Journal of Neurochemistry

JOURNAL OF NEUROCHEMISTRY | 2016 | 139 | 349–368 doi: 10.1111/jnc.13768

ORIGINAL ARTICLE

Preferential binding of a stable G3BP ribonucleoprotein complex to intron-retaining transcripts in mouse brain and modulation of their expression in the cerebellum

Sophie Martin,* Nicolas Bellora,†,‡ Juan González-Vallinas,† Manuel Irimia,§ Karim Chebli,* Marion de Toledo,* Monika Raabe,¶ Eduardo Eyras,†,** Henning Urlaub,¶ Ben J. Blencowe§ and Jamal Tazi*

*Institut de Génétique Moléculaire de Montpellier, CNRS UMR5535, Montpellier, France

†Computational Genomics Group Universitat Pompeu Fabra PRBB, Barcelona, Spain

‡Laboratorio de Microbiología Aplicada y Biotecnología, Instituto Andino-Patagónico de Tecnologías Biológicas y Geoambientales (IPATEC), CONICET – UNComahue, Bariloche, Argentina

§The Donnelly Centre, University of Toronto, Toronto, Ontario, Canada

¶Max Planck Institute for Biophysical Chemistry, Göttingen, Germany

**Catalan Institution for Research and Advanced Studies (ICREA), Passeig Lluís Companys 23, Barcelona, Spain

Abstract

Neuronal granules play an important role in the localization and transport of translationally silenced messenger ribonucleoproteins in neurons. Among the factors associated with these granules, the RNA-binding protein G3BP1 (stress-granules assembly factor) is involved in neuronal plasticity and is induced in Alzheimer's disease. We immunopurified a stable complex containing G3BP1 from mouse brain and performed high-throughput sequencing and cross-linking immunoprecipitation to identify the associated RNAs. The G3BP-complex contained the deubiquitinating protease USP10, CtBP1 and the RNA-binding proteins Caprin-1, G3BP2a and splicing factor proline and glutamine rich, or PSF. The G3BP-complex binds preferentially to transcripts that retain introns, and to non-coding sequences like 3'-untranslated region and long non-coding RNAs. Specific transcripts with retained introns appear to be enriched in the cerebellum compared to the rest of the brain and G3BP1 depletion decreased this intron retention in the cerebellum of G3BP1 knockout mice. Among the enriched transcripts, we

Received March 15, 2016; revised manuscript received August 2, 2016; accepted August 2, 2016.

Address correspondence and reprint requests to Prof. Jamal Tazi, Institute of Molecular Genetics, 1919 Route de Mende, 34293 Montpellier, France. E-mail: jamal.tazi@igmm.cnrs.fr

Abbreviations used: AMPA, α-amino-3-hydroxy-5-methyl-4-isoxazolepropionic acid; *Btbd9*, BTB/POZ domain containing 9; *Cadm2*, cell adhesion molecule 2; CAGE, cap analysis gene expression; CBC, cap-binding complex; CDS, coding sequence; CNS, central nervous system; CtBP1, C-terminal binding protein 1; *Ctnnd2*, catenin delta 2; DAVID, database for annotation, visualization and integrated discovery; DHPG, (S-3,5-dihydroxyphenylglycine; EASE, expression analysis systematic explorer; *Eif3m*, eukaryotic translation initiation factor 3, subunit M; eIF4E, eukaryotic translation initiation factor 4E; FUS/TLS, fused in sarcoma/translocated in liposarcoma; G3BP, RasGAP SH3 domain binding protein; *Gas5*, growth arrest-specific 5; GEM, genome multitool; GO, gene ontology; *Gria2*, glutamate receptor, ionotropic, ampa2; Grm, glutamate receptor, metabotropic; HITS-CLIP, high-throughput sequencing and crosslinking immunoprecipitation; *Hprt1*, hypoxanthine phosphoribosyltransferase 1; KO, knock-out; *Kpna1*, karyopherin alpha 1; lncRNA, long non-coding RNA; LTD, long-term depression; *Malat1*, metastasis-associated lung adenocarcinoma transcript 1; *Meg3*, maternally expressed 3; mGluR, metabotropic glutamate receptor; miRNA, microRNA; mRNP, messenger ribonucleoprotein; NMD, nonsense-mediated mRNA decay; NTF2, nuclear transport factor 2; PSF, polypyrimidine tract-binding protein-associated-splicing factor; PTBP1, polypyrimidine tract-binding protein 1; PTC, pre-termination codon; *Rbfox1*, RNA-binding protein, Fox-1 homolog; RNA-BP, ribonucleic acid-binding protein; RNP, ribonucleoprotein; RRM, ribonucleic acid recognition motif; *Sae1*, SUMO-activating enzyme subunit 1; SELEX, systematic evolution of ligands by exponential enrichment; SFPQ, splicing factor proline and glutamine rich; SG, stress granule; SINE, short interspersed nuclear elements; snoRNA, small nucleolar RNA; TDP-43, TAR DNA-binding protein 43; *Trpm3*, transient receptor potential cation channel, subfamily M, member 3; U2AF, U2 auxiliary factor; USP10, ubiquitin specific peptidase 10; UTR, untranslated region; WT, wild-type.

found an overrepresentation of genes involved in synaptic transmission, especially glutamate-related neuronal transmission. Notably, G3BP1 seems to repress the expression of the mature *Grm5* (metabotropic glutamate receptor 5) transcript, by promoting the retention of an intron in the immature transcript in the cerebellum. Our results suggest that G3BP is involved in a new functional mechanism to regulate non-coding RNAs including intron-retaining transcripts, and thus have broad implications for neuronal gene regulation, where intron retention is widespread.

Keywords: cerebellum, G3BP, HITS-CLIP, retained intron, stress granules.
J. Neurochem. (2016) **139**, 349–368.

Stress granules (SGs) are non-membranous cytoplasmic foci formed as a cellular protective response to environmental stress, such as elevated temperature, oxidative stress, hypoxia, osmotic shock, UV irradiation, glucose deprivation, or viral infection (Thomas *et al.* 2011). They can be induced chemically by treatment with compounds like sodium arsenite, which triggers oxidative stress. The primary effect of SGs is to accumulate stalled translation arrested messenger ribonucleoprotein (mRNPs) (Kedersha and Anderson 2009), providing a mechanism to sequester untranslated mRNA from the translational machinery until the stress has cleared. The dynamic shuttling of SG-associated mRNAs and proteins has suggested that SGs may also be sites of mRNA triage at which untranslated mRNAs are sorted and processed for either reinitiation, degradation, or packaging into stable non-polysomal mRNP complexes (Anderson and Kedersha 2009). The dynamic assembly of SGs is achieved at least in part by aggregation of specific RNA-binding proteins (RNA-BPs) that act downstream of translation repression. RasGAP SH3 domain binding protein (G3BP) is an important component of the assembly of these SGs (Tourrière *et al.* 2003) where it localizes in large cytoplasmic granules containing poly(A)+ RNAs .

G3BP is an evolutionarily conserved RNA-BP that was initially characterized through its interaction with a Ras-GTPase activating protein (RasGAP p120; Parker *et al.* 1996), but this interaction was revisited (Annibaldi *et al.* 2011). The G3BP family includes two members in mammals, G3BP1 (referred to as G3BP) and G3BP2 (Kennedy *et al.* 2001). Both proteins co-localize in SGs, when cells are subjected to stress (Kobayashi *et al.* 2012). The two proteins are encoded by genes on human chromosomes 5 and 4 and mouse chromosomes 11 and 5, respectively. They have an identical intron/exon structure and most likely have arisen from whole genome duplications at the base of vertebrates. The G3BP2 pre-mRNA is alternatively spliced to give G3BP2a and a shorter isoform G3BP2b lacking 33 amino acids in the central region (482 and 449 amino acids, respectively) (Kennedy *et al.* 1996, 2001). G3BP1 and G3BP2 proteins orthologs are highly conserved between human and mouse (~ 95% sequence identity), and also share high sequence similarity between the two paralogs (more than 60% identity). The N-terminal nuclear transport factor 2 (NTF2) domain which is the most highly conserved domain in G3BPs, is found in other factors including NTF2 itself and the highly related export receptor TAP and its co-factor, p15, which co-operatively function in nuclear mRNA export (Suyama *et al.* 2000; Stutz and Izaurralde 2003; Kristensen 2015). The crystal structure of this domain has been determined for *Drosophila* and human G3BP (Vognsen *et al.* 2011; Vognsen and Kristensen 2012) and is shown to have similarities with the other NTF2 domains. Accordingly, the NTF2 domain of G3BP influences the cellular localization of the protein and its oligomerization with itself or with other partners, as it is suggested to permit the formation of dimers and trimers of G3BP (Kent *et al.* 1996; Tourrière *et al.* 2003). The central region of G3BP proteins comprises a less conserved acid-rich domain, followed by proline-rich (PxxP) motifs, in varying numbers depending on the G3BP isoform. Finally, the G3BP C-termini comprises two motifs associated with RNA binding: the canonical RNA-recognition motif with conserved RNP1 and RNP2 motifs, followed by an arginine-glycine rich box.

G3BP was initially identified based on an intrinsic endonuclease activity thought to regulate mRNA stability in SGs (Tourrière *et al.* 2001). However, a plethora of additional cellular functions has been attributed to G3BP including its role as a DNA/RNA helicase VIII (Costa *et al.* 1999), as a regulator of the activity of ubiquitin protease through its association with ubiquitin specific peptidase 10 (USP10) (Soncini *et al.* 2001), or controlling viral infection, in particular vaccinia viruses where it acts as a partner of Caprin-1, a well-conserved cytoplasmic phosphoprotein that is needed for normal progression through the G1-S phase of the cell cycle (Solomon *et al.* 2007). However, diverse systems have been used to demonstrate these activities and as a result, it is not clear whether all these activities are required for G3BP function(s) *in vivo*.

Disruption of G3BP in mice has proven that both developmental growth and survival are critically dependent on G3BP levels, as homozygous null mutations of the *G3bp1* gene in 129/Sv mice induce embryonic lethality (Zekri *et al.* 2005). However, viable homozygous knockout mice with a mixed genetic background (50% 129/Sv / 50% Balb/c) were recently generated, allowing for a functional characterization

of the G3BP1 knockout (KO) mice (Martin *et al.* 2013). These mice displayed behavioral defects linked to the central nervous system (CNS) associated with an ataxia phenotype, demonstrating a major role for G3BP in the CNS. Consistently, G3BP deficiency leads to altered neuronal plasticity and calcium homeostasis, establishing a direct link between SG formation and neurodegenerative diseases (Martin *et al.* 2013). However, how G3BP can mediate these functions in the brain is presently unknown.

Here, we have used stringent conditions to immuno-purify G3BP ribonucleoprotein complexes from mouse brain and established their composition. High-Throughput Sequencing and Crosslinking Immunoprecipitation (HITS-CLIP) of G3BP demonstrates that both G3BP1 and G3BP2 are largely associated with non-coding intronic sequences from unspliced transcripts. The expression levels of G3BP targets in the cerebrum and the cerebellum of wild-type (WT) and G3BP1 KO mice reveal unexpected regulation of gene expression in the brain through association of a G3BP-containing complex with intronic sequences. Thus, our results suggest new mechanisms to explain the phenotypes associated with G3BP1 deficiency, namely the ataxia and the exacerbated neuronal responses.

Materials and methods

Animals and ethics statement

All animal procedures were executed according to the European Directive 2010/63/UE. The mice were maintained under pathogen-free conditions in our animal facility (E34-172-16), and the experiments were conducted by authorized personnel. The study plan was approved by the Institutional Review Board at the Animal Facility of the Institut de Génétique Moléculaire de Montpellier and the Regional Ethics Committee for Animal Experimentation of Languedoc-Roussillon (agreement no. CEEA-LR-1061).

WT and KO mice (5–8-week-old, combination male/female in same ratio) from a mixed Balb/c/129/Sv genetic background (Martin *et al.* 2013) were used in all the experiments.

CLIP and HITS-CLIP

Cross-linking and immunoprecipitation experiments were performed as previously described (Wang *et al.* 2009; Macias *et al.* 2012). We performed UV crosslink on WT and KO mouse (5–8-week-old) brain or cerebellum homogenates followed by immunoprecipitation with an anti-G3BP1 antibody (Abnova, Taipei, Taiwan) or a control IgG. Following highly stringent washes, the samples boiled in NuPAGE loading buffer were separated on NuPAGE (10% or 4–12% gradient) (Life Technologies, Grand Island, NY, USA) and transferred onto nitrocellulose. Proteins were analyzed from the gels or the nitrocellulose membranes. Radioactively labeled RNAs were extracted from nitrocellulose and pieces of the membrane were treated with proteinase K, followed by phenol-chloroform extraction and purification on a polyacrylamide gel. Ligation of RNA linkers and RT-PCR amplification steps were performed at the Ultrasequencing Unit from Montpellier GenoMix (MGX, Montpellier, France), for the first experiment, and GATC Biotech (Konsantz, Germany) for the second experiment, using Illumina Small RNA Kit sequencing (v1.5) (Illumina, San Diego, CA, USA).

Proteomics

Proteins were revealed on sodium dodecyl sulfate–polyacrylamide gel electrophoresis (SDS-PAGE) or NuPAGE (Tris-Glycine 4–12% acrylamide gradient; Life Technologies) using silver staining (Silver staining kit, # 24612; Pierce, Rockford, IL, USA).

