Summer 2013

Fragile X Mental Retardation Protein: Self-Regulation and miRNA Pathway Involvement

Anna Blice-Baum

FRAGILE X MENTAL RETARDATION PROTEIN: SELF-REGULATION AND miRNA PATHWAY INVOLVEMENT

A Dissertation

Submitted to the Bayer School of Natural and Environmental Sciences

Duquesne University

In partial fulfillment of the requirements for

the degree of Doctor of Philosophy

By

Anna C. Blice-Baum

August 2013

FRAGILE X MENTAL RETARDATION PROTEIN: SELF-REGULATION AND

miRNA PATHWAY INVOLVEMENT

By

Anna C. Blice-Baum

Approved June 28, 2013

Mihaela-Rita Mihailescu, Ph.D. Associate Professor of Chemistry and Biochemistry (Committee Chair)	Jeffry M. Madura, Ph.D. Professor of Chemistry and Biochemistry (Committee Member)
Michael Cascio, Ph.D. Associate Professor of Chemistry and Biochemistry (Committee Member)	Subha R. Das, Ph.D. Associate Professor of Chemistry Carnegie Mellon University (External Reviewer)
Ralph A. Wheeler, Ph.D. Chair, Department of Chemistry and Biochemistry Professor of Chemistry and Biochemistry	David W. Seybert, Ph.D. Dean, Bayer School of Natural and Environmental Sciences Professor of Chemistry and Biochemistry

ABSTRACT

FRAGILE X MENTAL RETARDATION PROTEIN: SELF-REGULATION AND miRNA PATHWAY INVOLVEMENT

By

Anna Blice-Baum

August 2013

Dissertation supervised by Dr. Mihaela-Rita Mihailescu

Fragile X syndrome, the most common form of inherited mental impairment in humans, affects 1 of 4000 males and 1 of 8000 females. It is caused by the absence of the fragile X mental retardation protein (FMRP), resulting from a CGG trinucleotide repeat expansion in the 5'-untranslated region (UTR) of the *fragile x mental retardation-1 (FMR1)* gene, and subsequent translational silencing of FMRP. FMRP, a proposed translational regulator of neuronal messenger RNA (mRNA) targets, has three RNA binding domains: two K-homology domains (KH1 and KH2) and one arginine-glycine-glycine (RGG) box domain. FMRP RGG box has been shown to bind with high affinity to G-quadruplex forming mRNAs. G-quadruplexes are formed by stacked G-quartets bonded by Hoogsteen base pairing and stabilized by monocations. FMRP undergoes alternative splicing, including the alternative splice site at exon 15, giving rise to FMRP

minor isoforms, truncated within close proximity of the RGG box domain. The binding of FMRP to a proposed G quadruplex structure in the coding region of its own mRNA named FBS has been proposed to affect mRNA splicing events for FMRP minor isoforms. In this study we used biophysical methods to directly demonstrate the folding of *FMR1* FBS into two specific G-quadruplexes and analyze its binding by the FMRP isoforms. Additionally, we analyzed the binding of an FMRP mutant in which Ser500 was replaced with Asp500 (ISOP), mimicking FMRP phosphorylation. We showed that the minor splice isoforms bind more tightly to the FBS mRNA, suggesting a negative feedback loop of FMRP binding to its mRNA to regulate alternate splicing.

FMRP associates directly with the Ago1 protein, a key component in the microRNA (miRNA) pathway. Interestingly, one of the FMRP mRNA targets, the microtubule associated protein 1B (MAP1B) mRNA, has a G-quadruplex structure in its 5'-UTR shown to be bound by the FMRP RGG box, and a potential binding site for the miRNA let-7b in its 3'-UTR. In this study we investigated the binding of the let-7b miRNA to this sequence within MAP1B mRNA by using biophysical methods. Dr. Yue Feng at Emory University confirmed the translation regulation of let7b miRNA on MAP1B mRNA.

DEDICATION

I dedicate this document and all that it represents to my mom, Christine, and my dad, Mark. You both have been an inspiration in academics and life in general throughout my whole existence. Mom, you are the strongest woman I know. I am so proud to be your daughter and your SAI sister, and I will be so proud to be among the female Blice-Baums with doctorates! Dad, I will always be your little girl, and I will always accept being spoiled by you! I cherish the advice you give me, and I will never stop asking you questions about my finances and life in general.

I also dedicate this document to my brother, Steve. You keep me grounded and make me realize that I can always be better. You always find something about which to make fun of me, but you also compliment me when you find it appropriate. I look up to you as my big brother, and I love you very much!

Thank you all for your support in the completion of my goal in getting my Ph.D!

ACKNOWLEDGEMENTS

I would first like to acknowledge my advisor and mentor, Dr. Mihaela-Rita Mihailescu. You have pushed and nourished me as a scientist during my graduate studies at Duquesne University. You have guided me on my way to becoming an independent thinker in the sciences and urged me to desire the best in myself, both in my current research and in my future endeavors. I will always strive to teach and mentor as well as you have with me as your student.

Next, I would like to thank my committee members, Dr. Jeffry Madura, Dr. Michael Cascio, and my outside reader Dr. Subha Das from Carnegie Mellon University. Dr. Madura, your mind is amazing, yet you always manage to explain things in a way that makes sense to even a first year graduate student. Thank you for all of your tough love and support and for all of the shining letters of recommendation I have asked you for over the years. Dr. Cascio, your charismatic attitude about teaching and research makes me excited to learn new things about biochemistry. You are an inspiration for keeping research and teaching fun. Dr. Das, thank you very much for being willing to serve as my outside reader for my Ph.D. dissertation. Your vast knowledge of nucleic acids research helped me to make my dissertation the best it could be.

I also want to thank all of the former labmates with whom I have shared space and resources over the years I have been at Duquesne University, especially Dr. Timothy Evans, Dr. Sumangala Shetty, and Valerie Schrott, M.S. You are my friends, and you made going to work and school every day worth it. I was so proud to see you graduate successfully, and now it is my turn. Additionally, I need to acknowledge my current

labmates who also make coming to work worth it every day, Damian McAninch, Snezana Stefanovic, and Yang Zhang. You ask me questions, and you help me feel like I deserve this degree whenever I can answer them and direct you on the right path toward the answer. Thank you also to all of my friends in the Duquesne University Department of Chemistry and Biochemistry, especially Joe and Kim Rosmus, Marc Neiswonger, Nina Zyvith, Logan Miller, and Rathna Veeramachaneni, and former student, Dr. Courtney Watkins. It's so great to talk to you about science and other things on a daily basis! I must also give special thanks to the undergraduate researchers whom I have mentored over the years: Sara Katrancha, Emily Spitzer, Carley Chwal, Ashley Heinaman, and Jessica Rabuck.

I have to give special thanks to the guys downstairs in the instrument shop who fix our instruments every time we need something fixed: Dan Bodnar, Lance Crosby, Ben Lauterbach, and former employee Dave Hardesty. More special thanks is deserved by Ian Welsh, who works right across the hall from me and is always calm when I have a problem. Sorry I was always the bearer of bad news whenever the double DI water filter needed to be changed!

Thank you to my collaborators, Dr. Yue Feng from Emory University and Dr. Bassem Hassan from Katholieke Universiteit Leuven. I also would like to acknowledge the Bayer School of Natural and Environmental Sciences for funding through the Bayer Fellowship during the 2012 year and to the National Institutes of Health grant NIH #3R15GM074660-02A1S1.

TABLE OF CONTENTS

Page

LIST OF FIGURES

Page

LIST OF EQUATIONS

Page

LIST OF ABBREVIATIONS

2-AP	2-aminopurine
AGO	Argonaut family of proteins
AMP	ampicillin
CD	circular dichroism
CGG	cytosine-guanine-guanine
CHL	chloramphenicol
DNA	deoxyribonucleic acid
D_2O	deuterium oxide
Exp5-Ran	exportin 5 and Ran-binding protein complex
FBS	Fragile X mental retardation protein binding sequence
FMR1	Fragile X mental retardation gene
FMRP	Fragile X mental retardation protein
FXS	Fragile X syndrome
IPTG	isopropyl-β-D-1-thiogalactopyranoside
ISO	isoform
KH	K-homology domain
KO	knockout
MALDI-TOF-MS	matrix-assisted laser desorption/ionization-time of flight-mass spectrometry
MAP1B	microtubule associated protein 1B
mGluR	metabotropic glutamate receptor 5

miRISC	miRNA induced silencing complex
miRNA	microRNA
mRNA	messenger ribonucleic acid
mRNP	messenger ribonucleoprotein
NES	nuclear export signal
NLS	nuclear localization signal
NMR	nuclear magnetic resonance spectroscopy
PAGE	polyacrylamide gel electrophoresis
PSD95	post-synaptic density protein 95
RGG box	arginine-glycine-glycine box domain
S3F	semaphorin 3F protein
SDS	sodium dodecyl sulfate
UTR	untranslated region
UV-vis spectroscopy	ultra-violet-visible spectroscopy

CHAPTER 1: INTRODUCTION

1.1 Fragile X Syndrome: Background

In 1943, J. Purdon Martin and Julia Bell discovered a form of mental disability that was associated with the X chromosome (*1*). In 1969 and 1970, Herbert Lubs and Frederick Hecht located the irregular chromosome structure responsible for the mental retardation on the X chromosome and named it the "fragile site" (*2*). Over two decades later, the specific gene defect found on the fragile X mental retardation gene (*FMR1*) was discovered in patients with fragile X syndrome (FXS) (*3*). FXS is the most common form of inherited mental retardation in humans, affecting approximately one in 3000 males and one in 5000 females (*4*). Diagnosed individuals display impaired mental function, hyperactivity, and a 30% chance of autism, as well as physical abnormalities, such as elongated face and hands, large ears, and macroorchidism in males (*5*).

1.2 Gene mutation of Fragile X Syndrome

In the vast majority of cases, the phenotype of FXS described above is caused by the absence of a single protein required for normal neuronal function, named the fragile X mental retardation protein (FMRP) (*6, 7*). Transcriptional silencing of FMRP is caused by the expansion of a region of cytosine-guanine-guanine (CGG) repeats in the 5-untranslated region (5'-UTR) of the *FMR1* gene by an unknown mechanism. When this region expands to over 200 repeats, the cytosines become hypermethylated (*7*). In normal X chromosome gene regulation, methylation of cytosine causes inactivation of transcription. Thus, the unnatural hypermethylation of the CGG repeat expansion region

in FXS patients is thought to block normal transcription factors, causing the transcriptional silence of the *FMR1* gene. There have been extensive studies of FMRP given that the loss of the protein has been implicated in FXS (*8, 9*), but its role in normal brain function has not yet been fully explained, and the exact mechanism of how its loss leads to mental retardation is not known.

1.3 Fragile X Mental Retardation Protein

FMRP is an RNA binding protein expressed in most tissues, but found predominantly in the dendritic spines of neurons and the testes (*6*). The protein contains two different types of RNA binding domains: two K-homology domains and one arginine-glycine-glycine (RGG) box domain, as well as a nuclear localization signal (NLS) and a nuclear export signal (NES), which allow the protein to shuttle between the

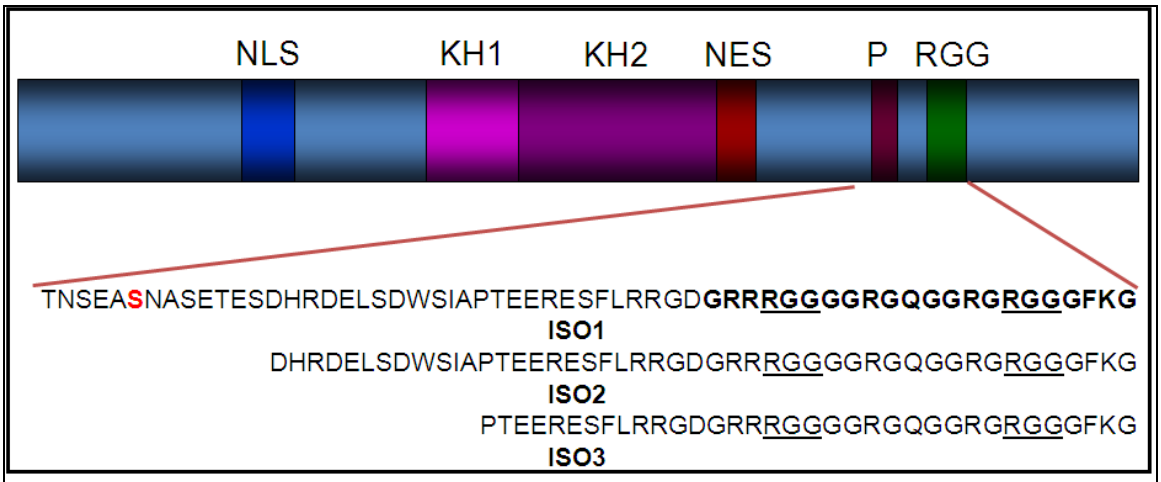


Figure 1.1 A schematic representation of the full-length FMRP, which shows the nuclear localization signal (NLS), the two K-homology domains (KH1 and KH2), the nuclear export signal (NES), the site of phosphorylation (P) and the RGG box (RGG). Isoforms 1 – 3, resulting from the alternative splicing at exon 15 of *FMR1* mRNA are also illustrated. The phosphorylation of serine 500 (red) has been shown to be biologically relevant. Adapted from (*10*).

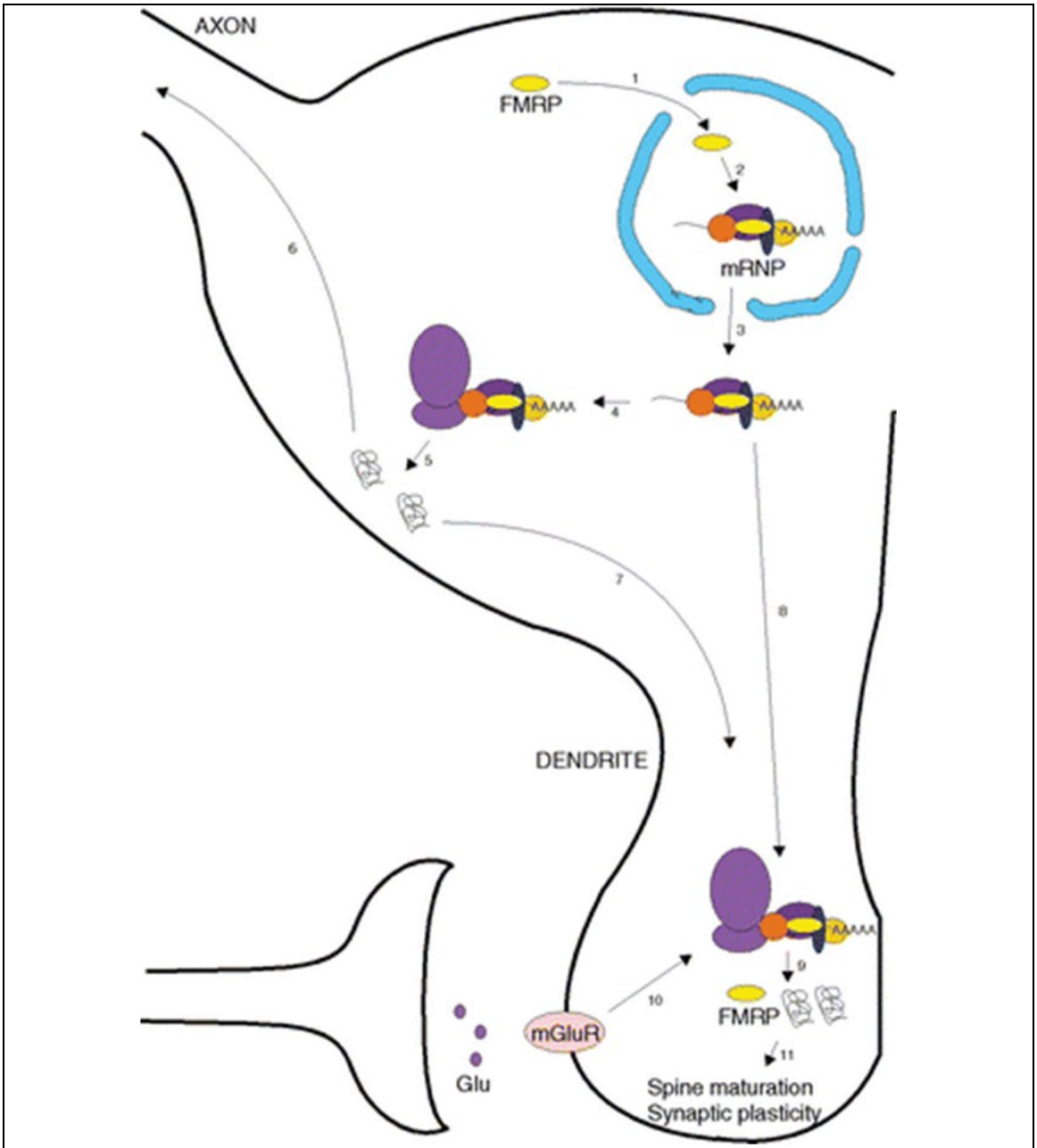


Figure 1.2 A proposed mechanism of FMRP function in neuronal cells. FMRP is transported into the nucleus of a neuronal cell by way of its NLS (1), wherein it associates with cellular mRNAs and proteins forming larger ribonucleoprotein (RNP) complexes (2). that are transported out of the nucleus into the cytoplasm via the NES (3). Once in the cytoplasm, FMRP can either transport its bound mRNA to ribosomes (4) for protein production (5) in response to mGluR stimulation and then transport the protein products to neurites (6, 7), or it can transport mRNAs to dendrites (8) and then allow for protein production of its own message or its mRNA targets (9) in response to mGluR stimulation (10). (*6*) Reprinted with permission from the publisher.

nucleus and cytoplasm of a cell (*6, 11, 12*). A schematic of the full-length FMRP as well as the splice variants of ISO1, ISO2, and ISO3 produced by alternative splicing at exon 15 are shown in figure 1.1 (*10*). In addition, FMRP has been shown to undergo the posttranslational modifications of phosphorylation and arginine methylation (*13, 14*). Since it binds to mRNAs as part of large ribonucleoprotein (mRNP) complexes, FMRP is postulated to be involved in regulating the translation of specific neuronal messenger RNA (mRNA) targets in response to synaptic output, although the mechanism of this is not well understood (*15*). The proposed mechanism of FMRP function as an RNA binding protein begins with FMRP being transported into the nucleus of a neuronal cell by way of its NLS, wherein it associates with cellular mRNAs and proteins in order to form larger ribonucleoprotein (RNP) complexes (figure 1.2). These RNPs are then transported out of the nucleus into the cytoplasm where FMRP can either transport its bound mRNA to ribosomes for protein production in response to mGluR stimulation and then transport the protein products to neurites, or it can transport mRNAs to neurites and then allow for protein production in response to mGluR stimulation (*6*). FMRP is known to bind to about 4% of all neuronal messages, although the detailed mechanisms by which it recognizes its targets are not fully understood (*16*).

Several FMRP isoforms can be produced through alternative splicing events that involve the inclusion/skipping of exons 12 and 14, as well as three acceptor sites at exon 15 and two at exon 17 (*17, 18*). Figure 1.3 is a schematic of the splicing patterns that are possible due to alternative splicing in the *FMR1* gene (*18*). In mouse models of the *FMR1* gene, it was found that although mRNA transcripts of isoforms 1 – 12 vary in concentration by as much as two orders of magnitude, all of the transcripts were

associated with polyribosomes, which is consistent with the translation of the *FMR1* mRNA transcripts. The distribution of the isoforms of *FMR1* is consistent with the different developmental stages of the cells investigated, emphasizing the implications of losing all isoforms in FXS. The three isoforms investigated in this study are FMRP isoforms 1, 2, and 3 (ISO1 – ISO3), which are created using the three acceptor sites at exon 15 and inclusive of exons 12 and 14 (figure 1.3). All three isoforms contain the KH domains as well as the RGG box domain, which has been shown to bind to G-quadruplex forming mRNA targets (figure 1.1).

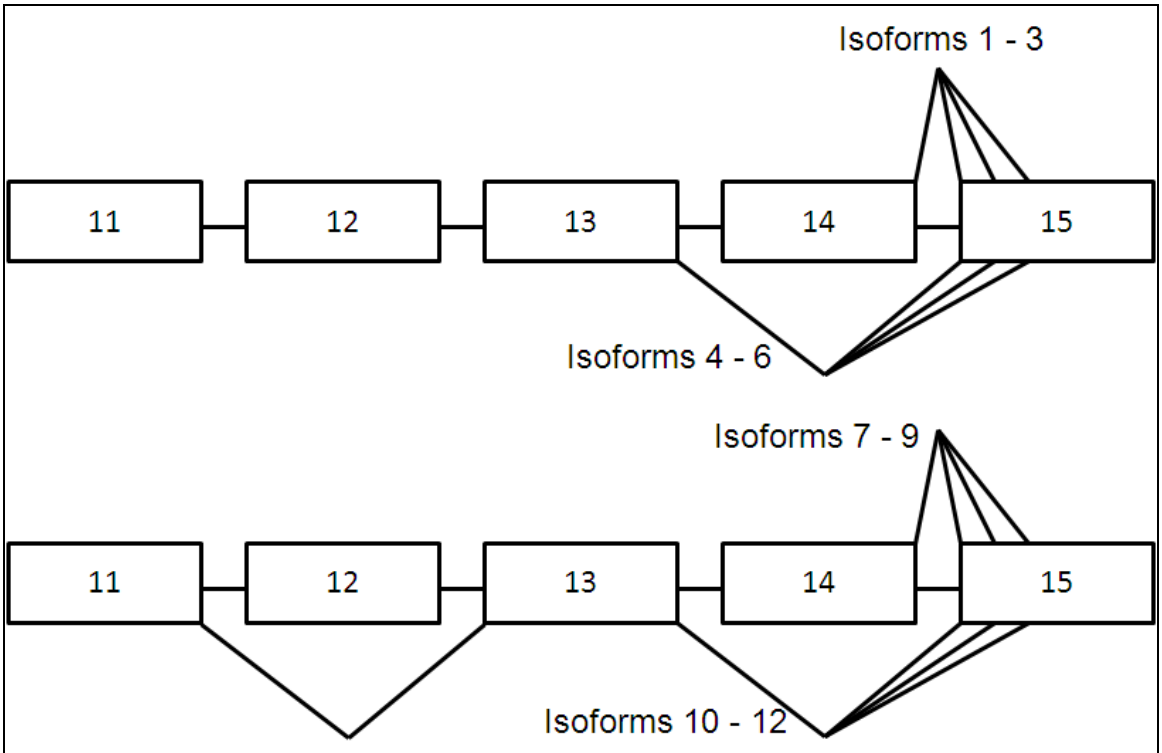


Figure 1.3 Schematic representation of the alternate splicing of FMR1 mRNA into isoforms 1 through 12: adapted from (*18*).

1.4 G-quadruplex structures

G-quadruplexes were first discovered in 1962 when Gellert and colleagues assembled X-ray diffraction data on a tetrameric helical structure formed by guanylic acid (*19*), after which it was revealed that similarly structured poly-guanosine oligomers adopted the same structure in DNA and RNA (*20, 21*). The best characterized DNA sequences that have been shown to form G-quadruplex structures are the intramolecular telomeric sequences, which are found at the ends of eukaryotic chromosomes (*22-24*) and whose purpose is to block telomere elongation by cancer-associated telomerase (*25, 26*). The role of G-quadruplex formation in genes seems to be mainly in gene regulation as they are predominantly found in promoter regions of genes at the transcriptional level (*26*). G-quadruplexes may be more prevalent in RNA than in DNA because RNA is single-stranded more often than DNA, allowing for the formation of G-quadruplex structures due to the availability of bases that are not involved in Watson-Crick base pairing (*26, 27*). Almost 3000 mRNAs have regions in their 5'-UTR that potentially form G-quadruplex structures, leading to the belief that the presence of this secondary structure is important in gene expression at the translational level (*26, 28*). G-rich RNA and DNA sequences have also been implicated in the alternative splicing of several genes (*29*). G-quadruplex structures are characterized by two or more stacks of planar G-quartets that are held together by Hoogsteen base pairing and stabilized by monocations, the potassium ion (K^+) providing the most thermodynamic stability (figure 1.4 A) (*30, 31*). Depending on the direction of each strand of RNA involved in the interaction, G-quadruplex structures can be described as parallel or antiparallel (figure 1.4 B). An intramolecular G-quadruplex may form from one strand of DNA or RNA folding in on

itself, while an intermolecular G-quadruplex may form when more than one strand of DNA or RNA interacts with each other. Intramolecular G-quadruplex RNA sequences are described as containing four distinct runs of two or more guanine bases that can form stable G-quadruplex structures by themselves from a single strand of RNA (*32*).

