Staufen2 mediated RNA recognition and localization requires combinatorial action of multiple domains

Simone Heber [1,2], Imre Gaspar[3], Jan-Niklas Tants[4], Johannes Günther[4], Sandra M. Fernandez Moya[5], Robert Janowski[2], Anne Ephrussi[3], Michael Sattler[2,4], Dierk Niessing[1,2]

1) Institute of Pharmaceutical Biotechnology, Ulm University
2) Institute of Structural Biology, Helmholtz Zentrum München
3) Developmental Biology Unit, European Molecular Biology Laboratory
4) Center for Integrated Protein Science Munich at Chair of Biomolecular NMR Spectroscopy, Department Chemistry, Technische Universität München
5) Biomedical Center Munich, Department of Cell Biology, LMU München

Summary

Proteins of the Staufen (Stau) family are core factors of mRNA localization particles in a number of metazoan organisms and cell types, including *Drosophila* oocytes and mouse neurons. They consist of three to four double-stranded RNA binding domains (dsRBDs) and a C-terminal dsRBD-like domain. Mouse Staufen2 (mStau2) like *Drosophila* Stau (dmStau) contains four dsRBDs. Existing data suggest that dsRBDs 3-4 are necessary and sufficient for binding to Stau-target mRNAs. In contrast, dsRBDs 1-2 have been suggested to be non-canonical dsRBDs, fulfilling other functions such as protein-protein interaction. Here, we show that dsRBDs 1 and 2 of mStau2 bind RNA with similar affinities and kinetics as dsRBDs 3 and 4. While RNA binding by these tandem domains is transient and characterized by high on- and off-rates, an mStau2 fragment with all four dsRBDs recognizes its target RNAs with high stability. Rescue experiments with dsRBD1-2 mutant versions of mStau2 in otherwise *staufen*-deficient *Drosophila* oocytes confirmed the physiological relevance of these findings. Full-length mStau2 partially rescued Stau-dependent mRNA localization, whereas a rescue with dsRBD1-2 mutant mStau2 failed. In summary, our data show that the dsRBDs 1-2 play essential roles in the mRNA recognition and function of Stau-family proteins of different species. The combinatorial binding of two tandem-dsRBDs allows for the cooperative recognition of more complex secondary structures and thus contributes to our understanding of how Stau proteins achieve selectivity in RNA binding.

Introduction

mRNA localization is an essential mechanism for a range of cellular processes, including embryonic development, cell differentiation and migration, as well as neuronal plasticity[1]. For active transport of mRNAs along the cellular cytoskeleton, ribonucleoprotein particles (RNPs) are formed. Such mRNA-containing RNPs (mRNPs) contain motor proteins, RNA binding proteins, helicases, and translational regulators[2].

In the mature nervous system, neuronal mRNA localization to pre- and postsynaptic areas followed by local translation has been implicated in memory and learning[3, 4]. For instance, dendritically localized RNAs produce proteins with synaptic functions such as Ca^{2+}/calmodulin kinase II (CaMKII), the cytoskeletal protein Arc or microtubule-associated protein 2 (MAP2), and AMPA or NMDA receptors.

The RNA binding protein Staufen (Stau) was originally identified in *Drosophila* as an mRNA transport factor required to establish the anterior-posterior axis of the embryo[5, 6]. Together with proteins of the exon-junction complex (EJC) and the translational repressor Bruno it binds to the *oskar* mRNA, which is transported from the nurse cells to the oocyte and then localized to its posterior pole[7]. During neurogenesis, the asymmetric segregation of *prospero* mRNA into the ganglion mother cell requires Stau function as well[8].

In mice, the two Staufen-paralogs Stau1 and 2 share about 50 % protein-sequence identity and have both been implicated in mRNA localization and RNA-dependent control of gene expression[9, 10, 11]. Whereas Stau1 is ubiquitously expressed and required for Staufen-mediated decay (SMD) of its target mRNAs via UPF1 interaction, Stau2 expression is enriched in heart and brain[12, 13, 14, 15, 16]. The two paralogs Stau1 and Stau2 were reported to bind distinct, yet overlapping, sets of target mRNAs[10, 17], indicating distinct but possibly complementary functions. Consistent with this hypothesis is the observation that although both paralogs appear to mediate degradation of RNAs, only Stau2 seems to also stabilize a subset its target mRNAs[18].

A transcriptome-wide analysis of *Drosophila* Stau (dmStau) targets suggested certain RNA-secondary structure elements as Stau-recognized structures (SRSs)[19]. A subsequent study in mice used immunoprecipitation- and microarray-based experiments to identify the *Regulator of G-Protein*

Signaling 4 (*Rgs4*) mRNA as an mStau2 regulated transcript and found two predicted SRS stem-loops in its 3'UTR[18].

All Stau-family proteins contain multiple so-called double-stranded RNA binding domains (dsRBD). Whereas mStau1 contains three dsRBDs, a tubulin-binding domain (TBD) and one C-terminal non-canonical dsRBD-like domain, mStau2 and dmStau have four dsRBDs, followed by a tubulin-binding domain (TBD) and a C-terminal, non-canonical dsRBD-like domain[20].

For all mammalian Stau proteins, the dsRBD3 and dsRBD4 are thought to be required and sufficient for full target mRNA binding[11, 12, 21], whereas dsRBDs 1, 2 and 5 are often referred to as pseudo-RBDs, which retained the fold but not activity of canonical dsRBDs[21]. The longest isoform of Stau2, Stau2[62], is most similar to dmStau, both possessing all five dsRBDs. Stau2[62] shuttles between nucleus and cytoplasm and has been proposed to transport RNAs from the nucleus to distal dendrites[22]. Because Stau dsRBDs only seem to interact with the backbone of RNA[23] and do not undergo recognizable sequence-specific interactions[24], one of the unresolved questions is how specific RNA binding can be achieved by dsRBD 3-4.

Here we show that in mStau2 the non-canonical dsRBDs 1 and 2 exhibit RNA binding activity of equal affinity and kinetic properties as the known RNA binding dsRBDs 3-4. Mutational analyses and biophysical characterization of RNA binding revealed that dsRBD1-2 have to act in concert with dsRBD3-4 to allow for stable, high-affinity RNA binding. Using *Drosophila* as model system, we demonstrate the importance of RNA binding by dsRBDs 1-2 for Stau function *in vivo* and show that mStau2 partially can substitute for dmStau function during early *Drosophila* development. The requirement of a combination of two dsRBD-tandem domains and thus binding to two stem-loops allows recognition of combinations of secondary structure and thus a much more complex readout for specific binding. This observation might help to explain how Stau proteins can bind selectively to their RNA targets *in vivo*.

Materials & Methods

Molecular cloning

DNA sequences of interest were amplified by polymerase chain reaction (PCR) from template plasmids. Cloning was performed with the In-Fusion HD Cloning kit (Clontech) according to the manufacturer's protocol. Point mutations or deletions were introduced by 3-point PCR with overlap-extension[25] or with the QuikChange II Site-directed mutagenesis kit (Agilent technologies), according to the manufacturer's instructions.

For *in vivo* experiments in *D. melanogaster*, rsEGFP2 was fused with mStau2 or with dmStau, via cloning into pBlueScript-KS. First, the primers pBSKS-rsEGFP2 FW and rsEGFP+3C RV were used to create an rsEGFP2 sequence with a pBlueScript-KS 5' overhang for In-Fusion cloning and a PreScission protease cleavage site as 3' overhang, and the primers 3C+Stau2 FW and pBSKS-mStau2 RV or 3C+dmStau FW and pBSKS-dmStaufen RV to create a mStau2 or dmStau sequence, respectively, with a PreScission cleavage site 5' overhang and a pBlueScript-KS 3' overhang for In-Fusion cloning. The PCR products from these reactions served as templates for a third PCR with the primers pBSKS-rsEGFP2 FW and pBSKS-mStau2 RV or pBSKS-dmStau RV to create rsEGFP2-mStau2/dmStau sequences with 5' and 3' overhangs for InFusion (Clontech) cloning into BamHI/XbaI-linearized pBlueScript-KS, according to the manufacturer's protocol. The resulting plasmids served as templates for PCR with primers pUASp-rsEGFP2 FW and pUASp-mStau2 RV or pUASp-dmStau RV to amplify rsEGFP2-mStau2/dmStau sequences with 5' and 3' overhangs for InFusion cloning into the BamHI/XbaI-linearized pUASp-attB plasmid.

Protein expression

Full-length mStau2 protein

mStau2 FL was expressed as a HisSUMO-tagged fusion protein in High Five insect cells. After cloning in pFastBacDual, recombinant baculovirus was produced with the Bac-to-Bac Expression System (Invitrogen) in Sf21 insect cells as described by the manufacturer's protocol.

Truncated mStau2 dsRBD protein constructs
mStau2 proteins were expressed after cloning into the expression vector pOPINS3C as fusion proteins with HisSUMO-tag in *E. coli* Rosetta cells using autoinduction ZY-medium[26].

Isotope-labeled proteins for NMR
Uniformly ^{15}N- or ^{15}N,^{13}C-labeled proteins for NMR experiments were expressed in ^{15}N-M9 minimal medium (1x ^{15}N-labeled M9 salt solution, 0.2 % (^{13}C-) glucose, 1 mM $MgSO_4$, 0.3 mM $CaCl_2$, 1 µg/L biotin, 1 µg/L thiamine, 1x trace metals) supplemented with antibiotics. 100 ml precultures were grown o/n at 37 °C, 150 rpm and used to inoculate 1 L pre-warmed M9 minimal medium. Cultures were grown at 37 °C, 150 rpm to OD_{600nm} = 0.6. Protein expression was induced with 0.25 mM IPTG and cultures were cooled for protein expression o/n at 18 °C.

Protein purification

Full-length mStau2 protein
High Five cell pellets containing His-SUMO-tagged mStau2 FL were lysed by sonication in lysis buffer (1x PBS, 880 mM NaCl, 400 mM arginine, 2 mM DTT, 10 mM imidazole). The lysate was cleared by centrifugation and the soluble protein fraction was purified by Ni-IMAC on HisTrap FF (GE). Bound protein was eluted with 200 mM imidazole after extensive washing with 15 CV lysis buffer. The protein was dialyzed in low salt buffer (40 mM Bis-Tris pH 7, 150 mM NaCl, 50 mM arginine, 2 mM DTT) o/n before further purification on a HiTrap Heparin HP column (GE) and size exclusion chromatography (SEC) on Superdex200 Increase (GE).

Truncated mStau2 dsRBD protein constructs
E. coli Rosetta cell pellets containing His-SUMO-tagged fusion proteins were lysed by sonication in lysis buffer (1x PBS, 880 mM NaCl, 400 mM arginine, 2 mM DTT, 10 mM imidazole). The lysate was cleared by centrifugation and the soluble protein fraction was purified by Ni-IMAC on HisTrap FF (GE). Bound protein was eluted with 200 mM imidazole after extensive washing with 15 CV lysis buffer. For fusion-tag removal, the protein was digested o/n with PreScission protease upon dialysis in low salt buffer (40 mM Bis-Tris pH 7, 150 mM NaCl, 50 mM arginine, 2 mM DTT). The protein was purified with a

second, subtractive Ni-IMAC, affinity chromatography on a HiTrap Heparin HP column (GE) and size exclusion chromatography (SEC) on Superdex 75 (GE). Size exclusion chromatography was performed in minimal buffer (40 mM Bis-Tris pH 7, 150 mM NaCl, 2 mM DTT) or the buffer required for downstream applications.

Small scale RNA *in vitro* transcriptions

RNAs for EMSAs were produced by small scale *in vitro* transcriptions with the MegaShortScript Transcription kit (Ambion) according to the manufacturer's protocol. HPLC-purified primers (Eurofins) were used as templates. In order to produce partially double-stranded template DNA, FW (T7prom) and RV primers were annealed after unfolding at 60 °C for 5 min by slow cooling to RT.