For mass spectrometry analysis, a first experiment permitted to sequence all the peptides in G3BP IP from WT mouse brain and in non-specific IgG IP. A second experiment reproduced the results after G3BP IP in WT and KO brain extracts, subjected or not to RNAse treatment. Proteins in silver-stained gel were cut out and in-gel digested with endoproteinase trypsin. Peptides were extracted from the gel and analyzed by LC-MSMS and an Orbitrap XL under standard conditions. Proteins were identified after database search against the NCBI non-redundant database, using Mascot as search engine.

Western blotting

Mouse brain extracts were prepared as in CLIP experiments [UV cross-linking of brain homogenates (except for conditions without crosslink), followed by lysis in the following buffer (50 mM Tris-HCl, pH 7.4; 100 mM NaCl; 1 mM $MgCl_2$; 0.1 mM $CaCl_2$; 1% NP-40; 0.5% sodium deoxycholate; 0.1% SDS; protease inhibitor and ANTI-RNAse), immunoprecipitation using protein G-coupled magnetic beads (Dynabeads; Invitrogen, Carlsbad, CA, USA) coupled to antibody, and washes in high salt buffer (50 mM Tris-HCl, pH 7.4; 1 M NaCl; 1 mM EDTA; 1% NP-40; 0.5% sodium deoxycholate; 0.1% SDS)]. For analysis of proteins in the cerebrum and the cerebellum, dissected parts of the brain were frozen in liquid nitrogen, grinded, then lysed and sonicated in lysis buffer. Samples were boiled in Laemmli buffer and resolved on Novex NuPAGE gels (Life Technologies). Proteins were transferred onto nitrocellulose membranes and probed with the following antibodies: mouse anti-G3BP1 (Abnova), 1 : 3000 or rabbit anti-G3BP1 (Novus) 1 : 10 000, rabbit anti-G3BP2 (Novus, Littleton, CO, USA) 1 : 10 000, rabbit anti-C-terminal binding protein (CtBP) (Santa Cruz Biotechnology, Santa Cruz, CA, USA) 1 : 1000, rabbit anti-splicing factor proline and glutamine rich (SFPQ) (Thermo Scientific, Waltham, MA, USA) 1 : 1000, anti-metabotropic glutamate receptor (mGluR)1, anti-mGluR5, and rabbit anti-Caprin-1 provided by Y. Wang, 1 : 1000. The blots were developed using an enhanced chemiluminescence technique (Pierce).

Mapping of high-throughput sequencing data to the mouse genome

For each sample, we filtered out reads shorter than 21 nucleotides to avoid ambiguous mapping locations, as any random sequence shorter than 21 nt is highly likely to be found in the genome. We did a two-step mapping, first using as reference the mouse mm9 genome (NCBI Build 37 assembly, July 2007) using Bowtie v0.12.7 (Langmead *et al.* 2009) allowing up to two mismatches. Furthermore, to avoid missing any tags in mature transcripts, split reads that could correspond to exon junctions were obtained using the genome multitool (GEM) split mapper (build 592) from the GEM library (Marco-Sola *et al.* 2012) using the following filters

for split reads: same strand, same chromosome, and a maximum of 100 000 bases apart, based on a intron length estimation in mm9. This was performed on the remaining reads not aligned with bowtie.

Pipeline for one of the samples

bowtie -c mm9_reference brain_WT1.fa -v 2 –all –best -f –tryhard > Brain_WT1_result

gem-split-mapper -I mm9_index –filter-matches same-chromosome, same-strand,maximum-distance=100000 -i Brain_WT1.fa -o Brain_WT1_splitmapped.gem

The results from bowtie and GEM were converted to BED format and joined with custom scripts.

Clusters annotation to Ensembl database

Uniquely read-tags mapped to the same strand were clustered according to their overlapping genomic positions with Pyicos (Althammer *et al.* 2011). Annotations of gene classes (biotypes) and location of genic structural regions were downloaded from Ensembl 62 database, NCBI37/mm9 assembly. Clusters were classified according to the overlap with the gene class (protein-coding, rRNA, miRNA, etc .). Clusters that do not fall in such regions were classified as intergenic. Protein coding-associated clusters were classified into exonic or intronic, and exonic ones into 5'-untranslated region (UTR), coding sequence (CDS), or 3'UTR. Because of the nature of gene structure annotation, a given cluster may fall in more than one category, e.g. miRNA and intronic. Clusters were visualized in the UCSC Genome Browser by uploading the tags in bed format as custom tracks.

Gene ontology annotations

We assigned the gene ontology (GO) terms using expression analysis systematic explorer (EASE) (Hosack *et al.* 2003), to compute the overrepresented functional categories in 'Biological Process,' 'Cellular Component,' and 'Molecular Function' systems according to database for annotation, visualization and integrated discovery database. EASE scores (modified Fisher's exact test probabilities) were computed for our input gene lists compared to mouse genome and also compared to a data set regrouping all cap analysis gene expression transcripts identified in mouse brain (Fantom 4) correlated with RefGene annotations, to avoid a bias toward brain expressed genes in our CLIP experiment. Both analyses gave the same results.

Analysis of clusters position

Clusters from the G3BP-complex were analyzed for the CLIP experiments performed with wild-type or G3BP1 KO mouse brain (representing G3BP1 and 2 HITS-CLIP and G3BP2 HITS-CLIP, respectively). Introns with clusters were classified and distributed according to their position in the transcripts (for the transcripts with a minimum of 10 introns) or their size (classes of different length bins), and the percentage of introns with clusters from HITS-CLIP were plotted according to these positions/sizes. For analysis along mRNA or intron length, mRNAs or introns were divided into 10 bins, from 5' to 3', to average the different transcripts/introns lengths. The frequencies of clusters or introns with clusters were plotted along these mRNAs or introns lengths.

PCR

G3BP IP was performed as for the CLIP experiments, but without RNase treatment of the brain lysates (except some conditions as indicated in the text). RNAs were directly recovered from proteinase K treatment of the beads resuspended in lysis buffer after IP washes, by Tri Reagent phenol extraction (Sigma-Aldrich, St. Louis, MO, USA). Same volumes of resuspended RNA were reverse-transcribed. In the case of total extracts, cerebellum from WT and KO brains were isolated, and lysates of cerebrum and cerebellum were obtained directly in Tri Reagent, using a FastPrep instrument (MP Biomedicals, Santa Ana, CA, USA). A quantity of 1.5 μg of RNA was reverse transcribed. Reverse transcriptions were performed using random hexamer primers and First Strand cDNA Synthesis Kit (GE Healthcare Life Sciences, Pittsburgh, PA, USA), for 1 h at 37°C. At least three independent experiments with individuals from different litters were carried out. In the case of real-time quantitative PCR, amplifications of purified cDNAs using Taq platinum polymerase (Life Technologies) and Sybr Green mix were followed in real-time using a Roche LightCycler 480 (Roche, Basel, Switzerland). Data were analyzed with the Roche LightCycler480 software. Absolute RNA concentrations were recovered from standard curves plotted for each pair of primer oligonucleotides used, and relative analysis were performed as described in the results. The oligonucleotides were designed with Geneious software (see Appendix S1), and concerning the real-time PCR experiments, they were chosen to produce amplicons of equivalent size and to use the same annealing temperature. Furthermore, the establishment of standard curves permitted to normalize the efficiency of each primer pair and thus rule out possible experimental PCR biases toward specific amplicons. Concerning IP, the result was expressed as an enrichment of the tested transcripts in specific G3BP IP relative to their quantitation in IgG non-specific IP (log10 scale histograms). Concerning the relative analyses between brain regions or mice genotypes, the reference samples (cerebrum or WT) were plotted as 1. Hypoxanthine phosphoribosyltransferase 1 (*Hprt1*) gene was used as a normalizer. Non-quantitative PCRs were carried out on Mastercycler Gradient 96 thermocycler (Eppendorf, Hamburg, Germany) for 28 cycles (of 30 s denaturation at 95°C, 30 s annealing at 60°C and 30 s amplification at 72°C using Taq Platinum polymerase), and amplification products were analyzed by 2% agarose gel electrophoresis and visualized by ethidium bromide staining.

Data from quantitative PCRs were analyzed with two-sided Student's *t*-tests (after assessment of equal variances with Fisher tests). When Gaussian distributions were not assessed, non-parametric Mann–Whitney tests were used to confirm the statistical difference of the data with a p-value < 0.05.

Data access

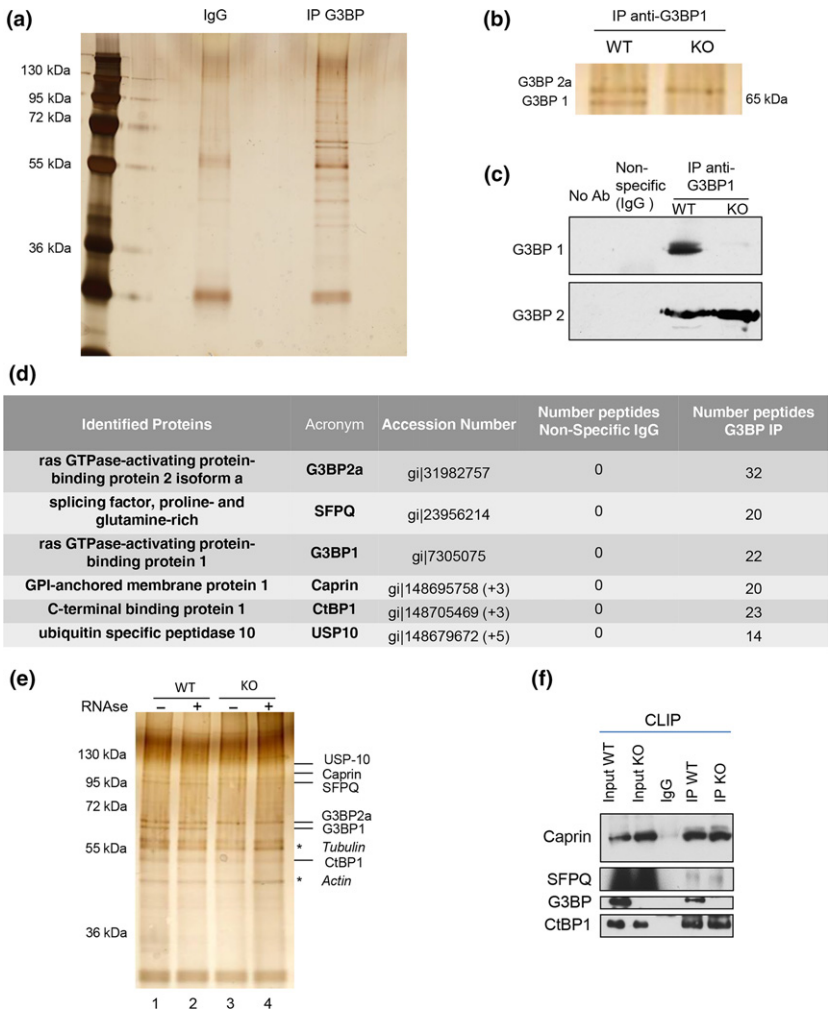
The sequencing data are available on a URL link G3BP HITS-CLIP data http://www.igmm.cnrs.fr/jtl/ (login "igmm", password "igmm").

Results

Proteomic analysis of G3BP immunopurified complex

Considering that G3BP1 KO mice present defects in synaptic plasticity and neuronal calcium homeostasis, we wanted to

identify the endogenous RNA targets of G3BP1 in mouse brain. We used HITS-CLIP which relies on covalent cross-link induced by UV irradiation between the target RNA and the RNA-binding protein *in vivo* followed by immunopurification (Licatalosi *et al.* 2008; Darnell 2010). Highly stringent conditions of immunoprecipitation used during the CLIP protocol (several washes at 1M NaCl), followed by size separation on denaturing SDS gels of the immune complex, were used to considerably reduce the amount of background RNAs as well as of co-precipitating proteins. Silver staining of a denaturing gel revealed that G3BP1 was efficiently immunoprecipitated, but a few other proteins were also immunoprecipitated (Fig. 1a). Performing the experiment with another antibody targeting another epitope in the protein gave a similar protein profile (Figure S1a), minimizing the possibility of cross-reacting proteins. More striking was a doublet of protein bands running at the size of G3BP. Mass spectrometry analysis and western blotting revealed that the lower band of the doublet corresponds to G3BP1 and the upper band corresponds to G3BP2a, one of the two



Identified Proteins	Acronym	Accession Number	Number peptides Non-Specific IgG	Number peptides G3BP IP
ras GTPase-activating protein-binding protein 2 isoform a	G3BP2a	gi|31982757	0	32
splicing factor, proline- and glutamine-rich	SFPQ	gi|23956214	0	20
ras GTPase-activating protein-binding protein 1	G3BP1	gi|7305075	0	22
GPI-anchored membrane protein 1	Caprin	gi|148695758 (+3)	0	20
C-terminal binding protein 1	CtBP1	gi|148705469 (+3)	0	23
ubiquitin specific peptidase 10	USP10	gi|148679672 (+5)	0	14

Fig. 1 Proteomic analysis of G3BP immunopurification reveals stably associated partners. (a) Silver staining of proteins recovered by G3BP1 IP (IP G3BP) or control IgG (IgG) from extracts of wild-type (WT) mice brain analyzed by SDS-PAGE. (b) Comparison of silver staining of proteins recovered by G3BP1 IP (IP G3BP) from brain extracts of wild-type (WT) and G3BP1-knockout mice (KO) analyzed by SDS-PAGE. (c) Immunoblotting of proteins recovered by G3BP1 IP (IP G3BP) from brain extracts of wild-type (WT) and G3BP1 knockout mice (KO) with anti-G3BP1 (panel G3BP1) and anti-G3BP2 (panel G3BP2). (d) Peptides identified by mass spectrometry after immunopurification of G3BP-complex using anti-G3BP1 antibodies. (e) Silver staining of proteins recovered by G3BP1 IP from WT (lanes 1 and 2) and KO (lanes 3 and 4) brain extracts treated (lanes 2 and 4) or not (lanes 1 and 3) with RNase (RNases A/T1 mix, Ambion). The proteins identified by mass spectrometry analysis are indicated in the right of the panel. Actin and Tubulin (*) are contaminating proteins as they were recovered in control IgG IP. (f) Western Blot confirming IP of splicing factor proline and glutamine rich (SFPQ) (panel SFPQ), Caprin-1 (panel Caprin-1) and CtBP1 (panel CtBP1) under cross-linking immunoprecipitation condition using G3BP1 antibodies to IP G3BP from brains of WT and KO mice.

