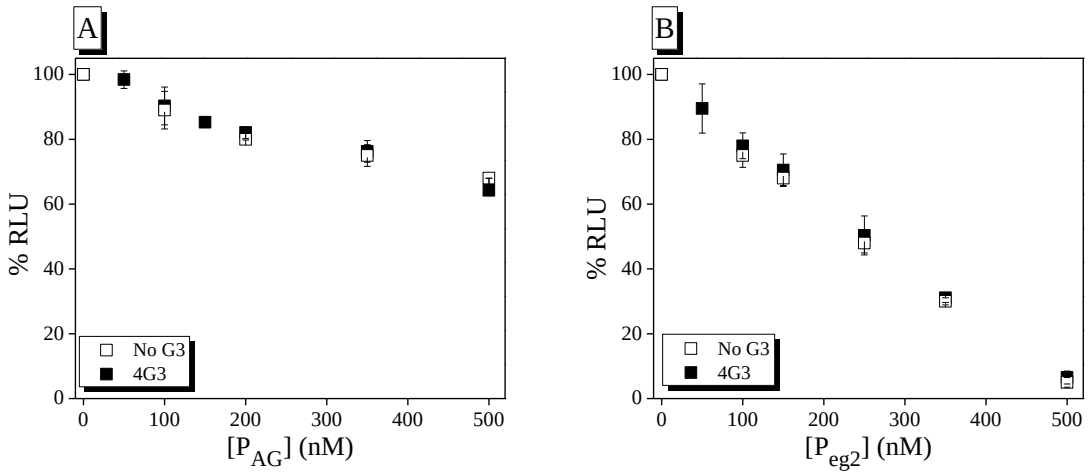


Figure 4.2: Effect of PNAs on luciferase mRNA translation. Dose-response curves for P_{AG} (**A**) and P_{eg2} (**B**) against 4G3 target (filled squares) and No G3 control (open squares) mRNAs. [mRNA] = 10 nM and PNA-RNA samples were incubated for one hr at 37 °C prior to adding to lysate. Data are normalized to 100% for the samples that lacked any PNA.

4.3.4 Elucidating Off-Target Binding by P_{eg2}

To understand the contributions of off-target binding to the translation inhibition recorded for $\mathbf{P_{eg2}}$, we examined the entire sequence of the **No G3** control transcript to identify potential C-rich sequences that could sequester the G-rich $\mathbf{P_{eg2}}$ in unintended duplexes, thus simultaneously decreasing its potency—by withholding active molecules from the quadruplex target, and its specificity, provided the hybrid duplexes stall processing ribosomes. (The entire sequence for the control and reporter transcripts, with relevant sites highlighted, is presented in Figure C2-C3.) Chart 4.2 presents the sequences of potential C-rich off-target sites and their positions relative to the start codon. Interestingly, all identified sites occur within the coding sequence for the firefly luciferase reporter, implying that their potential contributions to off-target binding would be present in both the control and reporter systems.

Chart 4.2: Sequences of C-rich stretches that could potentiate off-target binding and inhibition. (DNA sequence written 5'-3'. Positions given are after the start codon.)

Off target	Sequence	Position
OT1	CGCCCGC	305-312
OT2	ACCTCCCG	513-521
OT3	CCCTTCC	623-630
OT4	CCAACCC	862-869
OT5	GGCCCCCGCT	1344-1354
OT6	ACACCCCAAC	1377-1387
OT7	TCCCGCCGCC	1437-1447

A competition assay was then employed to identify the most culpable off target site(s) by incubating 200 nM $\mathbf{P_{eg2}}$—the IC_{50} for translation inhibition (Figure 4.2B)—with an equimolar amount of synthetic DNA strands (called **OTx**) corresponding to the potential off-target sites for 1 h to allow hybridization. The preformed hybrids were subsequently incubated with the **No G3** control transcript for 1 h prior to addition of the cell lysate. Figure 4.3A summarizes the extent of translation inhibition induced by $\mathbf{P_{eg2}}$ in the presence (black bars) or absence (red bar) of the competing 'off-target' DNA strands. While all competing DNA strands diminish translation inhibition by the probe (Figure 4.3A, black bars), we observe near-complete abrogation of the inhibitory effect of $\mathbf{P_{eg2}}$ by **OT7**. Importantly, all **OTx** DNA strands showed moderate or negligible effects on **No G3** translation (Figure 4.3B), demonstrating that any inhibition observed in Figure 4.3A was due to residual unhybridized probe molecules.

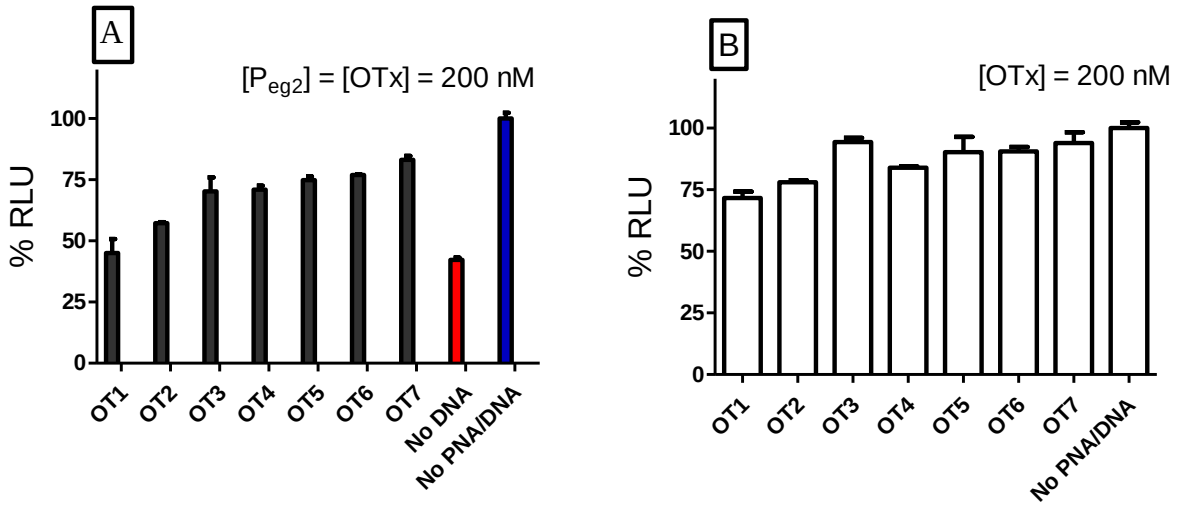


Figure 4.3: Luciferase competition experiments. Luciferase expression from No G3 mRNA samples treated with P_{eg2} in the presence (black bars) or absence (red bar) of competing OTx DNA oligomers (**A**). Luciferase expression from No G3 mRNA samples treated with OTx oligomers alone (**B**). PNA-DNA samples were incubated for 1 h at 37 ˚C prior to addition of mRNA transcript. Cocktail was incubated for another 1 h prior to addition of lysates. Results are normalized to signal from samples without DNA/PNA treatment. [mRNA] = 10 nM; where present, [PNA]/[DNA] = 200 nM

Further, to determine if the remediation of the probe's (**P_{eg2}**) non-specific inhibition by **OT7** is a titratable effect, we incubated 200 nM **P_{eg2}** with increasing concentrations of the competing DNA strand. Our results show that while the un-encumbered probe displays significant inhibitory effects on translation of the **No G3** transcript (Figure 4.4, red bar), increasing concentrations of **OT7** gradually improve the probe's inertness to the control transcript (Figure 4.4, filled bars). Importantly, full recovery of specificity was achieved at 120 nM **OT7** (Figure 4.4), demonstrating the potential to mitigate the probe's non-specific effect with sub-stoichiometric amounts of a competing DNA strand.

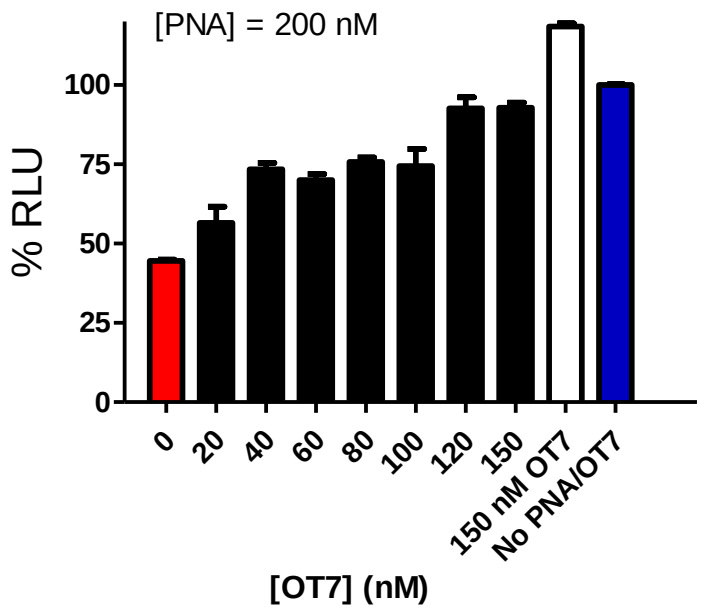


Figure 4.4: Luciferase competition experiments. Luciferase expression from No G3 mRNA samples treated with P_{eg2} in the absence (red bar) or presence (black bars) of 20 – 150 nM OTx DNA. PNA-DNA samples were incubated for 1 h at 37 ˚C prior to addition of mRNA transcript. Cocktail was incubated for another 1 h prior to addition of lysates. Results are normalized to signal from samples without DNA/PNA treatment. [mRNA] = 10 nM; where present, [PNA] = 200 nM

We next determined if **OT7** could extricate the non-specific effects of $\mathbf{P_{eg2}}$ in the context of the **4G3** reporter transcript. Equimolar amounts (200 nM) of the probe and competing DNA strand were preincubated for 1 hour prior to addition of the **4G3** transcript (for 1 h) and subsequent introduction of the cell lysate. Figure 4.5 shows that the presence of **OT7** diminishes but does not completely eliminate the inhibitory effect of $\mathbf{P_{eg2}}$ on translation of the **4G3** reporter (Figure 4.5). Importantly, $\mathbf{P_{eg2}}$, in the presence of **OT7**, displays only a negligible effect on translation of the **No G3** control. These results, taken together, identify **OT7** as a potential source of off-target

inhibition by the PNA probe, and suggest that strategies to discourage hybridization to this site might improve specificity of translation inhibition.

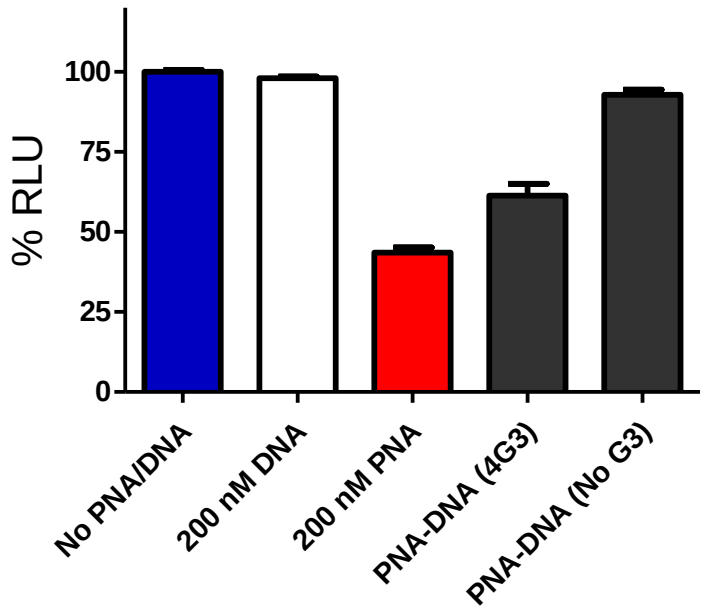


Figure 4.5: Effects of PNA-DNA hybrids on mRNA translation. Luciferase expression from 4G3 mRNA samples treated with P_{eg2} in the absence (red bar) or presence (black bar) of OTx DNA. PNA-DNA samples were incubated for 1 h at 37 ˚C prior to addition of mRNA transcript. Cocktail was incubated for another 1 h prior to addition of lysates. Results are normalized to signal from samples without DNA/PNA treatment. [mRNA] = 10 nM; where present, [PNA]/DNA = 200 nM

4.3.5 Introduction of γ-modifications to Tune Specificity

To obviate the need for a competing DNA strand to tune the specificity of $\mathbf{P_{eg2}}$, we introduced γ-modifications at strategic positions of the oligomer to drive left-handed helical pre-organization and potentially discourage hybridization to the C-rich off-target sequences identified to facilitate inhibition of the control transcript. The modifications were made by substituting the G-monomers

with derivatives bearing the same nucleobase on a hydroxymethyl-modified PNA backbone[25]. NMR and CD spectroscopic analyses previously revealed that helical induction potentiated by γ-modification in PNA oligomers is unidirectional, with effects originating from the PNA C-terminus[24].

We therefore made substitutions at the three G residues proximal to the C-terminus (as in **^{4,5,6}D-P$_{eg2}$**) or at the C-terminal G residue of each G_3 tract (as in **^{3,6}D-P$_{eg2}$**, Chart 4.1). CD spectra for these sequences showed that they were weakly pre-organized, displaying—with weak intensities (Figure C4)—peaks previously reported for a similar left-handed homologous probe (maxima at 240- and 275 nm; minima at 260- and 295 nm)[17]. Importantly, UV-melting experiments confirmed that both probes were able to form stable hybrid quadruplexes with the 4G3 target, consistent with a previous report showing that this recognition mode is sufficiently robust to accommodate even drastic modifications to the PNA backbone[17] (Table 4.1 and Figure C5).

Table 4.1. T_m values (°C) for different PNA/γPNA-4G3 RNA heteroquadruplexes.

PNA/γPNA	T_m (°C)
P$_{eg2}$	> 80
^{3,6}D-P$_{eg2}$	79.2 ± 0.2
^{4,5,6}D-P$_{eg2}$	72.3 ± 0.5

SPR direct-binding experiments were next used to examine the interactions between the modified probes and an immobilized C-rich DNA target. (Although a more ideal assay would have examined binding to immobilized RNA, this experiment was still expected to inform us about the potential for the modified probes to discriminate against C-rich targets.) Figure 4.6 presents sensorgrams

for binding of both the unmodified and modified probes to the immobilized C-rich DNA strand. We observed strong disparities in the binding responses between **P_{eg2}** and either of the modified probes (Figure 4.6). More modest differences were observed between the initial association rates, with **P_{eg2}** hybridizing 1.4 and 2.3 times faster than **^{3,6}D-P_{eg2}** and **^{4,5,6}D-P_{eg2}**, respectively (Figure 4.6). Further, we observed fast dissociation rates for all probes from the target, and this trend was more pronounced for those containing γ-modifications. This result is consistent with our previous report showing that the incorporation of abasic residues and γ-modifications to a homologous probe mitigated binding to an immobilized C-rich target, while also accelerating dissociation of any bound probe molecules from the target[17].

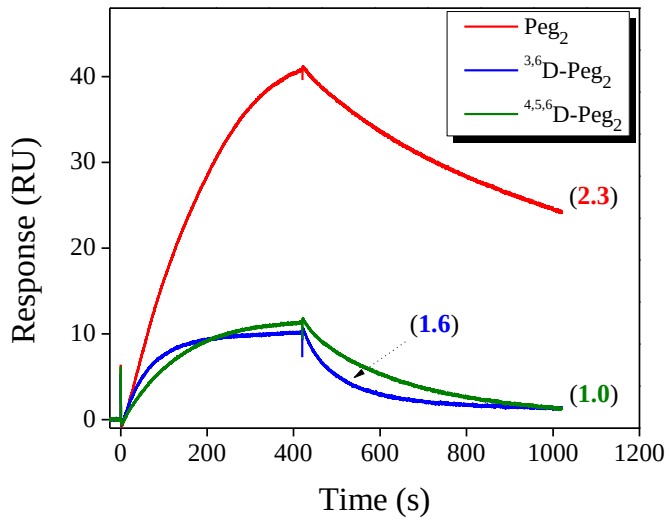


Figure 4.6: Sensorgrams for binding of P_{eg2} (red curve), ^{3,6}D-P_{eg2} (blue curve), and ^{4,5,6}D-P_{eg2} (green curve) to immobilized C-rich DNA (see methods). [PNA] or [γPNA] = 25 nM. Running buffer contained 100 mM KCl, 10 mM Tris-HCl (pH = 7.4), 0.1 mM Na_2EDTA, and P20 surfactant (0.1 % v/v). Relative initial association rates are presented in parenthesis.

We also sought to examine the extent to which the modified residues would frustrate hybrid quadruplex formation between the probes and RNA target. While UV-melting experiments (Figure C5) provided information on the thermal stability of the hybrids already formed with the target, we sought to examine if the modifications would decelerate quadruplex invasion. SPR direct-binding experiments were therefore performed to examine the relative initial association rates for PNA-RNA binding. (The modified SPR assay designed to enable this sort of analysis on immobilized RNA was reported in Chapter 2 and Ref 8.)

Figure 4.7 presents sensorgrams for quadruplex invasion by either the modified or unmodified homologous probes. We observe that invasion of the **4G3** quadruplex by $\mathbf{P_{eg2}}$ is 3 and 5 times faster than $^{3,6}\mathbf{D\text{-}P_{eg2}}$ and $^{4,5,6}\mathbf{D\text{-}P_{eg2}}$, respectively (Figure 4.7), suggesting that the γ-modifications also decelerate hybridization to the **4G3** QFS. This data is consistent with our previous finding that left-handed preorganization of a PNA probe impeded hybridization to the DNA QFS, although the incorporated modifications affected hetero-duplex formation (to a C-rich target) more severely[17]. Comparisons between these results and those in Figure 6 for binding to the C-rich target are not straightforward, since both backbone (DNA vs. RNA) and sequence disparities (G-rich vs C-rich) likely impact the results. However, it is noteworthy that we obtained faster off-rates with the C-rich target (Figure 4.6) than the G-rich target (Figure 4.7). Similar SPR experiments comparing hybridization to G-rich and C-rich RNA strands should be of utmost importance in future studies.