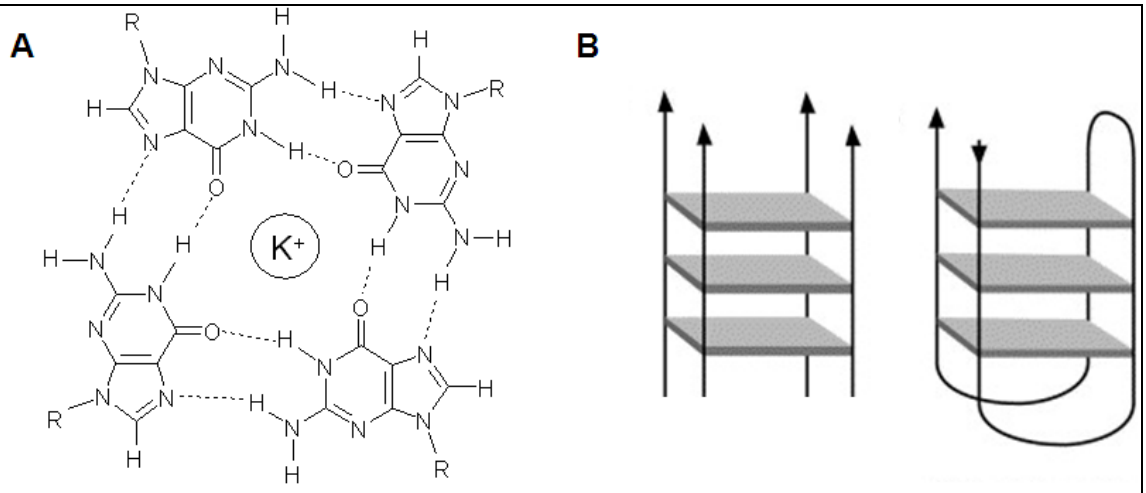


Figure 1.4 A structure of a G-quartet with Hoogsteen hydrogen bonding between guanines stabilized by K^+ ions. These quartets stack on top of one another to form a G-quadruplex RNA structure further stabilized by K^+ ions (A). Parallel versus antiparallel G-quadruplex structures (B).

Several biophysical methods have been used to characterize G-quadruplex formation in DNA and RNA. Some of these methods include nuclear magnetic resonance (NMR) spectroscopy, circular dichroism (CD) spectroscopy, gel electrophoresis, and x-ray crystallography. The imino protons involved in the characteristic Hoogsteen base-pairing in G-quadruplex formation resonate in the range 10 – 12 ppm in the ^{1}H-NMR spectrum in contrast to the range 12 – 15 ppm for imino protons involved in Watson-Crick base pairing (*33*). CD spectroscopy is used to characterize G-quadruplexes as parallel or antiparallel (Figure 1.4 B). Parallel G-quadruplexes possess a positive band at 265 nm and a negative band at 240 nm (*34, 35*)

while antiparallel G-quadruplexes typically possess two positive bands, one at 290 nm and one at 245 nm, and one negative band at 265 nm (*36*) (Figure 1.5). Interestingly, A-form RNA

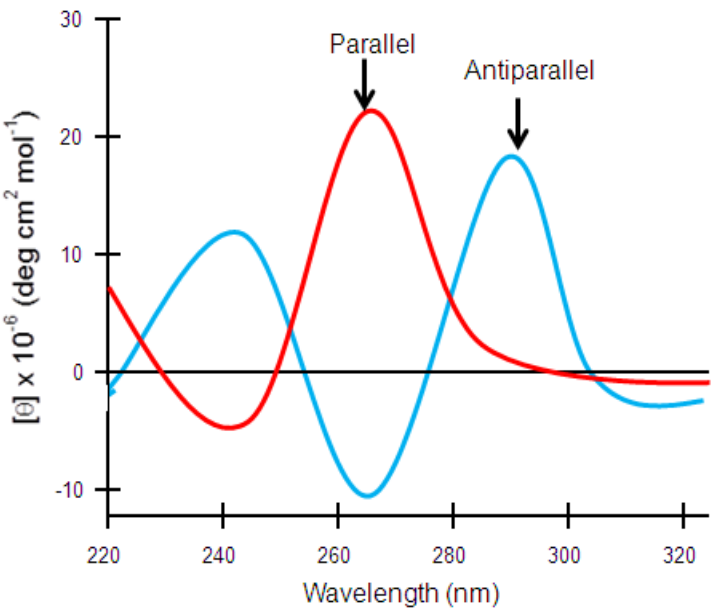


Figure 1.5 Example of CD spectra from parallel G-quadruplex formation with a positive band at 265 nm and a negative band at 240 nm and antiparallel G-quadruplex formation with positive bands at 290 nm and 245 nm and a negative band at 265 nm.

has a positive band at 260 nm, and B-DNA has a negative band at 245 nm and a positive band at 275 nm (*37*), both of which are in close proximity of the bands seen for Parallel G-quadruplex RNA formation. However, if CD spectroscopy experiments are performed in conjunction with ^{1}H-NMR spectroscopy showing characteristic peaks for G-quadruplex imino protons rather than Watson-Crick base-paired imino protons, then the CD spectra can be interpreted as characterizing parallel G-quadruplexes rather than any other secondary structure of RNA.

1.5 FMRP recognition of its mRNA targets

The RGG box domain of FMRP is characterized by two arginine-glycine-glycine repeats within close proximity of each other within the FMRP amino acid sequence (*38*). 44 human RNA binding proteins have been found to contain an RGG box (*39*), but there is no consensus sequence of the RNA binding domain. Several studies have shown that the RGG box domain of FMRP binds with high affinity to mRNA targets that form G-quadruplex structures (*15, 40, 41*).

It is thought that the RGG box domain is dynamic in nature, becoming structured upon binding with G-quadruplex forming mRNA. Recently, a solution structure of the RGG box domain peptide binding with a 36 nucleotide (nt) sequence sc1 RNA, a G-quadruplex model system, was solved by NMR spectroscopy (*42*). The NMR spectroscopy solution structure confirmed that the RGG box indeed becomes structured upon binding to the G-quadruplex formed by the RNA sequence and revealed that the RGG box binds to a junction between the G-quadruplex and a stem formed by the RNA (*10, 42*). However, since this is the only RGG box-RNA structure available, it is not clear if this mode of binding is unique to the synthetic sc1 RNA or a general mode of RNA recognition by the RGG box. Several specific neuronal gene targets of FMRP have been shown directly by biophysical methods to adopt G-quadruplex structures in their mRNA sequences and to be bound with high affinity and specificity by the isolated FMRP RGG box or the full-length FMRP isoforms (*10, 43-45*). These include the mRNAs that encode for the microtubule associated protein 1B (MAP1B), post synaptic density protein 95 (PSD95), and semaphorin 3F (S3F) (*43, 45-49*). MAP1B is a protein that is involved in controlling microtubule dynamics and synapse formation (*43, 50-52*).

PSD95 is a protein component of the postsynapse that controls synapse strength and stabilizes dendritic spines (*47, 53, 54*). S3F is involved in the maturation of neurons as well as the induction of growth cone collapse (*55, 56*). All of the protein products of these mRNA targets of FMRP are essential for the formation and maintenance of neuronal cells, and the misregulation of these proteins due to the lack of FMRP translation regulation in FXS patients may have implications in the development of mental disability. Interestingly, it has been shown that the FMRP RGG box binds with high affinity and specificity to a semaphorin 3F mRNA sequence that forms a G-quadruplex but does not include a stem structure (*45*). This finding suggests that the FMRP RGG box might have different modes of recognition of its mRNA G-quadruplex structures, considering that in the case of sc1 RNA the binding occurred at the junction between a G-quadruplex and stem structure.

In addition to the RGG box RNA binding domain, FMRP contains two KH domains (figure 1.1). KH domains were first discovered in the heterogeneous nuclear ribonucleoprotein K in 1994 and have since been identified in many RNA binding proteins to perform a multitude of cellular functions (*57*). The KH domain recognizes RNA and single-stranded DNA sequences to perform its function, and two or more KH domains are commonly found within the same protein. The KH domains of FMRP have been shown to bind to RNAs that form kissing-complexes (*58*), however, no consensus sequence was determined, and no FMRP mRNA targets that harbor kissing complexes have been identified. A kissing complex in RNA is formed by Watson-Crick base pairing between the loops of two RNA hairpin structures. An FXS patient was found to

harbor a missense mutation (I304N) within the second KH domain of FMRP, suggesting that the KH domain is relevant for the protein function (*58, 59*).

Additionally, a patient recently diagnosed with FXS was found to have normal levels of FMRP (unpublished data, Dr. Bassem Hassan, Katholieke Universiteit Leuven). However, the analysis of the *FMR1* gene revealed a frameshift mutation that essentially abolishes the RGG box domain, truncates the C-terminus of the protein by creating a premature stop codon, and changes the identity of the remaining residues after the mutation position. The elimination of the RGG box domain in this patient who expresses normal levels of FMRP illustrates the importance of this RNA binding domain in the normal function of FMRP.

1.6 FMRP and the miRNA pathway

Recent lines of evidence connect FMRP directly with the microRNA pathway (*7, 60-65*). The miRNA pathway is shown in figure 1.6 (*66, 67*). miRNAs are short, noncoding RNA sequences encoded by the genome of a cell, with over 1,400 miRNAs being identified in humans (*68*). miRNAs are 20 to 25 nucleotide single stranded RNAs that have as a precursor longer strand RNAs processed and cleaved by the RNases Drosha and Dicer (*66*). Primary miRNAs (pri-miRNAs) are initially transcribed by RNA polymerase II and cleaved in the nucleus by the Drosha-DGCR8 complex to form pre-miRNAs, which are approximately 70 nucleotides in length. After the pre-miRNA is exported from the nucleus by the Exp5-Ran complex, Dicer further cleaves the pre-miRNA in the cytoplasm to produce a miRNA duplex and finally the single miRNA strand, which is incorporated into the RNA-induced silencing complex (RISC), a multi-

protein complex that involves the Argonaut class of proteins (*7, 66, 69, 70*). miRNAs exert their translation regulation function by identifying and binding to target mRNA sequences (*66, 67, 71*). The evidence of FMRP association with the RISC suggests that the miRNA pathway may be involved in the exertion of the translation regulator function

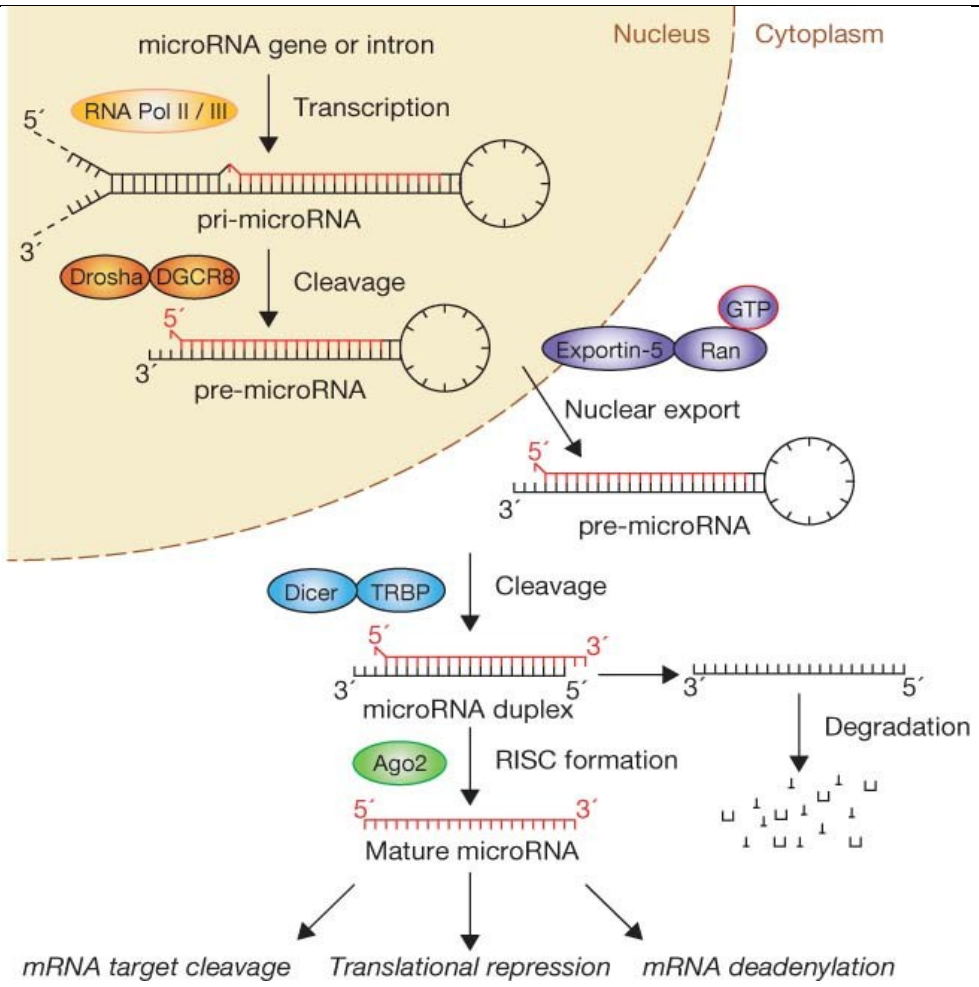


Figure 1.6 Schematic representation of the miRNA pathway. In the nucleus, the miRNA gene is transcribed into pri-miRNA, which is cleaved by the Dicer-Drosha complex into pre-miRNA. Pre-miRNA is transported into the cytoplasm by Exp5-Ran complex, is further cleaved by Dicer and bound by an Argonaut protein (Ago) in the RNA induced silencing complex (RISC) to be used in either gene silencing (partial complementarity) or mRNA cleavage (perfect or near perfect complementarity). Adapted and reprinted with permission from the publisher from (*67*).

of FMRP (*7*). Given this connection, it is reasonable to inquire if the translation regulation function of FMRP is mediated by the miRNA pathway. Interestingly, one of the FMRP mRNA targets, the microtubule associated protein 1B (MAP1B) mRNA has in its 5'-untranslated region (5'-UTR) a G-quadruplex structure that has been shown to be bound with high affinity by the FMRP RGG box and in its 3'-UTR has a potential binding site for the miRNA let-7b. In collaboration with Dr. Yue Feng from Emory University, we investigated the *in vitro* binding of let-7b miRNA to the sequence within the 3'-UTR of MAP1B mRNA (MAP1B seq1 RNA) by biophysical methods (Chapter 5). The *in vivo* interactions between let-7b and the MAP1B 3'-UTR were also analyzed by complementary experiments in normal and FMRP knockout (KO) mice, performed in our collaborator's, Dr. Yue Feng, laboratory at Emory University.

1.7 The FMRP binding sequence of *FMR1* mRNA

FMRP has also been shown to interact with its own mRNA by binding to a 100 nucleotide G-rich region named the FMRP binding sequence (FBS), which was proposed to fold into two distinct G-quadruplex structures (*41, 72, 73*). This finding prompted the hypothesis that FMRP might use an autoregulatory loop to regulate its own translation (*41*). However, in a subsequent study, it was shown that the FMRP interactions with FBS do not affect the *FMR1* mRNA stability and translation. The purine rich FBS region has been found instead to be a potent exonic splicing enhancer whose function is dependent on the presence of the G-rich region (*72, 73*). FBS is located in the proximity of the three different acceptor sites at exon 15, and FMRP binding to FBS has been found to control the splicing events at exon 15. Thus, an overexpression of FMRP ISO1 decreased the

usage of the exon 15 first acceptor site, concomitant with an increased usage of the exon 15 second and third acceptor sites (*72*). This direct involvement of FMRP in regulating the production of its minor isoforms created by the usage of exon 15 acceptor sites 2 and 3, which include ISO2 and ISO3, could add a new layer of regulation to the FMRP translation regulator function, as these minor isoforms lack the major phosphorylation site at position 500 (figure 1.1, red). The FMRP phosphorylation/dephosphorylation events have been shown to play a major role in the association of FMRP with the miRNA pathway, which might assist the protein in exerting its translation regulator function at least for a subset of its mRNA targets that contain miRNA binding sites. Phosphorylated FMRP has been found associated with the RNA induced silencing complex (RISC), the microRNA miR-125a, and the postsynaptic density 95 (PSD95) mRNA, whereas its dephosphorylation triggered by synaptic input leads to the dissociation of the RISC complex from PSD95 mRNA, allowing for the translation of the PSD95 protein (*47*). The fact that FMRP isoforms ISO2 and ISO3 lack the major site of phosphorylation raises the possibility that these minor isoforms might control the timing and duration of the "on" state for the translation of specific mRNA targets. Thus, their production has to be tightly regulated, and this may be achieved by a feedback mechanism involving the FBS exonic splicing enhancer.

It is of great interest to directly characterize the presence of a G-quadruplex structure within the FBS sequence of *FMR1* mRNA by biophysical methods to show that FMRP indeed binds to its own mRNA in a G-quadruplex dependent manner, and this was accomplished in the first part of this study (chapter 3). Furthermore, it is also relevant to quantify the binding of major and minor isoforms of FMRP to this sequence by

biophysical methods as it has been shown *in vivo* that the major isoform binding to this sequence influences the production of the minor splice variants of the protein. This was accomplished in the second part of this study, the results being summarized in chapter 4.

1.8 Specific Aims of the research

Specific Aim I: Direct characterization by biophysical methods of G-quadruplex forming mRNA sequence(s) located in the RGG box coding region of *FMR1* mRNA.

Specific Aim II: Characterization of the binding between several FMRP isoforms and the G-quadruplex formed by *FMR1* mRNA by biophysical methods.

Specific Aim III: Characterization of the binding of let7b miRNA to the 3'-UTR of a known FMRP mRNA target, MAP1B mRNA.

1.9 Relevance of the research

The research performed in this project will increase our understanding of the role played by FMRP in controlling the alternate splicing of its own mRNA. It is important to directly prove that the G-rich FBS sequence of *FMR1* mRNA forms a G-quadruplex to which FMRP binds with high affinity. The fact that FMRP binds to its own mRNA in order to regulate the alternate splicing into major and minor isoforms may have implications in other regulatory processes modulated by FMRP as the minor splice

isoforms lack the site of phosphorylation in exon 15. Although the phosphorylation of FMRP ISO1 has not been shown to affect its binding to the FBS sequence of *FMR1* mRNA in this study, it has been shown and proposed to have importance in other aspects of protein function. We have shown that the minor FMRP isoforms, ISO2 and ISO3, bind more tightly to the FBS sequence of *FMR1* mRNA, indicating that neuronal cells utilize the negative feedback loop of FMRP binding to its own mRNA to regulate the alternate splicing into minor isoforms.

Furthermore, the role of FMRP as a translational regulator is not fully understood, and the results of the second part of this study could help elucidate how FMRP functions in the presence of target mRNAs that contain both G-quadruplex structures and complementary binding sequences for miRNAs. This research could be relevant beyond the scope of fragile X syndrome to investigate how other proteins exhibit their translational regulatory functions in the presence of miRNA.

CHAPTER 2: MATERIALS AND METHODS

2.1 Synthesis of RNA *in vitro*

Unlabeled FBS_67 RNA, FBSsh RNA, FBS_Q1 RNA and FBS_Q2 RNA (table 2.1) oligonucleotides were synthesized *in vitro* off synthetic DNA templates (Trilink Biotechnologies, Inc.) using T7 RNA polymerase produced in-house (*74*). Oligonucleotides were purified by 15% or 20% 8 M urea denaturing polyacrylamide gel

Table 2.1 Unlabeled RNA sequences

RNA name	RNA Sequence
FBS_67 RNA	5′- GGA CGG CGG CGU GGA GGG GGA GGA AGA GGA CAA GGA GGA AGA GGA CGU GGA GGA GGC UUC AAA GGA A -3′
FBS_Q1 RNA	5′- GGA GGG GGA GGA AGA -3′
FBS_Q2 RNA	5′- GGA CAA GGA GGA AGA GGA C -3′
FBSsh RNA	5′- GGC GUG GAG GGG GAG GAA GAG GAC AAG GAG GAA GAG GAC GUG -3′

electrophoresis (PAGE), eluted from the gel pieces using electrophoretic elution, and extensively dialyzed against 10 mM cacodylic acid, pH 6.5 (*74*). A fluorescent oligonucleotide was designed for FBSsh in which the highly fluorescent adenine analog 2-aminopurine (2AP) replaced the adenine at position 14 of FBSsh RNA to construct FBSsh_14AP (Dharmacon, Inc.). Two unlabeled RNA sequences, MAP1B seq1 RNA and let-7b miRNA, were synthesized by Dharmacon, Inc. as well since the sequences do not begin with two guanines, a requirement for *in vitro* T7 RNA polymerase

transcription. Samples of FBS_67, Q1, and Q2 RNA sequences were annealed by heating at 95 °C in the presence or absence of KCl and slow cooling to 25 °C for 30 minutes. All samples of FBSsh and FBSsh_14AP RNA were prepared by incubation in the presence or absence of KCl at 25 °C for 20 minutes.

2.2 RGG box peptide synthesis

The FMRP RGG box peptide with the sequence N-RRGDGRRRGGGG RGQGGRGRGGGFKGNDDHSR-C was chemically synthesized by the Peptide Synthesis Unit at the University of Pittsburgh Center for Biotechnology and Bioengineering.

2.3 Expression of recombinant FMRP isoforms

The recombinant pET21a-FMRP plasmid encoding FMRP isoform 1 (ISO1) fused with a C-terminal 6X histidine tag was a kind gift from Dr. Bernhard Laggerbauer (Institut für Molekularbiologie und Tumorforschung, Philipps-Universität Marburg, Marburg, Germany). The truncations of the FMR1 gene encoding for FMRP ISO1 to create the genes encoding for ISO2 and ISO3 were performed by GenScript USA, Inc. and confirmed by sequencing at the University of Pittsburgh Genomics and Proteomics Core. In order to recombinantly express, purify, and dialyze ISO2 and ISO3, we used the previously developed protocol in our laboratory for the expression of FMRP ISO1 (*75*). In brief, plasmids were transformed into Rosetta2(DE3)pLysS *E. coli* cells. All media used in the cell growth consisted of Luria-Bertani (LB, Fisher Scientific) media containing 200 μg/mL ampicillin (AMP, MP Biomedical) and 15 μg/mL

chloramphenicol (CHL, MP Biomedical). Cells were incubated at 37 °C until the target absorbance at 600 nm of 0.8 – 1.0 was reached, and protein expression was induced by adding 1 mM isopropyl β-D-1-thiogalactopyranoside (IPTG) and incubating cells at 250 rpm and 25 °C for 12 hours. Cells were harvested, lysed, and purified using Ni-NTA Superflow resin (Qiagen) as described (*75*). Purified proteins were concentrated using dialysis tubing filled with polyethylene glycol (PEG) 20,000 and dialyzed into a buffer devoid of K^+, Na^+, or imidazole. Final protein buffer consisted of 5% glycerol, 1 mM EDTA, and 300 mM LiCl. The concentration of FMRP isoforms was determined at A_{280} by using the molar extinction coefficients of 46370 $M^{-1}cm^{-1}$ for ISO1 and ISO2 and 40680 $M^{-1}cm^{-1}$ for ISO3 (*76, 77*). The presence of the isoforms was analyzed using a 10% tris-glycine sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) and visualized by Coomassie blue staining.

2.4 Mass Spectrometry Analysis of FMRP ISO2 and ISO3

The identity of FMRP ISO2 and ISO3 was confirmed using peptide mass fingerprinting (Genomics and Proteomics Core Laboratories, University of Pittsburgh). FMRP ISO2 and ISO3 from a 10% SDS-PAGE gel was excised and trypsin digested followed by analysis via matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOF-MS). Five MALDI-TOF-MS characterized amino acid sequences matched those found in FMRP isoforms, and the one sequence corresponding to exon 15 was truncated according to the splice variations of ISO2 and ISO3.

2.5 Design of the phosphomimetic FMRP isoform 1, ISOP

FMRP ISO1 has been shown to be phosphorylated at 3 serine sites in exon 15, serine 500 being shown to be biologically relevant (*47, 78*). Sara Katrancha, an undergraduate in the Mihailescu lab, designed a phosphomimetic version of FMRP ISO1 in which serine 500 was replaced by aspartic acid in order to mimic phosphorylation on serine 500 (*79*). This phosphomimetic FMRP ISO1 (ISOP) mutation was performed by GenScript USA, Inc. In order to recombinantly express, purify, and dialyze ISOP, the previously developed protocol (*75*) for the expression of FMRP ISO1 was used as described in section 2.4. The concentration of FMRP ISOP was determined at A_{280} by using the molar extinction coefficients of 46370 $M^{-1}cm^{-1}$ (*76*). The presence of FMRP ISOP was analyzed using a 10% tris-glycine sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) and visualized by Coomassie blue staining.

