UNIVERSITY OF CALIFORNIA, SAN DIEGO

Dissecting the RNA Binding Specificity of Fragile X Mental Retardation Protein

A Thesis submitted in partial satisfaction of the requirements for the degree

Master of Science

in

Chemistry

by

Youssi Momen Athar

Committee in charge:

Professor Simpson Joseph, Chair
Professor Gourisankar Ghosh
Professor Thomas Hermann

2016

The Thesis of Youssi Momen Athar is approved and it is acceptable in quality and form for publication on microfilm and electronically:

_____

_____

_____
Chair

University of California, San Diego
2016

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

Conducting research has been a pleasure thanks to my mentor Professor Simpson Joseph. I most appreciate my scientific discussions with Dr. Joseph and his candid guidance during my budding scientific career.

Everyone in the Joseph lab has been a collaborative colleague and a friend. I particularly acknowledge Xinying Shi, Dr. Norman Zhu and Lila Mouakkad for their direct contributions to my research. Xinying initially cloned our human FMRP. Norm developed the expression and purification procedures for the full-length human FMRP, NT-hFMRP and NT-dFMRP constructs that I have continued to employ in my project. Lila established the fluorescence anisotropy method which I have adopted in my own project.

I also thank my former colleagues in Professor James Bowie's group at UCLA. Dr. Bowie gave me an opportunity to learn biochemistry research in his lab as a post-baccalaureate research assistant. Dr. Bowie and my former colleagues at UCLA helped me develop research skills which I carried over into my research here in the Joseph lab.

Finally, I am eternally grateful to my parents and my family for encouraging me in all my academic endeavors.

VITA

2016: Master of Science, Chemistry

University of California, San Diego

2011-2014: Research Assistant

University of California, Los Angeles

2011: Bachelor of Arts, Molecular and Cell Biology

University of California, Berkeley

ABSTRACT OF THE THESIS

Dissecting the RNA Binding Specificity of Fragile X Mental Retardation Protein

by

Youssi Momen Athar

Master of Science in Chemistry

University of California, San Diego, 2016

Professor Simpson Joseph, Chair

FMRP binds and regulates translation of neuronal mRNAs. Loss of FMRP causes aberrant expression of proteins, which results in Fragile X syndrome. Three promising RNA sequence/structure motifs have been proposed to specifically bind FMRP's RNA-binding domains. The KH1/2 domain-binding motifs have not been thoroughly tested for direct binding to FMRP. An in vitro selected RNA G-quadruplex (GQ) has been shown to directly bind the RGG motif, but it is unclear whether this specific GQ structure is an accurate model of the RNA GQ structures FMRP may bind in vivo.

Here we quantified binding of human FMRP variants containing different RNA-binding domains to a set of model RNAs representing the three proposed RNA motifs. We find human FMRP variants containing the KH1 and KH2 domains bind $(UGGA)_4$ with sub-micromolar affinity, while the RGG domain alone binds $(UGGA)_4$ with tenfold lower affinity. Surprisingly, FMRP cannot bind $(GACG)_4$ at all. Both the $PolyG_{18}$ and UG_4U GQ model RNAs bound with nanomolar affinities specifically to the RGG motif. The RGG domain alone bound with equal or higher affinity to both RNAs compared to the full-length and N-terminus truncated variants, suggesting that the RNA GQ is recognized and bound by the RGG motif with little contribution from the KH1 and KH2 domains.

Chapter 1: Introduction

1.1 Fragile X Syndrome is caused by the FMR1 gene

Fragile X syndrome (FXS) is the most common form of inherited intellectual disability (ID). FXS is an inherited X-linked condition and according to the National Fragile X Foundation, FXS affects approximately 1 in 4000 males and 1 in 6000 females worldwide[1]. Over 100,000 individuals in the U.S. suffer FXS with costs of treatment exceeding $200 million per year[2].

While ID is a common symptom amongst FXS patients, severity of the symptom can vary. Some patients display symptoms observed in patients with autism spectrum disorder (ASD). ASD symptoms may include impaired cognitive function and unusual behavior, but more specifically difficulties with speech, social anxiety, seizures, and attention-deficit/hyperactivity disorder (ADHD)[3,4]. Patients with FXS may display physical abnormalities as well, such as elongated facial features and ears, flat feet, and macroorchidism[5].

Postmortem examinations of FXS patients' brains revealed the physiological abnormality behind the disease. The FXS patients' neurons contained abnormally dense and immature dendritic spines[6]. Genetic studies revealed FXS patients had a long expansion of cytosine-guanine-guanine (CGG) trinucleotide repeats at the 5' untranslated region (UTR) of the same gene within the X chromosome, then identified as the Fragile X mental retardation gene (*FMR1*)[7,8]. The CGG expansion causes FXS by two major

mechanisms involving silencing of the *FMR1* gene: (1) hypermethylation of the DNA and (2) hybridization of the encoded mRNA to the *FMR1* gene's complementary CGG array to repress transcription of the gene[9]. While it is unclear how these two mechanisms control transcription of the *FMR1* gene, it is evident they effectively diminish the expression of the encoded Fragile X mental retardation protein (FMRP). Overall, the studies suggest FXS is caused by FMRP deficiency in the neuron.

Later an isoleucine 304 to asparagine point mutation (I304N) within the K-Homology 2 (KH2) domain was identified in patients suffering a severe form of FXS[10]. Investigation of the I304N variant revealed a decrease in FMRP's stability, affinity for target RNAs and association with polyribosomes[10,11].

Interestingly, the length of the CGG repeat expansion can cause a different disease in those with shorter repeat expansions. While normal individuals have 6-54 CGG repeats and individuals with FXS have over 200 CGG repeats, individuals with 55-200 CGG repeats can develop pre-mutation diseases such as Fragile X-associated tremor/ataxia syndrome (FXTAS) and Fragile X-related primary ovarian insufficiency (FXPOI)[5]. Unlike FXS symptoms that manifest as ID in childhood, FXTAS symptoms manifest in adulthood and result in neurodegenerative disorders that cause involuntary movements, imbalance, and symptoms generally associated with Parkinson's and Alzheimer's disease[12]. FXPOI causes early-onset menopause in women before the age of 40 as well as dysfunctional ovaries which results in infertility[13]. Studies of FXTAS suggest the pre-mutation number of CGG repeats results in

elevated levels of the *FMR1* messenger RNA (mRNA) levels but decreased FMRP levels[14,15].

1.2 Architecture of FMRP

FMRP contains several highly conserved domains (Figure 1.1). Amongst these, four are RNA-binding domains. There are three K-Homology (KH) domains — KH0, KH1 and KH2 — and an arginine-glycine-glycine (RGG) box domain, all four of which are hypothesized to mediate FMRP binding to target mRNAs[16].

The KH domains were first identified as nucleic acid binding domains in heterogeneous nuclear ribonucleoprotein K (hnRNP K), and have since been found in many eukaryotic (Type I) and prokaryotic (Type II) proteins[17]. A 1.9 Å resolution crystal structure of the human FMRP KH domains confirms both KH1 and KH2 domains' topologies match the typical Type I found in several eukaryotic nucleic acid binding proteins[11]. A GXXG loop located in FMRP KH

Figure 1.1: Architecture of human FMRP. The relative position of each domain is depicted throughout the primary sequence above. The Agenet 1 and (cyan) Agenet 2 (orange) domains are also known as the Tudor 1 and Tudor 2 domains, respectively.

domains is also found in the KH-type splicing regulatory protein (KSRP), further suggesting FMRP KH domains play a role in specific RNA binding[18].

Indeed, patients suffering a severe form of FXS often had a I304N mutation in the KH2 domain. Even with a normal span of CGG repeats at the *FMR1* 5' UTR, the I304N point mutation was sufficient to cause FXS[10]. Subsequent studies of the I304N FMRP variant revealed the point mutation abolished FMRP's ability to bind target mRNAs and associate with polysomes[19,20]. Inspection of the KH2 domain's structure shows isoleucine 304 resides within a hydrophobic region of the KH2 domain. Substitution of the hydrophobic isoleucine for a polar asparagine residue could therefore destabilize the protein's structure[11]. It is still unknown through which mechanism(s) the mutation causes FXS.

Another FXS patient was later found to have a glycine 266 to glutamate mutation within the KH1 domain[21]. The G266E KH1 mutant FMRP was also unable to associate with polysomes and bind target mRNAs. While only two point mutations within the KH domains have been identified, it is possible that other KH1 and KH2 mutations that similarly disrupt FMRP stability or function may cause FXS as well.

In addition to the three KH domains, FMRP also contains a RGG domain composed of two tandem RGG motifs[22]. An NMR structure of the human FMRP RGG motif bound to the in vitro-selected *sc1* RNA G-quadruplex (GQ) reveals how the disordered RGG motif is stabilized upon binding to the GQ structure[23]. A subsequent crystal structure reveals in atomic detail how each RGG motif

amino acid interacts with the *sc1* GQ nucleotides[24]. While the *sc1* RNA was selected as a stable RGG motif binding partner, it is possible that the RGG motif mediates FMRP binding to GQ structures in target mRNAs. And considering the ubiquity of GQ structures amongst mRNAs, the presence of the RGG motif within FMRP hints at FMRP's potential to bind a vast pool of target mRNAs in vivo.

Besides FMRP's RNA binding domains, FMRP contains a multifunctional domain spanning its N-terminal region known as the N-terminal domain of FMRP (NDF). The NDF is defined as the beginning of FMRP up to the boundary of the KH1 domain, and studies of the NDF suggests it mediates assembly of both FMRP-protein interaction as well as stable FMRP homodimers[25,26]. The NMR structure of the NDF revealed the presence of two tandem Agenet/Tudor domains that have been observed to bind trimethylated lysines on histones via hydrophobic patches[26,27]. There is in fact evidence of FMRP involvement in gametogenesis and the DNA damage response (DDR) via binding of its tandem Tudor domains to chromatin[28].

A recent 3.0 Å X-ray crystal structure of a stable region of the NDF spanning residues 1 to 209 of human FMRP (FMRPΔ) revealed a novel KH0 domain upstream of the KH1 and KH2 domains[29]. The KH0 domain adopts a Type I topology similar to KH1 and KH2 domains, but shows three key differences from other KH domains: (1) the conserved GXXG present in KH domains is replaced with a single lysine reside, (2) the KH0 is missing the conserved hydrophobic residues of the IGXXGXXI motif, and (3) there is no

positively charged groove for RNA binding. These three differences suggest the KH0 domain interacts with putative binding partners in a different manner from other KH domains.

In the same study, FMRPΔ was found to exist as stable homodimers. This is in agreement with previous studies showing the NDF mediates FMRP dimerization. However, this study revealed three possible dimer interfaces and attempted to identify the true dimer interface. Small-angle X-ray scattering (SAXS) analysis of FMRPΔ hinted at an ensemble of dimers. Subsequent minimal ensemble search (MES) to fit the SAXS data suggested all three dimer interfaces are somehow involved in FMRP oligomerization, although the authors propose the Agenet 2 domain contains the primary dimer interface (Figure 1.2). It remains unclear which of the proposed interfaces is primarily stabilizing FMRPΔ dimerization.

Figure 1.2: 3.0 Å crystal structure of human FMRPΔ depicting the Agenet 2 dimer interface. The dimerization interface between the the Agenet 2 domains (magenta and orange) of two monomers is proposed to be the primary dimer interface despite the small area of the dimer interface relative to the size of human FMRP. The Agenet 1 and KH0 domains are highlighted in cyan and blue, respectively. Protein Data Bank entry code 4OVA.

FMRP also contains both a nuclear localization signal (NLS) and a nuclear export signal (NES)[30]. In light of FMRP's RNA-binding domains, FMRP may also shuttle target mRNAs from the nucleus to the cytosol. Interestingly, an arginine 138 mutation to glutamine (R138Q) was identified in a patient suffering developmental delays[31]. More recently, FMRP was found to function in gametogenesis and the DDR within the cell nucleus. Indeed, the R138Q mutation disrupted FMRP's binding to nucleosomes, effectively impairing its role in DDR[28].

1.3 FMRP binds specific sequence or structure motifs in target mRNAs

FXS has been traced to the absence of the Fragile X mental retardation protein (FMRP) in neurons. Loss of FMRP in turn has been linked to the deregulation of translation of specific mRNAs by the ribosome[32,33]. This suggests FMRP is a RNA-binding protein that must be able to recognize its target mRNA and coordinate with the ribosome to regulate their translation. While FMRP has been shown to bind the ribosome directly, it is still unclear how the FMRP, the target mRNA and the ribosome coordinate spatiotemporally to suppress translation of the mRNA[34,35]. Most endeavors to identify FMRP mRNA targets have employed high-throughput methods such as HITS-CLIP, PAR-CLIP, RNAcompete and TRIBE. Unfortunately, the proposed catalog of target RNAs, as well as RNA sequence and structure motifs, overlap poorly between each study[34,36-38]. Analysis of the various proposed RNAs and RNA recognition

elements yielded two relatively promising RNA recognition elements: the GACR (R is a purine, A or G) and the WGGA (W is an A or U) sequences[40]. The GACR sequences were enriched for an FMRP variant composed solely of the KH1 and KH2 RNA binding domains[38]. The WGGA sequences were proposed to bind specifically to the KH1 domain of FMRP[36].

A more recent study has proposed FMRP primarily binds and represses the mRNA encoding diacylglycerol kinase subunit kappa (DGKκ), and that loss of this single interaction is sufficient for causing FXS-like symptoms in mice[39]. The proposition that FMRP's role in FXS pathophysiology is dictated primarily through loss of binding solely to the DGKκ mRNA boldly refutes the field's working hypothesis: FMRP binds a target set of mRNAs by recognizing either a sequence motif or a structural motif, and represses their translation by the ribosome locally.

Despite the controversy, one RNA motif that FMRP has been confirmed to bind experimentally is the RNA G-quadruplex (GQ)[40,41]. In fact, the SELEX-derived GQ-assembling RNA *sc1* was shown to bind FMRP's arginine-glycine-glycine (RGG) box motif with high affinity (Figure 1.3)[23,24]. However, being selected under optimized conditions in vitro, the *sc1* RNA GQ may not be representative of any RNA GQ structures that FMRP may bind in vivo[42]. It is also unclear if the FMRP RGG motif is sufficient for specifically binding GQ structures in target mRNAs. Altogether it is unknown how FMRP specifically binds the GQ and if the RNA GQ is a general method by which FMRP recognizes target mRNAs.

To answer these questions, we sought to quantify binding affinity between FMRP and various model RNAs. A PolyG_{18} RNA was originally used to serve as a model RNA GQ. Later we employed a six nucleotide RNA molecule UG_4U that has been shown to assemble a stable intermolecular $(UG_4U)_4$ GQ structure in potassium, but remain primarily as a monomer in lithium, as another model RNA GQ (Figure 1.4)[43,44]. Being able to switch a RNA between a GQ and a hexanucleotide would allow FMRP's affinity for a RNA GQ to be determined more definitively.