Fig. 2 High-throughput sequencing and cross-linking immunoprecipitation (HITS-CLIP) of G3BP-complex in mouse brain reveals RNAs involved in synaptic transmission. (a) Autoradiogram of the RNA-G3BP complex from the CLIP experiment. Extracts were prepared from UV and non-UV cross-linked WT brain and RNA was partially digested using high (+++) or low (+) RNase concentrations (RNases A/T1 mix, Ambion). Complexes were immunopurified using anti-G3BP1 antibody and RNAs were 5'end-labeled using Polynucleotide Kinase. After size separation by running on a denaturing NuPAGE, complexes were transferred onto a nitrocellulose membrane. No radioactive signal was detected when brain was not cross-linked, leading to disruption of protein-RNAs association (lane a). No signal was detected either when using a control IgG (lane d). A complex running around 20–30 kDa above G3BP molecular weight is associated to low RNAse treatment (lane c band no. 2), and it shifts toward G3BP size with high RNAse treatment (because of the lower size of the degraded RNAs) (lane b band no. 3). Interestingly, this treatment also revealed two other complexes (bands no. 1 and no. 4), one of them running below the molecular size of G3BP protein (band no. 4). The brackets permit to better visualize the size of the complexes smears. Western blot of proteins immunoprecipitated from the beads (bottom) confirms the presence of G3BP only when using an antibody against G3BP in the IP. RNAs present in the four different complexes indicated by numbers 1–4 were extracted by cutting out the nitrocellulose, digestion with proteinase K and phenol chloroform extraction, then run on a polyacrylamide gel subjected to autoradiography. (b) RNAs associated to the low RNase treatment complex range from 50 to 100 nucleotides in length (lane 2), while the three complexes obtained from high RNase treatment all contain RNAs 20–50 nucleotides long (lane 1, 3, 4). RNAs obtained from the CLIP experiment following low RNAse treatment (band no. 2 in a and lane 2 in b) and obtained from two independent CLIP experiments were subjected to Illumina high sequencing after RT-PCR, mapped and clustered. (c) Gene ontology terms were attributed to the transcripts reproducibly identified in the replicate HITS-CLIP experiments to compute the overrepresented functions in 'Biological Process', 'Molecular Function', and 'Cell Component' compared to all the transcripts in mouse brain. Expression analysis systematic explorer (EASE) score (p-value of enrichment) was plotted for the four most enriched classes in each of the three categories. (d, e) Validation by immunoprecipitation of RNAs associated to G3BP-complex followed by real-time PCR (d) or traditional PCR (e) of some targets identified in HITS-CLIP. The results are plotted as the fold enrichment of RNAs in specific G3BP IP compare to a non-specific IgG IP (d) hypoxanthine phosphoribosyltransferase 1 (*Hprt1*) and *Gusb* are used as negative controls, growth arrest-specific 5 (*Gas5*) as a positive control. RNAs from input total sample and IP with G3BP antibody or control IgG was analyzed by RT-PCR; the last lane corresponds to a negative control of IP sample RNAs not subjected to reverse transcription in order to check that we do not amplify DNA sequences. The sequences targeted by PCR were mostly 3'-untranslated region, or exon junction sequences of mature transcripts.

products of *G3bp2* gene obtained by alternative splicing (Fig. 1b and c). Furthermore, this protein was also immunoprecipitated by the anti-G3BP1 antibody, from extracts derived from brains of G3BP1 KO mice, which do not express G3BP1 (Fig. 1b and c). Since western blots performed with the same antibody failed to detect G3BP2a but only G3BP1, we assume that the epitope recognized by the antibody resides in a region of G3BP2a that is sensitive to denaturation (Fig. 1c). G3BP2a was also immunoprecipitated using an antibody directed against another epitope of G3BP1 (Figure S1a), preventing us from analyzing the G3BP1 complex alone. Given that both G3BP1 and G3BP2a co-localize in neurons and in SGs formed under arsenite treatment in neurons (Figure S1b), and that there is evidence of interaction between both proteins (Matsuki *et al.* 2013), we postulate that G3BP1, G3BP2, and the associated proteins are part of the same complex. However, it is too speculative at this stage to rule out the possibility that G3BP1 and G3BP2 exist in separate complexes with the same binding partners, and we will thus characterize the 'G3BP-complex', as a mixture of G3BP1 and G3BP2 RNP complexes.

Mass spectrometry analysis of the complex (Fig. 1d) revealed well-known G3BP1 partners: named as GPI-anchored membrane protein 1 (Caprin-1) (Solomon *et al.* 2007a,b) and USP10 (Soncini *et al.* 2001). We also identified CtBP1 and SFPQ (splicing factor proline and glutamine rich, or PSF), proteins with functions both in the nucleus and the cytoplasm (Chanas-Sacré *et al.* 1999; Nardini *et al.* 2003; Kanai *et al.* 2004; Peng *et al.* 2006; Hübler *et al.* 2012). These proteins failed to immunoprecipitate with non-specific IgGs, used here as negative control, unlike tubulin or actin, supporting a stable association with G3BP 1 and 2. These results were reproduced in independent experiments. The presence of these proteins was further validated by immunoblotting, as they were detected in G3BP IP but not in control IgG IP (Fig. 1f). Association of these proteins with G3BP-complex was not sensitive to RNase treatment, implying that it involves protein-protein interactions (Fig. 1e, compare lane 1 and 2). The same analysis performed with extracts from brains of G3BP1 KO mice, allowed identification of the same set of proteins (Fig. 1e, lanes 3 and 4), indicating that both G3BP1 and G3BP2a are able to associate with the same partners. The immunopurified complex from extracts of WT mouse brain is therefore called G3BP-complex.

HITS-CLIP analysis of the G3BP-complex

HITS-CLIP analysis was performed for the G3BP-complex (G3BP1 and G3BP2a in the WT mice) as well as for G3BP2a alone, which was immunopurified from brains of G3BP1 KO mice. Labeling of RNA within the G3BP-complex after low RNase treatment according to HITS-CLIP protocol (see Materials and Methods) and analysis of ^{32P}RNA:G3BP complex on SDS-PAGE revealed a labeled band above the size of G3BP (Fig. 2a, lane c, band no. 2). This protein-RNA complex was resolved into three complexes of different sizes

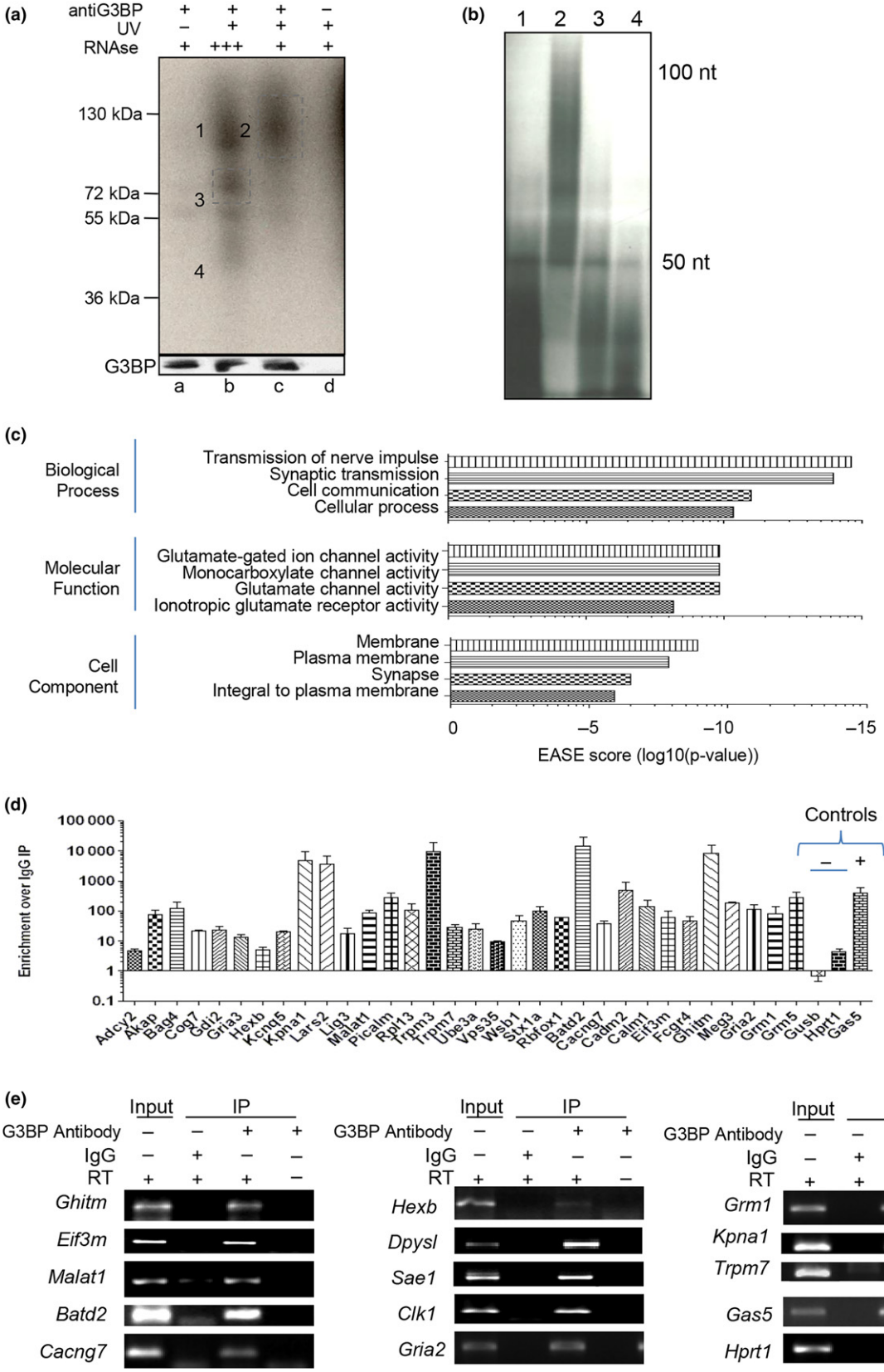

(a)
antiG3BP + + + –
UV – + + +
RNAse + +++ + +
130 kDa
72 kDa
55 kDa
36 kDa
G3BP
1
2
3
4
a b c d
(b)
1 2 3 4
100 nt
50 nt
(c)
Biological Process
Transmission of nerve impulse
Synaptic transmission
Cell communication
Cellular process
Molecular Function
Glutamate-gated ion channel activity
Monocarboxylate channel activity
Glutamate channel activity
Ionotropic glutamate receptor activity
Cell Component
Membrane
Plasma membrane
Synapse
Integral to plasma membrane
0 –5 –10 –15
EASE score (log10(p-value))
(d)
Enrichment over IgG IP
100 000
10 000
1000
100
10
1
0.1
Controls
– +
Adcy2 Akap Bag4 Cog7 Gdi2 Gria3 Hexb Kcnq5 Kpna1 Lars2 Lig3 Malat1 Ptcalm Rpl13 Trpm3 Trpm7 Ube3a Vps35 Wsb1 Stx1a Rbfox1 Batd2 Cacng7 Cadm2 Calm1 Eif3m Fcgrt Ghitm Meg3 Gria2 Grm1 Grm5 Gusb Hprt1 Gas5
(e)
Input IP
G3BP Antibody – – + +
IgG – + – –
RT + + + –
Ghitm
Eif3m
Malat1
Batd2
Cacng7
Hexb
Dpysl
Sae1
Clk1
Gria2
Grm1
Kpna1
Trpm7
Gas5
Hprt1

