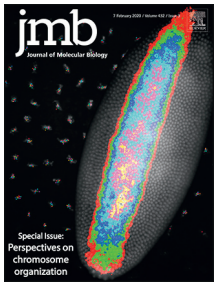

RNA-Binding Specificity of the Human Fragile X Mental Retardation Protein

Youssi M. Athar and Simpson Joseph

Department of Chemistry and Biochemistry, *University of California, San Diego, La Jolla, CA92093-0314, USA*

Correspondence to Simpson Joseph: *sjoseph@ucsd.edu*
https://doi.org/10.1016/j.jmb.2020.04.021
Edited by Ruben L. Gonzalez

Abstract

Fragile X syndrome is the most common form of inherited intellectual disability and is caused by a deficiency of the fragile X mental retardation protein (FMRP) in neurons. FMRP regulates the translation of numerous mRNAs within dendritic synapses, but how FMRP recognizes these target mRNAs remains unknown. FMRP has KH0, KH1, KH2, and RGG domains, which are thought to bind to specific RNA recognition elements (RREs). Several studies used high-throughput methods to identify various RREs in mRNAs that FMRP may bind to *in vivo*. However, there is little overlap in the mRNA targets identified by each study, suggesting that the RNA-binding specificity of FMRP is still unknown. To determine the specificity of FMRP for the RREs, we performed quantitative *in vitro*RNA binding studies with various constructs of human FMRP. Unexpectedly, our studies show that the KH domains do not bind to the previously identified RREs. To further investigate the RNA-binding specificity of FMRP, we developed a new method called *M*otif *Id*entification by *A*nalysis of *S*imple sequences (MIDAS) to identify single-stranded RNA sequences bound by KH domains. We find that the FMRP KH0, KH1, and KH2 domains bind weakly to the single-stranded RNA sequences suggesting that they may have evolved to bind more complex RNA structures. Additionally, we find that the RGG motif of human FMRP binds with a high affinity to an RNAG-quadruplex structure that lacks single-stranded loops, double-stranded stems, or junctions.

© 2020 Elsevier Ltd. All rights reserved.

Introduction

Fragile X syndrome (FXS) is the most common form of inherited intellectual disability. FXS is primarily caused by the deficiency of the fragile X mental retardation protein (FMRP) in neuronal cells [1]. The expansion of CGG trinucleotide repeats in the 5′-untranslated region of the *FMR1* gene leads to gene silencing and FMRP deficiency [2–5]. FMRP has been proposed to regulate local protein synthesis within dendritic synapses [6–8]. FMRP deficiency results in the formation of dendrites with long, thin, and immature spines possibly caused by the deregulation of translation of specific neuronal mRNAs [9–11]. However, how FMRP binds to specific neuronal mRNAs to regulate their translation is unknown.

FMRP contains several conserved domains. Among the conserved domains, four are putative RNA-binding domains. Three K-homology (KH) domains, KH0, KH1, and KH2, and an arginine–glycine–glycine (RGG) motif are hypothesized to mediate FMRP binding to mRNAs [12,13]. The KH domain was first identified as a nucleic acid-binding domain in heterogeneous nuclear ribonucleoprotein K and has since been found in many eukaryotic (type I) and prokaryotic (type II) proteins [14,15]. A GXXG loop in the KH1 and KH2 domains of FMRP is conserved in many RNA-bindingKH domains, such as the KH-type splicing regulatory protein (KSRP) and neuro-oncological ventral antigen (Nova-1 and 2) proteins, further suggesting that FMRP KH domains play a role in binding-specific RNAs [16,17]. A third KH domain was recently discovered upstream of the KH1 domain, termed KH0 [18,19]. Unlike KH1 and KH2, KH0 does not contain the GXXG loop, which suggests KH0 may not bind RNA or bind a different RNA motif. Mutations in the KH1 and KH2 domains of FMRP also result in FXS, indicating that the RNA-binding function of FMRP is essential for normal health [20,21]. The KH2 domain

© 2020 Elsevier Ltd. All rights reserved.

Please cite this article as: Y. M. Athar and S. Joseph, RNA-Binding Specificity of the Human Fragile X Mental Retardation Protein, Journal of Molecular Biology, https://doi.org/10.1016/j.jmb.2020.04.021

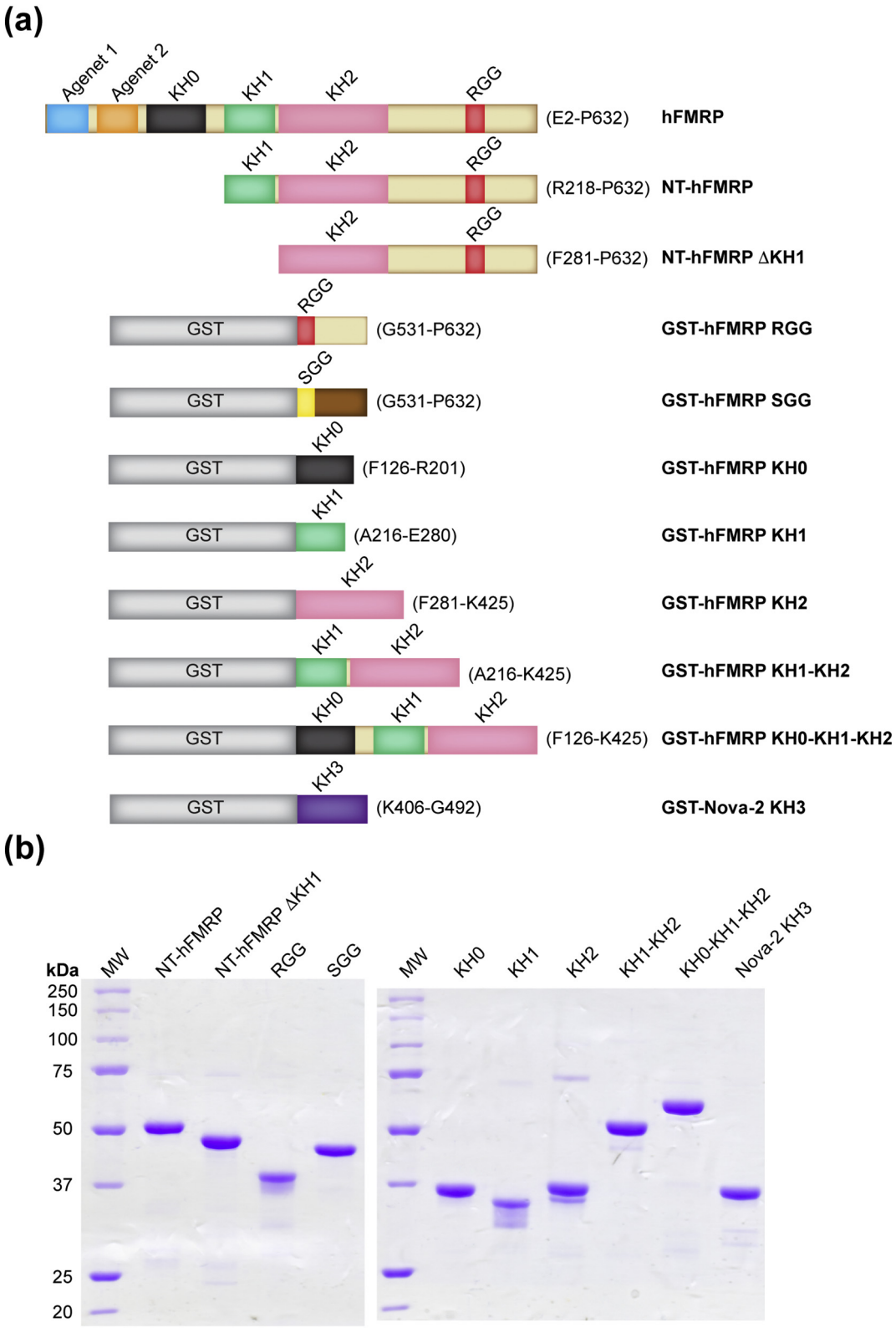

Figure 1 (*legend on next page*)

Please cite this article as: Y. M. Athar and S. Joseph, RNA-Binding Specificity of the Human Fragile X Mental Retardation Protein, Journal of Molecular Biology, https://doi.org/10.1016/j.jmb.2020.04.021

was shown to bind to an in vitro-selected kissing complex RNA called KC2 [22]. Additionally, the RGG motif binds to an *in vitro*-selectedG-quadruplex (GQ) forming RNA called Sc1 [23]. An NMR structure of the human FMRP RGG peptide bound to Sc1 RNA GQ reveals how the disordered RGG motif is stabilized upon binding to the GQ structure [24]. A subsequent crystal structure reveals in atomic detail the interaction of each amino acid within the RGG motif with the GQ nucleotides of Sc1 [25]. Interestingly, the RGG peptide interacts not with the GQ region, but binds to the junction between the GQ anda double-stranded stem in SC1. While Sc1 was selected as a high-affinity binding partner for theRGG motif, it suggested that the RGG motif mediates FMRP binding to GQ structures in mRNAs. Indeed, the FMRP RGG peptide has been shown to bind to various GQ structures found in mRNAs with a range of binding affinities [26–30]. Given the ubiquity of potential GQ structures throughout the transcriptome, the presence of the RGG motif within FMRP hints at FMRP's potential to bind a vast pool of target mRNAs *in vivo*.

The identification of FMRP target mRNAs is crucial for understanding the molecular function of FMRP and has been the focus of numerous studies. Most endeavors to identify FMRP target mRNA motifs have employed high-throughput methods such as coimmunoprecipitation of mRNA-FMRP complexes (Co-IP), cross-linking and immunoprecipitation combined with high-throughput sequencing (HITS-CLIP), photoactivatable ribonucleoside-enhancedcross-linking and immunoprecipitation (PAR-CLIP), RNAcompete, and targets of RNA-binding proteins identified by editing (TRIBE) [31–35]. Unfortunately, the proposed catalog of RNA sequence and structural motifs, or RNA-recognition elements (RREs), overlap poorly between each study [36]. Analysis of the proposed RREs yielded two promising RREs: clustered WGGA (W = A/U) sequences and the GACR (R = A/G) sequence [36]. The WGGA clusters were originally proposed to bind specifically to the KH1 domain of FMRP [31]. A subsequent study proposed that WGGA clusters could form GQ structures and therefore bind the RGG motif of FMRP instead [36]. The GACR sequences were enriched in RNAcompete for an FMRP fragment composed solely of the KH1 and KH2 RNA binding domains, suggesting that the GACR motif binds a KH domain [35].

Here, we use fluorescence anisotropy to quantify binding affinities between FMRP and RNAs of a defined length and sequence. We modeled the RNAs after proposed FMRP RREs [31,35,36]. Unexpectedly, our studies show that the KH0, KH1, and KH2 domains do not bind to the WGGA, ACUK, and GACR motifs. Previous studies showed that many KH domains recognize only four to five nucleotides in single-stranded RNAs (ssRNAs) [16,17,37–39]. This inspired us to develop a new method to identify RNA sequences that bind to KH domains, which we named *Motif Identification by Analysis of Simple sequences* (MIDAS). We synthesized a pool of RNAs having five randomized nucleotides flanked by fixed sequences of eight nucleotides at the 5′ and 3′ ends. We used the randomized RNA pool to identify RNA molecules that bind to the KH domains of FMRP (the pool will contain 4^5 = 1024 unique sequences). Surprisingly, the FMRP KH domains bind weakly to the selected RNA sequences suggesting that the KH domains ofFMRP may not interact with simple RREs as the canonical KH domains. Instead, the FMRP KH domains may have evolved to recognize more complex RNA motifs. Additionally, consistent with previous studies, we find that the RGG motif of human FMRP binds to RNA GQ structures with a 100-fold range in affinities showing that FMRP can discriminate between the different GQ structures [26–30,40,41].

