Molecular Cell

The Solution Structure of FUS Bound to RNA Reveals a Bipartite Mode of RNA Recognition with Both Sequence and Shape Specificity

Graphical Abstract
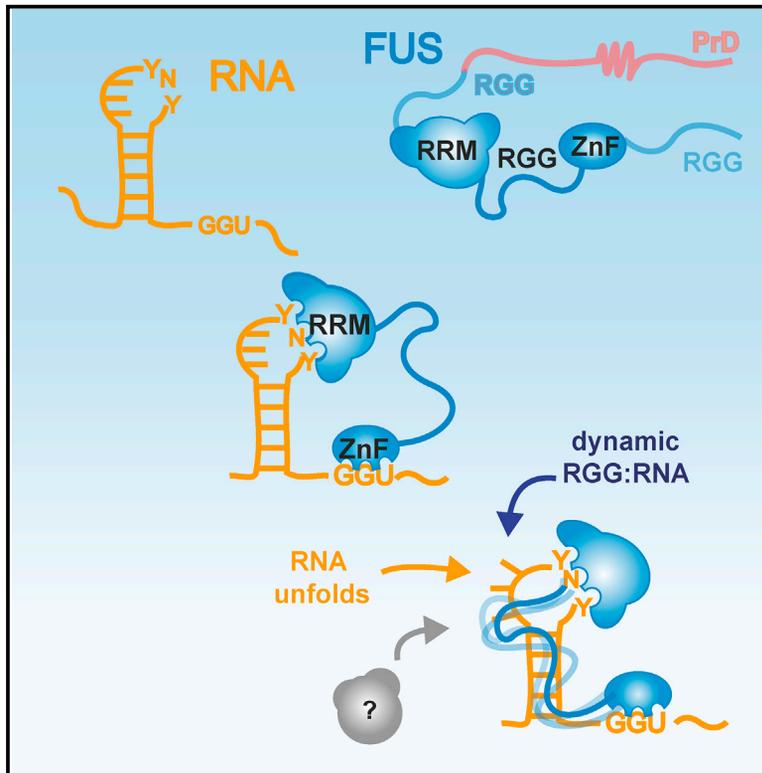
Authors

Fionna E. Loughlin, Peter J. Lukavsky, Tamara Kazeeva, ..., Magdalini Polymenidou, Marc-David Ruepp, Frédéric H.-T. Allain

Correspondence

fionna.loughlin@monash.edu (F.E.L.), allain@mol.biol.ethz.ch (F.H.-T.A.)

In Brief

FUS is an RNA binding protein associated with several neurodegenerative diseases, for which mode of nucleic acid binding has been elusive. Loughlin et al. solved the solution structure of FUS bound to RNA, revealing a sequence-specific recognition for a GGU motif and an unusual shape recognition of a stem loop by two separate domains.

Highlights

- Structures of FUS ZnF bound to GGU and FUS RRM bound to an RNA stem loop were solved

- The RGG motif destabilizes the RNA structure, suggesting a role in remodeling RNA

- The RGG motif increases RNA binding without forming an ordered structure

- RNA binding by the ZnF and RRM domains contribute to FUS-mediated splicing functions

CellPress

Molecular Cell

Article

CellPress

The Solution Structure of FUS Bound to RNA Reveals a Bipartite Mode of RNA Recognition with Both Sequence and Shape Specificity

Fionna E. Loughlin,[1,2,*] Peter J. Lukavsky,[1,7] Tamara Kazeeva,[1] Stefan Reber,[3,4,6] Eva-Maria Hock,[5] Martino Colombo,[3,4] Christine Von Schroetter,[1] Phillip Pauli,[1] Antoine Cléry,[1] Oliver Mühlemann,[3] Magdalini Polymenidou,[5] Marc-David Ruepp,[3,6] and Frédéric H.-T. Allain[1,8,*]

[1]Department of Biology, Institute of Molecular Biology and Biophysics, ETH Zürich, 8093 Zürich, Switzerland
[2]Monash Biomedicine Discovery Institute, Department of Biochemistry and Molecular Biology, Monash University, Clayton, 3800 VIC, Australia
[3]Department of Chemistry and Biochemistry, University of Bern, 3012 Bern, Switzerland
[4]Graduate School for Cellular and Biomedical Sciences, University of Bern, 3012 Bern, Switzerland
[5]Institute of Molecular Life Sciences, University of Zurich, 8057 Zurich, Switzerland
[6]Present address: United Kingdom Dementia Research Institute Centre, Department of Basic and Clinical Neuroscience, Institute of Psychiatry, Psychology and Neuroscience, Maurice Wohl Clinical Neuroscience Institute, King's College London, London SE5 9NU, UK
[7]Present address: Central European Institute of Technology, Masaryk University, Kamenice 753/5, 62500 Brno, Czech Republic
[8]Lead Contact
*Correspondence: fionna.loughlin@monash.edu (F.E.L.), allain@mol.biol.ethz.ch (F.H.-T.A.)
https://doi.org/10.1016/j.molcel.2018.11.012

SUMMARY

Fused in sarcoma (FUS) is an RNA binding protein involved in regulating many aspects of RNA processing and linked to several neurodegenerative diseases. Transcriptomics studies indicate that FUS binds a large variety of RNA motifs, suggesting that FUS RNA binding might be quite complex.

Here, we present solution structures of FUS zinc finger (ZnF) and RNA recognition motif (RRM) domains bound to RNA. These structures show a bipartite binding mode of FUS comprising of sequence-specific recognition of a NGGU motif via the ZnF and an unusual shape recognition of a stem-loop RNA via the RRM. In addition, sequence-independent interactions via the RGG repeats significantly increase binding affinity and promote destabilization of structured RNA conformation, enabling additional binding. We further show that disruption of the RRM and ZnF domains abolishes FUS function in splicing. Altogether, our results rationalize why deciphering the RNA binding mode of FUS has been so challenging.

INTRODUCTION

Fused in sarcoma (FUS) is a multi-functional heterogeneous nuclear ribonucleoprotein (hnRNP) involved in many aspects of RNA homeostasis, including transcription, splicing, and miRNA biogenesis (Schwartz et al., 2012, 2013). A component of many RNP granules, including stress granules and neuronal transport granules (Kanai et al., 2004), FUS is particularly adept at undergoing phase separation, a process thought to contribute to RNP

granule integrity in cells (Burke et al., 2015; Murakami et al., 2015; Patel et al., 2015).

The two neurodegenerative diseases amyotrophic sclerosis (ALS) and fronto-temporal lobar degeneration (FTLD) show neuropathological protein aggregates containing FUS, and it is hypothesized that mis-regulation of RNA processing could play a major role in these diseases (Ling et al., 2013; Ramaswami et al., 2013). ALS is a devastating disease affecting motor neurons, resulting in fatal muscular atrophy, and over 50 point mutations in FUS have been discovered in the familial forms of ALS, especially in the nuclear localization sequence (NLS) (Guerrero et al., 2016; Shang and Huang, 2016). FTLD is the second most common dementia following Alzheimer's disease and is characterized by progressive changes in behavior, personality, and language (Neumann et al., 2009). In FTLD patients, FUS protein also accumulates in aggregates in the absence of mutations along with two related FET protein family members (Neumann et al., 2011): EWS (Ewing's sarcoma protein) and TAF15 (TATA box binding protein associated factor). All three proteins show decreased levels of arginine methylation within RGG regions, preventing correct nuclear localization and increasing aggregation propensity. In contrast, decreased arginine methylation is not observed in ALS-FUS, indicating that these two diseases may proceed through distinct molecular mechanisms (Dormann et al., 2012; Hofweber et al., 2018).

ALS-linked mutations are almost absent from the folded RNA binding domains (RBDs), and RBD mutants diminish the toxicity of FUS bearing ALS-linked mutant in a fly model of ALS (Daigle et al., 2013). Accordingly, detailed characterization of FUS interaction with RNA is important to fully understand FUS function and to potentially suggest strategies for alleviating FUS pathogenicity. FUS shows widespread co-transcriptional binding to RNA, associating with a large variety of RNA classes (Ishigaki et al., 2012; Nakaya et al., 2013; Zhou et al., 2013). Binding data showed that FUS can bind many different RNAs promiscuously

CellPress

in a length-dependent manner (Ozdilek et al., 2017; Wang et al., 2015), yet a robust specificity has remained elusive.

FUS possesses an N-terminal prion-like region, which is involved in protein interactions and liquid-liquid phase separation (Burke et al., 2015; Murray et al., 2017). The C-terminal region is implicated in RNA binding and consists of two globular domains, an RNA recognition motif (RRM) and a zinc finger (ZnF), which are interspersed between three regions rich in RGG repeats. A C-terminal PY-NLS directs nucleo-cytoplasmic localization (Dormann et al., 2012). EWS and TAF15 have similar domain structures and appear to cross regulate the expression levels between them. The ZnF belongs to the small family of RanB2 ZnFs comprised of zinc ribbon-like domains with two crossed β-hairpins and a zinc atom bound by four cysteines. A subset of these domains has been shown to bind RNA (Iko et al., 2004; Nguyen et al., 2011). The role of the RRM in RNA recognition is less clear, with contradictory reports on its capacity to bind RNA (Iko et al., 2004; Liu et al., 2013). The FUS RRM has a classic β1α1β2β3α2β4 fold with the addition of a β-hairpin (β'β'') between α1 and β2, which is unique among RRMs and conserved in FET family members (Liu et al., 2013) and a non-canonical RNP1, a signature sequence of the RRM involved in RNA binding. The RGG-rich regions of FUS have a dual role: they can bind RNA, such as promoter-associated prD (Schwartz et al., 2013) or structured telomeric G-quadruplexes (Takahama et al., 2013, 2015), but when methylated, can also bind proteins, such as SMN (survival motor neuron) and can also independently promote phase separation (Guo et al., 2018; Hofweber et al., 2018; Qamar et al., 2018; Yoshizawa et al., 2018).

In order to unravel the contributions of the various domains of FUS for RNA binding, we have investigated RNA binding with a FUS construct comprised of the RRM, the ZnF, and the intervening RGG repeat, to many of the proposed RNA motifs or targets. Our nuclear magnetic resonance spectroscopy (NMR) chemical shift perturbation experiments revealed that both the ZnF and the RRM bind RNA and that each domain has a distinct RNA binding preference. The ZnF specifically binds GGU-containing RNAs, whereas the RRM binds with a broader specificity. The structure of the RRM bound to a stem-loop RNA revealed a very unusual binding mode for an RRM, binding the 3' side of the RNA loop with limited sequence specificity using its β sheet and the C-terminal helix and contacting the RNA stem by insertion of its unusual β-hairpin (β'β'') into the RNA major groove. Additionally, we found that the RGG repeats contribute significantly to the affinity for RNA and disrupts base stacking present in the 5' side of the RNA loop. Mutation of the RNA binding surface of the RRM and ZnF show that both domains can contribute to minor intron splicing, whereas the ZnF contributes to autoregulation of FUS and cross regulation of TAF15 expression levels. These results provide the first structural insights into how this enigmatic protein FUS recognizes RNA.

RESULTS

The RRM and ZnF of FUS Have Different RNA Binding Specificities

We expressed the region of FUS (269–454), encompassing the RRM, the ZnF, and the intervening RGG region (RGG2) in order to analyze the RNA binding specificity of FUS (Figures 1A and S1). The ^{1}H-^{15}N-HSQC (heteronuclear single quantum correlation) of this construct revealed dispersed cross-peaks corresponding to the folded RRM and ZnF, along with cross-peaks with little chemical shift dispersion centered at around 8 ppm from RGG2. Low and negative [1]H-^{15}N nuclear Overhauser effect (NOE) further confirmed that RGG2 region is highly dynamic and disordered (Figure S1A).

Several different RNA motifs have been proposed as targets for FUS, whereas other studies propose that FUS binds promiscuously. We titrated FUS RRM-RGG2-ZnF (269–454) with several RNAs, including stem-loop RNAs (Hoell et al., 2011); CG-rich (Ray et al., 2013), AUU-rich (Hoell et al., 2011), and a GGU-containing RNA (Lagier-Tourenne et al., 2012; Lerga et al., 2001; Rogelj et al., 2012; Masuda et al., 2015); and with single-stranded DNA (ssDNA) motifs (Tan et al., 2012; Figures S1B–S1D). NMR chemical shift differences were monitored to identify which regions are involved in binding the various RNA targets. Titration of the stem-loop RNA from *SON* and *hnRNPA2/B1* pre-mRNA, including the UA-Y motif at the stem-loop junction (Hoell et al., 2011), shows binding by the RRM, but not to the ZnF (Figures 1B, top, and S1B). The RRM also binds single-stranded RNA (ssRNA) of widely different sequences, including CG- and AUU-rich RNAs, suggesting that the RRM has a broad binding specificity (Figures 1B and S1C). Within the RRM, the same residues are involved in RNA binding, indicating that, for all three RNAs, the same RNA binding surface is involved, but not necessarily in the same manner (Figure 1C). This is illustrated by the chemical shift perturbations of specific residues, such as N284, which show clear differences in magnitude and direction of shifts with GC- versus AUU-rich RNAs, and T317, which shows larger chemical shift perturbations when binding RNA stem loops (Figures 1C, S1B, and S1C). In contrast, upon addition of the UGGUA motif, chemical shift perturbations are only observed in the ZnF (Figure 1B), indicating that binding is sequence specific. The RGG-rich interdomain linker shows little chemical shift change for all RNAs tested, yet it does undergo a reduction in reorientation motions when FUS is bound to the stem loop as shown by [1]H-^{15}N NOE (Figure S1A, bottom). Together, these data establish that the FUS RRM and ZnF each bind in a very distinct manner in terms of both structure and sequence. By inspecting the sequences surrounding the stem-loop motif bound by FUS in photoactivatable ribonucleoside-enhanced cross-linking and immunoprecipitation (PAR-CLIP) (Hoell et al., 2011), we identified the presence of neighboring GGU motifs upstream or downstream in numerous pre-mRNAs, including *SON* and *hnRNPA2/B1*. Binding of FUS (269–454) to the *SON* pre-mRNA containing the stem loop and a downstream GGU revealed that both RRM and ZnF domains are bound simultaneously (see chemical shift perturbation of T317 and K427; Figures 1B, 1C, and S1E, top). Similar results were obtained with hnRNPA2/B1 stem loop, which has an upstream GGU motif (Figure S1E, bottom), strongly suggesting that each domain of FUS can bind their respective motifs when both are present in the same RNA.

Structural Basis for RNA Recognition by the ZnF of FUS

In order to understand the molecular basis of RNA recognition by FUS, we determined the structures of the two folded RNA
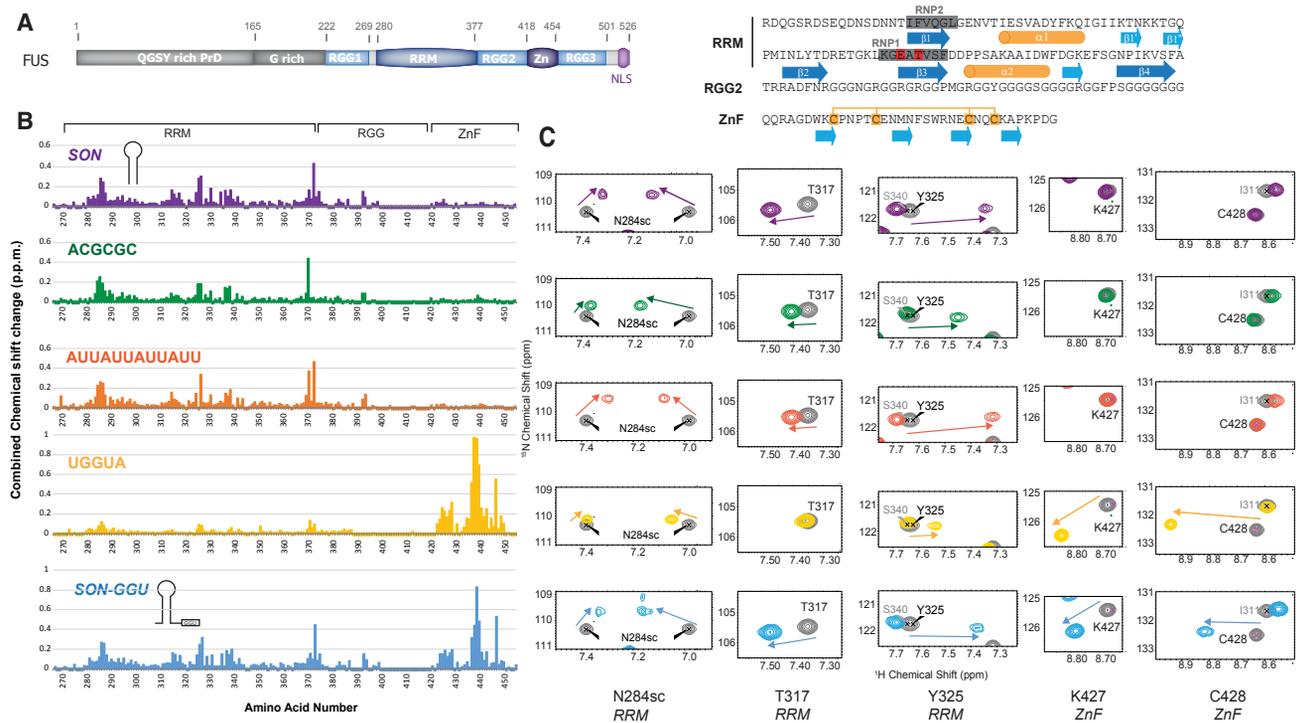
Figure 1. Mapping RNA Binding Sites in FUS RRM-RGG2-ZnF

(A) Domain composition of FUS (left) and sequence of RRM-RGG2-ZnF (zinc finger; right).

(B) Combined chemical shift perturbations of RRM-RGG2-ZnF on binding to an RNA stem loop from SON pre-mRNA (purple), ACGCGC (green), AUUAUUAUUAUU (orange), UGGUA (yellow), and part of SON containing a stem loop and a GGU sequence downstream.

(C) Inserts of overlaid ^{1}H-^{15}N-HSQC spectra showing chemical shift perturbations of individual cross-peaks from representative backbone amides and one side chain (N284) of FUS in their free (gray) and RNA-bound state (colored similarly to B).

See also Figure S1.

binding domains (ZnF and RRM) bound to one of their targets using NMR spectroscopy. ^{1}H-^{15}N HSQC spectra of individual domains bound to their target RNAs (UGGUG for the ZnF and a stem loop for the RRM) were almost identical to FUS (269–454) bound to the same RNAs (Figures S1F, S1H, and S1I), and therefore, we used spectra of the individual complexes to determine the 3D structures.

We first determined the structure of the ZnF (amino acids 418–454) bound to UGGUG (Figure 2). Using 74 intermolecular NOE-based distance constraints (Figures 2B and S2A; Table S1), we could calculate a precise structural ensemble of the complex (Figure 2C; Table 1). The ZnF binds only UGGU (Figure 2D), as we do not detect contacts to G_5 (Figure 2C). G_2 and G_3 stack on each side of F438 (Figure 2C), consistent with the large upfield shifts of this side chain (Figure S2B). The N-terminal residues Q420-Q421-R422 fold upon binding to interact with G_2, resulting in G_2 being sandwiched between R422 and F438 and recognized by four hydrogen bonds with D425 and W440 side chains and S439 main chain (Figure 2E). G_3 stacks with F438 and U_4, with its Hoogsteen edge recognized by R441 side chain and its Watson-Crick edge by M436 and F438 main chains (Figure 2F). Both the exchangeable imino resonances of G_2 and G_3 are observable, consistent with direct hydrogen bonding, and each gave rise to several intermolecular NOEs (Figure S2A, middle). In comparison, the two flanking uracils are more solvent

exposed. U_4 contacts the side chains N435 and N445, both of which recognize U_4 in a sequence-specific manner (Figure 2F). The hydrogen bonds are consistent with the observation of an additional protected imino group in the 1D spectrum (Figure S2A, middle), which could arise from the imino of U_4. U_4 also stacks with Q446 consistent with upfield shifts of the NH2 groups of this residue (Figure 2A). In contrast, U_1 contacts the ZnF by stacking with W440, in a sequence-independent interaction. The presence of U_1 changes the NMR exchange regime from fast to intermediate in ^{1}H-^{15}N HSQC titrations, demonstrating a direct contribution to binding (Figure S2D). Overall, the structure is consistent with a sequence-specific recognition of a NGGU motif, where N can be any ribonucleotide (A, C, T, or G).