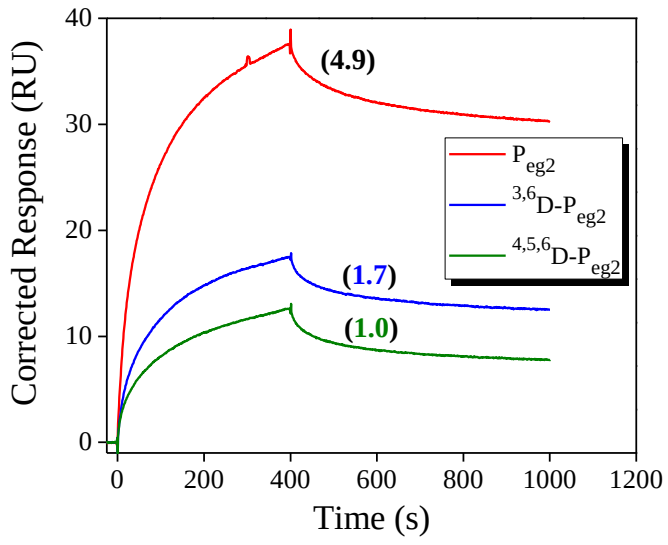


Figure 4.7: Sensorgrams for binding of P_{eg2} (red curve), ^{3,6}D-P_{eg2} (blue curve), and ^{4,5,6}D-P_{eg2} (green curve) to immobilized 4G3 RNA (see methods). [PNA] or [γPNA] = 25 nM. Running buffer contained 100 mM KCl, 10 mM Tris-HCl (pH = 7.4), 0.1 mM Na_2EDTA, and P20 surfactant (0.1 % v/v). Relative initial association rates are presented in parenthesis.

4.3.6 Effects of ^{3,6}D-P_{eg2} and ^{4,5,6}D-P_{eg2} on Translation

We next sought to examine if the modifications introduced to discourage hybridization to C-rich targets would enhance the probes' specificities in translation inhibition. Both the **4G3** reporter and **No G3** control transcripts were titrated with increasing concentrations of each probe. Our results show that both probes still exhibit non-specific repressive effects on translation, with overlapping dose-response curves against both transcripts (Figure 4.8). These results suggest that the modified probes still do not discriminate against the off-target sequences that mediate non-specific binding and inhibition.

It is possible that the modifications introduced to $\mathbf{P_{eg2}}$ are insufficient to discourage hybridization to C-rich off-targets in the context of the *in vitro* transcribed RNA. Although the competition

experiments described above could be broadly applied to elucidate the impact of probe modifications on specificity, they are unlikely to be useful to evaluate the γPNAs in this context, since *left-handed* helical preorganization of the probes—as is the case with **^{3,6}D-P$_{eg2}$** and **^{4,5,6}D-P$_{eg2}$**—would impede the ability of the competing *right-handed* DNA strand to sequester active probe molecules from the mRNA transcript. Whether additional modifications to further bias recognition against C-rich targets will improve specificity remains to be seen. Together, our results show that the inhibitory effects on translation recorded for these homologous probes cannot, at this point, be ascribed to invasion of the **4G3** quadruplex in the 5'-UTR.

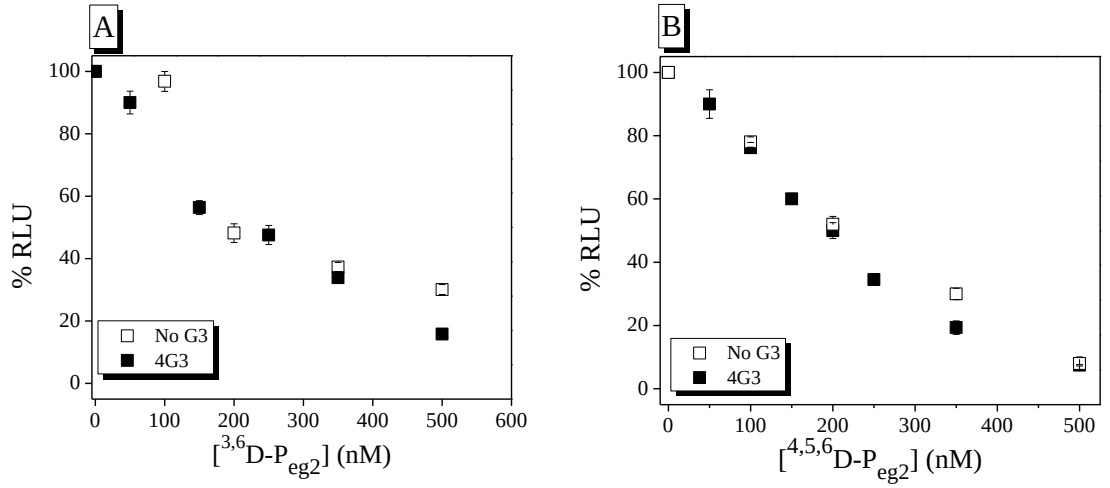


Figure 4.8: Effects of γPNAs on mRNA translation. Dose response curves for ^{3,6}D-P$_{eg2}$ (**A**) or ^{4,5,6}D-P$_{eg2}$ (**B**) oligomers against 4Gx target (filled squares) or control (open squares) reporters. [mRNA] = 10 nM, and RNA-γPNA samples were incubated for one hr at 37 °C prior to adding to lysate. Data are normalized to 100% for the samples that lacked any γPNA probe.

4.4 DISCUSSION

The data presented in this chapter show that homologous PNA probes can invade a stable RNA quadruplex, yielding stable hybrid RNA-PNA heteroquadruplexes. We also observed translation inhibition by these probes, with more potent—albeit nonspecific—repressive effects recorded when the intervening, non-bonding residues were replaced with abasic units. The non-specific effects are mediated, at least partly—by C-rich stretches within the coding sequence of the luciferase. Specific inhibition is restored by competing DNA strands that sequester probe molecules from potential off-target sites. However, we find that helical preorganization of the probe by incorporation of γ-modified (**^{D}G)** monomers, in the absence of the competing DNA strands, is insufficient to engineer specificity. These results, taken together, suggest that while homologous probes may have the potential to form stable hybrid quadruplexes with a G-rich target, additional modifications—beyond those reported here—would be required to tune specificity and understand the regulatory impact of the resulting hybrids.

While our data are inconclusive on the regulatory effects induced by PNA-RNA heteroquadruplexes, the viability of this nucleic acid recognition mode for effecting biochemical and/or biological regulation has been demonstrated in previous reports. We previously showed that hybrid quadruplexes formed between a homologous PNA probe and QFS-laden DNA templates could stall primer extension by DNA pol η, with the major stop site occurring before the PNA-DNA heteroquadruplexes[30]. Evidently, these hybrids leverage their remarkable thermodynamic stabilities relative to the DNA quadruplex to stall processing and/or translocation by the polymerase. Further, Watson-Crick-based duplex hybrids formed by replacing the homologous probes with complementary PNAs *enhanced* primer extension, demonstrating that the

repressive effects were specific to hybrid quadruplex formation, and not due just to binding at the QFS motif[30].

In a related example, Ito *et al* demonstrated that a small RNA oligomer homologous to G-rich inserts in the 5'-UTR, coding sequence, or both of an eGFP reporter could inhibit translation by forming RNA-RNA intermolecular quadruplexes in HeLa cells[28]. The authors further showed that the potency of inhibition was dependent on both the number of binding sites available to the probe and, in the cases where only one obvious site was available—the position of the resulting intermolecular quadruplex[28]. Importantly, the fidelity of recognition—or at least, specificity of inhibition was demonstrated by the inertness of the probe to a control reporter devoid of the G-rich insert[28]. This result stands in contradistinction to ours, as we record significant nonspecific effects with the homologous PNA. Comparisons between these data sets are difficult, since different reporter systems (firefly luciferase[8] vs eGFP[28]) and assay contexts (cell lysates[8] vs HeLa cells[28]) are utilized. However, it is possible that the higher affinity of PNA over RNA for complementary RNA sequences confers on the former a tolerance for mismatched off-target sequences that will not be recognized by the latter.

The intransigence of the nonspecific effects in our system is surprising, as even modifications to induce left-handed preorganization do not alter either the potency or specificity of inhibition. While additional modifications to further pre-organize the probes could be incorporated, it is also possible that the non-specific effects arise not from C-rich off-targets, but from G-rich, non-quadruplex-forming stretches also present in the luciferase coding sequence (Figure C2-B3). These sites, expected to be more accessible than the folded **4G3** QFS, could simultaneously sequester probe molecules from the QFS target (in the reporter transcript) and produce nonspecific inhibition in the control assay.

Also, although competition experiments demonstrate the culpability of the C-rich stretches in mediating non-specific binding, it is possible that the hybrid duplexes formed between $\mathbf{P_{eg2}}$ and the **OTx** fragments enhance specificity by mitigating binding of the probes to the G-rich stretches. This hypothesis is supported by a previous report showing that PNA-RNA hybrid duplexes—the structures formed from PNA hybridization to the ostensible C-rich off-targets—formed distal to the 5'-end of the message display only weak effects on translation[31]. (**OT7** occurs ~ 1,440 nucleotides from the start codon.) A more informative experiment then would involve the introduction of 'masking' strands that veil the suspected off-target sites to examine if they improve specific action by the PNA probe. However, such assays would be useful only to the extent that the 'masks' have sufficient affinities for their targets to compete favorably with potential displacement by the PNA.

In conclusion, a regulatory role for the RNA-PNA hybrid quadruplex cannot be assigned at this time, as additional efforts will be required to improve the probe's specificity. However, the differential repressive effects of the $\mathbf{P_{eg2}}$**-OT7** hybrid on translation of the reporter and control transcripts suggest that it might be possible to de-conflate specific, heteroquadruplex-mediated effects from nonspecific inhibition in the context of the reporter transcript. It will also be important to demonstrate hybrid (heteroquadruplex) formation *in vivo*, and the myriad strategies—such as backbone modification with guanidinium groups[32], conjugation to cell-penetrating peptides[33], or formulation into polymeric nanoparticles[34]—being developed for intracellular PNA delivery should facilitate this effort.

4.5 REFERENCES

[1] Haeusler, A. R., Donnelly, C. J., Periz, G., Simko, E. A., Shaw, P. G., Kim, M. S., Maragakis, N. J., Troncoso, J. C., Pandey, A., Sattler, R., Rothstein, J. D., and Wang, J. (2014) C9orf72 nucleotide repeat structures initiate molecular cascades of disease, *Nature 507*, 195-200.

[2] Green, J. J., Ying, L., Klenerman, D., and Balasubramanian, S. (2003) Kinetics of unfolding the human telomeric DNA quadruplex using a PNA trap, *J. Am. Chem. Soc. 125*, 3763–3767.

[3] Datta, B., and Armitage, B. A. (2001) Hybridization of PNA to Structured DNA Targets: Quadruplex Invasion and the Overhang Effect, *J. Am. Chem. Soc. 123*, 9612–9619.

[4] Amato, J., Pagano, B., Borbone, N., Oliviero, G., Gabelica, V., Pauw, E. D., D'Errico, S., Piccialli, V., Varra, M., Giancola, C., Piccialli, G., and Mayol, L. (2011) Targeting G-quadruplex structure in the human c-Kit promoter with short PNA sequences, *Bioconjug. Chem. 22*, 654-663.

[5] Kumar, N., Patowary, A., Sivasubbu, S., Petersen, M., and Maiti, S. (2008) Silencing c-MYC Expression by Targeting Quadruplex in P1 Promoter Using Locked Nucleic Acid Trap, *Biochemistry 47*, 13179-13188.

[6] Rouleau, S. G., Beaudoin, J. D., Bisaillon, M., and Perreault, J. P. (2015) Small antisense oligonucleotides against G-quadruplexes: specific mRNA translational switches, *Nucleic Acids Res. 43*, 595-606.

[7] Murat, P., Zhong, J., Lekieffre, L., Cowieson, N. P., Clancy, J. L., Preiss, T., Balasubramanian, S., Khanna, R., and Tellam, J. (2014) G-quadruplexes regulate Epstein-Barr virus-encoded nuclear antigen 1 mRNA translation, *Nat. Chem. Biol. 10*, 358-364.

[8] Oyaghire, S. N., Cherubim, C. J., Telmer, C. A., Martinez, J. A., Bruchez, M. P., and Armitage, B. A. (2016) RNA G-Quadruplex Invasion and Translation Inhibition by Antisense gamma-Peptide Nucleic Acid Oligomers, *Biochemistry 55*, 1977-1988.

[9] Gupta, A., Lee, L., Roy, S., Tanious, F. A., Wilson, W. D., Ly, D. H., and Armitage, B. A. (2013) Strand invasion of DNA quadruplexes by PNA: comparison of homologous and complementary hybridization, *ChemBioChem 14*, 1476–1484.

[10] Amato, J., Gabelica, V., Borbone, N., Rosu, F., De Pauw, E., Oliviero, G., Piccialli, G., and Mayol, L. (2008) A short C-rich PNA fragment capable to form novel G-quadruplex-PNA complexes, *Nucleic Acids Symp. Ser. (Oxf.) 52*, 167-168.

[11] Amato, J., Oliviero, G., De Pauw, E., and Gabelica, V. (2009) Hybridization of short complementary PNAs to G-quadruplex forming oligonucleotides: An electrospray mass spectrometry study, *Biopolymers 91*, 244-255.

[12] Watson, J. D., and Crick, F. H. C. (1953) Molecular Structure of Nucleic Acids- A Structure for Deoxyribose Nucleic Acid, *Nature 171*, 737-738.

[13] Datta, B., Schmitt, C., and Armitage, B. A. (2003) Formation of a PNA 2 - DNA 2 Hybrid Quadruplex, *J. Am. Chem. Soc. 125*, 4111-4118.

[14] Marin, V. L., and Armitage, B. A. (2005) RNA Guanine Quadruplex Invasion by Complementary and Homologous PNA Probes, *J. Am. Chem. Soc. 127*, 8032–8033.

[15] Marin, V. L., and Armitage, B. A. (2006) Hybridization of Complementary and Homologous Peptide Nucleic Acid Oligomers to a Guanine Quadruplex-Forming RNA, *Biochemistry 45*, 1745–1754.

[16] Roy, S., Tanious, F. A., Wilson, W. D., Ly, D. H., and Armitage, B. A. (2007) High-affinity homologous peptide nucleic acid probes for targeting a quadruplex-forming sequence from a MYC promoter element, *Biochemistry 46*, 10433–10443.
[17] Lusvarghi, S., Murphy, C. T., Roy, S., Tanious, F. A., Sacui, I., Wilson, W. D., Ly, D. H., and Armitage, B. A. (2009) Loop and backbone modifications of peptide nucleic acid improve g-quadruplex binding selectivity, *J. Am. Chem. Soc. 131*, 18415-18424.
[18] Roy, S., Zanotti, K. J., Murphy, C. T., Tanious, F. A., Wilson, W. D., Ly, D. H., and Armitage, B. A. (2011) Kinetic discrimination in recognition of DNA quadruplex targets by guanine-rich heteroquadruplex-forming PNA probes, *Chem. Commun. 47*, 8524-8526.
[19] Kormuth, K. A., Woolford, J. L., Jr., and Armitage, B. A. (2016) Homologous PNA Hybridization to Noncanonical DNA G-Quadruplexes, *Biochemistry 55*, 1749-1757.
[20] Englund, E. A., Xu, Q., Witschi, M. A., and Appella, D. H. (2006) PNA-DNA duplexes, triplexes, and quadruplexes are stabilized with trans-cyclopentane units, *J. Am. Chem. Soc. 128*, 16456-16457.
[21] Gaynutdinov, T. I., Englund, E. A., Appella, D. H., Onyshchenko, M. I., Neumann, R. D., and Panyutin, I. G. (2015) G-quadruplex formation between G-rich PNA and homologous sequences in oligonucleotides and supercoiled plasmid DNA, *Nucleic Acid Ther. 25*, 78-84.
[22] Paul, A., Sengupta, P., Krishnan, Y., and Ladame, S. (2008) Combining G-quadruplex targeting motifs on a single peptide nucleic acid scaffold: a hybrid (3+1) PNA-DNA bimolecular quadruplex, *Chem. Eur. J. 14*, 8682-8689.
[23] Ishizuka, T., Yang, J., Komiyama, M., and Xu, Y. (2012) G-rich sequence-specific recognition and scission of human genome by PNA/DNA hybrid G-quadruplex formation, *Angew. Chem. Intl. Ed. Engl. 51*, 7198-7202.
[24] Dragulescu-Andrasi, A., Rapireddy, S., Frezza, B. M., Gayathri, C., Gil, R. R., and Ly, D. H. (2006) A Simple γ-Backbone Modification Preorganizes Peptide Nucleic Acid into a Helical Structure, *J. Am. Chem. Soc. 2006*, 10258–10267.
[25] Sahu, B., Sacui, I., Rapireddy, S., Zanotti, K. J., Bahal, R., Armitage, B. A., and Ly, D. H. (2011) Synthesis and characterization of conformationally preorganized, (R)-diethylene glycol-containing gamma-peptide nucleic acids with superior hybridization properties and water solubility, *J. Org. Chem. 76*, 5614-5627.
[26] Amblard, M., Fehrentz, J. A., Martinez, J., and Subra, G. (2006) Methods and protocols of modern solid phase Peptide synthesis, *Mol. Biotechnol. 33*, 239-254.
[27] Canady, T. D., Telmer, C. A., Oyaghire, S. N., Armitage, B. A., and Bruchez, M. P. (2015) In Vitro Reversible Translation Control Using gammaPNA Probes, *J. Am. Chem. Soc. 137*, 10268-10275.
[28] Ito, K., Go, S., Komiyama, M., and Xu, Y. (2011) Inhibition of translation by small RNA-stabilized mRNA structures in human cells, *J. Am. Chem. Soc. 133*, 19153-19159.
[29] Jackson, R. J., Hellen, C. U., and Pestova, T. V. (2010) The mechanism of eukaryotic translation initiation and principles of its regulation, *Nat. Rev. Mol. Cell. Biol. 11*, 113-127.
[30] Murphy, C. T., Gupta, A., Armitage, B. A., and Opresko, P. L. (2014) Hybridization of G-quadruplex-forming peptide nucleic acids to guanine-rich DNA templates inhibits DNA polymerase eta extension, *Biochemistry 53*, 5315-5322.

