PNAS

Supporting Information for

SRSF1 interactome determined by proximity labeling reveals direct interaction with spliceosomal RNA helicase DDX23

Danilo Segovia[1,2], Dexter W. Adams[3,4], Nickolas Hoffman[1], Polona Safaric Tepes[1], Tse-Luen Wee[1], Paolo Cifani[1], Leemor Joshua-Tor[3], Adrian R. Krainer[1]
1 - Cold Spring Harbor Laboratory, 1 Bungtown Road, Cold Spring Harbor, NY 11724, USA
2 - Graduate Program in Molecular and Cellular Biology, Stony Brook University, 100 Nicolls Rd, Stony Brook, NY 11794, USA
3 - W. M. Keck Structural Biology Laboratory, Howard Hughes Medical Institute, Cold Spring Harbor, NY 11724, USA
4 - Graduate Program in Genetics, Stony Brook University, 100 Nicolls Rd, Stony Brook, NY 11794, USA
*Adrian R. Krainer
Cold Spring Harbor Laboratory, 1 Bungtown Road, Cold Spring Harbor, NY 11724, USA
516-367-8417
Email: krainer@cshl.edu

This PDF file includes:

Supplementary Methods.

Plasmids.

For proximity labeling, we generated the BS construct by cloning SRSF1 cDNA in the mammalian retroviral vector myc-BioID2-pBABE puro (Addgene #80900) (1). We generated the SB construct by cloning SRSF1 cDNA downstream of BioID2 in the MCS-BioID2-HA pBabe-puro (Addgene #120308). To generate the BG construct, we subcloned EGFP-3XNLS from pEGFP-C1 EGFP-3XNLS (Addgene #58468) into myc-BioID2-pBABE puro. To generate Dox-inducible cell lines, we cloned T7-SRSF1 or the constructs containing the BioID2 fusions into the third-generation lentiviral vector pCW57.1 (Addgene #41393).

For BiFC, we used the dual expression plasmid pKK-BiFC-Venus (Addgene #105804) (2). We cloned SRSF1 cDNA downstream of the N-terminus of mVenus, and the different gene candidates upstream of the C-terminus of mVenus. For the analysis of interactions between additional SR protein family members and DDX23, we cloned the cDNA of each SR protein into the vector harboring the DDX23-mV-Ct fusion. All these constructs were generated using Gibson assembly (NEB).

For recombinant protein production, we used the His-SUMO-SRSF1 plasmid for SRSF1 expression, and pCDF-CLK1 for CLK1 co-expression (3). We generated a plasmid expressing His-SUMO-RRMs by subcloning into pLIC-SGC1 (Addgene #39187). For expression of isolated domains of SRSF1, we subcloned RRM1, RRM2 and RS domain into pNH-TrxT, and expressed the proteins as a fusion with His-TrxA.

For the generation of DDX23 constructs, we amplified the cDNA from a HeLa-cell-derived library and we cloned it into pNH-TrxT (Addgene #26106) using ligation-independent cloning (4). We generated the deletion mutants by PCR, and the 9GA mutant was synthesized by Genscript. We then subcloned the cDNA from DDX23-9GA into pNH-TrxT. For expression of the minimal 37-residue DDX23 region that interacts with SRSF1, we designed the TrxA fusion construct and cloned it into pET-28a (+). DEAD to AAAD mutants were generated by site directed mutagenesis.

For expression of DDX23 in human cells, we cloned the cDNA of FL-DDX23 or the deletion constructs into pDarmo.CMVT_v1 (#133072) (5) using Gibson assembly (NEB). For direct visualization of protein localization in live cells, we cloned mVenus upstream of DDX23.

Generation of cell lines for proximity labeling.

For retrovirus production, we used HEK-293 Ampho Phoenix cells. We seeded 1.5-2$\times10^6$ cells in 6-cm plates in DMEM, and after 18 h we transfected them with 10 μg of the BioID2 vectors (BS, SB, BG) mixed with 10 μL of Lipofectamine 2000 in 500 μL opti-MEM. 18 h after transfection, we replaced the medium, and after 48 h we collected the supernatants containing retrovirus.

The day before retroviral transduction, we seeded 2×10^5 HeLa cells in 6-cm dishes. For transduction, we used a 1/3 dilution of freshly produced retrovirus, supplemented with 4 μg/mL polybrene. We performed the transduction for 24 h, and started puromycin selection 48 h post-transduction, using puromycin at 2 μg/mL.

Sample preparation for MS analysis.

For MS analysis, we used MS-grade reagents and solvents. We performed the streptavidin pulldown as described in Methods. We then washed the paramagnetic beads with 200 μL of 100

mM aqueous ammonium bicarbonate pH 8.4 (ABC). We then reduced the bound proteins by performing an incubation with 20 μL of 3 mM TCEP/100 mM ABC for 20 min at room temperature under constant agitation. We alkylated reduced Cys-residues by adding 20 μL of 100 mM CEMTS/50% acetonitrile and incubating for 20 min at room temperature under constant agitation. After dilution with 260 μL 100 mM ABC, we added 1 μg of sequencing-grade modified porcine trypsin (Promega) to each sample, and we performed the proteolysis overnight at 37 °C under constant agitation. The following day, we quenched the proteolysis by adding 1% v/v MS grade TFA, and we transferred the reaction solutions into clean tubes. We then washed the beads with 100 μL of 50% ACN, and we pooled the flowthroughs with the respective eluates. We then lyophilized the eluates by vacuum centrifugation until the volume was approximately 50 μL, and we added 200 μL of 0.1% TFA to each sample to improve retention on C18. We desalted the peptides by C18 solid-phase extraction using Pierce C18 columns (5 mg loading capacity), and we eluted them using 66% ACN, and then completely lyophilized them. We resuspended the peptide pellets in 30 μL of 5% DMSO/0.1% formate by sonication, and we used 2 μL of each sample for LC/MS.

LC/MS analysis.

We loaded the peptides on a 25 cm x 75 μm ID column packed with Reprosil 1.9 C18 silica particles and resolved them on a 90-min 5-35% acetonitrile gradient in water (0.1% formate) at 200 nL/min. We performed the chromatography using a Thermo EasyLC1200 nano chromatograph. We ionized the eluting peptides by electrospray (2200V) using a 10 μm ID silica emitter (Fossil IonTech) kept at 2200V compared to the mass spectrometer's inlet. We set the mass spectrometer to collect 120,000 resolution precursor scans (m/z 380-2000 Th) every 3 sec, interspersed with fragmentation spectra recorded in data-dependent mode with a dynamic exclusion of 30 sec (10 ppm mass tolerance). We selected precursor ions using the quadrupole for HCD fragmentation at stepped 28,33,38% normalized energy (max injection time 35 ms). We collected spectra in the linear trap at "normal" scan rate (unit resolution), with first mass locked to 100 Th.

We searched the mass spectra using the Mascot scoring function within ProteomeDiscoverer (v.2.4), with the mass tolerance set at 5 ppm for MS1, and 0.5 Da for MS2. We matched the spectra against the UniProt human database and a database of common contaminants (Crap). We set K and N-terminus biotinylation, M-oxidation and N/Q-deamidation as variable modifications. We set CEMTS adducts on C as a fixed modification. We filtered peptide-spectral matches to maintain FDR<1% using Percolator.

We extracted ion chromatograms (XIC) of precursor ions intensities and we integrated them for label-free quantification across samples. We used the respective area as a quantitative metric for peptide relative quantification. We expressed protein quantification as the sum of the XIC area of all peptides assigned to a given protein. Overall, we identified 2879 proteins at FDR<0.01. We omitted from the analysis non-quantified proteins and proteins detected in a blank injection performed before the experiment. To permit ratio calculation, we imputed missing values with a value close to the minimum empirical quantification recorded (used as proxy for noise level). Only proteins with empirical quantification in at least 2 replicates were used for quantification. To assess differential enrichment in the samples, we calculated ratios of the average LFQ intensity for each protein across samples.

We deposited the mass spectrometry proteomics data set to the ProteomeXchange Consortium via the PRIDE partner repository with the dataset identifier PXD048180.

Expression and purification of recombinant proteins.

We used *E. coli* Rosetta cells for expression of SRSF1 constructs. Constructs that included SRSF1's RS domain were co-expressed with CLK1 kinase (3). We initially grew transformed bacterial cultures in LB medium at 37 °C until they reached an optical density (OD) of 0.6 – 0.8. We then transferred the bacterial cultures into a low-temperature incubator (18 °C). Once the cultures

cooled down, we added 0.5 mM IPTG to induce recombinant protein expression, and continued the incubation at 18 °C for 20 h. We then collected the cells by centrifugation at 4000g for 20 min and resuspended them in lysis buffer (25 mM Tris-HCl pH 7.5, 500 mM NaCl, 10 mM NaF, 1× protease inhibitor cocktail). Thereafter, lysates were kept on ice or in a cold room at 4 °C. We lysed the cells by sonication (40% output, 5 s ON, 5 s OFF) on ice. We added 50 units of benzonase nuclease to the extracts, and fractionated them by centrifugation at 27,000g, for 30 min at 4 °C. We recovered the soluble fraction and filtered it through a 0.45 μm cellulose acetate filter. For affinity purification, we incubated the soluble fraction with complete His-tag Purification resin (Roche), for 90 min at 4 °C. We washed the beads with 10 column volumes (CV) of lysis buffer supplemented with 20 mM imidazole, and then eluted the proteins in low-salt lysis buffer (50 mM NaCl) with 300 mM imidazole. We purified the proteins by anion exchange chromatography (Q Sepharose), incubating the IMAC elution pool with Q Sepharose beads equilibrated in 25 mM Tris-HCl pH 7.5, 50 mM NaCl for 12 hs. We washed the beads with equilibration buffer (10 CV) and eluted the bound proteins in a gradient of 100-500 mM NaCl. We pooled the eluted fractions and further purified the proteins by size exclusion chromatography on a Superdex 200 Increase column (Cytiva), in SEC buffer (25 mM Tris-HCl pH 7.5, 500 mM NaCl, 5 mM $MgCl_2$, 1 mM DTT). We selected protein-containing fractions by SDS-PAGE analysis, pooled them, and supplemented them with 10% glycerol for storage at -80 °C. To obtain tag-free SRSF1, we performed an additional step after the IMAC purification, consisting of Ulp1 digestion (IMAC purified Ulp1, 1 mg per 20 mg of protein extract, at 4 °C), followed by a second IMAC purification in which we collected the flowthrough.

