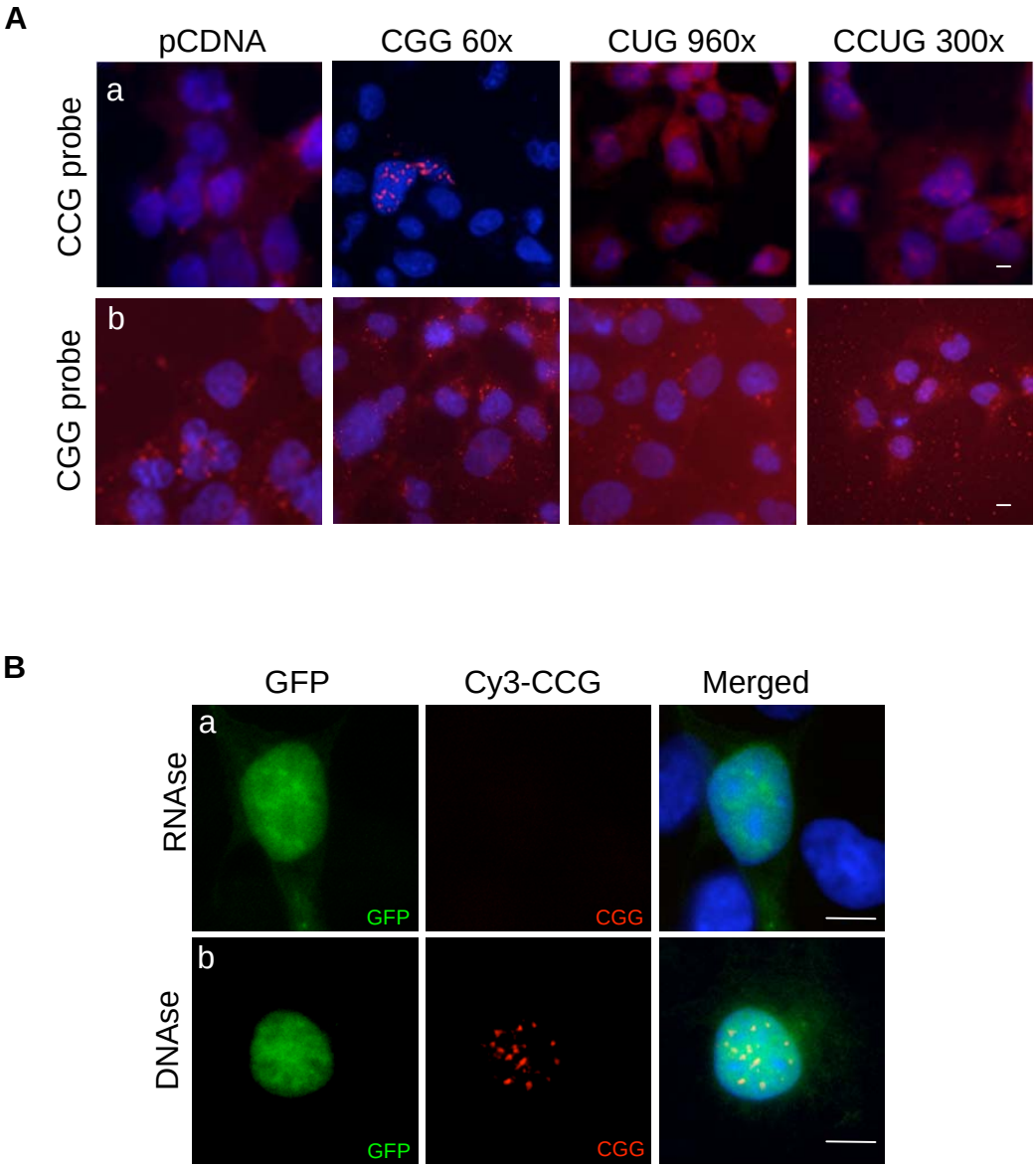
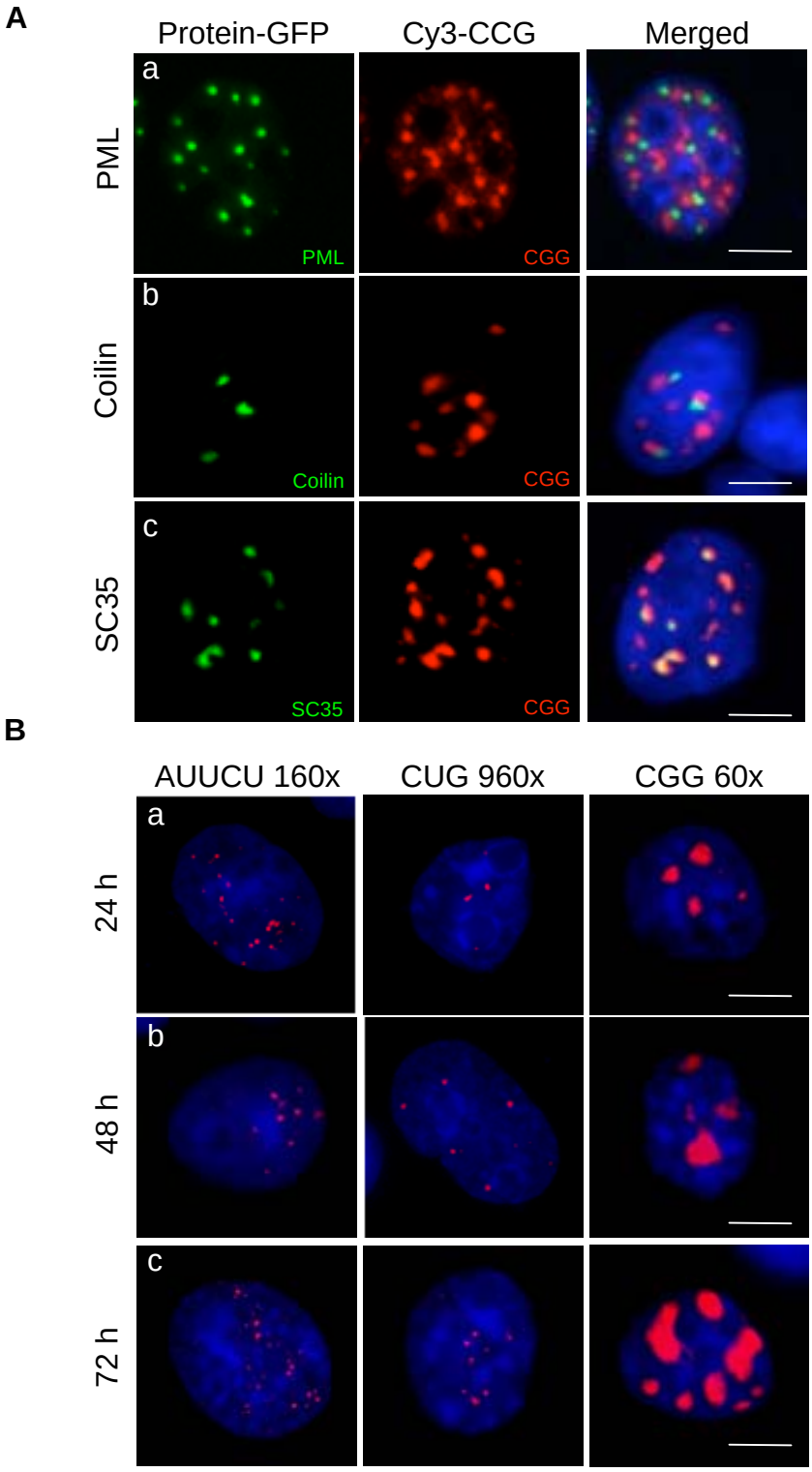