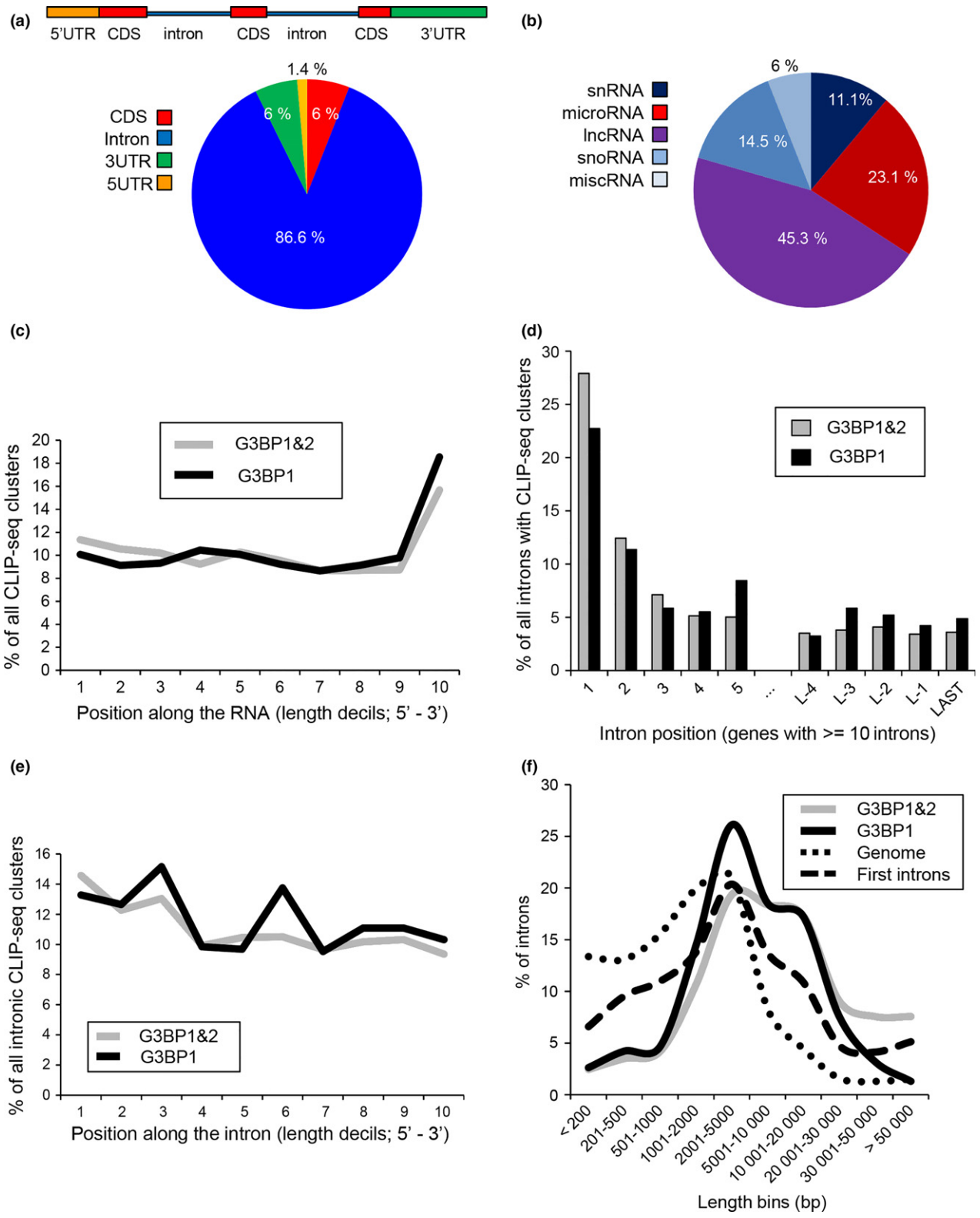
Fig. 3 G3BP clusters partition within protein-coding and non-coding regions. (a) G3BP-complex is associated to a majority of intronic sequences of protein-coding genes. Clusters then fall into coding sequence and 3'-untranslated region (UTR) with few in 5'UTR. (b) Distribution of clusters within known non-coding RNAs reveals a majority of long non-coding RNAs as well as miRNAs. (c) The representation of the percentage of clusters along the mature mRNAs reveals an enrichment toward the 3' end of the mRNAs. Clusters from CLIP performed in WT (G3BP1&2) and clusters from WT CLIP after filtering out the genes identified in knock-out (KO) CLIP to remove G3BP2-associated RNAs (thus called G3BP1) are represented. (d) Distribution of introns with clusters of G3BP1&2 or G3BP1 (transcripts with at least 10 introns). Approximately, 25% of intronic G3BP 1 or 2 CLIP clusters are within the first intron of transcripts. (e) Percentage of intronic clusters from G3BP CLIP along the introns, from 5' to 3' of the intron. The average positions in introns of different lengths were obtained by dividing the introns into 10 decil bins. Along the introns, the clusters are not biased toward a specific end. Distributions of clusters of G3BP1 and G3BP2 and clusters of G3BP1 only are represented. (f) Distribution of the percentages of introns based on intron size within clusters from G3BP1 and 2 CLIP, G3BP1 CLIP, total genomic introns, or first introns of genes. Introns with G3BPs CLIP clusters and first introns of transcripts present the same distribution of size, with a peak between 2000–5000 bp length. Bins of different intron lengths were determined and plotted in the x axis. MiscRNA, miscellaneous RNA.

when the extract is submitted to high RNase treatment (Fig. 2a, lane b, bands no. 1, 3, 4). In addition to a band at the size of G3BP (Fig. 2a, lane b, band no. 3), two other protein bands were also detected; one running slower than G3BP (band no. 1) and one running faster (band no. 4). The slower migrating band no. 1 is consistent with the size of the two RNA-binding proteins, Caprin-1 and/or SFPQ (105 kDa and 95 kDa, respectively), identified by mass spectrometry as strong partners of G3BP. The fast migrating band might correspond to another RNA-binding protein with a molecular weight around 40 kDa which is consistent with the size of CtBP1 even if no RNA-binding activity has yet been identified for this protein. Using even more stringent conditions for the CLIP (2M NaCl, 2M NaCl and 2% NP40, or 1M urea), we failed to resolve a unique band at the size of G3BP without the other complexes (Figure S1c). After high RNAse treatment, all the conditions tested led to the same profile. The three protein bands that were resolved following high RNase treatment were associated with RNAs of the same size, ranging between 20 and 50 nucleotides (Fig. 2b, lanes 1, 3 and 4); in contrast, the size of the RNAs associated with the band in the low RNase-treated sample ranged between 50 and 100 nucleotides (Fig. 2b, lane 2). This analysis confirmed that different RNA-binding proteins of different sizes are, indeed, associated with small sized RNAs after high RNase treatment. We decided to sequence the RNAs which were less degraded, after low RNase treatment, to be able to uniquely map the reads of a correct size (Wang *et al.* 2009) (band no. 2 in Fig. 2a and lane no. 2 in Fig. 2b), corresponding to G3BP-complex-associated RNAs.

For an exhaustive and precise analysis of G3BP-complex-associated RNAs, we performed two independent replicate HITS-CLIP experiments using WT mouse brain (RNAs sequenced from band no. 2, Fig. 2a, lane c). Reads were filtered then mapped to the genome, and uniquely mapped reads (59 800 in one experiment and 53 183 in the second experiment) were clustered according to genomic position. Furthermore, we used the GEM split mapper (build 592) from the GEM library (Marco-Sola *et al.* 2012) to detect any exon–exon junction on the remaining reads, in order not to discard tags that would fall into junctions in mature transcripts (see Materials and Methods). Data from the two alignment methods were combined and confronted to Ensembl 62 database to identify their precise location. 39 671 different clusters in one experiment and 19 631 in the second were identified, mapping respectively to 5800 and 4890 different protein-coding or non-coding genes (1 or more clusters per gene). 3034 transcripts were in common between the two data sets and were selected for further analysis.

We performed a motif enrichment analysis using a kmer-enrichment approach to look for a possible consensus binding site for the complex in WT brain (Appendix S1). First, we observed that around 8% of the clusters have at least one long, overrepresented motif (p-value < 0.001), which is part of a transposable element of the short interspersed nuclear elements-Alu-B1 family (Figure S2a), and which is found at the center of the clusters (Figure S2b). We performed the kmer analysis after filtering out the repeated elements, and derived several possible binding sequences, consistent with the fact that G3BP-complex contained different RNA-binding proteins (RBPs) (Table S1 and Figure S2c, 1–2). We also scanned the clusters for the presence of previously identified consensus sequences for G3BP1 (systematic evolution of ligands by exponential enrichment motif; Tourrière *et al.* 2001) or G3BP2 (RNAcompete; Ray *et al.* 2013), and as for the previous sequences, these motifs were present but not highly enriched in the clusters from G3BP-complex HITS-CLIP over the background (Figure S2c, 3–4).

GO terms were attributed to the G3BP-complex targets using EASE (Hosack *et al.* 2003). Annotation enrichment was ascertained by attributing an EASE score for each GO classification, relative to the mouse genome and also more specifically to all the transcripts identified in mouse brain in Fantom 4 cap analysis gene expression database, to avoid potential enrichment bias toward brain functions because of the fact that HITS-CLIP was performed in brain. Both analyses gave the same results and indicated that G3BP-complex target genes are involved in synaptic transmission,



in particular glutamate-mediated transmission (Fig. 2c). We experimentally validated some of the identified targets using both real-time PCR and traditional PCR (Fig. 2d and e). Real-time PCR data were expressed as enrichment of the transcripts in specific IP over a non-specific IP performed with IgG. Thresholds of immunoprecipitation efficiency were

determined by including positive and negative controls. As a positive control, we used the non-coding RNA growth arrest-specific 5 (*Gas5*), a well-known G3BP target (Zekri *et al.* 2005) that was identified among the G3BP HITS-CLIP tags, and as a negative control *Hprt1*, a housekeeping gene not present in G3BP HITS-CLIP. Thus, several protein-coding transcripts involved in synaptic transmission like *Kpna1*, *Trpm3* and some involved in plasticity through glutamate signaling like the metabotropic receptor mGluR1 or the α-amino-3-hydroxy-5-methyl-4-isoxazolepropionic acid receptor subunit Gria2 transcripts were highly enriched in G3BP-complex. We also confirmed the presence of other transcripts involved in more general cell metabolism like eukaryotic translation initiation factor 3, subunit M (*Eif3m*), SUMO-activating enzyme subunit 1 (*Sae1*) or long non-coding RNAs (lncRNA) such as the brain-enriched *Malat1* transcript or *Meg3* originating from an imprinted non-coding RNA gene.

The G3BP-complex preferentially binds non-coding sequences (intronic and 3'UTR) of target protein-coding genes

We determined the position of clusters within the protein-coding genes, including 5'UTR, 3'UTR, CDS, and introns (Fig. 3a). Surprisingly, more than 80% of the clusters fall within introns. We did not expect so many intronic sequences as G3BP proteins are mainly detected in the cytoplasm of neurons both in brain slices or in cultured neurons ((Martin *et al.* 2013) and Figure S1b). We observed that 6% of the clusters mapped to CDS, and 6% to 3'UTR sequences, consistent with a possible role for the G3BP-complex in the regulation of the stability/translation of some transcripts. Few clusters (1.4%) map to 5'UTR sequences. Since in vertebrates, introns are larger than other gene regions, we calculated the percentages of the different genes identified instead of clusters to avoid a bias as a result of more clusters in introns of one particular transcript than in shorter exons. Again, we still found a large majority of genes for which the clusters mapped to introns (72.7%), 12.6% in CDS, 11.6% in 3'UTR, and fewer (3.1%) in 5'UTR. In the fraction of non-coding transcripts (3% of total mapped clusters), a majority (45.3%) were lncRNAs, 23.1% were microRNAs (miRs) and 14.5% were snoRNAs (Fig. 3b). miRNAs and snoRNAs can also be processed from introns (like C/D box snoRNAs) (Hirose and Steitz 2001) and a large fraction could be part of host protein-coding and non-coding genes. The G3BP-complex is thus largely associated with non-coding regions of protein-coding genes, which might indicate a role of this complex in the control of the stability of these RNAs, by preventing their engagement in translation (see Discussion).

Consistent with the structural similarities between G3BP1 and G3BP2 and the identity of associated partners in neurons, analysis of G3BP2a HITS-CLIP from brain of G3BP1 KO mice gave rise to a similar distribution of tags along the transcripts as was observed for G3BP in WT brain (Figure S3). A total of 75.15% of protein-coding tags mapped to intronic regions and the distribution into the transcripts classes or along the transcripts were comparable to the WT, with yet a higher percentage of snoRNAs among the non-coding transcripts (Figure S3b). However, among the 3034 transcripts associated to the G3BP-complex in the WT, only 21% were identified in association with G3BP2 in the KO. This supports either the hypothesis of the existence of separate G3BP1 and G3BP2 complexes, or suggests a difference in the binding and composition of a single G3BP-complex in the absence of G3BP1. The non-redundancy of the proteins established from the G3BP1 KO phenotype might thus be linked to the identity of the transcripts they bind.

To investigate where the clusters of tags of G3BP-complex fall within different regions of the target genes, we first plotted the frequency of all clusters along the length of mRNA (mature RNA, CDS and UTR) by dividing each mRNA into 10 decil bins from 5' to 3', which allows the comparison of transcripts of different lengths. For both samples (G3BP-complex and G3BP2a-complex HITS-CLIP tags), we found an enrichment of tag clusters at the 3'-most decil, presumably because of the relatively high frequency of G3BP-complex binding at the 3'UTR (Fig. 3c).

Second, we looked only at the clusters located within introns (considering genes with at least 10 introns), and we found a different distribution along the gene with bias from 5' to 3' (Figure S4a), and a significant preference for the first intron (Fig. 3d), whereas there was no clear bias of position along each intron (Fig. 3e). Furthermore, the size distribution of the introns containing G3BP-complex tags followed closely the size distribution of the first intron in the whole genome (Fig. 3f), suggesting that G3BP-complex binding does not show a specific preference for neither longer nor shorter introns. Thus, the G3BP-complex seems to have a preference to bind non-coding sequences in the intron or in the 3'UTR. A motif analysis performed on the restricted set of intronic clusters revealed an enrichment of C- and CTG-rich sequences (Figure S5).

In order to assess whether G3BP target genes were highly expressed in the brain and could therefore easily contaminate the G3BP-complex, we analyzed their distribution among all expressed genes in neural and non-neural tissues, using available RNAseq data from neural samples (including whole brains, cerebellum, retina, and isolated neurons) and non-neural samples (including liver, kidney, heart, muscle, testis, T-cells, and different cell lines) (See Methods). Specifically, we classified genes depending on the transcripts expression levels, and we obtained the same distribution of G3BP-complex-associated transcripts in the different classes, with no particular enrichment in the highly expressed genes (both in neural and non-neural samples) (Figure S4b). This suggests that G3BP-complex targets are not neural- (nor non-neural-), specifically highly expressed genes.

