
Article:
Castelli, L.M., Lin, Y.-H., Sanchez-Martinez, A. et al. (19 more authors) (2023) A cell-penetrant peptide blocking C9ORF72-repeat RNA nuclear export reduces the neurotoxic effects of dipeptide repeat proteins. Science Translational Medicine, 15 (685). ISSN 1946-6234

https://doi.org/10.1126/scitranslmed.abo3823

A cell-penetrant peptide blocking *C9ORF72*-repeat RNA nuclear export reduces the neurotoxic effects of dipeptide repeat proteins

Lydia M. Castelli[1†], Ya-Hui Lin[1†], Alvaro Sanchez-Martinez[2†], Aytaç Gül[1], Kamallia Mohd Imran[1], Adrian Higginbottom[1], Santosh Kumar Upadhyay[3], Nóra M. Márkus[1‡], Raquel Rua Martins[1], Johnathan Cooper-Knock[1], Claire Montmasson[1§], Rebecca Cohen[1‖], Amy Walton[1], Claudia S. Bauer[1], Kurt J. De Vos[1], Richard J. Mead[1], Mimoun Azzouz[1], Cyril Dominguez[3], Laura Ferraiuolo[1¶], Pamela J. Shaw[1¶], Alexander J. Whitworth[2¶], Guillaume M. Hautbergue[1¶*]

Affiliations:

[1]Sheffield Institute for Translational Neuroscience (SITraN), University of Sheffield; 385 Glossop Road, Sheffield S10 2HQ, United Kingdom.

[2]MRC Mitochondrial Biology Unit, University of Cambridge; Cambridge Biomedical Campus, Hills Road, Cambridge CB2 0XY, United Kingdom.

[3]Leicester Institute of Structural & Chemical Biology and Department of Molecular & Cell Biology, University of Leicester; Leicester LE1 7RH, United Kingdom.

[†]These authors contributed equally to this work.

‡Present address: Mission Therapeutics, The Glenn Berge Building, Babraham Research Campus, Cambridge, CB22 3FH, United Kingdom.

§Present address: INSERM U1270; Institut du Fer à Moulin, 75005 Paris, France.

‖Present address: Manchester Centre for Genomic Medicine, Manchester University NHS Foundation Trust; M13 9WL, United Kingdom.

¶Senior authors

*Corresponding author. Email: g.hautbergue@sheffield.ac.uk (G.M.H.)

Overline: NEURODEGENERATIVE DISEASE

One Sentence Summary:

Blocking nuclear export of *C9ORF72*-repeat transcripts conferred neuroprotection in patient-derived motor neurons and a fly model of ALS/FTD.

Abstract:

Hexanucleotide repeat expansions in *C9ORF72* are the most common genetic cause of familial amyotrophic lateral sclerosis (ALS) and frontotemporal dementia (FTD). Studies have shown that the hexanucleotide expansions cause the non-canonical translation of *C9ORF72* transcripts into neurotoxic dipeptide repeat proteins (DPRs) that contribute to neurodegeneration. Here, we show that a cell-penetrant peptide blocked the nuclear export of *C9ORF72*-repeat transcripts in HEK293T cells by competing with the interaction between SR-rich splicing factor 1 (SRSF1) and nuclear export factor 1 (NXF1). The cell-penetrant peptide also blocked the translation of toxic DPRs in neurons differentiated from induced neural progenitor cells (iNPC) which were derived from individuals carrying *C9ORF72*-linked ALS mutations. This peptide also increased survival of iNPC-differentiated C9ORF72-ALS motor neurons co-cultured with astrocytes. Oral administration of the cell-penetrant peptide reduced DPR translation and rescued locomotor deficits in a *Drosophila* model of mutant C9ORF72-mediated ALS/FTD. Intrathecal injection of this peptide into the brains of ALS/FTD mice carrying a *C9ORF72* mutation resulted in reduced expression of DPRs in mouse brain. These findings demonstrate that disrupting the production of DPRs in cellular and animal models of ALS/FTD might be a strategy to ameliorate neurodegeneration in these diseases.

INTRODUCTION

Polymorphic GGGGCC (G4C2) hexanucleotide repeat expansions in the *C9ORF72* gene cause the most common genetic forms of amyotrophic lateral sclerosis (ALS) and frontotemporal dementia (FTD) *(1, 2)*. These fatal adult-onset neurodegenerative disorders, which can co-manifest in the same individuals or families, lead to muscle weakness linked to progressive

paralysis and altered cognitive and personality features associated with neurodegenerative changes in the frontal and temporal lobes of the brain *(3)*. A growing body of evidence in cell and animal models suggests that toxic functions of the RNA through repeat-associated non-AUG (RAN) translation of dipeptide repeat proteins (DPRs), produced in all reading frames from sense and antisense transcripts, might be one of the main drivers of disease pathophysiology *(4)*. We recently showed that sequestration of SR-rich splicing factor 1 (SRSF1) on hexanucleotide-repeat RNA and interaction with nuclear export factor 1 (NXF1) triggers the nuclear export of sense and antisense *C9ORF72*-repeat transcripts and subsequent RAN translation of cytotoxic DPRs *(5)*. Concordantly, depletion of SRSF1 confers neuroprotection in C9ORF72-ALS patient-derived motor neurons and in a *Drosophila* model *(5)*.

No effective neuroprotective treatment is currently available for ALS or FTD. The standard of care involves treatment with riluzole, which inhibits glutamate signaling, for ALS *(6)* and serotonin reuptake inhibitors such as citalopram for FTD *(7)*. However, these treatments have limited efficacy. A Phase 1 clinical trial based on repeat intrathecal injections of an antisense-oligonucleotide (BIIB078) to downregulate the expression of pathological *C9ORF72*-repeat transcripts *(8)* did not show the expected results in C9ORF72-ALS patients. A further trial testing a stereo pure antisense-oligonucleotide with increased efficacy and selectivity is currently in progress *(9)*. However, only pathological sense transcripts are targeted via this genetic therapy approach. The development of drugs for neurodegenerative disorders is further hindered by the difficulty in producing molecules that can cross the blood brain barrier and by the lack of complete knowledge of disease-causing mechanisms. Over 50 conventional drug-based clinical trials for ALS have failed in the past 20 years *(10)*.

Peptides of biological interest fused to a protein transduction domain mediating intracellular delivery have emerged as potential therapeutic compounds in pre-clinical models of cerebral ischemia, Duchenne muscular dystrophy, cardiac disease and cancer *(11)*. We and others have also shown that these cell permeable/penetrating peptides (CPPs) form useful molecular tools to specifically target biological pathways in mammalian cells *(12, 13)*. Here, we evaluated the potential efficacy of an SRSF1-based cell permeable peptide to inhibit the interaction of SRSF1 with NXF1, resulting in prevention of the nuclear export of pathological *C9ORF72*-repeat transcripts in cell, fly and mouse models of ALS/FTD.

RESULTS

Engineering of a cell permeable peptide that blocks SRSF1/NXF1-mediated nuclear export of *C9ORF72*-repeat transcripts

We previously reported that arginines 90, 93, 117 and 118 in the unstructured amino-acid 89-120 linker sequence between the two RNA recognition motifs of SRSF1 (**Fig. S1A**) mediate the interaction with NXF1 both in pull down assays in vitro and by co-immunoprecipitation in human embryonic kidney (HEK) cells *(14)*. Pull-down assays with a recombinant GST-SRSF1 fusion protein and ^{35}S-methionine radiolabeled NXF1 proteins showed that SRSF1 aa 89-120 comprising the linker sequence is sufficient for interaction with the amino-terminal RNA-binding region of NXF1, which encompasses amino-acid 1-198 (**Fig. S1B-C**). We hypothesized that the introduction of a CPP that occupies the minimal NXF1-binding site, would compete with the endogenous SRSF1NXF1 interaction to inhibit the nuclear export of pathological *C9ORF72*-repeat transcripts and the downstream cytoplasmic RAN translation of toxic DPRs. To generate SRSF1-CPPs, wild-type SRSF1 amino-acid 89-120 WT-CPP) and the same peptide bearing

arginine 90,93,117,118 to alanine mutations (m4-CPP) were fused to a V5 tag and the protein transduction domain of the HIV1 TAT protein (amino-acid 47-57). A peptide lacking the SRSF1 sequence was also used as a negative control (Ctrl-CPP) (**Fig. 1A**). CPPs were custom synthesized and resuspended in phosphate-buffered saline (PBS). The experimental molar extinction coefficients of the cell permeable peptides were determined using serial dilutions to further allow for accurate and reproducible quantification of CPPs when new batches of peptides were prepared (**Fig. S2**).

Immunofluorescence imaging indicated that the V5-tagged CPPs were efficiently delivered into human HEK293T cells when added to the cell culture medium for 72 hours at 1 or 10 μM (**Fig. 1B, Fig. S3**). Quantification using high content automated immunofluorescence confocal microscopy further indicated that CPPs were delivered and maintained within the nucleus and cytoplasm of >95% HEK293T cells treated with 1 μM up to 10 μM CPPs (**Fig. 1C**). Furthermore, WT-CPP specifically reduced the co-immunoprecipitation of endogenous SRSF1 with FLAG-tagged NXF1, in a dose-dependent manner, interacting in turn with NXF1 in HEK293T cells (**Fig. 1D**). Addition of WT-CPP to the culture media thus disrupted the SRSF1:NXF1 interaction in human cells in vitro in agreement with our hypothesis. To precisely characterize the molecular mechanism of action, we performed pull-down assays using bacterially-expressed proteins purified in high salt and in the presence of RNase to eliminate the detection of false positive interactions bridged by RNA molecules. These assays showed that WT-CPP interacted directly with NXF1 in a dose-dependent manner (**Fig. 1E**) directly competing with the SRSF1:NXF1 interaction (**Fig. 1F**). As a specificity control, mutation of the 4 arginines involved in the interaction with NXF1, reduced the binding of m4-CPP to NXF1 by a factor of approximately 10-fold (**Fig. 1E**) and an approximate 10-fold higher concentration of

m4-CPP was thus required to disrupt the SRSF1:NXF1 interaction (**Fig. 1F**), indicating that arginines 90,93,117,118 of SRSF1 play a critical role in the binding of WT-CPP to NXF1. Isothermal titration calorimetry assays (ITC) further quantified the affinity of WT-CPP and of the minimal NXF1-interacting sequence of SRSF1 comprising aa89-120 (**Fig. S4A**) with the SRSF1/RNA-binding site of NXF1 (amino-acid 1-198**, Fig. S1B-C**). The heat released during binding to NXF1 is shown for WT-CPP and SRSF1 amino-acid 89-120 peptide in **Fig. S4B-C**. Transformation of the thermodynamic parameters acquired from the heat-released titration experiments into protein-ligand binding titration curves further indicated that the affinity of SRSF1 WT-CPP compared to the SRSF1 amino-acid 89-120 peptide was increased by 5-fold (**Table S1**). This is likely due to the presence of additional arginine residues in the TAT protein transduction domain of WT-CPP, advantageously conferring an increased potency for disrupting the SRSF1:NXF1 interaction. We concluded that binding of WT-CPP to NXF1 involved arginines 90,93,117,118 within the SRSF1 amino-acid 89-120 sequence because mutations of these in m4-CPP led to approximately a 10-fold reduction in binding NXF1 (Fig. 1E) and inhibition of the SRSF1:NXF1 interaction (Fig. 1F), whereas arginines in the TAT sequence were likely to enhance the binding of SRSF1-CPP compared to the SRSF1 amino-acid 89-120 sequence alone because an SRSF1 amino-acid 89-120 peptide fused to the V5-TAT domain had higher affinity for NXF1 (Table S1).

In addition, we also demonstrated that WT-CPP competed for the interaction between full length mammalian SRSF1 synthesized in rabbit reticulocytes and recombinant NXF1 (**Fig. S5**). Taken together, these data demonstrated that SRSF1-CPPs are delivered intracellularly and inhibit the SRSF1:NXF1 interaction in human cells through direct competition.

SRSF1 WT-CPP-mediated inhibition of *C9ORF72*-repeat RNA nuclear export is cytoprotective in HEK293T cells

We next tested whether WT-CPP would inhibit the SRSF1:NXF1-dependent nuclear export and RAN translation of *C9ORF72*-repeat transcripts in human cell models of C9ORF72-ALS/FTD. To facilitate the quantification of RAN-translated DPRs, we generated uninterrupted repeat constructs lacking canonical AUG start codons and comprising 45 sense G4C2 or 43 antisense G2C4 hexanucleotide repeats, with downstream V5 tags in the three coding frames to allow detection of all DPR species in a single western blot assay. Plasmid maps of the engineered sense and antisense reporter plasmids are presented in **Figs. S6** and **S7, respectively**. Addition of increasing concentrations of WT-CPP to the medium of HEK293T cells transfected with $G4C2_{45}$-3xV5 or $G2C4_{43}$-3xV5 repeat transcript constructs led to dose-dependent inhibition of nuclear export with nuclear increase associated with concomitant cytoplasmic decrease of *C9ORF72*-repeat transcripts, whereas the Ctrl-CPP had no effect (**Fig. 2A-B**). Western blotting analysis of total, nuclear and cytoplasmic fractions confirmed the absence of detectable nuclear contamination in the cytoplasmic samples probed with the chromatin-associated SSRP1 factor and the predominantly cytoplasmic HSPA14 heat shock protein (**Fig. S8A-B**). The WT-CPP-induced nuclear export inhibition of *C9ORF72*-repeat transcripts correlated with a dose-dependent reduction of the RAN-translation of V5-tagged sense and antisense DPRs with an IC_{50} concentration of approximately 0.5 μM (Fig. 2C-D, quantification in Fig. 2G-H). m4-CPP also inhibited the production of sense and antisense DPRs, but to a lesser extent, in agreement with its reduced affinity for NXF1 (Fig. 2C-F). Moreover, we validated that individual expression of sense poly-GA, GP, GR and antisense poly-GP, PR, PA DPRs was similarly reduced under the same experimental conditions (**Fig. 2G-H**). The dose-dependent inhibition of V5-tagged DPRs

further correlated with increased cell proliferation as measured in MTT assays undertaken with the same conditions as in Fig. 2A-H, with 0.5 μM WT-CPP and 1 μM m4-CPP respectively, rescuing the DPR-induced cytotoxicity in the C9ORF72-ALS/FTD cell models (**Fig. 2I-J**). Taken together, our data show that SRSF1-CPPs efficiently inhibited the nuclear export and subsequent RAN translation of both sense and antisense reporter *C9ORF72*-repeat transcripts in human HEK293T cell models of ALS/FTD.

The regulatory splicing and translational functions of SRSF1 are preserved upon treatment with SRSF1 WT-CPP

In addition to its redundant/cooperative role in the mRNA nuclear export pathway *(14-16)*, SRSF1 serves additional nuclear and cytoplasmic functions including splicing *(17, 18)* and translation regulation *(19-21)*. Expression of total, nuclear and cytoplasmic SRSF1 is not altered in HEK293T cells treated with CPPs, suggesting that these other cellular functions of SRSF1 are maintained (**Fig. 3A**). To investigate this further, we transfected HEK293T cells with either Ctrl-RNAi, SRSF1-RNAi, T7-tagged SRSF1 WT or T7-tagged SRSF1-NRS, a hyperphosphorylated mutant which harbors a nuclear retention signal (NRS) and is constitutively retained in the nucleus *(20, 22, 23)*. Depletion of SRSF1 or treatment with CPPs did not induce cytotoxicity, however, expression of the SRSF1-NRS mutant specifically affected cell proliferation (**Fig. 3B**). Depletion of SRSF1 and hyperphosphorylation of SRSF1-NRS was validated by western blot analysis under the same conditions (**Fig. 3C**). The phosphorylation of SRSF1 is linked to co-transcriptional splicing, whereas dephosphorylation allows subsequent interactions with NXF1 for the nuclear export of mRNAs *(24)*. We thus investigated the SRSF1-dependent alternative splicing of *CLK4* exons 4-5 *(20)* upon depletion of SRSF1, supplementation with CPPs or

expression of phospho-SRSF1-NRS. As expected for a dominant mutant with an enforced link to splicing, SRSF1-NRS led to alteration of the alternative splicing of the *CLK4* transcript, whereas depletion of SRSF1 or treatment with CPPs had no effect on the alternative splicing of *CLK4* (**Fig. 3D**).

In addition, we showed that neither Ctrl-CPP or WT-CPP affected global protein synthesis using pulse puromycin-labeling of translating polypeptide chains prior to western blotting with an anti-puromycin antibody (**Fig. 4A**; quantification in **Fig. 4B**). Moreover, partial depletion of SRSF1 or treatment with CPPs did not alter the translation of reporter poly-GR, GP and GA DPRs expressed independently of the G4C2-repeat sequence and RAN translation context (**Figs. 4C-E**). This was in contrast to the control SRSF1-RNAi or WT-CPP mediated inhibition of the nuclear export of G4C2-repeat transcripts, which led to subsequent inhibition of the cytoplasmic production of DPRs (**Fig. 4F**). Expression of DPRs was further quantified in **Figs. 4G-J** respectively. We concluded that CPPs did not affect global protein synthesis or the translation of reporter DPRs.

Data presented in figures 2, 3 and 4 demonstrated that SRSF1-WT CPP specifically inhibited the nuclear export of *C9ORF72*-repeat transcripts and the subsequent RAN translation of DPRs in human reporter cell lines, whereas it had no effect on splicing or translation.

SRSF1-CPP inhibits repeat-associated non-AUG translation in rat primary neurons

To investigate the relevance of these findings in neurons, primary cultures of rat cortical neurons were treated with 1, 5 and 10 μM CPPs for 72 hours and compared with untreated controls. Automated V5-immunofluorescence microscopy using the *Opera Phenix* high content imaging

system showed intracellular delivery of SRSF1 WT-CPP with increased V5 signal in a dose-dependent manner at 5 and 10 μM in magnified confocal planes of 1-2 neurons (**Fig. 5A**). Quantification further indicated that the CPPs were also efficiently delivered and maintained within >90% primary cortical neurons when added to the medium for 72 hours at 5-10 μM (**Fig. 5B**). Higher concentrations of CPPs appeared to be required for efficient delivery in neurons compared to HEK293T cells (Fig. 1C) potentially due to the presence of myelin sheaths which restrict uptake. We next generated lentivirus expressing the $G4C2_{45}$-3xV5 or $G2C4_{43}$-3xV5 repeat transcripts and transduced cortical neuron cultures with these constructs for 16 hours prior to replacing the media with fresh medium supplemented with either Ctrl-CPP or SRSF1-CPPs. Western blot analysis of V5-tagged DPRs conducted 72 hours post CPP addition indicated that WT-CPP, and m4-CPP to a lesser extent, inhibited the production of sense and antisense DPRs with a pronounced effect at 5 μM (**Fig. 5C**, V5-DPRs panel). Higher concentrations of CPPs led to increased intracellular delivery of V5-tagged CPPs in the neuronal extracts (Fig. 5C, V5-CPPs panel). Lentiviral-mediated depletion of SRSF1 was further used as a positive control *(5)* showing that both depletion of SRSF1 or addition of 5 μM WT-CPP similarly inhibited the expression of sense and antisense DPRs (Fig. 5C, V5-DPRs panel). Immunofluorescence microscopy staining of DPRs in rat cortical neurons cultured under the same conditions but treated with 5 μM of Ctrl-CPP or WT-CPP custom synthesized without the V5 tag (to allow for detection of the V5-tagged DPRs) validated the results from the previous western blot analysis, showing that WT-CPP treatment led to a reduction in both the intensity and the number of neurons with a V5-DPR signal (**Fig. S9A**). Quantification further showed significant reduction in both sense and antisense V5-DPR positive neurons treated with either lentivirus-SRSF1-RNAi or WT-CPP (**Fig. S9B**, $p<0.05$ or $p<0.0001$), despite underestimation of the SRSF1 depletion or

WT-CPP effects, given that neurons with successful SRSF1-targeted reduction of DPRs remain counted as DPR-positive cells. Overall, these data show proof-of-concept for intracellular delivery and functionality of the SRSF1 WT-CPP in a primary neuron model of C9ORF72-ALS/FTD.

SRSF1-CPP-mediated inhibition of *C9ORF72*-repeat RNA nuclear export confers neuroprotection in C9ORF72-ALS/FTD patient-derived motor neurons

The neuroprotective potential of SRSF1 WT-CPP was further tested in C9ORF72-ALS patient-derived motor neurons that harbored hexanucleotide repeat expansions in intron-1 of the *C9ORF72* gene. Fibroblasts from healthy control individuals and ALS patients carrying C9ORF72 mutations (**Table S2**) were reprogrammed into induced neural progenitor (iNPC) cells that were further differentiated into induced astrocytes, neurons or motor neurons (iAstrocytes, iNeurons or iMNs respectively) *(5, 25, 26)*. It was not practically feasible to differentiate enough iMNs from iNeuron cultures for experiments involving quantification of DPRs or subcellular fractionation to measure the nuclear export of pathological *C9ORF7*2-repeat transcripts. These assays were therefore conducted in control and C9ORF72-ALS patient-derived iNeurons, which can yield several million cells with a homogeneous population of mature neurons staining positive for the markers TUJ1/ MAP2 (90%) and NeuN (>85 %) (**Fig. 6A-B**). Moreover, we have recently characterized the transcriptomes of these iNeuron lines showing that the differentiation protocol principally drives cells to the neuronal lineage, with very limited contamination from astrocyte or oligodendrocyte markers *(27)*. The C9ORF72-ALS iNeurons also exhibited TDP-43 aggregation/proteinopathy, one of the molecular hallmarks of most cases of ALS including C9ORF72-ALS. TDP-43 aggregates were detected using a TDP-43 antibody

directed against phosphorylation of Serine 409,which recognizes the major phosphorylated and ubiquitinated form of TDP-43 in cytoplasmic inclusions (**Fig. 6C**). Ctrl-CPP and WT-CPP were next added to the cultures of iNeurons at 10 μM for 72h. We showed that, similar to experiments in HEK293T cells, the SRSF1-dependent alternative splicing of *CLK4* exons 4-5 was not affected upon WT-CPP treatment of healthy control iNeurons (**Fig. 6D**). Total nuclear and cytoplasmic fractions were prepared from iNeurons derived from either three healthy control individuals or three patients with C9ORF72-ALS. The quality of the subcellular fractionation was shown by western blots using antibodies directed against the nuclear chromatin-associated SSRP1 factor and the cytoplasmic pan neuronal TUJ1 marker, showing absence of detectable nuclear and cytoplasmic contamination in the cytoplasmic and nuclear fractions respectively (**Fig. S10**). We next quantified the nuclear export of pathological *C9ORF72*-repeat pre-mRNAs retaining G4C2 expansions in intron-1 and of wild type intron-1-spliced transcripts using qRT-PCR assays with a forward primer annealing in exon-1 and a reverse oligonucleotide annealing either in intron-1 or exon-2 respectively. Similar to the lentiviral-mediated depletion of SRSF1 *(5)*, a 72-hour addition of 10 μM WT-CPP led to selective and efficient nuclear export inhibition of *C9ORF72*-repeat transcripts in patient-derived iNeurons showing increased nuclear and concomitant cytoplasmic reduction (**Fig. 6E**), whereas there was no effect on the expression of wild-type *C9ORF72* transcripts (**Fig. 6F**). Consistent with the SRSF1-targeted nuclear export inhibition of pathological repeat transcripts, mesoscale discovery (MSD) based ELISA assays showed that both the lentiviral-mediated depletion of SRSF1 *(5)* and the addition of SRSF1 WT-CPP were associated with a significant reduction in the cytoplasmic RAN translation of poly-GP DPRs encoded by sense and antisense *C9ORF72*-repeat transcripts under the same iNeuron cultures and conditions (**Fig. 6G**, $p<0.01$, standard curve **Fig. S11**). The survival of C9ORF72-

ALS-derived iNeurons or iMNs was not affected compared to neurons differentiated from healthy control individuals, however, co-cultures with C9ORF72-ALS patient-derived iAstrocytes showed reduced survival of iMNs in vitro *(5, 25)*. Approximately 70% of differentiated iMNs were positive for the spinal non-phosphorylated neurofilament marker SMI-32 (**Fig. 6H**). Co-culture experiments further showed that a 72-hour medium supplementation of 10 μM SRSF1 WT-CPP promoted the survival of the same lines of C9ORF72-ALS iMNs in co-cultures with C9ORF72-ALS iAstrocytes whilst it had no detrimental cytotoxic effects on healthy control motor neurons, similar to the depletion of SRSF1 which was mediated via lentiviral expression of SRSF1-miRNA in both iAstrocytes and iMNs (**Fig. 6I**; quantification in **Fig. 6J**). Taken together, these data show that SRSF1 WT-CPP inhibited the nuclear export of pathological *C9ORF72*-repeat transcripts and the subsequent RAN translation of DPRs, conferring neuroprotection in vitro for C9ORF72-ALS patient-derived co-cultures of astrocytes and motor neurons.

SRSF1-CPP-mediated inhibition of C9ORF72-repeat RNA nuclear export confers neuroprotection in a *Drosophila* model

We next sought to test the therapeutic potential of SRSF1 WT-CPP in vivo in a *Drosophila* model of disease *(28)* as well as in a mouse model of C9ORF72-ALS/FTD *(29, 30)*. The NXF1 family of proteins is conserved across species. Relevant to this study, patches of negatively-charged acidic residues, which can mediate interactions with the positively-charged arginine moieties of the SRSF1 WT-CPP are located in the amino-terminal RNA/SRSF1-binding sites (**Fig. S12**).