Results

Isolating the RNA-binding domains of FMRP

To elucidate how FMRP specifically binds its target RNAs, we isolated specific regions of FMRP and quantified their binding affinity to a catalog of RNAs modeled after reported FMRP RREs. We sought to both test the validity of the proposed RREs within RNAs and clarify the roles of the KH0, KH1, KH2, and RGG domains in binding the RREs. Weexpressed and purified nine human FMRP (hFMRP) variants from *Escherichia coli* for our studies (Figure1). The N-terminal truncated human NT-FMRP construct was cloned by deleting the amino acids immediately upstream of the KH1 domain from the full-length sequences (NT-hFMRP

Figure 1. Design and purification of human FMRP constructs. (a) hFMRP is the full-length human FMRP isoform 1, spanning E2-P632 as indicated. The different truncation constructs of hFMRP that were made are illustrated. GST-hFMRP SGG is a fusion between the GST and the RGG motif-containing sequence of hFMRP spanning G531–P632 where the 16 arginines spanning the RGG motif to the C terminus were mutated to serines, illustrated using a different color for the C terminus. (b) SDS-polyacrylamide gel showing the purified proteins illustrated above. Lanes are molecular weight standards (MW), NT-hFMRP (47 kDa), NT-hFMRP ΔKH1 (40 kDa), GST-hFMRP RGG (40 kDa), GST-hFMRP SGG (39 kDa), GST-hFMRP KH0 (38 kDa), GST-hFMRP KH1 (36 kDa), GST-hFMRP KH2 (45 kDa), GST-hFMRP KH1–KH2 (52 kDa), GST-hFMRP KH0–KH1–KH2 (63 kDa), and GST-Nova-2 KH3 (38 kDa).

Please cite this article as: Y. M. Athar and S. Joseph, RNA-Binding Specificity of the Human Fragile X Mental Retardation Protein, Journal of Molecular Biology, https://doi.org/10.1016/j.jmb.2020.04.021

encodes Arg218 to Pro632). This truncation eliminates the FMRP N-terminal domain, which contains the tandem Agenet 1 and Agenet 2 domains as well as the KH0 domain; the N-terminal domain is also responsible for FMRP dimerization [18]. The NT-hFMRP construct conserves the three proposed RNA-binding domains [42–44]. We also generated a NT-hFMRP△KH1 construct encoding Phe281 to Pro632 to determine the binding specificity of the KH2 domain. Additionally, we made individual KH0, KH1, KH2, KH1–KH2, and KH0–KH1–KH2 fused downstream of the glutathione S-transferase (GST) tag. The fusion proteins are large enough to generate clear shifts in fluorescence anisotropy should the KH motifs bind the fluorescein-labeled RNAs.

To evaluate the importance of the RGG motif in binding RNA GQ structures, we generated an hFMRP construct encoding Gly531 of the RGG domain to the C-terminal Pro632 and fused it downstream of the GST tag to generate a GST-hFMRPRGG fusion protein. Finally, we mutated all of the arginines to serines in the hRGG-encoding region, generating a GST-hFMRPSGG fusion protein. Together these FMRP constructs should delineate the contributions of the KH domains and the RGG motif toward binding to RNAs.

Design of model RNAs

We designed a set of RNAs modeled after three promising FMRP RREs, GQ structure, GACR clusters, and WGGA clusters, to determine if and how FMRP specifically binds these sequences (Table 1). Apart from the UG_4U 6-merRNA, we standardized our model RNAs to 18 nucleotides in length to control for length-dependent binding affinities; M1 served as an 8-merRNA size marker for UG_4U. We used $PolyG_{17}$ and UG_4U as two different model GQ RNAs to quantify and compare FMRP binding affinities for different RNA GQ structures [45–47]. UG_4U is noteworthy because the X-ray crystal structure showed that it forms a highly stable, four stacked G-tetrad structure [46], and we wanted to test whether the RGG domain could bind to a GQ structure that does not contain single-stranded regions, junctions, or double-stranded stems.

We generated $(GACG)_4$ and $(UGGA)_4$18-mer RNAs containing four tandem repeats of GACG and UGGA to test FMRP binding to the GACR cluster and WGGA cluster, respectively. The GACR array was reported to bind the KH1 and KH2 domains of FMRP [35]. The WGGA array was originally reported to bind the KH1 domain but was later hypothesized to assemble an intramolecular GQ structure that may bind the FMRP RGG motif [31,36]. We also tested FMRP binding to ACUU and UGGA RNAs to determine if a single ACUK (K = G/U) or WGGA motif is sufficient for binding to FMRP, respectively [31]. Finally, we used two 18-mer RNAs with different base compositions that cannot assemble any secondary structures as control RNAs, denoted $PolyC_{17}$, and CR1. All RNAs had a 3′-terminal uridine to prevent quenching of the fluorescein dye attached to the 3′ end [48]. The fluorescein dye was used to visualize the RNAs on gels and for performing fluorescence anisotropy binding studies.

Table 1. Nomenclature and sequences of the RNAs

RNA	Sequence (5′ to 3′)
$PolyG_{17}$	GGGGGGGGGGGGGGGGGU
UG_4U	UGGGGU
$(UGGA)_4$	UUGGAUGGAUGGAUGGAU
$(GACG)_4$	UGACGGACGGACGGACGU
UGGm^{6}ACU	UGGACUUGGm^{6}ACUUGGACU
$PolyC_{17}$	CCCCCCCCCCCCCCCCCU
ACUU	CCCCCCCACUUCCCCCCU
UGGA	CCCCCCCUGGACCCCCCU
CR1	GCUAUCCAGAUUCUGAUU

Folding of model RNAs

We first tested whether our model RNAs assemble into folded structures by denaturing and native polyacrylamide gel electrophoresis (PAGE) analyses (Figure 2). We tested whether UG_4U assembles the reported 24-mer intermolecular GQ composed of four stacked G-tetrads using two methods: (1) a denaturing urea PAGE in the presence of KCl, NaCl or LiCl as described previously [49], and (2) a native PAGE comparing UG_4U migration to the 8-mer M1, 18-mer $PolyC_{17,}$ and 18-mer CR1 RNAs, which remain unfolded. The denaturing PAGE showed UG_4U migrating as two different species, with the presence of the slow-migrating species being K^+-dependent (Figure 2(a)). The UG_4U GQ assembly was remarkably stable as it endured urea denaturing conditions even with Li^+. The native PAGE showed UG_4U migrating much slower than the 8-nucleotide M1 RNA and slightly slower than the 18-mer RNAs (Figure 2(c)). Together, both PAGE analyses indicate that UG_4U indeed assembles the tetrameric GQ structure. All 18-nucleotide RNAs migrated as expected on a denaturing PAGE (Figure 2(b)). However, on a native PAGE, $PolyG_{17}$ migrated much slower than the control 18-nucleotide RNAs as a heterogeneous array (Figure 2(c)). This suggests the $PolyG_{17}$ array may comprise a mix of 36-nucleotide intermolecular GQ structures in different registers. This is consistent with a recent report that showed that $PolyG_{40}$ forms stable GQ structures [47].

The $(GACG)_4$ and $(UGGA)_4$ RNAs migrated differently from the other RNAs. The $(GACG)_4$RNA migrated as two species, one migrating slowly at a similar rate as $PolyG_{17}$ and the other migrating faster

Please cite this article as: Y. M. Athar and S. Joseph, RNA-Binding Specificity of the Human Fragile X Mental Retardation Protein, Journal of Molecular Biology, https://doi.org/10.1016/j.jmb.2020.04.021

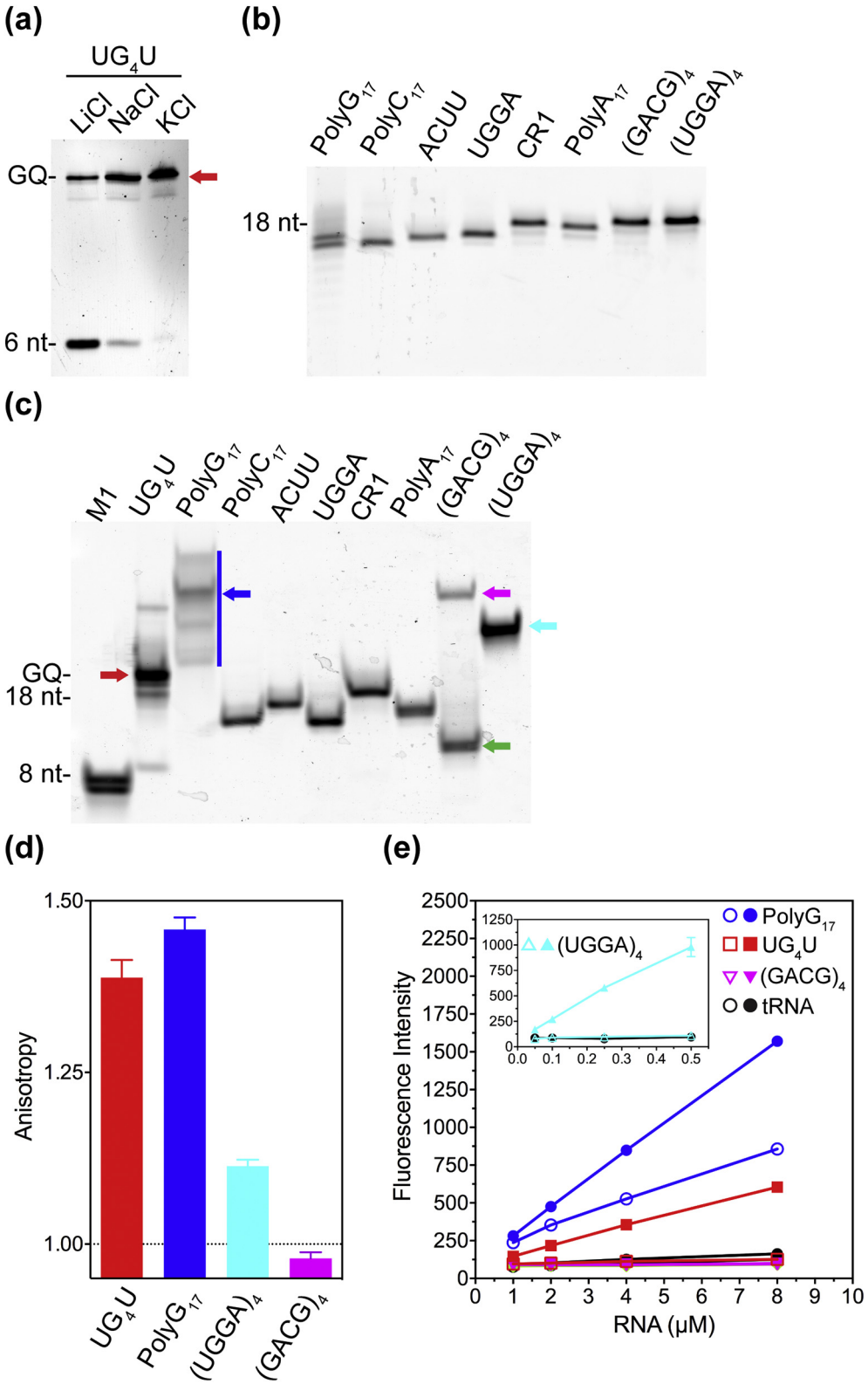

Figure 2 (*legend on next page*)

Please cite this article as: Y. M. Athar and S. Joseph, RNA-Binding Specificity of the Human Fragile X Mental Retardation Protein, Journal of Molecular Biology, https://doi.org/10.1016/j.jmb.2020.04.021

than the unstructured 18-nucleotide control RNAs (Figure 2(c)). Unlike $PolyG_{17}$, $(GACG)_4$ does not contain four repeats of at least two consecutive guanines required to assemble an intramolecular GQ structure [50,51]. The slower $(GACG)_4$ species may be a stable dimer structure formed by intermolecular base pairs (Supplemental Figure 1). The faster $(GACG)_4$ species could then be the monomeric 18-mer stem–loop structure, which would be expected to migrate faster than the unfolded control RNAs.

The $(UGGA)_4$RNA is predicted to assemble an intramolecular GQ structure composed of two stacked G-tetrads [52,53]. It is indeed migrating at a similar rate as one of the $PolyG_{17}$ species within the $PolyG_{17}$ array (Figure 2(c)). To verify whether $(UGGA)_4$ is forming a GQ structure, we used the BG4 antibody that binds specifically to GQ structures [54]. The binding of the BG4 antibody to the fluorescein-labeled RNAs was monitored by fluorescence anisotropy, which should increase because the RNA-BG4 antibody complex will tumble slower than the free RNAs. We observed a substantial increase in anisotropy with UG_4U and $PolyG_{17}$, which is consistent with the fact that these RNAs form GQ structures (Figure 2(d)). The $(UGGA)_4$RNA showed a modest increase in anisotropy, suggesting that the BG4 antibody binds poorly to the GQ structure formed by $(UGGA)_4$. Alternatively, the change in anisotropy is small because the fluorescein attached to the RNA is dynamic in the $(UGGA)_4$-RNA-BG4 antibody complex and does not correlate with the tumbling rate of the complex. No change in anisotropy was observed with $(GACG)_4$, indicating that this sequence does not form a GQ structure.

