The *Drosophila* RNA binding protein Hrp48 binds a specific RNA sequence of the *msl-2* mRNA 3' UTR to regulate translation

Andrea Lomoschitz [a,1], Julia Meyer [a,b,1], Tanit Guitart [c], Miroslav Krepl [d], Karine Lapouge [e], Clara Hayn [a], Kristian Schweimer [b], Bernd Simon [a,f], Jiří Šponer [d,g], Fátima Gebauer [c,h], Janosch Hennig [a,b,*]

[a] *Molecular Systems Biology Unit, European Molecular Biology Laboratory (EMBL), 69117 Heidelberg, Germany*
[b] *Department of Biochemistry IV – Biophysical Chemistry, University of Bayreuth, 95447 Bayreuth, Germany*
[c] *Centre for Genomic Regulation (CRG), The Barcelona Institute of Science and Technology, Dr Aiguader 88, 08003 Barcelona, Spain*
[d] *Institute of Biophysics of the Czech Academy of Sciences, Kralovopolska 135, 612 65 Brno, Czech Republic*
[e] *Protein Expression and Purification Core Facility, European Molecular Biology Laboratory Heidelberg, 69117 Heidelberg, Germany*
[f] *Molecular Biology and Biophysics – University of Connecticut Health Center, Farmington, CT 06030, USA*
[g] *Regional Centre of Advanced Technologies and Materials, Czech Advanced Technology and Research Institute, Palacký University Olomouc, Olomouc 783 71, Czech Republic*
[h] *Universitat Pompeu Fabra (UPF), 08003 Barcelona, Spain*

A B S T R A C T

Repression of *msl-2* mRNA translation is essential for viability of *Drosophila melanogaster* females to prevent hypertranscription of both X chromosomes. This translational control event is coordinated by the female-specific protein Sex-lethal (Sxl) which recruits the RNA binding proteins Unr and Hrp48 to the 3' untranslated region (UTR) of the *msl-2* transcript and represses translation initiation. The mechanism exerted by Hrp48 during translation repression and its interaction with *msl-2* are not well understood. Here we investigate the RNA binding specificity and affinity of the tandem RNA recognition motifs of Hrp48. Using NMR spectroscopy, molecular dynamics simulations and isothermal titration calorimetry, we identified the exact region of *msl-2* 3' UTR recognized by Hrp48. Additional biophysical experiments and translation assays give further insights into complex formation of Hrp48, Unr, Sxl and RNA. Our results show that Hrp48 binds independent of Sxl and Unr downstream of the E and F binding sites of Sxl and Unr to *msl-2*.

1. Introduction

In sexually reproducing organisms the number of X chromosomes across sexes are often inequal. Without further regulation, this would lead to unbalanced expression of X-linked genes [22]. Dosage compensation mechanisms have evolved that, through regulation of transcription and translation, enable uniform expression levels of X-linked genes [78]. In *Drosophila melanogaster*, transcription of the single male X chromosome is upregulated two-fold to equalize the expression levels to that of both female X chromosomes [35]. This hyper-transcription is achieved by the multi-subunit dosage compensation complex (DCC) [69], consisting of five proteins (Mle, Msl1–3, and Mof), of which Msl2 is the rate-limiting component [10,37,40] and one of the two long non-coding RNAs *roX1* and *roX2* [62]. Hyper-transcription would have lethal consequences in females. Here, a mechanism has evolved, which inhibits the assembly of the DCC by translation repression of *msl-2* mRNA [48]. This is orchestrated by the female-specific RNA-binding protein (RBP) Sex-lethal (Sxl) in multiple regulatory steps of gene expression [7,32,64]. First, Sxl regulates the alternative splicing of a 5' untranslated region (UTR) facultative intron of the *msl*-2-pre-mRNA [28,63] followed by nuclear retention of *msl-2* transcripts [33,34]. Second, transcripts that escape to the cytoplasm are repressed at the level of translation initiation by Sxl binding to uridine-rich motifs in both the 5' and 3' UTRs of *msl-2* mRNA (designated as A- to F- sites, Fig. 1A-C, [8,9,31,33,49,61]). At the 3' UTR, in a PABP-dependent translational control mechanism [24], Sxl together with the RBPs
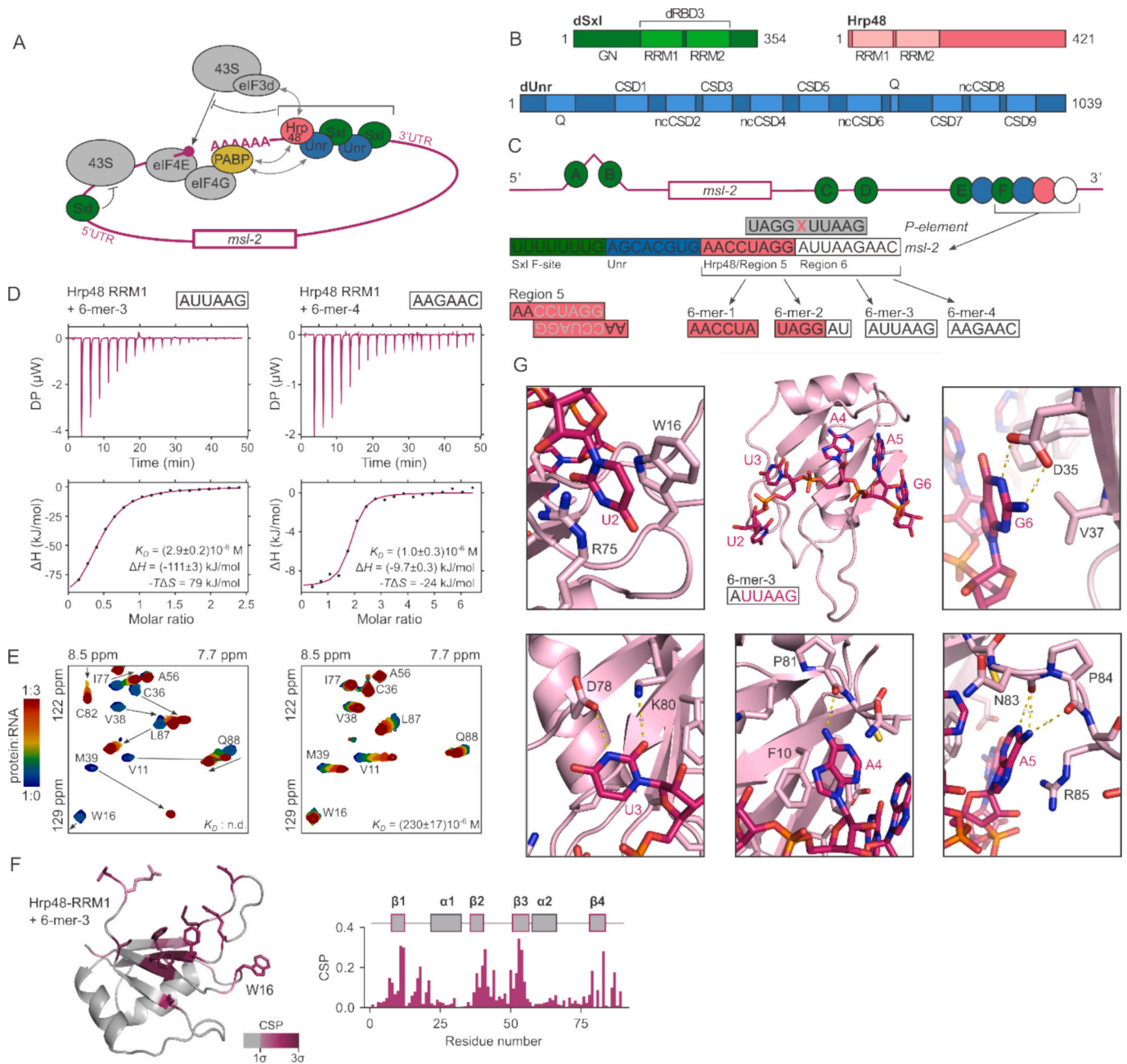
Fig. 1. Structure and RNA binding of Hrp48-RRM1. **A:** Model for translation repression of *msl-2* 3' and 5' UTR mediated control. Sxl binds to both UTRs and recruits Unr to a specific RNA region of the 3' UTR, where it inhibits along with Hrp48 the recruitment of the 43S preinitiation complex (PIC). Hrp48 supposedly binds the 43S *via* interaction with the eIF3d subunit. Those PICs that escaped the 3' UTR regulation are inhibited by Sxl bound at the 5' UTR, which prevents scanning of the start codon by the 43S. **B:** Domain arrangement of Sxl, Hrp48 and Unr constructs used in this study. Amino acid numbers are indicated (GN: glycine-asparagine rich region, RRM: RNA recognition motif, dRBD3: former nomenclature of a region encompassing both RRMs of Sxl, residues 123–294, Q: glutamine-rich domain, CSD: cold shock domain, ncCSD: non-canonical cold shock domain, [41]). **C:** Schematic model of the full-length *msl-2* mRNA with reported binding sites for Sxl (green), Hrp48 (red) and Unr (blue). Sites B, E and F are required for optimal translational repression [31,32]. The white circle marks the region we studied in addition to the earlier identified binding site. Region 5 is the sequence suggested previously as the *msl-2* Hrp48 binding site [79]. It can self-associate through duplex formation in isolation (Fig. S5C). Hrp48 *P-element* binding site shows high similarity to a region overlapping Region 5 and 6 of *msl-2* (grey). The 6-mer constructs were used in this study to refine the Hrp48 binding site. **D:** Isothermal titration calorimetry (ITC) data of Hrp48-RRM1 titrated with 6-mer-3 and 6-mer-4. **E:** A zoomed-in region of the $^{15}N,^{1}H$-HSQC NMR titration experiments of ^{15}N-labeled Hrp48-RRM1 and 6-mer-3 and 6-mer-4. **F:** Mapping of the chemical shift perturbation (CSP) data of the RRM1–6-mer-3 titration on the RRM1 crystal structure (left panel) and sequence (right panel). The threshold for mapping a CSP onto the structure was one standard deviation of the CSP values. **G:** MD-derived Hrp48-RRM1-RNA complex with 6-mer-3 RNA. The overall structure is surrounded by zoom-ins of single nucleotides being base-specifically recognized by RRM1 residues. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Heterogeneous nuclear ribonucleoprotein 48 (Hrp48, also known as Hrb27C) and Upstream-of-N-ras (Unr) inhibits ribosomal recruitment to the mRNA (Fig. 1A, B, [1,2,23,24,36,79]). Sxl binds to the E- and F-site of *msl-2* 3' UTR cooperatively with Unr which acts as a necessary co-repressor (Fig. 1C, [39]). As part of the repression complex, Hrp48 binds to the initiation factor eIF3d *via* direct contacts [79]. Ribosomal pre-initiation complexes that could potentially escape the 3' UTR translational repression are inhibited by Sxl bound to the B-site at the 5' UTR which interferes with ribosomal scanning and recognition of the AUG initiation codon (Fig. 1A, [9]).

Despite advances in unraveling the complexity of this translational control mechanism, how Hrp48 interacts with *msl-2* 3' UTR and re-presses translation is not well understood. The exact RNA recognition mode of Hrp48 and whether it establishes cooperative contacts with Sxl and Unr are not known.

In this study, we investigate the RNA binding activity of Hrp48 in isolation and as part of a multiprotein complex involving Sxl and Unr bound to an extended site F *msl-2*-derived RNA fragment. We used NMR spectroscopy and molecular dynamics simulations to refine the cognate RNA motif within the 3' UTR, which partially overlaps with the previously reported motif [79]. Furthermore, we determined binding affinities of the two structured RNA recognition motifs (RRM) of Hrp48, which jointly bind *msl-2*. NMR ^{15}N spin relaxation and titration data suggest that the two domains bind RNA cooperatively and tumble together in the RNA bound state. *In vitro* translation assays support our findings on the binding site at the functional level. We report the crystal structure of Hrp48 RRM1 at 1.2 Å resolution and validate the structure prediction model of RRM2 by NMR spectroscopy. Further biochemical and biophysical data on Hrp48, Unr, Sxl and *msl-2* suggest no interaction of the proteins in the absence of RNA. Notably, the three proteins bind *msl-2* simultaneously and we established a protocol to reproducibly form the quaternary complex. Our results suggest that Hrp48 binds *msl-2* mRNA independently of the Sxl-Unr moiety in absence of its intrinsically disordered regions (IDRs).

2. Results

2.1. Structure and RNA binding of Hrp48-RRM1

Initially, we used the divide-and-conquer approach and investigated each RRM domain of Hrp48 in isolation with regard to their structure and RNA binding properties. To this end, we conducted isothermal

Table 1
Detailed parameters and results of ITC measurements.