Wild type G4C2x3 and G4C2x36 *Drosophila*, which expresses 36 uninterrupted G4C2 repeat transcripts and sense DPRs *(28)*, were fed with food supplemented with 10 μM Ctrl-CPP, WT-CPP and m4-CPP (**Fig. 7A**). Whole-fly, nuclear and cytoplasmic fractions prepared from untreated or CPP-treated animals showed absence of detectable nuclear Histone H3 contamination in the cytoplasmic fractions and cytoplasmic tubulin in the nuclear samples (**Fig. 7B**), thus qualitatively validating the sub-cellular fractionation process. The expression of histone H3 or tubulin was not directly comparable from one *Drosophila* CPP-treated condition to another as only equivalent proportions, rather than equal amounts of total, nuclear or cytoplasmic proteins, were loaded to check the quality of the fractionation. Nevertheless, we demonstrated that SRSF1 WT-CPP did not affect the expression of cellular proteins using total protein stain analysis and western blot analysis from gels loaded with the same amounts of proteins (**Fig. S13**). In contrast, the oral administration of WT-CPP led to a specific reduction of cytoplasmic G4C2x36 repeat transcripts normalized to total G4C2-repeat and control loading Tub84b transcripts, whereas m4-CPP showed a less pronounced effect (**Fig. 7C**), indicating that similar to the cell models, WT-CPP inhibits the nuclear export of *C9ORF72*-repeat transcripts in a *Drosophila* model of disease. This correlated with a reduction in the neurodegeneration-associated motor deficits observed in C9ORF72-ALS fly larvae (**Fig. 7D**) and adult flies (**Fig. 7E**) as well as an inhibition of the RAN translation of poly-GA, GP, GR DPRs (**Fig. 7F**). Moreover, the cell permeable peptides had no detrimental effects on the behavior of healthy control flies bearing 3 G4C2 repeats (Figs 7D-E), indicating that the CPPs were safe and well-tolerated in vivo in *Drosophila*. We were unable to validate intracellular delivery of the CPPs in *Drosophila* brains using V5 immunofluorescence microscopy, likely due to very small amounts of CPP uptake and circulation via the oral administration route. However, using dot blots, we

were able to confirm that Ctrl-CPP and WT-CPP were delivered to the heads of adult flies treated with 60 μM CPPs (**Fig. 7G**). Taken together, our data demonstrate that oral administration of SRSF1 WT-CPP led to a reduction of cytoplasmic G4C2x36 transcripts leading to an inhibition of RAN translation of DPRs and mitigated the DPR-associated locomotor deficits in a C9ORF72-ALS *Drosophila* model of disease.

We next tested the proof-of-principle in mammals through single intrathecal injections of Ctrl-CPP and WT-CPP into the cerebrospinal fluid of BAC-transgenic C9ORF72-ALS/FTD mice expressing approximately 500 pathological G4C2 repeats from a human *C9ORF72*-linked ALS gene (C9-500 Tg) *(29, 30)*. In compliance with the 3R principles (Replacement, Reduction and Refinement) regulating our UK Home Office licence for animal experimentation, we performed our pilot study in 1-month old (P30) asymptomatic mice which readily express DPRs, the major readout for therapeutic target engagement. Mice were injected via the *cisterna magna* with 5 μl of Ctrl-CPP or SRSF1 WT-CPP at 1 mM (corresponding to ~5 μM brain exposure) and were sacrificed 4 days later. Using immunofluorescence microscopy, we confirmed that CPPs were delivered into the cerebellum and to a less extent to the brain stem (**Fig. 8A**). Compared with untreated or Ctrl-CPP-treated C9-500 Tg mice, injections of WT-CPP led to a 70% reduction in the expression of poly-GP DPRs quantified by MSD-ELISA in cerebellum/brain stem homogenates 96 hours post-delivery (**Fig. 8B**, standard curve **Fig. S11**). No adverse events were observed and no weight loss was noted in the brain and peripheral organs of the CPP-treated animals (**Fig. 8C**). Taken together, our animal data show promising therapeutic potential of SRSF1 WT-CPP in vivo, demonstrating proof-of-principle of translatability in cell models, *Drosophila* and a mammalian model of C9ORF72-ALS/FTD.

SRSF1 WT-CPP specifically targets NXF1 in human HEK293T cells, patient-derived neurons and mouse brains

NXF1 belongs to a family of proteins which includes NXF2 and NXF3 (**Fig. S14**), two other related members expressed in a tissue-specific manner, principally in the testis. To investigate potential off-target effects of SRSF1 WT-CPP on NXF2 and NXF3, we performed pull down assays using recombinant GST-fusion proteins and a range of CPP concentrations varying from 0.125 up to 8 μM. As previously described in Fig. 1E, WT-CPP directly and robustly interacted with GST-NXF1:p15 (**Fig. S15A**), whereas it did not or very poorly interacted with GST-NXF2:p15 (**Fig. S15B**) or GST-NXF3 (**Fig. S15C**). We further wanted to investigate the expression of the NXF1-3 proteins in our human cell and mouse models. However, western blots with the purified recombinant GST-NXF1-3 proteins indicated that the commercially-available NXF2 and NXF3 antibodies present specificity issues (**Fig. S15D**). We thus performed qRT-PCR analysis to quantify the expression of *NXF1*, *NXF2* and *NXF3* mRNAs in HEK293T cells, patient-derived neurons, mouse testis and mouse brain. Although, as expected, all NXF1-3 mRNAs were expressed in the testis, we show here that NXF2 and NXF3 mRNAs were not expressed in the human cell models or mouse brain tissue, confirming that NXF1 is the target of WT-CPP (**Fig. S15E**).

DISCUSSION

In this study, we report targeting the nuclear export of pathological *C9ORF72*-repeat transcripts through treatment with a cell-penetrant peptide. This cell-penetrant peptide inhibited the SRSF1-bound repeat RNA:NXF1 interaction and the subsequent RAN translation of neurotoxic DPRs to confer neuroprotection in preclinical patient-derived neurons. It also ameliorated DPR-associated

motor deficits in a *Drosophila* model of ALS/FTD. In a proof-of-concept study, intrathecal injection of SRSF1-CPP into C9ORF72-ALS/FTD mouse brains aimed to validate target engagement and inhibition of DPRs. Moreover, no short-term toxicity effects of the CPPs were observed in control healthy or patient-derived motor neurons/astrocytes, or *Drosophila* or mice. This study also validated SRSF1 as a therapeutic target in C9ORF72-ALS/FTD, in agreement with our previous SRSF1-RNAi mediated gene therapy approach *(5)*. Our current study further reinforces the concept that the nuclear export of mRNAs *(31, 32)* can be manipulated therapeutically in neurodegenerative disease models. We also recently reported that targeting the SRSF1-dependent nuclear export function does not affect the global transcriptome or mRNA nuclear export pathway *(27)*.

The reduction in the expression of DPRs conferred by oral administration of SRSF1 WT-CPP indicated efficient brain delivery in the C9ORF72-ALS/FTD *Drosophila* model. Previous studies have also reported that TAT-based CPPs are able to cross the blood-brain barrier and be delivered to the central nervous system in mice *(33, 34)*. Oral administration of a poly-Q binding peptide 1 (QBP1)-based CPP, which interacts with polyglutamine stretches and inhibits the aggregation of poly-Q proteins, suppressed neurodegeneration-associated premature death and inclusion body formation in a poly-Q *Drosophila* model *(35)*, while improving the poly-Q phenotype of a mouse model through intra-cerebroventricular injection *(36)*. Similarly, a cell-permeable peptide bearing an RNA-binding motif of nucleolin, which binds poly-Q-encoded CAG-repeat RNA and sense *C9ORF72*-repeat reporter transcripts, rescued the rough eye neurodegenerative phenotype in a *Drosophila* model of disease *(37, 38)*. Our study further demonstrates successful intracellular delivery and efficacy of the SRSF1 CPP in rat primary cortical neurons and C9ORF72-ALS patient-derived neurons and motor neurons. We also

observed that SRSF1 WT-CPP is functional in mouse brains for at least 4 days following intrathecal delivery, providing a 70% reduction in the expression of poly-GP DPRs and a proof-of-principle of efficacy in mammals.

Limitations of the current study include the lack of a full therapeutic evaluation of SRSF1 WT-CPP in C9ORF72-ALS/FTD mice. Rapid turnover of peptides typically constitutes one of the main challenges associated with their therapeutic efficacy. This may require future chemical modifications of WT-CPP to increase its in vivo stability. Biodistribution in the brain and spinal cord is another key determinant which will require optimization and validation. The determination of the pharmacokinetic and pharmacodynamic parameters of WT-CPP will be essential to define the optimal repeat dosing interval and dose that would allow evaluating the long-term therapeutic efficacy and safety of WT-CPP in C9ORF72-ALS/FTD mice.

Intranasal delivery of SRSF1-CPP would potentially constitute an attractive route of administration in repeat-dosing studies. Relevant to this, it was reported that daily intranasal co-injections of L-penetratin, a cell-permeable peptide and insulin or extendin-4, a glucagon-like peptide-1 (GLP-1) receptor agonist, ameliorate the progressive cognitive dysfunction phenotype in a senescence-accelerated mouse prone-8 (SAMP8) model *(39, 40)*. Taken together, our findings highlight promising prospects for the development of CPP as an alternative to conventional small molecule inhibitors that frequently have poor blood-brain barrier penetration. Moreover, the development of CPPs for disrupting the nuclear export of other microsatellite repeat transcripts, which are involved in over 40 neurological disorders *(41)*, could provide future potential agents for dissecting the nature of the pathogenic mechanisms linked to the formation of intranuclear RNA foci and RAN translation of polymeric repeat proteins.

MATERIALS AND METHODS

Study design

The overall aim of this study was to engineer a cell penetrant peptide, which would competitively inhibit the interaction of SRSF1-bound *C9ORF72*-repeat RNAs with the nuclear export receptor NXF1, in order to reduce the nuclear export of pathological *C9ORF7*2-repeat transcripts and their subsequent RAN translation into neurotoxic DPRs. This multidisciplinary study used purified recombinant proteins, custom-synthesized RNA in pull down and isothermal calorimetry assays. Custom-synthesized peptides were used in in vitro biochemical assays, and in vitro and in vivo models of C9ORF72-ALS/FTD including reporter human HEK293T cells, primary rat neurons, patient derived neurons and motor neurons, *Drosophila* and mice. Intracellular delivery of peptides and motor neuron survival were quantified by high-content immunofluorescence microscopy automated imaging. Co-immunoprecipitation assays were used to evaluate the effects on CPPs on the interaction of SRSF1 with NXF1 in HEK293T cells. Nuclear export of *C9ORF72*-repeat transcripts was quantified following qRT-PCR analysis of total, nuclear and cytoplasmic fractions. Western blotting was used to investigate protein expression levels. MTT colorimetric assays were used to quantify cell proliferation. In-house MSD-ELISA were used to quantify poly-GP DPRs. Locomotor function was assessed in *Drosophila* using crawling assays in larvae and climbing indexes in adult flies. No study size calculations or randomization were carried out. All samples were quantified in a blinded manner for microscopy and mouse experiments. No cell or animal samples were excluded. Several operators were involved in performing and/or cross-validating replicate experiments for some qRT–PCR, western blot, microscopy and MSD-ELISA data (Fig. 1C; 2A-B; 5B-C; 6D-G; 6J; 7F-G; 8B; S9).

Mouse studies were carried out in accordance with the UK Animals (Scientific Procedures) Act 1986 under a UK Home Office project license number P608B5AD3. The license underwent review and approval by the University of Sheffield Animal Welfare and Ethical Review Body. Mouse maintenance and day to day care was conducted according to the Home Office Code of Practice for Housing and Care of Animals Used in Scientific Procedures. The ARRIVE 2.0 guidelines were followed in conducting in vivo experiments *(42)*.

Informed consent was obtained from all subjects before collection of fibroblast biopsies (Study number STH16573, Research Ethics Committee reference 12/YH/0330).

Differentiation of patient-derived neurons, motor neurons and astrocytes

Human patient derived iAstrocytes, iNeurons and iMotor Neurons were differentiated from induced Neural Progenitor Cells (iNPCs) as described previously *(5, 25)*.

Co-cultures of patient-derived astrocytes and motor neurons

Differentiation of iMotor Neurons (iMNs). Human patient and control-derived neurons (iNeurons) were differentiated from induced neural progenitor cells (iNPCs) using a modified version of protocol *(25)* as previously described *(5)*. In brief, 100,000 iNPCs were plated in a 6-well plate coated with fibronectin (Millipore) and expanded to 70-80% confluence. Once they reached this confluence, iNPC medium was replaced with neuron differentiation medium (DMEM/F-12 with glutamax supplemented with 1% N2, 2% B27 (Gibco) containing 2.5 μM of DAPT (Tocris) to determine differentiation towards neuronal lineage on day 1. On day 3, the neuron differentiation medium was supplemented with 1 μM retinoic acid (Sigma), 0.5 μM smoothened agonist (SAG) (Millipore) and 2.5 μM forskolin (Sigma) for 7 days until Day 10. This protocol leads to typical yields of 70% β-III tubulin (Tuj1) positive cells. To obtain iMotor

Neurons (iMN), ~ 5,000 iNeurons per well were re-plated on 96-well plates coated with fibronectin and maintained in iNeuron differentiation medium (containing retinoic acid, SAG and forskolin) supplemented with BDNF, CNTF and GDNF (all at 20 ng/ml) for the last 14 days of differentiation.

Differentiation of iAstrocytes. Human patient-derived astrocytes (iAstrocytes) were differentiated from iNPCs as previously described (*5, 14*) and cultured in DMEM glutamax (*Gibco*) with 10% FBS (*Sigma*) and 0.02% N2 (*Invitrogen*) for 5 days. Cells were maintained in a 37˚C incubator with 5% CO_2.

Co-cultures of patient-derived iMNs and iAstrocytes. iAstrocytes were lifted at day 5 of differentiation and ~5,000 iAstrocytes were re-plated on iMNs at day 20 of differentiation. Co-cultured iMNs and iAstrocytes were maintained in neuron differentiation medium with BDNF, GDNF and CTNF (all at 20 ng/ml) for 4 days. 12 h after the start of co-cultures (on day 21), 1 or 10 μM CPP was added to the medium and iMNs/ iAstrocytes were imaged for 72 h at days 22, 23, 24. For SRSF1 knockdown, iMNs and iAstrocytes were separately transduced 48h prior to co-culture with lentivirus (LV) expressing control or SRSF1-RNAi co-expressing GFP *(5)* at an MOI of 5 at day 18 of iMN differentiation and at day 3 of iAstrocyte differentiation (diagrammatic representation in Fig. S16).

Mice

BAC transgenic C9ORF72-ALS/FTD mice expressing approximately 500 G4C2 repeats from a a human *C9ORF72*-linked ALS gene (FVB/NJ-Tg(C9orf72)500Lpwr/J) were created by Prof Laura Ranum's group *(29, 30)* and obtained from the Jackson Laboratory (Stock No: 029099, also known

as C9-500). Mice were housed in groups of 2 to 5 with access to food and water *ad libitum* (Envigo, standard rodent diet 2018). The temperature was maintained at 21°C with a 12-hour light/dark cycle. A plastic house was provided in each cage with sawdust (Datesand) to cover the floor of the cages and nesting material provided (Datesand paper wool).

Statistics

For all normally-distributed data involving more than 2 experimental groups, one-way and two-way ANOVA (analysis of variance) with correction for the recommended multiple comparison tests (*GraphPad Prism*) were used. Immunofluorescence microscopy analysis of DPR-positive neurons used Fisher's exact test to compare 2 groups. The *Drosophila* climbing assays are not normally-distributed and were analyzed using a non-parametric Kruskal-Wallis test with Dunn's correction for multiple comparisons. No randomization was used in the animal studies as small numbers of animals were used. Data were plotted using GraphPad Prism versions 8-9. Significance is indicated as follows; NS: non-significant, $p \geq 0.05$; * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$; **** $p < 0.0001$. DPR-positive neuron counts, DPR quantification and weight in mouse data, *Drosophila* crawling and climbing assays were analyzed in a blinded manner.

Supplementary Materials

Supplementary Materials and Methods

Fig. S1 to S16

Table S1 to Table S5

Data Files S1

References (*43*-*48*)

MDAR

References

1. M. DeJesus-Hernandez, I. R. Mackenzie, B. F. Boeve, A. L. Boxer, M. Baker, N. J. Rutherford, A. M. Nicholson, N. A. Finch, H. Flynn, J. Adamson, N. Kouri, A. Wojtas, P. Sengdy, G.-Y. R. Hsiung, A. Karydas, W. W. Seeley, K. A. Josephs, G. Coppola, D. H. Geschwind, Z. K. Wszolek, H. Feldman, D. S. Knopman, R. C. Petersen, B. L. Miller, D. W. Dickson, K. B. Boylan, N. R. Graff-Radford, R. Rademakers, Expanded GGGGCC Hexanucleotide Repeat in Noncoding Region of C9ORF72 Causes Chromosome 9p-Linked FTD and ALS, *Neuron* **72**, 245–256 (2011).

2. A. E. Renton, E. Majounie, A. Waite, J. Simón-Sánchez, S. Rollinson, J. R. Gibbs, J. C. Schymick, H. Laaksovirta, J. C. van Swieten, L. Myllykangas, H. Kalimo, A. Paetau, Y. Abramzon, A. M. Remes, A. Kaganovich, S. W. Scholz, J. Duckworth, J. Ding, D. W. Harmer, D. G. Hernandez, J. O. Johnson, K. Mok, M. Ryten, D. Trabzuni, R. J. Guerreiro, R. W. Orrell, J. Neal, A. Murray, J. Pearson, I. E. Jansen, D. Sondervan, H. Seelaar, D. Blake, K. Young, N. Halliwell, J. B. Callister, G. Toulson, A. Richardson, A. Gerhard, J. Snowden, D. Mann, D. Neary, M. A. Nalls, T. Peuralinna, L. Jansson, V.-M. Isoviita, A.-L. Kaivorinne, M. Hölttä-Vuori, E. Ikonen, R. Sulkava, M. Benatar, J. Wuu, A. Chio, G. Restagno, G. Borghero, M. Sabatelli, ITALSGEN Consortium, D. Heckerman, E.

Rogaeva, L. Zinman, J. D. Rothstein, M. Sendtner, C. Drepper, E. E. Eichler, C. Alkan, Z. Abdullaev, S. D. Pack, A. Dutra, E. Pak, J. Hardy, A. Singleton, N. M. Williams, P. Heutink, S. Pickering-Brown, H. R. Morris, P. J. Tienari, B. J. Traynor, A hexanucleotide repeat expansion in C9ORF72 is the cause of chromosome 9p21-linked ALS-FTD, *Neuron* **72**, 257–268 (2011).

3. J. R. Burrell, G. M. Halliday, J. J. Kril, L. M. Ittner, J. Götz, M. C. Kiernan, J. R. Hodges, The frontotemporal dementia-motor neuron disease continuum, *Lancet* **388**, 919–931 (2016).

4. R. Balendra, A. M. Isaacs, C9orf72-mediated ALS and FTD: multiple pathways to disease, *Nature Reviews Neurology* **14**, 544–558 (2018).

5. G. M. Hautbergue, L. M. Castelli, L. Ferraiuolo, A. Sanchez-Martinez, J. Cooper-Knock, A. Higginbottom, Y.-H. Lin, C. S. Bauer, J. E. Dodd, M. A. Myszczynska, S. M. Alam, P. Garneret, J. S. Chandran, E. Karyka, M. J. Stopford, E. F. Smith, J. Kirby, K. Meyer, B. K. Kaspar, A. M. Isaacs, S. F. El-Khamisy, K. J. De Vos, K. Ning, M. Azzouz, A. J. Whitworth, P. J. Shaw, SRSF1-dependent nuclear export inhibition of C9ORF72 repeat transcripts prevents neurodegeneration and associated motor deficits, *Nat Commun* **8**, 16063 (2017).

6. G. Bensimon, L. Lacomblez, V. Meininger, A controlled trial of riluzole in amyotrophic lateral sclerosis. ALS/Riluzole Study Group, *N Engl J Med* **330**, 585–591 (1994).

7. M. F. Mendez, Frontotemporal dementia: therapeutic interventions, *Front Neurol Neurosci* **24**, 168–178 (2009).

8. K. M. Schoch, T. M. Miller, Antisense Oligonucleotides: Translation from Mouse Models to Human Neurodegenerative Diseases, *Neuron* **94**, 1056–1070 (2017).

9. Y. Liu, J.-C. Dodart, H. Tran, S. Berkovitch, M. Braun, M. Byrne, A. F. Durbin, X. S. Hu, N. Iwamoto, H. G. Jang, P. Kandasamy, F. Liu, K. Longo, J. Ruschel, J. Shelke, H. Yang, Y. Yin, A. Donner, Z. Zhong, C. Vargeese, R. H. Brown, Variant-selective stereopure oligonucleotides protect against pathologies associated with C9orf72-repeat expansion in preclinical models, *Nat Commun* **12**, 847–15 (2021).

10. D. Petrov, C. Mansfield, A. Moussy, O. Hermine, ALS Clinical Trials Review: 20 Years of Failure. Are We Any Closer to Registering a New Treatment? *Front Aging Neurosci* **9**, 68 (2017).

11. G. Guidotti, L. Brambilla, D. Rossi, Cell-Penetrating Peptides: From Basic Research to Clinics, *Trends Pharmacol. Sci.* **38**, 406–424 (2017).

12. I. E. Gallouzi, J. A. Steitz, Delineation of mRNA export pathways by the use of cell-permeable peptides, *Science* **294**, 1895–1901 (2001).

13. G. M. Hautbergue, M.-L. Hung, A. P. Golovanov, L.-Y. Lian, S. A. Wilson, Mutually exclusive interactions drive handover of mRNA from export adaptors to TAP, *Proc. Natl. Acad. Sci. U.S.A.* **105**, 5154–5159 (2008).

14. A. M. Tintaru, G. M. Hautbergue, A. M. Hounslow, M.-L. Hung, L.-Y. Lian, C. J. Craven, S. A. Wilson, Structural and functional analysis of RNA and TAP binding to SF2/ASF, *EMBO Rep.* **8**, 756–762 (2007).

15. Y. Q. Huang, R. Gattoni, J. Stevenin, J. A. Steitz, SR splicing factors serve as adapter proteins for TAP-dependent mRNA export, *Molecular Cell* **11**, 837–843 (2003).

16. M. Müller-McNicoll, V. Botti, A. M. de Jesus Domingues, H. Brandl, O. D. Schwich, M. C. Steiner, T. Curk, I. Poser, K. Zarnack, K. M. Neugebauer, SR proteins are NXF1 adaptors that link alternative RNA processing to mRNA export, *Genes & Development* **30**, 553–566 (2016).

17. J. F. Caceres, A. R. Krainer, Functional analysis of pre-mRNA splicing factor SF2/ASF structural domains, *EMBO J.* **12**, 4715–4726 (1993).

18. J. C. Long, J. F. Cáceres, The SR protein family of splicing factors: master regulators of gene expression, *Biochem. J.* **417**, 15–27 (2009).

19. J. R. Sanford, N. K. Gray, K. Beckmann, J. F. Cáceres, A novel role for shuttling SR proteins in mRNA translation, *Genes & Development* **18**, 755–768 (2004).

20. M. M. Maslon, S. R. Heras, N. Bellora, E. Eyras, J. F. Cáceres, The translational landscape of the splicing factor SRSF1 and its role in mitosis, *Elife* **2014**, e02028 (2014).

21. F. Haward, M. M. Maslon, P. L. Yeyati, N. Bellora, J. N. Hansen, S. Aitken, J. Lawson, A. von Kriegsheim, D. Wachten, P. Mill, I. R. Adams, J. F. Cáceres, Nucleo-cytoplasmic shuttling of splicing factor srsf1 is required for development and cilia function, *Elife* **10**, e65104 (2021).

22. D. Cazalla, J. Zhu, L. Manche, E. Huber, A. R. Krainer, J. F. Cáceres, Nuclear export and retention signals in the RS domain of SR proteins, *Molecular and Cellular Biology* **22**, 6871–6882 (2002).

23. S. Lin, R. Xiao, P. Sun, X. Xu, X.-D. Fu, Dephosphorylation-dependent sorting of SR splicing factors during mRNP maturation, *Molecular Cell* **20**, 413–425 (2005).

24. Y. Huang, T. A. Yario, J. A. Steitz, A molecular link between SR protein dephosphorylation and mRNA export, *Proc. Natl. Acad. Sci. U.S.A.* **101**, 9666–9670 (2004).

25. K. Meyer, L. Ferraiuolo, C. J. Miranda, S. Likhite, S. McElroy, S. Renusch, D. Ditsworth, C. Lagier-Tourenne, R. A. Smith, J. Ravits, A. H. Burghes, P. J. Shaw, D. W. Cleveland, S. J. Kolb, B. K. Kaspar, Direct conversion of patient fibroblasts demonstrates non-cell autonomous toxicity of astrocytes to motor neurons in familial and sporadic ALS, *Proc. Natl. Acad. Sci. U.S.A.* **111**, 829–832 (2014).

26. N. Gatto, C. Dos Santos Souza, A. C. Shaw, S. M. Bell, M. A. Myszczynska, S. Powers, K. Meyer, L. M. Castelli, E. Karyka, H. Mortiboys, M. Azzouz, G. M. Hautbergue, N. M. Márkus, P. J. Shaw, L. Ferraiuolo, Directly converted astrocytes retain the ageing features of the donor fibroblasts and elucidate the astrocytic contribution to human CNS health and disease, *Aging Cell* **20**, e13281 (2021).

27. L. M. Castelli, L. Cutillo, C. D. S. Souza, A. Sanchez-Martinez, I. Granata, Y.-H. Lin, M. A. Myszczynska, P. R. Heath, M. R. Livesey, K. Ning, M. Azzouz, P. J. Shaw, M. R. Guarracino, A. J. Whitworth, L. Ferraiuolo, M. Milo, G. M. Hautbergue, SRSF1-dependent

inhibition of C9ORF72-repeat RNA nuclear export: genome-wide mechanisms for neuroprotection in amyotrophic lateral sclerosis, *Mol Neurodegener* **16**, 53–24 (2021).

28. S. Mizielinska, S. Gronke, T. Niccoli, C. E. Ridler, E. L. Clayton, A. Devoy, T. Moens, F. E. Norona, I. O. C. Woollacott, J. Pietrzyk, K. Cleverley, A. J. Nicoll, S. Pickering-Brown, J. Dols, M. Cabecinha, O. Hendrich, P. Fratta, E. M. C. Fisher, L. Partridge, A. M. Isaacs, C9orf72 repeat expansions cause neurodegeneration in Drosophila through arginine-rich proteins, *Science* **345**, 1192–1194 (2014).

29. Y. Liu, A. Pattamatta, T. Zu, T. Reid, O. Bardhi, D. R. Borchelt, A. T. Yachnis, L. P. W. Ranum, C9orf72 BAC Mouse Model with Motor Deficits and Neurodegenerative Features of ALS/FTD, *Neuron* **90**, 521–534 (2016).

30. L. Nguyen, L. A. Laboissonniere, S. Guo, F. Pilotto, O. Scheidegger, A. Oestmann, J. W. Hammond, H. Li, A. Hyysalo, R. Peltola, A. Pattamatta, T. Zu, M. H. Voutilainen, H. A. Gelbard, S. Saxena, L. P. W. Ranum, Survival and Motor Phenotypes in FVB C9-500 ALS/FTD BAC Transgenic Mice Reproduced by Multiple Labs, *Neuron* **108**, 784–796 (2020).

31. C. G. Heath, N. Viphakone, S. A. Wilson, The role of TREX in gene expression and disease, *Biochem. J.* **473**, 2911–2935 (2016).

32. G. M. Hautbergue, RNA Nuclear Export: From Neurological Disorders to Cancer, *Adv. Exp. Med. Biol.* **1007**, 89–109 (2017).

