4516–4526 *Nucleic Acids Research, 2014, Vol. 42, No. 7* *Published online 26 January 2014*
doi:10.1093/nar/gku073

Functional signature for the recognition of specific target mRNAs by human Staufen1 protein

Susana de Lucas[1,2], Juan Carlos Oliveros[3], Mónica Chagoyen[4] and Juan Ortín[1,2,*]

[1]Departamento de Biología Molecular y Celular, Centro Nacional de Biotecnología (CSIC), C/Darwin 3, Campus Cantoblanco, 28049 Madrid, Spain, [2]Ciber de Enfermedades Respiratorias (ISCIII), Mallorca, Spain, [3]Servicio de Genómica Computacional, Centro Nacional de Biotecnología (CSIC), C/Darwin 3, Campus Cantoblanco, 28049 Madrid, Spain and [4]Bioinformática de Sistemas, Centro Nacional de Biotecnología (CSIC), C/Darwin 3, Campus Cantoblanco, 28049 Madrid, Spain

Received July 21, 2013; Revised December 23, 2013; Accepted December 27, 2013

ABSTRACT

Cellular messenger RNAs (mRNAs) are associated to proteins in the form of ribonucleoprotein particles. The double-stranded RNA-binding (DRB) proteins play important roles in mRNA synthesis, modification, activity and decay. Staufen is a DRB protein involved in the localized translation of specific mRNAs during *Drosophila* early development. The human Staufen1 (hStau1) forms RNA granules that contain translation regulation proteins as well as cytoskeleton and motor proteins to allow the movement of the granule on microtubules, but the mechanisms of hStau1-RNA recognition are still unclear. Here we used a combination of affinity chromatography, RNAse-protection, deep-sequencing and bioinformatic analyses to identify mRNAs differentially associated to hStau1 or a mutant protein unable to bind RNA and, in this way, defined a collection of mRNAs specifically associated to wt hStau1. A common sequence signature consisting of two opposite-polarity Alu motifs was present in the hStau1-associated mRNAs and was shown to be sufficient for binding to hStau1 and hStau1-dependent stimulation of protein expression. Our results unravel how hStau1 identifies a wide spectrum of cellular target mRNAs to control their localization, expression and fate.

INTRODUCTION

All along their life cycle, from initial synthesis in the nucleus to final degradation in the cytoplasm, messenger RNAs (mRNAs) are associated to many proteins to form ribonucleoprotein particles [for reviews see (1–3)]. Many of these proteins belong to the hnRNP and SR families and participate in the regulation of essentially all steps of the mRNA biological activity. In addition, other class of RNA-binding proteins, which contain a conserved double-stranded RNA (dsRNA)-binding domain (DRBD), also play important roles in mRNA synthesis, activity and decay (4). This class includes nuclear proteins like NFAR and RHA involved in transcription and the ADARs responsible for RNA editing, as well as cytoplasmic proteins like PKR, PCT and Dicer, which are involved in the regulation of gene expression and stability. For most members of this family, the interaction with dsRNA is not sequence-specific as documented by the atomic structure of dsRNA bound to various RBDs (5,6). The DRB protein Staufen was initially identified in *Drosophila melanogaster* as a maternal factor essential for the proper localization of *bicoid* and *oskar* mRNAs during the formation of the anteroposterior axis (7,8). The corresponding mammalian homologues Staufen1 and Staufen2 participate in various aspects of the mRNA life cycle. Thus, human Staufen1 (hStau1) (9,10) has a role in processes such as mRNA transport and localized translation in polarized cells (11–13), stimulation of translation of specific mRNAs (14) and mRNA decay through a mechanism called Staufen-mediated decay (SMD), in which specific mRNAs are targeted to degradation by a protein complex including hStau1 (15,16). In addition, hStau1 has a role in infectious diseases, as it is important for the replication of RNA viruses such as HIV and influenza virus (17–20). The participation of hStau1 in the processes indicated above takes place in the context of large ribonucleoprotein complexes or RNA granules that are functionally diverse and show a dynamic protein and RNA composition [for reviews see (21–23)]. These include (i) RNA transport granules, structures that contain cytoskeleton and motor proteins as well as regulators of translation such as PABP and FMRP, suggesting their role in RNA transport and localized translation (11–13,24), (ii) stress granules (25,26) and (iii) hStau1-dependent RNA decay complexes (15,16).

*To whom correspondence should be addressed. Tel: +34915854557; Fax: +34915854506; Email: jortin@cnb.csic.es

The observation that dmStaufen DRBD3 binds dsRNA without any sequence specificity (5), and yet specific fly mRNAs, like *bicoid*, *oskar* and *prospero*, are targets of dmStaufen indicates that the RNA signature required for dmStaufen recognition is still unclear. Likewise, hStau1 binds RNA *in vitro* without apparent sequence specificity (9), but a large number of human mRNAs have been found associated to hStau1 intracellular complexes *in vivo*, as reported by transcriptomic analyses of hStau1 immunopurified samples (15,27) (S. de Lucas, unpublished results). However, for many of these hStau1-associated mRNAs the question remains whether they are directly recognized by hStau1 or rather by other RNA-binding protein(s) present in the purified intracellular complexes. In this report we have analysed by deep-sequencing the RNAs present in affinity-purified hStau1 intracellular complexes and compared with those associated to analogous complexes containing a mutant hStau1 protein defective in RNA binding. Bioinformatic analyses of the sequence data identified mRNAs preferentially associated to wild-type hStau1 complexes. On the other hand, deep-sequencing of the RNAs associated to RNAse-treated purified hStau1 complexes indicated that most of the protected sequences mapped to the 3′-untranslated regions (UTRs) of the associated mRNAs and identified a sequence signature containing at least two opposite-polarity motifs analogous to human Alu sequences. Functional analyses verified that this signature is important for mRNA association to hStau1 complexes and plays a role in the efficient expression of the mRNAs.

MATERIALS AND METHODS

Biological materials

The HEK293T (28) cell line was cultured as previously described (13) and transfected with DNA plasmids using the calcium phosphate method (29). Plasmids pChStaufen-TAP, pChMut-TAP and pC-TAP as well as rabbit and rat antibodies specific for hStau1 have been previously described (13,18).

Protein analyses

The tandem affinity purification (TAP) of recombinant hStaufen1-TAP an Mut-TAP proteins has been previously described (13,18). Briefly, cell extracts were prepared in 50 mM Tris-HCl pH 7.5, 100 mM NaCl, 5 mM EDTA (TNE) buffer containing 0.5% NP-40, 1 mM dithiothreitol (DTT) in the presence of protease and RNAse inhibitors for 30 min at 4°C. The lysates obtained were incubated with IgG-Sepharose (GE Healthcare) for 12 h at 4°C, washed in IPP-150 (150 mM NaCl, 10 mM Tris ± HCl, 0.1% Igepal, pH 8.0) buffer and digested with 1 U of Tobacco Etch Virus (TEV) protease per 10^7 cells for 3 h at room temperature. Supernatants were incubated with Calmodulin-agarose (Stratagene) resin for 12 h at 4°C. In the corresponding samples, RNAse T1 digestion was carried out for 1 h at room temperature before Calmodulin resin elution. After extensive washings, protein complexes were eluted in a buffer containing 3 mM EGTA (13). The purified proteins were analysed by electrophoresis in 10% polyacrylamide gel and either western blot or silver staining.

RNA purification, cDNA synthesis and deep sequencing

The RNAs associated to purified complexes were obtained by treatment of Calmodulin-agarose resin with 0.2 mg/ml Proteinase K and 0.5% sodium dodecyl sulphate in TNE for 1 h at 37°C. After phenol extraction, RNAs were precipitated with 2 ethanol volumes and 20 μg of glycogen (Roche) and resuspended in diethyl pyrocarbonate (DEPC)-treated H_2O. Before cDNA synthesis, each RNA preparation was monitored using the Agilent 2100 Bioanalyzer (Agilent technologies). RNA samples of hStau1-TAP (treated or not with RNAse T1) and the untreated hStau1-Mut (100 ng each) were used to generate cDNA libraries according to Illumina mRNA-seq protocol at Fasteris SA Company (Plan-Les-Ouates, Switzerland), but the poly (A+) RNA purification step was skipped. Furthermore, mRNA fragmentation was omitted in the T1-RNAse–treated hStau1-associated RNA, as it already contained RNAs ~100 nt in size. cDNA libraries were subjected to ends repair reaction, Illumina adapter ligation and further purified from polyacrylamide gels. This purification step excluded the identification of putative small RNAs present in the samples. cDNA libraries were sequenced at Fasteris SA in an Illumina HiSeq2000 sequencer. Reads length were 100 bp in a single-read run cycle.