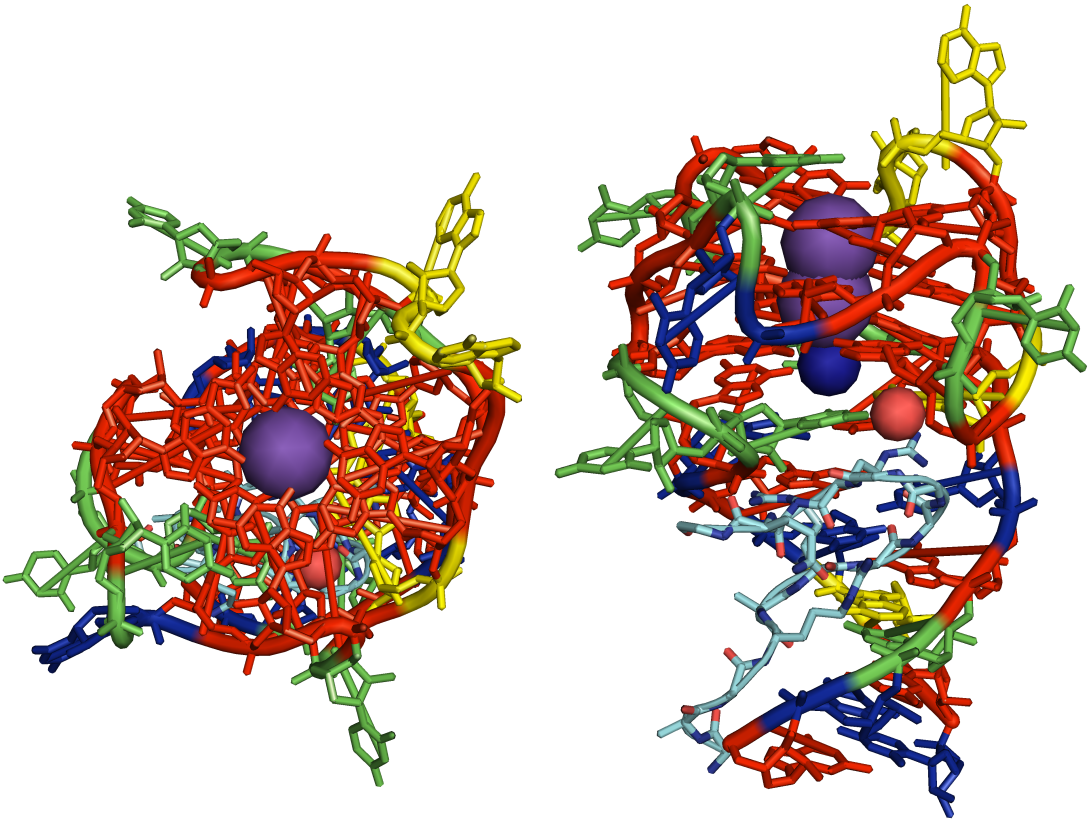

Figure 1.3: 2.8 Å crystal structure of cesium bound human FMRP RGG motif bound to *sc1* RNA. Top view (left) reveals the guanosines (red) are assembling into G tetrads. Side view (right) shows two K^+ ions (purple spheres) and a Cs^+ ion (indigo sphere) each coordinating to O6 of the tetrad guanines. The Cs^+ coordination suggests K^+ is not specifically required at this site. Uridines, adenosines and cytidines are depicted in green, yellow and blue, respectively. The RGG motif (cyan) binds the *sc1* RNA below the GQ structure. The oxygen atom (red sphere) may mediate binding between the RGG motif and *sc1*. Protein Data Bank entry code 5DEA.

1.4 FMRP represses translation of target mRNAs

FMRP was found to be mainly expressed in the mammalian brain's neurons and reproductive organs[45-47]. In each case, FMRP was found primarily involved within the cytoplasm where it interacted with translating polysomes and messenger ribonucleoprotein (mRNP) complexes[19,48,49].

It was later confirmed that FMRP is indeed repressing translation of target mRNAs locally at the dendritic spines of neurons[50-52]. Early studies of translation repression by FMRP established that FMRP could inhibit target mRNA translation dose-dependently whereas the I304N mutant identified in the FXS patient could not[50,51].

While FMRP is understood to repress translation of RNAs, the mechanism(s) by which FMRP interacts with the target mRNA and the translating ribosome is still unknown. One theory is FMRP hinders translation initiation by binding to the target RNA and recruiting the cytoplasmic FMRP-interacting protein (CYFIP1) which then binds the eukaryotic initiation factor 4E (eIF4E) at the mRNA 5' 7-methylguanosine (m^7G) cap. Binding of CYFIP1 to eIF4E in turn prevents binding of eukaryotic initiation factor 4G (eIF4G) to eIF4E, and the eukaryotic initiation factor 4F (eIF4F) complex is prevented from assembling at the m^7G cap[53].

Interestingly however, most of the cytoplasmic FMRP has been found to associate with polysomes that are actively translating mRNAs. Indeed, FMRP has been found to bind target mRNAs associated with stalled

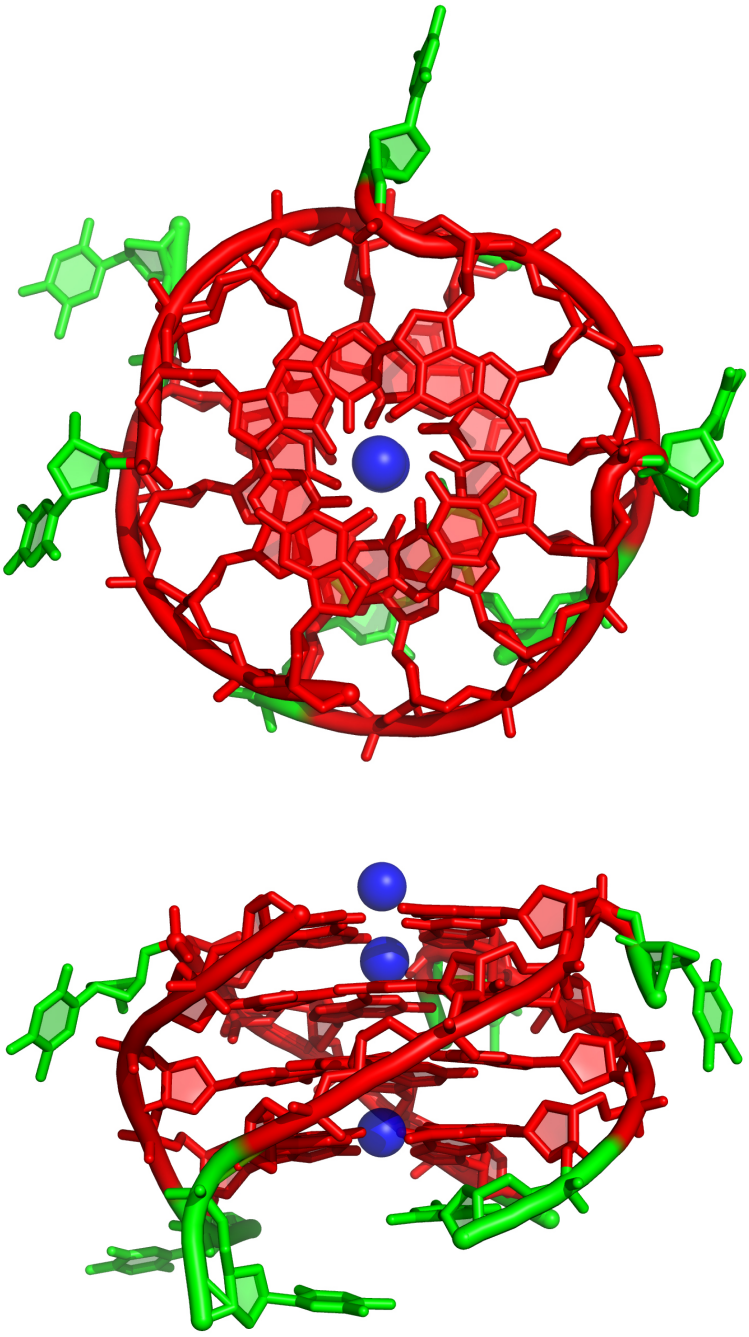

Figure 1.4: 0.61 Å crystal structure of $(UG_4U)_4$. Top view (top) and side view (bottom) of four UG_4U RNA molecules assembling into the intermolecular G-quadruplex $(UG_4U)_4$. Uridines are depicted in green and guanosines in red. The blue spheres passing through the axis of the helix represent Sr^{2+} ions which coordinate to the O6 of every other guanine base plane within the quadruplex. Protein Data Bank entry code 1J8G.

ribosomes[34,54]. Additionally, FMRP has been shown to interact with the ribosome even after translation initiation is blocked[55]. Altogether, FMRP is suggested to repress translation by binding the target mRNA and stalling the ribosome during elongation.

Recently FMRP was shown to both repress translation of target mRNAs and bind directly to the ribosome in a manner consistent with the model of elongation repression. A cryo-electron microscopy (cryo-EM) model of *Drosophila* FMRP (dFMRP) bound to the *Drosophila* 80S ribosome depicts the FMRP binding at a site within the ribosome that would prevent binding of tRNA and translation factors[35]. Additionally, the FMRP-ribosome structure reveals the KH1 and KH2 domains binding near the ribosome's peptidyl site and the RGG box domain of FMRP residing near the aminoacyl site, presumably available for binding target mRNA motifs (i.e. RNA GQ structures). Together with previous studies showing (1) FMRP primarily binds actively translating ribosomes (polysomes) and (2) FMRP binds RNA GQ structures through its RGG box domain, the cryo-EM structure suggests FMRP may tether itself to the ribosome and the mRNA via its KH and RGG domains, respectively, to stall the ribosome during elongation[33-35,41,56,57].

FMRP is also suggested to repress translation via RNA interference (RNAi). FMRP has been shown to associate with the RNA-induced silencing complex (RISC) proteins Dicer and Argonaute 2 (Ago2) as well as specific microRNAs (miRNAs)[58-60]. In fact, FMRP has been reported to assemble the Ago2 and miRNA-125a inhibitory complex on the postsynaptic density protein

95 (PSD-95) mRNA through phosphorylation of FMRP[61,62]. The PSD-95 RNAi case may be just one of many RNAi translation repression pathways FMRP employs. It is possible that FMRP can recruit from a catalog of miRNA complexes to specifically inhibit a variety of mRNAs via RNAi.

Post-translational modification of FMRP can also modulate FMRP activity. Serine 500 of human FMRP must be phosphorylated for FMRP to bind the ribosome and inhibit translation[55]. Additionally, methylation of four specific arginines within the RGG box domain of FMRP reduces FMRP binding to GQ-assembling target mRNAs[63,64]. Regarding the aforementioned case of PSD-95 assembly by phosphorylated FMRP, signaling of the metabotropic glutamate receptor (mGluR) has been shown to decrease FMRP phosphorylation and disassemble the miRNA-125a-Ago2-PSD-95 inhibitory complex on the target mRNA[62].

FMRP has been found to repress translation of target mRNAs through different mechanisms and during different stages (Figure 1.5)[65]. It is possible FMRP employs all of these methods to specifically repress a wide catalog of target mRNAs under different conditions within the cell.

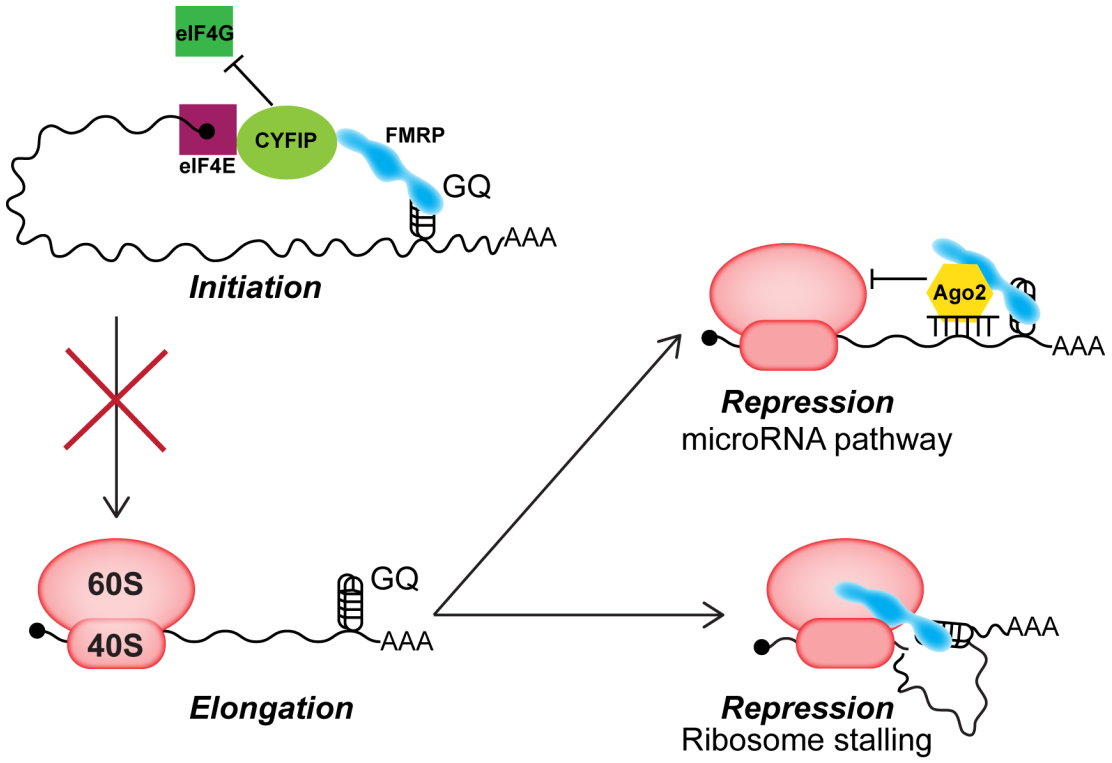

Figure 1.5: Models for translational repression by FMRP binding the target mRNA G-quadruplex. FMRP can inhibit translation initiation by binding mRNA and recruiting CYFIP, which prevents assembly of eIF4E and eIF4G into eIF4F at the 5' m^7G cap (above). Alternatively, FMRP can inhibit translation elongation via direct binding to the ribosome to stall elongation (right), or via a Ago2/RISC mediator in the microRNA pathway (above).