33. L.-L. Zou, J.-L. Ma, T. Wang, T.-B. Yang, C.-B. Liu, Cell-penetrating Peptide-mediated therapeutic molecule delivery into the central nervous system, *Curr Neuropharmacol* **11**, 197–208 (2013).

34. S. Stalmans, N. Bracke, E. Wynendaele, B. Gevaert, K. Peremans, C. Burvenich, I. Polis, B. De Spiegeleer, Cell-Penetrating Peptides Selectively Cross the Blood-Brain Barrier In Vivo, *PLoS ONE* **10**, e0139652 (2015).

35. H. A. Popiel, Y. Nagai, N. Fujikake, T. Toda, Protein Transduction Domain-mediated Delivery of QBP1 Suppresses Polyglutamine-induced Neurodegeneration In Vivo, *Molecular Therapy* **15**, 303–309 (2007).

36. H. A. Popiel, Y. Nagai, N. Fujikake, T. Toda, Delivery of the aggregate inhibitor peptide QBP1 into the mouse brain using PTDs and its therapeutic effect on polyglutamine disease mice, *Neuroscience Letters* **449**, 87–92 (2009).

37. Q. Zhang, Z. S. Chen, Y. An, H. Liu, Y. Hou, W. Li, K. F. Lau, A. C. Koon, J. C. K. Ngo, H. Y. E. Chan, A peptidylic inhibitor for neutralizing expanded CAG RNA-induced nucleolar stress in polyglutamine diseases, *RNA* **24**, 486–498 (2018).

38. Q. Zhang, Y. An, Z. S. Chen, A. C. Koon, K. F. Lau, J. C. K. Ngo, H. Y. E. Chan, A Peptidylic Inhibitor for Neutralizing r(GGGGCC)exp-Associated Neurodegeneration in C9ALS-FTD, *Mol Ther Nucleic Acids* **16**, 172–185 (2019).

39. N. Kamei, T. Shingaki, Y. Kanayama, M. Tanaka, R. Zochi, K. Hasegawa, Y. Watanabe, M. Takeda-Morishita, Visualization and Quantitative Assessment of the Brain Distribution

of Insulin through Nose-to-Brain Delivery Based on the Cell-Penetrating Peptide Noncovalent Strategy, *Mol Pharm* **13**, 1004–1011 (2016).

40. N. Kamei, N. Okada, T. Ikeda, H. Choi, Y. Fujiwara, H. Okumura, M. Takeda-Morishita, Effective nose-to-brain delivery of exendin-4 via coadministration with cell-penetrating peptides for improving progressive cognitive dysfunction, *Sci Rep* **8**, 17641–14 (2018).

41. L. M. Castelli, W.-P. Huang, Y.-H. Lin, K.-Y. Chang, G. M. Hautbergue, Mechanisms of repeat-associated non-AUG translation in neurological microsatellite expansion disorders, *Biochem. Soc. Trans.* **49**, 775–792 (2021).

42. N. Percie du Sert, V. Hurst, A. Ahluwalia, S. Alam, M. T. Avey, M. Baker, W. J. Browne, A. Clark, I. C. Cuthill, U. Dirnagl, M. Emerson, P. Garner, S. T. Holgate, D. W. Howells, N. A. Karp, S. E. Lazic, K. Lidster, C. J. MacCallum, M. Macleod, E. J. Pearl, O. H. Petersen, F. Rawle, P. Reynolds, K. Rooney, E. S. Sena, S. D. Silberberg, T. Steckler, H. Würbel, The ARRIVE guidelines 2.0: Updated guidelines for reporting animal research, *PLoS Biol* **18**, e3000410 (2020).

43. J. C. Greene, A. J. Whitworth, I. Kuo, L. A. Andrews, M. B. Feany, L. J. Pallanck, Mitochondrial pathology and apoptotic muscle degeneration in Drosophila parkin mutants, *Proc. Natl. Acad. Sci. U.S.A.* **100**, 4078–4083 (2003)

44. M. Janowski, M. Kuzma-Kozakiewicz, D. Binder, H.-J. Habisch, A. Habich, B. Lukomska, K. Domanska-Janik, A. C. Ludolph, A. Storch, Neurotransplantation in mice: the concorde-like position ensures minimal cell leakage and widespread distribution of cells transplanted into the cisterna magna, *Neuroscience Letters* **430**, 169–174 (2008).

45. N. Deglon, J. L. Tseng, J. C. Bensadoun, A. D. Zurn, Y. Arsenijevic, L. P. De Almeida, R. Zufferey, D. Trono, P. Aebischer, Self-inactivating lentiviral vectors with enhanced transgene expression as potential gene transfer system in Parkinson's disease, *Human Gene Therapy* **11**, 179–190 (2000).

46. A. Bachi, I. C. Braun, J. P. Rodrigues, N. Panté, K. Ribbeck, C. von Kobbe, U. Kutay, M. Wilm, D. Görlich, M. Carmo-Fonseca, E. Izaurralde, The C-terminal domain of TAP interacts with the nuclear pore complex and promotes export of specific CTE-bearing RNA substrates, *RNA* **6**, 136–158 (2000).

47. H. L. Wiegand, G. A. Coburn, Y. Zeng, Y. Kang, H. P. Bogerd, B. R. Cullen, Formation of Tap/NXT1 Heterodimers Activates Tap-Dependent Nuclear mRNA Export by Enhancing Recruitment to Nuclear Pore Complexes, *Molecular and Cellular Biology* **22**, 245–256 (2002).

48. I. C. Braun, A. Herold, M. Rode, E. Izaurralde, Nuclear Export of mRNA by TAP/NXF1 Requires Two Nucleoporin-Binding Sites but Not p15, *Molecular and Cellular Biology* **22**, 5405–5418 (2002).

Acknowledgments:

We acknowledge Adrian Isaacs (University College London) for his kind support with developing our in-house poly-GP MSD-ELISA and for providing the G4C2x3 and G4C2x36 *Drosophila* lines. Stuart Wilson (University of Sheffield) and Javier Cáceres (University of Edinburgh) are acknowledged for kindly donating NXF2/NXF3 and

SRSF1-NRS plasmids, respectively. Laura Ranum (University of Florida) is acknowledged for useful discussions on the C9ORF72-ALS/FTD mouse model. We gratefully thank the patients with ALS and their families for participating in our research and donating biological samples to the NIHR Sheffield Biomedical Research Centre: Translational Neuroscience for chronic neurological disorders (IS-BRC-1215-20017).

Funding:

This study was funded by the Medical Research Council (MRC) New Investigator research grant MR/R024162/1 and the Biotechnology and Biological Sciences Research Council (BBSRC) grants BB/S005277/1 (to GMH); BB/S005579/1 (to CD); MND Association grant Hautbergue/Apr16/846-791 (to GMH, AJW, LF, PJS);

MRC core funding MC_UU_00028/6 and ERC Starting grant DYNAMITO: 309742 (to AJW);

Academy of Medical Sciences Springboard Award SBF002/1142 and MRC grant MR/W00416X/1 (to LF);

Governmental Turkish Doctoral Scholarship (to AG);

LifeArc/ MND Association grant (to KMI, GMH);

NIHR Senior Investigator NF-SI-0617-10077, MND Association AMBRoSIA MNDAOct15/972-797 Medical Research Council (MRC) Centres of Excellence for Neurodegeneration Pathfinder award MR/SOO4920/1 (to PJS);

Thierry Latran Foundation FTLAAP2016/ Astrocyte secretome (to LF, GMH);

Alzheimer's Research UK ARUK-PG2018B-005, European Research Council ERC Advanced Award 294745, MRC DPFS 129016 (to MA);

MRC grants MR/M013251/1 and MR/S025979/1 (to KJDV)

Author contributions:

GMH conceptualized the overall study. GMH, AJW, LF, CD, RM, MA designed the methodologies. PJS genetically characterized human fibroblasts. LF established the iNPC reprogramming and differentiation protocols. LMC performed the HEK cell assays; YHL, NMM, RRM performed the assays with patient-derived neurons and motor neurons . LMC, YHL, GMH performed the in vitro nuclear RNA export assays, YHL, AG performed the pull-down assays. ASM performed the *Drosophila* experiments and LMC, GMH performed the dot blot assays. AG, KMI performed the mouse experiments. SKU performed ITC assays. LMC, AG, AH performed the intracellular delivery imaging. PJS, AH, YHL, LMC, AG performed the poly-GP DPR MSD ELISA. CSB, JCK, RC, CM, AW, KDV performed the primary neuron assays.

GMH wrote the original draft of the manuscript.

All authors contributed to reviewing and editing the manuscript.

Competing interests: GMH, MA, AJW and PJS are coinventors on patents granted in the USA (US10/801027) and Europe (EP3430143) for the use of "inhibitors of SRSF1 to treat neurodegenerative disorders" (WO2017207979A1). The other authors declare no competing interests.

Data and materials availability:

All data are available in the main text or the supplementary materials. Plasmids generated in this study can be provided on request under a University of Sheffield Materials Transfer Agreement. Patient-derived cell lines generated at the University of Sheffield can be provided for collaborative projects under a University of Sheffield Materials Transfer Agreement as stated in the ethics consent form. Human cells obtained from the Coriell biorepository (AG8620) can be purchased directly from Coriell. *Drosophila* lines are available from the Bloomington *Drosophila* Stock Center (Research Resource Identifiers (RRID) are provided in the Supplementary Materials and Methods).

Figures legends

Fig. 1. **WT-CPP inhibits the interaction of endogenous SRSF1 with NXF1 by direct competition** (**A**) Cell permeable peptide sequences were composed of SRSF1 amino-acid 89-120 bearing or not bearing alanine substitutions of 4 arginines interacted with NXF1, a V5 tag and the TAT protein transduction domain (TAT-PTD). (**B**) Shown is immunofluorescence imaging microscopy of HEK293T cells cultured in media supplemented with or without WT-CPP for 72 hours. WT-CPP is detected in the green channel using anti-V5 staining. DAPI was used to delineate the nucleus, whereas vimentin staining in the red channel was used to indicate the cytoplasm. Scale bar: 20 μm. (**C**) Shown is high content automated confocal immunofluorescence imaging (*Opera Phenix*) quantification of V5-CPP positive HEK293T cells cultured with each CPP (n=18 fields; 2

wells; 9 fields/well; 1,200 - 3,000 cells/well). (**D**) Co-immunoprecipitation assays from HEK293T cells transfected for 72 h with either FLAG control or FLAG-NXF1 expressing plasmids and increasing concentrations of either Ctrl-CPP or WT-CPP (0-5 μM). 12% and 20% SDS-PAGE gels were run to respectively analyze the expression of NXF1 (71 kDa) or SRSF1 (28 kDa) and CPPs (Ctrl-CPP: 3 kDa; SRSF1-CPP: 6.1 kDa). (**E**) Shown is a pull down assay using recombinant GST and GST-NXF1:p15 fusions with increasing concentrations of Ctrl-CPP, WT-CPP and m4-CPP. Binding reactions were analyzed by SDS-PAGE stained with Coomassie blue and western immunoblotting using anti-V5 and NXF1 antibodies. Note that GST-NXF1 was co-expressed in bacteria with p15/NXT1, a co-factor which improves the expression of NXF1 through heterodimerization with the NTF2-like domain (aa371-551) of NXF1. (**F**) Competitive binding assays using purified GST, GST-NXF1 and SRSF1 11-196 recombinant proteins expressed in *E. coli* together with increasing concentrations of Ctrl-CPP, WT-CPP or m4-CPP. Binding reactions were analyzed by SDS-PAGE stained with Coomassie blue and western immunoblotting using anti-V5, NXF1 and SRSF1 antibodies. (Full length SRSF1 is not expressed in *E. coli* and could not be used; Fig. S1C showed that amino-acid 89-120 of SRSF1 constitute the NXF1-binding site).

Fig. 2. **WT-CPPs inhibit the nuclear export of reporter *C9ORF72*-repeat transcripts rescuing DPR-associated cytotoxicity in HEK293T cells.** (**A**) Total, nuclear and cytoplasmic expression of sense G4C2$_{45}$-3xV5 repeat transcripts from 72h-transfected HEK293T cells were quantified by qRT-PCR following normalization to *U1* snRNA expression and to 100% in untreated cells (mean ± SEM; one-way ANOVA with Tukey's

correction for multiple comparisons, NS: non-significant, *: p<0.05, ***: p<0.001, ****: p<0.0001; N=3 biological triplicates). (**B**) Total, nuclear and cytoplasmic expression of antisense G2C4$_{43}$-3xV5 repeat transcripts from 72h-transfected cells were quantified by qRT-PCR following normalization to *U1* snRNA expression and to 100% in untreated cells (mean ± SEM; one-way ANOVA with Tukey's correction for multiple comparisons, NS: non-significant, *: p<0.05, ***: p<0.001, ****: p<0.0001; N=3 biological triplicates). (**C, D**) Shown are western blots from cells transfected for 72h with no DPR control or G4C2$_{45}$-3xV5 and G2C4$_{43}$-3xV5 plasmids for 6 hours prior to medium change with media supplemented with increasing concentrations of Ctrl-CPP, m4-CPP or WT-CPP. Blots were probed for V5 and α-tubulin (N=3). (**E**) Intensities of V5-tagged sense DPR bands in panel C were quantified relative to α-tubulin loading control expression (Mean ± SEM; one-way ANOVA with Tukey's correction for multiple comparisons; *: p<0.05, ***: p<0.001, ****:p<0.0001; N=3 biological triplicates). (**F**) Intensities of V5-tagged antisense DPR bands in panel D were quantified relative to α-tubulin loading control expression (Mean ± SEM; one-way ANOVA; *: p<0.05, ****:p<0.0001; N=3 biological triplicates). (**G-H**) Shown is western blot analysis from cells transfected for 72h with no DPR control or G4C2$_{45}$-3xV5 and G2C4$_{43}$-3xV5 plasmids and treated with Ctrl-CPP, m4-CPP or WT-CPP as in (C, D). Blots were probed for poly GR, GP and GA DPRs along with SRSF1, NXF1 and α-tubulin. (**I-J**) MTT cell proliferation assays for cells transfected and treated with CPPs as in panels A-H were performed (mean ± SEM; one-way ANOVA with Tukey's correction for multiple comparisons, **: p<0.01, ***: p<0.001, ****: p<0.0001; N=3 biological triplicates).

Fig. 3. **WT-CPPs do not alter the splicing function of SRSF1** (**A**) Shown are western blots from HEK293T cells treated with 1 μM CPPs for 72h, a concentration that led to efficient nuclear export of *C9ORF7*2-repeat transcripts in figures 2A-B. Blots were probed for SRSF1, NXF1, the nuclear chromatin-associated SSRP1 factor and the cytoplasmic marker TUJ1. (**B**) MTT cell proliferation assays for cells transfected with the indicated plasmids or treated with 1 μM CPPs for 72h (mean ± SEM; one-way ANOVA with Tukey's correction for multiple comparisons, NS: non-significant, ****: $p<0.0001$; N=3 biological triplicates). (**C**) Western blots from transfected or CPP-treated cells (as in panel B) were probed for SRSF1, phospho-SR epitope proteins, SRPK1 or α-tubulin. (**D**) SRSF1-dependent alternative splicing assay is shown. HEK293T cells were treated with 1 μM CPPs or transfected with SRSF1-WT and SRSF1-NRS plasmids for 72 hours prior to whole-cell RNA extraction and qRT-PCR quantification of *CLK4* alternatively-spliced exons 4-5 and constitutively-spliced exon 2 transcript isoforms. Bar chart shows a ratio plotted as % of relative concentrations of the *CLK4* transcript isoforms experiments following normalization to *U1* snRNA expression levels (mean ± SEM; one-way ANOVA with Tukey's correction for multiple comparisons, NS: non-significant, **: $p<0.01$; N=3 biological triplicates).

Fig. 4. **WT-CPPs do not affect protein synthesis or translation of codon-optimized DPRs.** (**A**) HEK293T cells were seeded onto tissue culture plates and treated or not treated 24h later with 1 μM CPPs. Puromycin, a protein synthesis inhibitor resembling an aminoacylated tRNA which is incorporated into actively-translating polypeptide chains, was added for 30 minutes immediately after addition of CPPs (time 0) or after 24, 48 or 72h of culture.

Puromycylated proteins were analyzed by western immunoblotting with an anti-puromycin antibody; tubulin was used as a loading control (N=3). (**B**) Intensities of puromycylated lanes in panel A were quantified relative to their α-tubulin loading control expression (Mean ± SEM; one-way ANOVA with Tukey's correction for multiple comparisons; NS: non-significant, **: $p<0.01$; N=3 biological triplicates). (**C-F**) Shown are western blots of cells transfected with no DPR control, AUG-driven codon-optimized poly-GR-V5, GP-V5 or GA-V5 DPRs biosynthesized independently of the G4C2 hexanucleotide repeat expansion (C-E) or RAN-translated $G4C2_{45}$-3xV5 (F) plasmids for 72h (N=3). Blots are probed for V5 to detect the V5-tagged DPRs, SRSF1 and tubulin. (**G-J**) Intensities of V5-tagged DPR bands in panel C-F are respectively quantified in panels G-J, relative to their α-tubulin loading control expression (Mean ± SEM; one-way ANOVA with Tukey's correction for multiple comparisons; **: $p<0.01$, ***: $p<0.001$, ****: $p<0.0001$; N=3 biological triplicates).

Fig. 5. **WT-CPPs block RAN-translation of DPRs in rat primary neurons** (**A**) High content immunofluorescence imaging microscopy of primary rat cortical neurons cultured in media supplemented with 5 or 10 μM WT-CPP for 72 hours is shown. CPPs were detected in the green channel using anti-V5 staining. DAPI was used to delineate the nucleus and TUJ1 staining (in the red channel) was used to indicate the cytoplasm in confocal planes. Scale bar: 10 μm. Note that the resolution of enlarged high content images was lower than that of conventional light microscopy but was used for the unbiased quantification of the intensity signals in thousands of cells. (**B**) Shown is high content automated confocal immunofluorescence imaging (*Opera Phenix*) quantification of V5-CPP positive neurons

cultured with each CPP (n=18 fields; 2 wells; 9 fields/well; 1,600 - 4,000 neurons/well). (**C**) Western blots from cultured rat cortical neurons transduced for 16 hours with $G4C2_{45}$-3xV5, $G2C4_{43}$-3xV5, Ctrl-RNAi or SRSF1-RNAi expressing lentivirus were treated for 72h with 1 or 5 μM Ctrl-CPP, m4-CPP or WT-CPP are shown. Blots run on 12% SDS-PAGE were probed for V5 antibody to detect the V5-tagged DPRs, SRSF1 and α-tubulin. 20% SDS-PAGE gels were run to detect V5-tagged CPPs. * denotes a non-specific cross-reaction of the antibody.

Fig. 6. **WT-CPPs confer neuroprotection in ALS patient-derived iNeurons.** (**A**) Shown are representative immunofluorescence microscopy images of patient-derived iNeurons stained with DAPI and one of the neuronal markers TUJ1, MAP2 and NeuN. Scale bar: 50 μm. (**B**) High content automated imaging of neuronal markers. Bar chart represents % of TUJ1, NeuN or MAP-2 positive iNeurons ± SEM. Each dot represents a donor line (2 Ctrl or 3 C9-ALS lines) with 3 technical replicates averaged for each dot. (**C**) Representative immunofluorescence microscopy images of C9ORF72-ALS patient-derived iNeurons stained with DAPI and antibodies against the neuronal marker TUJ1 and phospho-TDP-43 to visualize cytoplasmic inclusions. Scale bar: 50 μm. (**D**) In the SRSF1-dependent alternative splicing assay, 3 lines of control iNeurons were treated with 10 μM CPPs for 72 hours prior to whole-cell RNA extraction and qRT-PCR quantification of *CLK4* alternatively-spliced exons 4-5 and constitutively-spliced exon 2 transcript isoforms. Bar chart shows a ratio plotted as % of relative concentrations of the *CLK4* transcript isoforms experiments following normalization to *U1* snRNA expression (mean ± SEM; one-way ANOVA with Tukey's correction for multiple comparisons, NS: non-significant, $p>0.05$;

N=3 biological triplicates). (**E-F**) qRT-PCR quantification of total, nuclear and cytoplasmic expression of pathological *C9ORF72*-repeat transcripts (E) and wild type intron-1-spliced *C9ORF72* transcripts (F) in 3 lines each of control and C9ORF72-ALS patient-derived iNeurons treated with 10 μM of Ctrl-CPP or WT-CPP (mean ± SEM; one-way ANOVA with Tukey's correction for multiple comparisons, NS: non-significant, *: $p<0.05$, **: $p<0.01$, ****: $p<0.0001$; N=3 biological triplicates). (**G**) MSD-ELISA quantification of poly-GP DPRs in Ctrl and C9ORF72-ALS patient-derived iNeurons treated with either PBS, 10 μM Ctrl-CPP or WT-CPP and 5 MOI (Multiplicity Of Infection) of Ctrl-RNAi or SRSF1-RNAi expressing lentivirus (mean ± SEM; one-way ANOVA with Tukey's correction for multiple comparisons; NS: non-significant, **: $p<0.01$; N=3 independent triplicate lines). Poly-GP DPRs were quantified against a standard curve established with a GPx7 peptide and steady-state amounts were normalized to 100% for the PBS-treated cultures (Fig. S11). (**H**) Shown are representative immunofluorescence microscopy images of C9ORF72-ALS patient-derived motor neurons stained with DAPI and an anti SMI-32 antibody (left panel). Scale bar: 50 μm. High content automated imaging of neuronal markers. Bar chart represents % of SMI-32 positive motor neurons ± SEM (right panel). Each dot represents one Ctrl donor line or C9-ALS donor line with 3 technical replicates. (**I**) High content imaging analysis of healthy control and C9ORF72-ALS iMotor Neuron (iMN) survival in co-cultures with control or C9ORF72-ALS iAstrocytes. For lentiviral-mediated depletion of SRSF1, iAstrocyes and iMNs were separately transduced with lentivirus-SRSF1-RNAi 48 hours prior to co-culture for a period of 72 hours. 10 μM CPPs were added to the medium of the co-cultures for 72 hours. Scale bar: 50 μm. (**J**) Automated quantification of motor neuron survival after 72 hours in

co-culture with iAstrocytes (iMN count ± SEM; two-way ANOVA with Tukey's correction for multiple comparisons, NS: non-significant, ****: $p<0.0001$, N=3 independent triplicate lines). Data from 3 different healthy control cell lines were pooled together.

Fig. 7. WT-CPP inhibits RAN-translation of DPRs in vivo and mitigates locomotor deficits in a *Drosophila* repeat expansion model of C9ORF72-ALS/FTD. (**A**) The timeline for developmental stages of *Drosophila* larvae and adult flies fed on food supplemented with 10 μM CPPs is illustrated (**B**). Shown are representative western blots of total, nuclear and cytoplasmic fractions isolated from $G4C2_{36}$ third instar larvae treated or not treated with Ctrl-CPP, WT-CPP or m4-CPP were probed for nuclear Histone H3 and cytoplasmic tubulin proteins. Genotype: *da*-GAL4>*G4C2*$_{36}$. (**C**) Total and cytoplasmic expression of sense $G4C2_{36}$ repeat transcripts from $G4C2_{36}$ third instar larvae treated or not treated with Ctrl-CPP, WT-CPP or m4-CPP for 5 days were quantified by qRT-PCR following normalization to Tub84b mRNA expression. Plot represents normalized Cytoplasmic/Total ratios (mean ± SEM; one-way ANOVA with Sidak's correction for multiple comparisons, NS: non-significant, *: $p<0.05$, **: $p<0.01$, N=3 biological triplicates). Genotype: *da*-GAL4>*G4C2*$_{36}$. (**D**) Crawling assays in $G4C2_3$ healthy control or $G4C2_{36}$ third instar larvae fed or not fed with 10 μM Ctrl-CPP, WT-CPP or m4-CPP for 5 days (mean ± SEM; one-way ANOVA with Sidak's correction for multiple comparisons, NS: non-significant, ***: $p<0.001$). Numbers of animals used in the assays are indicated at the bottom of each bar. Genotype: *nSyb*-GAL4>*G4C2*$_3$, *nSyb*-GAL4>*G4C2*$_{36}$. (**E**) Climbing assays in $G4C2_3$ healthy control or $G4C2_{36}$ 2 day old adult *Drosophila* fed or not fed with 10 μM Ctrl-CPP, WT-CPP or m4-CPP for 10 days (mean ± 95% CI normalized to control; Kruskal-Wallis

test with Dunn's correction for multiple comparisons, NS: non-significant, *: $p<0.05$). Numbers of animals used in the assays are indicated at the bottom of each bar. Genotypes: *D42*-GAL4>*G4C2*$_3$, *D42*-GAL4>*G4C2*$_{36}$. (**F**) Dot blot assays from G4C2$_3$ healthy control or G4C2$_{36}$ 2-day-old adult heads from *Drosophila* fed or not fed with 10 μM Ctrl-CPP, WT-CPP or m4-CPP for 10 days were probed for poly GA, GP and GR DPRs as well as for the loading control actin. Genotypes: *D42*-GAL4>*G4C2*$_3$, *D42*-GAL4>*G4C2*$_{36}$. (**G**) Dot blot assays from adult heads of *Drosophila* fed with food supplemented or not with 60 μM of Ctrl-CPP or WT-CPP for 10 days. Anti-V5 and actin antibodies were used to probe for V5-tagged CPPs and as a loading control, respectively.

Fig. 8. **WT-CPP inhibits RAN-translation of DPRs in the brains of C9ORF72-ALS/FTD mice**. 5 μl of Ctrl-CPP and WT-CPP at 1 mM (~5 μM brain exposure) were injected into the cisterna magna of non-transgenic (NTg) and transgenic C9ORF72-ALS/FTD (C9-500) mice at post-natal day 30. Animals were sacrificed 4 days post-treatment (**A**) Shown are representative immunofluorescence microscopy images of untreated (UT) and CPP-injected C9-500 cerebellum/brainstem sections stained with DAPI and V5 antibody to detect V5-tagged CPPs. The no V5 antibody control shows that the staining signal is specific for the V5 antibody. Ctrl-CPP exhibits a more diffuse pattern in congruence with an inactive peptide which does not interact with any molecular target and is prone to diffusion and quick turnover. Arrows delineate the CSF space between the cerebellum (Cb) and the brain stem (BS). Scale bar: 50 μm. (**B**) Meso Scale Discovery (MSD)-ELISA quantification of poly-GP DPRs in non-transgenic (NTg) and transgenic C9ORF72-ALS/FTD (C9-500) mice that received intrathecal injections or not (-) of Ctrl-CPP and

WT-CPP at post-natal day 30. Animals were sacrificed 4 days post-delivery. Bar chart indicates poly-GP DPRs in cerebellum/brain stem homogenates (mean ± SEM; one-way ANOVA with Tukey's correction for multiple comparisons; **: $p<0.01$, ***: $p<0.001$; N=3 mice/group). Poly-GP DPRs were quantified against a standard curve established with a GPx7 peptide and steady-state amounts were normalized to 100% for the untreated C9-500 mice. (**C**) Weights of CNS and peripheral organs from NTg and C9-500 Tg mice treated or not with Ctrl-CPP or WT-CPP. Bar chart shows weight in grams (mean ± SEM; one-way ANOVA with Tukey's correction for multiple comparisons; NS: non-significant; **: $p<0.01$, ***: $p<0.001$; N=3 mouse organs/group).