2.6 UV spectroscopy

The melting curves of the RNA oligonucleotides were collected on a Varian Cary 3E spectrophotometer outfitted with a Peltier temperature control cell holder. Experiments were carried out in a 10 mm path length 200 µL quartz cuvette (Starna Cells) using samples prepared as follows in 10 mM cacodylic acid, pH 6.5 up to a final volume of 200 µL. FBS_67 RNA, FBS_Q1 RNA and FBS_Q2 RNA were all prepared by boiling the samples for 10 minutes and allowing them to cool to 25 °C for 30 minutes. FBSsh RNA was prepared and incubated at 25 °C. Samples were heated from 25 °C to 95 °C at a rate of 0.2 °C per minute with points recorded every 1 °C. Samples and reference cells were covered with 200 µL of mineral oil to prevent evaporation of

aqueous solutions at high temperatures. Spectral absorbencies were monitored at either 295 nm or 305 nm, wavelengths that have been identified as being sensitive to G-quadruplex dissociation, depending on the absorbance of the RNA at different concentrations (*80*). All UV thermal denaturation experiments were performed in duplicate. In order to determine the inter- or intramolecular conformation of the G-quadruplex, melting temperatures of various concentrations of RNA between 3 and 50 μM were recorded. Duplex denaturation between MAP1B seq1 RNA and let-7b miRNA was monitored at 275 nm (*81*).

2.7 Circular Dichroism Spectroscopy

CD experiments were performed on a J-810 spectrapolarimeter at 25 °C using a 200 μL quartz cuvette with a 1 mm path length (Starna Cells). Parallel G-quadruplex formation of the RNA oligonucleotides at 10 μM in 10 mM cacodylic acid, pH 6.5, was observed as KCl was titrated in increasing concentrations from a 2 M stock, by monitoring the change in molar ellipticity at 240 nm and 265 nm. Changes in molar ellipticity at 240 nm and 265 nm were monitored for FBSsh RNA as RGG box peptide was titrated into the sample from a 500 μM stock. Each spectrum was scanned 7 times from 200 to 350 nm with a one second response time and a 2 nm bandwidth.

2.8 1D ^{1}H-NMR spectroscopy

The one-dimensional (1D) ^{1}H-NMR spectra of the RNA oligonucleotides were acquired at 25 °C on a 500 MHz Bruker AVANCE spectrometer. Water suppression was carried out by using the Watergate pulse sequence (*82*). Maximum concentrations of

RNA oligonucleotides (200 μM to 1 mM) were prepared in 10 mM cacodylic acid, pH 6.5, in a 90:10 ratio of H_2O:D_2O. G-quadruplex formation was observed by titrating increasing concentrations of KCl from a 2 M stock to each sample and allowing samples to equilibrate for 10 minutes at 25 °C. Samples that were boiled were removed from the NMR tube prior to boiling and replaced after they were cooled to 25 °C. The D_2O exchange experiment was carried out by lyophilizing the FBS_Q2 RNA sample and re-suspending it in 250 μL of pure D_2O, with spectra measured in the interval 10 minutes to 5 days at 25 °C.

The 1D ^{1}H-NMR spectra of MAP1B seq1 RNA and let-7b miRNA duplex formation was observed on a Bruker 900 MHz US^2 spectrometer equipped with a cryoprobe for increased sensitivity at the University of Pittsburgh NMR Core Facility.

2.9 2D ^{1}H-^{1}H NOESY

Two-dimensional (2D) homonuclear NOESY experiments with mixing times of 50 ms were recorded at 25 °C in 90%:10% H_2O:D_2O (*82, 83*). Data sets were processed using XWIN-NMR (Bruker).

2.10 Native polyacrylamide gel electrophoresis (native PAGE)

Non-denaturing PAGE experiments on all RNA oligonucleotides were performed using acrylamide from J.T. Baker and vertical gel apparatus from BioRad. Gels (15 % for FBSsh and FBS_67 RNA and 20% for FBS_Q1 and FBS_Q2 RNA) were visualized by UV shadowing at 254 nm (*84*) or staining in ethidium bromide or SYBR gold and visualized on an AlphaImager (AlphaInnotech). FBS_67 RNA, FBS_Q1 RNA, and

FBS_Q2 RNA were prepared by heating to 95 °C for 5 minutes and allowing samples to cool to 25 °C for 30 minutes. Samples of FBSsh RNA were prepared by incubation at 25 °C for 20 minutes. Gels were run at 4 °C from 40 to 75 V for 4 to 6 hours. FMRP RGG box binding experiments were performed by preparing the RNA sequences as described above and then incubating samples with RGG box peptide in different ratios at 25 °C for an additional 30 minutes. Samples for native PAGE experiments on MAP1B seq1 RNA and let-7b miRNA duplex formation were mixed and allowed to incubate at 25 °C for 20 minutes, loaded onto a 20% gel, and run at 75 V for 6 hours at 25 °C.

2.11 Fluorescence spectroscopy

Steady-state fluorescence spectroscopy experiments of FBSsh_14AP were performed on a Horiba Jobin Yvon Fluoromax-3 and accompanying software fitted with a 150 W ozone-free xenon arc lamp. Experiments were performed in a 150 µL sample volume, 3 mm path-length quartz cuvette (Starna Cells). Excitation wavelength was set to 310 nm, the emission spectrum was recorded in the range of 330 – 450 nm, and the bandpass for excitation and emission monochromators were both set to 3 nm. For binding experiments, the temperature was set to 25 °C, and increasing concentrations of FMRP isoforms (ISO1, ISO2, ISO3, and ISOP) were titrated in 15 nM increments to a fixed RNA concentration of 150 nM. A ratio of 5:1 BSA:RNA was added to the RNA sample before titration began. Emission values were corrected for free protein emission of both BSA and FMRP isoforms, and the data was normalized to free RNA fluorescence intensity monitored at 371 nm. Experiments were performed in triplicate, and the K_d of each individual experiment was calculated by fitting the data to equation 2.1 and

averaged with the other two with the error representing the standard deviation of the average K_d. Data shown is representative of individual experimental data with the average K_d shown. I_F and I_B represent steady-state fluorescence intensities of free and bound RNA, respectively, $[RNA]_t$ is the total fixed RNA concentration, and $[P]_t$ is the total FMRP isoform concentration. The protein-RNA complex dissociation constant, K_d, was determined for each experiment by fitting the binding curve with equation 2.1. Reported errors are standard uncertainties of the averaged data from the best-fit theoretical curves.

$$F = 1 + \left[\frac{I_B}{I_F} - 1\right] \cdot \frac{(K_d + [P]_t + [RNA]_t) - \sqrt{(K_d + [P]_t + [RNA]_t)^2 - 4 \cdot [P]_t \cdot [RNA]_t}}{2 \cdot [RNA]_t} \quad \textbf{2.1}$$

CHAPTER 3: CHARACTERIZATION OF FBS RNA G-QUADRUPLEX STRUCTURE(S)

3.1 Characterization of FBS_67 RNA

FMRP has been shown to interact with its own mRNA by binding to a 100 nucleotide G-rich region named the FMRP binding sequence (FBS), which was proposed to fold into two distinct G-quadruplex structures (*41, 72, 73*) based on potassium-dependent stops of reverse-transcriptase that disappeared when mutations were introduced to FBS RNA. The purine rich FBS region has been found instead to be a potent exonic splicing enhancer whose function is dependent on the presence of the G-rich secondary structures (*72, 73*). The FBS region of *FMR1* mRNA is located near the three acceptor sites of exon 15, and FMRP binding to this region has been shown to control splicing events at exon 15. This direct involvement of FMRP in regulating the production of its minor isoforms created by the usage of exon 15 acceptor sites 2 and 3, which include ISO2 and ISO3, could add a new layer of regulation to the FMRP translation regulator function, as these isoforms lack the major phosphorylation site at amino acid position 500. It is of great interest to directly characterize the presence of (a) G-quadruplex structure(s) within the FBS sequence of *FMR1* mRNA by biophysical methods to determine if FMRP indeed binds to its own mRNA in a G-quadruplex dependent manner. Moreover, it is important to quantify the binding of major and minor isoforms of FMRP to this sequence by biophysical methods, as it has been shown *in vivo* that the major isoform binding to this sequence influences the production of the minor splice variants of the protein (*72*).

In this study, biophysical methods were employed to directly confirm the existence and characterize the fold of the G-quadruplex structures in *FMR1* mRNA. Initially, we truncated FBS RNA to a 67 nt fragment (position 1590 – 1657 within the *FMR1* gene), named FBS_67 RNA, which retained the G-rich region proposed to fold into two G-quadruplex structures. We chose this 67 nt sequence because it retained the two proposed G-quadruplexes plus flanking nucleotides from the original 100 nt RNA and because a shorter 67 nt RNA sequence had a higher chance of being successfully transcribed *in vitro* compared to the original 100 nt sequence. FBS_67 RNA was

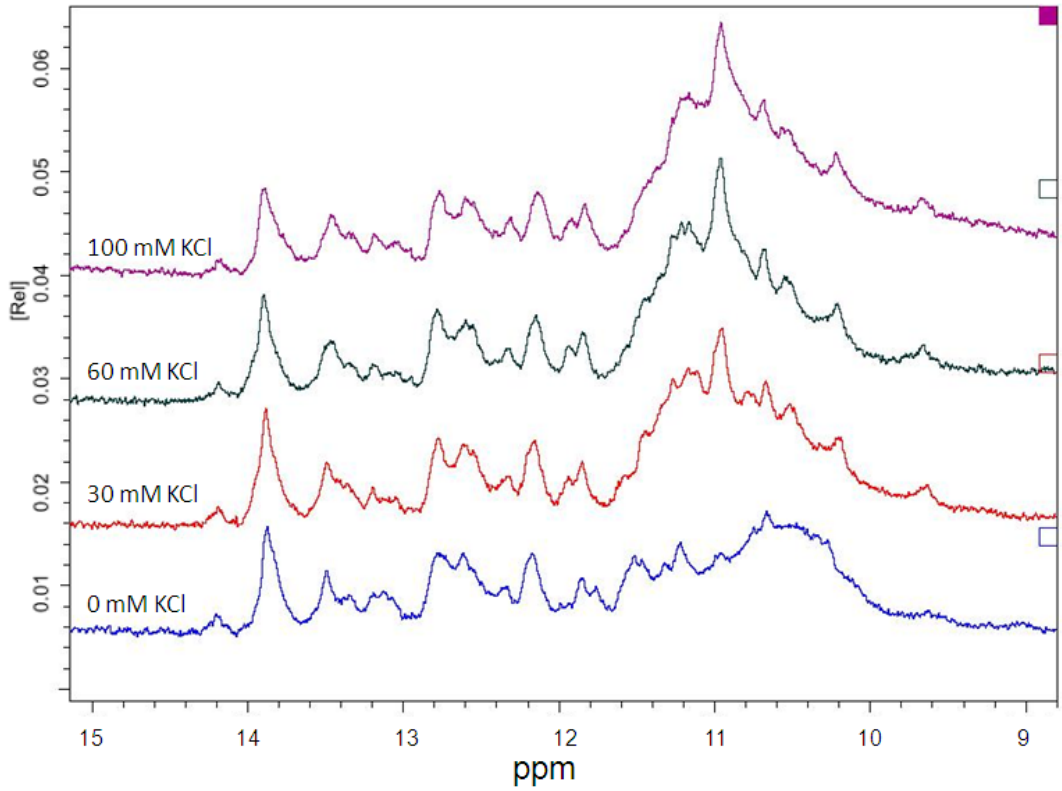


Figure 3.1 The ^{1}H-NMR spectra of FBS_67 RNA confirms the formation of a G-quadruplex in the presence of increasing concentrations of KCl as seen by the increasing of the intensity of the resonances between 10 and 12 ppm. Canonical base pairing, as seen by resonances between 11.5 and 15 ppm, are also consistently present even as the G-quadruplex resonance intensities increase.

produced by *in vitro* transcription reactions off a synthetic DNA template, and 1D ^{1}H-NMR spectroscopy was employed to analyze G-quadruplex formation in this sequence in the absence and presence of KCl. A group of imino proton resonances centered around 11 ppm, which correspond to guanine imino protons involved in G-quartet formation, were observed even in the absence of K^+ ions (Figure 3.1, bottom spectrum). Imino proton resonances were also observed in the region 12 – 15 ppm, corresponding to guanine and uracil imino protons involved in Watson-Crick base pair formation (*85*), indicating also the presence of a duplex and potentially multiple conformations in the structure of FBS_67 RNA. As the KCl concentration was increased in the range 0 – 100 mM, the resonances corresponding to the imino protons involved in G-quadruplex formation increased in intensity and became somewhat more well-defined while the resonances corresponding to canonical base pairing remained constant. These results indicate unambiguously that one or more G-quadruplex structures that are stabilized by K^+ ions are present in FBS_67 RNA.

Circular dichroism (CD) spectroscopy was employed next to gain information about the G-quadruplex fold of FBS_67 RNA. A positive band whose intensity increased upon KCl titration at 265 nm and a negative band at 240 nm were observed, signatures of parallel-type G-quadruplex structures (figure 3.2) (*34, 35*).

Next, we performed non-denaturing polyacrylamide gel electrophoresis (native PAGE) in the presence of increasing KCl concentrations and at different FBS_67 RNA concentrations in order to determine if this sequence formed multiple conformations that migrated at different rates throughout the gel. Several bands were observed for FBS_67 RNA at all KCl concentrations investigated in the range 0 – 100 mM and at all RNA

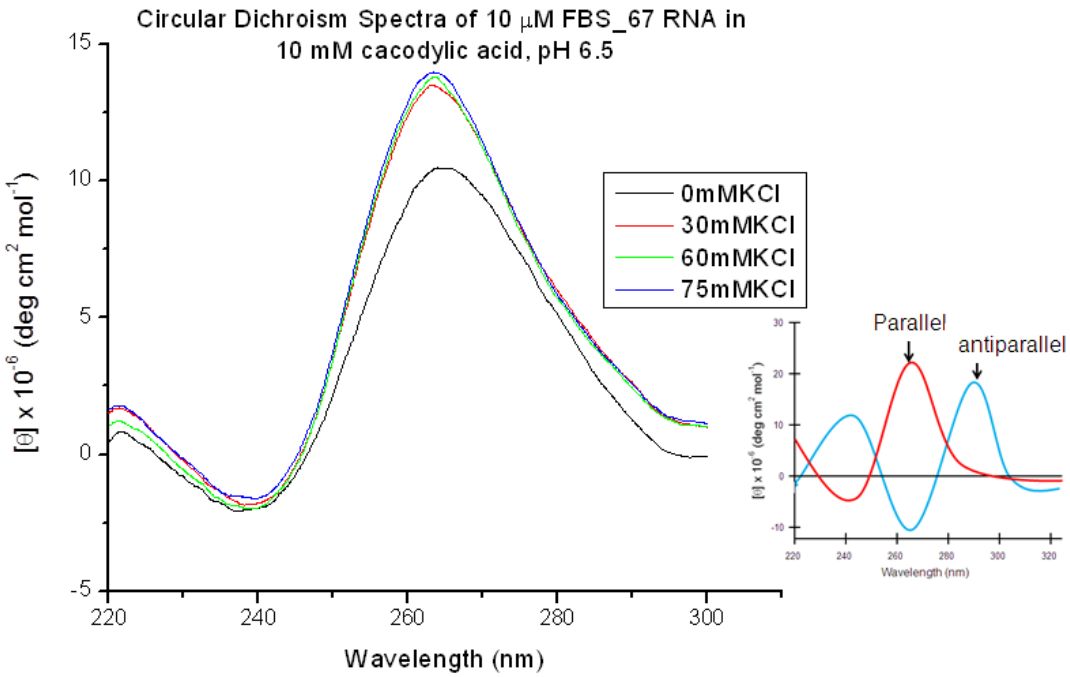
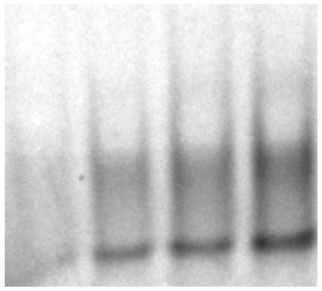
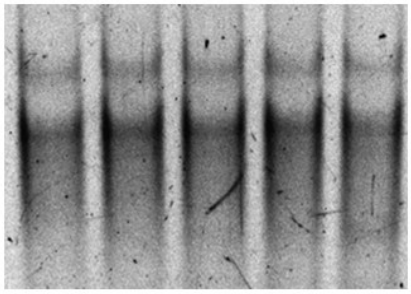


Figure 3.2 Circular dichroism (CD) spectra of FBS_67 RNA upon addition of KCl up to 75 mM. The presence of the positive peak at 265 nm and of the negative peak at 240 nm indicates the formation of a parallel G-quadruplex in the presence of K^+ ions.

concentrations investigated in the range 5 – 20 μM (figure 3.3), confirming that FBS_67 RNA adopts multiple conformations as suggested by the 1D ^{1}H-NMR spectra. The formation of alternate G-quadruplex structures is not uncommon when working with isolated RNA sequences, and in an attempt to solve this problem, two shorter fragments, FBS_Q1 RNA (15 nt, position 1602 – 1616 within the *FMR1* gene) and FBS_Q2 RNA (19 nt, position 1617 – 1635 within the *FMR1* gene), whose sequences were predicted to adopt G- quadruplex structures, were produced by *in vitro* transcription reactions (table 2.1).





RNA (μM)	5	10	15	20
KCl (mM)	50	50	50	50

RNA (μM)	20	20	20	20	20
KCl (mM)	0	25	50	75	100

Figure 3.3 (A) Native polyacrylamide gel electrophoresis (PAGE) of FBS_67 RNA at different RNA concentrations in a fixed concentration of 50 mM KCl. Several conformations exist at each concentration of RNA. (B) Native PAGE of FBS_67 RNA at a fixed concentration of 20 μM in the presence of increasing concentrations of KCl. Several conformations exist in the gel at all KCl concentrations.

3.2 Characterization of FBS_Q1 RNA and FBS_Q2 RNA

Once again, 1D ^{1}H-NMR was employed to investigate G-quadruplex formation within each of the two short RNA sequences, FBS_Q1 RNA and FBS_Q2 RNA. As seen in figure 3.4, even in the absence of KCl, resonances are present in the region 10 – 12 ppm, corresponding to imino protons of guanines involved in G-quartet formation, indicating the presence of a G-quadruplex structure in FBS_Q1 RNA. Upon the titration of KCl up to 25 mM, these imino proton resonances become sharper, indicating that the FBS_Q1 G-quadruplex is stabilized by K^{+} ions. Since the annealing of this sample results in even sharper resonances (figure 3.4, top spectrum), all further experiments were performed with annealed FBS_Q1 RNA samples. No imino proton resonances corresponding to Watson-Crick base pairs were present in the region 12 – 15 ppm, indicating the absence of alternate duplex conformation, which was expected as the

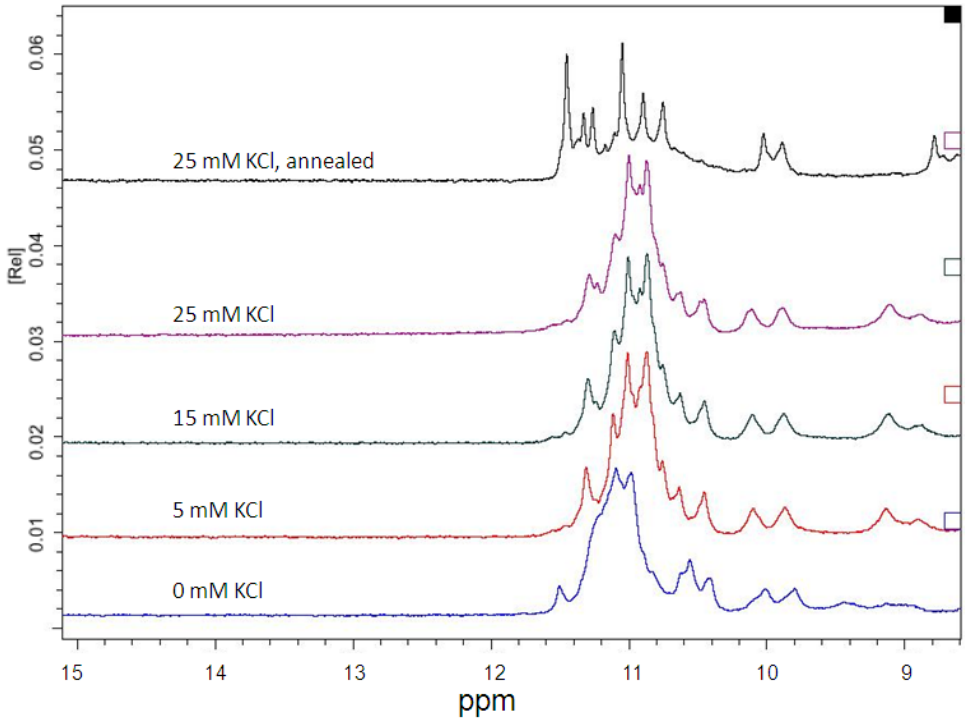


Figure 3.4 1D ^{1}H-NMR spectra of FBS_Q1 RNA in the presence of increasing KCl concentrations. The G-quadruplex (10 – 12 ppm) forms in 0 mM KCl, but the imino proton resonances increase in intensity and sharpness upon the addition of KCl. Amino proton resonances are also present between 8.6 and 10 ppm, and all resonances become sharper once the sample is boiled for 5 minutes and annealed.

sequence is made up entirely of guanines and adenines, which are incapable of forming Watson-Crick base pairs. It is interesting to note the presence of two unusually sharp and downfield guanine amino resonances at 9.9 ppm and 10.0 ppm. Sharp and downfield shifted guanine amino proton resonances observed in G-quadruplex forming sequences containing 5'-GGAGG-3' stretches have been attributed to the presence of an A:(G:G:G:G):A hexad in the structure in which two guanines have both of their amino protons hydrogen bonded, one to a neighboring guanine in the G-quartet and the second to the N7 of the adenine to form a hexad (figure 3.5) (*86-90*). A second set of sharp and

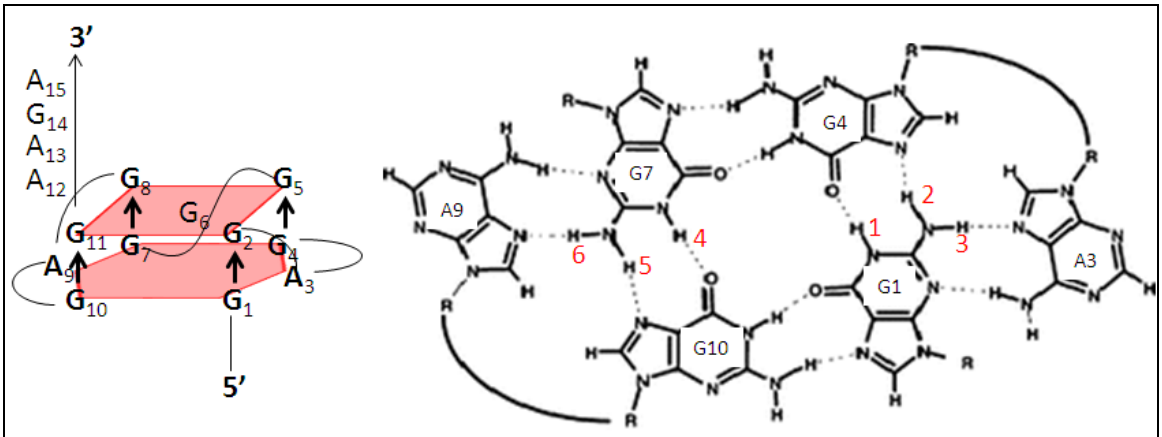


Figure 3.5 An RNA hexad is formed by two adenines that are in the plane with the G-quartet. The formation of a hexad stabilizes the amino protons that are involved in Hoogsteen base pairing within the G-quartet.

downfield shifted amino protons were observed at 8.60 ppm and 8.80 ppm, and a broader one at 8.73 ppm. 2D ^{1}H-^{1}H homonuclear Overhauser enhancement NMR spectroscopy (NOESY) experiments revealed strong NOE cross peaks between the two hydrogen bonded amino protons in each of the two guanines involved in the formation of the hexad: for the first guanine (10.0 ppm; 8.8 ppm), (11.5 ppm; 10.0 ppm) and (11.5 ppm; 8.8 ppm) and for the second guanine (9.9 ppm; 8.6 ppm), (11.4 ppm; 9.9 ppm) and (11.4 ppm; 8.6 ppm) (figure 3.6). These results are consistent with the presence of a hexad structure in which both amino protons of two guanines are involved in hydrogen bonding. Furthermore, proton-deuterium exchange experiments revealed that the same set of imino protons (at 11.5 ppm and 11.3 ppm) and amino proton pairs (at 10.0 ppm, 8.8 ppm and 9.9 ppm, 8.6 ppm, respectively) for which strong NOE cross peaks were observed in the NOESY experiments exchange very slowly (hours or days) as compared to the rest of the imino protons present in the spectrum (seconds or minutes) (figure 3.7). This exchange

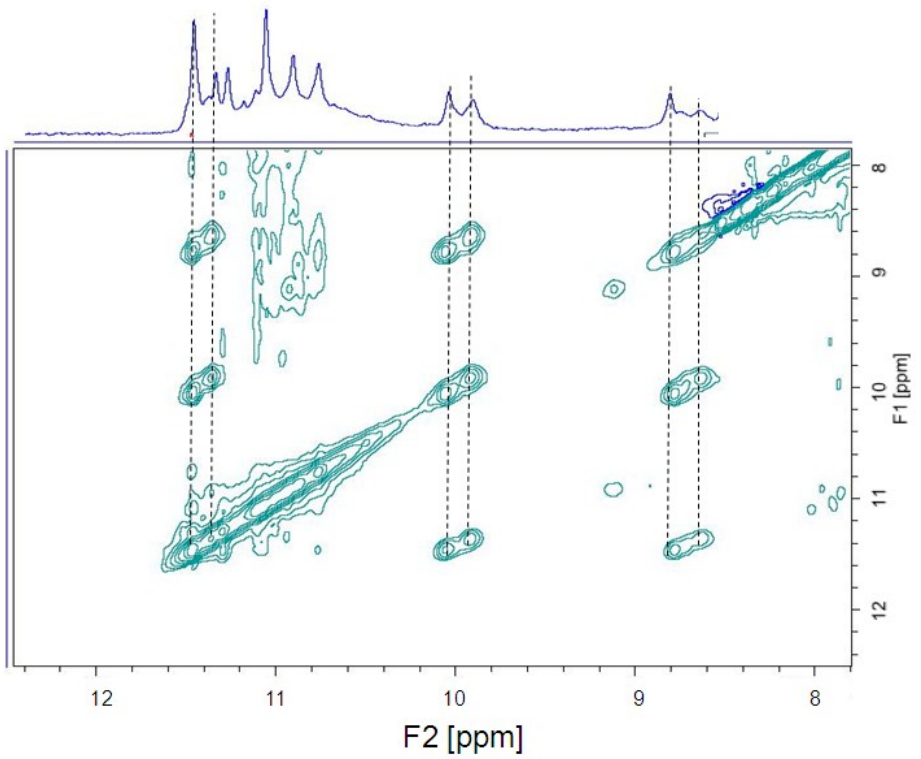


Figure 3.6 2D ^{1}H-^{1}H-NOESY of FBS_Q1 RNA in the presence of 25 mM KCl. The dotted lines show the relationship between the NOEs and the 1D ^{1}H-NMR spectrum of FBS_Q1 RNA.

with the solvent, which requires the base pairs to open transiently, is slowed down considerably when both amino protons of some of the guanines are involved in hydrogen bonding, as predicted in the hexad structure (figure 3.5). Taken together the 1D ^{1}H-NMR spectra, the 2D ^{1}H-^{1}H NOESY spectrum, and the1D deuterium exchange NMR spectra strongly suggest that the FBS_Q1 RNA sequence adopts a G-quadruplex structure that contains an A:(G:G:G:G):A hexad.