The NGGU motif is a short motif present within slightly longer CLIP-derived GUGGU and systematic evolution of ligand by exponential enrichment (SELEX)-derived GGUG motifs. Natural abundance ^{1}H-^{13}C HSQC of RNA do not show chemical shift perturbations in the flanking guanines G_0 in $\underline{G}$UGGU and G_5 in UGGU$\underline{G}$ on binding ZnF, and no additional intermolecular NOE to these nucleotides were detected (Figure S2C; data not shown). Interestingly, chemical shift perturbations were observed for A_5 of UGGUA at low temperature, suggesting a potential interaction (Figure S2E); however, again no intermolecular NOEs could be detected. Isothermal Titration Calorimetry (ITC) measurements showed similar binding affinities of the ZnF for
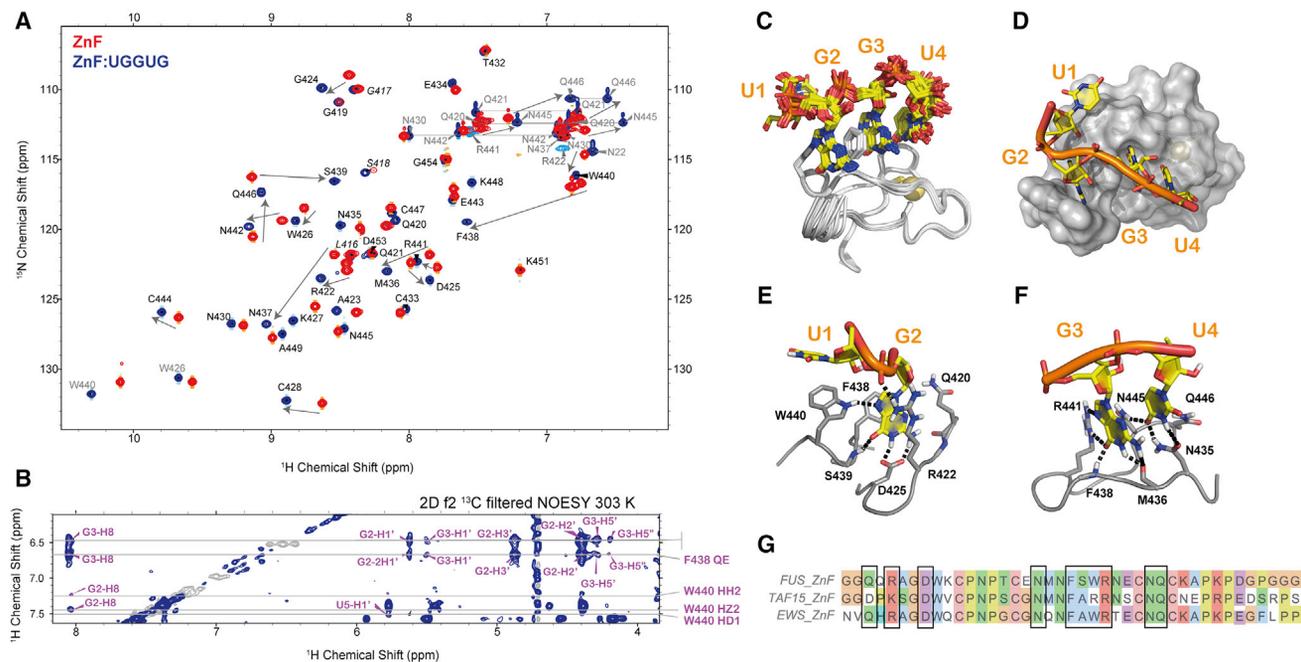
Figure 2. Structure of FUS ZnF Bound to UGGUG

(A) ^{1}H-^{15}N-HSQC of FUS ZnF free (red) and bound to UGGUG (blue).

(B) Section of ^{13}C-filtered 2D NOESY showing intermolecular NOEs between resonances of FUS Trp 440 and Phe 438 and RNA resonances.

(C) Superimposition of the 20 conformers forming the final NMR ensemble.

(D) Surface representation FUS ZnF bound to the RNA.

(E and F) Details of the intermolecular contacts of FUS ZnF and four RNA nucleotides (E) for U_1; G_2 and (F) G_3; U_4.

(G) Sequence alignments of the ZnF of the three members of the human FET protein family: FUS; EWS; and TAF15. Boxed show the residues interacting with the RNA.

See also Figure S2 and Table S1.

both UGGUG and UGGUA, in agreement with the final guanine not being bound in a sequence-specific manner (Figure S2F; Table S3). Our structure is in excellent agreement with a previously published binding assay (Nguyen et al., 2011), and the mode of recognition is likely to be conserved in EWS and TAF15 ZnFs, considering the high-sequence conservation between residues involved in interacting with RNA (Figure 2G).

Stem-Loop RNA Recognition by FUS RRM

As shown above, the RRM binds RNA with a broad specificity (Figure 1). However, the best quality NMR spectra (Figure 3A) and largest chemical shift differences (Figure 3B) were obtained with the RRM bound to the stem loop from hnRNP A2/B1 pre-mRNA (Hoell et al., 2011). Using 70 NOE-derived intermolecular distance constraints found between the RRM and the hnRNPA2/B1 stem loop and were confirmed in a complex with a deletion (ΔA11) in the loop region of the RNA (Figures S3A–S3C; Table S2), we obtained a precise structural ensemble (Table 1; Figure 3C). The RRM binds primarily four nucleotides (A_{12}, U_{13}, U_{14}, and C_{15}) in the 3′ side of the nine-nucleotide loop (Figure 3D). The 5′ side of the loop shows extensive intramolecular stacking (Figure 3D), reflective of sequential intra-RNA NOEs in this region. The binding site of the RRM is rather unexpected, because the proposed binding motif identified by PAR-CLIP included a UA motif in the 5′ side of the loop (Hoell et al., 2011),

which, at position U_8A_9, is not bound by the RRM in this structure. In addition to binding the RNA loop, the β-hairpin (β'β'') of the RRM inserts into the major groove of the stem (Figure 3D), contacting the phosphate backbone of both strands with the side chains K315 and K316 (Figure 3E).

The mode of binding shows little sequence specificity, although there are many intermolecular contacts (hydrophobic and hydrogen bonds) to both the bases and the sugar-phosphate backbone. All the phosphate groups between A_{12} and C_{15} are contacted by the short flexible C-terminal tail of the RRM that folds into a single helical turn and, together with the structured regions, make up the core RRM. Using this tail, A_{12}, U_{13}, U_{14}, and C_{15} phosphate oxygen are involved in hydrogen bonding by N376, R372, and R371 side chains and A373 main chain (Figure 3F), thus sandwiching the RNA on the β sheet surface. The bases of A_{12}, U_{13}, U_{14}, and C_{15} show stacking interactions and a limited number of hydrogen bonds. A_{12} interacts with Y325 and the edge of β2 (Figure 3G); U_{13} stacks between R328 and K334, with hydrogen bonds to R328 and T326 main chains (Figure 3H); U_{14} stacks between F288 and R372 with hydrogen bonds to T370 (Figure 3I); and C_{15} stacks between T286, N284, and R371, with hydrogen bonds to N285 and D283 main chains (Figure 3J). Many of these hydrogen bonds are not highly sequence specific, involving the O2 carbonyl group of all three pyrimidines, which cannot discriminate cytosines and uracils. Altogether, the

CellPress

Table 1. NMR Ensemble Statistics of FUS:RNA Complexes

	ZnF:UGGUG	RRM:hnRNPA2/B1
NMR Restraints		
Distance restraints	794	2,955
Protein		
Intramolecular	668	2,596
Intraresidual	152	517
Sequential ($|i - j| = 1$)	180	575
Medium range ($1 < |i - j| < 5$)	131	509
Long range ($|i - j| \geq 5$)	199	927
Hydrogen bonds[a]	6	30
RNA		
Intramolecular	52	324
Intraresidual	43	203
Sequential ($|i - j| = 1$)	9	103
Medium range ($1 < |i - j| < 5$)	0	0
Long range ($|i - j| \geq 5$)	0	4
Hydrogen bonds[a]	0	14
Complex		
Intermolecular	74	73
Long range ($|i - j| \geq 5$)	68	70
Hydrogen bonds[a]	6	3
Torsion Angles[b]	25	–
backbone	20	140
RNA		
Sugar pucker (DELTA)	5	40
Backbone in A form stem	–	88
Energy Statistics[c]		
Average distance constraint violations		
0.3–0.4 Å	0.5 ± 0.5	4.8 ± 2.0
>0.4 Å	0.0 ± 0.0	1.1 ± 1.4
Maximal (A)	0.30 ± 0.05	0.44 ± 0.08
Average angle constraint violations		
<5 degree	0.4 ± 0.6	4.2 ± 1.5
>5 degree	0.0 ± 0.0	0.2 ± 0.4
Maximal (degree)	0.19 ± 0.33	2.54 ± 3.71
Mean AMBER Violation Energy		
Constraint (kcal mol^{-1})	18.5 ± 1.9	109.0 ± 9.1
Distance (kcal mol^{-1})	18.5 ± 1.9	108.2 ± 8.4
Torsion (kcal mol^{-1})	0.0 ± 0.0	0.8 ± 1.3
Mean AMBER Energy (kcal mol^{-1})	−3,074.2 ± 4.7	−7,424.5 ± 17.7
Mean Deviation from ideal covalent geometry		
Bond Length (A)	0.0044 ± 0.0001	0.0041 ± 0.0000
Bond Angle (degrees)	1.357 ± 0.016	1.438 ± 0.011
Ramachandran Plot Statistics[c,d,e]		
Residues in most favored regions (%)	85.5 ± 2.1	85.7 ± 1.5

Table 1. Continued

	ZnF:UGGUG	RRM:hnRNPA2/B1
Residues in additionally allowed regions (%)	14.5 ± 2.1	13.5 ± 1.2
Residues in generously allowed regions (%)	0.0 ± 0.0	0.0 ± 0.0
Residues in disallowed regions (%)	0.0 ± 0.0	–
RMSD to Mean Structure Statistics[c,d]		
Protein		
Backbone atoms	0.23 ± 0.08	0.23 ± 0.06
Heavy atoms	0.41 ± 0.08	0.54 ± 0.08
RNA		
Backbone atoms	0.79 ± 0.25	0.72 ± 0.21
Heavy atoms	0.71 ± 0.22	0.74 ± 0.22
Complex		
Backbone atoms	0.52 ± 0.15	0.90 ± 0.38
Heavy atoms	0.55 ± 0.10	0.99 ± 0.35

RMSD, root-mean-square deviation. [a]Hydrogen bond constraints were identified from slow exchanging amide protons in D2O and imino protons in H2O.
[b]Torsion angle based on HNHA (ZnF) or TALOS+ (RRM); sugar puckers based on homonuclear TOCSY and DQF-COSY; RNA backbone constraints in A form stem (hnRNPA2/B1 RNA stem loop) based on standard A form geometry ± 20°.
[c]ZnF:UGGUG ZnF: 422-453, chain ID: A (sequence range: 418–454); RNA: 1–4, chain ID: B (sequence range: 1–5).
[d]RRM:hnRNPA2/B1 stem-loop RNA RRM: 284–372, chain ID: B (sequence range: 276–377) RNA: 4–20, chain ID: A (sequence range: 1–23).
[e]Ramachandran plot, as defined by the program Procheck (Laskowski et al., 1996).

structure suggests that the RRM binds a NYNY single-stranded quartet (where Y is a C or U and N any nucleotide A, C, G, and U).

Although the main RNA binding surface in FUS is located on the RRM β sheet surface, the path of the RNA is very atypical. In a canonical mode of RNA binding by RRMs, the single-stranded nucleotides are spread on the surface of the β sheet like rings on fingers, so that each base (ring) interacts predominantly with one β strand (finger), the nucleotides one to four interacting with β4, β1, β3, and β2, respectively, across the β sheet (Maris et al., 2005). Here, the four nucleotides bound by FUS RRM form a tight curve encircling T338 and N323 with A_{12}, U_{13}, U_{14}, and C_{15} bound in individual pockets (Figure 3D). Loop residues A_{12} and U_{13} bind near the end of β2, and U_{14} and C_{15} bind over β1, with no bases interacting with β3. Whereas in canonical RRMs, two aromatics in β1 and β3 stack with RNA bases and a third aromatic from β3 inserts between the two sugars, here, the only canonical aromatic in β1 (F288) stacks with U_{14} and Y325 (from β2) inserts between the sugars of U_{13} and U_{14} (Figure 3H). The base of C_{15} interacts with the N-terminal region of β1, stacking with T286 (Figure 3J). Overall, the binding topology of the single-stranded RNA over the β sheet of the RRM is unprecedented and reflects the lack of the two aromatic side chains, which are commonly found on β3 of an RRM.
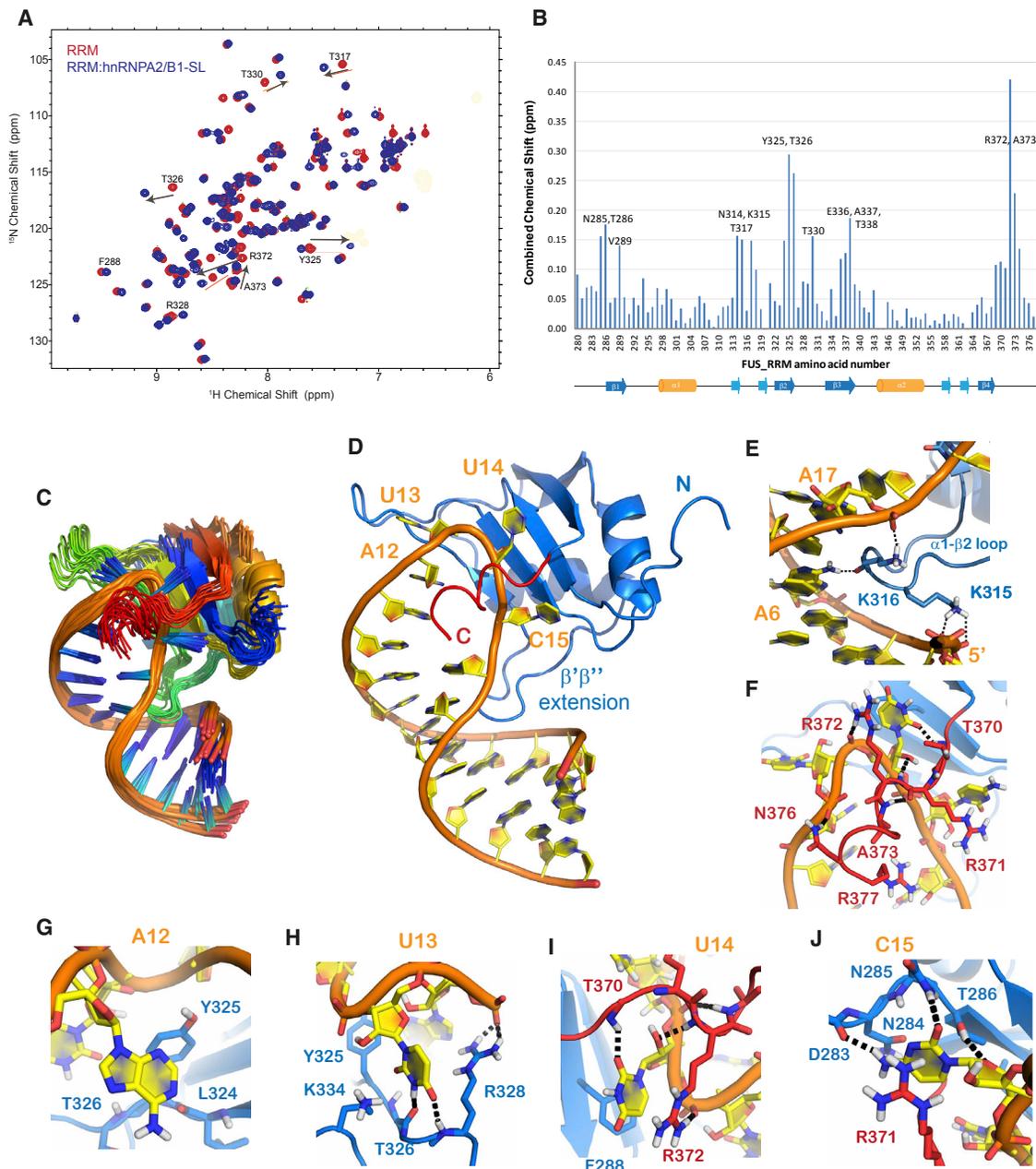
Figure 3. Structure of RRM Bound to Stem-Loop RNA from hnRNPA2/B1

(A) ^{1}H-^{15}N-HSQC of FUS RRM (280–377) free (red) and bound to the RNA stem loop of hnRNP A2/B1 pre-mRNA (blue).

(B) Combined chemical shift perturbation of FUS RRM on binding RNA stem loop from hnRNPA2/B1.

(C) Superposition of the 20 conformers forming the final NMR ensemble.

(D) Structure of the complex with the RRM ribbon shown in blue, the C-terminal tail in red, and the RNA in yellow.

(E and F) Details of the RNA recognition by two extensions of the RRM core: (E) $\beta'-\beta''$ hairpin of the α1-β2 loop and (F) the C-terminal helix.

(G–J) Details of the intermolecular contacts in the binding pockets for four RNA nucleotides on the RRM β sheet: (G) A_{12}; (H) U_{13}; (I) U_{14}; and (J) C_{15}.

See also Figure S3 and Table S2.

The canonical β sheet surface of FUS-RRM is supplemented in the FET family by a β-hairpin between α1 and β2 ($\beta'\beta''$), which extends the RNA binding surface. In our structure, this β-hairpin inserts into the major groove of the RNA stem with three side chains (N314, K315, and K316) interacting with three phosphate

oxygens of the stem (Figures 3E and S3D). The same structural feature and residues are present in the RRM of TAF15 and EWS, suggesting that this mode of RNA recognition is likely to be conserved throughout the FET family. The most C-terminal residue in this construct R377 represents the start of the RGG2

CellPress
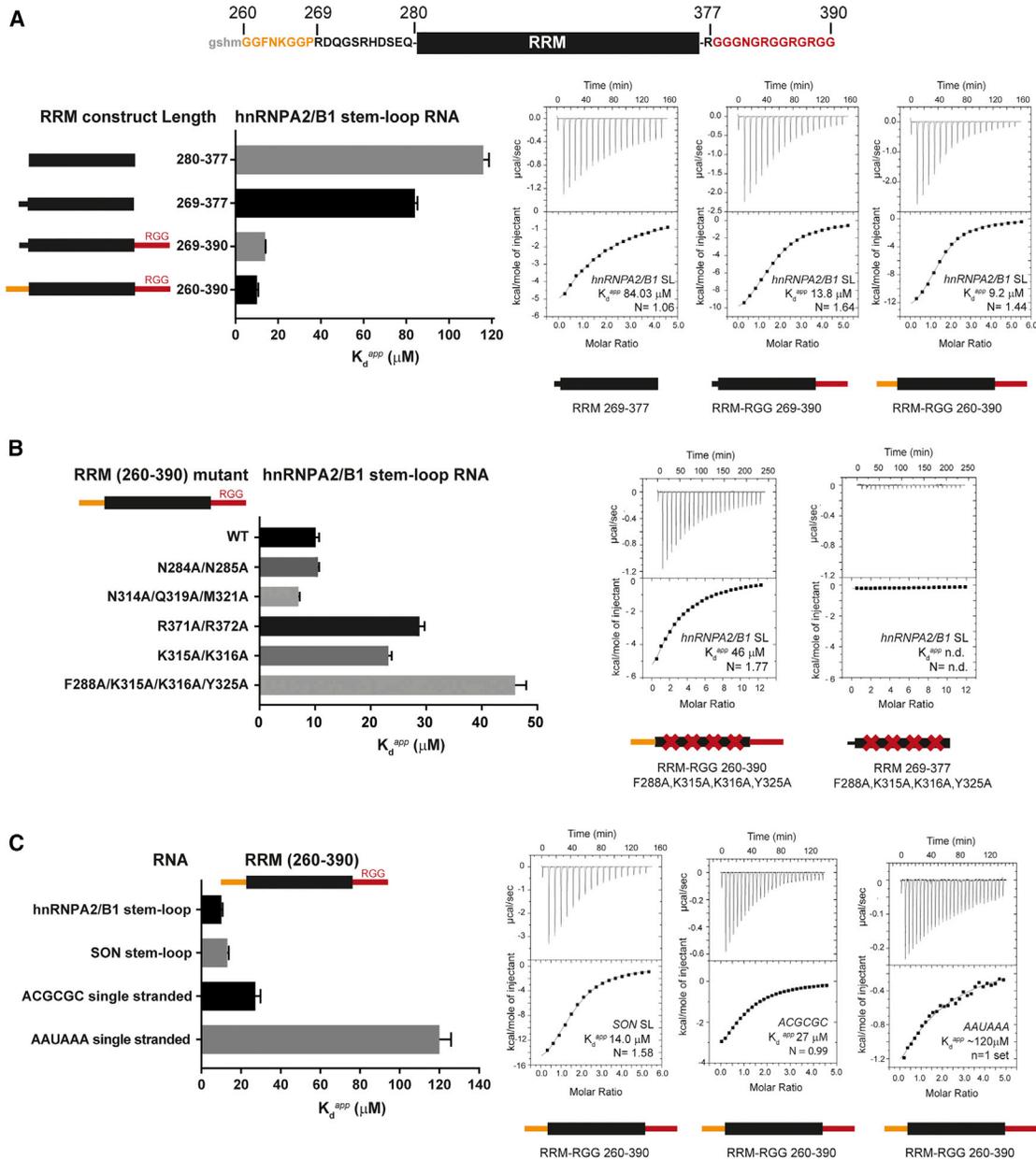
Figure 4. Binding Affinities of FUS RRM Constructs and Mutants with RNA

Binding affinities of RRM constructs and RNA showing Kdapp (left) and representative ITC measurements (right) showing injection profiles top, and binding isotherms, bottom.