Large scale RNA *in vitro* transcriptions

SRS2 and SRS2+5Δloop RNA were purchased from IBA (Göttingen). Other RNAs needed in large amounts for NMR experiments were produced by large scale *in vitro* transcription. As template, 4 µM HPLC-purified FW (T7prom) primer and 3.4 µM HPLC-purified RV primer were annealed after unfolding at 60 °C for 5 min by slow cooling to RT in 34 mM $MgCl_2$ in a total volume of 594 µL. This DNA template mixture was used for a 5 mL *in vitro* transcription reaction containing, in addition to the template, 4 mM of each NTP, a template-specifically optimized concentration of $MgCl_2$ (see below), 80 mg/mL PEG8000 and 0.5 mg/mL T7 RNA polymerase in 1x TRX buffer (40 mM Tris/HCl pH_{RT} 8.0, 1 mM spermidine, 0.1 ‰ Triton X-100, 5 mM DTT). The reaction was incubated for 2 h at 37°C. The reaction was stopped by removal of precipitants by centrifugation at 48,384 g in 5 min and subsequent RNA precipitation with 0.1 V 3 M NaOAc and 3 V absolute ethanol at -20 °C o/n.

The optimal $MgCl_2$ concentration for each RNA was determined beforehand by $MgCl_2$ screening in 50 µL reactions containing 4 – 60 mM $MgCl_2$. Quality and quantity of RNA in each $MgCl_2$ concentration were examined by 8 % 1x-urea-PAGE.

PAGE purification of RNA

RNA was pelleted by centrifugation at 48,384 g, 4 °C, in 30 min, air-dried and subsequently dissolved in 1x denaturing RNA loading dye. The RNA was

purified by 8 % 1x TBE- 8 M urea PAGE in an Owl sequencing chamber (Thermo Fisher Scientific) in 1x TBE running buffer at constant 300 V for 17-20 h. RNA bands were visualized by UV shadowing and the desired band was excised from the gel and extracted by electroelution in a Whatman Elutrap electroelution system (GE Healthcare) at constant 200 V in 1x TBE for 8 h. Eluted RNA was collected each hour. Eluted RNA was dialyzed against 5 M NaCl at 4 °C o/n and subsequently twice against RNase-free water at 4 °C o/n before drying in a Concentrator Plus SpeedVac (Eppendorf).

Radioactive labeling of RNA

RNAs for electrophoretic mobility shift assays (EMSA) were labeled radioactively for sensitive detection. *In vitro* transcribed RNA was 5' dephosphorylated in 20 µL reactions containing 10 pmol RNA, 1x Tango Buffer with BSA (Thermo Fisher), 2 U FastAP thermosensitive alkaline phosphatase (Thermo Fisher) and 20 U of the RNase inhibitor SUPERaseIn (Thermo Fisher). After incubation at 37 °C for 30 min, the dephosphorylated RNA was phenol/chloroform extracted and precipitated with 0.1 V 3 M NaOAc, 3 V absolute ethanol and subsequent chilling at -20 °C for ≥ 15 min.

For radioactive labeling, 10 pmol dephosphorylated RNA or chemically synthesized RNA were 5'-phosphorylated with ^{32}P from γ-^{32}P ATP (Hartmann Analytic) in a 20 µL reaction with T4 polynucleotide kinase (New England Biolabs) in 1x buffer A. The labeling reaction was incubated at 37 °C for 30 min and subsequently stopped at 72 °C for 10 min.

Remaining free nucleotides were removed by purification on a NucAway™ Spin column (Ambion) according to the manufacturer's instructions. Eluted radiolabeled RNA was diluted to a final concentration of 100 nM in RNase-free H_2O and stored at -20 °C.

RNA for annealing assays was labeled accordingly but purified by urea-PAGE. RNA was resolved on an 8 % urea-Midi Gel, detected with radiographic films, excised from the gel and eluted in 300 µL Crush'n'Soak buffer o/n. Subsequently, RNA was precipitated with 3 V 100 % EtOH, redissolved in 100 µL RNase-free H_2O to a final concentration of 100 nM and stored at -20 °C.

Electrophoretic mobility shift assay (EMSA)

For EMSAs with short RNAs (<100 nt), protein at the indicated final concentration was mixed with 5 nM radiolabeled RNA in RNase-free protein buffer supplemented with 4 % glycerol and 30 µg/mL yeast tRNA as competitor in a final volume of 20 µL. In order to allow protein-RNA complexes to form, the mixtures were incubated for > 20 min at RT.

Separation of protein-RNA complexes was performed by native PAGE on 6 % polyacrylamide 1x TBE gels in 40 min at constant 110 V in 1x TBE running buffer. Subsequently, the gels were fixed in 30 % (v/v) methanol, 10 % (v/v) acetic acid for 10 min before drying in a vacuum gel drier (BioRad). Visualization of radioactivity occurred after exposure of radiograph films (Kodak) in a Protec Optimax developer (Hohmann) or by PhosphorImaging with a Fujifilm FLA-3000.

In the case of long, unlabeled RNA (>100 nt), 10-100 nM RNA were used and separation of protein-RNA complexes was performed by 1-1.5 % agarose gel electrophoresis. Visualization of RNA was achieved by GelRed (Biotium) staining. Fluorescence was visualized with a Fusion SL imaging system (Vilber Lourmat) by UV at 254 nm.

Biotinylation of RNA

Ligand RNA for binding studies by Surface Plasmon Resonance was biotinylated to allow immobilization on a streptavidin-coated surface. For biotinylation, the Pierce™ RNA 3' End Biotinylation Kit (Thermo Fisher Scientific) was used according to the manufacturer's protocol. 50 pmol RNA were used per 30 µL reaction. After extraction and precipitation, the RNA was redissolved in 100 µL RNase-free water.

Surface Plasmon Resonance (SPR)

Surface Plasmon Resonance (SPR) experiments were performed with a BIACORE 3000 system (GE Healthcare). To assess protein-RNA interactions, biotinylated RNA in a volume of 60 µL was streptavidin-captured on a SA-Chip (GE Healthcare) surface at a flow rate of 10 µL/min after three consecutive 1 min conditioning injections of 50 mM NaOH, 1 M NaCl.

Full-length mStau interacted strongly with the blank SA-Chip surface. Thus, to assess the interaction of RNA with mStau FL, the protein was diluted in HBS-

EP buffer and covalently amine-coupled to a CM5-Chip (GE Healthcare) according to the manufacturer's instructions.

Kinetic analysis of protein-RNA interactions was performed at a flow rate of 30 µL/min in HBS-EP buffer (10 mM HEPES pH 7.4, 150 mM NaCl, 3 mM EDTA, 0.005 % surfactant P20). Analyte protein or, in case of mStau2 FL, RNA in HBS-EP buffer at the indicated concentrations was injected for 4 to 5 min to allow for association, subsequent dissociation was allowed for 10 to 15 min in HBS-EP buffer. To remove any residual bound protein, two 1 min regeneration injections with 1 M NaCl were performed.

Data were analyzed in the BIAevaluation software (GE Healthcare). After double-referencing of obtained binding curves against the signal in a ligand-free reference channel or, in the case of full-length mStau2, against a HisSUMO-coupled reference channel and a buffer run, average values for the analyte response at equilibrium were calculated. Steady state binding curves were obtained by plotting the response at equilibrium against analyte concentration and curve fitting with the Origin Pro 9.0 (OriginLab) software using the two-site binding or Hill1 fits available in the software. All experiments were performed at least in triplicate and the results of the $n \geq 3$ experiments averaged. Kinetic fits, if applicable, were performed with the BIAevaluation software (GE Healthcare) using the bivalent analyte model available in the software.

Circular Dichroism (CD) spectroscopy

CD spectra were collected with a Jasco J-715 spectropolarimeter from 260 nm to 190 nm in continuous scanning mode with a scanning speed of 50 nm/min at a bandwidth of 1 nm. 5 scans were collected per measurement and response time was 8 s. The measurements were performed at 20 °C. Spectra were analyzed with SpectraManager (Jasco).

Structure determination of SRS2 RNA

Crystallization experiments were performed at the X-ray Crystallography Platform at Helmholtz Zentrum München. Initial crystals of *Rgs4* SRS2 were obtained in 80 mM NaCl or 80 mM KCl, 20 mM $BaCl_2$, 40 mM Na cacodylate pH 6.5, 40 % MPD, 12 mM spermine after 4 days at RT using the sitting-drop vapor-diffusion method. Crystals could be reproduced in at 34-46 % MPD.

Crystals were harvested with cryogenic loops and flash-frozen in liquid nitrogen. Native datasets were collected at the Swiss Light Source (SLS) synchrotron, beamline PXIII. Anomalous data for phasing were collected at the European Synchrotron radiation Facility (ESRF) at beam line ID 23-2. Data were indexed and integrated using XDS and scaled via XSCALE. Structure factor amplitudes were obtained with Truncate (CCP4 package)[27]. The structure was solved by MAD phasing with Barium from the mother liquor, using the AutoRickshaw web server[28]. The structure was completed by iterative manual building in COOT and refinement with RefMac5 (CCP4 package)[27]. All crystallographic software was used from the SBGRID software bundle5. Images of the crystal structure were prepared with PyMol (Version 1.7; Schrodinger; http://www.pymol.org/).

Nuclear Magnetic Resonance (NMR)

RNA and RNA-protein complexes were dialyzed to 150 mM NaCl, 20 mM sodium phosphate buffer pH 7.0, 5 mM DTT prior to analysis and 5-10 % D2O was added for locking. Measurements were performed at 298 K on Bruker AVIII 600, AVIII 800 or AVIII 900 NMR spectrometers with cryogenic (TCI) triple resonance gradient probes. Data were processed with Topspin 3.0 or Topspin 3.5 and analyzed with Sparky 3[29] and CcpNmr Analysis[30]. RNA assignment of imino groups was based on ^{1}H,^{1}H-NOESY spectra; an initial protein backbone assignment was made with HNCACB spectra. Titration experiments with the single Stau2 dsRBDs 1 and 2 as well as the tandem domain dsRBD1-2 were performed at 50 µM protein concentration. After snap-cooling, the RNA ligand was added in molar ratios of 0.5, 1.0, 2.0 and 3.0 to the protein.

Fly strains and transgenesis

Heterozygous combination of the stau[D3][31] (FBal0016165) and stau[R9][5] (FBal0032815) alleles was used generate females lacking endogenous *dmstaufen*. The ◻Tub67C::GFP^{m6}-Staufen[32] (FBal0091177) transgene was used as source of GFP-dmStau in the female germ-line. Expression of the UASp-GFP-mStau2 transgenes was driven with one copy of oskar-Gal4[33] (FBtp0083699) in the female germ-line. *w^{1118}* (FBal0018186) was used as the wild-type control. All UASp-GFP-mStau2 transgenes were generated by

subcloning wild-type or mutant mStau2 coding sequences 3' to the cds of rsEGFP2 into the pUASp-attB trangenesis vector[34]. The UASp-rsEGFP2-mStau2 vectors were injected into embryos of y[1] M{vas-int.Dm}ZH-2A w[*]; PBac{y[+]-attP-3B}VK00033 (FBti0076453) females to facilitate psiC31 mediated insertion into the same locus on the 3L chromosome arm. All stocks were raised on normal cornmeal agar at 25 °C.

Single molecule fluorescent hybridization (smFISH)

42 and 24 different ssDNA oligonucleotides were labelled enzymatically with Atto532-ddUTP and Atto633-ddUTP, respectively, as described in references 35 and 36, to generate osk42x532[35] and bcd24x633 (Supplementary Table) probe-sets for smFISH. *Drosophila* ovaries were dissected and processed for smFISH analysis as described in references 35, 36.

Microscopy and image analysis

Drosophila egg-chambers mounted onto slides in Vectashield were imaged on a Leica TCS SP8 confocal laser scanning microscope using a 20x dry (NA=0.75) objective for imaging the RNA distribution and a 63x oil immersion (NA=1.4) objective to obtain high resolution images for co-localization analysis of *oskar* or *bicoid* mRNA and GFP-Stau. Analyses of RNA distribution within stage 9 and stage 10-11 oocytes were carried out as described in reference 37. Briefly, the outlines of the oocytes and the anteroposterior (AP) axis were manually specified and the smFISH signal was redistributed into a 100x100 matrix. Each column of this matrix represents the relative amount of signal found under 1% of the AP axis length with anterior on the left (column 1) and posterior on the right (column 100). Such matrices are then averaged to obtain a mean and the variability of the RNA localization during a certain stage of oogenesis. Moreover, descriptors such as the center of mass (relative to the geometric center of the ooctye) and the amount of RNA localizing to the posterior domain (defined on the minimum two-fold enrichment of the signal over what is expected) were extracted and compared statistically using a Kruskal-Wallis test followed by pair-wise Mann Whitney U test against the *stau* null condition.

