RAN Translation: Fragile X in the Running

Kaalak Reddy[1,2] and Christopher E. Pearson[1,2,*]
[1]Program of Genetics & Genome Biology, The Hospital for Sick Children, Toronto, ON M5G 1L7, Canada
[2]Department of Molecular Genetics, University of Toronto, Toronto, ON M5G 1L7, Canada
*Correspondence: cepearson.sickkids@gmail.com
http://dx.doi.org/10.1016/j.neuron.2013.04.034

In this issue of *Neuron*, Todd et al. (2013) reveal that noncanonical repeat associated non-AUG (RAN) translation occurs on nonexpanded (CGG)30–50 and premutation (CGG)59–160 repeats, associated with the *FMR1* gene, suggesting that the polyglycine and polyalanine products might have natural and pathogenic roles.

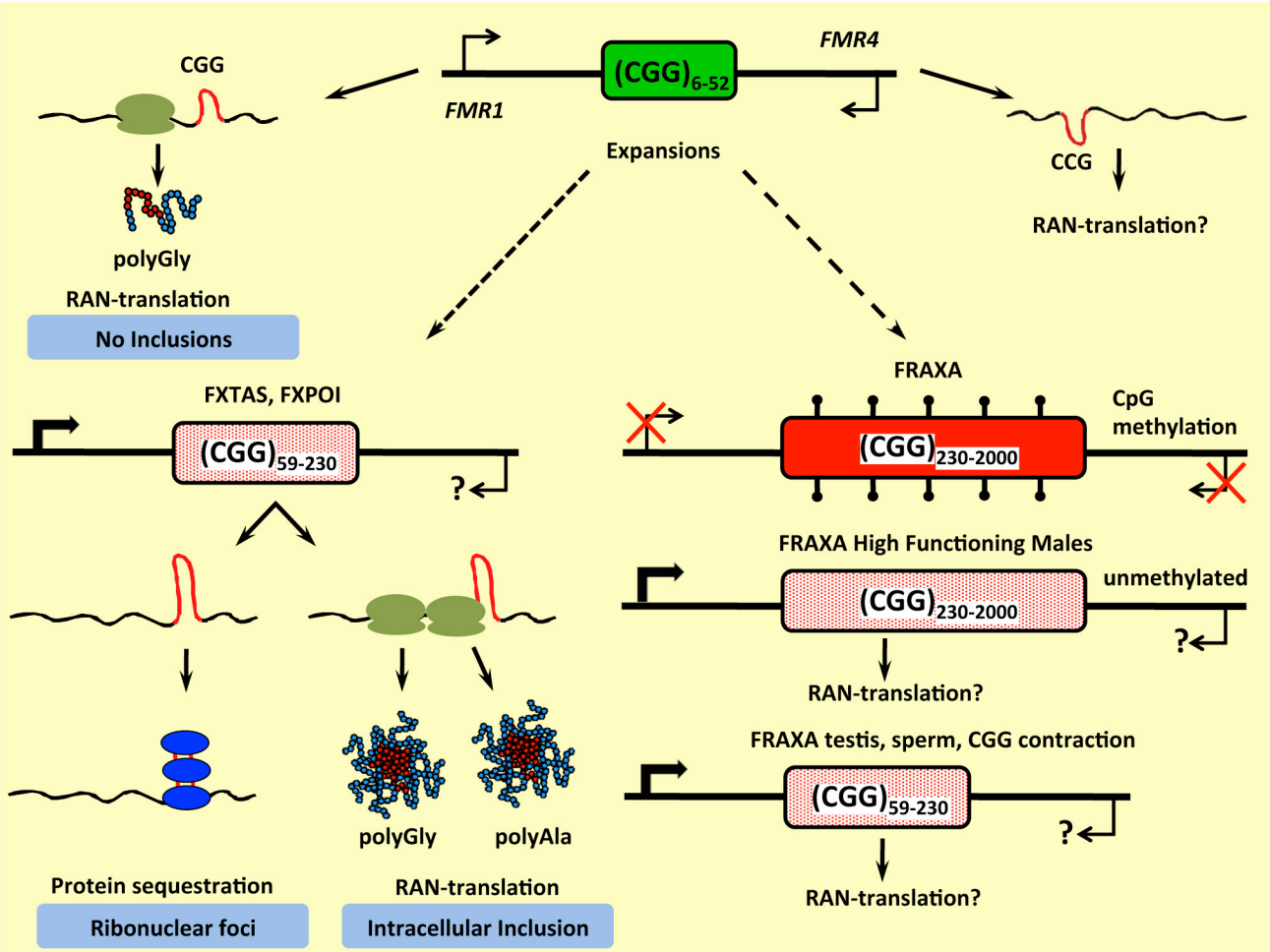
> "Well, in our country," said Alice, still panting a little, "you'd generally get to somewhere else—if you run very fast for a long time, as we've been doing."
>
> "A slow sort of country!" said the Queen. "Now, here, you see, it takes all the running you can do, to keep in the same place. If you want to get somewhere else, you must run at least twice as fast as that!"
>
> —Lewis Carroll, *Through the Looking-Glass*