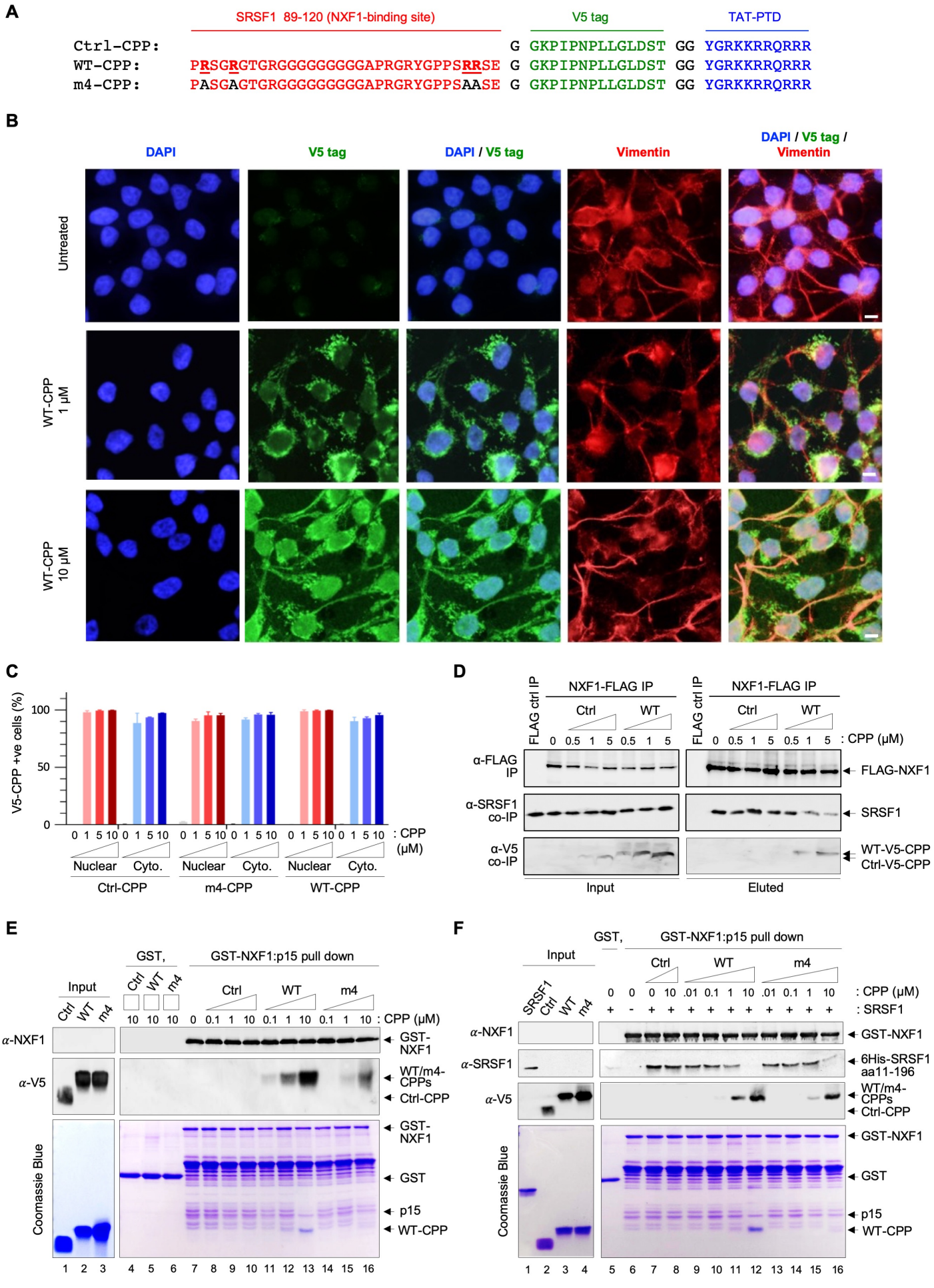

A
SRSF1 89-120 (NXF1-binding site)
V5 tag
TAT-PTD
Ctrl-CPP: G GKPIPNPLLGLDST GG YGRKKRRQRRR
WT-CPP: PRSGRGTGRGGGGGGGGGAPRGRYGPPSRRSE G GKPIPNPLLGLDST GG YGRKKRRQRRR
m4-CPP: PASGAGTGRGGGGGGGGGAPRGRYGPPSAASE G GKPIPNPLLGLDST GG YGRKKRRQRRR
B
DAPI
V5 tag
DAPI / V5 tag
Vimentin
DAPI / V5 tag / Vimentin
Untreated
WT-CPP 1 μM
WT-CPP 10 μM
C
V5-CPP +ve cells (%)
100
50
0
0 1 5 10 0 1 5 10 0 1 5 10 0 1 5 10 0 1 5 10 0 1 5 10 : CPP (μM)
Nuclear
Cyto.
Ctrl-CPP
m4-CPP
WT-CPP
D
FLAG ctrl IP
NXF1-FLAG IP
Ctrl
WT
0 0.5 1 5 0.5 1 5 : CPP (μM)
α-FLAG IP
α-SRSF1 co-IP
α-V5 co-IP
FLAG-NXF1
SRSF1
WT-V5-CPP
Ctrl-V5-CPP
Input
Eluted
E
Input
GST,
GST-NXF1:p15 pull down
Ctrl
WT
m4
10 10 10 0 0.1 1 10 0.1 1 10 0.1 1 10 : CPP (μM)
α-NXF1
α-V5
Coomassie Blue
GST-NXF1
WT/m4-CPPs
Ctrl-CPP
GST
p15
WT-CPP
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16
F
Input
GST,
GST-NXF1:p15 pull down
SRSF1
0 0 0 10 .01 0.1 1 10 .01 0.1 1 10 : CPP (μM)
+ - + + + + + + + + + + : SRSF1
α-NXF1
α-SRSF1
α-V5
Coomassie Blue
GST-NXF1
6His-SRSF1 aa11-196
WT/m4-CPPs
Ctrl-CPP
GST
p15
WT-CPP
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16

Figure 1

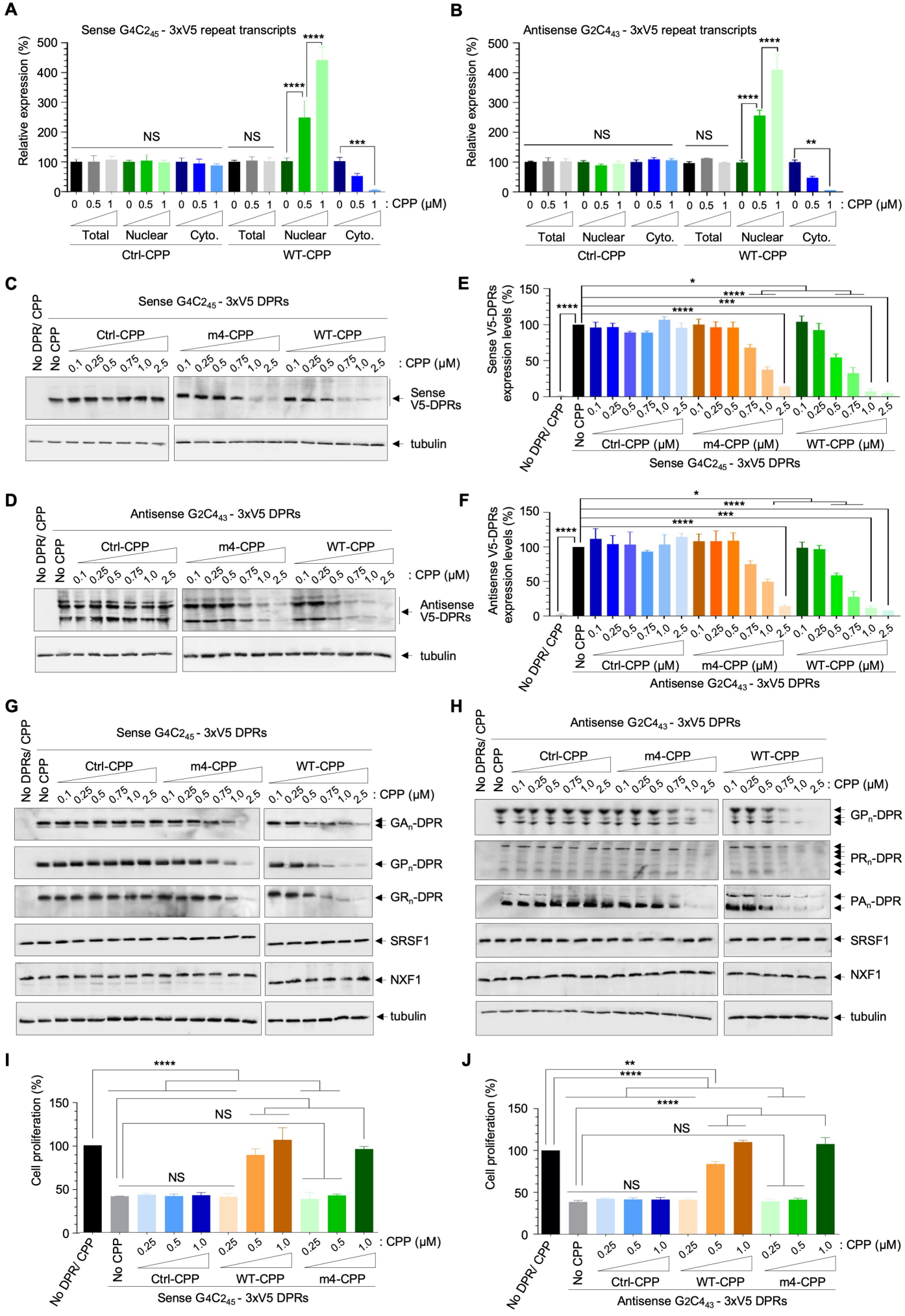

A
Sense G4C245 - 3xV5 repeat transcripts
Relative expression (%)
NS
NS
****
****
***
: CPP (μM)
Total
Nuclear
Cyto.
Ctrl-CPP
WT-CPP
B
Antisense G2C443 - 3xV5 repeat transcripts
**
C
Sense G4C245 - 3xV5 DPRs
No DPR/ CPP
No CPP
Ctrl-CPP
m4-CPP
WT-CPP
Sense V5-DPRs
tubulin
D
Antisense G2C443 - 3xV5 DPRs
Antisense V5-DPRs
E
Sense V5-DPRs expression levels (%)
Ctrl-CPP (μM)
m4-CPP (μM)
WT-CPP (μM)
F
Antisense V5-DPRs expression levels (%)
G
No DPRs/ CPP
GAn-DPR
GPn-DPR
GRn-DPR
SRSF1
NXF1
H
PRn-DPR
PAn-DPR
I
Cell proliferation (%)
J

Figure 2

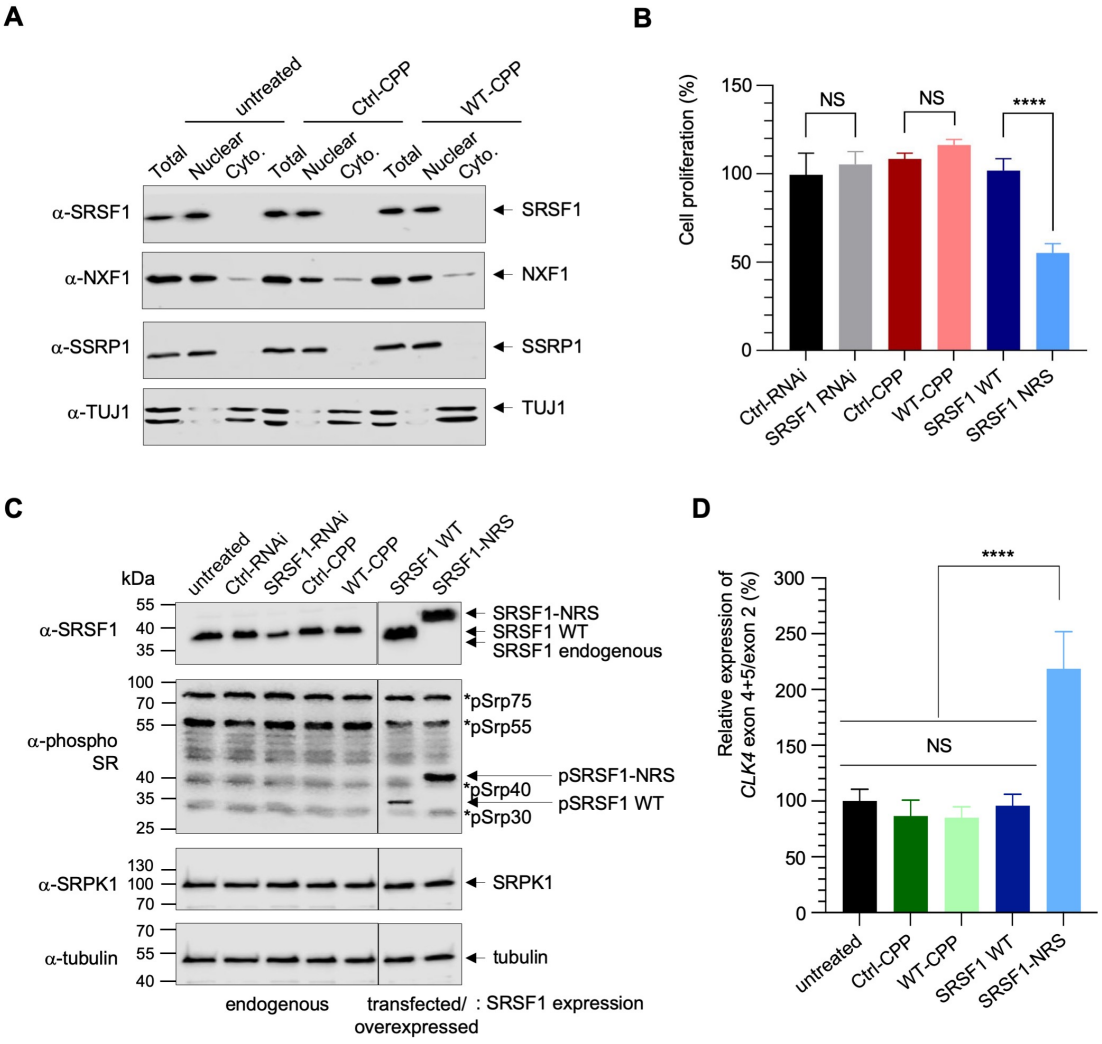

A
untreated
Ctrl-CPP
WT-CPP
Total
Nuclear
Cyto.
α-SRSF1
SRSF1
α-NXF1
NXF1
α-SSRP1
SSRP1
α-TUJ1
TUJ1
B
Cell proliferation (%)
150
100
50
0
NS
NS
****
Ctrl-RNAi
SRSF1 RNAi
Ctrl-CPP
WT-CPP
SRSF1 WT
SRSF1 NRS
C
kDa
untreated
Ctrl-RNAi
SRSF1-RNAi
Ctrl-CPP
WT-CPP
SRSF1 WT
SRSF1-NRS
α-SRSF1
55
40
35
SRSF1-NRS
SRSF1 WT
SRSF1 endogenous
α-phospho SR
100
70
55
40
35
25
*pSrp75
*pSrp55
pSRSF1-NRS
*pSrp40
pSRSF1 WT
*pSrp30
α-SRPK1
130
100
70
SRPK1
α-tubulin
70
55
40
tubulin
endogenous
transfected/ overexpressed
: SRSF1 expression
D
Relative expression of CLK4 exon 4+5/exon 2 (%)
300
250
200
150
100
50
0
****
NS
untreated
Ctrl-CPP
WT-CPP
SRSF1 WT
SRSF1-NRS

Figure 3

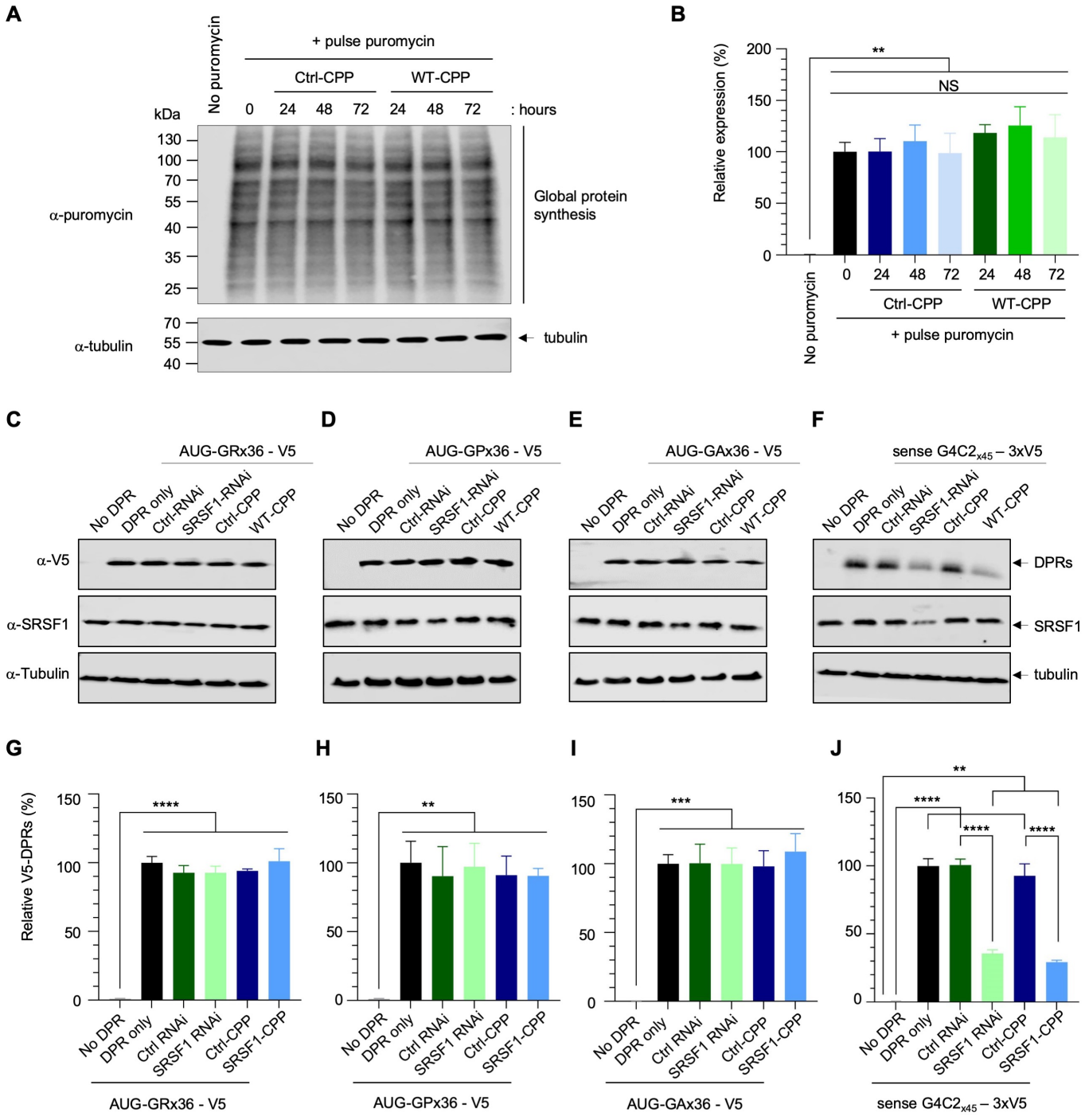

A
No puromycin
+ pulse puromycin
Ctrl-CPP
WT-CPP
kDa
0 24 48 72 24 48 72 : hours
130
100
70
55
40
35
25
α-puromycin
Global protein synthesis
70
55
40
α-tubulin
tubulin
B
Relative expression (%)
200
150
100
50
0
**
NS
0 24 48 72 24 48 72
Ctrl-CPP
WT-CPP
No puromycin
+ pulse puromycin
C
AUG-GRx36 - V5
D
AUG-GPx36 - V5
E
AUG-GAx36 - V5
F
sense G4C2x45 – 3xV5
No DPR
DPR only
Ctrl-RNAi
SRSF1-RNAi
Ctrl-CPP
WT-CPP
α-V5
α-SRSF1
α-Tubulin
DPRs
SRSF1
tubulin
G
H
I
J
Relative V5-DPRs (%)
150
100
50
0
****
**
***
No DPR
DPR only
Ctrl RNAi
SRSF1 RNAi
Ctrl-CPP
SRSF1-CPP
AUG-GRx36 - V5
AUG-GPx36 - V5
AUG-GAx36 - V5
sense G4C2x45 – 3xV5

Figure 4

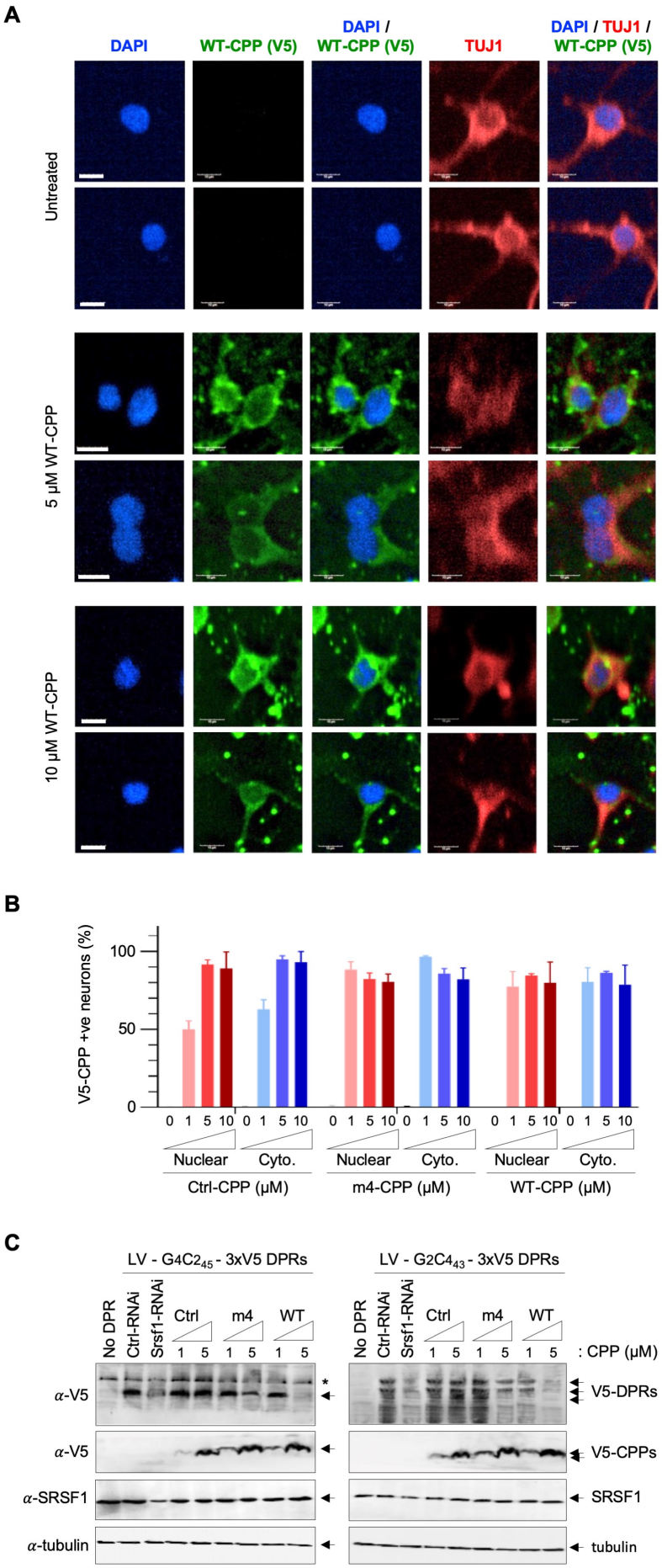

A
DAPI
WT-CPP (V5)
DAPI / WT-CPP (V5)
TUJ1
DAPI / TUJ1 / WT-CPP (V5)
Untreated
5 µM WT-CPP
10 µM WT-CPP
B
V5-CPP +ve neurons (%)
100
50
0
0 1 5 10 0 1 5 10 0 1 5 10 0 1 5 10 0 1 5 10 0 1 5 10
Nuclear
Cyto.
Ctrl-CPP (µM)
m4-CPP (µM)
WT-CPP (µM)
C
LV - G4C2 45 - 3xV5 DPRs
LV - G2C4 43 - 3xV5 DPRs
No DPR
Ctrl-RNAi
Srsf1-RNAi
Ctrl
m4
WT
1 5
: CPP (µM)
α-V5
α-V5
α-SRSF1
α-tubulin
V5-DPRs
V5-CPPs
SRSF1
tubulin