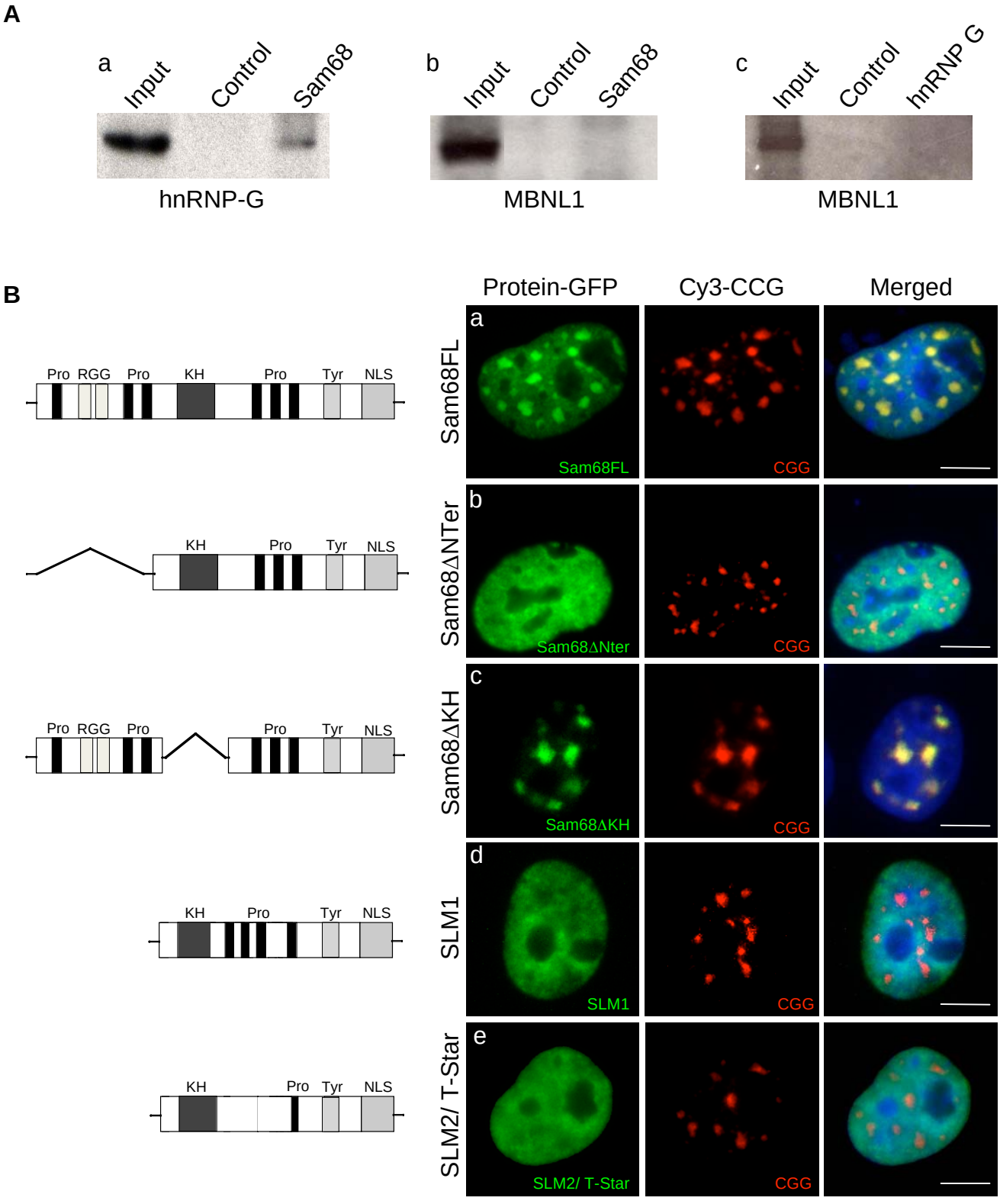


