Minireview

Targeting fragile X

Ilse Gantois and R Frank Kooy

Address: Department of Medical Genetics, University of Antwerp, Universiteitsplein 1, 2610 Antwerp, Belgium.

Correspondence: R Frank Kooy. E-mail: Frank.Kooy@ua.ac.be

Published: 25 April 2002

Genome ***Biology*** 2002, **3(5)**:reviews1014.1–1014.5

The electronic version of this article is the complete one and can be found online at http://genomebiology.com/2002/3/5/reviews/1014

Abstract

Ten years after the identification of the gene responsible for fragile X syndrome, recent studies have revealed a list of mRNAs bound by the fragile X gene product and have identified specific sequences required for the interaction between the fragile X protein and its targets. These results are a breakthrough in understanding why absence of the fragile X protein leads to mental retardation.

With a prevalence of 1 in 5,000, fragile X syndrome is the most common form of hereditary mental retardation [1]. Patients suffer from mild to severe mental retardation, mild facial dismorphologies and macroorchidism (enlarged testicles). In addition, behavioral problems like autism and hyperactivity may be associated with the syndrome [2]. In 1991, defects in the *FMR1* gene (fragile X mental retardation gene 1) were identified as the cause of the syndrome [3]. An expansion of a CGG repeat in the 5′ untranslated region (UTR) of the *FMR1* gene causes transcriptional silencing, resulting in absence of the fragile X protein, FMRP. It was soon discovered that FMRP can bind RNA and plays a role in the transport of mRNAs from the nucleus to ribosomes. Several recent studies [4-6] have provided new insights into the motifs responsible for RNA binding and revealed a list of mRNA targets bound by FMRP.

RNA binding by FMRP

The *FMR1* gene is ubiquitously expressed, with particularly high expression in neurons but barely detectable expression in glia [7]. The *FMR1* gene product can be up to 631 amino acids long and has five functional domains: two types of RNA-binding structures, namely two KH domains (KH denotes heterogeneous nuclear ribonucleoprotein K homology) and an RGG box (containing repeats of an Arg-Gly-Gly motif - RGG, in the single-letter amino-acid code); a nuclear localization signal (NLS); a nuclear export signal (NES); and two coiled-coil domains. FMRP is predominantly cytoplasmic but shuttles between the cytoplasm and nucleus, a process mediated by the nuclear import and nuclear export signals. FMRP can bind 4% of all brain mRNAs, including its own mRNA, presumably with the aid of its RNA-binding domains [8]. In the nucleus, the FMRP protein forms messenger ribonucleoprotein (mRNP) precursor particles along with other proteins, including nucleolin (a known component of other mRNPs) and the fragile X-related proteins FXR1P and FXR2P [9,10]. The *FXR1* and *FXR2* genes are autosomal paralogs of *FMR1*, and both gene products have high overall sequence similarity to FMRP, especially within the functional domains [11]. The FMRP-mRNP particles are transported out of the nucleus into the cytoplasm and to actively translating ribosomes [12,13]. It has been hypothesized that, in neurons, FMRP transports specific mRNAs along the dendrites towards actively translating ribosomes near the synapses. Which neuronal RNA targets are bound and why absence of the FMRP protein causes mental retardation remained unknown until recently, however.

Initially, the KH domains were thought to be crucial for the function of FMRP, because an exceptional patient had been described with a point mutation (Ile304 → Asn) in the second KH domain [14]. This patient had a more severe phenotype than any other patient lacking FMRP, including profound mental retardation and excessive macroorchidism. Experiments *in vitro* demonstrated that the FMRP KH domains interact with high affinity with various mRNAs,

including those of *FXR1* and *FXR2* [11,15]. FMRP is able to suppress translation of mRNAs *in vitro*, presumably through its KH domains [16,17]. Mutation of FMRP at position 304 abrogates its association with ribosomes and the mutated protein does not suppress translation.