(A) ITC measurements of several RRM constructs of different lengths and hnRNPA2/B1 stem-loop RNA.

(B) ITC measurements of RRM-RGG$_{\times 3}$ (260–390) and RRM mutants with hnRNPA2/B1 stem loop.

(C) ITC measurements of RRM-RGG$_{\times 3}$ constructs with RNA stem loop from hnRNPA2/B1, SON, and short ssRNA ACGCGC and AAUAAA.

See also Table S3.

region, and its position suggests that RGG2 may have further interactions with the stem loop.

Role of FUS RGG2 in RNA Interactions

To investigate the role of RGG2, we included 3 RGG repeats (RGG$_{\times 3}$ 378–390) C-terminal to the RRM and for which resonances could be assigned. The addition of these residues led

to a large increase in affinity with the K_d^{app} of 13.9 μM compared to 84 μM along with a large increase in binding enthalpy. In contrast, addition of ten N-terminal residues (260–269) shows only a marginal increase in affinity (Figure 4A; Table S3). This effect is observed for two different stem loops, but not for binding to short single-stranded RNA (Table S3; data not shown). When monitoring the RNA resonances of two stem-loop RNAs,
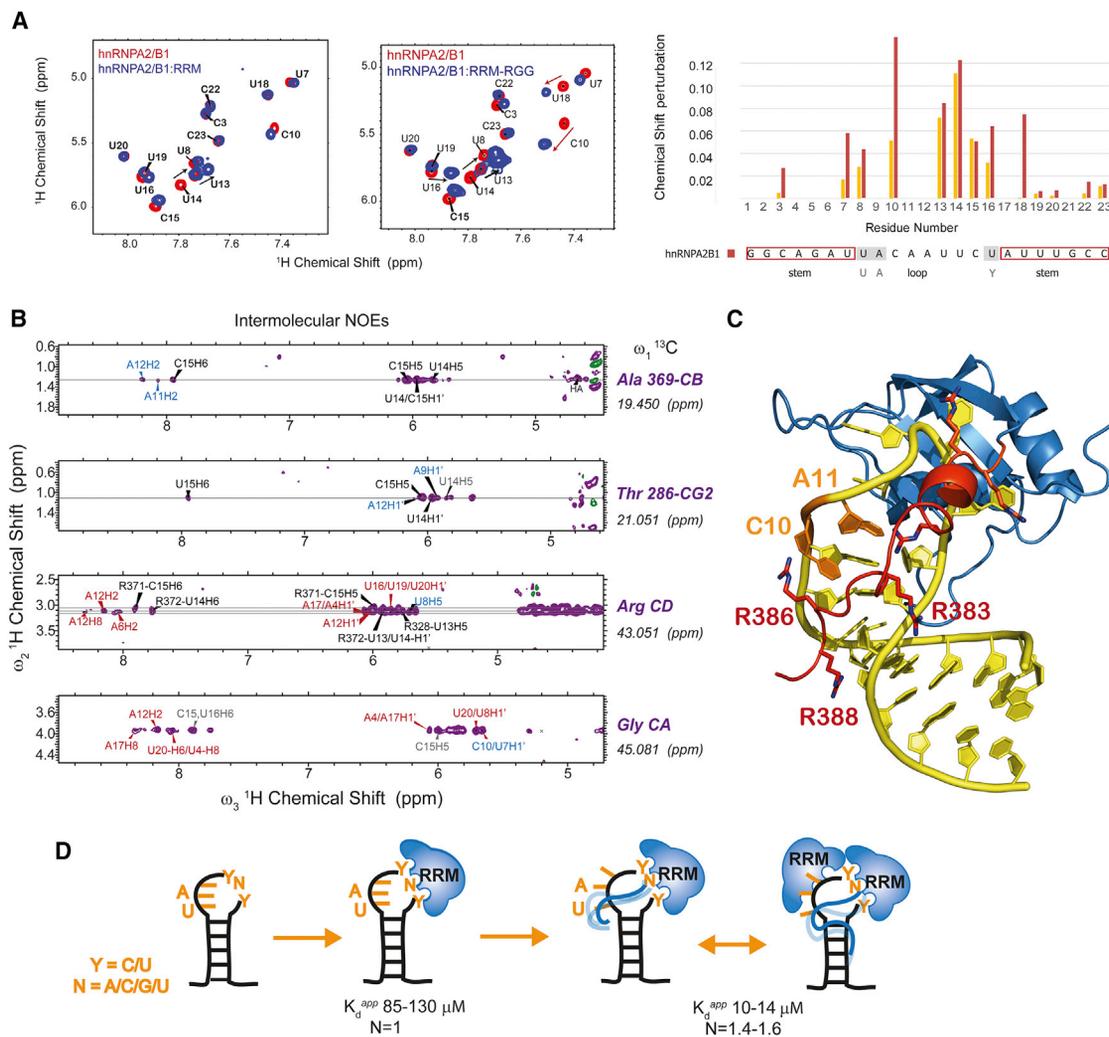
Figure 5. Effect of RGG Repeats on RNA Binding by RRM

(A) ^{1}H-^{1}H 2D TOCSYs showing pyrimidine H5-H6 correlation of the RNA stem loops free (red) and in complex with RRM (blue) with combined chemical shifts mapped onto the RNA sequence. Changes on main RNA binding site are shown in back arrows, and additional changes are marked with red arrows.

(B) 3D ^{13}C-filtered, ^{13}C-edited NOESY of RRM$_{RGG \times 3}$ (260–390) with the RNA stem loop of hnRNPA2/B1 showing intermolecular NOEs in common with the core RRM (280–377):stem loop (black), arising from the RGG motif (red) and an additional minor binding site (blue).

(C) NMR-based model of RRM$_{RGG \times 3}$ (280–390) and the RNA stem loop of hnRNPA2/B1 showing one conformer.

(D) Proposed schematic model of the RRM binding to RNA stem loops and role of the RGG2.

See also Figure S4.

we observe additional chemical shift changes indicative of an unfolding of the bases of the nucleotides in the 5′ part of the loop (U8, C10, U18) when bound by RRM-RGG$_{\times 3}$ (260–390) compared to RRM (280–377; Figures 5A and S4A). This suggests that the RGG repeats in the presence of the RRM can remodel the RNA structure.

Characteristic intermolecular NOEs and backbone chemical shift perturbations of the RRM-RGG$_{\times 3}$:SL RNA complex show that the RRM binds RNA in a similar manner to the core RRM construct at the AUUC in the 3′ loop region of the stem loop (Figures 5B, black labels, and S4B). Also additional intermolecular NOE involving the RGG repeats are detected (Figure 5B, red labels). Surprisingly, the RGG region shows little chemical shift perturbation in the backbone amide resonances

on binding RNA despite the large contribution to binding affinity. Heteronuclear NOE analysis indicates that the RGG$_{\times 3}$ is locally flexible with modest decrease in reorientational motions when bound to an RNA stem loop (Figure S4C). This is consistent with intermolecular NOEs detected from arginine and glycine residues of RGG2 showing multiple points of contact to the RNA (Figure 5B, red labels). To obtain insights into this interaction, we used restraints from the core RRM-SL complex combined with several RGG-RNA intermolecular NOEs to model the complex. The resulting models suggest how the RGG repeats interact with the 5′ side of the loop and the minor groove of the stem (Figures 5C and S4D). As such, the stem-loop:RRM structure provides a platform for interactions with RGG motifs.

To test the significance of shape recognition of FUS-RRM:SL complex, we measured binding affinities of a series of protein mutants of RRM-RGG$_{\times 3}$ (Table S3; Figure 4B). R371A and R372A resulted in a 3-fold decrease in binding affinity, confirming a contribution to the loop recognition, and K315A and K316A to a 2- to 3-fold decrease, confirming a contribution for binding the stem major groove. In contrast, mutation of residues arising from the opposite face of the β'-β" hairpin (N314A, Q319A, and M321A) showed little effect. Mutation of both the β sheet surface and the β'-β" hairpin (F288A/Y325A/K315A/K316A) gave a 4- to 5-fold binding decrease compared to wild-type. Introducing the same four mutations in the core RRM (residues 269–377) alone resulted in no detectable binding (Figure 4B). Comparing binding affinities of the core RRM (K_d^{app} 85 μM) with the RRM-RGG$_{\times 3}$ F288A/Y325A/K315A/K316A mutant (K_d^{app} 45 μM) suggests synergistic binding by the RRM and the RGG motifs, in good agreement with our NMR-based model and recent binding studies with a promoter-associated noncoding RNA sequence (Ozdilek et al., 2017).

On further analysis, we identified additional weak intermolecular NOEs consistent with an additional minor binding event by the RRM three nucleotides upstream of the main binding site: U$_8$; A$_9$; and C$_{10}$ (Figure 5B, blue labels). An additional binding event is consistent with the change in stoichiometry that we observe in our ITC measurements changing from 1 to ~1.5 with RRM-RGG$_{\times 3}$ (Figure 4; Table S3), with the latter measurements representing a mixture of 1:1 and 1:2 RNA:protein complexes as depicted in Figure 5D. Similar results were obtained when binding the *SON* stem loop (Figure S4). Overall, we propose a mechanism for FUS binding by which the RRM binds NYNY at the 3′ side of loop and the major groove of the stem, and RGG$_{\times 3}$ dynamically interacts with and unfolds the 5′ side of the loop. In our samples, unfolding of the 5′ loop enables another molecule of FUS to partially bind (Figure 5D), but in a cellular environment, this unfolding of the loop could enable binding of additional proteins to the RNA. As such, we speculate that the proposed UA motif at the 5′ side of the loop (Hoell et al., 2011) may originate from the binding of an unknown sequence-specific RNA binding protein. Further investigations would be required to identify this protein.

FUS Can Bind a YNY Stem-Loop-GGK Bipartite Motif *In Vivo*

We then asked whether FUS recognizes bipartite RNA elements *in vivo*. Using our structure-derived RNA motifs, we searched in CLIP datasets for FUS for the presence of stem loops containing a YNY motif (Y = C/U; N = A/C/G/U) at the 3′ end of a stem loop that is followed by a GGK motif (K = U/G) within a 30-nt distance (Figure 6A). We searched FUS cross-linking immunoprecipitation with high throughput sequencing (CLIP-seq) in HeLa cells (Zhou et al., 2013), PAR-CLIP in HEK293 cells (Hoell et al., 2011), and CLIP-seq in human brain and mouse neurons differentiated from embryonic stem cells (Lagier-Tourenne et al., 2012; Nakaya et al., 2013). We found in all datasets a small but significant increase of these motifs within FUS CLIP tags compared to control in which the sequences were randomized, maintaining original nucleotide composition (Figure 6B), suggesting that FUS could recognize such a bipartite motif *in vivo* as well as *in vitro* (Figures S1E and S5A).

RNA Binding by FUS ZnF Is Crucial to Splicing, whereas Binding by the RRM Is Less Critical

FUS widely regulates alternative splicing, including pre-mRNA with long introns and its own transcript via intron retention, leading to nonsense-mediated decay (Lagier-Tourenne et al., 2012; Zhou et al., 2013). It was recently shown that FUS regulates minor intron splicing and that loss of FUS and its binding partner U11 snRNP from the nucleus could contribute to pathological mechanisms in ALS (Reber et al., 2016). We therefore tested the functional role of RNA binding by the RRM and the ZnF in splicing assays, including minor intron excision from a skeletal-muscle sodium-channel (SCN4), which is stimulated by FUS, the autoregulation of FUS mRNA, and the cross regulation of TAF15 mRNA. We therefore used mutants affecting RNA binding of the RRM (F288A, Y325A, K315A, and K316A) and the ZnF (F438A, R441A, and N435A; Figure 6C). These mutations reduced binding but retain the native fold of the RRM and the ZnF (Figures S5B and S5C).

FUS knockdown induced deficiency in the splicing of the SCN4A minigene in HeLa cells, which could be rescued by ectopic expression of wild-type FUS, but not by FUS in which both the RRM and the ZnF were simultaneously mutated (Figure 6D). This demonstrates that both domains are essential for the role of FUS in minor intron splicing stimulation. Next, we tested the contribution of the individual domains. Mutants in the RRM alone had little effect, whereas mutations in the ZnF had an intermediate effect on splicing, suggesting that the ZnF has a dominant role but that the RRM and ZnF act synergistically in promoting excision of the intron. Importantly, these differences in splicing activity do not arise from different knockdown efficiencies, as all samples have a comparable knockdown of endogenous FUS ranging from 1% to 3% (Figure S5D) and neither from differences in ectopic expression, as there is no significant difference between the expression levels of FUS wild type and mutants (Figure S5E).

Next, we analyzed the effect of RNA binding by the RRM and the ZnF on the autoregulation of FUS and cross regulation of TAF15 mRNA (Figures 6E and 6F). Ectopic expression of FUS in HeLa cells resulted in a modest reduction in endogenous FUS mRNA. This effect was not observed with FUS carrying mutations in the ZnF (Figure 6E). To measure cross regulation of TAF15 mRNA, FUS knockout cells were used. Knockout of FUS in HeLa cells shows increased levels of TAF15 mRNA, and ectopic expression of wild-type FUS can efficiently restore repression (Figure 6F). FUS with the mutated ZnF can no longer repress TAF15. In contrast, the mutated RRM has little effect, and mutation of both domains has an intermediate effect, suggesting that the RRM and ZnF might not act synergistically for this intron retention event (Figure 6F). In both assays, the mRNA level changes are unlikely to arise from differences in overexpression efficiencies, because no significant differences between exogenous FUS levels were detected in any case except between wild-type FUS and the double mutant in the endogenous FUS mRNA level assay (Figures S6F and S6G). Together, these experiments show that the RNA interactions by the globular domains of FUS are important for FUS function in post-transcriptional regulation of gene expression.

CellPress
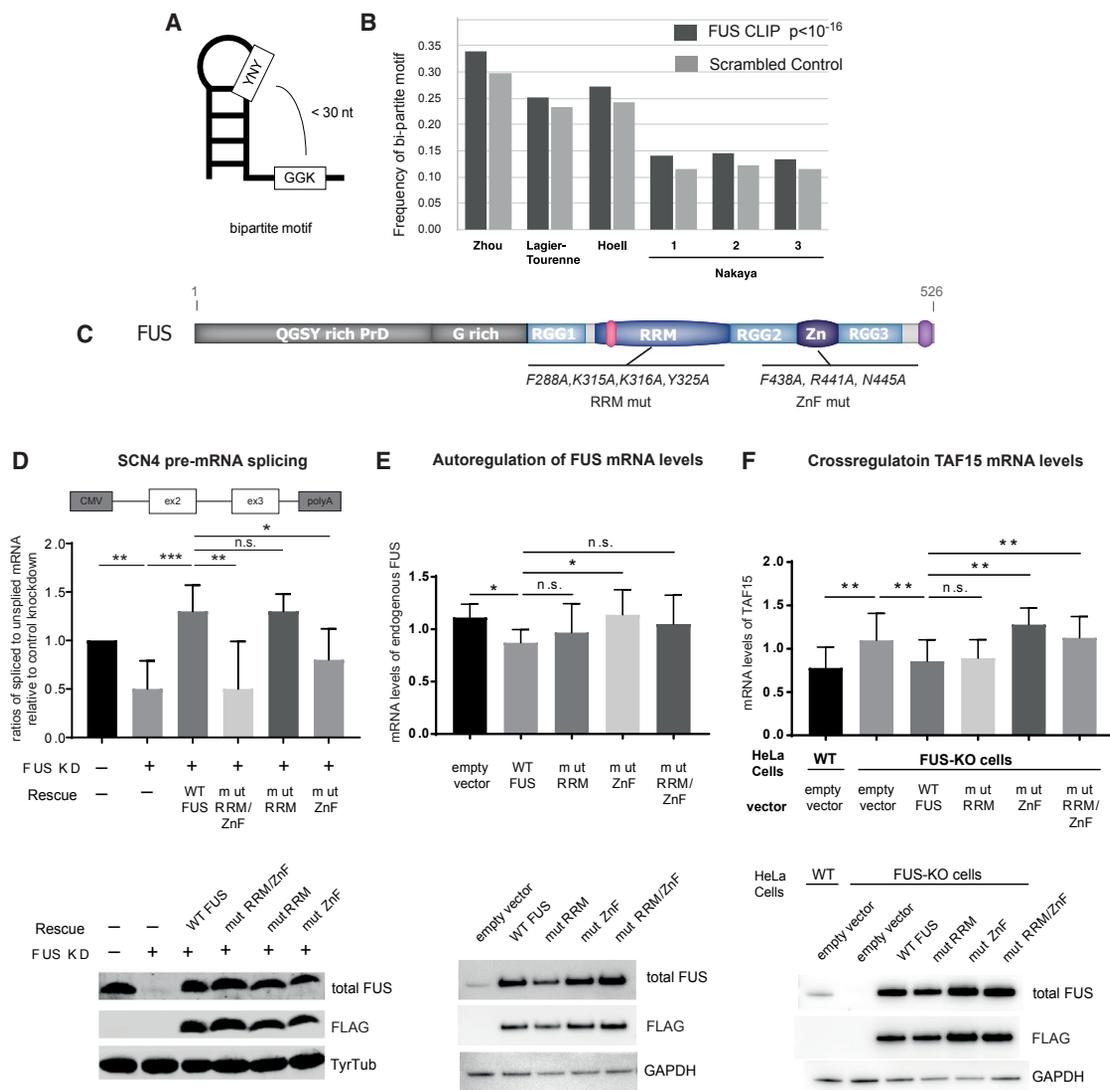
Figure 6. Functional Role of the Bipartite RNA Motif Present in FUS

(A) The structure of the motif investigated and its enrichment in FUS-associated regions of public datasets. The motif is composed of a stem-loop structure with a YNY tri-nucleotide (Y, C/U; N, A/C/G/U) at the 3′ end of the loop in combination with a GGK (K, G/U) sequence within 30 bp from the loop.

(B) The enrichment for this motif has been evaluated in several FUS CLIP studies to address whether it could play a functional role for FUS binding *in vivo*. As a control, these sequences have been randomized, maintaining the original nucleotide composition. In all studies, we could determine a significant (p^{-16}), albeit limited, enrichment for the motif, suggesting a possible functional role.

(C) Schematic of FUS protein showing the mutations used in this study.

(D) Splicing of SCN4 minigene: schematic representation of the SCN4A minigene (top). qRT-PCR results indicating the ratio of spliced to unspliced mRNA under control knockdown, FUS knockdown, and FUS knockdown rescued with RNAi-resistant cDNA expressing various constructs are shown (upper panel). Example immunoblots of FUS protein levels under control knockdown (lane 1), FUS knockdown (lane 2), and different FUS-FLAG rescue conditions (lanes 3–6) analyzed by western blot are shown. HeLa cell extracts were subjected to SDS-PAGE and western blotting with anti-FUS (upper row) and anti-FLAG (middle row). TyrTub (lower row) was used as loading control (lower panel).

(E) FUS mRNA levels: qRT-PCR results indicating the relative endogenous FUS mRNA levels normalized to five housekeeping genes upon different FUS overexpression conditions (upper panel). Example immunoblots show similar overexpression levels of the different constructs (lower panel).

(F) Same as (E) for TAF15 mRNA.

See also Figure S5.