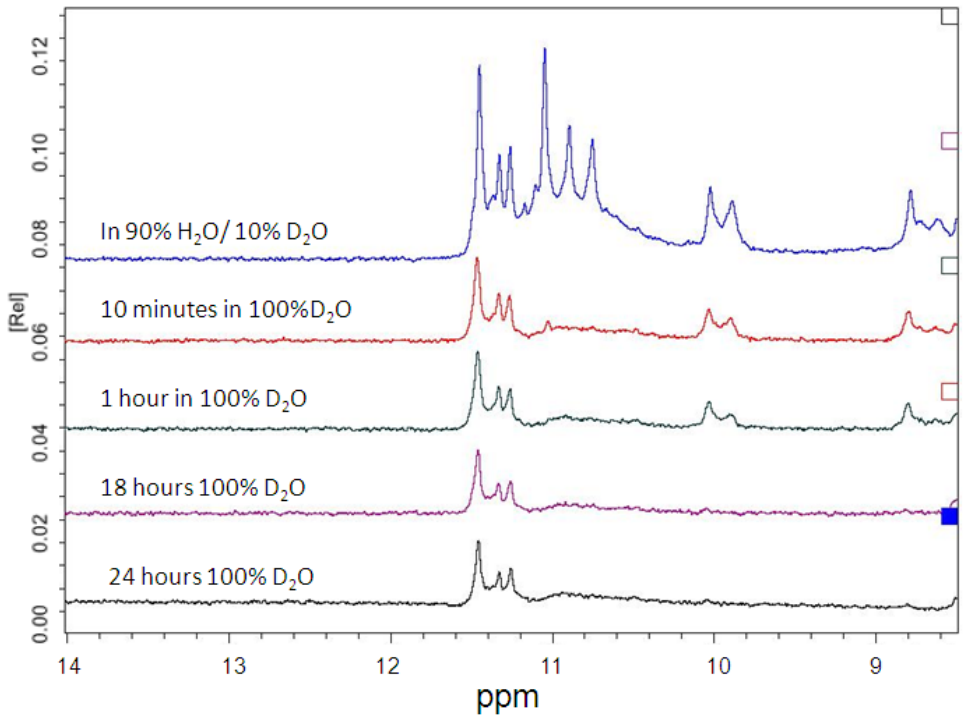


Figure 3.7 1D ^{1}H-NMR spectra of FBS_Q1 RNA, which was lyophilized and re-suspended in deuterium oxide (D_2O).

When FBS_Q2 was analyzed by 1D ^{1}H-NMR spectroscopy, a set of resonances corresponding to the imino protons of guanines and uracils involved in Watson-Crick base pairs were observed in the region 12 – 14 ppm at all KCl concentrations investigated, whereas the characteristic G-quadruplex imino proton resonances centered around 11 ppm were not observed until 25 mm KCl had been titrated into the sample (figure 3.8). The G-quadruplex imino proton resonances increased in intensity as more KCl was titrated up to 60 mM KCl in FBS_Q2 RNA, but this titration did not affect the intensity of the Watson-Crick base pairs, suggesting that FBS_Q2 RNA adopts alternate conformations involving Watson-Crick base pairing independent of the G-quadruplex structure forming upon the addition of KCl.

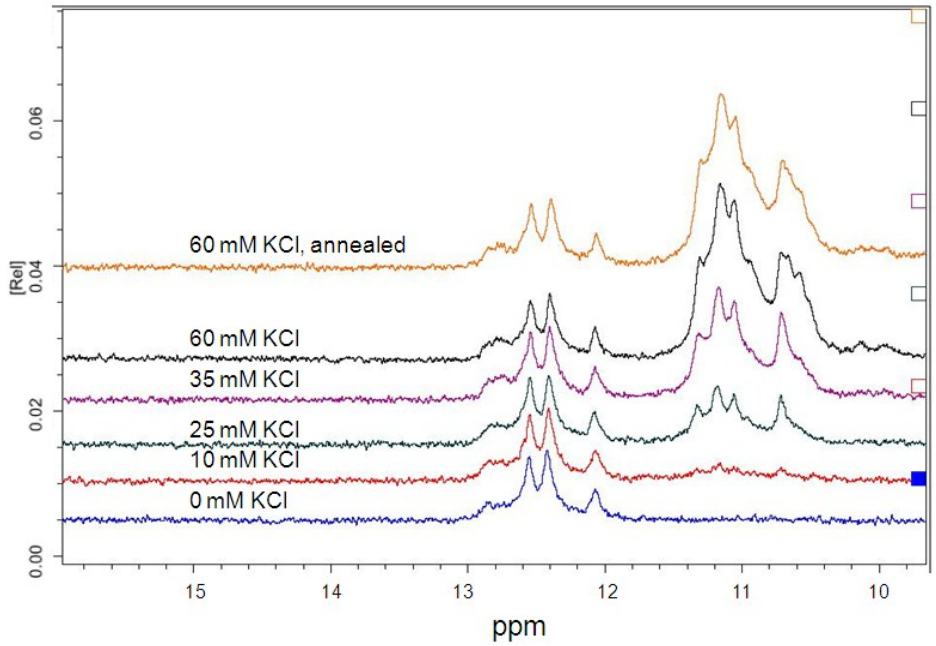
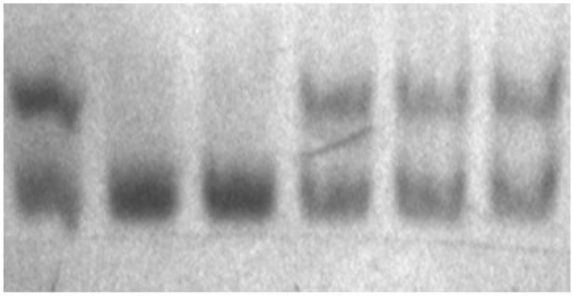


Figure 3.8 1D ^{1}H-NMR spectra of FBS_Q2 RNA at varying concentrations of KCl. The G-quadruplex region of the spectra between 10 and 12 ppm appears and increases upon increasing concentrations of KCl up to 60 mM. Watson-Crick base pairing can be seen as imino proton resonances between 12 and 14 ppm that do not change upon KCl titration.

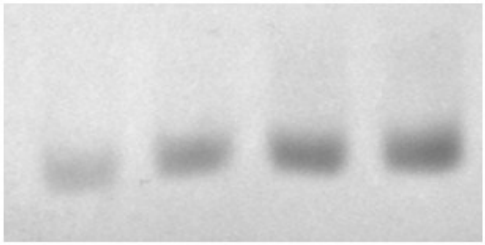
To investigate if each of the two sequences, FBS_Q1 RNA and FBS_Q2 RNA, exists in a single conformation at various KCl concentrations, native PAGE was employed. FBS_Q1 RNA was observed to form a single band on the native PAGE at 25 mM KCl and 50 mM KCl while it was observed to form two distinct conformations at 0 mM KCl as well as at KCl concentrations in the range 75 – 150 mM (figure 3.9 A). At 25 mM KCl, FBS_Q1 RNA was observed to form a single conformation by native PAGE at various concentrations of RNA in the range 10 – 40 μM (figure 3.9 B). The secondary band at higher concentrations of KCl seen in figure 3.9 (A) could represent secondary

A



RNA (μM)	30	30	30	30	30	30
KCl (mM)	0	25	50	75	100	150

B

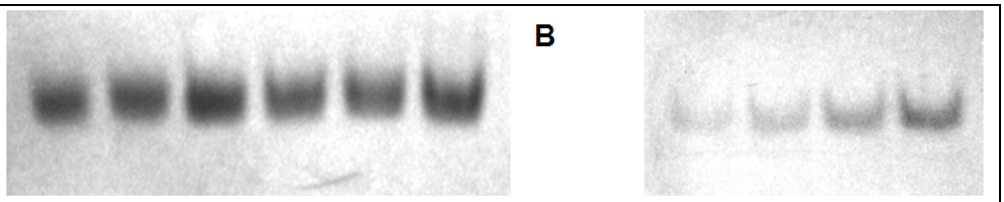


RNA (μM)	10	20	30	40
KCl (mM)	25	25	25	25

Figure 3.9 (A) Native PAGE of FBS_Q1 RNA in the presence of increasing KCl concentrations. The sequence forms more than one conformation at KCl concentrations above 50 mM, but it appears to form only one conformation in the presence of 25 mM KCl and 50 mM KCl. (B) Increasing concentrations of FBS_Q1 RNA forming one conformation in 25 mM KCl.

structure in which the hexad formed by FBS_Q1 RNA stacks upon another hexad in order to form a dimer upon increasing KCl concentrations, a phenomenon commonly observed in hexad forming sequences (*86*).

Although the ^{1}H-NMR spectroscopy results suggested that the sequence of FBS_Q2 RNA formed more than one conformation, one of which involved Watson-Crick base pairing, this RNA migrates as a single band in native PAGE experiments at all KCl



RNA (μM)	30	30	30	30	30	30	RNA (μM)	10	20	30	40
KCl (mM)	0	25	50	75	100	150	KCl (mM)	25	25	25	25

Figure 3.10 (A) Native PAGE of FBS_Q2 RNA in the presence of increasing KCl concentrations and (B) increasing concentrations of RNA at a fixed concentration of 25 mM KCl. The sequence forms only one conformation at KCl concentrations in the range 0 mM to 150 mM and RNA concentrations in the range 10 – 40 μM.

concentrations in the range 0 – 150 mM (figure 3.10). Thus, it is possible that the alternate conformations adopted by FBS_Q2 RNA cannot be distinguished based on their migration pattern through native PAGE experiments.

To gain information about the G-quadruplex fold of FBS_Q1 RNA and FBS_Q2 RNA, CD spectroscopy was once again utilized (figure 3.11). As expected from the ^{1}H-NMR data that showed a G-quadruplex forming in the absence of KCl, FBS_Q1 RNA has the signature positive band at 265 nm and a negative band at 240 nm of a parallel G-quadruplex structure in the absence of KCl. Even upon the titration of increasing KCl concentrations, the signature bands of parallel G-quadruplex formation remain constant (figure 3.11 A). The ^{1}H-NMR spectra of FBS_Q2 RNA showed the absence of resonances in the G-quadruplex region in the absence of KCl and the increase in G-quadruplex character upon the titration of KCl into the sequence. Similarly, the CD spectrum of FBS_Q2 RNA indicates that a parallel G-quadruplex is forming only upon the titration of KCl up to 25 mM with a positive band at 265 nm and a negative band at 240 nm (figure 3.11 B). In the absence of KCl, the positive band is centered at 270 nm, a

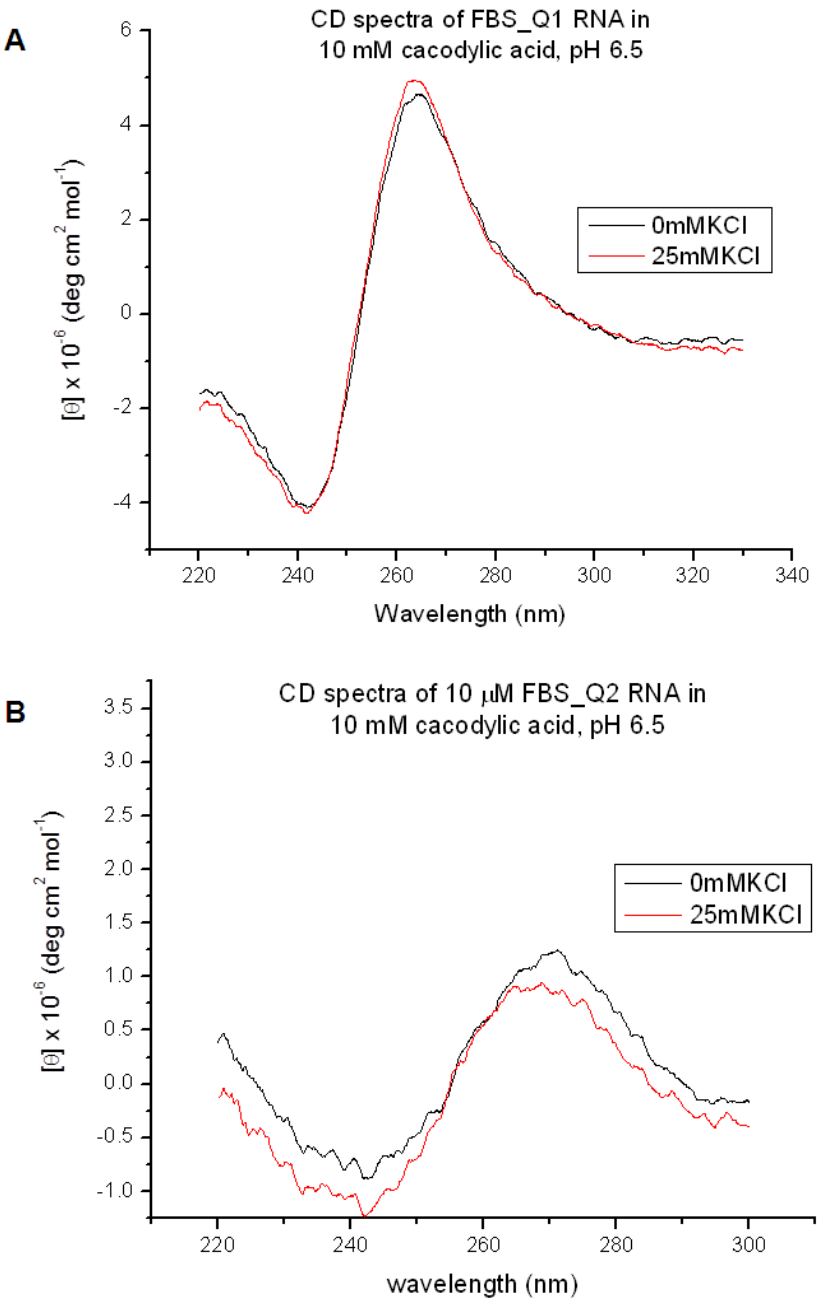
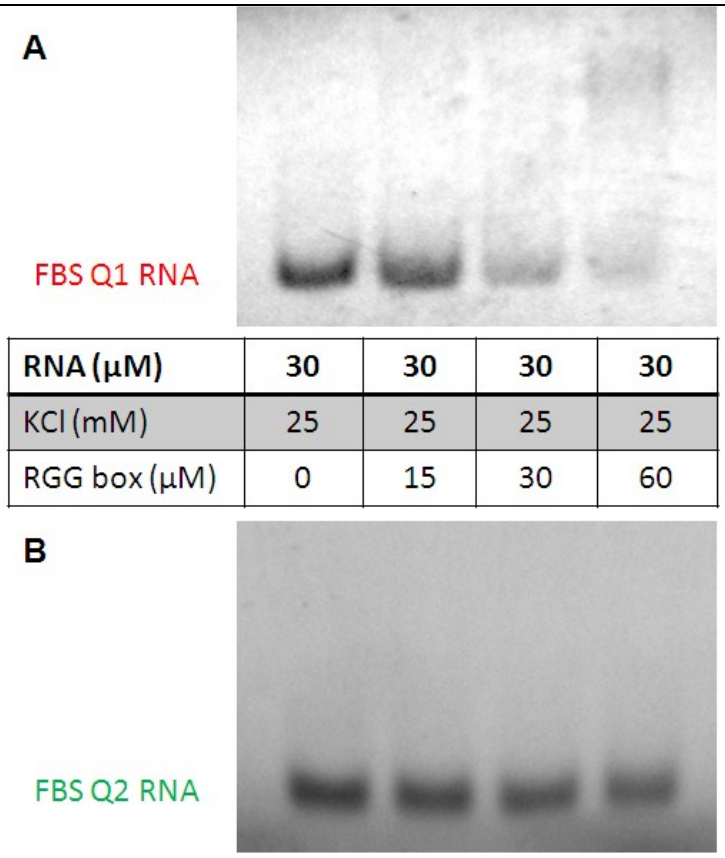


Figure 3.11 Circular dichroism (CD) spectra of FBS_Q1 RNA and FBS_Q2 RNA. (A) FBS_Q1 RNA upon titration of 25 mM KCl shows the positive band at 265 nm and the negative band at 240 nm, indicating the formation of a parallel G-quadruplex in the absence and presence of K^+ ions. (B) CD spectra of FBS_Q2 RNA upon the titration of 25 mM KCl shows a positive peak at 265 nm and a negative band at 240 nm, indicating that FBS_Q2 RNA forms a G-quadruplex upon the addition of K^+ ions. At 0 mM KCl, the positive band is observed at 270 nm, a signature of A-form duplex formation.

signature of A-form duplex formation within RNA (*91*), consistent with the ^{1}H-NMR data in figure 3.8.

To determine if the formation of parallel G-quadruplexes within the short sequences of FBS_Q1 RNA and FBS_Q2 RNA caused them to be targets of FMRP RGG box binding, native PAGE was once again employed (Figure 3.12). The synthetic FMRP



RNA (μM)	30	30	30	30
KCl (mM)	25	25	25	25
RGG box (μM)	0	15	30	60

Figure 3.12 Native PAGE of (A) FBS_Q1 RNA and (B) FBS_Q2 RNA in the presence of increasing concentrations of RGG box in 25 mM KCl.

RGG box peptide has been shown to have high binding affinity to G-quadruplex forming RNA sequences (*43-45*). The binding between RNA and the FMRP RGG box is

observed as a disappearance of the free RNA band, as the RGG box peptide is positively charged, causing the RNA-peptide complex to migrate towards the cathode, and a distinct band representing the complex will not be observed. When the FMRP RGG box is titrated into FBS_Q1 RNA up to a 2:1 ratio of peptide:RNA, the free RNA band disappears, indicating that the RGG box has high binding affinity for FBS_Q1 RNA (Figure 3.12 A). However, when the RGG box peptide is titrated into FBS_Q2 RNA in the same ratios, the free RNA band remains at nearly the same intensity, indicating that the RGG box does not have high binding affinity for FBS_Q2 RNA alone (Figure 3.12 B). This could be explained by the fact that the ^{1}H-NMR spectra of FBS_Q2 RNA suggested that multiple conformations existed in this sequence, but all possible conformations migrated at the same rate along the native PAGE. One of these conformations may not have high binding affinity for the RGG box peptide, but it is not possible to distinguish between these conformations in the binding gel.

Although FBS_Q2 RNA did not have high binding affinity for the RGG box peptide, the G-quadruplex formed by both FBS_Q1 and FBS_Q2 are proposed to exist simultaneously and adjacent to one another within the FBS sequence of *FMR1* mRNA. The short FBS_Q1 RNA was found to form a G-quadruplex that also contains an A:(G:G:G:G):A hexad. However, it is not clear if this hexad formation is induced by the short FBS_Q1 RNA sequence or if it is also present in a longer FBS stretch. To test this, FBS_Q1 RNA and FBS_Q2 RNA were combined into a 42 nt sequence named FBSsh RNA, which retained both proposed G-quadruplexes but was shorter than the initial FBS_67 RNA, making it more likely that this sequence would form a single conformation (Table 2.1).

3.3 Characterization of FBSsh RNA

To investigate the formation of a G-quadruplex within the FBSsh RNA sequence, 1D ^{1}H-NMR was once again employed. As seen in figure 3.13, even at 0 mM KCl, a broad resonance is present in the region centered at 11 ppm, corresponding to imino protons of guanines involved in G-quartet formation, whereas, upon the addition of KCl up to 25 mM, more defined resonances develop under this broad envelope, indicating that the G-quadruplex structures of FBSsh RNA are further stabilized in the presence of K^{+} ions. Nonetheless, at all KCl concentrations investigated, the G-quadruplex imino proton resonances remain broad, which could be due to the presence of multiple G-quadruplex structures in FBSsh RNA sequence. This may be because the imino proton resonances from each of the proposed G-quadruplexes resonate in the same region of the ^{1}H-NMR spectrum, and if two exist simultaneously, their resonances will overlap in the region centered at 11 ppm. Interestingly, no sharp amino proton resonances, signatures of the A:(G:G:G:G):A hexad, were observed around 10 ppm, indicating that the hexad structure is not formed in FBSsh RNA, likely being induced in the isolated FBS_Q1 RNA by the short length of the sequence. Additionally, in contrast with the 1D ^{1}H-NMR spectra of FBS_Q2 RNA, no resonances appear in the Watson-Crick base pair region of the ^{1}H-NMR spectrum of FBSsh RNA, indicating that this sequence does not form any alternate conformations that do not involve G-quadruplexes. This confirmed that FBS_Q1 and FBS_Q2 RNA sequences became more stable G-quadruplexes once incorporated into a longer sequence with each other. Thus, all subsequent characterization experiments were performed with the 42 nt FBSsh RNA folded in the presence of 25 mM KCl.