Quantitative reverse transcriptase-polymerase chain reaction

Quantitative reverse transcriptase-polymerase chain reactions (qRT-qPCRs) were used to determine mRNA and 28s ribosomal RNA (rRNA) levels in both total cell RNAs and RNAs associated to TAP-purified hStau1 complexes. First, cDNA was generated using the High-Capacity cDNA Reverse Transcription kit from Applied Biosystems. cDNA was then diluted to perform qPCR with SYBRGreen PCR Master Mix reagents (Applied Biosystems) and specific primers to detect mRNAs and 28s rRNA (Supplementary Table S3).

Plasmid construction

The 3′-UTRs of MMACHC, RAB2B and FCF1 mRNAs were amplified using Titan One RT-PCR System (Roche Applied Biosystems) following the manufacturer instructions. The specific primers used (Supplementary Table S3) contained an Xho I site (Forward primer) or a Kpn I site (Reverse primer) to facilitate the posterior cloning of RT-PCR products into the Xho I/Kpn I sites of pEGFP-C1 plasmid (Clontech Laboratories). This procedure generated the corresponding plasmids pEGFP-C-MMACHC, pEGFP-C-RAB2B and pEGFP-C-FCF1. To remove or invert the protected Alu repeats in the 3′-UTR of GFP-MMACHC or GFP-FCF1 constructs, specific restriction enzyme sites were inserted flanking each of the Alu sequences using the Quick Change Mutagenesis Kit from Stratagene. Plasmids were then digested and religation performed with T4 polynucleotide ligase to obtain deletion mutants or clones with randomly

inserted Alu repeats that were then sequenced to determine the Alu polarity.

Bioinformatic analyses

Deep sequencing data analysis. The output of the Illumina sequencer was composed of reads up to 100 nt long, some of which contained fragments of Illumina adapters' sequences. These adapter sequences were removed and reads of 50–100 nt long were merged in a single fastq file per sample. Raw reads were aligned to representative human rRNA sequences (28S, 18S, 5.8S and 5S) with BWA version 0.5.9rc1 (30) with default parameters. Those aligned with ribosomal sequences were not considered for ulterior analysis steps. Filtered reads were aligned to human genome (hg19) with BWA version 0.5.9rc1 with default parameters. After alignment, read counts that matched gene regions were normalized by the trimmed mean of M values method of edgeR package (31). Gene annotations and coordinates were obtained from UCSC database using the 'RefGene' table. Similarly, lincRNAs were identified using the UCSC lincRNA Transcripts database. Normalized counts per million of hStau1-wt and hStau1-mut1 samples were compared and Fold Changes calculated, considering the hStau1-wt sample as control. Common dispersion and exact test for a negative binomial distribution were used to determine the statistical significance of the results as described in edgeR. Finally, *P*-values were adjusted by FDR (32). Values of FDR < 0.05 and Fold Change < -2 were used as numerical thresholds for defining Stau (wt)-specific mRNAs. FIESTA viewer (http://bioinfogp.cnb.csic.es/tools/FIESTA) was used to visualize scatter plots, integrate all numerical results (counts per million, Fold Changes and FDR) and to evaluate the effect of applying different numerical thresholds to the RNAseq results. Integrative Genomics Viewer (IGV) 2.1 (33) was used to visualize reads aligned to human gene structures. The *read_distribution* program from RSeQC package (34) was used to classify and count aligned reads into specific genomic regions: coding exons, 5′-UTR, 3′-UTR and introns.

Gene ontology classification

The differentially represented mRNAs were analysed by DAVID Functional Annotation Bioinformatic Microarray Analysis tool from the National Institute of Allergy and Infectious Diseases (NIAID-NIH) (35).

Sequence alignment

To determine the presence of common sequences within the identified protected regions, the reverse complement of each protected sequence was obtained. Both protected sequences and their reverse complement sequences were aligned with MUSCLE (36). A preliminary short motif (~130 nt) was extracted from the alignment. The presence of such motif in the protected sequences and corresponding 3′-UTR was further assessed with Infernal (37). Visual inspection of the preliminary motif confirmed the presence of the characteristic poly-A segment joining FLAM and FRAM monomers in the Alu primary structure. To confirm the motif as a segment of a generalized Alu, protected sequences were analysed with RepeatMasker (http://www.repeatmasker.org). Alu repeat subfamilies were then identified in both protected sequences and corresponding 3′-UTR sequences with RepeatMasker. A diagram showing the localization of Alu repeats and protected sequences in 3′-UTR was created with DOG 1.0 (38). A final full-length alignment of representative human Alu sequences (from RepeatMasker) and protected sequences in standard orientation was obtained with MUSCLE, using full-length Alu repeats in 3′-UTR sequences for optimal alignment. A Neighbourhood joining tree of protected sequences was obtained with Jalview (39).

RESULTS

To identify mRNAs specifically associated to hStau1 and decipher the hStau1-specific RNA recognition site we used the experimental strategy described in Figure 1, which involved (i) affinity purification of either wt or RNA binding-defective hStau1 intracellular complexes, (ii) deep-sequencing of the associated RNAs, (iii) bioinformatic filtering for differential mRNA presence in wt-versus-mutant hStau1 complexes and resistance to RNAse, (iv) search for common features in the specific protected RNAs and (v) experimental validation and testing of functional relevance.

Identification of human mRNAs specifically associated to hStau1

The intracellular complexes containing either wt hStau1 or a RBD3 mutant protein, which binds RNA with ~10-fold less affinity (mut1) and is expressed to levels similar to wt protein (Figure 2A; <10% difference in four independent experiments) (13), were purified using the TAP tag two-step affinity chromatography process. As previously reported, a complex pattern of proteins was found associated to hStau1 (13) (Figure 2B, hStau1), while only background signal was observed for a control purification in which only the TAP tag was expressed (Figure 2B, Crtl). Furthermore, similar protein composition was found in complexes containing RBD3 mutant hStau1 (Figure 2B, hMut1). To identify the RNA species interacting with hStau1 protein and gain insight on the potential RNA interaction regions, an RNAse T1 treatment was included before final elution of either wt or mut1 hStau1-TAP protein, to remove RNA sequences not protected within the complexes. The protein complexes obtained after washing away the RNAse showed a protein pattern similar to untreated hStau1 complexes, as revealed by silver staining (Figure 2B). Essentially no RNA was detected associated to the control TAP complexes and reduced amounts were found associated to mut1- as compared with wt-hStau1. The size of RNAs associated to hStau1 complexes was in the range of hundreds of nt, but it was reduced to ~100 nt after RNAse T1 treatment (Figure 2C). Deep-sequencing was performed by the Illumina mRNA seq protocol on the RNA samples associated to wt hStau1 (treated or not with RNAse) and to mut1 hStau1, but the RNA

