Autism-Misregulated eIF4G Microexons Control Synaptic Translation and Higher Order Cognitive Functions

Highlights
- Autism-disrupted eIF4G microexons mediate activity-dependent responses
- eIF4G microexons suppress the expression of critical synaptic proteins
- eIF4G microexon-deficient mice display social behavior and memory deficits
- eIF4G microexons coalesce with neuronal granule components and stall ribosomes

Authors
Thomas Gonatopoulos-Pournatzis, Rieko Niibori, Eric W. Salter, ..., Melanie A. Woodin, Sabine P. Cordes, Benjamin J. Blencowe

Correspondence
t.gonatopoulos@gmail.com (T.G.-P.), cordes@lunenfeld.ca (S.P.C.), b.blencowe@utoronto.ca (B.J.B.)

In Brief
Gonatopoulos-Pournatzis et al. demonstrate that conserved, autism-disrupted microexons in eIF4G translation initiation factors regulate the neuronal proteome and control higher order cognitive functions. The microexons function as a translational brake and elicit ribosome stalling on transcripts encoding synaptic proteins through their propensity to coalesce with FMRP and other neuronal granule components.
Autism-Misregulated eIF4G Microexons Control Synaptic Translation and Higher Order Cognitive Functions

Thomas Gonatopoulos-Pournatzis,1,* Rieko Niibori,2,16 Eric W. Salter,2,3,16 Robert J. Weatheritt,1,4,5,16 Brian Tsang,6,7,16 Shaghayegh Farhangmehr,1,6,16 Xinyi Liang,9 Ulrich Braunschweig,1 Jonathan Roth,1,2,8 Shen Zhang,2 Tyler Henderson,2,8 Eesha Sharma,1,8 Mathieu Quesnel-Vallières,1,8 Jon Permanyer,10 Stefan Maier,2 John Georgiou,2 Manuel Irimia,10,11,12 Nahum Sonenberg,13,14 Julie D. Forman-Kay,6,7 Anne-Claude Gingras,2,8 Graham L. Collingridge,2,3,15 Melanie A. Woodin,8 Sabine P. Cordes,2,8,16 and Benjamin J. Blencowe1,8,17,*

1Donnelly Centre, University of Toronto, Toronto, ON M5S 3E1, Canada
2Lunenfeld-Tanenbaum Research Institute, Mount Sinai Hospital, Toronto, ON M5G 1X5, Canada
3Department of Physiology, University of Toronto, Toronto, ON M5S 1A8, Canada
4EMBL Australia, Garvan Institute of Medical Research, Darlinghurst, NSW 2010, Australia
5St. Vincent’s Clinical School, University of New South Wales, Sydney, NSW 2010, Australia
6Program in Molecular Medicine, The Hospital for Sick Children, Toronto, ON M5G 0A4, Canada
7Department of Biochemistry, University of Toronto, Toronto, ON M5S 1A8, Canada
8Department of Molecular Genetics, University of Toronto, Toronto, ON M5S 1A8, Canada
9Department of Cell and Systems Biology, University of Toronto, Toronto, ON M5S 3G5, Canada
10Centre for Genomic Regulation, Barcelona 08003, Spain
11Universitat Pompeu Fabra (UPF), Barcelona 08002, Spain
12ICREA, Barcelona 08010, Spain
13Goodman Cancer Research Center, McGill University, Montreal, QC H3A 1A3, Canada
14Department of Biochemistry, McGill University, Montreal, QC H3G 1Y6, Canada
15Tanz Centre for Research in Neurodegenerative Diseases, University of Toronto, Toronto, ON M5S 1A8, Canada
16These authors contributed equally
17Lead Contact
*Correspondence: t.gonatopoulos@gmail.com (T.G.-P.), cordes@lunenfeld.ca (S.P.C.), b.blencowe@utoronto.ca (B.J.B.)
https://doi.org/10.1016/j.molcel.2020.01.006

SUMMARY

Microexons represent the most highly conserved class of alternative splicing, yet their functions are poorly understood. Here, we focus on closely related neuronal microexons overlapping prion-like domains in the translation initiation factors, eIF4G1 and eIF4G3, the splicing of which is activity dependent and frequently disrupted in autism. CRISPR-Cas9 deletion of these microexons selectively upregulates synaptic proteins that control neuronal activity and plasticity and further triggers a gene expression program mirroring that of activated neurons. Mice lacking the Eif4g1 microexon display social behavior, learning, and memory deficits, accompanied by altered hippocampal synaptic plasticity. We provide evidence that the eIF4G microexons function as a translational brake by causing ribosome stalling, through their propensity to promote the coalescence of cytoplasmic granule components associated with translation repression, including the fragile X mental retardation protein FMRP. The results thus reveal an autism-disrupted mechanism by which alternative splicing specializes neuronal translation to control higher order cognitive functioning.

INTRODUCTION

Alternative splicing has critical roles in numerous fundamental biological processes (Baralle and Giudice, 2017; Scotti and Swanson, 2016; Ule and Blencowe, 2019; Vuong et al., 2016). This is especially the case in the mammalian brain, which expresses among the most complex and conserved alternative splicing programs (Barbosa-Morais et al., 2012; Merkin et al., 2012), the disruption of which plays a causative or contributing role to different neurological diseases and disorders (Licatalosi and Darnell, 2010; Nussbacher et al., 2015; Quesnel-Vallières et al., 2019; Sibley et al., 2016). However, only a small number of neural-regulated alternative splicing events have been functionally characterized (Raj and Blencowe, 2015; Ustianenko et al., 2017; Vuong et al., 2016). We previously detected a program of 3–27 nucleotide neuronal microexons that is primarily controlled by the serine/arginine repetitive matrix protein 4 (SRRM4) (also known as nSR100; Calarco et al., 2009; Gonatopoulos-Pournatzis et al., 2018; Irimia et al., 2014; Quesnel-Vallières et al., 2015, 2016). Neuronal microexons are highly conserved at the levels of both sequence and regulatory characteristics. The majority are located within and preserve open reading frames, and they invariably are located on protein surfaces, where they participate in the regulation of protein-protein interactions (Dergai et al., 2010; Irimia et al., 2014; Li et al., 2015; Ohnishi et al., 2014; Toffolo et al., 2014). Approximately one-third of neuronal...
microexons show increased skipping, concordant with decreased levels of expression of SRRM4, in the brains of individuals that primarily have idiopathic forms of autism spectrum disorder (ASD) (Irimia et al., 2014). Mice haploinsufficient for SRRM4 recapitulate microexon misregulation and display multiple ASD-like phenotypes, including altered social behavior, increased sensitivity to environmental stimuli, and altered synaptic spine density and transmission (Quesnel-Vallières et al., 2015, 2016). These observations, and the finding that microexons are enriched in genes with genetic links to ASD (Irimia et al., 2014), have prompted important questions, including: what are their individual roles, which contribute to ASD-associated phenotypes, and what are the mechanisms by which they function?

In addition to altered splicing, misregulation of translation is a recurring mechanism associated with ASD (Bagni and Zukin, 2019; Borrie et al., 2017; Holt and Schuman, 2013; Jung et al., 2014; Richter et al., 2015; Santini and Klann, 2014). In fragile X syndrome (FXS), the most common form of syndromic autism, repeat expansions in the FMR1 gene reduce the translation-repressive activity of its product, FMRP. This results in altered synaptic protein synthesis, plasticity, and ASD-associated phenotypes (Bakker et al., 1994; Damell et al., 2011; Feng et al., 1995; Huber et al., 2002). Tuberous sclerosis, a genetic disorder caused by loss-of-function mutations in TSC1 or TSC2 genes, is frequently characterized by autistic-like phenotypes that can be rescued by pharmacomodulation of mTOR signaling, which controls translation initiation (Auerbach et al., 2011; Borrie et al., 2017; Ehninger et al., 2008). The role of altered translational control in ASD is further supported by the discovery of mutations in PTEN and the cap binding initiation factor eIF4E (Neves-Pereira et al., 2009; Zhou and Parada, 2012), and from the observation of ASD-like phenotypes in mouse models with genetic alterations that result in increased cap-dependent translation in neurons (Gkogkas et al., 2013; Santini et al., 2013). An important question is whether ASD-associated altered splicing patterns, including those involving microexons, might also impact neuronal translation to result in autistic phenotypes. In this regard, a more general question is whether regulated alternative splicing contributes to the control of the translational outputs that underlie neurodevelopmental and behavioral phenotypes.

We have addressed these questions through investigation of previously uncharacterized, ASD-misregulated microexons in the eukaryotic translation initiation factors 4 gamma, eIF4G1 and eIF4G3, paralogs. eIF4G proteins interact with eIF4E and the eIF4A DEAD-box RNA helicase to form a heterotrimeric eIF4F complex required for translation initiation (Gingras et al., 1999). Deletion of the eIF4G microexons specifically upregulates the translation of numerous proteins that control synaptic transmission and neuronal activity, including the obligatory NMDA receptor subunit GluN1, which controls calcium influx in neurons, synaptic plasticity, and memory formation. These changes contribute to an activated neuronal state and strengthen synaptic connectivity. Consistently, mice deficient of the eIF4G1 micro exon show altered social behavior, memory, and learning deficits. Inclusion of the microexons promotes ribosome stalling on translationally repressed synaptic protein transcripts that significantly overlap binding targets of FMRP, whereas their skipping, which is activity-dependent and observed in the brains of ASD subjects, relieves this repression. Consistent with these observations, the eIF4G1 microexon promotes the coalescence of components of neuronal granules associated with translation repression, including FMRP. Collectively, the results demonstrate a critical function for individual alternative microexons in the specialization of neuronal translation and animal behavior. They further suggest that disruption of protein synthesis outputs, controlled by microexon-dependent changes in multivalent interactions with neuronal granule components, may commonly arise in ASD.

RESULTS

A Neuronal Microexon Switch in eIF4G Translation Initiation Factors

To investigate possible roles for alternative splicing in the functional specialization of translation, we performed an analysis of RNA sequencing (RNA-seq) data from ~30 diverse cell and tissue types in human and mouse for regulated splicing events that potentially affect ribosome assembly and core ribosomal proteins. This reveals 227 alternative splicing events, 32 of which are conserved in mammals (Table S1) and several of which display pronounced neuronal-specific differential regulation (Figure 1A). Among the latter events are previously reported microexons in the small ribosomal subunit protein 24 (RPS24) and cytoplasmic polyadenylation element binding protein 4 (CPEB4) (Figures S1A and S1B; Gupta and Warner, 2014; Parras et al., 2018). However, two of the events correspond to previously uncharacterized microexons. These are each predicted to insert seven amino acids with the closely related sequences, GGFRSLQ and GGFRPQ, proximal to the N-termini of eIF4G1 and eIF4G3, respectively (Figure 1B).

The eIF4G microexons individually display conservation at the levels of amino acid sequence and neuronal-differential splicing across vertebrates spanning at least 450 Ma of evolution (Figures 1C and S1B–S1F). In contrast, a third eIF4G paralog, eIF4G2/DAP5/p97, lacks the N-terminal microexon-containing region found in eIF4G1 and eIF4G3. eIF4G1 and eIF4G3 act as scaffolds in the formation of productive initiation complexes (see Introduction; Gingras et al., 1999). Given their high degree of conservation, we hypothesized that differential regulation of the eIF4G1 and eIF4G3 microexons may be important for controlling neuronal-specific translation and consequently brain function.

A Regulatory Network Genetically Associated with Neurological Disorders Controls eIF4G Microexon Splicing

To explore possible roles of the eIF4G microexons, we initially investigated their regulation. Neuronal microexons are frequently controlled by a complex consisting of SRRM4, SRSF11, and RNPS1, whereas RBFOX regulates an overlapping although distinct and smaller subset of microexons (Gonatopoulos-Pournatzis et al., 2018; Li et al., 2015). Consistent with these findings, the expression level of Srrm4 closely tracks increased levels of splicing of both exons during glutamatergic neuronal differentiation (Figures 1C, S1A, and S1B). Moreover, both exons show decreased splicing in the brains of Srrm4-knockout mice.
(Figure S2A), and ectopic expression of Srrm4 is sufficient to activate microexon splicing in mouse embryonic stem cells (mESCs) (Figure 2A). Furthermore, knockdown of endogenous Srrm4, Srrm3 (a paralog of Srrm4), Rnps1, Srsf11, and RBfox2 all result in skipping of the EIF4G microexons in N2A cells, whereas knockdown of Ptbp1, a widely acting repressor of neuronal alternative splicing (Vuong et al., 2016), increases the inclusion levels of these exons (Figures 2B, S2B, and S2C).

To assess whether Srrm4, Srrm3, Rnps1, and Srsf11 directly activate EIF4G1 and EIF4G3 microexon splicing, we performed an individual nucleotide resolution crosslinking and immunoprecipitation coupled to sequencing (iCLIP-seq) analysis of these factors. Srrm4, Srrm3, and Srsf11, but not GFP (analyzed as a negative control), form binding peaks proximally upstream of the EIF4G microexons (Figures 2D and S2E). Furthermore, consistent with its distinct binding map and mechanism of action (Lee et al., 2016; Lovci et al., 2013; Weyn-Vanhentenryck et al., 2014), RBfox forms spatially distinct binding peaks downstream of the EIF4G1 and EIF4G3 microexons (Figures 2D and S2E).

Chromatin regulators, such as Ep300, control Srrm4 expression (Gonatopoulos-Pournatzis et al., 2018). Consistently, its ablation also affects EIF4G1 microexon splicing (Figure S2C). Notably, sequencing studies have identified mutations or risk alleles in SRRM4, SRRM3, SRSF11, RNPS1, RBFOX1, and EP300 that are linked to neurological disorders, including autism and intellectual disability (C.Yuen et al., 2017; Iossifov et al., 2014; Lim et al., 2017; Lucariello et al., 2016; Nguyen et al., 2013; De Rubeis et al., 2014; S.W. Scherer and R. Yuen, personal communication). Collectively, these results provide evidence that a network of chromatin and splicing factors impacted in ASD regulate the splicing of EIF4G microexons.

**EIF4G Microexon Splicing Is Activity Dependent and Disrupted in Autistic Brains**

SRRM4 levels decrease upon induced neuronal activity, leading to the proposal that diverse genetic alterations resulting in increased neuronal activity may frequently disrupt microexon splicing in autistic patients (Quesnel-Vallières et al., 2016). Moreover, RBFOX subcellular localization and thus its splicing function are also regulated by neuronal activity (Lee et al., 2009). Accordingly, we next asked whether neuronal activity impacts the splicing of the EIF4G microexons. Indeed, depolarization of neurons with KCl, or treatment with the ionotropic glutamate receptor agonists kainate and domoic acid, leads to a rapid and marked decrease in splicing of the EIF4G1 and EIF4G3 microexons in a transcription-dependent manner (Figures 2E, S2F, and S2G). These results thus suggest that the EIF4G microexons are dynamically regulated in response to neuronal activity.

Analysis of RNA-seq data reveals disrupted splicing of the EIF4G1 exon in a subset of autistic patients (Figure S2H; Irimia et al., 2014). Expanding this analysis to a larger cohort (58 ASD and matched control subjects) using PsychENCODE RNA-seq (Gandal et al., 2018) confirmed a significant reduction in splicing of the EIF4G1 microexon in ASD compared to control subjects ($p < 3.89 \times 10^{-4}$, Mann-Whitney U test; Figure 2F). In contrast, the splicing levels of the EIF4G3 microexon are not significantly
Disruption of EIF4G1 microexon splicing was confirmed by RT-PCR analysis of a subset of available RNA samples from the RNA-seq profiled ASD and control subjects (Figure 2G). Notably, the EIF4G1 microexon ranks as the third most strongly disrupted of 35 microexons with detected significant reductions in splicing levels in the brains of ASD versus control subjects (Figure S2J). Collectively, these data suggest that disruption of EIF4G1 microexon splicing may contribute to molecular changes and phenotypes associated with autism.

eIF4G Microexons Regulate the Expression of Synaptic Proteins

To investigate whether the eIF4G microexons function to control neuronal translation, CRISPR-Cas9 editing was used to delete both exons in mESCs, which were subsequently differentiated into neurons. Because eIF4G1 and eIF4G3 possess similar biochemical activities (Gingras et al., 1999), double microexon-knockout (DKO) cell lines were generated to eliminate potential compensatory mechanisms. The DKO neurons were subjected to isobaric labeling and quantitative mass spectrometry (Figures 3A, 3B, and S3A–S3D). Relative to control neurons, ~350 proteins (representing ~10% of total detected proteins) display significant changes (false discovery rate [FDR] < 0.01) in steady-state levels in the DKO neurons, with more accentuated differences at a later time point of neuronal differentiation when functional synapses are formed (Figures 3C and S3E; Table S2). Notably, proteins with increased expression in the DKO neurons are significantly enriched in synaptic functions (Figures 3C and 3D; p < 6.18 × 10^{-6}; Fisher’s exact test). These include the obligatory subunit of the ionotropic glutamate NMDA receptor, GluN1 (encoded by Grin1), which is critical for major forms of synaptic plasticity and controls learning and memory (Tsien et al., 1996; Volianskis et al., 2015). Critical postsynaptic components of inhibitory synapses, including Gephyrin and Neuroli-gin-2 (Tyagarajan and Fritschy, 2014), also display increased protein levels in the DKO neurons (Figure S3F; Table S2). Increased expression of these proteins was confirmed by western blotting (Figures 3E and 3F). These results thus reveal that the eIF4G microexons control the expression of numerous synaptic proteins, including NMDA receptors that govern activity-dependent responses underlying learning and memory.