Exp. name	Replicates	Conc. syringe (μM)	Conc. cell (μM)	N (sites)	N (error)	Kd (μM)	Kd (μM, error)	ΔH (kJ/mol)	−TΔS (kJ/mol)	Injection no.	Injection vol. (μL)	Shown in figure
Hrp48-RRM1 +6-mer-1	2	6-mer-1 766–800	RRM1 30.0	2.4 3.1	45·10⁻³ 121·10⁻³	1.5 3.5	0.4 1.2	−7.2 ± 0.3 −5.3 ± 0.4	−25.6 −25.4	19	2.0	S1
Hrp48-RRM1 + 6-mer-2	3	6-mer-2 189–550	RRM1 17.5–20.0	0.8 0.8 0.9	35·10⁻³ 76·10⁻³ 94·10⁻³	8.7 4.0 7.4	1.3 2.2 2.9	−76 ± 6 −65 ± 12 −85 ± 16	48 35 56	19/25	2.0/1.5	S1
Hrp48-RRM1 + 6-mer-3	4	6-mer-3 235–601	RRM1 23.0–25.0	0.8 0.6 0.4 0.7	17·10⁻³ 18·10⁻³ 13·10⁻³ 17·10⁻³	5.0 6.1 2.6 3.4	0.6 0.5 0.4 0.4	−105 ± 5 −100 ± 5 −107 ± 6 −74 ± 3	75 71 76 43	19/25	2.0/1.5	1
Hrp48-RRM1 + 6-mer-4	3	6-mer-4 650–1400	RRM1 17.5–30.0	1.7 1.6 2.4	32·10⁻³ 12·10⁻³ 37·10⁻³	1.0 0.8 3.0	0.3 0.1 0.5	−9.7 ± 0.3 −10.4 ± 0.1 −7.3 ± 0.2	−24.0 −23.9 −23.7	19/25	2.0/1.5	1
Hrp48-RRM12 + Region 5	1	Region 5 1000	RRM12 60.0	No binding detectable by ITC.						19	2.0	S5
Hrp48-RRM12 + 10-mer-WT	2	10-mer 171	RRM12 22.0	0.6 0.6	2.2·10⁻³ 1.21·10⁻³	0.02 0.02	2.9·10⁻³ 1.6·10⁻³	−275 ± 2.6 −282 ± 1.5	232 238	25	1.5	3
Hrp48-RRM12 + 9-mer	2	9-mer 182	RRM12 21.6	0.7 0.7	3.2·10⁻³ 2.5·10⁻³	0.02 0.01	3.7·10⁻³ 3.2·10⁻³	−235 ± 0.3 −226 ± 0.7	191 181	25	1.5	3
Hrp48-RRM12 + 10-mer-CC	2	10-mer-CC 190	RRM12 21.6	No binding detectable by ITC.						25	1.5	S5
Hrp48-RRM12 + GG-30-mer	3	GG-30-mer	RRM12	Complex binding mode of multiple binding sites.						19/25	2.0/1.5	5
Hrp48-RRM12 +GG-30-mer–Sxl-dRBD3–Unr-CSD12 complex	4	complex 283–300	RRM12 19.0–29.3	0.5 0.6 0.5	18·10⁻³ 14·10⁻³ 9.7·10⁻³	0.9 1.8 1.9	0.2 0.3 0.3	−124 ± 7 −126 ± 6 −120 ± 4	90 94 128	19/25	2.0/1.5	5

titration calorimetry (ITC) and ^{15}N,^{1}H-HSQC NMR experiments to observe chemical shift perturbations of the protein resonances upon titration with RNA. The previously reported *msl-2* binding site of Hrp48 (AACCUAGG) is downstream and adjacent to the Unr binding site close to the Sxl F site and has been termed Region 5 (Fig. 1C, [79]). Overlapping with this Region 5 but shifted by four nucleotides towards the 3' end is a sequence (UAGGAUUAAG), which is highly similar to a sequence reported earlier to be the Hrp48 binding site on *P-element* mRNA (UAGGUUAAG, Fig. 1C, [74,75]. Thus, we divided the 17 nucleotides just downstream of the Sxl (F site)-Unr binding region into 6-mer RNA oligonucleotides to assess binding of the single RRM domains while preventing formation of double-stranded RNA (dsRNA) by the palindromic motif of Region 5 (Fig. 1C). This resulted into four 6-mers with the following sequences: 6-mer-1: AACCUA, 6-mer-2: UAGGAU, 6-mer-3: AUUAAG and 6-mer-4: AAGAAC (Fig. 1C). We also added a U_6-mer as a control (6-mer-5, Fig. S1), a sequence unrelated to the other 6-mers. In our ITC experiments, we tested RNA binding of RRM1 for each 6-mer. Binding of the control U_6-mer (6-mer-5) was not strong enough to be detected by ITC. From ITC-derived dissociation constants (K_D) the optimal motif for RRM1 could not be identified as they were in the same order of magnitude for all four 6-mers (Fig. 1D, S1A—C, Table 1). Interestingly, a clear difference in the enthalpic and entropic contributions to the affinity of binding could be observed between 6-mer-1/4 and 6-mer-2/3 (Fig. 1D and S1A, B Table 1). This could be explained by the different stoichiometries we observed during the titration, which suggest that two 6-mer-1/4 molecules bind to one RRM1, whereas the stoichiometry for 6-mer-2/3 and RRM1 is rather 1:1. However, this does not explain the negative ΔTS values for 6-mer-1/4 and we rather refrain from speculations about potential models explaining this data.

We then used NMR titration experiments to further study the RNA binding to complement the ITC results. Based on the chemical shift perturbations (CSPs), RRM1 binds to all five hexamers, including the control 6-mer-5, but with different affinities (Fig. 1E and S1D—H). For the 6-mer-1, 6-mer-4 and 6-mer-5 titrations, CSPs indicate that RNA binding is in the fast exchange regime on the NMR time scale, whereas for 6-mer-2 and 6-mer-3 we could observe strong CSPs and intensity changes, characteristic for an intermediate-to-slow exchange regime [86]. CSPs in the intermediate-to-slow exchange regime do not allow for reliable determination of dissociation constants. However, slow exchange usually indicates stronger binding with low micromolar affinities [50]. Thus, from the NMR data, one could qualitatively conclude that 6-mer-2 and 6-mer-3 are bound stronger by Hrp48 than 6-mer-1 and 6-mer-4, with 6-mer-3 featuring the highest number (6-mer-3: 24 and 6-mer-2: 18) of peaks with CSPs in the slow exchange regime. This is exemplified by tryptophan W16 located in the loop region between β1 and α1 (Fig. 1F) since it shows the highest CSP upon 6-mer-3 binding and is in the intermediate-to-slow exchange regime (Fig. S1D—H). Aromatic residues located in this region are likely to be involved in RNA binding which already has been shown for other RRM domains including Fox-1 [5,6]. Another example is RRM2 of the splicing-factor RBM39 which exploits residues in the β1- α1 loop-region as extended RNA binding interface [15]. Thus, we concluded from the NMR titrations that 6-mer-3 is the best Hrp48 binder since this ligand leads to a bound state that is closest to a single low energy complex structure due to the high number of CSPs in the slow-intermediate exchange regime. Also, 6-mer-3 induces the strongest CSPs over most residues (Fig. S1I). Backbone chemical shift assignment of RRM1 enabled us to map the CSPs onto the crystal structure of RRM1 (Fig. 1F) we determined at 1.2 Å resolution (Table 2, PDB: 9EN7) using molecular replacement based on the hnRNP A1-RRM1 structure (PDB: 1HA1, [72]) since hnRNP A1 is a putative mammalian homolog of Hrp48 which has 62 % sequence similarity and 43 % sequence identity (Fig. S2). The structure of Hrp48-RRM1 adopts the canonical βαββαβ topology with four β-strands folding into an antiparallel β-sheet which is packed against two α-helices (Fig. 1F, [65]). The CSPs show that the RNA binds along the four β-strands, resembling

Table 2
Data collection and refinement statistics.

	Hrp48_RRM1
Wavelength	
Resolution range	31.82–1.191 (1.233–1.191)
Space group	P 21 21 21
Unit cell	38.344 45.4996 57.026 90 90 90
Total reflections	416,881 (41304)
Unique reflections	32,496 (856)
Multiplicity	12.8 (13.1)
Completeness (%)	91.18 (26.81)
Mean I/sigma(I)	18.30 (1.27)
Wilson B-factor	14.59
R-merge	0.06499 (1.536)
R-meas	0.06773 (1.597)
R-pim	0.01881 (0.4347)
CC1/2	1 (0.609)
CC*	1 (0.87)
Reflections used in refinement	29,733 (857)
Reflections used for R-free	2000 (58)
R-work	0.1627 (0.2193)
R-free	0.1617 (0.2151)
CC(work)	0.972 (0.855)
CC(free)	0.976 (0.840)
Number of non-hydrogen atoms	803
Macromolecules	670
Ligands	28
Solvent	121
Protein residues	84
RMS(bonds)	0.006
RMS(angles)	0.88
Ramachandran favored (%)	98.78
Ramachandran allowed (%)	1.22
Ramachandran outliers (%)	0.00
Rotamer outliers (%)	0.00
Clashscore	2.98
Average B-factor	24.44
Macromolecules	21.65
Solvent	66.07

Statistics for the highest-resolution shell are shown in parentheses.

the canonical RNA binding mode of RRM domains.

Interestingly, peaks corresponding to residues located C-terminal to β-strand 4 just before the linker (P81-R85) between RRM1 and RRM2 starts, exhibit also strong CSPs. These may be induced by direct RNA binding or by allosteric effects.

As we could not determine an experimental RRM1-RNA complex structure, we utilized unbiased atomistic simulations using a recently rescaled RNA force field that allows observing the full spontaneous binding process of single-strand RNAs to proteins [51]. We applied this method to study the binding of 6-mer-3 (UUAAG) to RRM1, generating a large structural ensemble (~110 μs in total; see Methods). The calculated ensemble was subsequently analyzed and all the observed bindings annotated against the NMR CSP data. A single representative RRM1–6-mer-3 complex binding motif, which excellently reflected the CSP data while also demonstrating long-term stability in MD simulations, was finally selected for further analysis (Fig. 1G). In this structure, direct protein-RNA contacts are established where U2 (numbering according to the appearance within the full length 6-mer-3) is sandwiched between the indole ring of W16 and guanidinium group of R75. The arginine is oriented in a way that also allows formation of base-specific hydrogen bonds or electrostatic interaction with the downstream phosphate group during thermal fluctuations. The U3 is also recognized base-specifically by D78 and K80, forming hydrogen bonds, whereas adenosine 4 stacks with F10 and is base-specifically recognized by hydrogen-bonding with the backbone carbonyl of P81. Adenosine 5 forms a base-specific hydrogen bond network with the backbone carbonyls of N83 and P84 while the base stacks with the guanidinium group of R85. G6 is also recognized base-specifically by hydrogen bonds with the side chain of D35 and K8 while V37 forms hydrophobic interactions with the purine ring. Thus, all bases of 6-mer-3 can be contacted in a base-specific

fashion, providing the structural basis for Hrp48's highly sequence-specific RNA binding. To experimentally validate the aforementioned simulated residue-specific interactions between RRM1 and 6-mer-3, we generated five RRM1-mutants and tested their binding affinities towards 6-mer-3 by NMR-titration experiments (Fig. S3A—F). Each RRM1 variant was designed to affect binding to one specific nucleotide by introducing one or two point-mutations without impairing the tertiary structure of the RRM domain. In addition, only amino acids that are not canonical RRM binding residues and presumably interact with their side-chain rather than establishing contacts through their protein backbone were selected for mutagenesis. CSP analyses of the NMR titration data resulted in a significant global decrease of binding affinity in all tested mutants compared to wild-type RRM1 (Fig. S3G), indicating that the mutated amino acid residues are directly involved in RNA binding.
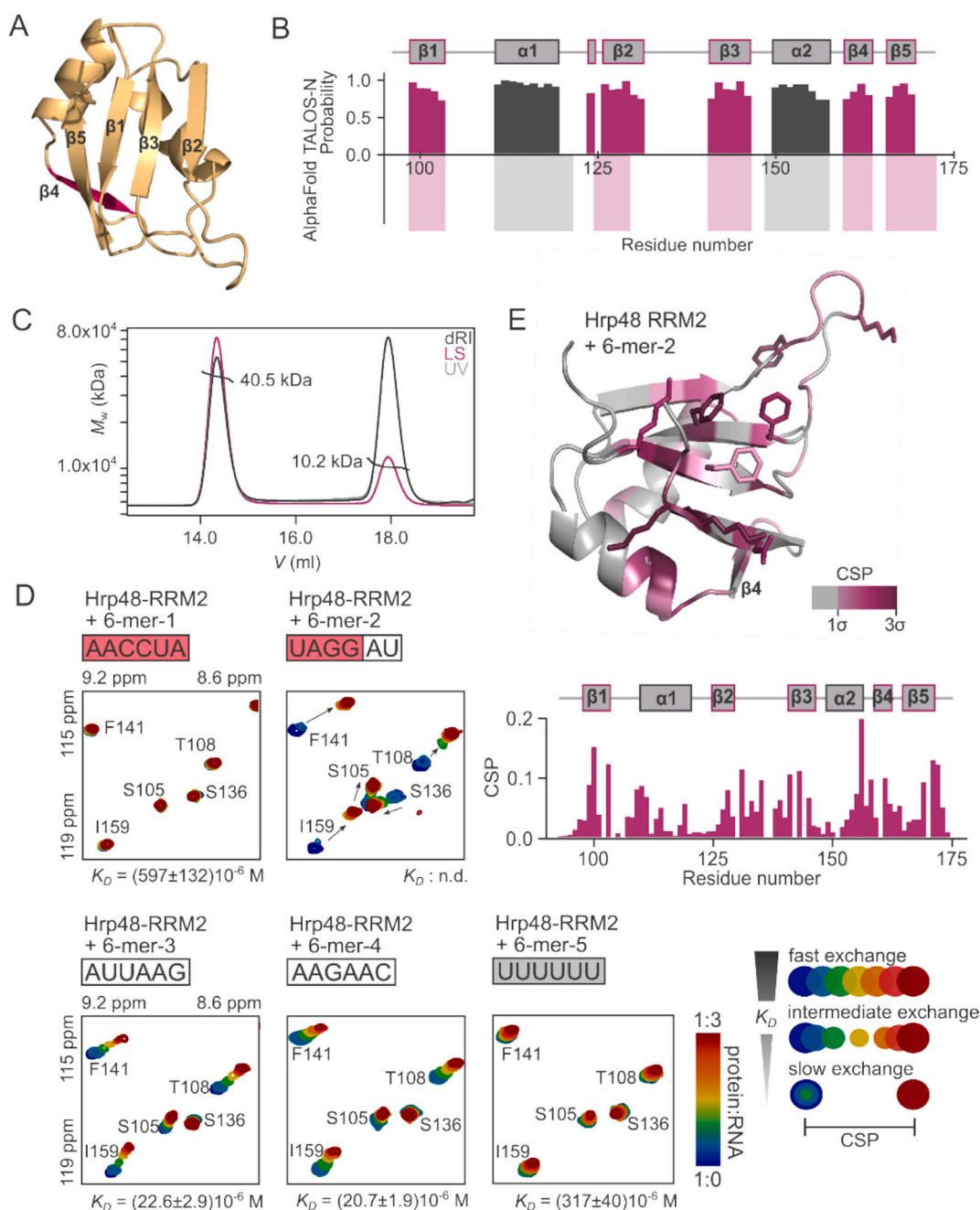

Fig. 2. Structure and RNA binding of Hrp48-RRM2. **A:** AlphaFold2 (AF2) model of RRM2 features a non-canonical RRM fold with an additional β-strand, depicted in magenta (β4) [45]. **B:** The βαββαβ fold prediction of AF2 is confirmed by NMR chemical shift assignment data used in TALOS+ [73]. The lower part of the secondary structure plot shows the secondary structure predicted by AF2. **C:** Size exclusion chromatography coupled with multi-angle light scattering (SEC-MALS) of RRM2 reveals a strong oligomerization tendency of RRM2. Dark grey: differential refractive index (dRI), purple: light scattering (LS), light grey: UV absorption (UV). **D:** Zoomed-in regions of the ^{15}N,^{1}H-HSQC NMR spectra of ^{15}N-labeled Hrp48-RRM2 titrated with the five different 6-mer RNA oligos to three-fold excess. RRM2 does not bind 6-mer-1 and weakly binds 6-mer-5. 6-mer-3 and 6-mer-4 binding is in the fast exchange regime on the NMR time scale, and 6-mer-2 in the intermediate-slow exchange regime. Schematic representation of CSP patterns indicating the different exchange regimes observable in NMR titrations. **E:** Mapping the CSPs on the structure of RRM2 reveals a mostly canonical RRM-RNA recognition mode. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