Co-localization between the mRNAs and GFP-Stau was assayed as described in reference 38. Briefly, images deconvoluted using Huygens

Essential were segmented, and nearest neighbor pairs between *oskar* or *bicoid* mRNPs and GFP-Stau particles were established. To determine the number of mRNA molecules in an mRNP and to normalize GFP-Stau signal intensity, we fitted multiple Gaussian functions to the corresponding signal intensity distributions taken from the nurse cells (as in references 39 and 35 using the mixtools package in R (https://cran.r-project.org/web/packages/mixtools/index.html). The μ value of Gaussian fit that described the largest portion of the distribution (for *oskar* mRNPs ~ 60%, for *bicoid* mRNPs ~80%, GFP-Stau >85%) was taken as the signal intensity of a unit (for mRNPs the intensity of a signal mRNA molecule). The chosen μ value was always the smallest among the μ values of the fitted Gaussians. Raw signal intensities were normalized with the determined unit values. RNPs were clustered based on this normalized intensity under the following rule: $[2^i:2^{i+1}]$, i ◘ [0:8], i.e. 1, 2:3, 4:7, 8:15 etc. The observed nearest neighbor co-localization frequencies were computed for each of the clusters and were compared to the expected co-localization frequencies (governed by the object-densities, determined in randomized object localizations[38]. Linear correlation between RNA content and GFP-Stau intensity was established (R^2>0.9 in all cases, except between [F40A F157A] and bicoid, R^2=0.61). The slopes of the fitted lines were compared pair-wise using least-squares means analysis[40]. All statistical analyses were carried out in R[41] using RStudio (www.rstudio.com). All graphs were plotted by the ggplot2 library in R[42].

Results

mStau2 binds directly to SRS motifs in the *Rgs4* 3'UTR

To probe a potential direct interaction between mStau2 and the *Rgs4* mRNA, we performed *in vitro* binding experiments with mStau2 and the two previously predicted SRS motifs of the *Rgs4* 3'UTR (Figure S1 A). EMSAs with full-length mStau2 showed binding with apparent equilibrium dissociation constants (K_d) in the low µM range for *Rgs4* SRS1 as well as for SRS2 (Figure 1 A). The entire 3'UTR of *Rgs4* mRNA was bound by full-length Stau2 with higher affinity (Figure S1 B). Deletion of SRS1 and SRS2 did not reduce RNA binding, indicating that regions other than the SRS motifs contribute to Stau2 binding. Since EMSAs with very long RNAs do not yield very precise results, we also performed experiments with a 3'UTR fragment consisting of 634 bases of the *Rgs4* 3'UTR (Rgs4-mini) that contains both predicted SRSs. EMSAs with the previously reported RNA binding dsRBDs 3-4 using either wild-type *Rgs4*-mini RNA or a mutant version, in which SRS1 and SRS2 were deleted, showed similar affinities (Figure S1 C).

These observations indicated that other cryptic SRSs might be present in the *Rgs4* 3'UTR. In the *Rgs4*-mini RNA another region was predicted to fold into a stable imperfect stem-loop with 26 paired bases interrupted by two bulges. Although this stem-loop is longer and predicted to be more stable, it could still serve as cryptic SRS. Surprisingly mStau2 bound to this stem loop with a K_D in the hundred nanomolar range (Figure 1 A). This stem loop is termed SRS* (Figure S1 A).

The Stau2 tandem domain dsRBD1-2 binds dsRNA

Next, we tested whether binding is indeed only mediated by dsRBD3-4, as indicated by previous studies[12, 21]. Surprisingly, in EMSA experiments mStau2 dsRBD1-2 bound the SRS2 RNA with an affinity comparable to that of dsRBD3-4 (Figure 1 B). This finding shows that dsRBDs 1 and 2 are not inactive pseudo-RBDs as previously suggested but contribute to RNA recognition of Stau2. Binding of mStau2 dsRBD1-2 to SRS2 RNA was further confirmed by NMR titration experiments (Figure S2). Upon addition of the stem-loop RNA to the tandem domain dsRBD1-2, chemical shift changes and differential line-broadening of NMR signals in the protein and the RNA are

observed. To investigate the binding interface of mStau2 dsRBD1-2 on SRS2 RNA, imino signals of the unbound RNA were compared to the respective resonances when bound to mStau2 at equimolar ratio. Significant line broadening was observed where imino signals of four base pairs close to the stem terminus (U4, U27, U29, G30) were most strongly affected in NMR spectrum of the complex (Figure 1 C), while other imino signals were less affected. The differential line-broadening indicates binding kinetics in the intermediate exchange regime at the chemical shift time scale[43, 44].

To unambiguously confirm the dsRNA fold of *Rgs4* SRS2 we solved its crystal structure at 1.73 Å resolution (Figure 1 D). The RNA adopts a typical A-form double-stranded helix, characterized by a wide and shallow minor groove and a deep and narrow major groove. Whereas electron density in the stem region of both molecules is very well defined, the density map in the loop region is poor (Figure S3), indicating flexibility of the RNA in this region. Imino signals of base pairs are observed in imino ^{1}H,^{1}H-NOESY spectra and are consistent with the base pairing observed in the crystal structure (Figure S4). Our NMR data and the crystal structure thus confirm that SRS2 folds into a canonical stem-loop structure.

Role of the length of dsRNA for Stau2 tandem domain binding

To test the effect of stem-loop length on RNA binding, the SRS2 stem was extended by five base-pairs (SRS2+5, see Figure S1 A). EMSAs with dsRBD1-2 or with dsRBD3-4 showed significantly improved binding (Figure 1 E), indicating that the length of the stem has great influence on the affinity. In NMR titration experiments with dsRBD1-2, amino acids affected upon SRS2+5 RNA binding seem identical to SRS2 binding (Figure 1 F, compare to Figure S2). 1D imino traces of the RNA, however, revealed line broadening at substoichiometric concentrations for SRS2+5 dsRBD1-2 binding (Figure 1 F, compare with Figure 1 C). Strong line broadening of imino signals can point to dynamic binding comprising sliding, as previously shown for other dsRBDs[45, 46, 47].

To determine whether the loop region of the RNA is required for stem-loop recognition by mStau2 tandem domains, we tested the SRS2-stem elongated

by 5 base pairs but lacking its loop (SRS2+5Δloop). This elongated stem was bound by both tandem domains dsRBD3-4 and dsRBD1-2 with an affinity similar to the original SRS2 stem-loop (Figure S5 A, compare with Figure 1 B and E), indicating that the loop region is not essential for RNA recognition.

After showing that the affinity of mStau2 to RNA correlates with the length of dsRNA, we aimed to define the minimal length of the RNA stem required for recognition by mStau2 tandem domains. Both tandem domains dsRBD1-2 and dsRBD3-4 bound to RNA stem-loops comprising stems of 11 bp, 9 bp and 7 bp with similar affinities, with apparent dissociation constants (K_Ds) in the micromolar concentration range (Figure S5 B-D). Only when the stem was decreased to 6 bp, was binding almost abolished (Figure S5 E). Thus, a stem of 7 bp appears to be the minimal length required for recognition by Stau2 tandem domains.

Kinetics of Stau2 RNA binding

In order to understand the kinetics of mStau2 binding, surface plasmon resonance (SPR) experiments with biotin-labeled SRS* RNA or SRS2+5 RNA coupled to a streptavidin sensor chip surface were performed. For the tandem domains dsRBD1-2 and dsRBD3-4, rapid binding and dissociation kinetics were observed for both RNAs already at the lowest tested concentration of 10 nM (Figure 2 A and B). Because of the fast kinetics, the on- and off-rates could not be properly quantified. However, the steady-state binding is best described by a two-site binding fit for dsRBD1-2 with a K_D1 of 130 nM for the SRS2+5 RNA and of 25 nM for the SRS* RNA (Figure 2 A,Tables 1 and 2). K_D2 could not be determined because binding was not saturated at the highest measured concentration of 1 µM. Because mStau2 tends to oligomerize at low micro-molecular concentrations, even higher concentration ranges could not be tested.

Also, dsRBD3-4 bound with similar properties, yielding a K_D1 of 18 nM for SRS2+5 RNA and of 9 nM for the SRS* RNA (Figure 2 B; Tables 1 and 2). As with dsRBD1-2, K_D2 were in the µM range and could not be determined. Together these findings confirm the RNA binding activities of dsRBD1-2 and of dsRBD3-4, with similar binding properties. The observed fast kinetics for

K_D1 explain why in EMSAs no high-affinity band shifts were observed (K_D1) and only binding in the µM range (K_D2) was seen.

Interestingly, when the SPR experiments were repeated with Stau2 dsRBD1-4, the binding kinetics changed dramatically with both RNAs. Binding as well as dissociation occurred at much slower rates, indicating that the formed complexes are stable (Figure 2 C; Tables 1 and 2). Steady-state affinities could no longer be described by a two-site binding fit, most likely due to higher order binding events by Stau2's four dsRBDs. However, using a Hill-fit yielded K_Ds of 357 nM and 330 nM for the SRS2+5 RNA and the SRS*, respectively. Together with observed Hill coefficients of $n \geq 1.7$, these data indicate cooperative binding by all four RBDs, which results in the formation of stable mStau2-RNA complexes.

Finally, we assessed binding of the SRS2+5 and SRS* RNAs to full-length mStau2. Consistent with our RNA binding experiments with dsRBD1-4 (Figure 2C), in both cases stable complexes were formed. Steady state binding was described by two-site binding fits with nanomolar affinities of K_D1 = 1.3 nM and K_D2 = 185 nM for SRS2+5 and of K_D1 = 10.6 nM and K_D2 = 195 nM for the SRS* RNA (Figure 2 D, Table 3).

The individual dsRBDs 1 and 2 bind RNA dynamically

In order to obtain structural insights into RNA binding preferences of dsRBDs 1 and 2, $^{1}H,^{15}N$-HSQC NMR spectra of the individual dsRBDs and of the tandem dsRBD1-2 were measured. The spectra of both individual domains showed that they are well-folded (Figure S6). The two spectra also nicely matched with the NMR spectrum of the tandem domain dsRBD1-2. This indicated that in context of the tandem domain the individual structures of dsRBDs 1 and 2 are not altered and do not undergo significant interactions with each other. However, for dsRBD1 more signals than expected were observed in the 2D correlation spectrum, indicating the presence of multiple conformations. A number of intense NMR signals in the center of the dsRBD1 spectrum indicated a high degree of flexibility. Only partial chemical shift assignment was obtained, which indicated the presence of several

conformational states likely due to the high abundance of prolines in its sequence.

Upon titration of SRS2+5 RNA to each of the isolated domains, chemical shift perturbations (CSPs) and line broadening was observed (Figure 3 A, C). Residues affected by RNA binding to the isolated dsRBDs were similar to those affected in titration experiments with the dsRBD1-2 tandem domain (compare Figure 1 F), suggesting that both domains bind the RNA simultaneously and rather independently.

1D imino RNA traces upon protein binding revealed line broadening at sub-stoichiometric concentrations for dsRBD2 (Figure 3 D) similar to what was observed for the tandem domain dsRBD1-2 (Figure 1 E). Interestingly, for dsRBD1, less line-broadening is observed for imino signals (Figure 3B) at a protein:RNA ratio of 1:0.5. This indicates that the two dsRBDs bind RNA with different dynamics and possibly reduced affinity of dsRBD1 compared to dsRBD2.

dsRBD1 binds RNA significantly more weakly than dsRBD2

NMR data suggested differences in the RNA binding modes of dsRBDs 1 and 2. In order to understand the respective contribution of each dsRBD, we performed SPR experiments with each domain in isolation. Surprisingly, at protein concentrations up to 1 µM no RNA binding was observed for dsRBD1 (Figure 4 A, B, Tables 1, 2). In contrast, dsRBD2 bound to SRS2+5 RNA and to SRS* RNA with K_Ds of 650 nM and of 829 nM, respectively (Figure 4 C,D; Tables 1, 2). This binding was observed with fast on- and off-rates, similar to the tandem domain (Figure 2 A, B). Furthermore, dsRBD2 shows no sign of cooperativity, as indicated by Hill coefficients close to 1 (Figure 4 C, D). The lack of detectable RNA binding by SPR experiments with dsRBD1 compared to the detected interaction in NMR titrations can be explained by the much higher RNA concentrations used in the NMR experiments (50 µM).