Typically, DNA is transcribed to make an RNA, which is translated to make a protein. This translation almost always requires a signal that says, "Translation begins here." This signal is an ATG motif. Moreover, typically one RNA encodes only one protein. In this issue of *Neuron*, Todd et al. (2013) reveal that the noncanonical process of repeat associated non-AUG (RAN) translation occurs on the CGG repeat tract present in the 5′ UTR of *FMR1*. Expanded CGG repeats in *FMR1* have been associated with three diseases including fragile X mental retardation type A (FRAXA), fragile X-associated premature ovarian insufficiency (FXPOI), and fragile X-associated tremor ataxia syndrome (FXTAS) (Oostra and Willemsen, 2009; López Castel et al., 2010) (Figure 1). In unaffected individuals, the repeat is (CGG)6–52, in premutation individuals it is (CGG)59–200, and in FRAXA individuals it is (CGG)230–2,000 and aberrantly methylated at CpG sites coincident with loss of transcription and FMRP protein. Individuals with premutation lengths show increased levels of the *FMR1* transcript and the CCG-containing *FMR4* transcript from the strand complementary to *FMR1* (Pearson et al., 2005; Oostra and Willemsen, 2009; López Castel et al., 2010) and can present with FXTAS or FXPOI, both of which are thought to be caused by toxic RNAs, protein binding, and sequestration, similar to myotonic dystrophy (Jin et al., 2003; Lu et al., 2012). The findings of Todd et al. (2013) now reveal a novel pathogenic path for *FMR1*-related diseases.

RAN translation was recently demonstrated by the Ranum laboratory to be a novel form of translation in which there is an absence of an initiating AUG motif and initiation begins within all reading frames of an expanded (CAG)N repeat (Zu et al., 2011; Pearson, 2011). Such events occurred in tissues of myotonic dystrophy type 1 (DM1) and spinocerebellar ataxia type 8 (SCA8) patients. Similar findings were recently reported for the *C9orf72* (GGGGCC)N repeat in tissues of amyotrophic lateral sclerosis and frontotemoral dementia patients, yielding a series of dipeptide repeats (Mori et al., 2013; Ash et al., 2013). The excitement of this newly discovered RAN translation has kept many scientists running, much as Alice did to keep up with the Queen in Lewis Carroll's *Through the Looking-Glass*. Todd et al. (2013) now demonstrate similar but not identical phenomena at the *FMR1* CGG tract, whereby translation of the repeat occurred in the …C GGC GGC GG… or the …CG GCG GCG G… reading frames, yielding peptides of polyglycine or polyalanine. Polyarginine from the …CGG CGG CGG… reading frame was undetectable.