[31] Doyle, D. F., Braasch, D. A., Simmons, C. G., Janowski, B. A., and Corey, D. R. (2001) Inhibition of Gene Expression Inside Cells by Peptide Nucleic Acids: Effect of mRNA Target Sequence, Mismatched Bases, and PNA Length, *Biochemistry 40*, 53-64.
[32] Sahu, B., Chenna, V., Lathrop, K. L., Thomas, S. M., Zon, G., Livak, K. J., and Ly, D. H. (2009) Synthesis of Conformationally Preorganized and Cell-Permeable Guanidine-Based γ-Peptide Nucleic Acids (γGPNAs), *J. Org. Chem. 74*, 1509–1516.
[33] Copolovici, D. M., Langel, K., Eriste, E., and Langel, Ü. (2014) Cell-Penetrating Peptides: Design, Synthesis, and Applications, *ACS Nano 8*, 1972–1994.
[34] Gupta, A., Bahal, R., Gupta, M., Glazer, P. M., and Saltzman, W. M. (2016) Nanotechnology for delivery of peptide nucleic acids (PNAs), *J. Control. Release* 10.1016/j.jconrel.2016.01.005.

CHAPTER 5

TOWARDS CHIMERIC RECOGNITION OF GUANINE-QUADRUPLEXES

5.1 INTRODUCTION

The broad and varying regulatory roles[1-4] of guanine (G)-quadruplexes in important biological processes (reviewed in Chapter 1) incentivize the development of specific ligands able to recognize these structures and modulate their biochemical and/or biological functions. Such reagents would find use as tools for probing or verifying quadruplex function, thus obviating the need for quadruplex-destabilizing mutations[5] or sequence deletions[6]—with the attendant loss of sequence context—to functionally annotate QFS motifs. Towards this end, several parallel strategies for quadruplex recognition are being developed, among which *shape-selective recognition* using small-molecule ligands[7], and *sequence-selective recognition* using oligomeric ligands[8] that anneal to the primary structure of the quadruplex hold great promise.

The isolation of telomestatin from *Streptomyces anulatus* and demonstration of its inhibitory effect on telomerase[9] triggered initial interest in designing similar quadruplex-binding ligands[10] to act as anti-proliferative agents through modulation of telomerase activity. A common strategy is to design molecules with planar aromatic surfaces expected to be π-complementary to the terminal G-quartet(s) of a quadruplex structure[7, 11]. However, while this strategy can sometimes provide drug-like molecules with useful affinities for their targets[7], achieving binding selectivity is often difficult[12], since the quadruplex structure—especially in RNA[13]—is not sufficiently polymorphic to provide the selection pressure needed to isolate specific binders.

Selecting a single quadruplex target is likely to be achieved by designing molecules that recognize structural features unique to the target. For example, the discovery that G-quadruplex structures can tolerate bulges within the G_x tracts[14], and that long intervening loop nucleotides do not preclude quadruplex folding[15] introduces a diversity to quadruplex structures that might be helpful

for selective targeting. However, while such structural variations provide an opportunity to screen molecules *in silico* for binding to specific targets, progress via this approach is limited by several factors, not least of which is the paucity of comprehensive structural data on putative quadruplexes in biologically-relevant regions of the genome.

Sequence-selective strategies obviate the need for detailed structural characterization of the quadruplex target, requiring only that the primary structure (sequence) of the QFS motif be known prior to ligand design. We have reported that complementary[16-21] or homologous[17-20, 22-26] peptide nucleic acid (PNA)[27, 28] or gamma-peptide nucleic acid (γPNA)[29, 30] oligomers can invade stable guanine quadruplexes—and, in some examples,[20, 21] modulate their biochemical functions. Hybridization by the former molecules yields hybrid duplex structures[16-21], while the latter result in heteroquadruplexes[17-20, 22-26], both of which are more stable than the quadruplex targets.

However, the broad application of this strategy for quadruplex/QFS targeting remains stymied by the stringent criteria requisite for selective quadruplex recognition. At least two levels of selectivity are required of any hybridization probe: (1) the oligomer, whether complementary or homologous, should be able to discriminate against off target sequences presenting binding sites that are partially complementary to the probe. Such sites, even if thermodynamically disfavored, might act as kinetic traps that decelerate binding to the QFS target. (2) The oligomer should also be able to select for a single quadruplex, while evading similar structures that diverge from the target only in areas such as loop and/or overhang composition/length. The consequences of promiscuity in this context could be more sinister, since recognition of off-target QFS motifs—reported to be overrepresented in regions of the genome associated with regulation[31-33], could potentiate unintended regulation with potential deleterious downstream effects.

Molecular engineering efforts to satisfy the above criteria are slow in accruing, but show promise towards the intended goals. We previously showed that incorporation of abasic units in the intervening region between the G_3 tracts of a homologous PNA disfavored binding to a C-rich target, with minimal perturbation in the binding affinity for a DNA QFS target[24]. Discrimination against the C-rich target was further improved by substituting the unmodified G-residues of the PNA with γ-modified G-monomers that induce left-handed helical pre-organization[24]. While the ability of such modifications to induce selective binding to QFS motifs instead of C-rich off-targets in the context of a biochemical or biological assay is yet to be reported, these results demonstrate the potential to achieve such selectivity, at least in buffer solutions.

We also observed kinetic discrimination in hybridization of a homologous probe to different DNA G-quadruplexes, with preferential (faster) binding to a parallel folded target[25]. However, while these results are interesting, their extrapolation to inform selective recognition of RNA G-quadruplexes—which predominantly adopt parallel folds[13]—remains to be seen.

More recently, we reported that recognition of overhanging bases adjacent to an RNA G-quadruplex was crucial for quadruplex invasion and translation inhibition by a complementary γPNA oligomer[21]. Evidently, overhang recognition conferred upon the probe a kinetic advantage that was absent in γPNAs directed to other regions of the QFS, even though they possessed the requisite complementary nucleobases[21].

This dependence on overhang recognition for invasion highlights an interesting potential strategy for achieving selective recognition of G-quadruplexes using PNA or γPNA oligomers: it might be possible to tether an overhang-recognizing complementary domain to a quadruplex-invading homologous domain (Scheme 5.1, *vide infra*), thus creating a chimeric molecule able to

simultaneously form a hybrid duplex and quadruplex, respectively, with the target. The putative chimeric probe should preserve the orientational preferences previously reported for PNA-DNA/RNA hybrid duplexes[34] and quadruplexes[22] in order to ensure synergy between the two binding domains. Further, for discrimination between different quadruplex targets, the binding energy of the probe should be sufficiently delocalized across both domains to ensure that off-target quadruplexes do not remain annealed to the probe by binding to one domain.

This chapter summarizes our efforts to achieve chimeric recognition of a G-rich sequence using a PNA oligomer possessing both homologous and complementary domains. We observe that truncated probes possessing the isolated (homologous or complementary, only) domains can form the expected hybrid quadruplex or duplex structures, respectively. However, the contributions of both in the context of the chimeric structure remain nebulous. Importantly, titration curves generated by SPR competition analysis show that the chimeric molecule is bound more readily by a target presenting binding sites for both domains of the probe, with less binding to a sequence presenting only one domain. These results illustrate the potential to achieve selective quadruplex targeting using chimeric probes that recognize unique features of the target in addition to the QFS.

5.2 MATERIALS AND METHODS

5.2.1 RNA/DNA Oligomers

RNA and DNA oligomers used (Chart 5.1) were obtained from Integrated DNA Technologies (www.idtdna.com). The DNA capture strand used in SPR competition experiments is modified with biotin on the 5'-end (Chart 5.1).

5.2.2 UV Melting Experiments

Thermal melting experiments were performed on a Varian Cary 300 spectrophotometer equipped with a thermoelectrically-controlled multicell holder. Samples were prepared in a buffer containing 10 mM Tris-HCl (pH 7.4), 0.1 mM Na/Li_2EDTA, and different concentrations of KCl/LiCl. The solutions were incubated at 95 °C for 5 min and cooled to 15 °C at a rate of 1 °C /min. Subsequently, the annealed samples were incubated at 15 °C for 5 min, and a heating ramp was applied at 1 °C/min up to 95 °C. Melting curves were generated by monitoring absorbance values at 260 nm (for hetero-duplexes) and 295 nm (for quadruplexes) every 0.5 °C. Where possible, the melting temperature (T_m) was determined from a first-derivative plot of the respective melting curve. Each reported melting temperature is the average of three independent experiments.

5.2.3 Circular Dichroism (CD) Spectropolarimetry

CD spectra were obtained on a Jasco J-715 circular dichroism spectropolarimeter equipped with a water-circulating temperature controller. Samples were prepared in a buffer containing 10 mM Tris-HCl (pH 7.4), 0.1 mM Na/Li_2EDTA, and different concentrations of KCl/LiCl. The samples were incubated at 95 °C for 5 min and cooled slowly to room temperature. All spectra were collected at 37 °C. Each spectrum was the average of 6 consecutive scans (200 – 360 nm) collected

at a scan rate of 100 nm/min and baseline corrected. The CD spectra for 3 different samples were collected and averaged.

5.2.4 Surface Plasmon Resonance (SPR) Experiments

All SPR experiments were performed on a Biacore T100 instrument equipped with a four-channel CM5 sensor chip (GE Health Care). The sensor chip was coated with a carboxylmethyl dextran matrix that allows further functionalization with streptavidin via a standard NHS-EDC coupling procedure. Covalent immobilization of streptavidin was continued until 7000 response units (RU) of the protein were captured on each of the four channels (flow cells). The final step of the sensor design involved non-covalent immobilization of a biotinylated capture strand on the streptavidin-modified surface.

The sensor surface was primed by a buffer injection [100 mM KCl/LiCl, 10 mM Tris-HCl (pH 7.4), 3 mM Na/Li_2EDTA, and 0.005 % v/v P-20 surfactant] prior to all experiments. For the competition assay[23], 1000 RU of a DNA capture strand was immobilized on the surface of the chip. The flow cell containing the capture strand was subsequently calibrated for the free PNA concentration by injecting a series of solutions containing 10 – 150 nM of the free probe for 100 s. The slopes for the sensorgrams were obtained over a 10 s window (beginning 30 s post-injection) and plotted against the concentration of free PNA to obtain a linear regression curve. (Where the quality of the sensorgrams precluded such analysis, the sensor response at 60 s was used to construct the calibration curve.)

A fixed concentration of each PNA probe was then incubated with increasing concentrations of the pre-annealed RNA target at 25 °C for 2 hours prior to injection over the sensor surface. The slopes for the resulting sensorgrams were then used to obtain the amounts of free PNA in

equilibrium with the hybrid using the previously-generated regression curve (equation 1). Further, the concentration of bound PNA was calculated by assuming that the system obeys the mass conservation condition (equation 2). Finally, the fraction of probe molecules bound by each RNA was determined from equation 3.

$$[PNA]_{free} = \frac{Slope_{40-50\,s}}{M} \quad \boldsymbol{Eqn\ 1}$$

$$M = slope\ of\ calibration\ curve$$

$$[PNA]_{Total} = [PNA]_{free} + [PNA]_{bound} \quad \boldsymbol{Eqn\ 2}$$

$$Fraction_{\ bound} = \frac{[PNA]_{bound}}{[PNA]_{Total}} \quad \boldsymbol{Eqn\ 3}$$

5.3 RESULTS

5.3.1 Target Selection and Probe Design

The target studied here (**PM RNA**, Chart 5.1) is derived from part of a G-rich RNA sequence present in the 5'-untranslated region of the human vascular endothelial growth factor (VEGF) mRNA. The full length sequence is capable of adopting multiple two-tetrad intramolecular quadruplex conformations, each of which show disparate stimulatory effects on cap-independent translation initiation[35]. Also, quadruplex formation by this sequence modulates both the recruitment of- and binding affinity to the 40 S ribosomal subunit[36]. **PM RNA** features only the last two G-tracts of this biologically-relevant motif and thirteen nucleotides, which do not participate in quadruplex formation, but are immediately adjacent to the G-tracts in the full-length mRNA. This change was made to simplify both design and binding analysis of the bivalent PNA probe. We also examined hybridization to a control target (**MM RNA**, Chart 5.1) that features two G-tracts on its 5'end, but has a different sequence of adjacent nucleotides.

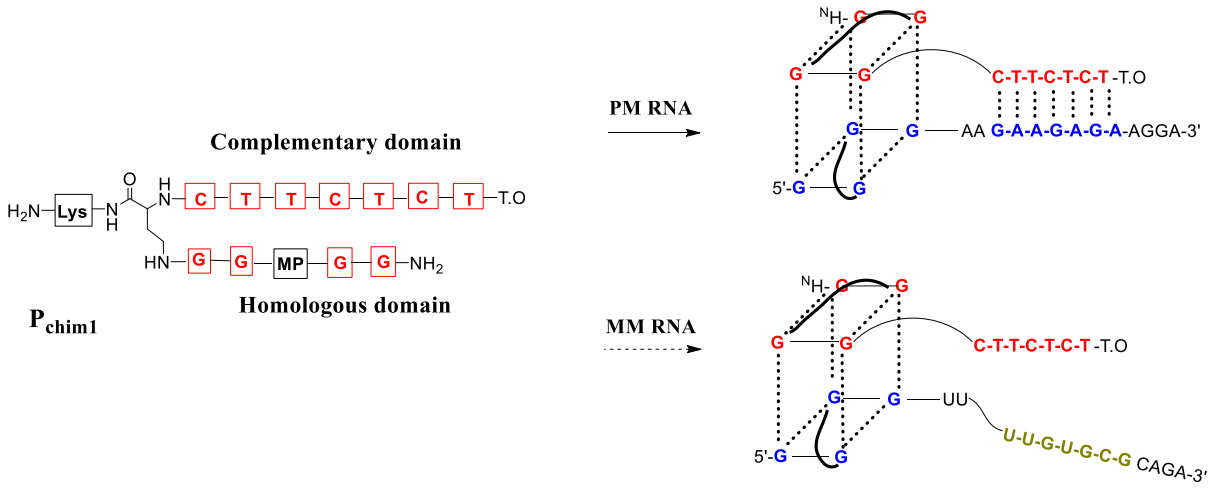
Chart 5.1: Sequences of PNA and RNA oligomers used in this study. (RNA sequence written 5'-3', and PNA written C-N. All PNA molecules have C-terminal lysines. Underlined Gs in RNA and PNA are predicted to participate in hybrid G tetrads. dab = diamino butyric acid.)

Oligomer	Sequence
PM RNA	GGAGGAAGAAGAGAAGGA
MM RNA	GGCGGUUUUGUGCGCAGA
P_{chim1}	GGegGG-dab-CTTCTCT-TO
P_{comp}	dab-CTTCTCT-TO
P_{homo}	GGegGG-dab
P_{chim2}	GegGegGG-$(eg)_2$-dab-eg-CTTCTCT-lys
DNA capture strand	ATTACCCCCCAAGAAGAGAATTA

The chimeric PNA probe (**P_{chim1}**, Chart 5.1) featured two different PNA binding domains tethered by an abasic linker. Both domains were expected to hybridize to **PM RNA** using different recognition modes (Scheme 5.1)—and, if designed properly, would be co-dependent on each other, whereby binding by one domain would be requisite for a successful interaction on the adjacent domain. The two different domains of **P_{chim1}** were designed to preserve the orientational preferences previously reported for hybrid duplexes and quadruplexes. (3'-end to N-terminus for hetero-duplexes[34] and 5'-end to N-terminus for hetero-quadruplexes[22].) A fluorogenic dye, thiazole orange was also conjugated to the probe to enable facile synthesis (*vide infra*). An abasic unit was inserted in one of the G_2 tracts in the homologous domain of the probe to create **P_{chim2}** (Chart 5.1). This modification was made to improve the probe's discrimination against an off-target sequence (discussed later).

Further, whereas hybridization to **PM RNA** was expected to engage both recognition domains of the PNA probe, the corresponding reaction with **MM RNA**—if it occurred—was expected to be mediated only by the homologous domain of **P_{chim1}** (Scheme 5.1). Control oligomers capable of forming either heteroduplex (**P_{comp}**) or heteroquadruplex (**P_{homo}**) structures, but not both, with the **PM RNA** were also tested to examine the advantage conferred by chimeric binding to the thermal stability of the hybrids and/or specificity of binding.

Scheme 5.1

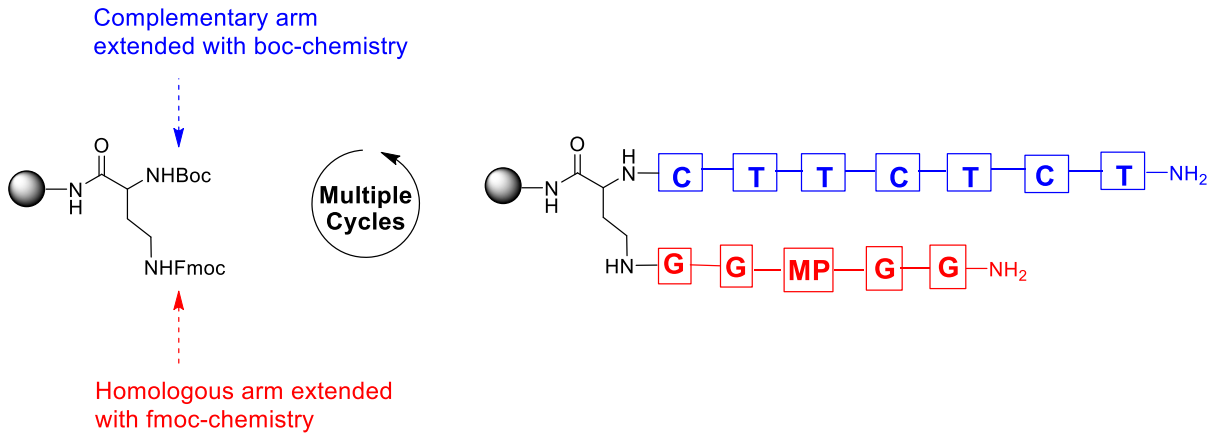


Hybridization of $\mathbf{P_{chim1}}$ using either complementary or homologous domains, or both

5.3.2 Synthesis of Chimeric PNA Oligomers

We initially surmised that the chimeric probes could be obtained using an orthogonal synthetic strategy that allowed the complementary and homologous domains to be concatenated to the boc- and fmoc-amino groups (respectively) of an orthogonally-protected unnatural amino acid scaffold (Scheme 5.2). However the challenges of fmoc-deprotection and subsequent oligomer extension on the homologous domain mitigated progress via this approach. Therefore subsequent efforts were directed towards introducing synthetic and structural modifications that eliminated the need for fmoc-protected monomers and also improved the coupling efficiency of building blocks to the deprotected amino groups.

