

# A short conserved motif in ALYREF directs cap- and EJC-dependent assembly of export complexes on spliced mRNAs

Agnieszka M. Gromadzka, Anna-Lena Steckelberg, Kusum K. Singh, Kay Hofmann and Niels H. Gehring\*

Institute for Genetics, University of Cologne, D-50674 Cologne, Germany

Received November 17, 2014; Revised December 18, 2015; Accepted January 04, 2016

## ABSTRACT

**The export of messenger RNAs (mRNAs) is the final of several nuclear posttranscriptional steps of gene expression. The formation of export-competent mRNPs involves the recruitment of export factors that are assumed to facilitate transport of the mature mRNAs. Using *in vitro* splicing assays, we show that a core set of export factors, including ALYREF, UAP56 and DDX39, readily associate with the spliced RNAs in an EJC (exon junction complex)- and cap-dependent manner. In order to elucidate how ALYREF and other export adaptors mediate mRNA export, we conducted a computational analysis and discovered four short, conserved, linear motifs present in RNA-binding proteins. We show that mutation in one of the new motifs (WxHD) in an unstructured region of ALYREF reduced RNA binding and abolished the interaction with eIF4A3 and CBP80. Additionally, the mutation impaired proper localization to nuclear speckles and export of a spliced reporter mRNA. Our results reveal important details of the orchestrated recruitment of export factors during the formation of export competent mRNPs.**

## INTRODUCTION

In metazoan cells, synthesis of mature messenger RNAs (mRNA) requires several steps of processing, which are accompanied by the deposition of proteins onto the mRNA to form messenger ribonucleoprotein complexes (mRNPs). Importantly, the composition of mRNPs influences gene expression both in the nucleus and in the cytoplasm and undergoes constant remodelling (1,2). In eukaryotes transcription and translation, the two major steps of gene expression

are spatially separated from each other by the nuclear membrane. Therefore, the assembly, nuclear export and subsequent transport of mature mRNPs are prerequisites for the efficient biosynthesis of proteins.

In mammalian cells, splicing plays an important role in mRNP remodelling and thereby influences later steps of gene expression. Consequently, the deposition of export components occurs mainly during splicing (3–5). This is thought to result at least in part from the splicing-dependent deposition of exon junction complexes (EJC), which interact with some of the export factors (6–8). The mammalian TREX complex, which is composed of UAP56, ALYREF and the mammalian THO complex, plays a pivotal role in coordinating splicing with mRNA export (9). ALYREF (also referred to as Aly, REF; in yeast: Yra1p) and its interaction partner UAP56 (BAT1; in yeast: Sub2) were described as components of the EJC (6). Interestingly, UAP56 is also an essential splicing factor, required for the first ATP-dependent step of splicing and spliceosome interaction with the branch point (10,11). A single UAP56 gene is present in yeast and fruit flies, whereas mammalian cells also express DDX39, an additional UAP56-related helicase (12,13). Furthermore, in mammalian cells only the simultaneous depletion of UAP56 and DDX39 causes a substantial retention of mRNAs in the nucleus (14), suggesting a redundant function of these two proteins (12).

According to current models, ALYREF is deposited onto mRNAs through its interaction partner UAP56 (3,15). The formation of export-competent mRNPs involves the splicing-dependent assembly of the THO/TREX complex, of which UAP56 and ALYREF are integral components (15,16). Moreover, the 5' cap facilitates the recruitment of export factors to intron-containing and intronless mRNA (16–18). Eventually, the binding of the general export receptor NXF1 (Tap) and its heterodimerisation partner NXT1 (p15) to the mRNP initiates the translocation through the

\*To whom correspondence should be addressed. Tel: +49 (0)221 470 3873; Fax: +49 (0)221 470 5264; Email: ngehring@uni-koeln.de  
Present address:

Agnieszka M. Gromadzka, Department of Neurodegenerative Diseases, Institute of Neurology, UCL, London, UK.

Anna-Lena Steckelberg, Department of Biochemistry and Molecular Genetics, University of Colorado Denver School of Medicine, Aurora, Colorado, USA.

Kusum K. Singh, Department of Biosciences and bioengineering, Indian Institute of Technology, Guwahati, India.

nuclear pore (19–21). Several export factors have been identified in mammalian cells, but their mode of interaction with mRNAs and the determinants for their recruitment to the specific mRNP is not understood entirely.

Up to date, several additional mRNA export adaptors have been identified to modulate the process of mRNA export: CHTOP, CIP29, LUZP4 and UIF. Interestingly, CHTOP, LUZP4 and UIF are found to directly interact with UAP56 via a short conserved sequence referred to as UBM (UAP56 binding motif) (22). CHTOP enhances the ATPase activity of UAP56 (23). LUZP4 was found to complement ALYREF knock down *in vivo* and UIF was reported to influence export through binding with NXF1 (22,24). Finally, CIP29 (yeast homologue: Tho1p) interacts with UAP56 and ALYREF, forming a trimeric complex (25,26). Interestingly, CIP29 does not have an UBM. Moreover, it has been suggested that the family of export factors shares a common motif structure, enabling the interaction with UAP56 and possibly other mRNP components to enter the mRNA export pathway (22,24). Thus, the list of export adaptors is not yet complete, and the systematic approach to investigate these is yet to be found.

Short-linear motifs (SLiMs), also referred to as ELM (eukaryotic linear motifs), are short (3–10 amino acid long), conserved sequences situated in intrinsically disordered regions, lacking secondary and tertiary structures. SLiMs mediate low-affinity interactions between proteins and also between RNA and proteins (e.g. RGG and YGG motifs, RS domains, GY-rich regions, etc.) (27–30). Although little is known about the function of low complexity regions in RNA-binding proteins, short linear motifs might be involved in the dynamic rearrangement of mRNPs composition during mRNA export and the assembly of a variety of export adaptors.

In this study, we investigated the interactions of mammalian export components with different RNA substrates. We find that export components UAP56 and ALYREF are recruited to spliced mRNA in an EJC- and cap-dependent manner. We identified four short, conserved, linear motifs present in proteins involved in mRNP processing and export, providing a systematic view into the family of export adaptors. Finally, we show that short linear motif in ALYREF (WxHD) is required for its *in vitro* binding to spliced RNA and for efficient export of spliced mRNAs *in vivo*. Together, these findings expand our view on how nuclear mRNA processing during splicing orchestrates the recruitment of export factors to the RNA and provide insight into the role of short linear motifs in this process.

## MATERIALS AND METHODS

### Plasmids and cloning

DDX39 and UAP56 were obtained by reverse transcription (RT)-PCR using HeLa cells RNA. ALYREF was amplified by PCR from plasmid DNA (Origene). ALYREF deletion mutants: 1–106, 106-ter, 1–187, RRM,  $\Delta$ N15,  $\Delta$ C53 and  $\Delta$ N15/ $\Delta$ C53 were generated using PCR. ALYREF 87 WQHD/DQAK 90 was generated by site-directed mutagenesis. Expression vectors for FLAG CBP80, FLAG Y14 and FLAG SELOR were described previously (31) (Figures 1 and 2).

Inserts were transferred into various vectors in order to serve specific purposes. For protein expression in *in vitro* splicing experiments, constructs were inserted into the pCI-neo vector (Promega) with an N-terminal FLAG (DYKDDDDK) tag (Figures 1–3 and 5). ALYREF deletion mutants and the wild-type were cloned into pCI-FLAG-emGFP or into pCI-FLAG constructs for *in vitro* splicing (Figure 3). For localization studies, constructs were cloned into pCI-FLAG-emGFP (Figure 5E). For immunoprecipitation studies, constructs were cloned into pCDNA5 FRT/TO (Life Technologies) vector with an N-terminal triple FLAG-tag (Figure 5D). For export assays, pCI-Firefly and pCMV128 NanoLuc were used. pCMV128 NanoLuc was generated by inserting the NanoLuc sequence (Promega) into the NotI and BamHI sites of pCMV128 (32) (Figure 6).

Templates for *in vitro* transcription (MINX, AdML) of radiolabelled RNA were cloned into pGEM-4Z (Promega). MINX was described previously (31). AdML and AdML $\Delta$ i were generated by PCR amplification of a synthetic DNA (Figures 1–3 and 5). All constructs were verified by DNA sequencing.

### Bacterial expression

Recombinant PYM (rPYM) was obtained as described previously (31). Briefly, full-length PYM fused to a C-terminal Strep-tag was expressed in BL21 (DE3) bacteria and purified on StrepTactin columns. Purified PYM was desalted by ultrafiltration and diluted in buffer E.

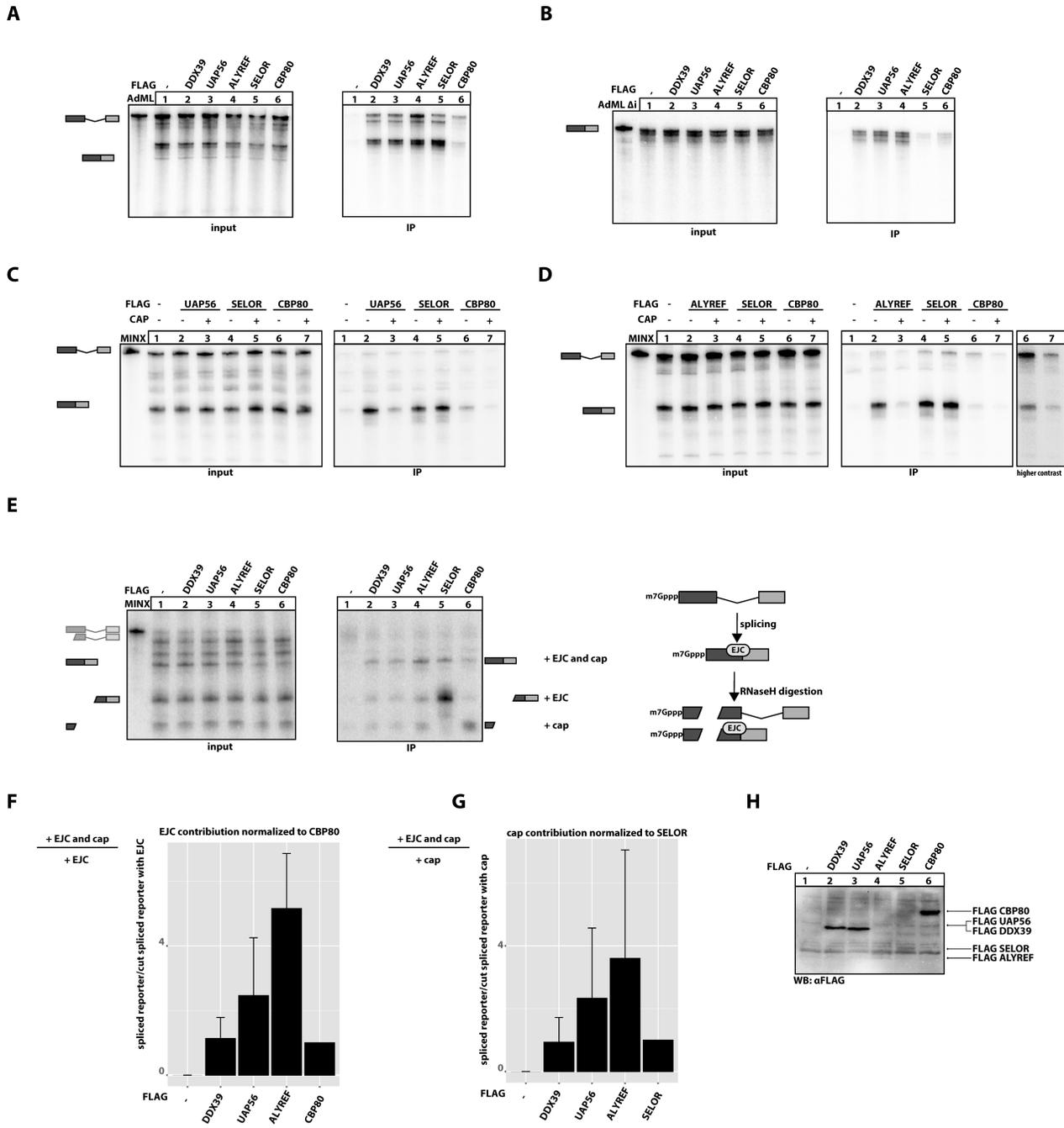
### Cell culture and plasmid transfections