To further validate the formation of GQ structures, we used the *N*-methyl mesoporphyrin IX (NMM) dye, which shows enhanced fluorescence when it intercalates into GQ structures [55,56]. We examined the presence of GQ structures in $PolyG_{17}$, UG_4U, $(UGGA)_4$, $(GACG)_4$, and tRNA reconstituted in buffer containing either KCl or LiCl with the NMM dye. The tRNA served as a negative control. Consistent with the native gel analysis and the BG4 antibody assay, we observed increased fluorescence intensity with $PolyG_{17}$, UG_4U, and $(UGGA)_4$ but not with $(GACG)_4$ and tRNA (Figure 2(e)). The increase in the fluorescence intensity was higher when the RNAs were in KCl than in LiCl, demonstrating that $PolyG_{17}$, UG_4U, and $(UGGA)_4$ form GQ structures.

FMRP binds to different GQ RNAs

Little is known regarding how FMRP specifically binds its target mRNAs. A previous study suggested that the KH1 and KH2 domains of *Drosophila*FMRP bind to the ribosome, while the RGG motif is available to bind the target mRNA [57]. Previous reports also showed that the FMRP RGG motif binds to the *in vitro* selected Sc1 RNA GQ structure and to a variety of GQ structures from potential mRNA targets of FMRP [23,25–30,40,41,58]. However, it is unknown whether FMRP can bind with a high affinity to a GQ structure that does not have any single-stranded regions, double-stranded stems, or junctions. We employed fluorescence anisotropy to determine the equilibrium dissociation constants (K_D) of several human FMRP variants binding to RNAs that assemble different GQ structures or contain RREs that have been reported to bind FMRP (Table 1).

We observed that NT-hFMRP binds to $PolyG_{17}$ with a K_D of 14 ± 2 nM and UG_4U with a K_D of 110 ± 16 nM (Figure 3(a), Table 2). Additionally, NT-hFMRP binds to $(UGGA)_4$ with a K_D of 710 ± 50 nM. Since all three RNAs form GQ structures, our results show that NT-hFMRP binds to the different GQ structures with a 100-fold range in binding affinities. Interestingly, NT-hFMRP failed to bind to the $(GACG)_4$RNA, which was modeled after the GACR array reported to bind to the KH domains of FMRP (Figure 3(a)). The four tandem GACG repeats may be insufficient for binding to FMRP. The $(GACG)_4$RNA could assemble a mixed population of GC stem–loop monomers and double-stranded dimers. FMRP's inability to bind either the GC stem–loop monomer or dimer would suggest FMRP does not bind double-strandedRNA assemblies even if they are G-rich. NT-hFMRP did not bind to RNAs having a single UGGA or ACUU, which cannot form GQ structures (Figure 3(a)). Finally, despite lacking the KH1 domain, NT-hFMRPΔKH1 showed a binding pattern similar to NT-hFMRP (Figure 3(b), Table 2).

Figure 2. Analysis of the structures formed by the model RNAs. (a) Urea PAGE of UG_4U denatured and then refolded in the presence of 100 mM LiCl, 100 mM NaCl, or 100 mM KCl. (b) Urea PAGE of the denatured 18-mer model RNAs. (c) Native PAGE of the model RNAs. GQ species are marked with red arrows and a blue bar; $(GACG)_4$ hairpin monomer and dimer species are marked with green and magenta arrows, respectively; $(UGGA)_4$ structure is marked with a cyan arrow. (d) Fold increase in fluorescence anisotropy of 5 nM fluorescein-labeled RNAs upon the addition of 430 nM anti-GQ BG4 scFV antibody. All RNAs are labeled at 3′ end with fluorescein. The standard deviations from two experiments are shown. (e) Increase in fluorescence intensity of NMM with increasing concentrations of unlabeled RNA. The inset shows the data for the $(UGGA)_4$ RNA performed at lower concentrations of RNA. The closed and open symbols show RNAs in KCl and LiCl, respectively. The standard deviations from three experiments are shown.

Please cite this article as: Y. M. Athar and S. Joseph, RNA-Binding Specificity of the Human Fragile X Mental Retardation Protein, Journal of Molecular Biology, https://doi.org/10.1016/j.jmb.2020.04.021

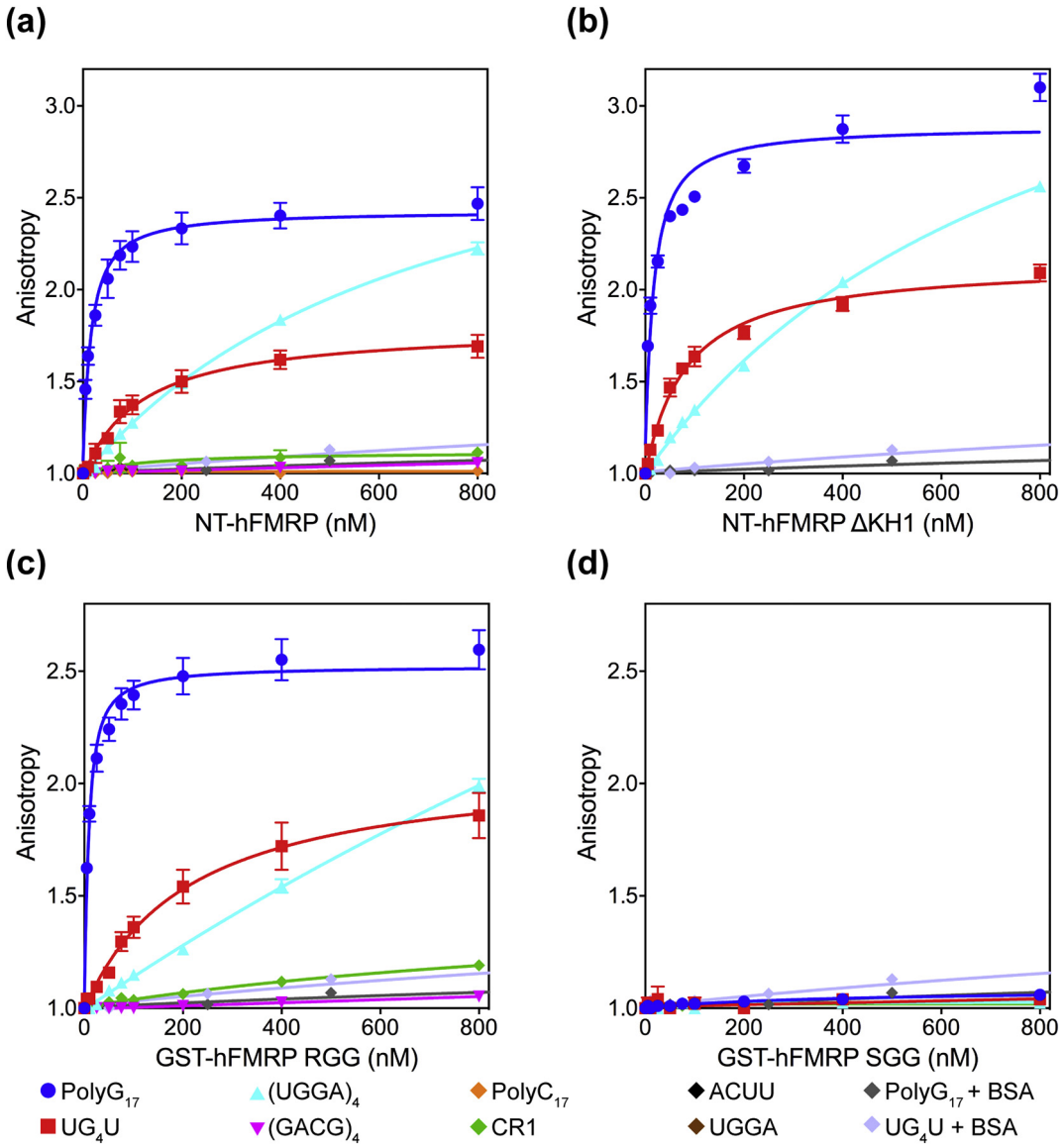

Figure 3. FMRP binding to model RNAs. Fold increase in fluorescence anisotropy of 5 nM fluorescein-labeled RNAs upon the addition of 0–800 nM of (a) NT-hFMRP, (b) NT-hFMRP ΔKH1, (c) GST-hFMRP RGG, and (d) GST-hFMRP SGG. Bovine serum albumin (BSA) binding to the GQ-assembling RNAs was also measured to test for non-specific protein binding by the two RNAs. Binding trials were performed at least three times, with error bars depicting standard deviation. The symbols indicate the identity of the different RNAs.

FMRP RGG motif specifically binds to RNA GQ structures

To determine whether the RGG motif is responsible for binding to the RNA GQ structures, we tested the binding of our model RNAs to GST-hFMRP RGG. Consistent with previous studies [23,25–30,40,41,58], the RGG motif was sufficient for binding $PolyG_{17}$ and UG_4U GQ structures (Figure 3(c), Table 2). In fact, GST-hFMRPRGG bound $PolyG_{17}$ with a K_D of 8.6 ± 1.2 nM and UG_4U with a K_D of 210 ± 18 nM, similar affinities as NT-hFMRP. The nanomolar affinity of NT-hFMRP for UG_4U shows that the RGG domain can bind to GQ structures that do not contain single-stranded loops, double-stranded stems, or junctions. Interestingly, GST-hFMRPRGG bound $(UGGA)_4$ with a K_D of >4000 nM, a significantly lower affinity than NT-hFMRP. This suggests that the region from the KH2 to the RGG domain may be necessary for binding to $(UGGA)_4$. To confirm that the RGG motif is required for binding to GQ RNAs, we made a GST-hFMRPSGG construct with all 16 arginines in the RGG motif and C-terminal domain changed to serines. The GST-hFMRPSGG protein did not bind

Please cite this article as: Y. M. Athar and S. Joseph, RNA-Binding Specificity of the Human Fragile X Mental Retardation Protein, Journal of Molecular Biology, https://doi.org/10.1016/j.jmb.2020.04.021

Table 2. Equilibrium dissociation constant, K_D (nM) of FMRP binding to RNAs

RNA	NT-hFMRP	NT-hFMRP ΔKH1	RGG	SGG
UG_4U	110 ± 16	80 ± 7	210 ± 18	No binding
$PolyG_{17}$	14 ± 2	17 ± 3	8.6 ± 1.2	No binding
$(UGGA)_4$	710 ± 50	840 ± 39	>4000	No binding
$(GACG)_4$	No binding	Not tested	No binding	Not tested
$UGGm^6ACU$	1250 ± 610	Not tested	Not tested	Not tested

to $PolyG_{17}$ and UG_4U GQ structures, demonstrating that the RGG motif is essential for binding to GQ RNAs (Figure 3(d)).

Finally, we tested whether the FMRP RGG motif is binding to the GQ assemblies of $PolyG_{17}$ and UG_4U, or unfolded forms of the RNAs. To assess specific binding to GQ assemblies, we compared the binding efficiency of GST-hFMRPRGG to $PolyG_{17}$ and UG_4U in the presence of 75 mM KCl *versus* 75 mM LiCl. If FMRP could bind unfolded forms of $PolyG_{17}$ and UG_4U, then we would observe increased binding efficiency in LiCl, where both RNAs favor unfolded states. Conversely, if FMRP specifically binds the folded form, then we would observe decreased binding efficiency due to a drop in the folded population of RNA. Indeed, substituting Li^+ in place of K^+ decreased the net anisotropy change for $PolyG_{17}$, reflecting a smaller population of the RNA binding to FMRP (Supplemental Figure 2a). Additionally, GST-hFMRPRGG bound to $PolyG_{17}$ with a 50-fold lower affinity in the presence of Li^+ (K_D = 480 ± 62 nM), consistent with a recent report [40]. In the case of UG_4U, we observed only a slight decrease in anisotropy with Li^+ (Supplemental Figure 2b). This may be because Li^+ does not significantly destabilize the UG_4U GQ assembly. While we see some destabilization by Li^+ in our urea-PAGE, the urea may be catalyzing it. In both RNA cases, binding was not completely depleted in Li^+. This suggests the cation substitution did not destabilize the protein. Instead, the protein was tightly binding to the remaining population of stable $PolyG_{17}$ and UG_4U GQ assemblies. The concomitant drop in the extent of the anisotropy shift with the drop in the GQ population supports the conclusion that the RGG motif specifically binds to RNA GQ structures.