DISCUSSION

The ZnF Provides Sequence Specificity to FUS

We have elucidated how FUS ZnF is bound to UGGU. The two central guanines are deeply buried in sequence-specific pockets and U$_4$ is recognized by side chains that could also accommodate a G, and U$_1$ is not sequence specifically recognized. The consensus sequence bound by FUS ZnF is NGGK. Conservation of structure and side chains within EWS and TAF15 indicate that this mode of recognition is conserved within this family (Figure 2G).
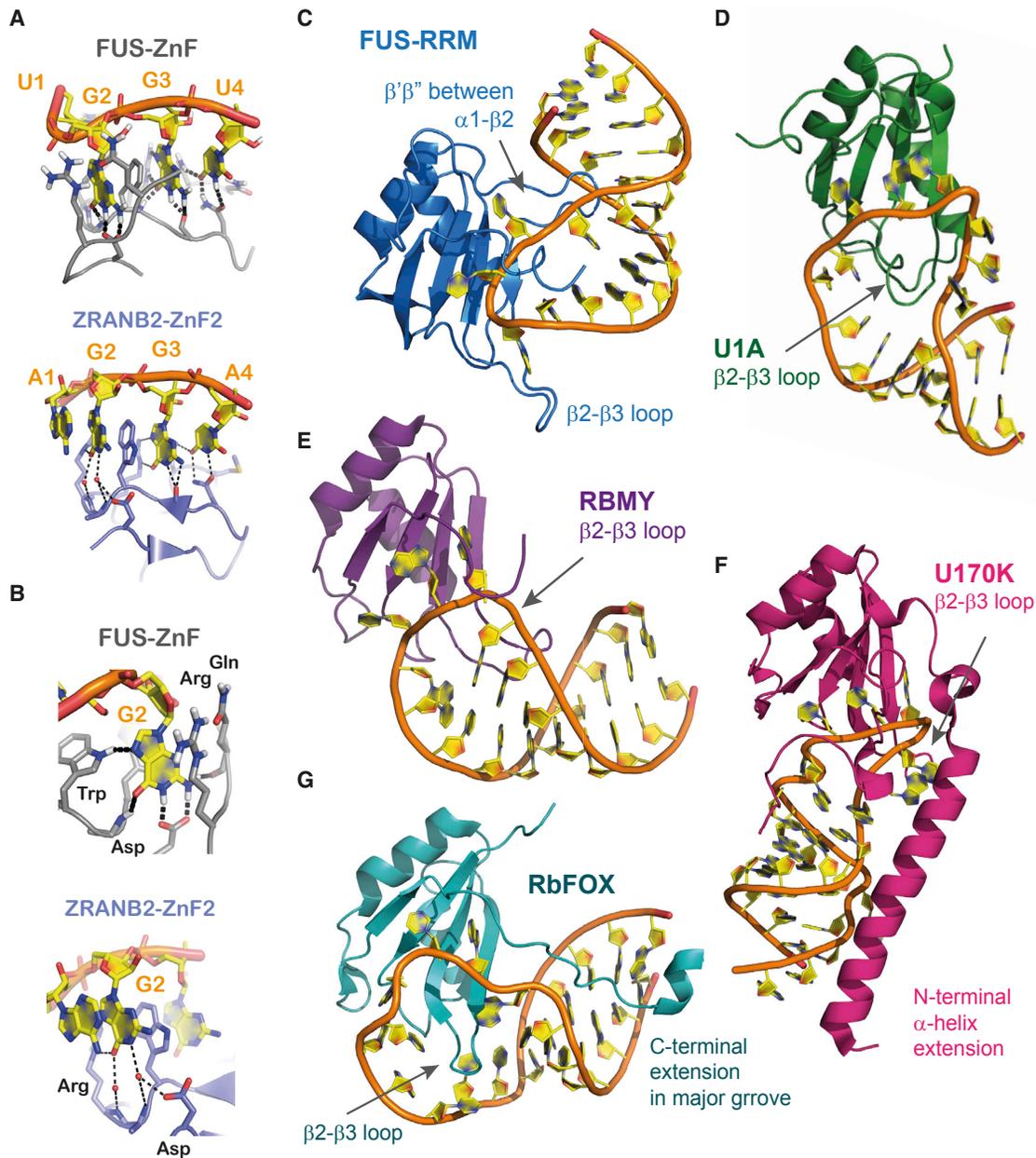
Figure 7. Structural Comparison of FUS RNA Complexes

(A and B) Comparison between (A) ZnF of FUS and ZnF2 of ZRANB2 bound to RNA showing placement of NGGU and (B) direct and water-mediated hydrogen bonding of G2.

(C–G) RRM domains bound to stem-loop RNAs.

(C) FUS RRM bound to a stem loop from hnRNPA2/B1 pre-mRNA.

(D) U1A RRM1 bound to U1 snRNA stem loop II (PDB: 1URN).

(E) RBMY RRM bound to SELEX-derived stem loop (PDB: 2FY1).

(F) U170K RRM bound to U1 snRNA stem loop I (PDB: 4PKD).

(G) RbFox1 RRM bound to GCAUG motif in pre-miR 20b RNA showing interaction of the C-terminal linker with the RNA stem.

RanB2 ZnFs are a small family of ZnFs with small ribbon-like domains consisting of two crossed β-hairpins, and a subset of these domains bind RNA (Nguyen et al., 2011). The only structural data available of this ZnF family in complex with RNA come from ZnF2 of the alternative splicing factor ZRANB2 (Loughlin et al., 2009), which contains a central WARR motif that binds a GGU motif. Despite two changes from Trp to Phe and Arg to Trp, FUS-ZnF binds the central GGU in a very similar manner (Figure 7A). Stacking of G_2 and G_3 with a Phe (rather than a Trp) by FUS ZnF enables direct hydrogen bonds between G_2

CellPress

rather than water-mediated hydrogen bonds, as observed with ZRANB2 (Figure 7B). The Trp replacing Arg conserved one hydrogen bond to G_2 and introduced stacking with the U_1 base. In the FUS complex, additional residues Q420 to R422 are present and fold upon RNA binding, expanding the interface to envelop G_2. TAF15 ZnF possesses a Phe-Ala-Arg-Arg motif, thus combining residues from both FUS and ZRANB2. Thus, despite an adjustment of some hydrogen bonds, the principle of RNA recognition of GGU in this family of ZnF is conserved.

A GGU recognition motif has been detected in several CLIP experiments (Ishigaki et al., 2012; Lagier-Tourenne et al., 2012). Our solution structure of FUS ZnF bound to RNA and our numerous binding studies with various RNA sequences clearly show that the ZnF of FUS is responsible for binding this motif. It is also notable that a GGUA motif has been identified in both the TAF15 CLIP dataset (Kapeli et al., 2016) as well as with the in vitro method RNA compete (Ray et al., 2013), confirming that the sequence specificity mediated by ZnF extends to the FET family in vivo. Finally, our splicing assays clearly show that the FUS-ZnF is functionally essential for its role in post-transcription gene regulation (Figure 6).

The FUS RRM Binds Stem-Loop RNAs in an Unusual Manner and with Highly Degenerate Specificity

The structure of the FUS RRM bound to the stem-loop RNA of hnRNPA2/B1 pre-mRNA reveals three individual binding pockets on the β sheet surface, as expected for an RRM (Figure 3). However, the path taken by these nucleotides is unusual, with the three nucleotides forming a tight turn rather than a straight line. Most contacts are of non-sequence-specific nature, with hydrophobic interactions and contacts to the phosphate backbone. The hydrogen bonds to the bases suggest a highly degenerate sequence specificity for a NYNY quartet. The tight turn in the path of the RNA upon binding to the RRM could predispose FUS-RRM to binding stem-loop RNAs because the RNA would be pre-formed in such a conformation. In addition, the contact in the major groove of the RNA stem by the β-hairpin (β'β") unique to the RRM of the FET family further contributes to the recognition of such structures. The RNA recognition by the RRM is therefore more shape specific than sequence specific.

Only few RRMs show specific binding to stem loops, and among them are four structures with a single RRM bound to RNA, namely U1-A, U2-B", RBMY, and U1-70K (Kondo et al., 2015; Oubridge et al., 1994; Price et al., 1998; Skrisovska et al., 2007). Compared to these complexes, the way FUS RRM binds to the stem-loop RNA is markedly different (Figures 7C–7F). The most obvious difference is on the position of the RNA stem relative the RRM, which is rotated 180 degrees from U1-70K and 90 degrees with respect to the others (Figures 7D–7F). This originates primarily from interaction of the β-hairpin (β'β") with the RNA stem in FUS RRM, whereas in the other proteins, the interaction with the stem is mediated by the β2−β3 loop.

FUS-RRM does not exclusively bind stem-loop structures. When testing various single-stranded sequences, like AU repeats or GC repeats, chemical shift changes showed that the

RRM remained the primary RNA binding domain. This is in agreement with the very degenerated binding consensus sequence that we deciphered from the structure. The RRM interaction with single-stranded RNAs does possess an inherent sequence preference as shown by comparing affinity of ACGCGC (K_d 27 μM) and AAUAAA (K_d 120 μM; Table S3), with an exact definition of the RRM-RNA specificity requiring further investigation.

Role of the Disordered RGG-rich Interdomain Linker in RNA Binding

RGG boxes are multifunctional intrinsically disordered regions that are prevalent in RNA binding proteins. They are generally assumed to bind RNA through non-specific charge interactions or specifically bind and stabilize G quadruplexes, on which they could fold upon binding (Castello et al., 2016; Phan et al., 2011; Vasilyev et al., 2015). They are also known to be able to melt nucleic acids structure (Thandapani et al., 2013). For FUS, we observed a dynamic, "fuzzy" interaction between the RGG repeat and RNA. The RGG repeat does not adopt an ordered fold but destabilizes an RNA structure—the stacked 5′ end of the loop (Figure 5). Despite remaining disordered, the RGG repeat greatly increases FUS RNA binding affinity and could promote the binding of another protein. We speculate that the UA-Y motif identified by PAR-CLIP (Hoell et al., 2011) could then originate from the binding of an unknown partner protein and not from FUS itself.

Positively charged tails are found to greatly enhance binding affinity to structured RNA by folding and interacting with the stem (Amarasinghe et al., 2000; De Guzman et al., 1998). More recently, positively charged residues C-terminal to the Rbfox1 RRM (Chen et al., 2016) were found to increase binding of the core RRM to pre-miR20b by interacting in the major groove of the stem (Figure 7G). In the case of Rbfox1, the C-terminal region destabilizes the stem. What we found with FUS is therefore reminiscent to what was seen with Rbfox1, except that the FUS RGG interaction might also be crucial for recruiting additional proteins.

RNA recognition has been traditionally thought to be achieved by globular proteins and domains. In contrast, genome-wide studies capturing all RNA binding proteins in the cells (Castello et al., 2012) have found a surprisingly large enrichment in intrinsically disordered regions (IDRs). Our work shows how the intrinsically disordered RGG repeats in FUS could enhance RNA binding affinity and contribute to the re-modeling of RNA structure without adopting an ordered conformation. Conversely, it has been shown that RNA forms less ordered structures in vivo compared to in vitro, largely thought to be due to the presence of RNA binding proteins (Spitale et al., 2015). Intrinsically disordered RNA binding regions, such as RGG motifs, could contribute to this both by directly binding RNA and also by destabilizing RNA structure, enabling additional folded RNA binding domains to bind. It will be interesting to see whether this is a widespread mode of interaction by RGG-rich intrinsically disordered regions of RBPs.

Synergistic RNA Binding by the Multiple Domains of FUS

We have shown how the individual domains of FUS can each bind RNA and with which specificity. Do these domains bind RNA synergistically or independently? Chemical shift mapping

CellPress

of FUS RRM-RGG2-ZnF (Figures 1B and S1E) show that both the RRM and the ZnF domains can bind RNAs containing a stem loop and a GGU motif. Electrophoretic mobility shift assay (EMSA) also indicates improved binding compared to the individual domains, potentially reflecting the active contribution of RGG2 (Figure S5A). Mutational analysis shows that RNA binding by the RRM and ZnF do act synergistically in splicing of SCN4 minigene, whereas no synergistic effect was observed in regulation of FUS and TAF15 mRNA levels, suggesting this effect could be context dependent. SHAPE-Map and *ex vivo* SHAPE analysis of Xist long non-coding RNA cross referenced with CLIP data mapped direct FUS binding sites to regions of single-stranded RNA surrounded by structured RNA (Smola et al., 2016), with some of the sites, including stem loops, adjacent to a GGU motif. Interestingly, this region of Xist was shown to undergo changes in RNA structure *in vivo* compared to *ex vivo*, consistent with our proposal that FUS can remodel RNA structures.

Finally, RNA binding by FUS is important in facilitating efficient liquid-liquid phase separation into membrane-less compartments like liquid droplets (Burke et al., 2015), the aging of which could lead to aggregation of FUS in ALS and FTLD patients (Patel et al., 2015). The modular nature of FUS RNA binding and the weak RNA binding affinity of the folded domains shown here are perfectly consistent with the weak multivalent interactions known to facilitate phase transition (Li et al., 2012). This multivalent RNA binding, together with the disordered regions of FUS, are well suited to play a role in the formation of the different phases, and the RRM and ZnF have recently been shown to contribute to RNA-mediated phase separation of FUS (Maharana et al., 2018). Furthermore, the role of RGG regions in destabilizing structured regions of RNA in addition to direct binding may further facilitate this process. The importance and relative roles of the different RNA binding domains in this process remain to be explored.

In conclusion, we deciphered the first structural basis for RNA recognition by FUS. This provides an important step toward understanding in more detail the role of RNA in both FUS loss-of-function and toxic gain-of-function mechanisms in order to shed light on the role of FUS in ALS and FTLD pathology.

STAR★METHODS

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

- KEY RESOURCES TABLE
- CONTACT FOR REAGENT AND RESOURCE SHARING
- METHOD DETAILS
  - Expression and Purification FUS
  - RNA Preparation
  - NMR Spectroscopy
  - NMR Titrations and Complex Preparation
  - NMR Measurements and Assignments
  - Structure Calculation and Refinement
  - RRM-RGGx3:hnRNPA2/B1 Stem-Loop Model
  - Isothermal Titration Calorimetry
  - Band Shift Assays

- Searching for YNY Motif in FUS CLIP Datasets
- Minor Intron Splicing Assay
- FUS and TAF15 mRNA Regulation
- DATA AND SOFTWARE AVAILABILITY

REFERENCES

Amarasinghe, G.K., De Guzman, R.N., Turner, R.B., Chancellor, K.J., Wu, Z.R., and Summers, M.F. (2000). NMR structure of the HIV-1 nucleocapsid protein bound to stem-loop SL2 of the psi-RNA packaging signal. Implications for genome recognition. J. Mol. Biol. *301*, 491–511.

Burke, K.A., Janke, A.M., Rhine, C.L., and Fawzi, N.L. (2015). Residue-by-residue view of in vitro FUS granules that bind the C-terminal domain of RNA polymerase II. Mol. Cell *60*, 231–241.

Case, D.A., Cheatham, T.E., 3rd, Darden, T., Gohlke, H., Luo, R., Merz, K.M., Jr., Onufriev, A., Simmerling, C., Wang, B., and Woods, R.J. (2005). The Amber biomolecular simulation programs. J. Comput. Chem. *26*, 1668–1688.

Castello, A., Fischer, B., Eichelbaum, K., Horos, R., Beckmann, B.M., Strein, C., Davey, N.E., Humphreys, D.T., Preiss, T., Steinmetz, L.M., et al. (2012). Insights into RNA biology from an atlas of mammalian mRNA-binding proteins. Cell *149*, 1393–1406.

Castello, A., Fischer, B., Frese, C.K., Horos, R., Alleaume, A.M., Foehr, S., Curk, T., Krijgsveld, J., and Hentze, M.W. (2016). Comprehensive identification of RNA-binding domains in human cells. Mol. Cell *63*, 696–710.

Cathala, G., and Brunel, C. (1990). Use of n-butanol for efficient recovery of minute amounts of small RNA fragments and branched nucleotides from dilute solutions. Nucleic Acids Res. *18*, 201.

Chen, Y., Zubovic, L., Yang, F., Godin, K., Pavelitz, T., Castellanos, J., Macchi, P., and Varani, G. (2016). Rbfox proteins regulate microRNA biogenesis by

CellPress

sequence-specific binding to their precursors and target downstream Dicer. Nucleic Acids Res. 44, 4381–4395.

Daigle, J.G., Lanson, N.A., Jr., Smith, R.B., Casci, I., Maltare, A., Monaghan, J., Nichols, C.D., Kryndushkin, D., Shewmaker, F., and Pandey, U.B. (2013). RNA-binding ability of FUS regulates neurodegeneration, cytoplasmic mislocalization and incorporation into stress granules associated with FUS carrying ALS-linked mutations. Hum. Mol. Genet. 22, 1193–1205.

De Guzman, R.N., Wu, Z.R., Stalling, C.C., Pappalardo, L., Borer, P.N., and Summers, M.F. (1998). Structure of the HIV-1 nucleocapsid protein bound to the SL3 psi-RNA recognition element. Science 279, 384–388.

Dormann, D., Madl, T., Valori, C.F., Bentmann, E., Tahirovic, S., Abou-Ajram, C., Kremmer, E., Ansorge, O., Mackenzie, I.R., Neumann, M., and Haass, C. (2012). Arginine methylation next to the PY-NLS modulates Transportin binding and nuclear import of FUS. EMBO J. 31, 4258–4275.

Duss, O., Maris, C., von Schroetter, C., and Allain, F.H. (2010). A fast, efficient and sequence-independent method for flexible multiple segmental isotope labeling of RNA using ribozyme and RNase H cleavage. Nucleic Acids Res. 38, e188.

Ebrahimi, M., Rossi, P., Rogers, C., and Harbison, G.S. (2001). Dependence of 13C NMR chemical shifts on conformations of rna nucleosides and nucleotides. J. Magn. Reson. 150, 1–9.

Guerrero, E.N., Wang, H., Mitra, J., Hegde, P.M., Stowell, S.E., Liachko, N.F., Kraemer, B.C., Garruto, R.M., Rao, K.S., and Hegde, M.L. (2016). TDP-43/FUS in motor neuron disease: complexity and challenges. Prog. Neurobiol. 145-146, 78–97.

Güntert, P., and Buchner, L. (2015). Combined automated NOE assignment and structure calculation with CYANA. J. Biomol. NMR 62, 453–471.

Guo, L., Kim, H.J., Wang, H., Monaghan, J., Freyermuth, F., Sung, J.C., O'Donovan, K., Fare, C.M., Diaz, Z., Singh, N., et al. (2018). Nuclear-import receptors reverse aberrant phase transitions of RNA-binding proteins with prion-like domains. Cell 173, 677–692.e20.

Herrmann, T., Güntert, P., and Wüthrich, K. (2002). Protein NMR structure determination with automated NOE-identification in the NOESY spectra using the new software ATNOS. J. Biomol. NMR 24, 171–189.

Hoell, J.I., Larsson, E., Runge, S., Nusbaum, J.D., Duggimpudi, S., Farazi, T.A., Hafner, M., Borkhardt, A., Sander, C., and Tuschl, T. (2011). RNA targets of wild-type and mutant FET family proteins. Nat. Struct. Mol. Biol. 18, 1428–1431.

Hofweber, M., Hutten, S., Bourgeois, B., Spreitzer, E., Niedner-Boblenz, A., Schifferer, M., Ruepp, M.D., Simons, M., Niessing, D., Madl, T., et al. (2018). Phase separation of FUS is suppressed by its nuclear import receptor and arginine methylation. Cell 173, 706–719.e13.

Iko, Y., Kodama, T.S., Kasai, N., Oyama, T., Morita, E.H., Muto, T., Okumura, M., Fujii, R., Takumi, T., Tate, S., and Morikawa, K. (2004). Domain architectures and characterization of an RNA-binding protein, TLS. J. Biol. Chem. 279, 44834–44840.

Ishigaki, S., Masuda, A., Fujioka, Y., Iguchi, Y., Katsuno, M., Shibata, A., Urano, F., Sobue, G., and Ohno, K. (2012). Position-dependent FUS-RNA interactions regulate alternative splicing events and transcriptions. Sci. Rep. 2, 529.

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

Kapeli, K., Pratt, G.A., Vu, A.Q., Hutt, K.R., Martinez, F.J., Sundararaman, B., Batra, R., Freese, P., Lambert, N.J., Huelga, S.C., et al. (2016). Distinct and shared functions of ALS-associated proteins TDP-43, FUS and TAF15 revealed by multisystem analyses. Nat. Commun. 7, 12143.

Kondo, Y., Oubridge, C., van Roon, A.M., and Nagai, K. (2015). Crystal structure of human U1 snRNP, a small nuclear ribonucleoprotein particle, reveals the mechanism of 5′ splice site recognition. eLife 4, e04986.

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

Laskowski, R.A., Rullmannn, J.A., MacArthur, M.W., Kaptein, R., and Thornton, J.M. (1996). AQUA and PROCHECK-NMR: programs for checking the quality of protein structures solved by NMR. J. Biomol. NMR 8, 477–486.

Lee, W., Tonelli, M., and Markley, J.L. (2015). NMRFAM-SPARKY: enhanced software for biomolecular NMR spectroscopy. Bioinformatics 31, 1325–1327.

Lerga, A., Hallier, M., Delva, L., Orvain, C., Gallais, I., Marie, J., and Moreau-Gachelin, F. (2001). Identification of an RNA binding specificity for the potential splicing factor TLS. J. Biol. Chem. 276, 6807–6816.