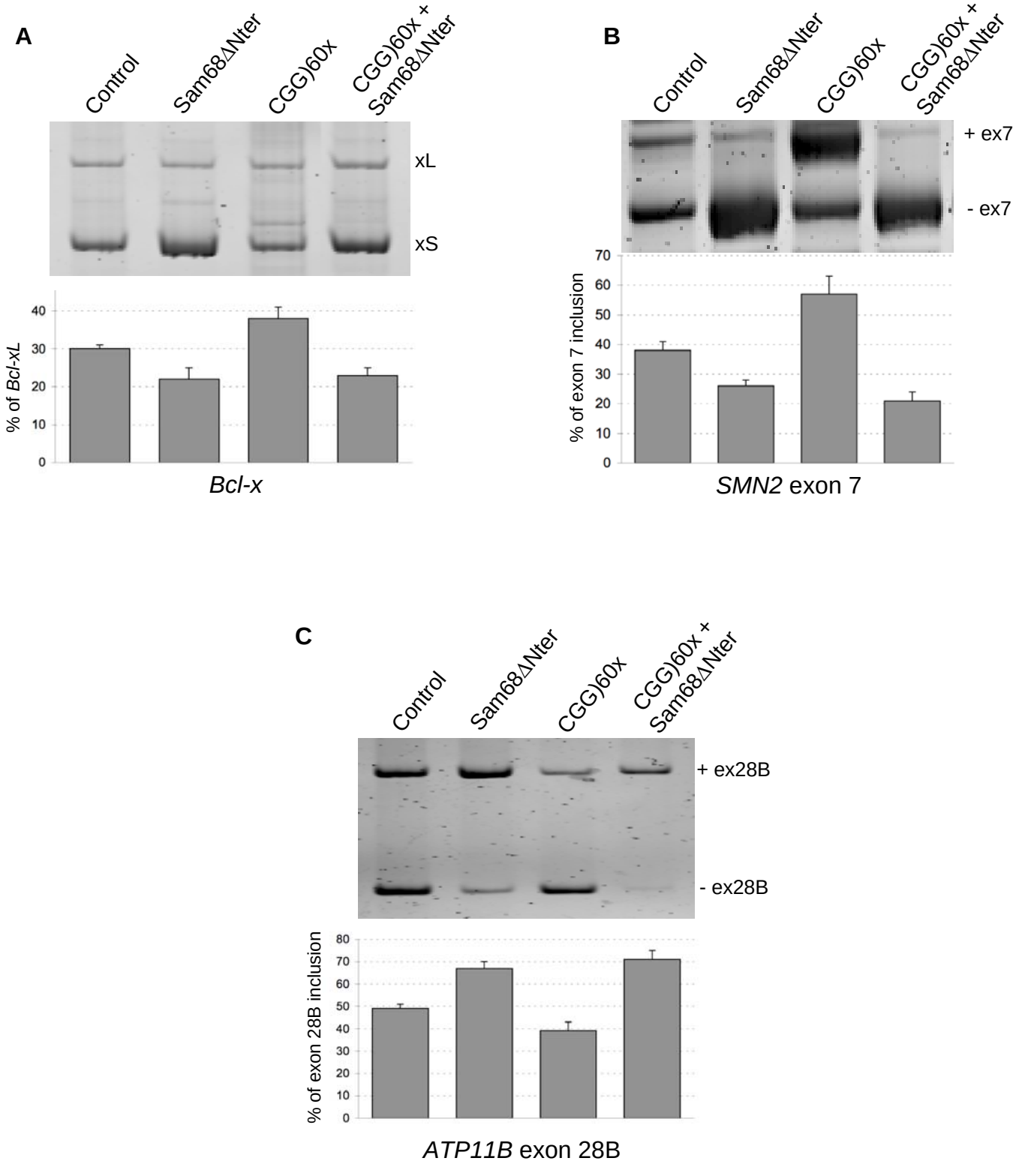
Supplementary Figure S1: (**A**) Aggregates of CGG repeats are specifically recognized by a (CCG)8x-Cy3 DNA oligonucleotide probe. COS7 cells were co-transfected with plasmids expressing either 60 CGG, 960 CUG or 300 CCUG repeats and analyzed 24 hours after transfection by RNA FISH. (a) CCG)8x-Cy3 DNA probe hybridized and recognized only CGG repeats and not CUG or CCUG repeats, while (b) (CGG)8x-Cy3 DNA probe recognized none of these repeats.
(**B**) Expanded CGG repeats form intranuclear RNA aggregates that are sensitive to RNAse but not to DNAse treatment. COS7 cells were co-transfected with a plasmid expressing GFP and a plasmid expressing 60 CGG repeats and analyzed 24 hours after transfection by RNA FISH preceded by RNAse (a) or DNAse (b) treatment. For RNAse treatment, cells were incubated in presence of 0.01g/ml DNAse-free RNAse (Sigma) after permeabilization. For DNAse treatment, cells were incubated in 1 U/㎕ RNAse-free DNAse I recombinant (Roche) with 1X incubation buffer. In all supplementary figures, the nuclei were counterstained with DAPI, one representative experiment from at least three separate experiments is shown, and scale bars represent 10 µm.