The KH domains of FMRP do not bind to the WGGA and GACR motifs

Our studies showed that NT-hFMRP binds with low affinity to $(UGGA)_4$RNA and does not bind to $(GACG)_4$ RNA. Furthermore, we showed that $(UGGA)_4$ forms a GQ structure suggesting that the RGG motif mediates the interaction of NT-hFMRP with $(UGGA)_4$. To confirm that the KH domains of FMRP do not bind to $(UGGA)_4$ and $(GACG)_4$ RNAs, we performed binding studies with the KH domains of FMRP fused to the GST tag (Figure 1). GST-hFMRP KH0, GST-hFMRP KH1, GST-hFMRP KH2, GST-hFMRP KH1–KH2, and GST-hFMRP KH0–KH1–KH2 did not bind to $(UGGA)_4$ and $(GACG)_4$ RNAs, demonstrating that the KH domains of FMRP do not bind to the WGGA and GACR motifs (Supplemental Figure 3).

FMRP does not bind with high affinity to m^6A RNA

Recent studies indicated that FMRP preferentially binds to mRNAs with the N^6-methyladenosine (m^6A) modification [59–61]. We, therefore, tested the binding of NT-hFMRP to the $UGGm^6ACU$ model RNA having a single m^6A modification in the middle of the sequence. UGGACU is the consensus sequence for m^6A modification in mRNAs and is sufficient for binding to YTH-domain proteins [62–64]. Surprisingly, our results show that NT-hFMRP binds poorly to the m^6A-modifiedRNA suggesting that at best, FMRP can only form a low-affinity complex with m^6A-containing mRNAs (Supplemental Figure 4).

The KH domains of FMRP do not bind to short ssRNA motifs

The above results indicated that the KH domains of FMRP do not bind to the previously identified RREs. Biochemical and structural studies have shown that KH domains generally recognize 4 to 5 nucleotides in ssRNAs [16,17,37–39]. For example, the KH3 domain of Nova-2 protein binds to the UCACC sequence [37,38]. Structural studies showed that the UCACC binds to a cleft formed by the invariant and variable loops of Nova-2 KH3. Similarly, the KH3 and KH4 domains of ZBP1 bind to ACAC and CGGAC sequences, respectively [39]. Thus, KH domains typically recognize short, ssRNA sequences.

The realization that KH domains recognize 4 to 5 nucleotides motivated us to develop a method called MIDAS to determine the RNA-binding specificity of the FMRP KH domains (Figure 4(a)). We synthesized a pool of RNA molecules having 5 nucleotides randomized in the middle and flanked by 8 nucleotides of fixed sequences in the 5′ and 3′ ends (the pool will have 4^5 = 1024 unique sequences). The sequence of the fixed nucleotides was explicitly designed to avoid any of the known KH domain binding RREs. The RNA pool was labeled at the 5′-end with ^{32}P and incubated with 10 μM protein having one or more KH domain. The RNA•protein complex was separated from free RNA using a native gel (Figure 4(b)). The band corresponding to the RNA•protein complex was excised, the RNA

Please cite this article as: Y. M. Athar and S. Joseph, RNA-Binding Specificity of the Human Fragile X Mental Retardation Protein, Journal of Molecular Biology, https://doi.org/10.1016/j.jmb.2020.04.021

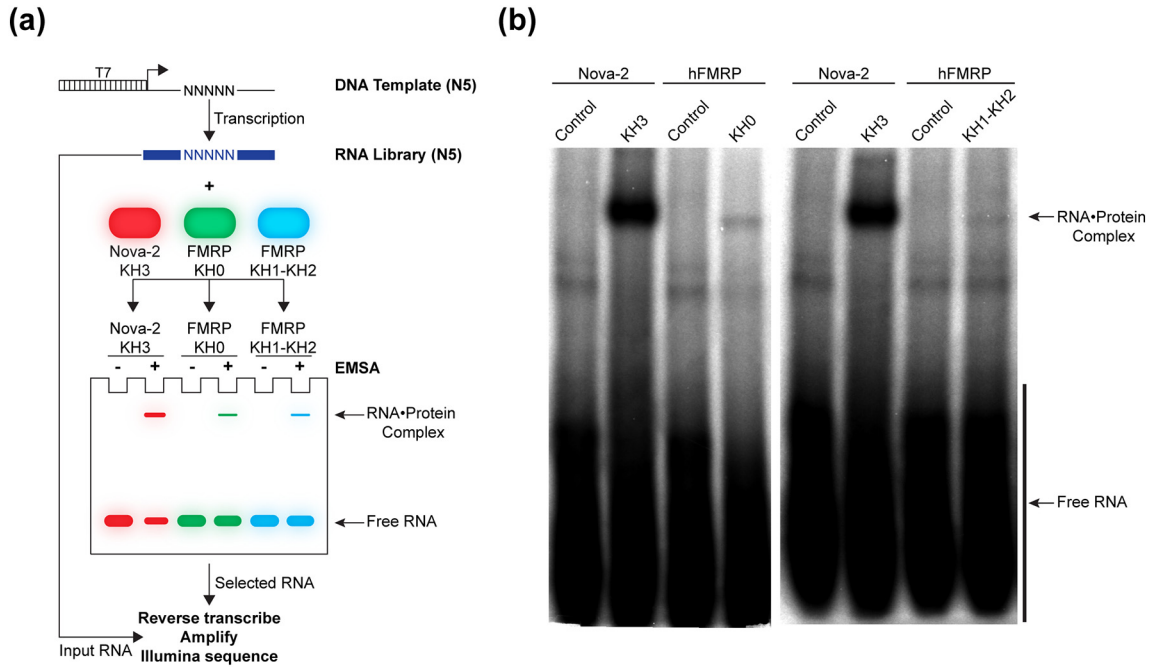

Figure 4. Illustration of MIDAS method. (a) Cartoon diagram of the MIDAS process. Purified GST-Nova-2 KH3, GST-hFMRP KH0, and GST-hFMRP KH1–KH2 are individually incubated with the RNA library containing a central N5 sequence. The RNAs bound by each protein are separated by an EMSA and purified from the gel. The input RNA sequences and the selected RNA sequences are then reverse transcribed, amplified, and Illumina sequenced. (b) Autoradiograph of EMSA. Contrast-enhanced image showing the free RNA library migrates as an array (black bar) and the RNAs bound by Nova-2 KH 3, FMRP KH0, and FMRP KH1–KH2 migrate similar distances as distinct bands. Control lanes are the MIDAS N5 RNA library incubated without protein.

purified, and the sequences of the purified RNA molecules were identified by next-generation sequencing.

To validate the MIDAS method, we used the well-studiedNova-2 KH3 domain [37,38]. As shown in Figure 4(b), Nova-2 KH3 binds to some of the RNA pool molecules resulting in the slower migrating RNA•protein complex band in the native gel. Surprisingly, we observed very faint RNA•protein complex bands with both FMRP KH0 and FMRP KH1–KH2, suggesting very weak binding to the RNA pool molecules. Additionally, we observed no binding to the RNA pool with the individual FMRP KH1 and FMRP KH2 domains (data not shown). We analyzed the RNA sequences that bind to Nova-2 KH3. Interestingly, the highly enriched sequences are UUUCA, UUUAA, UUCAC, UUCAU, and UCAAA (Table 3 and Supplemental Table1). The previously studied UCACC sequence was ranked much lower than the highly enriched sequences [37,38]. To further validate the MIDAS method, we synthesized RNA molecules corresponding to the highly enriched sequences and performed equilibrium binding studies (Figure 5(a)). We find excellent agreement between the enrichment factor and the K_D for Nova-2 KH3 binding to the selected RNA molecules, which shows that the MIDAS method is

Table 3. MIDAS enrichment factor (EF) and binding affinity for RNAs

	RNA	EF[a]	K_D (nM)
Nova-2 KH3	UUUCA	17.3	179 ± 23
	UUCAC	14.7	739 ± 82
	UUUAA	17.5	1038 ± 120
	UCACC	2.8	2991 ± 450
	CCACC	1.1	6991 ± 1848
FMRP KH0	UCACC	0.9	>10,000[b]
	CCACC	0.6	>10,000[b]
	GACUC	0.2	>10,000[b]
FMRP KH1–KH2	UUCAC	0.2	>10,000[b]
	UUUCA	0.1	>10,000[b]
	UCACC	0.5	>10,000[b]
	CCACC	0.3	>10,000[b]
	GACUC	0.3	>10,000[b]
	UUCAC	0.3	>10,000[b]
	UUUCA	0.1	>10,000[b]

[a] Enrichment factors were normalized to the relative abundance of selected RNAs from the EMSA gel.
[b] Estimated affinity based on EMSA selection using 10 μM FMRP KH domains.

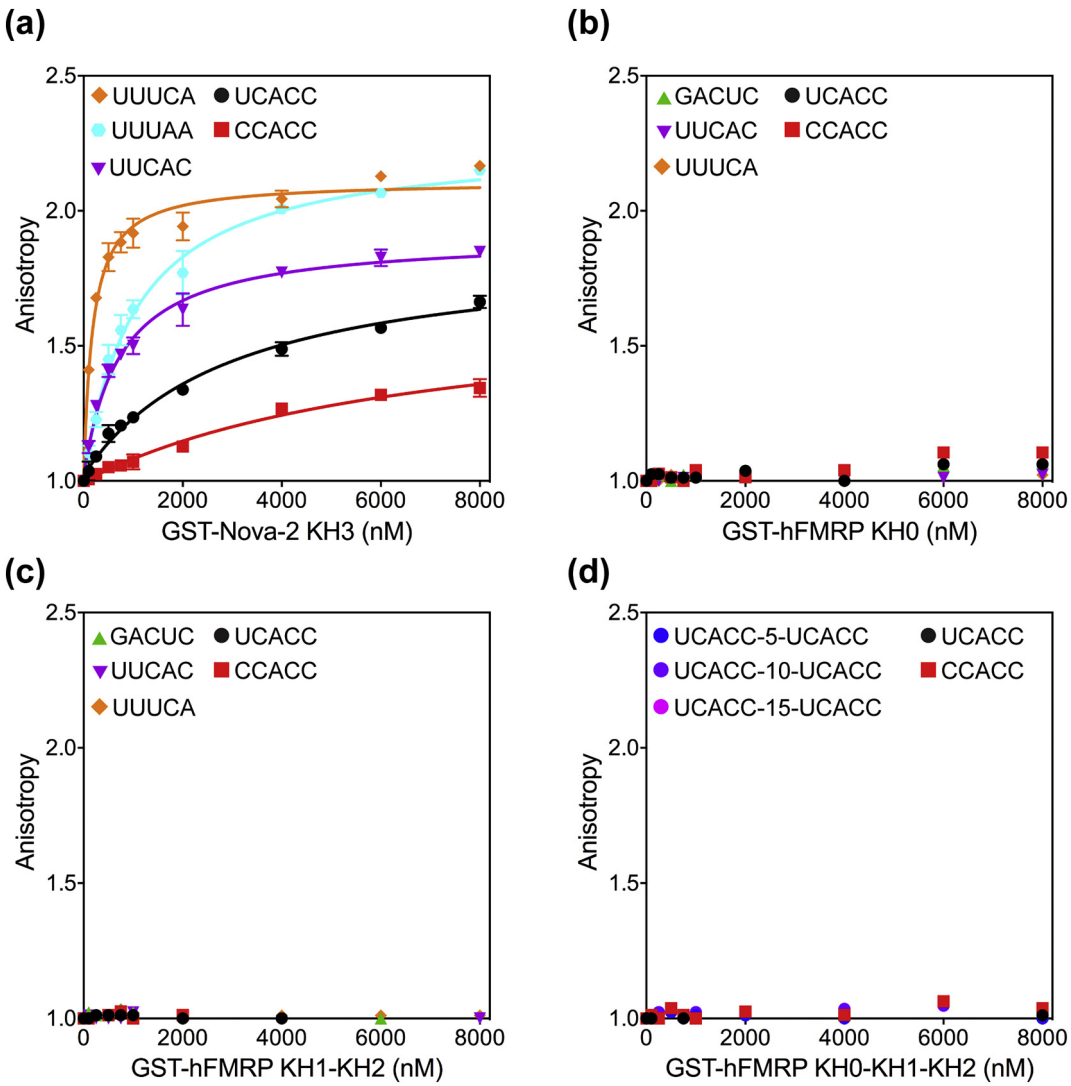

Figure 5. KH domains binding to RNAs identified by MIDAS. Fold increase in fluorescence anisotropy of 5 nM fluorescein-labeled RNAs upon the addition of 0–8000 nM of (a) GST-Nova-2 KH3, (b) GST-hFMRP KH0, (c) GST-hFMRP KH1–KH2, and (d) GST-hFMRP KH1–KH2-KH3. Binding trials were performed at least three times, with error bars depicting standard deviation. The symbols indicate the identity of the different RNAs tested for each protein construct.

capable of identifying authentic single-stranded RREs (Table 3).