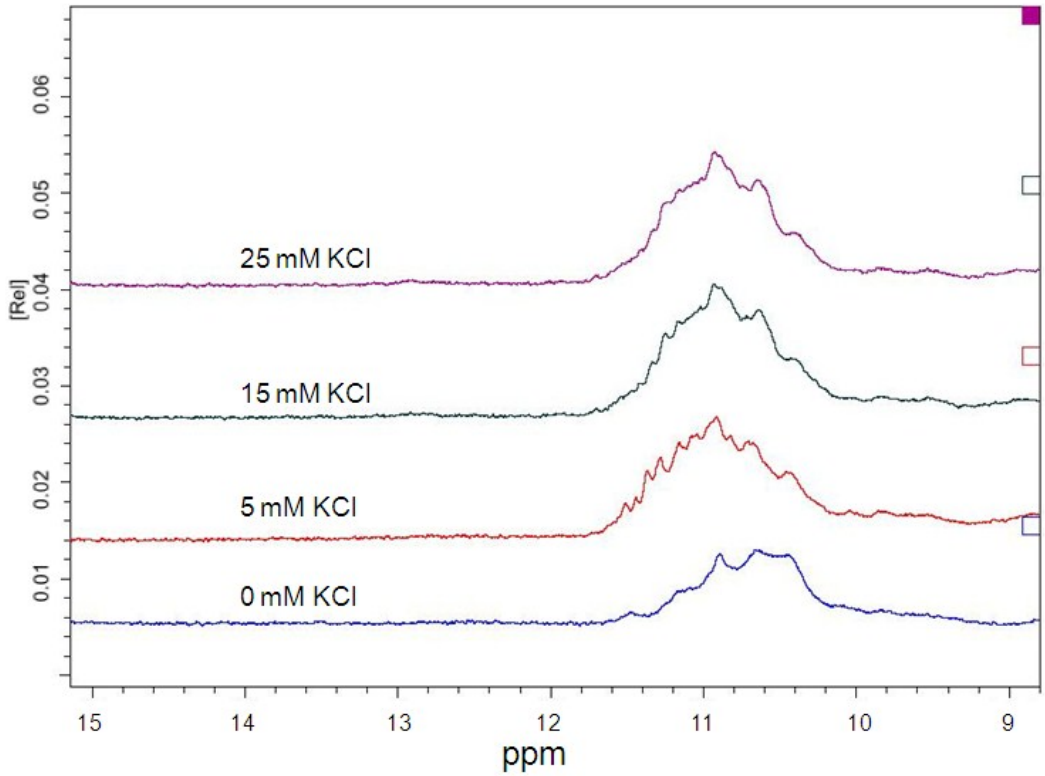
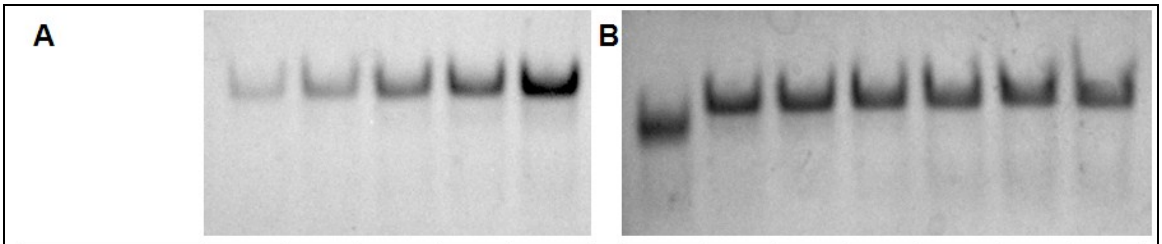


Figure 3.13 1D ^{1}H-NMR spectra of FBSsh RNA at varying concentrations of KCl. Resonances are present in the G-quadruplex imino proton region between 10 and 12 ppm even at 0 mM KCl, their intensity increasing with increasing KCl concentrations. The resonances are broad in nature, indicating that this sequence is dynamic and may contain two distinct G-quadruplexes, FBS_Q1 RNA and FBS_Q2 RNA.

To confirm that the 42 nt sequence of FBSsh RNA exists in a single conformation native PAGE was employed once again (figure 3.14). At 0 mM KCl, FBSsh RNA migrates in the gel as a single band, which changes position with increasing salt concentrations up to 150 mM KCl (figure 3.14 B), consistent with the 1D ^{1}H-NMR spectra. Similarly, a single band is present when several concentrations of FBSsh RNA in the range 5 – 30 μM were run on a native PAGE in a fixed concentration of 25 mM KCl (figure 3.14 A).



RNA (μM)	5	10	15	20	30	20	20	20	20	20	20	20
KCl (mM)	25	25	25	25	25	0	5	15	25	50	100	150

Figure 3.14 Native PAGE of FBSsh RNA in the presence of (A) varying concentrations of RNA and (B) varying concentrations of KCl.

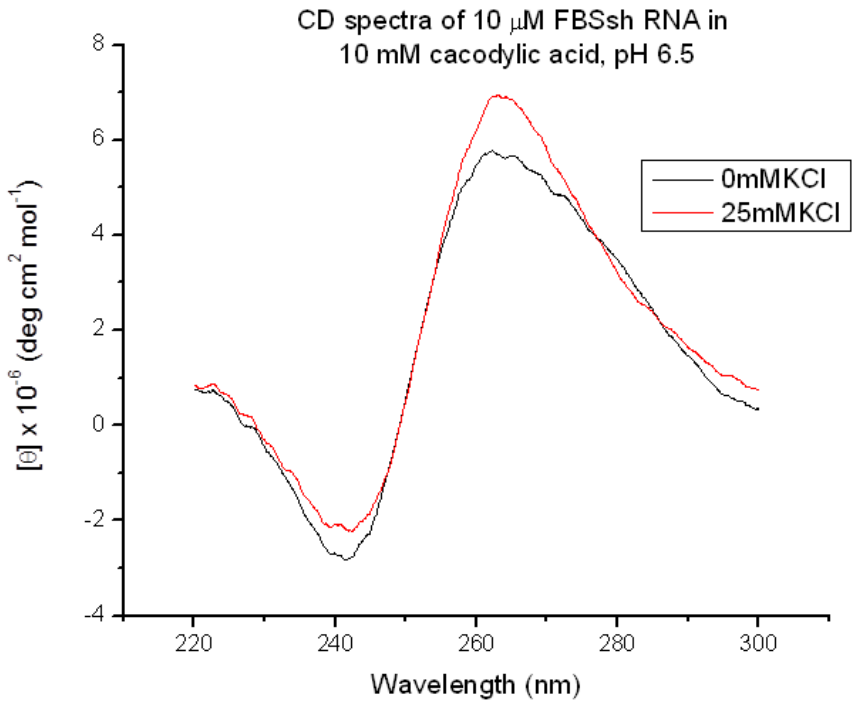


Figure 3.15 Circular dichroism (CD) spectra of FBSsh RNA upon addition of 25 mM KCl. The increasing positive band at 265 nm and the negative band at 240 nm indicates the formation of (a) parallel G-quadruplex(es) in the presence of K^+ ions.

To gain additional information about the fold of the G-quadruplexes within FBSsh RNA, we used CD spectroscopy. The CD spectrum of FBSsh RNA, shown in figure 3.15, shows a negative band at 240 nm and a positive band at 265 nm that increase in intensity upon the titration of KCl from 0 to 25 mM. This result, which indicates that one or more parallel G-quadruplex structures are present in FBSsh RNA, is consistent with the ^{1}H-NMR spectra of FBSsh RNA, which showed an increase in the intensities of the G-quartet imino proton resonances upon the titration of K^{+} ions.

Next, we employed thermodynamic methods to determine if FBSsh RNA folds into an intermolecular or intramolecular G-quadruplex structure. Thus, UV thermal denaturation was used to determine the melting temperatures of the FBSsh G-quadruplex structures at various RNA concentrations in the presence of 25 mM KCl. If a species contains *n* number of strands, its melting temperature, T_m, depends on the total RNA concentration, c_T (*92*). In the case of intermolecular G-quadruplex formation, the plot of 1/Tm (°C-1) versus ln[RNA] is a straight line fitting equation 3.1, where R is the Boltzmann constant (1.986 cal K^{-1} mol^{-1}) and $\Delta H°_{vH}$ and $\Delta S°_{vH}$ are the two Van't Hoff free energy components enthalpy and entropy, respectively.

$$\frac{1}{T_m} = \frac{R(n-1)}{\Delta H°_{vH}} \ln c_T + \frac{\Delta S°_{vH} - (n-1)R\ln 2 + R\ln n}{\Delta H°_{vH}} \qquad \textbf{3.1}$$

However, if a species is intramolecular, the T_m does not depend upon RNA concentration. For intramolecular species, n = 1 strand, T_m is independent of c_T, and the slope of the straight line resulting from plotting $1/T_m$ ($°C^{-1}$) versus ln[RNA] is zero, simplifying equation 3.1 to equation 3.2.

$$\frac{1}{T_m} = \frac{\Delta S^\circ_{vH}}{\Delta H^\circ_{vH}}$$

3.2

The T_m of each of the transitions in the UV thermal denaturation of FBSsh RNA at each RNA concentration was calculated by fitting the curves to equation 3.3 and rearranging the calculated ΔH°_{vH} and ΔS°_{vH} into equation 3.2 to determine T_m.

$$A(T) = \frac{A_U + A_F e^{-\Delta H^\circ/RT} e^{\Delta S^\circ/R}}{e^{-\Delta H^\circ/RT} e^{\Delta S^\circ/R} + 1}$$

3.3

In equation 3.3, A_U is the calculated absorbance of denatured RNA while A_F is the calculated absorbance of folded RNA. The experiments were recorded at 295 nm, the wavelength that has been shown to be sensitive to G-quadruplex denaturation (*80*). The UV thermal denaturation of FBSsh RNA at 295 nm in the presence of 25 mM KCl revealed two distinct hypochromic transitions characteristic of the denaturing of a G-quadruplex, the first with a $T_m \sim 43$ °C and the second with a $T_m \sim 70$ °C (figure 3.17 A). This result is consistent with the presence of two distinct G-quadruplex structures in FBSsh RNA, which was expected considering that the sequence of this RNA has the potential to form two separate G-quadruplexes. Figure 3.16 shows a model of the most likely conformation of two individual parallel G-quadruplexes proposed to form by FBSsh RNA as the combination of FBS_Q1 RNA and FBS_Q2 RNA.

Increasing concentrations of FBSsh RNA in the range 3 to 30 µM were thermally denatured, monitoring the decrease in absorbance at 295 nm. Each of the two transitions were fitted to equation 3.3 for each experiment at each concentration of RNA (figure 3.17

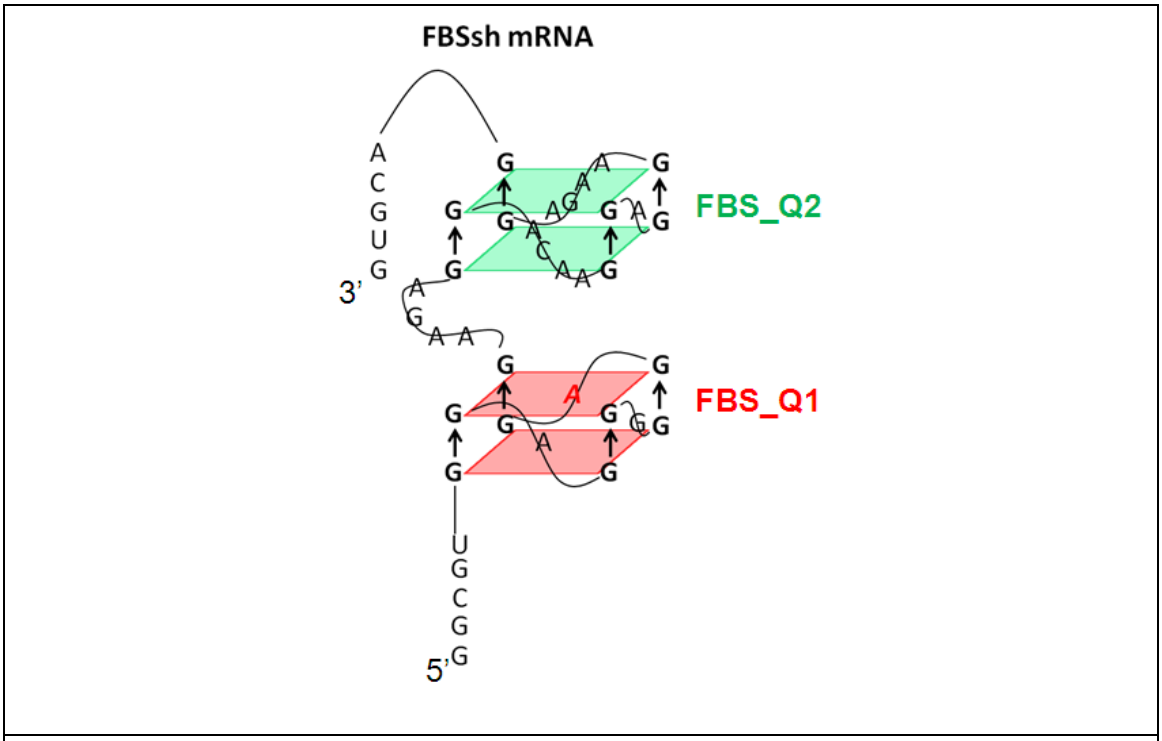


Figure 3.16 Model of FBSsh RNA as a combination of two G-quadruplexes formed by FBS_Q1 RNA and FBS_Q2 RNA.

B and C). The calculated melting points (equation 3.2) of each transition were plotted against the RNA concentration (figure 3.17 D) to show that each of the two melting points of FBSsh RNA were independent of the RNA concentration. A plot of the inverse melting temperature versus $\ln c_T$ (figure 3.17 E) also yields two linear curves with a slope of 0, indicating that n = 1 for each transition and that FBSsh RNA forms two intramolecular G-quadruplexes.

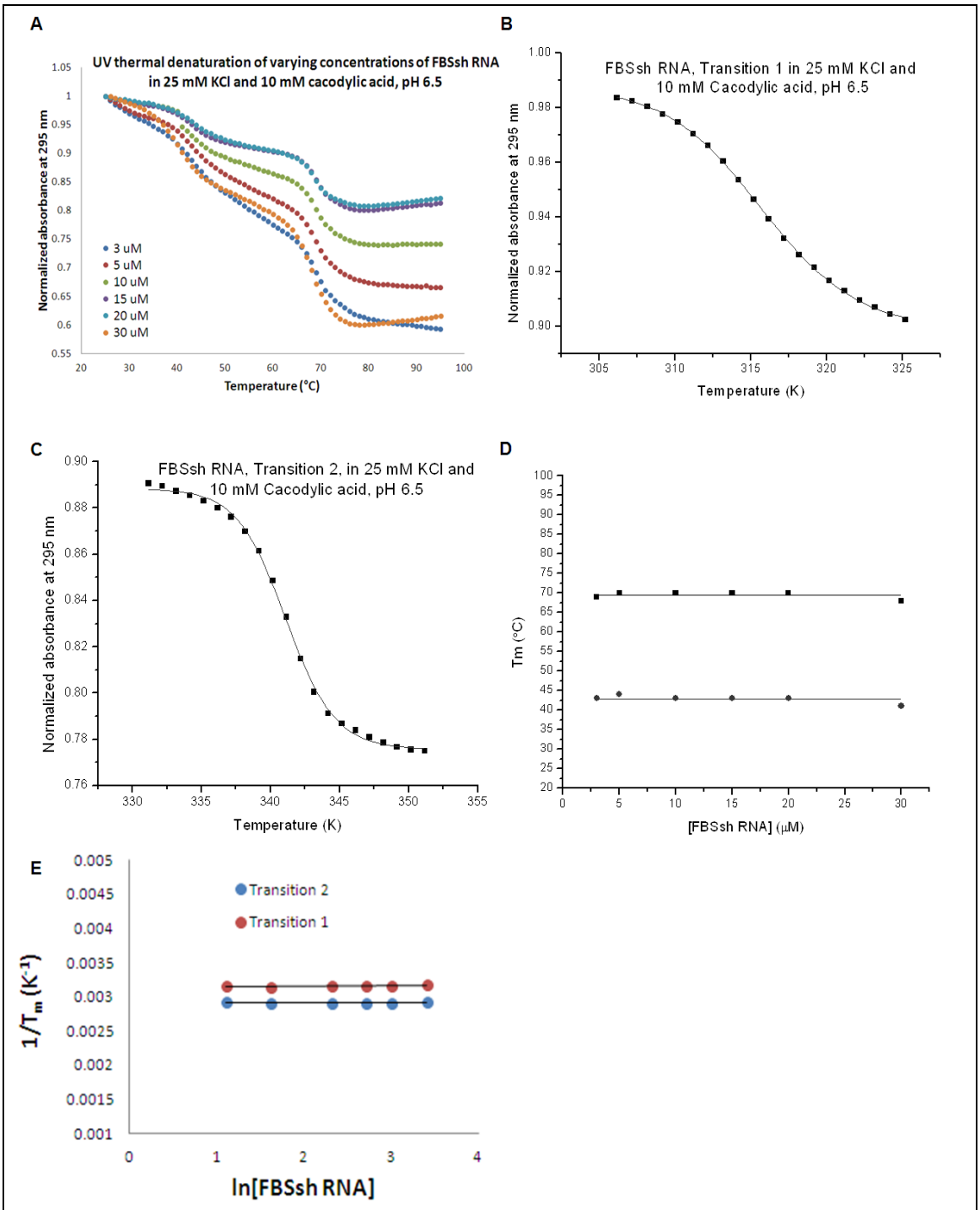


Figure 3.17 A) UV thermal denaturation curve of several concentrations of FBSsh RNA from 3 to 30 μM in 10 mM cacodylic acid, pH 6.5, containing 25 mM KCl showing two hypochromic transitions. The G-quadruplex transitions best fit to equation 3.3 for B) transition 1 and C) transition 2. D) The melting temperatures of both G-quadruplex structures of FBSsh RNA are independent of the RNA concentration, and E) the plot of 1/Tm vs ln[FBSsh RNA] yield straight lines at each transition with a slope equal to zero.

The thermodynamic parameters (standard enthalpy, entropy, and free energy), shown in table 3.1, for the formation of the FBSsh RNA G-quadruplex structures in 25 mM KCl were determined by fitting the G-quadruplex dissociation transitions present in the UV melting curves to equation 3.3, which assumes a two-state model (figure 3.17 B and C) (*35*).

Table 3.1 Thermodynamic Parameters for G-quartet Structure Formation

Molecule	Tm (°C)	ΔH°f (kcal/mol)	ΔS°f (cal/mol•K)	ΔG°f (kcal/mol)
FBSsh, Transition 1	42.6 ± 0.1	-61.9 ± 0.1	-196.3 ± 0.1	-3.45 ± 0.01
FBSsh, Transition 2	68.8 ± 1.0	-105.8 ± 0.1	-305.9 ± 1.3	-13.6 ± 0.3
FBSsh_14AP	69.6 ± 0.1	-88.7 ± 0.3	-258.8 ± 0.8	-11.6 ± 0.1

Based solely on the UV thermal denaturation experiments, it was not possible to assign the two hypochromic transitions to the FBS_Q1 and FBS_Q2 G-quadruplexes within FBSsh RNA. To solve this problem, we constructed a fluorescently labeled RNA that reports only on the melting of the first quadruplex in the sequence, FBS_Q1 RNA, by replacing the adenine at position 14 of FBSsh RNA with 2-aminopurine (2AP) (FBSsh_14AP RNA), shown in red in figure 3.16. 2AP is a highly fluorescent analog of adenine, which is sensitive to changes in its microenvironment (*93, 94*). Its substitution has also been shown not to perturb the formation of G-quadruplexes (*10, 43*). In contrast to UV thermal denaturation in which the change in absorbance as a function of temperature has contributions from all bases in the sequence, thermal denaturation using fluorescence spectroscopy monitors only the changes in the steady-state fluorescence of

the 2AP reporter. Because the 2AP fluorophore is located within FBS_Q1, the first G-quadruplex of FBSsh RNA, in these experiments, only the melting of this G-quadruplex will be observed. FBSsh_14AP RNA was melted in the presence of 25 mM KCl (figure 3.18), and at each temperature point the steady-state fluorescence emission at 371 nm was corrected to account for the known dependence of the free 2AP emission at 371 nm upon temperature (section 2.11). As expected, a single transition was

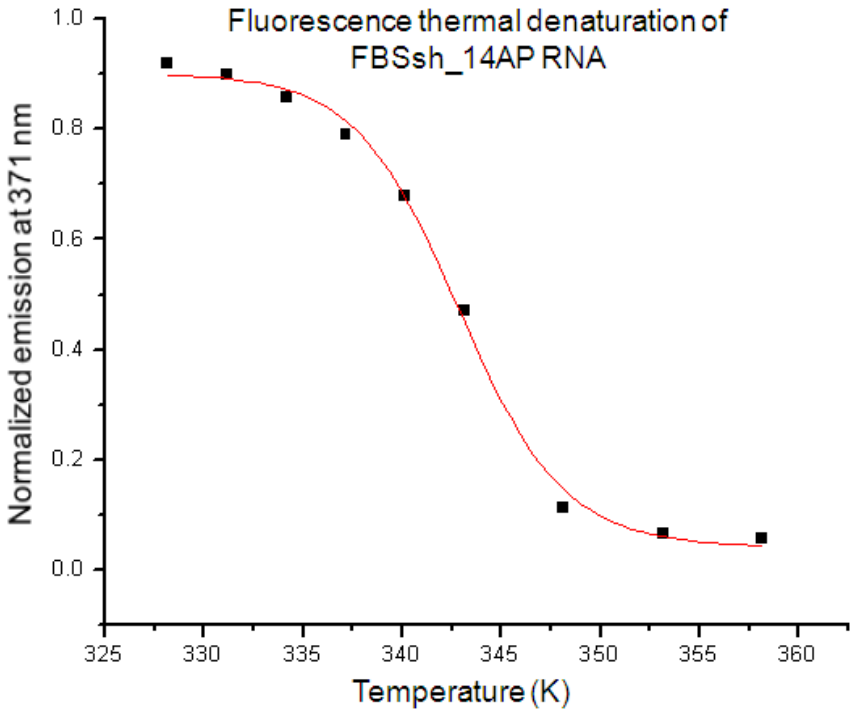


Figure 3.18 Fluorescence spectroscopy thermal denaturation of FBSsh_14AP RNA in the presence of 25 mM KCl. The transition seen here corresponds to transition 2 from the UV thermal denaturation of unlabeled FBSsh RNA.

observed for FBSsh_14AP, and upon fitting it with equation 3.3, a melting point of 69.4 ± 0.3 °C was calculated (equation 3.2). This result is in very good agreement with the T_m

value of 68.8 ± 1.0 °C determined for transition 2 by the UV thermal denaturation experiments. Thus, we assign transition 2 to the melting of the FBS_Q1 G-quadruplex and transition 1 to the melting of the FBS_Q2 G-quadruplex within FBSsh RNA. These results are consistent with our knowledge that the length of the loops connecting the G-quartet planes affects the stability of the G-quadruplex structure, with shorter loops forming tighter, more thermodynamically stable structures (*35*). Since each of the connecting loops in the FBS_Q1 have a single nucleotide, this structure was expected to be more thermodynamically stable and thus have a higher melting point than the second G-quadruplex, FBS_Q2, which has longer connecting loops on two sides of its structure (figure 3.16).

Next, the number of K^+ ions coordinated to each of the G-quadruplexes was determined by performing UV thermal denaturation at a fixed FBSsh RNA concentration of 10 μM and variable KCl concentrations in the range 5 – 150 mM. To calculate ΔG° of each of the two transitions at each KCl concentration, the curves of each of the two transitions were again fitted to equation 3.3, and ΔG° was calculated using the Gibbs free energy equation where T was set to 298.15 K. The resulting ΔG° at each KCl concentration and was plotted against log[KCl] according to equation 3.4 (figure 3.19 A and B).

$$\Delta n = \frac{d \ln K_{eq}}{d \ln [K^+]} = -\frac{\Delta\Delta G^\circ}{2.3RT\Delta \log [K^+]} \qquad \textbf{3.4}$$

The number of K^+ ions coordinated to each G-quadruplex was calculated from the negative slope of the plots (figure 3.19 A and B) to be 3 K^+ ions for each G-quadruplex in

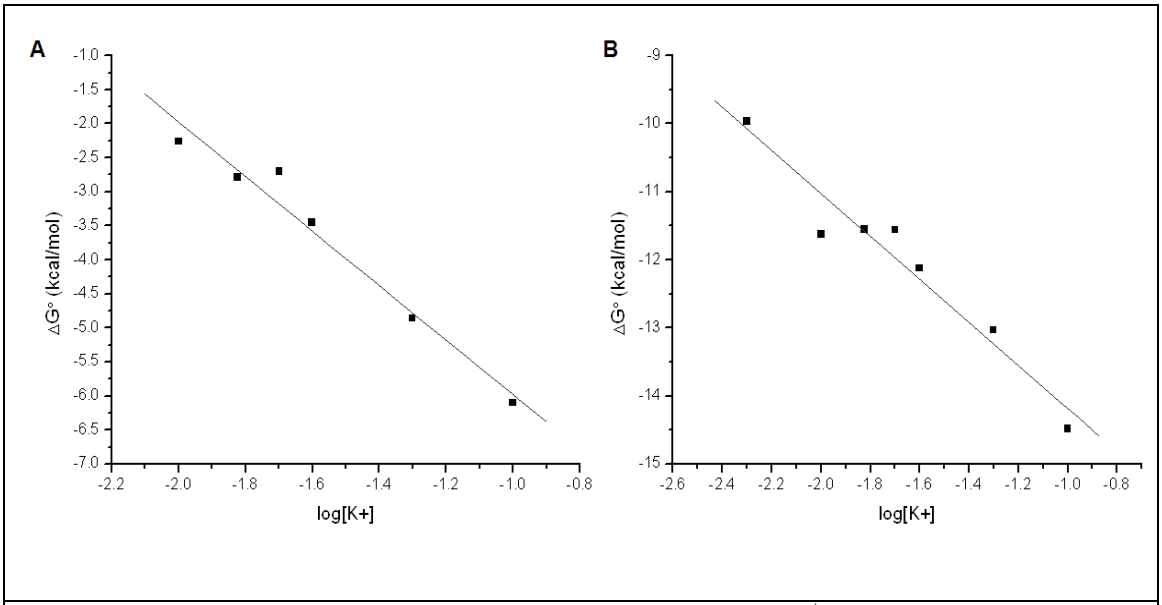


Figure 3.19 Plots of ΔG° as a function of the logarithm of K^+ ion concentration for (A) transition 1 and (B) transition 2 of FBSsh RNA. The number, n, of K^+ ion equivalents released upon the unfolding of FBSsh RNA G-quadruplexes was determined from the slope to be 3 K^+ ions for each transition.