Supplementary Figure S2: (**A**) CGG RNA aggregates co-localize partially with nuclear speckles in transfected cells. COS7 cells were co-transfected with a plasmid expressing 60 CGG repeats and a plasmid expressing either GFP-PML1 (a), GFP-coilin (b) or GFP-SC35 (c) and analyzed 24 hours after transfection by RNA FISH. The accumulation of CGG repeats in speckles suggests that CGG-repeat containing mRNAs are recognized as abberant and are blocked in their nuclear processing. Similar observations were reported for expanded CUG-repeat aggregates (Holt and al., 2007). (**B**) COS7 cells were transfected with a plasmid expressing either 60 CGG, 960 CUG or 160 AUUCU repeats and analyzed by RNA FISH either 24 (a), 48 (b) or 72 (c) hours after transfection. In contrast to AUUCU, CCUG or CUG aggregates (0.1 to 0.2 µm), in transfected cells, CGG aggregates are bigger (1 to 5 µm) and enlarge upon time. Scale bar, 10 µm.

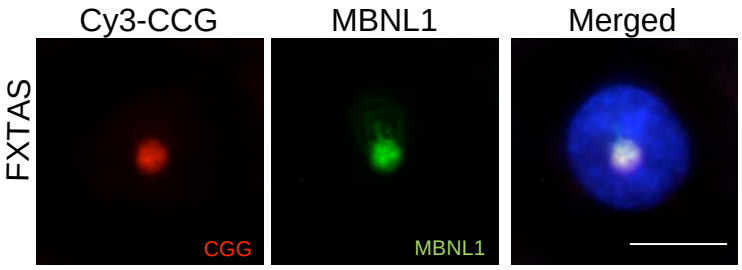


Supplementary Figure S3: (**A**) CO7 cells were lyzed and Sam68 (a and b) or hnRNP G (c) were immunoprecipitated in presence of RNase A, 150 mM NaCl, 20 mM HEPES, 2 mM MgCl2, 1 mM DTT. Immunoprecipitated proteins were tested for presence of endogenous hnRNP G (a) or MBNL1 (b and c) by western-blotting. (**B**) Sam68 co-localization within CGG RNA aggregates required its Sam68 N-terminal part. COS7 cells were co-transfected with a plasmid expressing 60 CGG repeats and a plasmid expressing GFP-Sam68 (a), GFP-Sam68 deleted for its N-terminal part (Sam68ΔNter) (b), GFP-Sam68 deleted of its KH RNA binding domain (c), GFP-Sam68 paralogs SLM1 (d) or SLM2 (e) and analyzed 24 hours after transfection by RNA FISH. Scale bar, 10 µm.

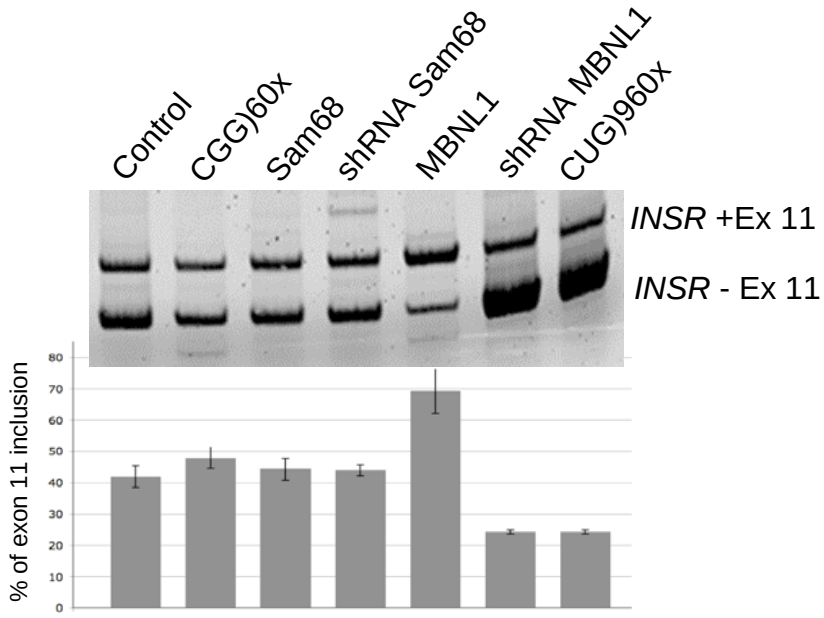


Supplementary Figure S4: Expression of a mutant of Sam68 (Sam68ΔNter), which is not sequestered within CGG aggregates, rescued the deleterious effects of CGG repeats on splicing. COS7 cells were co-transfected with a *Bcl-x* minigene (**A**), a *SMN2* exon 7 minigene (**B**), or an *ATP11B* exon 28B minigene (**C**) and a plasmid expressing either no CGG repeat (Control), Sam68ΔNter, 60 CGG repeats or Sam68ΔNter + 60 CGG repeats. 24 hours after transfection RNA was extracted and analyzed by RT–PCR.

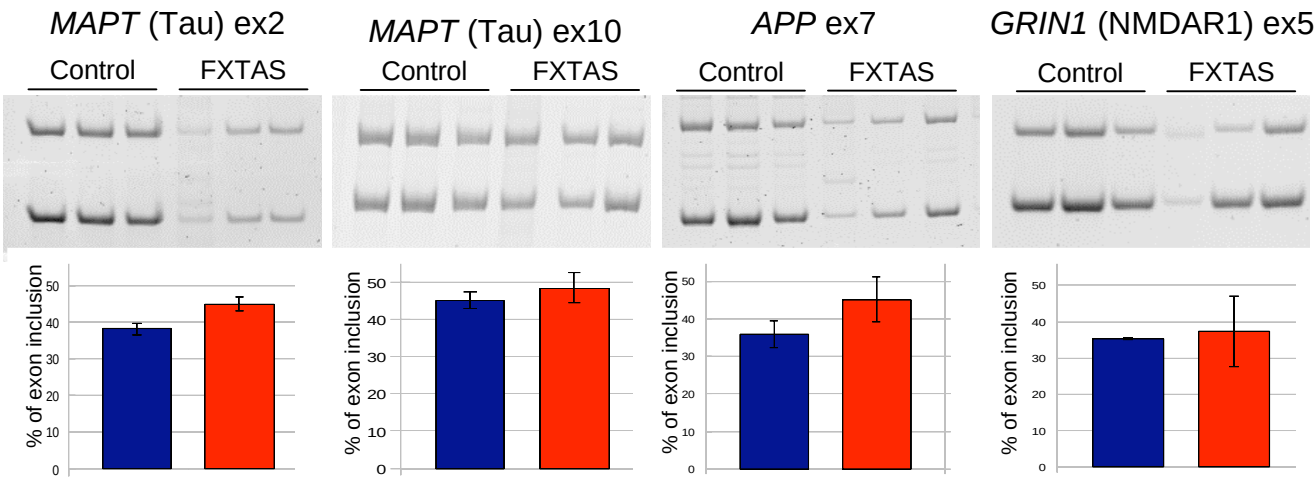
A


Cy3-CCG
MBNL1
Merged
FXTAS
CGG
MBNL1

B


Control
CGG)60x
Sam68
shRNA Sam68
MBNL1
shRNA MBNL1
CUG)960x
INSR +Ex 11
INSR - Ex 11
% of exon 11 inclusion
0
10
20
30
40
50
60
70
80

C


MAPT (Tau) ex2
MAPT (Tau) ex10
APP ex7
GRIN1 (NMDAR1) ex5
Control
FXTAS
% of exon inclusion
0
10
20
30
40
50