Analysis of the RNA molecules that bind to the FMRP KH0 and FMRP KH1–KH2 domains indicated that the enriched sequences are UCACC and CCACC. However, the overall enrichment factors for FMRP KH0 and FMRP KH1–KH2 were significantly lower compared to Nova-2 KH3 (Table 3). Binding studies showed that FMRP KH0 and KH1–KH2 do not bind to UCACC at the standard concentration range used in our anisotropy binding assay (Figure 5(b) and (c)). It was not feasible to test higher concentrations of proteins with this assay. Therefore, we used an electrophoretic mobility shift assay (EMSA) to estimate the binding affinity of FMRP KH0 and KH1–KH2 for UCACC. The EMSA indeed showed that hFMRP KH0 and KH1–KH2 bind to UCACC, but the affinity is weaker than that of Nova-2 KH3 binding to UCACC (K_D = 3 μM) (Supplemental Figure 5). This suggests that the MIDAS method can identify even low-affinityRNA ligands.

To test whether the weak binding may be the result of having a single RRE, we synthesized RNAs having two UCACC sequences separated by 5, 10, or 15 nucleotides. The distance between the UCACC sequence was varied because we do not know what the optimal spacing is for the multiple KH domains in FMRP to bind to the RNA. Additionally,

Please cite this article as: Y. M. Athar and S. Joseph, RNA-Binding Specificity of the Human Fragile X Mental Retardation Protein, Journal of Molecular Biology, https://doi.org/10.1016/j.jmb.2020.04.021

ARTICLE IN PRESS

we used the GST-hFMRP KH0–KH1–KH2 construct to evaluate potentially synergistic binding when all three KH domains of hFMRP are present. Binding studies showed that hFMRP KH0–KH1–KH2 does not bind to the RNAs having the two UCACC (Figure5(d)). Our results indicate that the KH domains of FMRP do not bind to simple RREs, as proposed by previous studies. Instead, the KH domains of FMRP have likely evolved to bind to a more complex RNA or RNA–protein structure that is not represented in our MIDAS RNA pool. Indeed, FMRP is known to interact with the ribosome, and the KH domains may bind to a unique structure in the ribosome [33,43,57,65–70].

Discussion

FMRP has been estimated to regulate the translation of nearly 4% of the mRNAs in the human fetal brain [12]. However, identifying the RRE of FMRP in these mRNAs has remained challenging. Early RNA*in vitro* selection studies showed that the KH2 motif binds to kissing complex RNAs, whereas the RGG motif binds to a GQ-formingRNA [22,23]. In addition, high-throughput approaches have been used to determine the RREs of FMRP. For example, the WGGA and the ACUK motifs were shown to be enriched in FMRP mRNA targets by the PAR-CLIP method [31]. The WGGA and ACUK motifs were proposed to bind to the KH1 and KH2 domains of FMRP, respectively. In contrast to the above study, the GACR motif was identified using the KH1 and KH2 domains of FMRP by the *in vitro* technique called RNAcompete [35]. Further in-depth analysis of the FMRP target datasets indicated that the GACR motif is highly enriched, the WGGA motif is modestly enriched, and the ACUK motif is not enriched [36]. Moreover, the WGGA motifs are clustered together in the mRNA targets suggesting that they are enriched in the FMRP dataset because they form GQ structures [36].

In the PAR-CLIP method, 4-thiouridine is incorporated into nascent mRNA transcripts to improve the efficiency of cross-linking proteins that bind to mRNAs. However, as noted previously [36], this may bias the selection of mRNA sequences having more uridine residues, which may explain the enrichment of the ACUK motif in the FMRP target dataset. Consistent with this interpretation, our studies show that FMRP does not bind to an RNA having a single ACUU sequence. Furthermore, the PAR-CLIP method used full-lengthFMRP to identify mRNA targets, which may bias the selection of mRNAs having GQ structures because of the presence of the RGG domain. Consistent with this explanation, our results show that an RNA with four repeats of the WGGA motif forms a GQ structure and binds to the RGG domain but not to the KH domains of hFMRP.

Surprisingly, our studies show that FMRP does notbind to $(GACG)_4$, which was identified by the RNAcompete method and appears to be highly enriched in the FMRP target datasets [35,36]. The discrepancy between our data and the previous RNAcompete data may be explained by the size of the FMRP protein or the length of the RNA used in both studies. We used NT-hFMRP and several deletion constructs to determine K_D. In contrast, the RNAcompete method used a KH1–KH2 fragment, which may not have the same structure and the RNA-binding specificity as the larger protein fragment. Alternatively, we used 18-nucleotide RNA for our binding studies, whereas the RNAcompete used 30- to 41-nucleotide RNAs, which may fold into more complex structures that could bind to the KH1–KH2 fragment.

Interestingly, we show that FMRP has the highest binding affinity for $PolyG_{17}$ and UG_4U RNAs, whichform stable GQ structures. Furthermore, the RGGdomain of FMRP is sufficient for binding to $PolyG_{17}$ and UG_4U. These results are consistent with previous studies that showed the RGG domain of FMRP binds specifically to GQ structures [23–30,40,58]. Taken together, our results show that the KH0, KH1, and KH2 domains do not bind to the WGGA, ACUK, and GACR motifs, whereas the RGG domain binds with high affinity to RNA GQ structures. We developed the MIDAS method to identify in an unbiased manner single-stranded RREs that could bind to the KH0, KH1, and KH2 domains of hFMRP. Although the MIDAS method was able to identify RREs that bind with high affinity to the Nova-2 KH3 domain, we were only able to identify RREs that bind with weak affinity to the KH0 and KH1–KH2 domains of FMRP. The inability to identify any RREs that bind to the KH domains of hFMRP using the MIDAS method is consistent with our data showing that they do not bind to model RNAs having the previously identified RREs. One possibility for our unexpected results is that the phosphorylation of hFMRP at Ser500 is essential for the KH domains to bind to ssRNAs with high affinity [71,72]. Although Ser500 is not located within the KH domains, FMRP phosphorylation may induce some structural rearrangement in the protein that increases the binding affinity of the KH domains for ssRNAs [71].

A second possibility is that the KH domains of hFMRP recognize more complex RNA motifs than represented in our randomized RNA pool. Indeed, Darnell and co-workers showed that an *in vitro* selection experiment performed with an RNA pool having 25 randomized nucleotides did not yield any RNA binders for the hFMRP KH1–KH2 domains [22]. They obtained RNA binders only when they used an RNA pool with 52 randomized nucleotides. The *in vitro* selected RNA binders formed a kissing complex RNA motif, which binds to the KH2 domain

Please cite this article as: Y. M. Athar and S. Joseph, RNA-Binding Specificity of the Human Fragile X Mental Retardation Protein, Journal of Molecular Biology, https://doi.org/10.1016/j.jmb.2020.04.021

ARTICLE IN PRESS

of FMRP, most likely in a different mode than simple RREs binding to the canonical KH domains. These results suggest that that the KH domains of FMRP are more specialized than the canonical KH domains that recognize 4–5 nucleotides in ssRNA. The evolutionary changes in the sequence of the KH domains of FMRP may inhibit their binding to simple ssRNA sequences so that FMRP may bind to a unique structure such as the ribosome [33,43,57,65–70].

Materials and Methods

Protein expression and purification

The human FMR1 isoform 1 gene was assembled from *E. coli*codon-optimized gene blocks (IDT). The gene encoding the full-length isoform 1 human FMRP (hFMRP) spanning residues E2-P632 was subcloned into the LIC expression vector pMCSG7 (DNASU plasmid repository) with the 5′ TEV cleavage site deleted, conferring a N-terminal hexahistidine tag. The N-terminal truncated human FMRP (NT-hFMRP) spanning residues R218-P632 was subcloned into the LIC expression vector pMCSG7, conferring a N-terminal hexahistidine tag with the 5′ TEV cleavage site deleted. The NT-hFMRP ΔKH1 construct spanning residues F281–P362 was subsequently generated from NT-hFMRP via PCR deletion of the KH1 domain. The human FMRP RGG domain (GST-hFMRP RGG) spanning residues G531–P632 was subcloned into the LIC expression vector pMCSG10 (DNASU), which confers an N-terminalhexahistidine-GST fusion tag that is cleavable using TEV protease. The mutant SGG domain (GST-hFMRP SGG) was assembled from an *E. coli*codon-optimized gene block where all 16-arginine residues of the RGG region were mutated to serines. The resulting SGG fragment was subcloned into the LIC expression vector pMCSG10. The human Nova-2 KH3 domain (GST-Nova-2 KH3) spanning residues K406–G492 was subcloned into the LIC expression vector pMCSG10. The human FMRP KH0 domain (GST-hFMRP KH0) spanning residues F126–R201, human FMRP KH1 domain (GST-hFMRP KH1) spanning residues A216–E280, human FMRP KH2 domain (GST-hFMRP KH2) spanning residues F281–K425, human FMRP tandem KH1–KH2 domains (GST-hFMRP KH1–KH2) spanning residues A216–K425, and human FMRP tandem KH0–KH1–KH2 domains (GST-hFMRP KH0–KH1–KH2) spanning residues F126–K425 were also subcloned into the LIC expression vector pMCSG10.

The NT-hFMRP, NT-hFMRP ΔKH1, GST-hFMRPRGG,GST-hFMRPSGG,GST-Nova-2 KH3, and the five GST-hFMRPKH construct expression plasmids were all transformed in *E. coli* BL21(DE3) cells (Novagen). NT-hFMRP and NT-hFMRP ΔKH1 were purified using Ni-NTA affinity chromatography (Qiagen). GST-hFMRPRGG,GST-hFMRPSGG, and the five GST-hFMRPKH constructs were purified using glutathione affinity chromatography (GE Healthcare). All proteins were further purified using Superdex 75 16/60 or Superdex 200 16/60 (GE Healthcare) gel filtration chromatography and stored in their respective running buffers. Gel filtration buffer composition varied between FMRP constructs. NT-hFMRP and NT-hFMRP ΔKH1 gelfiltration running buffer contained 25 mM Tris (pH 7.4), 150 mM KCl, and 1 mM DTT. GST-hFMRPRGG and GST-hFMRPSGG gel filtration running buffer contained 50 mM Tris (pH 7.5) and 1 mM DTT. GST-Nova-2 KH3 gel filtration running buffer contained 10 mM Hepes (pH 7.5), 100 mM KCl, 5 mM $MgCl_2$, and 1 mM DTT. All five GST-hFMRPKH domain constructs were purified in gel filtration running buffer containing 50 mM Tris (pH 7.4), 150 mM NaCl, and 1 mM DTT.