G3BP-complex targets are intron-retaining transcripts

The large content of sequences originating from introns in the G3BP HITS-CLIP tags could represent either transcripts of protein-coding genes with retained intronic sequences, non-coding transcripts originating from the sense strand (as the transcription strand was taken into account when mapping the reads to the genome) or premature messenger RNAs (pre-mRNAs). In order to distinguish between these different possibilities, we studied more closely a really restrictive subset of transcripts specifically associated with G3BP-complex and important for neuronal function, which were representative of the clusters distribution over the whole genome as shown in Fig. 3: most of the clusters mapped to intronic regions with a high density of reads toward 5' biased introns (Figure S6). Real-time PCR was used to amplify different regions across the selected genes: (i) the exon–intron junction that preceded the intron with the highest density of clusters in replicate experiments, to avoid any artifact because of amplification that may result from degradation products of the intron, (ii) exon–intron junction of another large intron of the transcript, and (iii) a portion of the 3'UTR (Figure S6 and Appendix S1). Efficiencies of each primers pair were tested and taken into account to quantify each amplified product, to avoid enrichment biases toward one of the amplicons (that would be attributable to PCR efficiency, See Methods). For all five transcripts, amplifications were performed from immunopurified G3BP-complex and efficiencies of enrichment were established relative to the background amount associated with the antibodies in the control immunoprecipitations with IgGs. *Gas5* exon1 amplification was used as positive control, whereas amplification from intronic and exonic sequences of *Hprt1* (housekeeping gene not identified in HITS-CLIP) was used as negative control. As *Hprt1* sequences were slightly enriched with G3BP antibody relative to IgG antibody, we further arbitrarily set this enrichment value as a threshold to consider if a sequence was readily associated with G3BP-complex over background (i.e. ratio G3BP IP/IgG IP ≥ 10) (Fig. 4a). Globally, the different regions of the transcripts tested were efficiently amplified, suggesting that pre-mRNA and/or incompletely spliced transcripts rather than mature mRNA or small RNAs were associated with the G3BP-complex. Concerning the cell adhesion molecule 2 (*Cadm2*) and *Btbd9* transcripts, however, one intronic region was amplified only slightly above the considered background levels, indicating some specificity of association/regulation by G3BP-complex depending on the transcript. We also considered the expression of these different transcripts regions in brain (Fig. 4b). For two transcripts, *Fgf14* and *Btbd9,* amplifications of the retained intronic sequences were comparable to amplification of the 3'UTR, suggesting that the premature transcripts are up-regulated over mature transcripts; whereas for RNA-binding protein, Fox-1 homolog (*Rbfox1*), *Cadm2* and *Ctnnd2* transcripts, the amplification of 3'UTR was higher than intron sequences, indicating expression of both mature and premature transcripts at comparable levels. Traditional (not quantitative) PCR further validated the presence of these intronic sequences in the brain and in the G3BP-complex (Fig. 4c). Altogether, the G3BP-complex in the brain appears to be associated with premature transcripts containing at least two introns rather than excised introns or transcribed small transcripts from the same locus.

G3BP1 influences the expression level of retained intron sequences of target transcripts in the cerebellum

The fact that G3BP1 deficiency in mice was associated with an ataxia phenotype, suggested that G3BP1 plays an important role in the cerebellum. To determine whether this phenotype could be mediated by alteration in the expression of G3BP-complex target transcripts with retained introns, we analyzed the expression of these transcripts in isolated cerebellum and cerebrum. First, real-time PCR was used as before to quantify different regions of each transcript from the cerebellum and cerebrum of WT mice. Each quantitation involved RNA extracted from at least three separate brain samples and triplicates from cerebrum normalized to *Hprt1* were set as 1. Figure 5(a) shows that the levels of transcripts with retained intronic sequences were significantly higher in the cerebellum compared to cerebrum, albeit with differences between the four G3BP target genes tested. *Rbfox1* demonstrated a clear example where the two retained introns tested were amplified to the same level and were 2.5 times higher in the cerebellum compared to the cerebrum, whereas no difference was observed for the 3'UTR (which accounts for the total of mature and premature transcript), suggesting that most *Rbfox1* transcripts in the cerebellum have retained introns. For all the three other genes tested, the most 5' intron was always more expressed than the other regions of the transcripts, which may suggest that the preferential binding of G3BP-complex to this intron allows its stabilization. Consistent with this suggestion, quantitation of the same transcripts from KO mice demonstrated that the retained introns were expressed to comparable levels in the cerebrum of WT and KO mice, although they were less expressed in the cerebellum of KO compared to WT mice (Fig. 5b). For the *Rbfox1* transcript the levels of all amplified regions were lower in KO cerebellum compared to WT, whereas for all the other transcripts, the first intron was the most affected in the KO cerebellum. Indeed, in the case of *Fgf14*, *Cadm2* and *Ctnnd2*, although there was no difference between WT and KO in the cerebrum, there was a decrease in the level of the most 5' intron in the cerebellum of KO mice. Thus, the preferential binding of G3BP to the first intron could be a way to stabilize premature transcripts in the cerebellum.

Accordingly, HITS-CLIP performed with isolated WT cerebellum also revealed 80% of clusters mapping to intronic

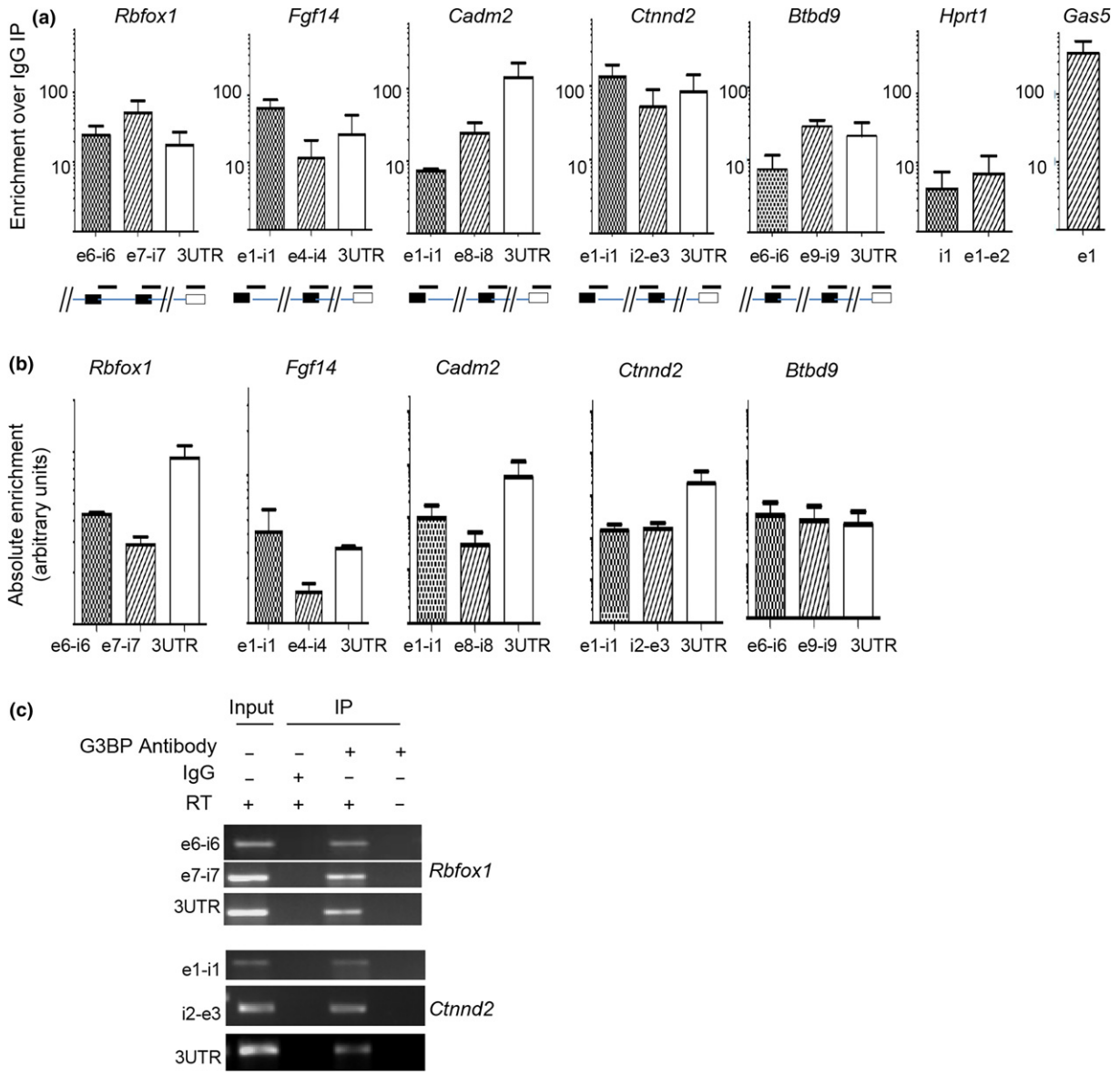


Fig. 4 Immunoprecipitation of intron-retaining transcripts in G3BP-complex. A few transcripts were selected to study specific premature sequences in each of them (See Figure S4). (a) Three regions were examined by real-time PCR in immunopurified transcripts from G3BP IP (exon–intron junction adjacent to the intron with the highest density of clusters in replicate experiments, exon–intron junction downstream, or 3'-untranslated region sequence). The fold enrichment in G3BP IP over non-specific IgG IP is given for each transcript section tested. Hypoxanthine phosphoribosyltransferase 1 (*Hprt1*) and growth arrest-specific 5 (*Gas5*) are used as negative and positive threshold controls of enrichment, respectively. X axis: regions targeted by PCR, using specific primers, with: e: exon; i: intron, e-i: exon-intron junction; numbers: exon or intron number of the transcript as established in NCBI database (See Figure S4). (b) Input enrichments (total brain extract) of the same transcripts regions. (c) Examples of RT-PCR products (non-quantitative) from input total extract, non-specific IgG IP, G3BP IP and negative control of PCR without reverse transcription (G3BP IP RNA) run on an agarose gel.

sequences in the cerebellum (Figure S7a). To investigate why the G3BP targets retaining introns are specifically regulated in the cerebellum, we analyzed the expression of the proteins of the complex in the cerebrum and the cerebellum of WT and G3BP1 KO mice by western blot (Figure S7b). Interestingly, the G3BP partner Caprin-1 is less expressed in the cerebellum compared to cerebrum, whereas USP10 has opposite expression profile and SFPQ has the same expression level in both cerebrum and cerebellum. Furthermore, in the absence of G3BP1, the level of USP10 is increased in the cerebrum, while decreased in the cerebellum. Recently, it has been shown that the binding of Caprin-1/USP10 to G3BP is mutually exclusive: Caprin-1 binding promotes, but USP10 binding inhibits, SG aggregation (Kedersha *et al.* 2016). Of