For expression of DDX23 constructs we used *E. coli* Rosetta cells grown in Terrific Broth. We grew the cultures at 30 °C until OD = 0.8, and we induced the protein expression with 0.5 mM IPTG at 18 °C for 20 h. We then collected the cells by centrifugation at 4000 g for 20 min, and resuspended them in lysis buffer (50 mM Tris-HCl pH 7.5, 2 M LiCl, 5 mM $MgCl_2$, 1 mM DTT, 1× protease inhibitor cocktail). We continued the purification protocol similarly to the SRSF1 purification, except that we eluted the proteins in high-salt lysis buffer (500 mM NaCl), and we removed the His-tag by TEV protease digestion (IMAC purified TEV_{SH}, 1 mg per 20 mg of protein extract, at 4 °C), for 16 h in digestion buffer (50 mM Tris-HCl pH 7.5, 500 mM NaCl, 5 mM $MgCl_2$, 1 mM DTT). After digestion, we performed an additional IMAC purification, collecting the flowthrough containing the digested proteins. We further purified the digested proteins on a Heparin column and eluted the bound proteins in a 500 mM-1000 mM NaCl gradient. Finally, we purified the proteins by size exclusion chromatography on a Superdex 200 Increase column, in SEC buffer.

For expression of the minimal His-TrxA-DDX23 (S41-R77) we used *E. coli* Rosetta cells grown in LB. We grew the bacterial cultures at 37 °C until OD = 0.8, and then induced protein expression with 0.5 mM IPTG at 18 °C for 20 h. We then collected the cells by centrifugation at 4000 g for 20 min, and resuspended the cells in lysis buffer (25 mM Tris-HCl pH 7.5, 500 mM NaCl, 1× protease inhibitor cocktail). We performed affinity purification similar to the SRSF1 purification, followed by size exclusion chromatography on a Superdex 200 Increase column, in SEC buffer. We selected protein-containing fractions by SDS-PAGE analysis, and we supplemented them with 10% glycerol before storing them at -80°C. For the obtention of the phosphorylated version of this protein, we co-expressed CLK1 and followed an equivalent expression and purification protocol.

Protein extraction from mammalian cells.

We washed cell monolayers with PBS and lysed them in 20 mM Tris HCl pH 8, 200 mM NaCl, 0.1% NP40, 2 mM $MgCl_2$, 100 mM NaF, 2 mM EDTA supplemented with protease inhibitor cocktail. We sonicated the cell lysates in a sonication bath, for 10 cycles of 30s, alternating with 30 s resting time. We then centrifuged the samples for 15 min at 20,000 g, at 4 °C. We then mixed the supernatants with Laemmli buffer and analyzed them by SDS-PAGE/Western blotting. For

detection of proteins, we used the following antibodies: anti Myc Tag (Cell signaling, 9B11), anti HA tag (Thermo Fisher 26183), anti SRSF1 (AK-96, CSHL), and anti α-tubulin (Sigma T9026).

Immunofluorescence and imaging.

For immunofluorescence staining, we seeded HeLa cells in coverslips and incubated them at 37 °C for 24 h. We then fixed the cells using 4% formaldehyde in PBS for 10 min at room temperature. After washing with PBS, we permeabilized the fixed cells with 0.2% Triton X100 in PBS, and after a new set of PBS washes, we blocked the slides for 1 h using 1% BSA, 22.52 mg/mL glycine in PBST. After blocking, we incubated the slides for 1 h with 100 μL of anti-Myc Tag / HA-Tag antibody solution (1:200) (Cell Signaling Technology or Thermo Fisher) diluted in 1% BSA in PBST. We then washed the slides with PBS and incubated them with anti-mouse secondary antibody, labeled with Alexa Fluor 488 (Invitrogen # A32723), at 10 μg/mL. After a 1-h incubation, we washed the slides with PBS, stained with DAPI, and visualized the slides by fluorescence microscopy. We imaged the fluorescence emitted by Alexa Flour 488 (510 nm emission) using the GFP channel, and the emission at 461 nm using the DAPI channel, in an Echo Revolve fluorescence microscope (Echo).

RNA extraction form mammalian cells and RT-PCR for splicing analysis.

We washed cell monolayers with PBS, lysed them using 1 mL Trizol, and collected the samples in fresh tubes. For RNA extraction, we added 200 μL of chloroform and vortexed the samples. Immediately after, we centrifuged the samples at 12,000 g for 15 min at 4 °C. We then transferred the upper aqueous phase into fresh tubes and precipitated the RNA by adding 600 μL of 100% isopropanol. We vortexed the samples and incubated them at -20 °C for 30 min. Then, we centrifuged the samples at 12,000 g for 10 min at 4 °C. We washed the RNA pellets with 70% EtOH, air dried them, and resuspended them in RNAse-free water. We proceeded to cDNA preparation using oligo dT and ImProm-II reverse transcriptase as per the manufacturer's instructions (Promega). We analyzed the alternative splicing pattern of *PEA15* by radioactive RT-PCR, as described (6).

***In vitro* transcription**

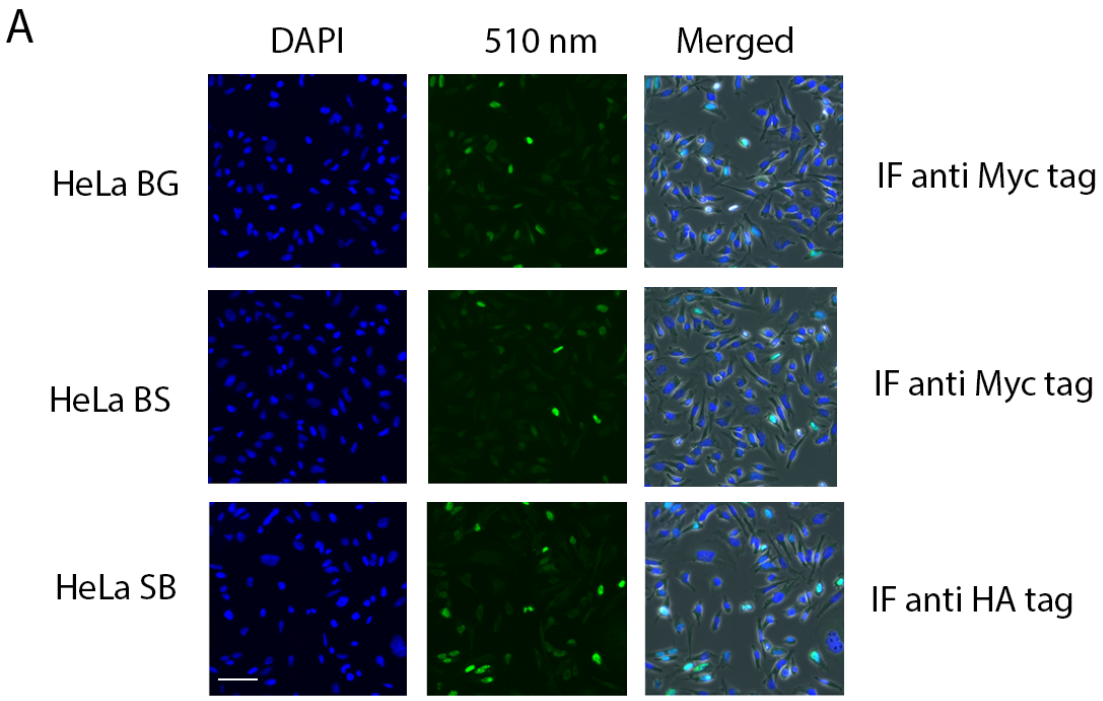
We digested 50 μg of the template DNA (Addgene #11244) with BamHI (NEB) for 3 h at 37 °C. After heat inactivation of the restriction enzyme at 70 °C for 20 min, we treated the digested DNA with 25 units of alkaline phosphatase for 1 h at 37 °C and purified the linear DNA template by agarose gel electrophoresis. For *in vitro* transcription, we used a Promega T7 RNA polymerase kit. Transcription reactions in 20 μL used Transcription Optimized buffer (Promega) supplemented with 10 mM DTT, 0.5 μL RNAsin, 2 μL of rNTP nucleotide mix (5 mM ATP, 5 mM CTP, 1 mM GTP and 0.5 mM UTP), 1 μL of 10 mM G-cap analog, 2.5 μL of [α^{32}P] UTP (10 mCi/μL), 1 μL of T7 RNA polymerase, and 2 μg of DNA template. We incubated the transcription reaction for 2.5 h at 37 °C. Immediately after, we digested the template DNA with RQ1 DNAse in the presence of 1 mM $CaCl_2$ for 30 min at 37°C, and diluted the sample with 25 μL of RNA sample buffer. We purified the RNA by denaturing PAGE and elution overnight in 10 mM Tris-HCl pH 7.4, 1 mM EDTA, 0.5% SDS buffer. Finally, we extracted the RNA using acid phenol chloroform, pH 4.2, followed by ethanol precipitation.

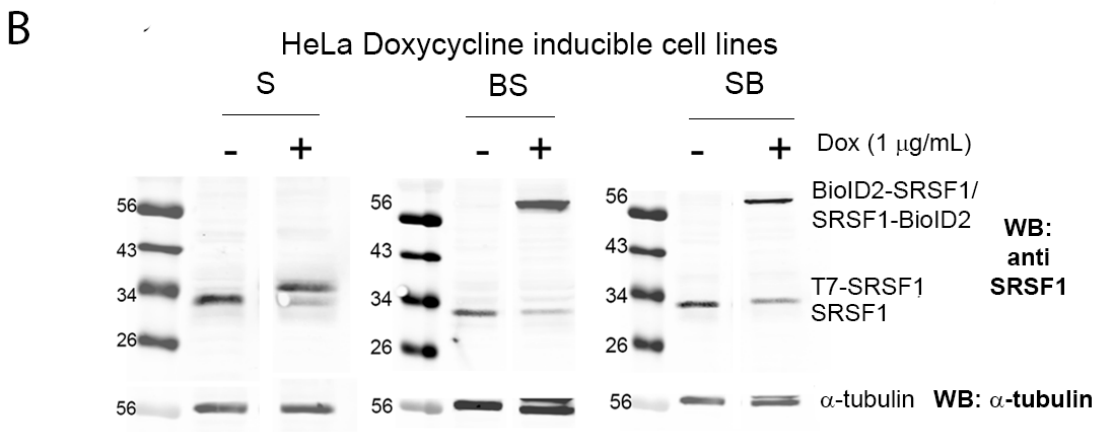
Post-translational modification analysis by MS

We reduced samples of SEC-purified His-TrxA-DDX23 (S41-R77) or His-TrxA-DDX23 (S41-R77) co-expressed with CLK1, with 3 mM TCEP (tris(2-carboxyethyl)) and alkylated them with 10 mM MMTS (methyl methanethiosulfonate). We then performed an overnight digestion with LysC at 37 °C and quenched the reactions with TFA before LC/MS/MS analysis.