FBSsh RNA. One of each of these K^+ ions could be assigned to being coordinated between the two planes formed by each of the G-quadruplexes. The location of the other two ions may be interpreted in several different ways. Three K^+ ions are released upon the first transition of FBSsh RNA, which may indicate that one of these K^+ ions is serving to stabilize stacking interactions between the two individual G-quadruplexes. Also, K^+ ions may be coordinating above and below the two stacked G-quartet planes, serving to further stabilize the two G-quadruplexes. The coordination of more than one K^+ ion to an RNA sequence proposed to form a two layer G-quadruplex has been observed before (*43*). Furthermore, a recent crystal structure of a G-quadruplex was elucidated in which the K^+ ions were found to coordinate between the planes of the G-quadruplex, and they were also shown to play a role in stabilizing the loops of the G-

quadruplex (*95*). Thus, it is possible that the two K^+ ions that are not coordinating between the two planes formed by FBS_Q1 in the second transition are serving to help stabilize the short loops in the sequence or the longer strand that is no longer forming a G-quadruplex at the higher temperatures. However, much higher concentrations of RNA would be needed to perform high resolution structural studies on FBSsh RNA in order to determine the actual location of the K^+ ions coordinated to the two G-quadruplexes formed by the sequence.

Our direct characterization of the FBSsh RNA sequence to form two distinct parallel G-quadruplex structures supports the indirect evidence that Moine and colleagues collected indicating that two G-quadruplexes existed within the FBS region of the *FMR1* mRNA (*72*). The formation of two G-quadruplexes within the *FMR1* gene is important because the binding of FMRP to this region located close to the three acceptor sites in exon 15 directly affects the splicing of FMRP into its minor splice isoforms, including FMRP ISO2 and ISO3, which lack the major site of phosphorylation. Considering that the phosphorylation/dephosphorylation of FMRP is tightly connected with its translation regulator function and that the minor splice isoforms ISO2 and ISO3 lack this phosphorylation site, it is essential to investigate the binding affinities of the different FMRP splice isoforms, ISO1, ISO2, ISO3, and ISOP, to the FBSsh sequence of the *FMR1* gene (Chapter 4) to determine if the production of these isoforms is controlled by an auto-regulatory loop.

CHAPTER 4: CHARACTERIZATION OF THE BINDING OF FMRP ISOFORMS TO FBS RNA

In chapter 3, we demonstrated that FBSsh RNA forms two specific G-quadruplexes in the presence of KCl: FBS_Q1 RNA and FBS_Q2 RNA. Next, we analyzed the binding to FBSsh RNA of several FMRP isoforms, ISO1, ISO2, and ISO3 and the phosphomimetic ISOP. FMRP ISO1, ISO2, and ISO3 have been chosen because ISO1 is the longest isoform of FMRP, and ISO2 and ISO3 are alternatively spliced at exon 15 within close proximity of the RGG box domain (figure 1.1, chapter 1). This proximity is important as the RGG box peptide binds with high affinity with G-quadruplex forming RNAs, and the truncation close to the domain has been shown to affect the binding of FMRP to G-quadruplex forming target RNA sequences (*10*). Phosphorylation of FMRP occurs at serine 500 of FMRP ISO1, which is also within close proximity of the RGG box domain, this site being eliminated in FMRP ISO2 and ISO3. The phosphorylation of FMRP has been shown to be important in mediating its translation regulator function (*47, 96*), and the alternate splicing into minor isoforms of FMRP, ISO2 and ISO3, prevent FMRP from being re-phosphorylated, possibly remaining in a prolonged "on" state for translation of its mRNA targets.

4.1 Interactions between FBSsh RNA and the FMRP RGG box peptide

It has been shown that the FMRP RGG box peptide binds with high affinity and specificity to G-quadruplex forming RNA sequences (*42-45*). Furthermore, it has been

shown that super-stoichiometric concentrations of RGG box peptide unwind G-quadruplex forming RNA sequences (*43, 44*).

To demonstrate that the RGG box peptide binds to the G-quadruplex forming FBSsh RNA, native PAGE was first utilized. 10 μM FBSsh RNA was incubated with varying concentrations of FMRP RGG box peptide in the range 2.5 – 20 μM at 25°C in the presence of 25 mM KCl, and 10 mM cacodylic acid, pH 6.5. The samples were run on a 15% non-denaturing polyacrylamide gel, which contained 25 mM KCl in the gel as well as in the running buffer for 6 hours at 75 V (figure 4.1). The RGG box peptide is

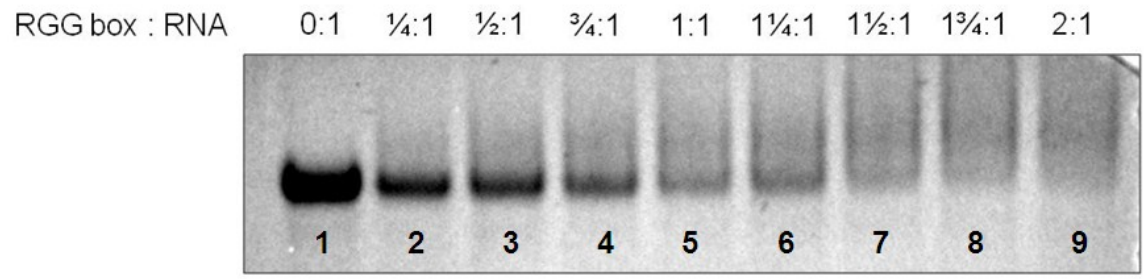


Figure 4.1 Native PAGE of FBSsh RNA incubated with increasing ratios of the FMRP RGG box peptide. Because the peptide is highly positively charged, the complex formation is quantified by the disappearance of the free RNA band. The free RNA band disappears completely at a 2:1 ratio of RGG box peptide: RNA.

positively charged and thus will not enter the gel as a free peptide. Because of the positive nature of the RGG box peptide and since FBSsh RNA is only 42 nt, the complex formed upon the binding of the RGG box to FBSsh RNA has a less negative charge and will not be visible as a distinct tight band in the gel (figure 4.1). The binding of the FMRP RGG box to FBSsh RNA is thus quantified as the disappearance of the free RNA band. As seen in figure 4.1 lanes 2 – 9, the free FBSsh band decreases in intensity upon

increasing the concentration of the RGG box in the sample, indicating the formation of the FBSsh-FMRP RGG box complex. A diffused band appears above the free RNA band in lanes 2 – 9 in the gel, and its intensity becomes higher as the RGG box concentration increases, indicating the formation of the RNA-RGG box complex, but the broadness of the band renders its intensity difficult to quantify.

It has been shown previously that excess RGG box peptide unwinds G-quadruplex forming RNA sequences (*43, 44*). To determine if excess FMRP RGG box peptide unwinds FBSsh RNA, increasing ratios of FMRP RGG box peptide were titrated into a fixed concentration of 10 μM FBSsh RNA, and the spectral changes were monitored through CD spectroscopy. A shown in figure 4.2, no spectral changes

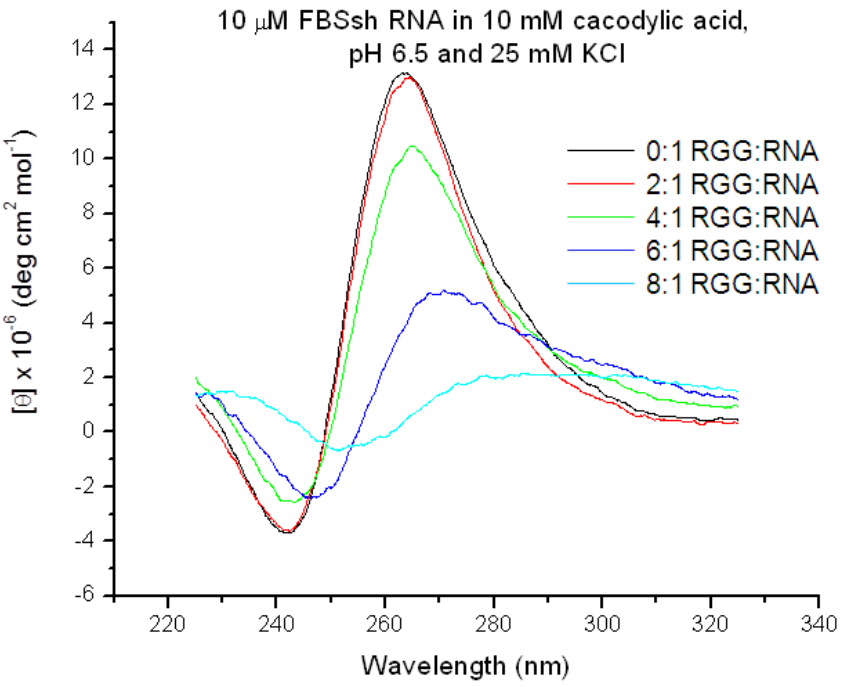


Figure 4.2 Spectral changes in the CD spectrum of FBSsh RNA upon the addition of up to 8:1 RGG box:RNA ratios.

occurred when the FMRP RGG box was present in a 1:1 (not shown) or 1:2 ratio to FBSsh RNA. However in ratios higher than 4:1, the RGG box unwinds the FBSsh G-quadruplex structures. This result is similar to other reports in which the FMRP RGG box unwinds the G quadruplex structures of its mRNA targets (*43, 44*). Because two G-quadruplexes form within FBSsh RNA, it is not possible to distinguish between the unwinding of one quadruplex versus the other by CD spectroscopy. Regardless of this impediment, the unwinding of FBSsh RNA by the FMRP RGG box may have important biological implications with regard to the alternative splicing at exon 15 of *FMR1* mRNA in that unwinding of *FMR1* mRNA by super-soichiometric amounts of FMRP isoforms may allow the sequence to be recognized by proteins that may facilitate the alternative splicing, such as hnRNP (*97*). Furthermore, the FMRP RGG box is an unstructured RNA binding domain, and the FMRP RGG box peptide is unstructured as well, so we did not expect to see any contribution to the CD spectrum from the peptide alone. If the peptide became structured upon binding to the G-quadruplexes formed in FBSsh RNA, we would still not expect for any contributions to interfere with the unwinding experiment as amino acid secondary structures are typically observed between 140 and 190 nm in the CD spectrum, not within the range of parallel G-quadruplex formation.

4.2 Interactions between FBSsh RNA and different full-length FMRP isoforms

To obtain quantitative information about the interactions between the different FMRP isoforms ISO1, ISO2, ISO3 and ISOP and FBSsh RNA, we employed fluorescence spectroscopy. The FBSsh RNA used in the study, named FBSsh_14AP,

was constructed by replacing the adenine at position 14 located within the G-quadruplex formed by FBS_Q1 RNA by 2-aminopurine (2AP, figure 3.17). UV thermal denaturation and CD spectroscopy of FBSsh_14AP were performed, ensuring that the insertion of 2-AP does not affect the secondary structure and stability of the labeled RNA. 2-AP is an analog of adenine that is highly sensitive to changes in its microenvironment, and it can replace adenine without affecting the structure of DNA and RNA (*93, 94*).

Recombinant FMRP isoforms ISO1, ISO2, ISO3, and ISOP were expressed in *E. coli* cells and purified by nickel affinity column purification as described in chapter 2 (*75*). The identities of ISO1, ISO2, and ISO3 were confirmed by peptide mass fingerprinting (Genomics and Proteomics Core Laboratories, University of Pittsburgh) (*10*). The phosphomimetic mutant, ISOP, in which serine 500 was mutated to aspartic

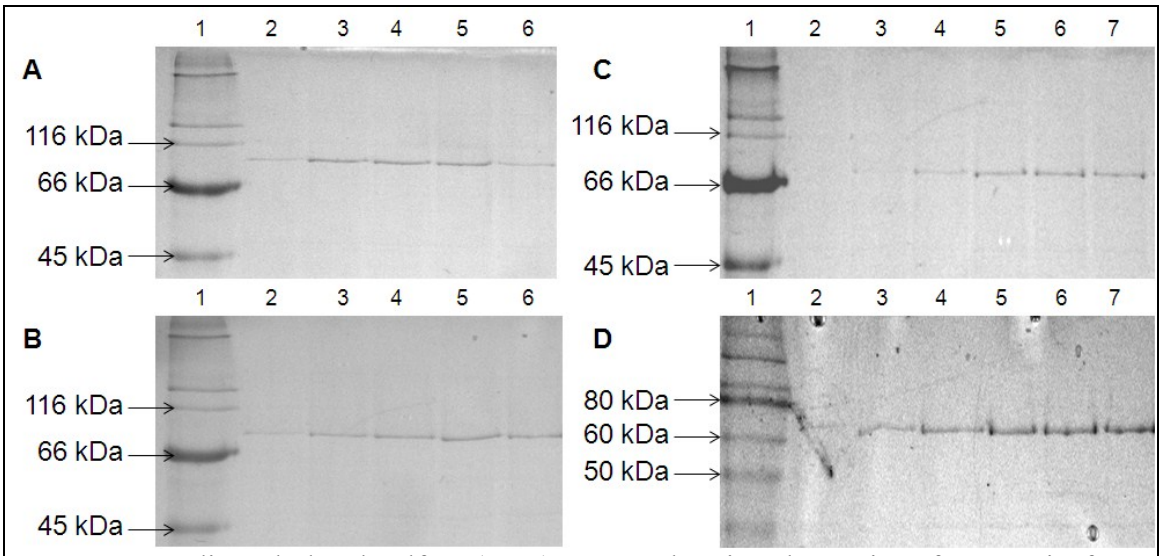


Figure 4.3 Sodium dodecyl sulfate (SDS) PAGE showing the purity of FMRP isoforms ISO1 (A), ISO2 (B), ISO3 (C), and ISOP (D). In all of the gels, lane 1 represents a protein ladder, and lanes 2 through 6 and 2 through 7 in (A) and (B) and (C) and(D), respectively, represent the eluate from the 500 mM imidazole elution step from the purification step in the protocol in chapter 2 (*75*).

acid, was designed by Sara Katrancha, an undergraduate in the Mihailescu laboratory, and it was expressed and purified using the same protocol for the other three FMRP isoforms. The purity of each of the FMRP isoforms was verified by SDS PAGE (figure 4.3). As stated earlier, the phosphorylation of FMRP at serine 500, the major site of phosphorylation, has been shown to be biologically relevant to the translation regulator function of FMRP, and the mutation of a serine to aspartic acid or glutamic acid mimics the phosphorylation of proteins in recombinant proteins (*79*).

The binding of the FMRP isoforms to a fixed concentration of FBSsh_14AP RNA was measured by monitoring the 2-AP steady-state fluorescence intensity changes upon the titration of increasing amounts of each isoform, ISO1, ISO2, ISO3, and the phosphomimetic isoform, ISOP. A decrease in the steady-state fluorescence of the 2-AP reporter was observed upon the addition of each FMRP isoform. This change was observed previously in FMRP steady-state fluorescence binding experiments on S3Fsh_8AP RNA (*10, 75*), a fragment of the mRNA encoding the Semaphorin 3F protein, a known target of FMRP. Each individual binding curve was fitted with equation 2.1 (chapter 2), determining the dissociation constant, K_d, of each isoform binding to FBSsh_14AP RNA (figure 4.4). These experiments were performed in triplicate for each isoform, and the reported errors represent the standard deviation of the average K_d of each curve fit of the data to the equation. As seen in table 4.1, the K_d values for FMRP ISO1 and ISOP are within error of each other, indicating that the binding of FMRP ISO1 to its own mRNA sequence is not affected by the mutation of Ser500 to aspartic acid, mimicking the posttranslational modification of phosphorylation. The phosphorylation of Ser500 has been shown to be important in the exertion of the FMRP translation regulator

function on some of its target mRNAs (*47, 96*). Although the mutation we have made on FMRP ISO1 has been shown to mimic phosphorylation in many instances, it would be interesting to see how *in vitro* phosphorylation of FMRP ISO1 would affect the binding

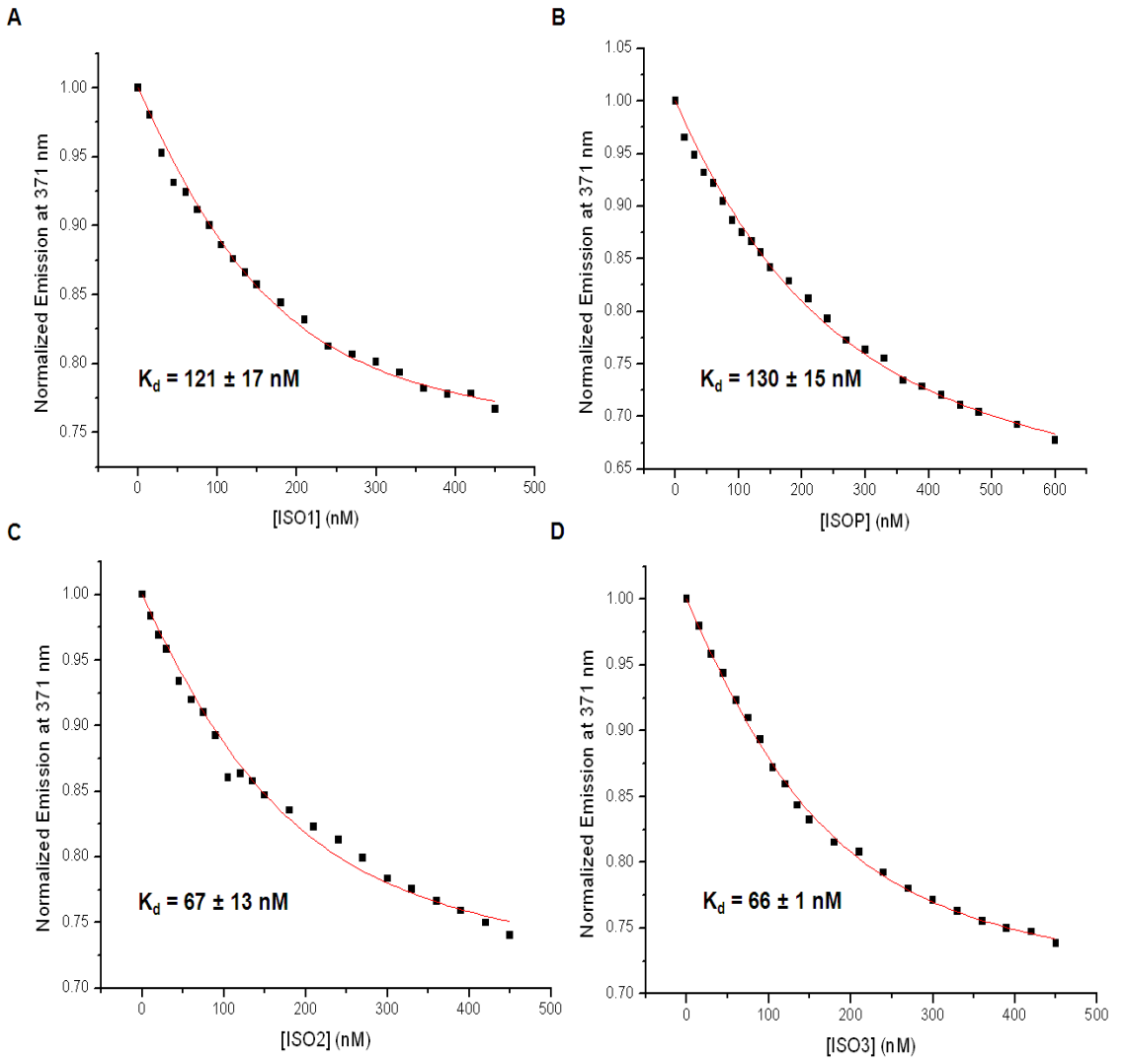


Figure 4.4 FMRP isoforms ISO1, ISO2, ISO3, and ISOP were titrated into free FBSsh_14AP RNA in the presence of 750 nM BSA and 25 mM KCl. (A) FMRP ISO1 and (B) ISOP bind FBSsh_14AP RNA within error of each other, and (C) FMRP ISO2 and (D) ISO3 bind FBSsh_14AP RNA within error of each other and more tightly than ISO1 and ISOP. Each plot represents one experimental data fit to the equation, and the reported K_d values represent the average of three K_d values and an error representing the standard deviation of the three values.

of the protein to FBSsh RNA. The phosphomimetic protein did not show a distinctly different binding affinity for FBSsh RNA, but because the steady-state fluorescence binding assay is highly sensitive, there may be a difference between the mutated protein and the *in vitro* phosphorylated protein. Because of the low yield of recombinant FMRP isoforms, however, *in vitro* phosphorylation on FMRP ISO1 has proven to be difficult to perform in our laboratory. Based on our data, the phosphomimetic FMRP ISO1, ISOP, does not bind differently to FBSsh RNA than the dephosphorylated ISO1. No other quantitative data is currently available about the binding of FMRP ISOP to other G-quadruplex forming mRNA targets to determine if the phosphomimetic FMRP ISO1 binds differently to other FMRP RNA targets than ISO1 or if the mutation does not affect the protein like *in vitro* phosphorylation would in the binding assays. Thus, it remains to be seen if this trend is valid for mRNAs whose translation is regulated directly by FMRP.

Table 4.1 Results of the fluorescence binding assays of FMRP isoforms 1, 2, and 3 and ISOP to G-quadruplex forming FBSsh_14AP RNA

FMRP isoform	Kd (nM)
ISO1	120 ± 17
ISO2	67 ± 13
ISO3	66 ± 1
ISOP	130 ± 15

The K_d values for FMRP ISO2 and ISO3 binding to FBSsh_14AP RNA were also within error of each other yet lower than the K_d values of ISO1 and ISOP binding to FBSsh_14AP RNA (table 4.1), indicating that the truncation of the protein in the proximity of the RGG box results in tighter binding to *FMR1* mRNA. The K_d values measured for each FMRP isoform binding to FBSsh_14AP RNA were in the nanomolar range, similar to values determined for their binding to the G-quadruplex forming S3Fsh_8AP RNA, another FMRP target mRNA (*10*).

We performed control experiments in which the steady-state fluorescence of a non-specific RNA, DLS_9AP RNA, was monitored as FMRP isoforms were added to a

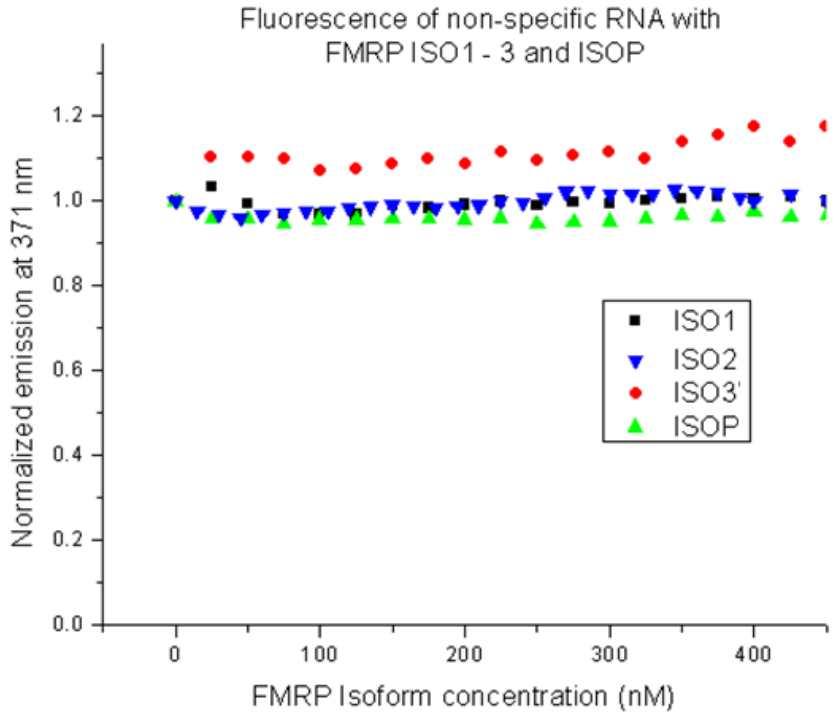


Figure 4.5 Steady state fluorescence spectroscopy emission changes at 371 nm of a 2-AP reporter located in the terminal loop of a 16 nt control hairpin as FMRP ISO1 - ISO3 or ISOP are titrated into a fixed concentration of the RNA (150 nm). The sequence of DLS_9AP RNA is 5'- UCA CGG CG2AP GCU GUG A -3'.

fixed concentration of 150 nM RNA (figure 4.6). The 2-AP is located in the terminal loop of the 16 nt DLS_9AP RNA. The addition of the FMRP isoforms to the non-specific RNA did not have an effect on the fluorescence intensity of the reporter up to triple the concentration, and even then FMRP ISO3 caused the emission of DLS_9AP RNA to increase slightly (*10*).