Functions and Pathogenesis of CGG RAN Translation

What is the role, if any, of RAN-translated products in pathogenesis and can this be distinguished from the pathogenic contribution of the CGG RNA? Separating the individual pathogenic contributions of the expanded RNA from that of RAN-translated peptides is a tough challenge, as expression of one without the other is limited by the codons encoding the homopolymers polyGly, polyAla, and polyArg, all of whose codons are redundant enough to have some repetitive nature. Identification of a factor required for RAN translation that is not required for RNA toxicity would ideally separate the contributions of each. Todd et al. (2013) show that transfection of a CGG repeat into human cells reduced cell viability that depended upon both repeat length and translation levels. In *Drosophila*, RAN translation of an expanded CGG RNA caused retinal degeneration. A polyGly-mediated effect was supported by enhancing or suppressing protein quality control pathways, which modulated retinal degeneration in (CGG)90 GFP-expressing flies. Thus, both toxic CGG RNA and toxic RAN proteins may contribute to disease pathogenesis. An interesting clue came from the analysis of two mouse models of FXTAS with slightly different constructs. One line, in which premutation CGG repeats integrated into the 5′ UTR of the mouse fmr1, showed more polyGly inclusions. High levels of inclusions were present in the CNS, whereas inclusions were less frequent in a second knockin mouse that had retained a TAA stop codon just 5′ of the integrated CGG tract, which is predicted to reduce RAN translation.



			(repeat unit)n=					
Disease / locus	Gene	Repeat	Normal	Premutation	Expanded / Disease	Bidirectional transcription	Pathogenesis	RAN-translation
DM1: myotonic dystrophy type 1	*DMPK*	(CTG)•(CAG)	5-37	34-90		Yes	GOF RNA, RAN?	Yes
DM1: myotonic dystrophy type 1	*DM1-AS*	(CAG)•(CTG)	5-37	34-90	90-6500	Yes	?, RAN?	Yes
SCA8: spinocerebellar ataxia type 8	*ATXN8*	(CAG)•(CTG)	2-130	45-109	>110	Yes	GOF, RAN?	Yes
SCA8: spinocerebellar ataxia type 8	*ATXN8OS*	(CTG)•(CAG)	2-130	45-109	>110	Yes	RNA, RAN?	Yes
ALS-FTD: amyotrophic lateral sclerosis-frontotemporal dementia	*C9orf72*	(GGGGCC)n•(GGCCCC)n	2-22	-	23-4400	Yes	GOF RNA?LOF RNA? Ran?	Yes
FRAXA: fragile X syndrome	*FMR1*	(CGG)n•(CCG)n	6-52	59-230	230-2000	Yes	LOF, RNA, LOF-RAN?	?
FXTAS: fragile X tremor/ataxia syndrome	*FMR1*	(CGG)n•(CCG)n	6-52	59-230	59-230	Yes	GOF, RNA, RAN?	Yes
FXPOI: fragile X associated premature ovarian failure	*FMR1*	(CGG)n•(CCG)n	6-52	59-230	59-230	Yes	GOF, RNA, RAN?	?
fragile X A locus (FRAXA?, FXTAS?, FXPOI?)	*FMR4*	(CGG)n•(CCG)n	6-52	59-230	59-230	Yes	?, RAN?	?
FRAXE:fragile X syndrome	*FMR2*	(CCG)n•(CGG)n	4-39	31-61	200-900	?	LOF	?
FRAXF: no confirmed disease association	*FAM11A*	(CGG)n•(CCG)n	7-40	ND	306-1008	?	ND	?
FRA10A: no confirmed disease association	*FRA10AC1*	(CCG)n•(CGG)n	8-14	ND	>100	?	ND	?
FRA11B: Jacobsen syndrome	*CBL2*	(CCG)n•(CGG)n	11	80	100-1000	No	LOF, RAN?	?
FRA16A: no confirmed disease association	-	(CCG)n•(CGG)n	16-49	ND	1000-1900	-	ND	?

Figure 1. RAN Translation in CGG Repeats
RAN translation occurs in nonexpanded and expanded premutation CGG repeats associated with FXTAS, FXPOI, and the testis and sperm of FRAXA males. Table: RAN-translatable disease-associated genes and rare fragile sites whose expanded repeats may undergo RAN translation.