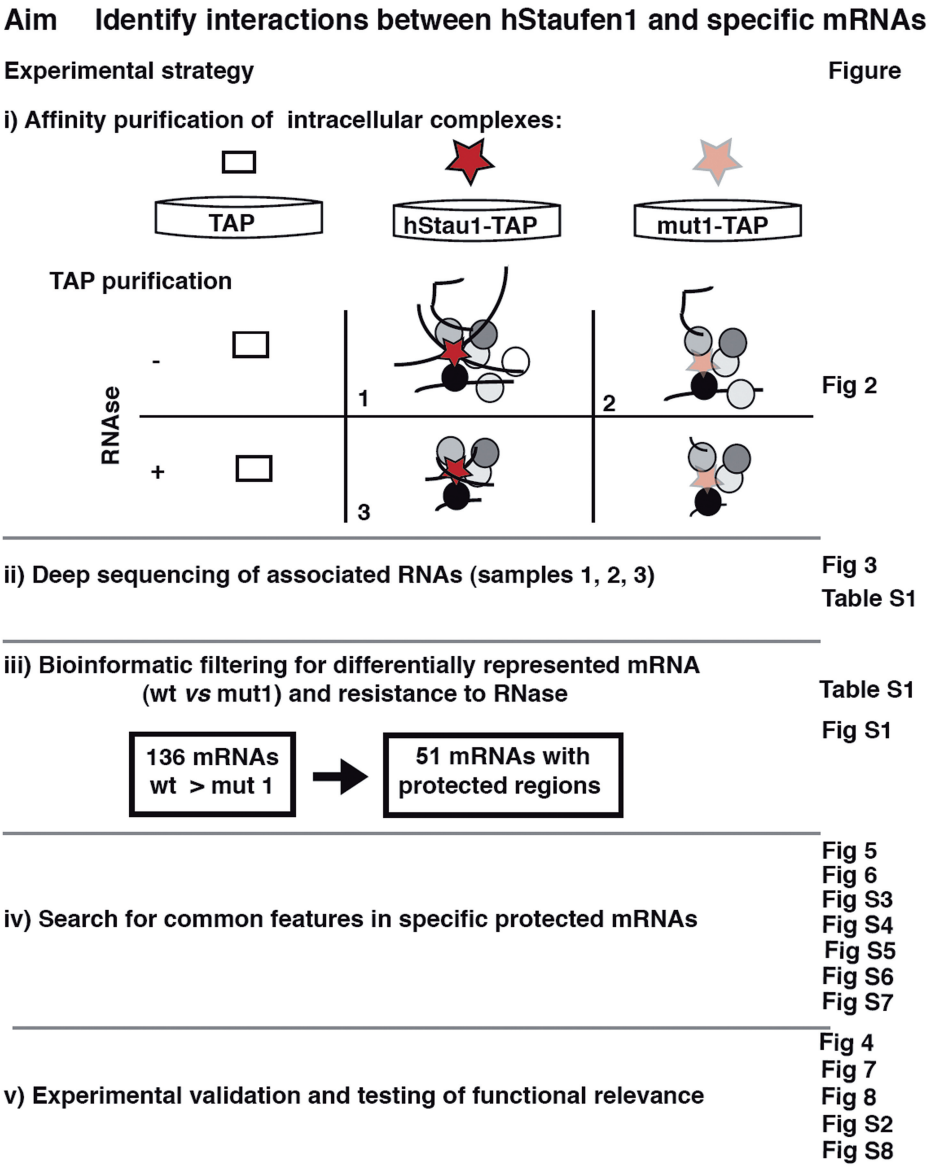

Figure 1. Diagram of the experimental approach. The main objective of the work and the various steps of the experimental approach are presented. The corresponding items in the paper are referred to at the right.

samples obtained from TAP-expressing controls or associated to mut1 hStau1 and treated with RNAse did not contain enough RNA to be processed.

The distribution of sequences corresponding to rRNAs and non-rRNAs are shown in Figure 3A. Consistent with the presence of ribosomal subunits in hStau1 complexes (13), rRNAs were highly represented in wt hStau1 complexes and this was also the case in mut1-associated RNAs, suggesting that the mutation in hStau1-RBD3 is not sufficient to abolish ribosome binding to RNA complexes. However, RNAse treatment reduced ~10-fold the percentage of rRNAs sequences detected, indicating that the association of rRNAs to the granule is more susceptible to RNAse treatment than the association of mRNAs. As rRNAs are normally more RNAse-resistant than mRNAs, these results suggest that RNAse treatment releases ribosomes bound to the complex through mRNA-mediated interactions. To further analyse hStau1-associated RNA, the rRNA reads were filtered out and the remaining sequence data were aligned to the human genome. A total of 3534 and 3733 mRNAs were detected in the wt and mut1 hStau1-associated RNAs, respectively, considering 10 as a minimal cut-off value for the number of reads per mRNA. A considerable over-representation of sequences corresponding to 5′-UTRs, 3′-UTRs and coding exons was observed for the reads from wt or mut1 hStau1-associated RNA (Figure 3B), whereas intron sequences were under-represented, consistent with the cytoplasmic localization of hStau1 complexes.

To compare the number of reads aligned to RNAs associated to wt or mut1 hStau1 we used the coverage

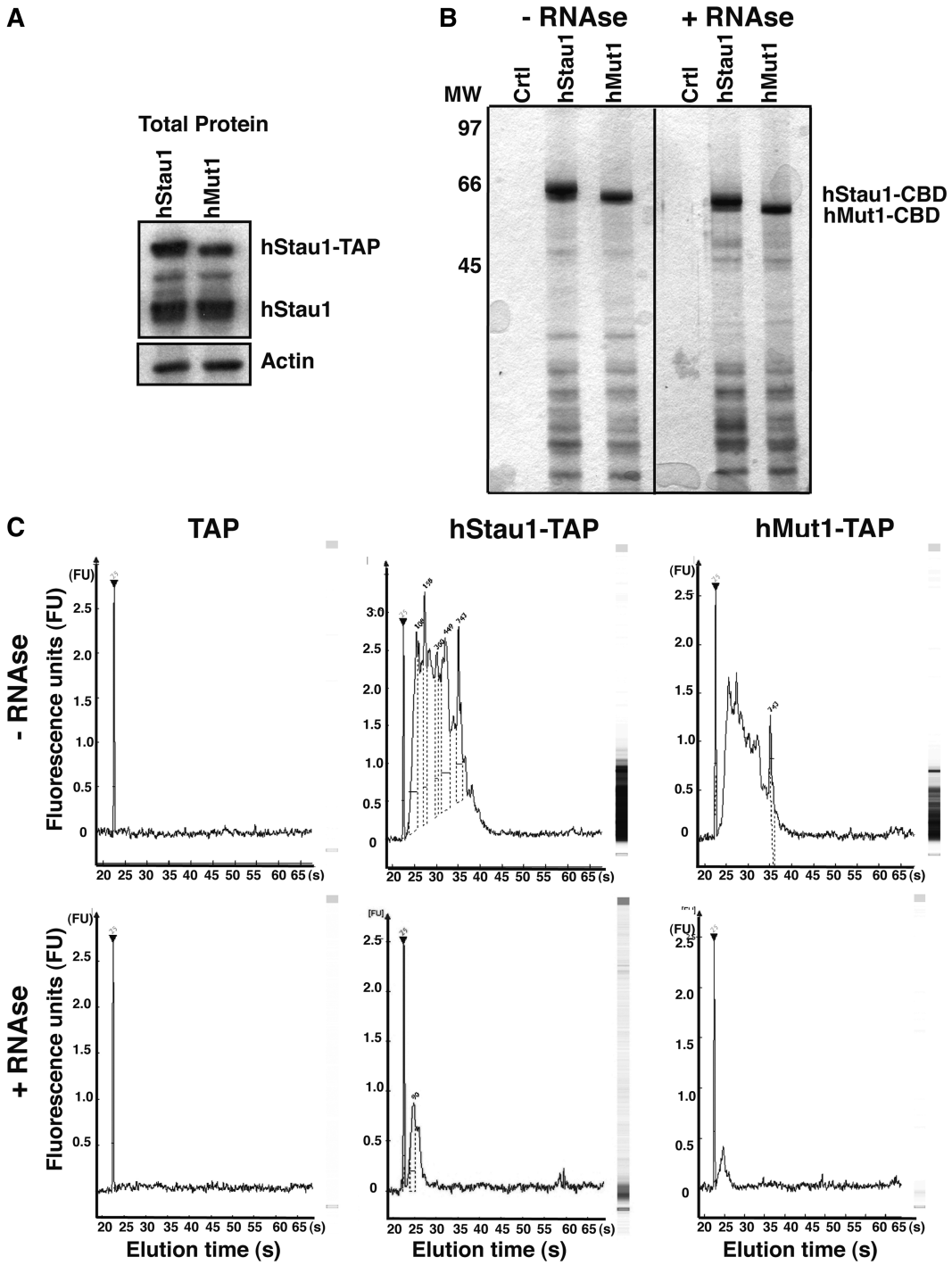

Figure 2. Protein and RNA patterns after TAP purification of hStau1 complexes. Cultures of HEK293T cells were transfected with plasmids expressing wt-hStau1-TAP, mut1-hStau1-TAP or the TAP tag as a control. The corresponding intracellular complexes were purified by affinity chromatography including or not and RNAse T1 treatment. (**A**) Western blot analysis of hStau1-TAP and Mut-TAP expression levels in total cell extracts. (**B**) Analysis by silver staining of the purified complexes. The size of the molecular weight markers is indicated to the left. (**C**) Analysis of the RNAs associated to each complex by Bioanalyser.