2.2. Biophysical characterization of Hrp48-RRM2

Next, we wanted to assess the structure and RNA binding properties of Hrp48-RRM2. Crystallization trials with or without RNA of RRM2 were not successful. Therefore, we used AlphaFold2 to predict the structure of RRM2 [45]. Interestingly, the predicted structure revealed an additional β-strand between the last α-helix and β-strand, deviating from the canonical βαββαβ fold of RRM2 (Fig. 2A). Secondary structure prediction based on our NMR chemical shift assignment of RRM2 using TALOS+ [73] confirmed the presence of this additional β-strand (β4) in solution (Fig. 2B). During purification and analysis of NMR data we observed that RRM2 tends to oligomerize. Size-exclusion chromatography coupled with multi-angle light scattering (SEC–MALS) confirmed that this isolated domain forms tetramers in an isolated *in vitro* context at 20 μM and above (Fig. 2C). ITC data suffered from noise at these low concentrations, indicating generally lower binding affinity of RRM2

compared to RRM1. However, ITC could not be performed at higher concentrations due to the oligomerization tendency, thus we utilized NMR for titration experiments to assess the RNA binding specificity and affinity of RRM2. Upon titration with 6-mer-1 and control RNA 6-mer-5 no or only weak CSPs could be observed (Fig. 2D and S4). Titration with 6-mer-3 and 6-mer-4 resulted in CSPs in the fast and fast-to-intermediate exchange regime, whereas 6-mer-2 titration induced CSPs in the intermediate-to-slow exchange regime (Fig. 2D and S4). The latter could not be fitted to obtain a K_D value but it qualitatively indicates that 6-mer-2 binds strongest to RRM2. Thus, for RRM2, 6-mer-2 is the best binder. Mapping the CSPs onto the RRM2 structural model shows that the canonical β-strands are involved, typical for RRM domains (Fig. 2E). Interestingly, also peaks corresponding to residues of the non-canonical β4-strand and α2-helix exhibit strong CSPs upon titration with 6-mer-2. This might be due to direct interaction with RNA or due to allosteric effects.
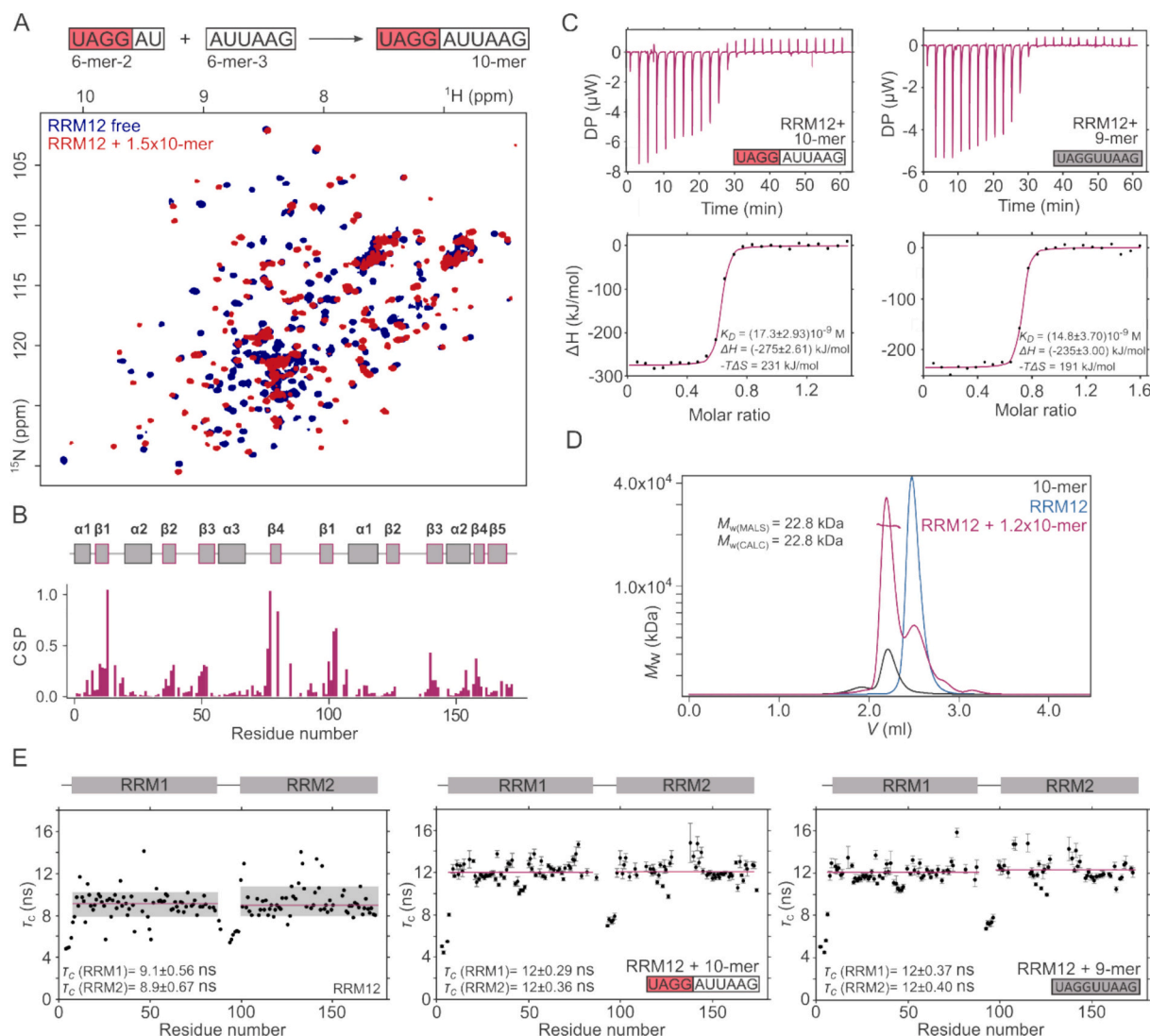

Fig. 3. RNA binding and dynamics of Hrp48-RRM12. **A:** 10-mer RNA consisting of 6-mer-2 and 6-mer-3 sequences (upper panel) was used in ^{15}N,^{1}H-HSQC NMR titration experiment for RRM12 (lower panel). Peaks are in the slow-exchange regime; additional titration points are shown in Fig. S5A. **B:** CSP plot of RRM12–10-mer NMR titration. Both RRM domains are involved in the interaction and residues in the β-strands are the most affected, including β4 of RRM2. **C:** ITC of Hrp48-RRM12 with 10-mer (left panel) and 9-mer (right panel), respectively. **D:** The 10-mer and RRM12 forms a 1:1 complex. SEC-MALS UV absorption chromatograms of free RRM12 (grey), 10-mer (blue) and their complex (purple). **E:** Rotational correlation times of RRM12 residues in the apo state (left panel), bound to 10-mer (middle panel) and bound to 9-mer (right panel). The domains tumble jointly in an RNA bound state. Left panel: grey: one standard deviation of the mean value of τ_c of each residue within a single RRM domain. Purple: Mean value of τ_c of residues within the standard deviation. Values outside the standard deviation are considered as outliers and were excluded for calculating the mean. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

We conclude that both RRM domains of Hrp48 recognize *msl-2* and prefer binding to 6-mer-2 and 6-mer-3 which together embrace a 10-mer recognition sequence (Fig. 3A). RRM2 binds upstream of RRM1 which is also consistent with the binding behavior of RRM tandem domains of other proteins [20,38,71]. Furthermore, the recognition site of Hrp48 on the *msl-2* mRNA is very similar to the Hrp48 binding site on *P-element* mRNA which strengthens the validity of our results [74,75]. Next, we assessed binding of the tandem RRM12 construct of Hrp48 to *msl-2* mRNA.

2.3. The tandem domain RRM12 of Hrp48 binds msl-2 cooperatively

For binding studies of the tandem domain RRM12 of Hrp48, we used 10-mer RNA (UAGGAUUAAG, Fig. 3A). This sequence is four nucleotides downstream of the start of the previously reported Region 5 [79], and consequently does not include the palindromic motif, allowing meaningful *in vitro* investigations without the risk of forming dsRNA. Indeed, we could confirm self-complementary base pairing of Region 5 by 1D-NMR which showed strong peaks in the imino region, indicative of base pairing (Fig. S5A). With this 10-mer RNA, we wanted to test whether a tandem RRM12 construct would bind stronger and performed an NMR titration and ITC experiments to determine the binding affinity (Fig. 3A-C). In the NMR titration, resonances corresponding to residues of both domains exhibited strong CSPs in the slow exchange regime indicating strong binding in the nanomolar range (Fig. 3A, S5B). Both domains exhibit considerably larger CSPs upon RNA binding than the isolated domains. This is confirmed by ITC experiments from which we could determine a dissociation constant of 17.3 ± 2.9 nM. Thus, RRM12 binds around 100-times stronger to this 10-mer RNA than the individual domains to their best 6-mer motifs. As such, both RRM domains bind to single-stranded RNA simultaneously and synergistically. We could also show that RRM12 forms a 1:1 complex with the 10-mer in solution by SEC-MALS (Fig. 3D, S5C), which together with the NMR data confirms that both domains bind one RNA 10-mer.

Having identified the optimal RNA binder for RRM12, we attempted to obtain a crystal structure of an Hrp48-RRM12-RNA complex. However, as for the isolated RRM domains, we could not obtain crystals with RNA despite extensive crystallization trials. Instead, we assessed the dynamics of the tandem domain through NMR ^{15}N spin relaxation experiments. In the absence of RNA both domains have similar global rotational correlation times τ_c of 9.1 ± 0.6 ns for RRM1 and 8.9 ± 0.7 ns for RRM2. This indicates that both domains tumble mostly independently from each other (Fig. 3E), as it can be estimated that the rotational correlation time is approximately 0.6 times the molecular weight of a globular domain [68]. In case of fully independent tumbling of both domains, a τ_c of around 6 ns would be expected. Thus, the short linker influences the tumbling of each domain or weak unspecific interactions between both domains could lead to an elevated rotational correlation time. The linker residues exhibit much lower rotational correlation times, indicating a high degree of flexibility, which does not change upon addition of 10-mer RNA. Thus, the linker is not involved in RNA binding. However, in presence of RNA the rotational correlation times of both domains are 12 ± 0.3 ns for RRM1 and 12 ± 0.4 ns for RRM2. This clearly shows that both domains tumble together and do not move freely with respect to each other but adopt a fixed orientation in RNA-bound state. Consistent with the NMR titration data of the isolated RRM1 domain, we observed elevated τ_c values and significant CSPs for assigned residues located in between β-strand 4 and the linker residues, confirming that these residues are not flexible but involved in RNA contacts.

Since the Hrp48 binding site of *msl-2* mRNA is not identical to the recognition sequence on the *P-element* mRNA, we investigated whether the difference (one additional A in the center part of the *msl-2* mRNA binding sequence, UAGG**A**UUAAG) has an impact on the dynamics of the tandem domain RRM12 in the RNA bound state. Given that NMR relaxation data for both complexes resulted in almost identical rotational correlation times per residue, we conclude that the additional adenosine of the 10-mer does not lead to increased flexibility between both RRM domains (Fig. 3E).

To test whether the additional adenosine has any impact on the binding affinity of the RRM domains towards the RNA we conducted ITC experiments by titrating 9-mer RNA to RRM12 (Fig. 3C, right panel) and in addition used a 10-mer mutant sequence in which the two nucleotides in the center part of the 10-mer wild-type (10-mer-WT) are mutated to CC (UAGG**CC**UAAG, 10-mer-CC, Fig. S5D). The K_D value resulting from the 9-mer titration is with 14.8 ± 3.7 nM in the same range as for the 10-mer-WT experiment (17.3 ± 2.9 nM) which is consistent with the NMR relaxation data where the additional A does not induce changes in backbone dynamics of Hrp48. Interestingly, the ITC experiment for the 10-mer-CC titration suffered from severe noise, hence calculation of the binding affinity was not feasible (Fig. S5D, right panel). However, since the experimental setup was identical to 10-mer-WT and 9-mer titration experiments, we conclude that the RRM12 has a significantly weaker affinity for 10-mer-CC than for the 10-mer-WT. However, it must be considered, that the 10-mer-CC is partly self-complementary and could form dsRNA, which would interfere with Hrp48 binding. Nevertheless, the result is consistent with our MD-derived structural model in which the central U (Fig. 1G, U2 in 6-mer-3) is specifically recognized by R75. Substitution to a C would abolish these specific contacts and weaken the affinity.

Accordingly, we hypothesized that the adjacent U (U6 in 10-mer-WT) downstream of the additional A (A5 in 10-mer-WT) could have an important role for Hrp48-RRM12 binding. To validate this hypothesis, we used shortened RNA constructs of a length of 4 nucleotides (4-mer-1–4-mer-6, Fig. S6A) for NMR titration experiments. All 4-mers together encompass the entire 10-mer sequence and the previously identified Region 5 (Fig. S6A). The results confirmed that only 4-mer-3 and 4-mer-6, which correspond to the 5' and the 3' end of our already confirmed 10-mer-WT binding sequence show stronger CSPs in comparison to all other 4-mer NMR titration experiments (Fig. S6B, C). 4-mer-4 and 4-mer-5 that consist of sequences which include the central A and U, show lower CSPs compared to 4-mer-3 and 4-mer-6 which indicates that the central region binds weaker. We then used 4-mer-3 and 4-mer-6 for NMR titrations to the single RRM domains to check whether a 4-nucleotide sequence is sufficient for efficient RNA binding (Fig. S7A-E). CSP analyses revealed that RNA still binds both RRM domains but the affinity decreased significantly compared to 6-mer binding (Fig. S7F). From these data we conclude that the RRM domains mainly bind to the 5' and the 3' end of the 10-mer-WT binding site, whereby four nucleotides are not sufficient for efficient binding, thus the central U is bound by Hrp48 whereas the central A is dispensable.