This stop codon was not present in the former mouse model. Since nuclear inclusions containing RAN-translated products were only evident in human FXTAS brains and not controls without CGG expansions, Todd et al. (2013) propose that, although RAN translation occurs in both cases, only the expanded CGG repeat leads to protein aggregation and inclusion formation. Identifying age-matched asymptomatic individuals positive for premutation lengths but negative for RAN-peptide inclusions would further support the proposal of Todd et al. (2013) that inclusions are pathogenic to FXTAS. Evidence for or against the pathogenic contribution of RAN inclusions in any of the repeat diseases is needed.

Interestingly, RAN translation arose from nonexpanded lengths of r(CGG)30 repeats. This leads to two important questions. First, what roles do RAN translation products have, if any, in the unaffected population? Such nonpathogenic roles

might be the reason that CGG repeats were evolutionarily retained, even though they can be unstable and lead to disease. Second, what is the clinical effect when RAN-translation products are lost? Fragile X syndrome has long been thought to be due to a loss of the FMRP protein, mediated by loss of *FMR1* transcription, which would also lead to loss of RAN-translation products. A unique female hemizygous for the FRAXA locus due to a large deletion encompassing *FMR1* of one X chromosome and on the other X chromosome a microdeletion in *FMR1* including all CGG repeats plus some flanking sequences, leaving transcription and translation start sites intact, displayed FMRP expression and a non-FRAXA presentation (Grønskov et al., 1997). This individual would be devoid of *FMR1* RAN-translation products. Whether the other symptoms of the patient (hearing problems and some deficits in perceptual spatial skills) were the result of the loss of *FMR1* RAN-translation products or the very large deletion on the other X chromosome is not known. A FRAXA individual without CGG expansions but having an FMRP-inactivating point mutation, I304N, that still produces a (CGG)25-containing mRNA and presumably RAN-translation products, argues that there is little phenotypic contribution of the loss of RAN translation to the typical FRAXA presentation (Oostra and Willemsen, 2009; Pearson et al., 2005). The natural functions of RAN-translation products deserve attention (Pearson, 2011).

Mechanism(s) of CGG RAN Translation

Using *Drosophila*, CGG mouse models, human cell lines, and patient tissues, Todd et al. (2013) observed RAN translation on both nonexpanded and disease-associated premutation CGG lengths in two of the three possible reading frames (GGC, polyGly and GCG, polyAla) involving distinct RAN-translation mechanisms for each frame. RAN translation leading to polyGly required an upstream AUG-like codon (such as GUG) and was sensitive to stop codons. PolyAla formation arose independent of AUG-like codons, was insensitive to stop codons, and was repeat length dependent. Thus, there are multiple mechanisms governing CGG RAN translation. Placement of a stop codon at −12 bp, but not at −21 bp, upstream of the CGG blocked RAN translation of polyGly, suggesting that RAN translation initiates between 21 and 12 bases 5′ of the CGG repeat (Figure 1). Interestingly, this same region of the *FMR1* transcript had previously been identified to function as an internal ribosome entry site (IRES) and a site of ribosome pausing (Chiang et al., 2001). Similar to the reduced translation of FMRP from the premutation *FMR1* CGG repeat compared to the nonexpanded sizes (Pearson et al., 2005; Oostra and Willemsen, 2009), RAN translation in the polyGly frame may also be decreased with longer tracts (Todd et al., 2013).

RAN translation of (CAG)N repeats required the formation of an intrastrand hairpin in the r(CAG)n RNA (Zu et al., 2011), where RNA structure may impact initiation of RAN translation (Pearson, 2011). *FMR1* r(CGG)19–47 repeats were shown to form hairpins, complex multibranched or G-quadruplex structures (Napierala et al., 2005; Khateb et al., 2007). Interestingly, the *C9orf72* repeat, which also undergoes RAN translation (Mori et al., 2013; Ash et al., 2013), was recently demonstrated to form G-quadruplexes (Reddy et al., 2013). The precise RNA structures and contexts (AUG-like motifs) involved in RAN translation of each repeat sequence, and possibly each reading frame, needs to be elucidated.

Where Else Might CGG RAN Translation Occur?