HEK 293 cell lysates expressing FLAG-tagged proteins for *in vitro* splicing assays were obtained as described previously (33). For imaging experiments, HeLa cells were transfected with 250 ng of FLAG-emGFP constructs and 250 ng of mCherry-SRSF2 in 6-well plates using 2X BES-buffered saline (BBS)-based calcium phosphate precipitation (34). For luciferase assays HeLa cells were transfected with 2  $\mu$ g of FLAG-emGFP constructs, 500 ng of pCMV128 NanoLuc and 50 ng of pCI-Firefly constructs in 6-well plates using 2X BES-buffered saline (BBS)-based calcium phosphate precipitation (34).

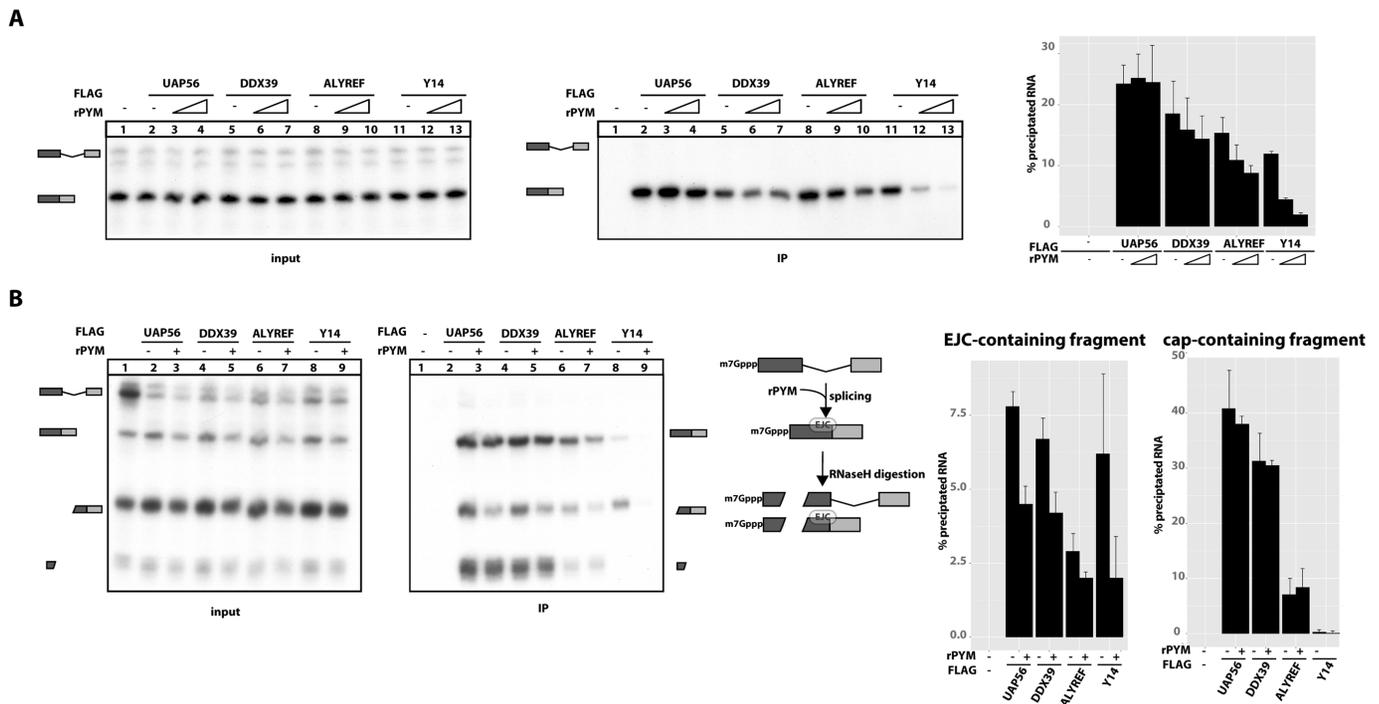
The Flp-In T-Rex System (Life Technologies) was used to generate tetracycline-inducible, stable mammalian cell lines. Transgene expression was induced by adding 1  $\mu$ g/ml doxycycline to the cell culture medium 72 h prior to harvesting. For immunoprecipitation experiments, HeLa cells were cultured and transfected as described previously (35).

### *In vitro* splicing assay

*In vitro* splicing assays were performed as described previously (33). Briefly, radiolabelled RNA substrates were incubated under splicing conditions in HeLa nuclear extract supplemented with whole-cell extracts from HEK293 cells expressing FLAG-tagged proteins. After splicing, RNA-protein complexes were immunoprecipitated via the FLAG-tagged proteins. The RNA was purified and subsequently analysed by denaturing gel electrophoresis and phosphorimaging.



**Figure 1.** Export factors binding to mRNA during splicing. (A) *In vitro* splicing reactions of AdML substrate were supplemented with extracts expressing the indicated FLAG-tagged export factors. SELOR (eIF4A3-binding domain of BTZ) and CBP80 were used as controls, unused FLAG-tag served as a negative control. FLAG-containing mRNPs were immunoprecipitated with anti-FLAG beads and co-precipitated RNA resolved on a denaturing PAGE. A total of 10% of each splicing reaction was used as input. Schematic representations of splicing products are depicted on the left side of the autoradiograph. (B) *In vitro* splicing reactions were performed as in (A) using intronless AdML substrate (AdML Δi). (C) *In vitro* splicing reactions were performed as in (A) were supplemented with 10 μM of cap analogue (m<sup>7</sup>GpppG) as indicated and mRNPs were immunoprecipitated and RNAs resolved on denaturing PAGE. FLAG-tagged SELOR and CBP80 served as negative and positive controls, respectively. (D) *In vitro* splicing reaction of extract expressing FLAG-tagged ALYREF supplemented with cap analogue was performed as in (C). A high-contrast picture of lanes 6 and 7 is shown as separate panel. (E) *In vitro* splicing reactions were performed as in (A) using MINX substrate followed by incomplete RNase H digestion of splicing products and immunoprecipitation of mRNPs. The splicing and digestion scheme is represented on the left side of the autoradiograph. Expression of FLAG-tagged proteins in HEK 293 extracts used in panels A–D was determined by representative immunoblot analysis using a FLAG-antibody. (F) Quantification of three independent, biological repetitions of the experiment presented in (E). The EJC contribution to the binding of export components is a calculated ratio of intensities of two bands as depicted on the left side of the graph. EJC contribution to the CBP80 binding serves as a negative control and was set to 1. (G) Quantification of three independent, biological repetitions of the experiment presented in (E). The EJC contribution to the binding of export components is a calculated ratio of intensities of two bands as depicted on the left side of the graph. The cap contribution to the SELOR binding serves as a negative control and was set to 1. (H) Expression of FLAG-tagged proteins in HEK 293 extracts used in (A–G) was determined by immunoblot analysis using a FLAG-antibody.



**Figure 2.** Export factors binding to capped, spliced mRNA is influenced by EJC disassembly by PYM. (A) Splicing reactions were performed as described in Figure 1C in the presence of increasing amounts (0, 1 and 2.5  $\mu$ g) of recombinant PYM (rPYM) as indicated by the triangle. mRNPs were immunoprecipitated using FLAG affinity agarose and analysed by denaturing PAGE. Three independent repetitions of the experiment presented in this panel were quantified and presented on the graph. (B) Splicing reactions and oligonucleotide directed RNase H digestions (performed as described in Figure 1E) were carried out in the presence or absence of rPYM as described in Figure 2A. mRNPs were immunoprecipitated using FLAG affinity agarose and analysed by denaturing PAGE. Three independent repetitions of the experiment presented in this panel were quantified and presented on the graphs.

For the cap competition assays shown in Figure 1C and D, splicing reactions were supplemented with 10  $\mu$ M of m<sup>7</sup>GpppG RNA cap analogue. After splicing, FLAG-tagged proteins were immunoprecipitated and associated RNA was isolated as described above.

For RNase H digestion of *in vitro* splicing assays (Figure 1E–G and Supplementary Figure S1) 1  $\mu$ M of a cDNA oligonucleotide was added after the splicing reaction was completed. Digestion was performed for 15 min at 30°C. After splicing and digestion, RNA-protein complexes were immunoprecipitated via the FLAG-tagged protein. RNA was isolated and analysed as described above.

Oligonucleotides used for the oligo-directed digestion of spliced RNAs are listed below.

Name	Sequence
MINX	5' GAGTGGGCGAGC 3'
AdMI -36	5' CGGAAGAGAGTG 3'
AdMI -58	5' CGCGACGTCTCG 3'
AdMI -70	5' ACCTGCAGGCCA 3'

In the splicing assay with rPYM, splicing reaction was performed as described above in addition of 1 and 2.5  $\mu$ g of rPYM (Figure 6A, two concentrations). In Figure 6B, 2.5  $\mu$ g of rPYM was used.