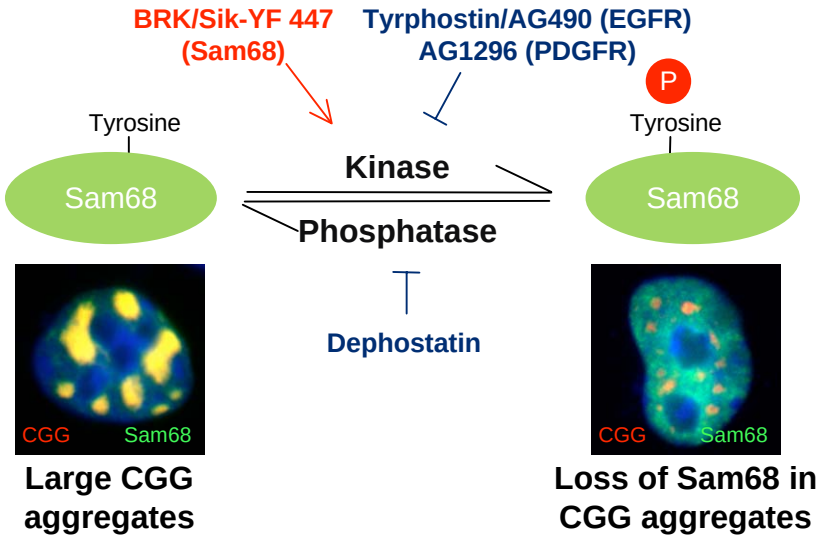
Supplementary Figure S5:

(**A**) MBNL1 co-localizes within CGG aggregates in FXTAS patients. Brain sections of FXTAS patients were analyzed by RNA FISH using an (CCG)8x-Cy3 probe combined with immunofluorescence against MBNL1 (polyclonal antibody, gift from Charles Thornton) using an Alexa488 secondary antibody. Scale bar, 10 µm. MBNL1 co-localization with CGG aggregates questions whether MBNL1 is sequestered within CGG aggregates and whether FXTAS and DM diseases share a similar pathogenetic mechanism.

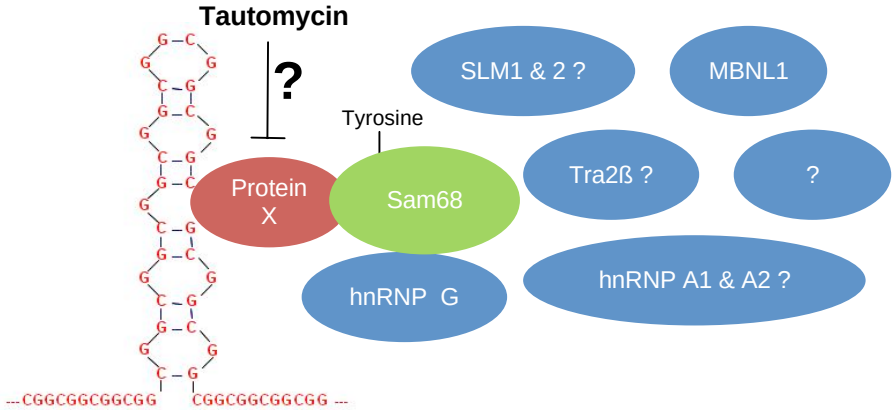
(**B**) Splicing activity of MBNL1 is not altered by CGG repeats. CO7 cells were co-transfected with a minigene containing the exon 11 of the insulin receptor and a plasmid expressing either Sam68, MBNL1, shRNA directed against Sam68 or MBNL1, 60 CGG or 960 CUG repeats. Alternative splicing of the exon 11 of the insulin receptor minigene was analyzed 72 hours after transfection with similar RT-PCR protocol than Ho et al., 2004. Splicing of the exon 11 of the insulin receptor is regulated by MBNL1 and by CCUG expanded repeats. The absence of regulation by CGG repeats indicates that, while MBNL1 co-localize within both CGG and CUG aggregates, MBNL1 is immobilized and consequently loses its splicing regulatory function only in presence of CUG repeats.

(**C**) Splicing activity of MBNL1 is not altered in FXTAS patients. Alternative splicing of *MAPT*, *APP* and *GRIN1* pre-mRNAs was analyzed in control and FXTAS brain samples with similar RT-PCR protocol than Jiang et al., 2004. The splicing of *MAPT*, *APP* and *GRIN1* splicing is altered in DM patients, which is consistent with a sequestration of MBNL1 within CUG aggregates (Jiang et al., 2004). In contrast, we found that the alternative splicing of *MAPT*, *APP* and GRIN1 pre-mRNAs is not altered in FXTAS patients, suggesting that while MBNL1 co-localize with CGG aggregates in FXTAS, MBNL1 is not sequestered within these aggregates, and consequently does not lose its splicing function.

A



B

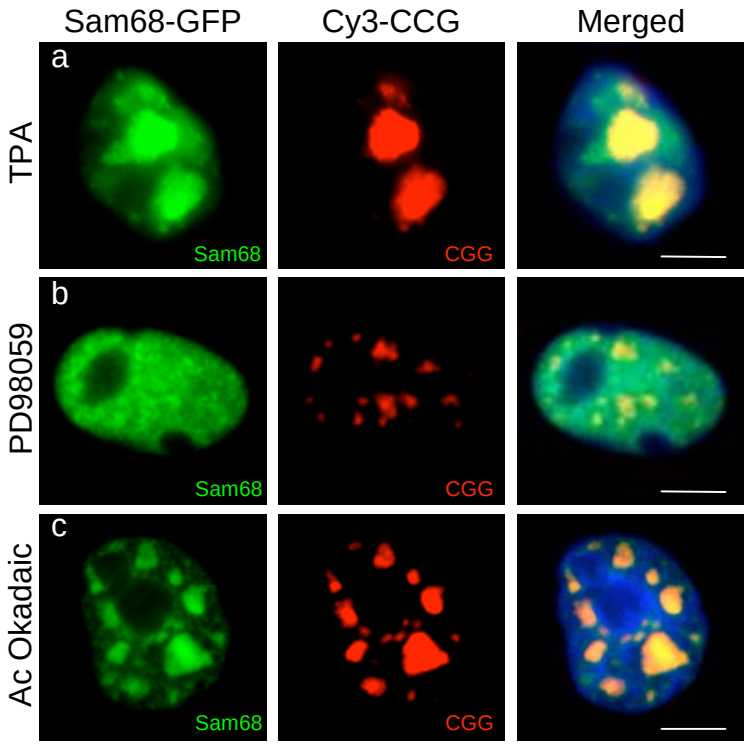


Supplementary Figure S6:

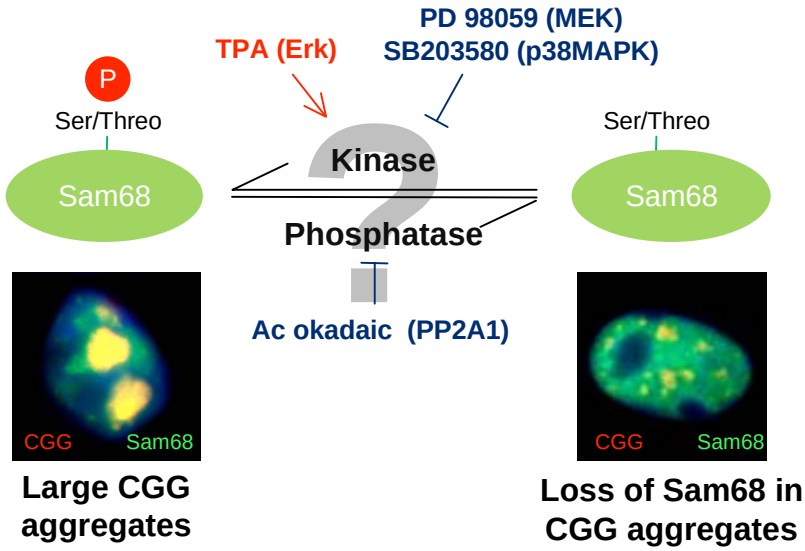
(**A**) Model of Sam68 Tyrosine phosphorylation and its presence within CGG aggregates.

(**B**) Model of CGG RNA aggregate formation. An unknown protein binds to expanded CGG repeats and recruits Sam68 through protein-protein interactions which involves the N-terminal part of Sam68. Then, Sam68 recruits through protein-protein inteactions hnRNP G and other unidentified proteins, which recruit MBNL1. Preliminary experiments show that proteins (SLM1, SLM2, hnRNP A and Tra2ß) known to interact with Sam68 are recruited within CGG aggregates. We propose that the late and indirect recruitment of MBNL1 at the edge of the CGG aggregates may explain why MBNL1 is not immobilized within CGG aggregates and why MBNL1 splicing function is not lost in FXTAS patients. In contrast, we propose that Sam68 is recruited early on, and within the core of the CGG aggregates, resulting in its sequestration and loss of function in FXTAS patients.

A


Sam68-GFP
Cy3-CCG
Merged
TPA
PD98059
Ac Okadaic
a
b
c
Sam68
CGG

B


TPA (Erk)
PD 98059 (MEK)
SB203580 (p38MAPK)
P
Ser/Threo
Sam68
Kinase
Phosphatase
Ac okadaic (PP2A1)
CGG
Sam68
Large CGG aggregates
Loss of Sam68 in CGG aggregates

Supplementary Figure S7: Sam68 has been reported as a substrate of the ERK-serine/threonine kinase, and stimulation of the Ras-MAP kinase pathway by 12-O-tetradecanoyl phorbol-13-acetate (TPA) treatment enhances Sam68 splicing activity (Matter et al., 2002). We thus tested the effect of serine/threonine phosphorylation on Sam68/ CGG aggregate formation.

(**A**) COS7 cells were co-transfected with a plasmid expressing GFP-Sam68 and a plasmid expressing 60 CGG repeats and incubated in presence of 40 ng/ml of TPA (a), 50 µM of PD98059 (b), 4 nM of acid okadaic (c) and analyzed 24 hours after transfection by RNA FISH. Activation of the ERK-serine/threonine kinase pathway by TPA enhanced Sam68 recruitment within CGG repeats and the formation of giant CGG aggregates at early time points after transfection. Consistent with a role of serine/threonine phosphorylation for Sam68 recruitment within CGG aggregates, treatment with PD98059 and SB203580, which are inhibitors of MEK and P38 serine/threonine kinases respectively, reduced Sam68 co-localization with CGG RNA aggregates from 80% to ~30%. In contrast, inhibition of serine/threonine phosphatases using okadaic acid promoted formation of giant aggregates as early as 24 hours after transfection. Scale bar, 10 µm.

(**B**) Model of Sam68 recruitment within CGG aggregates according regulation of its serine/ threonine phosphorylation. Tyrosine and serine/threonine phosphorylation of Sam68 have antagonistic effects on its activity. Stimulation of the Ras-MAP kinase pathway by TPA treatment enhances Sam68 splicing activity (Matter et al., 2002), whereas activation of the EGF-BRK/Sik pathway alters the RNA-binding ability and localization of Sam68 protein (Haegebarth et al., 2004; Lukong et al., 2005). Consistent with opposite effects on Sam68 RNA-binding ability, serine/threonine and tyrosine phosphorylation have antagonistic effects on Sam68 co-localization within CGG aggregates. Whereas activation of the serine/threonine MAPK pathway stimulates Sam68 recruitment and formation of giant CGG aggregates, the BRK/Sik tyrosine-kinase pathway reduces Sam68 localization with CGG RNA aggregates. Overall these results suggest that serine/threonine phosphorylation enhances formation of large CGG aggregates, probably by stimulating the RNA- and protein-binding capacity of Sam68 (Matter et al., 2002). In contrast, tyrosine phosphorylation reduces Sam68 recruitment within CGG aggregates, probably by reducing Sam68 RNA- and protein-binding ability (Derry et al., 2000; Haegebarth et al., 2004; Lukong et al., 2005).