Li, P., Banjade, S., Cheng, H.C., Kim, S., Chen, B., Guo, L., Llaguno, M., Hollingsworth, J.V., King, D.S., Banani, S.F., et al. (2012). Phase transitions in the assembly of multivalent signalling proteins. Nature 483, 336–340.

Ling, S.C., Polymenidou, M., and Cleveland, D.W. (2013). Converging mechanisms in ALS and FTD: disrupted RNA and protein homeostasis. Neuron 79, 416–438.

Liu, X., Niu, C., Ren, J., Zhang, J., Xie, X., Zhu, H., Feng, W., and Gong, W. (2013). The RRM domain of human fused in sarcoma protein reveals a noncanonical nucleic acid binding site. Biochim. Biophys. Acta 1832, 375–385.

Loughlin, F.E., Mansfield, R.E., Vaz, P.M., McGrath, A.P., Setiyaputra, S., Gamsjaeger, R., Chen, E.S., Morris, B.J., Guss, J.M., and Mackay, J.P. (2009). The zinc fingers of the SR-like protein ZRANB2 are single-stranded RNA-binding domains that recognize 5′ splice site-like sequences. Proc. Natl. Acad. Sci. USA 106, 5581–5586.

Maharana, S., Wang, J., Papadopoulos, D.K., Richter, D., Pozniakovsky, A., Poser, I., Bickle, M., Rizk, S., Guillén-Boixet, J., Franzmann, T.M., et al. (2018). RNA buffers the phase separation behavior of prion-like RNA binding proteins. Science 360, 918–921.

Maris, C., Dominguez, C., and Allain, F.H. (2005). The RNA recognition motif, a plastic RNA-binding platform to regulate post-transcriptional gene expression. FEBS J. 272, 2118–2131.

Masuda, A., Takeda, J., Okuno, T., Okamoto, T., Ohkawara, B., Ito, M., Ishigaki, S., Sobue, G., and Ohno, K. (2015). Position-specific binding of FUS to nascent RNA regulates mRNA length. Genes Dev. 29, 1045–1057.

Murakami, T., Qamar, S., Lin, J.Q., Schierle, G.S., Rees, E., Miyashita, A., Costa, A.R., Dodd, R.B., Chan, F.T., Michel, C.H., et al. (2015). ALS/FTD mutation-induced phase transition of FUS liquid droplets and reversible hydrogels into irreversible hydrogels impairs RNP granule function. Neuron 88, 678–690.

Murray, D.T., Kato, M., Lin, Y., Thurber, K.R., Hung, I., McKnight, S.L., and Tycko, R. (2017). Structure of FUS protein fibrils and its relevance to self-assembly and phase separation of low-complexity domains. Cell 171, 615–627.e16.

Nakaya, T., Alexiou, P., Maragkakis, M., Chang, A., and Mourelatos, Z. (2013). FUS regulates genes coding for RNA-binding proteins in neurons by binding to their highly conserved introns. RNA 19, 498–509.

Neumann, M., Rademakers, R., Roeber, S., Baker, M., Kretzschmar, H.A., and Mackenzie, I.R. (2009). A new subtype of frontotemporal lobar degeneration with FUS pathology. Brain 132, 2922–2931.

Neumann, M., Bentmann, E., Dormann, D., Jawaid, A., DeJesus-Hernandez, M., Ansorge, O., Roeber, S., Kretzschmar, H.A., Munoz, D.G., Kusaka, H., et al. (2011). FET proteins TAF15 and EWS are selective markers that distinguish FTLD with FUS pathology from amyotrophic lateral sclerosis with FUS mutations. Brain 134, 2595–2609.

Nguyen, C.D., Mansfield, R.E., Leung, W., Vaz, P.M., Loughlin, F.E., Grant, R.P., and Mackay, J.P. (2011). Characterization of a family of RanBP2-type zinc fingers that can recognize single-stranded RNA. J. Mol. Biol. 407, 273–283.

Oubridge, C., Ito, N., Evans, P.R., Teo, C.H., and Nagai, K. (1994). Crystal structure at 1.92 A resolution of the RNA-binding domain of the U1A spliceosomal protein complexed with an RNA hairpin. Nature 372, 432–438.

Ozdilek, B.A., Thompson, V.F., Ahmed, N.S., White, C.I., Batey, R.T., and Schwartz, J.C. (2017). Intrinsically disordered RGG/RG domains mediate degenerate specificity in RNA binding. Nucleic Acids Res. 45, 7984–7996.

CellPress

Patel, A., Lee, H.O., Jawerth, L., Maharana, S., Jahnel, M., Hein, M.Y., Stoynov, S., Mahamid, J., Saha, S., Franzmann, T.M., et al. (2015). A liquid-to-solid phase transition of the ALS protein FUS accelerated by disease mutation. Cell *162*, 1066–1077.

Peterson, R.D., Theimer, C.A., Wu, H., and Feigon, J. (2004). New applications of 2D filtered/edited NOESY for assignment and structure elucidation of RNA and RNA-protein complexes. J. Biomol. NMR *28*, 59–67.

Phan, A.T., Kuryavyi, V., Darnell, J.C., Serganov, A., Majumdar, A., Ilin, S., Raslin, T., Polonskaia, A., Chen, C., Clain, D., et al. (2011). Structure-function studies of FMRP RGG peptide recognition of an RNA duplex-quadruplex junction. Nat. Struct. Mol. Biol. *18*, 796–804.

Price, S.R., Evans, P.R., and Nagai, K. (1998). Crystal structure of the spliceosomal U2B"-U2A' protein complex bound to a fragment of U2 small nuclear RNA. Nature *394*, 645–650.

Qamar, S., Wang, G., Randle, S.J., Ruggeri, F.S., Varela, J.A., Lin, J.Q., Phillips, E.C., Miyashita, A., Williams, D., Ströhl, F., et al. (2018). FUS phase separation is modulated by a molecular chaperone and methylation of arginine cation-π interactions. Cell *173*, 720–734.e15.

Raczynska, K.D., Ruepp, M.D., Brzek, A., Reber, S., Romeo, V., Rindlisbacher, B., Heller, M., Szweykowska-Kulinska, Z., Jarmolowski, A., and Schümperli, D. (2015). FUS/TLS contributes to replication-dependent histone gene expression by interaction with U7 snRNPs and histone-specific transcription factors. Nucleic Acids Res. *43*, 9711–9728.

Ramaswami, M., Taylor, J.P., and Parker, R. (2013). Altered ribostasis: RNA-protein granules in degenerative disorders. Cell *154*, 727–736.

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

Reber, S., Stettler, J., Filosa, G., Colombo, M., Jutzi, D., Lenzken, S.C., Schweingruber, C., Bruggmann, R., Bachi, A., Barabino, S.M., et al. (2016). Minor intron splicing is regulated by FUS and affected by ALS-associated FUS mutants. EMBO J. *35*, 1504–1521.

Reber, S., Mechtersheimer, J., Nasif, S., Benitez, J.A., Colombo, M., Domanski, M., Jutzi, D., Hedlund, E., and Ruepp, M.D. (2018). CRISPR-Trap: a clean approach for the generation of gene knockouts and gene replacements in human cells. Mol. Biol. Cell *29*, 75–83.

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

Schwartz, J.C., Ebmeier, C.C., Podell, E.R., Heimiller, J., Taatjes, D.J., and Cech, T.R. (2012). FUS binds the CTD of RNA polymerase II and regulates its phosphorylation at Ser2. Genes Dev. *26*, 2690–2695.

Schwartz, J.C., Wang, X., Podell, E.R., and Cech, T.R. (2013). RNA seeds higher-order assembly of FUS protein. Cell Rep. *5*, 918–925.

Shang, Y., and Huang, E.J. (2016). Mechanisms of FUS mutations in familial amyotrophic lateral sclerosis. Brain Res. *1647*, 65–78.

Shen, Y., Delaglio, F., Cornilescu, G., and Bax, A. (2009). TALOS+: a hybrid method for predicting protein backbone torsion angles from NMR chemical shifts. J. Biomol. NMR *44*, 213–223.

Skrisovska, L., Bourgeois, C.F., Stefl, R., Grellscheid, S.N., Kister, L., Wenter, P., Elliott, D.J., Stevenin, J., and Allain, F.H. (2007). The testis-specific human protein RBMY recognizes RNA through a novel mode of interaction. EMBO Rep. *8*, 372–379.

Smola, M.J., Christy, T.W., Inoue, K., Nicholson, C.O., Friedersdorf, M., Keene, J.D., Lee, D.M., Calabrese, J.M., and Weeks, K.M. (2016). SHAPE reveals transcript-wide interactions, complex structural domains, and protein interactions across the Xist lncRNA in living cells. Proc. Natl. Acad. Sci. USA *113*, 10322–10327.

Spitale, R.C., Flynn, R.A., Zhang, Q.C., Crisalli, P., Lee, B., Jung, J.W., Kuchelmeister, H.Y., Batista, P.J., Torre, E.A., Kool, E.T., and Chang, H.Y. (2015). Structural imprints in vivo decode RNA regulatory mechanisms. Nature *519*, 486–490.

Takahama, K., Takada, A., Tada, S., Shimizu, M., Sayama, K., Kurokawa, R., and Oyoshi, T. (2013). Regulation of telomere length by G-quadruplex telomere DNA- and TERRA-binding protein TLS/FUS. Chem. Biol. *20*, 341–350.

Takahama, K., Miyawaki, A., Shitara, T., Mitsuya, K., Morikawa, M., Hagihara, M., Kino, K., Yamamoto, A., and Oyoshi, T. (2015). G-quadruplex DNA- and RNA-specific-binding proteins engineered from the RGG domain of TLS/FUS. ACS Chem. Biol. *10*, 2564–2569.

Tan, A.Y., Riley, T.R., Coady, T., Bussemaker, H.J., and Manley, J.L. (2012). TLS/FUS (translocated in liposarcoma/fused in sarcoma) regulates target gene transcription via single-stranded DNA response elements. Proc. Natl. Acad. Sci. USA *109*, 6030–6035.

Thandapani, P., O'Connor, T.R., Bailey, T.L., and Richard, S. (2013). Defining the RGG/RG motif. Mol. Cell *50*, 613–623.

Vandesompele, J., De Preter, K., Pattyn, F., Poppe, B., Van Roy, N., De Paepe, A., and Speleman, F. (2002). Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. Genome Biol *3*, RESEARCH0034.

Vasilyev, N., Polonskaia, A., Darnell, J.C., Darnell, R.B., Patel, D.J., and Serganov, A. (2015). Crystal structure reveals specific recognition of a G-quadruplex RNA by a β-turn in the RGG motif of FMRP. Proc. Natl. Acad. Sci. USA *112*, E5391–E5400.

Wang, X., Schwartz, J.C., and Cech, T.R. (2015). Nucleic acid-binding specificity of human FUS protein. Nucleic Acids Res. *43*, 7535–7543.

Yoshizawa, T., Ali, R., Jiou, J., Fung, H.Y.J., Burke, K.A., Kim, S.J., Lin, Y., Peeples, W.B., Saltzberg, D., Soniat, M., et al. (2018). Nuclear import receptor inhibits phase separation of FUS through binding to multiple sites. Cell *173*, 693–705.e22.

Zhou, Y., Liu, S., Liu, G., Oztürk, A., and Hicks, G.G. (2013). ALS-associated FUS mutations result in compromised FUS alternative splicing and autoregulation. PLoS Genet. *9*, e1003895.

CellPress

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
Rabbit anti-FUS	described in Raczynska et al., 2015	N/A
anti-Flag M2 antibody	Sigma-Aldrich	F1804; RRID: AB_262044
mouse anti-tyrosine tubulin	Sigma-Aldrich	T9028; RRID: AB_261811
Bacterial and Virus Strains		
Escherichia coli BL21(DE3) codon plus RIL	Stratagene	230240
Critical Commercial Assays		
QuikChange Lightning Multi Site-Directed Mutagenesis Kit	Agilent	210515
SuperScript III First-Strand Synthesis SuperMix	Life Technologies	18080400
Deposited Data		
ZnF:UGGUG	PBD	PDB: 6G99
RRM:RNA stem-loop	PDB	PDB: 6GBM
ZnF:UGGUG Chemical Shifts	BMRB	BMRB: 34258
RRM:RNA stem-loop Chemical Shifts	BMRB	BMRB: 34259
Experimental Models: Cell Lines		
FUS KO cells HeLa	Reber et al., 2016	N/A
Oligonucleotides		
Oligonucleotides used for structural studies	This study	Table S4
Primers for mutation and mRNA levels and splicing assays	This study	Table S5 and S6
Recombinant DNA		
pET28a -TEV-RRM-RGG2-ZnF (269–454)	This study	N/A
pET28a -TEV-RRM-RGG2-ZnF (280–454)	This study	N/A
pET28a -TEV-RRM (280–377)	This study	N/A
pET28a -TEV-RRM (269–377)	This study	N/A
pET28a -TEV-RRM$_{RGGx3}$ (269–390)	This study	N/A
pET28a -TEV-RRM $_{RGGx3}$ (260–390)	This study	N/A
pGEX6P-ZnF (422–454)	This study	N/A
pGEX-6P-ZnF (418–454)	This study	N/A
pET28a-RRM-RGG2-ZnF (269–454) F288A,K315A,K316A, Y325A, F348A, R441A, N445A	This study	N/A
pcDNA6F-FUS-GSG15-FLAG	Reber et al., 2016	N/A
pcDNA6F-FUS-GSG15-FLAG F288A,K315A,K316A,Y325A, F348A, R441A, N445A	This study	N/A
SCN4 minigene	Raczynska et al., 2015	N/A
Software and Algorithms		
Topspin 3.2	Bruker	N/A
Sparky / NMR-FAM	Lee et al., 2015	http://nmrfam.wisc.edu/software/
Unio ATNOS CANDID	Herrmann et al., 2002	N/A
CYANA 3.96	Güntert and Buchner, 2015	L.A. Systems
AMBER 12	Case et al., 2005	N/A
PYMOL	Schrodinger	https://pymol.org/2/

CellPress

CONTACT FOR REAGENT AND RESOURCE SHARING

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Frederic Allain (allain@mol.biol.ethz.ch).

METHOD DETAILS

Expression and Purification FUS
RRM, RRM-RGG-ZnF, RRM$_{RGGx3}$, and ZnF Subcloning

DNA fragments encoding RRM, RRM-$_{RGGx3}$ and RRM-RGG2-ZnF were subcloned using *NdeI* and *XhoI* restriction sites into pET28a plasmid which had been modified to include to TEV restriction site. DNA fragments encoding FUS ZnF were subcloned into pGEX6P-1 using BamHI and EcoRI.

ZnF Protein Expression and Purification

ZnF fragments transformed into BL21 codon-plus cells (RIL) and expressed in M9 minimal media supplemented with ^{15}N ammonium chloride (1g/L) and or ^{13}C glucose (1g/L) thiamine and trace elements in the presence of 50 μg/mL Carbenicillin and 34 μg/mL Chloramphenicol Typically 2L of culture were grown at 37°C to log phase (OD$_{600}$ 0.6-0.8) then shifted to 25°C, media was supplemented with zinc chloride to 0.1mM expression was induced with isopropyl β-D-1-thiogalactopyranoside to 0.5mM overnight. Cells were harvested by centrifugation and resuspended in lysis buffer of 50 mM Tris-HCl (pH 8.0), 150 mM NaCl, 2 mM β-mercaptoethanol, 0.5% Triton X-100, 5 mM MgCl$_2$, 0.1 mM ZnCl$_2$ supplemented with 10 mg Lysozyme and Complete-EDTA-Free Protease inhibitor (Roche). Cell were lysed by two freeze thaw cycles, the addition of DNase I at room temperature, followed by sonication. Lysate was clarified by centrifugation at 38,000 x g.

The clarified lysate was loaded onto glutathione-Sepharose 4B beads (GE Healthcare) pre-equilibrated with the lysis buffer for 30 min at 4°C. Beads were washed with 50 mM Tris-HCl (pH 8), 300 mM NaCl, 10% Glycerol 1mM β-mercaptoethanol and eluted with 50 mM Tris-HCl (pH 8.0) 150mM NaCl 20 mM glutathione 1mM β-mercaptoethanol. The fusion protein was dialysed into cleavage buffer at 4°C 50 mM Tris-Cl pH 7.0 at RT (7.5 at 4°C) 150 mM NaCl, 1 mM DTT, 0.001% Triton X-100 and cleaved with PreScission (purified in house) leaving 5 additional N-terminal residues (GPLGS). Cleaved solution was adjusted to pH 8.0 and passed over glutathione-Sepharose 4B beads to removed cleaved GST tag and cleaved protein concentrated and further purified by size exclusion chromatography on a Sephadex 75 10/30 column (GE healthcare) in 50 mM HEPES pH 7.2, 150 mM NaCl, 1 mM DTT, 5 μM ZnCl$_2$ and protein was concentrated and stored in this buffer. Final protein was analyzed for nucleic acid contamination using A$_{260nm}$/A$_{280nm}$ and concentration was estimated using A$_{280nm}$ using a theoretical extinction coefficient of 4523 M^{-1}cm^{-1} and stored at −80°C. Before use, ZnF protein was dialysed into NMR buffer 20 mM Na$_2$HPO$_4$/ NaH$_2$PO$_4$, 1 mM β-mercaptoethanol, 5 μM ZnCl$_2$.

RRM and RRM-RGG2-ZnF Protein Expression and Purification

RRM, RRM$_{RGGx3}$ and RRM-RGG2-ZnF constructs were expressed as per ZnF in the presence of 50 μg/mL Kanamycin and 34 μg/mL Chloramphenicol. Protein expression was induced at OD$_{600}$ 0.6-0.8 with 1 mM with isopropyl β-D-1-thiogalactopyranoside at 22°C overnight. Cell were harvested by centrifugation and resuspended in lysis buffer of 50 mM Na$_2$HPO$_4$/ NaH$_2$PO$_4$ pH 8, 1M NaCl, 2 mM β-mercaptoethanol, 0.5% Triton X-100, 0.1 mM ZnCl$_2$ supplemented with 5 mM MgCl$_2$ 10 mg Lysozyme and Complete-EDTA-Free Protease inhibitor (Roche). Cell were lysed by two freeze thaw cycles, the addition of DNase I at room temperature, followed by sonication. Lysate was clarified by centrifugation at 38,000 x g.

The clarified lysate was supplemented with imidazole to 5 mM then loaded onto Ni-NTA beads (QIAGEN) pre-equilibrated with the lysis buffer. Beads were washed stepwise with 50 mM Na$_2$HPO$_4$/ NaH$_2$PO$_4$ pH 8, 1M NaCl, 1 mM β-mercaptoethanol 20 mM imidazole, followed by 50 mM imidazole, then eluted with 50 mM Na$_2$HPO$_4$/ NaH$_2$PO$_4$ pH 8, 1M NaCl, 1 mM β-mercaptoethanol 300 mM imidazole. The fusion protein was dialysed into TEV cleavage buffer 50 mM Tris-Cl pH 8.0 50 mM NaCl 2 mM β-mercaptoethanol and cleaved with TEV (purified in house) at 4°C overnight leaving 4 additional N-terminal residues (GSHM). Cleaved solution was adjusted to 1M NaCl, 10 mM imidazole before and passing over Ni-NTA beads to remove His tag concentrated and stored at −80°C.

Cleaved protein of RRM constructs were incubated with SuperaseIn RNase inhibitor (Ambion) for 30 minutes and further purified by size exclusion chromatography on a Sephadex 75 10/30 column (GE healthcare) in, 50 mM Na$_2$HPO$_4$/NaH$_2$PO$_4$ pH 7, 150 mM NaCl, 1 mM DTT. Final protein purity was checked by SDS gels (Figure S6) and analyzed for nucleic acid contamination using A$_{260nm}$/A$_{280nm}$ and concentration was estimated using A$_{280nm}$ by calculating with the theoretical extinction coefficient of 8480 M^{-1}cm^{-1} (RRM, RRM$_{RGG3}$) and 20970 M^{-1}cm^{-1} (RRM-RGG2-ZnF) and stored at −80°C.

Before use, proteins were dialyzed into NMR/ITC buffer:

- RRM: 20 mM Na$_2$HPO$_4$/ NaH$_2$PO$_4$, 1 mM β-mercaptoethanol.
- RRM$_{RGGx3}$: 20 mM Na$_2$HPO$_4$/ NaH$_2$PO$_4$, 40 mM NaCl, 1 mM β-mercaptoethanol.
- RRM-RGG2-ZnF: 20 mM Na$_2$HPO$_4$/ NaH$_2$PO$_4$, 1 mM β-mercaptoethanol, 5 μM ZnCl$_2$.
- ITC buffer (all samples): 20 mM Na$_2$HPO$_4$/ NaH$_2$PO$_4$, 40 mM NaCl, 1 mM β-mercaptoethanol.