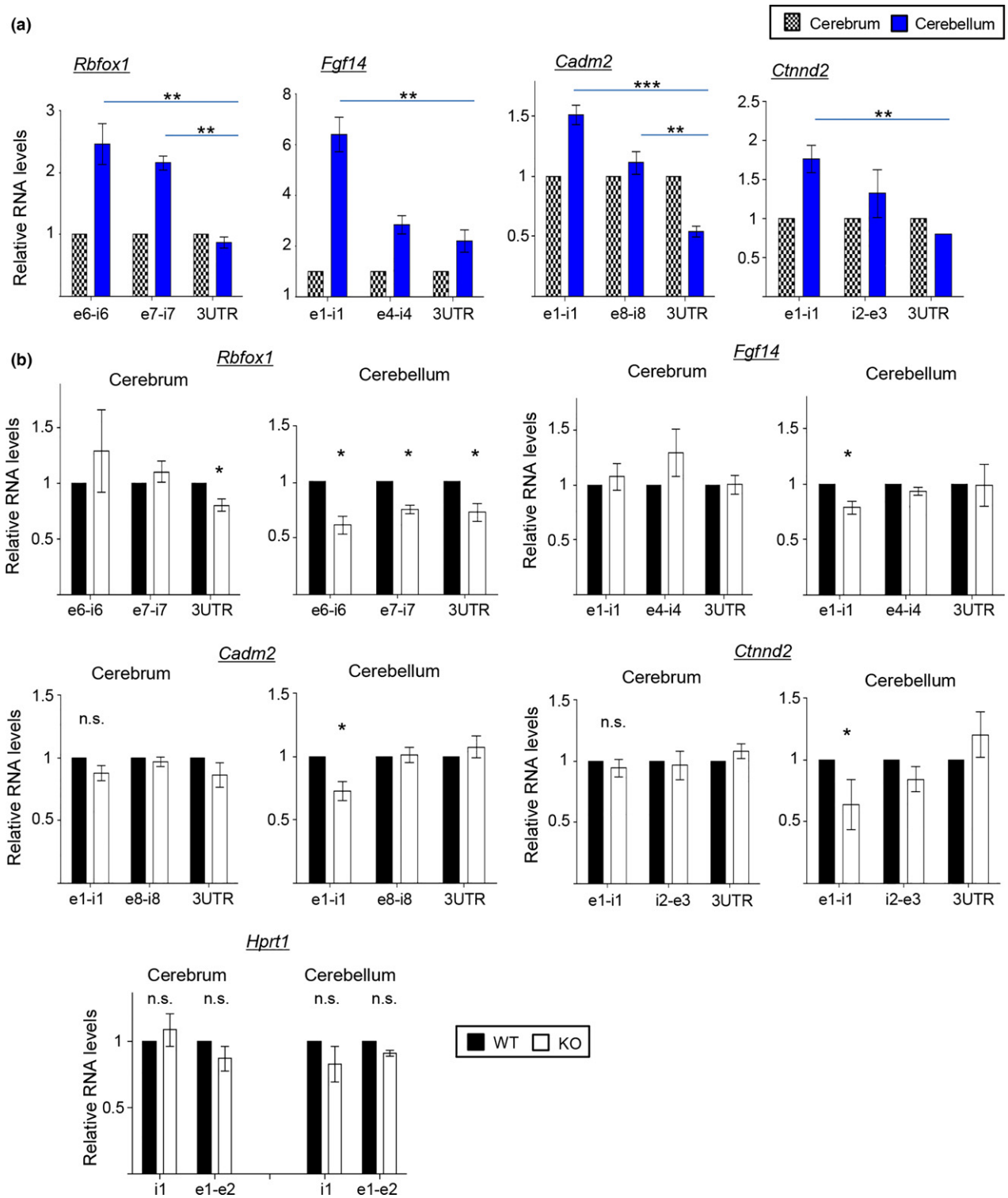


Fig. 5 Stabilization of intronic sequences in the cerebellum of WT mice, influenced by G3BP1. (a) Relative expression levels of the transcript sequences tested (same regions as in Figure 4) in cerebellum (blue bars) relative to cerebrum (hatched bars, plotted as 1), in brains of wild-type mice. (b) Relative expression levels of the tested sequences in knock-out (KO) (white bars) relative to WT (black bars, plotted as 1) in the cerebrum or the cerebellum, for each gene. Hypoxanthine phosphoribosyltransferase 1 (*Hprt1*) was used as a control of charge, and neither exonic nor intronic sequences showed any difference in expression levels in all the conditions tested. Results are represented as means ± SEM of relative RNA levels of at least three independent experiments. The results were analyzed with Student's *t*-tests: *Significant difference with p-value < 0.05; **p-value < 0.005; ***p-value < 0.0005; n.s., not significant.

Fig. 6 G3BP1 is involved in the regulation of metabotropic glutamate receptors 1/5 (mGluR1/5) transcripts. Three genes identified in cross-linking immunoprecipitation were selected based on their function in glutamate-linked synaptic transmission and the phenotype of G3BP1 knock-out (KO) mice, which show enhanced mGluR5-long-term depression in the hippocampus: the ionotropic α-amino-3-hydroxy-5-methyl-4-isoxazolepropionic acid glutamate receptor subunit GluR2 (*Gria2* transcript), and the metabotropic mGluRs from mGluR1/5 family: *Grm5 and Grm1* transcripts. Different sequences were tested: two intronic, 3'-untranslated region (UTR) sequences, and a sequence spanning exon–exon junction to amplify the mature transcript with excised intron. (a) Enrichment in G3BP IP over non-specific IgG IP assessed by RT-qPCR. (b) Absolute enrichment in input brain samples. It is unclear why the exon–exon junction is better enriched than the 3'UTR, this could be because of alternative polyadenylation sites which are not yet described. (c) Relative RNA levels of transcripts containing different regions of *Grm5* (i3, i4, 3'UTR and e3-e4) and *Grm1* (i2-e3, i3-e4, 3'UTR and e4-e5) genes in the cerebellum (blue bars) compared to the cerebrum (hatched bars) [(a), left panels], in WT. Relative RNA levels of the same regions of *Grm5* and *Grm1* in KO (white bars) compared to WT (black bars), in the cerebrum and the cerebellum [(b), right upper and lower panels]. Results represent means ± SEM of relative RNA levels of at least three independent experiments. The results were analyzed with Student's *t*-tests: *Significant difference with p-value < 0.05; **p-value < 0.005. (d) Expression of mGluR1 and mGluR5 in the cerebrum and the cerebellum from WT and G3BP1 KO mouse. Total homogenates were prepared from cerebrum and cerebellum of each genotype, and equal amounts of each protein lysates (confirmed by loading control actin) were analyzed by SDS–polyacrylamide gel electrophoresis, followed by immunoblotting. The western blot is representative of three independent experiments with different mice of each genotype, n = 3. Calbindin is used as a marker of the cerebellum.

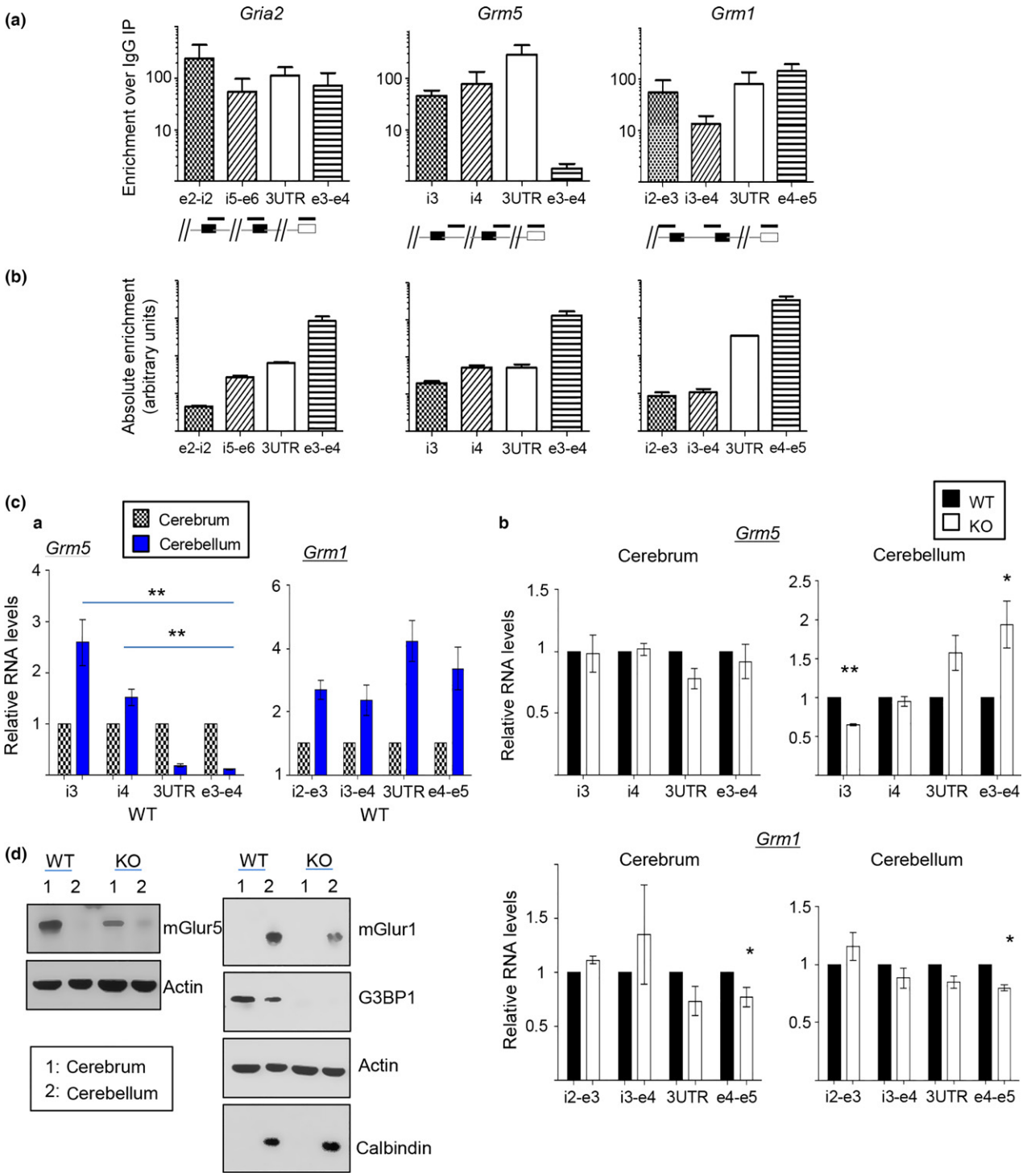
note also, the protein CtBP1 seems to be less abundant in the cerebellum after G3BP depletion (Figure S7b). Altogether these results suggest that a different composition of the G3BP-complex may lead to its specific role in the cerebellum and it could account for the ataxia phenotype associated with G3BP1 depletion.

G3BP1 is involved in the regulation of the expression of premature transcripts involved in glutamate signaling

G3BP1-deficiency in neurons leads to an increase in intracellular calcium release in response to (S)-3,5-dihydroxyphenylglycine, a selective agonist of group I metabotropic glutamate receptors, and leads to an enhancement of mGluR5 long-term depression (LTD) (Martin *et al.* 2013). Since at least three target genes identified by HITS-CLIP, namely *Gria2* (ionotropic glutamate receptor subunit 2 GluR2, internalized during LTD), *Grm1*, and *Grm5* (metabotropic glutamate receptors from the mGluR1/5 family with important functions in the cerebellum and the hippocampus, respectively), are involved in glutamate-linked synaptic transmission, the effect of G3BP1 depletion was further evaluated on the expression of these three transcripts both in the cerebellum and the cerebrum. The mGluR1/5 family members are interesting candidates because mGluR1 is highly expressed in the cerebellum in Purkinje cells, whereas mGluR5 expression is strongly decreased in the cerebellum in adult (Casabona *et al.* 1997). First, we validated the association of the three transcripts with the G3BP-complex in the brain (Fig. 6a, structure of the genes from UCSC with the clusters Figure S6b). As expected from the previous analysis, G3BP was associated with intron-retaining species of the three transcripts, even if inputs of brain total lysates reveal the preferential enrichment of mature transcripts (Fig. 6b). For *Grm5* transcripts, the G3BP-complex appears to be exclusively associated with premature transcripts, as RNAs that contain spliced Exon3-Exon4 failed to co-purify with the G3BP-complex (Fig. 6a, panel *Grm5*), while this region was readily amplified from the input of total lysates.

As expected from their regulated expression pattern, we observed very low levels of the 3'UTR or mature sequences of *Grm5* transcript in WT cerebellum relative to WT cerebrum, whereas the *Grm1* mature transcript was enriched in the cerebellum (Fig. 6c, compare *Grm5* and *Grm1* panels). In contrast, the intronic sequences of *Grm5* were enriched over the other regions in the cerebellum compared to the rest of the brain, suggesting a stabilization of the premature transcript. Again, the most 5' intron was overrepresented compared to other regions of the transcripts. Interestingly, in the absence of G3BP1, the levels of premature *Grm5* transcript in the cerebellum were reduced, with a concomitant increase of the mature transcript (Fig. 6d, panel *Grm5*, Cerebellum), suggesting the importance of control at the level of the premature transcript to regulate the expression of the mature form. Concerning *Grm1*, the pattern of expression in KO relative to WT was rather reversed, with a decrease of the mature transcript sequence in KO (Fig. 6d panel *Grm1*, Cerebellum). However, this decrease, observed in cerebrum and cerebellum, is not associated to a significant alteration of the intronic sequences.

These results further support the notion that G3BP1 and its protein partners are involved in the regulation of the expression levels of intron-retaining transcripts in the cerebellum. The regulation of the expression of the premature and mature *Grm5* and *Grm1* transcripts suggests a control of their translation and/or degradation. Interestingly, the mGluR1 protein, expressed in the cerebellum, was decreased in the cerebellum of G3BP1 KO mice (consistent with the mature transcript down-regulation), while the mGluR5 protein, expressed in the cerebrum, was slightly decreased in the cerebrum but started to be expressed in the G3BP1 KO cerebellum (Fig. 6d). These results do suggest that the G3BP-complex is involved at least in part in the



control of translation of intron-retaining transcripts. The subtle changes in the levels of the protein, however, also suggest that the transcripts may be degraded in the absence of G3BP1.

The opposite G3BP-mediated translational regulation of *Grm1* and *Grm5* transcripts by the G3BP-complex is surprising and could be as a result of differences in binding specificity by the other proteins of the complex. Western blot analysis revealed that Caprin-1 and USP10 have opposite expression levels in the cerebrum and cerebellum (Figure S7b); Caprin-1 is more expressed in the cerebrum, whereas USP10 is more expressed in the cerebellum. Since Caprin-1 and USP10 have been recently found in separate complexes with G3BP and binding of each protein is

mutually exclusive (Kedersha *et al.* 2016), it can be hypothesized that the cerebrum is enriched in G3BP/Caprin-1 complex, whereas cerebellum in enriched in G3BP/USP10 complex. These complexes can mediate differential binding/regulation to *Grm1* and *Grm5* transcripts and thereby influence their fate in opposite manner in the cerebrum and cerebellum. G3BP/Caprin-1 complex favors expression of *Grm5* but represses expression of *Grm1* in the cerebrum, whereas G3BP/USP10 favors expression of *Grm1* but represses *Grm5* in the cerebellum. Strikingly, G3BP depletion led to up-regulation of USP10 and down-regulation of Caprin-1 in the cerebrum but opposite regulation in the cerebellum (Figure S7b). Thus, the little expression of *Grm5* seen in the cerebellum and its down-regulation in the cerebrum of G3BP1 KO mice (Fig. 6d), can be correlated with changes in the expression of Caprin-1 and USP10, which might associate with G3BP2 (left after G3BP1 depletion). Further analyses are required to provide direct evidence for these mutually exclusive interactions of Caprin-1 and USP10 with G3BP1/2 in the mouse brain.