Consistent with deletion of the eIF4G1 and eIF4G3 microexons impacting the expression of the affected proteins at the translational level, RNA-seq profiling of the DKO neurons does not reveal appreciable changes in steady-state levels of the
Figure 3. eIF4G Microexons Control the Expression of Synaptic Receptors Linked to the Regulation of Neuronal Activity

(A) Overview of experimental strategy for the generation of wild-type (WT) or double Eif4g1 and Eif4g3 microexon-deficient (DKO) mESC-derived neurons.

(B) RT-PCR assays monitoring Eif4g1 and Eif4g3 microexon splicing during neuronal differentiation of a WT and DKO CGR8 mESC clonal line.

(C) Proteins differentially expressed between WT and DKO neurons as determined by quantitative mass spectrometry (MS). Three biological replicates of each genotype were analyzed, and the abundance ratio of differentially expressed proteins is shown (left panel). Analysis of Gene Ontology term enrichment among proteins whose expression is repressed by eIF4G microexons (right panel) is shown.

(D) Percentage of eIF4G microexon-regulated proteins that overlap synaptic proteins (Pirooznia et al., 2012) compared to MS-detected proteins not regulated by eIF4G microexons (p < 0.001; Fisher’s exact test).

(E and F) Western blots (E) and quantification of relative expression (F) of GluN1, Gephyrin, and Neuroligin-2 protein levels in WT and DKO neurons. Immuno-blotting of Actin and Gapdh are shown to control for loading and recovery. Error bars, SD; *p < 0.05; **p < 0.01; two-tailed t test.

(G) Log2 mRNA abundance plotted for WT and DKO mESC-derived neurons. Genes with significant differential expression (FDR < 0.05) are indicated by black dots, and representative gene names are shown. Activity-dependent genes are indicated in red.

(legend continued on next page)
corresponding mRNAs (Figure 3G; Table S3). Notably, however, genes with transcript level changes significantly overlap those regulated by neuronal activity, as determined by comparison with gene expression changes detected after KCl depolarization of primary neurons (Figures 3G and 3H; p < 7.6 \times 10^{-5}; Fisher’s exact test). Many of the changes involve increased expression of immediate-early response genes, such as cFos, JunB, Egr1, and Arc (Table S3), which is a hallmark of activity-dependent responses (Ebert and Greenberg, 2013). qRT-PCR assays confirmed six out of seven analyzed transcript level changes detected by RNA-seq (Figures 3I and S3G). Moreover, KCl depolarization of neurons and treatment with glutamate receptor agonists results in an even stronger induction of activity-dependent genes in the DKO neurons than in wild-type neurons (Figures 3J and S3H). Collectively, these data provide evidence that the eIF4G microexons control translation of synaptic proteins, including NMDA receptors, to impact activity-dependent gene expression patterns. Moreover, because the eIF4G microexons themselves are regulated by neuronal activity (Figures 2E, S2F, and S2G), the results further suggest that their splicing levels are integral to the control of activity-dependent responses in neurons.

Deletion of the Eif4g1 Microexon Results in Social Behavior, Learning, and Memory Deficits

To investigate the function of the eIF4G1 and eIF4G3 microexons in vivo, we used CRISPR-Cas9 editing to generate mice with each microexon individually deleted (ΔMIC). We crossed heterozygous Eif4g1+/ΔMIC or Eif4g3+/ΔMIC animals to generate mice homozygously deleted for each microexon (Eif4g1ΔMIC/ΔMIC or Eif4g3ΔMIC/ΔMIC; Figure 4A). PCR and Sanger sequencing confirmed successful deletions (Figure S4A; data not shown). Both homozygous deletion strains are born with expected Mendelian ratios and do not display obvious morphological phenotypes. For the rest of the study, we focused on the Eif4g1ΔMIC/ΔMIC mice, because Eif4g3ΔMIC/ΔMIC mice do not display obvious behavioral phenotypes (data not shown). This may be due to the reduced (~2-fold) levels of Eif4g3 relative to Eif4g1 expression or because eIF4G3 may have a greater degree of functional redundancy than eIF4G1.

Eif4g1ΔMIC/ΔMIC mice have normal weight, motor coordination, locomotion, habituation, anxiety, aggression, hearing, and response to light stimuli (Figures 4B–S4I). However, they display social behavior abnormalities in the three-chamber apparatus. In this apparatus, wild-type mice generally display a preference for unfamiliar mice over an inanimate object (social preference test) or unfamiliar over familiar mice (social novelty test). Eif4g1ΔMIC/ΔMIC mice do not display an appreciable difference with wild-type mice in the social preference test (Figure S4J). However, in contrast to wild-type mice, they lack a social preference in the novelty test (Figures 4B and S4K). Moreover, using the reciprocal interaction test for sociability, which measures the amount of time a mouse spends interacting with an age- and sex-matched mouse of the same genotype, the Eif4g1ΔMIC/ΔMIC mice interact significantly less often than do wild-type mice (p < 0.0384; two-tailed t test; Figure 4C). Taken together with the finding that the Eif4G1 microexon is among the most strongly skipped in the brains of ASD individuals (Figure S2J), these results suggest that its misregulation may contribute to behavioral phenotypes associated with autism in humans.

Western blotting analysis of synaptosomal preparations from the brains of Eif4g1ΔMIC/ΔMIC mice demonstrates that deletion of the Eif4G1 microexon increases the expression of critical synaptic proteins, including GluN1 (Figure 4D), thus confirming in vivo the results from analyzing cultured DKO neurons (Figures 3E and 3F). Given the importance of GluN1 in the control of synaptic plasticity, we investigated whether the Eif4g1ΔMIC/ΔMIC mice display altered learning and memory. Notably, the mutant mice have impaired episodic memory, as measured by the contextual fear-conditioning test, which associates a training (environmental) context with a foot shock stimulus and requires both hippocampal- and amygdala-dependent memory formation (Figure 4E; p < 0.01; two-way ANOVA). However, no appreciable difference is observed between Eif4g1ΔMIC/ΔMIC and wild-type mice when assessing solely amygdala-dependent cued fear conditioning, a test that associates the foot shock with an auditory stimulus, regardless of environment (Figure 4E). These data strongly suggest that mice deficient of the eIF4G1 microexon have impaired hippocampal memory consolidation.

eIF4G1 Microexon Controls Synaptic Transmission and Plasticity

Memory formation involves activity-dependent changes in the strength of specific synaptic connections. Newly acquired memory is temporarily stored in hippocampal circuits by activity-dependent strengthening of individual synapses through processes known as hippocampal long-term potentiation (LTP). Although long-lasting forms of LTP are dependent on new protein synthesis (Costa-Mattioli et al., 2009), activity-dependent mechanisms that serve to maintain long-term synaptic connectivity remain poorly understood. We therefore investigated whether the eIF4G1 microexon controls synaptic function and plasticity in hippocampal slices from wild-type and Eif4g1ΔMIC/ΔMIC littermates. Whole cell voltage-clamp recording in CA1 hippocampal neurons reveals a significant increase in the amplitude of spontaneous inhibitory postsynaptic currents (sIPSCs) compared to wild-type neurons (Figure 5A; p < 1 \times 10^{-5}; Mann-Whitney U test), suggesting that the mutant neurons have increased inhibitory synaptic conductance, as would be expected due to increased expression of inhibitory synaptic proteins, such as Gephyrin and Neuroligin-2 (Figure 4D). This effect is specific to sIPSC amplitude because no appreciable difference in the frequency and amplitude of miniature (m) and spontaneous (s) excitatory postsynaptic currents (EPSCs), or in the...
frequency of m/sIPSCs, is observed (Figures 5A, 5B, S5A, and S5B). In addition, no differences in intrinsic excitability or in resting membrane potential are detected between wild-type and mutant neurons (Figures S5 C and S5D). Furthermore, paired-pulse facilitation, a presynaptic form of short-term plasticity, is unchanged, as is the input-output relationship of field excitatory postsynaptic potentials (fEPSPs), which suggests that there are no differences in basal excitatory synaptic transmission between wild-type and Eif4g1<sup>ΔMIC/ΔMIC</sup> mice at hippocampal (CA3-CA1 neuron) synapses (Figures 5 C and S5E).

We then asked whether the eIF4G1 microexon impacts activity-dependent changes in CA3-CA1 synaptic connections. eIF4G1 microexon-deficient neurons display a significant facilitation of LTP after induction with multiple tetani (i.e., 4 × 100 Hz; 1-s

---

### Figure 4. Altered Social Behavior and Impaired Memory in eIF4G1 Microexon-Deficient Mice

(A) RT-PCR (upper) and western blot (lower) monitoring Eif4g1 microexon splicing and eIF4G1 expression from cerebral cortices of P1 WT and Eif4g1<sup>ΔMIC/ΔMIC</sup> mice (n = 3).

(B) Three-chamber social novelty test showing the duration of time that WT and Eif4g1<sup>ΔMIC/ΔMIC</sup> male mice spent in each chamber; n > 10 mice per genotype. Error bars, SEM; ***p < 0.001; two-way ANOVA and Tukey’s post hoc test.

(C) Reciprocal social interaction test showing duration of time for direct nose-to-nose interactions and random touching frequencies. n ≥ 10 male mice per genotype. Error bars, SEM; **p < 0.01; *p < 0.05; two-tailed t test.

(D) Western blots (upper panel) and quantification of relative expression (lower panel) of GluN1, Gephyrin, and Neuroligin-2 protein levels in synaptosomal preparations from WT and Eif4g1<sup>ΔMIC/ΔMIC</sup> hippocampal tissues. Gapdh immunoblot is shown to control for loading and recovery. Error bars, SD; *p < 0.05; two-tailed t test.

(E) Schematic representation of fear-conditioning paradigm for testing associative learning and memory (upper panel). Percentage of contextual freezing time of WT, Eif4g1<sup>ΔMIC/ΔMIC</sup>, or Eif4g1<sup>ΔMIC/ΔMIC</sup> mice displayed within a 5-min period 1 day after the acquisition of a noxious stimulus is shown (foot shock; middle panel). Percentage of freezing time of WT, Eif4g1<sup>ΔMIC/ΔMIC</sup>, or Eif4g1<sup>ΔMIC/ΔMIC</sup> mice displayed within a 5-min period, 2 days after the acquisition of a stimulus is shown (foot shock), associated with a tone. The tone was presented during the 3<sup>rd</sup> min of the test period (right panel). n ≥ 10 mice per genotype; Error bars, SEM; **p < 0.01; two-way ANOVA and Tukey’s post hoc test (see Supplemental Information for details).
Figure 5. Loss of Eif4g1 Microexon Impacts Synaptic Function and Plasticity
(A and B) Representative traces (left), cumulative distribution, and bar graph plots (right) of sIPSCs (A) and sEPSCs (B) from CA1 hippocampal pyramidal neurons in acute slices from WT and Eif4g1ΔMIC/MIC mice. Eif4g1ΔMIC/MIC mice display increased amplitude of inhibitory synaptic transmission. n > 18 slices.
(C) Activation of LTP via a milder induction protocol (100 Hz; 1-s electrical pulse) that engages the protein-synthesis-dependent form of LTP (Figures 5D and 5E; p < 0.05; two-tailed t test). These results may be due to increased expression of GluN1 and activity-dependent genes in eIF4G microexon-deficient neurons (Figure 3). Activation of LTP via a milder induction protocol (100 Hz; 1-s electrical pulse) that engages the protein-synthesis-dependent form of LTP (Figures 5D and 5E; p < 0.05; two-tailed t test). These results may be due to increased expression of GluN1 and activity-dependent genes in eIF4G microexon-deficient neurons (Figure 3). Activation of LTP via a milder induction protocol (100 Hz; 1-s electrical pulse) that engages the protein-synthesis-dependent form of LTP (Figures 5D and 5E; p < 0.05; two-tailed t test). These results may be due to increased expression of GluN1 and activity-dependent genes in eIF4G microexon-deficient neurons (Figure 3).

Eif4g1 Microexons Promote Associations with Cytoplasmic mRNP Granule Components
To investigate the mechanism by which the eIF4G microexons regulate translation and behavior, we asked whether they encode or overlap amino acid sequence features predictive of functional properties. Both microexons are embedded within eukaryotic-conserved, intrinsically disordered, low-complexity regions identified as “prion-like domains” (Figures 6A and 6A). Prion-like domains have a propensity for phase separation.
Figure 6. eIF4G Microexon Promotes Interactions with mRNP Cytoplasmic Granule Components and Ribosomal Stalling

(A) Left panel shows predicted prion-like amino acid composition (PLAAC) (Lancaster et al., 2014) for eIF4G1. Right panel shows predicted disordered regions (DISOPRED3) (Jones and Cozzetto, 2015) for human eIF4G1. Known domains and microexon location in eIF4G1 are indicated above the panels. Similar results are shown for eIF4G3 in Figure S6A.

(B) Peptide count ratios from BioID-MS for eIF4G1 and eIF4G1+MIC variants expressed in N2A cells. Proteins previously identified as cytoplasmic mRNP granule components (Youn et al., 2018) are indicated in blue. A significant fold increase in association of granule components with the eIF4G1+MIC relative to the eIF4G1ΔMIC variant is indicated (p < 8.9 x 10^-3; Mann-Whitney U test). Dot plot of spectral counts represents the relative abundance of proximal proteins detected for the eIF4G1 splice variants. BFDR, Bayesian false discovery rate.

(legend continued on next page)
in vitro (Banani et al., 2017; Franzmann and Alberti, 2019; Vernon and Forman-Kay, 2019) and contribute to the formation of cellular biomolecular condensates, such as cytoplasmic ribonucleoprotein granules in neurons that function in localized translation and the stabilization of memory formation (Gomes and Shorter, 2019; Li et al., 2013; Mittag and Parker, 2018; Ramaswami et al., 2013; Si and Kandel, 2016). The eIF4G microexon sequences (consensus: GGGFXnxG; Figure 1B) resemble RNA-binding RGG motifs that enhance phase separation via multivalent electrostatic, hydrogen bonding and pi interactions during the formation of granules (Chong et al., 2018; Thandapani et al., 2013). Notably, all of the conserved residues in the microexon sequence contain pi groups (i.e., backbone/side-chain amide, aromatic, and guanidinium) that can contribute to multivalent planar pi-pi stacking interactions with RNA bases, or with similar low-complexity disordered protein regions, to promote phase separation (Nott et al., 2015; Vernon et al., 2018).

Thus, inclusion of the microexon is predicted to increase the propensity for phase separation of eIF4G.

To investigate whether the microexons promote phase separation, we generated recombinant proteins comprising the N-terminal disordered region (amino acids 1–200) of eIF4G1, with or without the microexon (eIF4G11–200+MIC and eIF4G11–200ΔMIC; Figure S6B). These proteins were analyzed for RNA binding and phase separation in vitro. Compared to eIF4G11–200ΔMIC, the eIF4G11–200+MIC protein has a modest increase in affinity for RNA, as indicated by fluorescence polarization assays (Figure S6C). Both isoforms phase separate in the presence of RNA, forming liquid droplets (Figure S6D). Notably, the eIF4G11–200+MIC protein also displays a significantly increased propensity for phase separation compared to eIF4G11–200ΔMIC, as quantified by turbidity measurements (p < 0.001; two-way ANOVA; Figure S6D). These results provide evidence that the eIF4G1 microexon can promote multivalent interactions within phase-separated protein-RNA condensates.

To investigate whether the eIF4G microexons function in the formation of cytoplasmic granules, we employed proximity biotin labeling (i.e., via tagging with the promiscuous biotinylase BirA*), followed by streptavidin capture and mass spectrometry (BioID-MS), to identify microexon-dependent changes in proteins that are proximal to eIF4G1. N2A Flp-In lines expressing N-terminal BirA*-tagged eIF4G1, with and without its microexon (eIF4G1+MIC and eIF4G1ΔMIC), were generated, and captured peptides were quantified for differential enrichment. A significant enrichment for peptides corresponding to proteins with known associations with cytoplasmic mRNP granules (Youn et al., 2018) is observed for the eIF4G1+MIC-proximal interactome, compared to the eIF4G1ΔMIC-proximal interactome (p < 8.9 × 10−6; Mann-Whitney U test; Figure 6B; Table S4). Moreover, the eIF4G1+MIC-proximal interactome comprises a higher fraction of intrinsically disordered amino acid residues relative to the eIF4G1ΔMIC-proximal interactome (Figure S6E), as well as a higher fraction of proteins that contain disordered prion-like domains (Figure S6F). The increased association of eIF4G1+MIC with numerous cytoplasmic granule components, including Fxr1, Ataxin-2, Larp1, and Staun2, was validated by co-immunoprecipitation-western blot analysis using lysates pre-treated with RNase (Figure S6G). Moreover, a comparison of the intracellular localization of mCherry-eIF4G1+MIC and mCherry-eIF4G1ΔMIC proteins reveals that the microexon increases the propensity of eIF4G1 to associate with cytoplasmic foci (p < 0.0352; two-way ANOVA; Figure S6H). These data, together with the results from the in vitro phase separation assays, are consistent with a role for the eIF4G1 microexon in promoting the formation of neuronal granules.

