# DIGITAL ACCESS TO SCHOLARSHIP AT HARVARD

# Evidence that a consensus element found in naturally intronless mRNAs promotes mRNA export

| Citation | Lei, Haixin, Bo Zhai, Shanye Yin, Steve Gygi, and Robin Reed. 2012. Evidence that a consensus element found in naturally intronless mrnas promotes mrna export. Nucleic Acids Research 41(4): 2517-2525. |
|---|---|
| Published Version | doi:10.1093/nar/gks1314 |
| Accessed | February 19, 2015 11:59:05 AM EST |
| Citable Link | http://nrs.harvard.edu/urn-3:HUL.InstRepos:10589783 |
| Terms of Use | This article was downloaded from Harvard University's DASH repository, and is made available under the terms and conditions applicable to Other Posted Material, as set forth at http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#LAA |

*(Article begins on next page)*

# Evidence that a consensus element found in naturally intronless mRNAs promotes mRNA export

## Haixin Lei, Bo Zhai, Shanye Yin, Steve Gygi and Robin Reed*

Department of Cell Biology, Harvard Medical School, 240 Longwood Avenue, Boston, MA 02115, USA

## ABSTRACT

**We previously showed that mRNAs synthesized from three genes that naturally lack introns contain a portion of their coding sequence, known as a cytoplasmic accumulation region (CAR), which is essential for stable accumulation of the intronless mRNAs in the cytoplasm. The CAR in each mRNA is unexpectedly large, ranging in size from ∼160 to 285 nt. Here, we identified one or more copies of a 10-nt consensus sequence in each CAR. To determine whether this element (designated CAR-E) functions in cytoplasmic accumulation of intronless mRNA, we multimerized the most conserved CAR-E and inserted it upstream of β-globin cDNA, which is normally retained/degraded in the nucleus. Significantly, the tandem CAR-E, but not its anti-sense counterpart, rescued cytoplasmic accumulation of β-globin cDNA transcripts. Moreover, dinucleotide mutations in the CAR-E abolished this rescue. We show that the CAR-E, but not the mutant CAR-E, associates with components of the TREX mRNA export machinery, the Prp19 complex and U2AF2. Moreover, knockdown of these factors results in nuclear retention of the intronless mRNAs. Together, these data suggest that the CAR-E promotes export of intronless mRNA by sequence-dependent recruitment of the mRNA export machinery.**

## INTRODUCTION

Different steps in gene expression such as transcription, RNA processing, mRNA export and mRNA surveillance are carried out by specific machineries, which are extensively coupled to one another both functionally and physically (1–4). For example, the conserved TREX mRNA export complex associates with the spliceosome and is recruited to mRNA during splicing (5,6). Consistent with this splicing-dependent recruitment of TREX, spliced mRNA typically accumulates in the cytoplasm more efficiently than its cDNA counterpart (6,7). β-globin provides a particularly striking example of this coupling, as the cDNA is largely retained/degraded in the nucleus, whereas the spliced mRNA is rapidly and efficiently exported (8–11). The TREX complex, conserved from yeast to human, contains the multi-subunit THO complex and the proteins UAP56, Aly and CIP29 (6,7, 12,13). The human TREX complex is recruited to the 5′-end of mRNAs during splicing via an interaction between Aly and the cap-binding complex (14). In yeast, the TREX complex is recruited to mRNA during transcription and 3′-end formation (15–17). Recently, studies in both yeast and human revealed that the Prp19 complex is also involved in a network of coupled interactions involving transcription, splicing and mRNA export (18–20). The Prp19 complex contains a number of proteins conserved from yeast to human, including Prp19, Cdc5l, Xab2 and Crnkl1 (21). The splicing factor U2AF2 was recently found to associate with the Prp19 complex and function in coupling transcription to splicing (19).