CellPress

RNA Preparation

Short RNA oligonucleotide (Table S4) were purchased from Dharmacon and deprotected in supplied deprotection buffer, 65°C 60 min, lyophilized and resuspended in water twice, before resuspending in D_2O or the appropriate buffer.

Longer RNAs were produced using *in vitro* RNA transcription (Table S4). Double stranded DNA templates were prepared by annealing two DNA oligonucleotides or cloning sequence of interest into pUC18 (stem-loop from hnRNPA2/B1) or pUC19 (stem-loop from *SON*) and linearized by cleavage with BbsI at 37°C overnight. RNA was transcribed for 4 hr at 37°C using with T7 RNA polymerase (purified in house) in 40 mM Tric-Cl pH 8.0 1 mM spermidine, 0.01% Triton-X, 5 mM DTT, with $MgCl_2$ and rNTP concentrations optimized for each transcript and the reaction stopped by the addition of EDTA. For ^{15}N-^{13}C labeled RNA stem-loops $^{15}N^{13}C$ labeled rNTPs purified in house were used. Reaction mixtures were then purified on denaturing anion exchange HPLC (Duss et al., 2010) and isolated using butanol extraction (Cathala and Brunel, 1990), resuspended in H_2O lyophilized. RNA pellets were then resuspended in 10 μM in H_2O, heated for 5 min at 65°C then snapped cooled in liquid nitrogen followed by lyophilization. RNA pellets were then resuspended in NMR buffer.

NMR Spectroscopy

All NMR Spectroscopy measurements were performed using Bruker AVIII 500 MHz, AVIII 700 MHz and Avance 900 MHz spectrometers equipped with cryoprobes. The data were processed using Topspin 3.1 (Bruker) and analyzed with Sparky (http://www.cgl.ucsf.edu/home/sparky/) and NMR-FAM-SPARKY (Lee et al., 2015).

NMR Titrations and Complex Preparation

NMR titrations of protein RNA complexes were performed by adding unlabeled concentrated RNA (1-5 mM) to ^{15}N or ^{13}C-^{15}N labeled protein (0.2-1.5 mM) in NMR buffer (or vice versa) and protein monitored by ^1H-^{15}N-HSQC, and RNA monitored by 1D and ^1H-^1H-TOCSY (tm 50 ms) at 303 K and 283 K. The ZnF:UGGUG complex was formed in a similar manner with an additional incubation step at 37°C to facilitate binding. Sample measurements were then performed at 303 K, 283 K or 278 K. For complex formation of FUS-RRM, ^{15}N labeled protein was titrated into $^{15}N^{13}C$ hnRNPA2/B1 stem-loop to monitor chemical shift changes in sugar, base, and imino resonances monitored by ^1H-^{13}C-HSQC and ^1H-^{15}N-HSQCs. Sample was lyophilized before resuspending in D_2O. Low salt buffers were required for RRM-RGG2-ZnF and ZnF ^1H-^{15}N HSQC titrations in order to avoid intermediate exchange of the ZnF resonances when bound to RNA.

NMR Measurements and Assignments

RRM-RGG2-ZnF Free and in Complex with RNAs

Sequence specific backbone and side chain assignments of protein were achieved using a classical approach. All measurements were at 303 K unless otherwise indicated. Sequence specific backbone assignments of RRM-RGG2-ZnF in free and bound to hnRNPA2/B1-GGU RNA were achieved using 2D ^1H-^{15}N HSQC, 2D ^1H-^{13}C-HSQC, 3D HNCA, 3D CBCACONH, 3D HNCACB, 3D HNCO, 3D HNCACO including 159 residues of 180 non proline residues (88%). The assignment of some stretches of sequential glycines in RGG2 was hampered due to overlap. Assignment of all other RNA bound forms RRM-RGG2-ZnF:RNA complexes were achieved by following peaks in NMR titrations. The heteronuclear ^1H-^{15}N values were measured employing water flip-back pulse.

ZnF:UGGUG Complex

Protein sequence specific backbone and side chain assignments of FUS-ZnF bound to UGGUG (^{13}C-^{15}N-ZnF:unlabelledRNA 1:1.3) were achieved using 3D HNCA, 3D H(C)CH-TOCSY (tm 25 ms) at 303 K and 3D ^1H-^{15}N NOESY and 3D ^1H-^{13}C NOESY (tm 120 ms) and 2D NOESY at 303 K and 278 K. Intramolecular ZnF constraints FUS-ZnF bound to UGGUG (ZnF:RNA 1:1.3) were taken from NOESY experiments at 303 K. Backbone PHI angles restraints were taken from 3D HNHA.

RNA was assigned from 2D TOCSY (tm 25 and 50 ms), DQF-COSY, 2D NOESY (tm 120 ms), natural abundance ^{13}C-HSQCs and ^1H-^{31}P COSY from a sample of ^{15}N-ZnF:unlabeled UGGUG (ZnF: RNA ratio 1:0.9) combined with a and 2D F_2 filtered NOESY (tm 120 ms) with a sample of ^{15}N-^{13}C ZnF in complex with unlabeled RNA (1:1). Sugar puckers in the complex were identified from 2D ^1H-^1H- TOCSY (tm 25 ms) and DQF-COSY, syn or anti conformations were identified from NOE patterns of H6 or H8 resonances. Intermolecular NOEs were identified from 2D NOESY and 2D F_2 ^{13}C-filtered NOESY (Peterson et al., 2004) in D_2O and 2D NOESY optimized for imino detection and 3D ^{15}N-NOESY with ^{15}N carrier frequency adjusted for optimal Arg Nε excitation both at 278 K. Intermolecular hydrogen bond constraints were derived from slow exchanging imino resonances and preliminary structures.

FUS RRM: RNA Stem-Loop hnRNPA2/B1 Complex

Protein sequence specific backbone and side chain assignments of RRM (280-377): hnRNPA2/B1 RNA stem-loop (^{13}C-^{15}N RRM:^{15}N or unlabeled RNA 1:1.2) were achieved using 3D HNCA, 3D H(C)CH-TOCSY and 3D (H)CCH-TOCSY (tm 25 ms) at 303 K. Intramolecular RRM constraints were incorporated from 3D ^1H-^{15}N NOESY and 3D ^1H-^{13}C NOESY (tm 120 ms) and 2D NOESY at 303 K. Protein dihedral backbone constraints derived from TALOS+ (Shen et al., 2009).

Stem-loop RNA in free and RRM bound forms was assigned using ^1H-^{13}C HSQCs, DQF-COSY, 2D TOCSY (tm 25 and 50 ms), 3D (H)CCH-TOCSY and 3D ^1H-^{13}C NOESYs in D_2O. Slow exchanging imino and amino resonances were assigned from

^{1}H-^{15}N- HSQCs 2D NOESY (tm 150 ms) in H$_2$O at 283K. RNA base pair constraints were derived from slow exchanging imino and amino resonances and cross-strand NOEs in RNA helical regions. Sugar puckers were identified by 2D TOCSY (tm 25ms) and confirmed by can1-can2 analysis of ^{13}C sugar chemical shifts (Ebrahimi et al., 2001) and syn or anti conformations identified from NOE patterns of H6 or H8 resonances. Intermolecular NOEs were identified in 2D F$_2$ ^{13}C-filtered NOESY and 3D ^{13}C-(F1 edited, F3 filtered) NOESY spectra with either the RRM ^{15}N-^{13}C labeled and RNA unlabeled or RRM ^{15}N labeled and RNA ^{13}C-^{15}N labeled and a 2D NOESY (tm 80ms) in D$_2$O with ^{15}N labeled RRM and unlabeled RNA. Detection and assignment of intermolecular NOEs were also confirmed and clarified using sample of RRM bound to hnRNPA2/B1 ΔA11 in which A11 was not present. This residue does not contact the protein and intermolecular protein-RNA contacts are conserved between the two complexes.

FUS RRM-$_{RGGx3}$: RNA Stem-Loop Complexes

Protein sequence specific backbone and side chain assignments of RRM (260-390) were achieved using 3D HNCA, CBCACONH, HNCA. HNCACO and HCCH-TOCSY. In complex with RNA stem-loops from hnRNPA2/B1 and SON, intermolecular NOEs were detected using 2D F$_2$ ^{13}C-filtered NOESY and 3D ^{13}C-(F1 edited, F3 filtered) NOESY spectra in D$_2$O. The heteronuclear {^{1}H}-^{15}N NOE values were measured as proposed employing water flip-back pulse.

Structure Calculation and Refinement

The resonance assignments of each protein bound to RNA was used as input for automatic peak picking and NOESY assignment using ATNOS-CANDID (Herrmann et al., 2002). Resulting peak lists were checked and supplemented manually. RNA and intermolecular NOESY peaks were picked and assigned manually and calibrated from H5-H6 peaks of pyrimidines. Protein peaks were then re-assigned with NOEASSIGN module of CYANA 3.96 (Güntert and Buchner, 2015) and manually checked.

For the ZnF:UGGUG complex, initial calculations were run in the absence of zinc ion. Once the zinc binding residues were identified from preliminary structures, loose SG-SG constraints were utilized. Automatic NOE assignment was run for residues 425-454 keeping an initial set of mannual assignments. Constraints from residues 418-424 were then added in subsequent rounds of calculations in the presence of RNA. Intermolecular hydrogen bonds were identified from slow exchanging imino resonances and preliminary structure calculations and explicit intermolecular hydrogen bonds and zinc constraints were only added in final structure calculations and refinement.

For the RRM:stem-loop complex, initial complex calculations, unambiguous intermolecular NOEs to the RRM (280–370) were included positioning the nucleotides, followed by inclusion of intermolecular NOEs to the C-terminal extension (R371–R377). Intermolecular NOE with ambiguous assignments were then included as ambiguous restraints in CYANA and assigned based on preliminary calculations. A number of intermolecular NOEs from A$_{12}$ incompatible with this structure were observed from RRM:hnRNPA2/B1 SL complex (listed in SI Table S2) in low salt buffer used in samples for structure determination. To further confirm the final intermolecular constraints, we used intermolecular NOEs from RRM:ΔA11-hnRNPA2/B1 complex using and stemloop with A11 deleted which showed a single set of intermolecular NOEs. The additional A$_{12}$ intermolecular NOEs present with RRM:hnRNPA2/B1 appear to be a subset of those detected in the RRM$_{RGGx3}$:hnRNPA2/B1 complex which was interpreted as a second binding event as shown with ITC data.

Final structure calculations in CYANA, restraints included intra protein, RNA and intermolecular NOEs, protein dihedral backbone constraints, intra protein hydrogen bond and RNA base pair constraints, sugar pucker, and syn or anti conformations identified from NOE patterns of H6 or H8 resonances and for stem-loop RNAs loose (± 30°) backbone dihedral backbone constraints based on standard A form geometry. For ZnF:UGGUG complex zinc ion was included using a CYSZ residue. Of 500 structures, the 50 lowest energy structures were selected for refinement with SANDER module of AMBER12 (Case et al., 2005) using ff12SB force field with implicit solvent and 20 were selected.

RRM-RGGx3:hnRNPA2/B1 Stem-Loop Model

A model of RRM-$_{RGGx3}$: hnRNPA2/B1 was calculated using intermolecular NOEs identified from RGG extension in combination with constraints from the core RRM:hnRNPA2/B1 stem-loop complex. Structures were calculated in CYANA and refined in Amber 12.

Isothermal Titration Calorimetry

ITC experiments were performed on a VP-ITC microcalorimeter (Microcal). Protein and RNA were extensively dialyzed in ITC buffer 20 mM Na$_2$HPO$_4$/ NaH$_2$PO$_4$, 40 mM NaCl, 1 mM β-mercaptoethanol. RRM, RRM-RGG$_{x3}$ or ZnF (400-800 μM) was titrated into RNA (20-40 μM) using a single 2 μL followed by 12 μL injections every 600 s at 30°C with a stirring rate of 307 rpm. Raw data were integrated was analyzed according to a' single set of sites' model in Origin 7.0. The data measured did not support reliable modelling as two independent sites.

Band Shift Assays

Radiolabeled Band Shifts

Purified RNA from *in vitro* RNA transcription reactions were dephosphorylated Antarctic phosphastase (NEB), buffer exchanges then 5′ end labeled with T4 polynucleotide kinase using (New England Biolabs) with γ-^{32}P-ATP (5000Ci/mmol) according to the manufacturer's instructions. Binding reactions were set up in a total volume of 30 μL and comprising of a constant concentration of ^{32}P-labeled probe (0.1 pmole), increasing concentrations of protein, in a buffer consisting of 20 mM HEPES 7.9 50 mM KCl

5 mM $MgCl_2$ 5% Glycerol,0.03 mg mL^{-1} heparin. Binding reactions were setup on ice followed by an incubation at 4°C for 30 min after which 15 μL of each sample was loaded onto a pre- run 8% native polyacrylamide gel made up in 0.5 x Tris-Borate and electrophosed at 100 V for 1 hr, then dried and analyzed using a PhosphorImager (Amersham Biosciences).

Non-radiolabeled Band Shifts
Binding reactions were setup in a similar manner except the RNA concentrations was 10 μM in a total volume was 20 μL of which 10 μL was loaded onto the gel. The RNA was visualized using 0.1% toluidine blue.

Searching for YNY Motif in FUS CLIP Datasets
For this analysis, we made use of processed data as provided on the relative GEO page of each study. These data represent genomic intervals determined as putative FUS binding sites on the relative mRNA by peak calling procedures. These regions were converted to the latest human genome assembly hg38, using UCSC liftOver utility. The nucleotide sequence has been retrieved by querying the genome with Samtools (v1.2). The presence of hairpin loop in these sequences has been determined via RNAfold (v2.1.9), with local setup and no-lonely-pairs option (RNALfold -noLP) in order to detect reliable stem structures instead of large-scale secondary structures. The plot shows the percentage of regions in each study that contain at least one bipartite motif. This quantity is influenced by many factors, most notably the average length of the peaks. For this reason, we compared and tested our results against a scrambled version of the nucleotide sequences as control.

Minor Intron Splicing Assay
The SCN4A minor intron splicing reporter assay with related RT-qPCR primers were described previously (Reber et al., 2016) and determination of endogenous FUS mRNA levels with related RT-qPCR primers are described in Reber et al. (2018). Quantification of FUS protein levels were normalized against tubulin or polyclonal rabbit anti-actin (20-33) [Sigma, A 5060]. The pcDNA6F-FUS-GSG15-FLAG plasmid described in Reber et al., 2016 was used to create the vectors coding for the FUS RNA-binding mutants (mutRRM, mutZNF and mutRRM/ZNF respectively) with the QuickChange Lightning Multi Site-Directed Mutagenesis Kit (Agilent) using primers in Table S5.

The SCN4A minor intron splicing reporter assay is done as follow. 2.5×10^5 HeLa cells, maintained in a humidified incubator at 37°C and 5% CO_2, where seeded per well of a 6-well plate in DMEM$^{+/+}$ (day 0; DMEM$^{+/+}$: Dulbecco's modified Eagle's medium (DMEM, (ThermoFisher Scientific, 32500035)) supplemented with 10% fetal calf serum (FCS (Amimed, Bioconcept, 2-01F30-I)), penicillin (100 IU/mL), and streptomycin (100 μg/mL)). On day 1, the cells were co-transfected with 400 ng of the SCN4A reporter plasmid, 500 ng pSUPuro plasmid and 600-800 ng (to allow for equal expression) of the corresponding FUS rescue plasmid using Dogtor (OZ Biosciences, DT51000)) according to the manufacturer's instructions. On day 2, the cells were split into a T25 flask and selection was started using 2 μg/ml puromycin (Santa Cruz Biotechnology: sc-108071A. Puromycin selection was maintained until day 4. On day 5, half of the cells were harvested using Trizol for subsequent standard RNA isolation. The purified RNA was DNase treated using the TURBO DNA-free Kit (ThermoFisher Scientific, AM1907) according to the manufacturer's manual. Reverse transcription of total RNA was performed using the AffinityScript Multiple Temperature cDNA Synthesis Kit (Agilent Technologies, #200436) according to the manufacturer's manual. RT-qPCR was performed using 3 μl cDNA, 1 x MESA GREEN qPCR Mastermix Plus for SYBRR Assay No ROX (Eurogentec, 05-SY2X-03+NRWOU) and each 8 μl forward and reverse primer in a total volume of 15 μl per reaction. Samples were measured in duplicates in a Rotorgene6000 (Corbett). The following cycling conditions were used: 95°C, 5 min; 95°C, 15 s; 60°C 1 min; 40 cycles. A melting curve was recorded from a temperature gradient from 65°C to 95°C, 5 s/°C. Analysis was performed using the Rotor-Gene 6000 Series Software V1.7. The other half of the cells were re-suspended in SDS-loading buffer for subsequent SDS-PAGE western blot analysis. Hereto, 1×10^5 cell equivalents were separated on a 8% SDS-PAGE and transferred to a nitrocellulose membrane (Whatman, Optitran BA-S 85) using a TE77 ECL Semi-Dry Transfer Unit (Amersham Biosciences). Membranes were cut in two pieces and blocked with 5% non-fat dry milk in 0.1% Tween in TBS and subsequently incubated for 2 hr at room temperature with the primary antibodies mouse anti-FLAG (Sigma-Aldrich, F1804), rabbit anti-FUS (homemade; Raczynska et al., 2015) or mouse anti-tyrosine tubulin (Sigma-Aldrich, T9028) and rabbit anti-actin (Sigma, A 5060) respectively. Thereafter, the membranes were washed 5 × 5 min with 0.1% Tween in TBS and subsequently incubated with the fluorescence-labeled secondary antibodies (IRDye 800CW Goat anti-Rabbit IgG (H + L) (LI-COR, 925-32211) and IRDye 680LT Goat anti-mouse IgG (H + L) (LI-COR, 926-68020)) for 1.5 h at room temperature. The dried membranes were analyzed with the Odyssey Infrared Imaging System (LI-COR). See Table S6 for qPCR primers.

FUS and TAF15 mRNA Regulation
HeLa wild-type and FUS knockout cells were grown in DMEM-F12 (Life Technologies) with 10% FBS (Life Technologies) and 1% penicillin/streptomycin (GIBCO). Transient plasmid transfections of pcDNA6F-FUS-GSG15-FLAG plasmid with RRM/ZnF mutations were achieved using Lipofectamine2000 Transfection Reagent (Life Technologies) according to manufacturer's protocol with 1ng/μl DNA. Cells were harvested in Trizol for standard RNA extraction or RIPA buffer for immunoblot analysis. After RNA extraction, cDNA was prepared using SuperScript III First-Strand Synthesis SuperMix (Life Technologies). qPCR was performed in duplicates from 6 biological replicates using FastStart Universal SYBR Green Master (Roche) and the primer sets (Table S5).

qPCR data were analyzed according to Vandesompele et al. (2002) with their provided excel macro tool. The statistical significance of two groups of results was determined by a two-tailed, paired t test and quantification graphs are displayed as mean ± SD. For

CellPress

protein analysis the concentration was adjusted based on BCA assay (Thermofisher) and lysates were boiled in loading buffer with reducing agent before loading the samples on Bolt 12% Bis-Tris gels. Gels were transferred onto nitrocellulose membranes using iBlot 2, which were blocked with 5% non-fat skimmed powder milk in PBS-Tween and probed with primary antibodies ON (anti-FUS, Bethyl A300-293A, 1:10 000; anti-GAPDH, Abcam, 1:5000; anti-FLAG, FG4R, Thermofisher, 1:500) followed by secondary HRP-conjugated goat anti mouse or rabbit IgG antibodies (1:5000, 1:10000, respectively Jackson Laboratories). Immunoreactivity was visualized by chemiluminescence (GE Healthcare).

DATA AND SOFTWARE AVAILABILITY

The accession number for ZnF:UGGUG reported in this paper is PDB: 6G99 and for the RRM:RNA stem-loop is PDB: 6GBM.

The accession number for ZnF:UGGUG chemical shifts reported in this paper is BMRB: 34258 and for RRM:RNA stem-loop chemical shifts is BMRB: 34259.