It came as a surprise, therefore, that a series of recent papers [4-6] revealed that the RGG box, not the KH domains, is responsible for the binding of a specific subset of mRNAs to FMRP. RGG boxes are usually found within proteins in combination with other RNA-binding domains [18]. Darnell *et al.* [4] identified which RNAs were preferentially bound by FMRP, by screening a pool of 96-mer RNAs containing 52 bases of random sequence. After multiple rounds of selection a limited set of RNAs was identified with a common motif of $[DWGGN_{(0\text{-}2)}]_4$ (where D is any nucleotide except C, W is U or A, and N is any nucleotide). At the same time, Schaeffer *et al.* [5] determined the minimal site within the *FMR1* mRNA necessary for FMRP binding. They concluded that a 100-nucleotide fragment in the 3′ coding region of the *FMR1* gene is responsible for the specific binding of FMRP with its own mRNA, in contrast to earlier experiments in which FMRP was reported to bind to the 3′ UTR of the *FMR1* mRNA [19].

G-quartet structures

Both the random RNAs with the repeated DWGG motif that are preferentially bound by FMRP and the 100 nucleotides at the 3′ end of the *FMR1* gene necessary for FMRP binding have the ability to form planar structures, so-called G quartets. G quartets are hydrogen-bonded structures formed from four guanosine residues in a square-planar array stabilized preferentially by K^+ (Figure 1). G-quartet structures have been observed in telomeric sequences and stabilize chromosome ends; they dimerize the RNA genome of the human immunodeficiency virus (HIV) and are also involved in the site-specific recombination required for immunoglobulin maturation [20]. The presence of a stable G-quartet structure near the 3′ end of the coding sequence of *FMR1* was investigated by reverse transcriptase (RT)-coupled PCR [5]. Adding K^+ to the RT buffer caused an abrupt pause in reverse transcription before and within the purine-rich region and this block was resolved when Na^+ was added. Binding of FMRP to a G-quartet-containing sequence was strongly reduced when Na^+ and Li^+ cations were used compared with K^+. Surprisingly, the mRNA-binding site of FMRP that binds the G quartet overlaps with the RGG box, which is responsible for the binding of specific mRNAs [4,5]. When translated, therefore, the product of the G-quartet sequence, which is the RGG box of FMRP, is able to bind to its own mRNA.

G-quartet-containing RNAs were also shown to be targets of FMRP *in vivo*. Using microarrays, Brown *et al.* [6] identified from total murine brain mRNAs associated with mRNPs that