We confirmed these findings by EMSA experiments, in which dsRBD1 did not bind to SRS2+5 RNA (Figure S7 A). Also, dsRBD2 bound SRS2+5 RNA much more weakly than the tandem domain dsRBD1-2, as binding was observed only at concentrations >10 µM (Figure S7 A).

The two domains are connected by a linker region of 19 amino acids. Thus, an alternative explanation to the cooperative binding of the dsRBDs could be that the linker region together with one dsRBD provides cooperative RNA binding. We tested this possibility by performing EMSAs either with a dsRBD1-linker fragment or with a fragment consisting of linker-dsRBD2. In neither of these cases did we observe any improved binding to SRS2+5 RNA (Figure S7 B). Also, mixing the two individual dsRBDs with linker did not improve RNA binding activity of the tandem domain. Together these results indicate that the two domains bind dsRNA with some cooperativity that is mediated by the presence of the linker, which itself does not contribute to the RNA recognition.

Mutations in dsRBD1 moderately impair RNA binding of dsRBD1-2

For further verification of the observed binding properties and to allow for functional *in vivo* studies of the RNA binding activity of Stau2 dsRBD1-2, RNA binding mutants of the dsRBDs 1 and 2 were designed. For dsRBD1, mutants were designed based on the NMR titration experiments and multiple sequence alignments with dmStau (Figure S8 A). A partial assignment allowed for the identification of residues with chemical shift perturbations upon RNA titration, pointing at their location within or close to the binding interface. These residues map to the predicted end of helix α_1, loop 2 and the beginning of helix α_2, which are the regions that mediate RNA binding in a canonical dsRBD (Figure S 8). Conserved dsRBD residues close or within these regions were chosen for mutation. We mutated glutamate in helix α_1 (E15), histidine in loop 2 (H36), lysines from the conserved KKxxK motif (K59 and K60) and phenylalanine in the beta strand β_2 (F40). Mutation of these residues in dsRBD3 from *D. melanogaster* to alanines had been shown to abolish RNA binding completely[23].

The dsRBD1-2 tandem domain with a range of mutations in dsRBD1 were tested for binding to SRS2+5 (Figure S10; Table 5). For the mutations E15A, H36A, F40A, K59A, K60A and K59A K60A binding kinetics were fast. Except for E15A the steady-state binding curves are best described by Hill-fits with Hill coefficients $n \approx 1$, indicating non-cooperative binding. Whereas the

observed K_Ds of dsRBD1-2 H36A and K59A are similar to that of dsRBD2 alone, dsRBD1-2 mutations F40A, K60A and K59A K60A bind with even lower affinity than RBD2 alone. These results indicate that binding-activity of RBD1 was abolished by these mutations. The only exception was RBD1-2 E15A, where steady-state binding to SRS2+5 was fitted with a first order binding reaction and a K_D of 132±31 nM, indicating that RNA binding activity of RBD1 might be compromised but not completely abolished.

When the same RBD1 mutations in context of the dsRBD1-2 fragment were tested for binding to the SRS* RNA, the mutations H36A, F40A, K59A and K59A K60A behaved again very similarly to their binding to SRS2+5, confirming that RNA binding activity of RBD1 is abolished by these mutations (Figure S11; Table 4). However, RBD1-2 E15A binds to SRS* RNA similar to the wild-type protein, showing two-site binding with K_D1=15±11 nM and K_D2=405±260 nM (Figure S11 A, compare with Figure 2 A), indicating that RNA binding activity of RBD1 is not corrupted by this mutation. Also, RBD1-2 K60A bound SRS* RNA, unlike SRS2+5 RNA, with affinities similar to the wild-type protein (Figure S 11 D). This mutation possibly has a less drastic effect that can be compensated for by the longer RNA.

Mutations in dsRBD2 impair RNA binding of dsRBD1-2 fragment

Due to the lack of NMR assignments for RBD2, to design mutations in this domain we had to rely on sequence homology. A sequence alignment of twelve species was used to identify conserved, positively charged or aromatic residues for mutation (Figure S8 B). These residues, E99A, K106A, F157A and H169A, were individually mutated in the context of the dsRBD1-2 tandem domain, and subsequently tested for RNA binding by SPR. All mutants showed strongly decreased binding to SRS2+5 (Figure S12, Table 5) and fitting of binding curves indicated that two-site binding was lost in all mutants. All dsRBD2 mutations in the context of dsRBD1-2 were also tested for binding to SRS* RNA. Unlike SRS2+5, Stau2 dsRBD1-2 E99A and K106A bound SRS* with properties similar to the wild-type protein (Figure S13, Table 4), indicating that the effects of these mutations are less dramatic and can be compensated for by the longer RNA. Binding of dsRBD1-2 H169A, however,

was still strongly decreased, such that a K_D could not be determined. Binding of dsRBD1-2 F157A was again strongly impaired and no K_D could be determined, thus confirming the results obtained for SRS2+5.

Double mutations in RBD1 and RBD2 strongly impair RNA binding of the tandem RBD1-2 domain

Based on the SPR results for the single point mutations in RBD1 and RBD2, double-mutants were designed in the context of Stau2 dsRBD1-2. In RBD1, the mutation F40A was chosen because it had a strong effect on binding to both tested RNAs, its resonance shifted upon RNA titration in the ^{1}H, ^{15}N-HSQC spectra, and it is conserved in dmStau. In RBD2, the mutations F157A and H169A were chosen. F157A showed altered binding kinetics and H169A had the strongest effect on RNA binding of all tested RBD2 mutations. Both residues are conserved in dmStau. Correct folding of double-mutant proteins was verified by CD spectroscopy (Figure S9). As expected, in the SPR experiments all double-mutant versions of Stau2 RBD1-2 lacked binding to SRS2-5 and to SRS* RNAs (Figure S14, Tables 3-4).

Double mutations in dsRBD1 and dsRBD2 strongly impair RNA binding of Stau2 dsRBD1-4

To assess the contribution of dsRBD1-2 to RNA binding in the context of all four verified RNA binding dsRBDs, SPR experiments were performed with Stau2 dsRBD1-4. Comparison of the wild-type Stau2 dsRBD1-4 with a dsRBD1-4 fragment harboring the double mutation F40A F157A (Figure 5 A), showed that RNA binding by the mutant protein was significantly impaired (Figure 5 B and D; compare to Figure 2 C).

In contrast to wild-type dsRBD1-4, the Stau2 dsRBD1-4 F40A H169A bound SRS2+5 and SRS* RNA with fast binding kinetics, resembling the transient binding by dsRBD3-4 alone (Figure 5 C, E; Tables 3 and 4; compare with Figure 2 B). The steady-state binding is well described by a two-site binding fit with K_Ds resembling those of dsRBD3-4 alone (Figure 5 C; compare with Figure 2 B). Additionally, kinetic fits to the binding curves obtained at 1000 nM protein were performed. A bivalent analyte fit to the binding curve at 1000 nM shows that the rate-constants ka_1 and kd_1 are both significantly increased for

F40A F169A when compared to the wild-type protein (Figure S15). Taken together, this indicates that in RBD1-4 F40A H169A RNA binding is mediated by RBD3-4 alone. Since mutations in dsRBD1-2 impair the affinity of dsRBD1-4 and because the interactions become much more transient, our data indicate that for efficient and stable RNA binding of Stau2 all four dsRBDs have to act in concert.

Rescue experiments with *Drosophila* Stau$^{-/-}$ embryos confirm the *in vivo* importance of RNA binding by RBD1-2

To assess the relevance of our findings for the *in vivo* function of Staufen proteins, we took advantage of the well-studied role of Stau in early *Drosophila* development. Here Stau is required for intracellular transport of *oskar* mRNA within the oocyte for its localization and translational activation at the posterior pole of the oocyte[7]. We expressed GFP-dmStau, GFP-mStau2, GFP-mStau2[F40A H169A] and GFP-mStau2[F40A F157A] transgenes in the germline of *stau^{R9}/stau^{D3}* mutant females[5] that lack endogenous dmStau protein and characterized their ability to rescue the mutant phenotype. As expected, wild-type GFP-dmStau supported efficient accumulation of *oskar* at the posterior pole (Figure 6 A-D) and even hatching of the majority of the resulting larvae. Also, mStau2-expressing *stau^{R9}/stau^{D3}* mutant flies mutant showed localization of a moderate fraction of *oskar* mRNA to the posterior pole of stage 9 oocytes (Figure 6 A-D). At stage 10, the localization of *oskar* mRNA improved further (data not shown). However, no hatching larvae were observed. mStau2 bearing the double mutation [F40A H169A] or [F40A F157A] in its dsRBDs 1 and 2 was equally expressed (Figure S 16), but failed to localize *oskar* mRNA to the posterior pole beyond *stau* null levels (Figure 6 A-D).

Evualuation of the content of mRNPs revealed a clear correlation between dmStau protein and *oskar* mRNA copy number (Figure 6 E-F). To a lesser extent, GFP-mStau2 also showed this positive correlation (Figure 6 E-F), suggesting an interaction between the mammalian Stau protein and *oskar* mRNA.

In contrast, the mutant proteins largely failed to scale with *oskar* mRNA copy numbers (Figure 6 F). In the case of *bicoid* mRNPs, Stau copy number per

mRNA scaled like wild-type in the case of [F40A H169A], but failed to scale in the case of [F40A F157A] (Figure 6 G, H).

In summary, these rescue experiments confirm the importance of RNA binding by dsRBD 1-2 for the *in vivo* function of Stau proteins.

Discussion

Previous reports had identified dsRBDs 3 and 4 in Staufen proteins as the RNA binding domains[11, 12, 48], suggesting that dsRBDs 1 and 2 fulfill other functions. In our present work we have demonstrated that the mStau2 dsRBDs 1 and 2 also possess RNA binding activity. Furthermore, we could show that the two dsRBDs 1 and 2 work together as a tandem domain to achieve their full functionality. Our data further confirm that also dsRBD3-4 also act as a tandem domain. Our comparison of RNA binding affinities of single domains and tandem domains bearing mutations suggest that the first binding event with moderate affinity is achieved by the second dsRBD in each tandem domain, namely dsRBD2 and dsRBD4.

Based on our results, we propose a model in which sequential binding events lead to stable RNA recognition by Staufen (Figure 7). In this model, binding of the first tandem domain occurs initially at a random position, with dsRBD2 achieving the first interaction (Figure 7 A, left side). Our NMR data suggest that dsRBD2 unspecifically slides along the stem, thereby positioning dsRBD1 close to the dsRNA. Subsequently, dsRBD1 also binds, thereby increasing the affinity of the tandem domain to dsRNA (Figure 7A, right side). In the tandem domain dsRBD3-4 the second domain, dsRBD4, binds with higher affinity (Figure S17) and thus likely undergoes the first priming contact in a fashion similar to dsRBD2 (Figure 7 B, left side). Then, the other, free dsRBD of the tandem domain also joins the RNA-bound complex (Figure 7B, right side). Only when the two tandem domains dsRBD1-2 and dsRBD3-4 act together does the protein form a stable RNA complex. This interpretation is consistent with the stronger and more stable RNA binding of dsRBD1-4 and full-length mStau2.

To confirm *in vivo* that RNA binding by the dsRBD1-2 is important for the function of the full-length protein, we utilized the *Drosophila* oocyte as model system. In the germline of otherwise *stau* null flies, different variants of Stau were expressed and the rescue of the mutant phenotype assessed by analyzing *oskar* mRNA localization to the posterior pole. Surprisingly, in mStau2-expressing oocytes a moderate rescue of *oskar* mRNA localization was observed. In contrast, mStau2-rescue constructs bearing RNA binding mutations in dsRBD1-2 failed to rescue *oskar* localization. This observation confirms the importance of dsRBD1-2 for RNA binding and RNA localization *in vivo*.

This observation, together with the fact that the long isoform of mStau2 has the same number of dsRBDs, indicate that mStau2 might be the functional homolog of dmStau. The fact that mStau1 lacks the first dsRBD raises the question how this paralog achieves full binding. One option is that its mode of RNA binding is different enough from mStau2 to allow for strong and stable binding even with only two or three dsRBDs. An alternative could be that the reported dimerization of mStau1[48] allows for the joint action of dsRBDs in trans and thus full, stable RNA binding.