Figure 5

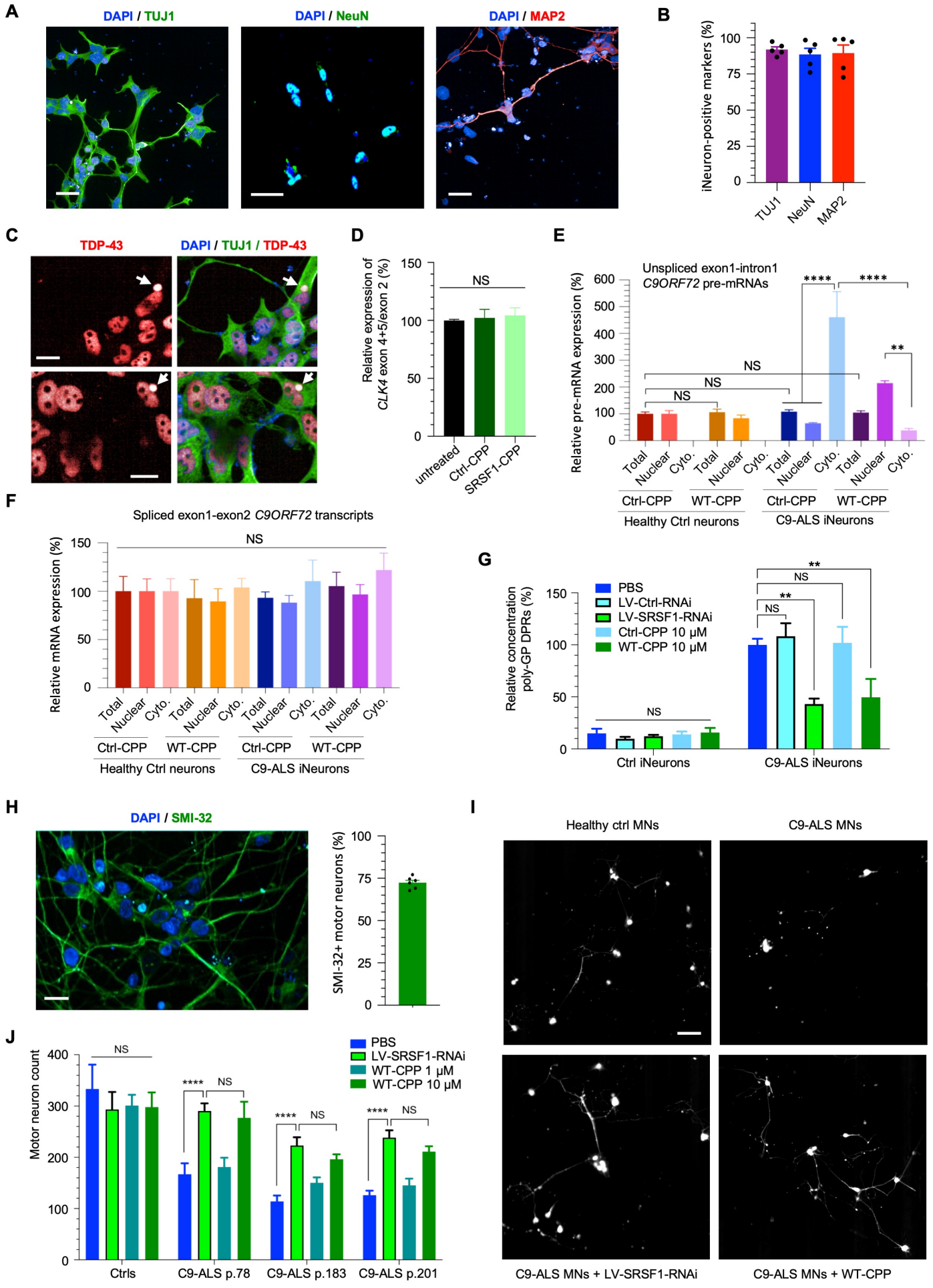

A
DAPI / TUJ1
DAPI / NeuN
DAPI / MAP2
B
iNeuron-positive markers (%)
100
75
50
25
0
TUJ1
NeuN
MAP2
C
TDP-43
DAPI / TUJ1 / TDP-43
D
Relative expression of CLK4 exon 4+5/exon 2 (%)
150
100
50
0
NS
untreated
Ctrl-CPP
SRSF1-CPP
E
Unspliced exon1-intron1 C9ORF72 pre-mRNAs
Relative pre-mRNA expression (%)
600
500
400
300
200
100
0
****
****
**
NS
NS
NS
Total
Nuclear
Cyto.
Ctrl-CPP
WT-CPP
Healthy Ctrl neurons
C9-ALS iNeurons
F
Spliced exon1-exon2 C9ORF72 transcripts
Relative mRNA expression (%)
150
100
50
0
NS
Total
Nuclear
Cyto.
Ctrl-CPP
WT-CPP
Healthy Ctrl neurons
C9-ALS iNeurons
G
Relative concentration poly-GP DPRs (%)
150
100
50
0
PBS
LV-Ctrl-RNAi
LV-SRSF1-RNAi
Ctrl-CPP 10 µM
WT-CPP 10 µM
NS
**
NS
**
NS
Ctrl iNeurons
C9-ALS iNeurons
H
DAPI / SMI-32
SMI-32+ motor neurons (%)
100
75
50
25
0
I
Healthy ctrl MNs
C9-ALS MNs
C9-ALS MNs + LV-SRSF1-RNAi
C9-ALS MNs + WT-CPP
J
Motor neuron count
400
300
200
100
0
PBS
LV-SRSF1-RNAi
WT-CPP 1 µM
WT-CPP 10 µM
NS
****
NS
****
NS
****
NS
Ctrls
C9-ALS p.78
C9-ALS p.183
C9-ALS p.201

Figure 6

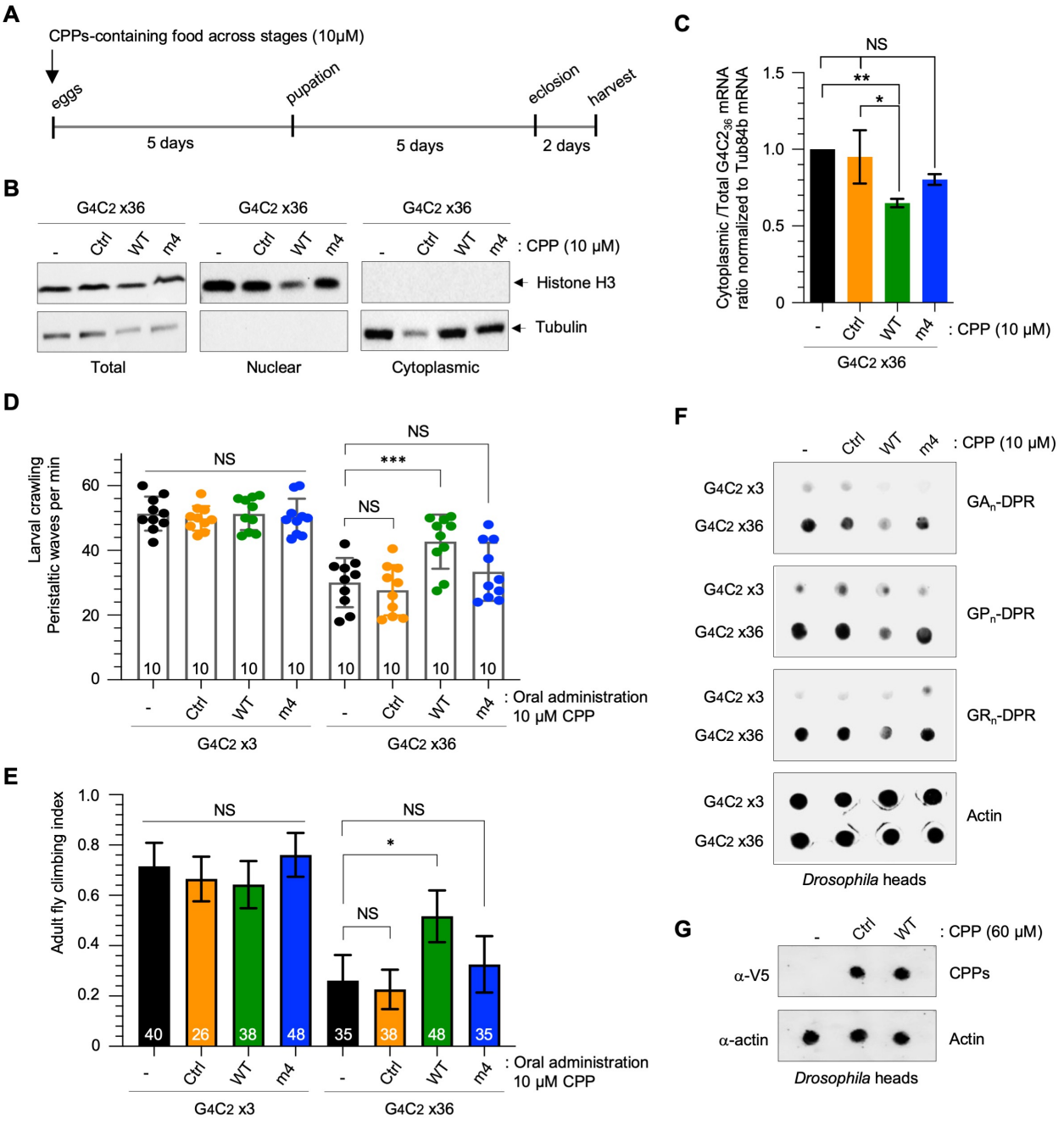

A
CPPs-containing food across stages (10μM)
eggs
pupation
eclosion
harvest
5 days
5 days
2 days
B
G4C2 x36
- Ctrl WT m4
: CPP (10 μM)
Histone H3
Tubulin
Total
Nuclear
Cytoplasmic
C
Cytoplasmic /Total G4C2$_{36}$ mRNA ratio normalized to Tub84b mRNA
NS
**
*
: CPP (10 μM)
G4C2 x36
D
Larval crawling Peristaltic waves per min
NS
***
: Oral administration 10 μM CPP
G4C2 x3
G4C2 x36
E
Adult fly climbing index
40 26 38 48 35 38 48 35
F
G4C2 x3
G4C2 x36
GA$_n$-DPR
GP$_n$-DPR
GR$_n$-DPR
Actin
Drosophila heads
G
: CPP (60 μM)
α-V5
CPPs
α-actin
Actin
Drosophila heads

Figure 7

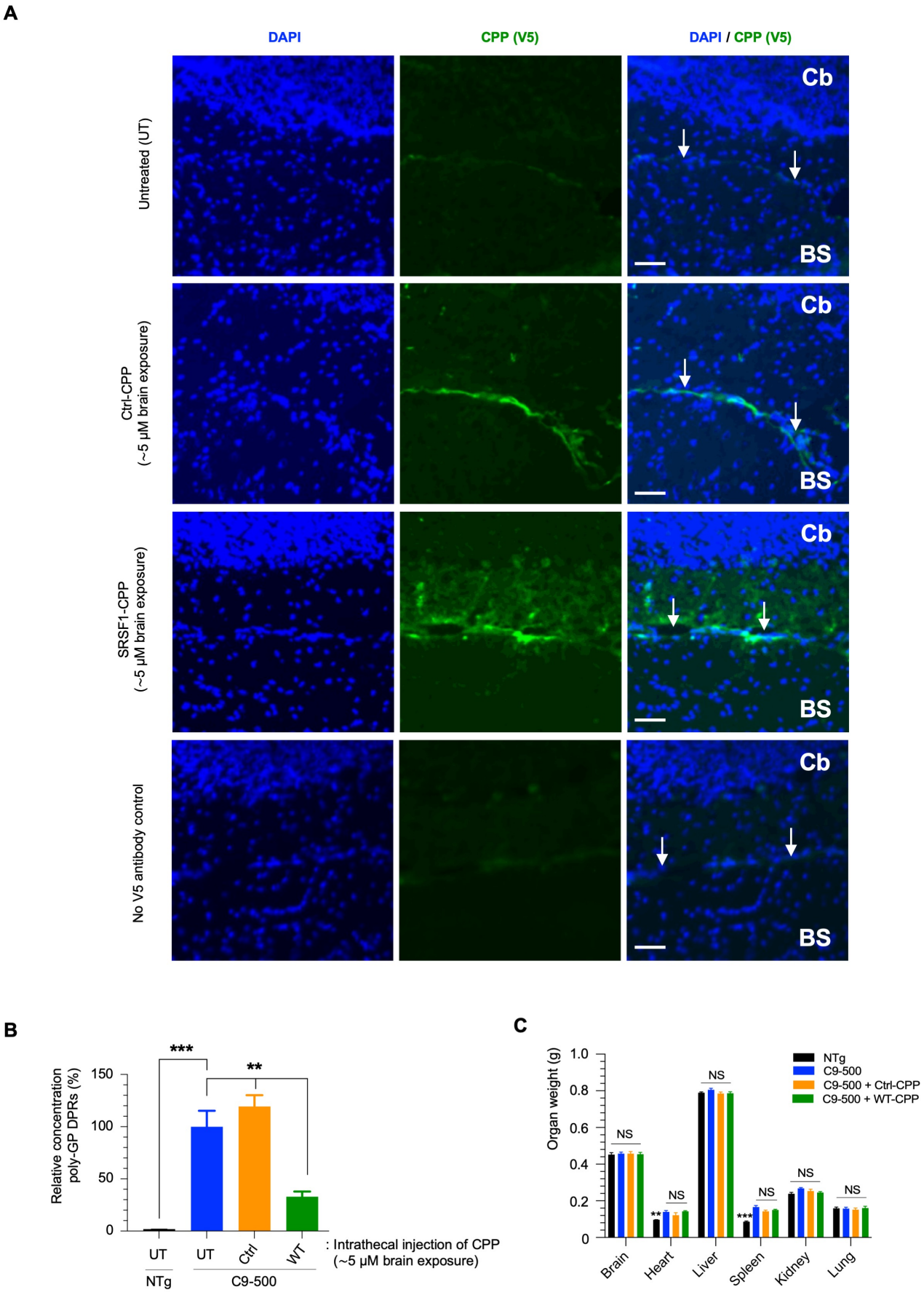

A
DAPI
CPP (V5)
DAPI / CPP (V5)
Untreated (UT)
Ctrl-CPP
(~5 μM brain exposure)
SRSF1-CPP
(~5 μM brain exposure)
No V5 antibody control
Cb
BS
B
Relative concentration
poly-GP DPRs (%)
150
100
50
0
***
**
UT
UT
Ctrl
WT
NTg
C9-500
: Intrathecal injection of CPP
(~5 μM brain exposure)
C
Organ weight (g)
1.0
0.8
0.6
0.4
0.2
0
NTg
C9-500
C9-500 + Ctrl-CPP
C9-500 + WT-CPP
NS
**
***
Brain
Heart
Liver
Spleen
Kidney
Lung

Figure 8

Supplementary Materials for

A cell-penetrant peptide blocking *C9ORF72*-repeat RNA nuclear export reduces the neurotoxic effects of dipeptide repeat proteins

Lydia M. Castelli[1†], Ya-Hui Lin[1†], Alvaro Sanchez-Martinez[2†], Aytaç Gül[1], Kamallia Mohd Imran[1], Adrian Higginbottom[1], Santosh Kumar Upadhyay[3], Nóra M. Márkus[1‡], Raquel Rua Martins[1], Johnathan Cooper-Knock[1], Claire Montmasson[1§], Rebecca Cohen[1‖], Amy Walton[1], Claudia S. Bauer[1], Kurt J. De Vos[1], Richard M. Mead[1], Mimoun Azzouz[1], Cyril Dominguez[3], Laura Ferraiuolo[1‖], Pamela J. Shaw[1‖], Alexander J. Whitworth[2‖], Guillaume M. Hautbergue[1‖]*

* Correspondence to: g.hautbergue@sheffield.ac.uk

This PDF file includes:

Supplementary Materials and Methods
Figs. S1 to S16
Tables S1 to S5
References (*5, 12-14, 20, 22, 28, 43-48*)

Supplementary Materials and Methods

Tissue culture of HEK cells and rat primary neurons

HEK293T cells (*ATCC* CRL-3216) were maintained in a 37°C incubator with 5% CO_2. HEK293T cells were cultured in Dulbecco's Modified Eagle Medium (Lonza) supplemented with 10% fetal bovine serum (FBS) (Biosera) and 5 U/ml Penstrep (Lonza). For experiments, cells were plated on either 24-well plates (50,000 cells / well for western blot analysis or 25,000 cells / well for MTT analysis), 6-well plates (200,000 cells / well) or 10 cm plates (1.5×10^6 cells per plate). Cells were transfected with 500-700 ng plasmid(s) / well for 24-well plates, 2 µg / well for 6-well plates or 15 µg plasmid for 10 cm plates using 3.5 µg PEI/ml media and one tenth medium volume OptiMEM for 72 h. A list of plasmid used in this study is provided in **Table S3**.

Primary rat cortical neurons were isolated at embryonic stage E18 from Sprague Dawley rat embryos (*Charles River*) and cultured on glass coverslips coated with poly-L-lysine in 12- or 24-well plates in neurobasal medium supplemented with B27 supplement (*Invitrogen*), 100 U/ml penicillin, 100 µg/ml streptomycin and 2 mM L-glutamine. Cortical neurons were transduced with 5 MOI in fresh culture medium overnight (16 h) prior to replacement with fresh medium containing CPP for 72 h. For Ctrl- and SRSF1-RNAi conditions, neurons were transduced with 5 MOI of LV-DPR and LV-RNAi viruses overnight (16 h) prior to replacement with fresh medium for 72 h.

Drosophila melanogaster stocks and husbandry

Flies were raised under standard conditions in a humidified, temperature-controlled incubator with a 12 h:12 h light:dark cycle at 25°C, on standard food consisting of agar, cornmeal, molasses, propionic acid and yeast. Transgene expression was induced using the ubiquitous da-GAL4 driver or tissue specific nSyb-GAL4 or D42-GAL4. The following strains were obtained from the Bloomington *Drosphila* Stock Center (RRID:SCR_006457): da-GAL4 (RRID:BDSC_55850), nSyb-GAL4 (RRID:BDSC_51635), D42-GAL4 (RRID:BDSC_8816). UAS-G4C2x3 and UAS-G4C2x36 lines *(28)* were a gift from Adrian M. Issacs (UCL Queen Square Institute of Neurology) and are available from the Bloomington *Drosophila* Stock Center (RRID:BDSC_58687 and RRID:BDSC_58688, respectively). For treatment with CPPs, crosses of the parental genotypes were set up in standard food containing 10 µM Ctrl, WT or m4 CPPs, providing the offspring exposure to the peptides throughout development and into the adult stage (5 days for larval crawling assays and 10 days for adult climbing assays).

Drosophila melanogaster locomotor assays.

The startle induced negative geotaxis (climbing) assay was performed using a counter-current apparatus. Briefly, twenty 2-day-old male adult flies treated throughout development or not with 10 µM Ctrl, WT and m4 CPPs were placed into the first chamber, tapped to the bottom, and given 10 s to climb a 19 cm distance. The weighted performance of several group of flies for each genotype was normalized to the maximum possible score and expressed as Climbing index *(43)***.** For larval crawling assays, wandering third instar larvae treated throughout development or not with 10 µM Ctrl, WT or m4 CPPs were placed in the centre of a 1% agar plate and left to acclimatise for 30 seconds, after which the number of peristaltic waves that occurred in the following minute were recorded.

Cisterna magna injections in mice

BAC transgenic mice at postnatal; (p) day 30 were injected directly in the cisterna magna with 5 µl of SRSF1 WT-CPP or Control-CPP at 1 mM in PBS using the method essentially as described in

(44). Mice were anesthetized with isofluorane and injected using a stereotaxic apparatus containing a 33-gauge Hamilton syringe mounted on an automated perfusion pump and the solution administered at 1 µl per minute. At p34 mice were perfused under terminal anesthesia with PBS and tissue collected and snap frozen for analysis. As a further control non-transgenic littermates and BAC transgenic mice were collected with no injection of CPP peptides. As this was a pilot study, only three female mice per group were used as the minimum to enable statistical testing. Mice were not randomized to groups, instead litters were distributed evenly across groups. The tissue analysis of DPR expression levels was conducted blinded.

Plasmids and cell permeable peptides (CPPs)

pcDNA3.1 plasmids bearing uninterrupted C9ORF72 hexanucleotide sense (GGGGCCx45) and antisense (CCCCGGx43) repeats were generated using annealed and concatemerized G4C2x15 Mung bean-blunted oligonucleotides cloned into the Klenow filled-in EcoRI site of pcDNA3.1-myc-HisA (*Invitrogen*) as described in *(5)*. A 3x V5 tag cassette with V5 and stop codons in the 3 frames was custom synthesised (ThermoFisher Scientific) and subcloned downstream of the repeat sequences using the NotI/Xba1 sites to allow detection of DPR proteins in all three reading frames (Figs. S5 and S6). The hexanucleotide sense and antisense repeats with 3x V5 tags were subsequently subcloned from the pcDNA3.1 plasmids into the lentiviral plasmid SIN-PGK-cPPT-GDNF-WHV *(45)* using the BamHI and XhoI restriction sites. Lentiviral control and SRSF1-miRNA constructs were generated in *(5)*. SRSF1 (amino-acids 11-196) was cloned into pET24b harbouring a 6His C-terminal tag. p3XFLAG-NXF1/TAP was generated in *(13)* using p3XFLAG-myc-CMV^{TM}-26 (*Sigma*). Codon-optimized sequences encoding poly-Gly-Pro, poly-Gly-Arg and poly-Gly-Ala x36 DPRs independently of G4C2 repeats were subcloned using BamHI/NotI into pCI-Neo-V5-N restricted with BclI/NotI. SRSF1 WT and SRSF1-NRS plasmids were kindly provided by Prof Javier Cáceres at the University of Edinburgh *(22)*. Full length TAP/NXF1 was amplified as a BamHI/XhoI PCR fragment and cloned into pGEX6P1 BamHI/XhoI (this study) prior to co-transformation into bacteria with pET9a-p15 *(13)* for recombinant purification of the GST-NXF1:p15 complex as in *(13)*. pGEX-6P1 plasmids encoding GST fusions of NXF2 and NXF3 were kindly provided by Prof Stuart Wilson at the University of Sheffield (unpublished). All cell permeable peptides used in this study were custom synthesised by *ThermoFisher Scientific* at >90% purity (10-14 mg scale) and resuspended in PBS at 1 mM prior to flash freezing of aliquots in liquid nitrogen and storage at -80°C. Design and sequence information is provided in Fig. 1A.

Lentivirus production

Lentivirus production was performed as described previously *(5)*. Briefly, twenty 10 cm dishes seeded with 3×10^6 HEK293T cells/dish were each transfected with 13 µg pCMVΔR8.92, 3.75 µg pM2G, 3 µg pRSV and 13 µg SIN-CMV-Ctrl-miRNA, SIN-CMV-SRSF1-miRNA, SIN-CMV-G4C2x45-3xV5 or SIN-CMV-G2C4x43-3xV5 using calcium phosphate transfection. Media was replaced after 12 h. After a further 48 h, the supernatant was collected, filtered through a 0.45 µm filter and centrifuged at 19,000 rpm for 90 min at 4°C using a SW28 rotor (Beckman). The viral pellet was re-suspended in PBS with 1% BSA and stored at −80'C. The biological titre of the miRNA containing virus was determined by transducing HeLa cells with 10^{-2}, 10^{-3} and 10^{-4} dilutions of the vector. 72 h post-transduction, the percentage of GFP positive cells was measured with a Fluorescent-Activated cell sorter (FACS, LSRII). For the repeat containing virus, HeLa cells were transduced with 10^{-2}, 10^{-3} and 10^{-4} dilution of the vector. 72 h post-transduction, cells were fixed and immunofluorescence stained with V5 antibody as detailed later to obtain a biological titre. The biological titre is expressed as the number of transducing units per ml (TU/ml)

and is calculated as follows: Vector titre = [(% positive cells × total number of cells) × dilution factor × 2].

Recombinant protein purification and GST Pull-down assay

E. coli BL21 (DE3)-RP cells were co-transformed with pGEX6P1-NXF1-3 and pET9a-p15 (*12*). All recombinant proteins were expressed in *E. coli* BL21 (DE3)-RP cells and induced with 0.4 mM of isopropyl-β-D-thiogalactoside (IPTG, Sigma) for 3h at 37˚C (SRSF1), overnight at room temperature for GST-NXF1:p15, GST-NXF2:p15 and GST-NXF3:p15. Recombinant GB1-6His-SRSF1 (11-196) was purified by IMAC chromatography on TALON/Cobalt beads (Clontech) with the following buffers (Lysis Buffer: 50 mM Tris pH8.0, 1M NaCl, 0.5% Triton X-100; Wash Buffer: 50 mM Tris pH8.0, 1M NaCl, 5 mM Imidazole; Elution Buffer: 20 mM Tris pH8.0, 0.5 M NaCl, 0.2 M Imidazole). Protein concentration was determined by Bradford protein assay (*BioRad*). For pull-down reaction, 0.25 g IPTG-induced GST-NXF1:p15 bacteria pellets was lysed by sonication in phosphate buffered saline (PBS) with 0.5% Triton X-100 and supplemented with protease inhibitor (Complete™, EDTA-free Protease Inhibitor Cocktail, *Roche*) and 2 mM phenylmethanesulfonyl fluoride (PMSF, Sigma), and immobilized on GST Sepharose 4B beads (GE Healthcare). For cell permeable peptide binding reactions, different concentrations of CPPs were incubated with immobilized GST-NXF1:p15 in 250 µl of PBS with 0.5 % Triton X-100 and RNase A (0.04 mg/ml) at 4℃ for 1 hr. For competition assay, 1.2 µg of purified SRSF1 (11-196)-6His was incubated with immobilized GST-NXF1:p15 in 250 µl of PBS containing 0.5 % Triton X-100 and RNase A (0.04 mg/ml) and different concentrations of CPPs at 4 ℃ for 1 hr. The GST-NXF1:p15-interacting proteins were eluted with GSH elution buffer (50 mM Tris, 100 mM NaCl, 40 mM reduced glutathione, pH 7.5) prior to SDS-PAGE electrophoresis and analysed by Coomassie blue staining and immunoblotting.

Isothermal Titration Calorimetry (ITC) measurements

The nucleotide sequence encoding NXF1 amino-acids 1-198 was codon optimized for expression in *E. coli*, custom synthesized by GenScript and cloned into pET-28a(+)-TEV vector (Cloning site: NheI/XhoI) containing a N-terminal hexahistidine (6His) tag followed by a tobacco etch virus (TEV) protease cleavage site. Cloned 6His-NXF1 1-198 was transformed into *E. coli* BL21(DE3) pLysY/Iq (New England Biolabs) cells and subsequent colonies were used for the initial culture growth at 37°C in TY media supplemented with Kanamycin antibiotic resistance (50 µg/ml). Expressed cells were harvested (4000 x g) and lysed into lysis buffer consisting of 50 mM Tris (pH 7.5), 500 mM NaCl, 5 mM imidazole and 5% (v/v) glycerol and frozen at -80° C in the presence of lysozyme. Frozen cells were lysed by sonication and the soluble protein fractions were separated by ultracentrifugation (20000 x g) for 30 min at 4°C. The fractions were filtered through a 0.45 µM syringe filter (Merck Millipore Inc). The filtered supernatant was loaded onto 2 ml of Ni-NTA Agarose affinity resin (Qiagen UK). The resin than washed with 40 ml of lysis buffer and 20 ml of wash buffer (50mM Tris, pH 7.5, 500mM NaCl, 25 mM imidazole and 5% (v/v) glycerol). The proteins were eluted by an elution buffer (50 mM Tris, pH 7.5, 500 mM NaCl, 250 mM imidazole and 5% (v/v) glycerol. The protein fractions were pooled and buffer exchanged to 20 mM sodium phosphate (pH 6.5), 300 mM NaCl using a PD10 desalting column (GE Healthcare). Protein purity was checked by NuPAGE Bis-Tris (Invitrogen) and concentration was determined by absorbance at 280 nm with extinction coefficients obtained using ProtParam tool (http://web.expasy.org/protparam). The protein was concentrated using a 3000 MWCO Amicon centrifugal filter units (MerckMillipore Inc). Synthetic FAM-labelled SRSF1 peptides (purity >99%) were purchased from *ThermoFisher Scientific UK*.