Supplemental Information

The Solution Structure of FUS Bound to RNA

Reveals a Bipartite Mode of RNA Recognition

with Both Sequence and Shape Specificity

Fionna E. Loughlin, Peter J. Lukavsky, Tamara Kazeeva, Stefan Reber, Eva-Maria Hock, Martino Colombo, Christine Von Schroetter, Phillip Pauli, Antoine Cléry, Oliver Mühlemann, Magdalini Polymenidou, Marc-David Ruepp, and Frédéric H.-T. Allain
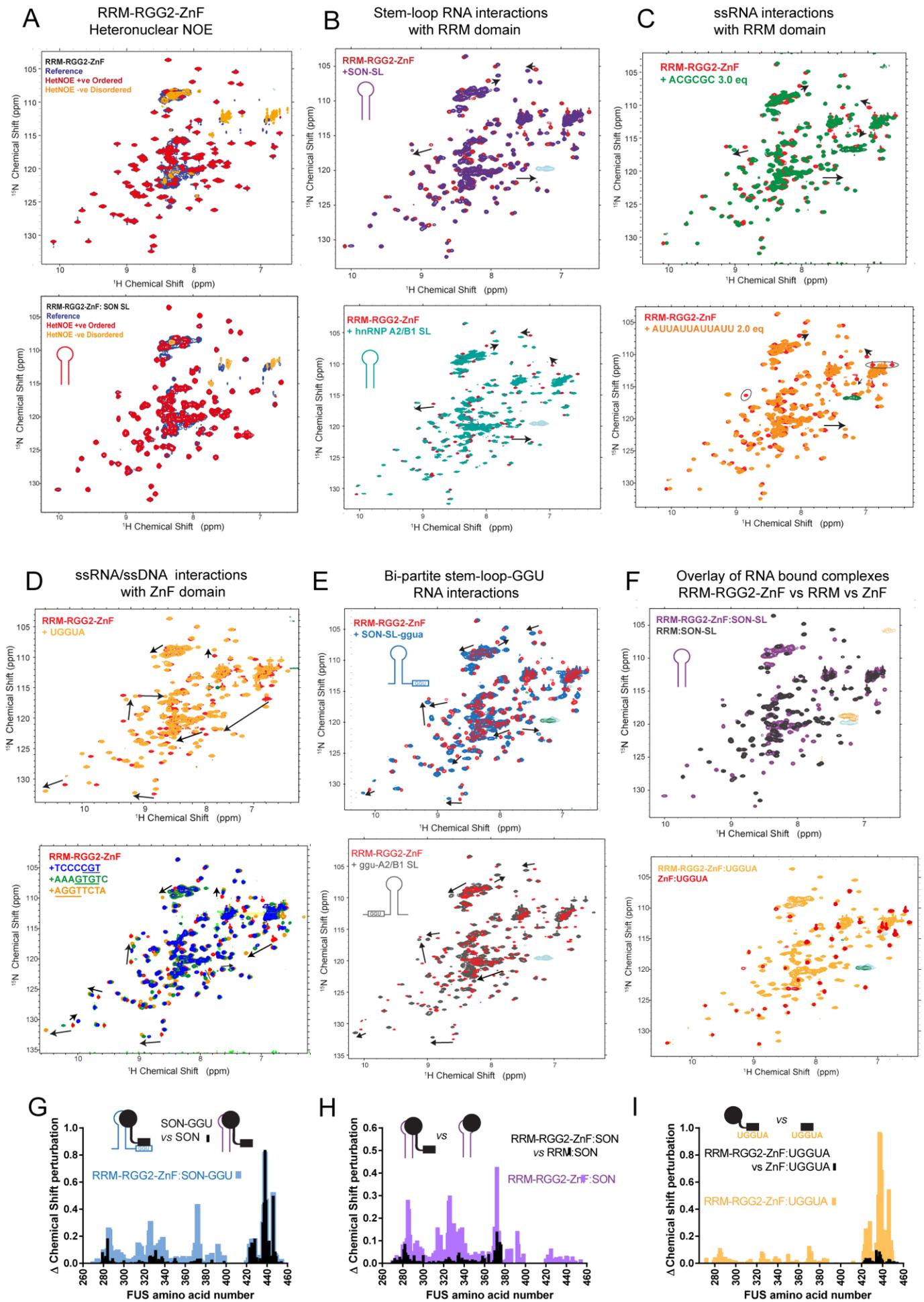
Figure S1. Mapping RNA binding sites in FUS RRM-RGG2-ZnF related to Figure 1.

A) ^{1}H-^{15}N Heteronuclear NOE free (*upper*) and bound to stem-loop of SON (*lower*). Reference (blue), NOE experiment (red-positive and yellow-negative). B-F) ^{1}H-^{15}N-HSQC of FUS RRM-RGG-ZnF free (red) and bound to RNA motifs (other color): B) Bound to stem-loops from SON (*upper,* purple) and hnRNPA2/B1 (*lower*, aqua) pre-mRNAs. C) Bound to single-stranded RNA ACGCGC (*upper*, green) and AUUAUUAUUAUU (*lower*, orange), D) Bound to single stranded RNA UGGUA (*upper*, yellow) and DNA (*lower*) TCCCCGT (blue), AAAGTGTC (green), AGGTTCTA (yellow), E) Bound to stem-loop RNA with an extension containing a GGU motif (SL+GGU) from SON (*upper*, light blue), and from hnRNPA2/B1 pre-mRNA (*lower*, grey), F) (*upper*) Overlay of RRM-RGG2-ZnF bound to SON stem-loop (purple) and of the isolated RRM bound to the SON SL (black). (*lower*) Overlay of RRM-RGG2-ZnF bound to UGGUA (yellow) and of the isolated ZnF bound UGGUA (red). G) Chemical shift difference between RRM-RGG2-ZnF bound to SON-GGU and SON (black) compared to the chemical shift perturbation of RRM-RGG2-ZnF on binding SON-GGU (blue). H) Chemical shift difference between RRM-RGG2-ZnF-SON SL and RRM:SON SL complexes (black) from overlay shown in F *upper* panel) compared to chemical shift perturbation of RRM-RGG2-ZnF binding to SON-SL RNA (purple, from spectra shown in B *upper* panel). I) Chemical shift differences between RRM-RGG2-ZnF:UGGUA and ZnF:UGGUA complexes (black, from overlay shown in B lower panel) compared to chemical shift perturbation of RRM-RGG2-ZnF in binding UGGUA RNA (yellow, shown in D upper panel).
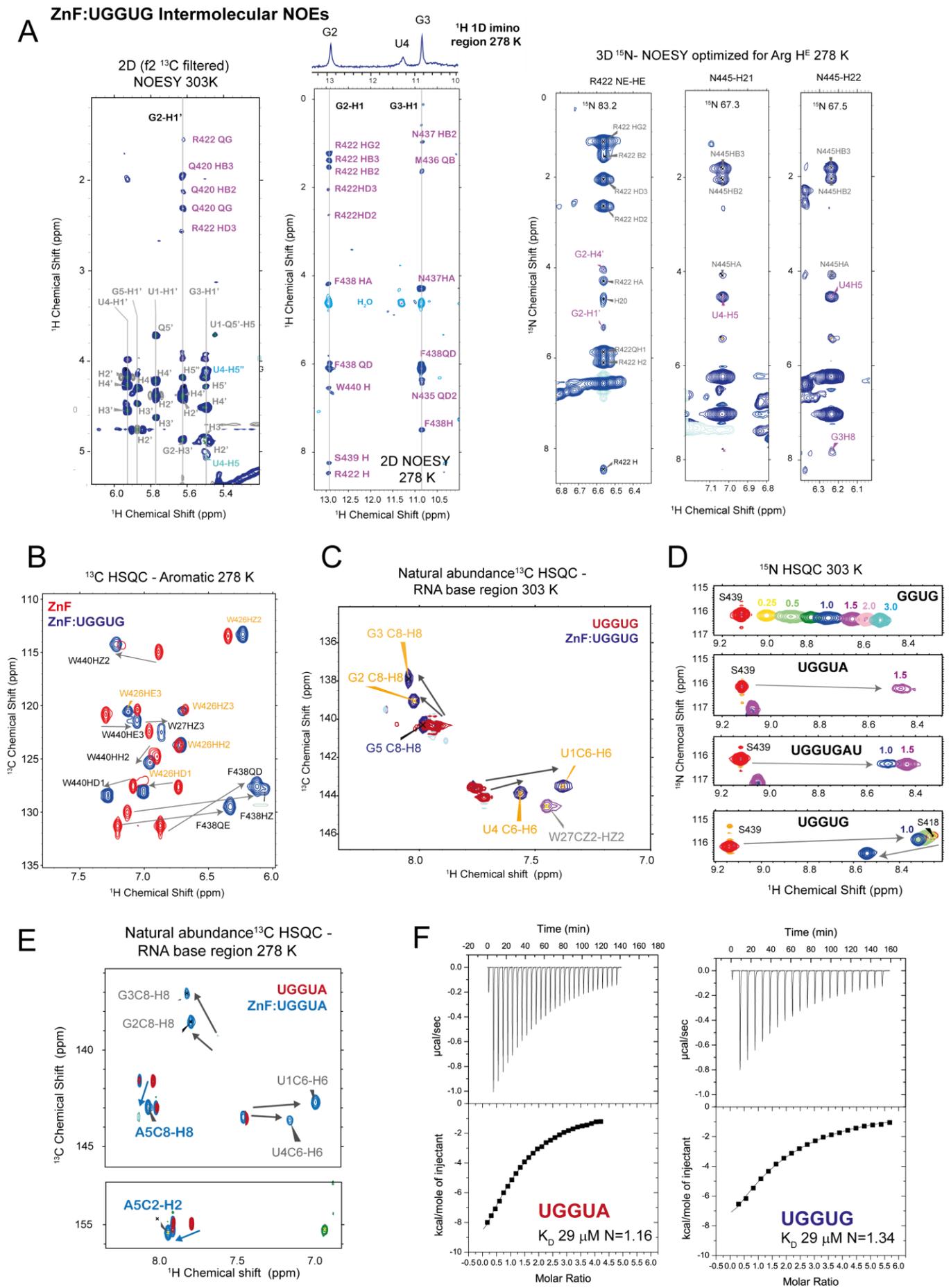
Figure S2. FUS ZnF binding to UGGUG, related to Figure 2.

A) Intermolecular NOEs in ZnF:UGGUG. Intermolecular NOEs between the RNA H1' resonances and the FUS resonances in a 2D f2 ^{13}C-filtered NOESY at 303K. (left) Intermolecular NOE from G_2 and G_3 imino H1 resonances and the FUS resonances in a 2D NOESY at 278K (middle). Cross sections from the 3D ^{1}H-^{15}N-NOESY at 278K of the complex showing NOEs from the exchangeable N_H/N_{H2} resonances of Arg 422 and Asn 445. Intermolecular NOEs are labelled in pink. B) ^{1}H-^{13}C-HSQC aromatic regions of Znf free (red) and bound to UGGUG (blue) at 283K. C) ^{1}H-^{13}C-HSQC spectra showing aromatic resonances of UGGUG RNA free (red) and FUS ZnF bound (blue). D) Effect of U_1 on chemical shift perturbations of ZnF as shows by representative residue S439. ZnF free (red) and in complex with 0.25 – 3.0 molar equivalents of ssRNA (colored). E) ^{1}H-^{13}C-HSQC spectra showing aromatic resonances of UGGUA RNA free (red) and bound to FUS ZnF (aqua) and resonance arising from ZnF (green). F) ITC measurements of ZnF titrated with UGGUG or UGGUA (right).
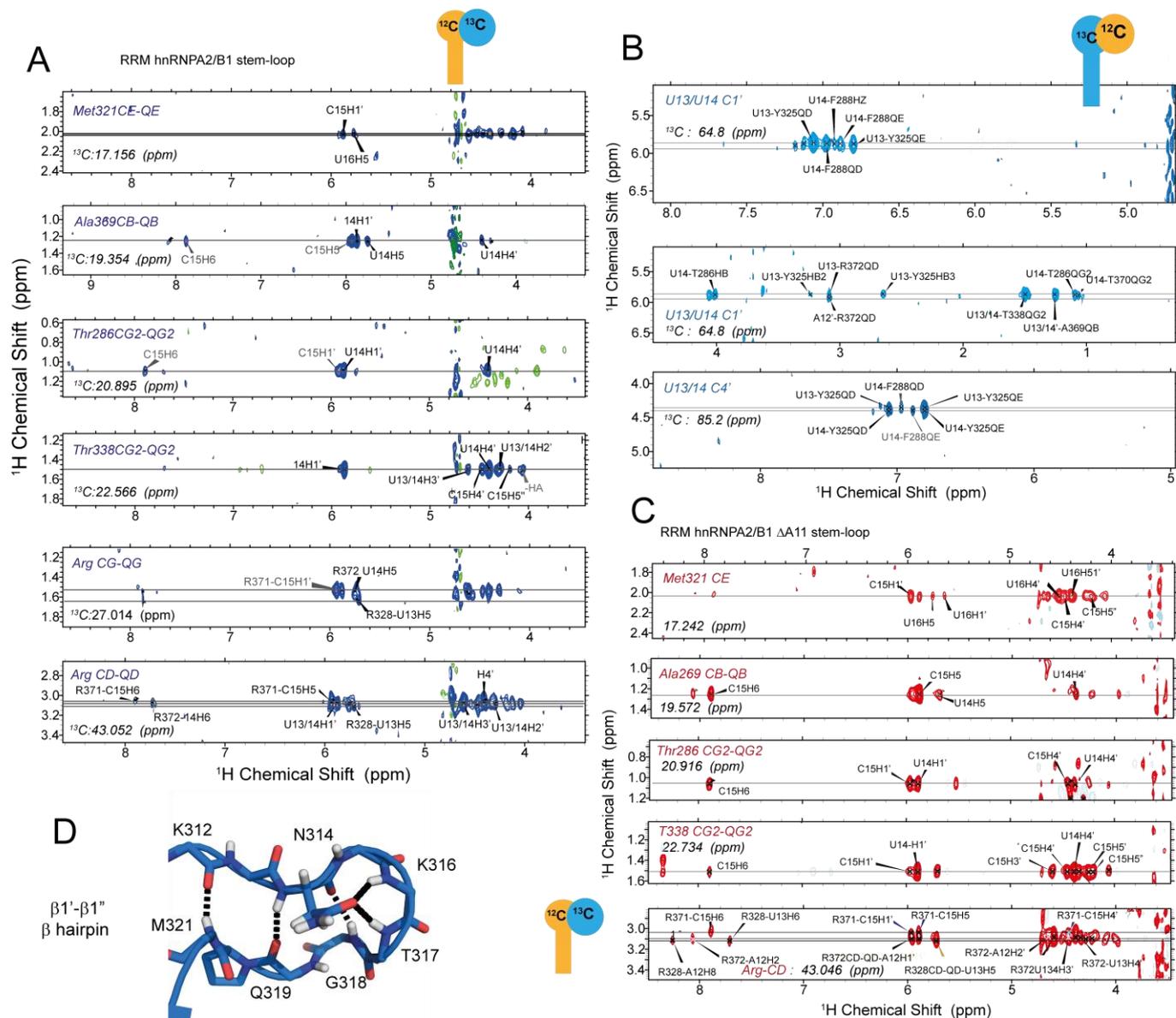
Figure S3. NMR spectra of RRM:RNA stem-loop hnRNPA2/B1, related to Figure 3.

A) Intermolecular NOEs 3D ^{13}C-filtered, ^{13}C-edited NOESY (t$_m$ 150-ms) ^{13}C^{15}N RRM: ^{1}H RNA. B) Intermolecular NOEs ^{13}C-filtered, ^{13}C-edited NOESY (t$_m$ 120-ms) in D2O 303K of ^{13}C, ^{15}N RNA: ^{15}N RRM. C) Intermolecular NOEs of RRM:RNA stem-loop hnRNPA2/B1 ΔA11, from 3D ^{13}C-filtered, ^{13}C-edited NOESY (t$_m$ 80-ms) ^{13}C^{15}N RRM: ^{1}H RNA. D) Hydrogen bonds present in β1'-β1" hairpin of FUS RRM.
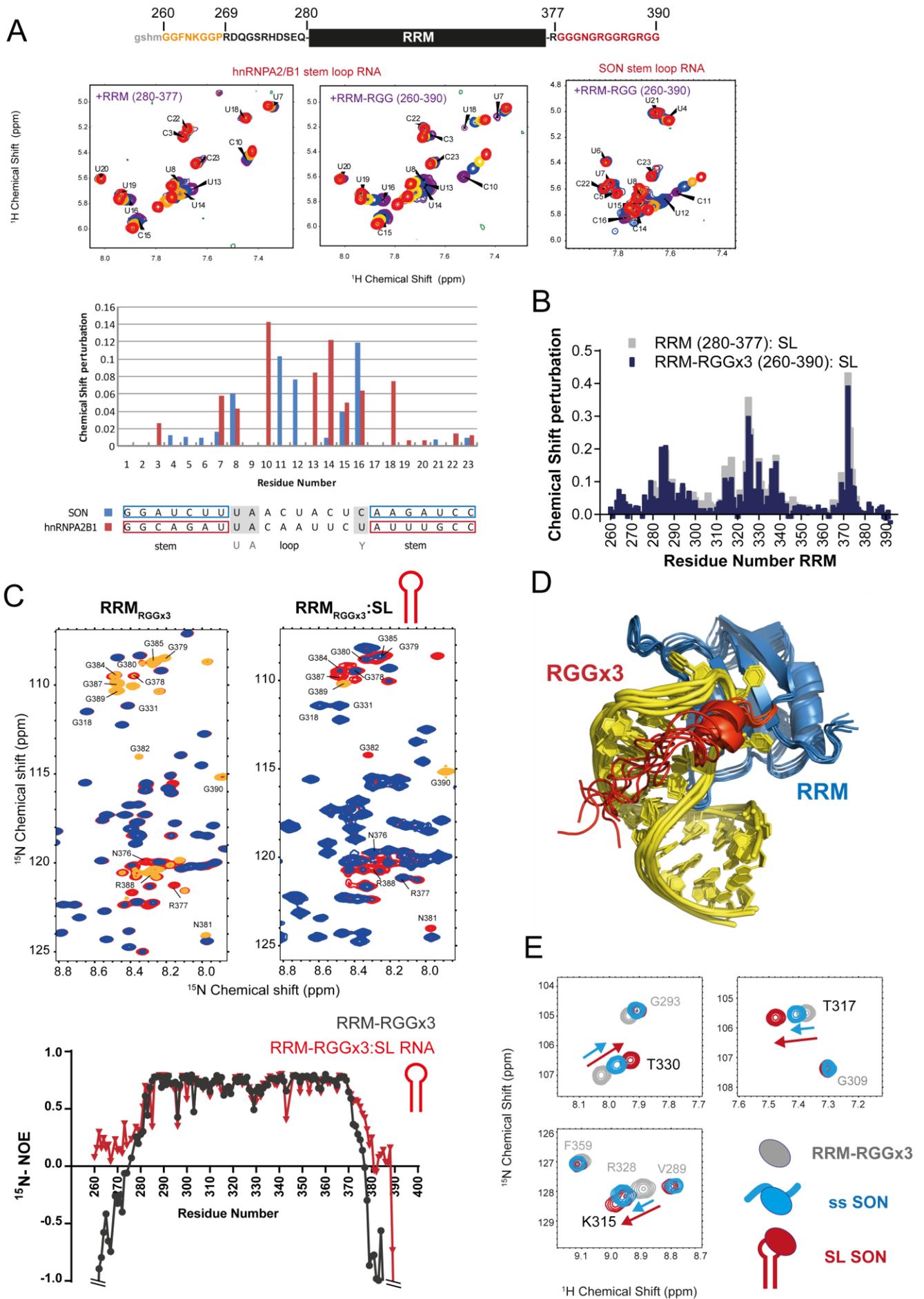
Figure S4. Effect of RGG repeats on RNA binding, related to Figure 5.

A) ^{1}H-^{1}H 2D TOCSY showing H5-H6 of pyrimidine of stem-loops free and on addition of RRM at molar equivalents 0.5 yellow, 1.0 blue and 2.0 purple and chemical shifts differences of H5 and H6. B) Chemical shift perturbation of RRM (280-377, grey) and RRMRGGx3 (260-390, blue) bound to hnRNPA2/B1 stem-loop in buffers 20 mM NaH$_2$PO$_4$/Na$_2$HPO$_4$ 2mM β-mercaptoethanol pH 6.5 and 20 mM NaH$_2$PO$_4$/Na$_2$HPO$_4$, 40 mM NaCl 2mM β-mercaptoethanol pH 6.5 respectively. C) {^{1}H}-15N Heteronuclear NOE at 310 K of RRM$_{RGGx3}$ free (left) and in complex with hnRNPA2/B1 stem loop. NOE (blue/yellow), reference experiment (red), with selected assignments resonances from RGG2 extension (amino acids 377-390) shown. D) NMR based model of RRM-RGGx3:hnRNPA2/B1 stem-loop showing 5 conformers showing C-terminal extension and RGG motifs (371-390, red), E) ^{15}N-HSQC of RRM-RGGx3 free (grey), bound to single stranded loop of SON (blue), and bound to SON stem-loop (red).