of the wt hStau1 sample as reference. In this way we identified 136 mRNAs (Supplementary Table S1) as well as 58 lincRNAs (Supplementary Table S2) under-represented in the mut1 hStau1-associated RNA (Fold Change < -2; FDR < 0.05), suggesting that the RNA binding capacity of hStau1 is important for their association to the RNA complexes (Supplementary Table S1). In contrast, 79 mRNAs were found enriched in the mut1 hStau1-associated RNA (Fold Change > 2; FDR < 0.05). These mRNAs would be associated to the mutant hStau1 complexes by other RNA binding proteins and were not analysed further (Supplementary Table S1). To ascertain the potential specificity of the wt hStau1-mRNA association, we compared their Gene Ontology classification

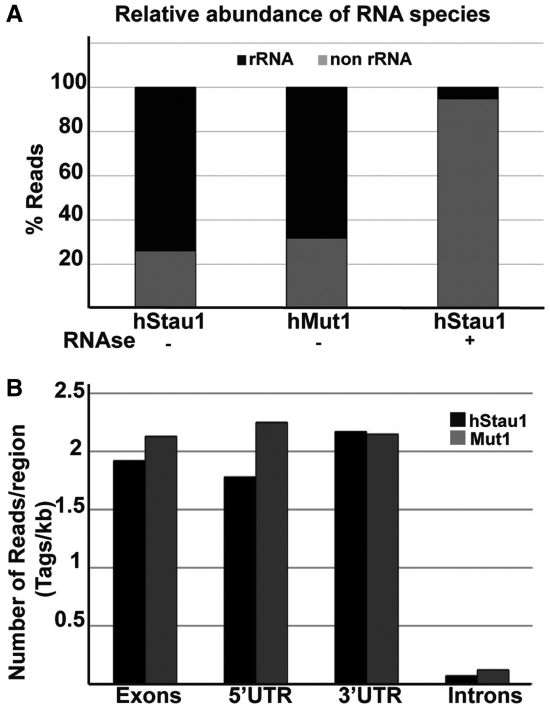

Figure 3. Gene assignment of RNAs associated to hStau1 complexes. The deep-sequencing data corresponding to wt- or mut1-hStau1 associated RNAs were filtered by comparison with the human genome. (**A**) Proportion of sequence reads corresponding to ribosomal or non-rRNAs. (**B**) Proportions of non-rRNAs mapping to 5′-, 3′-UTRs, coding exons or introns.

(40) with the general transcriptome. The results are presented in Supplementary Figure S1 and show that the mRNAs preferentially associated to wt hStau1 are enriched in specific functional, molecular function or cell compartment categories, like protein translation, ribosome or ribonucleoprotein complexes. These results suggest that the set of hStau1-associated mRNAs is not a random representation of the human transcriptome but rather a specific group of mRNAs. To experimentally validate the results obtained by deep sequencing, the presence of selected mRNAs in purified wt hStau1 complexes was ascertained by RT-qPCR, using mRNAs not detected in the hStau1 complexes as controls. As shown in Figure 4, the FCF1, BPNT, MAGT1 and MMACHC mRNAs were over-represented in wt hStau1 as compared with mut1-associated RNA samples, whereas the ANKRD10 mRNA used as a control did not associate to either complex.

A sequence signature is recognized in hStau1-associated mRNAs

Deep sequencing technology provides an extremely powerful tool to identify the mRNAs present in hStau1 complexes, and also to determine the specific regions of the mRNA involved in RNA–protein interactions, when combined with RNAse protection assays. The Integrative Genomics Viewer (IGV 2.1) (33,41) was used to visualize the localization of reads in each of the 136 mRNAs specifically associated to wt hStau1. Among these target mRNAs, 51 showed at least one well-populated protected

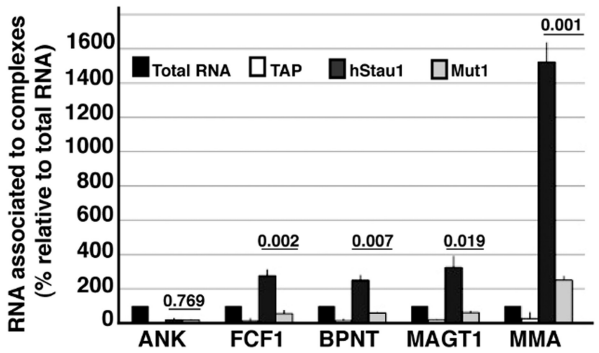

Figure 4. Verification of mRNA association to hStau1 complexes. Either total cell RNA of RNAs associated to wt-, mut1-hStau1 or TAP control purified complexes were used for RT-qPCR using primers specific for FCF1, BPNT1, MMACHC or MAGT as representative of mRNAs preferentially associated to wt-hStau1 complexes or ANKDR10 as a non-associated mRNA control. The concentration of each mRNA is compared with that found in total cell RNA as a reference. The values represent averages and standard deviations of three experiments.

region 70–300 nt in length when the hStau1-associated RNAse-treated RNA sample was analysed (see Supplementary Table S1 and Supplementary Figure S2 as an example), but the remaining 85 mRNAs did not contain sufficient protected reads to warrant a correct analysis and were not considered further. In contrast, among the 58 lincRNAs specifically associated to wt hStau1 only three contained a protected region (Supplementary Table S2; see example in Supplementary Figure S3). Whereas the sequence reads of these 51 mRNAs specifically associated to hStau1 localized mainly to coding exons and 3′-UTRs, the corresponding reads in the RNAse-treated RNA localized mainly to the 3′-UTRs (Figure 5A; see Supplementary Figure S2 as an example). All together, these results suggest that the RBD3-dependent interaction of these mRNAs to hStau1 complexes takes place through RNA signatures localized at their 3′-UTRs.

Alignment of the RNAse-protected sequences found in the 51 mRNAs specifically associated to wt hStau1 showed the presence of a conserved sequence motif ~130 nt in length with a high homology among the various protected regions (Figure 5B). Searches of the human genome using the conserved sequence motif identified Alu repetitive elements abundantly present in the human genome as the best hit (Supplementary Figure S4; see Supplementary Figure S5 for a clustering analysis of the protected sequences). Thirty-one of the hStau1-associated RNAs contained more than one of these sequence motifs (Supplementary Table S1; Figure 6A) and the most prevalent pattern is to find at least two Alu motifs in opposite orientations, a pattern which we call an Alu-signature (Figure 6B; Supplementary Figure S6) and this was also the case for one of the lincRNAs specifically associated to hStau1 complexes. Furthermore, such an Alu-signature was significantly more abundant in the wt- versus mut1-associated mRNAs (Supplementary Table S1 and Supplementary Figure S7). Of note, a search for Alu repetitive elements within the hStau1-associated RNAs

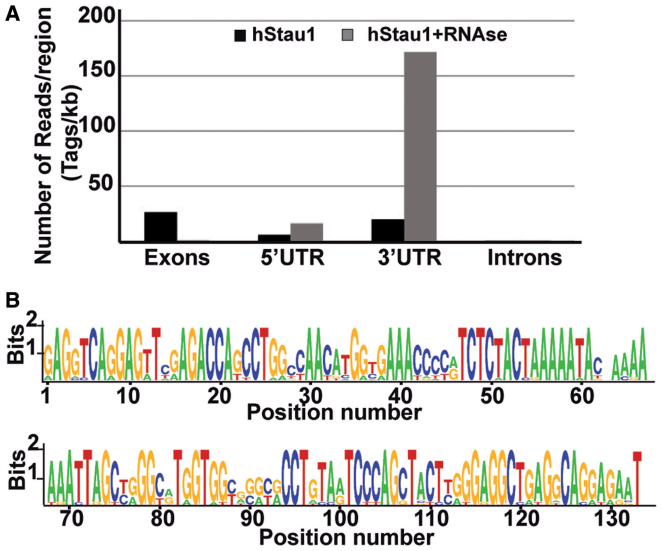

Figure 5. Comparative sequence analysis of the mRNAs specifically associated to wt-hStau1 complexes. (**A**) Localization of sequence reads in the mRNAs preferentially associated to wt-hStau1 complexes, either treated (grey columns) or untreated with RNAse (black columns). (**B**) Sequence motif found by alignment of the protected sequences present in the mRNAs preferentially associated to wt-hStau1 complexes (see Supplementary Figures S4, S5).