### Immunoblot Analysis and Immunoprecipitation

Immunoblot analysis and immunoprecipitation were performed as described previously (36). Briefly, Magnetic M2

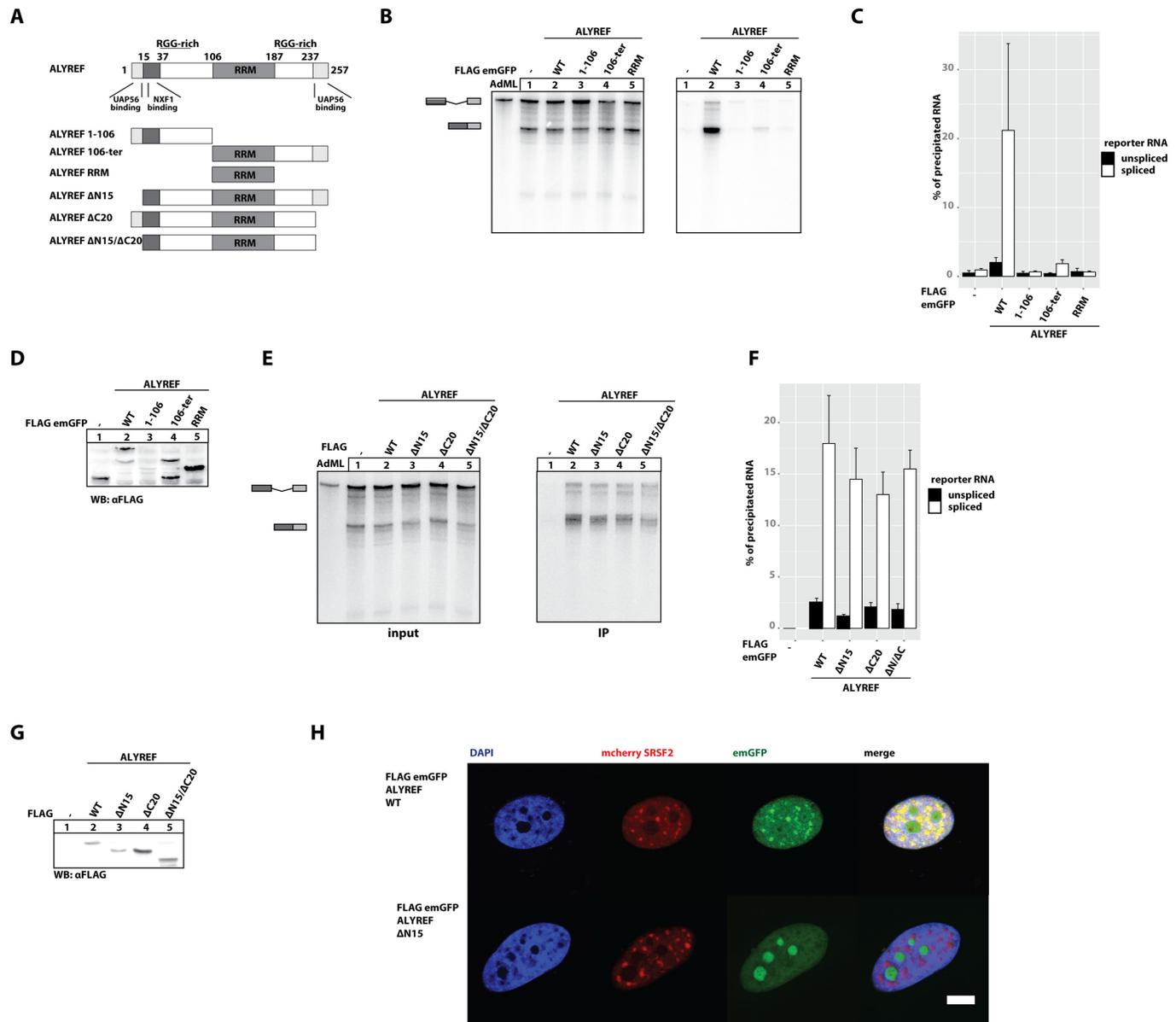
anti-FLAG beads (Sigma-Aldrich) were used to immunoprecipitate FLAG-tagged complexes from RNase A-treated (50  $\mu$ g/ml) HeLa cell lysates in lysis buffer (50 mM Tris [pH 7.2], 150 mM NaCl and 0.5% Triton X-100) supplemented with protease inhibitor (Sigma-Aldrich). Complexes were eluted with SDS-sample buffer, separated by SDS-PAGE and analysed by immunoblotting.

### Antibodies

FLAG-antibodies were from Sigma. The antibodies against UAP56 and ALYREF were generated by Thermo Scientific using peptides (UAP56: ENDVDNELLDYEDDEVET; ALYREF: CEELDAQLDAYNARMDTS) corresponding to amino acids 3–20 of human UAP56 and 241–257 of human ALYREF, respectively. The eIF4A3 polyclonal antibody was made by GenScript with an N-terminal peptide of eIF4A3 (37). The antibody against CBP80 was kindly provided by Elisa Izaurralde.

### Microscope image acquisition

Images were obtained at room temperature on an Olympus FV1000 confocal microscope using a 60x UPlanApo (NA 1.35) objective. Assembly, contrast and brightness adjustments of final image sets were performed using Photoshop CS4 (Adobe Systems, Inc.)



**Figure 3.** ALYREF binding to RNA *in vitro* is not mediated by its canonical RNA binding domain and is independent of UAP56 binding. (A) Schematic representation of ALYREF wild-type and mutants used in the study. (B) *In vitro* splicing reactions of AdML RNA were supplemented with extracts expressing FLAG-emGFP ALYREF truncation mutants (as in Figure 1A). Schematic representations of splicing and digestion products are depicted on the left side of the autoradiograph. (C) The results of three, independent, biological repetitions were quantified. Band intensities corresponding to the unspliced reporter RNA and spliced RNA are presented as black and white bars, respectively. (D) Expression of FLAG-tagged proteins in HEK 293 extracts used in (B) was determined by immunoblot analysis using a FLAG-antibody. (E) *In vitro* splicing reactions of AdML RNA were supplemented with extracts expressing FLAG-emGFP ALYREF truncation mutants as in (B). Schematic representations of splicing and digestion products are depicted on the left side of the autoradiograph. (F) The results of three independent biological repetitions were quantified and presented on the graph. Band intensities corresponding to the unspliced reporter RNA and spliced RNA are presented as black and white bars, respectively. (G) Expression of FLAG-tagged proteins in HEK 293 extracts used in (E) was determined by immunoblot analysis using a FLAG-antibody. (H) Localization of FLAG-emGFP ALYREF wild-type and mutant in HeLa cells. DAPI was used to stain nuclear DNA. Scale bar, 5  $\mu$ m.

### Localization studies

Cells were seeded onto coverslips 24 h post-transfection, fixed with 3.7% formaldehyde (Carl Roth) in phosphate buffered saline (PBS) after 48 h and permeabilised for 5 min with 0.5% Triton X-100 (Sigma-Aldrich) in PBS on ice. Nuclei were counterstained with DAPI (Carl Roth). Coverslips were mounted in DABCO: Moviol (Carl Roth).

### Luciferase assay

Cells were harvested 72 h post-transfection using 300  $\mu$ l Passive Lysis Buffer (Promega). Lysates (2  $\mu$ l) were analyzed using the NanoGlo DLR Assay (Promega) on a Centro XS3 LB960 microplate luminometer (Berthold Technologies).

## Motif discovery

For computational motif discovery, a set of ALYREF orthologues from various metazoan species was subjected to multiple alignment by the L-INS-I method of the MAFFT software (38). The resulting alignment was used for the construction of a generalised profile (39) and iterative database searches in the UNIPROT database, always including hits with a *P*-value better than 0.01 for subsequent iteration cycles. An analogous procedure was used for finding sequence conservation in BTZ, BCLAF1, THRAP3, UIF, CHTOP, SKAR, LUZP4 and PHAX. When it became clear (usually after 1–3 iteration cycles) that the sequence conservation was restricted to one or more short motif-size regions, the multiple alignments were cropped and (if necessary) split, continuing the iteration cycle with one multiple alignment per conserved motif. Iterative refinement was stopped when no new significant hits were found in the database. The sequence logo representation shown in Figure 4 was generated using the WebLogo software (40). Finally, the relationship of the motifs to each other was assessed by using the HH-SEARCH software (41).

## RESULTS

### Export complex components specifically bind to mRNA containing both cap and EJC

The recruitment of mRNA export factors to any given mRNA is a prerequisite for mRNA export and likely influences the efficiency and pathway of mRNA transport. Therefore, we aimed to understand how a subset of mRNA export factors interacts with mRNAs. We utilised a well-established *in vitro* splicing system, which has been shown to recapitulate the binding of different mRNP components to intron-containing (spliced) and intronless (non-spliceable) AdML and MINX RNA model substrates (33,35). Briefly, HeLa nuclear extract was supplemented with splicing-competent whole-cell extracts from HEK293 cells expressing FLAG-tagged export components (DDX39, UAP56 and ALYREF), FLAG-tagged SELOR (the eIF4A3-interacting domain of the EJC protein Barentsz) and FLAG-tagged CBP80. Subsequent to the splicing reaction, the FLAG-tagged proteins were immunoprecipitated together with their associated mRNAs (Figure 1A).