A

RRM | ZnF | RRM-ZnF

− 0.1 1.0 10 | 0.1 1.0 10 | 0.1 1.0 10 | µM

Complex

RNA

SON SL-GGU
GGU

B

FUS

1 | 269 | 454 | 526

QGSY rich PrD | G rich | RGG1 | RRM | RGG2 | Zn | RGG3

RRM-RGG2-ZnF mutant

F288A,K315A,K316A,Y325A RRM mut | *F438A, R441A, N445A* ZnF mut

RRM-RGG2-ZnF

WT | Mut

µM | 0.0 3.0 5.0 7.5 10 | 0.0 3.0 5.0 7.5 10

Complex

RNA

hnRNPA2/B1 GGU-SL
GGU

RRM-RGG2-ZnF

WT | Mut

µM | 0.0 3.0 5.0 7.5 10 | 0.0 3.0 5.0 7.5 10

Complex

RNA

SON SL-GGU
GGU

C

RRM-RGG2-ZnF-mut
+SON-GGU RNA

T330 | T317 | N284sc | C428

¹⁵N Chemical Shift (ppm)
¹H Chemical Shift (ppm)

WT
RRM-RGG2-ZnF
+SON-GGU RNA

mut
RRM-RGG2-ZnF
+SON-GGU RNA

RRM-N284sc

N284sc | N284sc

ZnF-C428

I311 | C428 | C428

¹⁵N Chemical Shift (ppm)
¹H Chemical Shift (ppm)

D

Endogenous FUS mRNA in SCN4 splicing assay

1.0

endogenous FUS mRNA levels normalized to 5.8S rRNA

0.03 | 0.02 | 0.01 | 0.01 | 0.01

FUS KD − + + + + +
Rescue − − WT FUS | mut RRM/ZnF | mut RRM | mut ZnF

E

Total FUS protein levels in SCN4 splicing assay

total FUS protein levels normalized to protein controls

FUS KD − + + + + +
Rescue − − WT FUS | mut RRM/ZnF | mut RRM | mut ZnF

F

Total exogenous FUS protein levels in Autoregulation of FUS mRNA levels assay

FLAG/GAPDH normalized

empty vector | WT FUS | mut RRM | mut ZnF | mut RRM/ZnF

G

Total FUS protein levels in Crossregulatoin TAF15 mRNA levels

FLAG/GAPDH normalized

HeLa Cells | WT | FUS-KO cells
vector | empty vector | empty vector | WT FUS | mut RRM | mut ZnF | mut RRM/ZnF

Figure S5. Domain contributions to binding of a bipartite motif RNA motif consisting of a stem-loop-GGU from SON pre-mRNA and functional analysis of RNA binding mutants, related to Figure 6

(A)Trace amounts of SON-ggu RNA was incubated with increasing amounts of protein RRM (269-377), ZnF (418-454) or RRM-RGG2-ZnF (269-454) and separated by electrophoresis. (-) indicate lane without protein. B) Schematic of FUS protein showing RRM and ZnF mutations used in this study (top). Band shift of RRM-RGG2-ZnF wt and mutant with GGU+SL RNAs stained with toluidine blue (bottom). C) Overlaid ^{15}N-HSQC RRM-RGG2-ZnF mutant (F288A, K315A, K316A, Y325A-F428AR441A, N445A), free (red) and bound to SON-GGU RNA (blue). Inserts of ^{1}H-^{15}N-HSQC spectra showing chemical shift perturbations of RRM-RGG2-ZnF WT and mutant binding to SON-GGU RNA (right). D) RT-qPCR results indicating the relative endogenous FUS mRNA levels normalized to 5.8S rRNA levels of control knockdown, FUS knockdown and different FUS rescue conditions in (the forward RT-qPCR primer binds in the 5'UTR of endogenous FUS mRNA which is not present in the different rescue plasmids coding for RNAi resistant FLAG-tagged FUS). Average and standard deviations of six biological replicates are shown. E) Quantification of total FUS protein relative to loading control in SCN4 minor intron splicing assays. Average and standard deviations of six biological replicates are shown. F) Relative overexpression levels of different exogenous FLAG-FUS constructs (empty vector control, WT-FUS, mut-RRM, mut-ZnF, mut-RRM/ZnF) in HeLa WT cells used for the endogenous FUS mRNA level assay (Fig. 6C) are shown normalized to loading control as mean with standard deviation for 5 biological replicates. G) Relative overexpression levels of different exogenous FLAG-FUS constructs (empty vector control, WT-FUS, mut-RRM, mut-ZnF, mut-RRM/ZnF) in HeLa WT and KO cells used for the eTAF15 mRNA level assay (Fig. 6D) are shown normalized to loading control as mean with standard deviation for 4 biological replicates.
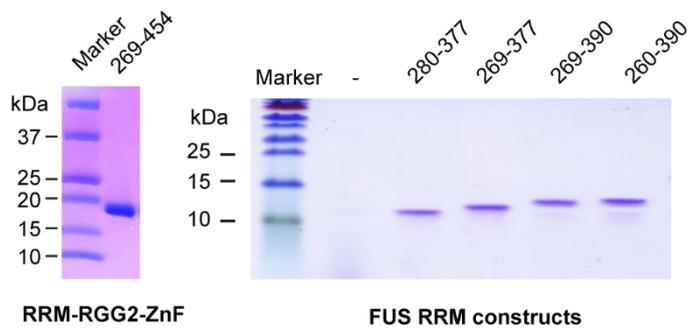
Supplementary Figure 6. SDS gel of various FUS constructs used in this study, related to STAR Method.

Table S1. Intermolecular NOEs from FUS ZnF:UGGUG at 303 K and 278 K, related to Figure 2.

RNA residue Number	Type	Atom	RRM Residue Number	Type	Atom	Distance
			303 K			
440	TRP	HD1	101	URA	H1'	5.00
440	TRP	HE1	101	URA	H1'	5.00
440	TRP	HZ2	101	URA	H1'	4.00
440	TRP	HH2	101	URA	H1'	6.00
440	TRP	HZ2	101	URA	H2'	4.00
440	TRP	HH2	101	URA	H2'	5.00
440	TRP	HZ3	101	URA	H5	5.00
440	TRP	HH2	101	URA	H5	6.00
420	GLN	QG	102	RGUA	H1'	4.00
420	GLN	HB2	102	RGUA	H1'	5.00
420	GLN	HB3	102	RGUA	H1'	5.00
422	ARG	QG	102	RGUA	H1'	5.00
422	ARG	HD3	102	RGUA	H1'	6.00
422	ARG	HD2	102	RGUA	H1'	6.00
438	PHE	QD	102	RGUA	H1'	6.00
438	PHE	QE	102	RGUA	H1'	5.00
438	PHE	HZ	102	RGUA	H1'	5.00
438	PHE	QE	102	RGUA	H2'	3.00
438	PHE	HZ	102	RGUA	H2'	3.00
438	PHE	QE	102	RGUA	H3'	4.00
438	PHE	QD	102	RGUA	H3'	6.00
438	PHE	HZ	102	RGUA	H3'	6.00
438	PHE	QD	102	RGUA	H8	6.00
438	PHE	QE	102	RGUA	H8	5.00
438	PHE	HB2	102	RGUA	H8	6.00
438	PHE	HB3	102	RGUA	H8	6.00
440	TRP	HZ2	102	RGUA	H8	5.00
440	TRP	HE1	102	RGUA	H8	4.00
440	TRP	HD1	102	RGUA	H8	5.00
438	PHE	HZ	103	RGUA	H51	5.00
438	PHE	HZ	103	RGUA	H52	5.00
438	PHE	QE	103	RGUA	H51	4.00
438	PHE	QE	103	RGUA	H52	5.00
438	PHE	QD	103	RGUA	H1'	6.00
438	PHE	QE	103	RGUA	H1'	6.00
438	PHE	QD	103	RGUA	H8	5.00
438	PHE	QE	103	RGUA	H8	4.00
438	PHE	HZ	103	RGUA	H8	6.00
445	ASN	QD2	104	URA	H5	5.00
			278 K			
440	TRP	HE1	102	RGUA	H8	5.00
440	TRP	HE1	102	RGUA	H1'	6.00
422	ARG	HE	102	RGUA	H1'	5.50
422	ARG	HE	102	RGUA	H2'	5.50
422	ARG	QH2	102	RGUA	H1'	6.00
422	ARG	QH2	102	RGUA	H4'	6.00
422	ARG	QH1	102	RGUA	H3'	6.00
439	SER	H	102	RGUA	H1	5.50
422	ARG	H	102	RGUA	H1	5.00
438	PHE	QD	102	RGUA	H1	4.00
438	PHE	HA	102	RGUA	H1	4.50
422	ARG	HB2	102	RGUA	H1	4.50
422	ARG	HB3	102	RGUA	H1	4.00
422	ARG	QG	102	RGUA	H1	4.00
422	ARG	HD2	102	RGUA	H1	5.50
422	ARG	HD3	102	RGUA	H1	6.00
438	PHE	H	103	RGUA	H1	3.50
435	ASN	HD21	103	RGUA	H1	5.50
435	ASN	HD22	103	RGUA	H1	5.00
437	ASN	HA	103	RGUA	H1	3.00

436	MET	QB	103	RGUA	H1	5.00
437	ASN	HB2	103	RGUA	H1	5.00
437	ASN	HB3	103	RGUA	H1	6.00
445	ASN	HD21	104	URA	H5	3.50
445	ASN	HD22	104	URA	H5	4.00
445	ASN	HD22	103	RGUA	H8	5.50
441	ARG	HE	103	RGUA	H1	6.00
441	ARG	HE	103	RGUA	H8	5.00
441	ARG	QH2	103	RGUA	H8	6.00

Table S2. Intermolecular NOEs from FUS RRM (280-377) in complex with hnRNP A2/B1 and Δ A11 RNA stem-loops, related to Figure 3.

RNA residue Number	Type	Atom	RRM Residue Number	Type	Atom	Distance
12	RADE	H8	325	TYR	QD	6.00
12	RADE	H8	325	TYR	QE	5.50
12	RADE	H8	328	ARG	QD	5.50
12	RADE	H8	328	ARG	QG	6.00
12	RADE	H2	325	TYR	QD	4.00
12	RADE	H2	325	TYR	QE	3.00
12	RADE	H2	323	ASN	HB2	5.00
12	RADE	H2	323	ASN	HB3	5.00
12	RADE	H1'	325	TYR	QD	5.00
12	RADE	H1'	325	TYR	QE	5.00
12	RADE	H1'	377	ARG	QD	6.00
12	RADE	H4'	377	ARG	QD	5.00
13	URA	H1'	325	TYR	QE	3.50
13	URA	H1'	325	TYR	QD	2.80
13	URA	H1'	325	TYR	HB2	5.50
13	URA	H1'	325	TYR	HB3	5.50
13	URA	H1'	334	LYS	HE3	6.00
13	URA	H1'	334	LYS	QD	6.00
13	URA	H4'	325	TYR	QE	3.50
13	URA	H4'	325	TYR	QD	4.50
13	URA	H3'	372	ARG	QD	3.00
14	URA	H3'	372	ARG	HB2	4.00
14	URA	H3'	372	ARG	HB3	4.00
14	URA	H3'	372	ARG	QG	3.50
13	URA	H5	328	ARG	QD	4.00
13	URA	H5	328	ARG	QG	4.00
13	URA	H5	328	ARG	QB	4.00
13	URA	H6	328	ARG	QD	6.00
14	URA	H5	288	PHE	QE	4.00
14	URA	H5	288	PHE	HZ	4.00
14	URA	H5	338	THR	QG2	6.00
14	URA	H1'	288	PHE	QD	3.00
14	URA	H1'	288	PHE	QE	4.50
14	URA	H1'	338	THR	QG2	3.00
14	URA	H1'	338	THR	HB	6.00
14	URA	H1'	369	ALA	QB	3.00
14	URA	H1'	369	ALA	HA	6.00
14	URA	H1'	286	THR	QG2	4.00
14	URA	H1'	286	THR	HB	6.00
14	URA	H1'	370	THR	QG2	5.00
14	URA	H1'	370	THR	HA	6.00
14	URA	H4'	325	TYR	QE	4.50
14	URA	H4'	325	TYR	QD	5.50
14	URA	H4'	338	THR	QG2	3.00
14	URA	H4'	286	THR	QG2	5.50
14	URA	H4'	369	ALA	QB	5.00
15	RCYT	H1'	286	THR	QG2	4.00
15	RCYT	H1'	321	MET	QE	4.00
15	RCYT	H1'	338	THR	QG2	6.00
15	RCYT	H2'	286	THR	QG2	5.00
15	RCYT	H2'	321	MET	QE	6.00
15	RCYT	H2'	371	ARG	QD	3.00
15	RCYT	H2'	371	ARG	QG	6.00
15	RCYT	H4'	321	MET	QE	5.00
15	RCYT	H4'	338	THR	QG2	4.00
15	RCYT	H4'	286	THR	QG2	4.50
15	RCYT	H5	369	ALA	QB	4.00
15	RCYT	H5	371	ARG	QD	6.00

15	RCYT	H5	371	ARG	QG	6.00
15	RCYT	H6	371	ARG	QD	5.50
15	RCYT	H6	371	ARG	QG	6.00
15	RCYT	H6	369	ALA	QB	4.00
15	RCYT	H6	286	THR	QG2	4.00
15	RCYT	H5'	338	THR	QG2	4.00
15	RCYT	H5"	338	THR	QG2	4.00
16	URA	H4'	321	MET	QE	5.00
16	URA	H5'	321	MET	QE	6.00
16	URA	H5"	321	MET	QE	6.00
16	URA	H5	321	MET	QE	6.00
16	URA	H6	321	MET	QE	6.00
16	URA	H1'	321	MET	QE	6.00
17	RADE	H8	316	LYS	QG	6.00

Additional Intermolecular NOEs present in hnRNP A2/B1 complex interpreted as arising from a minor form					
12	*RADE*	*H2*	*369*	*ALA*	*QB*
12	*RADE*	*H1'*	*369*	*ALA*	*QB*
12	*RADE*	*H2*	*370*	*THR*	*QG2*
13	*URA*	*H1'*	*321*	*MET*	*QE*

Table S3. ITC measurements of FUS constructs and RNA species in 20 mM Na₂HPO₄/NaH₂PO₄, 40 mM NaCl, 1 mM β-mercaptoethanol, pH 6.5 at 30 ºC, related to Figure 4. We have modelled these measurements as K_d^{app} and ΔG^{app} as a number of samples have multiple binding events (*)

FUS	amino acids	RNA	K_d^{app}, (µM)	$\Delta G^{app}=$ $\Delta H-T\Delta S$ (kcal/mol)	ΔH^{app} (kcal/mol)	$-T\Delta S$ (kcal/mol)	N
RRM	280-377	*hnRNPA2/B1* SL	116 ± 2.7	-4093	-1.1 ± 0.1	5.15	= 1
	269-377	*hnRNPA2/B1* SL	84.0 ± 1.2	-5648	-17.3± 0.3	11.7	1.06
	269-377 mut F288A/Y325A/K315A/K316A	*hnRNPA2/B1* SL	n.d.	n.d.	n.d.	n.d.	n.d.
RRM-RGGₓ₃	269-390	*hnRNPA2/B1* SL	13.9 ± 0.2*	-6744*	-13.0 ±0.7	6.30	1.64
	260-390	*hnRNPA2/B1* SL	10.1 ± 0.6*	-6941*	-15.5 ±0.4	8.58	1.44
	260-390	*SON* SL	13.1 ± 0.7*	-6759*	-19.3 ± 0.4	12.5	1.58
	260-390	ACGCGC	27.2 ± 2.6	-4691	-5.5 ± 0.5	0.82	0.99
	260-390	AAUAAA	120 ± 5.9	-8763	-5.8 ± 0.2	0.34	= 1
	260-390 mut N284A/N285A	*hnRNPA2/B1* SL	10.5 ± 0.2*	-6912*	-16.4 ± 0.4	3.6	1.74
	260-390 mut N314A/Q319A/M321A	*hnRNPA2/B1* SL	7.0 ± 0.2*	-7151*	-19.5 ± 0.4	12.4	1.65
	260-390 mut R371A/R372A	*hnRNPA2/B1* SL	28.8 ± 0.9*	-3024*	-4.6 ± 0.09	1.64	1.66
	260-390 mut K315A/K316A	*hnRNPA2/B1* SL	23.2 ± 0.6*	-6432*	-10.1 ±0.2	3.7	1.73
	260-390 mut F288A/Y325A/K315A/K316A	*hnRNPA2/B1* SL	46 ± 2*	-6024*	-16.2 ± 1.5	10.2	1.77
ZnF	318-454	UGGUA	29.8 ± 1.2	-31411	-18.9 ± 0.8	12.6	1.16
	318-454	UGGUG	29.5 ± 3.0	-32165	-19.2 ±0.2	12.9	1.34

Table S4: RNA and ssDNA sequences used for structural studies. Nucleotides included to stabilize the stem are underlined. Related to STAR Method

Name	Sequence	Preparation
SON – SL	GGAUCUUUAACUACUCAAGAUCC	T7 transcription
SON-GGU	GGATCTTTAACTACTCAAGATACTGAACATGACATGGTA	T7 transcription
hnRNPA2/B1 SL	<u>GG</u>CAGAUUACAAUUCUAUUU<u>GCC</u>	T7 transcription
hnRNPA2/B1 ΔA11 SL	<u>GG</u>CAGAUUACAUUCUAUUU<u>GCC</u>	
GGU-hnRNPA2/B1	GATTAGGTTTTGTGAGTAGACAGATTACAATTCTATTTTAA	T7 transcription
ssSON	UAACUACUC	Chemical synthesis
ACGCGC	ACGCGC	Chemical synthesis
AUU rich	AUUAUUAUUAUU	Chemical synthesis
GGUG	GGUG	Chemical synthesis
UGGUG	UGGUG	Chemical synthesis
UGGUA	UGGUA	Chemical synthesis
GUGGU	GUGGU	Chemical synthesis
AAUAAA	AAUAAA	Chemical synthesis
ssDNA1	TCCCCGT	Chemical synthesis
ssDNA2	AAAGTGTC	Chemical synthesis
ssDNA3	AGGTTCTA	Chemical synthesis

Table S5: Primers used for Splicing assays related to STAR Method

Name	Oligonucleotide sequence	assay
F288A	CAGACAACAACACCATCGCTGTGCAAGGCCTGGGTG	mutation
K315A;K316A	GCCAAATTAATCATGGGCTGTCCCGTTGCTGCGTTTGTCTTAATAATACCAATCTGCTTGAAGTAATCA	mutation
Y325A	5'-GGACAGCCCATGATTAATTTGGCCACAGACAGGGAAACTG-3'	mutation
F348A; R441A; N445A	GTTTAGGGGCCTTACACTGGGCGCATTCATTCGCCCAAGAGGCGTTCATATTCTCAC	mutation
TAF15:	GCCTCCTATGCAGCTCAGTC, GGTTGTAACCCCCTGTGCTA;	mRNA levels
FLAG-FUS:	CAAGGACGACGACGACAA, ACTGCTCTGCTGGGAATAGC;	mRNA levels
end.human FUS:	GTGGTGGCAGAGGAGGCTATGATCG, CAGGACAAAAAGCTGTTCCAG;	mRNA levels
NDRG2:	CAGGAGGTGCAGATCACAGA, AGTCCCACATCGTGGTAGGT;	mRNA levels
NDRG2:	AGTCCCACATCGTGGTAGGT;	mRNA levels
MAG:	GTCCTGTTCAGCAGCGACTT, AGCGTGTAGCTGTCCTTGGT;	mRNA levels
KCND3:	GGCAAGACCACCTCACTCAT, GCTGGACAGTGAGGGACTTC;	mRNA levels
ACTINB:	AGAAAATCTGGCACCACACC, AGAGGCGTACAGGGATAGCA;	mRNA levels
HMBS:	GGCAATGCGGCTGCA, GGGTACCCACGCGAATCAC;	mRNA levels
HPRT1:	TGACACTGGCAAAACAATGCA, GGTCCTTTTCACCAGCAAGCT;	mRNA levels
SDHA:	TGGGAACAAGAGGGCATCTG, CCACCACTGCATCAAATTCAT;	mRNA levels
UBC:	ATTTGGGTCGCGGTTCTT, TGCCTTGACATTCTCGATGGT;	mRNA levels

Table S6: qPCR Primers used for Splicing assays related to STAR Method

primer	5'-sequence-3'	Use
sybr SCN4A f	ACAAGGGCAAGGCCATCTTC	qPCR (both SCN4A assays)
sybr SCN4Aspl r	CATGCTGAACAGCGCATGG	qPCR (SCN4A spliced)
sybr SCN4Aun r	GGTCAAGGAAAGTGAGGAAGCAG	qPCR (SCN4A unspliced)
5.8S rRNA fwd	GGTGGATCACTCGGCTCGT	qPCR (normalizer)
5.8S rRNA rev	GCAAGTGCGTTCGAAGTGTC	qPCR (normalizer)
sybr FUS f	AGCGGTGTTGGAACTTCG	qPCR (specific for endogenous FUS mRNA)
sybr FUS r	GACTGCTCTGCTGGGAATAG	qPCR (specific for endogenous FUS mRNA