Presumably, CGG-mediated RAN translation and the production of toxic polypeptides/inclusions may form wherever an expanded CGG repeat-containing transcript is produced (Figure 1). The clinical presentation of *FMR1* mutations is modulated by CGG tract length and aberrant CpG methylation. FRAXA patients exhibit nearly complete methylation of their expanded CGG tract, whereas rare individuals classified as "high-functioning" males or "methylation mosaics" possess unmethylated CGG repeats expanded to disease lengths and display *FMR1* transcription and FMRP protein and are not affected (Pearson et al., 2005). Albeit, males with fully expanded and methylated repeats do show some *FMR1* transcript (Pearson et al., 2005; Oostra and Willemsen, 2009). Furthermore, unmethylated premutation (CGG)59–230 lengths, which are both transcribed and translated to FMRP, exist in the testes and sperm of males with full-mutation FRAXA, even though all other tissues display methylated disease-length alleles of (CGG)230–2,000 (Pearson et al., 2005). In all of the above instances, the *FMR1* and *FMR4* transcripts with fully expanded or premutant CGG and CCG tracts may be templates for RAN translation and their potentially toxic products. Other fragile site genes, each associated with expanded CGG repeats and their CCG-containing antisense transcripts, may also undergo RAN translation (Figure 1), as might any of the disease-associated expanded CAG/CTG repeats (Pearson, 2011).

Does RAN translation occur in all premutations, FXPOI tissues, high-functioning methylation mosaics with fully expanded CGG repeats, or is it limited to FXTAS patients (Figure 1)? Both unmethylated full mutation and premutation alleles show increased *FMR1* transcript levels (Pearson et al., 2005; Oostra and Willemsen, 2009). Might RAN translation of the CGG tract be sensitive to repeat length, in that premutation expansions associated with FXTAS, but not fully expanded (CGG) 230–2,000 tracts, are susceptible to this phenomenon? Might the very large CGG tracts be recalcitrant to RAN translation, as premutation lengths are recalcitrant to FMRP translation (Pearson et al., 2005; Oostra and Willemsen, 2009)?

That RAN translation occurs on CGG repeats of nonexpanded and premutation length transcripts broadens our understanding of both the naturally occurring *FMR1* gene products as well as the potential toxic entities that may be involved in FXTAS, FXPOI, and FRAXA. This adds *FMR1* to the growing list of disease-associated repeats that undergo RAN translation and those that might be (Figure 1), with some evidence of a pathogenic role, and may be common to the 40 repeat-associated diseases.

REFERENCES

Ash, P.E., Bieniek, K.F., Gendron, T.F., Caulfield, T., Lin, W.L., Dejesus-Hernandez, M., van Blitterswijk, M.M., Jansen-West, K., Paul, J.W., 3rd, Rademakers, R., et al. (2013). Neuron *77*, 639–646.

Chiang, P.W., Carpenter, L.E., and Hagerman, P.J. (2001). J. Biol. Chem. *276*, 37916–37921.

Grønskov, K., Hjalgrim, H., Bjerager, M.O., and Brøndum-Nielsen, K. (1997). Am. J. Hum. Genet. *61*, 961–967.

Jin, P., Zarnescu, D.C., Zhang, F., Pearson, C.E., Lucchesi, J.C., Moses, K., and Warren, S.T. (2003). Neuron *39*, 739–747.

Khateb, S., Weisman-Shomer, P., Hershco-Shani, I., Ludwig, A.L., and Fry, M. (2007). Nucleic Acids Res. *35*, 5775–5788.

López Castel, A., Cleary, J.D., and Pearson, C.E. (2010). Nat. Rev. Mol. Cell Biol. *11*, 165–170.

Lu, C., Lin, L., Tan, H., Wu, H., Sherman, S.L., Gao, F., Jin, P., and Chen, D. (2012). Hum. Mol. Genet. *21*, 5039–5047.

Mori, K., Weng, S.M., Arzberger, T., May, S., Rentzsch, K., Kremmer, E., Schmid, B., Kretzschmar, H.A., Cruts, M., Van Broeckhoven, C., et al. (2013). Science *339*, 1335–1338.

Napierala, M., Michalowski, D., de Mezer, M., and Krzyzosiak, W.J. (2005). Nucleic Acids Res. *33*, 451–463.