Sliding as initial binding mode also occurs in other RNA binding proteins such as the *Drosophila* protein Loqs-PD, a member of the siRNA silencing pathway. Loqs-PD contains two canonical dsRBDs that show highly dynamic binding that involves sliding along RNA stems to scan for suitable secondary structures[46]. A similar activity was also reported for the human ortholog of Loqs-PD, TAR RNA binding protein (TRBP)[45]. In contrast to these examples, however, Stau2 involves two tandem domains with four dsRBDs for its sequential RNA-target recognition.

The feature of stable RNA binding is likely to be of great importance for transport of transcripts over longer distances. It is therefore not surprising that our rescue experiments of *Drosophila stau* mutants with Stau constructs required all four dsRBDs to be functional. The presented model (Figure 7) offers a mechanistic view on how Stau2 may recognize biological targets with high affinity and stability.

RNA binding proteins in higher eukaryotes very often contain multiple RNA binding domains[49]. It is thought that these act in a combinatorial fashion such

as we have shown for the dsRBDs of mStau2. However, for most of these multidomain proteins the manner in which they act cooperatively for function and specificity is not well understood. mStau2 contains two tandem domains, each of which can bind to a stem-loop. It is likely that the choice and combination of the stem-loops as well as their specific spacing determines the specificity of Stau binding for transport of selected mRNAs *in vivo*.

Acknowledgement

We thank Vera Roman for technical support and Michael A. Kiebler for fruitful discussions and support. We acknowledge the support by the Bavarian NMR center and the X-ray crystallography platform at the Helmholtz Zentrum München. This work was supported by the Deutsche Forschungs-gemeinschaft (FOR2333 to S.H., I.G., A.E., and D.N.).

Tables

Table 1: Stau2 binding to SRS*

Stau2 wt	binding	kinetics	K_D1 [nM]	K_D2 [nM]	Hill coefficient
dsRBD3-4	two-site	transient	9 ± 1	n.d. (>1000)	-
dsRBD1-2	Two-site	transient	25 ± 8	n.d. (>1000)	-
dsRBD1-4	Hill	stable	-	287 ± 127	1.7 ± 0.4
dsRBD2	Hill	transient	-	828 ± 29	1.1 ± 0.04
dsRBD1	no binding	-	-	-	-

Table 2: Stau2 binding to dsSRS2+5 RNA

Stau2 wt	binding	kinetics	K_D1 [nM]	K_D2 [nM]	Hill coefficient
dsRBD3-4	two-site	transient	18 ± 16	n.d. (>1000)	-
dsRBD1-2	two-site	transient	130 ± 30	n.d. (>1000)	-
dsRBD1-4	Hill	stable	-	330 ± 148	1.6 ± 0.2
dsRBD2	Hill	transient	-	650 ± 247	1.3 ± 0.3
dsRBD1	no binding	no binding	-	-	-

Table 3: RNA binding to Stau2 FL

RNA	binding	kinetics	**K_D1 [nM]**	**K_D2 [nM]**
SRS2+5	two-site	stable	1.3 ± 0.6	185 ± 66
SRS3	two-site	stable	10.6 ± 5	195 ± 103

Table 4: Stau2 dsRBD1-2 mutant binding to SRS* RNA

Stau2 dsRBD1-2 protein	binding	kinetics	K_D1 [nM]	K_D2 [nM]	Hill coefficient
wt	two-site	transient	25 ± 8	n.d. (>1000)	-
E15A	two-site	transient	15 ± 11	406 ± 359	-
H36A	Hill	transient	-	499 ± 149	0.9 ± 0.06
F40A	Hill	transient	-	769 ± 427	1 ± 0.1
K59A	Hill	transient	-	624 ± 142	0.9 ± 0.2
K60A	two-site	transient	36 ± 5	n.d. (>1000)	-
K59A K60A	Hill	transient	-	n.d. (>1000)	1 ± 0.1
E99A	two-site	transient	14 ± 0.7	850 ± 295	-
K106A	two-site	transient	30 ± 10	n.d. (>1000)	-
F157A	Hill	stable	-	n.d. (>1000)	1.2 ± 0.2
H169A	Hill	transient	-	n.d. (>1000)	1 ± 0.6
F40A F157A	no fit	transient	-	-	-
F40A H169A	no fit	transient	-	-	-

Table 5: Stau2 dsRBD1-2 mutant binding to dsSRS2+5 RNA

Stau2 dsRBD1-2 protein	binding	kinetics	K_D1 [nM]	K_D2 [nM]	Hill coefficient
wt	two-site	transient	130 ± 30	n.d. (>1000)	-
E15A	Hill	transient	-	132 ± 31	1 ± 0.1
H36A	Hill	transient	-	701 ± 386	1 ± 0.2
F40A	Hill	transient	-	n.d. (>1000)	1.2 ± 0.1
K59A	Hill	transient	-	509 ± 136	1.2 ± 0.2
K60A	Hill	transient	-	n.d. (>1000)	1 ± 0.1
K59A K60A	Hill	transient	-	n.d. (>1000)	1 ± 0.02
E99A	Hill	transient	-	635 ± 308	1.5 ± 0.3
K106A	no fit	transient	-	n.d. (>1000)	-
F157A	Hill	stable	-	497 ± 238	1.5 ± 0.2
H169A	no fit	transient	-	n.d. (>1000)	-
F40A F157A	no fit	stable	-	-	-
F40A H169A	no fit	transient	-	-	-

References

1. Buxbaum AR, Haimovich G, Singer RH. In the right place at the right time: visualizing and understanding mRNA localization. *Nat Rev Mol Cell Biol* **16**, 95-109 (2015).

2. Blower MD. Molecular insights into intracellular RNA localization. *International review of cell and molecular biology* **302**, 1-39 (2013).

3. Holt CE, Schuman EM. The central dogma decentralized: new perspectives on RNA function and local translation in neurons. *Neuron* **80**, 648-657 (2013).

4. Hutten S, Sharangdhar T, Kiebler M. Unmasking the messenger. *RNA Biol* **11**, 992-997 (2014).

5. St Johnston D, Beuchle D, Nusslein-Volhard C. Staufen, a gene required to localize maternal RNAs in the Drosophila egg. *Cell* **66**, 51-63 (1991).

6. St Johnston D, Brown NH, Gall JG, Jantsch M. A conserved double-stranded RNA-binding domain. *Proc Natl Acad Sci U S A* **89**, 10979-10983 (1992).

7. Lehmann R. Germ Plasm Biogenesis--An Oskar-Centric Perspective. *Curr Top Dev Biol* **116**, 679-707 (2016).

8. Li P, Yang X, Wasser M, Cai Y, Chia W. Inscuteable and Staufen mediate asymmetric localization and segregation of *prospero* RNA during *Drosophila* neuroblast cell divisions. *Cell* **90**, 437-447 (1997).

9. Buchner G, Bassi MT, Andolfi G, Ballabio A, Franco B. Identification of a novel homolog of the Drosophila staufen protein in the chromosome 8q13-q21.1 region. *Genomics* **62**, 113-118 (1999).

10. Furic L, Maher-Laporte M, DesGroseillers L. A genome-wide approach identifies distinct but overlapping subsets of cellular mRNAs associated with Staufen1- and Staufen2-containing ribonucleoprotein complexes. *RNA* **14**, 324-335 (2008).

11. Wickham L, Duchaine T, Luo M, Nabi IR, DesGroseillers L. Mammalian staufen is a double-stranded-RNA- and tubulin-binding protein which localizes to the rough endoplasmic reticulum. *Mol Cell Biol* **19**, 2220-2230 (1999).

12. Duchaine TF, Hemraj I, Furic L, Deitinghoff A, Kiebler MA, DesGroseillers L. Staufen2 isoforms localize to the somatodendritic domain of neurons and interact with different organelles. *J Cell Sci* **115**, 3285-3295 (2002).

13. Gong C, Maquat LE. lncRNAs transactivate STAU1-mediated mRNA decay by duplexing with 3' UTRs via Alu elements. *Nature* **470**, 284-288 (2011).

14. Kiebler MA*, et al.* The mammalian staufen protein localizes to the somatodendritic domain of cultured hippocampal neurons: implications for its involvement in mRNA transport. *J Neurosci* **19**, 288-297 (1999).

15. Kim YK, Furic L, Desgroseillers L, Maquat LE. Mammalian Staufen1 recruits Upf1 to specific mRNA 3'UTRs so as to elicit mRNA decay. *Cell* **120**, 195-208 (2005).

16. Monshausen M*, et al.* Two rat brain staufen isoforms differentially bind RNA. *J Neurochem* **76**, 155-165 (2001).

17. Kusek G, Campbell M, Doyle F, Tenenbaum SA, Kiebler M, Temple S. Asymmetric segregation of the double-stranded RNA binding protein Staufen2 during mammalian neural stem cell divisions promotes lineage progression. *Cell stem cell* **11**, 505-516 (2012).

18. Heraud-Farlow JE*, et al.* Staufen2 Regulates Neuronal Target RNAs. *Cell reports* **5**, 1511-1518 (2013).

19. Laver JD*, et al.* Genome-wide analysis of Staufen-associated mRNAs identifies secondary structures that confer target specificity. *Nucleic Acids Res* **41**, 9438-9460 (2013).

20. Park E, Maquat LE. Staufen-mediated mRNA decay. *Wiley interdisciplinary reviews RNA* **4**, 423-435 (2013).

21. Gleghorn ML, Maquat LE. 'Black sheep' that don't leave the double-stranded RNA-binding domain fold. *Trends Biochem Sci* **39**, 328-340 (2014).

22. Macchi P, Brownawell AM, Grunewald B, DesGroseillers L, Macara IG, Kiebler MA. The brain-specific double-stranded RNA-binding protein Staufen2: nucleolar accumulation and isoform-specific exportin-5-dependent export. *J Biol Chem* **279**, 31440-31444 (2004).

23. Ramos A*, et al.* RNA recognition by a Staufen double-stranded RNA-binding domain. *EMBO J* **19**, 997-1009 (2000).

24. Sugimoto Y*, et al.* hiCLIP reveals the in vivo atlas of mRNA secondary structures recognized by Staufen 1. *Nature* **519**, 491-494 (2015).

25. Ho SN, Hunt HD, Horton RM, Pullen JK, Pease LR. Site-directed mutagenesis by overlap extension using the polymerase chain reaction. *Gene* **77**, 51-59 (1989).

26. Studier FW. Protein production by auto-induction in high density shaking cultures. *Protein Expr Purif* **41**, 207-234 (2005).

27. Winn MD*, et al.* Overview of the CCP4 suite and current developments. *Acta Crystallogr D Biol Crystallogr* **67**, 235-242 (2011).

28. Panjikar S, Parthasarathy V, Lamzin VS, Weiss MS, Tucker PA. Auto-rickshaw: an automated crystal structure determination platform as an efficient tool for the validation of an X-ray diffraction experiment. *Acta Crystallogr D Biol Crystallogr* **61**, 449-457 (2005).

29. Kneller JM*, et al.* Continuous-flow optical pumping NMR in a closed circuit system. *J Magn Reson* **147**, 261-265 (2000).

30. Vranken WF*, et al.* The CCPN data model for NMR spectroscopy: development of a software pipeline. *Proteins* **59**, 687-696 (2005).

31. St Johnston D, Driever W, Berleth T, Richstein S, Nusslein-Volhard C. Multiple steps in the localization of bicoid RNA to the anterior pole of the Drosophila oocyte. *Development* **107 Suppl**, 13-19 (1989).

32. Schuldt AJ*, et al.* Miranda mediates asymmetric protein and RNA localization in the developing nervous system. *Genes Dev* **12**, 1847-1857 (1998).

33. Telley IA, Gaspar I, Ephrussi A, Surrey T. Aster migration determines the length scale of nuclear separation in the Drosophila syncytial embryo. *J Cell Biol* **197**, 887-895 (2012).

34. Bischof J, Bjorklund M, Furger E, Schertel C, Taipale J, Basler K. A versatile platform for creating a comprehensive UAS-ORFeome library in Drosophila. *Development* **140**, 2434-2442 (2013).