**eIF4G Microexons Promote Associations with FMRP and Ribosome Stalling**

An intriguing possibility is that splicing of the eIF4G microexons results in translational repression through their propensity to contribute to the formation of neuronal granules, consistent with known roles for these structures in translational control (Darnell and Richter, 2012; Li et al., 2013; Pimentel and Boccaccio, 2014; Ramaswami et al., 2013). To explore this possibility, we initially investigated the impact of the eIF4G microexons on global protein synthesis rates. Puromycin incorporation into nascent polypeptides reveals a modest but significant increase in the translation rates of the microexon DKO neurons (Figure 6C; p < 0.05; two-tailed t test). In contrast, the polysome profiles of wild-type and DKO neurons are essentially indistinguishable, and no appreciable differences are detected in the polysome distributions of transcripts encoding proteins that are upregulated upon deletion of eIF4G microexons (Figure S7A; refer to Figure 3).

Next, we performed ribosome profiling (Ingolia et al., 2011) in wild-type and DKO neurons to assess differences in the
translational engagement of mRNAs. Deep sequencing of ribosome-protected fragments (RPFs) reveals a high degree of reproducibility between replicates (Figure S7B) and the expected three-nucleotide periodicity of the RPF reads (Figure S7C). To compute translational engagement across the transcriptome, we compared the levels of RPF reads relative to the corresponding mRNA reads derived from parallel RNA-seq data. No appreciable differences in overall ribosome occupancy are observed between wild-type and DKO neurons (Figure S7D; Table S5).

We next investigated whether the eIF4G microexons impact ribosome pausing by measuring peaks of RPF reads relative to background densities of reads in the surrounding regions of the corresponding transcripts. Importantly, we detect a global increase in ribosomal stalling across coding sequences in wild-type versus DKO neurons (Figure S7E; p < 3.09 × 10^{-5}; Fisher’s exact test; Table S5). Transcripts encoding the synaptic and other proteins in Figure 3C that are translationally repressed by the eIF4G microexons display a significant increase in stalling compared to those of other genes (Figure 6D; p < 2.4 × 10^{-4}; Fisher’s exact test). To further test whether the eIF4G microexons induce ribosomal pausing, we performed in situ run-off assays in hippocampal neuronal cultures followed by puromycylation (i.e., ribopuromycylation; Graber et al., 2013, 2017). This experiment reveals a significant reduction in stalled ribosomes in the DKO neurons (Figure 6E; p < 0.78 × 10^{-10}; Mann Whitney U test). Collectively, these data strongly suggest that the eIF4G microexons control translation by promoting ribosome stalling.

To further investigate this proposed mechanism, we analyzed the ribosome profiling data for codon or motif enrichment at pausing sites. Consistent with previous observations, we observe increased pausing at stop and arginine codons (Rodnina, 2016) but with no appreciable differences between wild-type and DKO neurons (Figure S7F). However, we detect a strong enrichment for motifs corresponding to known binding sites (Ray et al., 2013) of the FXS-related proteins FXR1, FXR2, and FMRP downstream of pausing sites (Figure 6F; FDR < 0.001). These findings suggest that FMR-related proteins may participate in microexon-dependent ribosome pausing, consistent with our observation of eIF4G microexon-promoted granule component associations involving FMRP and accumulating evidence that cytoplasmic mRNP granules enriched for FMRP are associated with translationally stalled 80S ribosomes (Darnell and Richter, 2012; Darnell et al., 2011; El Fatimy et al., 2016; Graber et al., 2013; Richter and Coller, 2015). Further supporting a mechanistic connection between eIF4G microexons and FMRP, there is a significant enrichment (i.e., odds ratio = 2.6; p < 0.014; Fisher’s exact test) of FMRP-binding targets among transcripts encoding proteins that are upregulated in the microexon-DKO neurons (Figure S7G). Moreover, transcripts with ribosomal pausing sites in the DKO neurons and those in Fmr1-deficient neurons (Liu et al., 2018) significantly overlap (p < 1.392 × 10^{-11}; Fisher’s exact test; Figure S7H). Also of note, transcripts that exhibit increased ribosome stalling in the DKO neurons significantly overlap those known to have a dendritic localization (Figure S7I; p < 1.51 × 10^{-11}; Fisher’s exact test).

Because eIF4G microexons promote phase separation and the association of numerous cytoplasmic granule components, we asked whether they also enhance interactions with FMRP. Because FMRP is of relatively low abundance and not detected in our BioID-MS analysis, we performed immunoprecipitation of eIF4G1+MIC and eIF4G1ΔMIC splice variants followed by western blot analysis with anti-FMRP antibody. Importantly, the eIF4G1+MIC splice variant displays a 2-fold increased interaction with FMRP that is largely insensitive to RNase treatment (Figures 6G and S6G). Next, to assess whether the eIF4G microexons might directly impact FMRP granule formation, we incubated eIF4G1^1-200+MIC and eIF4G1^1-200ΔMIC proteins with the recombinant C-terminal disordered region of FMRP (residues 445–632). In this experiment, we compared either phosphorylated (at 8–10 sites) or unphosphorylated forms of FMRP^445-632, given previous observations that phospho-FMRP is preferentially associated with translational repression and stalled ribosomes and has a higher propensity to phase separate in vitro (Ceman et al., 2003; Tsang et al., 2019). Notably, eIF4G1^1-200+MIC has a higher propensity for phase separation with phospho-FMRP^445-632 than eIF4G1ΔMIC, as revealed by increased optical density and droplet formation (Figure 6H; p < 0.001; two-way ANOVA). Moreover, this property of the eIF4G1^1-200+MIC protein is dependent on FMRP phosphorylation because unphosphorylated FMRP^445-632 fails to phase separate (Figure 6H).

Taken together, the results provide evidence that the eIF4G microexons promote neuronal granule formation with critical regulators of synaptic translation, including FMRP, and that the resulting interactions likely lead to ribosomal stalling and dysregulation of the translation of synaptic proteins that are critical for proper cognitive functioning (Figure 7).

**DISCUSSION**

Altered splicing and translational control have emerged as recurring mechanisms underlying ASD and associated neurological disorders (Iakoucheva et al., 2019; Queensel-Vallières et al., 2019), yet how these two layers of regulation impact each other to control cognitive functioning has not been previously determined. In this study, we show that vertebrate-conserved, neuronal microexons in eIF4G1 and eIF4G3 control the translation of critical synaptic proteins and cognitive functioning. We further provide evidence that the eIF4G microexons normally repress protein synthesis by promoting multivalent interactions with cytoplasmic mRNP components, including the fragile X-linked translation regulator FMRP. Because FMRP is of relatively low abundance and not detected in our BioID-MS analysis, we performed immunoprecipitation of eIF4G1+MIC and eIF4G1ΔMIC splice variants followed by western blot analysis with anti-FMRP antibody. Importantly, the eIF4G1+MIC splice variant displays a 2-fold increased interaction with FMRP that is largely insensitive to RNase treatment (Figures 6G and S6G). Next, to assess whether the eIF4G microexons might directly impact FMRP granule formation, we incubated eIF4G1^1-200+MIC and eIF4G1^1-200ΔMIC proteins with the recombinant C-terminal disordered region of FMRP (residues 445–632). In this experiment, we compared either phosphorylated (at 8–10 sites) or unphosphorylated forms of FMRP^445-632, given previous observations that phospho-FMRP is preferentially associated with translational repression and stalled ribosomes and has a higher propensity to phase separate in vitro (Ceman et al., 2003; Tsang et al., 2019). Notably, eIF4G1^1-200+MIC has a higher propensity for phase separation with phospho-FMRP^445-632 than eIF4G1ΔMIC, as revealed by increased optical density and droplet formation (Figure 6H; p < 0.001; two-way ANOVA). Moreover, this property of the eIF4G1^1-200+MIC protein is dependent on FMRP phosphorylation because unphosphorylated FMRP^445-632 fails to phase separate (Figure 6H).

Taken together, the results provide evidence that the eIF4G microexons promote neuronal granule formation with critical regulators of synaptic translation, including FMRP, and that the resulting interactions likely lead to ribosomal stalling and dysregulation of the translation of synaptic proteins that are critical for proper cognitive functioning (Figure 7).
Our results further highlight a role of alternative splicing in controlling the phase separation propensities of proteins, with associated functional consequences. In particular, our data suggest that the eIF4G microexons impact translation by modulating the propensity of eIF4G to associate and coalesce with components of neuronal granules, including FMRP, in a manner that stalls ribosome elongation on synaptic protein transcripts. Consistent with this proposed mechanism, previous studies have shown that eIF4G associates with FMRP neuronal granules, which contain polyribosomes and multiple additional components identified in our BioID analysis (El Fatimy et al., 2016; Knowles et al., 1996) that have been linked to ribosomal pausing (Ceman et al., 2003; Darnell et al., 2011; Graber et al., 2013; Pimentel and Boccaccio, 2014). Interestingly, despite decreased ribosomal pausing and elevated puromycin incorporation in the Eif4g1/3 microexon DKO neurons, we did not detect a substantial shift in polysome profiles or overall differences in ribosome occupancy (Figure S7). This suggests that the eIF4G microexons not only induce ribosomal stalling but may also prevent new rounds of translation initiation, which also has been reported to occur in neuronal granules (Graber et al., 2013, 2017; Pimentel and Boccaccio, 2014). Also noteworthy in this regard is that LARP1, which is a cytoplasmic granule component that prevents cap-dependent translation initiation (Fonseca et al., 2015; Lahr et al., 2017; Wilbertz et al., 2019), additionally displays a preferential interaction with the microexon containing eIF4G isoforms (Figures 6B and S6G). Thus, the microexon-promoted associations of multiple granule components likely regulate translation of synaptic protein expression.

Neuronal microexons are significantly enriched for activity-responsive regulation compared to longer brain-specific alternative exons (Quesnel-Vallières et al., 2016), and specific functions for these exons are beginning to emerge. These include roles in chromatin regulation and transcription (e.g., Kdm1a and Mef2c; Rusconi et al., 2017; Zhu et al., 2005), neurite formation (e.g., Unc13b and protrudin; Ohnishi et al., 2014, 2017; Quesnel-Vallières et al., 2015) and animal behavior (Kdm1a and Cpeb4; Parras et al., 2018; Rusconi et al., 2016; Wang et al., 2015). Notably, a recent report has provided evidence that the microexon in Cpeb4 preferentially controls deadenylation and the expression of genes enriched in ASD-risk alleles, and altered ratios of expression of the corresponding Cpeb4 splice variants were associated with ASD-like phenotypes in mice (Parras et al., 2018). Although the mechanism by which the CPEB4 microexon affects translation is not known, like the eIF4G microexon, we observe that it is regulated by Srrm4 (Figures S7J and S7K;
Quesnel-Vallières et al., 2015). Interestingly, CPEB family proteins also have prion-like disordered domains that promote phase separation, and prion-like states of these proteins have been linked to memory stabilization (Si and Kandel, 2016). Taken together with our findings, it is apparent that the wider network of SRRM4-regulated neuronal microexons functions in the activity-dependent remodeling of the neuronal proteome to control animal behavior.

Although misregulation of Srm4-dependent microexons has been detected in 30% of analyzed ASD patients with idiopathic forms of ASD (Irinia et al., 2014), our findings have intriguing links with FXS. In particular, altered translational control of synaptic proteins as a consequence of eIF4G microexon skipping appears to occur through a mechanism related to that of disruption of the FMRP protein in FXS. It is worth noting, however, that, although the knockout of a single microexon in Eif4g1 in the present study results in social behavior and cognitive defects, it is probable that the more subtle perturbation of the splicing of this and many additional microexons controlled by Srm4 collectively contribute to ASD-associated phenotypes. As such, defining the specific functions of additional microexons controlled by Srm4 represents an important goal of future research. Our results further highlight the potential of modulating Srm4 and its target microexons as a possible future therapeutic strategy for ASD and related disorders.

STAR METHODS

Detailed methods are provided in the online version of this paper and include the following:

- **KEY RESOURCES TABLE**
- **LEAD CONTACT AND MATERIALS AVAILABILITY**
- **EXPERIMENTAL MODEL AND SUBJECT DETAILS**
  - Cell culture
  - In vitro neuronal differentiation
  - Primary neuronal culture
  - In vivo mouse studies
  - Housing and husbandry conditions for experimental animals
- **METHOD DETAILS**
  - siRNA and plasmid transfections
  - Polysome profiling
  - RNA extraction and RT-PCR assays
  - Quantitative RT-PCR
  - Ribopuromycylation
  - Fluorescence microscopy
  - Co-immunoprecipitation experiments
  - Puromycin labeling
  - Preparation of synaptoneurosomal fractions
  - Immunoblotting
  - Individual-nucleotide Resolution Cross-linking and Immunoprecipitation (iCLIP-Seq)
  - CRISPR/Cas9-directed deletion of Eif4g microexons in CGR8 cells
  - Quantitative proteomics
  - RNA sequencing
  - Ribosome profiling
- **QUANTIFICATION AND STATISTICAL ANALYSIS**
  - Identification of conserved tissue-specific translation factors
  - PsychENCODE data analysis
  - Computational detection of prion-like and intrinsically disordered regions in eIF4G proteins
  - Quantitative proteomics data analysis
  - Functional enrichment analysis
  - BioID mass spectrometry data analysis
  - BioID data visualization
  - iCLIP-Seq analysis
  - RNA sequencing and ribosome profiling data analysis
  - Overlap of paused genes in Eif4g DKO and Fmr1 KO neurons
  - Motif enrichment in Eif4g DKO paused genes
  - Quantification of stalled ribosomes from micrographs
  - Foci Quantification
- **DATA AND CODE AVAILABILITY**

SUPPLEMENTAL INFORMATION

Supplemental Information can be found online at https://doi.org/10.1016/j.molcel.2020.01.006.

ACKNOWLEDGMENTS

The authors gratefully acknowledge Jonathan Warner, a giant in the field of ribosome biosynthesis who sadly passed away in September, for inspiring discussions. M. Aregger and A. Nagpal are thanked for helpful comments on the manuscript. D. Torti, T. Durbic, and G. O’Hanlon in the Donnelly Sequencing Centre are gratefully acknowledged for generating RNA-seq data and L. Nutter in the Centre for Phenogenomics (TCP) for assistance with generating mutant mouse strains. We also thank J. Rai for assistance with molecular biology experiments; M. Palandra for assistance with biophysical experiments; R. Vernon for phase separation prediction insights; C. Ouzounis for protein sequence analysis; I. Voineagu for sharing RNA-Seq data; A. Leonelli and C. Dalrymple for mouse husbandry; L. Kelsey, Z. Berberovic, and A. Flenniken for performing ABR and EGR tests; Z. Dargaei and J. Pressey for contributions to the electrophysiology experiments; and A. Au, K. Meier-Ross, and K. Szederkenyi for assistance with microscopy experiments. T.G.-P. was supported by postdoctoral fellowships from EMBO, OIRM, and CIHR; R.J.W. by CIHR Postdoctoral, Marie Curie IOF, and Scrimshaw Foundation fellowships; B.T. by a CIHR Vanier Scholarship; and S.Z. by a Mitacs Elevate postdoctoral fellowship. This research was funded by grants from the Canadian Institutes of Health Research (B.J.B., S.P.C., A.-C.G., M.A.W., J.D.F.-K., and G.L.C.), Simon’s Foundation (B.J.B.), Canada First Research Excellence Fund Medicine by Design Program (B.J.B.), Canada Foundation for Innovation and Genome Canada (A.-C.G. and others), European Research Council (ERC) and the Spanish Ministry of Economy (M.I.), Brain Canada (G.L.C.), Camille Dan Chair Mt. Sinai Hospital (S.P.C.), and Canada Research Chairs Tier I (J.D.F.-K., S.P.C. and A.-C.G.). B.J.B. holds the Banbury Chair in Medical Research at the University of Toronto.
AUTHOR CONTRIBUTIONS

T.G.-P., performed molecular and cellular characterization of the regulation and function of microexons, with contributions from S.F., T.H., and E.S.; R.J.W., U.B., T.G.-P., and M.I. performed analyses of RNA-seq, ribo-seq, and ChIP-seq; B.T. and J.D.F.-K. generated and analyzed in vitro RNA binding and phase separation data; J.P. analyzed microexon regulatory conservation in zebrafish; S.Z., J.R., T.G.-P., R.J.W., U.B., S.M., and A.-C.G. generated and analyzed MS data; E.W.S., X.L., J.G., G.L.C., and M.A.W. generated and analyzed electrophysiology data; and R.N., T.G.-P., M.Q.-V., and S.P.C. performed and analyzed in vivo phenotypic data. N.S. contributed critical reagents and insights. T.G.-P., S.P.C., and B.J.B. designed the study with input from other authors. T.G.-P. and B.J.B. wrote the manuscript with input from other authors.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: January 21, 2019
Revised: August 15, 2019
Accepted: January 2, 2020
Published: January 29, 2020