1.5 FMRP and the mGluR-LTD pathway

Research into FXS therapeutics has implicated FMRP in the Group 1 metabotropic glutamate receptor-long term depression (mGluR-LTD) pathway involved in learning and memory[66]. To activate mGluR-LTD, certain proteins necessary for synaptic function must be synthesized[67]. Studies in Fmr1 knockout mice revealed a connection between excess activation of mGluR-LTD and increased internalization of the α-amino-3-hydroxy-5-methyl-4-isoxazolepropionic acid receptor (AMPAR) from excessive production of the membrane protein, suggesting FMRP normally represses translation of synaptic proteins required for mGluR-LTD[68]. With loss of regulation by FMRP, these proteins are overly expressed which creates an imbalance between long term depression (LTD) and long term potentiation (LTP). This imbalance in turn is thought to degrade neurological function[69].

1.6 Translating FMRP and FXS studies into animal models

The FMR1 gene is highly conserved amongst eukaryotes. Several orthologs exist even amongst model organisms such as *Drosophila melanogaster*, zebrafish and mouse. The *Drosophila* dFmr1 gene which encodes *Drosophila* FMRP (dFMRP) shares 56% amino acid sequence identity with human FMRP (hFMRP), including RNA binding domains that are 75%

identical. In fact, dFMRP can regulate translation of some of the same mRNAs as hFMRP[70-72].

The Gideon Dreyfuss group were the first to characterize dFmr1. They highlighted the RNA binding capacity of dFMRP by mutating the KH1 and KH2 in a manner analogous to the I304N mutant isolated from the patient with the aggravated form of FXS[10,73]. They isolated two mutations, an I244N in the KH1 domain and an I307N in the KH2 domain. While homozygous expression of dFMRP caused aberrant apoptosis in the fly's eyes, expression of either I244N or I307N caused less apoptosis. However, heterozygous expression of I244N and I307N rescues wild type phenotype. This suggests both mutations are unique loss-of-function mutations presumably from dFMRP's reduced RNA binding capacities. Further studies on dFMRP found that loss of dFMRP causes abnormal axon morphology within mushroom bodies as well as disrupted courtship behaviors, both similar symptoms found in FXS patients[70,74,75].

Like dFmr1, the mouse Fmr1 gene which encodes murine FMRP (mFMRP) shares 97% amino acid sequence identity with hFMRP. Deletion of Fmr1 resulted in several FXS-like symptoms within the knockout mice[76,77]. Unfortunately, completely replicating the full mutation-causing CGG expansion at the Fmr1 5' UTR has proven futile as transcription of the Fmr1 gene does not seem to be silenced from hypermethylation as FMR1 is[15]. Although, the large CGG expanse within the 5' UTR did similarly drop mFMRP levels[78].

Conditional knockout mice have been created more recently using Cre-Lox recombination. The Cre-Lox recombination method has empowered researchers to generate specific Fmr1 mutations only to desired cells (such as neurons) and study the effects in the mouse[79].

Both *Drosophila* and mice should prove invaluable in translating understanding at a molecular level to elucidating the effects of FMRP in FXS.

Chapter 2: Investigation of FMRP binding to model RNAs using fluorescence anisotropy

Introduction

FMRP has been shown to selectively bind target mRNAs to regulate their translation within neurons[17]. Insufficient expression of FMRP causes FXS through a loss in translation regulation of both pre- and post-synaptic mRNAs[32-34]. It is therefore imperative that we determine how FMRP specifically binds these target mRNAs to regulate their expression.

Efforts from various groups have generated catalogs of potential mRNA targets for FMRP[34,36-38,40]. While there is some overlap between the proposed catalogs of mRNA targets, few of the mRNA targets have been tested for direct binding to FMRP. Furthermore, it remains unclear how the KH and RGG domains of FMRP recognize and bind specific target mRNAs. The KH1 and KH2 domains have been reported to bind WGGA and ACUK sequences in mRNAs, but these two domains have also been observed to bind directly to the ribosome between the 60S and 40S subunits to repress translation of mRNAs[35,36]. On the other hand, the RGG motif has been shown to bind RNA GQ structures, but most of the studies use an in vitro selected RNA that may not be representative of RNA GQ structures FMRP may bind in vivo[23,24,33,41,42,57]. Whether or not the KH domains and the RGG motif bind the

ribosome and target mRNAs independently or in concert requires investigation.

In chapter 2, we describe how we (1) isolated full-length human FMRP (hFMRP) and various truncated human FMRP constructs and then (2) quantified their binding affinities to a set of fluorescently-labeled model RNAs using fluorescence anisotropy. The model RNAs represent RNA sequences or structural motifs previously reported to bind FMRP. Isolating specific regions of hFMRP and quantifying their binding affinity to each model RNA enabled us to both test the validity of the proposed binding sequences and clarify the roles of the KH1, KH2 and RGG domains in binding the model RNAs.

Results and Discussion

2.1 Isolation of the FMRP variants

Four human FMRP (hFMRP) variants and one *Drosophila* FMRP (dFMRP) variant were purified from *E. coli* expression strains for our studies (Figure 2.1). Efforts to biochemically characterize FMRP binding to RNA in our lab began with a previous student in our lab, Lila Mouakkad. Lila tested FMRP binding to RNA using a stable fragment of dFMRP missing the N-terminal region which we call NT-dFMRP (Figure 2.1E). As FXS is caused by human FMRP, we would ultimately need to characterize hFMRP binding to RNA to better understand the role FMRP binding to RNAs plays in FXS. Since then, we have successfully expressed and purified hFMRP from *E. coli* [35].

The full-length human FMRP (hFMRP) spans 632 amino acids and is a 72 kDa protein (Figure 2.1A). It contains a 46 residue Agenet 1 and a 50 residue Agenet 2 domain (also known as Tudor 1 and Tudor 2 domains, respectively), as well as a 76 residue K-Homology (KH) 0 domain within its N-terminal region. Downstream of the N-terminal region lies a 64 residue KH1 and a 124 residue KH2 domain. Further downstream lies the 25 residue arginine-arginine-glycine (RGG) box domain containing the hypothesized RNA GQ-binding RGG motif.

The N-terminus truncated human FMRP (NT-hFMRP) spans 415 amino acids and is a 47 kDa protein (Figure 2.1B). NT-hFMRP was cloned by truncating hFMRP immediately upstream of its KH1 domain. This truncation essentially eliminates the FMRP N-terminal region, which is thought to be responsible for FMRP dimerization, while conserving its three proposed RNA-binding domains.

We also wanted to generate a human FMRP construct containing the RGG domain and none of the other proposed RNA-binding domains to evaluate the importance of the RGG motif in binding RNA GQ structures. We truncated hFMRP immediately upstream of its RGG box domain and cloned it immediately downstream of a glutathione S-transferase (GST) tag to generate a GST-hRGG fusion construct that spans 352 amino acids and is a 40 kDa protein (Figure 2.1C). In addition to the usual benefits conferred by the GST tag, its fusion to the RGG domain makes the protein bulky enough to observe clear shifts in anisotropy upon binding the fluorescein-labeled RNAs.

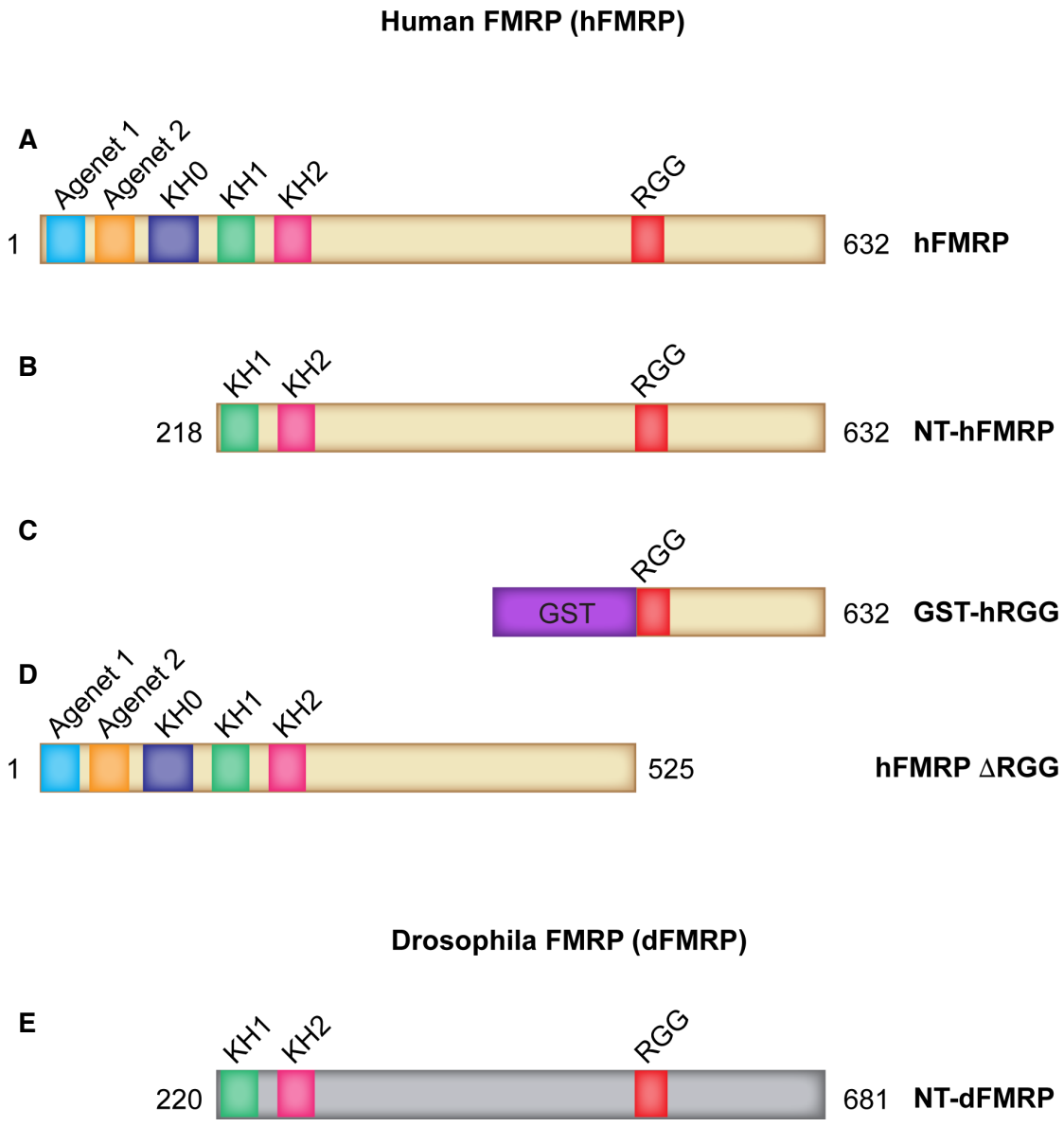

Figure 2.1: FMRP constructs. (A) hFMRP is the full length human FMRP. (B) NT-hFMRP is a N-terminus truncated hFMRP spanning R218-P632. (C) GST-hRGG is a fusion between the glutathione S-transferase and the RGG box domain-containing sequence of hFMRP spanning G531-P632. (D) hFMRP ΔRGG spans M1-F525. (E) NT-dFMRP is a N-terminus truncated dFMRP spanning G220-N681.

We are also planning to measure the binding affinity of a hFMRP construct missing the RGG domain (hFMRP ΔRGG) to our model RNAs (Figure 2.1D). This may also help us better understand how important the KH domains are in binding the model RNAs. Of particular interest would be the model RNAs that the KH domain-containing FMRP constructs bind with higher affinity compared to the GST-hRGG construct.

2.2 Isolation and characterization of the model RNAs

We first determined which model RNAs folded into GQ structures (Table 2.1). Apart from the UG_4U RNA, we standardized our model RNAs to 18 nucleotides in length; M1 5' UTR served as an 8-mer RNA size marker. We also tested the GACR (R = A or G) and WGGA (W = A or U) RNA motifs that were

Table 2.1: Model RNAs and their sequences.

RNA	RNA Sequence
$PolyG_{18}$	5' GGGGGGGGGGGGGGGGGU 3'-Fluorescein
UG_4U	5' UGGGGU 3'-Fluorescein
$(UGGA)_4$	5' UUGGAUGGAUGGAUGGAU 3'- Fluorescein
$(GACG)_4$	5' UGACGGACGGACGGACGU 3'- Fluorescein
$PolyC_{18}$	5' CCCCCCCCCCCCCCCCCU 3'-Fluorescein
$PolyC_{18}$ (ACUU)	5' CCCCCCCACUUCCCCCCU 3'-Fluorescein
$PolyC_{18}$ (UGGA)	5' CCCCCCCUGGACCCCCCU 3'-Fluorescein
CR1	5' GCUAUCCAGAUUCUGAUU 3'-Fluorescein
$PolyA_{18}$	5' AAAAAAAAAAAAAAAAAU 3'-Fluorescein
M1 5' UTR	5' GGUAGAUA 3'-Fluorescein

reported to bind FMRP by designing 18-mer RNAs containing four tandem repeats of each, $(GACG)_4$ and $(UGGA)_4$. The GACR motif is suspected of binding a KH domain of FMRP while the WGGA motif is suspected of assembling into an intramolecular GQ and binding the FMRP RGG motif[36,38,40]. The Control RNA 1 (CR1) and $PolyC_{18}$ RNAs would serve as negative control RNAs, while the $PolyC_{18}$(ACUU) and $PolyC_{18}$(UGGA) were designed to test if a single ACUK or WGGA motif is sufficient for binding FMRP[36]. $PolyA_{18}$ would serve as a stacked RNA marker for gel analyses.