In contrast to splicing-dependent mRNA export, little is known about how naturally intronless mRNAs in higher eukaryotes are exported in the absence of splicing. Most of the work has been done with naturally intronless viral mRNAs, revealing that these mRNAs bind to specific cellular or viral proteins that recruit the mRNA export machinery (22–26). In the case of naturally intronless cellular mRNAs, histone H2A contains an element that binds to SR proteins and recruits the mRNA export receptor Nxf1 (27–31). However, H2A mRNA is unique among cellular mRNAs, lacking a polyA tail. In *Drosophila*, the homolog of human U2AF2 associates with naturally intronless mRNAs and functions in their export (32).

To further understand the mechanism for exporting naturally intronless mRNAs, we previously examined three of these mRNAs, HSPB3, IFNα1 and IFNβ1 (10). We found that the TREX complex associates with the intronless mRNA and their export is blocked by

*To whom correspondence should be addressed. Tel: +1 617 834 2396; Fax +1 617 432 3091; Email: rreed@hms.harvard.edu
Present address:
Haixin Lei, Institute of Cancer Stem Cell, Cancer Center, Dalian Medical University, Dalian, Liaoning Province, People's Republic of China.

knockdown of TREX components (UAP56 and Thoc2) or Nxf1 (10). We also showed that a portion of the coding region, which we named cytoplasmic accumulation regions (CARs), is essential for stability/export of these mRNAs. Insertion of the CARs upstream of β-globin cDNA promotes the cytoplasmic accumulation of the cDNA transcript, whereas the antisense of the CARs does not (10). These data are consistent with a previous study showing that a portion of the coding region of c-Jun mRNA, which is also naturally intronless, has the same function as the CARs that we identified (33). For the purpose of this study, we refer to the c-Jun element as a CAR. In all of the intronless mRNAs, the size of the CAR is unexpectedly large, ranging from 162 to 402 nt. Further work is required to understand how CARs function in promoting cytoplasmic accumulation of naturally intronless mRNAs.

Here, we report the identification of a 10-nt consensus element C[CA]AG[ATC][TA][CG][CG]TG, designated CAR-E (<u>CAR</u> Element) found in the CARs in HSPB3, c-Jun, IFNα1 and IFNβ1 mRNAs. We show that tandem duplication of the CAR-E, but not the antisense or a mutated counterpart, promotes cytoplasmic accumulation of β-globin cDNA. Our data indicate that components of the TREX complex, the Prp19 complex and U2AF2 associate with the tandem CAR-E and function in mRNA export. Together, our data suggest that naturally intronless mRNAs bypass splicing-dependent export via specific elements that recruit the mRNA export machinery.

## MATERIALS AND METHODS

### Identification of the CAR-E

To identify sequences that function in export of naturally intronless mRNAs, we used the MEME algorithm (34) (http://meme.nbcr.net) to search for overrepresented motifs in the four CARs. Using several different approaches, we identified similar or overlapping motifs that were specifically enriched in the CARs but not in their antisense or in the sequences surrounding the CARs. For example, in one approach, we looked for 13-nt motifs, searching for any number of elements in the four CARs. In another approach, we looked for 10-nt motifs looking for any number of motifs using a combination of sense and antisense of the four CARs. The motifs generated via MEME were then further analyzed using FIMO (35) to search for the motifs in the antisense of the CARs or in size-matched sequences surrounding each CAR. Using a *P*-value output threshold set at $1 \times 10^{-4}$, the motif was not present in these negative control sequences. For further analysis, we used the most conserved motif (CCAGTTCCTG), which had the lowest *P*-value ($2.37 \times 10^{-6}$) in our MEME search.