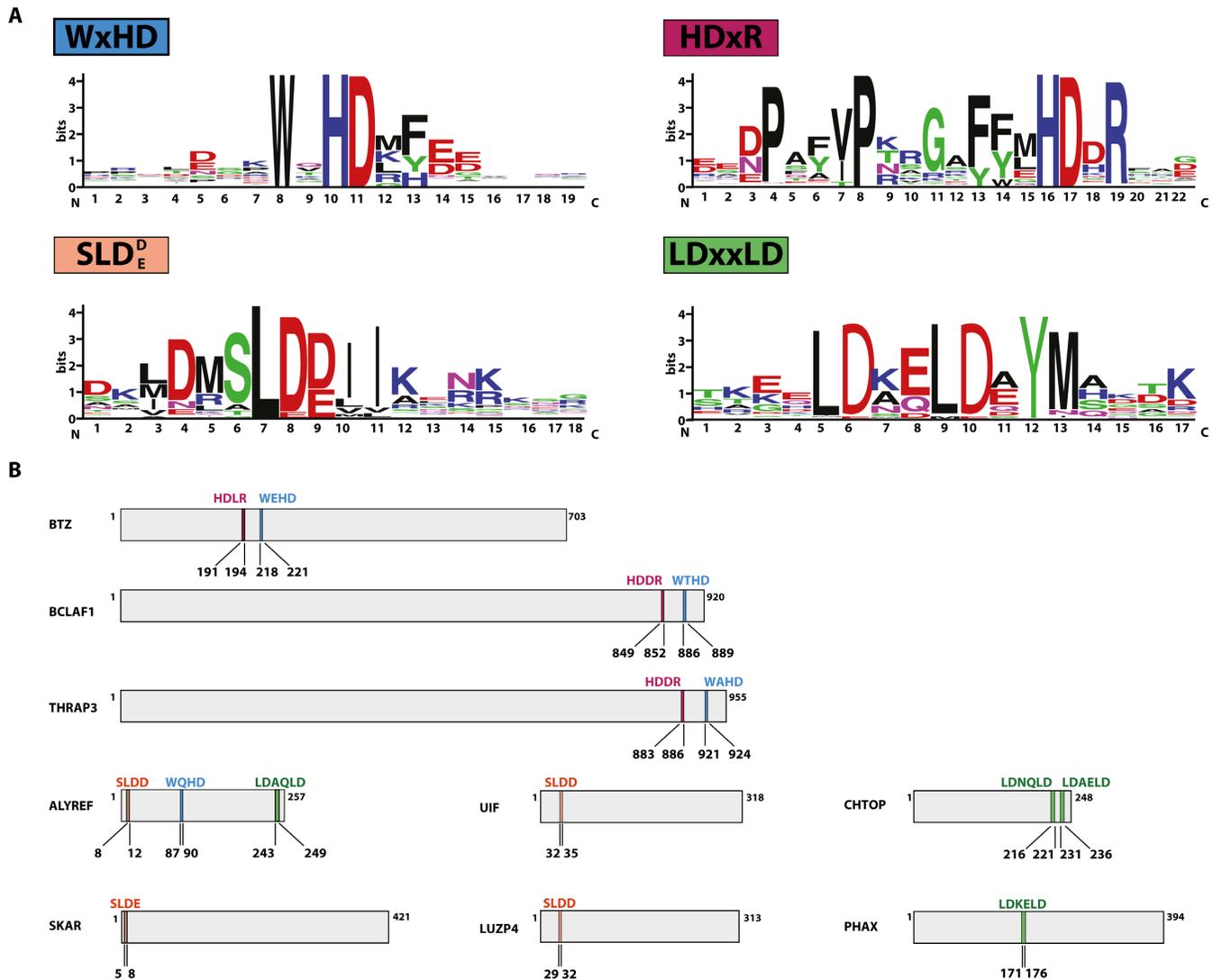
In this assay, ALYREF precipitated comparable amounts of spliced RNA as SELOR (the eIF4A3-interacting domain of Barentsz), indicating that it binds strongly to the spliced mRNA (Figure 1A, lanes 4 and 5). Barentsz (also referred to as BTZ, CASC3 or MLN51) is a component of EJC, known to bind spliced transcripts and therefore its EJC-binding domain SELOR was used as a positive control (35,42). Similar to ALYREF, UAP56 and DDX39 also bound spliced RNAs, but precipitated RNA less efficiently (Figure 1A, lanes 2 and 3). The binding pattern of the tested mRNA export components was different from the cap-binding protein CBP80. CBP80 bound preferentially to the cap and precipitated equal amounts of spliced and unspliced mRNA from *in vitro* splicing reactions (Figure 1A, lane 6). We also examined binding of export components to an intronless mRNA (AdML  $\Delta$ i), which was identical in se-

quence to spliced AdML RNA, but did not undergo splicing (Figure 1B). As expected, SELOR did not precipitate the intronless RNA (Figure 1B, lane 5). In contrast small amounts of AdML $\Delta$ i were precipitated by ALREF, UAP56 and DDX39, demonstrating their residual interaction with intronless RNAs. Nonetheless, these export factors appear to preferentially bind to the spliced mRNA, as reported previously (15–17). Notably, similar results were obtained with a MINX reporter RNA (data not shown).

Binding of export factors to the cap or the cap region of mRNAs has been previously demonstrated (16), but contradicting results were reported regarding the cap-dependent interaction of UAP56 and ALYREF with spliced mRNA (16,17). To investigate the role of the cap, we analysed if the presence of cap analogue (m<sup>7</sup>GpppG) interferes with binding of UAP56 and ALYREF to spliced mRNA (Figure 1C and D, respectively). In contrast to previous reports, the addition of cap analogue to the *in vitro* splicing reaction did not decrease splicing efficiency (17,43–44). Strikingly, binding to spliced RNA of both ALYREF and UAP56 was strongly impaired in the presence of cap analogue (compare lane 2 and 3 in Figure 1C and D), demonstrating that this interaction requires cap-binding proteins. This result is in line with the previously reported recruitment of the mRNA export machinery to the 5' end of mRNA (16).

To further examine the importance of 5' cap in binding of export components to the spliced RNA and to determine the position, where the core export components bind onto the RNA, we used oligonucleotide-directed RNase H cleavage of the mRNA after *in vitro* splicing assay. A similar assay was previously used to analyse the binding of the TREX complex to spliced RNAs (16). In our experiment, spliced and unspliced reporter mRNAs were cleaved at a position upstream of the EJC binding site in the first exon, which enabled us to separately study the binding of export factors to the cap or the EJC (Figure 1E, scheme). While the uncleaved spliced mRNA contains both the cap and EJC, the 5' fragment of spliced mRNA contains only the cap, but no EJC, and the 3' fragment of spliced mRNA contains the EJC, but not the cap (Figure 1E, scheme). Although DDX39, UAP56 and ALYREF efficiently bound to the full length spliced RNA, neither of the two fragments produced after RNase H digestion was efficiently precipitated (Figure 1E, lanes 2–4). In contrast, SELOR precipitated the 3' fragment with the EJC binding site (Figure 1E, lane 5), and CBP80 precipitated the cap-containing 5' fragment (Figure 1E, lane 6), in line with their expected binding preferences. The expression of the FLAG-tagged proteins used in panels A–E was determined with Western blotting (Figure 1H).

To quantify the effect of cooperative binding to spliced mRNAs containing both the EJC and the cap, three independent biological repetitions were performed and quantified. In Figure 1F we calculated the ratio between the uncleaved spliced mRNA, containing both the cap and the EJC, and the cleaved mRNA containing only the cap (Figure 1F, Supplementary Figure S1A for alternative graphs). This ratio was set to 1 for CBP80, which does not bind to the cleaved mRNA containing only the EJC (Figure 1F), and therefore binds to mRNA based only on the presence of the cap. All tested export factors showed an enrichment in cooperative binding to uncleaved spliced mRNA com-



**Figure 4.** Short, conserved motifs present in RNA-binding proteins. (A) WebLogo diagrams of short conserved motifs identified in RNA-binding proteins. (B) Schematic representation of RNA-binding proteins and the distribution of short, conserved motifs listed in (A) (as above: WxHD motif in blue, HDxR motif in purple, SLDD/E motif in orange and LDxxLD motif in green). The sequence of the motif is indicated above the bar representing the protein, the position of the motif is indicated below the bar.

pared to CBP80 (Figure 1F). Reciprocally, in Figure 1G we calculated the ratio between the uncleaved spliced mRNA and the cleaved mRNA containing only the EJC (Figure 1G, Supplementary Figure S1). This ratio was set to 1 for SELOR, which does not bind to the cleaved mRNA containing only the cap (Figure 1E), and therefore binds to mRNA based only on the presence of the EJC. All tested export components showed enrichment in cooperative binding to the uncleaved spliced mRNA compared to SELOR (Figure 1G). Notably, ALYREF showed the strongest enrichment in binding to spliced RNA for both parameters (>4-fold) (Figure 1F and G).

To validate that the observed differential precipitation of cleavage products is independent of the RNA sequence, we repeated the experiment using a different splicing substrate (AdML, Supplementary Figure S1B–D). As shown before for the MINX splicing substrate, DDX39, UAP56 and

ALYREF only precipitated RNA fragments that contained both the cap and an EJC-binding site. Moreover, differential precipitation of cleavage products is also independent of the RNA length, as demonstrated by the use of DNA oligos that target different positions in the spliced RNA (Supplementary Figure S1B–D). Interestingly, binding of ALYREF was stronger on long cap-containing mRNA fragments. This is in agreement with previous studies suggesting that an RNA of sufficient length can enter the mRNA export pathway more readily than short and structured RNAs (45).

It has been reported that the presence of either the cap or the EJC increases binding of export factors to mRNA. However, it was not known if the simultaneous presence of both EJC and cap confers cooperative binding. Our results suggest that 5' cap and EJC-binding site have an additive effect on the interaction of export factors with spliced mRNA.

### ALYREF binding to spliced mRNAs *in vitro* requires the EJC

To test if the association of ALYREF with spliced mRNA involves the EJC, we investigated the influence of PYM on binding of export components to mRNAs. PYM directly binds to the MAGOH-Y14 heterodimer to dissociate EJCs from mRNAs in the cytoplasm to promote recycling of EJCs (31,46). *In vitro*, the displacement of EJCs can be recapitulated by the addition of rPYM to splicing reactions (Figure 2B, scheme). We observed that UAP56 binding to spliced mRNAs was not affected by the addition of rPYM, whereas there was a slight decrease in RNA precipitation through DDX39 in the presence of rPYM (Figure 2A, lanes 1–7). In contrast, ALYREF binding to spliced mRNAs was significantly reduced after EJCs were disassembled (Figure 2A, lanes 8–10), albeit to a lesser extent than the binding of Y14, a core component of the EJC (Figure 2A, lanes 11–13). This observation suggests that ALYREF requires the interaction with an EJC to be recruited to spliced mRNAs.

To further validate that PYM-dependent reduction in RNA precipitation is mediated by the EJC, we performed a splicing assay followed by RNase H digestion (as in Figure 1E) in the presence of rPYM (Figure 2B). After splicing in the presence of rPYM and RNase H-directed RNA digestion, three main species of RNA are present: uncleaved, spliced RNA containing cap and exon-junction; cleaved, spliced RNA with exon-junction and cleaved, spliced RNA with the cap. As shown before, addition of rPYM to the splicing reaction led to a significant reduction in the precipitation of RNA fragments containing an EJC-binding site (Figure 2B). This effect was much more pronounced for RNA fragments that lacked the 5' cap (Figure 2B). Interestingly, in the presence of rPYM, UAP56, DDX39 and ALYREF all displayed an equally reduced ability to precipitated RNA fragments lacking the 5' cap, demonstrating that the interaction of the three export receptors with uncapped RNA species is mediated by the EJC. Taken together, our data show that displacement of EJCs through PYM reduces the ability of UAP56, DDX39 and ALYREF to precipitate spliced RNA and therefore establishes a role for the EJC in export receptor recruitment. However, in the presence of a 5' cap, individual export receptors display variable levels of EJC-dependence, with ALYREF being most and UAP56 least dependent on the EJC.