35. Gaspar I, Wippich F, Ephrussi A. Enzymatic production of single-molecule FISH and RNA capture probes. *RNA* **23**, 1582-1591 (2017).

36. Gaspar I, Ephrussi A. Ex vivo Ooplasmic Extract from Developing Drosophila Oocytes for Quantitative TIRF Microscopy Analysis. *Bio Protoc* **7**, (2017).

37. Gaspar I, Yu YV, Cotton SL, Kim DH, Ephrussi A, Welte MA. Klar ensures thermal robustness of oskar localization by restraining RNP motility. *J Cell Biol* **206**, 199-215 (2014).

38. Gaspar I, Sysoev V, Komissarov A, Ephrussi A. An RNA-binding atypical tropomyosin recruits kinesin-1 dynamically to oskar mRNPs. *EMBO J* **36**, 319-333 (2017).

39. Little SC, Sinsimer KS, Lee JJ, Wieschaus EF, Gavis ER. Independent and coordinate trafficking of single Drosophila germ plasm mRNAs. *Nat Cell Biol* **17**, 558-568 (2015).

40. Lenth RV. Least-Squares Means: TheRPackagelsmeans. *Journal of Statistical Software* **69**, (2016).

41. Team RDC. *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria (2008).

42. Wickham H. *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York (2016).

43. Göbl C, Madl T, Simon B, Sattler M. NMR approaches for structural analysis of multidomain proteins and complexes in solution. *Prog Nucl Magn Reson Spectrosc* **80**, 26-63 (2014).

44. Williamson RT, Buevich AV, Martin GE, Parella T. LR-HSQMBC: a sensitive NMR technique to probe very long-range heteronuclear coupling pathways. *J Org Chem* **79**, 3887-3894 (2014).

45. Koh HR, Kidwell MA, Ragunathan K, Doudna JA, Myong S. ATP-independent diffusion of double-stranded RNA binding proteins. *Proc Natl Acad Sci U S A* **110**, 151-156 (2013).

46. Tants JN*, et al.* Molecular basis for asymmetry sensing of siRNAs by the Drosophila Loqs-PD/Dcr-2 complex in RNA interference. *Nucleic Acids Res* **45**, 12536-12550 (2017).

47. Wang Z, Bhattacharya A, Ivanov DN. Identification of Small-Molecule Inhibitors of the HuR/RNA Interaction Using a Fluorescence Polarization Screening Assay Followed by NMR Validation. *PLoS One* **10**, e0138780 (2015).

48. Gleghorn ML, Gong C, Kielkopf CL, Maquat LE. Staufen1 dimerizes through a conserved motif and a degenerate dsRNA-binding domain to promote mRNA decay. *Nat Struct Mol Biol* **20**, 515-524 (2013).

49. Lunde BM, Moore C, Varani G. RNA-binding proteins: modular design for efficient function. *Nat Rev Mol Cell Biol* **8**, 479-490 (2007).

Figures - Heber et al.

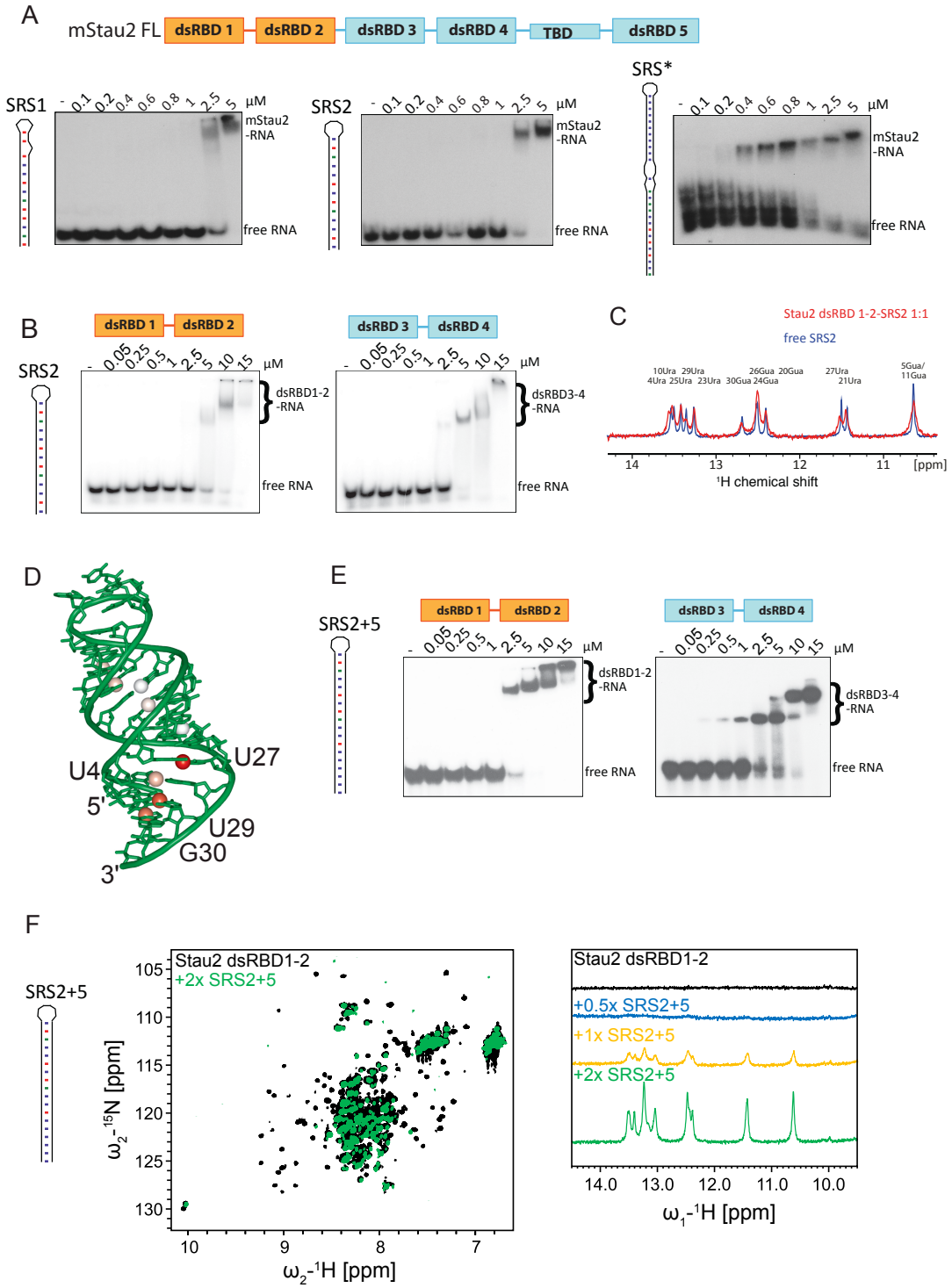

Figure 1: The RBD1-2 tandem domain of Stau2 binds RNA. A) EMSAs with full-length Stau2 (Stau2 FL) and different SRS RNAs from the 3'UTR of the *Rgs4* mRNA. **B)** EMSAs with Stau2 tandem domains dsRBD1-2 or dsRBD3-4 and SRS2. **C)** NMR titration experiments of Stau2 dsRBD1-2 with SRS2 RNA. **D)** Iminos showing significant line-broadening are indicated by colored (orange to red) spheres on the crystal structure of SRS2. **E)** EMSAs with Stau2 tandem-domains dsRBD1-2 or dsRBD3-4 and SRS2 RNA extended by five basepairs (SRS2+5). **F)** NMR titration experiments of Stau2 dsRBD1-2 with SRS2+5 RNA. Left: Overlay of $^{1}H,^{15}N$-HSQC spectra of dsRBD1-2 in absence and presence of 2x excess SRS2+5 RNA. Resonance shifts and line broadening of several signals are observed. Right: Comparison of 1D imino traces of SRS2+5 RNA at different stoichiometric ratios with dsRBD1-2. Strong line broadening of imino signals is observed in presence of protein.

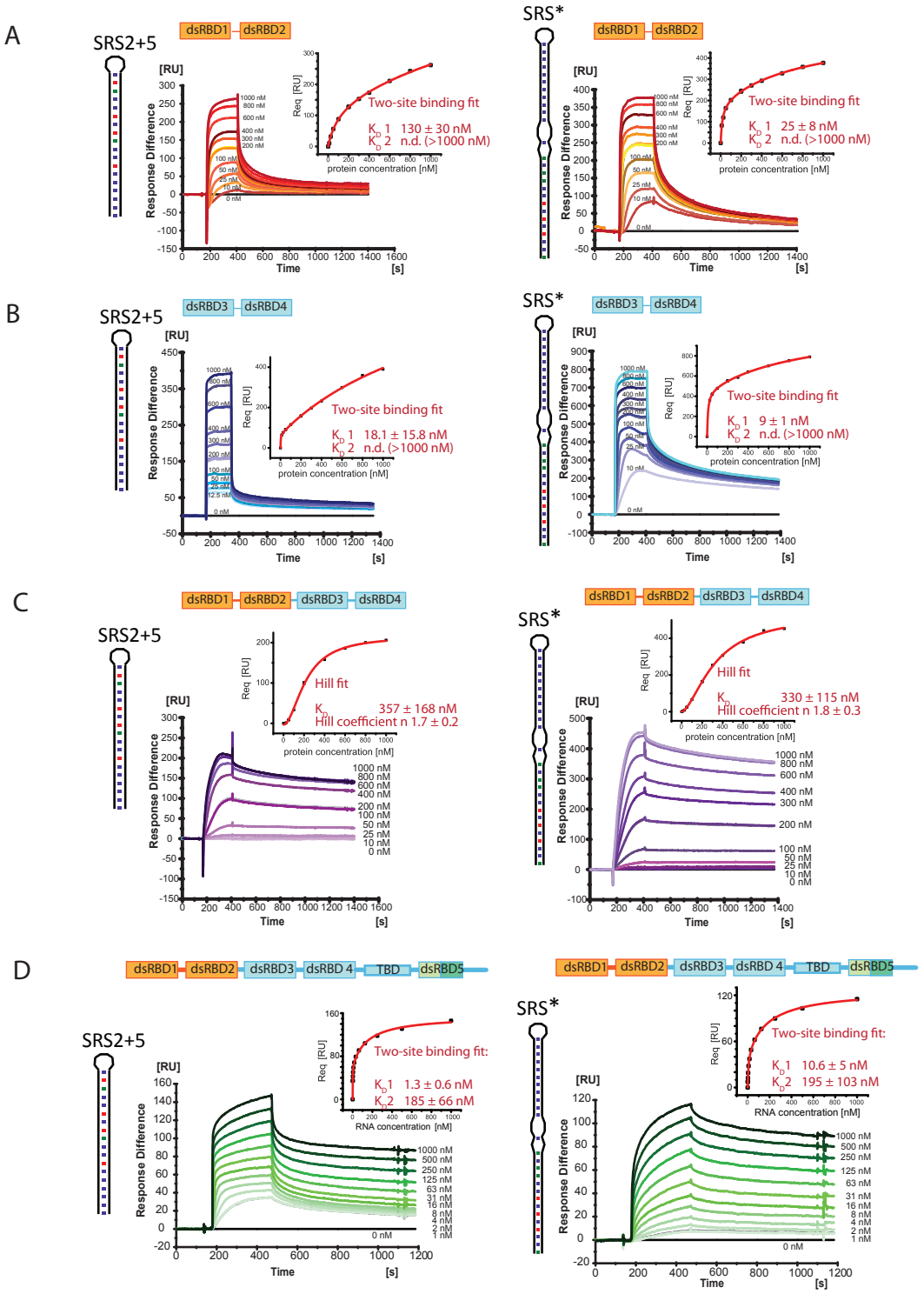

Figure 2: Surface Plasmon Resonance (SPR) sensorgrams and steady-state binding curves of Stau2 protein binding to SRS2+5 RNA (left) and SRS* RNA (right). A) Stau2 dsRBD1-2, and **B)** Stau2 dsRBD3-4 binding to surface-coupled SRS2+5 and SRS*. RNAs. The tandem domains dsRBD1-2 and dsRBD3-4 bind transiently with fast kinetics. The steady-state binding curves do not saturate up to 1 µM protein concentration but can be described by a two-site binding fit with K_D1 of 18 nM and 130 nM, respectively, for dsSRS2+5 and K_D1 of 9 nM and 25 nM, respectively, for SRS*. **C)** Stau2 dsRBD1-4 binding to surface-coupled SRS2+5 RNA and SRS* is stable with slower kinetics. The steady-state binding curve saturates at approximately 1 µM and is described by a Hill fit with an apparent overall K_D of 357 nM and a Hill coefficient n=1.7 for SRS2+5 and an apparent overall K_D of 330 nM and a Hill coefficient n=1.8 for SRS*, indicating positive cooperative binding. D) SRS2+5 and SRS* bind to surface-coupled Stau2 FL stably and with high affinity. The steady state binding curves can be described by a two-site binding fit with K_D1 of 1.3 nM and 10.6 nM and K_D2 of 185 nM and 195 nM, respectively.