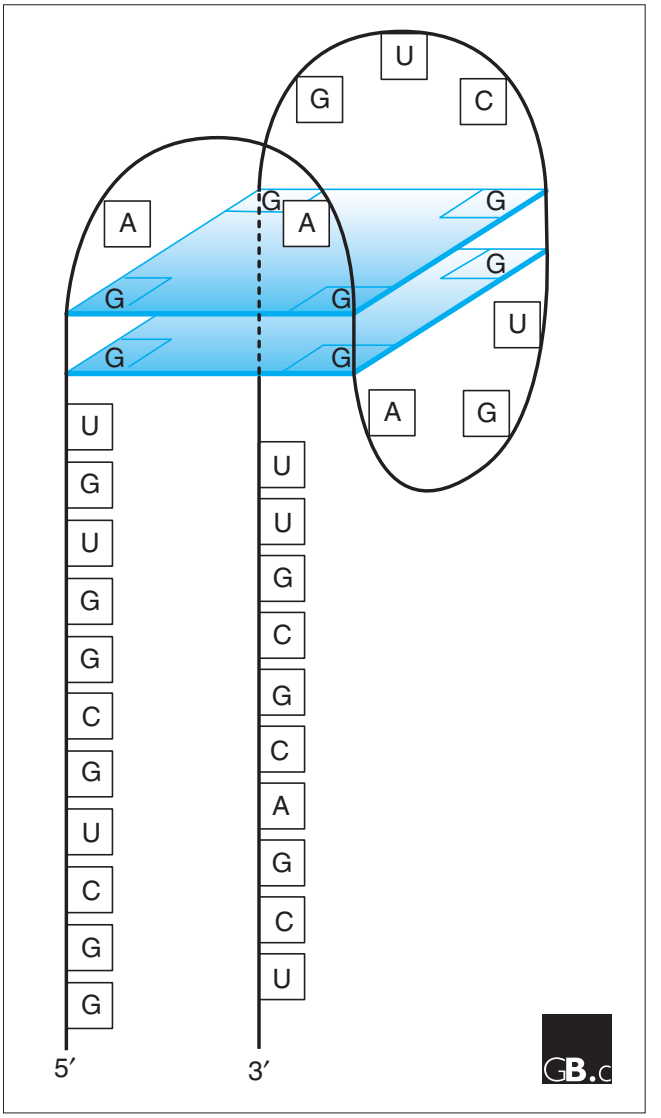

Figure 1
Structure of a G quartet. In this schematic representation, the sequence of *sc1* RNA is used to illustrate the three-dimensional structure of a G quartet; *sc1* is one of a series of six random RNAs specifically bound by FMRP [4]. Four guanines form hydrogen bonds with each other in a symmetrical square planar arrangement. The G-quartet structure is stabilized by cations such as K^+.

interact with actively translating ribosomes. FMRP-RNP particles were isolated from wild-type and knockout mouse brains and hybridized to microarrays containing 25,181 RNAs with sequences taken from the UniGene database of the National Center for Biotechnology Information [21] and 19,021 expressed sequence tags (ESTs) from the TIGR Gene Indices database [22]. These experiments confirmed that FMRP binds approximately 4% of total brain mRNA. After rigorous selection, Brown *et al.* [6] obtained 432 different mRNAs specifically associated with the mouse FMRP-RNP complexes. Subsequently, comparing polyribosome fractions

from normal human lymphoblastoid cells with those from cells isolated from fragile X patients on a microarray containing 35,000 human genes or ESTs, they identified 251 human mRNAs that show abnormal polyribosome profiles in fragile X patients. The set of genes whose mRNAs were found in the mRNP complexes in mRNAs from total mouse brain were compared with the set of human genes that gave an abnormal polyribosomal profile in the absence of the fragile X protein, using human orthologs of the mouse genes and the murine orthologs of the human genes, respectively. Of the genes found to be associated with the murine FMRP-RNP complexes, 14 were also identified in the abnormal polyribosomal fraction in the human study [6]. A G-quartet structure was identified in 8 of these 14 genes. Thus, FMRP within cells binds preferentially to mRNAs that contain G-quartet structures.

mRNA targets of the fragile X protein

Identifying the mRNA targets of FMRP in the cell may help us understand the clinical consequences of the absence of FMRP. Only around 0.01% of the sequences present in the UniGene database are predicted to have a G-quartet structure; thus only a minor fraction of all proteins may be transported through the cell by FMRP [4]. Many of the predicted G-quartet-containing genes identified have a role in synaptogenesis, a process that is impaired in fragile X syndrome [23]. The RNA that was the most enriched in the FMRP-RNP particles isolated from normal mouse brain, when compared with particles from fragile X mouse brain, codes for a guanine-nucleotide exchange factor related to the yeast secretion protein Sec-7 and is predicted to have a G-quartet structure [6]. Several other members of GTPase pathways, including genes encoding oligophrenin-1 (a Rho-GTPase-activating protein), αGDI (a guanine nucleotide dissociation inhibitor) and the guanine-nucleotide exchange factor ARHGEF6 have been found mutated in patients with non-specific mental retardation [24].

Of special interest are genes that have been isolated in more than one study. Examples of genes that were identified in the abnormal polyribosomal fractions in fragile X patients and that were present in the murine FMRP-RNP complexes are those encoding Munc13, a phorbol ester receptor involved in the maturation of Golgi vesicles and vesicle transport in neurons, and SAPAP4, which is involved in neuronal cell signaling [6]. From the list of 14 mRNAs containing a G quartet that bind FMRP, as identified by Darnell *et al.* [4], 5 were also identified in the microarray studies by Brown *et al.* [6], including those encoding NAP22, which is present in axon terminals and dendritic spines and plays a role in maturation or maintenance of synapses, and the microtubule-associated