### ALYREF binding to UAP56 is dispensable for its interaction with mRNA *in vitro*

The recruitment of ALYREF to mRNAs by UAP56 has been reported to link pre-mRNA splicing and mRNA export (3,22). Therefore, we sought to study the function of the UAP56-binding regions of ALYREF in our splicing assay.

ALYREF contains an RNA-recognition motif (RRM; 106–187 aa), two previously described UAP56-interaction domains (UBM; N-terminal 1–15 aa, C-terminal 237–257 aa), an NXF1-binding motif (15–37 aa) and two arginine-rich regions (RGG; 17–87 aa; 198–231 aa) (3,13,22). Notably, the N- and C-terminal domains, but not the RRM, were sufficient for murine REF binding to RNA *in vitro* (47,48).

We constructed three large truncation mutants to investigate the mode of ALYREF interaction with mRNA. We deleted either the N-terminal part (106-ter aa), the C-terminal part (1–106 aa), or both parts of ALYREF (106–187 aa, corresponding to the RRM) (Figure 3A) and tested the RNA binding of these variants in our splicing assay (Figure 3B and C for quantification). The expression of ALYREF 106-ter and ALYREF RRM was comparable to the expression of full length ALYREF (Figure 3D). In contrast, we observed lower expression levels of the mutant ALYREF 1–106, which might be caused by the deletion of all known structured regions of ALYREF. None of the deletion mutants precipitated spliced RNA, suggesting that neither the RGG-rich regions nor the RRM is able to bind RNA by itself (Figure 3B). The substantial deletion of the C-terminal region of ALYREF (ALYREF 1–187) resulted in a milder reduction of binding to spliced and unspliced RNA (Supplementary Figure S2A), suggesting the RRM might mediate a weak interaction of ALYREF with both RNA species.

To correlate UAP56 interaction and RNA binding of ALYREF, we used mutants of ALYREF previously described to abolish binding to UAP56 (22) in order to investigate their interaction with RNA in the splicing assay (Figure 3E and F for quantification). We observed only a mild decrease of RNA binding with ALYREF lacking either the N-terminal (Figure 3E, lane 3) or the C-terminal sequence (Figure 3E, lane 4) or both (Figure 3E), even though these mutants display a strongly decreased ability to interact with UAP56 (Supplementary Figure S2B).

In conclusion, we show that neither of the potential RNA-binding domains of ALYREF (RGG, RRM) can mediate RNA binding in isolation. Furthermore, the deletion of the UAP56 binding sites marginally reduced the capability of ALYREF to bind spliced RNA *in vitro*, opposing the prevailing model that UAP56 is required for the deposition of ALYREF on spliced mRNAs.

In living cells, the recruitment of ALYREF to mRNAs during splicing is presumably executed in nuclear speckles (4,48–49). We performed co-localization studies using FLAG-emGFP ALYREF as well as the UAP56-binding deficient mutant deltaN15, together with the splicing factor (SRSF2, mCherry fusion), a marker for nuclear speckles (Figure 3H), and confirmed their co-localization (upper panel). Strikingly, the ALYREF mutant lacking the N-terminal UAP56 binding domain does not show speckle localization, but instead forming aggregates in the nucleoli, suggesting that while UAP56 is not required for RNA-loading of ALYREF, it plays a role in the proper localization of ALYREF to nuclear speckles.

### Proteins involved in mRNA export and mRNA processing contain various, short, conserved motifs

Recently it was reported that many mRNA-interacting proteins are intrinsically disordered and enriched in repetitive amino acid motifs (28,50–51). ALYREF contains two presumably unstructured regions in the N- and C-terminus, which might harbour functionally important sequences. In order to identify motifs potentially involved in mRNA export, we performed a computational analysis using mRNP-

related proteins subjected to multiple alignment by the L-INS-I method (see Materials and Methods). Indeed, we identified four conserved motifs with the core sequences SLDD/E, LDxxLD, HDxR and WxHD (where x represents a variable amino acid) (Figure 4A). In addition to the short conserved core sequences, some highly conserved amino acids are found at other positions within the identified sequences (e.g. for HDxR: P is always present at the position -8 and -12).

All proteins containing the SLDD/E motif have been previously described to be involved in mRNA export, i.e. ALYREF, SKAR, UIF and LUZP4 (Figure 4B). The motif has been previously designated UBM (UAP56 binding motif), since it mediates the interaction with UAP56 (22,24,52).

The LDxxLD motif is present in ALYREF, PHAX and twice in CHTOP. In ALYREF this motif is part of the C-terminal UAP56-binding motif (22). CHTOP was found to be involved in mRNA export through the direct interaction with UAP56 (23). PHAX is involved in the export of snRNA and snoRNA through binding to the cap binding complex (CBC) (53,54). In contrast to mRNA export, which requires NXF1/NXT1, snRNA and snoRNA export involves CRM1 (a nuclear export receptor of the importin  $\beta$  family) and RanGTP. It is interesting to note that two conserved motifs required for UAP56-binding are found in proteins, which are implicated in the export of different classes of mRNAs. However, only ALYREF contains both motifs, whereas all other proteins have a single motif.

The SLDD/E and the LDxxLD motifs described above were previously reported to be variants of the same motif (22). Although we identified these sequences independently from each other, both motifs allow proteins to bind UAP56 and therefore execute the same molecular function.

Another short conserved motif present in ALYREF (WxHD) was found in BTZ, a component of the EJC core (8,42), and the two splicing factors BCLAF1 and THRAP3 (55,56). In the case of BTZ, BCLAF1 and THRAP3, the motif is preceded by another short motif HDxR. The function of these motifs remains unknown. However, the sequence WxHD of BTZ was previously identified to mediate binding to eIF4A3. Mutations within this motif (W218D; HD220AK) abolish the interaction with eIF4A3 and prevent BTZ from binding to mRNA and to other EJC components (35).

#### **ALYREF contains a short motif, conserved in other RNA-binding proteins, required for interaction with mRNA *in vitro***

Since ALYREF binding to mRNA does not involve the RRM and is independent of the presence of UAP56-binding motifs, we hypothesised that the conserved WxHD motif could contribute to mRNA binding of ALYREF (Figure 4).

Strikingly, a mutant of ALYREF, in which WQHD has been replaced by DQAK (Figure 5A) showed drastically reduced ability to co-immunoprecipitate mRNA generated by *in vitro* splicing (Figure 5B, compare lane 2 and 3), despite being expressed at similar levels (Figure 5C).

Since ALYREF preferentially interacted with spliced and capped RNAs *in vitro*, we wanted to study whether the WxHD motif is involved in the interaction with proteins

of spliced mRNPs. To this end, cell lines expressing either 3xFLAG ALYREF WT or WQHD mutant were subjected to FLAG co-immunoprecipitation experiments. In RNase-treated lysates 3xFLAG ALYREF WT precipitated endogenous eIF4A3, CBP80 and UAP56. However, the 3xFLAG ALYREF WQHD mutant precipitated only UAP56, whereas the interaction with CBP80 and eIF4A3 was lost (Figure 5D).

To further test whether mutation of the WxHD motif affects the subcellular localization of ALYREF, we studied its colocalization with the splicing factor SRSF2 using fluorescently labelled proteins (Figure 5E). Interestingly, we observed a partial mislocalisation of the WQHD mutant. Although partial co-localization with SRSF2 was detectable, a substantial amount of ALYREF formed aggregates that did not colocalise with SRSF2, suggesting that RNA binding of ALYREF plays an important role in the proper targeting of the protein to nuclear speckles.

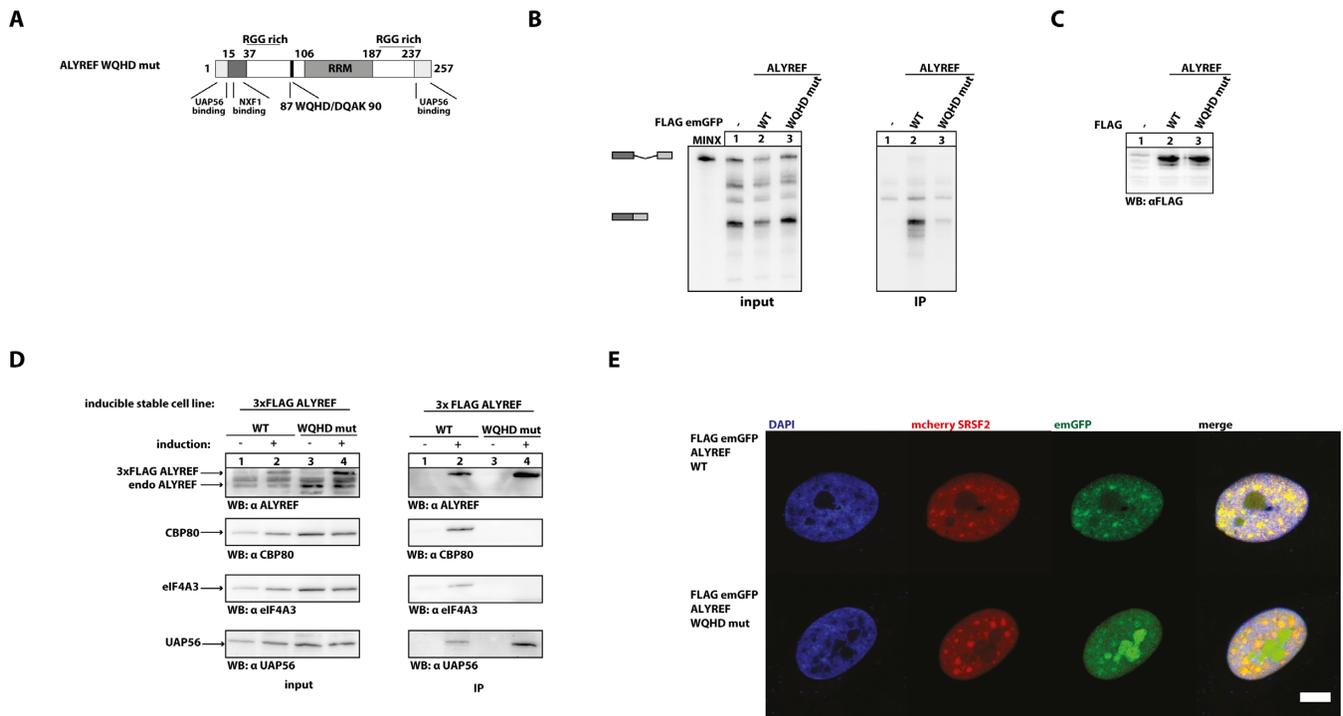
#### **The WQHD motif in ALYREF is required for efficient expression of spliced mRNAs *in vivo***

In order to test the biological significance of the identified WQHD motif in ALYREF, we studied the ability of ALYREF to promote the expression of a reporter mRNA in cultured cells (Figure 6). Briefly, HeLa cells were transfected with FLAG-emGFP ALYREF wild-type, WQHD mutant or FLAG-emGFP as negative control. Additionally, cells were cotransfected with two reporter constructs: pCI-FireflyLuc and pCMV128-NanoLuc. Importantly, FireflyLuc pre-mRNA is spliced in the nucleus (EJC is recruited), exported and translated (Figure 6A). In contrast, the open reading frame of the NanoLuciferase in pCMV128 NanoLuc is placed between two weak splice sites and can only be translated when the unspliced pre-mRNA is transported to the cytoplasm (57). If the pre-mRNA is spliced, the NanoLuc coding sequence is removed and no luciferase can be expressed (57).