Discussion

Identification of G3BP endogenous RNA targets from mouse brain revealed that G3BP associates predominantly with non-coding regions, intronic and 3'UTR, of coding genes. Proteomics analysis of mouse brain showed that G3BP is stably associated with specific partners; Caprin-1, USP10, CtBP1, and the splicing regulator SFPQ (splicing factor proline and glutamine rich, or PSF), which is also a major component of neuronal RNA granules in the cytoplasm (Kanai *et al.* 2004; Kiebler and Bassell 2006). As G3BP is a major component in the assembly of SGs, where specific protein-coding transcripts are sequestered in a translationally repressed state, we propose that the G3BP-complex is involved in the physiology of the neuron by favoring the accumulation of intron-retaining transcripts, resulting in their translational repression. Also, similar to SGs where transcripts are not degraded, the G3BP-complex could prevent these intron-retaining transcripts from degradation.

While our study did not address the function and the localization of these unspliced transcripts in neurons, a large number of intronic sequences have been reported in some dendritically localized mRNAs of primary rat and mouse hippocampal neurons, among which we found G3BP-complex targets (Buckley *et al.* 2011; Khaladkar *et al.* 2013). One hypothesis is thus that a larger than expected fraction of incompletely spliced transcripts could reach the cytoplasm. As there is good evidence that transport of mRNPs and translational regulation in neuron might be intimately coupled, it would suggest a mechanism whereby G3BP-complex may play an important role in the regulation of stability/translation and/or transport of transcripts with retained introns.

However, as our current study does not differentiate between nuclear and cytoplasmic targets, it is also possible that G3BP protein partners, or G3BP itself as it is proposed to shuttle to the nucleus, interact with transcripts in the nucleus, and either repress splicing of nascent pre-mRNAs, or stabilize nuclear intron-retaining transcripts. A large-scale total RNA-Seq study revealed that many more introns were sequenced in the human brain compared to liver, and even more in the fetal compare to the adult brain (Ameur *et al.* 2011). This high detection of introns is supposed to be linked to splicing regulation, as a pattern of sequenced tags along the long introns could be established and revealed a 'saw-tooth' pattern from 5' to 3' of each intron. Interestingly, HITS-CLIP of fused in sarcoma/translocated in liposarcoma (FUS/TLS) and TAR DNA-binding protein 43 (TDP-43), two RNA-binding proteins involved in amyotrophic lateral sclerosis and frontotemporal lobar degeneration, showed that these two proteins bind long introns of transcripts involved in neuronal integrity and function (Polymenidou *et al.* 2011; Rogelj *et al.* 2012). FUS harbors this 'saw-tooth'-like pattern of binding and is suggested to bind pre-mRNAs co-transcriptionally until splicing is completed (Lagier-Tourenne *et al.* 2012). TDP-43, which binds a majority of introns and 3'UTR sequences as observed for G3BP-complex (Bhardwaj *et al.* 2013), was also suggested to prevent unproductive splicing events of transcripts involved in synaptic plasticity. Although FUS/TLS and TDP-43 have well-characterized roles in splicing regulation in the nucleus, they present dual functions with an important role in cytoplasmic granules aggregation. They are involved in SGs assembly regulation in link to G3BP (Aulas *et al.* 2012), suggesting the importance of these two factors, as well as G3BP, in mRNA processing regulation and repression in different subcellular compartments in neurons.

Furthermore, a mechanism of regulated intron retention is likely to play a major role in gene expression in neurons. For example, Huang and colleagues identified an alternative transcript of apolipoprotein E gene, which retains intron 3. This retention event was specific to neurons and not observed in glia or astrocytes. This transcript is retained in the nucleus, inducing a decrease in the level of apolipoprotein mRNA and the absence of translation. This intron retention is thus a way of regulating the production of apolipoprotein, and the authors showed that under injury, the RNA will then be processed into mature transcript (Xu *et al.* 2008). On a larger scale, the polypyrimidine tract-binding protein Ptbp1 was proposed to regulate the abundance of a number of neuronal transcripts through intron retention and nuclear retention (Yap *et al.* 2012), suggesting again that intron retention may be an important way to regulate gene expression spatiotemporally in the developing brain. Interestingly, the G3BP partner CtBP1, which is known mainly as a transcriptional co-repressor (Stankiewicz *et al.* 2014), seems to be down-regulated in the cerebellum of G3BP1 KO mice (Figure S7b),

suggesting a possible importance of this partner in intron retention in the nucleus.

Accumulation of intron-retaining transcripts suggests that they escape the nonsense-mediated decay (NMD) pathway, which is expected to induce degradation of transcripts harboring a premature stop codon. One way to escape NMD will be that these incompletely spliced pre-mRNAs are retained in the nucleus. Nuclear retention can be caused by interaction of splicing factors with the splice-site consensus sequences, when RNAs are not properly spliced. The splicing factor U2 auxiliary factor that binds to polypyrimidine tracts at the 3' splice site was shown to be involved in nuclear retention activity (Takemura *et al.* 2011). SFPQ also binds polypyrimidine tracts at the 3' splice site and could compete for the binding of U2 auxiliary factor (Peng *et al.* 2006; Takemura *et al.* 2011). The intron-retained transcripts can also escape NMD because they lack factors binding to the cap structure (like cap-binding complex, CBC complex) and/or are captured in translationally repressed particles. It is possible that G3BP-complex destabilizes the association of factors like CBC or eukaryotic translation initiation factor 4E with the cap structure to keep the RNP particle away from the translation machinery. Relevant in this context is the finding that the G3BP-complex preferentially binds the most 5' intron whose splicing is critically dependent on CBC. Previous studies have shown that CBC complex at the cap favor the recruitment of TAP/p15 for efficient export of the mRNP (Cheng *et al.* 2006). The structural similarities between the G3BP homodimer and TAP/p15 heterodimer, would intuitively suggest that they may replace each other during export. If that were the case, mRNPs containing TAP/p15 would be subjected to first round of translation and NMD but not G3BP-containing mRNPs.

Interestingly, we observed a high enrichment of introns in the cerebellum compared to the rest of the brain, suggesting that the proposed regulation of intron retention in the brain could even be more important in the cerebellum. In the case of *Grm5*, two intronic sequences tested are enriched in WT cerebellum compared to WT cerebrum; however, a mature sequence which does not contain one of these introns (primers targeting the adjacent exons) or the 3'UTR sequence (which permits to detect also the mature RNA) are preferentially reduced in the cerebellum compared to the rest of the brain. This is consistent with the fact that expression of *Grm5*, encoding the metabotropic Glutamate receptor 5 (mGluR5), is decreased in the cerebellum during development (Casabona *et al.* 1997). In the absence of G3BP1, however, these intronic sequences are decreased in the cerebellum, whereas the mature transcript is increased. As a negative control, G3BP1 does not affect the premature forms of *Grm1* transcript which is expressed as mature mRNA in the cerebellum, while it induces a decrease in mature mRNA in both the cerebrum and cerebellum. The finding that *Grm1* and *Grm5* transcripts (encoding mGluR1 and mGluR5) are targets of the G3BP-complex, and may be regulated at the level of their premature sequences, is particularly interesting in link to the phenotype associated to G3BP1 deficiency: enhanced mGluR-dependent LTD in the hippocampus and ataxia (Martin *et al.* 2013). The retention and subsequent removal of introns from premature *Grm5* transcript might be a mechanism allowing rapid and post-transcriptionally controlled translation in response to glutamate signaling, especially in the cerebellum. The finding that changes in intronic sequences stability can be correlated with changes in the expression of direct partners of G3BP, Caprin-1, and USP10 which have an opposite expression level in the cerebellum compared to the rest of the brain, implies that the dynamics and/or assembly of G3BP granules could play a direct role in the glutamate signaling. Consistent with this hypothesis are several reports showing that cytoplasmic intron retention plays a key role in calcium signaling (Glanzer *et al.* 2005; Bell *et al.* 2010; Buckley *et al.* 2011) and our finding that G3BP depletion impacts not only intronic sequences stability but also the expression of the mGluR1 and mGluR5 proteins.

Finally, we found that a large variety of non-coding RNAs were also associated to the G3BP-complex. In particular, long non-coding RNAs, snoRNAs and miRNAs are abundant in the brain relative to other tissues, and can encompass or originate from protein-coding gene intronic sequences and be transcribed from the same strand as the pre-mRNA (Qureshi and Mehler 2012). Interestingly, intronic miRNAs tend to be present in large and 5' biased introns in both human and mouse (Zhou and Lin 2008). If we consider only tags clusters associated with G3BP-complex that fall within introns, there is a clear 5'-to-3' bias with clear bias to be in the first intron. It is tempting to speculate that this enrichment may be linked to the production of small ncRNAs such as intronic miRNAs. The use of primers flanking the exon–intron junctions, and the enrichment of other introns, further support the finding that the G3BP-complex contains sequences derived from the pre-messenger transcripts rather than the non-coding RNAs themselves. Recently, an intron retained in a newly identified splice variant of hDKC1 (human dyskeratosis congenita 1) and not degraded by the NMD pathway, was shown to encode a snoRNA, whose expression would be tissue and time specific (Turano *et al.* 2013). From these observations, an appealing model emerges for post-trancriptional gene regulation, in which a fraction of intron-retaining transcripts participate in regulatory modulation. Processing of these transcripts to remove non-coding sequence could produce a translatable transcript in the cytoplasm in addition to potentially other intron-encoded RNAs that may further regulate either their own host transcript or a different gene's product. Since neuronal activity and synaptic signaling are known to affect mRNA transport and translation *in vivo*, analysis of their regulation promises to provide important information about the

maintenance and enhancement of brain tissue viability and function. We speculate that mechanisms surrounding G3BP RNA granules allow for synapse-specific modifications, thereby yielding molecular, structural, and functional reorganization of individual synapses that occur during neuronal development and synaptic plasticity, processes which may go awry in neurological diseases such as Alzheimer and ataxias.

Acknowledgments and conflict of interest disclosure

This work was supported by a Fondation pour la Recherche Médicale (FRM) grant (Equipe FRM 2011 – no. DEQ20111223745). S.M. was supported by a graduate fellowship from the Ministère Délégué à la Recherche et aux Technologies and FRM. J.G.V., N.B. and E.E. were supported by the Spanish Ministry of Science with grants BIO2011-23920 and CSD2009-00080. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript. The authors declare to have no conflict of interest.

All experiments were conducted in compliance with the ARRIVE guidelines.

Supporting information

Additional Supporting Information may be found online in the supporting information tab for this article:

Figure S1. (a) G3BP1 IP was performed with a different G3BP1 antibody, which also revealed additional proteins immunoprecipitated under CLIP conditions, as seen in a silver staining of the SDS-PAGE gel. (b) G3BP1 and G3BP2 co-localize in primary hippocampal neurons (i), as well as in stress granules formed under arsenite treatment (ii). DNA is counter-stained with Hoechst staining. Scale bar represents 5 μm. (c) Different highly stringent washes did not permit to eliminate the presence of the G3BP partners and the separation of three ribonucleoproteic complexes under high RNase treatment (a). The numbers 1 to 4 are indicative of the four complexes detailed in Figure 2A and B. In (b), different RNAse concentrations in normal CLIP washes conditions show G3BP-complex gradually shifting to high molecular weight as RNAse treatment decreases.

Figure S2. (a) Long motif overrepresented in G3BP-complex clusters, part of a SINE-Alu-B1 transposable element. (b) Around 8 % of the clusters possess this sequence with an occurrence > 0, with a peak in the center of the clusters. When looking for motifs overrepresented in G3BP-complex clusters compared to background which are not part of repeated sequences, several sequences are found, less enriched compared to the logo in Figure S2A, and less centered. (c) (1) Consensus from two 7mers: CACTCTG + GACTCTG. This motif is still part of Alu repeat elements, and may be missed by the RepeatMasker tool. (2) Consensus from two 7-mers: CCCTCCC + CCCACCC. (3) Consensus binding motif identified for G3BP2 by RNAcompete. (4) Center of consensus binding motif identified for G3BP1 by SELEX.

Figure S3. HITS-CLIP of G3BP2 in the G3BP1 KO mouse brain.

Figure S4. (a) Distribution of G3BP HITS-CLIP clusters within introns along the RNA, from 5' to 3', for G3BP1 and G3BP2, or G3BP1 only. The position is 5' biased (although not at the most 5' decile, presumably because the first intron of a transcript is often very long). (b) Expression levels of genes with G3BPs clusters compared to expression levels of the genes in neural or non-neural tissues determined from RNA-Seq experiments ('Neural': 6 RNAseq samples: 2 of whole brain, one of cerebellum, one retina, and isolated neurons (cerebellar granular neurons, neurons from Dorsal root ganglia), and non-neural tissues: 'Rest': 14 samples: 2 of Liver, 2 kidney, 2 Heart, Muscle, Myoblast 168h of differentiation, T cells, Testis, Lung fibroblasts, 3T3 cell line and 2× embryonic fibroblasts). Five groups of expression were made from these RNASeq data, based on average corrected cRPKMs (RPKMs with some extra corrections): (i) (nearly) no expression (av. cRPKM<2), (ii) low expression (2–10), (iii) moderate expression (10–25), (iv) high expression (25–50), (v) very high expression (> 50).

Figure S5. Consensus sequences analysis in introns.

Figure S6. Images from UCSC genome browser showing the structure of the gene and the distribution of the clusters for representative genes of G3BP-complex HITS-CLIP.

Figure S7. The G3BP-complex regulates immature transcripts in the cerebellum.

Appendix S1. Supplementary Materials and methods.

References

Althammer S., González-Vallinas J., Ballaré C., Beato M. and Eyras E. (2011) Pyicos: a versatile toolkit for the analysis of high-throughput sequencing data. *Bioinformatics* **27**, 3333–3340.