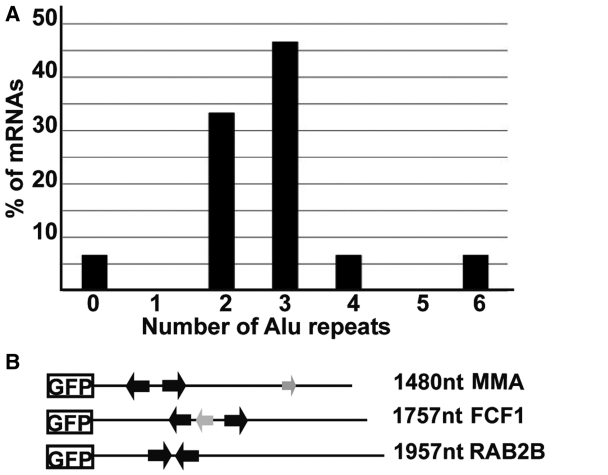

Figure 6. Number and patterns of Alu repeats in the mRNAs specifically associated to wt-hStau1 complexes. (**A**) Number of protected sequence motifs present in the 51 mRNAs preferentially associated to wt-hStau1 complexes (see Supplementary Figure S6). (**B**) Pattern of sequence motifs present in a selection of mRNAs preferentially associated to wt-hStau1 complexes (see Supplementary Figure S6). Arrows show the polarity of the motifs. Black indicates motifs protected from RNAse degradation within hStau1 complexes and grey indicates unprotected motifs.

revealed the presence of additional Alu sequences that were not RNAse-protected within the hStau1 intracellular complexes (Figure 6B; Supplementary Figure S6).

Functional relevance of the hStau1-specific sequence signature

To test the potential relevance of the sequence signature identified above for the functionality of the hStau1-associated mRNAs, the 3′-UTRs of a selected gene set including FCF1, MMACHC and RAB2B were fused downstream of a GFP reporter driven by a strong constitutive promoter (Figure 6B) and used for transfection experiments in human HEK293T cells. The consequences of including these 3′-UTRs for GFP mRNA association to hStau1 were analysed by over expression of hStau1-TAP and affinity purification of the intracellular complexes. A 5- to 7-fold increase of specific GFP mRNA association to hStau1 was observed for all 3′-UTRs tested when compared with a GFP control mRNA (Figure 7A). Under the conditions used the association of GFP mRNAs to hStau1 was 10- to 100-fold higher than that found for a TAP control (Supplementary Figure S8). To analyse the consequences of hStau1 over expression for the fate and expression of the various recombinant GFP mRNAs, the accumulation of GFP protein and mRNA were determined in total cell extracts. The results are presented in Figure 7 and indicate that the amounts of GFP mRNAs containing any of the selected 3′-UTRs did not change by increasing the cellular levels of hStau1 (Figure 7B). However, the accumulation of GFP protein was enhanced by 2- to 3-fold only when the selected 3′-UTRs were included (Figure 7C, D). These results indicate that the 3′-UTRs of the mRNAs associated to hStau1 are relevant for the binding to the complexes and for the expression of the encoded genes.

Next, the structure–function relationship of the sequence signature was analysed by mutation. The protected Alu motifs present in the 3′-UTRs of MMACHC and FCF1 were either deleted (Figure 8A) or inverted (Figure 8C) in the GFP constructs, and hStau1-association experiments similar to those described in Figure 7 were performed. In the case of GFP-MMACHC mRNA, deletion of either Alu motif or double deletion resulted in significantly reduced mRNA association to hStau1 complexes (Figure 8B). Inversion of Alu I also reduced association, but inversion of Alu II motif only caused a small, not significant, decrease. These results could mean that the third Alu motif, normally not protected (Figure 6B and Supplementary Figure S6), might rescue flipping of Alu II motif. Opposite results were obtained for the GFP-FCF1 construct. The association to hStau1 complexes was either unaltered or enhanced as compared with wt when Alu I or Alu II were deleted or flipped (Figure 8C, D).

DISCUSSION

The localization and regulated translation of mRNAs is a general mechanism to allow the spatio-temporal regulation of protein expression that is used in all Eukarya, from yeast to humans (42,43). It is important for a variety of processes, including embryonic development, cell differentiation and motility and synaptic plasticity (44–47). Although it was originally described for a small number of specific mRNAs, like ASH1 in yeast, *bicoid* in the fly, β-actin in fibroblasts or Vg1 in *Xenopus* (8,47–49), it has later been recognized as a mechanism affecting a

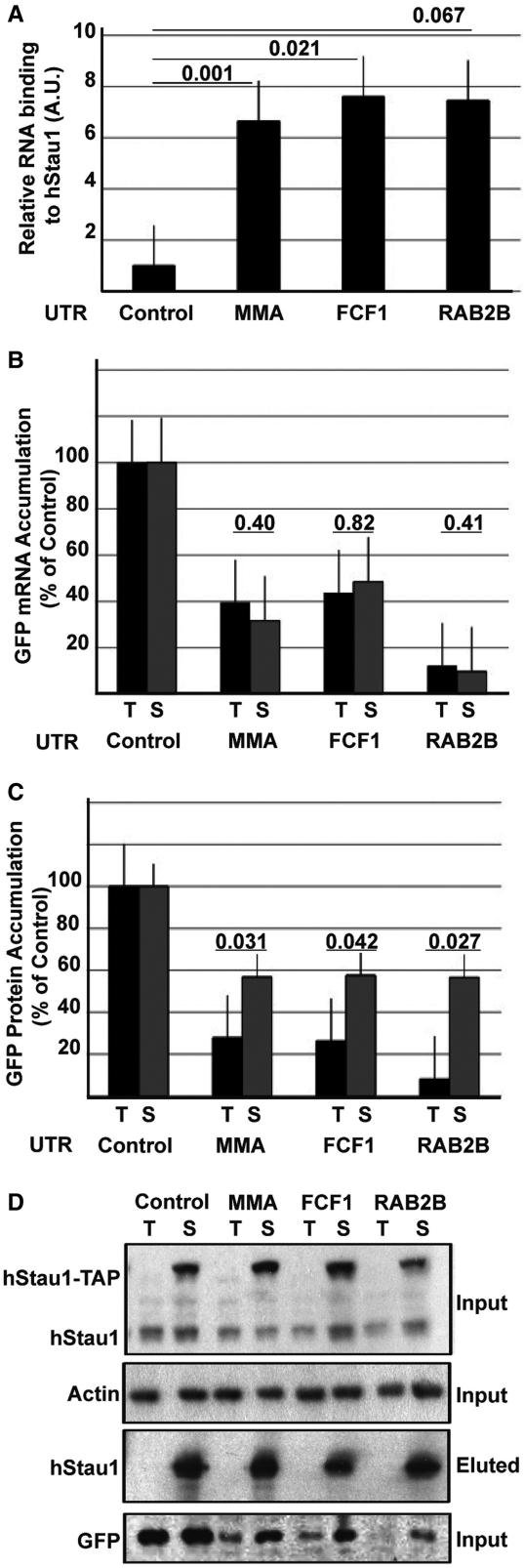

Figure 7. Association to hStau1 and gene expression of recombinant GFP-UTR mRNAs. The plasmids expressing the GFP ORF linked to MMACHC, FCF1 or RAB2B 3′-UTRs were cotransfected into HEK293T cells with either hStau1-TAP or TAP expressing plasmids. As control, an irrelevant UTR present in pEGFP-C1 plasmid (248 nt) was used. (**A**) The intracellular complexes were purified by affinity chromatography and RNA was isolated from either total cell extracts or purified complexes. The accumulation of 28S rRNA and GFP mRNA was determined by RT-qPCR. The amount of 28S rRNA was used to standardize the values of GFP mRNA in each sample and the ratio of GFP mRNA present in the purified complex to that present in the original cell extract was determined. The results are presented as relative values using those of GFP linked to a control UTR as reference and represent averages and standard deviations of 4–10 experiments (A.U., arbitrary units). (**B**) The accumulation of 28S rRNA and GFP mRNA was determined by RT-qPCR in total extracts of cells cotransfected with either a TAP (T) or a hStau1-TAP expressing plasmid (S). The amounts of GFP mRNA were standardized by the 28S RNA content, are referred to the values of control samples and represent averages and standard deviations of 4–8 experiments. (**C**) GFP protein accumulation was determined by western blot of total extracts of cells cotransfected with either a TAP (T) or an hStau1-TAP expressing plasmid (S). Western-blot signals were quantified with ImageJ. The amounts of GFP were standardized by the actin content of the extracts, are referred to the values of control samples and represent averages and standard deviations of 5–8 experiments. (**D**) Representative experiment showing the expression and purification of hStau1 complexes, as well as GFP expression. Endogenous and recombinant hStau1 protein expression is shown (first panel), using actin as a loading marker (second panel). The purified hStau1 complex is shown (third panel) as well as the accumulation of GFP (forth panel). The numbers in (A, B and C) denote the statistical significance as determined by the Student t-test.