### Constructs and antibodies

Wild-type (WT) β-globin containing its natural introns and β-globin cDNA plasmids were described (10). The tandem CAR-E and CAR-E$_{AS}$ β-globin cDNA plasmids were constructed by inserting a DNA sequence (5′ CCAG TTCCTG × 16 3′) into the Hind III site upstream of the cDNA. To generate the CAR-E mutant 1–5 constructs, oligonucleotides containing each dinucleotide mutation were synthesized and inserted into the HindIII site upstream of the β-globin cDNA construct. To generate the Slu7 and DDX3x plasmids, the cDNA and HA-tag were first amplified from Slu7 or DDX3 cDNA plasmids (36,37) by PCR with the following primers: Slu7-F-5′-TTT GGTACCATGTACCCATACGACGTCCCAGACTAC GCTTCAGCCACAGTTGTAGATGCA; Slu7-R-5′-AA ACTCGAGCTACTGTCCAAGGAAAGAGGCCAT; DDX3x-F-5′-TTTGGTACCATGTACCCATACGACG TCCCAGACTACGCTAGTCATGTGGCAGTGGAAA AT; DDX3x-R-5′-ATACTCGAGTCAGTTACCCCACC AGTCAACCCCCTG. The PCR products were then inserted into the pcDNA5/FRT/TO vector (Invitrogen) at the KpnI and XhoI sites to generate the cDNA versions of the constructs. The same PCR products were also inserted into the KpnI and XhoI sites of pcDNA5/ FRT/TO vector containing the tandem CAR-E. Constructs were verified by DNA sequencing. Polyclonal antibodies against Xab2 (Proteintech, 1:800), U2AF2 (Santa Cruz, 1:800), AQR (Bethyl, 1:2000) and monoclonal antibodies against the HA tag (Covance, 1:2000), EGFP (Origene, 1:1000) and tubulin (Sigma, 1:10 000) were used for western blots.

### Cell culture, transfection and HeLa nuclear microinjection

HeLa cells were cultured in Dulbecco's modified eagle medium supplemented with 10% fetal bovine serum. Transfection was carried out in MatTek plates using 1 μg of each plasmid and Lipofectamine 2000. For siRNA transfection, 1.25 μl of 40 μM siRNAs (Smartpool from Dharmacon) was added after mixing with 1 μl lipofectamine 2000. To determine the knockdown efficiency for CRNKL1 and ISY1, the following primers were used for RT–PCR: CRNKL1-RT-F-5′-TGAGGAC GTCGATGAGAGTG, CRNKL1-RT-R-5′-GCAGAGC TGGGAAATGAACT and ISY1-RT-F-5′-ACTGGTGC GAAGGAAGAAAA, ISY1-RT-R-5′-TCAAAATGGC AGTGCAAGTC. For microinjection of CMV constructs, a 10-μl mixture containing 2 μl 70-kDa Dextran (Molecular Probes) and 50 ng/μl of each plasmid DNA was microinjected into HeLa cell nuclei. For fluorescence *in situ* hybridization (FISH), samples were rinsed once with 1 × phosphate buffered saline (PBS) and fixed with 4% paraformaldehyde in PBS for 15 min, and then permeabilized in PBS containing 0.1% Triton X-100. After three rinses with 1 × PBS, the samples were rinsed twice with 1 × SSC/50% formamide, fluorescent probe was added and hybridization was performed overnight at 37°C. The FISH probe was a 70-nt DNA oligonucleotide complementary to the vector sequence downstream of the multiple cloning site and pre-labeled at the 5′-end with Alexa Fluor 546 NHS ester and HPLC-purified. Probe sequence: 5′-AAGGCACGGGGGAGGGGCAAACAA CAGATGGCTGGCAACTAGAAGGCACAGTCGAG GCTGATCAGCGGGT. Images were taken with a Nikon TE2000E Inverted Fluorescence Microscope.