Our hypothesis that UG_4U and $PolyG_{18}$ assemble into a GQ was tested by denaturing and native PAGE analyses (Figure 2.2). We confirmed both UG_4U and $PolyG_{18}$ assemble GQ structures by comparing their migrations on denaturing urea PAGE and native PAGE. UG_4U was determined to assemble into the reported 24-mer intermolecular GQ using two methods: (1) a denaturing PAGE in the presence of KCl, NaCl or LiCl as described previously by Kim, Cheong and Moore, and (2) native PAGE comparing UG_4U migration to the 8-mer M1 5' UTR and control 18-mer $PolyC_{18}$ and CR1 RNAs which are known to remain largely unfolded[44]. The denaturing PAGE showed UG_4U migrating as two different species, with the slower species being potassium-dependent (Figure 2.2A). The native PAGE showed UG_4U migrating much slower than M1 5' UTR and slightly slower than the 18-mer RNAs (Figure 2.2C). Together, both PAGE analyses suggest UG_4U can assemble into a robust RNA GQ. Similarly, $PolyG_{18}$ migrated much slower compared to the control RNAs, possibly as if it

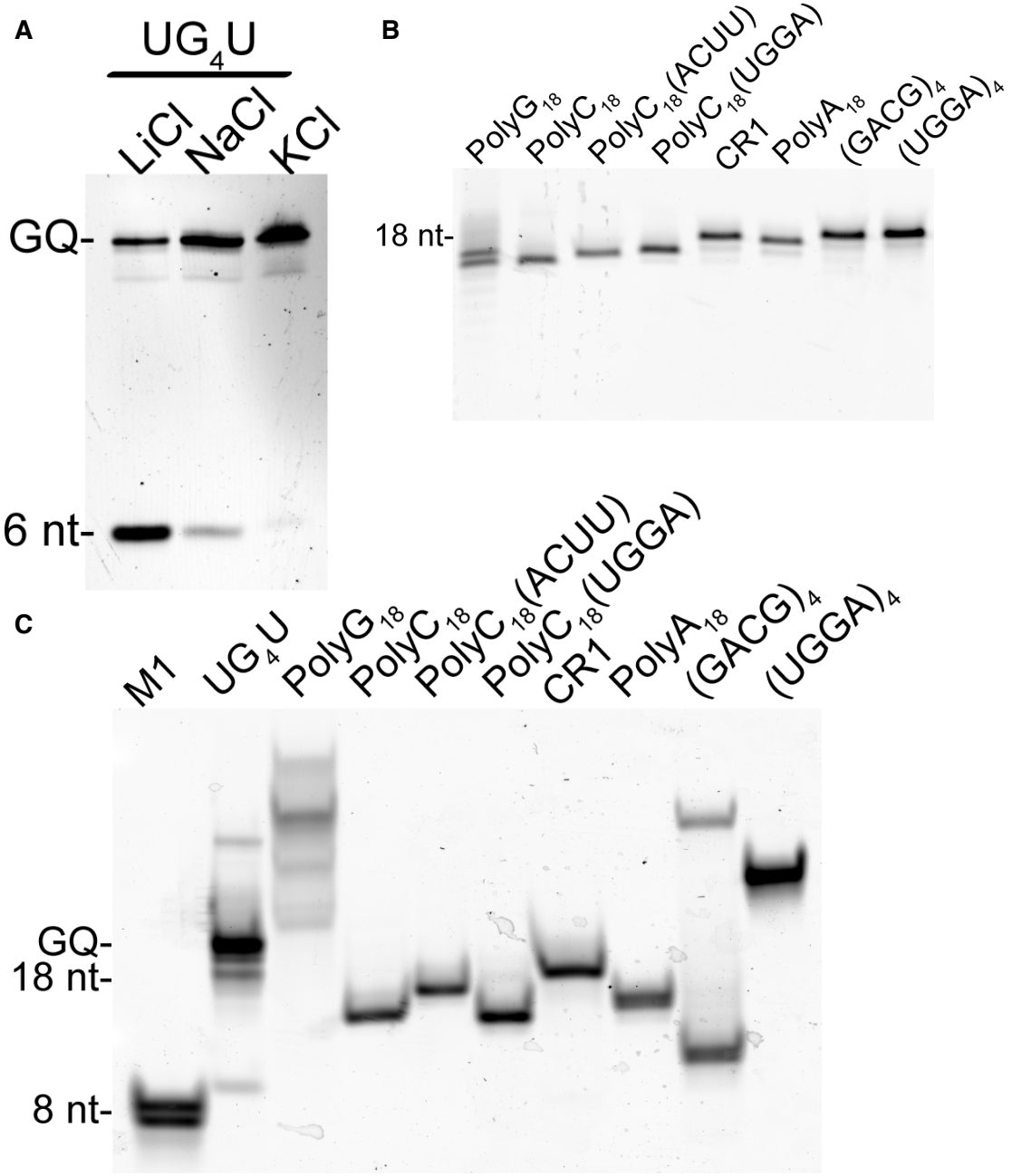

Figure 2.2: Polyacrylamide gel analysis of the model RNAs. (A) Urea PAGE of UG_4U denatured and then refolded in the presence of (left to right): 100 mM LiCl, 100 mM NaCl or 100 mM KCl. (B) Urea PAGE of the denatured 18-mer model RNAs (left to right): $PolyG_{18}$, $PolyC_{18}$, $PolyC_{18}(ACUU)$, $PolyC_{18}(UGGA)$, CR1, $PolyA_{18}$, $(GACG)_4$, and $(UGGA)_4$. (C) Native PAGE of the 18-mer model RNAs (left to right): M1 5' UTR, UG_4U, $PolyG_{18}$, $PolyC_{18}$, $PolyC_{18}(ACUU)$, $PolyC_{18}(UGGA)$, CR1, $PolyA_{18}$, $(GACG)_4$, and $(UGGA)_4$. All RNAs were 3' labeled with fluorescein and visualized using a ChemiDoc XRS^+ (A) or a Typhoon FLA 9500 (B and C).

were a 36-mer RNA, suggesting it may assemble a stable intermolecular GQ structure (Figure 2.2B and C).

The $(GACG)_4$ and $(UGGA)_4$ RNAs reported to bind FMRP's KH and RGG domains, respectively, migrated differently from the other RNAs. The $(GACG)_4$ RNA migrated as two species, one migrating slowly at a similar rate as $PolyG_{18}$ and the other migrating faster than the unstructured $PolyC_{18}$ RNAs (Figure 2.2A). The slower $(GACG)_4$ species may be a stable dimer structure, similar in size and shape to the predominant $PolyG_{18}$ species, while the faster $(GACG)_4$ species may be a compact, folded monomer structure able to migrate faster than the unfolded $PolyC_{18}$ RNAs. Unlike $PolyG_{18}$ however, the QGRS mapper does not predict $(GACG)_4$ to fold into an intramolecular GQ[80]. Therefore, if $(GACG)_4$ is indeed assembling into a GQ, it would probably be a dimeric 36-mer intermolecular GQ that QGRS cannot predict. The faster $(GACG)_4$ species could be a stem-loop structure which would be expected to migrate faster than the unfolded 18-mer control RNAs[81]. The $(UGGA)_4$ RNA is predicted by QGRS to assemble an intramolecular GQ like $PolyG_{18}$, and it is migrating at a similar rate as a secondary $PolyG_{18}$ species below the primary species. Considering both $PolyG_{18}$ and $(UGGA)_4$ are 18-mer RNAs capable of assembling into intramolecular GQ structures, it is possible the secondary $PolyG_{18}$ band and the $(UGGA)_4$ band are intramolecular GQ structures, while the primary $PolyG_{18}$ band is a dimeric 36-mer intermolecular GQ.

2.3 Quantification of FMRP binding affinity for G-quadruplex RNAs

While little is known regarding how FMRP specifically binds target mRNAs, reports of FMRP binding to the RNA GQ warrants further investigation. A previous study showed that the KH1 and KH2 domains of *Drosophila* FMRP bind to the ribosome while the RGG domain is exposed in solution and available for binding to mRNA[5,12]. Therefore, we tested whether the FMRP RGG motif can specifically bind an RNA GQ. We quantified the equilibrium dissociation constant (K_D) of human FMRP and *Drosophila* FMRP binding to fluorescein-labeled GQ-assembling RNAs using fluorescence anisotropy (Table 2.1). We measured binding affinities of several human FMRP variants and *Drosophila* FMRP to RNAs that either assemble GQ structures — PolyG_{18} and UG_4U — or contain motifs that have been reported to bind FMRP — $(GACG)_4$ and $(UGGA)_4$ (Figure 2.3). We used four 18-mer RNAs with different base compositions that cannot assemble GQ structures — PolyC_{18}, PolyC_{18}(ACUU), PolyC_{18}(UGGA) and CR1 — as control RNAs (Table 2.1). The ACUU and UGGA sequences were inserted into the PolyC_{18} negative control RNA to test whether the two reported sequences could confer binding to FMRP; the CR1 sequence contains all four bases that cannot assemble GQ and does not contain any ACUK or WGGA sequences, and served as a control RNA sequence[6].

We found that hFMRP and NT-hFMRP specifically bind PolyG_{18} and UG_4U RNAs with nanomolar affinities, and cannot bind any of the control RNAs (Figure 2.3A and B, Table 2.2). Also, NT-hFMRP bound both PolyG_{18} and UG_4U

with approximately ten-fold higher affinity than hFMRP. It is possible the equilibrium between monomeric and dimeric hFMRP reduces the binding affinity via blockage of the RNA GQ to the FMRP RGG motif. With NT-hFMRP existing as a monomer in solution, its RGG motif may be more accessible, explaining the higher binding affinity. Interestingly, all the FMRP variants failed to bind the $(GACG)_4$ RNA which was reported to bind a KH domain of FMRP. It is possible the four tandem GACG repeats is insufficient for binding to FMRP directly, or FMRP associates with GACR-rich sequences through a secondary protein. Also, if $(GACG)_4$ is indeed folding into a dimeric 36-mer intermolecular GQ, it seems FMRP cannot bind it at all, despite being able to bind the tetrameric 24-mer intermolecular $(UG_4U)_4$ GQ. On the other hand, both hFMRP and NT-hFMRP bound $(UGGA)_4$ with similar affinities. The $(UGGA)_4$ RNA is predicted by QGRS to assemble into an intramolecular RNA GQ which the native PAGE supports (Figure 2.2A). However, unlike the case with $PolyG_{18}$ and UG_4U, the dimeric full-length hFMRP seems to bind $(UGGA)_4$ with similar affinity as the monomeric NT-hFMRP. It would be interesting if $(UGGA)_4$ assembles a unique structure, GQ or otherwise, that stabilizes the full-length hFMRP as a monomer such that its affinity is like that of the monomeric NT-hFMRP.

NT-dFMRP could not bind $PolyG_{18}$, UG_4U and $(UGGA)_4$ with as high affinity as NT-hFMRP could (Figure 2.3C, Table 2.2). We realized hFMRP and NT-hFMRP binding to RNA was done in 75 mM KCl while NT-dFMRP binding to RNA was done in 300 mM KCl (NT-dFMRP precipitates out of solution in 75 mM KCl). Therefore, we compared NT-hFMRP binding affinity to GQ model

RNAs $PolyG_{18}$ and UG_4U in 75 mM KCl, 150 mM KCl and 300 mM KCl (Figure 2.4). We found that increasing KCl concentration from 75 mM to 300 mM decreased hFMRP affinity for $PolyG_{18}$ and UG_4U significantly, explaining NT-dFMRP's weaker binding to the GQ RNAs. This also suggests electrostatic interactions are involved in FMRP binding to RNA GQ structures.

To determine whether the RGG motif is specifically responsible for binding the RNA GQ, we also tested binding of the RNAs to a GST-fused human RGG motif (GST-hRGG) composed solely of the hFMRP C-terminus up to the RGG motif and an N-terminal GST fusion tag (Figure 2.3D). Consistent with previous studies, the RGG motif is sufficient for binding $PolyG_{18}$ and UG_4U. In fact, it bound $PolyG_{18}$ and UG_4U with similar affinity as NT-hFMRP (Figure 2.3D, Table 2.2). Interestingly, GST-hRGG bound $(UGGA)_4$ with significantly lower affinity than hFMRP and NT-hFMRP, suggesting the KH domains may be working together with the RGG motif to bind $(UGGA)_4$. We plan to test binding of $PolyG_{18}$, UG_4U and $(UGGA)_4$ to a hFMRP variant missing the RGG motif (hFMRP ΔRGG). This test will further elucidate how the RGG motif and the KH domains bind the various structures assembled by the $PolyG_{18}$, UG_4U and $(UGGA)_4$ RNAs.

Finally, we determined if the FMRP proteins are binding a GQ conformation of $PolyG_{18}$ and UG_4U, and not an unfolded population of the RNAs. To assess binding to a GQ structure, we measured binding of GST-hRGG to $PolyG_{18}$ and UG_4U in the presence of LiCl or KCl (Figure 2.5). Indeed, simply having lithium ions in place of potassium ions decreased the shift in

anisotropy, consistent with $PolyG_{18}$ and UG_4U favoring non-GQ conformations in lithium. Altogether, the concomitant drop in anisotropy change with the drop in the GQ RNA population supports our hypothesis that the RGG motif specifically binds RNA GQ structure.

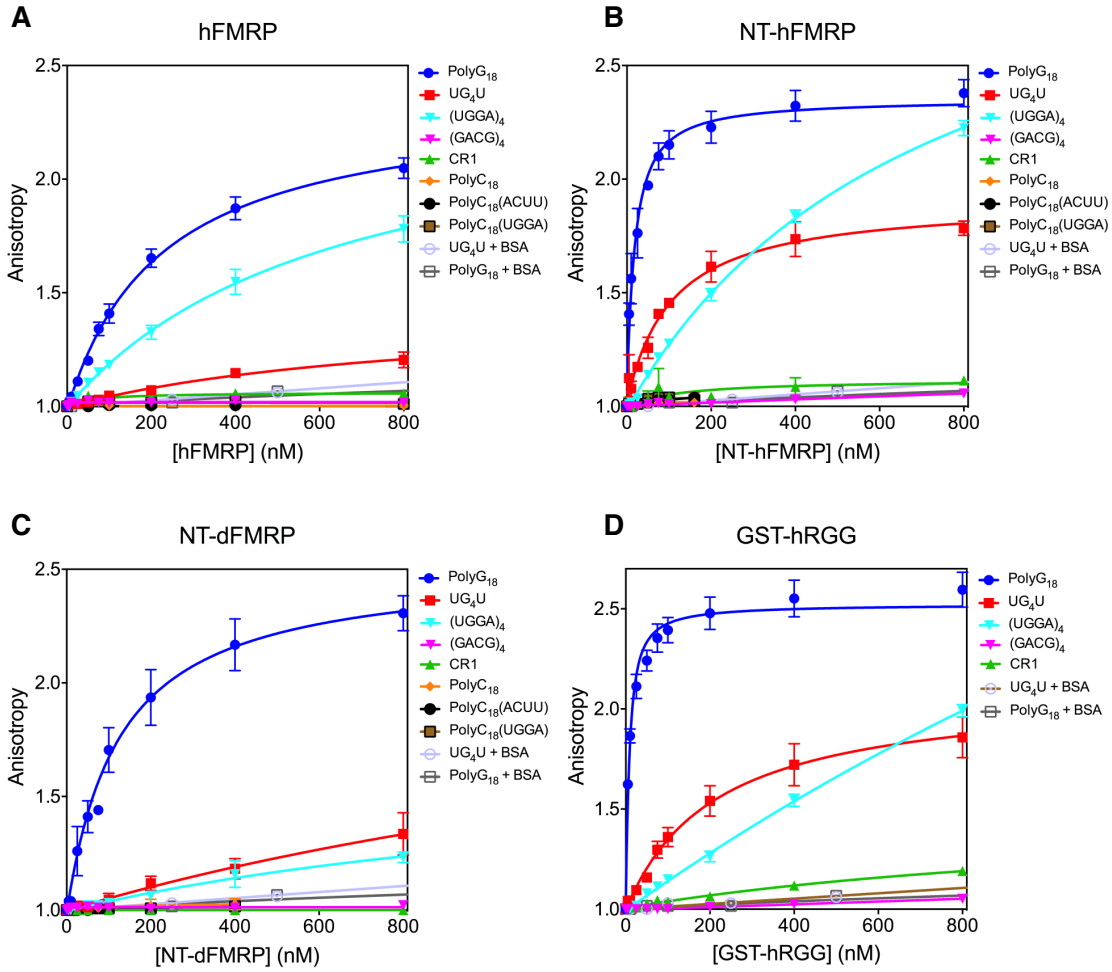

Figure 2.3: Fluorescence anisotropy measurements of FMRP binding to model RNAs. 0-800 nM of (A) Full-length hFMRP, (B) NT-hFMRP, (C) NT-dFMRP and (D) GST-hRGG titrated against 5 nM fluorescein-labeled RNAs. Bovine serum albumin (BSA) binding to the GQ-assembling RNAs was also measured to test for inherent protein binding qualities in the two RNAs. All FMRP binding trials were performed at least three times, with error bars depicting standard deviation between trials.