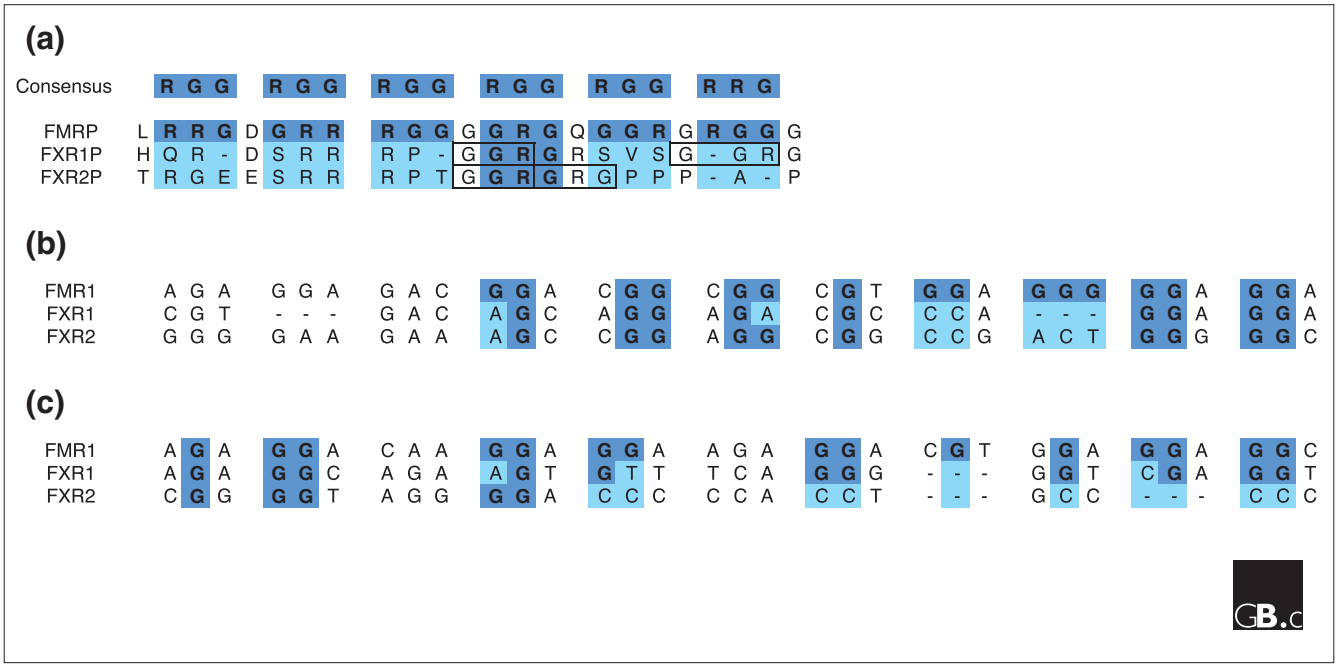

Figure 2
Comparison of consensus sequences in the human members of the fragile X gene family. **(a)** Protein sequence alignment of the RGG boxes of FMRP, FXR1P and FXR2P. The consensus sequence is indicated in dark blue and differences are highlighted in light blue. We used the RGG-box consensus motif determined by Burd and Dreyfuss [18] and the alignment of the three proteins generated by Zhang *et al.* [11]. Other possible RGG motifs for FXR1P and FXR2P are boxed. **(b, c)** Transcript sequence comparison of the purine-rich region between *FMR1* and the autosomal paralogs *FXR1* and *FXR2*. Guanines conserved between human, mouse and rat *FMR1* genes are indicated in dark blue; some of these are not conserved in *FXR1* and *FXR2* (highlighted in light blue). The sequences in (c) continue from (b).

protein MAP1B, which is involved in transport within neurons and is the human ortholog of the *Drosophila futsch* gene [25]. *Drosophila* has only one homolog (*dFXR)* of the fragile X gene family (consisting of *FMR1*, *FXR1* and *FXR2*). dFXR acts as a translational repressor of *futsch*, thereby regulating the synaptic microtubule cytoskeleton. Association of dFXR with Futsch protein *in vivo* was demonstrated by immunoprecipitation. Flies with dFXR mutations display, among other symptoms, enlarged synaptic terminals, whereas double mutants deficient in both dFXR and Futsch have a normal phenotype, demonstrating the importance of the dFXR-Futsch interaction in flies. It remains to be resolved whether each of the identified FMR1 target genes is responsible for a small effect, or whether a few genes will turn out to be responsible for the majority of clinical symptoms of fragile X patients.