2.4. Validation of Hrp48's RNA binding by mutational analysis in translation assays

To further validate the refined Hrp48 binding site of *msl-2* mRNA, we performed a mutational analysis and tested the ability of the mutant mRNAs to be repressed by Sxl using translation assays in *D. melanogaster* embryo extracts. Several RNA constructs were tested for *in vitro* translation activity in Luciferase reporter assays (Fig. 4, methods). We used CU-repeats to mutate specific parts of the 3' UTR including 5m (mutation of Region 5), 6m (mutation of Region 6) or fragments within (Fig. 4A). We also investigated the impact on translation repression upon substituting the entire region (Region 5 and 6, designated 56m). The latter has the strongest de-repression effect compared at a RBD4/RNA concentration ratio of 10 (Fig. 4A). Mutation of the 10-mer sequence (10m) de-repressed translation to a lesser extent compared to 56m but still had an important effect compared to the wild-type sequence. Interestingly, mutation of the last 6 nucleotides of the 10-mer sequence (6m) resulted in similar translation de-repression than for the whole 10-mer, indicating that Hrp48-RRM1 is more important for translation repression than binding of RRM2. This hypothesis is
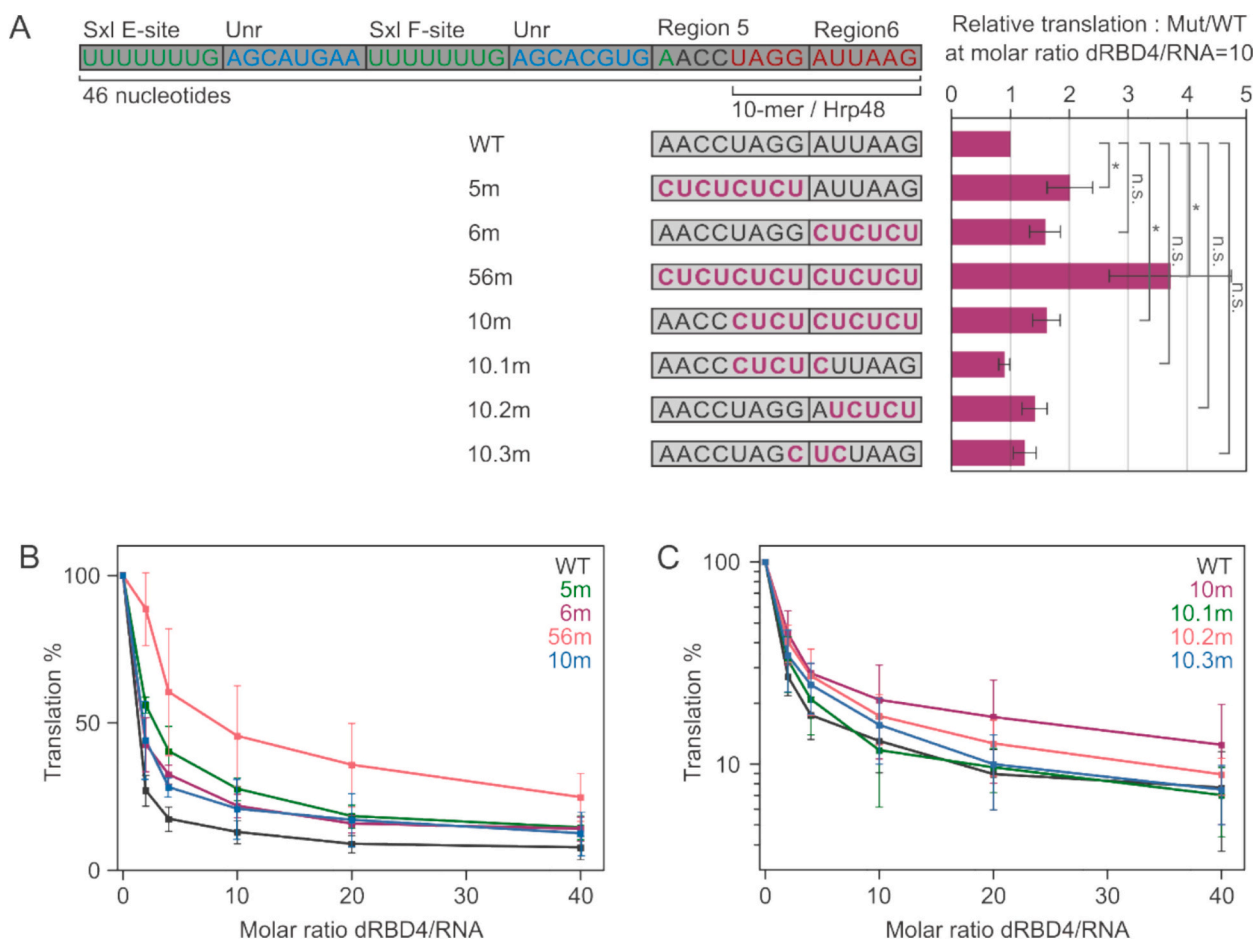
Fig. 4. The Hrp48 RNA-binding site is necessary for translation repression. **A:** *In vitro* translation assays were performed with Firefly luciferase reporters appended to a segment of the 3' UTR of *msl*-2 containing the minimal Sxl (green) and Unr (blue) binding sites, and derivatives of the Hrp48-binding region. The CU-repeat substitutions are indicated with bold pink fonts. The bar graph shows the relative translation of the luciferase reporter at a Sxl/RNA concentration ratio of 10. Co-translated *Renilla* luciferase was used as an internal control and as a reference for normalization of the reporter Firefly luciferase signal, and the data was referred to the wild type (WT) signal activity. Data represent the average of three independent experiments; error bars represent standard deviation. *$p < 0.05$. **B–C:** *In vitro* translation assay results for all Sxl/RNA concentration ratios used in the experiment. The data was plotted as relative translation taking as 100 % the initial activity with no Sxl added to the extract. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

supported by the results of the 10.1m luciferase assay. Here, translation repression of the reporter is as efficient as in the wild-type sequence. In addition, CSPs of the NMR titration experiments of RRM1 and RRM2 with 6-mer RNA indicate generally stronger binding of RRM1 compared to RRM2 which would also strengthen the aforementioned hypothesis. The translation repression assay with 10.3m resulted in a slightly de-repressed translation. However, this weak change in translation efficiency compared to the wild-type indicates that the central part of the 10-mer is not essential for translation repression. Mutating the complete sequence (56m), in which both adenosines at the 5′ end were also exchanged with a CU-repeat, had the largest translation de-repression effect. These two adenosines are in contact with Sxl which would explain the elevated translation de-repression effect for 5m and 56m [39]. Thus, we confirmed that the 10-mer motif is the optimal Hrp48 binding site and its mutation has an impact on translation repression.

2.5. Complex formation of Hrp48 with Sxl and Unr is RNA dependent

Having obtained mechanistic insight into the Hrp48 binding mode and base specificity, we wanted to assess whether Hrp48-RRM12 forms a complex with Sxl and Unr and whether Hrp48-RRM12 binds to an extended RNA, including the F site of Sxl and Unr binding sites. However, we tested first whether Hrp48 interacts with Sxl or Unr in the absence of RNA. It has been shown previously that Sxl-dRBD3 and Unr-

CSD1 do not interact without RNA, but protein-protein contacts between the two proteins are established upon RNA binding [39]. We expressed and purified unlabeled Sxl-dRBD3 and different constructs of Unr. The domain boundaries of Unr were based on previous work in our lab, where four additional non-canonical CSDs were identified in between the five predicted canonical CSDs (Fig. 1B, [41]). $^{15}N,^1H$-HSQC spectra were recorded of Hrp48-RRM12 alone, and Hrp48-RRM12 together with one equivalent Sxl or Unr (Fig. S8A—D). We also performed the reverse experiment with labeled Sxl-dRBD3 and unlabeled Hrp48-RRM12 (Fig. S8E). Due to the lack of chemical shift perturbations or significant intensity decrease in all titrations, we concluded that Sxl-dRBD3 and the tested Unr constructs do not interact with Hrp48-RRM12 in the absence of RNA. The small decrease in intensity for some of the peaks might be a result of unspecific interactions or aggregation of the proteins. Overlay of SEC–MALS chromatograms of the separate proteins and all three proteins together further confirmed that the constructs we later used for complex formation, Sxl-dRBD3, Unr-CSD12 and Hrp48-RRM12 do not interact in the absence of RNA (Fig. 5A).

With the intention of testing whether Sxl, Unr and Hrp48 jointly interact with *msl-2*, which has been proposed earlier [79], we designed two 30-mer RNA constructs (WT-30-mer and GG-30-mer) based on published data for Sxl and Unr and our findings of the Hrp48 binding site [39] (Fig. 5B). This site combines the F site to which Sxl binds, followed by the Unr and the 10-mer Hrp48 binding site. In the 30-mer-GG, we
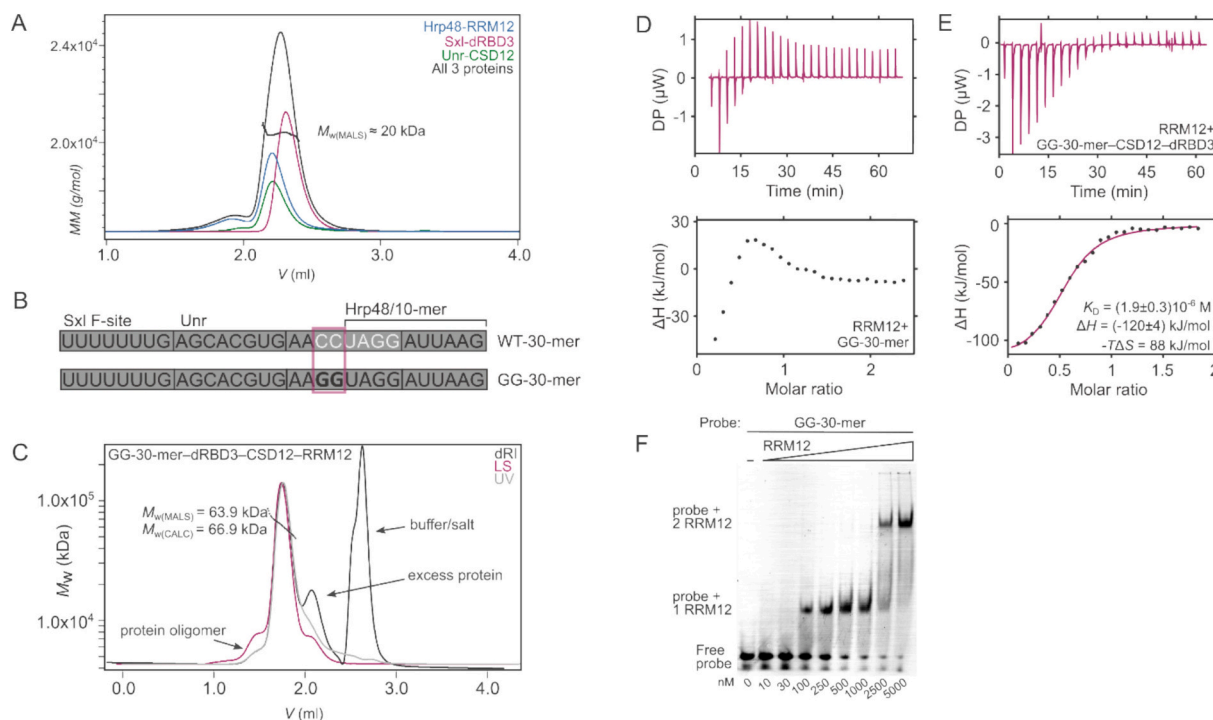
Fig. 5. Complex formation of Hrp48, Sxl, Unr and RNA. **A:** Hrp48-RRM12, Sxl-dRBD3 and Unr-CSD12 do not interact in the absence of RNA. SEC–MALS UV absorption chromatograms of the proteins Hrp48-RRM12 (blue), Sxl-dRBD3 (purple), Unr-CSD12 (green) and the proteins at the same concentration injected at once (grey). The molecular weight is around 20 kDa, which corresponds to the M_w of the free protein constructs. **B:** The RNA constructs used for complex formation. Both constructs contain one binding site for each protein. The GG-30-mer is a mutant in which the palindromic sequence is abolished by substituting the indicated (framed) CC into a GG pair to decouple the dimerization from protein binding in biophysical experiments. We also confirmed that this substitution has no effect on translation repression (Fig. S8F). **C:** SEC-MALS of the GG-30-mer-Hrp48-RRM12-Unr-CSD12-Sxl-dRBD3 complex. (LS), dark grey: differential refractive index (dRI), purple: light scattering, light grey: UV absorption. The molecular weight was calculated for a 1:1:1:1 complex. **D – E:** ITC of GG-30-mer titrated into Hrp48-RRM12 (D) and of GG-30-mer-Unr-CSD12-Sxl-dRBD3 titrated into Hrp48-RRM12 (E). **F:** Electrophoretic mobility shift assay using the GG-30-mer and increasing amounts of Hrp48-RRM12. At higher concentration of Hrp48 an additional shift appears, indicating that two Hrp48-RRM12 molecules bind one RNA. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

introduced a mutation to avoid the palindromic sequence and preclude formation of dsRNA (Fig. 5B). A negative effect of this mutant on the functionality was not detectable in translation assays (Fig. S8F). For complex formation studies involving RNA, we chose shortened constructs of Sxl (Sxl-dRBD3), Unr (Unr-CSD12) and Hrp48 (Hrp48-RRM12), that we know interact with *msl-2* mRNA.

To determine the best conditions for complex formation, we injected the two proteins Sxl-dRBD3 and Unr-CSD12 together with the GG-30-mer, with different molar ratios. We obtained the highest amounts of complex at a 1:2:3 ratio (RNA:dRBD3:CSD12, Fig. S8G). Then we optimized the Hrp48 amount, and accordingly we used two-fold excess with respect to the RNA (Fig. 5C). The resulting main peak of the chromatogram corresponds to the complex. The discrepancy of the calculated and determined molecular weight could be explained by low resolution and similar retention times of the overlapping peaks (Fig. 5C). While pre-forming the ternary complex with WT-30-mer in absence of Hrp48-RRM12, we found that the molecular weight is almost double of what was expected (Fig. S8I). This could be explained by dsRNA formation due to the palindromic sequence located within the WT-30-mer. For this reason, we used GG-30-mer in the ITC and electrophoretic mobility shift assays, to avoid additional interactions complicating the biophysical characterization of the complex. However, upon addition of Hrp48-RRM12, we could detect the monomeric complex, thus binding of Hrp48-RRM12 resolves the formed dsRNA structure of the WT-30-mer and forms a stable quaternary complex with 30-mer RNA, Sxl-dRBD3 and Unr-CSD12 (Fig. S8H, I).