For Isothermal titration calorimetry (ITC), protein samples were thoroughly dialyzed overnight into 20 mM sodium phosphate (pH 6.5), 300 mM NaCl. ITC measurements were performed on a ITC 200 microcalorimeter (Malvern Instrument) at 298K. The titration cell was filled with 25 μM NXF1 1-198 and the syringe was filled with 1 mM peptide (Table S1). After an initial delay of 360 sec, a first injection of 0.4 μl followed by 13 injections of 3 μl of peptides were carried from a rotating syringe at a speed of 750 rpm to the titrant cell containing NXF1 1-198. The delay between each injection was 180 sec and the ITC data were recorded with high sensitivity mode. The data were analysed using the *Origin* software.

Co-immunoprecipitation

Cells were split into two 10 cm plates/treatment and transfected with 15 μg p3xFLAG or p3xFLAG/NXF1. 6 h post-transfection, media was replaced with fresh warm media containing 0.5, 1 or 5 μM Ctrl-CPP or SRSF1 WT-CPP. Proteins were extracted from HEK293T cells 48 h post-transfection. Cells were washed in ice cold PBS, scraped into 500 μl ice cold lysis buffer, passed through a 25G gauge needle 10 times and left to lyse on ice for 10 min. Lysed cells were cleared by centrifugation at 17,000 *g* at 4°C for 5 min and protein extracts were quantified using Bradford Reagent. 2 mg total protein in 1 ml lysis buffer was incubated with 20 μl anti-FLAG M2 affinity gel (Sigma A2220) (which had been blocked overnight with 1% BSA in IP lysis buffer) for 2 h at 4°C on a rotating wheel. Beads were washed 5 times with lysis buffer and eluted in 50 μl IP lysis buffer supplemented with 100 $\mu g.ml^{-1}$ 3xFLAG peptide (Sigma F4799) for 30 min at 4°C on a rotating wheel. 30 μg total protein and 15 μl eluates were subjected to western immunoblotting using FLAG, SRSF1, V5 and α-tubulin antibodies (described in **Table S4**).

Western blot analysis

HEK293T cells were cultured in a 24-well plate and transfected with mammalian plasmids expressing $G4C2_{45}$-3xV5, $G2C4_{43}$-3xV5, poly GA_{36}-V5, poly GP_{36}-V5, poly GR_{36}-V5, T7-tagged SRSF1-WT or T7-tagged SRSF1-NRS. 6 h post transfection, media was replaced with fresh warm media containing 0, 0.1, 0.25, 0.5, 0.75, 1 or 2.5 μM Ctrl-CPP, m4-CPP or WT-CPP. Proteins were extracted 72 h post-transfection. For puromycin study, HEK293T cells were treated with 1 μM Ctrl-CPP or WT-CPP for 24 h, 48 h or 72 h prior to lysis. Puromycin (1 μM) was added to the cells 30 min prior to lysis. Cells were washed in ice-cold phosphate-buffered saline (PBS) and scraped into ice-cold lysis buffer (50 mM Hepes pH7.5, 150 mM NaCl, 10% glycerol, 0.5% Triton X-100, 1 mM EDTA, 1 mM DTT, protease inhibitor cocktail (Merck)). Cells were lysed on ice for 10 min followed by centrifugation at 17,000 *g* at 4°C for 5 min. Protein extracts were quantified using Bradford Reagent (BioRAD), resolved by SDS–PAGE, electroblotted onto nitrocellulose membrane and probed using the relevant primary antibody.

Cultured cortical neurons were transduced as detailed below in the immunofluorescence microscopy section. 16 h post transduction, media was replaced with 1-5 μM of Ctrl-CPP, m4-CPP or WT-CPP (Fig. 5). 72 h post-transfection, cortical neurons were washed in ice-cold PBS and scraped into ice-cold lysis as detailed above. Cells were passed through a 25G needle 10 times to improve lysis efficiency and left to lyse on ice for 10 min followed by centrifugation at 17,000 *g* at 4°C for 5 min. Protein extracts were quantified using Bradford Reagent, resolved by SDS-PAGE, electroblotted onto nitrocellulose membrane and probed using the relevant primary antibody.

Heads from *D. melanogaster* and third instar larvae bearing 3 or 36 G4C2 repeats fed throughout development with either Ctrl, m4 or WT CPP were frozen in liquid nitrogen and lysed in ice-cold

lysis buffer detailed above using Eppendorf micropestles. Lysed fly heads were incubated on ice for 10 min followed by centrifugation at 17,000 *g* for 10 min at 4 °C. Protein extracts were quantified using Bradford Reagent and subjected to dot-blot analysis. For this, 50 µg total protein extracts prepared in ice-cold lysis buffer were loaded onto a nitrocellulose membrane using a microfiltration apparatus (Biorad), sliced into strips and analysed by immunoblotting.

All antibody details are provided in **Table S4**. Uncropped western blot images can be provided on request.

MSD ELISA assays

iNeurons treated or not with CPPs (at 10 µM) or lentivirus (at MOI=5) were harvested at D10 of differentiation. Cells were lysed in RIPA buffer (2% SDS, 50 mM Tris, pH 8.0, 150 mM NaCl, 1% IGEPAL® CA-630 (Sigma), 0.5% sodium deoxycholate, Benzonase (250 units in 10 ml buffer), protease inhibitor cocktail (Merck), 2 mM PMSF). Cells were lysed at RT for 10 minutes and pass through 10 times in a syringe using a 25 G needle to disrupt cells and sheared genomic DNA. Mouse brains (approximately 0.1 g) were lysed in 400 µl reporter lysis buffer (*Promega*) supplemented with protease inhibitors and PMSF as above using a Precellys evolution homogenizer and 8x 1.4 mm Zirconium Oxide beads at 5500 rpm for 2 x 30 seconds, 10 minutes incubation on ice and another round of homogenization. Protein lysates were collected after centrifugation at 17,000 x g for 10 minutes. Protein extract concentrations were determined using a BCA assay (Bio-Rad). Same amounts of total proteins (50 µl at 1.5 mg/ml for iNeurons and 50 µl at 4 mg/ml for mouse brains) were then mixed at 1:1 ratio with EC buffer (5 mM NaH2PO4, 20mM Na2HPO4, 400 mM NaCl, 2.5 mM EDTA, 0.05% (w/v) CHAPS, 0.2% (w/v) BSA, 0.4% (w/v) Block ACE (Bio-Rad), 0.05% (w/v) NaN3) prior to be incubated in duplicates with a poly-GP capture antibody (custom synthesis, *Eurogentec*) coated on 96-well plates. A standard curve is also generated for each plate using serial dilutions of the poly-GPx7 peptide CGPGPGPGPGPGPGPGP in duplicates (0.125 to 40 ng/ml). After washing, a second biotinylated poly-GP detect antibody (generated by Eurogentec from immunization of a second rabbit) is used with a sulfo-tag streptavidin substrate that leads to generation of the electroluminescent signal read in the MesoScale Discovery instrument (following manufacturer's instructions). The duplicate absorbance readings after blank background subtraction were averaged and converted to concentration unit through the standard curve. Expression levels of poly-GP DPRs was calculated by the percentage of the C9ORF72-untreated iNeuron or mouse samples.

Nuclear/Cytoplasmic fractionation for qRT-PCR and western immunoblotting from HEK293T cells, patient-derived iNeurons and *D. melanogaster* samples (RNA nuclear export assays)

HEK293T cells were cultured in 6-well plates and transfected with pcDNA3.1/G4C2$_{45}$-3xV5 or G2C4$_{43}$-3xV5. 6 h post-transfection, the media was replaced with fresh warm media containing 0, 0.5 or 1 µM of Ctrl or WT-CPP. For patient-derived neurons, induced neural progenitor cells (iNPCs) were differentiated to iNeurons using a modified version of the protocol (*25*) as previously described (*5*) and below (see co-cultures of patient-derived astrocytes and motor neurons). 10 µM CPPs were added into differentiating iNeurons on day 7 for 72 h prior to cell harvest. Induced motor neurons (iMNs) cannot be generated in a sufficient number for this assay.

For the cytoplasmic fractionation of cells, 3 wells (HEK293T cells) or 6 wells (iNeurons) of a 6-well plate were collected in DEPC PBS using a cut tip and pelleted by centrifugation at 400 *g* for 5 min. Cell pellets were quickly washed with hypotonic lysis buffer (10 mM HEPES pH 7.9, 1.5 mM $MgCl_2$, 10 mM KCl, 0.5 mM DTT) and lysed for 10 min on ice in hypotonic lysis buffer

containing 0.16 U μl^{-1} Ribosafe RNase inhibitors (Bioline), 2 mM PMSF (Sigma) and Protease Inhibitor Cocktail tablets, EDTA free (Sigma). For the cytoplasmic fractionation of *D. melanogaster*, fifteen third instar larvae were homogenised in 400 µl of hypotonic lysis buffer containing 1 µl/ml RNase OUT recombinant ribonuclease inhibitor (Thermo Fisher Scientific), 2 mM PMSF and complete protease inhibitor cocktail tablets, EDTA free. All lysates underwent differential centrifugation at 4°C of (1,500 *g*, 3 min, 3,500 *g*, 8 min, and 17,000 *g*, 1 min) for cells and (100 *g*, 2 min, 1,500 *g*, 3 min and 17,000 *g*, 1 min) for *D. melanogaster*, transferring the supernatants to fresh tubes after each centrifugation. Nuclear pellets obtained after centrifugation at 1,500 *g* for 3 min were lysed in Reporter lysis buffer (Promega) for 10 min on ice before drawing up/down with a 25G needle ten times and centrifugation at 17,000 *g*, 5 min, 4°C.

Total fractions were collected from 1 well (HEK293T cells) or 3 wells (iNeurons) of a 6-well plate in Reporter lysis buffer containing 16 U μl^{-1} Ribosafe RNase inhibitors, 2 mM PMSF and protease inhibitors prior to lysis for 10 min on ice before centrifugation at 17,000 *g*, 5 min, 4°C. For *D. melanogaster,* total fractions were isolated from five third instar larvae in RIPA lysis buffer [50 mM Tris·HCl, 150 mM NaCl, 1 mM EDTA, 1% Triton X-100, 0.5% SDS] with 1 mM PMSF and protease inhibitor mixture (Roche).

Equal volumes of total, nuclear and cytoplasmic lysates were subjected to western immunoblotting using the nuclear markers SSRP1 (HEK293T cells and iNeurons) or histone H3 (*D.* melanogaster) and the cytoplasmic markers HSPA14 (HEK293T cells), β-Tubulin III/TUJ1 (iNeurons) or tubulin (*D. melanogaster*). Western blot with HEK293T cell extracts were additionally probed to the antibodies SRSF1 and NXF1 (**Table S4**).

Total RNA extraction and qRT-PCR quantification

HEK293T cells were cultured in 6-well plates and either transfected with SRSF1-WT/SRSF1-NRS or treated with 0 or 1 µM of Ctrl or WT-CPP for 72 h. Total fractions were collected from 1 well (HEK293T cells) of a 6-well plate in Reporter lysis buffer (Promega) containing 16 U μl^{-1} Ribosafe RNase inhibitors, 2 mM PMSF and protease inhibitors and passed through a 25G needle 10 times prior to lysis for 10 min on ice before centrifugation at 17,000 *g*, 5 min, 4°C. Total and fractionated extracts were added to PureZOL™ (BioRAD) for HEK293T cells and iNeurons or TRI Reagent LS (Sigma-Aldrich) for *D. melanogaster* to extract the RNA. Briefly, total and nuclear lysates was cleared by centrifugation for 10 min at 12,000 *g* at 4°C. One fifth the volume of chloroform was added and tubes were vigorously shaken for 15 s. After 10 min incubation at room temperature, tubes were centrifuged 12,000 *g*, 10 min, 4°C and the upper phase collected. RNA was precipitated overnight at -20°C with one tenth volume of 3M sodium acetate, 1 µl glycogen (5 µg μl^{-1}, Ambion) and equal volume isopropanol prior to be subsequently pelleted at 17,000 *g,* 20 min, 4°C. Pellets were washed with 70% DEPC ethanol and re-suspended in DEPC water. All RNA samples were treated with DNaseI (Roche) and quantified using a Nanodrop (NanoDropTechnologies).

Following quantification, 2 µg (HEK293T cells and iNeurons) or 500 ng (*D.* melanogaster) RNA was converted to cDNA using BioScript Reverse Transcriptase (Bioline) for HEK293T cells and iNeurons or Maxima H Minus cDNA Synthesis Master Mix (Thermo Fisher Scientific) for *D. melanogaster*. The sequences of human *U1 snRNA*, *CLK4*, *C9ORF72* or C9-RAN and *D. melanogaster* Tub84b and G4C2x36 primers used in this study are provided in **Table S5**. qRT-PCR reactions were performed in duplicate using the Brilliant III Ultra-Fast SYBR Green QPCR Master Mix (Agilent Technologies) on a C1000 Touch™ thermos Cycler using the CFX96™ Real-Time System (BioRAD) using an initial denaturation step, 45 cycles of amplification (95°C for 30 s; 60°C for 30 s; 72°C for 1 min) prior to recording melting curves. qRT-PCR data were analysed

using CFX Manager™ software (Version 3.1) (BioRAD) and quantification performed using the comparative C_T method and GraphPad Prism (Version 7).

MTT cell proliferation assay

HEK293T cells were split into 24-well plates (25,000 cells per well) and transfected with mammalian plasmids expressing G4C2$_{45}$-3xV5 or G2C4$_{43}$-3xV5, SRSF1 WT or SRSF1-NRS. Where indicated, 6 h post transfection, media was replaced with fresh warm media containing 0, 0.25, 0.5 or 1 μM of Ctrl-CPP, m4-CPP or WT-CPP. 72 h post transfection, 250 mg Thiazolyl Blue Tetrazolum Bromide reagent (MTT) was added to each well and incubated in the dark at 37°C for 1 h. Cells were subsequently lysed with an equal volume of MTT lysis buffer (20% SDS, 50% dimethylformamide (DMF)) and incubated in the dark, shaking at room temperature for 1 h. Absorbance at 595 nm was assessed with a PHERAstar FS (BMG Labtech). Absorbance data were retrieved using a PHERAstar MARS plate reader (BMG Labtech).

Immunofluorescence microscopy

For peptide uptake immunofluorescence assays, HEK293T cells (2,000 cells per well of a 96-well optical plate) and primary rat cortical neurons (4,500,000 neurons per 24-well plate) were cultured for 1 and 5 days respectively. For HEK cells, Ctrl-CPP, m4-CPP or WT-CPP were added to aliquots of warm media at concentrations of 0, 1, 5 or 10 μM prior to media change. Cultured cortical neurons were transduced as detailed above. 16 h post lentiviral transduction, media was replaced 5 μM Ctrl-CPP or WT-CPP custom synthesised without the V5 tag to allow for V5 immunofluorescence detection of V5-tagged DPRs (Fig. S9). Immunofluorescence staining of HEK293T cells and rat cortical neurons was performed 72 h after the addition of CPPs as described previously *(5)* with the exception that cells were blocked with 4% goat serum in PBS for 2 h at room temperature and incubated overnight at 4°C with the V5, TUJ1 or Vimentin antibodies in PBS containing 4% goat serum. Cells were washed three times with PBS containing 4% goat serum and incubated for 1 h with TUJ1 antibody in PBS containing 4% goat serum. Cells were then washed three times with PBS containing 4% goat serum and incubated for 1 h in 4% goat serum with Alexa Fluor 488 to detect V5, Alexa Fluor 647 to detect TUJ1 or Alexa Fluor 594 for Vimentin. Cells were subsequently stained with Hoechst 33342 for 10 min at room temperature, washed 3 times with PBS and imaged using either a standard light microscope (*Olympus* BX53) or the *Opera Phenix* high content screening system (Perkin Elmer) for quantification. For any well a minimum of 9 random fields using the 20x objective were captured in 3 focal planes, meaning an excess of 3000 cells in total were obtained for any well to calculate percent uptake. Quantification was assessed using the *Columbus Image* data storage and analysis system (Perkin Elmer) software, nuclei were identified using Hoechst (blue channel), and cytoplasm was masked using TUJ1 or vimentin (red channel) and peptide uptake was measured (green channel) in both nucleus and cytoplasm. For representative images of cellular distribution, cells were captured using 40x objective.

For experiments involving transduction of lentivirus expressing DPRs, primary rat cortical neurons were isolated and cultured in 24-well plates with coverslips (4,500,000 cells per plate) for 5 days. Cells were transduced as detailed above (MOI 5). 16 h post transduction, the media was replaced with fresh warm media containing 5 μM of ctrl-CPP or WT-CPP. Immunofluorescence of cortical neurons was performed 7 days (with half medium containing CPPs replaced every 48 h) after transduction as described previously (*5*). The experiments were repeated three times and

within each experiment, each condition was examined with >10 fields of view with a minimum of 100 cells. All analysis was performed blinded to experimental condition.

For immunofluorescence microscopy of mouse brain tissue, the 4 % PFA post-fixed brain was incubated in 20 % sucrose at 4°C for 24 hours prior to be incubated in 30% sucrose at 4°C for 12 hours. The resulting tissue was mounted in OCT by freezing in liquid nitrogen and iso-pentane. The brain tissue was sectioned into 8 µm thickness slices using a cryostat at -21°C and directly fixed on charged slides and left to dry for 1 hour before storage at -20°C. Prior to staining, the slides were warmed up for 30 minutes at RT and rehydrated with PBS for 5 minutes. The sections were then blocked for 1.5 hours at RT with blocking buffer (0.5% Triton X-100, 5% BSA, 2% heat-inactivated goat serum, 1% FBS in PBS). Primary anti-V5 antibody was diluted in PBS with 0.1 % Triton X-100 and 1% heat-inactivated goat serum at a 1:500 dilution prior to incubating sections overnight in primary antibody. Following day, the sections were washed 3 times (10 minutes per wash) with buffer (PBS containing 0.1% Triton X-100 and 1% heat-inactivated goat serum). Once the washing was completed, secondary anti-mouse ALEXA Fluor 488 antibody was diluted in same buffer as for primary antibody at a 1:1000 dilution. The sections were then incubated in secondary antibody for 1 hour at RT prior to washing 3 times with PBS (10 minutes per wash). After the washing step, the sections were dried, wiping excess liquid around the edge of section and one drop of Hardset Vectashield medium with DAPI was used to mount the coverslips. Coverslips were slowly lowered over the sections so that the Vectashield mounting medium was spread evenly over all the sections.

Supplementary Figures

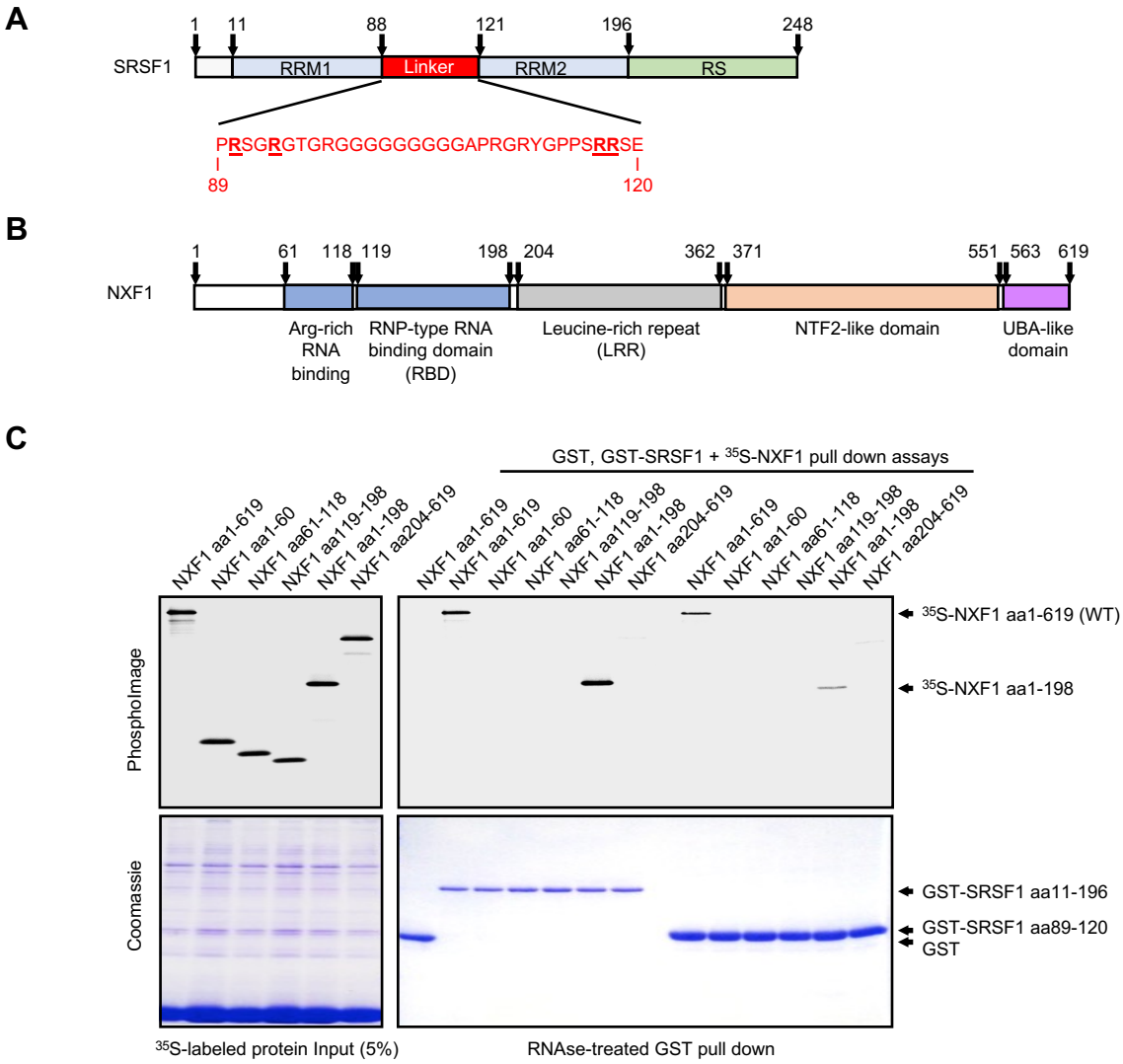


Fig. S1. The unstructured linker region of SRSF1 comprising amino-acids 89-120 interacts with the RNA binding domain of NXF1. (**A**) Schematic diagram of SRSF1 highlighting the RNA recognition motifs (RRM), the NXF1-interacting linker region and the RS domain which is rich in serine and arginine residues. Numbers depict amino-acid (aa) position. (**B**) Schematic diagram of NXF1 highlighting the RNA binding regions comprising an arginine-rich RNA-binding domain and the RNP-type RNA binding domain (aa1-198) *(13)*, a leucine-rich repeat domain (aa 204-362), a NTF2-like domain (aa 371-551) which heterodimerizes with p15/NXT1 *(46, 47)* and a UBA-like domain (aa 563-619) which associates with nucleoporins together with the NTF2-like scaffold *(48)*. The p15 protein heterodimerizes and stabilises NXF1 to stimulate the NXF1-dependent RNA nuclear export *(47)*. It will therefore be co-expressed with NXF1 in the *in vitro* GST-NXF1 pull down assays performed in this study. (**C**) Recombinant bacterially-expressed GST and GST-NXF1:p15 pull down assays were performed in presence of RNase and mammalian recombinant ^{35}S-labelled NXF1 protein full length (WT) or various domains expressed in rabbit reticulocytes.

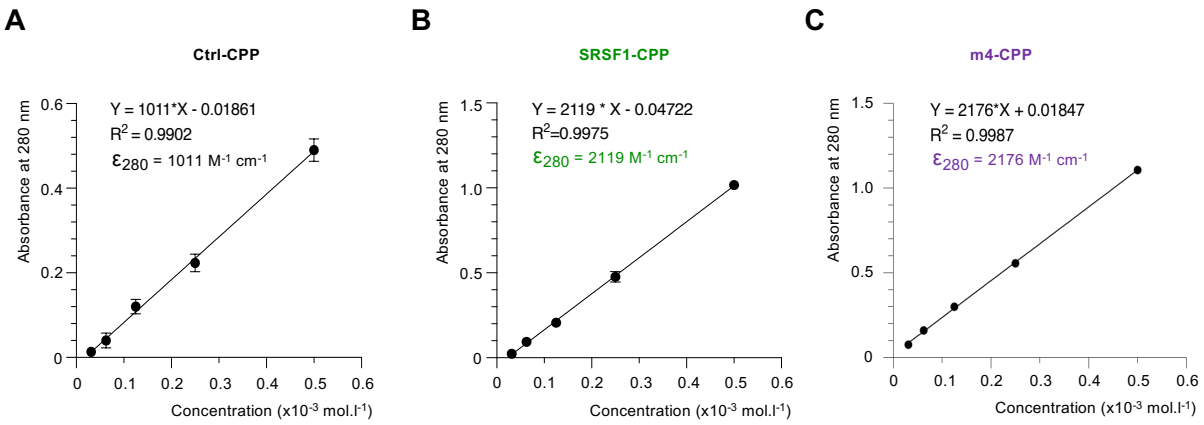


Fig. S2. Quantification of cell permeable peptides. The cell permeable peptides were custom synthesized by *ThermoFisher Scientifc* at >90% purity and resuspended in PBS at 1 mM. Absorbance of 1:2 serial dilutions of the CPPs (0.5 to 0.03125 mM) was measured at 280 nm in triplicate. According to the Beer-Lambert Law A = ε x l x C (A: Absorbance, ε: molar extinction coefficient, l: optical path, C: concentration), the slope of the graph A = f(C) corresponds to ε x l where l = 1 cm. Error bars represent standard deviation. The experimental molar extinction coefficients were determined in PBS for Ctrl-CPP (**A**, ε_{280} = 1011 M^{-1} cm^{-1}), WT-CPP (**B**, ε_{280} = 2119 M^{-1} cm^{-1}) and m4-CPP (**C**, ε_{280} = 2176 M^{-1} cm^{-1}). In comparison, the theoretical value computed in water (https://web.expasy.org/protparam/) were respectively 1490, 2980 and 2980 M^{-1} cm^{-1}. We used the determined experimental molar extinction coefficients and measured the absorbance at 280 nm of stock CPP solutions to accurately quantify their concentrations across batches.