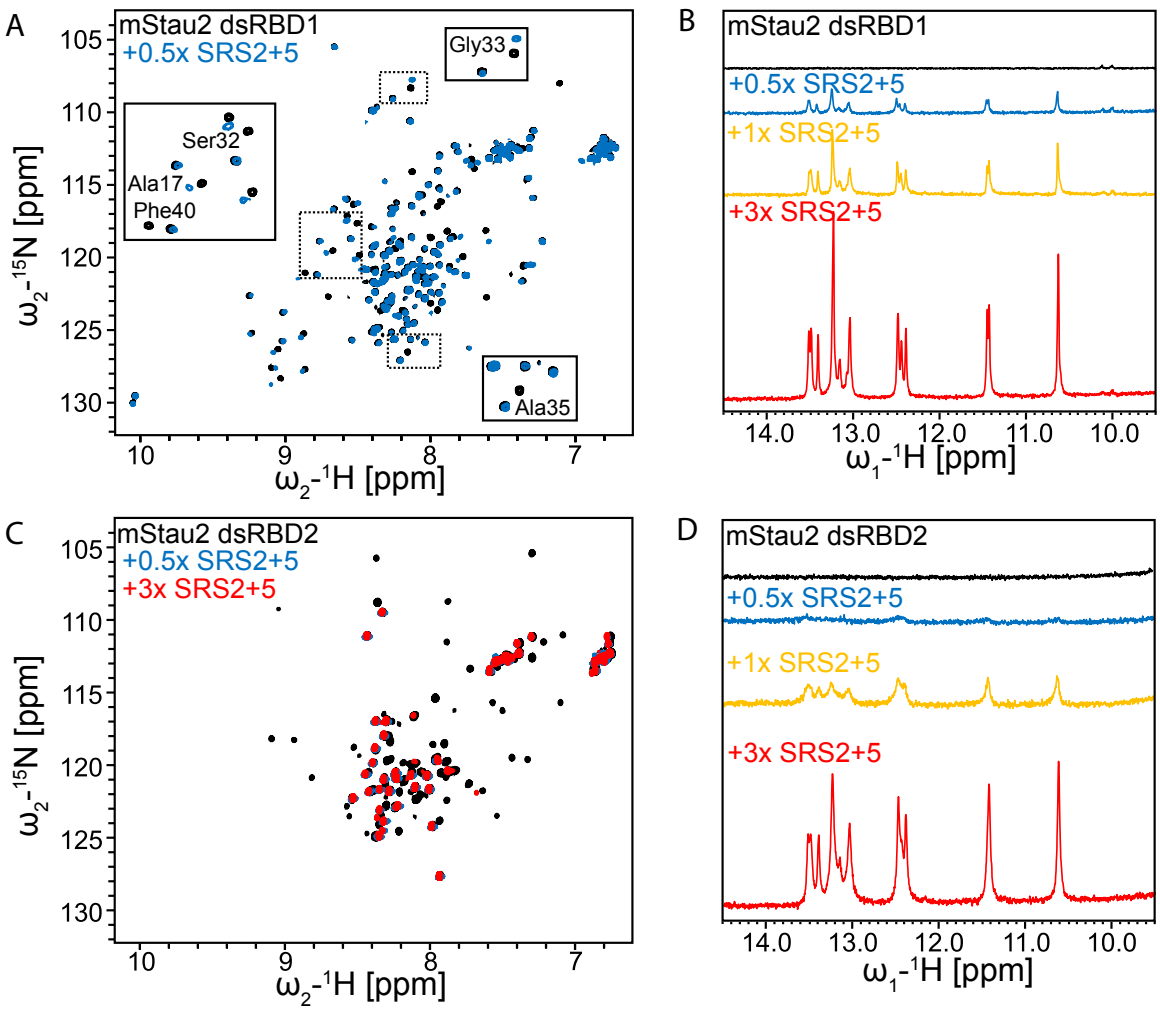

Figure 3: NMR titration experiments of Stau2 dsRBD1 (A and B) and dsRBD2 (C and D) with SRS2+5 RNA. A) and **C)** Overlay of ^{1}H, ^{15}N-HSQC spectra of dsRBD1 (A) or dsRBD2 (C) in absence and presence of SRS2+5 RNA. Resonance shifts and line broadening of several signals are observed for both domains. **B)** and **D)** Comparison of 1D imino traces of SRS2+5 RNA at different stoichiometric ratios with dsRBD1 (B) or dsRBD2 (D). Strong line broadening of imino signals is observed in presence of dsRBD2 but not dsRBD1, pointing at differences in binding.

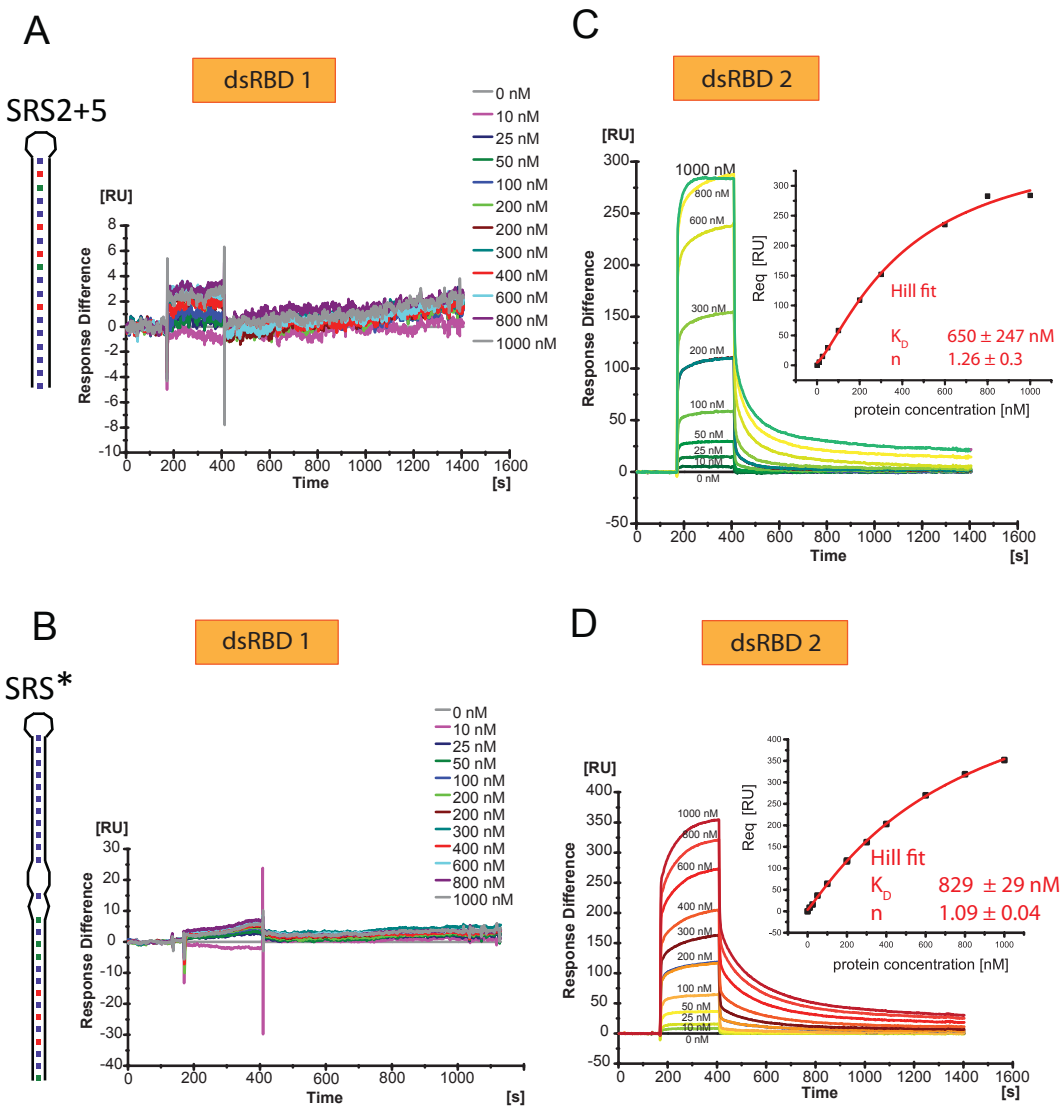

Figure 4: SPR sensorgrams and steady-state binding curves for Stau2 dsRBD1 or Stau2 dsRBD2 binding to SRS2+5 or SRS*. At micro-molar concentrations, Stau2 dsRBD1 alone binds neither to **(A)** SRS2+5 nor to **(B)** SRS*. In contrast, dsRBD2 binds to **(C)** SRS2+5 and to **(D)** SRS* with fast kinetics in a non-cooperative fashion with K_Ds of 650 nM or 829 nM, respectively. Steady-state binding curves are described by a Hill fit.

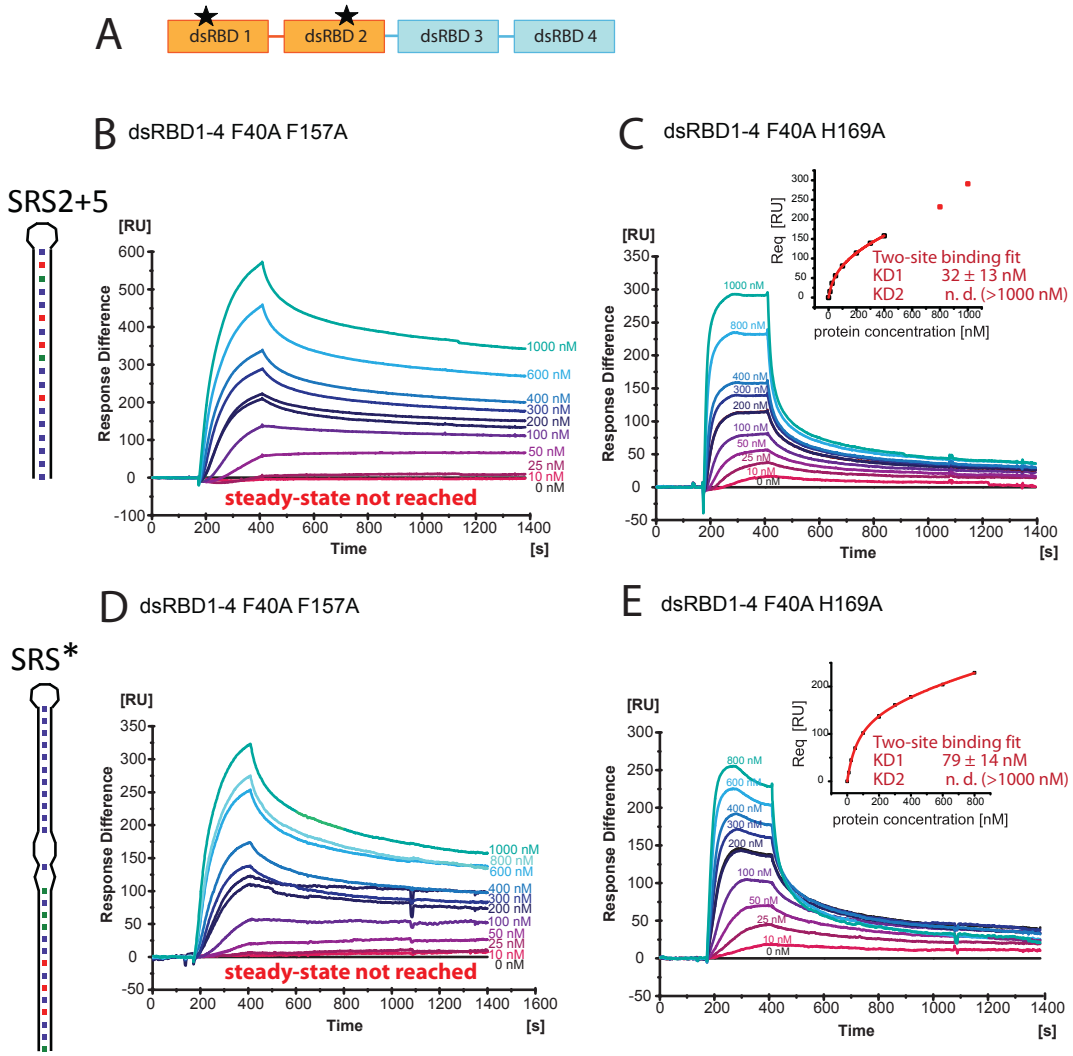

Figure 5: Binding experiments with mutant versions of dsRBD1-4 confirm a contribution of dsRBDs1-2 to RNA binding. (A) Schematic drawing of Stau2 with its two mutations in dsRBDs1-2. **B, C)** SPR experiments with Stau2 dsRBD1-4 double-mutants binding to SRS2+5 RNA and (C,D) to SRS*. Binding to (A) SRS2+5 and (B) SRS* is strongly decreased for the F40A F157A mutant as compared to Stau2 dsRBD1-4 wild type. Binding of the F40A H169A mutant shows dramatically altered kinetics as compared to dsRBD1-4 wild type and resembles binding by the tandem-domain dsRBD3-4 alone.