REFERENCES


# STAR METHODS

## KEY RESOURCES TABLE

<table>
<thead>
<tr>
<th>REAGENT or RESOURCE</th>
<th>SOURCE</th>
<th>IDENTIFIER</th>
</tr>
</thead>
<tbody>
<tr>
<td>Antibodies</td>
<td></td>
<td></td>
</tr>
<tr>
<td>anti-eIF4G1</td>
<td>Cell Signaling</td>
<td>2498; RRID:AB_2096025</td>
</tr>
<tr>
<td>anti-eIF4G3</td>
<td>Gift by Nahum Sonenberg</td>
<td>N/A</td>
</tr>
<tr>
<td>anti-NMDAR1</td>
<td>Thermo Fisher Scientific</td>
<td>PA3-102; RRID:AB_2112003</td>
</tr>
<tr>
<td>anti-GAPDH</td>
<td>Proteintech</td>
<td>10494-1-AP; RRID:AB_2263076</td>
</tr>
<tr>
<td>anti-Gephyrin</td>
<td>Synaptic Systems</td>
<td>147 111; RRID:AB_887719</td>
</tr>
<tr>
<td>anti-Neuroligin-2</td>
<td>Synaptic Systems</td>
<td>129 203</td>
</tr>
<tr>
<td>anti-FMRP</td>
<td>Cell Signaling</td>
<td>4317; RRID:AB_1903978</td>
</tr>
<tr>
<td>anti-FXR1</td>
<td>Santa Cruz Biotechnology</td>
<td>sc-374148; RRID:AB_10918113</td>
</tr>
<tr>
<td>anti-STAU2</td>
<td>Thermo Fisher Scientific</td>
<td>PA5-78473; RRID:AB_2736660</td>
</tr>
<tr>
<td>anti-eIF4E</td>
<td>Cell Signaling</td>
<td>9742; RRID:AB_823488</td>
</tr>
<tr>
<td>anti-LARP1</td>
<td>Abcam</td>
<td>ab86359; RRID:AB_1925169</td>
</tr>
<tr>
<td>anti-Ataxin-2</td>
<td>Proteintech</td>
<td>21776-1-AP; RRID:AB_10858483</td>
</tr>
<tr>
<td>anti-Flag M2</td>
<td>Sigma-Aldrich</td>
<td>F3165; RRID:AB_262044</td>
</tr>
<tr>
<td>anti-Puromycin 12D10</td>
<td>MD-Millipore</td>
<td>MABE343; RRID:AB_2566826</td>
</tr>
<tr>
<td>anti-RFP</td>
<td>Chromotek</td>
<td>[S8]; RRID:AB_2336064</td>
</tr>
<tr>
<td>anti-Actin</td>
<td>Abcam</td>
<td>ab3280; RRID:AB_303668</td>
</tr>
<tr>
<td>anti-TUJ1</td>
<td>Covance</td>
<td>MRB-435P; RRID:AB_663339</td>
</tr>
<tr>
<td>anti-MAP2</td>
<td>Abcam</td>
<td>ab5392; RRID:AB_2138153</td>
</tr>
<tr>
<td>Bacterial Strains</td>
<td></td>
<td></td>
</tr>
<tr>
<td>BL21-CodonPlus (DE3) Competent Cells</td>
<td>Agilent</td>
<td>230280</td>
</tr>
<tr>
<td>Subcloning Efficiency DH5x Competent Cells</td>
<td>Thermo Fisher Scientific</td>
<td>18265017</td>
</tr>
<tr>
<td>Chemicals, Peptides, and Recombinant Proteins</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Actinomycin D</td>
<td>Thermo Fisher Scientific</td>
<td>11805017</td>
</tr>
<tr>
<td>Kainic acid</td>
<td>Tocris</td>
<td>0222</td>
</tr>
<tr>
<td>Domoic acid</td>
<td>Tocris</td>
<td>0269</td>
</tr>
<tr>
<td>(S)-3,5-DHPG</td>
<td>Abcam</td>
<td>ab144484</td>
</tr>
<tr>
<td>Bicuculline methiodide</td>
<td>HelloBio</td>
<td>HB0893</td>
</tr>
<tr>
<td>Picrotoxin</td>
<td>HelloBio</td>
<td>HB0506</td>
</tr>
<tr>
<td>CGP 55845</td>
<td>HelloBio</td>
<td>HB0960</td>
</tr>
<tr>
<td>Bicuculline</td>
<td>Sigma-Aldrich</td>
<td>14340</td>
</tr>
<tr>
<td>6-cyano-7nitroquinoline-2,3-dione (CNQX)</td>
<td>Tocris</td>
<td>0190</td>
</tr>
<tr>
<td>DL-2-amino-5-phosphonovaleric acid (AP-5)</td>
<td>Tocris</td>
<td>0105</td>
</tr>
<tr>
<td>Tetrodotoxin (TTX)</td>
<td>Affix scientific</td>
<td>AF3015</td>
</tr>
<tr>
<td>Homoharringtonine</td>
<td>Sigma-Aldrich</td>
<td>SML1091</td>
</tr>
<tr>
<td>Puromycin</td>
<td>Sigma-Aldrich</td>
<td>P8833</td>
</tr>
<tr>
<td>Emetine</td>
<td>Sigma-Aldrich</td>
<td>E2375</td>
</tr>
<tr>
<td>Cycloheximide</td>
<td>Sigma-Aldrich</td>
<td>01810</td>
</tr>
<tr>
<td>Cas9 mRNA</td>
<td>Thermo Fisher Scientific</td>
<td>A29378</td>
</tr>
<tr>
<td>Casein Kinase II</td>
<td>New England Biolabs</td>
<td>P6010</td>
</tr>
<tr>
<td>Opti-MEM I Reduced Serum Medium</td>
<td>Gibco</td>
<td>31985070</td>
</tr>
<tr>
<td>Neurobasal-A medium</td>
<td>Thermo Fisher Scientific</td>
<td>10888-022</td>
</tr>
<tr>
<td>GlutaMAX Supplement</td>
<td>Thermo Fisher Scientific</td>
<td>35050-061</td>
</tr>
<tr>
<td>N2 Supplement</td>
<td>Thermo Fisher Scientific</td>
<td>17502001</td>
</tr>
<tr>
<td>B27 Supplement</td>
<td>Thermo Fisher Scientific</td>
<td>A3653401</td>
</tr>
</tbody>
</table>

(Continued on next page)
## REAGENT or RESOURCE SOURCE IDENTIFIER

<table>
<thead>
<tr>
<th>REAGENT or RESOURCE</th>
<th>SOURCE</th>
<th>IDENTIFIER</th>
</tr>
</thead>
<tbody>
<tr>
<td>CultureOne Supplement</td>
<td>Thermo Fisher Scientific</td>
<td>A3320201</td>
</tr>
<tr>
<td>Laminin</td>
<td>Sigma-Aldrich</td>
<td>L2020</td>
</tr>
<tr>
<td>Poly-D-Lysine</td>
<td>Sigma-Aldrich</td>
<td>6407</td>
</tr>
<tr>
<td>Poly-L-Lysine</td>
<td>Thermo Fisher Scientific</td>
<td>P8920</td>
</tr>
<tr>
<td>KO serum replacement</td>
<td>Thermo Fisher Scientific</td>
<td>10828010</td>
</tr>
<tr>
<td>DMEM (high glucose)</td>
<td>Sigma-Aldrich</td>
<td>D5796</td>
</tr>
<tr>
<td>FBS</td>
<td>GIBCO</td>
<td>12483-020</td>
</tr>
<tr>
<td>FBS premium</td>
<td>MULTICELL</td>
<td>920040</td>
</tr>
<tr>
<td>Sodium pyruvate</td>
<td>GIBCO</td>
<td>11360-070</td>
</tr>
<tr>
<td>MEM NEAA</td>
<td>GIBCO</td>
<td>11140-050</td>
</tr>
<tr>
<td>RNAiMAX</td>
<td>Thermo Fisher Scientific</td>
<td>13778150</td>
</tr>
<tr>
<td>X-tremeGENE HP DNA Transfection Reagent</td>
<td>Roche</td>
<td>06366236001</td>
</tr>
<tr>
<td>Lipofectamine 2000</td>
<td>Thermo Fisher Scientific</td>
<td>11668019</td>
</tr>
<tr>
<td>Dynabeads protein G</td>
<td>Thermo Fisher Scientific</td>
<td>10004D</td>
</tr>
<tr>
<td>Turbo DNase</td>
<td>Thermo Fisher Scientific</td>
<td>AM2238</td>
</tr>
<tr>
<td>RNase I</td>
<td>Thermo Fisher Scientific</td>
<td>AM2295</td>
</tr>
<tr>
<td>Benzonase Nuclease</td>
<td>Sigma-Aldrich</td>
<td>E826</td>
</tr>
<tr>
<td>His-SUMO-εIF4G1+MIC (aa 1-200)</td>
<td>This study</td>
<td>N/A</td>
</tr>
<tr>
<td>His-SUMO-εIF4G1ΔMIC (aa 1-193)</td>
<td>This study</td>
<td>N/A</td>
</tr>
<tr>
<td>His-SUMO-FMRP (aa 445-632)</td>
<td>(Tsang et al., 2019)</td>
<td>N/A</td>
</tr>
<tr>
<td>5′-FAM-sc1 RNA</td>
<td>(Tsang et al., 2019)</td>
<td>N/A</td>
</tr>
</tbody>
</table>

## Critical Commercial Assays

<table>
<thead>
<tr>
<th>REAGENT or RESOURCE</th>
<th>SOURCE</th>
<th>IDENTIFIER</th>
</tr>
</thead>
<tbody>
<tr>
<td>OneStep RT-PCR Kit</td>
<td>QIAGEN</td>
<td>210210</td>
</tr>
<tr>
<td>SensiFAST SYBR® No-ROX Kit</td>
<td>BIORAD</td>
<td>BIO-98005</td>
</tr>
<tr>
<td>Maxima First Strand cDNA Synthesis Kit for RT-qPCR</td>
<td>Thermo Fisher Scientific</td>
<td>K1671</td>
</tr>
<tr>
<td>SuperScriptIII</td>
<td>Thermo Fisher Scientific</td>
<td>18080044</td>
</tr>
<tr>
<td>RNaseasy</td>
<td>Thermo Fisher Scientific</td>
<td>87793</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>CLIP-Seq</th>
<th>This study</th>
<th>GEO: GSE112598 and SRP039559</th>
</tr>
</thead>
<tbody>
<tr>
<td>CLIP-Seq</td>
<td>(Weyn-Vanhentenryck et al., 2014; Gonatopoulos-Pournatzis et al., 2018)</td>
<td>GEO: GSE112598 and SRP039559</td>
</tr>
<tr>
<td>RNA-Seq</td>
<td>This study</td>
<td>GEO: GSE112599</td>
</tr>
<tr>
<td>Ribosome profiling</td>
<td>This study</td>
<td>GEO: GSE112599</td>
</tr>
<tr>
<td>iTRAQ proteomics</td>
<td>This study</td>
<td><a href="http://massive.ucsd.edu">http://massive.ucsd.edu</a> MSV000084658</td>
</tr>
<tr>
<td>Proximity ligation proteomics (BiolD) data</td>
<td>This study</td>
<td><a href="http://massive.ucsd.edu">http://massive.ucsd.edu</a> MSV000083349 and ProteomeXchange: PXD012421</td>
</tr>
</tbody>
</table>

## Experimental Models: Cell Lines

<table>
<thead>
<tr>
<th>EXPERIMENTAL MODELS: CELL LINES</th>
<th>SOURCE</th>
<th>IDENTIFIER</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mouse: C7GR8</td>
<td>ECACC</td>
<td>07032901</td>
</tr>
<tr>
<td>Mouse: N2A</td>
<td>ATCC</td>
<td>CCL-131</td>
</tr>
</tbody>
</table>

## Experimental Models: Organisms/Strains

<table>
<thead>
<tr>
<th>EXPERIMENTAL MODELS: ORGANISMS/STRAINS</th>
<th>SOURCE</th>
<th>IDENTIFIER</th>
</tr>
</thead>
<tbody>
<tr>
<td>C57BL/6J εIF4g1 microexon deficient mice (εIF4g1ΔMIC/ΔMIC)</td>
<td>This study</td>
<td>N/A</td>
</tr>
<tr>
<td>C57BL/6J εIF4g3 microexon deficient mice (εIF4g3ΔMIC/ΔMIC)</td>
<td>This study</td>
<td>N/A</td>
</tr>
<tr>
<td>C57BL/6J εIF4g1/3 microexon deficient mice (εIF4g1ΔMIC/ΔMIC &amp; εIF4g3ΔMIC/ΔMIC)</td>
<td>This study</td>
<td>N/A</td>
</tr>
</tbody>
</table>

(Continued on next page)
LEAD CONTACT AND MATERIALS AVAILABILITY

Further information and requests for resources, reagents and materials should be directed to and will be fulfilled by the Lead Contact, Benjamin J. Blencowe (b.blencowe@utoronto.ca). All unique/stable reagents generated in this study are available from the Lead Contact with a completed Materials Transfer Agreement.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Cell culture
Mouse neuroblastoma Neuro-2A (N2A) and human SH-SY5Y cells were grown in DMEM (high glucose; Sigma-Aldrich) supplemented with 10% FBS, sodium pyruvate, nonessential amino acids, and penicillin/streptomycin. Mouse embryonic stem cells (mESC) were grown in gelatin coated plates in GMEM supplemented with 100 μM β-mercaptoethanol, 0.1 mM nonessential amino acids, 2 mM sodium pyruvate, 2.0 mM L-glutamine, 5,000 units/mL penicillin/streptomycin, 1000 units/mL recombinant mouse LIF (all Thermo Fisher Scientific) and 15% ES fetal calf serum (ATCC). Cells were maintained at sub-confluent conditions. mESC-derived neurons were generated and cultured as described below. All cell lines were maintained at 37°C with 5% CO2. Cells were regularly monitored for absence of mycoplasma infection.

In vitro neuronal differentiation
In vitro differentiation of mESCs into glutamatergic neurons was performed as described previously (Gueroussov et al., 2015; Hubbard et al., 2013). The CGR8 mESC lines were maintained for at least 2 passages prior to differentiation and were differentiated into

<table>
<thead>
<tr>
<th>REAGENT or RESOURCE</th>
<th>SOURCE</th>
<th>IDENTIFIER</th>
</tr>
</thead>
<tbody>
<tr>
<td>Oligonucleotides</td>
<td>See STAR Methods</td>
<td>This study N/A</td>
</tr>
<tr>
<td>Software and Algorithms</td>
<td>vast-tools 1.0 (Tapial et al., 2017)</td>
<td><a href="https://github.com/vastgroup/vast-tools">https://github.com/vastgroup/vast-tools</a></td>
</tr>
<tr>
<td></td>
<td>Prism 8</td>
<td>GraphPad Prism</td>
</tr>
<tr>
<td></td>
<td>MS data storage and analysis: ProHits-LIMS</td>
<td>Liu et al., 2016</td>
</tr>
<tr>
<td></td>
<td>MS Data Significance Analysis of INTeractome analysis: SAINT</td>
<td>Teo et al., 2014</td>
</tr>
<tr>
<td></td>
<td>PausePred (Kumari et al., 2018)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Ribodiff (Zhong et al., 2017)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>ImageJ</td>
<td>NIH</td>
</tr>
<tr>
<td></td>
<td>Velocity</td>
<td>Quorum Technologies</td>
</tr>
<tr>
<td></td>
<td>Clampfit 10.7</td>
<td>Molecular Devices</td>
</tr>
<tr>
<td></td>
<td>Activity Monitor software</td>
<td>Med Associates Inc.</td>
</tr>
<tr>
<td></td>
<td>Ethovision XT software</td>
<td>Noldus</td>
</tr>
<tr>
<td></td>
<td>Video Freeze software</td>
<td>Med Associates Inc.</td>
</tr>
<tr>
<td></td>
<td>Trimmomatic (Bolger et al., 2014)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Tophat (Trapnell et al., 2009)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Proteome Discoverer 2.1.0.81 software</td>
<td>Thermo Scientific</td>
</tr>
<tr>
<td></td>
<td>g:Profiler (Reimand et al., 2016)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Cytoscape – Enrichment map</td>
<td>Isserlin et al., 2014</td>
</tr>
<tr>
<td></td>
<td>PLAAC</td>
<td>Lancaster et al., 2014</td>
</tr>
<tr>
<td></td>
<td>DISOPRED3</td>
<td>Jones and Cozzetto, 2015</td>
</tr>
</tbody>
</table>
neurons between 5–35 passages. Initially, $5 \times 10^6$ cells were transferred to a 15 cm non-adhesive plate (not coated) containing 30 mL of differentiation medium (GMEM containing 5% KnockOut Serum replacement (Thermo Fisher Scientific), 0.1 mM nonessential amino acids, 1 mM sodium pyruvate, 2 mM L-glutamine, 5,000 units/mL penicillin/streptomycin and without LIF). This point was designated as days *in vitro* (DIV) – 8. Half media changes were conducted every 48 hours by collecting cell aggregates to the center of the plates by swirling. The media was supplemented with 6 μM retinoic acid (Sigma-Aldrich) at DIV –4 and DIV –2. On DIV 0, cellular aggregates were dissociated with TrypLE Express for 5 minutes at 37°C. Trypsinization was halted with soybean trypsin inhibitor (Thermo Fisher Scientific) and after gentle dissociation by trituration, cells were pelleted for 5 minutes at 300 x g. Neural progenitors were washed in N2 medium (Neurobasal-A medium with 1x N2 vitamins, 2 mM glutamine and antibiotics (Thermo Fisher Scientific)), counted using a hemocytometer and plated at $1.5 \times 10^5$ cells/cm² in poly-D-Lysine/Laminin coated dishes. The plated neural progenitors were washed with N2 medium after 24 hours to remove residual serum and non-adherent cells. At DIV 2, N2 was replaced with B27 medium (Neurobasal-A supplemented with antibiotics, 2 mM glutamine and 1 x B27 vitamins (Thermo Fisher Scientific)). Subsequently, differentiating neurons underwent full medium changes with B27 on DIV 4, 7 and 10.