We loaded peptides via a 10 cm x 100 µm ID trap column packed with 5 µm aqua C18 particles (Phenomenex) onto a 30 cm x 75 µm ID analytical column packed with Reprosil 1.9 µm C18 silica particles (Dr. Maisch), and resolved them on a 5-35% acetonitrile gradient in water (0.1% formic acid) at 200 nL/min. We ionized eluting peptides by electrospray (2200V) and transferred them to an Orbitrap Fusion Lumos Tribrid mass spectrometer (Thermo). We set the MS to collect 120K resolution precursor scans (m/z 380-2000 Th). We selected precursor ions between 380-900 m/z from HCD fragmentation at stepped 30,35,40% normalized energy and ETD fragmentation with charge-dependent ETD parameters. We selected precursor ions between 900-2000 m/z for HCD fragmentation only at stepped 30,35,40% normalized collision energy. We selected precursors in data-dependent mode for fragmentation one time before excluding them for 30 sec. We collected spectra in the Orbitrap at 30,000 resolution, with the first mass locked to 100 Th.

We searched for peptides using the PEAK's scoring function, with mass tolerance set at 10 ppm for MS1, and 0.02 Da for MS2. We matched the spectra against the E. coli database concatenated with the sequence of His-TrxA-DDX23 and CLK1 kinase, as well as a database of common contaminants (cRAP). We set M-oxidation, N-term/K-acetylation, and STY-phosphorylation as variable modifications. We allowed a maximum of eight variable modifications per peptide. Peptide-spectral matches were filtered to maintain FDR<1%. We integrated extracted ion chromatograms of precursor-ion intensities for label-free quantification across samples.


A
DAPI
510 nm
Merged
HeLa BG
IF anti Myc tag
HeLa BS
IF anti Myc tag
HeLa SB
IF anti HA tag


B
HeLa Doxycycline inducible cell lines
S
BS
SB
- + - + - +
Dox (1 μg/mL)
56
43
34
26
56
BioID2-SRSF1/
SRSF1-BioID2
T7-SRSF1
SRSF1
WB: anti SRSF1
α-tubulin
WB: α-tubulin

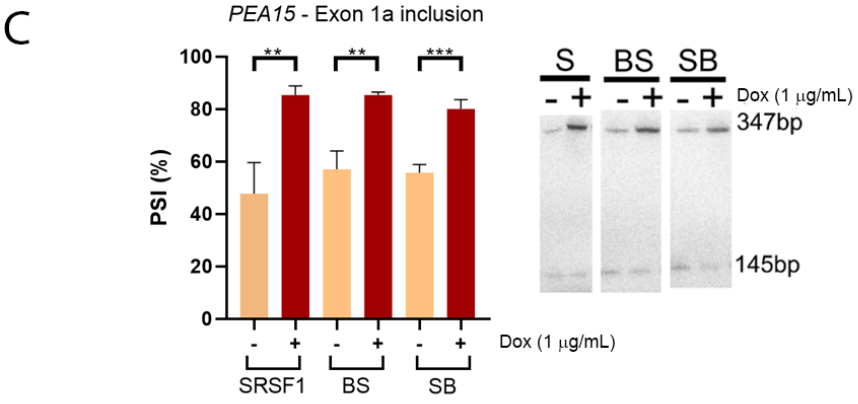
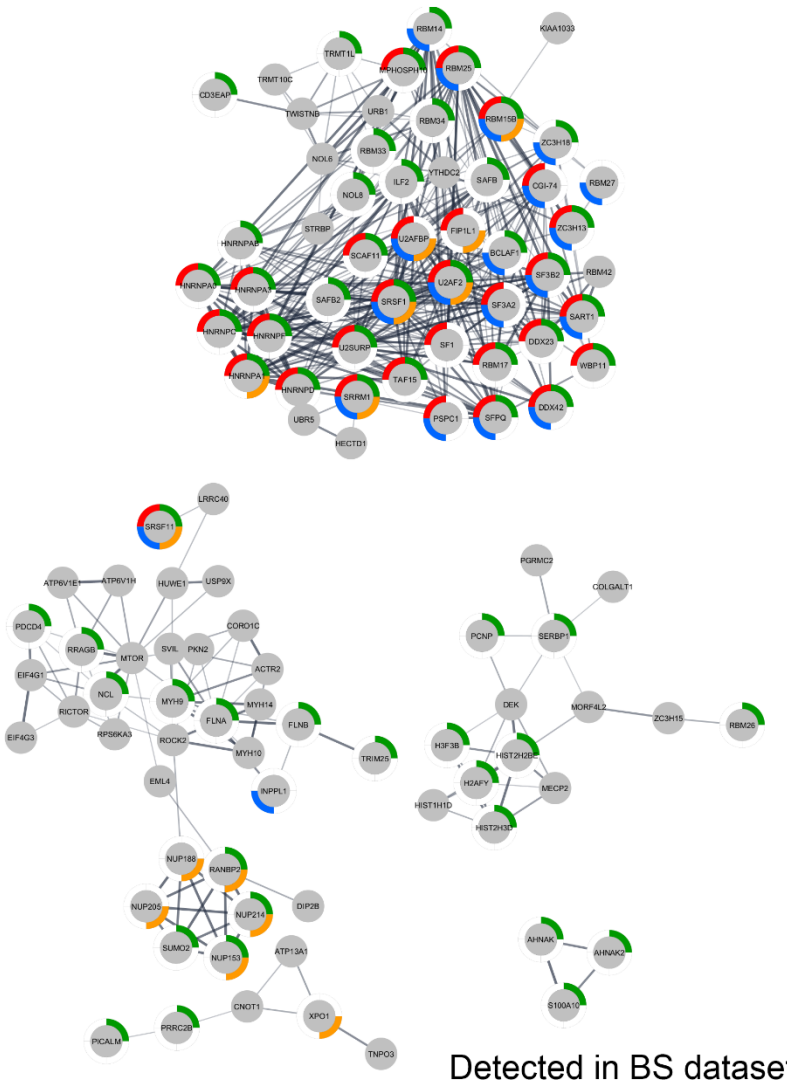
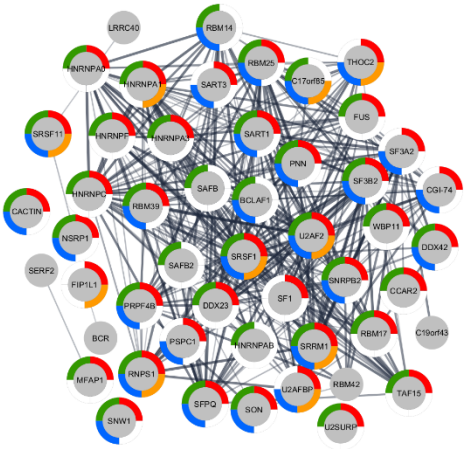

C
PEA15 - Exon 1a inclusion
PSI (%)
100
80
60
40
20
0
**
**
***
- + - + - +
SRSF1
BS
SB
Dox (1 μg/mL)
S
BS
SB
- + - + - +
Dox (1 μg/mL)
347bp
145bp

Fig. S1. BioID2 fusions do not alter SRSF1 localization, SRSF1 autoregulation nor its function in pre mRNA splicing. A. IF of HeLa cell lines used for proximity labeling. Myc or HA antibody was used to detect the BioID2 fusions of BG, BS and SB. DAPI staining shows nuclei, and predominantly nuclear localization can be observed for all BioID2 fusions. Scale bar represents ~30 μm. B. Western blotting analysis of Dox-inducible cell lines overexpressing either T7-SRSF1 (S), BioID2-SRSF1 (BS) or SRSF1-BioID2 (SB). Overexpression of SRSF1 was achieved by adding the indicated amount of Doxycycline, for 72 h. Upon overexpression of SRSF1, regardless of the tag used, we detect a decrease in the expression of endogenous SRSF1. This shows that expression of the fusion proteins regulates endogenous SRSF1 protein levels, and therefore, the fusion proteins engage in SRSF1's autoregulation mechanisms. C. RT-PCR analysis of splicing of *PEA15* exon 1a upon SRSF1 overexpression. An unpaired t-test was performed to compare the Percent Spliced In (PSI) in the presence or absence of Doxycycline (SRSF1 overexpression), in three different biological replicates. ** = $P \leq 0.01$; *** = $P \leq 0.001$. Error bars represent standard deviation. Cell lines expressing the BioID2 fusions can regulate alternative splicing of this endogenous SRSF1 target.



Detected in BS dataset
(from 353 protein list)



RNA splicing
Ubl conjugation
mRNA transport
Nuclear speckles

Detected in SB dataset
(from 246 protein list)

Fig. S2. Protein networks generated with individual datasets, generated from the BioID2-SRSF1 enriched protein list (top), or the SRSF1-BioID2 enriched protein list (bottom).