FMRP ISO1 binding to the G-rich region within *FMR1* has been shown to influence the splicing events at exon 15 (*72*) by increasing the production of the isoforms ISO2 and ISO3, which lack the major phosphorylation site at position 500, and decreasing the production of FMRP ISO1. It has been shown previously that FMR1 -/- knockout mice, which produce *FMR1* mRNA but not any FMRP isoforms, do not produce *FMR1* mRNA that has been alternatively spliced at exon 15 (*72*). This is because FMRP is not available to bind to the FBS RNA sequence in order to control the alternative splicing into minor mRNA and protein isoforms. Our findings that FMRP ISO2 and ISO3 isoforms bind FBSsh RNA more tightly than FMRP ISO1 suggests that the existence of a feedback auto-regulatory loop for the production of the FMRP ISO2 and ISO3. When sufficient amounts of FMRP ISO2 and ISO3 are produced by FMRP ISO1 binding to the FBS RNA sequence, they can compete with FMRP ISO1 to bind to the same sequence, shutting down their own production.

As discussed earlier, in its phosphorylated state, FMRP has been shown to be associated with stalled polyribosomes (*96*), and in the case of the PSD-95 mRNA target with the RISC complex (*47*), suggesting that in this state FMRP prevents the translation of its mRNA targets. The FMRP dephosphorylation in response to synaptic input, has

been shown to rescue the PSD-95 mRNA translation (*47*). The FMRP ISO2 and ISO3 isoforms do not have the ability to be regulated by phosphorylation/dephosphorylation events. Thus, our findings that they have higher affinity for G-quadruplex forming FMRP targets than FMRP ISO1 (*10*) suggests a mechanism by which FMRP ISO2 and/or FMRP ISO3 prolong the "on" state for translation of these mRNA targets by competing with FMRP ISO1, which can turn "off" translation by becoming re-phosphorylated. The levels of the different FMRP isoforms have to be tightly regulated in the cell, as the expression of specific isoforms could control the timing of the translation of different FMRP mRNA targets. Our results suggesting that the expression of FMRP ISO2 and ISO3 is controlled by feedback inhibition are significant, contributing to our understanding of this complex mechanism of regulation.

CHAPTER 5: LET-7B miRNA INTERACTIONS WITH MAP1B mRNA

MAP1B mRNA has been shown to form a G-quadruplex structure within its 5'-UTR and is a known target of FMRP. After sequence analysis of MAP1B mRNA, it was discovered that MAP1B mRNA also contained a sequence within its 3'-UTR that potentially had high binding affinity with the miRNA let-7b, one of the first discovered miRNAs. Since miRNAs are known to regulate the translation of target mRNAs, it was of great interest to investigate the binding of let-7b with the proposed MAP1B mRNA sequence as FMRP may exert its translation repression function in conjunction with let-7b miRNA.

5.1 Let7b miRNA interactions with MAP1B sequence 1

The first step in the investigation of the translation regulation function of let-7b on MAP1B mRNA was to determine whether binding was occurring. In the 3'-UTR of the MAP1B mRNA, a sequence was identified as a potential binding site for let-7b miRNA which was a promising discovery considering that miRNAs most commonly exert their translation repression function by binding to sequences within the 3'-UTR regions of target mRNAs. The location of this binding sequence within a known FMRP mRNA target is interesting because FMRP has been shown to interact with the miRNA pathway. To determine if let-7b miRNA has a translation repression function on the well-known FMRP target MAP1B mRNA, the potential interactions of let-7b miRNA and MAP1B mRNA have to be analyzed, and this was the goal of this study. Thus, the interactions between let-7b miRNA and MAP1B sequence 1 (MAP1B seq1) RNA, shown

in figure 5.1, was characterized by using 1D ^{1}H-NMR spectroscopy, UV thermal denaturation, and non-denaturing polyacrylamide gel electrophoresis (native PAGE). The RNA sequences of let-7b miRNA and MAP1B seq1 RNA were both synthesized by Dharmacon (Thermo Scientific).

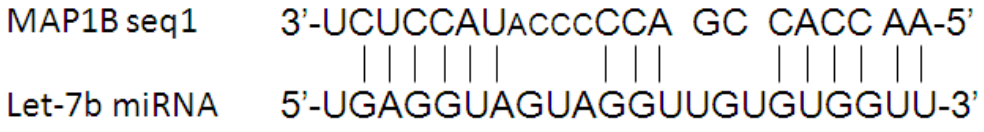


Figure 5.1 Proposed binding of let-7b miRNA to MAP1B seq1 RNA, located in the 3'-UTR of MAP1B mRNA.

First, the binding between let7b miRNA and MAP1B seq1 RNA was characterized by using 1D ^{1}H-NMR spectroscopy. Figure 5.1 shows the proposed interactions between let7b miRNA and MAP1B seq1 RNA, consisting of Watson-Crick and guanine-uracil wobble base pairing, and these interactions have distinct signatures in 1D ^{1}H-NMR spectroscopy in the imino proton resonance region between 10 and 15 ppm (*85*). If the imino protons of Gs and Us are not engaged in base pairing, their resonances will not be observed in the region 10 – 15 ppm due to their fast exchange with water protons. Figure 5.2 shows that let-7b miRNA and MAP1B seq1 RNA do not form Watson-Crick base-paired homodimers as no resonances are present in the imino proton region of their ^{1}H-NMR spectra. This finding is consistent with the secondary structure predictions we obtained by using RNAStructure 4.0 software, which did not find any regions of self-complementarity in these RNA sequences. In contrast, once let-7b and MAP1B seq1 were mixed in a 1:1 ratio at 25°C in 10 mM cacodylic acid, pH 6.5, the ^{1}H-

NMR spectrum reveals numerous resonances in the imino proton region, indicating that let-7b miRNA and MAP1B seq1 RNA bind by Watson-Crick (12 – 15 ppm and G-U wobble base pairing (10-12 ppm).

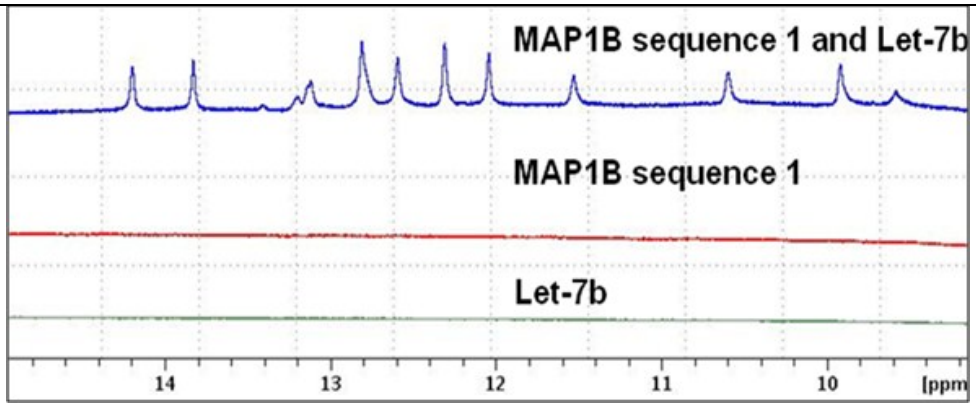


Figure 5.2. ^{1}H-NMR spectra of Let-7b miRNA (bottom green) and MAP1B seq1 RNA (middle red) alone as well as mixed in a 1:1 ratio (top blue).

To confirm that let-7b was able to bind to MAP1B seq1 RNA, we performed concentration dependent native PAGE (figure 5.3). Lanes 1 and 6 in figure 5.3 contain free MAP1B seq1 RNA and let-7b miRNA, respectively, used as controls, whereas lanes 2 – 5 contain mixtures of increasing concentrations of let-7b miRNA in a fixed concentration of MAP1B seq1 RNA in the following ratios: ¼:1 (lane 2), ½:1 (lane 3), ¾:1 (lane 4), and 1:1 (lane 5). MAP1B seq1 RNA migrates as a monomer (lane 1), but it is interesting to note that let-7b miRNA migrates through the gel as a larger molecular weight complex (lane 6). The monomeric single-stranded let-7b miRNA, 22 nt, should appear as a band migrating at the same rate through the gel as free MAP1B seq1 RNA based on the difference of only one nucleotide in their sequence length. This result indicates that let-7b miRNA forms some secondary structure that does not involve the

formation of canonical Watson-Crick base pairs, as such pairing was not observed in the imino proton region of its ^{1}H-NMR spectrum (figure 5.2, bottom spectrum). As the concentration of let7b miRNA increased in the let-7b: MAP1B seq1 mixtures, the band

Lane	1	2	3	4	5	6
Let7b miRNA (μM)	0	5	10	15	20	20
MAP1B seq1 RNA (μM)	20	20	20	20	20	0

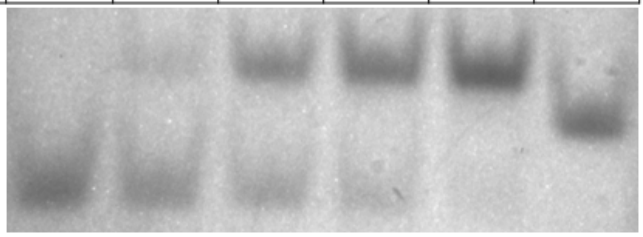


Figure 5.3. 20% native gel mobility assay of let-7b miRNA titrated into MAP1B seq1 RNA to a final ratio of one-to-one.

corresponding to the free MAP1B seq1 RNA decreased in intensity while the band corresponding to the let7b miRNA-MAP1B seq1 RNA duplex increased in intensity. At a 1:1 ratio of miRNA:RNA (lane 5), the free MAP1B seq1 RNA band completely disappeared, indicating that the let7b miRNA bound to MAP1B seq1 RNA in a 1:1 ratio.

To determine the stability of the duplex formed between let-7b miRNA and MAP1B seq1 RNA characterized by ^{1}H-NMR spectroscopy and native PAGE, UV thermal denaturation experiments were performed on the two individual sequences as well as on their complex formed by mixing them in a 1:1 ratio at 25 °C. As expected, the absorbance observed at 275 nm of MAP1B seq1 RNA did not change as the temperature increased from 20 ºC to 95 ºC (figure 5.4, purple data points), indicating that this

sequence exists as a stable monomer, consistent with the results of the ^{1}H-NMR spectroscopy and native PAGE experiments. In contrast, the absorbance observed at 275 nm of let-7b miRNA increased as the temperature increased from 20ºC to 95ºC, likely the result of the thermal denaturation of its secondary structure first suggested by the native PAGE experiment (figure 5.3, lane 6). When let-7b miRNA and MAP1B seq1 mRNA

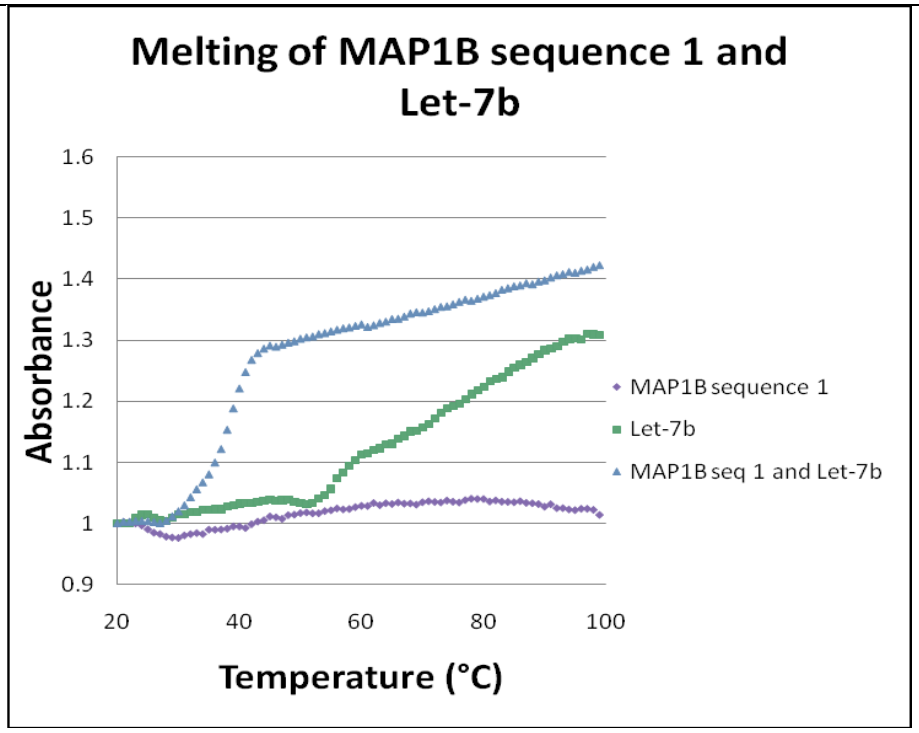


Figure 5.4. UV thermal denaturation of let-7b miRNA, MAP1B seq1 RNA, and a 1:1 mixture of let-7b miRNA and MAP1B seq1 RNA.

were mixed in a 1:1 ratio, their thermal denaturation profile observed at 275 nm showed a clear hyperchromic transition at 39 °C, characteristic of the denaturation of the RNA duplex upon increasing the temperature. This indicated that the duplex formed by let-7b miRNA and MAP1B seq1 RNA is stable up to physiological temperature, 37 °C.

Taken together, the results of 1H-NMR spectroscopy, native PAGE and UV thermal denaturation prove unambiguously that let-7b miRNA interacts by canonical base pairing with a sequence located in the MAP1B 3'-UTR *in vitro*, raising the possibility that these interactions might also be relevant *in vivo*. It is possible that FMRP, let-7b miRNA and the miRISC work together to regulate the translation of the protein product of MAP1B mRNA, as the interaction between FMRP and MAP1B mRNA (*43*) and the interaction between FMRP and the miRISC (*61, 65*) have been shown.

To test the hypothesis that let-7b miRNA interacts with the 3'-UTR of MAP1B mRNA to repress its translation and that this interaction is affected by FXS, *in vivo* experiments were performed by the Feng group at Emory University in Atlanta, GA.

In those studies, the presence of let-7b miRNA and MAP1B mRNA in the murine neonatal hippocampus cells used in the study were detected, and it was shown that the levels of let-7 miRNA were dynamic. Given the expectation that let-7b might act *in vivo* to suppress the translation of MAP1B mRNA, the levels of the MAP1B protein were measured and the levels of the MAP1B protein decreased while the let-7b miRNA levels increased, suggesting that let-7 miRNA may act to repress the translation of MAP1B mRNA.

Next the effect of the presence of FMRP on the production of let-7b in the murine neonatal hippocampus was tested by performing experiments on FMR1 knockout (KO) mice compared to wild-type (wt) mice. Let-7b miRNA production was shown to be reduced in the FMR1 KO mice compared to the wt mice compared to a control miRNA, miR-128, which was not proposed or shown to be affected by the presence or absence of FMRP.

Finally, the direct or indirect association of FMRP with let-7b miRNA as well as the translation repression of MAP1B by let-7b was explored. FMRP and let-7b miRNA were co-immunoprecipitated in wild-type and knock-out cells, and let-7b miRNA was shown to be pulled down with FMRP, indicating that FMRP and let-7b miRNA associate with each other, although the association could be indirect, mediated by the known association between FMRP and the miRNA pathway member, Ago2 (*65*). The binding of let-7b miRNA to MAP1B mRNA was shown to repress its translation in cells. Furthermore, luciferase reporter assays were performed to show that let-7b binding to the 3'-UTR of MAP1B mRNA added to the 3'-UTR of the luciferase gene indeed represses translation, consistent with the binding data shown in section 5.1. When the MAP1B 3'-UTR was absent from the luciferase gene, translation was not repressed as let-7b is not proposed to bind to or repress the translation of the luciferase gene.

We showed that let-7b miRNA binds to MAP1B seq1 in a 1:1 ratio, and FMRP has been shown to bind with high affinity to a G-quadruplex formed in the 5'-UTR of MAP1B mRNA, signifying that FMRP and the miRNA let-7b may work in conjunction to regulate the translation of MAP1B mRNA, being further supported by the interaction between FMRP and the Argonaut (Ago) protein, a key component of the miRNA pathway. Dr. Yue Feng and colleagues showed that FMRP interacts with let-7b miRNA, but it was not clear whether this interaction was direct or indirect through another component of the miRNA pathway since FMRP has already been shown to interact with Ago2 (*65*). Let-7b miRNA was expressed in lower quantities in FMRP KO mice, indicating that the presence of FMRP affects the overall production of this miRNA. The lower quantities of let-7b miRNA in FMRP KO mice may contribute to overall

translation misregulation of MAP1B mRNA in combination with the absence of FMRP. In the future it may be of interest to monitor the translation repression of MAP1B mRNA in the presence of both FMRP and let-7b compared to the presence of each individually by creating a luciferase gene containing both the 5'-UTR and 3'-UTR of MAP1B mRNA.

CHAPTER 6: CONCLUSIONS

6.1 Analysis of the formation of G-quadruplex structures within a G-rich sequence of the *FMR1* mRNA

A G-rich region of *FMR1* mRNA that is bound by FMRP has been analyzed by various biophysical methods, and two distinct G-quadruplex structures were shown to form within its structure. The initial sequence analyzed, FBS_67 RNA (67 nt), was shown to have a very dynamic structure, which involved the formation of several conformations as seen from its ^{1}H-NMR spectra as well as native PAGE experiments. The presence of imino proton resonances in the region centered around 11 ppm of the ^{1}H-NMR spectrum indicated that this sequence formed a G-quadruplex, even in the absence of KCl. However, imino proton resonances were also present in the region 12 – 15 ppm, corresponding to guanine and uracil involved in Watson-Crick base pairing, indicating the formation of an alternate stem-containing structure.

The long, dynamic FBS_67 RNA sequence was split into two short sequences, FBS_Q1 RNA (15 nt) and FBS_Q2 RNA (19 nt) (table 2.1), each postulated to form a G-quadruplex structure. FBS_Q1 RNA was found to adopt a G-quadruplex structure that also contained an A:(G:G:G:G):A hexad, a structure that has not been shown to occur naturally in RNA sequences but been shown to have some biological relevance as a secondary structure of an RNA aptamer against bovine prion protein (*98*). However, once this short FBS_Q1 RNA sequence was incorporated into the longer 42 nt FBSsh RNA, the hexad formation was lost, indicating that the hexad was an artifact caused by the short length of the isolated FBS_Q1 RNA sequence. Interestingly, even upon the

formation of the hexad in FBS_Q1 RNA, the FMRP RGG box peptide maintained high binding affinity to this RNA sequence. FBS_Q2 RNA was characterized to form a G-quadruplex only upon the addition of KCl, but its ^{1}H-NMR spectrum also contained imino proton resonances that correspond to guanine and uracil involved in Watson-Crick base pairing, indicating that this sequence formed more than one conformation *in vitro*. Furthermore, this sequence was found to have low binding affinity with the FMRP RGG box peptide in native PAGE experiments.

Finally, the two short sequences of FBS_Q1 and FBS_Q2 RNA were combined to form the 42 nt FBSsh RNA, which was successfully shown to form two distinct G-quadruplexes, which correspond to the individual FBS_Q1 RNA and FBS_Q2 RNA sequences as shown in figure 3.17, each of which released three K^+ ions upon thermal denaturation. This is significant, as although the G-rich sequence within the coding region of *FMR1* mRNA has been postulated to adopt G-quadruplex structures, their existence was never proved directly prior to these studies. It was important to show that two G-quadruplexes form in the coding region of exon 15 of the *FMR1* gene because it shows that FMRP controls the splicing of its own mRNA into its minor isoforms in a G-quadruplex-dependent manner as this RNA sequence is located close to the three acceptor sites in exon 15.

6.2 Analysis of the differences in binding activity of FMRP ISO1, ISO2, ISO3, and ISOP to the G-quadruplex forming sequences within *FMR1* mRNA

We have also obtained quantitative information about the binding of several FMRP isoforms to FBSsh RNA by fluorescence spectroscopy. The two shorter splice

isoforms of FMRP, ISO2 and ISO3, have a higher binding affinity for the FBSsh RNA than the longest isoform, ISO1 and its phosphomimetic mutant, ISOP. These results are consistent with the findings that a surplus of the major isoform FMRP ISO1 binding to its own mRNA switches the translation of FMRP from its major splice isoforms into its minor splice isoforms. This is significant because it indicates that neuronal cells utilize the negative feedback loop of FMRP binding to its own mRNA to regulate the alternate splicing into minor isoforms. Furthermore, the phosphomimetic mutant of FMRP, ISOP, was shown to bind with the same affinity as ISO1 to FBSsh RNA, indicating that the phosphorylation of serine 500 does not affect the mechanism of FMRP binding to its own mRNA and acting as an exonic splicing enhancer. The increase in production of minor isoforms may have implications for the function of phosphorylated versus dephosphorylated FMRP since ISO2 and ISO3 do not have the site of phosphorylation and would thus be present in a "prolonged state" of dephosphorylation. Although the phosphomimetic FMRP ISO1, ISOP, does not bind significantly differently to its own mRNA than unphosphorylated FMRP ISO1, the phosphorylation of FMRP is important in other aspects of FMRP translation regulator function, such as its interaction with the miRNA pathway with regard to PSD95 regulation (*47*). Phosphorylated FMRP has also been shown to interact with stalled ribosomes while dephosphorylated FMRP has been shown to interact with actively translating ribosomes(*96*). Our results that FMRP ISOP and ISO1 have similar binding affinity to its own mRNA, FMRP acting as an exonic splicing enhancer, support a model in which the phosphorylation of FMRP ISO1 is not important for its role as an exonic splicing enhancer even though it is important in its role as a translation regulator. However, despite our findings, *in vitro* phosphorylation of

FMRP ISO1 and subsequent steady-state fluorescence binding experiments with FBSsh RNA would be beneficial in comparison to the phosphomimetic FMRP ISOP to make sure this mutant is a valid model for quantitative binding experiments.

6.3 Analysis of the binding of let-7b miRNA with the 3'-UTR of MAP1B mRNA

We have shown that let-7b miRNA binds in a 1:1 ratio to a sequence in the 3'-UTR of MAP1B mRNA (MAP1B seq1), a known target of FMRP. FMRP has been shown to bind with high affinity to a G-quadruplex formed in the 5'-UTR of MAP1B mRNA. Furthermore, it has been shown that FMRP interacts directly with components of the miRNA pathway, specifically the Argonaut (Ago) protein, thus linking MAP1B mRNA to let-7b miRNA as a translational regulator target. It also suggests that let-7b miRNA is involved in the translation regulator function of FMRP on MAP1B mRNA, especially since let-7b miRNA appears to be produced in lower quantities in FMRP KO mice. The work done at Emory University in the laboratory of Dr. Yue Feng showed that FMRP may interact directly with let-7b miRNA or indirectly as part of the miRISC since it was shown to associate with the Ago protein.

6.4 Future Work

As mentioned previously, an important experiment to perform will be *in vitro* phosphorylation of FMRP ISO1 in order to compare binding affinities of the mutant FMRP ISOP with the *in vitro* phosphorylated FMRP ISO1. A potential problem with this experiment will be the loss of FMRP ISO1 in the process of phosphorylating the protein, as the protein tends to precipitate out of solution, and the protein is present in very low

concentrations already. It is possible that an *in vitro* phosphorylation assay may produce even lower yields of FMRP ISO1 or that the already low concentration of FMRP ISO1 may prevent the experiment from occurring efficiently in the first place. Furthermore, there has been some dispute over what kinase phosphorylates FMRP *in vivo (13, 99)*, and choosing a kinase for *in vitro* phosphorylation of FMRP ISO1 may prove to be problematic in this aspect as well.