Scheme 5.2



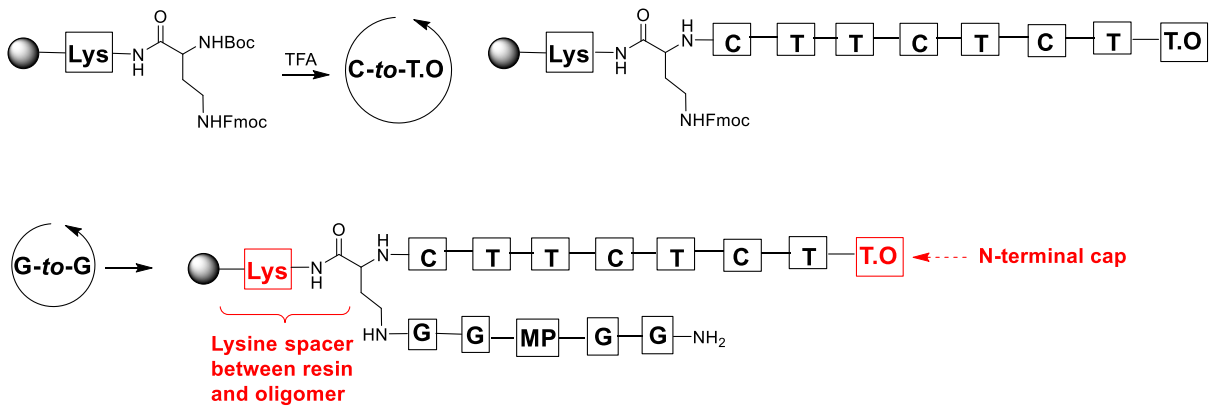
Potential synthetic route to chimeric PNA probe

The synthetic strategy outlined hitherto aimed to concatenate the complementary and homologous arms to the amino acid scaffold using boc- and fmoc- synthetic protocols exclusively (and respectively). We predicated this strategy on the orthogonality of both protecting groups that allowed oligomer extension exclusively on the unveiled amino group of the preceding monomer. However, we surmised that extension of the homologous arm could be accomplished using boc-chemistry if an acid-sturdy group was used to terminate the complementary domain. This N-terminal 'cap' would allow extension of the homologous arm (via boc-chemistry) without unwanted deprotection of the terminating group on the complementary arm (Scheme 5.3).

The second modification involved the incorporation of a lysine residue on the scaffold intervening between both arms of the chimeric probe. This structural modification installs the natural amino acid in order to increase the separation between the growing oligomer and the solid-phase resin. Decreased proximity to the resin was expected to minimize the possibility of hydrophobic

collapse—a phenomenon that has been reported to inhibit removal of protecting groups and/or restrict access to unveiled reactive moieties on the oligomer[37] (Scheme 5.3).

Scheme 5.3



Modified synthetic route to chimeric probe

5.3.3 Biophysical Characterization of PNA-RNA Chimera

5.3.3.1 UV-Melting

In order to determine the contribution from each recognition domain of the probe to the thermal stability of the complex formed with **PM RNA**, we monitored the melting transitions for samples containing the chimeric (or truncated) probe and target strand. (Characterization of the hybrids formed with **MM RNA** was not performed, but will be required to fully understand how the absence of one binding site on the target affects the thermal stability of the PNA-RNA complex.) Melting curves generated for samples containing $\mathbf{P_{chim1}}$ showed a hyperchromic transition (Figure 5.1, red circles) indicating the formation of a stable duplex with the target RNA. Interestingly, similar experiments performed with the truncated probe expected to form only a duplex with the target ($\mathbf{P_{comp}}$) showed similar melting transitions (Figure 5.1, blue squares).

One possible explanation for this result is that recognition by the complementary domain of the probe is only minimally dependent on the presence of a contributing heteroquadruplex substructure. However, extensive interpretation of this data is limited by the absence of comprehensive structural information on the PNA-RNA chimera. It is possible that a heteroquadruplex structure is absent in the **P_{chim1}**-RNA complex, or that such structure—if it exists, is formed by a second equivalent of the probe. Either of these outcomes would eliminate the potential contribution of heteroquadruplex formation to the thermal stability of the accompanying heteroduplex.

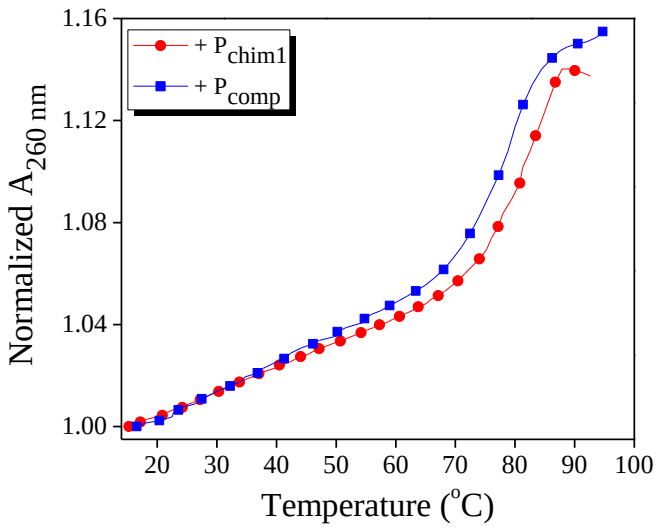


Figure 5.1: Biophysical characterization of PNA-RNA hybrids. UV-melting curves (260 nm) for PM RNA in the presence of P_{chim1} (red circles) and P_{comp} (blue squares). All samples contained 2.5 µM of both PNA and RNA, and were buffered in 100 mM KCl, 10 mM Tris-HCl (pH 7.4), and 0.1 mM Na_2EDTA

The contribution of the heteroquadruplex structure to the thermal stability of the chimeric hybrid was assessed by examining the complex formed between **P_{homo}** and **PM RNA**. Melting of the

P_{homo}-RNA heteroquadruplex elicited a hypochromic transition at a lower temperature (ΔT_m = -10 °C, Figure 5.2A) than that observed for the **P_{chim1}**-RNA chimeric structure (Figure 5.2B). While this result might suggest a greater dependence of the heteroquadruplex domain on chimeric binding, it is possible that the thermal stability enhancement observed for the chimera could be a direct consequence of effective concentration-mediated modulation in the stability of the complexes. Melting of the quadruplex substructure is expected to precede duplex unfolding (compare Figs. 5.1 and 5.2A). Therefore, to the extent that a heteroquadruplex exists in the **P_{chim1}**-RNA hybrid, its thermally-induced unfolding generates an intermediate where the two strands remain annealed due to the higher stability of the heteroduplex substructure. In contrast, melting of the **P_{homo}**-RNA heteroquadruplex generates two free strands.

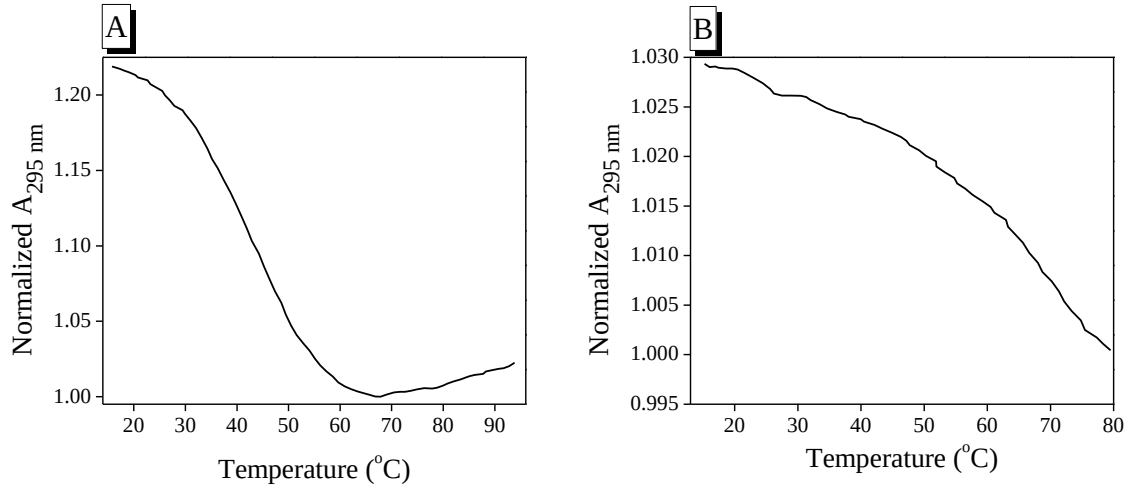


Figure 5.2: Biophysical characterization of PNA-RNA hybrids. UV-melting curves (295 nm) for PM RNA in the presence of P_{homo} (**A**) or P_{chim1} (**B**). All samples contained 2.5 µM of both PNA and RNA, and were buffered in 100 mM KCl, 10 mM Tris-HCl (pH 7.4), and 0.1 mM Na_2EDTA.

5.3.3.2 Circular Dichroism

Additional information on the nature of the interaction between the **PM RNA** target and PNA probes was obtained from CD spectra for the respective complexes. Samples containing the RNA strand alone produced a spectrum with an exciton band between 250 nm and 290 nm that reflects the helical organization of the target strand (Figure 5.3, open circles). Upon addition of the truncated homologous probe, $\mathbf{P_{homo}}$, negative and positive peaks appear at 242 nm and 265 nm, respectively (Figure 5.3, green triangles), indicating the formation of a PNA-RNA parallel quadruplex[18]. In the presence of the truncated complementary probe, $\mathbf{P_{comp}}$, a minimum and a maximum appear at 250 nm and 270 nm, respectively (Figure 5.3, blue squares), providing supporting evidence for the formation of an antiparallel PNA-RNA duplex[28].

The CD spectrum for samples containing equimolar amounts of the RNA target and chimeric probe ($\mathbf{P_{chim1}}$) is less lucid, since the minimum and maximum at 250 nm and 275 nm, respectively are signature peaks for both PNA-RNA antiparallel duplexes[28] and parallel quadruplexes[18, 22] (Figure 5.3, red circles). Taken together, while these data confirm that the isolated probe domains form the expected structures with the RNA target, the nature of their interactions with the target, once incorporated into the chimeric probe cannot be determined by this method.

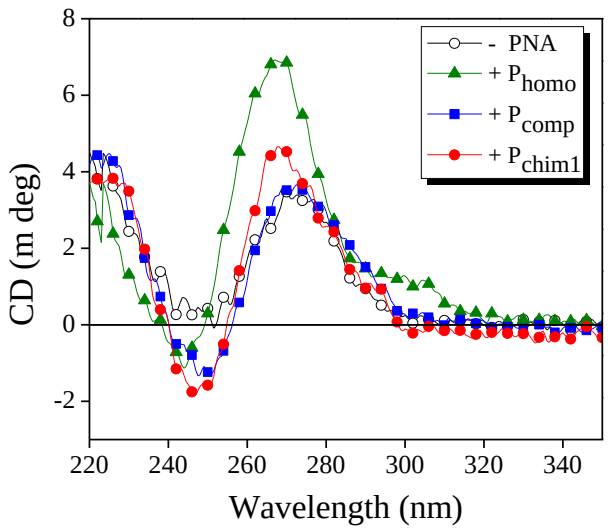


Figure 5.3: Biophysical characterization of PNA-RNA hybrids. CD spectra recorded for the PM RNA alone (open circles) and in the presence of P_{chim1} (red circles), P_{comp} (blue triangles), or P_{homo} (green squares).

5.3.4 Selectivity through Chimeric Recognition

We next utilized SPR competition experiments[23] to examine the ability of $\mathbf{P_{chim1}}$ to discriminate between **PM RNA** and the control target (**MM RNA**) possessing a scrambled sequence on the heteroduplex domain. In this experiment, a DNA capture strand immobilized on the SPR chip was used to quantify the amount of free $\mathbf{P_{chim1}}$ present in a mixture of the probe and either RNA target. (A calibration curve correlating the sensor response with free probe concentration is presented in Figure D1.)

Figure 5.4 presents results for the titration of $\mathbf{P_{chim1}}$ with increasing concentrations of either RNA target. (Sensorgrams are presented in Figure D2.) While probe consumption by **PM RNA** (black squares) was more pronounced and proceeded to a higher level than with **MM RNA** (red circles),

we observed that saturation in the binding isotherm for **P_{chim1}** occurs at similar concentrations of either RNA target (Figure 5.4), suggesting that this probe binds with similar affinities to both RNAs.

Additionally, hybridization by either RNA target did not exceed 80 % of the total **P_{chim1}** molecules present in solution (80 % for **PM RNA**; 50 % for **MM RNA**, Figure 5.4). While it is possible that the RNA targets lacked sufficient affinity to hybridize a higher fraction of free probe molecules, we have not ruled out the additional possibility that the RNA targets themselves bind to the immobilized DNA capture strand, thus occluding the binding sites for free probe molecules. In this event, low concentrations of free probe—likely to be the case at high [RNA]—would be difficult to detect if they cannot displace the bound RNA molecules. SPR direct-binding experiments to evaluate the inertness—or lack thereof— of the RNA targets to the immobilized DNA capture strand might yield insight on the extent to which inadvertent RNA binding to the capture strand impedes probe detection.

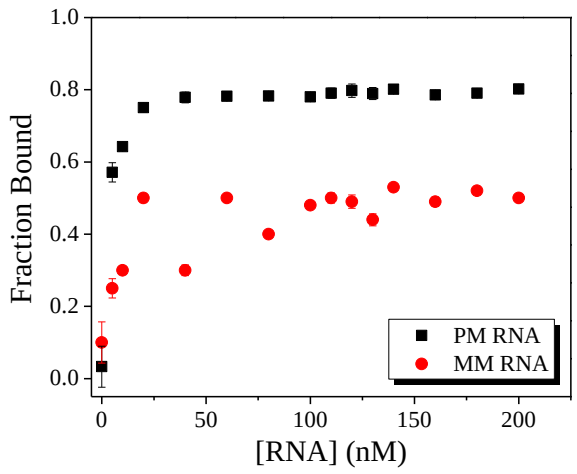


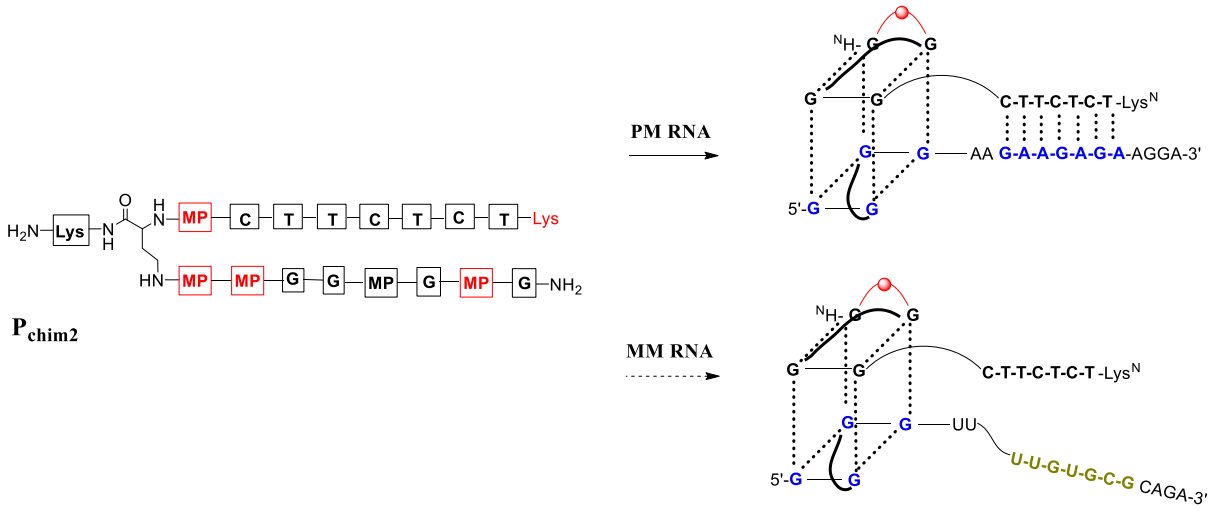
Figure 5.4: Selectivity of P_{chim1}. Titration of 120 nM P_{chim1} with 0 – 200 nM PM RNA (black squares) or MM RNA (red circles).

5.3.5 Structural Modifications to Improve Sequence Discrimination

The relatively high level (50 %) of binding to the control sequence prompted us to examine if additional modifications could be made in the probe to limit its interaction with **MM RNA**. Close inspection of this RNA sequence revealed that, although duplex formation with the probe was unlikely, the heteroquadruplex could be sufficiently stable to keep the **MM RNA** annealed to the probe. We therefore introduced an abasic residue (ethylene glycol) within one of the G_2 motifs on the homologous domain of the probe (**P_{chim2}**, Scheme 5.4). Successful binding by this modified domain would necessitate extrusion of the intervening residue away from the quadruplex core to form a bulge. It was not initially known whether this modification would entirely scuttle heteroquadruplex formation, since a thorough search of the literature revealed no examples of bulge-containing PNA-RNA/DNA heteroquadruplexes. However, bulges within G_3 tracts are tolerated well by DNA intramolecular quadruplexes[14]. Therefore we expected this modification to be tolerated in the PNA/RNA hybrid quadruplex, provided the G residues in the probe were still presented at appropriate intervals to form tetrads with the target.