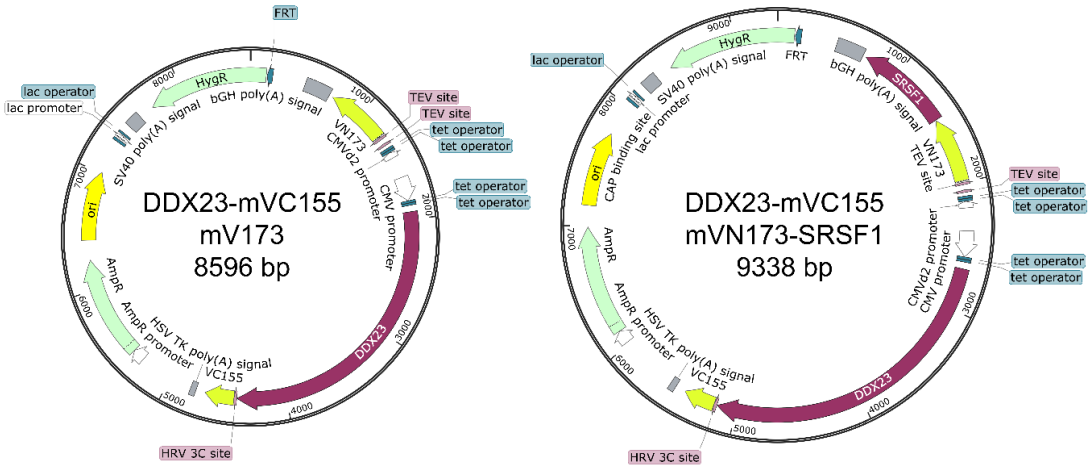
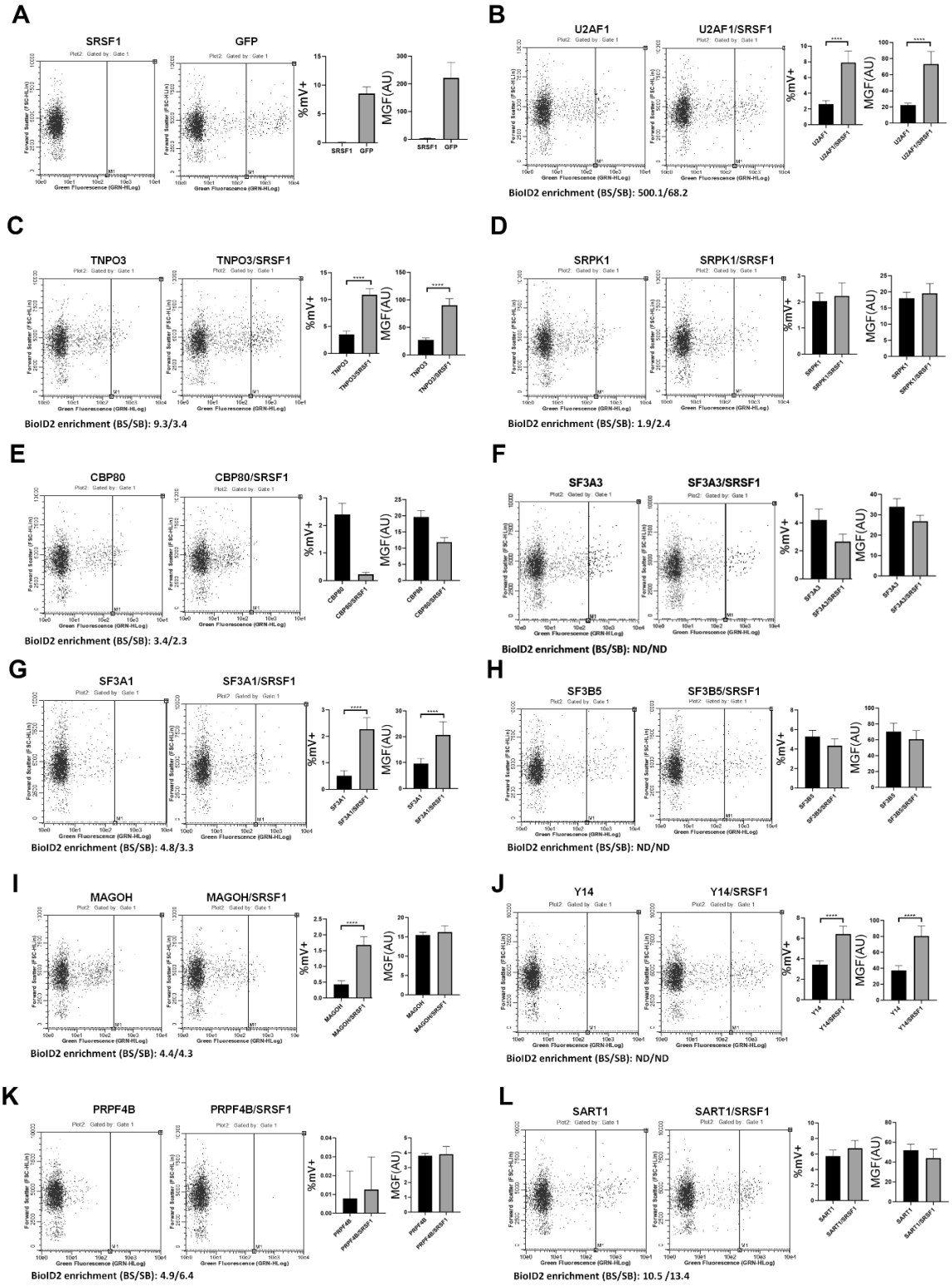


Fig. S3. Schematic representation of BiFC vectors. The fusion genes were cloned upstream of CMV or CMVd2 promoters. Images were generated with Snapgene software.


A
SRSF1
GFP
%mV+
MGF(AU)
B
U2AF1
U2AF1/SRSF1
BioID2 enrichment (BS/SB): 500.1/68.2
C
TNPO3
TNPO3/SRSF1
BioID2 enrichment (BS/SB): 9.3/3.4
D
SRPK1
SRPK1/SRSF1
BioID2 enrichment (BS/SB): 1.9/2.4
E
CBP80
CBP80/SRSF1
BioID2 enrichment (BS/SB): 3.4/2.3
F
SF3A3
SF3A3/SRSF1
BioID2 enrichment (BS/SB): ND/ND
G
SF3A1
SF3A1/SRSF1
BioID2 enrichment (BS/SB): 4.8/3.3
H
SF3B5
SF3B5/SRSF1
BioID2 enrichment (BS/SB): ND/ND
I
MAGOH
MAGOH/SRSF1
BioID2 enrichment (BS/SB): 4.4/4.3
J
Y14
Y14/SRSF1
BioID2 enrichment (BS/SB): ND/ND
K
PRPF4B
PRPF4B/SRSF1
BioID2 enrichment (BS/SB): 4.9/6.4
L
SART1
SART1/SRSF1
BioID2 enrichment (BS/SB): 10.5 /13.4

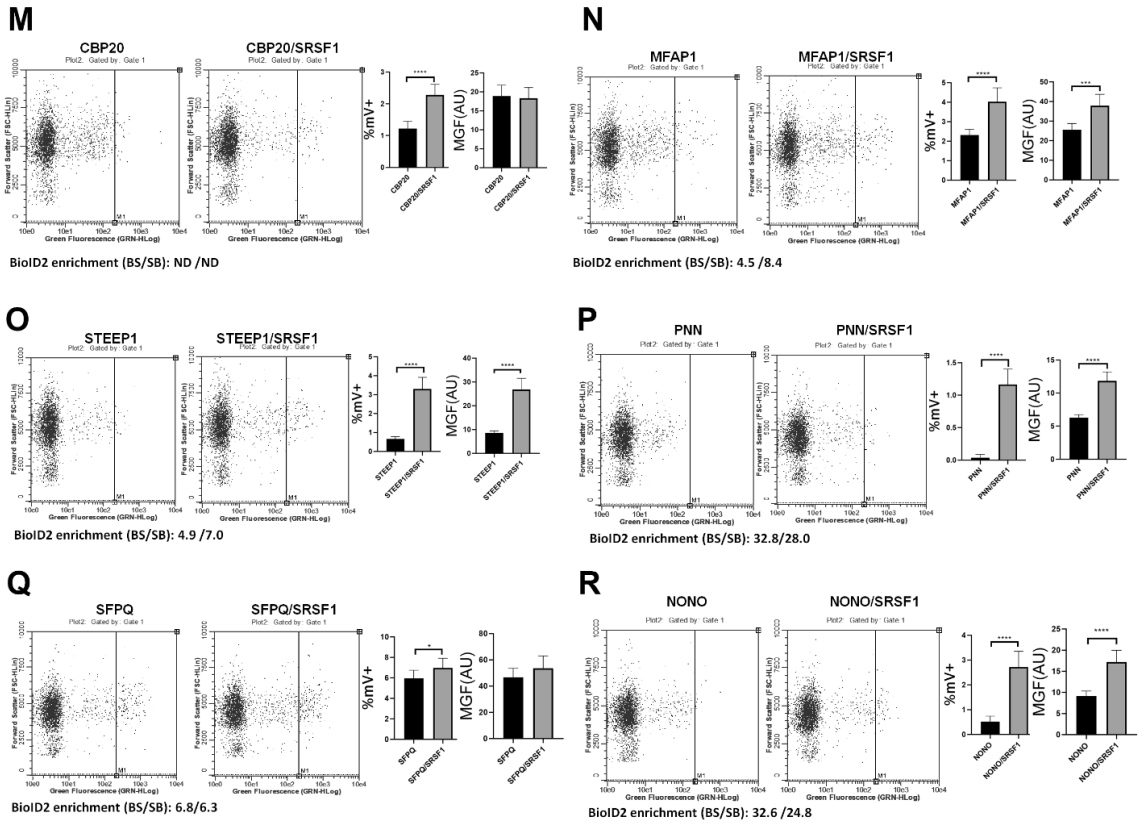


Fig. S4. Bimolecular fluorescence complementation validates some of the interactions detected by proximity labeling. A-R. BiFC results for the indicated proteins, represented as Forward scatter vs. green fluorescence plots. The vertical line indicates the threshold for cells that were considered mVenus positive. The bar graphs on the right of each panel show % mVenus-positive cells (%mV+) and mean green fluorescence (MGF). An unpaired t-test was performed to compare the %mV+ cells and MGF values between the test and control conditions. * = P ≤ 0.05; *** = P ≤ 0.001; **** = P≤ 0.0001. Error bars represent standard deviation. In several cases, no significant differences were observed or the % mVenus-positive cells decreased. Limitations of BiFC discussed in the main manuscript can affect the results observed for those interactions, therefore, we did not perform a statistical analysis on those samples.