Table 2.2: Binding affinities (K_D) between FMRP and the model RNAs.

	Protein			
RNA	**hFMRP**	**NT-hFMRP**	**NT-dFMRP**	**GST-hRGG**
UG_4U	847 ± 299 nM	104 ± 16 nM	> 2 μM	208 ± 32 nM
$PolyG_{18}$	192 ± 32 nM	16.4 ± 2.4 nM	121 ± 18 nM	6.7 ± 1.0 nM
$(UGGA)_4$	597 ± 84 nM	714 ± 50 nM	1.6 ± 0.9 μM	3.9 ± 0.9 μM
$(GACG)_4$	-	-	-	-

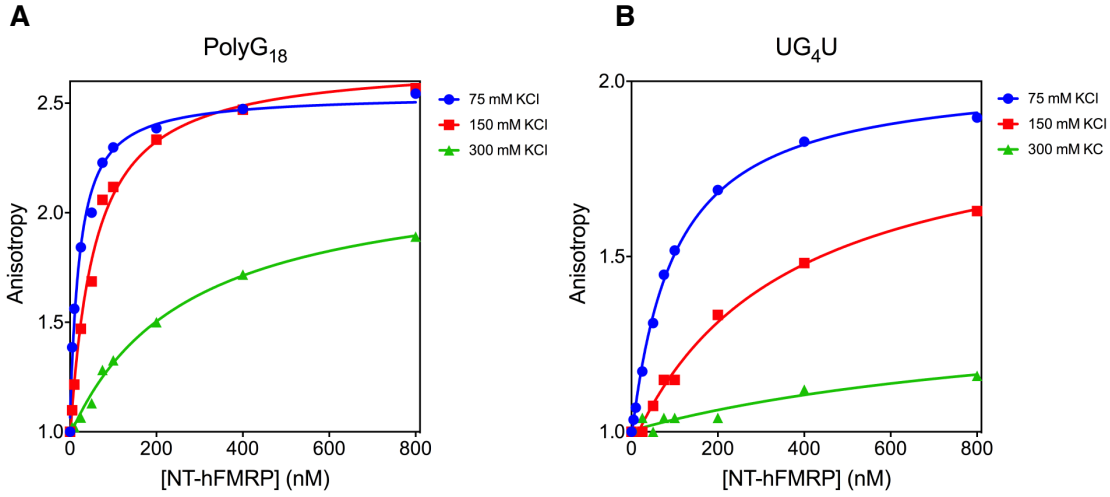

Figure 2.4: FMRP binding to PolyG$_{18}$ and UG$_4$U RNAs is affected by KCl concentration. (A) NT-hFMRP binding to PolyG$_{18}$ in 75 mM KCl, 150 mM KCl and 300 mM KCl. (B) NT-hFMRP binding to UG$_4$U in 75 mM KCl, 150 mM KCl and 300 mM KCl. Error bars denote standard deviation from three trials.

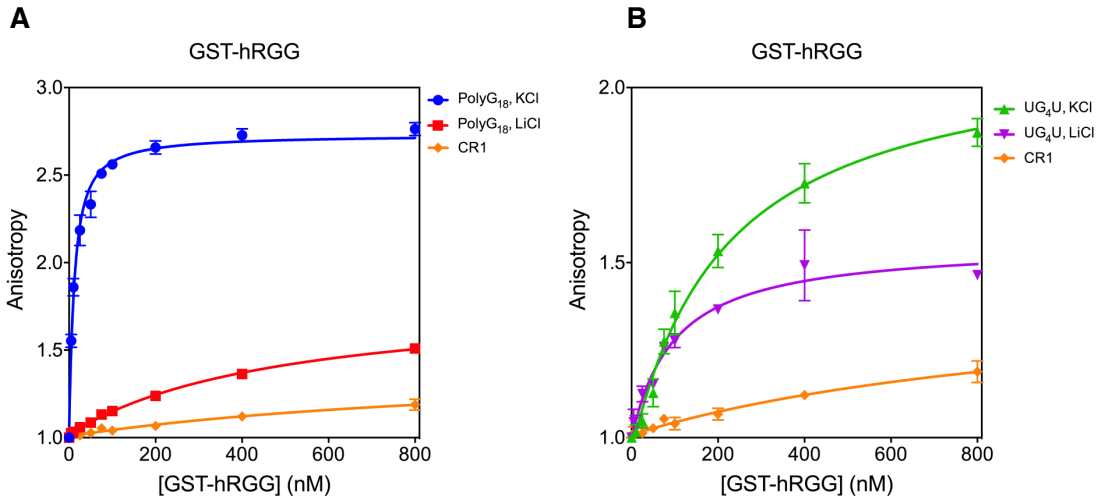

Figure 2.5: FMRP binding efficiency to PolyG$_{18}$ and UG$_4$U RNAs in KCl and LiCl. (A) GST-hRGG binding to PolyG$_{18}$ in 75 mM KCl and 75 mM LiCl. (B) GST-hRGG binding to UG$_4$U in 75 mM KCl and 75 mM LiCl. Error bars denote standard deviation from three trials.

Chapter 3: Conclusion and Future Directions

3.1 Test binding affinity of hFMRP ΔRGG to the model RNAs

The fluorescence anisotropy binding experiments of hFMRP, NT-hFMRP and GST-hRGG to the model RNAs provided some key insight into FMRP binding specificity for target RNAs. The results show: (1) monomeric FMRP variants binds the RNAs with higher affinity than dimeric FMRP variants, (2) the RGG motif is sufficient for specific binding to both the $PolyG_{18}$ and UG_4U GQ model RNAs, (3) FMRP binding to RNA GQ structures is indeed K^+-dependent, and (4) FMRP binding to RNA GQ structures is sensitive to salt concentration.

One question left unanswered is how hFMRP binds the $(UGGA)_4$ RNA. Unlike in the case with $PolyG_{18}$ and UG_4U, GST-hRGG bound $(UGGA)_4$ with significantly lower affinity than hFMRP and NT-hFMRP did. This suggests the KH domains are likely involved in binding the $(UGGA)_4$ RNA. We plan to test binding of hFMRP ΔRGG to $PolyG_{18}$, UG_4U and $(UGGA)_4$ RNAs. This study will (1) serve as a secondary test for RGG motif binding to the $PolyG_{18}$ and UG_4U GQ structures and (2) test whether the KH domains are mainly responsible for binding the $(UGGA)_4$ RNA structure.

3.2 Test each FMRP variant's ability to repress translation

While we have gained some understanding of how FMRP binds GQ RNA structures from the binding studies, we do not know if these structures are capable of regulating translation upon binding FMRP. We plan to test our model GQ structures' ability to repress translation upon binding to FMRP using an in vitro translation assay.

A recent study suggests RNA sequences known to assemble into GQ structures remain unfolded in eukaryotic cells[82]. While the study was done under steady state conditions, it does bring into question how transient these GQ structures are in vivo. The GQ may be a transient structure that is only available for binding FMRP under certain circumstances. Considering FMRP's proposed role in tightly regulating translation of various mRNAs in neurons, it makes sense for the GQ structures only being available for FMRP for short moments. As such, we plan to test FMRP's ability to repress translation of reporter mRNAs containing a model GQ structure within their 3' UTRs.

3.3 Isolate a human FMRP monomer for ribosome binding studies

FMRP has been found to be capable of both dimerizing with itself and heterodimerizing with both the Fragile X Related Proteins 1 and 2 (FXR1 and FXR2, respectively)[73,83,84]. A region within the KH0 domain is reported to be responsible for dimerization via coiled coil interaction[29,84]. Our lab's ribosome

binding studies with the various hFMRP and dFMRP protein variants show the full-length FMRP variants binding the 80S ribosome with much lower affinity than the NT-FMRP variants. One major difference between these two classes of FMRP is the presence of the N-terminal domain which contains the dimerization site. Considering the structure of NT-dFMRP bound to the 80S ribosome, it is unlikely a FMRP dimer can bind the ribosome[35]. We therefore hypothesize FMRP resides as a dimer but binds the ribosome as a monomer, adding another layer of regulation to FMRP activity in vivo.

To test our hypothesis that FMRP binds the ribosome as a monomer, we are working to isolate a hFMRP mutant that cannot dimerize. Considering the substantial evidence suggesting the key dimerization interface is a coiled coil interaction within the KH0 domain, we are attempting to disrupt this interaction by mutating key residues that may be stabilizing the coiled coil (Figure 3.1). Once we isolate a hFMRP monomer, we can proceed to solve a high-resolution structure of a human 80S ribosome-hFMRP complex by cryo-EM.

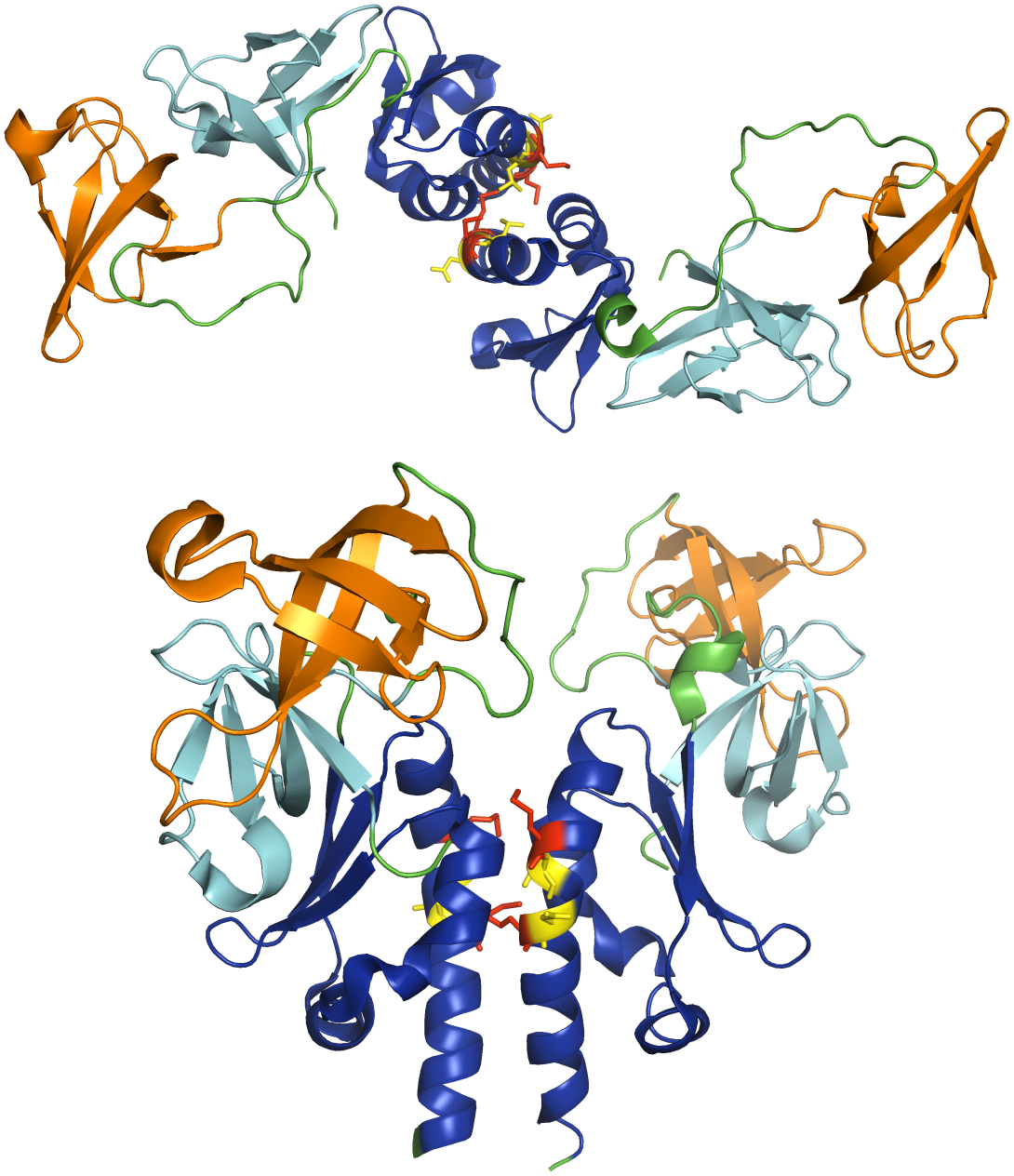

Figure 3.1: Human FMRPΔ 3.0 Å crystal structure depicting the KH0 dimer interface. Alternate views of the FMRPΔ structure depicting two monomers in adjacent unit cells potentially dimerizing through a coiled-coil interaction at the C-terminal helices of the KH0 domains depicted in blue. The key KH0 dimer interface residues M183 and M187 are highlighted in red, while residues L184 and D186 are highlighted in yellow. The Agenet 1 and Agenet 2 domains are highlighted in cyan and orange, respectively. The structure was accessed from the Protein Data Bank with entry code 4OVA.

Chapter 4: Materials and Methods

4.1 Expression and purification of human FMRP and Drosophila FMRP (hFMRP, NT-hFMRP, hFMRP ΔRGG, GST-hRGG domain, hFMRPΔ, and NT-dFMRP)

The human FMR1 isoform 1 gene was assembled from *E. coli* codon-optimized gene blocks (IDT). An extra methionine at amino acid position 162 was mistakenly incorporated (Table 4.1). The gene encoding the full-length isoform 1 human FMRP (hFMRP) spanning residues E2-P632 was subcloned into the SUMO protein tag expression vector pETHSUL. Then the Met162 was deleted by site-directed mutagenesis, and subsequently subcloned into the LIC expression vector pMCSG7 (DNASU plasmid repository) with the 5' TEV cleavage site excluded. The N-terminus truncated human FMRP (NT-hFMRP) spanning residues R218-P632 and the RGG motif deletion human FMRP (hFMRP ΔRGG) spanning residues E2-F525 were both generated from the hFMRP (pMCSG7) plasmid via PCR deletion mutagenesis. The human RGG domain (GST-hRGG) spanning residues G531-P632 was subcloned into the LIC expression vector pMCSG10 (DNASU) which confers a N-terminal hexahistidine-GST fusion tag that is cleavable using TEV protease. The N-terminal hFMRP (hFMRPΔ) spanning residues M1-Q208 was subcloned into pMCSG7 to produce a protein with an N-terminal hexahistidine tag and a downstream TEV cleavage site for tag removal.