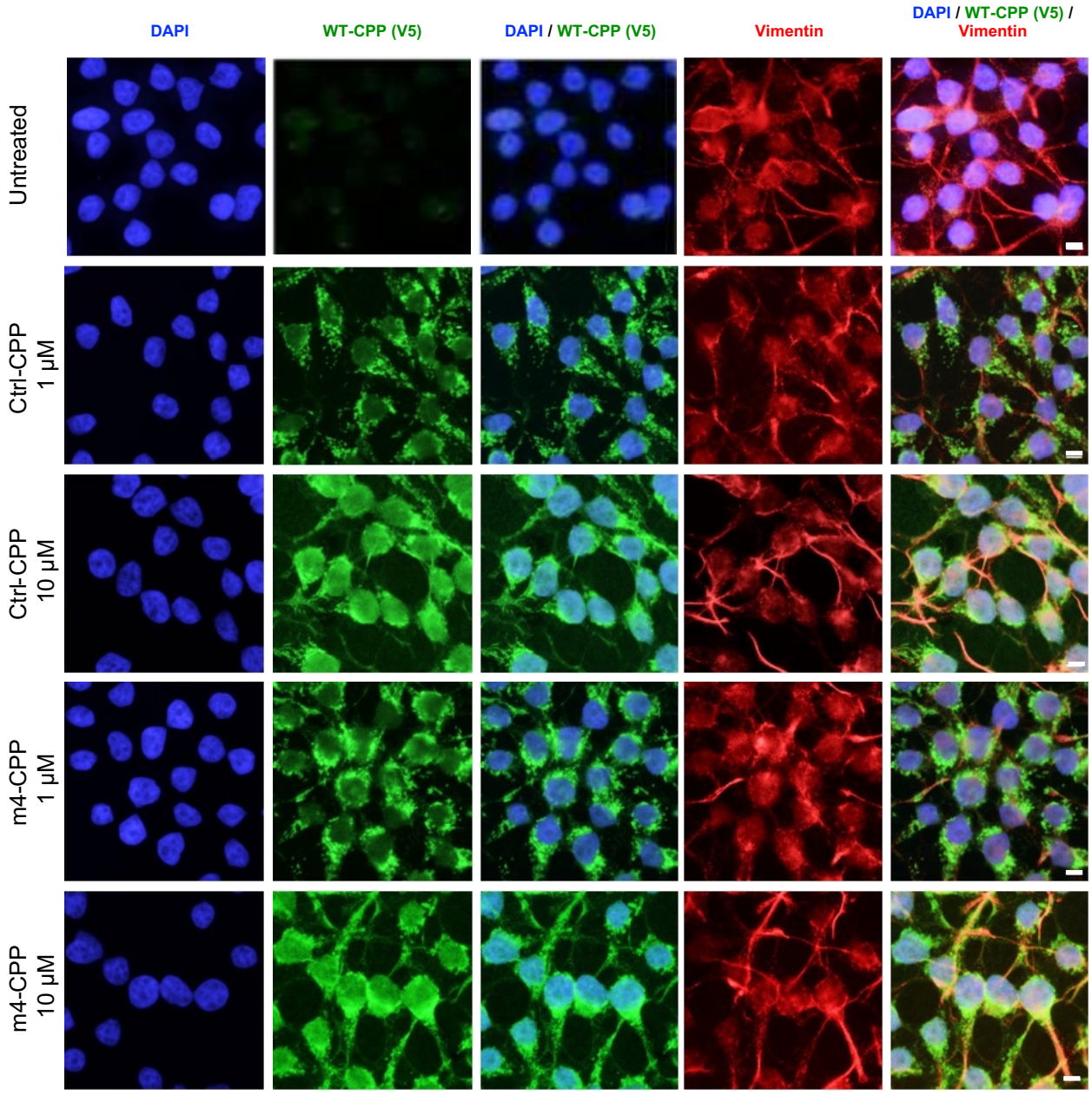
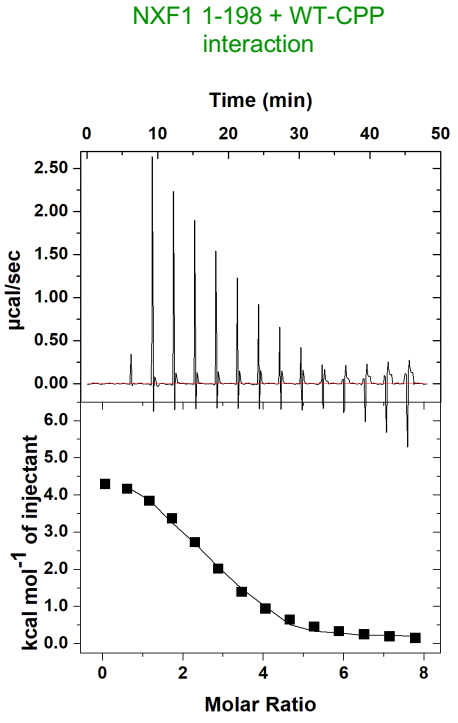


Fig. S3. Ctrl- and m4- CPPs are delivered intracellularly into HEK293T cells.
Immunofluorescence imaging microscopy of HEK293T cells cultured in media supplemented with or without CPPs for 72 hours. Ctrl- and m4-CPPs were detected in the green channel using anti-V5 staining. DAPI was used to delineate the nucleus while vimentin staining in the red channel was used to delineate the cytoplasm. Scale bar: 20 µm.

A

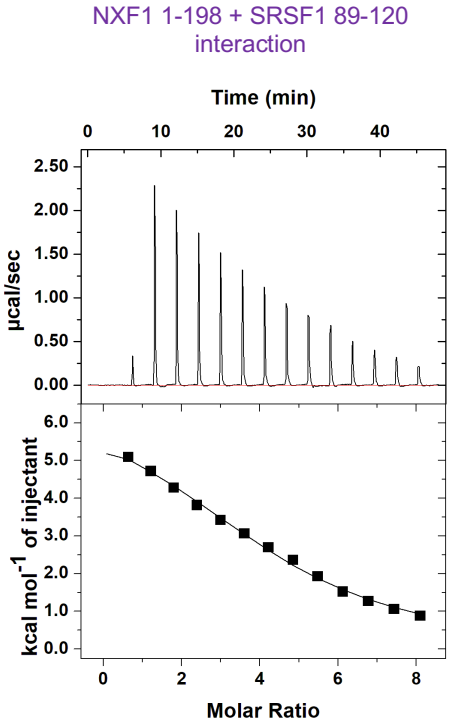
	SRSF1 89-120 (NXF1-binding site)	V5 tag	TAT-PTD
WT-CPP:	PRSGRGTGRGGGGGGGGGAPRGRYGPPSRRSE	GGKPIPNPLLGLDST	GG YGRKKRRQRRR
SRSF1 aa89-120:	PRSGRGTGRGGGGGGGGGAPRGRYGPPSRRSE		

B



$K_d = 11 \pm 3$ (µM)

C



$K_d = 52 \pm 7$ (µM)

Fig. S4. Affinity of NXF1 for WT-CPP and the minimal SRSF1-binding domain. (**A**) Minimal SRSF1 wild type (wt) interacting with the SRSF1-binding site of NXF1 (NXF1 aa 1-198). Underlined arginines 90, 93, 117, 118 within the SRSF1 amino acid 89-120 sequence are directly involved in the interactions with NXF1 (*14*). The presence of additional arginines in WT-CPP is highlighted in green in the TAT-PTD domain. (**B-C**) Affinity of the interactions between SRSF1 WT-CPP and the minimal binding site required for the interaction of SRSF1 with NXF1 and was measured using isothermal titration calorimetry (ITC) assays. Purified recombinant NXF1 aa 1-198, which corresponds to the SRSF1-binding site, was incubated with successive injections of WT-CPP (B) or SRSF1 amino acid 89-120 (C) peptides and the heat released from the interaction was plotted for each injection over time (top) and the molar ratio of injected peptide (bottom). Thermodynamic parameters are reported in Table S1.

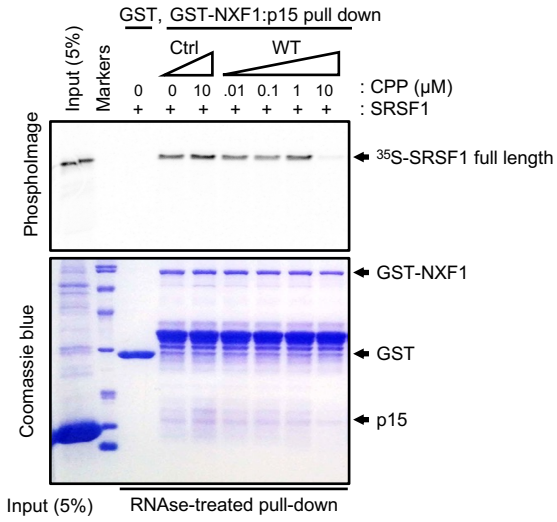


Fig. S5. SRSF1 WT-CPP competes with the interaction of SRSF1 full length and NXF1. Recombinant bacterially-expressed GST and GST-NXF1:p15 pull down assays were performed using ^{35}S-radiolabelled mammalian recombinant SRSF1 synthesised in rabbit reticulocytes in presence of increasing concentrations of Ctrl- or WT- CPPs.

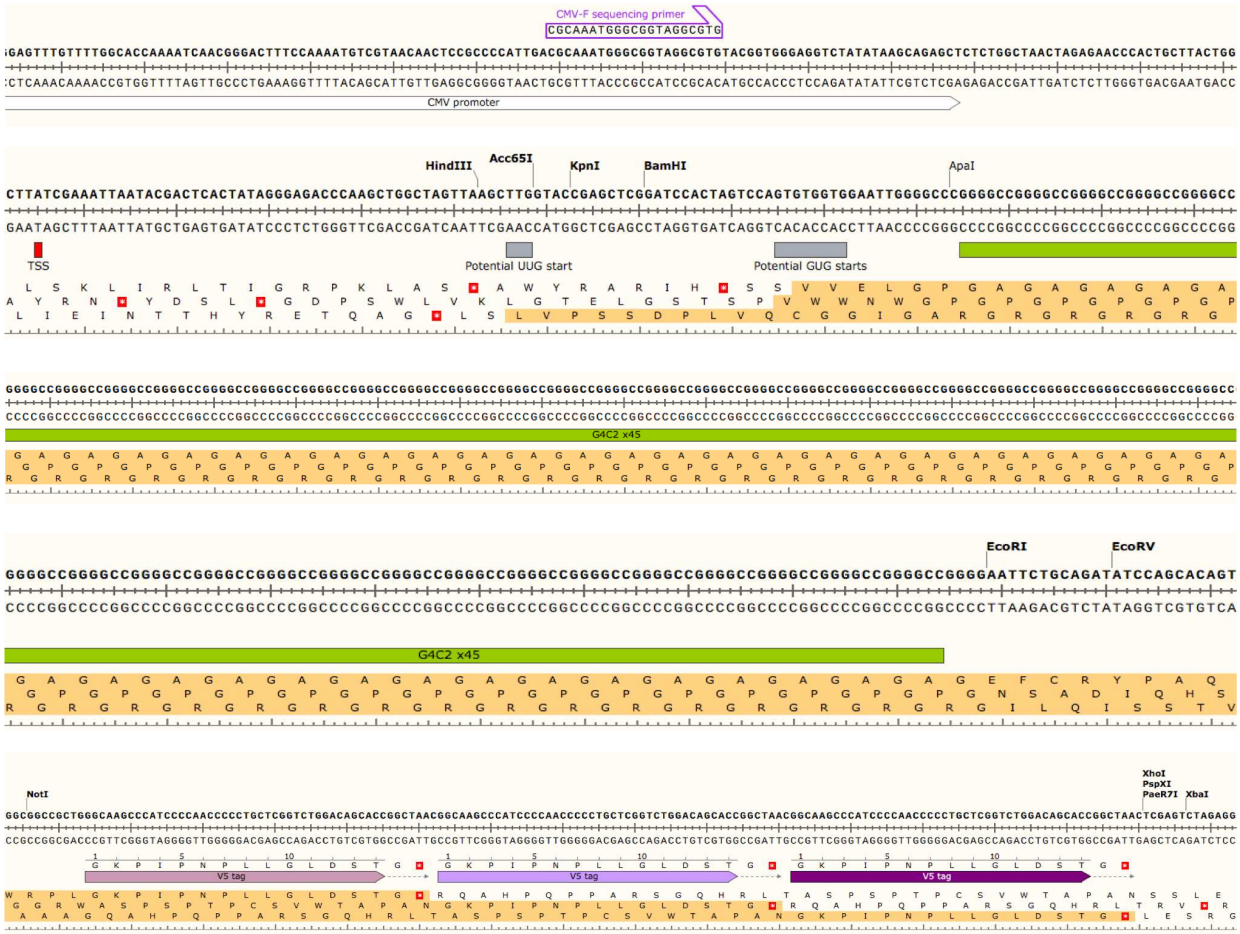


Fig. S6. Engineering RAN-dependent G4C2x45 repeat transcripts with 3xV5 tags in all open reading frames. Concatemerized and annealed G4C2/G2C4x15 DNA oligonucleotides blunted with Mung bean nuclease were cloned into the Klenow-filled site EcoRI of pcDNA3.1 to build pcDNA3.1-G4C2x45. A NotI/XbaI cassette encoding 3xV5 tags and 3 stop codons in all frames was digested from a synthetic custom-synthesized plasmid (*ThermoFisher Scientific*) and subcloned into the NotI/XbaI site of pcDNA3.1-G4C2x45. Sanger sequencing is available on request. Note the absence of canonical AUG start codons from the transcription start site (TSS) to investigate RAN-dependent translation of sense *C9ORF72*-repeat transcripts. The one letter amino-acid code sequences highlighted in orange indicate the DPRs respectively produced in the 3 reading frames.

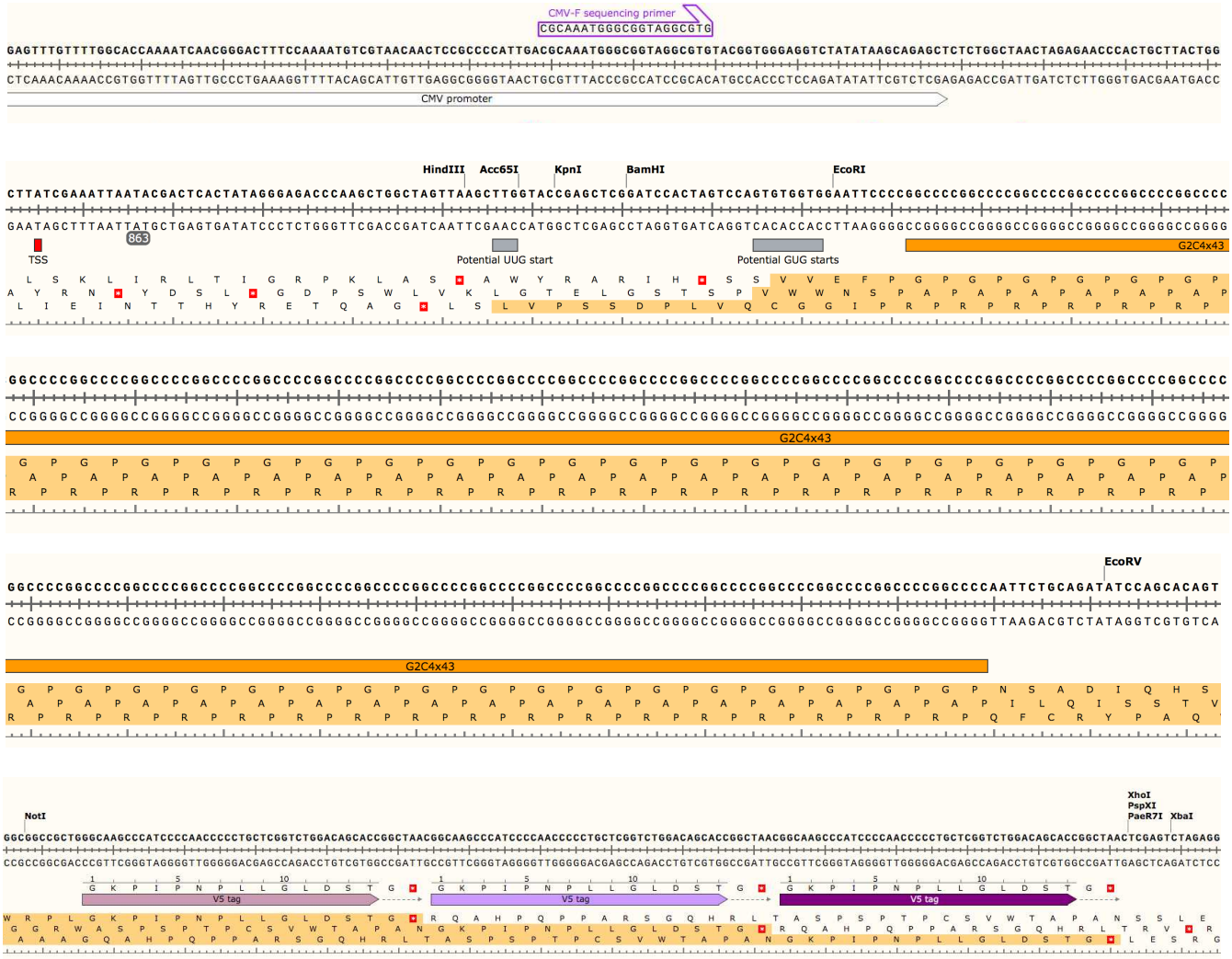


Fig. S7. Engineering RAN-dependent G2C4x43 repeat transcripts with 3xV5 tags in all open reading frames. Concatemerized and annealed G4C2/G2C4x15 DNA oligonucleotides blunted with Mung bean nuclease were cloned into the Klenow-filled site EcoRI of pcDNA3.1 to build pcDNA3.1-G2C4x43. A NotI/XbaI cassette encoding 3xV5 tags and 3 stop codons in all frames was digested from a synthetic custom-synthesized plasmid (*ThermoFisher Scientific*) and subcloned into the NotI/XbaI site of pcDNA3.1-G2C4x43. Sanger sequencing is available on request. Note the absence of canonical AUG start codons from the transcription start site (TSS) to investigate RAN-dependent translation of sense *C9ORF72*-repeat transcripts. The one letter amino-acid code sequences highlighted in orange indicate the DPRs respectively produced in the 3 reading frames.

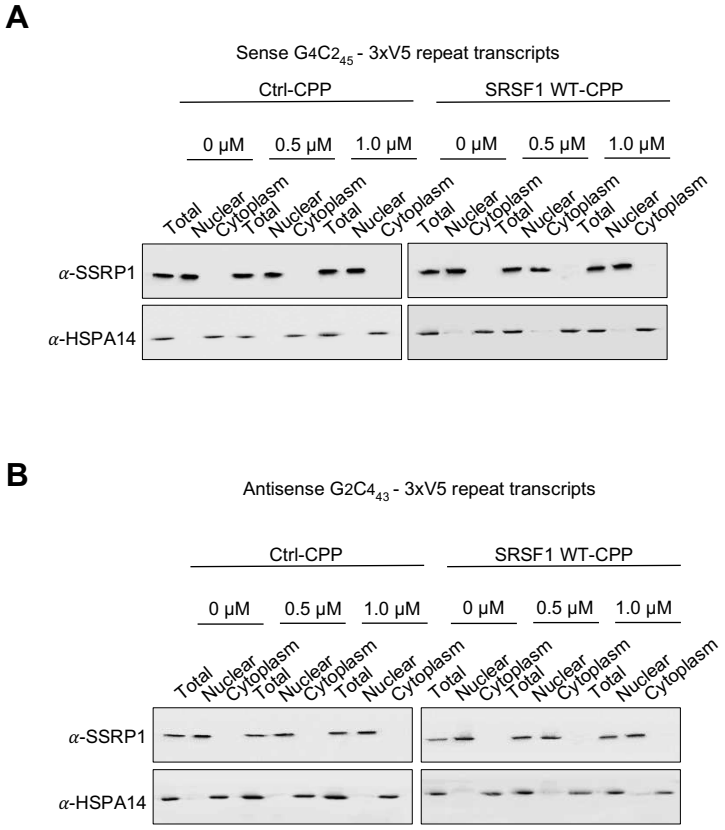


Fig. S8. Preparation of total, nuclear and cytoplasmic fractions for quantifying the nuclear export of reporter repeat transcripts in human cell models of C9ORF72-ALS. (**A, B**) Western blots of total, nuclear and cytoplasmic fractions isolated from HEK293T cells transfected for 72h with sense G4C2$_{45}$-3xV5 (A) or antisense G2C4$_{43}$-3xV5 (B) plasmids and treated with increasing concentrations of Ctrl or SRSF1 WT CPPs were probed with nuclear chromatin-associated SSRP1 factor and predominantly cytoplasmic heat-shock HSPA14 protein.

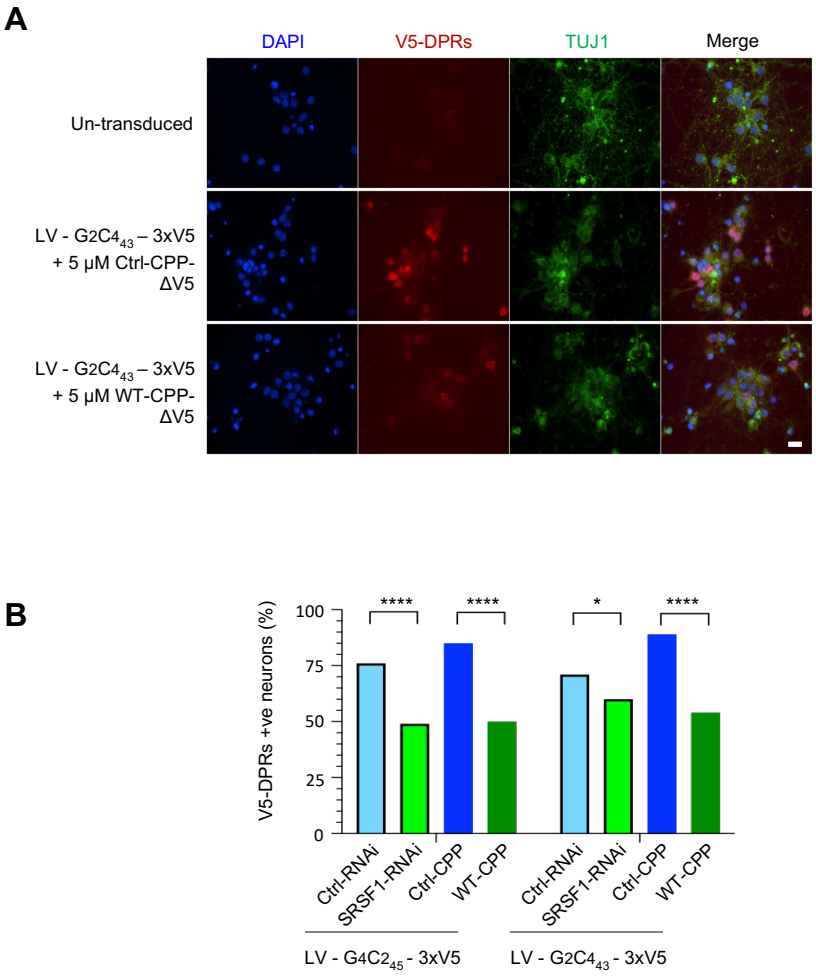


Fig. S9. SRSF1 WT-CPP inhibits RAN-translation of DPRs in primary rat cortical neurons. (**A**) Rat cortical primary neurons transduced or not with a lentivirus expressing G2C4$_{43}$-3xV5 repeat transcripts were treated or not with 5 µM Ctrl-CPP or WT-CPP lacking V5 tags (ΔV5) prior to V5 (in the red channel) and TUJ1 (in the green channel) immunofluorescence microscopy. DAPI was used to stain the nuclei. Data are shown for neurons transduced with virus expressing antisense DPRs. (**B**) Same experiments as in panel A but also including lentivirus expressing G4C2$_{45}$-3xV5 repeat transcripts, Ctrl-RNAi or SRSF1-RNAi. For each condition, V5-DPRs positive neurons were counted blinded in at least 10 microscopic fields for at least 100 neurons in biological triplicate experiments prior to use a Fisher's exact test to assess statistical significance of either SRSF1 depletion or WT-CPP treatment. Ctrl-RNAi versus SRSF1-RNAi conditions respectively generated p-values of 2.3E-05 (****) and 0.02 (*) for transductions with sense and antisense repeat transcripts. Ctrl-CPP versus SRSF1 WT-CPP conditions respectively provided p-values of 9.5E-09 (****) and 1.8E-11 (****) for transductions with sense and antisense repeat transcripts. Scale bar: 5 µm.

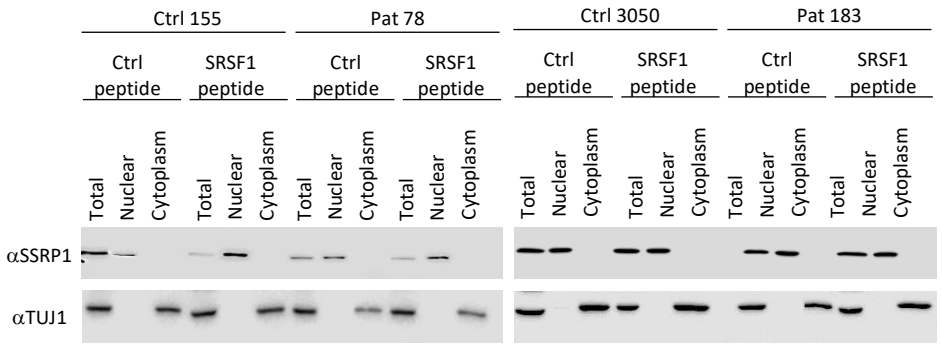


Fig. S10. Preparation of total, nuclear and cytoplasmic fractions for quantifying the nuclear export of *C9ORF72* transcripts in patient-derived neurons. Western blots of total, nuclear and cytoplasmic fractions isolated from patient-derived neurons treated for 72 h with 10 μM of Ctrl-CPP or SRSF1 WT-CPP were probed with nuclear chromatin-associated SSRP1 factor and cytoplasmic pan neuronal marker beat III tubulin/TUJ1.