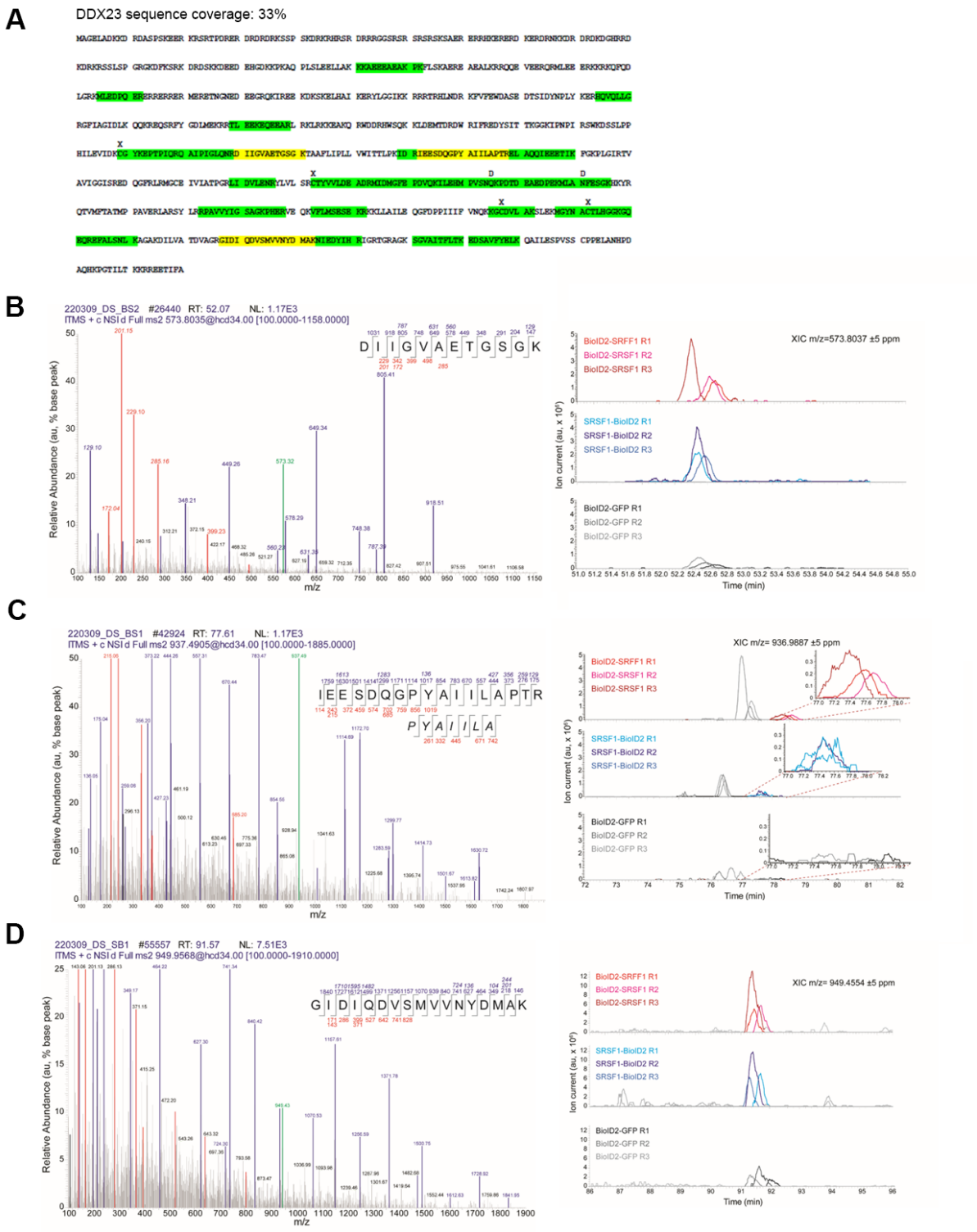


Fig. S5. Detection of DDX23 and label-free quantification. A. DDX23 primary sequence. Detected peptides are highlighted in green and yellow. Manually validated peptides are highlighted in yellow. D = N/Q deamidation, X = CEMTS adducts on C. B-D. Representative fragmentation

spectra of peptides are indicated in yellow. Peptide ions' extracted ion chromatograms (XIC) used for label-free quantification are shown for all constructs and all replicates.

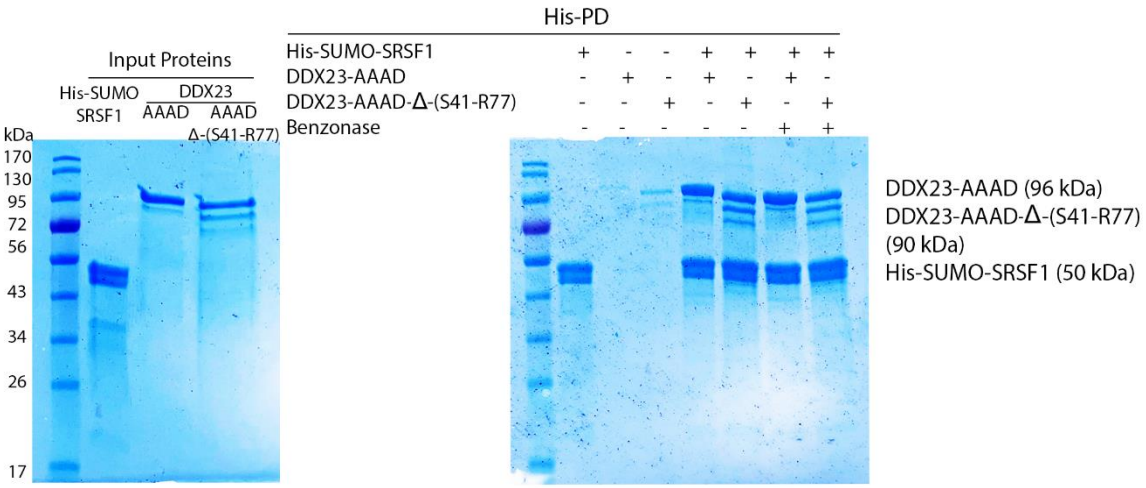


Fig S6. DDX23 inactive mutants (DEAD→AAAD) interact with SRSF1, even in the absence of the S41-R77 region. Representative His Pulldown assay using His-SUMO-SRSF1-RRMs purified from *E. coli* and untagged versions of mutant DDX23, as indicated in the figure.

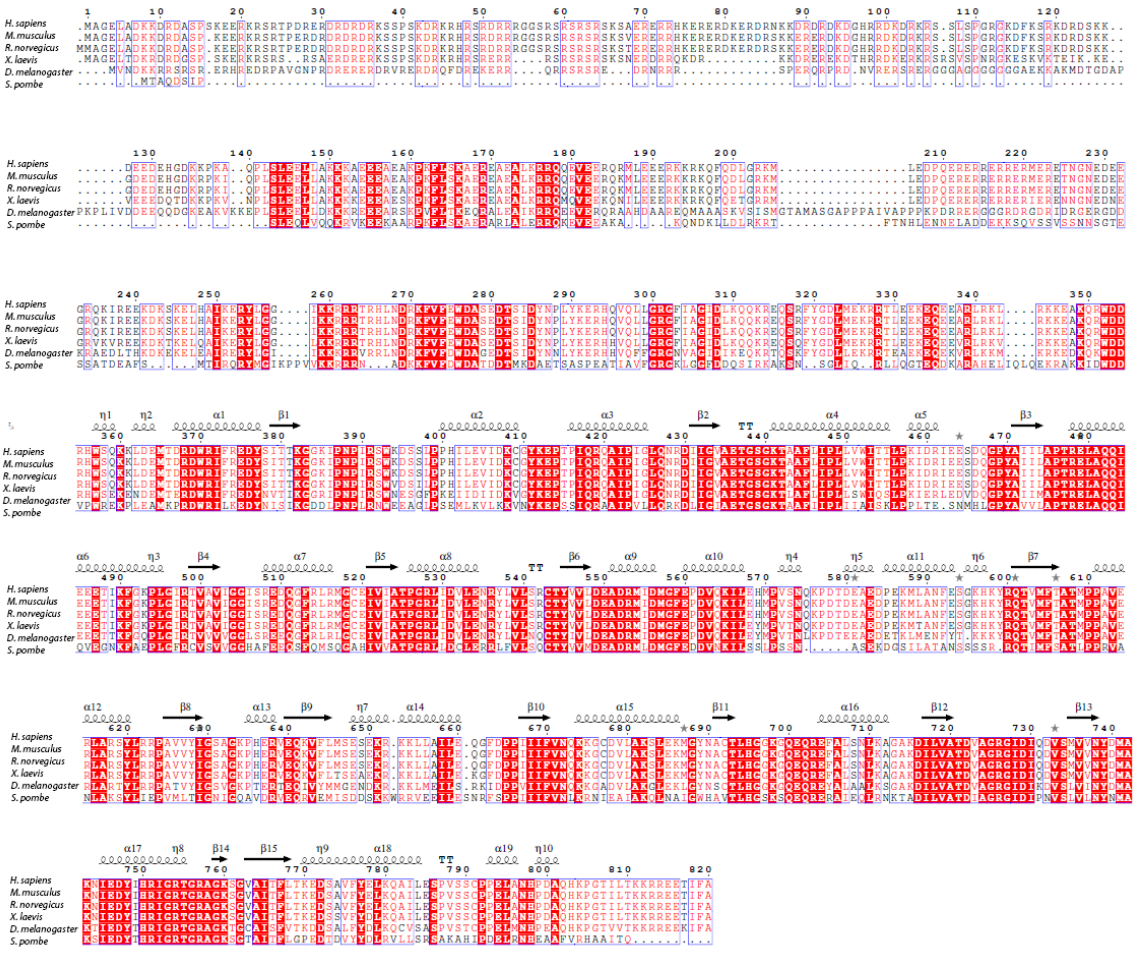


Fig S7. Sequence alignment of DDX23 and its orthologs. Secondary structures based on the crystal structure of human DDX23 (PDB: 4NHO) are depicted on top of the alignment.