**Primary neuronal culture**

Hippocampi of postnatal day 0 (P0) mice were dissected and dissociated using papain (Worthington) solution. Culturing surfaces were first coated with poly-L-Lysine solution in borax buffer and subsequently coated with Laminin solution in PBS. Hippocampal neurons were cultured in Neurobasal plus media system supplemented with CultureOne supplement, N2 supplement, Glutamax and Penicillin/Streptomycin (GIBCO) and half media changes were performed every 3 days until processing.

**In vivo mouse studies**

All mouse studies were conducted using male C57BL/6J mice. All experiments were conducted in compliance with the Animals for Research Act of Ontario and the Guidelines of the Canadian Council on Animal Care (CCAC). The Centre for Phenogenomics (TCP) Animal Care Committee reviewed and approved all procedures conducted on animals.

**Housing and husbandry conditions for experimental animals**

Animal room light cycles were scheduled on at 7 AM and off at 7 PM. Animal holding rooms are supplied with 100% fresh air. Supply and exhaust ventilation exchange rates were set at 10-15 air changes per hour. Room temperature was maintained at 20-22°C and the environmental conditions in the animal facility were monitored by a Siemens Building Automation System. All animals were appropriately housed in compliance with the CCAC Guidelines and the Ontario Ministry of Agriculture, Food and Rural Affairs (OMAFRA) Animals for Research Act. Animal care attendants performed daily husbandry and health checks on all animals. Weaning was performed at 21 days.

**METHOD DETAILS**

**siRNA and plasmid transfections**

N2A cells were transfected with 20 nM of siGENOME siRNA pools (Dharmacon) using RNAiMax (Thermo Fisher Scientific) as recommended by the manufacturer. A non-targeting siRNA pool (D-001206-13) was used as control. Cells were harvested 48 hours post transfection.

SH-SY5Y (ATCC CRL-2266) cells were plated on 6-well plates coated with Poly-L-Lysine (0.1% w/v) and transfected on the next day using X-tremeGENE HP DNA Transfection Reagent protocol version 8 (Roche Diagnostics) using a 3:1 DNA to transfection reagent ratio according to the manufacturer’s instructions.

**Polysome profiling**

Mouse brains were dissected in HBSS supplemented with 0.1 μg/ml cycloheximide on ice and immediately frozen on dry ice. Flash frozen cortices (10-15 mg) were homogenized in 500-750 μl polysome lysis buffer (100 mM KCl, 5 mM MgCl₂, 10 mM Tris-HCl, pH9, 1% Triton X-100, 1% sodium deoxycholate). Extracts were clarified by centrifugation at 2,500 g for 15 minutes at 4°C and cycloheximide and heparin were added to 0.1 μg/ml and 1 μg/ml, respectively. Cytoplasmic lysates with equal RNA content for control and mutant extracts were adjusted to 500 μl with polysome lysis buffer and then loaded onto 13 mL of a 20%–50% sucrose gradient solution in a polycarbonate ultracentrifuge tube. Following centrifugation at 151,000 g at 4°C for 2 hours in an SW41Ti (Beckman) rotor. The fractions were monitored by UV absorbance at 254 nm by ISCO UA-6 UV detector (Teledyne Isco).

**RNA extraction and RT-PCR assays**

RNA was extracted from cells using the QIAGEN RNeasy Mini Kit as recommended by the manufacturer. To assess Eif4g1 and Eif4g3 microexon splicing, forward and reverse primers were designed to anneal to the constitutively included exons upstream and downstream of each alternative exon, respectively. The primer sequences are provided below:

Hs eIF4G1 F: 5’-TCAGTACGCCACAAGCGAC-3’
Hs eIF4G1 R: 5’-AGCAGGGTAGACATGGGCAG-3’
RT-PCR assays were performed using the OneStep RT-PCR kit (QIAGEN) according to the manufacturer’s recommendations, except for RNA from Zebrafish tissues, which was first reverse transcribed into cDNA using oligodT priming with SuperScriptIII (Invitrogen). Reaction products were separated on 3.5% agarose gels. All amplification products shown in agarose gels correspond to expected sizes based on separation of size markers (GeneRuler 1 kb Plus DNA Ladder, Thermo Fisher Scientific). The Percent Spliced In (PSI) values were calculated using the ImageJ software. The intensity of the exon-included band was divided by the sum of the exon-included and exon-skipped bands. The result was multiplied by 100% to obtain the PSI value, which was rounded to the nearest whole integer.

Quantitative RT-PCR
Quantitative RT-PCR (qRT-PCR), was performed after generating first-strand cDNAs from 0.25-3 μg of total RNA using the Maxima H Minus First Strand cDNA synthesis Kit (Thermo Fisher Scientific), as per the manufacturer’s recommendations. The cDNA was diluted to 2 ng/μL. qPCR reactions were performed in a volume of 10 μL using 1 μL of diluted cDNA, 500 nM primers and 5 μL SensiFAST SYBR No-ROX Kit (BIOLINE), using a CFX96 Real-Time PCR Detection System (BIO-RAD) as per manufacturer’s recommendations. The primers used are provided below.

Ribopuromycylation
Labeling of stalled ribosomes was performed as described previously (Graber et al., 2013, 2017). Briefly, DIV 11 WT and DKO mouse hippocampal neurons were incubated in Neurobasal plus medium supplemented with 5 μM Homoharringtonine for 10 minutes to inhibit translation initiation. To tag stalled ribosomes, the medium was subsequently replaced by Neurobasal plus supplemented with 5 μM Homoharringtonine, 100 nM Puromycin and 200 μM Emetine (Sigma-Aldrich) for 5 minutes. Subsequently, cells were placed on ice and washed with HBSS solution supplemented with 0.0003% digitonin for 2 minutes. Neurons were washed with cold HBSS twice to remove Digitonin. Neurons were then fixed for 15 minutes in 4% paraformaldehyde and 10% sucrose in PBS.

Fluorescence microscopy
CGR8-derived neural precursors were plated on coverslips (coated with poly-D-Lysine/Laminin) in 6 well plates at 1 x 10⁶ cells/well. At DIV 8, cells were fixed in 4% paraformaldehyde/PBS. Fixed cells were permeabilized with 0.1% Triton X-100 in PBS for 10 minutes at room temperature and incubated in blocking solution (5% BSA, 1% normal goat serum and 0.2% Tween 20) for 90 minutes at room
temperature. Staining was performed with primary antibodies diluted in blocking solution for 90 minutes at 37°C (antibody dilutions: rabbit anti-TUJ1, Covance MRB-435P at 1:2,000; chicken anti-MAP2, Abcam ab5392 at 1:10,000; anti-Puromycin 12D10, MD-Millipore MABE343 at 1:1,000). Coverslips were washed three times with PBS containing 0.2% Tween 20 and incubated with secondary antibodies diluted in blocking buffer (anti-mouse IgG Alexa Fluor 488 at 1:1,000; anti-rabbit IgG Alexa Fluor 594 at 1:1,000) at room temperature for 60 minutes. Cells were washed five times with PBS containing 0.2% Tween 20 prior to mounting with Daco fluorescence fixing medium. DNA was labeled with Hoechst 33258 (Sigma-Aldrich) to indicate nuclei.

SH-SY5Y cells, 16-18 hours post-transfection, were fixed with 4% formaldehyde and mounted using Vectashield mounting media containing 4',6-Diamidino-2-phenylindole dihydrochloride (DAPI) (Vector Laboratories). Images were acquired using 20x (to view fields of cells) and 60x oil immersion objective lenses (Nikon D-eclipse C1 confocal system).

**Co-immunoprecipitation experiments**

Flag-tagged EIF4G1 ± microexon constructs were transiently transfected into N2A cells grown on 10cm plates using Lipofectamine 2000. After 48 hours, cells were washed in cold phosphate buffered saline (PBS) and lysed in 1 mL of lysis buffer (50 mM HEPES–KOH pH 8.0, 100 mM KCl, 2 mM EDTA, 0.1% NP40, 10% glycerol and supplemented with protease and phosphatase inhibitors). Lysates were subject to sonication (15 x 1 s pulses with 1 s in between at 30% power). For nuclease digestion, 10 μg RNase A, 25 Units RNase T1 and 75 Units of benzonase were added and lysates were incubated at 37°C with shaking for 10 minutes. Lysates were cleared in a microcentrifuge by spinning at 15,000 g for 20 minutes at 4°C. Anti-flag immunoprecipitation was performed using magnetic Dynabeads protein G (Thermo Fisher Scientific) complexed with anti-Flag M2 antibody (Sigma-Aldrich). Antibody was incubated with lysates for 1 hours at 4°C followed by incubation with washed Dynabeads protein G for 3 hours at 4°C with rotation. Following incubation, complexes were washed 5 times with lysis buffer. Elution was performed in 1x Laemmli buffer at 95°C for 5 minutes.

**Puromycin labeling**

CGR8-derived neurons from two independent WT and DKO clones (DIV 8) were treated with puromycin (0.1 μg/ml final) for 15 minutes, 30 minutes and 60 minutes. Cells were rinsed with PBS then extracted with lysis buffer supplemented with protease and phosphatase inhibitors. Immunoblots were prepared and developed as described below. Puromycin incorporation was assessed by measuring the immunoblot intensity of all protein bands after subtracting background.

**Preparation of synaptoneurosomal fractions**

Synaptoneurosomes were isolated using Syn-PER Synaptic Protein Extraction Reagent (Thermo Fisher Scientific) following the manufacturer’s recommendation. Briefly, dissected hippocampal tissues from WT and Eif4g1<sub>MIC</sub>/<sub>MIC</sub> mice were weighed and resuspended in 5x times volume of Syn-PER Reagent (500 μl per 100 mg brain tissue) supplemented with protease and phosphatase inhibitors. Dounce homogenization with 10 strokes was followed by centrifugation at 1,200 g for 10 minutes at 4°C. The supernatant was transferred to a fresh tube and centrifuged at 14,000 g for 10 minutes at 4°C. The supernatant was removed and the pellet re-suspended in Syn-PER Reagent containing 1x Laemmli buffer (100 μl for 100 mg of original brain tissue).

**Immunoblotting**

Cell lysates were mixed with Laemmli buffer and heated at 95°C for 5 min, separated on variable percentage SDS-PAGE gels, and transferred to PVDF membrane. Blots were incubated for one hour at room temperature with the following primary antibodies at the specified dilutions in 5% milk:

- Rabbit anti-eIF4G1 (Cell Signaling; #2898) at 1:1,000
- Rabbit anti-eIF4G3 (Sonenberg lab) at 1:2,000
- Rabbit anti-NMDAR1 (Thermo Fisher Scientific; PA3-102) at 1:500
- Rabbit anti-Gapdh (Proteintech; 10494-1-AP) at 1:2,000
- Rabbit anti-Gephyrin (Synaptic Systems; 147 111) at 1:2,000
- Rabbit anti-Neuroligin-2 (Synaptic Systems; 129 203) at 1:1,000
- Rabbit anti-FMRP (Cell signaling; #4317) at 1:1,000
- Mouse anti-FXR1 (Santa Cruz Biotechnology; B-2; sc-374148) at 1:500
- Rabbit anti-STAU2 (Thermo Fisher Scientific; PA5-78473) at 1:500
- Rabbit anti-eIF4E (Cell signaling; #9742) at 1:1,000
- Rabbit anti-LARP1 (Abcam; ab86359) at 1:2,000
- Rabbit anti-Ataxin-2 (Proteintech; 21776-1-AP) at 1:2,000
- Mouse anti-Flag M2 (Sigma-Aldrich; F3165) at 1:2,000
- Mouse anti-Puromycin 12D10 (MD-Millipore; MABE343) at 1:1,000
- Rat anti-RFP (Chromotek; [5F8]) at 1:500
- Mouse anti-Actin (Abcam; ab3280) at 1:5,000
Individual-nucleotide Resolution Cross-linking and Immunoprecipitation (iCLIP-Seq)

iCLIP was performed as described previously (Huuppertz et al., 2014). Srmr4, Srmr3 or GFP (as a negative control) were immunoprecipitated from CGR8 mESC-derived neurons induced for 36 hours with 2 μg/ml doxycycline to express Flag-tagged Srmr4, Srmr3 or GFP, respectively. Initially, we generated CGR8 mESC single cell clones expressing Flag-tagged Srmr4 or GFP using the PiggyBac transposase system. mESC clones were differentiated into cortical glutamatergic neurons as described above. Neurons (DIV 10) were crosslinked (0.20 J/cm²) at 254 nm with a Stratalinker 1800. Three independent replicates were used for generating iCLIP libraries. Lysates were treated with Turbo DNase (Thermo Fisher Scientific) and RNase I (Thermo Fisher Scientific) diluted in PBS 500 times for 5 minutes at 37°C to digest genomic DNA and trim RNA to short fragments of an optimal size range. RNA-protein complexes were purified using 100 μL of protein G Dynabeads (Thermo Fisher Scientific) and 10 μg of anti-Flag (Sigma-Aldrich) antibody. Following stringent high salt washes, the immunoprecipitated RNA was 3’ end dephosphorylated and ligated to a preadenylated adaptor at the 3’ end as described previously (Van Nostrand et al., 2016). After extensive additional washes and buffer exchange, the immunoprecipitated RNA was 5’ end-labeled using ATP [γ-32P]. The purified protein-RNA complexes were separated with SDS-PAGE and transferred to a nitrocellulose membrane (Protran). Protein was digested using proteinase K, the RNA was purified (using phenol/chloroform) and reverse transcribed into cDNA. The cDNA was size selected by resolving samples on a 6% TBE-Urea gel, circularized to add the adaptor to the 5’ end, digested at the internal BamHI site, and then PCR amplified using AccuPrime SuperMix I (Thermo Fisher Scientific). The final PCR libraries were purified by agarose gel electrophoresis using gel extraction columns (Qiagen). Eluted DNA was mixed at a ratio of 1:5:5 from the low, middle, and high fractions and submitted for sequencing.

For Flag-Srmr4 iCLIP-Seq the barcoded primers used were: Rt1clip, Rt9clip and Rt13clip. For Flag-Srmr3 iCLIP-Seq the barcoded primers used were: Rt10clip, Rt13clip and Rt14clip; and for Flag-GFP the barcoded primers used were Rt6clip, Rt14clip and Rt16clip:

Rt1clip: 5’/Phos/NNACCNNNAGATCGGAAGAGCGTCGTGgatcCTGAACCGC;
Rt6clip: 5’/Phos/NNCGCCNNNAGATCGGAAGAGCGTCGTGgatcCTGAACCGC;
Rt9clip: 5’/Phos/N NGCCANNNAGATCGGAAGAGCGTCGTGgatcCTGAACCGC;
Rt10clip: 5’/Phos/NNACCNNNAGATCGGAAGAGCGTCGTGgatcCTGAACCGC;
Rt13clip: 5’/Phos/NNGCCNNNAGATCGGAAGAGCGTCGTGgatcCTGAACCGC;
Rt14clip: 5’/Phos/NTGCCNNNAGATCGGAAGAGCGTCGTGgatcCTGAACCGC;
Rt16clip: 5’/Phos/NNTAANNNAGATCGGAAGAGCGTCGTGgatcCTGAACCGC.

CRISPR/Cas9-directed deletion of Eif4g1 and Eif4g3 microexons in CGR8 cells

CGR8 mESC lines deficient of both Eif4g1 and Eif4g3 microexons (DKO) were generated as described previously using the CRISPR-Cas9 system (Gueroussov et al., 2015; Ran et al., 2013). Briefly, 20-nt guide sequences were designed to have minimum off-target effects using the online CRISPR design tool provided by the Zhang laboratory (http://tools.genome-engineering.org). Two independent guides targeting sequences upstream (designated as U1: 5’-GTagttagtagtagatagagc-3’ and U2: 5’-CTCTTTTACTACTATCTAGG-3’) and downstream (designated as D1: 5’-CAAGTGGGATGTCACGGT-3’ and D2: 5’-CAGCATCCCTTCAAGTACCGG-3’) of Eif4g1 microexon as well as two independent guides targeting sequences upstream (designated as U1: 5’-GTGCCTCGTCCCGCTGAGTTGAGC-3’ and U2: 5’-CTTGGTAAACAGCCTACTGCTGAGCT-3’) and downstream (designated as D1: 5’-TAGGTACGCGCTGC TTCGAGGCGG-3’ and D2: 5’-CGGAGCGCCTACTAAGCTA-3’) of Eif4g3 microexon were chosen. The guide sequences were cloned individually into the pSpCas9(2B)-2A-Puro (PX459) vector (Addgene #48139), as described previously (Ran et al., 2013). For the simultaneous deletion of both Eif4g1 and Eif4g3 microexons, CGR8 cells were simultaneously transiently transfected with PX459-U1 and PX459-D1 or PX459-U2 and PX459-D2 plasmids using Lipofectamine 2000 (Thermo Fisher Scientific) following the manufacturer’s instructions. 24 hours post-transfection, 1 μg/ml Puromycin was applied to the media to select for transfected cells. Five days post-transfection the cells were plated on gelatin-coated 96-well plates at a concentration of 0.5 cells/well in order to achieve clonal mESC lines. Ten days later, the clones were expanded and screened for homozygous deletion of Eif4g1 and Eif4g3 microexons. For screening cell clones, gDNA was extracted using the Phusion Human Specimen Direct PCR Kit (Thermo Fisher Scientific) as per the manufacturer’s recommendations. Editing was detected by PCR amplification across Eif4g1 and Eif4g3 microexons using the following primers:

Eif4g1 forward: 5’-GGACTGTGCGTGGATAGGCTT-3’
Eif4g1 reverse: 5’-CGGAGCCTGTAAGCCTCTTC-3’
Eif4g3 forward: 5’-AGTGGGATGTCACGGTGCT-3’
Eif4g3 reverse: 5’-CAAGTGGGATGTCACGGTGCT-3’

Positive clones were analyzed further to confirm microexon deletion for both alleles by using reverse primers that recognize the microexon sequence and by Sanger sequencing.
Quantitative proteomics
Sample preparation
Three biological replicates of WT and double Eif4g1 and Eif4g3 micro exon deficient stem cell-derived neuros (DIV 3 and DIV 8) were dissolved in 2 mL 4% SDS with 4 μL of 500 × stock of protease inhibitor cocktail (Cat #P8340, Sigma-Aldrich). The samples were sonicated at 4°C using three 10 s bursts with 10 s pauses at 35% amplitude and then centrifuged at 20,000 g at 4°C for 50 minutes. Protein concentration was determined by a BCA assay (Cat #500-0119, Bio-Rad). The supernatants were stored at –80°C.