Another direction to take the future work of this project is the expression of other isoforms of FMRP since FMRP ISO1 – ISO3 are not the only isoforms expressed in human cells. The levels of *FMR1* mRNA transcripts in adult mouse brains have been identified, and the transcript levels of splice isoform 7 appear to be the highest, although the translation of each of these isoforms has not been directly shown (*18*). It has been shown that these transcript isoforms interact directly with polyribosomes, suggesting that they are being actively translated in neuronal cells (*18*). FMRP ISO7 lacks exon 12 but is inclusive of the site of phosphorylation. Since this isoform is suggested to be a major splice isoform of FMRP, the binding of ISO7 to the G-quadruplex forming FBSsh RNA may have the effect of increasing the minor splice isoforms ISO8 and ISO9 as a result of increased concentrations of ISO7, and the binding activity of ISO7 to FBSsh RNA should be investigated. The loss of exon 12 may have a major impact on the binding of FMRP to its mRNA targets, as the loss of the RNA binding domain KH1 may change the RNA binding properties of FMRP even if RNA binding is occurring through the RGG box domain.

Additionally, the gene analysis of the *FMR1* gene of a FXS patient who still produced FMRP revealed a frame shift mutation that essentially changes the RGG box

domain. I have been working on expressing and purifying this mutated FMRP, named FMRP G-insertion (FMRP G-ins), and I have worked on the development of the expression and purification of the protein. Optimization of the expression and purification of FMRP G-ins must be accomplished to investigate the effect of the mutation on FMRP binding with G-quadruplex forming mRNA targets. Binding studies have been done using the mutated FMRP RGG box peptide, and it is important to confirm similar results with the full-length FMRP G-ins.

Another interesting aspect of the association of FMRP and its target mRNAs with the miRNA pathway to investigate would be the association of both FMRP and let-7b miRNA with MAP1B mRNA. This could be done by creating a luciferase gene that contains both the 5'-UTR and 3'-UTR of MAP1B, as the 5'-UTR contains the G-quadruplex to which FMRP binds and the 3'-UTR contains the binding site for let-7b, which has been shown to repress MAP1B translation.

REFERENCES

1. Martin, J. P., and Bell, J. (1943) A Pedigree of Mental Defect Showing Sex-Linkage, *Journal of Neurology and Psychiatry 6*, 154-157.
2. Lubs, H. A. (1969) A marker X chromosome, *American Journal of Human Genetics 21*, 231-244.
3. Crawford, D. C., Acuna, J. M., and Sherman, S. L. (2001) FMR1 and the fragile X syndrome: human genome epidemiology review, *Genet Med 3*, 359-371.
4. Hawkins, M., Boyle, J., Wright, K. E., Elles, R., Ramsden, S. C., O'Grady, A., Sweeney, M., Barton, D. E., Burgess, T., Moore, M., Burns, C., Stacey, G., Gray, E., Metcalfe, P., and Hawkins, J. R. (2011) Preparation and validation of the first WHO international genetic reference panel for Fragile X syndrome, *Eur J Hum Genet 19*, 10-17.
5. Sofocleous, C., Kolialexi, A., and Mavrou, A. (2009) Molecular diagnosis of Fragile X syndrome, *Expert Review of Molecular Diagnostics 9*, 23-30.
6. O'Donnell, W. T., and Warren, S. T. (2002) A decade of molecular studies of fragile X syndrome, *Annu Rev Neurosci 25*, 315-338.
7. Jin, P., Zarnescu, D. C., Ceman, S., Nakamoto, M., Mowrey, J., Jongens, T. A., Nelson, D. L., Moses, K., and Warren, S. T. (2004) Biochemical and genetic interaction between the fragile X mental retardation protein and the microRNA pathway, *Nature Neuroscience 7*, 113-117.

8. Pieretti, M., Zhang, F. P., Fu, Y. H., Warren, S. T., Oostra, B. A., Caskey, C. T., and Nelson, D. L. (1991) Absence of expression of the FMR-1 gene in fragile X syndrome, *Cell 66*, 817-822.

9. Tassone, F., Hagerman, R. J., Ikle, D. N., Dyer, P. N., Lampe, M., Willemsen, R., Oostra, B. A., and Taylor, A. K. (1999) FMRP expression as a potential prognostic indicator in fragile X syndrome, *American Journal of Medical Genetics 84*, 250-261.

10. Evans, T. L., Blice-Baum, A. C., and Mihailescu, M. R. (2012) Analysis of the Fragile X mental retardation protein isoforms 1, 2 and 3 interactions with the G-quadruplex forming semaphorin 3F mRNA, *Molecular BioSystems 8*, 642-649.

11. Ashley, C. T., Sutcliffe, J. S., Kunst, C. B., Leiner, H. A., Eichler, E. E., Nelson, D. L., and Warren, S. T. (1993) Human and murine FMR-1: alternative splicing and translational initiation downstream of the CGG-repeat, *Nature Genetics 4*, 244-251.

12. Siomi, H., Siomi, M. C., Nussbaum, R. L., and Dreyfuss, G. (1993) The protein product of the fragile X gene, FMR1, has characteristics of an RNA-binding protein, *Cell 74*, 291-298.

13. Siomi, M. C., Higashijima, K., Ishizuka, A., and Siomi, H. (2002) Casein kinase II phosphorylates the fragile X mental retardation protein and modulates its biological properties, *Molecular and Cellular Biology 22*, 8438-8447.

14. Stetler, A., Winograd, C., Sayegh, J., Cheever, A., Patton, E., Zhang, X., Clarke, S., and Ceman, S. (2006) Identification and characterization of the methyl

arginines in the fragile X mental retardation protein Fmrp, *Human Molecular Genetics 15*, 87-96.

15. Darnell, J. C., Jensen, K. B., Jin, P., Brown, V., Warren, S. T., and Darnell, R. B. (2001) Fragile X mental retardation protein targets G quartet mRNAs important for neuronal function, *Cell 107*, 489-499.

16. Santoro, M. R., Bray, S. M., and Warren, S. T. (2011) Molecular mechanisms of fragile X syndrome: a twenty-year perspective, *Annual Review of Pathology 7*, 219-245.

17. Sittler, A., Devys, D., Weber, C., and Mandel, J. L. (1996) Alternative splicing of exon 14 determines nuclear or cytoplasmic localisation of fmr1 protein isoforms, *Human Molecular Genetics 5*, 95-102.

18. Brackett, D. M., Qing, F., Amieux, P. S., Sellers, D. L., Horner, P. J., and Morris, D. R. (2013) FMR1 transcript isoforms: association with polyribosomes; regional and developmental expression in mouse brain, *PloS One 8*, e58296.

19. Gellert, M., Lipsett, M. N., and Davies, D. R. (1962) Helix formation by guanylic acid, *Proceedings of the National Academy of Sciences of the United States of America 48*, 2013-2018.

20. Chantot, J. F. (1972) Nucleoside conformations. IX. A calorimetric study of gel formation by guanylic acids, *Archives of Biochemistry and Biophysics 153*, 347-356.

21. Ralph, R. K., Connors, W. J., and Khorana, H. G. (1962) Secondary Structure and Aggregation in Deoxyguanosine Oligonucleotides, *Journal of the American Chemical Society 84*, 2265-2266.

22. Mills, M., Lacroix, L., Arimondo, P. B., Leroy, J. L., Francois, J. C., Klump, H., and Mergny, J. L. (2002) Unusual DNA conformations: implications for telomeres, *Current Medicinal Chemistry 2*, 627-644.

23. Rhodes, D., Fairall, L., Simonsson, T., Court, R., and Chapman, L. (2002) Telomere architecture, *EMBO Reports 3*, 1139-1145.

24. Neidle, S., and Parkinson, G. N. (2003) The structure of telomeric DNA, *Current Opinion in Structural Biology 13*, 275-283.

25. Zaug, A. J., Podell, E. R., and Cech, T. R. (2005) Human POT1 disrupts telomeric G-quadruplexes allowing telomerase extension in vitro, *Proceedings of the National Academy of Sciences of the United States of America 102*, 10864-10869.

26. Bryan, T. M., and Baumann, P. (2010) G-quadruplexes: from guanine gels to chemotherapeutics, *Methods in Molecular Biology (Clifton, N.J 608*, 1-16.

27. Sacca, B., Lacroix, L., and Mergny, J. L. (2005) The effect of chemical modifications on the thermal stability of different G-quadruplex-forming oligonucleotides, *Nucleic Acids Research 33*, 1182-1192.

28. 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, *Nature Chemical Biology 3*, 218-221.

29. 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 Research 32*, 371-379.

30. 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.

31. Williamson, J. R. (1994) G-quartet structures in telomeric DNA, in *Annual Review of Biophysics and Biomolecular Structure*, pp 703-730.

32. Simonsson, T. (2001) G-quadruplex DNA structures--variations on a theme, *Biological Chemistry 382*, 621-628.

33. Adrian, M., Heddi, B., and Phan, A. T. (2012) NMR spectroscopy of G-quadruplexes, *Methods (San Diego, Calif) 57*, 11-24.

34. Dapic, V., Abdomerovic, V., Marrington, R., Peberdy, J., Rodger, A., Trent, J. O., and Bates, P. J. (2003) Biophysical and biological properties of quadruplex oligodeoxyribonucleotides, *Nucleic Acids Research 31*, 2097-2107.

35. Zanotti, K. J., Lackey, P. E., Evans, G. L., and Mihailescu, M. R. (2006) Thermodynamics of the fragile X mental retardation protein RGG box interactions with G quartet forming RNA, *Biochemistry 45*, 8319-8330.

36. Prakash, A., Kieken, F., Marky, L. A., and Borgstahl, G. E. Stabilization of a G-(2011) Quadruplex from Unfolding by Replication Protein A Using Potassium and the Porphyrin TMPyP4, *Journal of Nucleic Acids 2011*, 529828.

37. Ranjbar, B., and Gill, P. (2009) Circular dichroism techniques: biomolecular and nanostructural analyses- a review, *Chemical Biology & Drug Design 74*, 101-120.

38. Kiledjian, M., and Dreyfuss, G. (1992) Primary structure and binding activity of the hnRNP U protein: binding RNA through RGG box, *The EMBO Journal 11*, 2655-2664.

39. Corley, S. M., and Gready, J. E. (2008) Identification of the RGG box motif in Shadoo: RNA-binding and signaling roles?, *Bioinformatics and Biology Insights 2*, 383-400.

40. Brown, V., Jin, P., Ceman, S., Darnell, J. C., O'Donnell, W. T., Tenenbaum, S. A., Jin, X., Feng, Y., Wilkinson, K. D., Keene, J. D., Darnell, R. B., and Warren, S. T. (2001) Microarray identification of FMRP-associated brain mRNAs and altered mRNA translational profiles in fragile X syndrome, *Cell 107*, 477-487.

41. 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, *The EMBO Journal 20*, 4803-4813.

42. Phan, A. T., Kuryavyi, V., Darnell, J. C., Serganov, A., Majumdar, A., Ilin, S., Raslin, T., Polonskaia, A., Chen, C., Clain, D., Darnell, R. B., and Patel, D. J. (2011) Structure-function studies of FMRP RGG peptide recognition of an RNA duplex-quadruplex junction, *Nature Structural & Molecular Biology 18*, 796-804.

43. Menon, L., Mader, S. A., and Mihailescu, M. R. (2008) Fragile X mental retardation protein interactions with the microtubule associated protein 1B RNA, *RNA 14*, 1644-1655.

44. Menon, L., and Mihailescu, M. R. (2007) Interactions of the G quartet forming semaphorin 3F RNA with the RGG box domain of the fragile X protein family, *Nucleic Acids Research 35*, 5379-5392.

45. Bole, M., Menon, L., and Mihailescu, M. R. (2008) Fragile X mental retardation protein recognition of G quadruplex structure per se is sufficient for high affinity binding to RNA, *Molecular BioSystems 4*, 1212-1219.

46. Antar, L. N., Dictenberg, J. B., Plociniak, M., Afroz, R., and Bassell, G. J. (2005) Localization of FMRP-associated mRNA granules and requirement of microtubules for activity-dependent trafficking in hippocampal neurons, *Genes, Brain, and Behavior 4*, 350-359.

47. Muddashetty, R. S., Nalavadi, V. C., Gross, C., Yao, X., Xing, L., Laur, O., Warren, S. T., and Bassell, G. J. (2011) Reversible inhibition of PSD-95 mRNA translation by miR-125a, FMRP phosphorylation, and mGluR signaling, *Molecular Cell 42*, 673-688.

48. Subramanian, M., Rage, F., Tabet, R., Flatter, E., Mandel, J. L., and Moine, H. (2011) G-quadruplex RNA structure as a signal for neurite mRNA targeting, *EMBO Reports 12*, 697-704.

49. Zalfa, F., Eleuteri, B., Dickson, K. S., Mercaldo, V., De Rubeis, S., di Penta, A., Tabolacci, E., Chiurazzi, P., Neri, G., Grant, S. G., and Bagni, C. (2007) A new function for the fragile X mental retardation protein in regulation of PSD-95 mRNA stability, *Nature Neuroscience 10*, 578-587.

50. Zalfa, F., Giorgi, M., Primerano, B., Moro, A., Di Penta, A., Reis, S., Oostra, B., and Bagni, C. (2003) The fragile X syndrome protein FMRP associates with BC1 RNA and regulates the translation of specific mRNAs at synapses, *Cell 112*, 317-327.

51. Lu, R., Wang, H., Liang, Z., Ku, L., O'Donnell W, T., Li, W., Warren, S. T., and Feng, Y. (2004) The fragile X protein controls microtubule-associated protein 1B translation and microtubule stability in brain neuron development, *Proceedings of*

the National Academy of Sciences of the United States of America 101, 15201-15206.

52. Gonzalez-Billault, C., Jimenez-Mateos, E. M., Caceres, A., Diaz-Nido, J., Wandosell, F., and Avila, J. (2004) Microtubule-associated protein 1B function during normal development, regeneration, and pathological conditions in the nervous system, *Journal of Neurobiology 58*, 48-59.

53. Bingol, B., and Schuman, E. M. (2004) A proteasome-sensitive connection between PSD-95 and GluR1 endocytosis, *Neuropharmacology 47*, 755-763.

54. Colledge, M., Snyder, E. M., Crozier, R. A., Soderling, J. A., Jin, Y., Langeberg, L. K., Lu, H., Bear, M. F., and Scott, J. D. (2003) Ubiquitination regulates PSD-95 degradation and AMPA receptor surface expression, *Neuron 40*, 595-607.

55. Guan, F., Villegas, G., Teichman, J., Mundel, P., and Tufro, A. (2006) Autocrine class 3 semaphorin system regulates slit diaphragm proteins and podocyte survival, *Kidney International 69*, 1564-1569.

56. Nasarre, P., Constantin, B., Rouhaud, L., Harnois, T., Raymond, G., Drabkin, H. A., Bourmeyster, N., and Roche, J. (2003) Semaphorin SEMA3F and VEGF have opposing effects on cell attachment and spreading, *Neoplasia (New York, N.Y 5*, 83-92.

57. Valverde, R., Edwards, L., and Regan, L. (2008) Structure and function of KH domains, *The FEBS Journal 275*, 2712-2726.

58. Darnell, J. C., Fraser, C. E., Mostovetsky, O., Stefani, G., Jones, T. A., Eddy, S. R., and Darnell, R. B. (2005) Kissing complex RNAs mediate interaction between

the Fragile-X mental retardation protein KH2 domain and brain polyribosomes, *Genes & Development 19*, 903-918.

59. De Boulle, K., Verkerk, A. J., Reyniers, E., Vits, L., Hendrickx, J., Van Roy, B., Van den Bos, F., de Graaff, E., Oostra, B. A., and Willems, P. J. (1993) A point mutation in the FMR-1 gene associated with fragile X mental retardation, *Nature Genetics 3*, 31-35.

60. Chang, S., Wen, S., Chen, D., and Jin, P. (2009) Small regulatory RNAs in neurodevelopmental disorders, *Human Molecular Genetics 18*, R18-26.

61. Edbauer, D., Neilson, J. R., Foster, K. A., Wang, C. F., Seeburg, D. P., Batterton, M. N., Tada, T., Dolan, B. M., Sharp, P. A., and Sheng, M. (2010) Regulation of synaptic structure and function by FMRP-associated microRNAs miR-125b and miR-132, *Neuron 65*, 373-384.

62. Cheever, A., and Ceman, S. (2009) Translation regulation of mRNAs by the fragile X family of proteins through the microRNA pathway, *RNA Biology 6*.

63. Plante, I., Davidovic, L., Ouellet, D. L., Gobeil, L. A., Tremblay, S., Khandjian, E. W., and Provost, P. (2006) Dicer-Derived MicroRNAs Are Utilized by the Fragile X Mental Retardation Protein for Assembly on Target RNAs, *Journal of Biomedicine & Biotechnology 2006*, 64347.

64. Li, X., and Jin, P. (2009) Macro role(s) of microRNAs in fragile X syndrome?, *Neuromolecular Medicine 11*, 200-207.

65. Jin, P., Alisch, R. S., and Warren, S. T. (2004) RNA and microRNAs in fragile X mental retardation, *Nat Cell Biol 6*, 1048-1053.

66. Kim, V. N. (2005) Small RNAs: classification, biogenesis, and function, *Molecules and Cells 19*, 1-15.

67. Winter, J., Jung, S., Keller, S., Gregory, R. I., and Diederichs, S. (2009) Many roads to maturity: microRNA biogenesis pathways and their regulation, *Nat Cell Biol 11*, 228-234.

68. Baker, M. B., Bao, G., and Searles, C. D. (2012) In vitro quantification of specific microRNA using molecular beacons, *Nucleic Acids Research 40*, e13.

69. Cannell, I. G., Kong, Y. W., and Bushell, M. (2008) How do microRNAs regulate gene expression?, *Biochem Soc Trans 36*, 1224-1231.

70. Nelson, P. T., Hatzigeorgiou, A. G., and Mourelatos, Z. (2004) miRNP:mRNA association in polyribosomes in a human neuronal cell line, *RNA 10*, 387-394.

71. Kozomara, A., and Griffiths-Jones, S. (2011) miRBase: integrating microRNA annotation and deep-sequencing data, *Nucleic Acids Research 39*, D152-157.

72. 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 Research 36*, 4902-4912.

73. Schaeffer, C., Beaulande, M., Ehresmann, C., Ehresmann, B., and Moine, H. (2003) The RNA binding protein FMRP: new connections and missing links, *Biology of the Cell / Under the Auspices of the European Cell Biology Organization 95*, 221-228.

74. Milligan, J. F., and Uhlenbeck, O. C. (1989) Synthesis of small RNAs using T7 RNA polymerase, *Methods in Enzymology 180*, 51-62.

75. Evans, T. L., and Mihailescu, M. R. (2010) Recombinant bacterial expression and purification of human fragile X mental retardation protein isoform 1, *Protein Expression and Purification 74*, 242-247.

76. Gasteiger, E., Gattiker, A., Hoogland, C., Ivanyi, I., Appel, R. D., and Bairoch, A. (2003) ExPASy: The proteomics server for in-depth protein knowledge and analysis, *Nucleic Acids Research 31*, 3784-3788.

77. Valverde, R., Pozdnyakova, I., Kajander, T., Venkatraman, J., and Regan, L. (2007) Fragile X mental retardation syndrome: structure of the KH1-KH2 domains of fragile X mental retardation protein, *Structure 15*, 1090-1098.

78. Coffee, R. L., Jr., Williamson, A. J., Adkins, C. M., Gray, M. C., Page, T. L., and Broadie, K. (2011) In vivo neuronal function of the fragile X mental retardation protein is regulated by phosphorylation, *Human Molecular Genetics 21*, 900-915.

79. Tarrant, M. K., and Cole, P. A. (2009) The chemical biology of protein phosphorylation, *Annual Review of Biochemistry 78*, 797-825.

80. Mergny, J. L., Phan, A. T., and Lacroix, L. (1998) Following G-quartet formation by UV-spectroscopy, *FEBS Letters 435*, 74-78.

81. Borer, P. N., Dengler, B., Tinoco, I., Jr., and Uhlenbeck, O. C. (1974) Stability of ribonucleic acid double-stranded helices, *Journal of Molecular Biology 86*, 843-853.

82. Piotto, M., Saudek, V., and Sklenar, V. (1992) Gradient-tailored excitation for single-quantum NMR spectroscopy of aqueous solutions, *Journal of Biomolecular NMR 2*, 661-665.

83. Sklenar, V., Peterson, R. D., Rejante, M. R., and Feigon, J. (1993) Two- and three-dimensional HCN experiments for correlating base and sugar resonances in 15N,13C-labeled RNA oligonucleotides, *Journal of Biomolecular NMR 3*, 721-727.

84. Hendry, P., and Hannan, G. (1996) Detection and quantitation of unlabeled nucleic acids in polyacrylamide gels, *BioTechniques 20*, 258-264.

85. Furtig, B., Richter, C., Wohnert, J., and Schwalbe, H. (2003) NMR spectroscopy of RNA, *Chembiochem 4*, 936-962.

86. Lipay, J. M., and Mihailescu, M. R. (2009) NMR spectroscopy and kinetic studies of the quadruplex forming RNA r(UGGAGGU), *Molecular BioSystems 5*, 1347-1355.

87. Liu, H., Matsugami, A., Katahira, M., and Uesugi, S. (2002) A dimeric RNA quadruplex architecture comprised of two G:G(:A):G:G(:A) hexads, G:G:G:G tetrads and UUUU loops, *Journal of Molecular Biology 322*, 955-970.

88. Kettani, A., Gorin, A., Majumdar, A., Hermann, T., Skripkin, E., Zhao, H., Jones, R., and Patel, D. J. (2000) A dimeric DNA interface stabilized by stacked A.(G.G.G.G).A hexads and coordinated monovalent cations, *Journal of Molecular Biology 297*, 627-644.

89. Matsugami, A., Mashima, T., Nishikawa, F., Murakami, K., Nishikawa, S., Noda, K., Yokoyama, T., and Katahira, M. (2008) Structural analysis of r(GGA)4 found in RNA aptamer for bovine prion protein, *Nucleic Acids Symposium Series (2004)*, 179-180.

90. Mergny, J. L., De Cian, A., Amrane, S., and Webba da Silva, M. (2006) Kinetics of double-chain reversals bridging contiguous quartets in tetramolecular quadruplexes, *Nucleic Acids Research 34*, 2386-2397.

91. Kypr, J., Kejnovska, I., Renciuk, D., and Vorlickova, M. (2009) Circular dichroism and conformational polymorphism of DNA, *Nucleic Acids Research 37*, 1713-1725.

92. Hardin, C. C., Perry, A. G., and White, K. (2000) Thermodynamic and kinetic characterization of the dissociation and assembly of quadruplex nucleic acids, *Biopolymers 56*, 147-194.

93. Bharill, S., Sarkar, P., Ballin, J. D., Gryczynski, I., Wilson, G. M., and Gryczynski, Z. (2008) Fluorescence intensity decays of 2-aminopurine solutions: lifetime distribution approach, *Analytical Biochemistry 377*, 141-149.

94. Serrano-Andres, L., Merchan, M., and Borin, A. C. (2006) Adenine and 2-aminopurine: paradigms of modern theoretical photochemistry, *Proceedings of the National Academy of Sciences of the United States of America 103*, 8691-8696.

95. Wei, D., Parkinson, G. N., Reszka, A. P., and Neidle, S. (2012) Crystal structure of a c-kit promoter quadruplex reveals the structural role of metal ions and water molecules in maintaining loop conformation, *Nucleic Acids Research 40*, 4691-4700.

96. Ceman, S., O'Donnell, W. T., Reed, M., Patton, S., Pohl, J., and Warren, S. T. (2003) Phosphorylation influences the translation state of FMRP-associated polyribosomes, *Human Molecular Genetics 12*, 3295-3305.

97. Chen, M., and Manley, J. L. (2009) Mechanisms of alternative splicing regulation: insights from molecular and genomics approaches, *Nat Rev Mol Cell Biol 10*, 741-754.

98. Mashima, T., Matsugami, A., Nishikawa, F., Nishikawa, S., and Katahira, M. (2009) Unique quadruplex structure and interaction of an RNA aptamer against bovine prion protein, *Nucleic Acids Research 37*, 6249-6258.

99. Narayanan, U., Nalavadi, V., Nakamoto, M., Thomas, G., Ceman, S., Bassell, G. J., and Warren, S. T. (2008) S6K1 phosphorylates and regulates fragile X mental retardation protein (FMRP) with the neuronal protein synthesis-dependent mammalian target of rapamycin (mTOR) signaling cascade, *J Biol Chem 283*, 18478-18482.