Please, note that this is a putative model as we were unable to confirm direct serine/ threonine phosphorylation of Sam68 after drug treatments due to the poor quality of the anti- serine or anti-threonine antibodies that we tested.

Protein class	Protein identified by Maldi	24 hours
RNA BINDING PROTEIN	HnRNP A1	-
	HnRNPA1 like	
	HNRNPA2/B1	-
	HnRNP A3	
	HnRNP C	-
	HnRNP D	-
	HnRNPE1	-
	hnRNPG	+/-
	hnRNP M	
	Muscleblind like protein 1	+/-
	Splicing factor, arginine/serine rich 1	speckle
	Splicing factor, arginine/serine rich 4	speckle
	Splicing factor, arginine/serine rich 5	speckle
	Splicing factor, arginine/serine rich 6	speckle
	Splicing factor, arginine/serine rich 7	speckle
	Splicing factor, arginine/serine rich 10	speckle
	Spermatid perinuclear RNA-binding protein	-
	Nucleophosmin	-
RNA HELICASE	ATP-dependant RNA helicase DDX3	speckle
	ATP-dependant RNA helicase DDX5	speckle
	ATP-dependant RNA helicase DDX17	speckle
TRANSCRIPTION AND DNA BINDING FACTOR	Interleukin enhancer binding factor 2	
	Nuclear factor kappa B repressing factor	
	Metastasis-associated protein MTA2	
CHAPERONE AND DEGRADATION	HSP27	-
	Ubiquitin	-
CYTOSKELETON	Actin gamma	
	Keratine	
	Tubulin alpha	
	Tubulin alpha-1A	
	Tubulin beta chain	
	Microtubule associated protein 6	
	Tropomyosin 1 alpha chain	
VARIOUS	SH3 domain binding protein 5	
	Hemoglobin beta	
	GDP-mannose 4,6-dehydratase	
	ATP synthase subunit alpha, mitochondrial precursor	

Supplementary Table 1: Proteins found *in vitro* associated RNA containg 60 CGG repeats. Proteins were captured on streptavidin resin coupled to biotinylated *in vitro* transcribed RNA composed of 60 CGG and identified by MALDI-TOF analysis. Next, COS7 cells were co-transfected with a plasmid expressing 60 CGG repeats and a plasmid expressing the candidate protein tagged with GFP. Co-localization with CGG RNA aggregates was assayed by FISH 24 hours after transfection.

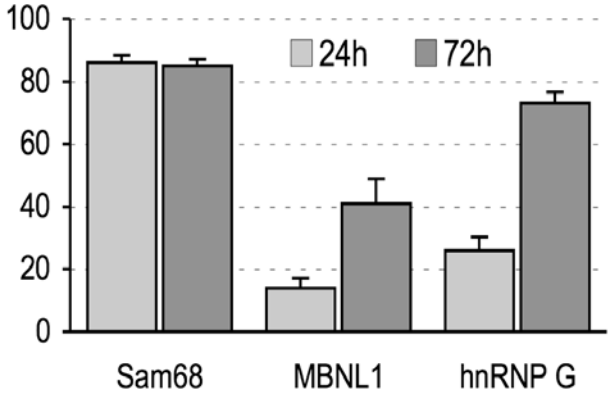
Tested proteins	Colocalization (24h)
ADAR1 (GFP)	-
ADAR2 (GFP)	-
Coilin (GFP)	-
Clk (GFP)	-
Cristalline (GFP)	-
CUGBP1 (GFP)	-
ETR3 (GFP)	-
DDX5 p68 (Ab)	-
Fox1 A2BP1 (Flag)	-
Hsp27 (GFP)	-
hnRNP A1 (GFP, Ab)	-
hnRNP A2/B1 (GFP, Ab)	-
hnRNP C (Ab)	-
HnRNP D (GFP)	-
hnRNP E (GFP)	-
HnRNP GT (GFP)	+/-
HnRNP G (GFP, Ab)	+/-
HnRNP K (GFP)	-
HnRNP I (Ab)	-
Jazz (GFP)	-
Kanadaptin (GFP)	-
mDAZL (GFP)	-
MIB2 (GFP)	-
MBNL1 (GFP, Ab)	+/-
MLN51 (YFP)	-
NF90 ILF3 (GFP)	-
NF110 (GFP)	-
Nucleophosmin (Ab)	-
PKR (Flag)	-
PP1 (GFP)	-
Puralpha (GFP, Ab)	-
QKI5 (GFP)	-

Tested proteins	Colocalization (24h)
RBMY (GFP)	-
RHA (GFP)	-
RZF (GFP)	-
Staufen1 (GFP)	-
Staufen2 (GFP)	-
SafB1 (GFP)	-
SafB2 (GFP)	-
Sam68 (GFP, Ab)	++
SMN1 (GFP)	-
SPNR (GFP)	-
SLM1 (GFP)	+/-
SLM2 TSTAR (GFP)	+/-
Sumo (GFP)	-
SUG1 (Ab)	-
TAP (GFP)	-
TDP43 (Flag)	-
Thap1 (GFP)	-
TLS (GFP)	-
Tra2β (GFP)	+
TRBP2 (GFP)	-
Ubiquitin (GFP, Ab)	-
YB1 (Ab)	-
Wig1 (GFP)	-

Supplementary Table 2: Proteins tested for co-localization within CGG RNA aggregates *in cellulo*. COS7 cells were co-transfected with a plasmid expressing 60 CGG repeats and a plasmid expressing the candidate tagged-protein followed by FISH 24 hours after transfection. We confirmed co-localization of the endogenous proteins by transfection of a plasmid expressing 60 repeats followed by FISH/ immunofluorescence 24 hours after transfection.