high proportion of cellular mRNAs (50,51). The Staufen protein is a key player in the localization and regulated translation of specific mRNAs in several organisms (8,13,16,49,52,53), but the sequence/structure signatures it uses to identify the target mRNAs are far from clear. Only recently the structures of the recognition signatures have been elucidated for *D. melanogaster* K10 and *oskar* mRNAs (54,55) and for a subset of human mRNAs that undergo SMD (16). In the present report we took an unbiased approach to uncover hStau1 target mRNAs based on the use of a mutant protein essentially unable to bind RNA. A combination of affinity purification, RNAse protection and deep-sequencing (Figure 1) allowed the identification of a long list of mRNAs specifically associated to hStau1 and the recognition of a minimal binding signature defined by two inverted Alu-like sequences located within the mRNA 3′-UTR (Figures 2–6). This is in contrast with the situation for mRNAs target of SMD, for which a single Alu-like sequence has been found in the 3′-UTR and the implication of Alu-containing lncRNAs has been reported (16). The experimental evidence presented here indicates that the specific mRNAs identified are associated to hStau1 (Figure 4) and that their 3′-UTRs are sufficient for association to hStau1 and for hStau1-dependent stimulation of protein expression *in vivo* (Figure 7). Furthermore, deletion or inversion of the Alu-like sequences protected in the 3′-UTR of MMACHC gene reduced hStau1 association (Figure 8). These results suggest that for this target mRNA both inverted Alu-like sequences are necessary and may act in *cis*, as their orientation is critical. The situation is different for the FCF1 target mRNA, as similar mutations did not affect hStau1 association (Figure 8). However, it should be mentioned that FCF1 3′-UTR contains an additional, non-protected Alu sequence between those protected (Figure 6B;

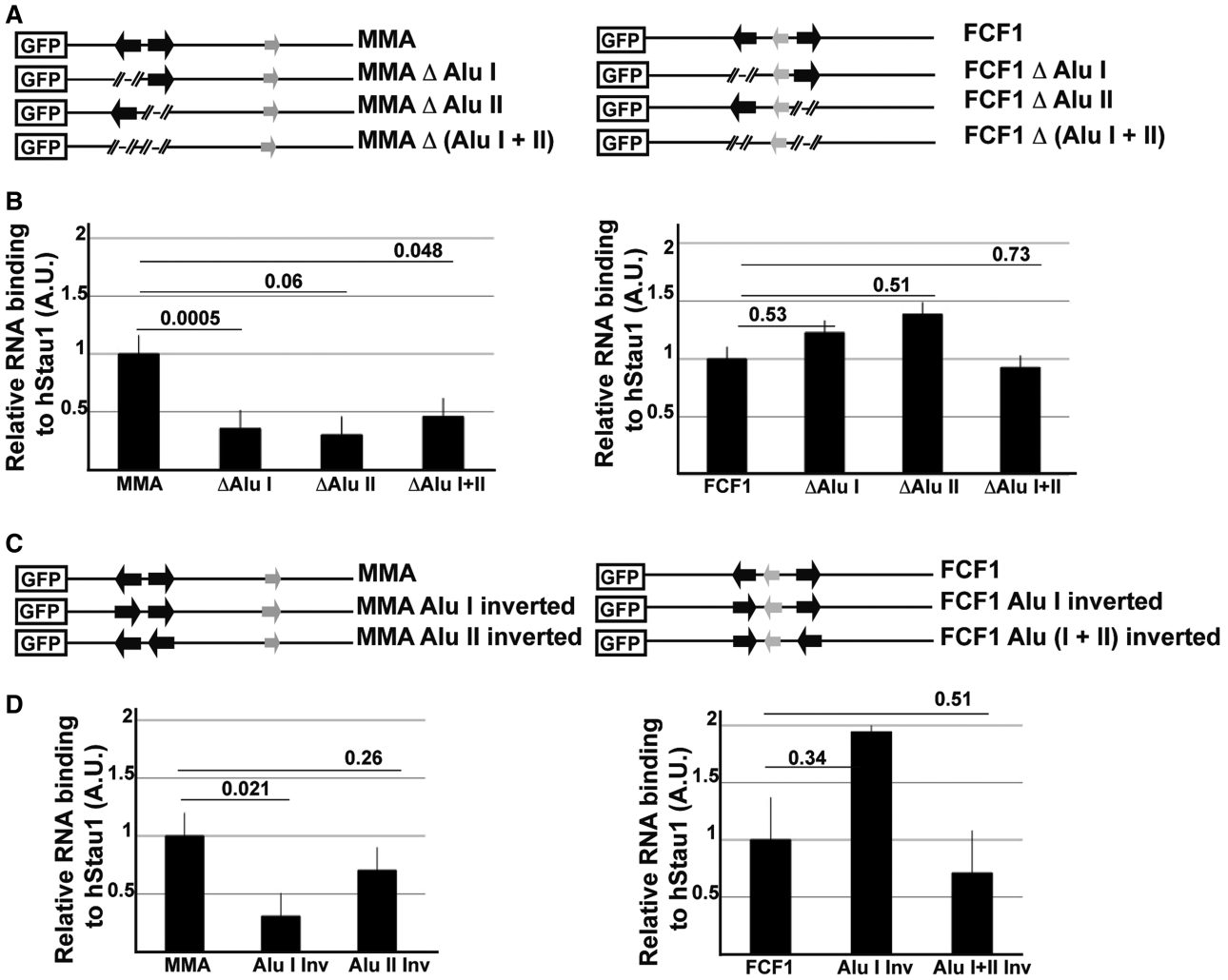

Figure 8. Structure and phenotype of signature mutants. (**A**) Structure of deletion mutants of recombinant GFP-3′-UTRs (see Figure 6B). (**B**) The relative GFP mRNA binding to hStau1 complexes was determined as indicated in Figure 7A and the results are presented using the values for the corresponding wt GFP-3′-UTR construct as a reference. The values represent averages and standard deviations of three experiments. (**C**) Structure of inversion mutants of recombinant GFP-3′-UTRs (see Figure 6B). (**D**) The relative GFP mRNA binding to hStau1 complexes was determined as indicated in Figure 7A and the results are presented using the values for the corresponding wt GFP-3′-UTR construct as a reference. The values represent averages and standard deviations of three experiments. The numbers in (B and D) denote the statistical significance as determined by the Student t-test.

Supplementary Figure S6). Such organization could hamper the formation in *cis* of the hStau1 signature and may require the participation of *trans*-acting lncRNAs containing Alu-like sequences, as reported for hStau1-dependent decay (16). A parallel screening of the deep-sequencing data using the lincRNA databases indicated the specific association of some of these RNAs to hStau1 complexes and a small fraction of these showed the presence of Alu-like sequences (Supplementary Table S2 and Supplementary Figure S3). Alternatively, formation of the sequence signature in FCF1 mRNA may involve intermolecular signals generated by *trans*-interaction of various mRNAs, as reported for *bicoid* mRNA in the fly (56).