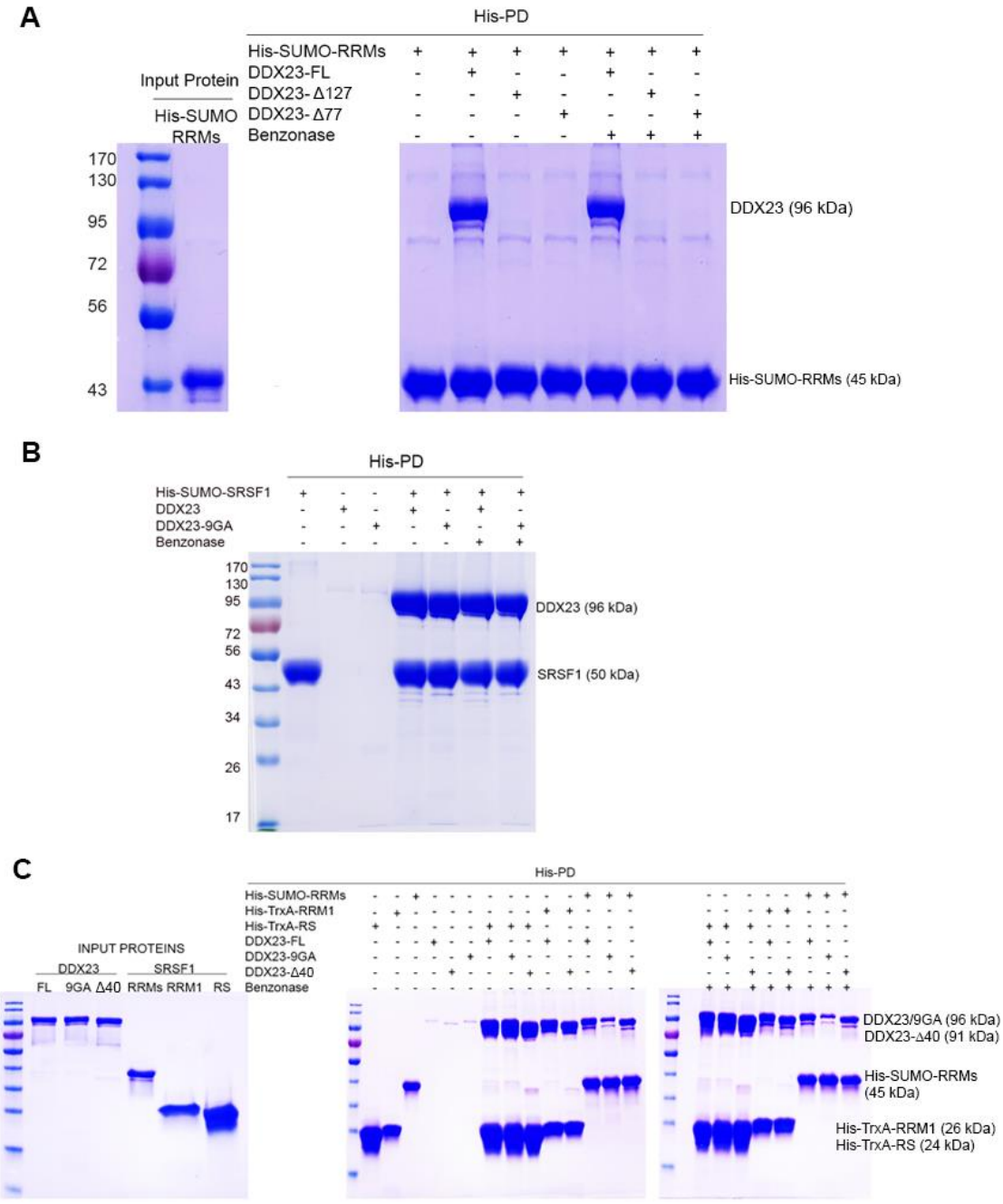


Fig. S8. DDX23 interacts with RRMs. The 9GA mutation On DDX23 does not affect its interaction with FL-SRSF1 but affects the interaction with RRM1. A. Representative His Pulldown assay using His-SUMO-SRSF1-RRMs purified from *E. coli* and untagged DDX23, or the indicated deletion mutants. B. Representative His Pulldown assay using His-SUMO-SRSF1 purified from *E. coli* (phosphorylated by co-expressed CLK1 kinase) and untagged DDX23, WT or 9GA-mutant, as indicated. A positive interaction can be observed for FL-SRSF1 and DDX23, irrespective of the presence of the 9GA mutation. C. Pulldown assays that demonstrate that the SRSF1 RS domain (phosphorylated by CLK1) binds to FL-DDX23 and DDX23-Δ40. In addition, RRM1 can also bind

to both FL-DDX23 and DDX23-Δ40. Decreased binding between His-SUMO-RRMs and the -9GA mutant is observed.

Fig. S9. Post-translational modifications (PTMs) detected by MS in His-TrxA-DDX23 chimeric protein co-expressed with CLK1 kinase in *E. coli*. Blue lines represent peptides identified, and colored letters represent PTMs.

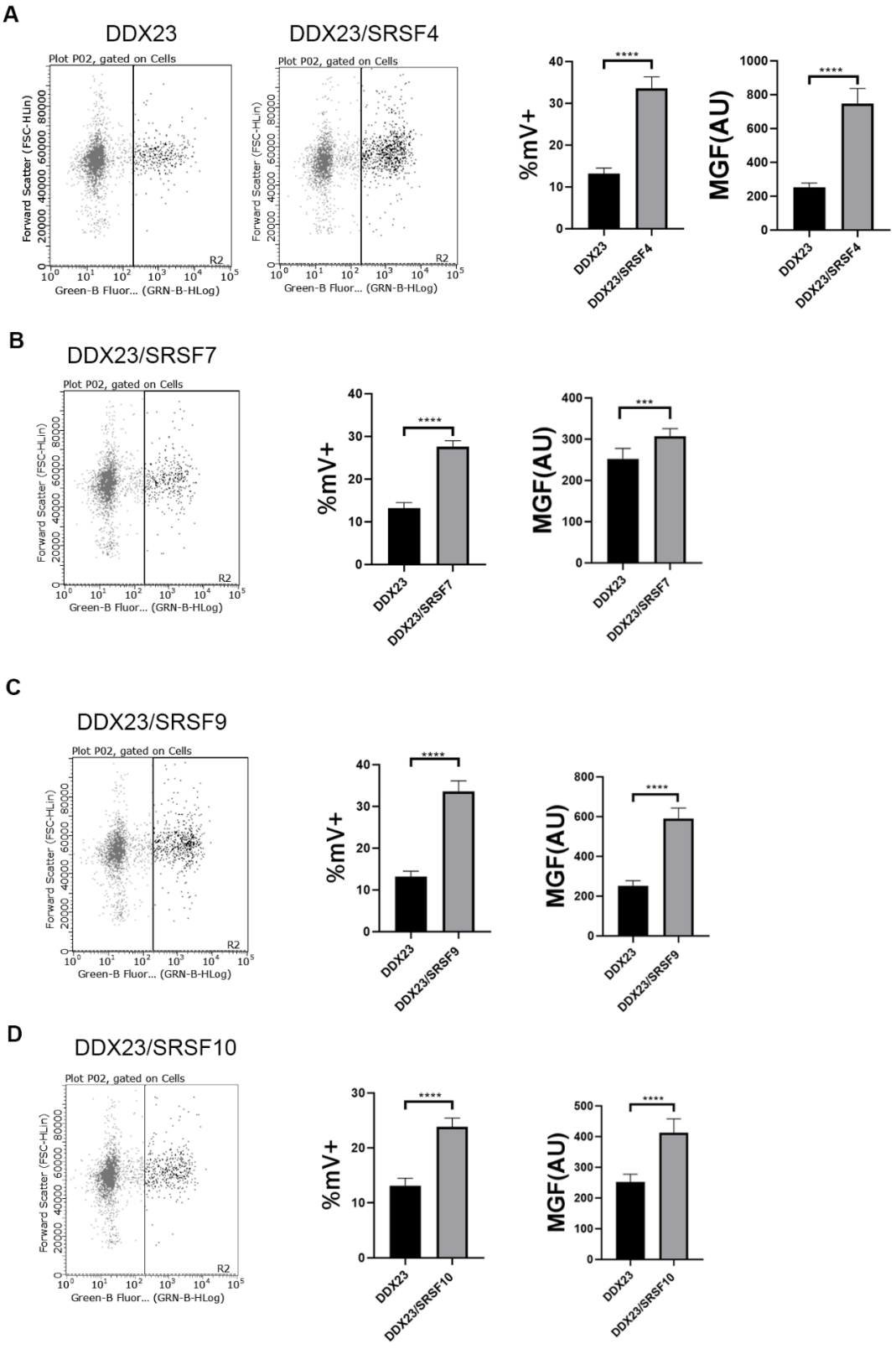

A
DDX23
DDX23/SRSF4
Plot P02, gated on Cells
Forward Scatter (FSC-HLin)
Green-B Fluor... (GRN-B-HLog)
R2
%mV+
MGF(AU)
****
DDX23
DDX23/SRSF4
B
DDX23/SRSF7
***
DDX23/SRSF7
C
DDX23/SRSF9
DDX23/SRSF9
D
DDX23/SRSF10
DDX23/SRSF10

Fig. S10. DDX23 interacts with additional SR proteins in BiFC assays. A-D. BiFC results for the indicated proteins, represented as Forward scatter vs. green fluorescence plots. Vertical line indicates the threshold for cells that were considered mVenus-positive. Bar graphs next to each panel represent % mVenus-positive cells (%mV+) and mean green fluorescence (MGF). An unpaired t-test was performed to compare the %mV+ cells and MGF values between the test and control conditions *** = P ≤ 0.001; **** = P≤ 0.0001. Error bars represent standard deviation.

Table S1. Complete list of genes encoding proteins that were enriched ≥5-fold in the proximity-labeling-derived datasets. Comparisons were made between BS and SB vs. the BG control. Values represent enrichment over control (fold change).

	Gene	BS/BG		Gene	BS/BG		Gene	BS/BG
1	*SRSF1*	4547.1	36	*KRT72*	29.5	71	*BCR*	16.5
2	*HAUS2*	950.7	37	*CDSN*	29.4	72	*ELAC2*	16.0
3	*AHNAK2*	796.2	38	*NCBP3*	29.1	73	*CLTC*	16.0
4	*U2AF1*	500.1	39	*TCERG1*	28.6	74	*SAFB2*	15.8
5	*MACROH2A1*	490.9	40	*KNOP1*	27.2	75	*TGM2*	15.6
6	*H2BU1*	444.2	41	*PPHLN1*	26.3	76	*DSG1*	15.6
7	*HMGXB4*	397.8	42	*SAFB*	26.1	77	*COLGALT1*	15.2
8	*RABL6*	374.8	43	*ATP5IF1*	24.8	78	*SRSF11*	15.0
9	*SERF2*	287.7	44	*H3-3A*	24.4	79	*PCF11*	14.6
10	*DNAJC8*	238.3	45	*PRRC2B*	23.7	80	*SAMD9*	14.5
11	*FIP1L1*	226.9	46	*RAB3GAP2*	23.0	81	*DIP2B*	14.5
12	*RASSF7*	213.8	47	*LMO7*	22.4	82	*CNOT1*	14.2
13	*TAF15*	176.3	48	*ITPR1*	22.0	83	*MTDH*	14.1
14	*BCLAF1*	143.7	49	*RP9*	21.5	84	*RBBP6*	13.8
15	*HNRNPA1*	140.4	50	*PC*	21.0	85	*EDF1*	13.8
16	*CACTIN*	127.9	51	*CCDC124*	20.8	86	*SF3B2*	13.4
17	*SF3A2*	125.6	52	*ITPR3*	20.6	87	*GBF1*	13.3
18	*POLR1G*	122.2	53	*SUB1*	20.4	88	*IGHG1*	12.9
19	*CERS2*	94.2	54	*ARHGEF1*	20.4	89	*ALB*	12.9
20	*SERBP1*	88.0	55	*CAVIN1*	20.2	90	*U2SURP*	12.8
21	*DVL1*	83.9	56	*SRRM1*	20.2	91	*SCAF11*	12.8
22	*LYZ*	70.6	57	*USP9X*	20.1	92	*MCCC1*	12.3
23	*RBM17*	59.1	58	*ATP6V1E1*	19.3	93	*BLVRA*	12.1
24	*PCNP*	51.6	59	*RNPS1*	19.2	94	*CCAR2*	12.0
25	*HNRNPA3*	50.0	60	*HNRNPAB*	19.1	95	*CCAR1*	11.9
26	*HACD2*	47.6	61	*MAP4K4*	18.9	96	*MINK1*	11.7
27	*HNRNPA2B1*	40.1	62	*CEP170*	18.6	97	*SRSF3*	11.6
28	*FUS*	38.7	63	*MPHOSPH10*	18.6	98	*RANBP2*	11.6
29	*RBM27*	35.5	64	*NCKAP1*	18.5	99	*ELP1*	11.4
30	*PSPC1*	33.8	65	*H3C15*	18.4	100	*KRT78*	11.3
31	*TCOF1*	32.8	66	*OAS3*	18.1	101	*DNAJC9*	11.2
32	*PNN*	32.8	67	*MTOR*	17.3	102	*SMG8*	11.2
33	*NONO*	32.7	68	*ARMCX2*	17.2	103	*RBM15B*	11.0
34	*PFKL*	30.9	69	*NOL8*	17.1	104	*S100A10*	10.9
35	*ZC3H18*	29.6	70	*RBM25*	16.7	105	*ACACA*	10.9