Two additional modifications were introduced in **P_{chim2}**: (1) we increased the separation between the homologous and complementary domains using ethylene glycol residues, in addition to the amino acid scaffold. This modification was expected to alleviate any steric limitations that might preclude synergistic binding by the probe. (2) The thiazole orange cap was replaced with a lysine residue to improve the solubility of the probe (Scheme 5.4).

Scheme 5.4



P_{chim2}

Hybridization of $\mathbf{P_{chim2}}$ using either complementary or homologous domains, or both

To examine the ability of $\mathbf{P_{chim2}}$ to discriminate between the **PM** and **MM** RNA targets, SPR competition experiments were again used to analyze the binding of the probe to the different RNA oligomers. Figure 5.5 presents results from titration of $\mathbf{P_{chim2}}$ with increasing RNA concentrations. (Sensorgrams for calibration/competition and calibration curve are presented in Figure D3-D4. The curves for $\mathbf{P_{chim2}}$ are much clearer than those for $\mathbf{P_{chim1}}$, presumably due to the improved aqueous solubility of the former.) We observe that $\mathbf{P_{chim2}}$ is consumed much more readily by **PM RNA** (black squares) than by the control sequence (red circles). As with $\mathbf{P_{chim1}}$ (Figure 5.4) we observe saturation in binding before molar equivalence is reached by **PM RNA**. While it is likely that RNA binding to the capture strand precludes detection of low concentrations of free probe, this sub-stoichiometric saturation might suggest that the chimeric hybrid is not a binary complex. Future work on this project should therefore seek methods to unambiguously verify the binding stoichiometry of the chimeric structure. Importantly, we observe greater discrimination between the target and control RNA strands with $\mathbf{P_{chim2}}$ than with $\mathbf{P_{chim1}}$.

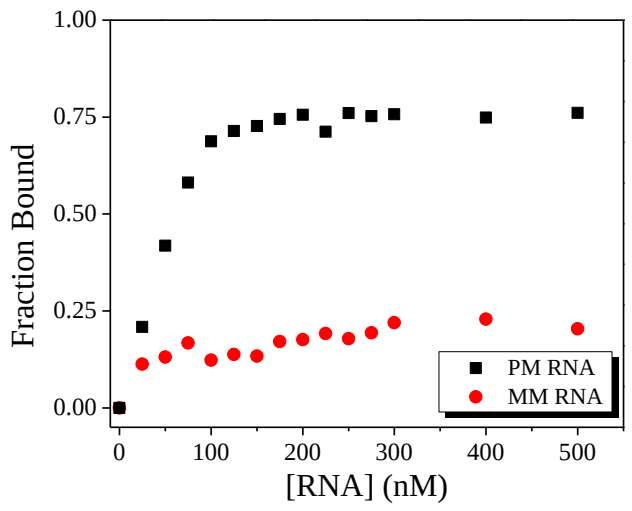


Figure 5.5: Selectivity of P_{chim2}. Titration of 170 nM P_{chim2} with 0 – 500 nM PM RNA (black squares) or MM RNA (red circles).

However, much remains to be learned about the mode of recognition adopted by $\mathbf{P_{chim2}}$. For instance, the specificity reported above is still attainable if hybridization is mediated by only the complementary domain of the probe. The incorporation of the ethylene glycol unit in one of the G_2 tracts of $\mathbf{P_{chim2}}$ was meant to destabilize the hetero-quadruplex domain, and increase its dependence on the flanking hetero-duplex. However, this structural modification has not been shown to be tolerated by isolated hetero-quadruplexes. Therefore additional characterization would be required to appraise the contribution of the homologous domain to overall target recognition.

5.4 DISCUSSION

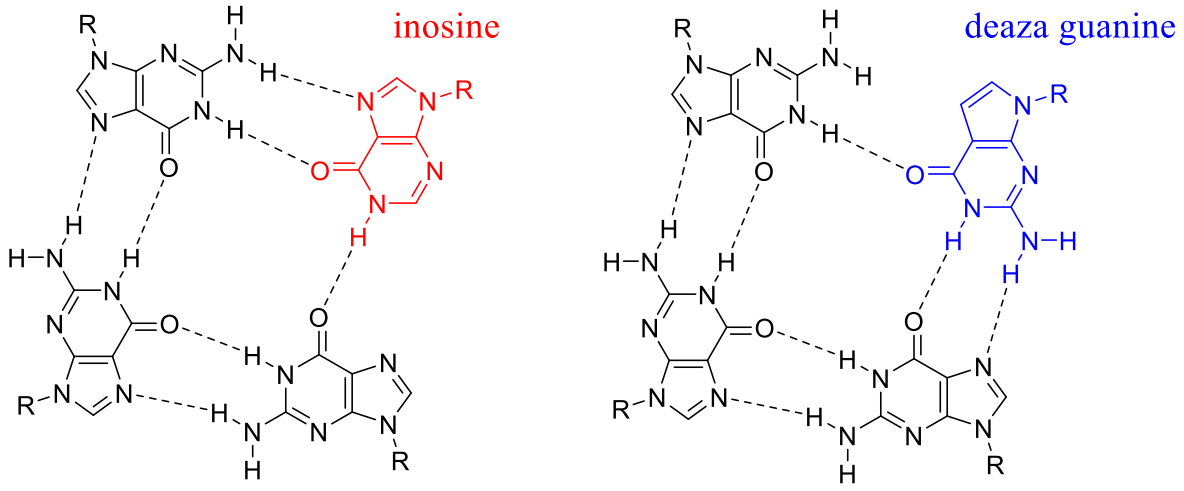
The results presented here show that a PNA oligomer possessing both homologous and complementary domains can form a stable hybrid with a corresponding RNA target presenting the appropriate binding sites. While hybridization has been verified for truncated probes representing the isolated binding domains, the contribution of each substructure to the overall binding energy of the chimeric probe has not been evaluated. Importantly, the PNA probe shows discrimination in recognition against an RNA target lacking one of the requisite binding sites—an effect that is improved by structural modifications intended to destabilize the interactions tethering the probe to the deficient RNA target.

The destabilizing modifications were made on the homologous domain of the probe, whereby an abasic residue within a G_2-tract would be extruded as a bulge in the resulting chimeric hybrid. To the extent that this modification is tolerated in the resulting heteroquadruplex substructure, it would be expected to decrease the contribution of this domain to the overall binding energy of the hybrids formed with either RNA target, thus increasing the probe's dependence on the complementary domain.

This enforced co-dependency is likely to be useful in achieving selective targeting of G-quadruplex motifs using PNA oligomers possessing homologous and complementary domains. The high affinities of even isolated (homologous[23] or complementary[18], only) PNA probes for QFS motifs make it likely that an inappropriately designed chimeric PNA would provide sufficient binding energy (on one domain) to tether the probe to an off-target sequence presenting only one of the requisite 'binding sites'. To circumvent this potential outcome, hybridization by the chimeric PNA to its target should depend on the simultaneous engagement of both binding domains in the resulting complex.

Although we have attempted to implement this strategy by incorporating abasic, bulging residues in the homologous domain, alternative strategies are possible. For example, inosine and deazaguanine residues could be incorporated in the homologous domain to deprive the hybrid tetrad of hydrogen bond donor and acceptor groups, respectively (Scheme 5). This modification might induce local destabilization of the tetrad that, if introduced at strategic positions—could propagate through the entire heteroquadruplex substructure. Additionally, truncation of the complementary domain might decrease the contribution of the heteroduplex substructure to the binding energy, thus increasing its dependence on the homologous domain. The extents to which these different 'levers' would have to be adjusted will likely depend on the structural features of the target, such as the number of G-tetrads and accessibility of its adjacent nucleotides.

Scheme 5.5



Incorporation of Inosine and deazaguanine residues to destabilize hybrid tetrad

Our results can also be compared to other strategies and/or demonstrations of chimeric recognition of QFS motifs previously reported in the literature. Paul *et al* have reported that a PNA-acridone conjugate was able to form a stable heteroquadruplex with a G-rich DNA target bearing three G_3

repeats[38]. The authors presented a binding model showing that each G residue of the PNA participated in hybrid tetrads formed with those of the DNA target, with stacking of the conjugated acridone ligand occurring on the terminal tetrad.[38] Further, the free energy for complex formation appeared to be contributed by both H-bonding (from PNA) and end-stacking (from acridone), since the PNA—absent the acridone moiety, formed a less stable hybrid with the target, while the unconjugated small molecule did not perturb the folding pattern or thermal stability of the target[38]. Interestingly, while the chimeric probe was able to assemble the hybrid quadruplex from an unfolded strand, intermolecular quadruplex formation by the DNA precluded binding by the probe, presumably because the binding energy supplied by H-bonding and end-stacking was insufficient to compensate for the energy investment required to displace the DNA strand supplying the last G_3 tract[38].

The limitation observed in the aforementioned report holds an important lesson for future work on this project: while our chimeric probes target a truncated RNA target predicted to be absent a stable quadruplex fold, it is possible that the design requirements for achieving simultaneous engagement of the two recognition domains will be affected by the stability of the quadruplex structure present in the target. For example, a more stable quadruplex might require the binding energy for our chimeric probe to be weighted more heavily on the homologous domain in order to facilitate quadruplex invasion. However, disproportionate localization of the binding affinity on one domain might compromise the probe's selectivity against off-targets, since one binding site might suffice for recognition.

In another example, Bhattacharyya *et al* have provided evidence that a DNA oligomer possessing both homologous and complementary domains was able to bind to an RNA target presenting the appropriate binding sites[39]. Evidently, hybridization was mediated by the formation of an RNA-

DNA hybrid quadruplex and duplex by the homologous and complementary domains, respectively, of the DNA probe[39]. Further, the probe was shown to induce specific repressive effects on translation when directed against an endogenous gene presenting the binding sites for the DNA in its 5-untranslated or coding regions, or both[39].

However, while the biochemical and biological effects of this chimeric probe are clear, the binding model to the target RNA is nebulous: the authors present biophysical evidence for formation of a parallel folded RNA-DNA hybrid quadruplex[39]. In contrast, their binding model aligns the 3'-end of the probe (containing the homologous domain) with the 5'-end of the RNA target[39]. Further, both chemical and enzymatic foot-printing data showed that quadruplex destabilization in LiCl did not trigger unfolding of the putative adjacent duplex structure, suggesting that the binding energy is disproportionately weighted on the complementary domain[39]. These contradictions therefore limit our ability to extrapolate useful guidelines from this report for improving our design of the chimeric PNA probes.

In conclusion, the chimeric probes presented in this chapter show discrimination against an RNA target lacking one of the requisite binding sites for the probes. The extent to which this selectivity is mediated by the simultaneous engagement of both domains of the same probe molecule has not been examined. However, the improvement in selectivity recorded upon putative destabilization of one of the recognition domains suggests that such synergistic effects can be measured by appropriate modifications to the probe. The ability of the probes to discriminate between target sequences capable of forming intramolecular quadruplexes is currently unknown, and should be a subject of future studies. It will also be interesting to examine the ability of these probes to induce effects on translation that are due to recognition of a specific quadruplex structure.

5.5 REFERENCES

[1] Bochman, M. L., Paeschke, K., and Zakian, V. A. (2012) DNA secondary structures: stability and function of G-quadruplex structures, *Nat. Rev. Genet. 13*, 770-780.

[2] Bugaut, A., and Balasubramanian, S. (2012) 5'-UTR RNA G-quadruplexes: translation regulation and targeting, *Nucleic Acids Res. 40*, 4727-4741.

[3] Millevoi, S., Moine, H., and Vagner, S. (2012) G-quadruplexes in RNA biology, *Wiley Interdiscip. Rev. RNA 3*, 495-507.

[4] Rhodes, D., and Lipps, H. J. (2015) G-quadruplexes and their regulatory roles in biology, *Nucleic Acids Res. 43*, 8627-8637.

[5] Morris, M. J., and Basu, S. (2009) An unusually stable G-quadruplex within the 5'-UTR of the MT3 matrix metalloproteinase mRNA represses translation in eukaryotic cells, *Biochemistry 48*, 5313-5319.

[6] Kumari, S., Bugaut, A., Huppert, J. L., and Balasubramanian, S. (2007) An RNA G-quadruplex in the 5' UTR of the NRAS proto-oncogene modulates translation, *Nat. Chem. Biol. 3*, 218-221.

[7] Zhang, S., Wu, Y., and Zhang, W. (2014) G-quadruplex structures and their interaction diversity with ligands, *ChemMedChem 9*, 899-911.

[8] Panyutin, I. G., Onyshchenko, M. I., Englund, E. A., Appella, D. H., and Neumann, R. D. (2012) Targeting DNA G-Quadruplex Structures with Peptide Nucleic Acids, *Curr. Pharm. Des. 18*, 1984-1991.

[9] Shin-ya, K., Wierzba, K., Matsuo, K., Ohtani, T., Yamada, Y., Furihata, K., Hayakawa, Y., and Seto, H. (2001) Telomestatin, a novel telomerase inhibitor from Streptomyces anulatus, *J. Am. Chem. Soc. 123*, 1262-1263.

[10] Tauchi, T., Shin-ya, K., Sashida, G., Sumi, M., Okabe, S., Ohyashiki, J. H., and Ohyashiki, K. (2006) Telomerase inhibition with a novel G-quadruplex-interactive agent, telomestatin: in vitro and in vivo studies in acute leukemia, *Oncogene 25*, 5719-5725.

[11] Han, F. X., Wheelhouse, R. T., and Hurley, L. H. (1999) Interactions of TMPyP4 and TMPyP2 with Quadruplex DNA. Structural Basis for the Differential Effects on Telomerase Inhibition, *J. Am. Chem. Soc. 121*, 3561–3570.

[12] Murat, P., Gormally, M. V., Sanders, D., Di Antonio, M., and Balasubramanian, S. (2013) Light-mediated in cell downregulation of G-quadruplex-containing genes using a photo-caged ligand, *Chem. Commun. 49*, 8453-8455.

[13] Zhang, A. Y., Bugaut, A., and Balasubramanian, S. (2011) A sequence-independent analysis of the loop length dependence of intramolecular RNA G-quadruplex stability and topology, *Biochemistry 50*, 7251-7258.

[14] Mukundan, V. T., and Phan, A. T. (2013) Bulges in G-quadruplexes: broadening the definition of G-quadruplex-forming sequences, *J. Am. Chem. Soc. 135*, 5017-5028.

[15] Guedin, A., Gros, J., Alberti, P., and Mergny, J. L. (2010) How long is too long? Effects of loop size on G-quadruplex stability, *Nucleic Acids Res. 38*, 7858-7868.

[16] Datta, B., and Armitage, B. A. (2001) Hybridization of PNA to Structured DNA Targets: Quadruplex Invasion and the Overhang Effect, *J. Am. Chem. Soc. 123*, 9612–9619.

[17] Marin, V. L., and Armitage, B. A. (2005) RNA Guanine Quadruplex Invasion by Complementary and Homologous PNA Probes, *J. Am. Chem. Soc. 127*, 8032–8033.

[18] Marin, V. L., and Armitage, B. A. (2006) Hybridization of Complementary and Homologous Peptide Nucleic Acid Oligomers to a Guanine Quadruplex-Forming RNA, *Biochemistry 45*, 1745–1754.
[19] Gupta, A., Lee, L., Roy, S., Tanious, F. A., Wilson, W. D., Ly, D. H., and Armitage, B. A. (2013) Strand invasion of DNA quadruplexes by PNA: comparison of homologous and complementary hybridization, *ChemBioChem 14*, 1476–1484.
[20] Murphy, C. T., Gupta, A., Armitage, B. A., and Opresko, P. L. (2014) Hybridization of G-quadruplex-forming peptide nucleic acids to guanine-rich DNA templates inhibits DNA polymerase eta extension, *Biochemistry 53*, 5315-5322.
[21] Oyaghire, S. N., Cherubim, C. J., Telmer, C. A., Martinez, J. A., Bruchez, M. P., and Armitage, B. A. (2016) RNA G-Quadruplex Invasion and Translation Inhibition by Antisense gamma-Peptide Nucleic Acid Oligomers, *Biochemistry 55*, 1977-1988.
[22] Datta, B., Schmitt, C., and Armitage, B. A. (2003) Formation of a PNA_2 - DNA_2 Hybrid Quadruplex, *J. Am. Chem. Soc. 125*, 4111-4118.
[23] Roy, S., Tanious, F. A., Wilson, W. D., Ly, D. H., and Armitage, B. A. (2007) High-affinity homologous peptide nucleic acid probes for targeting a quadruplex-forming sequence from a MYC promoter element, *Biochemistry 46*, 10433–10443.
[24] Lusvarghi, S., Murphy, C. T., Roy, S., Tanious, F. A., Sacui, I., Wilson, W. D., Ly, D. H., and Armitage, B. A. (2009) Loop and backbone modifications of peptide nucleic acid improve g-quadruplex binding selectivity, *J. Am. Chem. Soc. 131*, 18415-18424.
[25] Roy, S., Zanotti, K. J., Murphy, C. T., Tanious, F. A., Wilson, W. D., Ly, D. H., and Armitage, B. A. (2011) Kinetic discrimination in recognition of DNA quadruplex targets by guanine-rich heteroquadruplex-forming PNA probes, *Chem. Commun. 47*, 8524-8526.
[26] Kormuth, K. A., Woolford, J. L., Jr., and Armitage, B. A. (2016) Homologous PNA Hybridization to Noncanonical DNA G-Quadruplexes, *Biochemistry 55*, 1749-1757.
[27] Nielsen, P. E., Egholm, M., Berg, R. H., and Buchardt, O. (1991) Sequence-selective recognition of DNA by strand displacement with a thymine-substituted polyamide, *Science 254*, 1497-1500.
[28] Egholm, M., Buchardt, O., Christensen, L., Behrens, C., Freier, S. M., Driver, D. A., Berg, R. H., Kim, S. K., Norden, B., and Nielsen, P. E. (1993) PNA hybridizes to complementary oligonucleotides obeying the Watson–Crick hydrogen-bonding rules, *Nature 365*, 566-568.
[29] Dragulescu-Andrasi, A., Rapireddy, S., Frezza, B. M., Gayathri, C., Gil, R. R., and Ly, D. H. (2006) A Simple γ-Backbone Modification Preorganizes Peptide Nucleic Acid into a Helical Structure, *J. Am. Chem. Soc. 2006*, 10258–10267.
[30] Sahu, B., Sacui, I., Rapireddy, S., Zanotti, K. J., Bahal, R., Armitage, B. A., and Ly, D. H. (2011) Synthesis and characterization of conformationally preorganized, (R)-diethylene glycol-containing gamma-peptide nucleic acids with superior hybridization properties and water solubility, *J. Org. Chem. 76*, 5614-5627.
[31] Qin, Y., and Hurley, L. H. (2008) Structures, folding patterns, and functions of intramolecular DNA G-quadruplexes found in eukaryotic promoter regions, *Biochimie 90*, 1149-1171.
[32] Huppert, J. L., Bugaut, A., Kumari, S., and Balasubramanian, S. (2008) G-quadruplexes: the beginning and end of UTRs, *Nucleic Acids Res. 36*, 6260-6268.