The Alu sequences are members of the short interspersed elements (SINEs) that appeared in the primate ancestor some 65 Myears ago. They are typically ~300 nt long and arose as duplication of the 7SL signal recognition particle RNA. The human genome contains ~1 million Alu copies, mainly located at non-coding regions, which can be grouped in >200 subfamilies (57). They can be transcribed from an internal RNA pol III promoter or from nearby RNA pol II promoters and have important implications in genome evolution by gene rearrangements, mutations due to retrotransposition and induced recombination. In addition, important effects of Alu sequences have been reported in the modulation of gene expression, like alterations in transcription efficiency, alternative splicing, premature polyadenylation or mRNA editing (58). Here we show that a specific arrangement of Alu-like sequences can act as hStau1 binding site (Figure 7). It has been shown that a signature of two inverted Alu sequences reduces mRNA translation (59). Here we further report that such a signature determines

hStau1-dependent enhancement in gene expression, independent of mRNA stability (Figure 7). Our results confirm and extend the recent publication by Elbarbary *et al*. (60) and suggest that recognition of the Alu signature by hStau1 protein allows efficient mRNA nucleocytoplasmic transport and incorporation to the associated translation machinery.

Given that Staufen is a protein that has been partly conserved from the fly and *Xenopus* to humans, it might be considered unexpected to find that a combination of Alu sequences is used by hStau1 to identify target mRNAs, as these were only acquired recently in genome evolution. However, it should be noted that the specific mRNA binding sites of *D. melanogaster bicoid* and *oskar* mRNAs, *bona fide* targets for recognition by Staufen, contain stem-loop structures structurally similar to those expected from proximal inverted Alu sequences. In addition, a recent analysis of the mRNAs associated to Stau in *D. melanogaster* embryos, as detected by co-immunoprecipitation with recombinant or endogenous Stau protein, revealed structure motifs like mismatches, internal loops and bulges within RNA stems that could confer specificity to Stau binding (61). Such secondary structure motifs can also be predicted in the folded Alu-signatures described for hStau1 mRNA targets (data not shown). On the other hand, inspection of the 3′-UTRs of the mouse FCF1, MMACHC and RAB2B mRNAs indicates that the latter two genes also contain mouse SINE elements. Hence, our results suggest that mammalian Stau1 proteins have taken advantage of the presence of abundant SINE elements to secure recognition of specific mRNA targets.

ACCESSION NUMBERS

Deep-sequencing data have been deposited with NCBI-GEO database with accession code: GSE47226.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

We thank Yolanda Fernández and Noelia Zamarreño for technical assistance.

FUNDING

Spanish Ministry of Science and Innovation (*Ministerio de Ciencia e Innovación*) [BFU2010-17540/BMC] and Fundación Marcelino Botín. S. de L. was a postdoctoral fellow from FISS, Spanish Ministry of Science and Innovation (Ministerio de Ciencia e Innovación). Funding for open access charge: Spanish Ministry of Science and Innovation.

Conflict of interest statement. None declared.

REFERENCES

1. Dreyfuss,G., Kim,V.N. and Kataoka,N. (2002) Messenger-RNA-binding proteins and the messages they carry. *Nat. Rev. Mol. Cell Biol.*, **3**, 195–205.
2. Huang,Y. and Steitz,J.A. (2005) SRprises along a messenger's journey. *Mol. Cell*, **17**, 613–615.
3. Stutz,F. and Izaurralde,E. (2003) The interplay of nuclear mRNP assembly, mRNA surveillance and export. *Trends Cell Biol.*, **13**, 319–327.
4. Saunders,L.R. and Barber,G.N. (2003) The dsRNA binding protein family: critical roles, diverse cellular functions. *FASEB J.*, **17**, 961–983.
5. Ramos,A., Grunert,S., Adams,J., Micklem,D.R., Proctor,M.R., Freund,S., Bycroft,M., St,J.D. and Varani,G. (2000) RNA recognition by a Staufen double-stranded RNA-binding domain. *EMBO J*, **19**, 997–1009.
6. Ryter,J.M. and Schultz,S.C. (1998) Molecular basis of double-stranded RNA-protein interactions: structure of a dsRNA-binding domain complexed with dsRNA. *EMBO J.*, **17**, 7505–7513.
7. Riechmann,V. and Ephrussi,A. (2001) Axis formation during Drosophila oogenesis. *Curr. Opin. Genet. Dev.*, **11**, 374–383.
8. St Johnston,D., Beuchle,D. and Nusslein-Volhard,C. (1991) Staufen, a gene required to localize maternal RNAs in the Drosophila egg. *Cell*, **66**, 51–63.
9. Marión,R.M., Fortes,P., Beloso,A., Dotti,C. and Ortín,J. (1999) A human sequence homologue of staufen is an RNA-binding protein that localizes to the polysomes of the rough endoplasmic reticulum. *Mol. Cell. Biol.*, **19**, 2212–2219.
10. Wickham,L., Duchaine,T., Luo,M., Nabi,I.R. and DesGroseillers,L. (1999) Mammalian staufen is a double-stranded-RNA- and tubulin-binding protein which localizes to the rough endoplasmic reticulum. *Mol. Cell. Biol.*, **19**, 2220–2230.
11. Brendel,C., Rehbein,M., Kreienkamp,H.J., Buck,F., Richter,D. and Kindler,S. (2004) Characterization of Staufen 1 ribonucleoprotein complexes. *Biochem. J.*, **384**, 239–246.
12. Kanai,Y., Dohmae,N. and Hirokawa,N. (2004) Kinesin transports RNA: isolation and characterization of an RNA-transporting granule. *Neuron*, **43**, 513–525.
13. Villacé,P., Marión,R.M. and Ortín,J. (2004) The composition of Staufen-containing RNA granules from human cells indicate a role in the regulated transport and translation of messenger RNAs. *Nucleic Acids Res.*, **32**, 2411–2420.
14. Dugre-Brisson,S., Elvira,G., Boulay,K., Chatel-Chaix,L., Mouland,A.J. and DesGroseillers,L. (2005) Interaction of Staufen1 with the 5' end of mRNA facilitates translation of these RNAs. *Nucleic Acids Res.*, **33**, 4797–4812.
15. Kim,Y.K., Furic,L., Desgroseillers,L. and Maquat,L.E. (2005) Mammalian Staufen1 recruits Upf1 to specific mRNA 3′UTRs so as to elicit mRNA decay. *Cell*, **120**, 195–208.
16. Gong,C. and Maquat,L.E. (2011) lncRNAs transactivate STAU1-mediated mRNA decay by duplexing with 3′ UTRs via Alu elements. *Nature*, **470**, 284–288.
17. Chatel-Chaix,L., Clement,J.F., Martel,C., Beriault,V., Gatignol,A., DesGroseillers,L. and Mouland,A.J. (2004) Identification of Staufen in the human immunodeficiency virus type 1 Gag ribonucleoprotein complex and a role in generating infectious viral particles. *Mol. Cell. Biol.*, **24**, 2637–2648.
18. de Lucas,S., Peredo,J., Marion,R.M., Sanchez,C. and Ortin,J. (2010) Human Staufen1 protein interacts with influenza virus ribonucleoproteins and is required for efficient virus multiplication. *J. Virol.*, **84**, 7603–7612.
19. Mouland,A.J., Mercier,J., Luo,M., Bernier,L., DesGroseillers,L. and Cohen,E.A. (2000) The double-stranded RNA-binding protein Staufen is incorporated in human immunodeficiency virus type 1: evidence for a role in genomic RNA encapsidation. *J. Virol.*, **74**, 5441–5451.
20. Falcón,A.M., Fortes,P., Marión,R.M., Beloso,A. and Ortín,J. (1999) Interaction of influenza virus NS1 protein and the human homologue of Staufen in vivo and in vitro. *Nucleic Acids Res.*, **27**, 2241–2247.
21. Kiebler,M.A. and Bassell,G.J. (2006) Neuronal RNA granules: movers and makers. *Neuron*, **51**, 685–690.