Differences within the fragile X gene family

The identification of the RGG box as the site of RNA binding and the role of secondary RNA structures could perhaps explain some of the differences in function between *FMR1* and its autosomal paralogs *FXR1* and *FXR2*. We compared the RGG boxes of FMRP, FXR1P and FXR2P, and found that, although overall the sequences are reasonably well conserved, FMRP has six repeated blocks of the consensus motif RGG, whereas FXR1P and FXR2P each have only a single RGG motif (Figure 2a). FXR1P and FXR2P may have one additional RGG motif each when not aligned with FMRP (boxed sequences in Figure 2a) and FXR1P has two more motifs downstream of the RGG-box consensus sequence. Although the RGG-box sequence is poorly defined and the number of RGG-repeats necessary for RNA-protein interaction is not known, it seems unlikely that the RGG boxes of FMRP, FXR1P and FXR2P have identical RNA-binding capacities.

We also compared the G-quartet sequences of *FMR1*, *FXR1* and *FXR2* mRNAs (Figure 2b,c). Of the 31 guanines conserved between *FMR1* genes in human, mouse and rat RNA [5], only 20 were detectable in *FXR1* and 15 in *FXR2*. In *FXR2* in particular, however, extra guanines at other positions may compensate for the absence of these consensus guanines. It therefore remains to be verified whether *FXR1* and *FXR2* mRNAs are able to form a G-quartet structure similar to that of *FMR1* mRNA. Given the importance for the function of FMRP of interactions between the RGG box and the G-quartet structure, the differences in RGG-box sequence between FMRP, FXR1P and FXR2P could perhaps explain why FXR1P and FXR2P, which, unlike FMRP, are present in the mRNPs of fragile X patients, cannot fully compensate for the absence of the fragile X protein FMRP. Further characterization of the mRNA targets of the fragile X protein and possibly the identification of the mRNAs bound by FXR1P and FXR2P will aid our understanding of the complex mechanisms underlying fragile X syndrome.

Acknowledgements

Financial support for the fragile X research in Antwerp was obtained through grants of the Belgian National Fund for Scientific Research-Flanders (FWO), an Interuniversity Attaction Pole (IUAP-V) and the Fragile X Research Foundation (FRAXA).