Oostra, B.A., and Willemsen, R. (2009). Biochim. Biophys. Acta *1790*, 467–477.

Pearson, C.E. (2011). PLoS Genet. *7*, e1002018.

Pearson, C.E., Nichol Edamura, K., and Cleary, J.D. (2005). Nat. Rev. Genet. *6*, 729–742.

Reddy, K., Zamiri, B., Stanley, S.Y., Macgregor, R.B., Jr., and Pearson, C.E. (2013). J. Biol. Chem. *288*, 9860–9866.

Todd, P.K., Oh, S.Y., Krans, A., He, F., Sellier, C., Frazer, M., Renoux, A.J., Chen, K.C., Scaglione, K.M., Basrur, V., et al. (2013). Neuron *78*, this issue, 440–455.

Zu, T., Gibbens, B., Doty, N.S., Gomes-Pereira, M., Huguet, A., Stone, M.D., Margolis, J., Peterson, M., Markowski, T.W., Ingram, M.A., et al. (2011). Proc. Natl. Acad. Sci. USA *108*, 260–265.

Evidence for a Common Endocannabinoid-Related Pathomechanism in Autism Spectrum Disorders

Dilja D. Krueger[1,*] and Nils Brose[1,*]
[1]Department of Molecular Neurobiology, Max Planck Institute of Experimental Medicine, 37075 Göttingen, Germany
*Correspondence: krueger@em.mpg.de (D.D.K.), brose@em.mpg.de (N.B.)
http://dx.doi.org/10.1016/j.neuron.2013.04.030

In this issue of *Neuron*, Földy et al. (2013) report that endocannabinoid-mediated signaling at inhibitory synapses is dysregulated in mouse models of autism-associated Neuroligin-3 mutations. These findings carry implications regarding the pathophysiology of autism spectrum disorders and the development of treatment strategies.

The correct wiring of the brain during development is an extremely complex biological process, during which a staggering number of synapses with often very diverse characteristics have to be formed and maintained in a precise and delicate balance. Not surprisingly, therefore, numerous neurodevelopmental and psychiatric diseases appear to be disorders of aberrant synaptogenesis and synapse function, or "synaptopathies." Particularly in the context of autism spectrum disorders (ASDs), an ever-growing number of mutations in genes encoding synaptic proteins have been identified in affected individuals (Murdoch and State, 2013), and major research efforts are currently focusing on strategies to transform this knowledge base into viable treatment strategies.

However, the corresponding challenges are substantial. For example, very little is known about the role of ASD-related synaptic proteins in vivo, e.g., in neuronal circuits that control autism-relevant behavior. Second, many known ASD-related proteins are structural proteins with adhesion or scaffold functions and therefore poor targets for pharmacological intervention with small molecule drugs. Third, many ASD-related mutations lead to a loss of the corresponding protein so that no target for pharmacological intervention remains. Finally, each individual ASD-related mutation is rare, with the vast majority accounting for less than 1% of affected individuals each. In view of these difficulties, the focus in the field of ASD biology has shifted toward the identification of cellular protein-protein interactions or signaling pathways that are common to the various ASD-related proteins and therefore expected to be perturbed by a wide range of ASD-related mutations—with the hope that such pathways may represent more promising treatment targets than the ASD-linked proteins discovered so far.

One of the synaptic proteins associated with ASDs is Neuroligin-3 (NLGN3), a member of the Neuroligin family of postsynaptic cell adhesion molecules that interact with presynaptic Neurexins to control synapse development and function. Two distinct mutations in *NLGN3* have been linked to ASDs, a point mutation resulting in an R451C substitution in the Neurexin-binding domain (Jamain et al., 2003) and a deletion of the *NLGN3* gene (Sanders et al., 2011). Studies on the respective mouse models, a $Nlgn3^{R451C}$ knockin (KI) and a Nlgn3 knockout (KO), showed that both mutations cause ASD-related behavioral phenotypes (Radyushkin et al., 2009; Tabuchi et al., 2007) but have strikingly different effects on synapse and network function, with the $Nlgn3^{R451C}$ mutation resulting in a gain-of-function phenotype that is