22. Anderson,P. and Kedersha,N. (2006) RNA granules. *J. Cell. Biol.*, **172**, 803–808.
23. Thomas,M.G., Loschi,M., Desbats,M.A. and Boccaccio,G.L. (2011) RNA granules: the good, the bad and the ugly. *Cell. Signal.*, **23**, 324–334.
24. Krichevsky,A.M. and Kosik,K.S. (2001) Neuronal RNA granules: a link between RNA localization and stimulation-dependent translation. *Neuron*, **32**, 683–696.
25. Kedersha,N., Stoecklin,G., Ayodele,M., Yacono,P., Lykke-Andersen,J., Fritzler,M.J., Scheuner,D., Kaufman,R.J., Golan,D.E. and Anderson,P. (2005) Stress granules and processing bodies are dynamically linked sites of mRNP remodeling. *J. Cell Biol.*, **169**, 871–884.
26. Thomas,M.G., Martinez Tosar,L.J., Desbats,M.A., Leishman,C.C. and Boccaccio,G.L. (2009) Mammalian Staufen 1 is recruited to stress granules and impairs their assembly. *J. Cell Sci.*, **122**, 563–573.
27. Furic,L., Maher-Laporte,M. and DesGroseillers,L. (2008) A genome-wide approach identifies distinct but overlapping subsets of cellular mRNAs associated with Staufen1- and Staufen2-containing ribonucleoprotein complexes. *RNA*, **14**, 324–335.
28. DuBridge,R.B., Tang,P., Hsia,H.C., Leong,P.M., Miller,J.H. and Calos,M.P. (1987) Analysis of mutation in human cells by using an Epstein-Barr virus shuttle system. *Mol. Cell. Biol.*, **7**, 379–387.
29. Wigler,M., Pellicer,A., Silverstein,S., Axel,R., Urlaub,G. and Chasin,L. (1979) DNA-mediated transfer of the adenine phosphoribosyltransferase locus into mammalian cells. *Proc. Natl Acad. Sci. USA*, **76**, 1373–1376.
30. Li,H. and Durbin,R. (2009) Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics*, **25**, 1754–1760.
31. Robinson,M.D., McCarthy,D.J. and Smyth,G.K. (2010) edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics*, **26**, 139–140.
32. Benjamini,Y. and Hochberg,Y. (1995) Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. R. Stat. Soc.*, **57**, 289–300.
33. Robinson,J.T., Thorvaldsdottir,H., Winckler,W., Guttman,M., Lander,E.S., Getz,G. and Mesirov,J.P. (2011) Integrative genomics viewer. *Nat. Biotechnol.*, **29**, 24–26.
34. Wang,L., Wang,S. and Li,W. (2012) RSeQC: quality control of RNA-seq experiments. *Bioinformatics*, **28**, 2184–2185.
35. Huang,D.W., Sherman,B.T. and Lempicki,R.A. (2009) Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat. Protocols*, **4**, 44–57.
36. Edgar,R.C. (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.*, **32**, 1792–1797.
37. Nawrocki,E.P., Kolbe,D.L. and Eddy,S.R. (2009) Infernal 1.0: inference of RNA alignments. *Bioinformatics*, **25**, 1335–1337.
38. Ren,J., Wen,L., Gao,X., Jin,C., Xue,Y. and Yao,X. (2009) DOG 1.0: illustrator of protein domain structures. *Cell Res.*, **19**, 271–273.
39. Waterhouse,A.M., Procter,J.B., Martin,D.M., Clamp,M. and Barton,G.J. (2009) Jalview Version 2—a multiple sequence alignment editor and analysis workbench. *Bioinformatics*, **25**, 1189–1191.
40. Ashburner,M., Ball,C.A., Blake,J.A., Botstein,D., Butler,H., Cherry,J.M., Davis,A.P., Dolinski,K., Dwight,S.S., Eppig,J.T. *et al.* (2000) Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nat. Genet.*, **25**, 25–29.
41. Thorvaldsdottir,H., Robinson,J.T. and Mesirov,J.P. (2012) Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. *Brief. Bioinform.*, **14**, 178–192.
42. Pratt,C.A. and Mowry,K.L. (2013) Taking a cellular road-trip: mRNA transport and anchoring. *Curr. Opin. Cell Biol.*, **25**, 99–106.
43. Xing,L. and Bassell,G.J. (2013) mRNA localization: an orchestration of assembly, traffic and synthesis. *Traffic*, **14**, 2–14.
44. Broadus,J., Fuerstenberg,S. and Doe,C.Q. (1998) Staufen-dependent localization of prospero mRNA contributes to neuroblast daughter-cell fate. *Nature*, **391**, 792–795.
45. Johnstone,O. and Lasko,P. (2001) Translational regulation and RNA localization in Drosophila oocytes and embryos. *Annu. Rev. Genet.*, **35**, 365–406.
46. Kang,H. and Schuman,E.M. (1996) A requirement for local protein synthesis in neurotrophin-induced hippocampal synaptic plasticity. *Science*, **273**, 1402–1406.
47. Lawrence,J.B. and Singer,R.H. (1986) Intracellular localization of messenger RNAs for cytoskeletal proteins. *Cell*, **45**, 407–415.
48. Long,R.M., Singer,R.H., Meng,X., Gonzalez,I., Nasmyth,K. and Jansen,R.P. (1997) Mating type switching in yeast controlled by asymmetric localization of ASH1 mRNA. *Science*, **277**, 383–387.
49. Yoon,Y.J. and Mowry,K.L. (2004) Xenopus Staufen is a component of a ribonucleoprotein complex containing Vg1 RNA and kinesin. *Development*, **131**, 3035–3045.
50. Lecuyer,E., Yoshida,H., Parthasarathy,N., Alm,C., Babak,T., Cerovina,T., Hughes,T.R., Tomancak,P. and Krause,H.M. (2007) Global analysis of mRNA localization reveals a prominent role in organizing cellular architecture and function. *Cell*, **131**, 174–187.
51. Cajigas,I.J., Tushev,G., Will,T.J., Tom Dieck,S., Fuerst,N. and Schuman,E.M. (2012) The local transcriptome in the synaptic neuropil revealed by deep sequencing and high-resolution imaging. *Neuron*, **74**, 453–466.
52. Kiebler,M.A., Hemraj,I., Verkade,P., Kohrmann,M., Fortes,P., Marion,R.M., Ortin,J. and Dotti,C.G. (1999) The mammalian staufen protein localizes to the somatodendritic domain of cultured hippocampal neurons: implications for its involvement in mRNA transport. *J. Neurosci.*, **19**, 288–297.
53. LeGendre,J.B., Campbell,Z.T., Kroll-Conner,P., Anderson,P., Kimble,J. and Wickens,M. (2013) RNA targets and specificity of Staufen, a double-stranded RNA-binding protein in Caenorhabditis elegans. *J. Biol. Chem.*, **288**, 2532–2545.
54. Bullock,S.L., Ringel,I., Ish-Horowicz,D. and Lukavsky,P.J. (2010) A'-form RNA helices are required for cytoplasmic mRNA transport in Drosophila. *Nat. Struct. Mol. Biol.*, **17**, 703–709.
55. Ghosh,S., Marchand,V., Gaspar,I. and Ephrussi,A. (2012) Control of RNP motility and localization by a splicing-dependent structure in oskar mRNA. *Nat. Struct. Mol. Biol.*, **19**, 441–449.
56. Ferrandon,D., Koch,I., Westhof,E. and Nusslein,V.C. (1997) RNA-RNA interaction is required for the formation of specific bicoid mRNA 3' UTR-STAUFEN ribonucleoprotein particles. *EMBO J.*, **16**, 1751–1758.
57. Cordaux,R. and Batzer,M.A. (2009) The impact of retrotransposons on human genome evolution. *Nat. Rev. Genet.*, **10**, 691–703.
58. Walters,R.D., Kugel,J.F. and Goodrich,J.A. (2009) InvAluable junk: the cellular impact and function of Alu and B2 RNAs. *IUBMB Life*, **61**, 831–837.
59. Capshew,C.R., Dusenbury,K.L. and Hundley,H.A. (2012) Inverted Alu dsRNA structures do not affect localization but can alter translation efficiency of human mRNAs independent of RNA editing. *Nucleic Acids Res.*, **40**, 8637–8645.
60. Elbarbary,R.A., Li,W., Tian,B. and Maquat,L.E. (2013) STAU1 binding 3' UTR IRAlus complements nuclear retention to protect cells from PKR-mediated translational shutdown. *Genes Dev.*, **27**, 1495–1510.
61. Laver,J.D., Li,X., Ancevicius,K., Westwood,J.T., Smibert,C.A., Morris,Q.D. and Lipshitz,H.D. (2013) Genome-wide analysis of Staufen-associated mRNAs identifies secondary structures that confer target specificity. *Nucleic Acids Res.*, **41**, 9438–9460.