	Gene	BS/BG		Gene	BS/BG		Gene	BS/BG
106	*HNRNPA0*	10.8	142	*JUP*	8.7	178	*HMGB1*	7.5
107	*SART3*	10.7	143	*AUP1*	8.7	179	*HMGB2*	7.5
108	*EML4*	10.6	144	*URB1*	8.6	180	*STATH*	7.5
109	*TXN*	10.6	145	*EBP*	8.6	181	*RPS6KA3*	7.4
110	*SART1*	10.6	146	*FLNA*	8.6	182	*MACF1*	7.4
111	*PCCA*	10.5	147	*MAP1S*	8.6	183	*LMAN2*	7.4
112	*MYH14*	10.4	148	*TMED10*	8.5	184	*RBM14*	7.4
113	*SREK1IP1*	10.4	149	*PINX1*	8.5	185	*COPB1*	7.3
114	*RICTOR*	10.3	150	*HNRNPF*	8.5	186	*NKD2*	7.3
115	*RBM42*	10.3	151	*SSRP1*	8.5	187	*EPB41L1*	7.3
116	*KTN1*	10.2	152	*XPO7*	8.4	188	*ACO1*	7.2
117	*PDCD4*	10.2	153	*ZNF281*	8.4	189	*ASCC3*	7.2
118	*ATP2B4*	10.2	154	*NUP188*	8.3	190	*U2AF2*	7.2
119	*RBM39*	10.1	155	*LDHAL6B*	8.3	191	*HBS1L*	7.2
120	*RBM33*	10.0	156	*MTRR*	8.3	192	*GCN1*	7.2
121	*SBF1*	9.9	157	*HMGN4*	8.3	193	*EIF4G3*	7.1
122	*DPP3*	9.9	158	*AP2B1*	8.3	194	*TRMT1L*	7.1
123	*DDB1*	9.8	159	*TRABD*	8.2	195	*AHNAK*	7.1
124	*CYCS*	9.8	160	*LRRC40*	8.2	196	*ARMCX3*	7.0
125	*PKN2*	9.7	161	*SKIV2L*	8.2	197	*HNRNPC*	7.0
126	*ROCK2*	9.7	162	*TTC37*	8.1	198	*SF1*	7.0
127	*SNW1*	9.6	163	*SUMO2*	8.1	199	*ARHGAP18*	6.9
128	*MOGS*	9.5	164	*CARMIL1*	8.1	200	*ATP13A1*	6.9
129	*HUWE1*	9.4	165	*CAPN1*	8.1	201	*TRIM16*	6.9
130	*TNPO3*	9.3	166	*TSN*	8.0	202	*CCNT1*	6.9
131	*NSRP1*	9.3	167	*PICALM*	8.0	203	*POLE*	6.9
132	*MYH10*	9.2	168	*ECPAS*	7.9	204	*ILK*	6.9
133	*PPIG*	9.2	169	*VAPA*	7.9	205	*SFPQ*	6.8
134	*MECP2*	9.1	170	*DYNC1H1*	7.8	206	*RBM26*	6.8
135	*TRIR*	9.1	171	*YTHDC2*	7.8	207	*PLEC*	6.8
136	*SYNJ2*	9.0	172	*SUN1*	7.7	208	*LUC7L2*	6.8
137	*EMC7*	8.9	173	*EHD4*	7.7	209	*PTPRF*	6.8
138	*DDX42*	8.8	174	*CD151*	7.6	210	*LARP7*	6.8
139	*LYAR*	8.8	175	*FAR1*	7.6	211	*DEK*	6.7
140	*INF2*	8.8	176	*TMA7*	7.6	212	*S100A16*	6.7
141	*GBP1*	8.7	177	*TAF3*	7.6	213	*SPTBN1*	6.7

	Gene	BS/BG		Gene	BS/BG		Gene	BS/BG
214	*CLTC*	6.7	249	*SGPL1*	6.1	284	*NUP153*	5.7
215	*AP2A1*	6.7	250	*RHOG*	6.1	285	*H3BN98*	5.7
216	*NDUFB5*	6.6	251	*TRMT10C*	6.1	286	*CSNK2A1*	5.7
217	*PTPA*	6.6	252	*EHD2*	6.1	287	*ARL6IP5*	5.7
218	*HECTD1*	6.5	253	*SACM1L*	6.1	288	*NOL6*	5.7
219	*DSP*	6.5	254	*PRMT5*	6.1	289	*CCN1*	5.6
220	*EIF5B*	6.5	255	*SEC24C*	6.0	290	*CYC1*	5.6
221	*CARS1*	6.5	256	*FAM98A*	6.0	291	*XPO1*	5.6
222	*SVIL*	6.4	257	*GEMIN5*	6.0	292	*HNRNPD*	5.6
223	*STRBP*	6.4	258	*SQOR*	6.0	293	*GHITM*	5.6
224	*SSR1*	6.4	259	*AP3B1*	5.9	294	*STXBP1*	5.6
225	*UBR5*	6.4	260	*H1-10*	5.9	295	*NELFA*	5.6
226	*TRIM25*	6.4	261	*FANCI*	5.9	296	*UNC13D*	5.6
227	*INPPL1*	6.4	262	*MORF4L2*	5.9	297	*PGRMC2*	5.6
228	*MRPL1*	6.3	263	*RRAGB*	5.9	298	*UBE2M*	5.5
229	*MAP2K3*	6.3	264	*HSPB1*	5.9	299	*EXOC4*	5.5
230	*ATL3*	6.3	265	*NCAPG*	5.9	300	*POLR2B*	5.5
231	*ARHGAP1*	6.3	266	*TRIO*	5.9	301	*LPCAT1*	5.5
232	*RBM34*	6.3	267	*DHRS7B*	5.9	302	*RPN2*	5.5
233	*KMT2A*	6.3	268	*ATP6V1H*	5.9	303	*TOMM70*	5.5
234	*EPB41L2*	6.3	269	*PREB*	5.9	304	*PGM1*	5.5
235	*FADS2*	6.3	270	*ERGIC1*	5.8	305	*NUP205*	5.5
236	*ZC3H13*	6.3	271	*PRKCA*	5.8	306	*FLNB*	5.5
237	*SON*	6.2	272	*SCAMP3*	5.8	307	*POLR1F*	5.5
238	*FLNA*	6.2	273	*ILVBL*	5.8	308	*ACTR2*	5.5
239	*DHCR7*	6.2	274	*VTI1B*	5.8	309	*PODXL*	5.4
240	*NAA25*	6.2	275	*DHCR24*	5.8	310	*CXCR4*	5.4
241	*PRIM2*	6.2	276	*HMGA1*	5.8	311	*C7orf50*	5.4
242	*HS2ST1*	6.2	277	*CMBL*	5.7	312	*PLEC*	5.4
243	*MYH9*	6.2	278	*B4GALT1*	5.7	313	*MAP1B*	5.4
244	*UGT8*	6.1	279	*OSBPL8*	5.7	**314**	***DDX23***	**5.4**
245	*ZNF800*	6.1	280	*NRDC*	5.7	315	*COPS3*	5.4
246	*PPFIA1*	6.1	281	*DNAJB1*	5.7	316	*FASN*	5.4
247	*PLAA*	6.1	282	*ZC3H15*	5.7	317	*CAPN2*	5.4
248	*UPF2*	6.1	283	*EIF4G1*	5.7	318	*NSDHL*	5.3