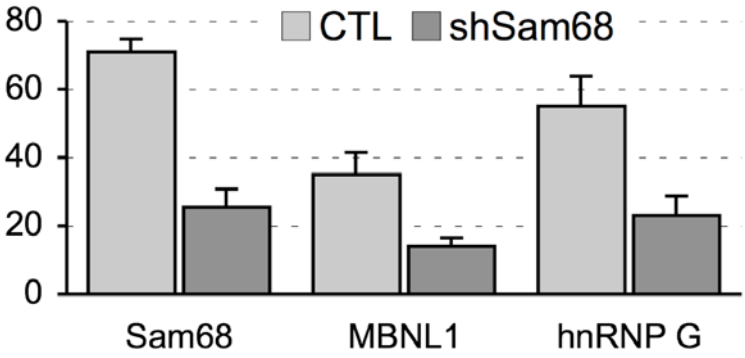
A

% of co-loc. with CGG aggregates		
	24 h	72 h
Sam68	86 +- 2.2	85 +- 1.9
MBNL1	14 +- 3.1	41 +- 7.9
hnRNP G	26 +- 4.3	73 +- 3.5



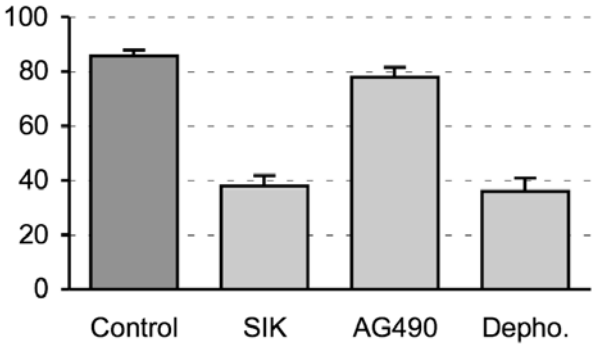
B

% of co-loc. with CGG aggregates		
	sh Control	shSam68
Sam68	72 +- 3.8	25 +- 5
MBNL1	35 +- 6.4	14 +- 2.5
hnRNP G	55 +- 8.7	23 +- 5.7



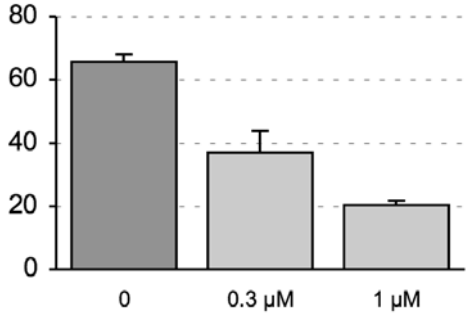
C

% of Sam68 co-localization with CGG aggregates (24h)	
Control	86 +- 2.2
SIK	38 +- 3.9
AG490	78 +- 3.4
Dephostatin	36 +- 4.9



D

% of CGG aggregates in transfected cells(24h)	
Control	66 +- 2.3
Tauto. 0.3 µM	37 +- 6.8
Tauto. 1 µM	20 +- 1.4



Supplementary Table 3:

(**A**) Percentage of endogenous Sam68, MBNL1 and hnRNP G co-localized within CGG RNA aggregates in transfected CO7 cells 24 or 72 hours after transfection. COS7 cells were co-transfected with a plasmid expressing 60 CGG repeats, co-localization with CGG RNA aggregates was assayed by FISH combined to immunofluorescence 24 or 72 hours after transfection. In all experiments: three independent transfections totalling a hundred cells were counted. Results are presented as mean +- SD.

(**B**) Percentage of endogenous Sam68, MBNL1 and of transfected GFP-hnRNP G co-localized within CGG RNA aggregates 72 hours after transfection in CO7 cells transfected with a plasmid expressing either a control shRNA (against LacZ) or a Sam68 shRNA.

(**C**) Percentage of endogenous Sam68 co-localized within CGG RNA aggregates in CO7 cells transfected with a plasmid expressing 60 CGG repeats and, either co-transfected with a plasmid expressing the SIK/BRK kinase, or treated with 10 µM of AG490 or 20 µM of dephostatin. co-localization was assayed by FISH/ immunofluorescence 24 hours after transfection.

(**D**) Percentage of cells containing CGG RNA aggregates in CO7 cells co-transfected with a plasmid expressing 60 CGG repeats and a plasmid expressing GFP. Cells were treated with 0.3 µM or 1 µM of Tautomycin. CGG aggregates formation was assayed by FISH 24 hours after transfection.

qPCR	
PO fwd	5'-GAAGTCACTGTGCCAGCCCA-3'
PO rev	5'-GAAGGTGTAATCCGTCTCCA-3'
ATP11B ex 28B fwd	5'-AATTTGCTGAAAGCCAGCAC-3'
ATP11B ex 28B rev	5'-GGGAAACAGCACATGGAGTC-3'
ATP11B ex 21 fwd	5'-GGAACAGCATGTAGACCCTCA-3'
ATP11B ex 22 rev	5'-GCCATTTCCAAGCAGAGATG-3'
SMN2 ex 7 fwd	5'-GCCTCCATTTCCTTCTGGAG-3'
SMN2 ex 7 rev	5'-ACCTTCCTTCTTTTTGATTTTGTCTAAA-3'
Plasmide CGG fwd	5'-GAACCCACTGGTTACTGGTTA-3'
Plasmide CGG rev	5'-AACGCTAGCCAGCTTGGGTC-3'
Plasmide NEO fwd	5'-GCTCTTCGTCCAGATCATCC-3'
Plasmide NEO rev	5'-TGCTCCTGCCGAGAAAGTAT-3'
PCR	
ATP11B ex 28 fwd	5'-CGACACCTCCACCCTACAAG-3'
ATP11B ex 29 rev	5'-CGTTCCAGCATTCTTCCAAC-3'
Minigene *ATP11B* fwd	5'-GCTCCGGATCGATCCTGAGAACT-3'
Minigene *ATP11B* rev	5'-GTAACCATTATAAGCTGCAA-3'
Minigene *SMN2* fwd	5′- GGTGTCCACTCCCAGTTCAA-3′
Minigene *SMN2* rev (Heinrich et al., 2009)	5′- GCCTCACCACCGTGCTGG-3'
Minigene *Bcl-x* fwd	5'-GGAGCTGGTGGTTGACTTTCT-3'
Minigene *Bcl-x* rev (Paronetto et al., 2007)	5'-TAGAAGGCACAGTCGAGG-3'

Supplementary Table 4: Sequences of the oligonucleotides used for PCR and qPCR.