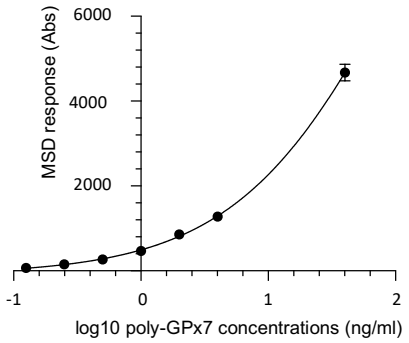


Fig. S11. Standard curve used for the quantification of poly-GP DPRs by MSD ELISA assays. Two sets of antibodies were generated in rabbits by *Eurogentec* following immunisation with a poly-GP7 peptide (CGPGPGPGPGPGPGP) coupled to KLH (Keyhole limpet hemocyanin) via the amino-terminal cysteine. Serial dilutions of the poly-GPx7 peptide (0.125 up to 40 ng/ml) were also used to produce a standard curve that allows extrapolation of the experimental poly-GP DPR concentrations quantified from patient-derived neurons and mouse tissue. A blank correction has been applied to retain the specific signal. The standard curve is performed in duplicate for each MSD ELISA assay. Bar chart represents mean MSD response ± SD.

Fig. S12. Inter-species conservation of NXF1 proteins. The primary sequence of the human NXF1 (hNXF1) protein was aligned with *Drosophila* and mouse NXF1 proteins (dNXF1 and

mNXF1 respectively) using the CLUSTAL Omega (1.2.4) multiple sequence alignment tool with the pre-set parameters of the EMBL-EBI server (https://www.ebi.ac.uk/Tools/msa/clustalo/). An "*" (asterisk) indicates positions which have a single fully conserved residue. A ":" (colon) indicates conservation between groups of strongly similar properties. A "." (period) indicates conservation between groups of weakly similar properties. 389 amino-acids are conserved among the 619 amino-acids of the full length human NXF1 protein (62.8%). Amino-acids, represented by the single-letter code, are also coloured based on their physicochemical properties. Acidic and basic residues are respectively highlighted in blue and magenta. Red residues correspond to small and hydrophobic, including the aromatic Y, amino-acids. Green labels depict hydroxyl, sulfhydryl, amine and G residues. It is noteworthy that patches of negatively-charged D/E amino-acids which are expected to interact with the positively charged arginine residues of the SRSF1-CPP are conserved across species within the SRSF1-binding site of NXF1 (aa 1-198).

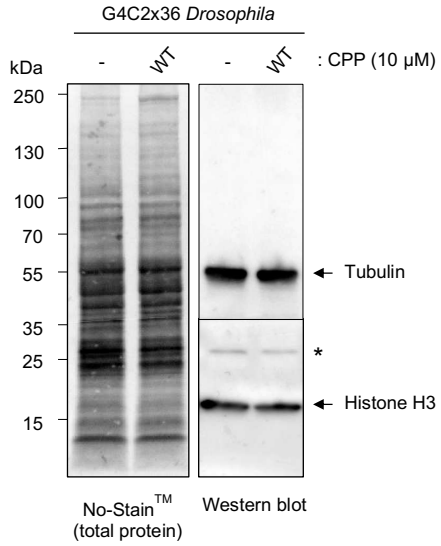


Fig. S13. SRSF1 WT-CPP does not affect protein steady-state levels in *Drosophila*. Total protein (30 μg) heads homogenates from adult flies fed with food supplemented with 10 μ M SRSF1 WT-CPP or not (-) were loaded onto SDS-PAGE prior to staining of total proteins with *No-Stain™* or immunoblotting with anti-tubulin and anti-histone H3 antibodies.

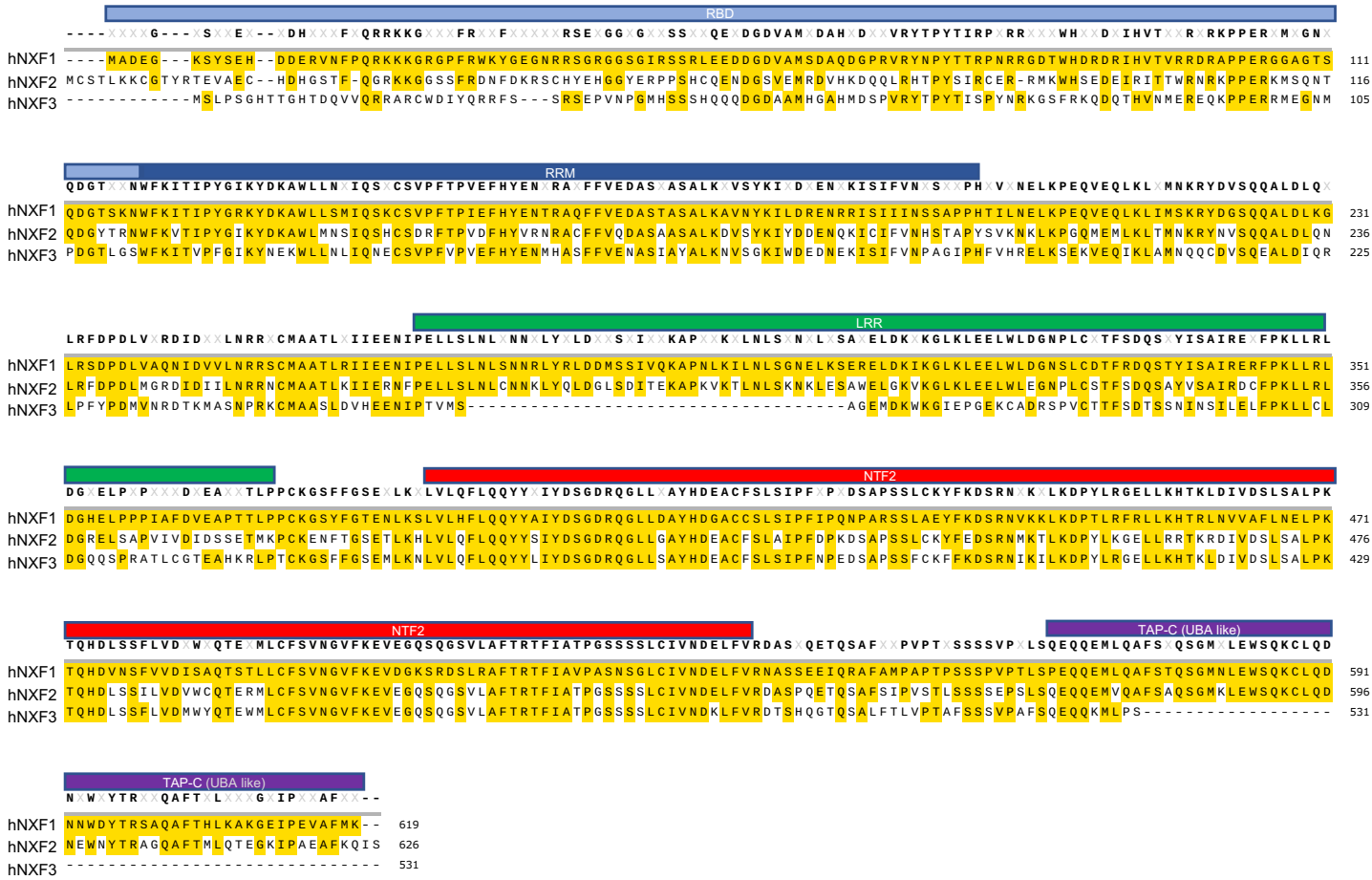


Fig. S14. Protein sequence alignment of human Nuclear RNA Export Factors NXF1-3. NXF1 (NCBI accession NM_006362.4), NXF2 (NCBI accession NM_022053.3) and NXF3 (NCBI accession NM_022052.1) protein sequences were aligned using the CLUSTAL Omega (1.2.4) multiple sequence alignment tool with the pre-set parameters of the EMBL-EBI server (https://www.ebi.ac.uk/Tools/msa/clustalo/) and the software *Snapgene* for the superposition of protein domains. Yellow boxes indicate identical single letter-coded amino acids. RBD: RNA-binding domain; ψRRM: pseudo RNA recognition motif; LRR: Leucine rich-repeat domain; NTF2-like: nuclear transport factor 2 like domain; TAP-C/ UBA-like: ubiquitin-associated–like domain. The NXF3 protein lacks the carboxyl-terminal UBA-like domain.

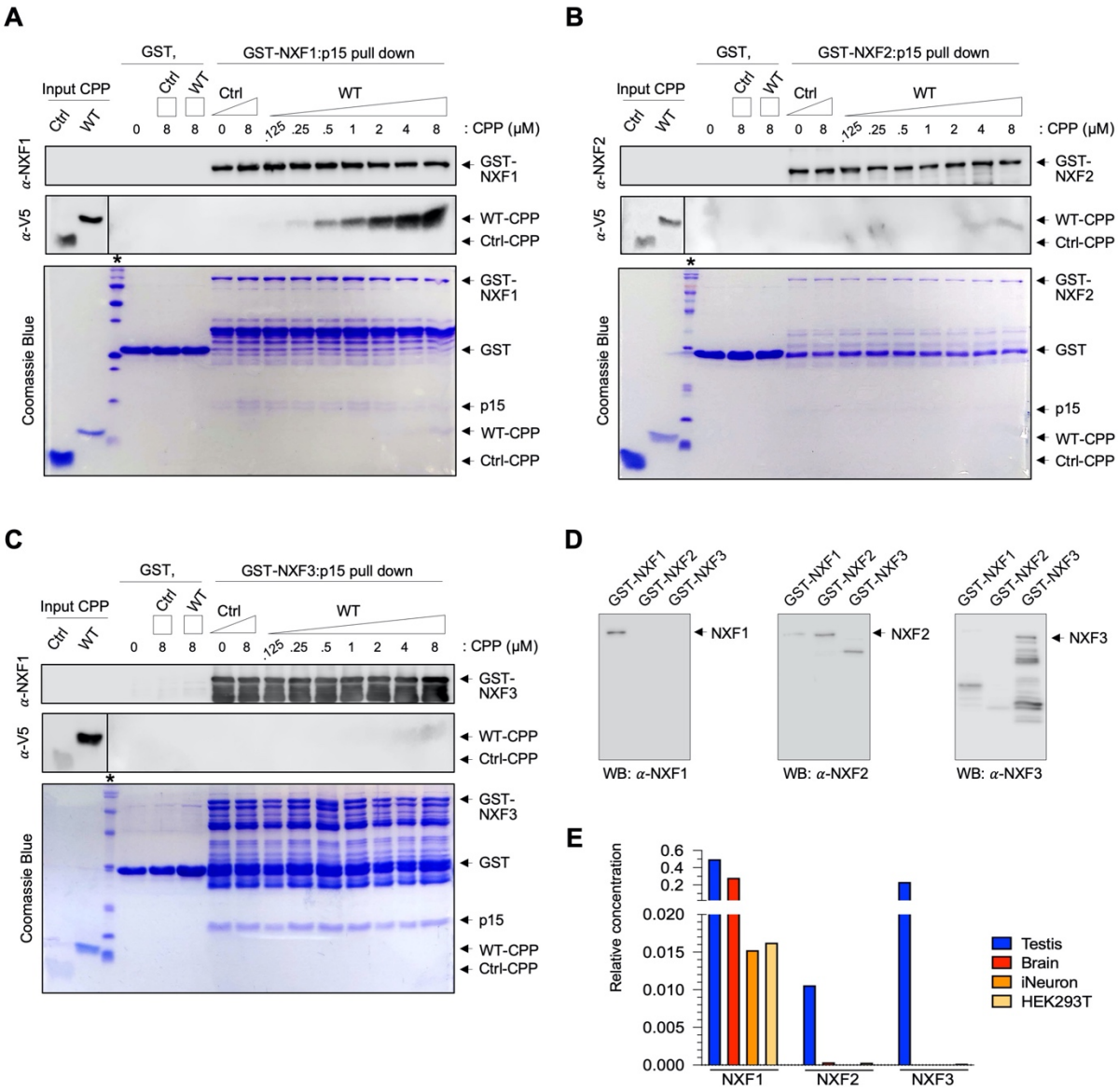


Fig. S15. WT-CPP only weakly interacts with the NXF1-related proteins, NXF2 and NFX3, which are not expressed in HEK293T, human-derived iNeurons and mouse brains. (**A-C**) Pull down assays using recombinant GST, GST-NXF1:p15 (A), GST-NXF2:p15 (B) and GST-NXF3:p15 (C) fusions with increasing concentrations of Ctrl-CPP, WT-CPP and m4-CPP. Binding reactions were analyzed by SDS-PAGE stained with Coomassie blue and western immunoblotting using anti-V5 and anti-NXF1 antibodies. Recombinant GST-NXF2 expression is lower while GST-NXF3 shows multiple GST-truncated protein products. The asterisk indicates the protein marker lane. (**D**) Purified recombinant GST-NXF1-3:15 proteins were analyzed by western immunoblotting using anti-NXF1, NXF2 and NXF3 antibodies. (**E**) The expression of *NXF1*, *NXF2* and *NXF3* mRNAs was investigated in mouse testis (positive control), mouse cortex homogenate, human-derived neurons and HEK293T cells using qRT-PCR. Chart shows relative concentrations of NXF1-3 mRNAs normalized to *U1 snRNAs*. Note that primers are different for mouse and human samples, and direct comparisons cannot be extrapolated.

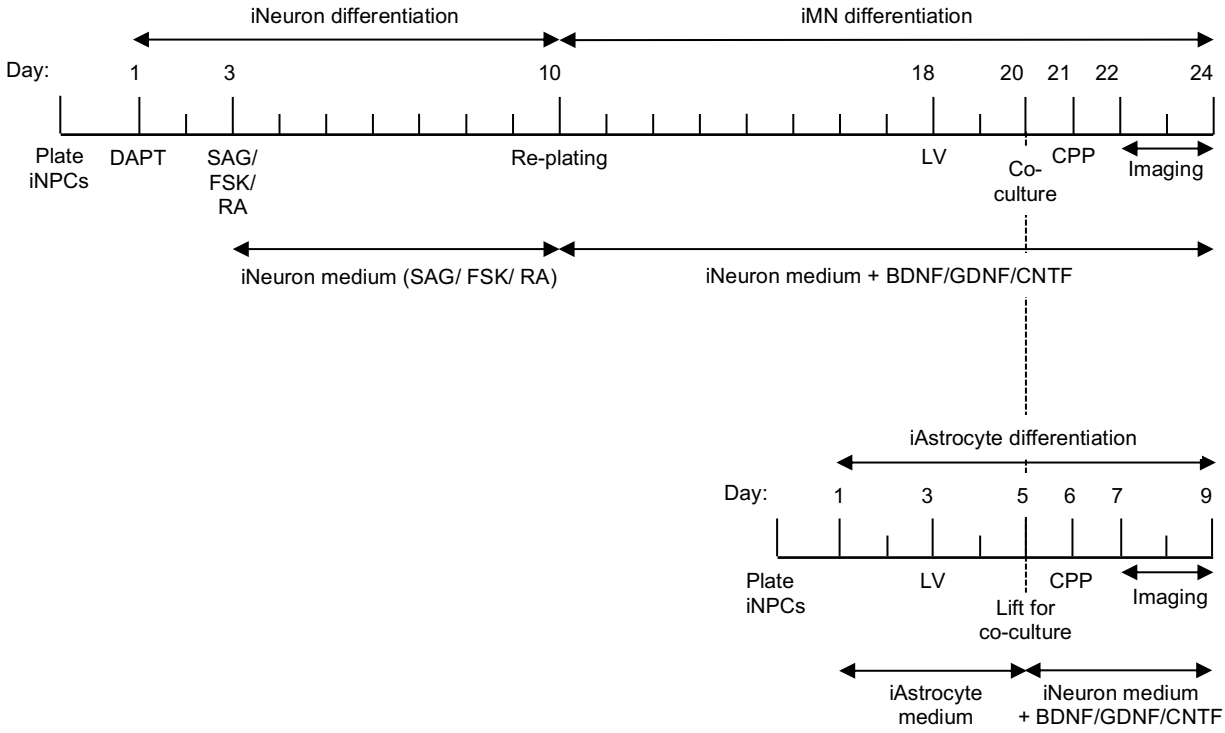


Fig. S16. Schematic representing the timelines for co-cultures of patient-derived iMNs and iAstrocytes.

Table S1. Thermodynamic parameters of NXF1 binding to SRSF1 WT and mutant peptides. Dissociation constants (K_d) and thermodynamic parameters were derived from ITC data (Fig. S4) for NXF1 (amino acid 1-198) in complex with WT-CPP ([*]) or SRSF1 amino acid 89-120 ([$]).

Protein	Peptide	K_d (μM)	ΔH (kcal/mol)	ΔG (kcal/mol)	-TΔS (kcal/mol)
$NXF1_{(1-198)}$	SRSF1-V5-TAT[*]	11 ± 3	4.75 ± 0.20	-6.77 ± 0.11	-11.51 ± 0.22
$NXF1_{(1-198)}$	SRSF1[$]	52 ± 7	6.76 ± 0.27	-5.84 ± 0.06	-12.60 ± 0.27

[*] PRSGRGTGRGGGGGGGGGAPRGRYGPPSRRSEGGKPIPNPLLGLDSTGGYGRKKRRQRRR
[$] PRSGRGTGRGGGGGGGGGAPRGRYGPPSRRSE

Table S2. List and characteristics of patient-derived lines used in this study.

Patient cell lines	Ethnicity	Gender	Cell type	Age at biopsy sample collection (years)
Ctrl 3050	Caucasian	Male	Healthy control	68
Ctrl 155	Caucasian	Male	Healthy control	40
Ctrl AG8620	Caucasian	Female	Healthy control	64
C9-ALS p78	Caucasian	Male	C9ORF72-ALS	66
C9-ALS p183	Caucasian	Male	C9ORF72-ALS	49
C9-ALS p201	Caucasian	Female	C9ORF72-ALS	66

Table S3. List of plasmids used in this study.

Plasmid name	Description	Tag(s)	Source
pcDNA3.1-G4C2x45-3xV5	Described in Fig. S6	3 x V5	This study
pcDNA3.1-G2C4x43-3xV5	Described in Fig. S7	3 x V5	This study
pcDNA3.1-polyGAx36-V5	Codon-optimized poly-GAx36 cloned as BamHI/NotI fragment into pCI-Neo-V5-N	V5	This study
pcDNA3.1-polyGPx36-V5	Codon-optimized poly-GPx36 cloned as BamHI/NotI fragment into pCI-Neo-V5-N	V5	This study
pcDNA3.1-polyGRx36-V5	Codon-optimized poly-GRx36 cloned as BamHI/NotI fragment into pCI-Neo-V5-N	V5	This study
pET9a-p15	Human *p15/NXT1* ORF cloned as NdeI/BamHI PCR fragment into pET9a	None	Ref. 13
pET24b-NXF1 1-619 (WT)	Human *NXF1* ORF cloned as NdeI/SalI PCR fragment into pET24b-GB1-6His	GB1+6His (5')	Ref. 13
pET24b-NXF1 1-60	Human *NXF1* codon 1-60 cloned as NdeI/XhoI PCR fragment into pET24b-GB1-6His	GB1+6His (5')	Ref. 13
pET24b-NXF1 61-118	Human *NXF1* codon 61-118 cloned as NdeI/XhoI PCR fragment into pET24b-GB1-6His	GB1+6His (5')	Ref. 13
pET24b-NXF1 119-198	Human *NXF1* codon 119-198 cloned as NdeI/XhoI PCR fragment into pET24b-GB1-6His	GB1+6His (5')	Ref. 13
pET24b-NXF1 1-198	Human *NXF1* codon 1-198 cloned as NdeI/XhoI PCR fragment into pET24b-GB1-6His	GB1+6His (5')	Ref. 13
pET24b-NXF1 204-619	Human *NXF1* codon 204-619 cloned as NdeI/SalI PCR fragment into pET24b-GB1-6His	GB1+6His (5')	Ref. 13
pET24b-SRSF1 1-248 (WT)	Human *SRSF1* ORF cloned as NdeI/XhoI PCR fragment into pET24b	6 x His	Ref. 14
pET24b-SRSF1 11-196	Human *SRSF1* codon 11-196 cloned as NdeI/XhoI PCR fragment into pET24b	6 x His	Ref. 14
pGEX6P1		GST	GE Healthcare
pGEX6P1-NXF1	Human *NXF1* ORF cloned as BamHI/NotI PCR fragment into pGEX6P1	GST (5')	This study
pGEX6P1-NXF2	Human *NXF2* ORF cloned as BamHI/XhoI PCR fragment into pGEX6P1	GST (5')	Prof SA Wilson
pGEX6P1-NXF3	Human *NXF3* ORF cloned as BamHI/XhoI PCR fragment into pGEX6P1	GST (5')	Prof SA Wilson
pGEX6P1-SRSF1 11-196	Human *SRSF1* codon 11-196 cloned as BamHI/XhoI PCR fragment into pGEX6P1	GST (5')	This study
pGEX6P1-SRSF1 89-120	Human *SRSF1* codon 89-120 cloned as BamHI/XhoI oligonucleotides into pGEX6P1	GST (5')	Ref. 14
p3X-FLAG-NXF1	Human *NXF1* ORF cloned as EcoRI/XbaI PCR fragment into p3x-FLAG	3x FLAG (5'). + Myc (3')	Ref .13
SRSF1-WT	Engineered in Prof J. Cáceres's group	T7	Ref. 22
SRSF1-NRS	Engineered in Prof J. Cáceres's group	T7	Ref. 22

Table S4. List of antibodies used in this study.

Antibody	Host Species	Dilution	Supplier
HSPA14	Rabbit	1:2,000	Abcam #ab108612
MAP2	Guinea pig	1:1,000	Synaptic Systems #188004
NeuN	Mouse	1:1,000	Biolegend #834502
NXF1	Mouse	1:2,000	Abcam #ab50609
NXF2	Rabbit	1:2,000	Proteintech #12185-1-AP
NXF3	Rabbit	1:500	Proteintech #13275-1-AP
SRSF1	Rabbit	1:1,000	Cell signaling #14902
SRPK1	Rabbit	1:2,000	Proteintech #14073-1-AP
SSRP1	Mouse	1:500	Abcam #ab26212
Phospho-Ser 409 TDP-43	Rabbit	1:1,000	Affinity Biosciences #AF7365
Phosphoepitope SR proteins (clone 1H4)	Mouse	1:1,000	Sigma MABE50
α-tubulin (clone DM1A)	Mouse	1:2,000	Insight #Sc32293
Class III β-tubulin (TUJ1)	Chicken	1:1,000 WB 1:1,000 IF	Millipore #AB9354
Actin (*D. melanogaster*)	Mouse	1:5,000	Abcam #ab8224
Histone H3 (*D. melanogaster*)	Rabbit	1:5,000	Abcam #ab1791
α-tubulin (*D. melanogaster*)	Mouse	1:5,000	Sigma-Aldrich #T9026
poly-Gly-Pro	Rabbit	1:2,000	Proteintech #24494-1-AP
poly-Gly-Arg	Rabbit	1:2,000	Proteintech #23978-1-AP
poly-Pro-Arg	Rabbit	1:2,000	Proteintech #23979-1-AP
poly-Pro-Ala	Rat	1:1,000	Merck clone 14E2 #MABN1790
poly-Gly-Ala	Mouse	1:500	kindly provided from Prof Dieter Edbauer
FLAG M2	Mouse	1:2,000	Sigma #F1804
Puromycin	Mouse	1:20,000	Merck MABE343
Vimentin	Chicken	1:5,000	
V5 clone D3H8Q [cortical neurons]	Rabbit	1:500	Cell Signaling #13202S
V5 [other experiments]	Mouse	1:5,000 WB 1:1,000 IF	Invitrogen #R96025
HRP-conjugated rabbit	Goat	1:5,000	Promega #W4021
HRP-conjugated mouse	Goat	1:5,000	Promega #W4011
HRP-conjugated chicken	Rabbit	1:5,000	Promega #G1351
HRP-conjugated rabbit (used in *D. melanogaster* assays)	Goat	1:5,000	Invitrogen #G-21234
HRP-conjugated mouse (used in *D. melanogaster* assays)	Goat	1:5,000	Abcam #ab6789
Anti-mouse Alexa Fluor 488	Goat	1:1,000	Invitrogen #A32723
Anti-chicken Alexa Fluor 647	Goat	1:1,000	Invitrogen #A32933

Table S5. List of qPCR primers used in this study.

Gene target	Sequences	Origin
Human *U1 snRNA*	Fwd: 5'-CCATGATCACGAAGGTGGTT-3' Rev: 5'-ATGCAGTCGAGTTTCCCACA-3'	Reference [5]
Mouse *U1 snRNA*	Fwd: 5'-CCATGATCACGAAGGTGGTT-3' Rev: 5'-CCACTACCACAAATTATGCAG-3'	This study
C9-RAN reporter	Fwd: 5'-GGGCCCTTCGAACCCCCGTC-3' Rev: 5'GGGAGGGGCAAACAACAGAT-3'	Reference [5]
Human *C9ORF72*	Exon-1 Fwd: 5'-TCAAACAGCGACAAGTTCCG-3' Exon-2 Rev: 5'-GTCGACATCACTGCATTCCA-3' Intron-1 Rev: 5'-GGAGAGAGGGTGGGAAAAAC-3'	Reference [5]
Human *CLK4*	Exon-2 Fwd: 5'-ATTCCAAAAGAACTCACTGTCC-3' Exon-2 Rev: 5'-ACTGGTGATGTGGTTTACAAGC-3' Exon-4+5 Fwd: 5'-AGATCCAGGAGTATAGAGG-3' Exon-4+5 Rev: 5'- TTCACCCAAAGTGTCCACG-3'	Reference [20]
Human *NXF1*	Fwd: 5'-CGTTGTCCTGAATCGCAGAAGC-3' Rev: 5'-GTTGGGTGCCTTCTGAACAATGC-3'	*Origene*
Human *NXF2*	Fwd: 5'-CTGCTTCTCCTTGGCTATTCCC-3' Rev: 5'-CGTTTTGTGCGCCTCAGCAGTT-3'	*Origene*
Human *NXF3*	Fwd: 5'-CCTTCTCGGATACCTCCAGCAA-3' Rev: 5'-GGTTGGTAACCTCTTGTGGGCT-3'	*Origene*
Mouse *NXF1*	Fwd: 5'-CCGAGAACCTAAAGAGTCTGGTC-3' Rev: 5'-TCGGACAGGATTCTGAGGGTTG-3'	*Origene*
Mouse *NXF2*	Fwd: 5'-CCAGAAAGCTCTGGACCTCGAA-3' Rev: 5'-CTGTAGTGTGGCAACCATGCAG-3'	*Origene*
Mouse *NXF3*	Fwd: 5'-ACCAACTCTCTGTGGTCCTGAAG-3' Rev: 5'-CTGCTTCAGGAACTGTAGGACC-3'	*Origene*
D. melanogaster Tub84b	Fwd: 5'-CTTCCTCATCTTCCACTCGTTC-3' Rev: 5'-ACTCCAGCTTGGACTTCTTG-3'	Reference [5]
D. melanogaster G4C2x36	Fwd: 5'-TTCCAACCTATGGAACTGATGA-3' Rev: 5'-GGTTTTCCTCATTAAAGGCATTC-3'	Reference [5]