Next, we wanted to assess whether Hrp48-RRM12 binding to RNA is synergistic with Sxl-dRBD3 and Unr-CSD12 binding to 30-mer RNA. To this end, we performed ITC experiments. GG-30-mer titrated into Hrp48-

RRM12 provided a complex, but reproducible ITC curve, which could indicate binding of multiple RRM12 domains on the long GG-30-mer (Fig. 5D). Another putative binding site of RRM12 on the GG-30-mer could be the Sxl F binding site, since our data of the 6-mer NMR titrations for RRM1 showed also weak interactions to the U_6 RNA (6-mer-5, Fig. S1E). However, it cannot be excluded that RRM12 could also bind to another part of the RNA. To check for multiple binding sites for Hrp48 on the GG-30-mer, we performed an electrophoretic mobility shift assay (EMSA), which resulted indeed in two sequential binding events (Fig. 5F). The first shift indicates binding of one Hrp48-RRM12 domain to the GG-30-mer and the second shift suggests that multiple Hrp48-RRM12 proteins could be bound to the RNA. We consider the second binding event as biologically irrelevant and as an *in vitro* artefact, as the affinity is very weak and this site is mostly occupied by Sxl and Unr. However, this could explain the complex ITC curve we obtained (Fig. 5D). ITC titrations of a pre-formed complex of GG-30-mer, Sxl-dRBD3 and Unr-CSD12 titrated into Hrp48-RRM12 abolished the second binding event and we could measure the K_D of Hrp48-RRM12 binding to this preformed complex. To our surprise, the affinity was about 100-fold weaker than when Hrp48-RRM12 interacts with the 10-mer RNA, which at the very least rules out synergistic binding of Hrp48-RRM12 to GG-30-mer in the presence of Sxl-dRBD3 and Unr-CSD12. However, it cannot be excluded that Hrp48's C-terminal IDR interacts with Sxl, Unr or both proteins. We could not include this IDR in our current studies as it mediated severe aggregation.

3. Discussion

In summary, we identified the binding site for the individual RRM

domains of Hrp48 by mapping the region of *msl-2* using 4-mers and 6-mers. Since the binding affinity upon 4-mer interaction decreased significantly compared to 6-mer binding, we concluded that the smallest possible binding motif for the single domains must be at least five nucleotides long. Furthermore, the recognition sequence for each RRM domain of Hrp48 includes a motif highly similar to the consensus binding sequence (UAGG) of RRM1 of hnRNP A1, a putative mammalian homolog to Hrp48 [12,14]. The crystal structure also shows the characteristic βαββαβ fold for Hrp48-RRM1, whereas RRM2 possesses an additional β-strand, which is also involved in RNA binding. This non-canonical RRM fold has also been observed for PTB RRM3 where it extends the length of the β-sheet to accommodate more nucleotides and thus increase sequence-specificity [18]. This could be similar for Hrp48, as CSPs were observed upon RNA binding for residues of this additional β-strand. Furthermore, the binding affinity of RRM2 towards 4-mer RNA decreases more than for RRM1, which could also indicate that more than four nucleotides are bound on the extended β-sheet of RRM2. On the other hand, CSPs on the additional β-strand could also occur due to allosteric effects. Additional β-strands have been reported for both RRM domains of TDP-43 which promote protein–protein interactions [53]. Other secondary structure elements as addition to the canonical RRM fold have been shown to increase plasticity and RNA specificity of RRMs (*e.g.* [17,25,84]). Preliminary data suggests that the self-interaction of Hrp48 through its RRM2 domain diminishes upon RNA binding, and that the RNA binding interface of RRM2 overlaps with its tetramerization interface. This might suggest a biological relevance of tetramerization, which might prevent unspecific RNA binding as in a tetramer the RNA binding interface of RRM2 is potentially buried. Dimerization or oligomerization tendencies of RRM domains to be biologically relevant has been shown for several other examples (*e.g.* [58,66,67]) The tetramerization tendency might also be a step in formation of higher oligomeric assemblies and eventually phase separation, as it has been shown in the context of *P-element* mRNA localization [13].

Both RRM domains of Hrp48 bind to 6-mer RNA motifs adjacent to each other, hence suggesting a continuous binding site of 10 nucleotides for the tandem RRM12 domain. The high affinity, simultaneous binding of RRM12 and the high similarity to the *P-element* mRNA binding site [74,75] also support the hypothesis of 10-mer RNA (UAGGAUUAAG) being the optimal sequence for Hrp48 binding on *msl-2*-mRNA. Additionally, RRM2 binds upstream of RRM1 which is also consistent with the binding behavior of RRM tandem domains of other proteins [20,38,71]. Studies about dynamics revealed that both RRM domains which are connected through a short seven nucleotide linker do not interact with each other in absence of RNA but tumble jointly in an RNA bound state. Experimental data strengthen this hypothesis, since 9-mer RNA, representing the *P-element* mRNA binding site which does not contain the central adenosine, has equal binding affinity to RRM12 as the 10-mer. Furthermore, the adenosine does not provide additional flexibility, since rotational correlation times of RRM12 bound to 10-mer-WT and 9-mer RNA are identical.

Activity assays confirmed that upon mutation of the binding site, translation is partially de-repressed which supports the validity of the binding site in a functional aspect and that Hrp48 contributes to translation repression of *msl-2*. In context of complex formation, Hrp48-RRM12 does not interact with Sxl and Unr in absence of RNA and also does not establish cooperative protein-protein contacts when bound to RNA. However, the results of complex formation studies only provide information about the RNA binding domains of the three proteins, thus, potential protein-protein interactions are still possible if they are established through contacts between Hrp48 IDRs or with other CSDs of Unr. A similar situation has been observed for the quaternary protein-RNA complex of Brain tumour (Brat), Pumilio (Pum) and Nanos, which together repress the translation of *hunchback* mRNA. This is an important step during *Drosophila* development, regulating the establishment of the anterior-posterior body axis. While the Pum-HD-Nanos-ZnF and Brat-NHL domains bind to adjacent RNA sequences [56,57,85],

only Pum-HD and Nanos-ZnF exhibit cooperativity, whereas Brat-NHL binds independent of the Pum-HD-Nanos-ZnF module to RNA [59]. However, also in this case it cannot be excluded that the IDRs of each protein would engage in further cooperative contacts. Despite the experimental difficulties often associated with IDRs (aggregation and precipitation), it will be essential for future studies to include these regions to obtain a complete picture of these complexes' mechanisms of action.

For the complex studied here, it needs to be considered that other proteins are also involved and could engage in further protein-protein contacts. Especially poly(A)-binding protein (PABP), which promotes closed-loop formation of the mRNA [82] and translation has been shown to take part in the *msl-2*-mRNA translation repression mechanism by interacting with the Sxl-Unr complex [24]. The interactions between Unr and PABP are direct [42] but the molecular mechanism behind PABP involvement remains elusive.

The Sxl-Unr-Hrp48 and the Pum-Nanos-Brat complexes are promising model systems to obtain insights into regulation of translation initiation and we are optimistic that cryo-EM or integrative structural biology [21] will enable high-resolution structure determination of such full-length complexes in the future.

4. Materials and methods

4.1. Protein cloning, expression and purification

The sequences of Hrp48 (UniProt code P48809), RRM1 (1–88), RRM2 (89–173), and the tandem RRM12 (1–173) constructs were cloned into the pETM- vector using restriction-free cloning [81]. Hrp48 constructs were expressed in *Escherichia coli* BL21 (DE3). Expression and purification of Sxl-dRBD3 (123–294) and all Unr constructs used in this study are described elsewhere [39,41]. Proteins were expressed in TB and LB medium or in M9 minimal medium for isotope labeling using $^{15}NH_4Cl$ for ^{15}N labeled samples, or $^{15}NH_4Cl$ and D-Glucose-$^{13}C_6$ for ^{15}N, ^{13}C labeled samples. Generally, precultures were grown in the same medium as the expression medium overnight at 37 °C. Cultures were inoculated with an OD_{600} of 0.04 and grown until the logarithmic growth phase was reached, followed by induction with 0.2 mM IPTG and overnight expression at 18 °C. The harvested cells were resuspended in ice cold lysis buffer (20 mM Hepes, 500 mM NaCl, 12 mM imidazole, 0.5 mM β-mercaptoethanol, pH 7.2) and then frozen and sonicated for further lysis. The thawed lysate was centrifuged for 30 mins at 18000 $\times g$ at 4 °C and the supernatant was syringe-filtered with 0.45 μm pore size filter. The cleared lysate was loaded three times on a Nickel-nitrilotriacetic acid (Ni-NTA) gravity flow column that was pre-conditioned in binding buffer (20 mM Hepes, 500 mM NaCl, 12 mM imidazole, pH 7.2). followed by adding the cleared lysate, washing with 20 column volumes of binding buffer and elution with 5 column volumes of elution buffer (20 mM Hepes, 500 mM NaCl, 200 mM Imidazole, pH 7.2). The His$_6$-tag was cleaved by TEV protease and 2 mM β-mercaptoethanol, incubated on ice for 1 h, and then dialyzed overnight at 4 °C into binding buffer to remove the β-mercaptoethanol. The sample was reloaded onto the Ni-NTA gravity flow column collecting the flow-through to remove the cleaved tag. For Unr CSD1, CSD12, and CSD123, the reverse Ni-affinity step was followed by a heparin purification. For this, the flow-through was diluted five times to lower the concentration, and then dialyzed against the low salt heparin binding buffer (10 mM MES, 50 mM NaCl, 2 mM DTT, pH 6.0) overnight. The protein was injected onto a HiTrap™ 5 mL Heparin HP, washed with heparin binding buffer and eluted with a heparin elution buffer (20 mM MES, 1500 mM NaCl, 2 mM DTT, pH 6.0) to remove bacterial RNA contamination. All proteins were further purified by size-exclusion chromatography (SEC) on a HiLoad 16/600 Superdex S75 or S200 pg column equilibrated with SEC buffer (20 mM MES, 200 mM NaCl, pH 6.5). Hrp48 RRM2 and RRM12 were kept at low concentrations (0.1–0.2 mg/mL) during all steps after the second Ni-NTA purification step to

avoid aggregation.

For Sxl-Unr-Hrp48-*msl-2* complex formation, the purified Sxl-dRBD3, Unr-CSD12 or CSD1 and Hrp48-RRM12 constructs and the WT-30-mer or GG-30-mer RNA were incubated on ice in a 2:3:2:1 ratio. The complex was concentrated with a 3.5 kDa cut-off concentrator unit. Depending on the required quantity, the final volume was ~300 μL for SEC purification on a Superdex 75 10/300 GL column, or ~1 mL for the HiLoad 16/600 Superdex 75 pg column. The identity of the complex peak was confirmed by UV absorption measurement at 260 and 280 nm and also using SEC-MALS weight determination.

After SEC, the pooled fractions were concentrated and the concentration was measured at 260 nm on a NanoDrop UV-VIS absorption spectrophotometer with an extinction coefficient of 325,000 (M^{-1} cm^{-1}), that is the sum of the extinction coefficients of the components at 260 nm.

4.2. Mutagenesis of RRM1

The introduction of point-mutations into RRM1 (1–88) was carried out according to the mutagenesis protocol of Edelheit et al. from 2009 [26]. Primers were designed that carried up to three point-mutations to alter the codon of the amino acid of interest into a GCG motif that codes for alanine. For each mutant, two single-primer PCR reactions were set up, containing 500 ng WT-RRM1 DNA template, 200 μM dNTPs, 0.02 U/mL Q5 polymerase, Q5 buffer and 40 pmol of either forward-primer or reverse-primer. The two PCR products were combined and reannealed. The DNA template was digested over night by addition of *Dpn*I and rCutsmart buffer at room temperature. The reaction was transformed into electrocompetent *E. coli* XL1-Blue cells, spread on LB plates with antibiotics and incubated over night at 37 °C. Clones were picked and sent for sequencing.

4.3. Isothermal titration calorimetry

ITC measurements were performed using a Malvern Panalytical MicroCal PEAQ-ITC calorimeter at 20 °C. For all experiments, the protein was loaded into the cell and RNA or protein–RNA complex into the syringe (Table 1). Prior, the samples were dialyzed against MES buffer (20 mM MES, 200 mM NaCl, 0.02 % NaN3, pH 6.5) buffer overnight, adjusted to the appropriate concentrations, centrifuged at 15000 rpm for 10 min, transferred to the final tubes for measurements and degassed for 5 min. The titrations were accomplished with 19–25 injections corresponding to 1.5–2 μL injection volumes. Each injection lasted 3 s followed by a 150 s delay. The sample stirring was set to 750 rpm with reference power was 10 μcal/s. Further details about the concentrations and experimental setup in individual titrations are listed in Table 1. The software MicroCal PEAQ-ITC analysis software (Malvern Panalytical) was used for data analysis.

4.4. NMR spectroscopy

All NMR spectra were recorded on Bruker Avance III NMR spectrometers with magnetic field strengths corresponding to proton Larmor frequencies of 600 MHz, 700 MHz, 800 MHz or 1 GHz equipped with a room temperature triple resonance probe head (700 MHz), or a cryogenic triple resonance gradient probe head (600, 800 MHz and 1 GHz) at 298 K. The NMR samples were measured in SEC buffer with 5 % D_2O for the deuterium lock. The multidimensional experiments were recorded using apodization weighted sampling [76]. Backbone resonance assignment of ^{13}C, ^{15}N-labeled Hrp48 constructs was achieved to a completion of 100 % (RRM1, excluding prolines), 94 % (RRM2), and 95 % (RRM12) in the free state and 79 % in the 10-mer bound state using ^1H,^{15}N-HSQC, HNCO, HN(*CA*)CO, HNCA, CBCA(*CO*)NH, HNCACB triple resonance experiments [70]. All NMR spectra were processed using NMRPipe [19], analyzed using CcpNmr Analysis [77,83], CARA [47], Sparky [55] and NMRViewJ [43]. Backbone torsion angles were

predicted from C_α and C_β chemical shifts for RRM2 using TALOS-N [73]. The NMR backbone chemical shift assignments are deposited at the Biological Magnetic Resonance Databank under the following accession codes: 52354 (Hrp48 RRM2), 52359 (Hrp48 RRM1), 52362 (Hrp48 RRM12 bound to 10-mer-WT RNA).

NMR titrations were performed by recording two-dimensional ^{15}N,^1H-HSQC spectra of the labeled protein and for each subsequent addition of the titrant protein or RNA, until reaching saturation (*i.e.* no further chemical shift perturbation could be observed) or to a 1:1 ratio. For NMR titration experiments, various protein concentrations were used: ^{15}N-labeled Hrp48-RRM1, different RRM1-mutants, RRM2, and RRM12 were titrated with different purchased RNA oligos, protein–protein interactions were tested by titrating ^{15}N-labeled Hrp48 RRM12 with unlabeled Sxl dRBD3, Unr CSD123, CSD456, and CSD789 (details described in Table 3). CcpNmr Analysis, Sparky and scripts used in our lab were used to trace chemical shift perturbations and determine dissociation constants (K_D) for interactions in the fast exchange regime. Individual K_D values of peaks were averaged to calculate global K_D values [86]. Errors were calculated from the individual fitting errors by error propagation. In the intermediate and slow exchange regime, the data fitting is erroneous and for comparative reasons the qualitative appearance of the spectra was assessed. Chemical shift perturbations were calculated according to

$$CSP = \sqrt{\frac{1}{2}\left[\delta_H^2 + (\alpha \bullet \delta_N)^2\right]}$$

where $\alpha = 0.2$ was used based on the ratio between peak dispersion in the ^1H and ^{15}N dimensions [86].