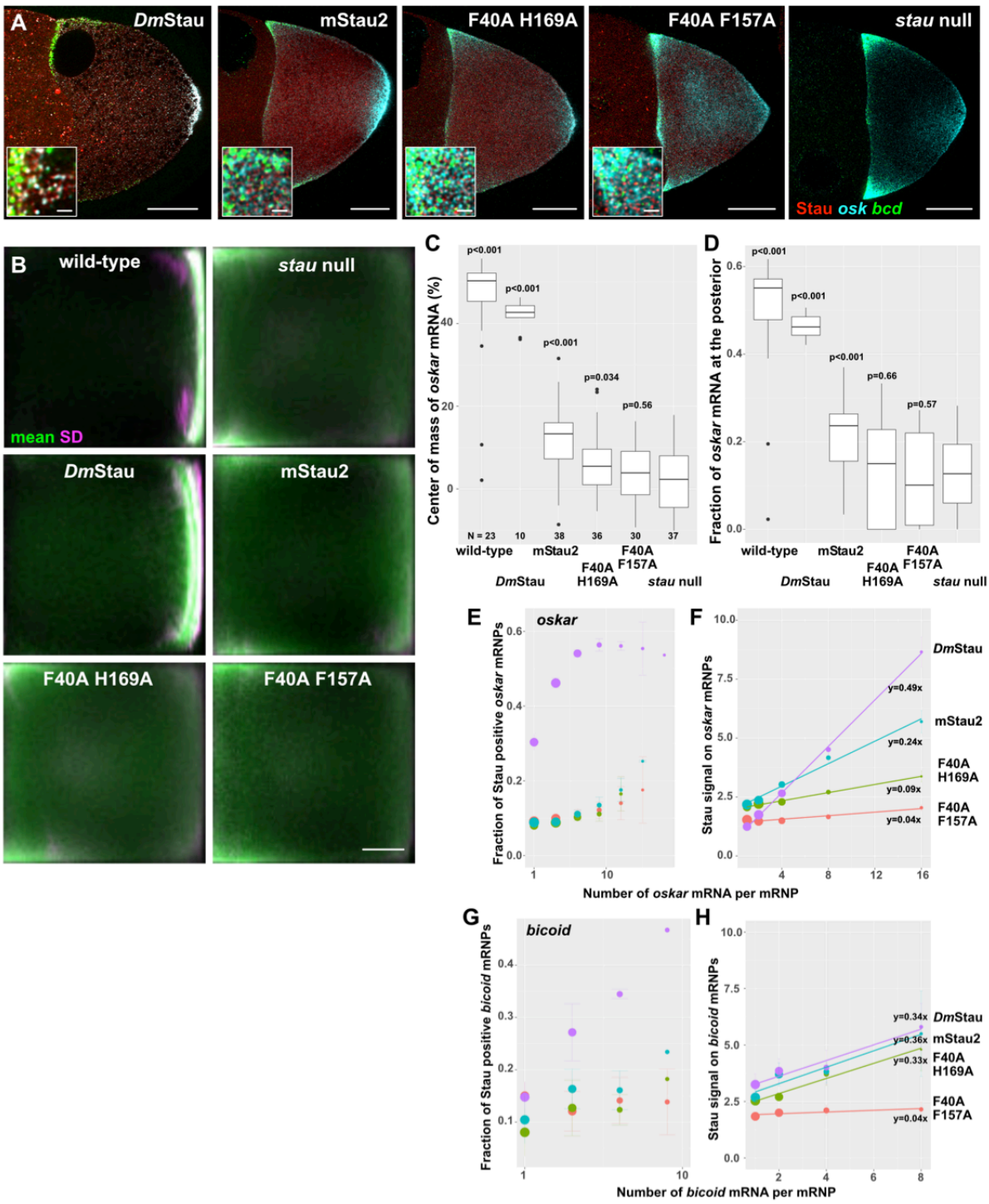

Figure 6: Functional interaction of mStau2 with *oskar* mRNA during *Drosophila* oogenesis. (A) Expression of GFP-DmStau, GFP-mStau2, GFP-mStau2[F40A H169A] and GFP-mStau2[F40A F157A] in the germline of $stau^{R9}/stau^{D3}$ mutant females that lack the dmStau protein. Transgenic GFP-Stau protein is shown in red, *bicoid* and *oskar* mRNA molecules are shown in green and cyan, respectively. In the dmStau-expressing oocytes *oskar* localizes almost exclusively to the posterior pole (right side of images) and *bicoid* to the anterior pole (left side of images) during stage 9 of oogenesis. In oocytes that lack Stau (*stau* null), *oskar* is found at both poles, enriching slightly more at the anterior, while *bicoid* localization is unaffected during stage 9. Insets show magnified regions of the upper anterior corner. Scale bar is 20 µm and 1 µm for the insets. **(B)** Typical localization of *oskar* mRNA in oocytes as a function of the expressed Stau protein. Using image transformation algorithms, the RNA signal within the oocytes was redistributed into a 100x100 square matrix. Such matrices of different oocytes could be statistically evaluated to obtain the average (green) and the variability (magenta) of the RNA distribution under a given condition. In wild-type oocytes (top left) most signal is found close to the posterior pole (on the right of the panels) by stage 9. In absence of Stau (stau null), there is a strong accumulation of

oskar mRNA at the anterior pole as well. Scale bar represents 20% of the length of the anteroposterior axis. **(C, D)** Evaluation of *oskar* mRNA localization by comparing the center of mass (relative to the geometric center at 0, **C**) and the fraction of *oskar* mRNA that localizes at the posterior pole (D) during stage 9. P-values show the result of pair-wise Mann-Whitney U tests versus the *stau* null condition (Bonferroni corrected alpha value is 0.01). Numbers in C show the number of different oocytes per condition. **(E-H)** Interaction of GFP-tagged Stau molecules with *oskar* (**E**, **F**) and *bicoid* (G. H) mRNPs. mRNPs were sorted by their mRNA content using quantitative smFISH. The fraction of Stau positive mRNPs (**E**, **G**) and the normalized GFP-Stau signal intensity (**F**, **H**) were plotted as a function of mRNA content of the mRNPs. The normalized GFP-Stau signal intensities were fitted linear models, the slope of which is indicated. In case of *oskar* mRNPs (**F**), all slopes are significantly different in pairwise comparisons ($p<0.0001$) with the exception of GFP-mStau2[F40A H169A] vs GFP-mStau2[F40A F157A] ($p=0.016$, $alpha_{corrected}=0.01$), indicating that wild-type mStau2 has higher affinity to *oskar* mRNA than its mutant counterparts. In the case of *bicoid* mRNPs (H), the slope of GFP-mStau2[F40A F157A] is significantly different from the other three ($p<0.01$), which do not differ significantly from each other ($p>0.9$).

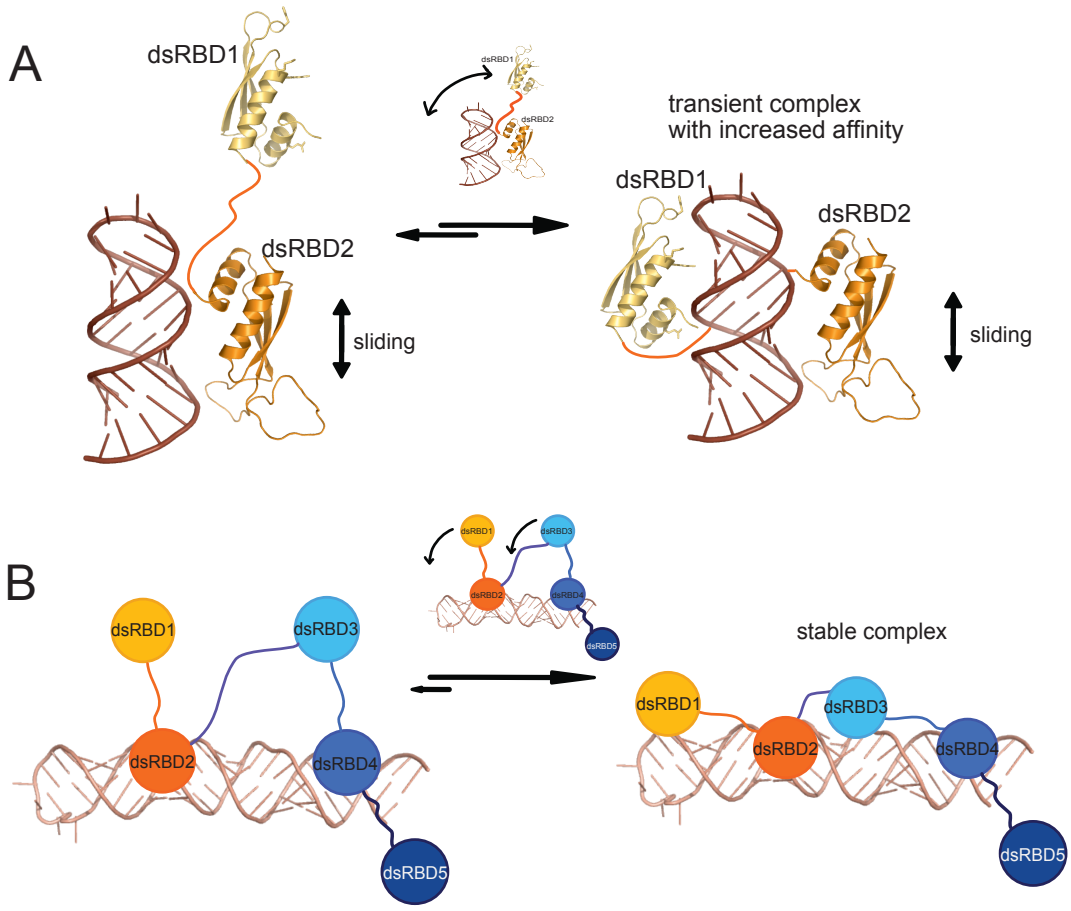

Figure 7: Model of the molecular recognition of dsRNA by Stau2. A) dsRNA recognition by the Stau2 tandem domain dsRBD1-2. dsRBD2 binds dsRNA promiscuously with moderate affinity and slides along the stem. Through this sliding, dsRBD2 positions dsRBD1 close to the dsRNA. When a suitable dsRNA structure is reached, dsRBD1 also binds, thereby strongly increasing the affinity of the tandem domain to dsRNA. **B)** Next to the tandem domain dsRBD1-2, dsRBD3-4 acts in a similar way. Here, dsRBD4 does the first promiscuous binding with moderate affinity. When dsRBD2 and dsRBD4 position dsRBD1 and dsRBD3, respectively, close to a suitable dsRNA, the respective domains also bind the dsRNA, thereby increasing affinity. Only when suitable dsRNA binding sites for both tandem domains are in sufficient spatial proximity can all four dsRBDs be bound and form a stable complex with the RNA target.