Consistent with previous results showing that the presence of EJC promotes export and translation in an ALYREF-dependent manner, overexpression of ALYREF led to an increase in Firefly luciferase (spliced transcript, with EJC) compared to NanoLuc (unspliced transcript, no EJC). This effect was much less pronounced upon overexpression of the WQHD mutant of ALYREF, indicating that the WQHD motif plays an important role in the export of spliced mRNAs to the cytoplasm.

## **DISCUSSION**

In this study, we have re-examined the role of the cap and the EJC for export factor recruitment in human cells. We show that an excess of free cap analogue ( $m^7GpppG$ ) dissociates UAP56 and ALYREF from spliced RNA (Figure 1C and D), in agreement with the known requirement of a 5' cap for their recruitment (17). We then exploited partial RNase H digestion to examine if simultaneous presence of both the 5' cap and the EJC is required for binding of export factors. Neither of the cleaved products containing either the cap or the EJC alone was precipitated to a similar extent as the uncleaved RNA, which contains both compo-



**Figure 5.** The WQHD motif in ALYREF is important for its interactions with spliced mRNA, eIF4A3 and CBP80. (A) The WQHD motif in ALYREF is placed in a low-complexity region. In the ALYREF point mutant used in this study, the WxHD motif is replaced by DQAK. (B) *In vitro* splicing reactions of MINX substrate were supplemented with extracts expressing FLAG-tagged ALYREF wild-type and DQAK mutant (as in Figure 1A). Splicing products are schematically represented on the left side of a gel. (C) Expression of FLAG-tagged ALYREF proteins in HEK 293 extracts used in (B) was determined by immunoblot analysis using a FLAG-antibody. (D) FLAG-immunoprecipitations of RNase A-treated lysates from stable cell lines with induced expression of 3x FLAG ALYREF WT and WQHD mutant. Co-immunoprecipitated CBP80, eIF4A3 and UAP56 were detected by immunoblotting using a CBP80-, eIF4A3- and UAP56-specific antibodies, respectively. A total of 5% of cell extracts were loaded as input. Extracts without induced expression of 3xFLAG-tagged proteins were used as controls. (E) Localization of FLAG-emGFP ALYREF mutants in HeLa cells. DAPI was used to stain nuclear DNA. Scale bar, 5  $\mu$ m.

nents (Figure 1E–G). We therefore conclude that recognition of spliced mRNAs by the core export factors requires cooperative binding to the cap and the EJC.

We also show that ALYREF binding to the spliced RNA is dependent on the presence of the EJC (Figure 2A and B). The dissociation of EJCs from RNA resulted in a significant decrease of ALYREF binding to spliced RNA, suggesting a functional connection between EJC and ALYREF. Furthermore, the EJC also contributes to binding of UAP56 and DDX39 to spliced mRNA. This contribution becomes only evident, when mRNA is cleaved into cap- and EJC-containing fragments (Figure 2B). These data suggest that UAP56 and DDX39 are more cap- than EJC-dependent, whereas ALYREF shows a more cooperative interaction with spliced mRNA.

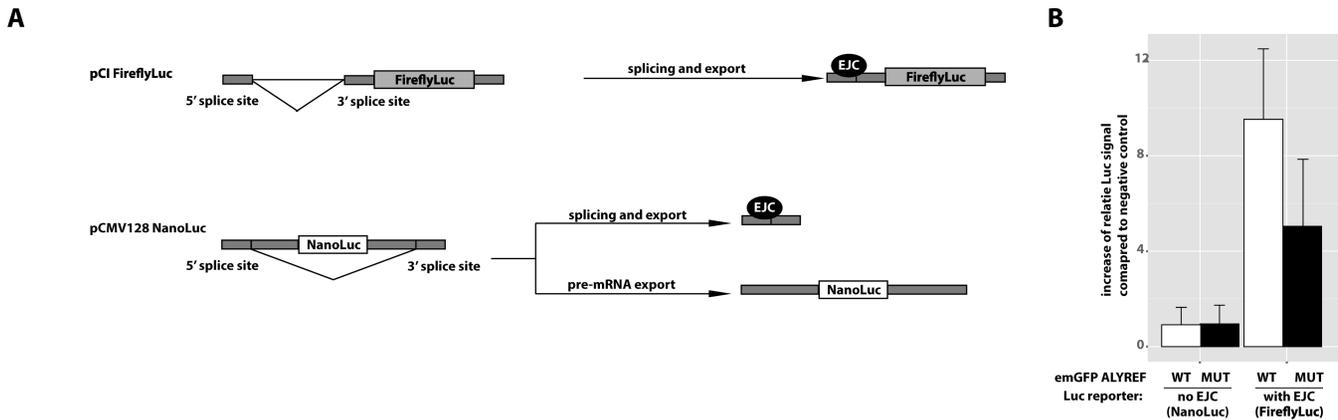
#### ALYREF binding to spliced mRNA is independent of UAP56

ALYREF contains different putative RNA-binding regions, of which RRM and RGG-rich regions of ALYREF were previously suggested as RNA-binding domains in living cells (58). However, we did not detect any interaction of the isolated RRM or RGG-rich regions with RNA. We also investigated the importance of the interaction of ALYREF and UAP56, since it was previously reported that ALYREF is loaded by UAP56 onto mRNA (3–4,15,18). Both, N-terminal and C-terminal UBM of ALYREF are involved in

UAP56 binding (21). In contrast to previous findings, mutants of ALYREF unable to bind to UAP56 still retained the ability to associate with spliced RNAs *in vitro*, indicating that the interaction of ALYREF with mRNA occurs in a UAP56-independent manner. However, we report that the N-terminal UAP56-interaction motif is required for localization to nuclear speckles (Figure 5E), highlighting the functional connection of spliceosomal components (e.g. UAP56) with export factors (e.g. ALYREF). Therefore, we hypothesize that recruitment of ALYREF to the spliceosome is an important function of UAP56 binding to ALYREF.

#### A novel mRNA binding motif WQHD is required for ALYREF deposition on spliced mRNA *in vitro*

Although several different models have been proposed in the literature, the exact mechanism of ALYREF deposition on spliced mRNA remains unknown. Here, we report the identification of a short motif (WxHD) in ALYREF, which is also present in four other proteins involved in mRNP formation. In ALYREF, the WQHD motif is located at position 87–91 in an unstructured region between the N-terminal RGG-like domain and the central RRM. Mutation of this motif leads to reduced binding of ALYREF to spliced RNA and impairs the interaction with CBP80 and eIF4A3. In contrast the interaction with UAP56 is not af-



**Figure 6.** The WQHD motif in ALYREF is required for efficient export of spliced mRNAs *in vivo*. (A) Schematic representation of luciferase reporters used in the export assay (B). (B) Quantifications of three independent biological repetitions of luciferase export assay.

ected (Figure 5). Hence, WxHD may represent a core sequence, which can promote binding to other mRNP components. Notably, the truncated ALYREF 1–106 protein contains the WxHD regions, but fails to interact with RNA. This indicates that the motif is necessary, but not sufficient for binding to our reporter RNA. The specificity of the interaction may be determined by the surrounding sequences and allow proteins with this motif to bind to a number of similar interaction partners.

The involvement of the WxHD motif in ALYREF for mRNA expression and likely mRNA export was shown using a luciferase-based mRNA export assay in living cells (Figure 6). Overexpression of wild-type ALYREF promoted the expression of spliced reporter RNA, suggesting that the WQHD motif is required for the interaction between the export adaptor ALYREF and the EJC, and thereby for efficient export of spliced transcripts. Unspliced RNAs lack an EJC, they are not expected to bind ALYREF, which explains why the mutation in ALYREF has no effect on the export of the unspliced transcript (pCMV128 NanoLuc). This does not contradict the results of tethering experiments, which show that ALYREF can increase the export of unspliced reporter RNA when it is artificially tethered with boxB sequences (57). We confirmed that in the absence of tethering, the mutation of the WQHD motif disrupts the interaction of ALYREF with the EJC, and therefore only affects export of spliced, but not unspliced transcripts (Figure 6B).

#### Novel short linear motifs are present in proteins involved in export of RNAs

In total, we identified four novel motifs that are present in proteins involved in mRNA export (ALYREF, CHTOP, LUZP4, PHAX, SKAR and UIF), which might belong to the growing family of SLiMs, also referred to as ELMs (28,59). SLiMs are short (3–10 amino acids) and present in disordered regions (also referred to as low-complexity regions, LC), and are expected to adopt a defined conformation only upon binding to their interaction partner (28). In contrast to protein–protein interactions mediated through folded domains, SLiMs likely interact with relatively low affinity with their binding partner, leading to transient inter-

action networks that can be regulated by post-translational modifications. Higher specificity and affinity are achieved by combining motifs within one protein, by additional interactions involving the flanking regions or by oligomerisation of several binding partners. Functionally, SLiMs either bind surfaces of the folded domains of their binding partner or intrinsically form complex mRNP granules (28,60). Our bioinformatic analysis suggests that various export adaptors bind to common interaction partners during the export of specific classes of mRNAs (22,24). The combination of different motifs may determine the specificity of a given export adaptor for certain mRNA targets, but may also be responsible for the observed plasticity of the network of mRNA export factors.

Notably, the function of two motifs (HDxR and WxHD) in two additional proteins – BCLAF1 (Bcl-2-associated transcription factor 1, also referred to as BTF) and THRAP3 (Thyroid hormone associated receptor 3, also referred to as TRAP150) has not been determined yet. BCLAF1 and THRAP3 were initially described as spliceosomal components (61), both were detected in the EJC-interactome and colocalise with EJC components in the cell (62,63). Moreover, RNA-like aggregates precipitated by biotinylated isoxazole contained proteins with WxHD motifs (e.g. ALYREF, THRAP3 and BTZ) (51,64). Notably, THRAP3 was identified as an NXF1 interaction partner and was reported to be involved in nuclear degradation of mRNA, suggesting a link between THRAP3 and RNA export machinery (55).