Measurements of longitudinal, R_1 and transverse, R_2 relaxation rates experiments for Hrp48-RRM12, Hrp48-RRM12–10-mer RNA complex and Hrp48-RRM12–9-mer RNA complex were acquired at proton Larmor frequencies of 700 MHz at 298 K using standard pulse sequences [46,54]. The sample concentration was 80 μM for RRM12 in the apo state and 200 μM for RRM12 in the RNA bound state. The 9-mer and 10-mer RNA were added at 1.2-fold excess. For RRM12 alone, only 2 points were measured because of the aggregation tendency of RRM12 without RNA at elevated concentrations. For the R_1 experiment, relaxation delays were 80 and 1200 ms and for the R_2 experiment 16 and 80 ms. For RNA-bound Hrp48-RRM12 each experiment was recorded with 6 and 9 relaxation delays. For R_1, delays of 80, 400, 800, 1200, 1600, and 2000 ms were used with 80, 800, 1200 and 2000 ms as duplicates. For R_2

Table 3
Relative and absolute concentrations used for NMR titrations.

Analyte protein in NMR tube	Conc. NMR tube (μM)	Titrants	Molar ratios
Hrp48-RRM1	20/50/100	6-mer-1, 6-mer-2, 6-mer-3, 6-mer-4, 6-mer-5, 4-mer-3, 4-mer-6	1.0:0.0–1.0:2.5 or 1.0:3.0
Hrp48-RRM1-W16A	30	6-mer-3	1.0:0.0–1.0:3.0
Hrp48-RRM1-W16A-R75A	30	6-mer-3	1.0:0.0–1.0:3.0
Hrp48-RRM1-D35A	30	6-mer-3	1.0:0.0–1.0:3.0
Hrp48-RRM1-D78A-K80A	30	6-mer-3	1.0:0.0–1.0:3.0
Hrp48-RRM1-R85A	30	6-mer-3	1.0:0.0–1.0:3.0
Hrp48-RRM2	20/50	6-mer-1, 6-mer-2, 6-mer-3, 6-mer-4, 6-mer-5, 4-mer-3, 4-mer-6	1.0:0.0–1.0:2.5 or 1.0:3.0
Hrp48-RRM12	100, 50	10-mer, Sxl-dRBD4, Unr-CSD123, Unr-CSD456, Unr-CSD789, 4-mers	1.0:0.0–1.0:2.5 1.0:0.0–1.0:1.0 or 1.0:1.5
Sxl-dRBD3	100	Hrp48-RRM12	1.0:0.0–1.0:2.5

delays, 16, 32, 48, 64, 80, 96, 112, 128 and 144 ms relaxation delays were used with 32 ms as duplicates. For peak integration, error estimation and exponential fitting NMRViewJ was used [43]. Calculations of the relaxation rates from two-point measurements were completed in spreadsheets and errors were estimated through calculation of the standard deviation of the mean value of τ_c of each residue within a single RRM domain. The mean value was than calculated again for residues with τ_c values within the standard deviation. Residues outside of this range were considered as outliers. The experimental rotational correlation times were calculated according to [29].

4.5. Crystallization, data collection, and structure determination

Crystallization trials were performed using 3-lens 288-well crystallization plates using the sitting drop method, in which two sample concentrations/conditions can be tested with the same crystallization buffer. For this, 5–6 different 96-condition commercial screens were used at 4 °C and at 20 °C. Each well contained 0.1 μL sample and 0.1 μL crystallization condition solution, set up with a Mosquito LCP liquid handling robot. Several conditions yielded crystals for Hrp48-RRM1. For the final crystal that gave the highest resolution in X-ray diffraction, RRM1 of 5 mg/mL concentration in the 200 mM NaCl, 30 mM NaPi, 2 mM DTT, pH 6.5. buffer was mixed in a 1:1 ratio (100 nL:100 nL) with 0.2 M K_2SO_4 and 20 % (*w/v*) PEG-3350 reservoir solution at 4 °C. The sitting drop vapor diffusion method was used, and rod-shaped crystals started to nucleate overnight and kept growing further for 7–10 days. Crystals were soaked in mother liquor supplemented with 30 % glycerol as a cryoprotectant prior to freezing.

Diffraction datasets were recorded at the P13 beamline at the German Electron Synchrotron (DESY), Hamburg, Germany. The crystals diffracted up to 1.15 Å resolution.

The structure of RRM1 was solved by *ab initio* molecular replacement with the human hnRNP A1 RRM1 structure (pdb 1HA1, 40 % sequence identity) as model using Phenix Phaser-MR [3]. The initial model was built using Phenix AutoBuild [80] and manual adjustments were executed with Coot [27]. The structure was further improved in iterative rounds of manual correction with Coot and restrained refinement with phenix.refine [4]. The crystal structure is deposited at the Protein Data Bank under the following accession code: 9EN7. Structural statistics are shown in Table 2.

4.6. Molecular dynamics simulation of the RRM1–6-mer-3 complex

To simulate the spontaneous binding process [51] of the 5'-UUAAG-3' motif of the 6-mer-3 RNA (AUUAAG) to the RRM1 domain, we have constructed a system with the RNA positioned ~20 Å away from the protein. Note that the 5'-terminal A_1 of the full 6-mer-3 was skipped to increase sampling efficiency of the simulations. The initial coordinates of the protein and RNA were obtained from the X-ray structure of the RRM1 domain (PDB: 9EN7) and by NAB (Nucleic Acid Builder), respectively [16]. The OL3 [87] and ff12SB [60] force fields were used to describe the RNA and protein, respectively. We have applied the recently introduced stafix approach (using scaling factor of 0.5) to eliminate spurious intramolecular RNA self-interactions occurring with the standard AMBER RNA simulation force field [51]. This enabled observation of the spontaneous RNA-protein binding process. The RNA and the protein were immersed in an octahedral box of SPC/E water molecules [11] with minimal distance of 14 Å between the solutes and the box border. We have added the K^+ and Cl^- ions [44] to neutralize the systems and obtain an ion concentration of ~0.15 M. Prior to the production simulations, the systems were minimized and equilibrated [52]. The production simulations were then performed in constant pressure ensemble. Monte-Carlo barostat and Langevin thermostat were used to control the pressure and temperature, respectively [16]. We have carried out nine 10-μs-long independent MD simulations, with different trajectories obtained by utilizing random seed numbers. The resulting

trajectories visualized the RNA at different stages of binding to RRM1. By careful visual analysis, we have subsequently manually selected a binding motif showing the RNA stably bound at a location close to all the protein residues which exhibited chemical shift perturbations in NMR experiments. In addition to the manual selection, we have also performed a selection using an automatic method involving a measurement of pairwise distances between the C1' atoms of the RNA and the nitrogens of all the protein amide groups for which the CSPs were experimentally measured upon RNA binding. Each pairwise distance was weighted by the experimentally measured magnitude of the corresponding CSP and all the distances were combined into a sum. The 1 % of the simulation frames with the lowest value of this sum were then selected from the entire combined simulation ensemble and k-means clustering analysis was performed on them using cpptraj, obtaining twenty clusters and their corresponding representative structures. Finally, these representative structures were sorted into a PDB ensemble by the value of the sum, from lowest to highest. We note the protein-RNA interface of the manually selected structure was virtually identical to the structures of the automatically selected ensemble. Long-term stability of the chosen binding motif was subsequently verified in two independent 10-μs-long MD simulations. Finally, the structure was used as our working model for the putative RRM1/6-mer-3 protein-RNA complex structure.

4.7. Size exclusion chromatography-multi angle light scattering (SEC-MALS)

For SEC-MALS experiments different columns were used: a Superdex 200 Increase 10/300 GL or a Superdex 200 Increase 5/150 GL gel-filtration column coupled to the MiniDAWN and Optilab MALS system from Wyatt Technology. For each complex the single components (proteins and RNA alone) were measured first followed by the complex. After mixing and incubation on ice, the samples were centrifuged for 10 mins at 15000 rpm and then 50 μL were loaded onto the Superdex 200 Increase 5/150 GL or 100 μL onto the Superdex 200 Increase 10/300 GL columns. The minimal concentration was 1.0 mg/mL for each sample. The experiments were performed at room temperature in SEC buffer filtered twice through 0.22 μm pore-size filter. Data analysis was performed using the Astra 7 software (Wyatt Technology).

4.8. In vitro translation assays

Wild type construct BLEF, as previously described [32], is composed of 69 nt of *msl-2* 5' UTR sequence including site B and 46 nt of the *msl-2* 3' UTR including sites E and F. For introducing mutations in the 3' UTR, primers were created by QuikChange Primer Design (https://www.agilent.com/) and used for restriction-free cloning to create changes in the Region 5 and Region 6. Sxl-dRBD4 (amino acids 122–301 of *Drosophila melanogaster* SXL) was expressed in *Escherichia coli* as an N-terminal GST-tagged fusion protein and purified as described [36]. The protein was dialyzed against buffer D (20 mM HEPES pH 8.0, 20 % glycerol,1 mM DTT, 0.01 % NP-40, 0.2 mM EDTA). BLEF mRNA derivatives were synthesized using T3 RNA polymerase (Ambion) and contained a 5' m^7GpppG cap and a poly(A) tail of 73 residues. mRNAs were purified by phenol-chloroform extraction and G_{50} columns (GE Heathcare). All mRNAs used in the same experiment were synthesized and quantified in parallel, and the concentration and quality confirmed by separation in agarose gels. *In vitro* translation assays were performed as described [31]. Briefly, 17 ng of *msl-2* Firefly reporter mRNA and 10 ng *Renilla* luciferase mRNA, used as an internal control, were incubated with increasing amounts of GST-dRBD4 in a final volume of 10 μL. The reaction contained 40 % *Drosophila* embryo extract, 60 μM amino acids, 16.8 mM creatine phosphate, 80 ng/μL creatine kinase, 24 mM HEPES pH 7.5, 0.6 mM Mg(OAc)$_2$ and 80 mM KOAc. The reaction was incubated at 25 °C for 90 min, and the Firefly and *Renilla* luciferase activities were measured using the Dual Luciferase kit (Promega).

4.9. Electrophoretic mobility shift assays

The GG-30-mer RNA oligo used for EMSAs was 3' end-labeled with pCp-Cy-5 (Cyanine 5). For this, the following reaction mixture was combined in 20 μL final volume: 100 pmol RNA, 200 pmol pCp-Cy-5, 2 μL T4 RNA ligase (10×, 20 U), 2 μL DMSO, 1 mM ATP, 1 mM DTT, 10 mM $MgCl_2$, 50 mM Tris-HCl (pH 7.5). The reaction was incubated at 16 °C overnight.

Subsequently the reaction was purified by NaOAc/EtOH precipitation. To the reaction mixture, 4 μL (0.2 × volume) of 3 M sodium-acetate, 1.5 μL of GlycoBlue Coprecipitant (Invitrogen), and 59 μL (2.5 × volume) of ethanol cooled to −20 °C was added and mixed by vortexing. The RNA was precipitated overnight at −70 °C, centrifuged at 4 °C and 13,000 rpm. For 30 min, washed two times with −20 °C 70 % ethanol and once with 100 % ethanol. The final amount and the labeling efficiency were measured by spectrophotometry using the NanoDrop™ OneC.

Electrophoretic mobility shift assays (EMSAs) were used to determine RNA-binding affinities in a semi-quantitative mode [30]. Recombinantly purified Hrp48-RRM12 was mixed with 20 nM Cy-5 labeled GG-30-mer probe in a buffer consisting of 50 mM Tris-HCl (pH 8.4), 100 mM NaCl, 1 mM EDTA, 250 mg/mL BSA, 10 % glycerol, 0.05 % Triton X-100 and 1 mM DTT in 12.5 μL reactions and incubated on ice for 30 min, protected from light. The concentration of proteins is indicated in the figure. The RNA – protein complexes were resolved on a 6 % native 1 × TBE polyacrylamide gel for 35 min at 200 V. The gels were imaged at a Typhoon Trio Imager 9000 (GE Healthcare).

CRediT authorship contribution statement

Andrea Lomoschitz: Writing – review & editing, Writing – original draft, Visualization, Methodology, Investigation, Formal analysis, Conceptualization. **Julia Meyer:** Writing – review & editing, Writing – original draft, Visualization, Validation, Methodology, Investigation, Formal analysis, Conceptualization. **Tanit Guitart:** Writing – review & editing, Methodology, Investigation, Formal analysis. **Miroslav Krepl:** Writing – review & editing, Methodology, Investigation, Formal analysis. **Karine Lapouge:** Writing – review & editing, Methodology, Investigation, Formal analysis. **Clara Hayn:** Writing – review & editing, Methodology, Investigation, Formal analysis. **Kristian Schweimer:** Writing – review & editing, Supervision, Methodology, Investigation, Formal analysis. **Bernd Simon:** Writing – review & editing, Supervision, Methodology, Investigation, Formal analysis. **Jiří Šponer:** Writing – review & editing, Supervision, Resources, Methodology. **Fátima Gebauer:** Writing – review & editing, Supervision, Resources, Methodology. **Janosch Hennig:** Writing – review & editing, Writing – original draft, Visualization, Supervision, Resources, Methodology, Funding acquisition, Formal analysis, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.bpc.2024.107346.

References

[1] I. Abaza, O. Coll, S. Patalano, F. Gebauer, Drosophila UNR is required for translational repression of male-specific lethal 2 mRNA during regulation of X-chromosome dosage compensation, Genes Dev. 20 (3) (2006) 380–389. https://pubmed.ncbi.nlm.nih.gov/16452509/. https://doi.org/10.1101/gad.371906.

[2] I. Abaza, F. Gebauer, Functional domains of drosophila UNR in translational control, RNA (New York, N.Y.) 14 (3) (2008) 482–490. https://pubmed.ncbi.nlm.nih.gov/18203923/. https://doi.org/10.1261/rna.802908.

[3] P.D. Adams, P.V. Afonine, G. Bunkóczi, V.B. Chen, I.W. Davis, N. Echols, et al., PHENIX: a comprehensive Python-based system for macromolecular structure solution, Acta Crystallogr. D Biol. Crystallogr. 66 (Pt 2) (2010) 213–221. https://pubmed.ncbi.nlm.nih.gov/20124702/. https://doi.org/10.1107/S0907444909052925.