	Gene	BS/BG		Gene	SB/BG		Gene	SB/BG
319	*SLC2A1*	5.3	1	*SRSF1*	5165.3	36	*RBM17*	49.2
320	*ILF2*	5.3	2	*MACROH2A1*	1276.6	37	*NOL8*	48.1
321	*DPYSL3*	5.3	3	*HAUS2*	849.2	38	*HMGXB4*	41.2
322	*RNF213*	5.3	4	*H2BU1*	741.7	39	*RABL6*	40.9
323	*CAMK2G*	5.3	5	*CINP*	472.5	40	*TCERG1*	39.5
324	*PSMC1*	5.3	6	*PPHLN1*	471.4	41	*KRT72*	35.2
325	*DST*	5.2	7	*LORICRIN*	334.1	42	*FUS*	31.7
326	*SCYL1*	5.2	8	*FIP1L1*	297.5	43	*DSG1*	30.1
327	*NCL*	5.2	9	*BCLAF1*	292.2	44	*SAFB2*	29.5
328	*DNTTIP2*	5.2	10	*NCBP3*	287.9	45	*TRIR*	29.5
329	*CORO1C*	5.2	11	*SERBP1*	255.6	46	*PNN*	28.1
330	*CDC45*	5.2	12	*AHNAK2*	217.0	47	*SRRM1*	27.7
331	*CTNNA1*	5.2	13	*TAF15*	205.2	48	*PCF11*	27.1
332	*WASHC4*	5.2	14	*SERF2*	196.1	49	*PCNP*	25.9
333	*WBP11*	5.2	15	*DVL1*	190.4	50	*DNAJC19*	25.4
334	*NUP214*	5.2	16	*HNRNPA1*	181.6	51	*ARMCX2*	25.2
335	*LPGAT1*	5.1	17	*CACTIN*	175.2	52	*NONO*	24.8
336	*GTF3C4*	5.1	18	*RASSF7*	169.8	53	*MAP4K4*	24.8
337	*COPB2*	5.1	19	*SF3A2*	156.5	54	*ALB*	24.8
338	*ACSL4*	5.1	20	*ATP5IF1*	133.5	55	*RBM15B*	23.8
339	*NFIB*	5.1	21	*DNAJC8*	110.3	56	*EDF1*	23.3
340	*SEC63*	5.1	22	*LYZ*	97.2	57	*PIP*	23.0
341	*SRP54*	5.1	23	*RBM27*	92.7	58	*ARHGEF1*	22.5
342	*KYNU*	5.1	24	*SAFB*	79.6	59	*ATP6V1E1*	22.4
343	*TMPO*	5.1	25	*RP9*	79.4	60	*IGHG1*	22.3
344	*CYFIP1*	5.1	26	*CDSN*	72.2	61	*COLGALT1*	22.3
345	*IPO9*	5.1	27	*U2AF1*	68.2	62	*OAS3*	22.3
346	*VAT1*	5.0	28	*RNPS1*	66.6	63	*PSPC1*	21.0
347	*ITGB1*	5.0	29	*HNRNPA3*	64.0	64	*TXN*	19.0
348	*UFL1*	5.0	30	*RBM25*	61.8	65	*RBBP6*	18.4
349	*H1-3*	5.0	31	*SF3B2*	60.0	66	*SNW1*	18.2
350	*KRT85*	5.0	32	*ZC3H18*	57.5	67	*SRSF11*	17.9
351	*CCDC47*	5.0	33	*H3C15*	54.3	68	*SUB1*	17.4
352	*HSD17B12*	5.0	34	*AZGP1*	53.1	69	*DDX42*	16.3
353	*TARS1*	5.0	35	*HNRNPA2B1*	51.3	70	*DCD*	16.1

	Gene	SB/BG		Gene	SB/BG		Gene	SB/BG
71	*PRRC2B*	15.8	106	*SF1*	11.4	141	*HNRNPF*	7.8
72	*U2SURP*	15.7	107	*GIMAP7*	11.4	142	*JUP*	7.7
73	*CSN1S2*	15.6	108	*ACOX3*	11.2	143	*S100A16*	7.7
74	*LMO7*	15.3	109	*KNOP1*	11.2	144	*ELP1*	7.6
75	*TCOF1*	15.1	110	*SRP54*	10.8	145	*PINX1*	7.6
76	*HACD2*	14.9	111	*SBF1*	10.8	146	*ELAC2*	7.5
77	*NSRP1*	14.7	112	*CACYBP*	10.7	147	*MELK*	7.4
78	*SRSF3*	14.6	113	*DSP*	10.7	148	*RAB3GAP2*	7.4
79	*ARMCX3*	13.7	114	*PC*	10.6	149	*MINK1*	7.4
80	*DNAH8*	13.6	115	*MOGS*	10.6	150	*SCAF1*	7.4
81	*ITPR1*	13.5	116	*TRABD*	10.4	151	*CARMIL1*	7.4
82	*SAMD9*	13.4	117	*PRIM2*	10.4	152	*SART3*	7.4
83	*SART1*	13.4	118	*ZFC3H1*	10.2	153	*WBP11*	7.3
84	*CCAR1*	13.4	119	*SON*	10.1	154	*CDK12*	7.3
85	*HNRNPAB*	13.2	120	*SREK1IP1*	9.8	155	*NCL*	7.1
86	*KRT78*	13.2	121	*CEP170*	9.8	156	*HSPA4L*	7.0
87	*SYNJ2*	13.1	122	*BLVRA*	9.6	157	*STEEP1*	7.0
88	*HNRNPA0*	12.9	123	*UFL1*	9.5	158	*MYH14*	7.0
89	*U2AF2*	12.9	124	*UGT8*	9.5	159	*ABCD1*	6.9
90	*CSN3*	12.7	125	*SMG8*	9.3	160	*TRMT1L*	6.9
91	*NCKAP1*	12.7	126	*UBE2M*	9.0	161	*DPP3*	6.9
92	*RBM33*	12.5	127	*SUN1*	8.9	162	*PCCA*	6.8
93	*CCDC124*	12.5	128	*DSC1*	8.9	163	*VAPA*	6.8
94	*CCAR2*	12.4	129	*RNPEP*	8.8	164	*RHOG*	6.7
95	*CSN2*	12.3	130	*DDB1*	8.6	165	*SVIL*	6.7
96	*LRRC40*	12.3	131	*HRNR*	8.6	166	*POLE*	6.6
97	*PPIG*	12.2	132	*MPHOSPH10*	8.5	167	*CSN1S1*	6.6
98	*CALML5*	12.1	133	*MFAP1*	8.4	168	*SMCHD1*	6.5
99	*RBM26*	12.0	134	*POLR1G*	8.3	169	*FANCI*	6.5
100	*BCR*	12.0	135	*EMC7*	8.3	170	*GNPAT*	6.5
101	*SCAF11*	11.9	136	*YTHDC2*	8.2	171	*CAVIN1*	6.5
102	*RICTOR*	11.9	137	*SQOR*	8.0	172	*PRPF4B*	6.5
103	*RBM39*	11.8	138	*ITPR3*	7.9	173	*JPH2*	6.4
104	*GOT1*	11.6	139	*EPB41L1*	7.9	174	*DVL3*	6.4
105	*NR2F2*	11.4	140	*FLNA*	7.9	175	*DIP2B*	6.4

	Gene	SB/BG		Gene	SB/BG
176	*SFPQ*	6.4	211	*ZC3H13*	5.6
177	*DHRS7B*	6.3	212	*PLCB3*	5.6
178	*THOC2*	6.3	213	*LUC7L2*	5.6
179	*NOL6*	6.3	214	*ACO1*	5.6
180	*NELFA*	6.3	215	*HNRNPC*	5.6
181	*PREB*	6.3	216	*RBM42*	5.5
182	*DPYSL3*	6.3	217	*ARHGAP5*	5.5
183	*CARS1*	6.2	218	*GBP1*	5.5
184	*KRT9*	6.2	219	*MAGT1*	5.5
185	*PPIP5K2*	6.2	220	*LMAN2*	5.4
186	*MAP2K3*	6.2	221	*NAA25*	5.4
187	*POLR2B*	6.1	222	*FLG2*	5.4
188	*UPF2*	6.1	223	*EIF4G1*	5.4
189	*INF2*	6.1	224	*SMC1A*	5.3
190	*MCCC1*	6.1	225	*GEMIN5*	5.3
191	*MICU2*	6.0	226	*HMGN4*	5.3
192	*SNRPB2*	6.0	**227**	***DDX23***	**5.3**
193	*ZNF638*	6.0	228	*NKD2*	5.3
194	*HLTF*	5.9	229	*HS2ST1*	5.3
195	*CMBL*	5.9	230	*AP2A1*	5.3
196	*TRIM25*	5.9	231	*TRAM1*	5.3
197	*CCNT1*	5..9	232	*PTPRF*	5.3
198	*KRT1*	5..9	233	*POLR1A*	5.2
199	*STXBP1*	5..9	234	*CCDC93*	5.2
200	*PRKCA*	5.9	235	*STRBP*	5.2
201	*INPPL1*	5.8	236	*XPC*	5.2
202	*PLAA*	5.8	237	*FLNA*	5.2
203	*KRT14*	5.8	238	*HBS1L*	5.2
204	*SCAMP3*	5.8	239	*TGM2*	5.2
205	*SGPL1*	5.7	240	*CYCS*	5.1
206	*FAM111B*	5.7	241	*KRT10*	5.1
207	*PKN2*	5.7	242	*MTRR*	5.1
208	*MAP1S*	5.7	243	*PPFIA1*	5.1
209	*RBM14*	5.7	244	*AGPS*	5.1
210	*CERS2*	5.6	245	*KAT7*	5.1
			246	*NKRF*	5.0

Dataset S1 (separate file).

Excel file containing quantitative proteomics results from Proteome Discoverer, and data analysis (detailed in the "Protocol" sheet, along with the description of each sheet). The following values are reported for each identified protein: False Discovery Rate (FDR) (High or Medium),UniProt accession number, a brief description of the protein, the name of the gene encoding the protein, the experimental q-value, Sum Pep score (protein score calculated as the negative log of the Posterior error probability (PEP) values of connected Peptide Spectrum Matches (PSMs)), coverage percentage (the percent calculated by dividing the number of amino acids in all found peptides by the total number of amino acids in the entire protein sequence), number of peptides, number of PSMs, number of unique peptides, number of amino acids, molecular weight, calculated isoelectric point (pI), Mascot score, number of peptides identified by Mascot and number of razor peptides (a peptide that has been assigned to the protein group with the largest number of total peptide identified). For each protein we also report the label free quantification (AUC) values recorded in two instrument blank runs (no injection) and in the different experimental replicates (BS1, BS2, BS3, SB1, SB2, SB3, BG1, BG2, BG3). 161 proteins with no LFQ value were removed (these proteins are not quantified because of their low abundance; therefore, they are unlikely to be enriched) and 12 proteins detected in the blank with intensity >1% than in average samples were removed (Filtering sheet). In addition, ratios between the BS or SB condition vs the BG control are reported in Imputation and Rations. Proteins sorted by total combined abundance are shown in the following sheet, followed by a list of the 1699 most abundant proteins, sorted by their abundance in BS or SB datasets respectively. Green color indicates proteins enriched 5 or more times in each dataset.

SI References

1. D. I. Kim *et al.*, An improved smaller biotin ligase for BioID proximity labeling. *Mol Biol Cell* **27**, 1188-1196 (2016).
2. R. J. Szczesny *et al.*, Versatile approach for functional analysis of human proteins and efficient stable cell line generation using FLP-mediated recombination system. *PLoS One* **13**, e0194887 (2018).
3. G. N. Maertens *et al.*, Structural basis for nuclear import of splicing factors by human Transportin 3. *Proc Natl Acad Sci U S A* **111**, 2728-2733 (2014).
4. P. Savitsky *et al.*, High-throughput production of human proteins for crystallization: the SGC experience. *J Struct Biol* **172**, 3-13 (2010).
5. K. B. Rogala *et al.*, Structural basis for the docking of mTORC1 on the lysosomal surface. *Science* **366**, 468-475 (2019).
6. O. Anczukow *et al.*, SRSF1-Regulated Alternative Splicing in Breast Cancer. *Mol Cell* **60**, 105-117 (2015).