#### Implications for mRNA export

It has been speculated that splicing and deposition of EJCs enhance recruitment of export components and thereby stimulate export of spliced mRNPs (7). The EJC protein BTZ is localised to the cytoplasm in steady state and therefore may only join the EJC after export. Hence, the surface patch of eIF4A3 that is used for interaction with BTZ is not occupied in the nucleus. Therefore, it is feasible that ALYREF docks onto eIF4A3 as part of the EJC in the nucleus after splicing is completed. However, a stable interaction of ALYREF with the mRNA requires the cap, which provides an additional binding site. This is in line with the

previous observation that export complexes are only located between cap and first EJC in mRNA containing many EJCs (16). Finally, the export-competent mRNP is transported to the nuclear pore by a transient interaction with NXF1. ALYREF and other export factors are disassembled from the mRNP and the eIF4A3 binding site can interact with BTZ in the cytoplasm. Alternatively, an EJC-independent recruitment of ALYREF to mature mRNAs may involve direct binding of the WxHD motif to RNA. In either case, it will be interesting to find out how the simultaneous interaction of ALYREF with two distinct regions of the mRNA is established.

*In vivo*, the assembly of export complexes on mature mRNA may involve additional components of the mRNP, which have not been tested in this study. It has been shown, for example, that ALYREF interacts with the 3' end processing factor Pcf11 (65), and UAP56 was co-purified with 3' end processing complexes (66).

### Implications for neurodegenerative disease

Several components of export machinery have been implicated to play a role in the mechanism of amyotrophic lateral sclerosis pathogenesis (ALS)(67). ALS leads to the progressive loss of motor neurons and neuromuscular junctions. Around 10% of the familial form of ALS is caused by the hexanucleotide expansion (GGGGCC) in the C9ORF72 gene, leading to faulty mRNAs and toxic dipeptides (68,69).

The recently published screening for hexanucleotide repeats expansions toxicity modifiers identified the components of export machinery as potential candidates. In fruit fly, the deficiency of ALYREF suppressed the toxicity of G4C2 expansions. This phenomenon might be explained by our observation that EJC-containing mRNAs are exported more efficiently when ALYREF is overexpressed (Figure 6). In contrast, mRNAs without EJCs (such as G4C2 repeats) are retained in the nucleus or not efficiently expressed (68,69). We postulate that ALYREF acts as part of a control mechanism that retains unspliced or faulty RNAs in the nucleus. Furthermore, ALYREF and its interaction with EJC through the low complexity region might also play a role as a regulator for RNA toxicity of C9ORF72 gene leading to neurodegeneration.

### SUPPLEMENTARY DATA

[Supplementary Data](#) are available at NAR Online.

### ACKNOWLEDGEMENT

We thank Heidi Thelen and Juliane Hancke for technical assistance, the Leptin, Schnetz and Uhlirova labs for sharing equipment. We thank all members of Gehring lab for useful discussions.

### FUNDING

International Graduate School in Development Health and Disease (fellowship to A.S.); Thyssen Stiftung (to K.K.S and N.H.G.); Deutsche Forschungsgemeinschaft [GE 2014/2-1 to N.H.G.]. Funding for open access charge: Deutsche Forschungsgemeinschaft.

*Conflict of interest statement.* None declared.

### REFERENCES

1. Swinburne, I.A., Meyer, C.A., Liu, X.S., Silver, P.A. and Brodsky, A.S. (2006) Genomic localization of RNA binding proteins reveals links between pre-mRNA processing and transcription. *Genome Res.*, **16**, 912–921.
2. Komili, S. and Silver, P.A. (2008) Coupling and coordination in gene expression processes: a systems biology view. *Nat. Rev. Genet.*, **9**, 38–48.
3. Luo, M.L., Zhou, Z., Magni, K., Christoforides, C., Rappsilber, J., Mann, M. and Reed, R. (2001) Pre-mRNA splicing and mRNA export linked by direct interactions between UAP56 and Aly. *Nature*, **413**, 644–647.
4. Zhou, Z., Luo, M., Straesser, K. and Katahira, J. (2000) The protein Aly links pre-messenger-RNA splicing to nuclear export in metazoans. *Nature*, **407**, 401–405.
5. Valencia, P., Dias, A.P. and Reed, R. (2008) Splicing promotes rapid and efficient mRNA export in mammalian cells. *Proc. Natl. Acad. Sci. U.S.A.*, **105**, 3386–3391.
6. Le Hir, H., Izaurralde, E., Maquat, L.E. and Moore, M.J. (2000) The spliceosome deposits multiple proteins 20–24 nucleotides upstream of mRNA exon-exon junctions. *EMBO J.*, **19**, 6860–6869.
7. Le Hir, H., Gatfield, D., Izaurralde, E. and Moore, M.J. (2001) The exon-exon junction complex provides a binding platform for factors involved in mRNA export and nonsense-mediated mRNA decay. *EMBO J.*, **20**, 4987–4997.
8. Tange, T.Ø., Shibuya, T., Jurica, M.S. and Moore, M.J. (2005) Biochemical analysis of the EJC reveals two new factors and a stable tetrameric protein core. *RNA*, **11**, 1869–1883.
9. Strässer, K., Masuda, S., Mason, P., Pfannstiel, J., Oppizzi, M., Rodriguez-Navarro, S., Rondón, A.G., Aguilera, A., Struhl, K., Reed, R. *et al.* (2002) TREX is a conserved complex coupling transcription with messenger RNA export. *Nature*, **417**, 304–308.
10. Fleckner, J., Zhang, M., Valcarcel, J. and Green, M.R. (1997) U2AF65 recruits a novel human DEAD box protein required for the U2 snRNP-branchpoint interaction. *Genes Dev.*, **11**, 1864–1872.
11. Libri, D., Graziani, N., Saguez, C. and Boulay, J. (2001) Multiple roles for the yeast SUB2/yUAP56 gene in splicing. *Genes Dev.*, **15**, 36–41.
12. Yamazaki, T., Fujiwara, N., Yukinaga, H., Ebisuya, M., Shiki, T., Kurihara, T., Kioka, N., Kambe, T., Nagao, M., Nishida, E. *et al.* (2010) The closely related RNA helicases, UAP56 and URH49, preferentially form distinct mRNA export machineries and coordinately regulate mitotic progression. *Mol. Biol. Cell*, **21**, 2953–2965.
13. Golovanov, A.P., Hautbergue, G.M., Tintaru, A.M., Lian, L.Y. and Wilson, S.A. (2006) The solution structure of REF2-I reveals interdomain interactions and regions involved in binding mRNA export factors and RNA. *RNA*, **12**, 1933–1948.
14. Kapadia, F., Pryor, A., Chang, T.-H.H. and Johnson, L.F. (2006) Nuclear localization of poly(A)+ mRNA following siRNA reduction of expression of the mammalian RNA helicases UAP56 and URH49. *Gene*, **384**, 37–44.
15. Masuda, S., Das, R., Cheng, H., Hurt, E., Dorman, N. and Reed, R. (2005) Recruitment of the human TREX complex to mRNA during splicing. *Genes Dev.*, **19**, 1512–1517.
16. Cheng, H., Dufu, K., Lee, C.-S.S., Hsu, J.L., Dias, A. and Reed, R. (2006) Human mRNA export machinery recruited to the 5' end of mRNA. *Cell*, **127**, 1389–1400.
17. Nojima, T., Hirose, T., Kimura, H. and Hagiwara, M. (2007) The interaction between cap-binding complex and RNA export factor is required for intronless mRNA export. *J. Biol. Chem.*, **282**, 15645–15651.
18. Taniguchi, I. and Ohno, M. (2008) ATP-dependent recruitment of export factor Aly/REF onto intronless mRNAs by RNA helicase UAP56. *Mol. Cell Biol.*, **28**, 601–608.
19. Viphakone, N., Hautbergue, G.M., Walsh, M., Chang, C.T., Holland, A., Folco, E.G., Reed, R. and Wilson, S.A. (2012) TREX exposes the RNA-binding domain of Nxf1 to enable mRNA export. *Nat. Commun.*, **3**, 1006.
20. Braun, I.C., Herold, A., Rode, M., Conti, E. and Izaurralde, E. (2001) Overexpression of TAP/p15 heterodimers bypasses nuclear retention