[4] P.V. Afonine, R.W. Grosse-Kunstleve, N. Echols, J.J. Headd, N.W. Moriarty, M. Mustyakimov, et al., Towards automated crystallographic structure refinement with phenix.Refine, Acta Crystallogr. D Biol. Crystallogr. 68 (Pt 4) (2012) 352–367. https://pubmed.ncbi.nlm.nih.gov/22505256/. https://doi.org/10.1107/S0907444912001308.

[5] S.D. Auweter, R. Fasan, L. Reymond, J.G. Underwood, D.L. Black, S. Pitsch, et al., Molecular basis of RNA recognition by the human alternative splicing factor Fox-1, EMBO J. 25 (1) (2006) 163–173, https://doi.org/10.1038/sj.emboj.7600918.

[6] S.D. Auweter, F.C. Oberstrass, F.H.-T. Allain, Sequence-specific binding of single-stranded RNA: is there a code for recognition? Nucleic Acids Res. 34 (17) (2006) 4943–4959. https://pubmed.ncbi.nlm.nih.gov/16982642/. https://doi.org/10.1093/nar/gkl620.

[7] G.J. Bashaw, B.S. Baker, The Msl-2 dosage compensation gene of drosophila encodes a putative DNA-binding protein whose expression is sex specifically regulated by sex-lethal, Development (Cambridge, England) 121 (10) (1995) 3245–3258. https://pubmed.ncbi.nlm.nih.gov/7588059/. https://doi.org/10.1242/dev.121.10.3245.

[8] G.J. Bashaw, B.S. Baker, The regulation of the Drosophila msl-2 gene reveals a function for sex-lethal in translational control, Cell 89 (5) (1997) 789–798. https://pubmed.ncbi.nlm.nih.gov/9182767/. https://doi.org/10.1016/s0092-8674(00)80262-7.

[9] K. Beckmann, M. Grskovic, F. Gebauer, M.W. Hentze, A dual inhibitory mechanism restricts msl-2 mRNA translation for dosage compensation in Drosophila, Cell 122 (4) (2005) 529–540. https://pubmed.ncbi.nlm.nih.gov/16122421/. https://doi.org/10.1016/j.cell.2005.06.011.

[10] J.M. Belote, J.C. Lucchesi, Male-specific lethal mutations of Drosophila melanogaster, Genetics 96 (1) (1980) 165–186, https://doi.org/10.1093/genetics/96.1.165.

[11] H.J.C. Berendsen, J.R. Grigera, T.P. Straatsma, The missing term in effective pair potentials, J. Phys. Chem. 91 (24) (1987) 6269–6271, https://doi.org/10.1021/j100308a038.

[12] I. Beusch, P. Barraud, A. Moursy, A. Cléry, F.H.T. Allain, Tandem hnRNP A1 RNA recognition motifs act in concert to repress the splicing of survival motor neuron exon 7, eLife 6 (2017), https://doi.org/10.7554/eLife.25736.

[13] M. Bose, M. Lampe, J. Mahamid, A. Ephrussi, Liquid-to-solid phase transition of oskar ribonucleoprotein granules is essential for their function in Drosophila embryonic development, Cell 185 (8) (2022) 1308–1324.e23. https://pubmed.ncbi.nlm.nih.gov/35325593/. https://doi.org/10.1016/j.cell.2022.02.022.

[14] C.G. Burd, G. Dreyfuss, RNA binding specificity of hnRNP A1: significance of hnRNP A1 high-affinity binding sites in pre-mRNA splicing, EMBO J. 13 (5) (1994) 1197–1204. https://pubmed.ncbi.nlm.nih.gov/7510636/. https://doi.org/10.1002/j.1460-2075.1994.tb06369.x.

[15] S. Campagne, D. Jutzi, F. Malard, M. Matoga, K. Romane, M. Feldmuller, et al., Molecular basis of RNA-binding and autoregulation by the cancer-associated splicing factor RBM39, Nat. Commun. 14 (1) (2023) 5366. https://www.nature.com/articles/s41467-023-40254-5. https://doi.org/10.1038/s41467-023-40254-5.

[16] D.A. Case, H.M. Aktulga, K. Belfon, D.S. Cerutti, G.A. Cisneros, V.W.D. Cruzeiro, et al., AmberTools, J. Chem. Inf. Model. 63 (20) (2023) 6183–6191, https://doi.org/10.1021/acs.jcim.3c01153.

[17] A. Cléry, M. Blatter, F.H.-T. Allain, RNA recognition motifs: boring? Not quite, Curr. Opin. Struct. Biol. 18 (3) (2008) 290–298. https://pubmed.ncbi.nlm.nih.gov/18515081/. https://doi.org/10.1016/j.sbi.2008.04.002.

[18] M.R. Conte, T. Grüne, J. Ghuman, G. Kelly, A. Ladas, S. Matthews, et al., Structure of tandem RNA recognition motifs from polypyrimidine tract binding protein reveals novel features of the RRM fold, EMBO J. 19 (12) (2000) 3132–3141. https://pubmed.ncbi.nlm.nih.gov/10856256/. https://doi.org/10.1093/emboj/19.12.3132.

[19] F. Delaglio, S. Grzesiek, G.W. Vuister, G. Zhu, J. Pfeifer, A. Bax, NMRPipe: a multidimensional spectral processing system based on UNIX pipes, J. Biomol. NMR 6 (3) (1995) 277–293. https://pubmed.ncbi.nlm.nih.gov/8520220/. https://doi.org/10.1007/BF00197809.

[20] R.C. Deo, J.B. Bonanno, N. Sonenberg, S.K. Burley, Recognition of polyadenylate RNA by the poly(a)-binding protein, Cell 98 (6) (1999) 835–845. https://pubmed.ncbi.nlm.nih.gov/10499800/. https://doi.org/10.1016/s0092-8674(00)81517-2.

[21] L. Dimitrova-Paternoga, P.K.A. Jagtap, P.-C. Chen, J. Hennig, Integrative structural biology of protein-RNA complexes, Structure (London, England 1993) 28 (1) (2020) 6–28. https://pubmed.ncbi.nlm.nih.gov/31864810/. https://doi.org/10.1016/j.str.2019.11.017.

[22] C.M. Disteche, Dosage compensation of the sex chromosomes, Annu. Rev. Genet. 46 (2012) 537–560. https://pubmed.ncbi.nlm.nih.gov/22974302/. https://doi.org/10.1146/annurev-genet-110711-155454.

[23] K. Duncan, M. Grskovic, C. Strein, K. Beckmann, R. Niggeweg, I. Abaza, et al., Sex-lethal imparts a sex-specific function to UNR by recruiting it to the msl-2 mRNA 3' UTR: translational repression for dosage compensation, Genes Dev. 20 (3) (2006) 368–379. https://pubmed.ncbi.nlm.nih.gov/16452508/. https://doi.org/10.1101/gad.371406.

[24] K.E. Duncan, C. Strein, M.W. Hentze, The SXL-UNR corepressor complex uses a PABP-mediated mechanism to inhibit ribosome recruitment to msl-2 mRNA, Mol. Cell 36 (4) (2009) 571–582. https://pubmed.ncbi.nlm.nih.gov/19941818/. https://doi.org/10.1016/j.molcel.2009.09.042.

[25] M.M. Duszczyk, H. Wischnewski, T. Kazeeva, R. Arora, F.E. Loughlin, C. von Schroetter, et al., The solution structure of dead end bound to AU-rich RNA reveals an unusual mode of tandem RRM-RNA recognition required for mRNA regulation, Nat. Commun. 13 (1) (2022) 5892. https://www.nature.com/articles/s41467-022-33552-x. https://doi.org/10.1038/s41467-022-33552-x.

[26] O. Edelheit, A. Hanukoglu, I. Hanukoglu, Simple and efficient site-directed mutagenesis using two single-primer reactions in parallel to generate mutants for protein structure-function studies, BMC Biotechnol. 9 (1) (2009) 61. https://bmcbiotechnol.biomedcentral.com/articles/10.1186/1472-6750-9-61. https://doi.org/10.1186/1472-6750-9-61.

[27] P. Emsley, B. Lohkamp, W.G. Scott, K. Cowtan, Features and development of coot, Acta Crystallogr. D Biol. Crystallogr. 66 (Pt 4) (2010) 486–501. https://pubmed.ncbi.nlm.nih.gov/20383002/. https://doi.org/10.1107/S0907444910007493.

[28] P. Förch, L. Merendino, C. Martínez, J. Valcárcel, Modulation of Msl-2 5' splice site recognition by sex-lethal, RNA (New York, N.Y.) 7 (9) (2001) 1185–1191. https://pubmed.ncbi.nlm.nih.gov/11565743/. https://doi.org/10.1017/s1355838201010536.

[29] D. Fushman, R. Weisemann, H. Thüring, H. Rüterjans, Backbone dynamics of ribonuclease T1 and its complex with 2'GMP studied by two-dimensional heteronuclear NMR spectroscopy, J. Biomol. NMR 4 (1) (1994) 61–78. https://link.springer.com/article/10.1007/BF00178336. https://doi.org/10.1007/BF00178336.

[30] M.M. Garner, A. Revzin, A gel electrophoresis method for quantifying the binding of proteins to specific DNA regions: application to components of the Escherichia coli lactose operon regulatory system, Nucleic Acids Res. 9 (13) (1981) 3047–3060. https://pubmed.ncbi.nlm.nih.gov/6269071/. https://doi.org/10.1093/nar/9.13.3047.

[31] F. Gebauer, D.F. Corona, T. Preiss, P.B. Becker, M.W. Hentze, Translational control of dosage compensation in Drosophila by sex-lethal: cooperative silencing via the 5' and 3' UTRs of msl-2 mRNA is independent of the poly(a) tail, EMBO J. 18 (21) (1999) 6146–6154. https://pubmed.ncbi.nlm.nih.gov/10545124/. https://doi.org/10.1093/emboj/18.21.6146.

[32] F. Gebauer, M. Grskovic, M.W. Hentze, Drosophila sex-lethal inhibits the stable association of the 40S ribosomal subunit with msl-2 mRNA, Mol. Cell 11 (5) (2003) 1397–1404. https://pubmed.ncbi.nlm.nih.gov/12769862/. https://doi.org/10.1016/s1097-2765(03)00176-x.

[33] F. Gebauer, L. Merendino, M.W. Hentze, J. Valcárcel, The drosophila splicing regulator sex-lethal directly inhibits translation of male-specific-lethal 2 mRNA, RNA (New York, N.Y.) 4 (2) (1998) 142–150. https://pubmed.ncbi.nlm.nih.gov/9570314/.

[34] A. Graindorge, C. Carré, F. Gebauer, Sex-lethal promotes nuclear retention of msl2 mRNA via interactions with the STAR protein HOW, Genes Dev. 27 (12) (2013) 1421–1433. https://pubmed.ncbi.nlm.nih.gov/23788626/. https://doi.org/10.1101/gad.214999.113.

[35] A. Graindorge, C. Militti, F. Gebauer, Posttranscriptional control of X-chromosome dosage compensation. Wiley interdisciplinary reviews, RNA 2 (4) (2011) 534–545. https://pubmed.ncbi.nlm.nih.gov/21957042/. https://doi.org/10.1002/wrna.75.

[36] M. Grskovic, M.W. Hentze, F. Gebauer, A co-repressor assembly nucleated by sex-lethal in the 3'UTR mediates translational control of Drosophila msl-2 mRNA, EMBO J. 22 (20) (2003) 5571–5581. https://pubmed.ncbi.nlm.nih.gov/14532129/. https://doi.org/10.1093/emboj/cdg539.

[37] W. Gu, P. Szauter, J.C. Lucchesi, Targeting of MOF, a putative histone acetyl transferase, to the X chromosome of Drosophila melanogaster, Dev. Genet. 22 (1) (1998) 56–64. https://pubmed.ncbi.nlm.nih.gov/9499580/, 10.1002/(SICI)1520-6408(1998)22:1<56:AID-DVG6>3.0.CO;2-6.

[38] N. Handa, O. Nureki, K. Kurimoto, I. Kim, H. Sakamoto, Y. Shimura, et al., Structural basis for recognition of the tra mRNA precursor by the sex-lethal protein, Nature 398 (6728) (1999) 579–585. https://www.nature.com/articles/19242. https://doi.org/10.1038/19242.

[39] J. Hennig, C. Militti, G.M. Popowicz, I. Wang, M. Sonntag, A. Geerlof, et al., Structural basis for the assembly of the Sxl-Unr translation regulatory complex, Nature 515 (7526) (2014) 287–290. https://pubmed.ncbi.nlm.nih.gov/25209665/. https://doi.org/10.1038/nature13693.

[40] A. Hilfiker, D. Hilfiker-Kleiner, A. Pannuti, J.C. Lucchesi, Mof, a putative acetyl transferase gene related to the Tip60 and MOZ human genes and to the SAS genes of yeast, is required for dosage compensation in Drosophila, EMBO J. 16 (8) (1997) 2054–2060. https://pubmed.ncbi.nlm.nih.gov/9155031/. https://doi.org/10.1093/emboj/16.8.2054.

[41] N.M. Hollmann, P.K. Ankush Jagtap, P. Masiewicz, T. Guitart, B. Simon, J. Provaznik, et al., Pseudo-Rna binding domains mediate rna structure specificity in upstream of N-Ras, Cell Rep. 32 (3) (2020) 107930.

[42] N.M. Hollmann, P.K.A. Jagtap, J.-B. Linse, P. Ullmann, M. Payr, B. Murciano, et al., Upstream of N-Ras C-terminal cold shock domains mediate poly(a) specificity in a novel RNA recognition mode and bind poly(a) binding protein, Nucleic Acids Res. 51 (4) (2023) 1895–1913, https://doi.org/10.1093/nar/gkac1277.

[43] B.A. Johnson, Using NMR view to visualize and analyze the NMR spectra of macromolecules, Methods Mol. Biol. (Clifton, N.J.) 278 (2004) 313–352. https://pubmed.ncbi.nlm.nih.gov/15318002/, 10.1385/1–59259–809-9:313.

[44] I.S. Joung, T.E. Cheatham, Determination of alkali and halide monovalent ion parameters for use in explicitly solvated biomolecular simulations, J. Phys. Chem. B. 112 (30) (2008) 9020–9041, https://doi.org/10.1021/jp8001614.

[45] J. Jumper, R. Evans, A. Pritzel, T. Green, M. Figurnov, O. Ronneberger, et al., Highly accurate protein structure prediction with AlphaFold, Nature 596 (7873) (2021) 583–589. https://pubmed.ncbi.nlm.nih.gov/34265844/. https://doi.org/10.1038/s41586-021-03819-2.