Ameur A., Zaghlool A., Halvardson J., Wetterbom A., Gyllensten U., Cavelier L. and Feuk L. (2011) Total RNA sequencing reveals nascent transcription and widespread co-transcriptional splicing in the human brain. *Nat. Struct. Mol. Biol.* **18**, 1435–1440.

Anderson P. and Kedersha N. (2009) Stress granules. *Curr. Biol.* **19**, R397–R398.

Annibaldi A., Dousse A., Martin S., Tazi J. and Widmann C. (2011) Revisiting G3BP1 as a RasGAP binding protein: sensitization of tumor cells to chemotherapy by the RasGAP 317-326 sequence does not involve G3BP1. *PLoS One*, **6**, e29024.

Aulas A., Stabile S. and Vande Velde C. (2012) Endogenous TDP-43, but not FUS, contributes to stress granule assembly via G3BP. *Mol. Neurodegener.* **7**, 54.

Bell T. J., Miyashiro K. Y., Sul J.-Y. *et al.* (2010) Intron retention facilitates splice variant diversity in calcium-activated big potassium channel populations. *Proc. Natl Acad. Sci. USA* **107**, 21152–21157.

Bhardwaj A., Myers M. P., Buratti E. and Baralle F. E. (2013) Characterizing TDP-43 interaction with its RNA targets. *Nucleic Acids Res.* **41**, 5062–5074.

Buckley P. T., Lee M. T., Sul J.-Y., Miyashiro K. Y., Bell T. J., Fisher S. A., Kim J. and Eberwine J. (2011) Cytoplasmic intron sequence-retaining transcripts can be dendritically targeted via ID element retrotransposons. *Neuron* **69**, 877–884.

Casabona G., Knöpfel T., Kuhn R., Gasparini F., Baumann P., Sortino M. A., Copani A. and Nicoletti F. (1997) Expression and coupling to polyphosphoinositide hydrolysis of group I metabotropic glutamate receptors in early postnatal and adult rat brain. *Eur. J. Neurosci.* **9**, 12–17.

Chanas-Sacré G., Mazy-Servais C., Wattiez R., Pirard S., Rogister B., Patton J. G., Belachew S., Malgrange B., Moonen G. and Leprince

P. (1999) Identification of PSF, the polypyrimidine tract-binding protein-associated splicing factor, as a developmentally regulated neuronal protein. *J. Neurosci. Res.* **57**, 62–73.

Cheng H., Dufu K., Lee C.-S., Hsu J. L., Dias A. and Reed R. (2006) Human mRNA export machinery recruited to the 5′ end of mRNA. *Cell* **127**, 1389–1400.

Costa M., Ochem A., Staub A. and Falaschi A. (1999) Human DNA helicase VIII: a DNA and RNA helicase corresponding to the G3BP protein, an element of the ras transduction pathway. *Nucleic Acids Res.* **27**, 817–821.

Darnell R. B. (2010) HITS-CLIP: panoramic views of protein-RNA regulation in living cells. *Wiley Interdiscip. Rev. RNA* **1**, 266–286.

Glanzer J., Miyashiro K. Y., Sul J.-Y., Barrett L., Belt B., Haydon P. and Eberwine J. (2005) RNA splicing capability of live neuronal dendrites. *Proc. Natl Acad. Sci. USA* **102**, 16859–16864.

Hirose T. and Steitz J. A. (2001) Position within the host intron is critical for efficient processing of box C/D snoRNAs in mammalian cells. *PNAS* **98**, 12914–12919.

Hosack D. A., Dennis G. J., Sherman B. T., Lane H. C. and Lempicki R. A. (2003) Identifying biological themes within lists of genes with EASE. *Genome Biol.* **4**, R70.

Hübler D., Rankovic M., Richter K., Lazarevic V., Altrock W. D., Fischer K.-D., Gundelfinger E. D. and Fejtova A. (2012) Differential spatial expression and subcellular localization of CtBP family members in rodent brain. *PLoS ONE* **7**, e39710.

Kanai Y., Dohmae N. and Hirokawa N. (2004) Kinesin transports RNA: isolation and characterization of an RNA-transporting granule. *Neuron* **43**, 513–525.

Kedersha N. and Anderson P. (2009) Regulation of translation by stress granules and processing bodies. *Prog. Mol. Biol. Transl. Sci.* **90**, 155–185.

Kedersha N., Panas M. D., Achorn C. A. *et al.* (2016) G3BP-Caprin1-USP10 complexes mediate stress granule condensation and associate with 40S subunits. *J. Cell Biol.* **212**, 845–860.

Kennedy D., Wood S. A., Ramsdale T., Tam P. P., Steiner K. A. and Mattick J. S. (1996) Identification of a mouse orthologue of the human ras-GAP-SH3-domain binding protein and structural confirmation that these proteins contain an RNA recognition motif. *Biomed. Pept. Proteins Nucleic Acids* **2**, 93–99.

Kennedy D., French J., Guitard E., Ru K., Tocque B. and Mattick J. (2001) Characterization of G3BPs: tissue specific expression, chromosomal localisation and rasGAP(120) binding studies. *J. Cell. Biochem.* **84**, 173–187.

Kent H. M., Clarkson W. D., Bullock T. L. and Stewart M. (1996) Crystallization and preliminary X-ray diffraction analysis of nuclear transport factor 2. *J. Struct. Biol.* **116**, 326–329.

Khaladkar M., Buckley P. T., Lee M. T., Francis C., Eghbal M. M., Chuong T., Suresh S., Kuhn B., Eberwine J. and Kim J. (2013) Subcellular RNA sequencing reveals broad presence of cytoplasmic intron-sequence retaining transcripts in mouse and rat neurons. *PLoS ONE* **8**, e76194.

Kiebler M. A. and Bassell G. J. (2006) Neuronal RNA granules: movers and makers. *Neuron* **51**, 685–690.

Kobayashi T., Winslow S., Sunesson L., Hellman U. and Larsson C. (2012) PKCα binds G3BP2 and regulates stress granule formation following cellular stress. *PLoS ONE* **7**, e35820.

Kristensen O. (2015) Crystal structure of the G3BP2 NTF2-like domain in complex with a canonical FGDF motif peptide. *Biochem. Biophys. Res. Commun.* **467**, 53–57.

Lagier-Tourenne C., Polymenidou M., Hutt K. R., Vu A. Q., Baughn M., Huelga S. C., Clutario K. M. *et al.* (2012) Divergent roles of ALS-linked proteins FUS/TLS and TDP-43 intersect in processing long pre-mRNAs. *Nat. Neurosci.* **15**, 1488–1497.

Langmead B., Trapnell C., Pop M. and Salzberg S. L. (2009) Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol.* **10**, R25.

Licatalosi D. D., Mele A., Fak J. J. *et al.* (2008) HITS-CLIP yields genome-wide insights into brain alternative RNA processing. *Nature* **456**, 464–469.

Macias S., Plass M., Stajuda A., Michlewski G., Eyras E. and Cáceres J. F. (2012) DGCR8 HITS-CLIP reveals novel functions for the Microprocessor. *Nat. Struct. Mol. Biol.* **60**, 873–885.

Marco-Sola S., Sammeth M., Guigó R. and Ribeca P. (2012) The GEM mapper: fast, accurate and versatile alignment by filtration. *Nat. Methods* **9**, 1185–1188.

Martin S., Zekri L., Metz A., Maurice T., Chebli K., Vignes M. and Tazi J. (2013) Deficiency of G3BP1, the stress granules assembly factor, results in abnormal synaptic plasticity and calcium homeostasis in neurons. *J. Neurochem.* **125**, 175–184.

Matsuki H., Takahashi M., Higuchi M., Makokha G. N., Oie M. and Fujii M. (2013) Both G3BP1 and G3BP2 contribute to stress granule formation. *Genes Cells* **18**, 135–146.

Nardini M., Spanò S., Cericola C., Pesce A., Massaro A., Millo E., Luini A., Corda D. and Bolognesi M. (2003) CtBP/BARS: a dual-function protein involved in transcription co-repression and Golgi membrane fission. *EMBO J.* **22**, 3122–3130.

Parker F., Maurier F., Delumeau I., Duchesne M., Faucher D., Debussche L., Dugue A., Schweighoffer F. and Tocque B. (1996) A Ras-GTPase-activating protein SH3-domain-binding protein. *Mol. Cell. Biol.* **16**, 2561–2569.

Peng R., Hawkins I., Link A. J. and Patton J. G. (2006) The splicing factor PSF is part of a large complex that assembles in the absence of pre-mRNA and contains all five snRNPs. *RNA Biol.* **3**, 69–76.

Polymenidou M., Lagier-Tourenne C., Hutt K. R. *et al.* (2011) Long pre-mRNA depletion and RNA missplicing contribute to neuronal vulnerability from loss of TDP-43. *Nat. Neurosci.* **14**, 459–468.

Qureshi I. A. and Mehler M. F. (2012) Emerging roles of non-coding RNAs in brain evolution, development, plasticity and disease. *Nat. Rev. Neurosci.* **13**, 528–541.

Ray D., Kazan H., Cook K. B., Weirauch M. T., Najafabadi H. S., Li X., Gueroussov S. *et al.* (2013) A compendium of RNA-binding motifs for decoding gene regulation. *Nature* **499**, 172–177.

Rogelj B., Easton L. E., Bogu G. K., Stanton L. W., Rot G., Curk T., Zupan B. *et al.* (2012) Widespread binding of FUS along nascent RNA regulates alternative splicing in the brain. *Sci. Rep.* **2**, 603.

Solomon S., Xu Y., Wang B., David M. D., Schubert P., Kennedy D. and Schrader J. W. (2007) Distinct structural features of caprin-1 mediate its interaction with G3BP-1 and its induction of phosphorylation of eukaryotic translation initiation factor 2alpha, entry to cytoplasmic stress granules, and selective interaction with a subset of mRNAs. *Mol. Cell. Biol.* **27**, 2324–2342.

Soncini C., Berdo I. and Draetta G. (2001) Ras-GAP SH3 domain binding protein (G3BP) is a modulator of USP10, a novel human ubiquitin specific protease. *Oncogene* **20**, 3869–3879.

Stankiewicz T. R., Gray J. J., Winter A. N. and Linseman D. A. (2014) C-terminal binding proteins: central players in development and disease. *Biomol. Concepts* **5**, 489–511.

Stutz F. and Izaurralde E. (2003) The interplay of nuclear mRNP assembly, mRNA surveillance and export. *Trends Cell Biol.* **13**, 319–327.

Suyama M., Doerks T., Braun I. C., Sattler M., Izaurralde E. and Bork P. (2000) Prediction of structural domains of TAP reveals details of its interaction with p15 and nucleoporins. *EMBO Rep.* **1**, 53–58.

Takemura R., Takeiwa T., Taniguchi I., McCloskey A. and Ohno M. (2011) Multiple factors in the early splicing complex are involved

in the nuclear retention of pre-mRNAs in mammalian cells. *Genes Cells* **16**, 1035–1049.

Thomas M. G., Loschi M., Desbats M. A. and Boccaccio G. L. (2011) RNA granules: the good, the bad and the ugly. *Cell. Signal.* **23**, 324–334.

Tourrière H., Gallouzi I. E., Chebli K., Capony J. P., Mouaikel J., van der Geer P. and Tazi J. (2001) RasGAP-associated endoribonuclease G3Bp: selective RNA degradation and phosphorylation-dependent localization. *Mol. Cell. Biol.* **21**, 7747–7760.

Tourrière H., Chebli K., Zekri L., Courselaud B., Blanchard J. M., Bertrand E. and Tazi J. (2003) The RasGAP-associated endoribonuclease G3BP assembles stress granules. *J. Cell Biol.* **160**, 823–831.

Turano M., Angrisani A., Di Maio N. and Furia M. (2013) Intron retention: a human DKC1 gene common splicing event. *Biochem. Cell Biol.* **91**, 506–512.

Vognsen T. and Kristensen O. (2012) Crystal structure of the Rasputin NTF2-like domain from Drosophila melanogaster. *Biochem. Biophys. Res. Commun.* **420**, 188–192.

Vognsen T., Möller I. R. and Kristensen O. (2011) Purification, crystallization and preliminary X-ray diffraction of the G3BP1 NTF2-like domain. *Acta Crystallogr. Sect. F Struct. Biol. Cryst. Commun.* **67**, 48–50.

Wang Z., Tollervey J., Briese M., Turner D. and Ule J. (2009) CLIP: construction of cDNA libraries for high-throughput sequencing from RNAs cross-linked to proteins in vivo. *Methods* **48**, 287–293.

Xu Q., Walker D., Bernardo A., Brodbeck J., Balestra M. E. and Huang Y. (2008) Intron-3 retention/splicing controls neuronal expression of apolipoprotein E in the CNS. *J. Neurosci.* **28**, 1452–1459.

Yap K., Lim Z. Q., Khandelia P., Friedman B. and Makeyev E. V. (2012) Coordinated regulation of neuronal mRNA steady-state levels through developmentally controlled intron retention. *Genes Dev.* **26**, 1209–1223.

Zekri L., Chebli K., Tourrière H., Nielsen F. C., Hansen T. V. O., Rami A. and Tazi J. (2005) Control of fetal growth and neonatal survival by the RasGAP-associated endoribonuclease G3BP. *Mol Cell Biol*, **25**, 8703–8716.

Zhou H. and Lin K. (2008) Excess of microRNAs in large and very 5' biased introns. *Biochem. Biophys. Res. Commun.* **368**, 709–715.