Table 4.1: Human FMR1 gene sequence.

Gene	Gene Sequence
Human FMR1 (extra Met162 codon in red)	ATGGAAGAACTGGTGGTTGAAGTGCGTGGTTCGAATGGCGC GTTTTACAAAGCCTTCGTGAAGGACGTTCATGAAGATAGCATT ACCGTTGCGTTTGAAAACAATTGGCAGCCGGATCGCCAAATC CCGTTTCACGACGTTCGTTTCCCGCCGCCGGTCGGTTACAAC AAGGACATCAACGAATCTGATGAAGTGGAAGTTTACAGTCGC GCGAACGAAAAAGAACCGTGCTGTTGGTGGCTGGCCAAAGT CCGTATGATTAAGGGCGAATTTTATGTGATCGAATACGCGGC CTGCGATGCGACCTATAATGAAATTGTCACGATCGAACGCCT GCGTAGCGTGAACCCGAATAAACCGGCCACCAAGGATACGTT CCATAAAATTAAGCTGGACGTGCCGGAAGATCTGCGTCAGAT GTGTGCAAAAGAAGCAGCTCACAAGGATTTTAAAAAGGCGGT CGGTGCCTTCTCTGTGACCTATATGGACCCGGAAAACTACCA GCTGGTGATTCTGTCCATCAATGAAGTTACCTCAAAACGCGC GCACATGCTGATTGATATGCACTTTCGCTCGCTGCGTACCAA ACTGAGCCTGATCATGCGTAACGAAGAAGCATCCAAGCAGCT GGAAAGTTCCCGCCAACTGGCTTCACGTTTTCATGAACAGTT CATTGTGCGCGAAGATCTGATGGGCCTGGCAATTGGTACCCA CGGCGCGAATATCCAGCAAGCCCGTAAAGTCCCGGGTGTGA CGGCTATTGATCTGGACGAAGATACCTGCACGTTCCATATCTA CGGCGAAGACCAAGATGCAGTTAAAAAGGCTCGCTCCTTTCT GGAATTCGCCGAAGATGTTATTCAGGTCCCGCGTAACCTGGT GGGTAAAGTTATCGGTAAAAATGGCAAGCTGATTCAAGAAATC GTTGATAAAAGCGGCGTCGTGCGCGTCCGTATTGAAGCAGAA AACGAAAAGAACGTGCCGCAGGAAGAAGAAATCATGCCGCC GAACTCACTGCCGTCGAACAATAGCCGCGTTGGTCCGAATGC TCCGGAAGAAAAGAAACATCTGGATATCAAGGAAAACAGCAC CCACTTTTCTCAGCCGAATAGTACGAAGGTGCAACGTGTCCT GGTGGCATCATCGGTTGTCGCTGGCGAATCTCAGAAACCGGA ACTGAAGGCCTGGCAGGGTATGGTTCCGTTTGTTTTCGTCGG CACCAAAGATAGTATCGCAAACGCTACGGTTCTGCTGGACTA TCATCTGAATTACCTGAAAGAAGTCGATCAGCTGCGCCTGGA ACGTCTGCAAATTGACGAACAGCTGCGCCAAATCGGTGCAAG CTCTCGTCCGCCGCCGAACCGTACCGATAAAGAAAAGTCTTA TGTGACGGATGACGGTCAGGGTATGGGTCGTGGCAGTCGTC CGTATCGCAATCGTGGTCATGGTCGTCGCGGTCCGGGTTACA CCTCCGGTACGAACTCTGAAGCAAGTAACGCTTCCGAAACCG AATCAGACCACCGCGATGAACTGTCAGATTGGTCGCTGGCGC CGACGGAAGAAGAACGTGAATCTTTTCTGCGTCGCGGTGACG GTCGTCGCCGTGGCGGTGGCGGTCGTGGTCAGGGCGGTCG CGGCCGTGGCGGTGGCTTCAAAGGCAATGATGACCATAGTC GTACCGATAATCGTCCGCGTAATCCGCGCGAAGCCAAAGGTC GTACCACGGATGGCTCGCTGCAGATTCGCGTGGACTGTAACA ATGAACGTAGCGTTCACACCAAAACGCTGCAAAACACCAGTT CCGAAGGTAGCCGCCTGCGTACGGGCAAAGATCGTAATCAG AAAAAGGAAAAACCGGACTCAGTGGATGGTCAACAACCGCTG GTCAATGGTGTGCCGTAA

The Drosophila *Fmr1* gene (GenBank ID code AF305881) was obtained from Prof. Gideon Dreyfuss (University of Pennsylvania). N-terminus truncated drosophila FMRP (NT-dFMRP) spanning residues 220 to 681 was subclonedinto pMCSG7 to produce a protein with an N-terminal hexahistidine tag and a downstream TEV cleavage site for tag removal.

The NT-hFMRP, GST-hRGG and NT-dFMRP expression plasmids were transformed in *E. coli* BL21(DE3) cells (Novagen) while the hFMRP expression plasmid was transformed in *E. coli* Rosetta 2 (DE3) pLysS cells (Novagen). Cells were grown in a 10 mL LB starter culture supplemented with 100 ug/mL ampicillin overnight; 25 ug/mL chloramphenicol was added when culturing Rosetta cells. 1 L LB media containing the appropriate antibiotics were inoculated with 3-5 mL of BL21 overnight starter culture, or 10 mL Rosetta 2 overnight starter culture, and outgrown at 37°C to OD_{600} between 0.5 and 0.8. The cultures were placed at 4°C for 20 minutes without shaking, then induced with 0.2 mM IPTG and grown for an additional 18-20 hours at 16-18°C. Bacterial cells were pelleted at 5000 rpm for 15 minutes in Beckman JLA-10.500 rotor.

NT-dFMRP expressing cells were re-suspended in lysis buffer containing 25 mM Tris pH 7.4, 0.5 M NaCl, 5 mM imidazole, 10 mM β-mercaptoethanol and 1 mM PMSF and then lysed by sonication (10 cycles of 8 seconds pulse with 30 seconds rest, repeated twice). The NT-hFMRP expressing cells were re-suspended in lysis buffer containing 25 mM Tris pH 7.4, 0.5 M NaCl, 5 mM imidazole, 10 mM β-mercaptoethanol and 1 mM PMSF, and then lysed by sonication (10 cycles of 8 seconds pulse with 30 seconds rest, repeated twice).

The hFMRP expressing cells were re-suspended in lysis buffer containing 25 mM Tris pH 7.4, 0.5 M NaCl, 5 mM imidazole, 10% glycerol, 10 mM β-mercaptoethanol and 1 mM PMSF, then lysed by French Press (2 passes with an applied pressure of 14000-20000 psi). Lysates were then clarified by centrifugation at 20,000 g for 60 minutes at 4°C and supernatant was incubated for 1 hour with pre-equilibrated Ni-NTA beads (Qiagen) at 4°C. The mixture was loaded onto a column and beads were washed with 10 column volumes of wash buffer containing 25 mM Tris pH 7.4, 0.5 M NaCl and 25 mM imidazole. Tagged FMRP was eluted in 1 column volume fractions with elution buffer containing 25 mM Tris pH 7.4, 0.5 M NaCl and 250 mM imidazole. Also, 10% glycerol is present in all hFMRP purification buffers. Fractions containing protein were identified by 12% SDS-PAGE. The elution fractions containing expressed FMRP were pooled and concentrated in a Centricon centrifugal filter (Millipore), and further purified on a Superdex 75 16/60 or Superdex 200 16/60 gel filtration column (GE Healthcare) in gel filtration buffer. Gel filtration buffer composition varied between FMRP constructs. NT-dFMRP gel filtration buffer contained 25 mM Tris pH 7.4, 0.3 M KCl and 1 mM DTT. NT-hFMRP gel filtration buffer contained 25 mM Tris pH 7.4, 0.25 M KCl and 1 mM DTT. Full-length hFMRP gel filtration buffer contained 25 mM Tris pH 7.4, 0.25 M KCl, 0.25 M imidazole and 10% glycerol.

GST-hRGG expressing cells were re-suspended in binding buffer containing PBS pH 7.3, 2 mM DTT and 1 mM PMSF, and then lysed by sonication (10 cycles of 8 seconds pulse with 30 seconds rest, repeated twice).

Lysate was then clarified by centrifugation at 20,000 g for 60 minutes at 4°C and supernatant was incubated for 1 hour with 2 mL of pre-equilibrated glutathione beads (GE Healthcare) at room temperature. The mixture was loaded onto a column and beads were washed with 10 column volumes of binding buffer. Tagged FMRP was eluted in 1 column volume fractions with fresh elution buffer containing 50 mM Tris pH 8.0 and 10 mM reduced glutathione. Fractions containing GST-hRGG were identified by 12% SDS-PAGE. The elution fractions containing the expressed protein were pooled and concentrated in a Centricon centrifugal filter (Millipore), and further purified on a Superdex 75 16/60 gel filtration column (GE Healthcare) in gel filtration buffer containing 50 mM Tris pH 7.5 and 1 mM DTT.

4.2 Purification and analysis of fluorescent RNAs

The $PolyG_{18}$, UG_4U, CR1, $PolyC_{18}$, $PolyC_{18}$(ACUU), and $PolyC_{18}$(UGGA), $(GACG)_4$, and $(UGGA)_4$ RNAs were purchased from Thermo Scientific/Dharmacon pre-modified with a fluorescein tag at the 3' end. Each RNA came with a 2'-ACE protecting group, so deprotection was required to restore the 2'-hydroxyl. First each RNA was resuspended in 400 uL of the provided 2'-Deprotection buffer (100 mM acetic acid-TEMED pH 3.8) by pipetting and then vortexing for 10 seconds. Each RNA suspension was then centrifuged for 10 seconds at room temperature, and incubated for 30 minutes in a 60°C water bath. Finally, each RNA suspension was thoroughly dried using

a SpeedVac which was covered with aluminum foil to minimize light exposure to the fluorescein tag. The deprotected RNA was resuspended in 100 uL of RNase-free water.

Each RNA was then purified under denaturing conditions by urea polyacrylamide gel electrophoresis (PAGE). First a 1.5 mm thick 10% polyacrylamide gel was prepared fresh and pre-run in 1X TBE for 30 minutes at 30 Watts. The RNA sample was mixed with equal volume 2X RNA loading solution (95% formamide, 20 mM EDTA, 0.05% xylene cyanol and 0.05% bromophenol blue), heated at 90°C for 2 minutes, and cooled on ice for 5 minutes. The RNA sample was then loaded onto the pre-run gel. The gel was run at room temperature for 2 hours at 30 Watts. The lights were turned off to minimize light exposure to the fluorescein tag.

The RNA was visualized by UV shadowing. The purified RNA band was excised from the gel with a clean razor, cut into pieces, and placed in a 1.5 mL tube. To extract the RNA from the gel pieces, 600 uL of RNA elution buffer (0.5 M sodium acetate pH 5.2, 0.1 mM EDTA pH 8.0, 0.1% SDS) was added to the tube containing the gel pieces. The tube was wrapped in aluminum foil, placed in an Eppendorf 5432 mixer, and shaken overnight at 4°C.

The RNA in solution was purified by three 600 uL chloroform extractions followed by precipitation in 1200 uL pure ethanol at -80°C for at least 1 hour. The RNA precipitate was recovered by centrifugation at 4°C for 30 minutes at 16000 x g. The supernatant was carefully removed and the RNA pellet was washed with 500 uL of -20°C 70% ethanol. The RNA pellet was centrifuged at

4°C for 2 minutes at 16000 x g. The ethanol supernatant was carefully removed and the RNA pellet was dried for 5 minutes in a SpeedVac without heat. The RNA pellet was finally dissolved in 20-40 uL of RNase-free water.

The RNA's concentration was measured by its absorbance at 260 nm using a spectrophotometer. The RNA was stored in several aliquots at -80°C.

The purity and quality of the purified RNA was assessed by the same urea PAGE method used to purify that RNA, except with 0.75 mm thickness.

4.3 PAGE analysis of labeled RNAs

Denaturing urea PAGE analysis of the UG_4U RNA was performed as previously described[44]. A 0.75 mm thick 16% polyacrylamide gel containing 6.9 M urea (MP Biomedicals) and 1X TBE was made using 40% (w/v) acrylamide:bisacrylamide (19:1) solution (Omnipur, Calbiochem). The gel was allowed to polymerize for at least one hour before use. The gel was pre-run at 20 Volts per centimeter (equals 600 V) at room temperature. 20 pmol UG_4U was reconstituted in 50 mM Tris pH 7.8 and either 100 mM KCl, 100 mM NaCl or 100 mM LiCl. The UG_4U samples were heated at 37°C for 90 minutes, then incubated at 25°C for 5 minutes in a thermal cycler. Then the 2X RNA loading dye was added to each RNA sample to a final 8 uL volume. The samples were heated at 60°C for 5 minutes and then incubated at 25°C for 5 minutes in a thermal cycler. Then the 8 uL samples were loaded onto the pre-run gel and run

for 4 hours at 600 V at room temperature in darkness. The gel was promptly imaged using a ChemiDoc XRS$^+$ (Bio-Rad).

For denaturing urea PAGE analysis of the 18-mer model RNAs, a 1 mm thick 15% polyacrylamide gel containing 6.9 M urea and 1X TBE was made using 40% (w/v) acrylamide:bisacrylamide (19:1) solution (Omnipur, Calbiochem). The gel was allowed to polymerize for at least one hour before use. The gel was pre-run at 25 Watts in 1X TBE running buffer for 45 minutes at room temperature. 10 uL samples containing 10 pmol Poly$G_{18,}$ 3 pmol PolyC_{18}, 2.2 pmol PolyC_{18}(ACUU), 2.2 pmol PolyC_{18}(UGGA), 6 pmol CR1, 4 pmol PolyA_{18}, 6 pmol $(GACG)_4$, or 10 pmol $(UGGA)_4$ in 20 mM Tris-HCl pH 7.45 and 100 mM KCl were prepared. The RNA samples were heated in an 80°C water bath for 2 minutes, then annealed by slowly cooling to room temperature in the water bath. Then 2X RNA loading dye containing 95% formamide, 20 mM EDTA, 0.05% xylene cyanol and 0.05% bromophenol blue was added to each sample to a final 20 uL volume. The samples were heated at 60°C for 5 minutes and then incubated at room temperature for 5 minutes. The 20 uL samples were then loaded onto the pre-run gel and run for 1 hour and 45 minutes at 25 Watts at room temperature in darkness. The gel was promptly imaged using a Typhoon FLA 9500 (GE Amersham).