[33] Chambers, V. S., Marsico, G., Boutell, J. M., Di Antonio, M., Smith, G. P., and Balasubramanian, S. (2015) High-throughput sequencing of DNA G-quadruplex structures in the human genome, *Nat. Biotechnol. 33*, 877-881.
[34] Egholm, M., Buchardt, O., Christensen, L., Behrens, C., Freier, S. M., Driver, D. A., Berg, R. H., Kim, S. K., Norden, B., and Nielsen, P. E. (1993) PNA hybridizes to complementary oligonucleotides obeying the Watson Crick hydrogen-bonding rules, *Nature 365*, 566-568.
[35] Morris, M. J., Negishi, Y., Pazsint, C., Schonhoft, J. D., and Basu, S. (2010) An RNA G-quadruplex is essential for cap-independent translation initiation in human VEGF IRES, *J. Am. Chem. Soc. 132*, 17831-17839.
[36] Bhattacharyya, D., Diamond, P., and Basu, S. (2015) An Independently folding RNA G-quadruplex domain directly recruits the 40S ribosomal subunit, *Biochemistry 54*, 1879-1885.
[37] Tam, J. P., and Lu, Y. (1995) Coupling Difficulty Associated with Interchain Clustering and Phase Transition in Solid Phase Peptide Synthesis, *J. Am. Chem. Soc. 117*, 12058–12063.
[38] Paul, A., Sengupta, P., Krishnan, Y., and Ladame, S. (2008) Combining G-quadruplex targeting motifs on a single peptide nucleic acid scaffold: a hybrid (3+1) PNA-DNA bimolecular quadruplex, *Chem. Eur. J. 14*, 8682-8689.
[39] Bhattacharyya, D., Nguyen, K., and Basu, S. (2014) Rationally induced RNA:DNA G-quadruplex structures elicit an anticancer effect by inhibiting endogenous eIF-4E expression, *Biochemistry 53*, 5461-5470.

APPENDIX

APPENDIX A

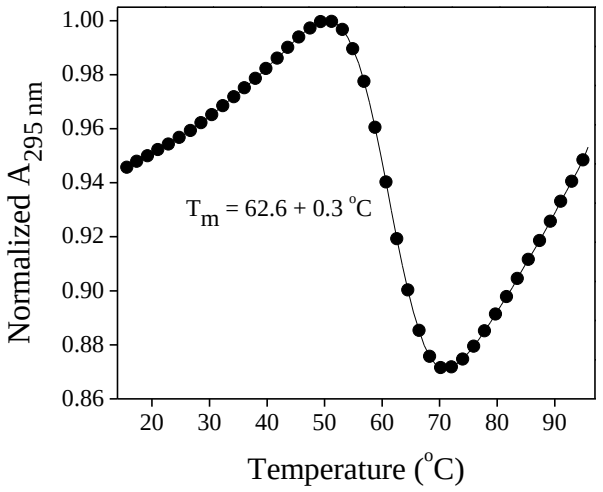


Figure A1. Quadruplex folding is intramolecular. UV-melting curve for 10 µM 4G3 RNA in 1 mM KCl.

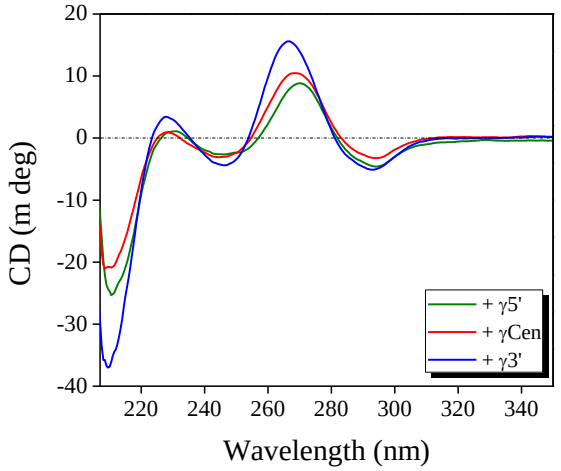


Figure A2. CD spectra of 4G3 RNA-γPNA hybrids. All samples contained 2 µM 4G3 RNA and 2 µM γPNA in 100 mM KCl.

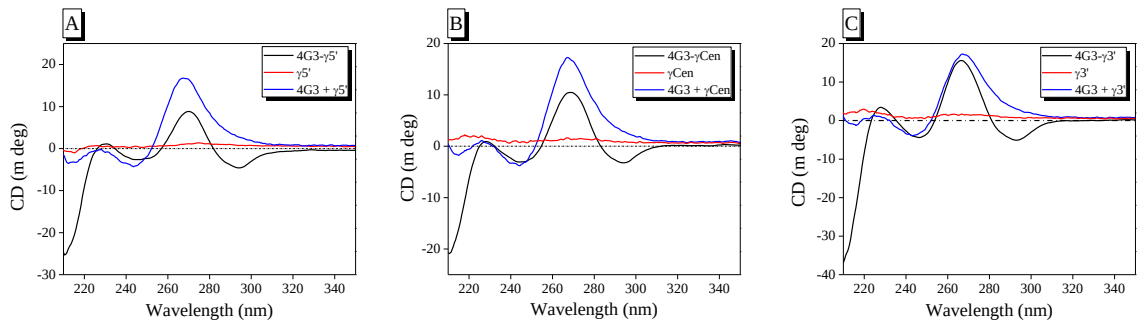


Figure A3. CD spectra for hybrid duplexes formed between 4G3 RNA and γ5' (**A**), γCen (**B**), or γ3' (**C**). Black curves represent spectra for hybrids; red curves represent spectra for γPNAs alone; blue curves represent mathematical aggregate of 4G3 RNA and γPNA spectra.

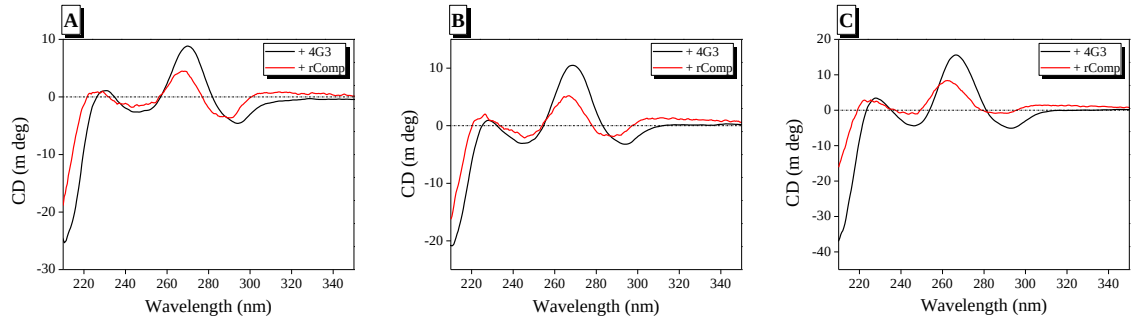


Figure A4. CD spectra for hybrid duplexes formed between 4G3 RNA (black curve) or rComp (red curve) and γ5' (**A**), γCen (**B**), and γ3' (**C**). All samples contained 2 μM γPNA and 2 μM RNA, and were buffered in 100 mM KCl, 10 mM Tris-HCl (pH 7.4), and 0.1 mM Na_2EDTA.

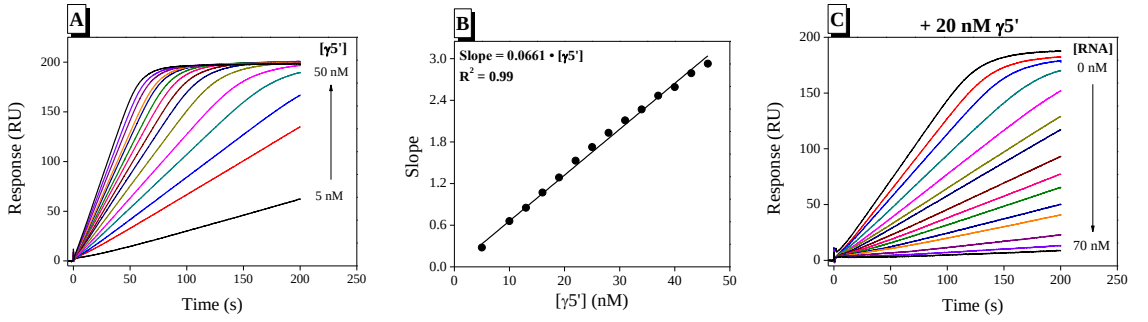


Figure A5. SPR competition experiments for γ5' + 4G3 RNA. Sensorgrams for 5 – 50 nM γ5' (**A**). Linear regression curve for slope against free γ5' concentration (**B**). Sensorgrams for titration of 20 nM γ5' with 0 – 70 nM 4G3 RNA (**C**). All samples were buffered in 100 mM KCl, 10 mM Tris-HCl (pH 7.4), and 0.1 mM Na_2EDTA.

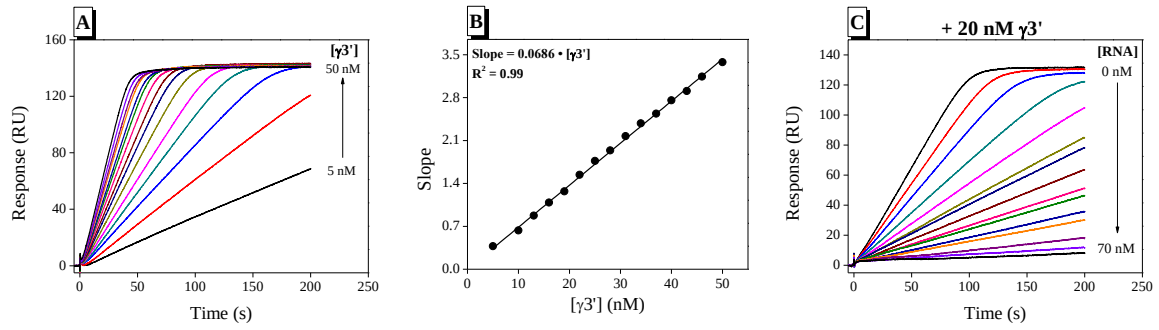


Figure A6. SPR competition experiments for γ3' + 4G3 RNA. Sensorgrams for 5 – 50 nM γ3' (**A**). Linear regression curve for slope against free γ3' concentration (**B**). Sensorgrams for titration of 20 nM γ3' with 0 – 70 nM 4G3 RNA (**C**). All samples were buffered in 100 mM KCl, 10 mM Tris-HCl (pH 7.4), and 0.1 mM Na_2EDTA.

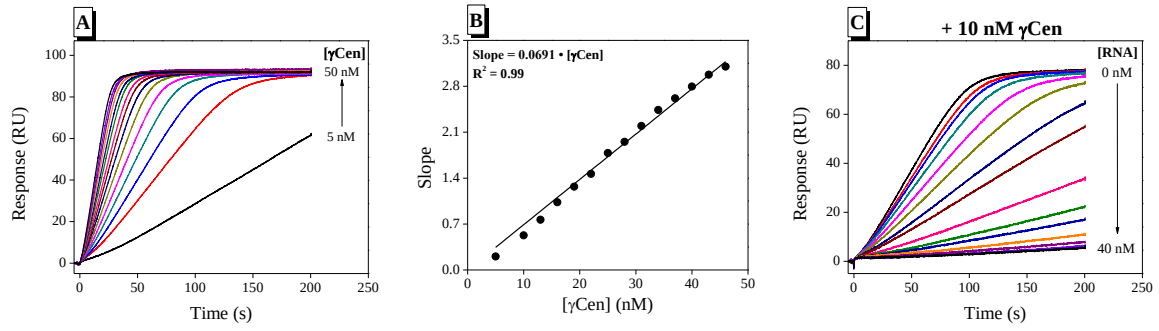


Figure A7. SPR competition experiments for γCen + 4G3 RNA. Sensorgrams for 5 – 50 nM γCen (**A**). Linear regression curve for slope against free γCen concentration (**B**). Sensorgrams for titration of 20 nM γCen with 0 – 70 nM 4G3 RNA (**C**). All samples were buffered in 100 mM KCl, 10 mM Tris-HCl (pH 7.4), and 0.1 mM Na_2EDTA.

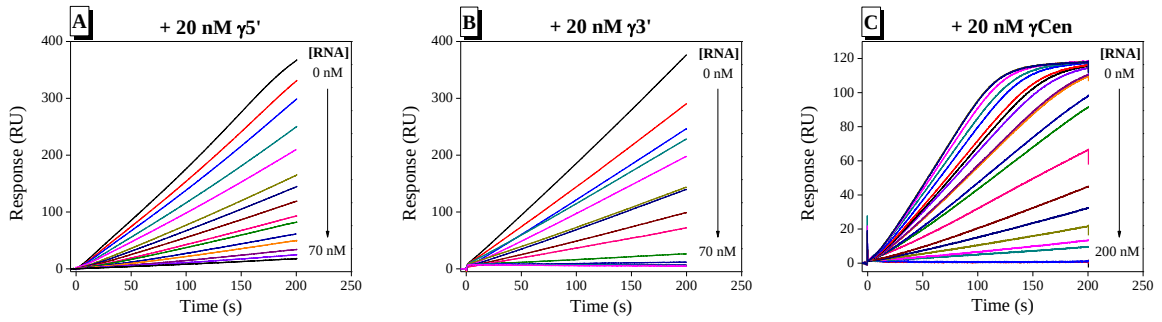


Figure A8. Sensorgrams for titration of γPNAs with rComp targets. 20 nM γ5' + 0 – 70 nM 5'-rComp (**A**). 20 nM γ3' + 0 – 70 nM 3'-rComp (**B**). 20 nM γCen + 0 – 200 nM Cen-rComp (**A**). All samples were buffered in 100 mM KCl, 10 mM Tris-HCl (pH 7.4), and 0.1 mM Na_2EDTA.

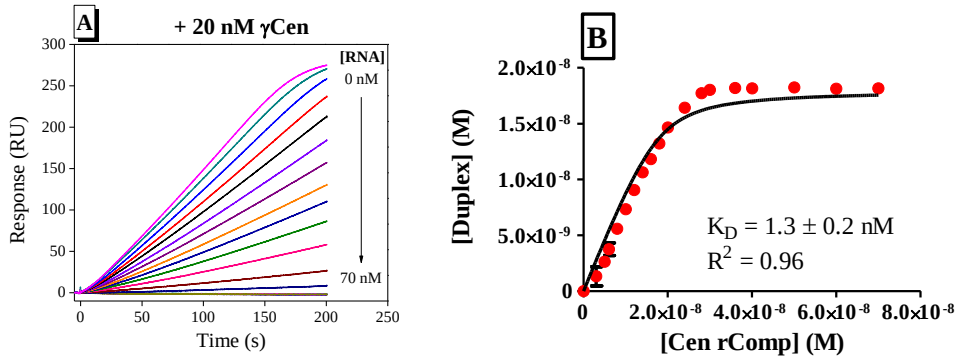


Figure A9. Characterization of γCen + Cen-rComp in 100 mM LiCl. Sensorgrams for titration of 20 nM γCen with 0 – 70 nM Cen-rComp (**A**). Binding isotherm for γCen + Cen-rComp (**B**).

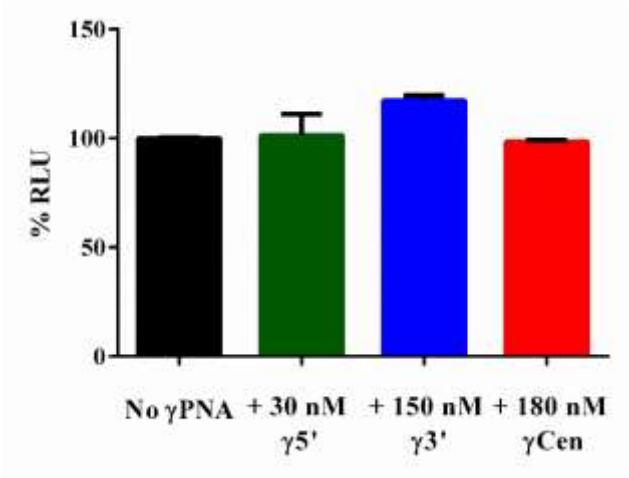


Figure A10. Effect of γPNAs on ribosomal translation efficiency. Cell lysates were incubated with 30 nM γ5' (green) or 150 nM γ3' (blue), or 180 nM γCen (red) prior to addition of control transcript (supplied by Promega).