- and stimulates nuclear mRNA export. *J. Biol. Chem.*, **276**, 20536–20543.
21. Hautbergue, G.M., Hung, M.L., Golovanov, A.P., Lian, L.Y. and Wilson, S.A. (2008) Mutually exclusive interactions drive handover of mRNA from export adaptors to TAP. *Proc. Natl. Acad. Sci. U.S.A.*, **105**, 5154–5159.
  22. Hautbergue, G.M., Hung, M.-L.L., Walsh, M.J., Snijders, A.P., Chang, C.-T.T., Jones, R., Ponting, C.P., Dickman, M.J. and Wilson, S.A. (2009) UIF, a New mRNA export adaptor that works together with REF/ALY, requires FACT for recruitment to mRNA. *Curr. Biol.*, **19**, 1918–1924.
  23. Chang, C.T., Hautbergue, G.M., Walsh, M.J., Viphakone, N., van Dijk, T.B., Philippen, S. and Wilson, S.A. (2013) Chtop is a component of the dynamic TREX mRNA export complex. *EMBO J.*, **32**, 473–486.
  24. Viphakone, N., Cumberbatch, M.G., Livingstone, M.J., Heath, P.R., Dickman, M.J., Catto, J.W. and Wilson, S.A. (2015) Lusz4 defines a new mRNA export pathway in cancer cells. *Nucleic Acids Res.*, **43**, 2353–2366.
  25. Jimeno, S., Rondon, A.G., Luna, R. and Aguilera, A. (2002) The yeast THO complex and mRNA export factors link RNA metabolism with transcription and genome instability. *EMBO J.*, **21**, 3526–3535.
  26. Dufu, K., Livingstone, M.J., Seebacher, J., Gygi, S.P., Wilson, S.A. and Reed, R. (2010) ATP is required for interactions between UAP56 and two conserved mRNA export proteins, Aly and CIP29, to assemble the TREX complex. *Genes Dev.*, **24**, 2043–2053.
  27. Calabretta, S. and Richard, S. (2015) Emerging roles of disordered sequences in RNA-binding proteins. *Trends Biochem. Sci.*, **40**, 662–672.
  28. Davey, N.E., Van Roey, K., Weatheritt, R.J., Toedt, G., Uyar, B., Altenberg, B., Budd, A., Diella, F., Dinkel, H. and Gibson, T.J. (2012) Attributes of short linear motifs. *Mol. BioSyst.*, **8**, 268–281.
  29. Jonas, S. and Izaurralde, E. (2013) The role of disordered protein regions in the assembly of decapping complexes and RNP granules. *Genes Dev.*, **27**, 2628–2641.
  30. Neduva, V. and Russell, R.B. (2005) Linear motifs: evolutionary interaction switches. *FEBS Lett.*, **579**, 3342–3345.
  31. Gehring, N.H., Lamprinaki, S., Kulozik, A.E. and Hentze, M.W. (2009) Disassembly of exon junction complexes by PYM. *Cell*, **137**, 536–548.
  32. Braun, I.C., Herold, A., Rode, M. and Izaurralde, E. (2002) Nuclear export of mRNA by TAP/NXF1 requires two nucleoporin-binding sites but not p15. *Mol. Cell Biol.*, **22**, 5405–5418.
  33. Steckelberg, A.L. and Gehring, N.H. (2014) Studying the composition of mRNPs in vitro using splicing-competent cell extracts. *Methods*, **65**, 342–349.
  34. Pari, G.S. and Xu, Y. (2004) Gene transfer into mammalian cells using calcium phosphate and DEAE-dextran. *Methods Mol. Biol.*, **245**, 25–32.
  35. Gehring, N.H., Lamprinaki, S., Hentze, M.W. and Kulozik, A.E. (2009) The hierarchy of exon-junction complex assembly by the spliceosome explains key features of mammalian nonsense-mediated mRNA decay. *PLoS Biology*, **7**, e1000120.
  36. Steckelberg, A.-L.L., Boehm, V., Gromadzka, A.M. and Gehring, N.H. (2012) CWC22 connects pre-mRNA splicing and exon junction complex assembly. *Cell Reports*, **2**, 454–461.
  37. Li, Q., Imataka, H., Morino, S., Rogers, G.W. Jr, Richter-Cook, N.J., Merrick, W.C. and Sonenberg, N. (1999) Eukaryotic translation initiation factor 4AIII (eIF4AIII) is functionally distinct from eIF4AI and eIF4AII. *Mol. Cell Biol.*, **19**, 7336–7346.
  38. Katoh, K., Misawa, K., Kuma, K. and Miyata, T. (2002) MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Res.*, **30**, 3059–3066.
  39. Bucher, P., Karplus, K., Moeri, N. and Hofmann, K. (1996) A flexible motif search technique based on generalized profiles. *Comput. Chem.*, **20**, 3–23.
  40. Crooks, G.E., Hon, G., Chandonia, J.M. and Brenner, S.E. (2004) WebLogo: a sequence logo generator. *Genome Res.*, **14**, 1188–1190.
  41. Soding, J., Biegert, A. and Lupas, A.N. (2005) The HHpred interactive server for protein homology detection and structure prediction. *Nucleic Acids Res.*, **33**, W244–248.
  42. Degot, S., Le Hir, H., Alpy, F., Kedinger, V., Stoll, I., Wendling, C., Seraphin, B., Rio, M.-C.C. and Tomasetto, C. (2004) Association of the breast cancer protein MLN51 with the exon junction complex via its speckle localizer and RNA binding module. *J. Biol. Chem.*, **279**, 33702–33715.
  43. Konarska, M.M., Padgett, R.A. and Sharp, P.A. (1984) Recognition of cap structure in splicing in vitro of mRNA precursors. *Cell*, **38**, 731–736.
  44. Ederly, I. and Sonenberg, N. (1985) Cap-dependent RNA splicing in a HeLa nuclear extract. *Proc. Natl. Acad. Sci. U.S.A.*, **82**, 7590–7594.
  45. Ohno, M., Segref, A., Kuersten, S. and Mattaj, I.W. (2002) Identity elements used in export of mRNAs. *Mol. Cell*, **9**, 659–671.
  46. Ghosh, S., Obrdlik, A., Marchand, V. and Ephrussi, A. (2014) The EJC binding and dissociating activity of PYM is regulated in Drosophila. *PLoS Genet.*, **10**, e1004455.
  47. Stutz, F., Bachi, A., Doerks, T., Braun, I.C., Séraphin, B., Wilm, M., Bork, P. and Izaurralde, E. (2000) REF, an evolutionary conserved family of hnRNP-like proteins, interacts with TAP/Mex67p and participates in mRNA nuclear export. *RNA*, **6**, 638–650.
  48. Rodrigues, J.P., Rode, M., Gatfield, D., Blencowe, B.J., Carmo-Fonseca, M. and Izaurralde, E. (2001) REF proteins mediate the export of spliced and unspliced mRNAs from the nucleus. *Proc. Natl. Acad. Sci. U.S.A.*, **98**, 1030–1035.
  49. Wichmann, I., Garcia-Lozano, J.R., Respaldiza, N., Gonzalez-Escribano, M.F. and Nuñez-Roldan, A. (1999) Autoantibodies to transcriptional regulation proteins DEK and ALY in a patient with systemic lupus erythematosus. *Hum. Immunol.*, **60**, 57–62.
  50. Jonas, S. and Izaurralde, E. (2013) The role of disordered protein regions in the assembly of decapping complexes and RNP granules. *Genes Dev.*, **27**, 2628–2641.
  51. Han, T.W., Kato, M., Xie, S., Wu, L.C., Mirzaei, H., Pei, J., Chen, M., Xie, Y., Allen, J., Xiao, G. *et al.* (2012) Cell-free formation of RNA granules: bound RNAs identify features and components of cellular assemblies. *Cell*, **149**, 768–779.
  52. Folco, E.G., Lee, C.-S.S., Dufu, K., Yamazaki, T. and Reed, R. (2012) The proteins PDIP3 and ZC11A associate with the human TREX complex in an ATP-dependent manner and function in mRNA export. *PLoS One*, **7**, e43804.
  53. McCloskey, A., Taniguchi, I., Shinmyozu, K. and Ohno, M. (2012) hnRNP C tetramer measures RNA length to classify RNA polymerase II transcripts for export. *Science*, **335**, 1643–1646.
  54. Ohno, M., Segref, A., Bachi, A., Wilm, M. and Mattaj, I.W. (2000) PHAX, a mediator of U snRNA nuclear export whose activity is regulated by phosphorylation. *Cell*, **101**, 187–198.
  55. Lee, K.-M.M., Hsu, I.-W.a. and Tarn, W.-Y.Y. (2010) TRAP150 activates pre-mRNA splicing and promotes nuclear mRNA degradation. *Nucleic Acids Res.*, **38**, 3340–3350.
  56. Richardson, R., Denis, C.L., Zhang, C.L., Nielsen, M.E., Chiang, Y.C., Kierkegaard, M., Wang, X., Lee, D.J., Andersen, J.S. and Yao, G. (2012) Mass spectrometric identification of proteins that interact through specific domains of the poly(A) binding protein. *Mol. Genet. Genomics*, **287**, 711–730.
  57. Wiegand, H.L., Lu, S. and Cullen, B.R. (2003) Exon junction complexes mediate the enhancing effect of splicing on mRNA expression. *Proc. Natl. Acad. Sci. U.S.A.*, **100**, 11327–11332.
  58. Strein, C., Alleaume, A.M., Rothbauer, U., Hentze, M.W. and Castello, A. (2014) A versatile assay for RNA-binding proteins in living cells. *RNA*, **20**, 721–731.
  59. Neduva, V. and Russell, R.B. (2006) Peptides mediating interaction networks: new leads at last. *Curr. Opin. Biotechnol.*, **17**, 465–471.
  60. Tompa, P. and Csermely, P. (2004) The role of structural disorder in the function of RNA and protein chaperones. *FASEB J.*, **18**, 1169–1175.
  61. Merz, C., Urlaub, H., Will, C.L. and Lührmann, R. (2007) Protein composition of human mRNPs spliced in vitro and differential requirements for mRNP protein recruitment. *RNA*, **13**, 116–128.
  62. Singh, G., Kucukural, A., Cenik, C., Leszyk, J.D., Shaffer, S.A., Weng, Z. and Moore, M.J. (2012) The cellular EJC interactome reveals higher-order mRNP structure and an EJC-SR protein nexus. *Cell*, **151**, 750–764.
  63. Varia, S., Potabathula, D., Deng, Z., Bubulya, A. and Bubulya, P.A. (2013) Btf and TRAP150 have distinct roles in regulating subcellular mRNA distribution. *Nucleus*, **4**, 229–240.
  64. Kwon, J., Xiang, S., Kato, M., Wu, L., Theodoropoulos, P., Wang, T., Kim, J., Yun, J., Xie, Y. and McKnight, S.L. (2014) Poly-dipeptides

- encoded by the C9orf72 repeats bind nucleoli, impede RNA biogenesis, and kill cells. *Science*, **345**, 1139–1145.
65. Johnson,S.A., Cubberley,G. and Bentley,D.L. (2009) Cotranscriptional recruitment of the mRNA export factor Yra1 by direct interaction with the 3' end processing factor Pef11. *Mol. Cell*, **33**, 215–226.
66. Shi,Y., Di Giammartino,D.C., Taylor,D., Sarkeshik,A., Rice,W.J., Yates,J.R. 3rd, Frank,J. and Manley,J.L. (2009) Molecular architecture of the human pre-mRNA 3' processing complex. *Mol. Cell*, **33**, 365–376.
67. Cooper-Knock,J., Walsh,M.J., Higginbottom,A., Robin Highley,J., Dickman,M.J., Edbauer,D., Ince,P.G., Wharton,S.B., Wilson,S.A., Kirby,J. *et al.* (2014) Sequestration of multiple RNA recognition motif-containing proteins by C9orf72 repeat expansions. *Brain*, **137**, 2040–2051.
68. Rossi,S., Serrano,A., Gerbino,V., Giorgi,A., Di Francesco,L., Nencini,M., Bozzo,F., Schinina,M.E., Bagni,C., Cestra,G. *et al.* (2015) Nuclear accumulation of mRNAs underlies G4C2-repeat-induced translational repression in a cellular model of C9orf72 ALS. *J. Cell Sci.*, **128**, 1787–1799.
69. Zhang,K., Donnelly,C.J., Haeusler,A.R., Grima,J.C., Machamer,J.B., Steinwald,P., Daley,E.L., Miller,S.J., Cunningham,K.M., Vidensky,S. *et al.* (2015) The C9orf72 repeat expansion disrupts nucleocytoplasmic transport. *Nature*, **525**, 56–61.