Denaturing urea PAGE analysis of the UG_4U RNA was performed as previously described[44]. A 0.75 mm thick 16% urea PAGE prepared similarly as for the 18-mer model RNAs; the gel was pre-run at 20 Volts per centimeter (equals 600 V) at room temperature. 20 pmol UG_4U was reconstituted in 50 mM

Tris pH 7.8 and either 100 mM KCl, 100 mM NaCl or 100 mM LiCl. The UG_4U samples were heated at 37°C for 90 minutes, then incubated at 25°C for 5 minutes in a thermal cycler. Then the 2X RNA loading dye was added to each RNA sample to a final 8 uL volume. The samples were heated at 60°C for 5 minutes and then incubated at 25°C for 5 minutes in a thermal cycler. Then the 8 uL samples were loaded onto the pre-run gel and run for 4 hours at 600 V at room temperature in darkness. The gel was promptly imaged using a ChemiDoc XRS^+ (Bio-Rad).

For native PAGE analysis of all the model RNAs, a 1 mm thick 15% native polyacrylamide gel containing 1X TBE was made using the same 40% (w/v) acrylamide:bisacrylamide (19:1) solution used for urea PAGE. The gel was allowed to polymerize for at least one hour before use. The gel was pre-run at 11 Watts in 1X TBE running buffer for 45 minutes at 4°C. 10 uL samples containing 10 pmol $PolyG_{18,}$ 3 pmol $PolyC_{18}$, 2.2 pmol $PolyC_{18}$(ACUU), 2.2 pmol $PolyC_{18}$(UGGA), 6 pmol CR1, 4 pmol $PolyA_{18}$, 6 pmol $(GACG)_4$, or 10 pmol $(UGGA)_4$ in 20 mM Tris-HCl pH 7.45 and 100 mM KCl were prepared. The RNA samples were heated in an 80°C water bath for 2 minutes, then annealed by slowly cooling to room temperature in the water bath. The samples were then placed on ice for 5 minutes before adding ice cold 30% (v/v) RNase-free glycerol to each RNA sample to a final 10% (v/v) glycerol. The 15 uL samples were promptly loaded onto the pre-run gel and run for 4 hours at 11 Watts at 4°C in darkness. The gel was promptly imaged using a Typhoon FLA 9500.

4.4 FMRP-RNA fluorescence anisotropy binding assay and equilibrium dissociation constant (K_D) determination

The various FMRP constructs were titrated against 5 nM fluorescein-labeled RNA. Fluorescence anisotropy was measured on a non-binding 96 well black bottom plate (Greiner) using a Tecan Safire 2 plate reader.

NT-dFMRP was diluted into dFMRP binding buffer (25 mM Tris pH 7.65, 300 mM KCl, 5 mM $MgCl_2$, 1 mM DTT, and 100 ng/uL total tRNA from *E. coli*) and mixed with dFMRP binding buffer supplemented with fluorescein-labeled RNA within the plate wells in a final 200 uL volume. Human FMRP binding assays were performed similarly, except in hFMRP binding buffer (25 mM Tris pH 7.65, 75 mM KCl, 5 mM $MgCl_2$, 1 mM DTT, and 100 ng/uL total tRNA from *E. coli*). The plate was incubated at room temperature in darkness for 30 minutes. Samples were excited at 470 nm and emission was measured at 520 nm with a 20 nm bandwidth; optimal signal gain was determined per read.

Bovine serum albumin was titrated against the FMRP-binding RNAs to test their intrinsic protein binding capacities under identical binding conditions.

To quantify the binding affinity between FMRP and the binding RNA, anisotropy data from each binding assay was fit to the equation below. The fluorescein-labeled RNA ligand concentration $[L]_t$ was fixed at 5 nM with variable FMRP concentration $[R]_t$ to calculate the equilibrium dissociation constant (K_D).

$$r([R]_t) = r_0 + c\left[\frac{K_D + [L]_t + [R]_t}{2} - \sqrt{\frac{(K_D + [L]_t + [R]_t)^2}{4} - [L]_t + [R]_t}\right]$$

References

1. Coffee, B., Keith, K., Albizua, I., Malone, T., Mowrey, J., Sherman, S.L. & Warren, S.T. (2009) Incidence of fragile X syndrome by newborn screening for methylated FMR1 DNA. *American Journal of Human Genetics* 85(4): 503-514.

2. Ouyang, L., Grosse, S., Raspa, M. & Bailey, D. (2010) Employment impact and financial burden for families of children with fragile X syndrome: findings from the National Fragile X Survey. *Journal of Intellectual Disability Research* 54(10): 918-928.

3. Garber, K.B., Visootak, J. & Warren, S.T. (2008) Fragile X syndrome. *European Journal of Human Genetics* 16: 666-672.

4. Maes, B., Fryns, J.P., Ghesquiere, P. & Borghgraef, M. (2000) Phenotypic checklist to screen for fragile X syndrome in people with mental retardation. *Mental Retardation* 38(3): 207-215.

5. Santoro, M.R., Bray, S.M. & Warren, S.T. (2012) Molecular Mechanisms of Fragile X Syndrome: A Twenty-Year Perspective. *Annu. Rev. Pathol. Mech. Dis.* 7: 219-245.

6. Irwin, S.A., Patel, B., Idupulapati, M., Harris, J.B., Crisostomo, R.A, Larsen, B.P., Kooy, F., Willems, P.J., Cras, P., Kozlowski, P.B., Swain, R.A., Weiler, I.J & Greenough, W.T. (2001) Abnormal dendritic spine characteristics in the temporal and visual cortices of patients with fragile-X syndrome: a quantitative examination. *American Journal of Medical Genetics* 98(2): 161-167.

7. Verkerk, A.J., Pieretti, M., Sutcliffe, J.S., Fu, Y.H., Kuhl, D.P, Pizzuti, A., Reiner, O., Richards, S., Victoria, M.F., Zhang, F., Eussen, B.E., Van Ommen, G.B., Blonden, L.A., Riggins, G.J., Chastain, J.L., Kunst, C.B., Galjaard, H., Caskey, C.T., Nelson, D.L., Oostra, B.A. & Warren, S.T. (1991) Identification of a gene (FMR-1) containing a CGG repeat coincident with a breakpoint cluster region exhibiting length variation in fragile X syndrome. *Cell* 65(5): 905-914.

8. Fu, Y.H., Kuhl, D.P., Pizzuti, A., Pieretti, M., Sutcliffe, J.S., Richards, S., Verkerk, A.J., Holden, J.J., Fenwick, R.G. Jr., Warren, S.T., Oostra, B.A., Nelson, D.L. & Caskey, C.T. (1991) Variation of the CGG repeat at the fragile X site results in genetic instability: resolution of the Sherman paradox. *Cell* 67(6): 1047-1058.

9. Colak, D., Zaninovic, N., Cohen, M.S., Rosenwaks, Z., Yang, W.Y., Gerhardt, J., Disney, M.D & Jaffrey, S.R. (2014) Promoter-bound trinucleotide repeat mRNA drives epigenetic silencing in fragile X syndrome. *Science* 343(6174): 1002-1005.

10. 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. & Willems, P.J. (1993) A point mutation in the FMR-1 gene associated with fragile X mental retardation. *Nat. Genet.* 3(1):31-35.

11. Valverde, R., Pozdnyakova, I., Kajander, T., Venkatraman, J. & Regan, L. (2007) Fragile X mental retardation syndrome: structure of the KH1-KH2 domains of fragile X mental retardation protein. *Structure* 15(9):1090-1098.

12. Greco, C.M., Berman, R.F., Martin, R.M., Tassone, F., Schwartz, P.H., Chang, A., Trapp, B.D., Iwahashi, C., Brunberg, J., Hessl, D., Becker, E.J., Papazian, J., Leehey, M.A., Hagerman, R.J. & Hagerman, P.J. (2006) Neuropathology of fragile X-associated tremor/ataxia syndrome (FXTAS). *Brain* 129(1): 243-255.

13. Sherman, S.L. (2000) Premature ovarian failure among fragile X permutation carriers: parent-of-origin effect? *American Journal of Human Genetics* 67(1): 11-13.

14. Kenneson, A., Zhang, F., Hagedorn, C.H. & Warren, S.T. (2001) Reduced FMRP and increased FMR1 transcription is proportionally associated with CGG repeat number in intermediate-length and premutation carriers. *Hum. Mol. Gen.* 10(14):1449-1454.

15. Brouwer, J.R., Mientjes, E.J., Bakker, C.E., Nieuwenhuizen, I.M., Severijnen, L.A. & Van der Linde, H. (2007) Elevated Fmr1 mRNA levels and reduced protein expression in a mouse model with an unmethylated Fragile X full mutation. *Experimental cell research* 313(2):244-253.

16. Ashley, C.T. Jr., Wilkinson, K.D., Reines, D. & Warren, S.T. (1993) FMR1 protein: conserved RNP family domains and selective RNA binding. *Science* 262(5133): 563-566.

17. Siomi, H., Matunis, M.J., Michael, W.M. & Dreyfuss, G. (1993) The pre-mRNA binding K protein contains a novel evolutionarily conserved motif. *Nucleic Acids Res.* 21(5): 1193-1198.

18. Hollingworth, D., Candel, A.M., Nicastro, G., Martin, S.R., Briata, P., Gherzi, R. & Ramos, A. (2012) KH domains with impaired nucleic acid

binding as a tool for functional analysis. *Nucleic Acids Res.* 40(14): 6873-6886.

19. Feng, Y., Absher, D., Eberhart, D.E., Brown, V., Malter, H.E. & Warren, S.T. (1997) FMRP associates with polyribosomes as an mRNA, and the I304N mutation of severe fragile X syndrome abolishes this association. *Mol. Cell* 1(1): 109-118.

20. Zang, J.B., Nosyreva, E.D., Spencer, C.M., Volk, L.J., Musunuru, K., Zhong, R., Stone, E.F., Yuva-Paylor, L.A., Huber, K.M., Paylor, R., Darnell, J.C. & Darnell, R.B. (2009) A mouse model of the human Fragile X syndrome I304N Mutation. *PLoS Genet.* 5(12): e1000758.

21. Myrick, L.K., Nakamoto-Kinoshita, M., Lindor, N.M., Kirmani, S., Cheng, X. & Warren, S.T. (2014) Fragile X syndrome due to a missense mutation. *European Journal of Human Genetics* 1-5.

22. Thandapani, P., O'Connor, T.R., Bailey, T.L. & Richard, S. (2013) Defining the RGG/RG motif. *Mol. Cell* 50(5): 613-623.

23. 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. & Patel, D.J. (2011) Structure-function studies of FMRP RGG peptide recognition of an RNA duplex-quadruplex junction. *Nat. Struct. Mol. Biol.* 18(7): 796-804.

24. Vasilyev, N., Polonskaia, A., Darnell, J.C., Darnell, R.B., Patel, D.J. & Serganov, A. (2015) Crystal structure reveals specific recognition of a G-quadruplex RNA by a β-turn in the RGG motif of FMRP. *Proc. Natl. Acad. Sci. U.S.A.* 112: 5391-5400.

25. Adinolfi, S., Ramos, A., Martin, S.R., Dal Piaz, F., Pucci, P., Bardoni, B., Mandel, J.L. & Pastore, A. (2003) The N-terminus of the fragile X mental retardation protein contains a novel domain involved in dimerization and RNA binding. *Biochemistry* 42(35): 10437-10444.

26. Ramos, A., Hollingworth, D., Adinolfi, S., Castets, M., Kelly, G., Frenkiel, T.A., Bardoni, B. & Pastore, A. (2006) The structure of the N-terminal domain of the fragile X mental retardation protein: a platform for protein-protein interaction. *Structure* 14(1): 21-31.

27. Maurer-Stroh, S., Dickens, N.J., Hughes-Davies, L., Kouzarides, T., Eisenhaber, F. & Ponting, C.P. (2003) The Tudor domain 'Royal Family': Tudor, plant Agenet, Chromo, PWWP and MBT domains. *Trends in Biochemical Sciences* 28(2): 69-74.

28. Alpatov, R., Lesch, B.J., Nakamoto-Kinoshita, M., Blanco, A., Chen, S., Stutzer, A., Armache, K.J., Simon, M.D., Xu, C., Ali, M., Murn, J., Prisic, S., Kutateladze, T.G., Vakoc, C.R., Min, J., Kingston, R.E., Fischle, W., Warren, S.T., Page, D.C. & Shi, Y. (2014) A Chromatin-Dependent Role of the Fragile X Mental Retardation Protein FMRP in the DNA Damage Response. *Cell* 157(4): 869-881.

29. Hu, Y., Chen, Z., Fu, Y., He, Q., Jiang, L., Zheng, J., Gao, Y., Mei, P., Chen, Z. & Ren, X. (2015) The amino-terminal structure of human fragile X mental retardation protein obtained using precipitant-immobilized imprinted polymers. *Nat. Commun.* 6: 6634.

30. Eberhart, D.E., Malter, H.E., Feng, Y. & Warren, S.T. (1996) The fragile X mental retardation protein is a ribonucleoprotein containing both nuclear localization and nuclear export signals. *Hum. Mol. Gen.* 5(8): 1083-1091.

31. Collins, S.C., Bray, S.M., Suhl, J.A., Cutler, D.J., Coffee, B., Zwick, M.E. & Warren, S.T. (2010) Identification of novel FMR1 variants by massively parallel sequencing in developmentally delayed males. *American Journal of Medical Genetics* 152A(10): 2512-2520.

32. Zhenzhong, L., Zhang, Y., Wilkinson, K.D. & Warren, S.T. (2001) The fragile X mental retardation protein inhibits translation via interacting with mRNA. *Nucleic Acids Res.* 29: 2276-2283.

33. Brown, V., Jin, P., Darnell, J.C., O'Donnell, W.T., Tenenbaum, S.A., Jin, X., Feng, Y., Wilkinson, K.D., Keene, J.D., Darnell, R.B. & Warren, S.T. (2001) Microarray Identification of FMRP-Associated brain mRNAs and Altered mRNA Translational Profiles in Fragile X Syndrome. *Cell* 107: 477-487.

34. Darnell, J.C., Van Driesche, S.J., Zhang, C., Hung, K.Y., Mele, A., Fraser, C.E., Stone, E.F., Chen, C., Fak, J.J., Chi, S.W., Licatalosi, D.D., Richter, J.D. & Darnell, R.B. (2011) FMRP stalls ribosomal translocation on mRNAs linked to synaptic function and autism. *Cell* 146: 247-261.