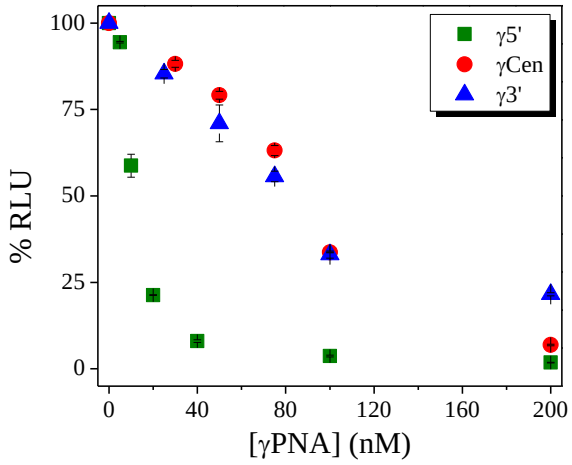


Figure A11. Comparison of antisense inhibition of 4G3 target by γ5' (green squares), γ3' (blue triangles), and γCen (red circles).

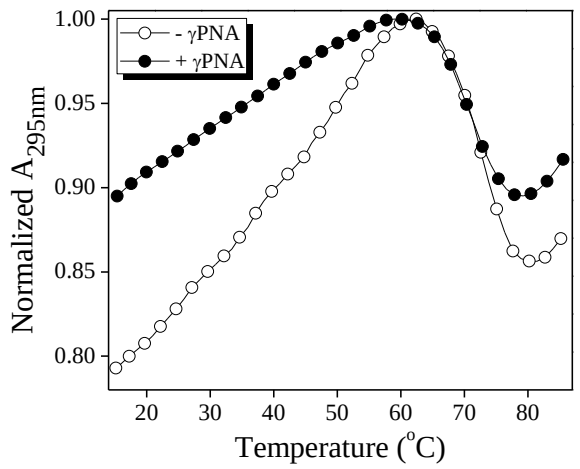


Figure A12. Melting profiles (295 nm) for 2 μM of 4G3 RNA in the absence (open circles) and presence (closed circles) of the γPNA capture strand. Each sample was buffered in 10 mM KCl, 10 mM Tris-HCl (pH 7.4), and 0.1 mM Na_2EDTA.

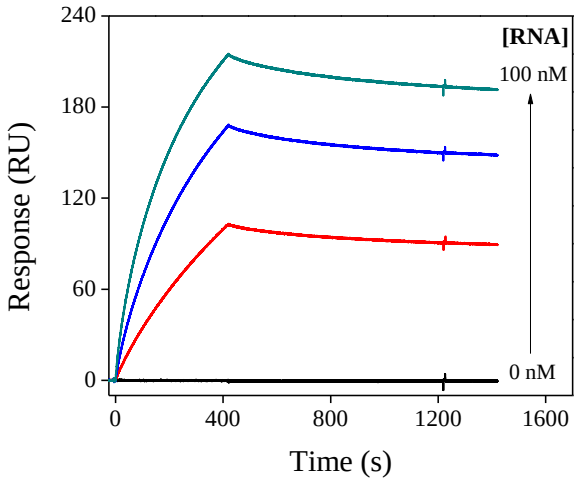


Figure A13. Sensorgrams for (0 – 100 nM) 4G3 RNA binding to γPNA capture strand. Running buffer contained 100 mM KCl, 10 mM Tris-HCl (pH 7.4), and 3 mM Na_2EDTA.

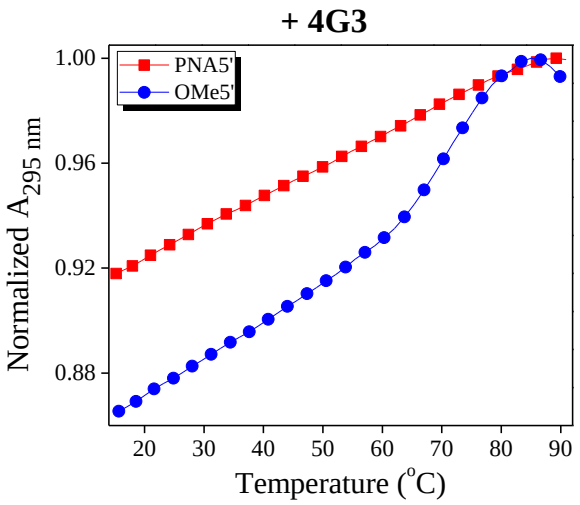


Figure A14. Biophysical characterization of 4G3-PNA5'/OMe5' hybrids. UV-melting curves at 295 nm for 2 µM 4G3 RNA in the presence of 2 µM PNA5' (red squares) or OMe5' (blue circles). Samples contained 100 mM KCl.

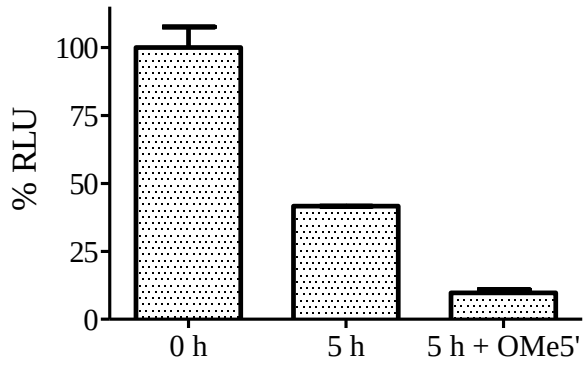


Figure A15. Effect of incubation time on 4G3 mRNA translation. [RNA] = 10 nM; Where present, [OMe5'] = 250 nM. Reactions were incubated at 37 °C for the indicated durations prior to addition to lysate.

APPENDIX B

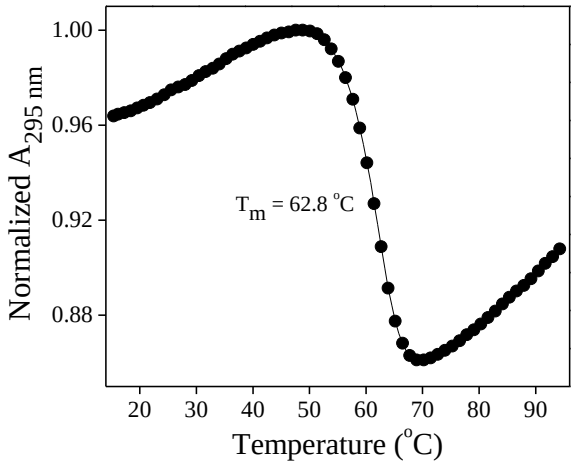


Figure B1. UV-melting curve for 10 µM 4G4 RNA target in 1 mM KCl.

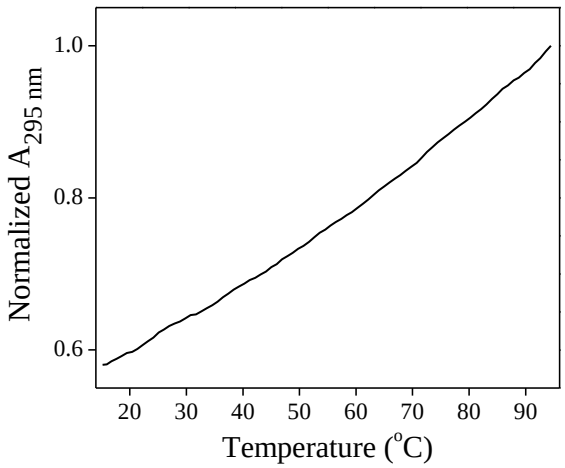


Figure B2. UV-melting curve at 295 nm for 2 µM 4G4 RNA in the presence of 2 µM γ4G4 in 100 mM KCl.

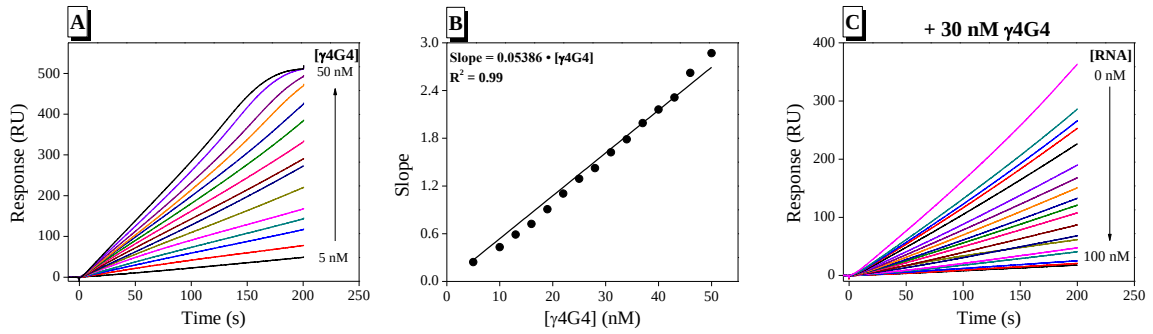


Figure B3. SPR competition experiments for γ4G4 + 4G4 RNA. Sensorgrams for 5 – 50 nM γ4G4 (**A**). Linear regression curve for slope against free γ4G4 concentration (**B**). Sensorgrams for titration of 20 nM γ4G4 with 0 – 100 nM 4G4 RNA (**C**). All samples were buffered in 100 mM KCl, 10 mM Tris-HCl (pH 7.4), and 0.1 mM Na_2EDTA.

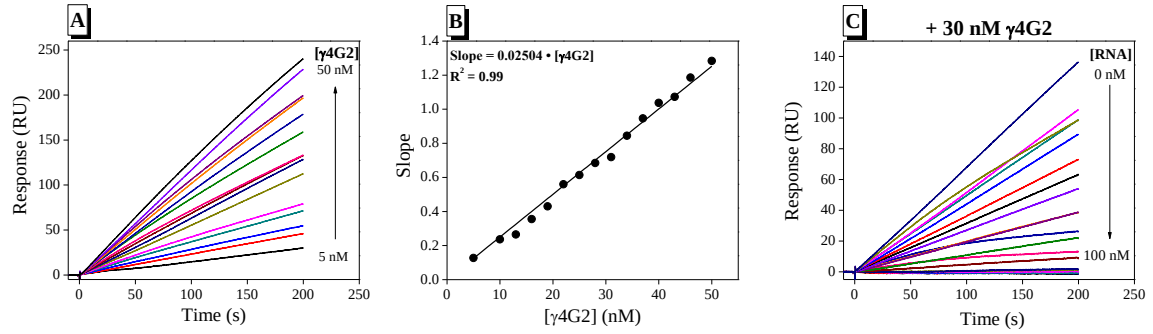


Figure B4. SPR competition experiments for γ4G2 + 4G2 RNA. Sensorgrams for 5 – 50 nM γ4G2 (**A**). Linear regression curve for slope against free γ4G2 concentration (**B**). Sensorgrams for titration of 20 nM γ4G2 with 0 – 100 nM 4G2 RNA (**C**). All samples were buffered in 100 mM KCl, 10 mM Tris-HCl (pH 7.4), and 0.1 mM Na_2EDTA.

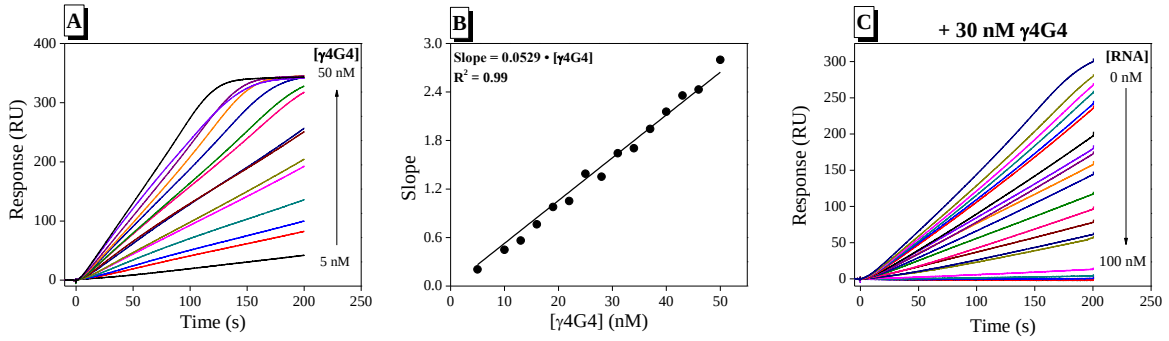


Figure B5. SPR competition experiments for γ4G4 + 4G4 rComp. Sensorgrams for 5 – 50 nM γ4G4 (**A**). Linear regression curve for slope against free γ4G4 concentration (**B**). Sensorgrams for titration of 20 nM γ4G4 with 0 – 100 nM 4G4 rComp (**C**). All samples were buffered in 100 mM KCl, 10 mM Tris-HCl (pH 7.4), and 0.1 mM Na_2EDTA.

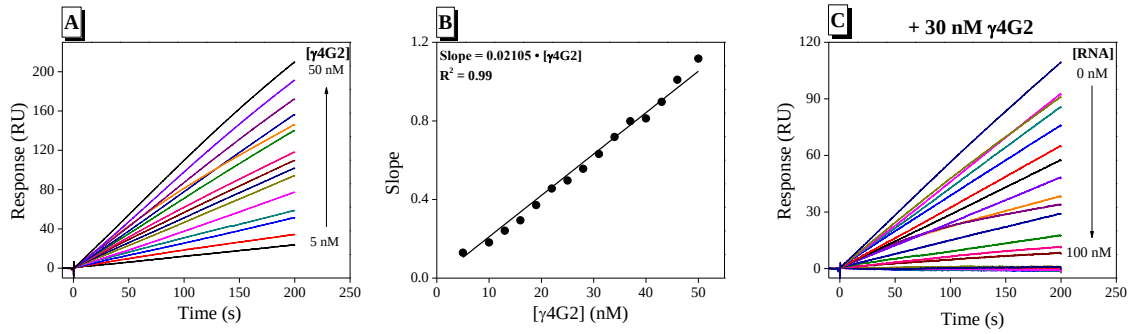


Figure B6. SPR competition experiments for γ4G2 + 4G2 rComp. Sensorgrams for 5 – 50 nM γ4G2 (**A**). Linear regression curve for slope against free γ4G2 concentration (**B**). Sensorgrams for titration of 20 nM γ4G2 with 0 – 100 nM 4G2 rComp (**C**). All samples were buffered in 100 mM KCl, 10 mM Tris-HCl (pH 7.4), and 0.1 mM Na_2EDTA.

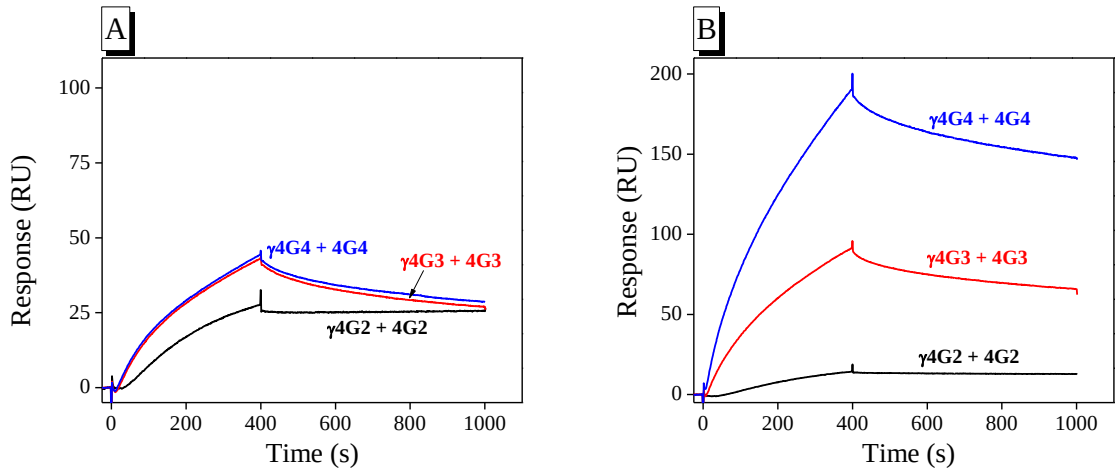


Figure B7. Sensorgrams for hybridization of 50 nM γPNAs to 4Gx RNA targets. γ4G4 + 4G4 (blue curve), γ4G2 + 4G2 (black curve), γ4G3 + 4G3 (red curve). Running buffer contained 10 mM Tris-HCl (pH 7.4), 3 mM Na_2EDTA, and 100 mM KCl (**A**) or 100 mM LiCl (**B**). All binding experiments were performed at 25 °C. Sensorgrams are corrected for the immobilization level of each RNA.

APPENDIX C

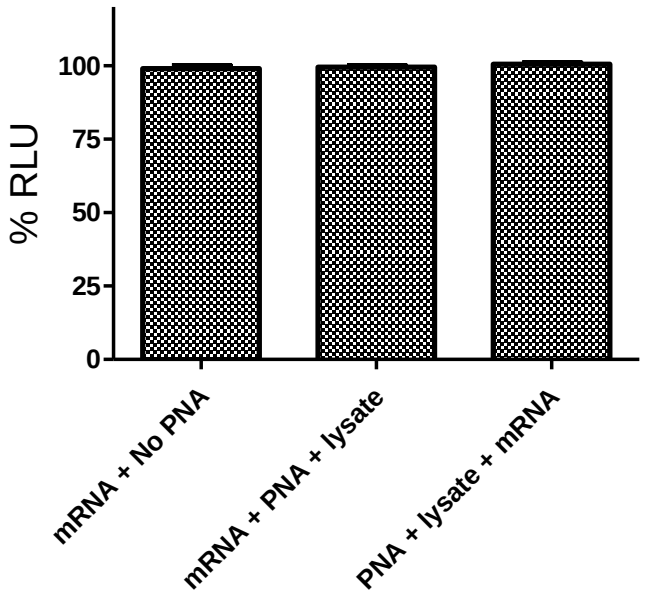


Figure C1. Effect of PNAs on ribosomal translation efficiency. Control transcript (Promega) was incubated with PNA prior to addition of cell lysates (middle bar). PNA was incubated with cell lysates prior to addition of control transcript (right bar). Samples are normalized to translation in the absence of PNA. [PNA] = 250 nM.