A filter-aided sample preparation (FASP) method was used for protein digestion (Wisniewski et al., 2009). 100 μg of protein from each sample was dissolved in 200 μL of a solution containing 4% SDS in 50 mM ammonium bicarbonate (pH 8). Then, 100 mM DTT (dissolved in 50 mM ammonium bicarbonate) was added to the sample and the mixture was heated at 90°C for 10 minutes for protein denaturation and reduction. The mixture was added to a 10 kDa centrifugal filter unit (Cat #VN01H02, Sartorius), followed by centrifugation at 14,000 g for 15 minutes. Alkylation was achieved by addition of 200 μL of a 20 mM iodoacetamide (IAA) solution in 50 mM ammonium bicarbonate directly on the membrane. The mixture was kept at room temperature for 20 minutes (in dark) and then centrifuged at 14,000 g for 15 minutes. The proteins on the membrane were sequentially washed 5 times with 150 μL of 8 M urea (dissolved in 100 mM TEAB), followed by three washes with 150 μL of 100 mM TEAB to remove urea. Finally, 150 μL of 100 mM TEAB was loaded on each filter and 4 μL of 1 μg/μL of trypsin solution (Cat #V5111, Promega) was added to each unit. The filter units were gently vortexed for 5 minutes to mix the trypsin and proteins, then transferred to a 37°C water bath for 12 hours for trypsic digestion. After digestion, the units were centrifuged at 15,000 g for 15 minutes, and the flow-through containing the peptides was collected. To increase peptide recovery from the membrane, the membrane was further washed with 100 mM TEAB and the flow-through from those two steps was then combined.

iTRAQ labeling and high pH fractionation
The peptides from each sample were labeled with iTRAQ 8 plex reagents (Cat #4390812, Sciex) according to the manufacturer’s instructions as follows: WT day 3, iTRAQ reagent 113; WT day 8, iTRAQ reagent 114; DKO day 3, iTRAQ reagent 116; DKO day 8, iTRAQ reagent 117. Differentially labeled peptides were mixed and then fractionated using the Pierce TM High pH Reversed-Phase Peptide Fractionation Kit (Cat #84686). Eight fractions were collected using: 10, 12.5, 15, 17.5, 20, 22.5, 25 and 50% acetonitrile. Those fractions were subsequently vacuum-centrifuged to dryness, resuspended in 10 μL of 5% formic acid and half of it was analyzed by LC-MS/MS.

Liquid Chromatography and Mass Spectrometry
LC-MS/MS was performed using an Eksigent 425 nano Ultra HPLC system and a Thermo Fisher Scientific Orbitrap Fusion Lumos mass spectrometer. Samples were loaded using an autosampler directly onto a home-made 75 μm ID × 15 cm packed tip column filled with C 18 particles (3 μm, Reprosil). Digested peptides were separated with a 112 minutes linear gradient from 4% solvent B (0.1% formic acid in acetonitrile) to 30% solvent B (0.1% formic acid in acetonitrile), at a flow rate of 200 nL/min. The run time was 180 minutes for each fraction, including sample loading and column reconditioning. Data-dependent acquisition was performed using the Xcalibur 4.0 software in positive ion mode at a spray voltage of 2.5 kV. Survey spectra were acquired in the Orbitrap with a resolution of 120,000 and a mass range from 350 to 1500 m/z. For HCD scans, the collision energy was set at 35, maximum inject time was 54 ms and the AGC target was 1.0e5. We used an isolation window of 0.7 m/z. Ions selected for MS/MS were dynamically excluded for 20 s after fragmentation.

RNA sequencing
RNA was extracted from two sets of WT and CRISPR edited DKO CGR8-derived neurons (DIV 8) and rRNA-depleted libraries were generated using the TruSeq Illumina library preparation kit (Epicenter) following manufacturer’s recommendation. The quantified pool was hybridized at a final concentration of 2.1 pM and sequenced by single-end 51 nucleotides reads on the Illumina HiSeq2500 platform using v4 SBS chemistry and resulting in an average of 48.5M reads per sample.

Ribosome profiling
Two sets of WT and DKO CGR8 clones were differentiated to cortical glutamatergic neurons (DIV 8) as described above. Cells were treated with 100 μg/mL cyclohexamide for 5 minutes. Ribosome profiling was performed using ARTseq Ribosome Profiling Kit (Epicenter) following the manufacturer’s recommendation. The quantified pool was hybridized at a final concentration of 2.1 pM and sequenced by single-end 51 nucleotides reads on the Illumina HiSeq2500 platform using v4 SBS chemistry and resulting in an average of 52.25M reads per sample.

Generation of a stable N2A Flp-In rTA3 cell lines
Generation of cell lines and sample preparation for BioID experiments was performed as previously described (Gonatopoulos-Pournatzis et al., 2018). Briefly, inducible N2A Flp-In rTA3 stable cell lines expressing either the eIF4G1+MIC or eIF4G1-MIC isoforms were generated by co-transfection of 200 ng of a rTA3 compatible pCDNA5/FRT plasmid (Thermo Fisher Scientific, V6010-20), with 2 μg of POG44 Flp-recombinase expression vector (Thermo Fisher Scientific, V600520). Cell lines with successful cDNA integration were selected and maintained using 200 μg/mL Hygromycin. Transgene expression was induced by addition of 2 μg/mL Doxycycline.
For BioID experiments, N2A Flp-In cell lines were grown to ~70% confluency in a 15 cm plate and expression was induced for 24 hours with 2 μg/mL Doxycycline, in the presence of 50 μM biotin. Cell pellets were collected from one 15 cm plate for each sample and stored at −80°C prior to BioID purification.

**BioID Sample Preparation for MS**

BioID experiments were performed essentially as described in Lambert et al. (2015). For BioID sample preparation, cell pellets for eIF4G1+MIC, eIF4G1ΔMIC, and control samples (BirA*-Flag-GFP, and BirA*-Flag-empty) were lysed using an ice cold RIPA buffer containing 50 mM Tris-HCl (pH 7.5), 150 mM NaCl, 1% Triton X-100, 1 mM EDTA, 1 mM EGTA and 0.1% SDS, with freshly added 0.5% sodium deoxycholate and protease inhibitor cocktail (Sigma-Aldrich P8340, 1:500) at a 1:10 pellet weight to volume ratio. Lysates were sonicated for 30 s using three 10 s bursts (35% amplitude) and incubated for 30 minutes at 4°C with 1 μL of Benzonase Nuclease (Sigma-Aldrich, E8263, 250U). Lysates were then cleared by centrifugation at 20,000 rcf. for 20 minutes at 4°C.

Following centrifugation, equal volumes of eIF4G1+MIC and eIF4G1ΔMIC lysates were transferred to tubes containing 60 μL of streptavidin–Sepharose bead slurry (GE Healthcare, Cat 17-5113-01). The lysate and streptavidin bead mix were incubated for 3 hours at 4°C with rotation to allow capture of biotinylated proteins. After purification, the streptavidin beads were pelleted (6000 rpm, 30 s), and the supernatant removed. For all samples, the streptavidin beads were resuspended in RIPA lysis buffer and transferred to a new microcentrifuge tube. A series of five sequential washes were then performed, once with RIPA lysis buffer, twice with TAP buffer (50 mM HEPES-NaOH (pH 8.0), 100 mM NaCl, 2 mM EDTA, and 10% glycerol), and twice with 50 mM ammonium bicarbonate (ABC) at pH 8 to remove non-biotinylated proteins and residual RIPA lysis buffer. Following the final wash, residual ABC was pipetted off the beads and replaced with 60 μL of ABC containing 1 μg of trypsin (Sigma-Aldrich, T6567). Samples were incubated overnight with rotation at 37°C. The supernatant was then removed to a new tube and the beads rinsed with an additional 60 μL of ABC buffer to collect the digested peptides (total volume of ~120 μL). A second trypsin digest was then performed with 0.5 μg of trypsin for another 4 hours at 37°C after which samples were acidified with formic acid to a final concentration of 2.5% and dried in a centrifugal evaporator.

**BioID mass spectrometry data acquisition**

BioID samples were analyzed by mass spectrometry in two biological replicates. For each sample, digested peptides were dissolved in 24 μL 5% formic acid with 5 μL of each sample directly loaded into a 15 cm 100 μm ID emitter tip packed in-house with 3.5 μm Reprosil C18. Peptides were eluted from the column over a 90 minutes gradient using a 425 NanoLC (Eksigent, Redwood, CA) and analyzed using a TripleTOF™ 6600 instrument (AB SCIEX, Concord, Ontario, Canada) operated in Data Dependent Acquisition (DDA) mode as previously described (Gonatopoulos-Pournatzis et al., 2018).

**Protein phase separation analysis**

**Recombinant protein expression and purification**

cDNA encoding the prion-like domain of eIF4G1 with and without the MIC was digested using BamHI and XhoI restriction enzymes and ligated into the multiple cloning sites of pET-SUMO vector (Thermo Fisher Scientific).

His-SUMO-eIF4G1+MIC (aa 1-200), His-SUMO-eIF4G1ΔMIC (aa 1-193), or His-SUMO-FMRP (aa 445-632) (Tsang et al., 2019) vectors were transformed into *Escherichia coli* BL21 (DE3) codon plus cells (Agilent) and grown at 37°C in LB medium. Protein expression was induced with 0.25 mM Isopropyl-D-1-thiogalactopyranoside (IPTG) at OD600nm of ~0.6 and grown overnight at 25°C. Cells were pelleted and lysed via sonication in a buffer containing 25 mM Tris pH 8.0, 6 M guanidinium hydrochloride (GdnHCl), 500 mM NaCl, 20 mM imidazole and 2 mM DTT. Cleared lysate was loaded onto a 5 mL HisTrap column (GE Healthcare) equilibrated in lysis buffer. The HisTrap column was extensively washed with 10 column volumes (CVs) of lysis buffer and the GdnHCl was removed by washing the column with 20 CVs of lysis buffer without GdnHCl. The protein was then eluted in 25 mM Tris pH 8.0, 500 mM NaCl, 300 mM imidazole and 2 mM DTT. Fractions containing the protein of interest were confirmed by Coomassie staining of SDS-PAGE gels and pooled together. The His-SUMO tag was cleaved off from the fusion protein using a SUMO protease, ULP, added to the pooled elutions. The cleavage reaction was incubated at room temperature for 3 hours and completion of the ULP reaction was confirmed using Coomassie stained SDS-PAGE gels. The sample was then concentrated using ultrafiltration to ~2 mL, filtered through a 1 μm filter and loaded onto a Superdex 75 HiLoad 16/600 gel filtration column (GE Healthcare) equilibrated with 50 mM Tris pH 8.0, 4 M GdnHCl, 150 mM NaCl and 2 mM DTT. Pure fractions containing the protein of interest were confirmed using Coomassie stained SDS-PAGE and concentrations were determined using a Bradford protein assay. Proteins were stored at −20°C and dialyzed into assay buffers when needed.

**In vitro Phosphorylation of FMRP**

Recombinant FMRP445-632 was phosphorylated *in vitro* as previously described (Tsang et al., 2019). Briefly, purified 100 μM of FMRP445-632 (~5-10 mL) was mixed with 5 μL Casein Kinase II (New England Biolabs) in a phosphorylation buffer containing 25 Tris pH 8.0, 100 mM KCl, 10 mM MgCl₂, 4 mM ATP, 0.5 mM EGTA and 2 mM DTT. The phosphorylation reaction was dialyzed against 4 L of the same buffer at room temperature and the reaction was quenched with the addition of GdnHCl. The quenched phosphorylation reaction was then concentrated to ~2 mL and passed through a Superdex 75 HiLoad 16/600 gel filtration column (GE Healthcare). On average, 8 to 10 phosphorylation sites were added as determined by mass spectrometry. The phosphorylated protein was dialyzed into assay buffers when needed.
**Fluorescence labeling of FMRP recombinant proteins**

Fluorescein (FITC) dye was added via a maleimide linkage to a single endogenous cysteine in FMRP<sup>445-632</sup> or pFMRP<sup>445-632</sup> (C584). First, purified FMRP<sup>445-632</sup> or pFMRP<sup>445-632</sup> was dialyzed into a buffer containing 25 mM Tris pH 7.5, 100 mM NaCl and 4 M GdnHCl. The proteins were passed through a Hi-Trap desalting column (GE-healthcare) to ensure that any residual reducing agents were removed before reacting the proteins with 5X excess FITC dye (Lumioprobe). The labeling reaction was incubated at room temperature for 3 hours and quenched with excess DTT. Unreacted dye was removed by dialysis and then passing the dialyzed protein through a Hi-Trap desalting column (GE-healthcare) equilibrated with 50 mM Tris pH 7.5, 100 mM NaCl, 4 M GdnHCl and 2mM DTT. Dye separation was confirmed by running the labeled samples on SDS-PAGE gel and using a fluorescence reader to detect any remaining free dye.

**Fluorescence polarization**

The 36-mer RNA oligo sc1 (5'-GCUGCGGUAGGAGAGUGGCAGCG-3') was labeled with a 5' 6-FAM dye (Sigma-Aldrich). Increasing concentrations of eIF4G1+MIC or eIF4G1ΔMIC were titrated into 25 nM of 5'6-FAM-sc1 in buffer containing 25 mM Na<sub>2</sub>PO<sub>4</sub> pH 7.4, 2 mM DTT and 0.01% NP-40. The protein-RNA mixture was equilibrated at room temperature for 1 hour and fluorescence polarization was monitored in a black 384-well plate (3820 Corning) with a SpectraMax i3x Multi-Mode Plate reader (Molecular Devices) at 25°C. Averaged values represent three independent biological replicates and dissociation constants (KD) was obtained fitting the data using a Hill plot binding model (Equation 1).

\[
F_p = \frac{\text{MaxFP} \times [\text{FMRP}^n]}{K_d^n + [\text{FMRP}]^n} \quad \text{(eq 1)}
\]

**Turbidity measurements**

30 μM of eIF4G1 with or without MIC was mixed with 5 μM sc1 RNA, 60 μM FMRP<sup>445-632</sup> or 60 μM pFMRP<sup>445-632</sup> in buffer containing 25 mM Na<sub>2</sub>PO<sub>4</sub> pH 7.4 and 2 mM DTT. 20 μL of the mixture was placed in a 384 clear bottom plate (Corning 3544) and incubated at room temperature for 3 minutes before reading the turbidity measurements at 600 nm using a SpectraMax i3x Multi-Mode Plate reader (Molecular Devices) at 25°C. All turbidity measurements were repeated with at least 8 independent replicates for each condition.

**Fluorescence microscopy imaging of phase-separated samples**

10 μM of FITC-pFMRP/FMRP<sup>445-632</sup> or 5'6FAM-sc1-RNA was mixed with 80 μM of eIF4G1 with or without the MIC to induce phase separation. All samples were prepared in buffer containing 25 mM Na<sub>2</sub>PO<sub>4</sub> pH 7.4 and 2 mM DTT. After mixing, samples were incubated at room temperature for 5 minutes and then half of the mixture was transferred to a 96 glass well plate (Eppendorf). Fluorescence images were acquired on a confocal Leica DMi8 microscope equipped with a Hamamatsu C9100-13 EM-CCD camera and 63x/1.4 (O) objective. FITC fluorescence was detected using a 491 nm (50 mW) laser. All images represent a single focal plane focused on the surface of the slide. Images were processed with Velocity software (Perkin-Elmer) and ImageJ (NIH).

**Eif4g1 and Eif4g3 microexon deletion mouse lines**

The *Eif4g1* and *Eif4g3* deletion mutants were generated by direct delivery of Cas9 reagents to C57BL/6J (The Jackson Laboratory, Stock 000664) mouse zygotes at TCP (Toronto, ON, Canada) as described previously (Gertsenstein and Nutter, 2018). Briefly, gRNAs with the desired spacer sequence (STAR Table 1) were synthesized by *in vitro* transcription from a PCR-derived template. A micro-injection mix of 20 ng/μL Cas9 mRNA (Thermo Fisher Scientific A29378) and 10 ng/μL each of four gRNAs, two on either side of the target site, for the individual target gene was microinjected into C57BL/6J zygotes. Injected zygotes were incubated in KSOMAA medium for 3 hours and then half of the mixture was transferred to a 96 glass well plate (Eppendorf). Fluorescence images were acquired on a confocal Leica DMi8 microscope equipped with a Hamamatsu C9100-13 EM-CCD camera and 63x/1.4 (O) objective. FITC fluorescence was detected using a 491 nm (50 mW) laser. All images represent a single focal plane focused on the surface of the slide. Images were processed with Velocity software (Perkin-Elmer) and ImageJ (NIH).