References

1. Kooy RF, Willemsen R, Oostra BA: **Fragile X syndrome at the turn of the century.** *Mol Med Today* 2000, **6:**193-199.
2. Hagerman RJ : **Physical and behavioral phenotype.** In *Fragile X Syndrome: Diagnosis, Treatment, and Research.* Edited by Hagerman RJ, Cronister A. Baltimore: Johns Hopkins University; 1996: 3-87.
3. Verkerk AJMH, Pieretti M, Sutcliffe JS, Fu Y-H, Kuhl DPA, Pizzuti A, Reiner O, Richards S, Victoria MF, Zhang F, *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* 1991, **65:**905-914.
4. Darnell J, Jensen K, Jin P, Brown V, Warren ST, Darnell RB: **Fragile X mental retardation protein targets G quartet mRNAs important for neuronal function.** *Cell* 2001, **107:**489-499.
5. Schaeffer C, Bardoni B, Mandel JL, Ehresmann B, Ehresmann C, Moine H: **The fragile X mental retardation protein binds specifically to its mRNA via a purine quartet motif.** *EMBO J* 2001, **20:**4803-4813.
6. Brown V, Jin P, Ceman S, Darnell J, O'Donnell WT, Tenenbaum SA, Jin X, Feng Y, Wilkinson KD, Keene JD, *et al.*: **Microarray identification of FMRP-associated brain mRNAs and altered mRNA translational profiles in fragile X syndrome.** *Cell* 2001, **107:**477-487.
7. Feng Y, Gutekunst C-A, Eberhart DE, Yi H, Warren ST, Hersch SM: **Fragile X mental retardation protein: nucleocytoplasmic shuttling and association with somatodendritic ribosomes.** *J Neurosci* 1997, **17:**1539-1547.
8. Ashley CT, Wilkinson KD, Reines D, Warren ST: **FMR-1 protein: conserved RNP family domains and selective RNA binding.** *Science* 1993, **262:**563-566.
9. Ceman S, Brown V, Warren ST: **Isolation of an FMRP-associated messenger ribonucleoprotein particle and identification of nucleolin and the fragile X-related proteins as components of the complex.** *Mol Cell Biol* 1999, **19:**7925-7932.
10. Willemsen R, Bontekoe C, Tamanini F, Galjaard H, Hoogeveen A, Oostra BA: **Association of FMRP with ribosomal precursor particles in the nucleolus.** *Biochem Biophys Res Commun* 1996, **225:**27-33.
11. Zhang Y, O'Connor JP, Siomi MC, Srinivasan S, Dutra A, Nussbaum RL, Dreyfuss G: **The fragile X mental retardation syndrome protein interacts with novel homologs FXR1 and FXR2.** *EMBO J* 1995, **14:**5358-5366.
12. Tamanini F, Bontekoe C, Bakker CE, van Unen L, Anar B, Willemsen R, Yoshida M, Galjaard H, Oostra BA, Hoogeveen AT: **Different targets for the fragile X-related proteins revealed by their distinct nuclear localizations.** *Hum Mol Genet* 1999, **8:**863-869.
13. Eberhart DE, Malter HE, Feng Y, Warren ST: **The fragile X mental retardation protein is a ribonucleoprotein containing both nuclear localization and nuclear export signals.** *Hum Mol Genet* 1996, **5:**1083-1091.
14. De Boulle K, Verkerk AJMH, Reyniers E, Vits L, Hendrickx J, van Roy B, van den Bos F, de Graaff E, Oostra BA, Willems PJ: **A point mutation in the FMR1 gene associated with fragile X mental retardation.** *Nat Genet* 1993, **3:**31-35.
15. Tamanini F, Meijer N, Verheij C, Willems PJ, Galjaard H, Oostra BA, Hoogeveen AT: **FMRP is associated to the ribosomes via RNA.** *Hum Mol Genet* 1996, **5:**809-813.
16. Laggerbauer B, Ostareck D, Keidel E-M, Ostareck-Lederer A, Fischer U: **Evidence that fragile X mental retardation protein is a negative regulator of translation.** *Hum Mol Genet* 2001, **10:**329-338.
17. Zhenzhong L, Zhang Y, Ku L, Wilkinson KD, Warren ST, Feng Y: **The fragile X mental retardation protein inhibits translation via interacting with mRNA.** *Nucleic Acids Res* 2001, **29:**2276-2283.
18. Burd CG, Dreyfuss G: **Conserved functions and diversity of functions of RNA binding proteins.** *Science* 1994, **265:**615-621.
19. Brown V, Small K, Lakkis L, Feng Y, Gunter C, Wilkinson KD, Warren ST: **Purified recombinant Fmrp exhibits selective**

RNA binding as an intrinsic property of the fragile X mental retardation protein. *J Biol Chem* 1998, **273:**15521-15527.
20. Williamson JR, Raghuraman MK, Cech TR: **Monovalent cation-induced structure of telomeric DNA: the G-quartet model.** *Cell* 1989, **59:**871-880.
21. **UniGene** [http://www.ncbi.nlm.nih.gov/UniGene/index.html]
22. **TIGR Gene Indices** [http://www.tigr.org/tdb/tgi.shtml]
23. Irwin SA, Patel B, Idupulapati M, Harris JB, Crisostomo RA, Larsen BP, Kooy F, Willems PJ, Cras P, Kozlowski PB, *et al.*: **Abnormal dendritic spine characteristics in the temporal and visual cortices of patients with fragile-X syndrome: a quantitative examination.** *Am J Med Genet* 2001, **98:**161-167.
24. Antonarakis SE, Van Aelst L: **Mind the GAP, Rho, Rab and GDI.** *Nat Genet* 1998, **19:**106-108.
25. Zhang YQ, Bailey AM, Matthies H-JG, Renden RB, Smith MA, Speese SD, Rubin GM, Broadie K: ***Drosophila* fragile X-related gene regulates the MAP1B homolog futsch to control synaptic structure and function.** *Cell* 2001, **107:**591-603.