GAGACCCAAGCTTTCAGATCCGCTAGCGCTACCGGGATCCAGCCACCATGGAAGACGCCAAAAACATA
AAGAAAGGCCCGGCGCCATTCTATCCTCTAGAGGATGGAACCGCTGGAGAGCAACTGCATAAGGCTAT
GAAGAGATACGCCCTGGTTCCTGGAACAATTGCTTTTACAGATGCACATATCGAGGTGAACATCACGT
ACGCGGAATACTTCGAAATGTCCGTTCGGTTGGCAGAAGCTATGAAACGATATGGGCTGAATACAAAT
CACAGAATCGTCGTATGCAGTGAAAACTCTCTTCAATTCTTTATGCCGGTGTTGGGCGCGTTATTTAT
CGGAGTTGCAGTTGCGCCCGCGAACGACATTTATAATGAACGTGAATTGCTCAACAGTATGAACATTT
CGCAGCCTACCGTAGTGTTTGTTTCCAAAAAGGGGTTGCAAAAAATTTTGAACGTGCAAAAAAAATTA
CCAATAATCCAGAAAATTATTATCATGGATTCTAAAACGGATTACCAGGGATTTCAGTCGATGTACAC
GTTCGTCACATCTCATCTACCTCCCGGTTTTAATGAATACGATTTTGTACCAGAGTCCTTTGATCGTG
ACAAAACAATTGCACTGATAATGAATTCCTCTGGATCTACTGGGTTACCTAAGGGTGTGGCCCTTCCG
CATAGAACTGCCTGCGTCAGATTCTCGCATGCCAGAGATCCTATTTTTGGCAATCAAATCATTCCGGA
TACTGCGATTTTAAGTGTTGTTCCATTCCATCACGGTTTTGGAATGTTTACTACACTCGGATATTTGA
TATGTGGATTTCGAGTCGTCTTAATGTATAGATTTGAAGAAGAGCTGTTTTTACGATCCCTTCAGGAT
TACAAAATTCAAAGTGCGTTGCTAGTACCAACCCTATTTTCATTCTTCGCCAAAAGCACTCTGATTGA
CAAATACGATTTATCTAATTTACACGAAATTGCTTCTGGGGGCGCACCTCTTTCGAAAGAAGTCGGGG
AAGCGGTTGCAAAACGCTTCCATCTTCCAGGGATACGACAAGGATATGGGCTCACTGAGACTACATCA
GCTATTCTGATTACACCCGAGGGGGATGATAAACCGGGCGCGGTCGGTAAAGTTGTTCCATTTTTTGA
AGCGAAGGTTGTGGATCTGGATACCGGGAAAACGCTGGGCGTTAATCAGAGAGGCGAATTATGTGTCA
GAGGACCTATGATTATGTCCGGTTATGTAAACAATCCGGAAGCGACCAACGCCTTGATTGACAAGGAT
GGATGGCTACATTCTGGAGACATAGCTTACTGGGACGAAGACGAACACTTCTTCATAGTTGACCGCTT
GAAGTCTTTAATTAAATACAAAGGATATCAGGTGGCCCCCGCTGAATTGGAATCGATATTGTTACAAC
ACCCCAACATCTTCGACGCGGGCGTGGCAGGTCTTCCCGACGATGACGCCGGTGAACTTCCCGCCGCC
GTTGTTGTTTTGGAGCACGGAAAGACGATGACGGAAAAAGAGATCGTGGATTACGTCGCCAGTCAAGT
AACAACCGCGAAAAAGTTGCGCGGAGGAGTTGTGTTTGTGGACGAAGTACCGAAAGGTCTTACCGGAA
AACTCGACGCAAGAAAAATCAGAGAGATCCTCATAAAGGCCAAGAAGGGCGGAAAGTCCAAATTGTAA
GCGGCCGCTCGAGCATGCATCTAGAGGGCCCTATTCTATAGTGTCACCTAAATGCTAGAGCTCCCAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAA

Figure C2: Sequence of No G3 control. Firefly luciferase coding sequence highlighted in yellow. Possible C-rich and G-rich off targets are highlighted in cyan and pink, respectively.

GAGACCCAAGCTTTCAGATCCGCTAGC**GGG**A**GGG**C**GGG**TCT**GGG**CGATCCAGCCACCATGGAAGACGC
CAAAAACATAAAGAAAGGCCCGGCGCCATTCTATCCTCTAGAGGATGGAACCGCTGGAGAGCAACTGC
ATAAGGCTATGAAGAGATACGCCCTGGTTCCTGGAACAATTGCTTTTACAGATGCACATATCGAGGTG
AACATCACGTACGCGGAATACTTCGAAATGTCCGTTCGGTTGGCAGAAGCTATGAAACGATATGGGCT
GAATACAAATCACAGAATCGTCGTATGCAGTGAAAACTCTCTTCAATTCTTTATGCCGGTGTTGGGCG
CGTTATTTATCGGAGTTGCAGTTGCGCCCGCGAACGACATTTATAATGAACGTGAATTGCTCAACAGT
ATGAACATTTCGCAGCCTACCGTAGTGTTTGTTTCCAAAAAGGGGTTGCAAAAAATTTTGAACGTGCA
AAAAAAATTACCAATAATCCAGAAAATTATTATCATGGATTCTAAAACGGATTACCAGGGATTTCAGT
CGATGTACACGTTCGTCACATCTCATCTACCTCCCGGTTTTAATGAATACGATTTTGTACCAGAGTCC
TTTGATCGTGACAAAACAATTGCACTGATAATGAATTCCTCTGGATCTACTGGGTTACCTAAGGGTGT
GGCCCTTCCGCATAGAACTGCCTGCGTCAGATTCTCGCATGCCAGAGATCCTATTTTTGGCAATCAAA
TCATTCCGGATACTGCGATTTTAAGTGTTGTTCCATTCCATCACGGTTTTGGAATGTTTACTACACTC
GGATATTTGATATGTGGATTTCGAGTCGTCTTAATGTATAGATTTGAAGAAGAGCTGTTTTTACGATC
CCTTCAGGATTACAAAATTCAAAGTGCGTTGCTAGTACCAACCCTATTTTCATTCTTCGCCAAAAGCA
CTCTGATTGACAAATACGATTTATCTAATTTACACGAAATTGCTTCTGGGGGCGCACCTCTTTCGAAA
GAAGTCGGGGAAGCGGTTGCAAAACGCTTCCATCTTCCAGGGATACGACAAGGATATGGGCTCACTGA
GACTACATCAGCTATTCTGATTACACCCGAGGGGGATGATAAACCGGGCGCGGTCGGTAAAGTTGTTC
CATTTTTTGAAGCGAAGGTTGTGGATCTGGATACCGGGAAAACGCTGGGCGTTAATCAGAGAGGCGAA
TTATGTGTCAGAGGACCTATGATTATGTCCGGTTATGTAAACAATCCGGAAGCGACCAACGCCTTGAT
TGACAAGGATGGATGGCTACATTCTGGAGACATAGCTTACTGGGACGAAGACGAACACTTCTTCATAG
TTGACCGCTTGAAGTCTTTAATTAAATACAAAGGATATCAGGTGGCCCCCGCTGAATTGGAATCGATA
TTGTTACAACACCCCAACATCTTCGACGCGGGCGTGGCAGGTCTTCCCGACGATGACGCCGGTGAACT
TCCCGCCGCCGTTGTTGTTTTGGAGCACGGAAAGACGATGACGGAAAAAGAGATCGTGGATTACGTCG
CCAGTCAAGTAACAACCGCGAAAAAGTTGCGCGGAGGAGTTGTGTTTGTGGACGAAGTACCGAAAGGT
CTTACCGGAAAACTCGACGCAAGAAAAATCAGAGAGATCCTCATAAAGGCCAAGAAGGGCGGAAAGTC
CAAATTGTAAGCGGCCGCTCGAGCATGCATCTAGAGGGCCCTATTCTATAGTGTCACCTAAATGCTAG
AGCTCCCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

Figure C3: Sequence of 4G3 reporter. Firefly luciferase coding sequence highlighted in yellow. Possible C-rich and G-rich off targets are highlighted in cyan and pink, respectively. G_3 tracts are in bold and underlined.

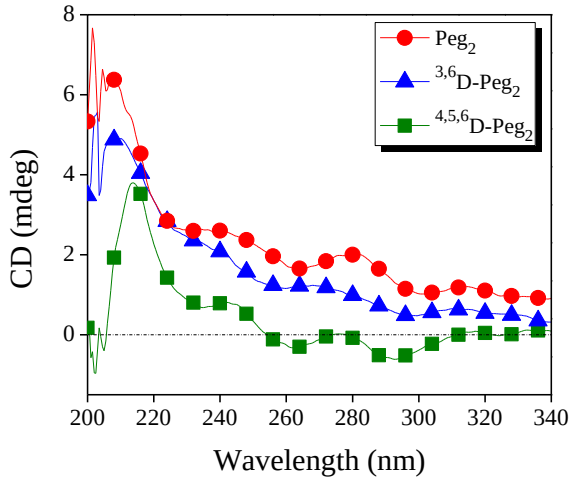


Figure C4. CD spectra for P_{eg2} (red circles), ^{3,6}D-P_{eg2} (blue triangles), and ^{4,5,6}D-P_{eg2} (green squares). Samples contained 5 µM PNA/γPNA buffered in 150 mM KCl, 10 mM Tris-HCl (pH = 7.4), and 0.1 mM Na_2EDTA

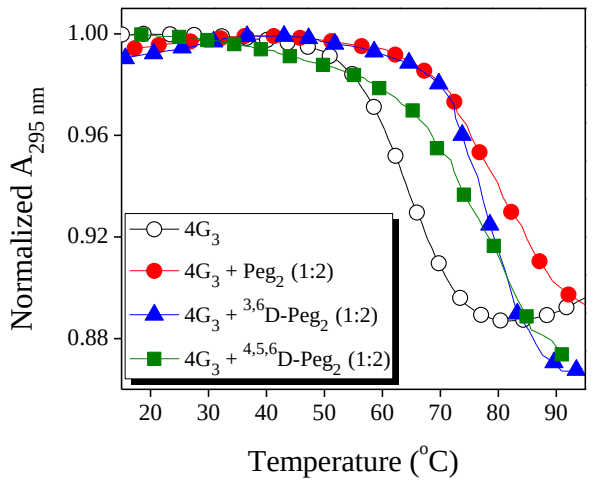


Figure C5. UV-melting for 4G3 RNA in the absence (open circles) or presence of P_{eg2} (red circles), ^{3,6}D-P_{eg2} (blue triangles), or ^{4,5,6}D-P_{eg2} (green squares). Samples contained 2.5 µM RNA and 5 µM PNA/γPNA buffered in 1 mM KCl, 10 mM Tris-HCl (pH = 7.4), and 0.1 mM Na_2EDTA.

APPENDIX D

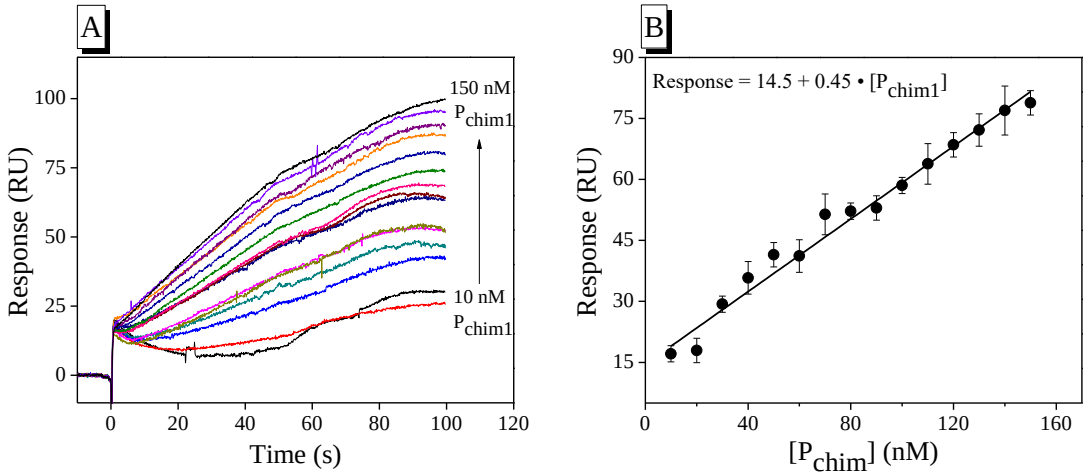


Figure D1: Calibration for SPR competition experiments. Sensorgrams for 10 – 150 nM P_{chim1} (**A**). Linear regression curve for concentration series (**B**).

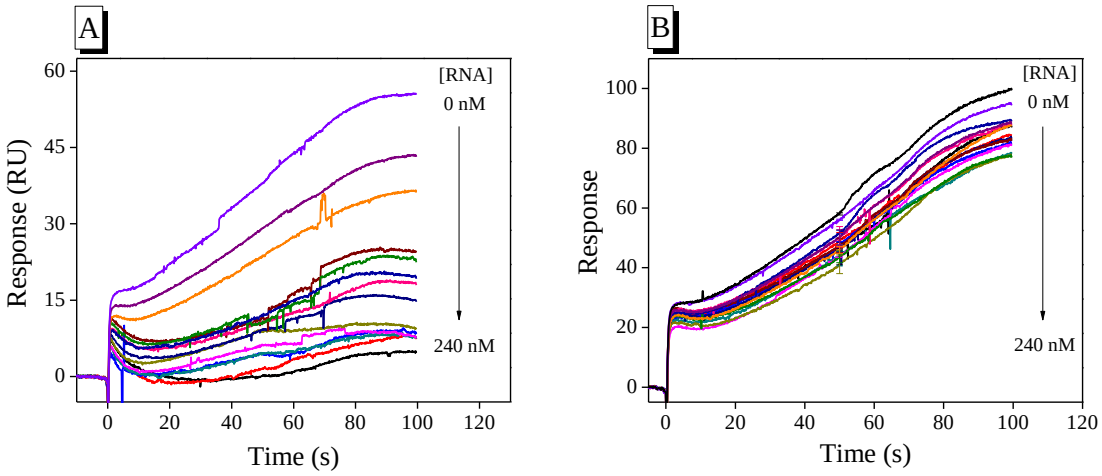


Figure D2: SPR competition experiments. Sensorgrams for titration of 120 nM P_{chim1} with 0 – 240 nM PM RNA (**A**) or MM RNA (**B**).

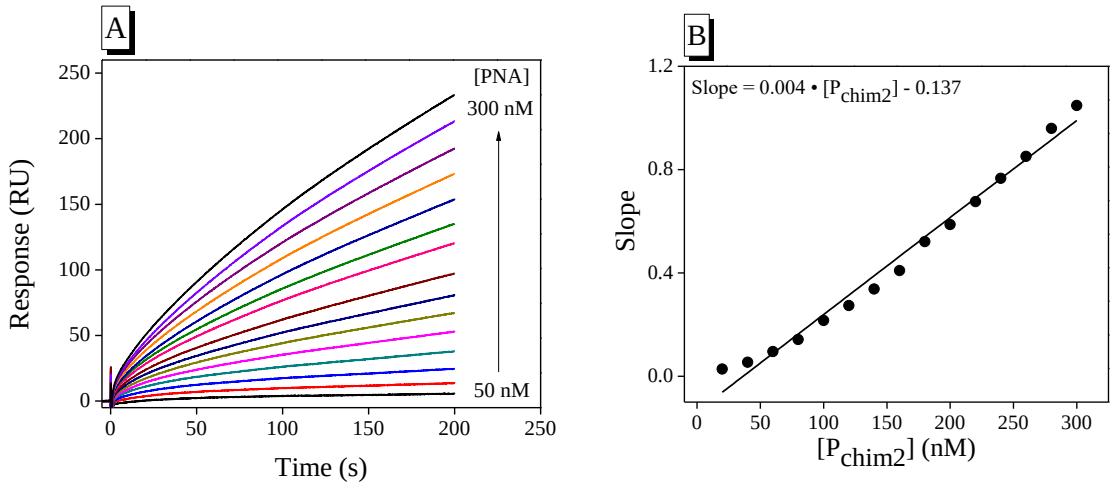


Figure D3: Calibration for SPR competition experiments. Sensorgrams for 10 – 150 nM P_{chim2} (**A**). Linear regression curve for concentration series (**B**).

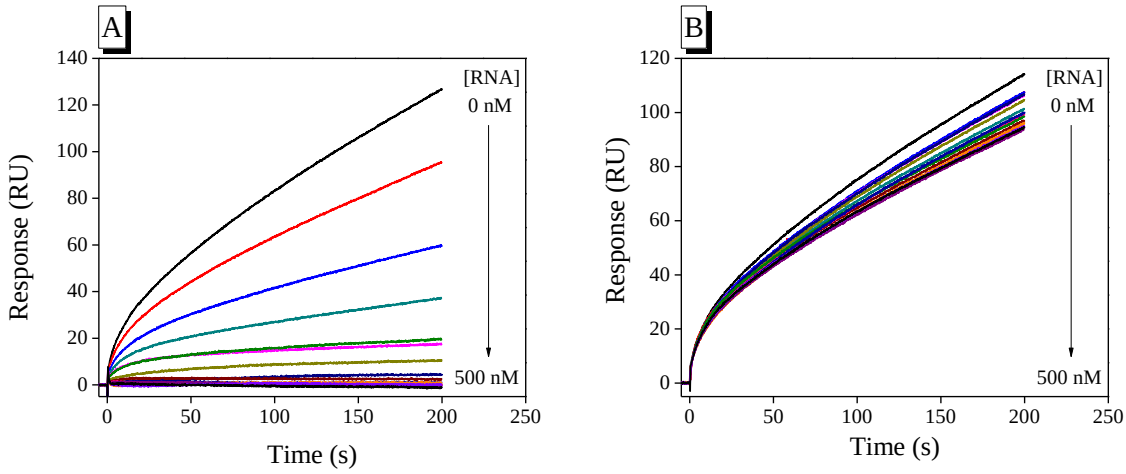


Figure D4: SPR competition experiments. Sensorgrams for titration of 170 nM P_{chim2} with 0 – 500 nM PM RNA (**A**) or MM RNA (**B**).