**Brain slice preparation and electrophysiological recordings**

Acute brain slices were prepared from male *Eif4g1<sup>ΔMIC</sup>*/*ΔMIC* mice (C57BL/6J background) and wild-type interleaved controls (P71-105) in accordance with a protocol approved by TCP Animal Care Committee (#20-0292H). The experimenter was blind to genotype during data acquisition and analysis. Isoflurane was administered, and the mouse was observed until a deep state of anesthesia was achieved, indicated by a loss of spinal reflexes. Subsequently, the mouse was decapitated and the brain was immediately submerged in ice-cold, oxygenated, ‘dissection’ artificial cerebrospinal fluid (aCSF) containing (in mM): 124 NaCl, 3 KCl, 24 NaHCO<sub>3</sub>, 1.25 NaH<sub>2</sub>PO<sub>4</sub>, 0 CaCl<sub>2</sub>, 2 MgCl<sub>2</sub> and 10 glucose. The cerebellum was removed with a scalpel, and coronal brain slices (400 μm) containing the dorsal hippocampus were prepared using a vibratome (VT1000S, Leica Biosystems). Slices were allowed to recover at room temperature for at least 2 hours in a submerged incubation chamber containing oxygenated recording aCSF, which was the same as dissection aCSF except for containing 2 mM CaCl<sub>2</sub> and 1 mM MgCl<sub>2</sub>. A slice was transferred to a submerged recording chamber continuously perfused with oxygenated aCSF (2.0 mL/min) at a temperature of 28°C. Synaptic responses were evoked...
STAR Table 1. Sequences used for production of Eif4g1 and Eif4g3 microexon deletions

<table>
<thead>
<tr>
<th>Name</th>
<th>Sequence (5’-3’)</th>
<th>Target site</th>
</tr>
</thead>
<tbody>
<tr>
<td>Eif4g1_gRNA_U5</td>
<td>TTACTACTATACGAGGAGTT</td>
<td>Chr16:20674569-20674588 (+1)</td>
</tr>
<tr>
<td>Eif4g1_gRNA_U3</td>
<td>AAAGCAGAATTTACCCCA</td>
<td>Chr16:20674745-20674764 (-1)</td>
</tr>
<tr>
<td>Eif4g1_gRNA_D5</td>
<td>CACGATCCTCTCAAGTCGGA</td>
<td>Chr16:20675015-20675034 (-1)</td>
</tr>
<tr>
<td>Eif4g3_gRNA_U5</td>
<td>CTTGTTAAGACGAGTCAGCT</td>
<td>Chr4:138095619-138095638 (-1)</td>
</tr>
<tr>
<td>Eif4g3_gRNA_U3</td>
<td>GTGCTGGGCTCCCGAGGGA</td>
<td>Chr4:138095727-138095746 (+1)</td>
</tr>
<tr>
<td>Eif4g3_gRNA_D5</td>
<td>CGGCAGCCCTTCAAGTGGA</td>
<td>Chr4:138096069-138096088 (-1)</td>
</tr>
<tr>
<td>Eif4g3_gRNA_D3</td>
<td>GTCACTGACACTCTAAGAC</td>
<td>Chr4:138096231-138096250 (+1)</td>
</tr>
<tr>
<td>Eif4g1_gRNA_D3</td>
<td>GATGTTATCTAGCTGTTAGG</td>
<td>Chr16:20675067-20675086 (+1)</td>
</tr>
<tr>
<td>Eif4g1_gRNA_D5</td>
<td>CACGATCCCTTCAAGTACCG</td>
<td>Chr16:20675015-20675034 (1)</td>
</tr>
<tr>
<td>Eif4g1_gRNA_U3</td>
<td>AAAGACGACAATTACCCCA</td>
<td>Chr16:20674745-20674764 (1)</td>
</tr>
<tr>
<td>Eif4g1_gRNA_D5</td>
<td>TTACTACTATACGAGGAGTT</td>
<td>Chr16:20674569-20674588 (1)</td>
</tr>
<tr>
<td>Eif4g1_gRNA_D3</td>
<td>GATGTTATCTAGCTGTTAGG</td>
<td>Chr16:20675067-20675086 (1)</td>
</tr>
<tr>
<td>Eif4g1_gRNA_U3</td>
<td>AAAGACGACAATTACCCCA</td>
<td>Chr16:20674745-20674764 (1)</td>
</tr>
<tr>
<td>Eif4g1_gRNA_D5</td>
<td>CACGATCCCTTCAAGTACCG</td>
<td>Chr16:20675015-20675034 (1)</td>
</tr>
<tr>
<td>Eif4g1_gRNA_U3</td>
<td>AAAGACGACAATTACCCCA</td>
<td>Chr16:20674745-20674764 (1)</td>
</tr>
<tr>
<td>Eif4g1_gRNA_D5</td>
<td>CACGATCCCTTCAAGTACCG</td>
<td>Chr16:20675015-20675034 (1)</td>
</tr>
<tr>
<td>Eif4g1_gRNA_U3</td>
<td>AAAGACGACAATTACCCCA</td>
<td>Chr16:20674745-20674764 (1)</td>
</tr>
</tbody>
</table>

from the Schaffer collateral pathway using a concentric bipolar platinum/iridium electrode (FHC) and stimulator (Multi Channel Systems, Model STG-4002). Field excitatory postsynaptic potentials (fEPSPs) were recorded with a borosilicate electrode (1.0 – 2.5 MΩ tip resistance) filled with aCSF placed in the stratum radiatum 150-200 μm from the stimulating electrode. Responses were collected using WinLTP software (http://winltp.com; (Anderson and Collingridge, 2007)) with low-pass filtering at 4 kHz using a MultiClamp 700B amplifier (Molecular Devices) and digitized at 20 or 40 kHz with a USB-6341 digitizer (National Instruments). At the beginning of each recording, an input-output curve was generated (stimulus intensities: 5, 10 – 70 μA, Δ = 10 μA). Additionally, a paired-pulse curve was generated by delivering two stimuli at varying paired pulse intervals (400, 200, 100, 50 and 40 ms) at a stimulus intensity of 25 μA; paired-pulses were delivered once every 30 s. After collecting basal synaptic properties, the baseline intensity was set to evoke a slope ~40% of the value at which a population spike was first observed. In all cases responses were evoked once every 30 s. A stable baseline was achieved for 30 minutes before delivering LTP induction stimuli. LTP was induced using 100 Hz stimulation for 1 s (tetanic stimulation), delivered either 1 time or 4 times with 5 minutes spacing in between each tetanus train. For DHPG-LTD experiments, bicuculline methiodide (10 μM; HelloBio), picrotoxin (50 μM; HelloBio) and CGP 55845 (100 nM; HelloBio) were included in the aCSF, as well as 2 mM MgCl₂ (instead of 1 mM), to facilitate LTD induction (Palmer et al., 1997). In these experiments, input-output or paired-pulse ratio curve were not performed, and basal responses were set to 1.25 mM sodium phosphate, 2 mM MgCl₂, 1 mM CaCl₂, 0.4 mM sodium ascorbate, and 3 mM sodium pyruvate, and saturated sectioning system in modified aCSF consisting of 180 mM sucrose, 25 mM sodium bicarbonate, 2.5 mM KCl, 10 mM HEPES, 0.2 mM EGTA, 4 mM ATP, 0.3 mM GTP and 10 mM phosphocreatine (pH 7.4, osmolarity: 300 mOsm). Signals were amplified using an Axon Instruments Multiclamp 200B and digitized using an Axon Instruments Digidata 1440A (Molecular Devices). The slope of the fEPSPs was taken from 20 – 60% of the maximum amplitude of each individual response. Data are presented as the average of 2 responses, normalized to the last 10 minutes of the baseline period.

Spontaneous synaptic transmission recordings

Mice were anesthetized with isoflurane, brains were quickly removed and sliced (300 μm thick coronal plane) with a Vibratome sectioning system in modified aCSF consisting of 180 mM sucrose, 25 mM sodium bicarbonate, 25 mM glucose, 2.5 mM KCl, 1.25 mM sodium phosphate, 1 mM MgCl₂, 1 mM CaCl₂, 0.4 mM sodium ascorbate, and 3 mM sodium pyruvate, and saturated with 95% O₂/5% CO₂ (pH 7.4, osmolarity: ~305 mOsm). Pipes were filled for 30 minutes in 35–37°C aCSF containing 125 mM NaCl, 25 mM glucose, 25 mM sodium bicarbonate, 2.5 mM KCl, 1.25 mM sodium phosphate, 1 mM MgCl₂, and 2 mM CaCl₂, and saturated with 95% O₂/5% CO₂ (pH 7.4, osmolarity: ~305 mOsm). Recordings were performed in whole-cell configuration and pipettes were pulled from thin-walled borosilicate glass (TW-150 F, World Precision Instruments) to resistances of 5–7 MΩ with a Sutter Instruments P-87 (Novato). Pipettes were filled with a Sutter Instruments P-87 (Novato). Pipettes were filled with an internal solution containing 130 mM potassium gluconate, 10 mM HEPES, 0.2 mM EGTA, 4 mM ATP, 0.3 mM GTP and 10 mM phosphocreatine (pH 7.4, osmolarity: 300 mOsm). Signals were amplified using an Axon Instruments Multiclamp 200B and digitized using an Axon Instruments Digidata 1440A (Molecular Devices). Recordings were initiated 10 minutes after membrane rupture. EPSCs were recorded by voltage-clamping the membrane potential at ~70 mV (the standard reversal potential of IPSCs) and inhibiting GABA A receptor mediated currents with 20 μM bicuculline (Sigma-Aldrich). IPSCs were recorded by voltage-clamping the membrane potential at +10 mV and inhibiting glutamatergic transmission with 10 μM 6-cyano-7nitroquinoxaline-2,3-dione (CNQX; TOCRIS) and 50 μM DL-2-amino-5-phosphonovaleric acid (AP-5; TOCRIS).
Miniature postsynaptic currents were recorded by the same method described above, with the addition of 1 μM tetrodotoxin (TTX; Affix scientific) to inhibit action potentials. All inhibitors were bath applied to slices at a perfusion rate of ~1 ml/min.

Recordings were analyzed offline using Clampfit 10.7 software (Molecular Devices; Sunnyvale, CA, USA). sEPSCs, mEPSCs, sIPSCs and mIPSCs were analyzed for amplitude and frequency. A separate template was created for each recording by averaging typical currents with varying detection thresholds. Data were extracted using template-based detection, and subsequently verified manually for precision.

Behavioral tests
Mice were housed on a 7:00am-7:00pm light cycle in TCP. Subject mice were handled for 1 minute on 3 consecutive days prior to first behavioral test and were allowed to acclimate to a behavioral test room for at least 30 minutes prior to tests. The age of mice was between 8 to 20 weeks old. For all tests a minimum of 8 females and 8 males were assessed per genotype.

Open field test
Mice were placed in an open field chamber (40 cm X 40 cm x 40 cm) made of plexiglass for 10 minutes. The open field chamber was in a sound attenuation box with dimmed light (75 lux). In order to test locomotor activity, mice were tracked using Activity Monitor software (Med Associates Inc.). The travel distance and time spent in center zone were analyzed.

Three-chamber social interaction test
A 60 cm x 40 cm three-chamber apparatus was located under dim light (5 lux). First, for habituation, mice explored all three chambers for 5 minutes, and then for the sociability test, once an object (an orange lid) in an empty cup had been placed in one of the side chambers and an age-, sex- and genetic background matched unfamiliar C57BL/6J mouse in the other side chamber, mice explored the chambers for 10 minutes. Finally, for the social novelty test, a novel mouse replaced the object and mice re-explored the chambers for 10 minutes. To exclude special preference, the chambers containing the object and the mouse were occasionally switched. Mice were video-tracked using the Ethovision XT software (Noldus). The time of direct interaction was measured as the amount of time the nose of the subject mouse was located within 3cm around the cup containing the stranger mouse, the familiar mouse or the object.

Reciprocal interaction
Mice were allowed to acclimate to a behavioral test room at least 30 minutes prior to tests. Two male stranger mice of the same genotype were placed together at the same time in a clean cage and allowed to interact for 5 minutes. The interactions were video recorded and then scored for the time of sniffing and touching manually in a blind condition. Statistical differences were assessed using unpaired t test.

Elevated zero maze test
Mice were placed in the middle of an open arm in the elevated zero maze (23 cm inner radius and 7 cm platform width) and their movements were video-tracked with Ethovision XT software (Noldus) for 5 minutes under dim lighting (5 lux). Time spent in open arms, the total distance traveled and the number of transitions between open and closed arms were analyzed.

Fear conditioning
During the first day mice were trained by a 2 minute period of acclimatizing to a new chamber, followed by one pairing of a tone (2800 Hz, 85 dB, 30 s) with a co-terminating foot-shock (0.75 mA for 2 s). The mice remained in the chamber for an additional 1 minute interval after the end of the last pairing, after which they were returned to their home cages. Contextual fear conditioning was assayed 24 hours after training by placing the animals in the conditioning context (same chamber where the training occurred) for a 5 minute period, during which the incidence of freezing (immobile except for respiration) was monitored using Video Freeze software (Med Associates Inc.). Auditory fear conditioning tests were performed 48 hours after training. The mice were placed into a new chamber (a triangle insert and a smooth-surfaced plastic sheet with 0.5% Acetic acid as an odor). After a 2 minute acclimation period to the new chamber (context), mice were exposed to the trained tone for 3 minutes. The fear conditioning was performed on male mice and the freezing behavior was monitored using Video Freeze software (Med Associates Inc.). Statistical analysis was based on two-way ANOVA and between-group comparisons by Tukey’s post hoc test.

Marble burying test
The test apparatus for marble burying was a cage (18 cm x 36 cm x 13 cm) with 20 glass marbles (4 rows x 5 columns) on 5 cm-depth bedding surface. Mice were allowed to explore the test box for 30 minutes in a darkened room, and the number of marbles that had 50% of their surface covered by bedding was quantified.

Startle response test
Mice were placed in a recording tube in a sound attenuation chamber of the SR-LAB System (San Diego Instruments). After first 5 startle exposures, mice were randomly presented with a total of 60 trials: pre-pulse alone, pre-pulse + startle, startle alone and no stimulus. Four pre-pulse intensities were presented: 70, 75, 80 and 85 db. Pre-pulse duration was 20 ms, and pre-pulse and startle intervals ranged from 50 to 120 ms. Startle intensity was set at 120 dB for 40 to 60 ms. Each type of trial was performed 6 times. Pre-pulse inhibition (PPI%) was calculated as follows: %PPI = 100 x (S – PPi_S)/S. S and PPi_S stand for the peak amplitude of startle and pre-pulse + startle, respectively. The global level of PPI was calculated as followed; 100 x [S - (PP1_S + PP2_S + PP3_S + PP4_S)/4]/S.
**Rotarod**
Mice were placed on the rotating rod of a rotar-rod apparatus (Harvard Apparatus) for 5 minutes. The rod was accelerating from 4 to 40 rpm. The time spent on the rod was measured for 3 trials per day for 3 consecutive days.

**Resident intruder test**
The subject male mice were isolated in a home cage individually for 14 days prior to tests. The singly housed mice with their home cages were transferred to a test room under dim light (10 lux). An age-matched intruder C57BL/6J male mouse was placed into the subject’s home cage for 10 minutes, during which time the aggressive behavior between the subject and the intruder mouse was video-recorded by Ethovision XT software (Noldus). The number of fights, latency to start fighting and frequency of fighting were quantified under experimenter-blinded condition.

**Auditory brainstem response (ABR)**
Mice were anesthetized and eye drops such as Viscotears were applied to avoid drying of the cornea. Needle electrodes were placed under the skin at the top/vertex of the head (active electrode) and overlying the left bulla (bony sheath enclosing the middle ear by the mastoid bone) (reference electrode) and right bulla (ground electrode). The animal was placed within a special acoustic chamber in a natural prone position, paws forward, facing the loudspeaker at a distance of 10 cm from the leading edge of the speaker to the mouse’s interaural axis. To record the ABR, a series of sound stimuli (6, 12, 18, 24, and 30 KHz) was presented to the mouse from the loudspeaker that included tone pips between the frequencies 6 and 30kHz and over levels of 0 - 85 dB (in 5dB increments).

**QUANTIFICATION AND STATISTICAL ANALYSIS**

**Identification of conserved tissue-specific translation factors**
Splicing events were defined as tissue-regulated based on a survey of RNA-Seq data from 32 human tissues or 30 mouse tissues with at least two independent samples representing each tissue (see Table S1). To comprehensively detect and quantify all AS events involving alternative exons, we used the vast-tools multi-module analysis pipeline, as previously described (Irimia et al., 2014; Tapial et al., 2017). Briefly, reads were initially mapped to genome assemblies using Bowtie, using –m 1 –c 2 parameters with reads that mapped to the genome discarded for AS quantifications. Unique EEJ (exon-exon junction) libraries were generated to derive measurements of exon inclusion levels using the metric “Percent Spliced In” (PSI). This utilized all hypothetically possible EEJ combinations from annotated and de novo splice sites, including cassette, mutually exclusive and microexon events (Irimia et al., 2014; Tapial et al., 2017).

Two criteria were used to score an AS event as tissue-dependent. First, at least 20 mapped reads were required to support the event. Second, a statistically robust PSI difference of > 20% between a given tissue and all other tissue samples was required, as measured using the diff module of vast-tools for differential PSI quantification (https://github.com/vastgroup). An event was considered conserved if it displayed differential splicing between equivalent tissues in both human and mouse samples. Only annotated genes within the GO term class “translation” (GO:0006412) were considered as “Translation Factors.” The heatmap was constructed using the R and after scaling by row (subtracting the mean (centering) and dividing by standard deviation (scaling)).