Purification of RNAs

The $PolyG_{17}$, UG_4U, CR1, $PolyC_{17}$, $PolyA_{17}$, ACUU, UGGA, $(GACG)_4$, $(UGGA)_4$, and UGGm^{6}ACU RNAs were purchased from Thermo Scientific/Dharmacon with a fluorescein tag at the 3′ end. For the NMM assay, $PolyG_{17}$ and UG_4U were also purchased without a fluorescein tag, and $(UGGA)_4$ was *in vitro* transcribed from annealed oligonucleotide template using T7 RNA polymerase. Each RNA was purified under standard denaturing conditions by urea PAGE and then extracted from the gel.

The RNA in solution was purified by chloroform extractions followed by ethanol precipitation. The purified RNAs were dissolved in RNase-free water. The RNA concentration was measured by absorbance at 260 nm using a spectrophotometer.

PAGE analysis of labeled RNAs

Denaturing urea PAGE analysis of the UG_4U RNA was performed as described previously [49]. UG_4U (20 pmol) 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 min, then incubated at 25 °C for 5 min in a thermal cycler. Then formamide loading solution was added to each RNA sample to a final 8 μl volume. The samples were heated at 60 °C for 5 min and then incubated at 25 °C for 5 min in a thermal cycler. Finally, the samples were run on a 16% urea polyacrylamide gel at room temperature for 4 h at 600 V.

For denaturing urea PAGE analysis of the 18-mer model RNAs, 10 pmol $PolyG_{17}$, 3 pmol $PolyC_{17}$, 2 pmol ACUU, 2 pmol UGGA, 6 pmol CR1, 4 pmol

Please cite this article as: Y. M. Athar and S. Joseph, RNA-Binding Specificity of the Human Fragile X Mental Retardation Protein, Journal of Molecular Biology, https://doi.org/10.1016/j.jmb.2020.04.021

ARTICLE IN PRESS

PolyA$_{17}$, 6 pmol (GACG)$_4$, and 10 pmol (UGGA)$_4$ in 20 mM Tris–HCl (pH 7.45) and 100 mM KCl were run on a 15% urea polyacrylamide gel at room temperature for 2 h at 25 W.

For native PAGE analysis of all the model RNAs, 10 pmol M1 5′ UTR, 50 pmol UG$_4$U, 10 pmol PolyG$_{17}$, 3 pmol PolyC$_{17}$, 2 pmol ACUU, 2 pmol UGGA, 6 pmol CR1, 4 pmol PolyA$_{17}$, 6 pmol (GACG)$_4$, and 10 pmol (UGGA)$_4$ in 20 mM Tris–HCl (pH 7.45), 100 mM KCl, and 5% (v/v) glycerol were run on a 15% native polyacrylamide gel containing 1× TBE. The gel was run at 4 °C for 4 h at 11 W.

BG4 antibody assay

Antibody BG4 scFv was purchased from Millipore Sigma. BG4 (100 nM) was mixed with 5 nM fluorescein-labeledRNA by diluting BG4 into fluorescence anisotropy buffer (20 mM Tris (pH 7.7), 75 mM KCl, 5 mM $MgCl_2$, and 100 ng/μl total tRNA from *E. coli*) followed by mixing with fluorescence anisotropy buffer supplemented with 5 nM fluorescein-labeledRNA directly in the well of a non-binding96-well black flat bottom plate (Greiner) in a final 200-μl volume. The plate was incubated at room temperature for 30 min. 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. Fluorescence anisotropy was measured on a Tecan Safire 2 plate reader in polarization mode. All experiments were performed a minimum of two or three times to determine the standard deviations.

NMM assay

NMM was purchased from Frontier Scientific. NMM stock solution was prepared by dissolving the stock in 0.2 N HCl to 5 mg/ml (equal to 8.61 μM). Fresh working solution was prepared by diluting the NMM stock solution to 400 μM in 10% (v/v) DMSO. Various concentrations of the RNAs were reconstituted in 158 μl solution containing 20 mM Tris (pH 7.7), 5 mM $MgCl_2$, 1 mM DTT, and either 100 mM KCl or 100 mM LiCl. The samples were then heated at 90 °C for 2 min and slowly cooled to room temperature over 1 h. Samples were briefly spun down and collected before adding 2 μl of 400 μM NMM to each 158 μl sample, to a final 5 μM NMM per sample. The samples were incubated at room temperature for 10 min before taking fluorescence measurements (JASCO FP-8500). Samples were excited at 400 nm and emission was scanned from 560 to 650 nm with a 5-nm excitation and emission bandwidths; automatic signal gain was used. The peak NMM fluorescence (610 nm) was plotted as a function of RNA concentration. All experiments were performed a minimum of two or three times to determine the standard deviations.

Fluorescence anisotropy RNA binding assay

The various FMRP constructs were titrated against 5 nM fluorescein-labeled RNAs. Fluorescence anisotropy was measured on a non-binding96-well black flat bottom plate (Greiner) using a Tecan Safire 2 plate reader in polarization mode. Some of the RNA fluorescence anisotropy measurements were also performed using a Tecan Spark plate reader.

FMRP was diluted into fluorescence anisotropy buffer supplemented with DTT (20 mM Tris (pH 7.7), 75 mM KCl, 5 mM $MgCl_2$, 100 ng/μl total tRNA from *E. coli*, and 1 mM DTT) and mixed with fluorescence anisotropy buffer supplemented with fluorescein-labeledRNA within the plate wells in a final 200-μl volume. The plate was incubated at room temperature for 30 min. 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.

To quantify the binding affinity between FMRP and RNA, the anisotropy data from each binding assay were normalized to the initial value without protein, plotted and fit to the quadratic equation below to calculate the equilibrium dissociation constant (K_D) as described previously [73]:

$$\frac{[P+RNA]}{[RNA]} = \frac{([P]+[RNA]+K_D)-\sqrt{([P]+[RNA]+K_D)^2-4[P][RNA]}}{2[P]}$$

where [P + RNA]/[RNA] is the anisotropy value, [RNA] is the RNA concentration, and [P] is the protein concentration. GraphPad Prism software (Graphpad Software Inc.) was used to perform the curve fits. All experiments were performed a minimum of two or three times to determine the standard deviations.

MIDAS: DNA template and primers

The MIDAS N5 DNA template sequence: 5′-mAmATGCGTANNNNNTGGATCCCTATAGT-GAGTCGTATTA-3′ was purchased from IDT and has two 2′-O-methyl substitutions (mA) at the 5′ terminus to reduce the amount of $n+1$ product during *in vitro* transcription. The five randomized positions were produced by hand mixing the nucleotides to give equal ratio of all four nucleotides. 18T7T sequence: 5′-TAATACGACTCACTATAG-3′. First-strand primer: 5′-GTCTCGTGGGCTCGGA-GATGTGTATAAGAGACAG AATGCGTA-3′. Second-strand primer: 5′-TCGTCGGCAGCGTCAGATGTGTATAAGAGA-CAGGGGATCCA-3′. Index 1 Read primer: 5′-CAAGCAGAAGACGGCATACGAGAT [i7]

Please cite this article as: Y. M. Athar and S. Joseph, RNA-Binding Specificity of the Human Fragile X Mental Retardation Protein, Journal of Molecular Biology, https://doi.org/10.1016/j.jmb.2020.04.021

ARTICLE IN PRESS

GTCTCGTGGGCTCGG-3′. Index 2 Read primer: 5′-AATGATACGGCGACCACCGAGATCTACAC [i5] TCGTCGGCAGCGTC-3′.

MIDAS: Separation of RNA•protein complex using a native gel

The MIDAS RNAs were generated by *in vitro* transcription of annealed 18T7T and MIDAS template DNAs using T7 RNA polymerase. The sequence of the 21-merMIDAS N5 RNA is 5′-GGGAUCCANNNN-NUACGCAUU-3′. The transcription products were purified under standard denaturing conditions by urea PAGE and then extracted from the gel. Each RNA was purified by chloroform extractions followed by ethanol precipitation. The purified RNAs were dissolved in RNase-free water. The RNA concentration was measured by absorbance at 260 nm using a spectrophotometer.

The MIDAS N5 RNA pool was 5′ dephosphorylated with alkaline phosphatase (NEB) and labeled with [γ-^{32}P]-ATP and T4 polynucleotide kinase (NEB). AnEMSA was used to detect and isolate protein-boundMIDAS RNAs from the pool. Briefly, in a 10-μl sample, 100,000 cpm of MIDAS N5 RNA was incubated with 10 μM protein (or protein storage buffer in control lane) in RNA binding buffer (50 mM Tris acetate (pH 7.7), 50 mM potassium acetate and 5 mM magnesium acetate, 1 mM DTT, 0.1 mg/ml bovine serum albumin, and 100 ng/μl total tRNA from *E. coli*) for 45 min at room temperature. Glycerol was added to 5% final concentration and the samples were run on an 8% (79:1 acrylamide: bisacrylamide) 0.5× TBE native polyacrylamide gel at 4 °C for 2 h and 30 min at a constant 300 V. The gel was exposed to a phosphor screen at −80 °C and imaged on a Typhoon FLA 9500 phosphorimager (GE) after 18 h. The gel was thawed at room temperature and the shifted band for each protein was extracted from the gel. Each selected RNA was purified by chloroform extractions followed by ethanol precipitation. The purified RNAs were dissolved in 35 μl RNase-free water.

MIDAS: Library preparation

The selected MIDAS RNAs were used to generate a cDNA library for next-generation sequencing. This was accomplished in three simple steps: (1) synthesize the first-strand cDNA by reverse transcription, (2) synthesize the second-strand by primer extension, and (3) amplify the dsDNA by PCR. First, 8 μl of the selected RNA and 2 μl of 10 μM MIDAS reverse transcription primer were combined and annealed by heating at 65 °C for 5 min and cooling to room temperature for 2 min before placing on ice. The 10 μl annealed RNA-primer was mixed with 10 μl 2× cDNA synthesis mix containing 2× FS buffer (100 mM Tris (pH) 8.3, 150 mM KCl, and 6 mM $MgCl_2$), 1 mM dNTPs, 10 mM DTT, and 10 U Superscript III reverse transcriptase (ThermoFisher). The reverse transcription was carried out in a thermal cycler: (1) 10 min at 25 °C, (2) 1 h and 30 min at 30 °C, (3) 15 min at 70 °C, (4) 37 °C hold for RNase H digestion. On step 4, the 20 μl first-strand samples were placed on ice for 1 min and spun down. Then 1 μl of 5 U/μl RNase H (NEB) was added, and the sample was incubated at 37 °C for 20 min.

To the 21 μl first-strand cDNA sample, we added: 3.5 μl of 100 mM NaCl, 6.4 μl of 100 mM $MgCl_2$, 10 μl of 10 mM dNTPs, 10 μl of 10 μM MIDAS cDNA Klenow primer, 2 μl of 5 U/μl Klenow fragment 3′ → 5′ exo$^-$ (NEB), and water to a final 100 μl volume. This made an ideal 100 μl primer extension reaction containing 10 mM Tris (pH 8.2), 15 mM KCl, 35 mM NaCl, 7 mM $MgCl_2$, and 1 mM DTT. The reaction proceeded for 30 min at room temperature and was then deactivated for 20 min at 75 °C. The second-strand product was then amplified by PCR with sample-specific Index 1 and Index 2 primers containing their unique i7 and i5 index sequences.

PCR was carried out using a KOD HS polymerase system (MilliporeSigma). PCRs (50 μl) were carried according to manufacturer protocol using 5 μl of the 100-μl second-strand product. The DNA was amplified over 25 cycles (annealing temperature of 60 °C and a 3 s extension time). The DNA library was finally cleaned up using an Amico Ultra-0.5 ml centrifugal filter with a 30,000-Da molecular weight cutoff. The 50-μl PCR product was diluted in 450 μl water and concentrated down to approximately 50 μl a total of three times. The DNA concentration was measured by absorbance at 260 nm using a spectrophotometer. The Nova-2 KH3, hFMRP KH0, hFMRP KH1–KH2, and total input DNA libraries were mixed in equimolar ratios and then sequenced with Illumina HiSeq 4000 at the UCSD IGM Genomics Center.