[46] L.E. Kay, D.A. Torchia, A. Bax, Backbone dynamics of proteins as studied by 15N inverse detected heteronuclear NMR spectroscopy: application to staphylococcal nuclease, Biochemistry 28 (23) (1989) 8972–8979. https://pubmed.ncbi.nlm.nih.gov/2690953/. https://doi.org/10.1021/bi00449a003.

[47] R. Keller, The Computer Aided Resonance Assignment Tutorial, 1st ed., Cantina Verl, Goldau, 2004.

[48] R.L. Kelley, I. Solovyeva, L.M. Lyman, R. Richman, V. Solovyev, M.I. Kuroda, Expression of msl-2 causes assembly of dosage compensation regulators on the X chromosomes and female lethality in Drosophila, Cell 81 (6) (1995) 867–877, https://doi.org/10.1016/0092-8674(95)90007-1.

[49] R.L. Kelley, J. Wang, L. Bell, M.I. Kuroda, Sex lethal controls dosage compensation in Drosophila by a non-splicing mechanism, Nature 387 (6629) (1997) 195–199. https://pubmed.ncbi.nlm.nih.gov/9144292/. https://doi.org/10.1038/387195a0.

[50] I.R. Kleckner, M.P. Foster, An introduction to NMR-based approaches for measuring protein dynamics, Biochim. Biophys. Acta 1814 (8) (2011) 942–968. https://pubmed.ncbi.nlm.nih.gov/21059410/. https://doi.org/10.1016/j.bbapap.2010.10.012.

[51] M. Krepl, P. Pokorná, V. Mlýnský, P. Stadlbauer, J. Šponer, Spontaneous binding of single-stranded RNAs to RRM proteins visualized by unbiased atomistic simulations with a rescaled RNA force field, Nucleic Acids Res. 50 (21) (2022) 12480–12496. https://pubmed.ncbi.nlm.nih.gov/36454011/. https://doi.org/10.1093/nar/gkac1106.

[52] M. Krepl, J. Vögele, H. Kruse, E. Duchardt-Ferner, J. Wöhnert, J. Sponer, An intricate balance of hydrogen bonding, ion atmosphere and dynamics facilitates a seamless uracil to cytosine substitution in the U-turn of the neomycin-sensing riboswitch, Nucleic Acids Res. 46 (13) (2018) 6528–6543. https://academic.oup.com/nar/article/46/13/6528/5035167?login=true. https://doi.org/10.1093/nar/gky490.

[53] P.-H. Kuo, L.G. Doudeva, Y.-T. Wang, C.-K.J. Shen, H.S. Yuan, Structural insights into TDP-43 in nucleic-acid binding and domain interactions, Nucleic Acids Res. 37 (6) (2009) 1799–1808. https://pubmed.ncbi.nlm.nih.gov/19174564/. https://doi.org/10.1093/nar/gkp013.

[54] N.-A. Lakomek, J. Ying, A. Bax, Measurement of ^{15}N relaxation rates in perdeuterated proteins by TROSY-based methods, J. Biomol. NMR 53 (3) (2012) 209–221. https://pubmed.ncbi.nlm.nih.gov/22689066/. https://doi.org/10.1007/s10858-012-9626-5.

[55] W. Lee, M. Tonelli, J.L. Markley, NMRFAM-SPARKY: Enhanced software for biomolecular NMR spectroscopy, Bioinformatics (Oxford, England) 31 (8) (2015) 1325–1327. https://pubmed.ncbi.nlm.nih.gov/25505092/. https://doi.org/10.1093/bioinformatics/btu830.

[56] I. Loedige, L. Jakob, T. Treiber, D. Ray, M. Stotz, N. Treiber, et al., The crystal structure of the NHL domain in complex with RNA reveals the molecular basis of Drosophila brain-tumor-mediated gene regulation, Cell Rep. 13 (6) (2015) 1206–1220. https://www.sciencedirect.com/science/article/pii/S2211124715011109. https://doi.org/10.1016/j.celrep.2015.09.068.

[57] I. Loedige, M. Stotz, S. Qamar, K. Kramer, J. Hennig, T. Schubert, et al., The NHL domain of BRAT is an RNA-binding domain that directly contacts the hunchback mRNA for regulation, Genes Dev. 28 (7) (2014) 749–764. https://pubmed.ncbi.nlm.nih.gov/24696456/. https://doi.org/10.1101/gad.236513.113.

[58] C.D. Mackereth, T. Madl, S. Bonnal, B. Simon, K. Zanier, A. Gasch, et al., Multi-domain conformational selection underlies pre-mRNA splicing regulation by U2AF, Nature 475 (7356) (2011) 408–411. https://www.nature.com/articles/nature10171. https://doi.org/10.1038/nature10171.

[59] J. Macošek, B. Simon, J.-B. Linse, P.K.A. Jagtap, S.L. Winter, J. Foot, et al., Structure and dynamics of the quaternary hunchback mRNA translation repression complex, Nucleic Acids Res. 49 (15) (2021) 8866–8885. https://pubmed.ncbi.nlm.nih.gov/34329466/. https://doi.org/10.1093/nar/gkab635.

[60] J.A. Maier, C. Martinez, K. Kasavajhala, L. Wickstrom, K.E. Hauser, C. Simmerling, ff14SB: improving the accuracy of protein side chain and backbone parameters from ff99SB, J. Chem. Theory Comput. 11 (8) (2015) 3696–3713, https://doi.org/10.1021/acs.jctc.5b00255.

[61] J. Medenbach, M. Seiler, M.W. Hentze, Translational control via protein-regulated upstream open reading frames, Cell 145 (6) (2011) 902–913. https://pubmed.ncbi.nlm.nih.gov/21663794/. https://doi.org/10.1016/j.cell.2011.05.005.

[62] V.H. Meller, B.P. Rattner, The roX genes encode redundant male-specific lethal transcripts required for targeting of the MSL complex, EMBO J. 21 (5) (2002) 1084–1091. https://pubmed.ncbi.nlm.nih.gov/11867536/. https://doi.org/10.1093/emboj/21.5.1084.

[63] L. Merendino, S. Guth, D. Bilbao, C. Martínez, J. Valcárcel, Inhibition of msl-2 splicing by sex-lethal reveals interaction between U2AF35 and the 3′ splice site AG, Nature 402 (6763) (1999) 838–841. https://pubmed.ncbi.nlm.nih.gov/10617208/. https://doi.org/10.1038/45602.

[64] R. Moschall, M. Gaik, J. Medenbach, Promiscuity in post-transcriptional control of gene expression: Drosophila sex-lethal and its regulatory partnerships, FEBS Lett. 591 (11) (2017) 1471–1488. https://pubmed.ncbi.nlm.nih.gov/28391641/. https://doi.org/10.1002/1873-3468.12652.

[65] K. Nagai, C. Oubridge, T.H. Jessen, J. Li, P.R. Evans, Crystal structure of the RNA-binding domain of the U1 small nuclear ribonucleoprotein a, Nature 348 (6301) (1990) 515–520. https://www.nature.com/articles/348515a0. https://doi.org/10.1038/348515a0.

[66] M. Pabis, G.M. Popowicz, R. Stehle, D. Fernández-Ramos, S. Asami, L. Warner, et al., HuR biological function involves RRM3-mediated dimerization and RNA binding by all three RRMs, Nucleic Acids Res. 47 (2) (2019) 1011–1029, https://doi.org/10.1093/nar/gky1138.

[67] N. Ripin, J. Boudet, M.M. Duszczyk, A. Hinniger, M. Faller, M. Krepl, et al., Molecular basis for AU-rich element recognition and dimerization by the HuR C-terminal RRM, Proc. Natl. Acad. Sci. U. S. A. 116 (8) (2019) 2935–2944, https://doi.org/10.1073/pnas.1808696116.

[68] P. Rossi, G.V.T. Swapna, Y.J. Huang, J.M. Aramini, C. Anklin, K. Conover, et al., A microscale protein NMR sample screening pipeline, J. Biomol. NMR 46 (1) (2010) 11–22. https://link.springer.com/article/10.1007/s10858-009-9386-z. https://doi.org/10.1007/s10858-009-9386-z.

[69] M. Samata, A. Akhtar, Dosage compensation of the X chromosome: a complex epigenetic assignment involving chromatin regulators and long noncoding RNAs, Annu. Rev. Biochem. 87 (2018) 323–350. https://pubmed.ncbi.nlm.nih.gov/29668306/. https://doi.org/10.1146/annurev-biochem-062917-011816.

[70] M. Sattler, J. Schleucher, C. Griesinger, Heteronuclear multidimensional NMR experiments for the structure determination of proteins in solution employing pulsed field gradients, Prog. Nucl. Magn. Reson. Spectrosc. 34 (2) (1999) 93–158. https://www.semanticscholar.org/paper/Heteronuclear-multidimensional-NMR-experiments-for-Sattler-Schleucher/70934f38a889888eeee3b044aae19cd1b7fb0c45.

[71] I.B. Schäfer, M. Yamashita, J.M. Schuller, S. Schüssler, P. Reichelt, M. Strauss, et al., Molecular basis for poly(a) RNP architecture and recognition by the Pan2-Pan3 Deadenylase, Cell 177 (6) (2019) 1619–1631.e21. https://pubmed.ncbi.nlm.nih.gov/31104843/. https://doi.org/10.1016/j.cell.2019.04.013.

[72] Y. Shamoo, U. Krueger, L.M. Rice, K.R. Williams, T.A. Steitz, Crystal structure of the two RNA binding domains of human hnRNP A1 at 1.75 a resolution, Nat. Struct. Biol. 4 (3) (1997) 215–222. https://pubmed.ncbi.nlm.nih.gov/9164463/. https://doi.org/10.1038/nsb0397-215.

[73] Y. Shen, F. Delaglio, G. Cornilescu, A. Bax, TALOS+: a hybrid method for predicting protein backbone torsion angles from NMR chemical shifts, J. Biomol. NMR 44 (4) (2009) 213–223. https://link.springer.com/article/10.1007/s10858-009-9333-z. https://doi.org/10.1007/s10858-009-9333-z.

[74] C.W. Siebel, L.D. Fresco, D.C. Rio, The mechanism of somatic inhibition of Drosophila P-element pre-mRNA splicing: multiprotein complexes at an exon pseudo-5′ splice site control U1 snRNP binding, Genes Dev. 6 (8) (1992) 1386–1401. https://pubmed.ncbi.nlm.nih.gov/1322855/. https://doi.org/10.1101/gad.6.8.1386.

[75] C.W. Siebel, R. Kanaar, D.C. Rio, Regulation of tissue-specific P-element pre-mRNA splicing requires the RNA-binding protein PSI, Genes Dev. 8 (14) (1994) 1713–1725. https://pubmed.ncbi.nlm.nih.gov/7958851/. https://doi.org/10.1101/gad.8.14.1713.

[76] B. Simon, H. Köstler, Improving the sensitivity of FT-NMR spectroscopy by apodization weighted sampling, J. Biomol. NMR 73 (3–4) (2019) 155–165. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6525709/. https://doi.org/10.1007/s10858-019-00243-7.

[77] S.P. Skinner, R.H. Fogh, W. Boucher, T.J. Ragan, L.G. Mureddu, G.W. Vuister, CcpNmr AnalysisAssign: a flexible platform for integrated NMR analysis, J. Biomol. NMR 66 (2) (2016) 111–124. https://pubmed.ncbi.nlm.nih.gov/27663422/. https://doi.org/10.1007/s10858-016-0060-y.

[78] T. Straub, P.B. Becker, Dosage compensation: the beginning and end of generalization, Nat. Rev. Genet. 8 (1) (2007) 47–57, https://doi.org/10.1038/nrg2013.

[79] E. Szostak, M. García-Beyaert, T. Guitart, A. Graindorge, O. Coll, F. Gebauer, Hrp48 and eIF3d contribute to msl-2 mRNA translational repression, Nucleic Acids Res. 46 (8) (2018) 4099–4113. https://pubmed.ncbi.nlm.nih.gov/29635389/. https://doi.org/10.1093/nar/gky246.

[80] T.C. Terwilliger, R.W. Grosse-Kunstleve, P.V. Afonine, N.W. Moriarty, P.H. Zwart, L.W. Hung, et al., Iterative model building, structure refinement and density modification with the PHENIX AutoBuild wizard, Acta Crystallogr. D Biol. Crystallogr. 64 (Pt 1) (2008) 61–69, https://doi.org/10.1107/S090744490705024X.

[81] F. van den Ent, J. Löwe, RF cloning: a restriction-free method for inserting target genes into plasmids, J. Biochem. Biophys. Methods 67 (1) (2006) 67–74. https://pubmed.ncbi.nlm.nih.gov/16480772/. https://doi.org/10.1016/j.jbbm.2005.12.008.

[82] Q. Vicens, J.S. Kieft, O.S. Rissland, Revisiting the closed-loop model and the nature of mRNA 5′-3′ communication, Mol. Cell 72 (5) (2018) 805–812. https://pubmed.ncbi.nlm.nih.gov/30526871/. https://doi.org/10.1016/j.molcel.2018.10.047.

[83] W.F. Vranken, W. Boucher, T.J. Stevens, R.H. Fogh, A. Pajon, M. Llinas, et al., The CCPN data model for NMR spectroscopy: development of a software pipeline, Proteins: Struct., Funct., Bioinf. 59 (4) (2005) 687–696. https://onlinelibrary.wiley.com/doi/full/10.1002/prot.20449. https://doi.org/10.1002/prot.20449.

[84] I. Wang, J. Hennig, P.K.A. Jagtap, M. Sonntag, J. Valcárcel, M. Sattler, Structure, dynamics and RNA binding of the multi-domain splicing factor TIA-1, Nucleic Acids Res. 42 (9) (2014) 5949–5966, https://doi.org/10.1093/nar/gku193.

[85] C.A. Weidmann, C. Qiu, R.M. Arvola, T.-F. Lou, J. Killingsworth, Z.T. Campbell, et al., Drosophila Nanos acts as a molecular clamp that modulates the RNA-binding and repression activities of Pumilio, eLife (2016) 5, https://doi.org/10.7554/eLife.17096.

[86] M.P. Williamson, Using chemical shift perturbation to characterise ligand binding, Prog. Nucl. Magn. Reson. Spectrosc. 73 (2013) 1–16. https://pubmed.ncbi.nlm.nih.gov/23962882/. https://doi.org/10.1016/j.pnmrs.2013.02.001.

[87] M. Zgarbová, M. Otyepka, J. Sponer, A. Mládek, P. Banáš, T.E. Cheatham, et al., Refinement of the Cornell et al. nucleic acids force field based on reference quantum chemical calculations of Glycosidic torsion profiles, J. Chem. Theory Comput. 7 (9) (2011) 2886–2902. https://pubmed.ncbi.nlm.nih.gov/21921995/. https://doi.org/10.1021/ct200162x.