35. Chen, E., Sharma, M.R., Shi, X., Agrawal, R.K. & Joseph, S. (2014) Fragile X Mental Retardation Protein Regulates Translation by Binding Directly to the Ribosome. *Mol. Cell* 54(3): 407-417.

36. Ascano, M., Mukherjee, N., Bandaru, P., Miller, J.B., Nusbaum, J.D., Corcoran, D.L., Langlois, C., Munschauer, M., Dewell, S., Hafner, M., Williams, Z., Ohler, U. & Tuschl, T. (2012) FMRP targets distinct mRNA sequence elements to regulate protein expression. *Nature* 492: 382-386.

37. McMahon, A.C., Rahman, R., Jin, H., Shen, J.L., Fieldsend, A., Luo, W. & Rosbash, M. (2016) TRIBE: Hijacking an RNA-Editing Enzyme to Identify Cell-Specific Targets of RNA-Binding Proteins. *Cell* 165: 1-12.

38. Ray, D., Kazan, H., Cook, K.B., Weirauch, M.T., Najafabadi, H.S., Li, X., Gueroussov, S., Albu, M., Zheng, H., Yang, A., Na, H., Irimia, M., Matzat, L.H., Dale, R.K., Smith, S.A., Yarosh, C.A., Kelly, S.M., Nabet, B., Mecenas, D., Li, W., Laishram, R.S., Qiao, M., Lipshitz, H.D., Piano, F., Corbett, A.H., Carstens, R.P., Frey, B.J., Anderson, R.A., Lynch, K.W., Penalva, L.O., Lei, E.P., Fraser, A.G., Blencowe, B.J., Morris, Q.D. & Hughes, T.R. (2013) A compendium of RNA-binding motifs for decoding gene regulation. *Nature* 499: 172-177.

39. Tabet, R., Moutin, E., Becker, J.A., Heintz, D., Fouillen, L., Flatter, E., Krezel, W., Alunni, V., Koebel, P., Dembele, D., Tassone, F., Bardoni, B., Mandel, J.L., Vitale, N., Muller, D., Merrer, J.L. & Moine, H. (2016) Fragile X Mental Retardation Protein (FMRP) controls diacylglycerol kinase activity in neurons. *Proc. Natl. Acad. Sci. U.S.A.* 113: 3619-3628.

40. Suhl, J.A., Chopra, P., Anderson, B.R., Bassell, G.J. & Warren, S.T. (2014) Analysis of FMRP mRNA target datasets reveals highly associated mRNAs mediated by G-quadruplex structures formed via clustered WGGA sequences. *Hum. Mol. Gen.* 23: 5479-5491.

41. Darnell, J.C., Jensen, K.B., Jin, P., Brown, V., Warrent S.T. & Darnell, R.B. (2001) Fragile X Mental Retardation Protein Targets G Quartet mRNAs Important for Neuronal Function. *Cell* 107: 489-499.

42. Ramos, A., Hollingworth, D. & Pastore, A. (2003) G-quartet-dependent recognition between the FMRP RGG box and RNA. *RNA* 9: 1198-1207.

43. Deng, J., Xiong, Y. & Sundaralingam, M. (2001) X-ray analysis of an RNA tetraplex $(UGGGGU)_4$ with divalent Sr^{2+} ions at subatomic resolution (0.61 Å). *Proc. Natl. Acad. Sci. U.S.A.* 98: 13665-13670.

44. Kim, J., Cheong, C. & Moore, P.B. (1991) Tetramerization of an RNA oligonucleotide containing a GGGG sequence. *Nature* 351: 331-332.

45. Abitbol, M., Menini, C., Delezoide, A.L., Rhyner, T., Vekemans, M. & Mallet, J. (1993) Nucleus basalis magnocellularis and hippocampus are the major sites of FMR-1 expression in the human fetal brain. *Nat. Genet.* 4(2): 147-153.

46. Devys, D., Lutz, Y., Rouyer, N., Bellocq, J.P. & Mandel, J.L. (1993) The FMR-1 protein is cytoplasmic, most abundant in neurons and appears normal in carriers of a fragile X premutation. *Nat. Genet.* 4(4): 335-340.

47. Tamanini, F., Willemsen, R., Van Unen, L., Bontekoe, C., Galjaard, H., Oostra, B.A. & Hoogeveen, A.T. (1997) Differential expression of FMR1, FXR1, and FXR2 proteins in human brain and testis. *Hum. Mol. Gen.* 6(8): 1315-1322.

48. Khandjian, E.W., Corbin, F., Woerly, S. & Rousseau, F. (1996) The fragile X mental retardation protein in associated with ribosomes. *Nat. Genet.* 12(1): 91-93.

49. Khandjian, E.W., Huot, M.E., Tremblay S., Davidovic, L., Mazroui, R. & Bardoni, B. (2004) Biochemical evidence for the association of fragile X mental retardation protein with brain polyribosomal ribonucleoparticles. *Proc. Natl. Acad. Sci. U.S.A.* 101(36):13357-13362.

50. Laggerbauer, B., Ostareck, D., Keidel, E.M., Ostareck-Lederer, A. & Fischer, U. (2001) Evidence that fragile X mental retardation protein is a negative regulator of translation. *Hum. Mol. Gen.* 10(4): 329- 338.

51. Li, Z., Zhang, Y., Ku, L., Wilkinson, K.D., Warren, S.T. & Feng, Y. (2001) The fragile X mental retardation protein inhibits translation via interacting with mRNA. *Nucleic Acids Res.* 29(11): 2276-2283.

52. Mazroui, R., Huot, M.E., Tremblay, S., Filion, C., Labelle, Y. & Khandjian, E.W. (2002) Trapping of messenger RNA by Fragile X mental retardation protein into cytoplasmic granules induces translation repression. *Hum. Mol. Gen.* 11(24): 3007-3017.

53. Napoli, I., Mercaldo, V., Boyl, P.P., Eleuteri, B., Zalfa, F., De Rubeis, S., Di Marino, D., Mohr, E., Massimi, M., Falconi, M., Witke, W., Costa-Mattioli, M., Sonenberg, N., Achsel, T. & Bagni, C. (2008) The fragile X syndrome protein represses activity- dependent translation through CYFIP1, a new 4E-BP. *Cell* 134(6): 1042-1054.

54. Stefani, G., Fraser, C.E., Darnell, J.C. & Darnell, R.B. (2004) Fragile X mental retardation protein is associated with translation polyribosomes in neuronal cells. *J. Neurosci.* 24(33): 7272-7276.

55. Ceman, S., O'Donnell, W.T., Reed, M., Patton, S., Pohl, J. & Warren, S.T. (2003) Phosphorylation influences the translation state of FMRP-associated polyribosomes. *Hum. Mol. Gen.* 12(24): 3295-3305.

56. Agrawal, R.K., Spahn, C.M., Penczek, P., Grassucci, R.A., Nierhaus, K.H. & Frank, J. (2000) Visualization of tRNA movements on the Escherichia coli 70S ribosome during the elongation cycle. *J. Cell Biol.* 150: 447–460.

57. Schaeffer, C., Bardoni, B., Mandel, J.L., Ehresmann, B., Ehresmann, C. & Moine, H. (2001) The fragile X mental retardation protein binds specifically to its mRNA via a purine quartet motif. *EMBO J*. 20: 4803–4813.

58. Caudy, A.A, Myers, M., Hannon, G.J. & Hammond, S.M. (2002) Fragile X-related protein and VIG associate with the RNA interference machinery. *Genes. Dev.* 16: 2491-2496.

59. Ishizuka, A., Siomi, M.C. & Siomi, H. (2002) A Drosophila fragile X protein interacts with components of RNAi and ribosomal proteins. *Genes. Dev.* 16: 2497-2508.

60. Plante, I., Davidovic, L., Ouellet, D.L., Gobeil, L.A., Tremblay, S., Khandjian, E.W. & Provost, P. (2006) Dicer-derived microRNAs are utilized by the fragile X mental retardation protein for assembly on target RNAs. *J. Biomed. Biotechnol.* 2006: 64347.

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. & Sheng, M. (2010) Regulation of synaptic structure and function by FMRP-associated microRNAs miR-125b and miR-132. *Neuron* 65: 373-384.

62. Muddashetty, R.S., Nalavadi, V.C., Gross, C., Yao, X., Xing, L., Laur, O., Warren, S.T. & Bassell, G.J. (2011) Reversible inhibition of PSD-95 mRNA translation by miR-125a, FMRP phosphorylation, and mGluR signaling. *Mol. Cell* 42(5): 673-688.

63. Stetler, A., Winograd, C., Sayegh, J., Cheever, A., Patton, E., Zhang, X., Clarke, S. & Ceman, S. (2006) Identification and characterization of the methyl arginines in the fragile X mental retardation protein Fmrp. *Hum. Mol. Gen.* 15(1): 87-96.

64. Blackwell, E., Zhang, X. & Ceman, S. (2010) Arginines of the RGG box regulate FMRP association with polyribosomes and mRNA. *Hum. Mol. Gen.* 19(7): 1314-1323.

65. Chen, E., & Joseph, S. (2015) Fragile X mental retardation protein: A paradigm for translational control by RNA-binding proteins. *Biochimie* 114: 147–154.

66. Huber, K.M., Kayser, M.S. & Bear, M.F. (2000) Role for rapid dendritic protein synthesis in hippocampal mGluR-dependent long-term depression. *Science* 288(5469): 1254-1257.

67. Weiler, I.J., Irwin, S.A., Klintsova, A.Y., Spencer, C.M., Brazelton, A.D., Miyashiro, K., Comery, T.A., Patel, B., Eberwine, J. & Greenough, W.T. (1997) Fragile X mental retardation protein is translated near synapses in response to neurotransmitter activation. *Proc. Natl. Acad. Sci. U.S.A.* 94(10): 5395-5400.

68. Snyder, E.M., Philpot, B.D., Huber, K.M., Dong, X., Fallon, J.R. & Bear, M.F. (2001) Internalization of ionotropic glutamate receptors in response to mGluR activation. *Nat. Neurosci.* 4(11): 1079-1085.

69. Sidorov, M.S., Auerbach, B.D. & Bear, M.F. (2013) Fragile X mental retardation protein and synaptic plasticity. *Molecular Brain* 6(15): 1-11.

70. Zhang, Y.Q., Bailey, A.M., Matthies, H.J., Renden, R.B., Smith, M.A., Speese, S.D., Rubin, G.M. & Broadie, K. (2001) *Drosophila* fragile X-related gene regulates the MAP1B homolog Futsch to control synaptic structure and function. *Cell* 107(5): 591-603.

71. Lee, A., Li, W., Bogert, B.A., Su, K. & Gao, F.B. (2003) Control of dendritic development by the *Drosophila* fragile X-related gene involves the small GTPase Rac1. *Development* 130(22): 5543-5552.

72. Reeve, S.P., Bassetto, L., Genova, G.K., Kleyner, Y., Leyssen, M., Jackson, F.R. & Hassan, B.A. (2005) The *Drosophila* fragile X mental retardation protein controls actin dynamics by directly regulating profilin in the brain. *Current Biology* 15(12): 1156-1163.

73. Wan, L., Dockendorff, T.C., Jongens, T.A. & Dreyfuss, G. (2000) Characterization of dFMR1, a *Drosophila melanogaster* Homolog of the Fragile X Mental Retardation Protein. *Mol. Cell. Biol.* 20(22): 8536-8547.

74. Morales, J., Hiesinger, P.R., Schroeder, A.J., Kume, K., Verstreken, P., Jackson, F.R., Nelson, D.L. & Hassan, B.A. (2002) Drosophila fragile X protein, DFXR, regulates neuronal morphology and function in the brain. *Neuron* 34(6): 961-972.

75. Tessier, C.R. & Broadie, K. (2008) Drosophila Fragile X Mental Retardation Protein Developmentally Regulates Activity-Dependent Axon Pruning. *Development* 135(8): 1547-1557.

76. Bakker, C.E., Verheij, C., Willemsen, R., Van der Helm, R., Oerlemans, F., Vermey, M., Bygrave, A., Hoogeveen, A.T. & Oostra, B.A. (1994)

Fmr1 knockout mice: a model to study fragile X mental retardation. *Cell* 78(1): 23-33.

77. Comery, T.A., Harris, J.B., Willems, P.J., Oostra, B.A., Irwin, S.A., Weiler, I.J. & Greenough, W.T. (1997) Abnormal dendritic spines in fragile X knockout mice: Maturation and pruning deficits. *Proc. Natl. Acad. Sci. U.S.A.* 94(10): 5401-5404.

78. Bontekoe, C.J., Bakker, C.E., Nieuwenhuizen, I.M., Van der Linde, H., Lans, H., De Lange, D., Hirst, M.C. & Oostra, B.A. (2001) Instability of a $(CGG)_{98}$ repeat in the Fmr1 promoter. *Hum. Mol. Gen.* 10(16): 1693-1699.

79. Mientjes, E.J., Nieuwenhuizen, I., Kirkpatrick, L., Zu, T., Hoogeveen-Westerveld, M., Severijnen, L., Rife, M., Willemsen, R., Nelson, D.L. & Oostra, B.A. (2006) The generation of a conditional Fmr1 knock out mouse to study FMRP function in vivo. *Neurobiology of Disease* 21(3): 549-555.

80. Kikin, O., D'Antonio, L. & Bagga, P.S. (2006) QGRS Mapper: a web-based server for predicting G-quadruplexes in nucleotide sequences. *Nucleic Acids Res.* 34(2): 676-682.

81. Stephenson, W., Asare-Okai, P.N., Chen, A.A., Keller, S., Santiago, R., Tenenbaum, S., Garcia, A.E., Fabris, D. & Li, P.T. (2013) The Essential Role of Stacking Adenines in a Two-Base-Pair RNA Kissing Complex. *J. Am. Chem. Soc.* 135: 5602−5611.

82. Guo, J.U. & Bartel, D.P. (2016) RNA G-quadruplexes are globally unfolded in eukaryotic cells and depleted in bacteria. *Science* 353(6306).

83. Zhang, Y., O'Connor, J.P., Siomi, M.C., Srinivasan, S., Dutra, A., Nussbaum, R.L. & Dreyfuss, G. (1995) The Fragile X Mental Retardation Syndrome protein interacts with novel homologs FXR1 and FXR2. *EMBO J.* 14(21): 5358-5366.

84. Siomi, M.C., Zhang, Y., Siomi, H. & Dreyfuss, G. (1996) Specific Sequences in the Fragile X Syndrome Protein FMR1 and the FXR Proteins Mediate Their Binding to 60S Ribosomal Subunits and the Interactions among Them. *Mol. Cell. Biol.*16(7): 3825-3832.