MIDAS: NGS data analysis

We obtained more than 4 million reads for each Illumina sequencing run. To identify and rank the favorite motifs of each protein, we used a custom Python script, which first filtered the data by sequence length to show only the N5 region sequences. Doing this revealed exactly 1024 unique 5-mer sequences. Counts from each of the 1024 sequences were compared between the protein-selected libraries and the total input library to generate a list of enrichment factors. To determine if any of the top enriched sequences had a significant number of counts coming from mutated sequences, the data were also filtered by sequence length to show the N5 region plus 3 bases upstream and 3 bases downstream. This revealed how many counts for each N5 motif came from the wild-type sequence

Please cite this article as: Y. M. Athar and S. Joseph, RNA-Binding Specificity of the Human Fragile X Mental Retardation Protein, Journal of Molecular Biology, https://doi.org/10.1016/j.jmb.2020.04.021

ARTICLE IN PRESS

and how many came from mutated sequences. More than 99% of the reads were the correct wild-type sequence and only counts coming from the wild-type sequences were used to calculate the enrichment factor. Identical sequencing data were obtained when the Illumina sequencing was repeated with another set of independent library preparations showing that the process is reproducible.

Acknowledgment

We thank Lila Mouakkad and Norman Zhu for performing some of the initial studies, Robert Konecny for writing the Python script to analyze the MIDAS NGS data, and Gourisankar Ghosh and members of the Joseph lab for useful comments on the manuscript. This publication includes data generated at the UC San Diego IGM Genomics Center utilizing an Illumina NovaSeq 6000 that was purchased with funding from a NIH SIG grant (#S10 OD026929).

Funding

This work was supported by the National Institutes of Health (R01GM114261 to S.J.). Funding for open-access charge: National Institutes of Health.

Declaration of Completing Interest

The authors declare that there is no conflict of interest.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jmb.2020.04.021.

Received 24 January 2020;
Received in revised form 16 April 2020;
Accepted 22 April 2020
Available online xxxx

Keywords:
fragile X syndrome;
fragile X mental retardation protein;
KH domain;
RGG domain;
RNA G-quadruplex

Abbreviations used:
FXS, fragile X syndrome; FMRP, fragile X mental retardation protein; KH, K-homology; RGG, arginine–glycine–glycine; MIDAS, *Motif Identification by Analysis of Simple sequences*; GST, glutathione S-transferase; NMM, *N-methyl mesoporphyrin IX*; EMSA, electrophoretic mobility shift assay.

References

[1] M.R. Santoro, S.M. Bray, S.T. Warren, Molecular mechanisms of fragile X syndrome: a twenty-year perspective, Annu. Rev. Pathol. 7 (2012) 219–245, https://doi.org/10.1146/annurev-pathol-011811-132457.

[2] B. Coffee, F. Zhang, S.T. Warren, D. Reines, Acetylated histones are associated with FMR1 in normal but not fragile X-syndrome cells, Nat. Genet. 22 (1999) 98–101, https://doi.org/10.1038/8807.

[3] D.L. Nelson, H.T. Orr, S.T. Warren, The unstable repeats—three evolving faces of neurological disease, Neuron. 77 (2013) 825–843, https://doi.org/10.1016/j.neuron.2013.02.022.

[4] J.S. Sutcliffe, D.L. Nelson, F. Zhang, M. Pieretti, C.T. Caskey, D. Saxe, S.T. Warren, DNA methylation represses FMR-1 transcription in fragile X syndrome, Hum. Mol. Genet. 1 (1992) 397–400.

[5] A.J. Verkerk, M. Pieretti, J.S. Sutcliffe, Y.H. Fu, D.P. Kuhl, A. Pizzuti, O. Reiner, S. Richards, et al., 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 (1991) 905–914 https://doi.org/0092-8674(91)90397-H (pii).

[6] G.J. Bassell, S.T. Warren, Fragile X syndrome: loss of local mRNA regulation alters synaptic development and function, Neuron. 60 (2008) 201–214 , https://doi.org/S0896-6273(08)00847-7 [pii] https://doi.org/10.1016/j.neuron.2008.10.004.

[7] B. Laggerbauer, D. Ostareck, E.M. Keidel, A. Ostareck-Lederer, U. Fischer, Evidence that fragile X mental retardation protein is a negative regulator of translation, Hum. Mol. Genet. 10 (2001) 329–338.

[8] Z. Li, Y. Zhang, L. Ku, K.D. Wilkinson, S.T. Warren, Y. Feng, The fragile X mental retardation protein inhibits translation via interacting with mRNA, Nucleic Acids Res. 29 (2001) 2276–2283.

[9] T.A. Comery, J.B. Harris, P.J. Willems, B.A. Oostra, S.A. Irwin, I.J. Weiler, W.T. Greenough, Abnormal dendritic spines in fragile X knockout mice: maturation and pruning deficits, Proc. Natl. Acad. Sci. U. S. A. 94 (1997) 5401–5404.

[10] A.W. Grossman, G.M. Aldridge, I.J. Weiler, W.T. Greenough, Local protein synthesis and spine morphogenesis: fragile X syndrome and beyond, J. Neurosci. 26 (2006) 7151–7155 , https://doi.org/26/27/7151 [pii] https://doi.org/10.1523/JNEUROSCI.1790-06.2006.

[11] S.A. Irwin, R. Galvez, W.T. Greenough, Dendritic spine structural anomalies in fragile-X mental retardation syndrome, Cereb. Cortex 10 (2000) 1038–1044.

[12] C.T. Ashley, K.D. Wilkinson, D. Reines, S.T. Warren, FMR1 protein: conserved RNP family domains and selective RNA binding, Science. 262 (1993) 563–566.

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

[14] H. Siomi, M.J. Matunis, W.M. Michael, G. Dreyfuss, The pre-mRNA binding K protein contains a novel evolutionarily conserved motif, Nucleic Acids Res. 21 (1993) 1193–1198.

ARTICLE IN PRESS

[15] R. Valverde, I. Pozdnyakova, T. Kajander, J. Venkatraman, L. Regan, Fragile X mental retardation syndrome: structure of the KH1–KH2 domains of fragile X mental retardation protein, Structure. 15 (2007) 1090–1098 , https://doi.org/S0969-2126(07)00280-8 [pii] https://doi.org/10.1016/j.str.2007.06.022.
[16] D. Hollingworth, A.M. Candel, G. Nicastro, S.R. Martin, P. Briata, R. Gherzi, A. Ramos, KH domains with impaired nucleic acid binding as a tool for functional analysis, Nucleic Acids Res. 40 (2012) 6873–6886 , https://doi.org/gks368 [pii] https://doi.org/10.1093/nar/gks368.
[17] G. Nicastro, I.A. Taylor, A. Ramos, KH–RNA interactions: back in the groove, Curr. Opin. Struct. Biol. 30 (2015) 63–70, https://doi.org/10.1016/j.sbi.2015.01.002.
[18] Y. Hu, Z. Chen, Y. Fu, Q. He, L. Jiang, J. Zheng, Y. Gao, P. Mei, et al., The amino-terminal structure of human fragile X mental retardation protein obtained using precipitant-immobilized imprinted polymers, Nat. Commun. 6 (2015) 6634, https://doi.org/10.1038/ncomms7634.
[19] L.K. Myrick, H. Hashimoto, X. Cheng, S.T. Warren, Human FMRP contains an integral tandem Agenet (Tudor) and KH motif in the amino terminal domain, Hum. Mol. Genet. 24 (2015) 1733–1740, https://doi.org/10.1093/hmg/ddu586.
[20] K. De Boulle, A.J. Verkerk, E. Reyniers, L. Vits, J. Hendrickx, B. Van Roy, F. Van den Bos, E. de Graaff, et al., A point mutation in the FMR-1 gene associated with fragile X mental retardation, Nat. Genet. 3 (1993) 31–35, https://doi.org/10.1038/ng0193-31.
[21] L.K. Myrick, M. Nakamoto-Kinoshita, N.M. Lindor, S. Kirmani, X. Cheng, S.T. Warren, Fragile X syndrome due to a missense mutation, Eur J Hum Genet (2014)https://doi.org/10.1038/ejhg.2013.311.
[22] J.C. Darnell, C.E. Fraser, O. Mostovetsky, G. Stefani, T.A. Jones, S.R. Eddy, R.B. Darnell, Kissing complex RNAs mediate interaction between the fragile-X mental retardation protein KH2 domain and brain polyribosomes, Genes Dev. 19 (2005) 903–918 , https://doi.org/gad.1276805 [pii] https://doi.org/10.1101/gad.1276805.
[23] J.C. Darnell, K.B. Jensen, P. Jin, V. Brown, S.T. Warren, R.B. Darnell, Fragile X mental retardation protein targets G quartet mRNAs important for neuronal function, Cell. 107 (2001) 489–499.
[24] A.T. Phan, V. Kuryavyi, J.C. Darnell, A. Serganov, A. Majumdar, S. Ilin, T. Raslin, A. Polonskaia, et al., Structure-function studies of FMRP RGG peptide recognition of an RNA duplex-quadruplex junction, Nat. Struct. Mol. Biol. 18 (2011) 796–804 , https://doi.org/nsmb.2064 [pii] https://doi.org/10.1038/nsmb.2064.
[25] N. Vasilyev, A. Polonskaia, J.C. Darnell, R.B. Darnell, D.J. Patel, A. Serganov, Crystal structure reveals specific recognition of a G-quadruplex RNA by a beta-turn in the RGG motif of FMRP, Proc Natl Acad Sci U A. 112 (2015) E5391–E5400, https://doi.org/10.1073/pnas.1515737112.
[26] A.C. Blice-Baum, M.R. Mihailescu, Biophysical characterization of G-quadruplex forming FMR1 mRNA and of its interactions with different fragile X mental retardation protein isoforms, RNA. 20 (2014) 103–114, https://doi.org/10.1261/rna.041442.113.
[27] L. Menon, M.R. Mihailescu, Interactions of the G quartet forming semaphorin 3F RNA with the RGG box domain of the fragile X protein family, Nucleic Acids Res. 35 (2007) 5379–5392, https://doi.org/10.1093/nar/gkm581.
[28] L. Menon, S.A. Mader, M.R. Mihailescu, Fragile X mental retardation protein interactions with the microtubule associated protein 1B RNA, RNA. 14 (2008) 1644–1655, https://doi.org/10.1261/rna.1100708.
[29] S. Stefanovic, B.A. DeMarco, A. Underwood, K.R. Williams, G.J. Bassell, M.R. Mihailescu, Fragile X mental retardation protein interactions with a G quadruplex structure in the 3′-untranslated region of NR2B mRNA, Mol. BioSyst. 11 (2015) 3222–3230, https://doi.org/10.1039/c5mb00423c.
[30] Y. Zhang, C.M. Gaetano, K.R. Williams, G.J. Bassell, M.R. Mihailescu, FMRP interacts with G-quadruplex structures in the 3′-UTR of its dendritic target Shank1 mRNA, RNA Biol. 11 (2014) 1364–1374, https://doi.org/10.1080/15476286.2014.996464.
[31] M. Ascano, N. Mukherjee, P. Bandaru, J.B. Miller, J.D. Nusbaum, D.L. Corcoran, C. Langlois, M. Munschauer, et al., FMRP targets distinct mRNA sequence elements to regulate protein expression, Nature. 492 (2012) 382–386, https://doi.org/10.1038/nature11737.
[32] V. Brown, P. Jin, S. Ceman, J.C. Darnell, W.T. O'Donnell, S. A. Tenenbaum, X. Jin, Y. Feng, et al., Microarray identification of FMRP-associated brain mRNAs and altered mRNA translational profiles in fragile X syndrome, Cell. 107 (2001) 477–487.
[33] J.C. Darnell, S.J. Van Driesche, C. Zhang, K.Y. Hung, A. Mele, C.E. Fraser, E.F. Stone, C. Chen, et al., FMRP stalls ribosomal translocation on mRNAs linked to synaptic function and autism, Cell. 146 (2011) 247–261 , https://doi.org/S0092-8674(11)00655-6 [pii] https://doi.org/10.1016/j.cell.2011.06.013.
[34] A.C. McMahon, R. Rahman, H. Jin, J.L. Shen, A. Fieldsend, W. Luo, M. Rosbash, TRIBE: hijacking an RNA-editing enzyme to identify cell-specific targets of RNA-binding proteins, Cell. 165 (2016) 742–753, https://doi.org/10.1016/j.cell.2016.03.007.
[35] D. Ray, H. Kazan, K.B. Cook, M.T. Weirauch, H.S. Najafabadi, X. Li, S. Gueroussov, M. Albu, et al., A compendium of RNA-binding motifs for decoding gene regulation, Nature 499 (2013) 172–177, https://doi.org/10.1038/nature12311.
[36] J.A. Suhl, P. Chopra, B.R. Anderson, G.J. Bassell, S.T. Warren, Analysis of FMRP mRNA target datasets reveals highly associated mRNAs mediated by G-quadruplex structures formed via clustered WGGA sequences, Hum. Mol. Genet. 23 (2014) 5479–5491, https://doi.org/10.1093/hmg/ddu272.
[37] K.B. Jensen, K. Musunuru, H.A. Lewis, S.K. Burley, R.B. Darnell, The tetranucleotide UCAY directs the specific recognition of RNA by the Nova K-homology 3 domain, Proc. Natl. Acad. Sci. U. S. A. 97 (2000) 5740–5745, https://doi.org/10.1073/pnas.090553997.
[38] H.A. Lewis, K. Musunuru, K.B. Jensen, C. Edo, H. Chen, R.B. Darnell, S.K. Burley, Sequence-specific RNA binding by a Nova KH domain: implications for paraneoplastic disease and the fragile X syndrome, Cell. 100 (2000) 323–332 https://doi.org/S0092-8674(00)80668-6 (pii).
[39] G. Nicastro, A.M. Candel, M. Uhl, A. Oregioni, D. Hollingworth, R. Backofen, S.R. Martin, A. Ramos, Mechanism of β-actin mRNA recognition by ZBP1, Cell Rep. 18 (2017) 1187–1199, https://doi.org/10.1016/j.celrep.2016.12.091.
[40] B.A. Ozdilek, V.F. Thompson, N.S. Ahmed, C.I. White, R.T. Batey, J.C. Schwartz, Intrinsically disordered RGG/RG domains mediate degenerate specificity in RNA binding, Nucleic Acids Res. 45 (2017) 7984–7996, https://doi.org/10.1093/nar/gkx460.
[41] K.J. Zanotti, P.E. Lackey, G.L. Evans, M.R. Mihailescu, Thermodynamics of the fragile X mental retardation protein RGG box interactions with G quartet forming RNA, Biochemistry. 45 (2006) 8319–8330, https://doi.org/10.1021/bi060209a.