**PsychENCODE data analysis**
Data for the Figures 2F and S2I-G were obtained from NIMH Repository & Genomic Resources, a centralized national biorepository for genetic studies of psychiatric disorders. RNA-Seq data from autism spectrum disorder and control post-mortem samples were downloaded from the UCLA-ASD’s PsychENCODE data storage. When fastq files were not available, bam files were converted to fastq using samtools (“samtools sort” followed by “samtools fixmate” and finally “samtools fastq” (with –t and –O tags)). Files were assessed using fastqc and adapters identified removed by Trimmomatics. Gene expression and alternative splicing quantification were performed using Whippet (https://github.com/timbitz/Whippet.jl) with default setting and an index (~suppress-low-tsl) based on Ensembl hg38. Microexons were extracted based on size. Median of controls was calculated and subtracted from each ASD sample to identify the most strongly disrupted microexons.

**Computational detection of prion-like and intrinsically disordered regions in eIF4G proteins**
The web-based prion prediction algorithm PLAAC (Prion-like amino acid composition) was applied to detect prion-related domains in human eIF4G1 and eIF4G3 (Lancaster et al., 2014) with amino acid distribution estimated using the setting (α 100) for weighting of background probabilities. PLAAC is also a predictor for phase separation (Vernon and Forman-Kay, 2019). To detect low complexity regions in human eIF4G1 and eIF4G3 we applied DISOPRED3 using the PSIPRED protein analysis workbench (Jones and Cozzetto, 2015). Figures 6A and S6A were generated as described in Bakthavachalu et al. (2018).

**Quantitative proteomics data analysis**
The Proteome Discoverer 2.1.0.81 software (Thermo Fisher Scientific) was used for analyses of RAW files. Searching was performed using The SEQUEST HT database search engine and the mouse proteome database (SwissProt TaxID = 10090, 2016-05-11) downloaded from ProteinCenter supplemented with a list of common contaminants derived from ProteinCenter (PD_Contaminants_2014_11). Both the forward and reversed databases were used for database searches in order to evaluate the false discovery rates (FDRs); 50,346 total entries were searched. Trypsin was selected as the digestion enzyme and a maximum number of two missed

Molecular Cell 77, 1176-1192.e1–e16, March 19, 2020  e13
cleavages were allowed. The mass tolerance of precursor ions and fragment ions were set as 10 ppm and 0.08 Da, respectively. The iTRAQ 8 plex (peptide N terminus and lysine) and carboxamidomethylation (cysteine) were set as the static modifications. Acetylation (Protein N-terminus), Oxidation (methionine) and deamination (Asparagine or Glutamine) were chosen as the dynamic modifications. The Percolator tool integrated in the Proteome Discoverer 2.1 software was used to validate the database search results based on the q-value. The identifications were filtered with peptide confidence value as high to obtain FDR less than 1% on the peptides level. Protein grouping was enabled, and the strict parsimony principle was applied.

Proteins identified and quantified in all three biological replicates were delivered to Benjamini-Hochberg (BH) FDR estimations, and those that passed the 1% BH-FDR threshold and with a log2 ratio < −0.25 or > 0.25 were considered as significant changes.

Mass spectrometry data generated in this study have been deposited at the ProteomeXchange consortium through partner MassIVE (https://massive.ucsd.edu/ProteoSAFe/static/massive.jsp). The data have been assigned ID MSV000084658.

**Heatmap of MS data**

Only detected proteins with a 1% pass rate were considered. The normalized abundance ratios (log(DKO/WT+1)) were calculated. The heatmap was generated using R from the abundance ratio of DKO/WT for each replicate.

**Functional enrichment analysis**

Functional enrichment analysis was performed using the functional enrichment analysis tool g:Profiler (https://biit.cs.ut.ee/gprofiler) (Reimand et al., 2016). Structured controlled vocabularies from Gene Ontology, as well as information from the curated KEGG and Reactome databases were included in the analysis. Only functional categories with more than three members and fewer than 2,000 members were included in the analysis. Significance values were assessed using the hypergeometric test with multiple testing correction by the Benjamini and Hochberg method. The Cytoscape plug-in Enrichment map (http://baderlab.org/Software/EnrichmentMap) (Isserlin et al., 2014) was used to visualize and arrange functional data. In enrichment maps, node size is proportional to the number of genes associated with the GO category, and edge width is proportional to the number of genes shared between GO categories. Node clusters are arranged to contain similar functional terms.

For enrichment analysis, rat genes enriched in the neuropil were extracted from data supplied by Cajigas et al. (2012). BioMart was used to identify one-to-one rat to mouse orthologs. Synaptic-related genes were downloaded from SynaptomeDB (http://metamoods.org/SynaptomeDB) (Pirooznia et al., 2012). Statistical enrichment was assessed using Fisher’s exact test and corrected for multiple hypothesis testing (when applicable).

**BioID mass spectrometry data analysis**

Mass spectrometry data was stored, searched and analyzed using the ProHIlts laboratory information management system (LIMS) platform (Liu et al., 2016). Briefly, the WIFF files were converted to MGF format using WIFF2MGF and subsequently converted to an mzML format using ProteoWizard (3.0.4468) and the AB SCIEX MS Data Converter (V1.3 beta). The mzML files were next searched using Mascot (v2.3.02) and Comet (2014.02 rev.2) against a protein database of 58,206 total proteins consisting of the NCBI mouse RefSeq database (v53, Sep 9th, 2015, forward and reverse sequences) supplemented with “common contaminants” from the Max Planck Institute (https://141.61.102.106:8081/share.cgi?ssid=0f2gfuB) and the Global Proteome Machine (https://www.thegpml.org/crap/index.html) as well as sequences from common fusion proteins and epitope tags. Search parameters were set for tryptic cleavages, were set for 2 missed cleavage sites per peptide, MS1 mass tolerance of 40 ppm with charges of 2+ to 4+ and an MS2 mass tolerance of ± 0.15 amu. Asparagine/glutamine deamidation and methionine oxidation were selected as variable modifications. The results from each search engine were jointly analyzed through the Trans-Proteomic Pipeline (TPP) (Deutsch et al., 2010) via the iProphet pipeline (Shteinberg et al., 2011). A minimum iProphet probability of 0.95 was required for protein identification with 2 unique peptides required for protein interaction scoring.

Significance Analysis of INTeractome (SAINTexpress version 3.6.1) (Teo et al., 2014) was used as a statistical tool to calculate the probability of each potential proximity interaction compared to control samples (BirA*-Flag-GFP, and BirA*-Flag-empty, each run in two biological replicates). For each biological replicate, counts for each protein in a given bait sample were analyzed independently against control samples before averaging of the score values and assessment of the Bayesian False Discovery Rates (BFDR) (Teo et al., 2014). High-confidence interactions are defined as those with FDR ≤ 5%.

Mass spectrometry data generated in this study, including the complete SAINTexpress results, have been deposited at the ProteomeXchange consortium through partner MassIVE (https://massive.ucsd.edu/ProteoSAFe/static/massive.jsp). The data have been assigned IDs MSV000083349 and PXD012421.

**BioID data visualization**

Spectral counts for both eIF4G1 splice variants BioID samples were calculated by averaging counts for both biological replicates and subtracting the average counts across all controls. To be included in the analysis, a protein was required to be identified with a minimum of 10 average counts and meeting the 5% BFDR threshold in at least one of the isoforms. For the bar plots, the proteins were sorted according to the average spectral count ratios across the eIF4G1+MIC and eIF4G1ΔMIC BioID experiments. Each protein was annotated according to whether it is an mRNP cytoplasmic granule protein based on recently published datasets (Youn et al., 2018) (Figure 6B), and significant differences in the spectral count ratios of proteins detected as associated with cytoplasmic granule components were compared using the Mann-Whitney U-test. BioID spectral count data for both eIF4G1 splicing variants was visualized
as a ‘dotplot’ (Figure 6B) representation using the “dotplot” generator tools available through ProHits-viz (Knight et al., 2017). Color intensity represents the averaged count data across replicates, the circle edge shows the respective BFDR thresholds as calculated by SAINTexpress, and the size of the circle is proportional to the relative counts between splice variants (after averaging counts between replicates) with the larger size representing the maximum value.

For all positions in a protein a score for intrinsic disorder was computed using IUPred (http://iupred.ence.hu). Amino acid residues with a score larger than 0.5 were considered disordered. Prion-like regions were identified using PLACC (prion-like amino acid composition) with amino acid distribution estimated using the setting (x 50) for weighting of background probabilities (Lancaster et al., 2014). Proteins preferentially associated with the eIF4G1+MIC isoform were defined as those that have at least a 1.25-fold difference in the ratio of spectral counts between the +MIC/ΔMIC splice variants. Proteins preferentially associated with eIF4G1ΔMIC were defined as those that have a ratio of spectral counts between the +MIC/ΔMIC splice variants of less than one.

**ICLIP-Seq analysis**

Analysis of ICLIP-Seq data was performed as previously described (Han et al., 2017). 51-nt raw reads that consisted of 3 random positions, a 4-nt multiplexing barcode, and another 2 random positions, followed by the cDNA sequence, were initially de-duplicated based on the first 45 nt. Reads were de-multiplexed and the random positions, barcodes, and any 3’-bases matching Illumina adapters were removed. Finally, reads shorter than 25 nt were filtered out and remaining reads trimmed to 35 nt. These steps were carried out using Trimmmomatic (Bolger et al., 2014). Remaining reads were mapped to the mouse genome/transcriptome (Ensembl annotation of NCBI37) using tophat (Trapnell et al., 2009) with default settings. To prevent false assignments of reads from repetitive regions, any reads with a mapping quality < 3 were removed from further analysis. Plots showing average crosslinking signal of events aligned to exon boundaries were generated as described after first reducing reads to their first position, which is adjacent to the crosslink position. A 21-bp running window average was used for plotting, and average signals across replicates are shown.

**RNA sequencing and ribosome profiling data analysis**

All fastq files were quality checked using FastQC. Trimmomatic was used to remove adaptors and low-quality reads. Reads were first aligned with Bowtie to human or mouse rRNA sequences, and matches were discarded (–v 3). Processed RNA-seq reads were aligned to the mouse genome (mm10) using whole genome alignment by STAR with 2-pass setting enabled and output in bam format. In addition, mRNA expression levels were calculated using kallisto (Bray et al., 2016) with transcripts downloaded from Ensembl mm10 and filtered for support level > 3. Kallisto was run on default settings with k = 25 and pseudobam output.

For differential mRNA expression analysis, read counts for each gene were calculated using vast-tools (irimia et al., 2014; Tapial et al., 2017) and analyzed with DESeq (Anders and Huber, 2010) (unless mentioned otherwise: adjusted p value < 0.05 and > 0.5 absolute fold change). Genes showing differential regulation during KCl-mediated activity dependence were identified by reanalyzing raw sequencing data from Quesnel-Vallières et al. (2016) using the vast-tools pipeline. Statistical enrichment of genes with transcript level changes in Eif4g DKO neurons with those regulated by KCl treatment was assessed using Fisher’s exact test.

To detect changes in translational efficiency per gene, RiboDiff (Zhong et al., 2017) was employed with default settings using combined read counts from Kallisto. The R package RiboSeqR (https://bioconductor.org/packages/release/bioc/html/riboSeqR.html) was used to identify reading frames and confirm the periodicity of the ribosome profiling data. Default settings were used in RiboSeqR with only canonical start and stop codons considered. To identify pausing sites, the algorithm PausePred (Kumari et al., 2018) was used with default settings. Only protein-coding genes within annotated transcripts were considered. Only genes with a 50% increase/decrease in number of pausing events were included in downstream analysis.

**Overlap of paused genes in Eif4g DKO and Fmr1 KO neurons**

Ribosome profiling data from Eif4g DKO neurons (this study) and Fmr1 KO adult hippocampal neural stem cells (Liu et al., 2018; GSE112502 samples) were analyzed using the PausePred (Kumari et al., 2018) algorithm (after trimming/rRNA removal and STAR alignment) with the same cutoffs as described above. Sites identified as paused in both datasets were compared to a background set of genes, corresponding to those genes detected in the ribosome profiling data (TPM > 1) and analyzed by PausePred.

High confidence FMRP targets were determined as those with at least 10 FMRP HITS-CLIP peaks in brain tissue (Maurin et al., 2018). Statistical enrichment was assessed using Fisher’s exact test and corrected for multiple hypothesis testing.

**Motif enrichment in Eif4g DKO paused genes**

High quality pause sites (score > 50) identified by PausePred (Kumari et al., 2018) were extracted and sequences 50 nucleotides upstream and downstream used in the analysis. CentriMO (Bailey and Machanick, 2012) was used to search for motif enrichment (–local). No hits (adj-p < 1e-3) were found when upstream sequences were analyzed. However, strong motif enrichments (adj-p < 1e-3) were identified when analyzing sequences downstream of the pause sites, both when using the default background (data not shown) and when using sequences upstream of the pause site as background (Figure 6F). In Figure 6F only recognition motifs corresponding to RNA binding proteins involved in translation regulation (GO:0006417) were analyzed (Ray et al., 2013).
Quantification of stalled ribosomes from micrographs

Images of 53 WT and 53 Eif4g microexon-DKO neurons from two independent preparations were captured using a Zeiss fluorescent microscope and Zen software. Images were analyzed in batches with “find maxima” function of ImageJ to detect and mark stalled ribosome granules. Reported counts were then manually corrected to remove counts from neighboring cells. Data were analyzed for significant differences using Mann-Whitney U test.

Foci Quantification

The transfected SH-SY5Y cells were binned into two groups: those showing diffuse cytoplasmic distribution of mCherry signal (“Dispersed”) and those with concentrated puncta-like mCherry signal (“Foci”). For each construct, over 400 transfected cells were analyzed (total cells: eIF4G1+MIC, 420 cells; eIF4G1ΔMIC 581 cells) and experiments were performed in triplicate. Data were analyzed using a two-way analysis of variance (ANOVA) and Tukey’s post hoc test.

Statistics for behavioral testing

One-way ANOVA, Two-way ANOVA, or two-tailed t test were performed for all the behavior analysis.

Related to 3-chamber sociability test (Figure S4J)

Two-way ANOVA chamber main effect, F (2, 120) = 228.3, p < 0.0001.
Tukey’s post hoc test:
- Male WT: Object versus Male WT: Mouse p = 0.0019,
- Male Eif4g1ΔMIC/ΔMIC: Object versus Male Eif4g1ΔMIC/ΔMIC: Mouse p = 0.0040,
- Female WT: Object versus Female WT: Mouse p = 0.0203,
- Female Eif4g1ΔMIC/ΔMIC: Object versus Female Eif4g1ΔMIC/ΔMIC: Mouse p = 0.0001.

Related to 3-chamber social novelty test (Figures 4B and S4K)

Two-way ANOVA chamber main effect F (2, 66) = 79.54, *** P < 0.0001.

Figure 4B: Tukey’s post hoc test:
- Male WT: Familial mouse versus Male WT: Stranger mouse p = 0.0196;
- Male Eif4g1ΔMIC/ΔMIC: Familial mouse versus Male Eif4g1ΔMIC/ΔMIC: Stranger mouse p = 0.9133,

Figure S4K: Tukey’s post hoc test:
- WT:Familiar mouse versus WT: Stranger mouse p = 0.0196,
- Eif4g1ΔMIC/+: Familial mouse versus Eif4g1ΔMIC/+: Stranger mouse p = 0.2095,
- Eif4g1ΔMIC/ΔMIC: Familial mouse versus Eif4g1ΔMIC/ΔMIC: Stranger mouse p = 0.9133.

Related to reciprocal interaction test (Figure 4C)

1 min: Male Eif4g1ΔMIC/ΔMIC versus Male Eif4g1ΔMIC/ΔMIC nose-to-nose sniffing; ** p = 0.0011,
1 min: Male WT versus Male WT touching frequency; p = 0.5570,
5 min: Male Eif4g1ΔMIC/ΔMIC versus Male Eif4g1ΔMIC/ΔMIC nose-to-nose sniffing; * p = 0.0384,
5 min: Male WT versus Male WT touching frequency; p = 0.17.

Related to fear conditioning test (Figure 4E)

Contextual fear conditioning: 2-way ANOVA genotype main effect, F (2, 126) = 132.1;
Tukey’s post hoc test:
- 1 min: Male WT versus 1 min: Male Eif4g1ΔMIC/ΔMIC *** p = 0.0036;
- 1 min: Male WT versus 1 min: Male Eif4g1ΔMIC/+ ** p = 0.0057;
- 2 min: Male WT versus 2 min: Male Eif4g1ΔMIC/ΔMIC ** p = 0.0035.
Cued fear conditioning: Two-way ANOVA time main effect, F (4, 165) = 13.74, p < 0.0001; genotype main effect F (2, 165) = 1.013, p = 0.3653.

DATA AND CODE AVAILABILITY

The accession numbers for Illumina iCLIP-Seq, RNA-Seq and Ribo-Seq data reported in this study are Gene Expression Omnibus (GEO): GSE141594, GSE141599 and GSE141599, respectively. Mass spectrometry data generated in this study have been deposited at the ProteomeXchange Consortium through partner MassIVE and have been assigned IDs MSV00083349 / ProteomeXchange: PXD012421 (BioID) and MSV00084658 (ITRAQ). Scripts are available upon request.