Please cite this article as: Y. M. Athar and S. Joseph, RNA-Binding Specificity of the Human Fragile X Mental Retardation Protein, Journal of Molecular Biology, https://doi.org/10.1016/j.jmb.2020.04.021

ARTICLE IN PRESS

[42] M.C. Siomi, H. Siomi, W.H. Sauer, S. Srinivasan, R.L. Nussbaum, G. Dreyfuss, FXR1, an autosomal homolog of the fragile X mental retardation gene, EMBO J. 14 (1995) 2401–2408.
[43] M.C. Siomi, Y. Zhang, H. Siomi, G. Dreyfuss, 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 (1996) 3825–3832.
[44] L. Wan, T.C. Dockendorff, T.A. Jongens, G. Dreyfuss, Characterization of dFMR1, a Drosophila melanogaster homolog of the fragile X mental retardation protein, Mol. Cell. Biol. 20 (2000) 8536–8547.
[45] C. Cheong, P.B. Moore, Solution structure of an unusually stable RNA tetraplex containing G- and U-quartet structures, Biochemistry. 31 (1992) 8406–8414.
[46] J. Deng, Y. Xiong, M. Sundaralingam, X-ray analysis of an RNA tetraplex (UGGGGU)(4) with divalent Sr(2+) ions at subatomic resolution (0.61 A), Proc Natl Acad Sci U A 98 (2001) 13665–13670, https://doi.org/10.1073/pnas.241374798.
[47] X. Wang, K.J. Goodrich, A.R. Gooding, H. Naeem, S. Archer, R.D. Paucek, D.T. Youmans, T.R. Cech, et al., Targeting of Polycomb repressive complex 2 to RNA by short repeats of consecutive guanines, Mol. Cell 65 (2017) 1056–1067 , e5 https://doi.org/10.1016/j.molcel.2017.02.003.
[48] N.G. Walter, J.M. Burke, Fluorescence assays to study structure, dynamics, and function of RNA and RNA-ligand complexes, Methods Enzym. 317 (2000) 409–440.
[49] J. Kim, C. Cheong, P.B. Moore, Tetramerization of an RNA oligonucleotide containing a GGGG sequence, Nature. 351 (1991) 331–332, https://doi.org/10.1038/351331a0.
[50] A. Bugaut, S. Balasubramanian, 5′-UTR RNA G-quadruplexes: translation regulation and targeting, Nucleic Acids Res. 40 (2012) 4727–4741, https://doi.org/10.1093/nar/gks068.
[51] G.W. Collie, S.M. Haider, S. Neidle, G.N. Parkinson, A crystallographic and modelling study of a human telomeric RNA (TERRA) quadruplex, Nucleic Acids Res. 38 (2010) 5569–5580, https://doi.org/10.1093/nar/gkq259.
[52] O. Kikin, L. D'Antonio, P.S. Bagga, QGRS mapper: a web-based server for predicting G-quadruplexes in nucleotide sequences, Nucleic Acids Res. 34 (2006) W676–W682, https://doi.org/10.1093/nar/gkl253.
[53] V.T. Mukundan, A.T. Phan, Bulges in G-quadruplexes: broadening the definition of G-quadruplex-forming sequences, J. Am. Chem. Soc. 135 (2013) 5017–5028, https://doi.org/10.1021/ja310251r.
[54] G. Biffi, M. Di Antonio, D. Tannahill, S. Balasubramanian, Visualization and selective chemical targeting of RNA G-quadruplex structures in the cytoplasm of human cells, Nat. Chem. 6 (2014) 75–80, https://doi.org/10.1038/nchem.1805.
[55] T. Endoh, A.B. Rode, S. Takahashi, Y. Kataoka, M. Kuwahara, N. Sugimoto, Real-time monitoring of G-quadruplex formation during transcription, Anal. Chem. 88 (2016) 1984–1989, https://doi.org/10.1021/acs.analchem.5b04396.
[56] P. Ivanov, E. O'Day, M.M. Emara, G. Wagner, J. Lieberman, P. Anderson, G-quadruplex structures contribute to the neuroprotective effects of angiogenin-induced tRNA fragments, Proc. Natl. Acad. Sci. U. S. A. 111 (2014) 18201–18206, https://doi.org/10.1073/pnas.1407361111.
[57] E. Chen, M.R. Sharma, X. Shi, R.K. Agrawal, S. Joseph, Fragile X mental retardation protein regulates translation by binding directly to the ribosome, Mol. Cell 54 (2014) 407–417, https://doi.org/10.1016/j.molcel.2014.03.023.
[58] A. Ramos, D. Hollingworth, A. Pastore, G-quartet-dependent recognition between the FMRP RGG box and RNA, RNA. 9 (2003) 1198–1207.
[59] A.E. Arguello, A.N. DeLiberto, R.E. Kleiner, RNA chemical proteomics reveals the *N*(6)-methyladenosine (m(6)A)-regulated protein–RNA Interactome, J. Am. Chem. Soc. 139 (2017) 17249–17252, https://doi.org/10.1021/jacs.7b09213.
[60] R.R. Edupuganti, S. Geiger, R.G.H. Lindeboom, H. Shi, P.J. Hsu, Z. Lu, S.Y. Wang, M.P.A. Baltissen, et al., *N*(6)-methyladenosine (m(6)A) recruits and repels proteins to regulate mRNA homeostasis, Nat. Struct. Mol. Biol. 24 (2017) 870–878, https://doi.org/10.1038/nsmb.3462.
[61] P.J. Hsu, H. Shi, A.C. Zhu, Z. Lu, N. Miller, B.M. Edens, Y.C. Ma, C. He, The RNA-binding protein FMRP facilitates the nuclear export of *N*6-methyladenosine-containing mRNAs, J. Biol. Chem. 294 (2019) 19889–19895, https://doi.org/10.1074/jbc.AC119.010078.
[62] K.D. Meyer, S.R. Jaffrey, Rethinking m(6)A readers, writers, and erasers, Annu. Rev. Cell Dev. Biol. 33 (2017) 319–342, https://doi.org/10.1146/annurev-cellbio-100616-060758.
[63] C. Xu, X. Wang, K. Liu, I.A. Roundtree, W. Tempel, Y. Li, Z. Lu, C. He, et al., Structural basis for selective binding of m6A RNA by the YTHDC1 YTH domain, Nat. Chem. Biol. 10 (2014) 927–929, https://doi.org/10.1038/nchembio.1654.
[64] T. Zhu, I.A. Roundtree, P. Wang, X. Wang, L. Wang, C. Sun, Y. Tian, J. Li, et al., Crystal structure of the YTH domain of YTHDF2 reveals mechanism for recognition of *N*6-methyladenosine, Cell Res. 24 (2014) 1493–1496, https://doi.org/10.1038/cr.2014.152.
[65] F. Corbin, M. Bouillon, A. Fortin, S. Morin, F. Rousseau, E.W. Khandjian, The fragile X mental retardation protein is associated with poly(A)+ mRNA in actively translating polyribosomes, Hum. Mol. Genet. 6 (1997) 1465–1472 https://doi.org/dda192 (pii).
[66] Y. Feng, D. Absher, D.E. Eberhart, V. Brown, H.E. Malter, S. T. Warren, FMRP associates with polyribosomes as an mRNP, and the I304N mutation of severe fragile X syndrome abolishes this association, Mol. Cell 1 (1997) 109–118.
[67] A. Ishizuka, M.C. Siomi, H. Siomi, A Drosophila fragile X protein interacts with components of RNAi and ribosomal proteins, Genes Dev. 16 (2002) 2497–2508, https://doi.org/10.1101/gad.1022002.
[68] E.W. Khandjian, F. Corbin, S. Woerly, F. Rousseau, The fragile X mental retardation protein is associated with ribosomes, Nat. Genet. 12 (1996) 91–93, https://doi.org/10.1038/ng0196-91.
[69] R. Mazroui, M.E. Huot, S. Tremblay, N. Boilard, Y. Labelle, E. W. Khandjian, Fragile X mental retardation protein determinants required for its association with polyribosomal mRNPs, Hum. Mol. Genet. 12 (2003) 3087–3096, https://doi.org/10.1093/hmg/ddg335 (ddg335 [pii]).
[70] D. Simsek, G.C. Tiu, R.A. Flynn, G.W. Byeon, K. Leppek, A. F. Xu, H.Y. Chang, M. Barna, The mammalian ribo-interactome reveals ribosome functional diversity and heterogeneity, Cell 169 (2017) 1051–1065 , e18 https://doi.org/10.1016/j.cell.2017.05.022.
[71] B. DeMarco, S. Stefanovic, A. Williams, K.R. Moss, B.R. Anderson, G.J. Bassell, M.R. Mihailescu, FMRP–G-quadruplex mRNA–miR-125a interactions: implications for miR-125a mediated translation regulation of PSD-95 mRNA, PLoS One 14 (2019), e0217275. https://doi.org/10.1371/journal.pone.0217275.
[72] R.S. Muddashetty, V.C. Nalavadi, C. Gross, X. Yao, L. Xing, O. Laur, S.T. Warren, G.J. Bassell, Reversible inhibition of

Please cite this article as: Y. M. Athar and S. Joseph, RNA-Binding Specificity of the Human Fragile X Mental Retardation Protein, Journal of Molecular Biology, https://doi.org/10.1016/j.jmb.2020.04.021

ARTICLE IN PRESS

PSD-95 mRNA translation by miR-125a, FMRP phosphorylation, and mGluR signaling, Mol. Cell 42 (2011) 673–688 , https://doi.org/S1097-2765(11)00337-6 [pii] https://doi.org/10.1016/j.molcel.2011.05.006.

[73] T.D. Pollard, A guide to simple and informative binding assays, Mol. Biol. Cell 21 (2010) 4061–4067, https://doi.org/10.1091/mbc.E10-08-0683.

Please cite this article as: Y. M. Athar and S. Joseph, RNA-Binding Specificity of the Human Fragile X Mental Retardation Protein, Journal of Molecular Biology, https://doi.org/10.1016/j.jmb.2020.04.021