## Purification of RNPs

Templates for *in vitro* transcription of the tandem CAR-E or mutant tandem CAR-E$_{m2}$ were amplified from the β-globin CAR-E and CAR-E$_{m2}$ cDNA plasmids using the following primers: F-5′-TGGAGGTCGCTGAGTA GTGC and R-5′-TTCCATGGTGGCGGCGGTACCAA . For *in vitro* transcription, 1 μg of the PCR product was transcribed with 50 units of T7 RNA polymerase (New England Biolabs) for 90 min in reaction mixtures containing $^{32}$P-UTP and biotin-16-UTP (Roche). For RNP assembly, 2 μg of CAR-E or CAR-E$_{m2}$ RNA was incubated for 2 h at 30°C in a 2-ml reaction mixture containing 600 μl of HeLa nuclear extract, 600 μl splicing dilution buffer (20 mM HEPES at pH 7.9, 100 mM KCl), 3.2 mM MgCl$_2$, 20 mM creatine phosphate di-Tris salt and 0.5 mM ATP (13). The mixtures were then loaded on Sephacryl S-500 columns and the peak fractions containing the RNP were pooled and purified by binding to streptavidin agarose (Thermo Scientific). Total proteins were eluted from equivalent amounts of $^{32}$P-labeled RNA from each mRNP and TCA precipitated. For mass spectrometry, total proteins were subjected to in-solution trypsin digestion overnight. Tryptic peptides were analyzed by nanoscale-microcapillary reversed phase liquid chromatography–tandem mass spectrometry (LC–MS/MS) essentially as described (38). MS/MS spectra were searched using the SEQUEST algorithm (39) against a concatenated forward/reverse mouse IPI (ver 3.60) database as described (40) with dynamic modification of methionine oxidation. All peptide matches were filtered based on mass deviation, charge, XCorr and dCn to a target peptide false discovery rate of 1% using Linear Discriminant Analysis to distinguish between forward and reverse hits as described previously (41). For Table 1 and Supplementary Table S1, the proteins were categorized based on their best-known function. We found similar numbers of peptides for CBP20 and CBP80 in the CAR-E versus CAR-E$_{m2}$ mRNP and used these proteins for peptide number comparisons with other proteins listed in Table 1. Contaminants such as keratin, tubulin, ribosomal and translation-related proteins were omitted. In addition, proteins >300 amino acids with total peptide numbers ≤3 were omitted.

## RESULTS

### A consensus element in CARs promotes cytoplasmic accumulation of RNA

To identify sequences that function in cytoplasmic accumulation of naturally intronless mRNAs, we used a variety of criteria to search for overrepresented motifs within the four CARs using the MEME algorithm (34) (see 'Materials and Methods' section). Our analysis led to the identification of a 10-nt consensus element, designated the CAR-E (CAR-Element) that is present in one to four copies in each CAR (Figure 1A). The positions of the CAR-E in each CAR are shown in Figure 1B. The sequence of each CAR-E (colored nucleotide) and non-conserved flanking regions are shown in Figure 1C.

**Table 1.** Proteins present in tandem CAR-E and CAR-E$_{m2}$ RNPs

| | CAR-E | | CAR-E$_{m2}$ | | Mol weight (kDa) |
|---|---|---|---|---|---|
| | Unique | Total | Unique | Total | |
| TREX complex | | | | | |
| THOC2 | 29 | 37 | 1 | 1 | 183 |
| THOC5 | 13 | 16 | | | 79 |
| THOC6 | 7 | 9 | | | 38 |
| THOC3 | 5 | 5 | | | 39 |
| THOC1 | 5 | 5 | | | 76 |
| UAP56/URH49 | 13 | 15 | | | 49 |
| Prp19 complex and U2AF | | | | | |
| Xab2 | 9 | 12 | 1 | 1 | 100 |
| AQR | 10 | 13 | 2 | 2 | 171 |
| CDC5L | 4 | 4 | 1 | 1 | 92 |
| Prp19 | 6 | 6 | | | 55 |
| U2AF2 | 7 | 11 | | | 54 |
| Cap-binding complex | | | | | |
| NCBP1 | 18 | 28 | 7 | 24 | 92 |
| NCBP2 | 3 | 3 | 3 | 6 | 18 |
| Examples of proteins enriched in CAR-E$_{m2}$ RNP | | | | | |
| PTBP1 | | | 10 | 40 | 57 |
| ADAR | 9 | 11 | 17 | 59 | 136 |
| DICER1 | 1 | 1 | 12 | 22 | 219 |
| KHSRP | | | 8 | 14 | 73 |
| NOLC1 | 1 | 1 | 8 | 13 | 74 |

A subset of proteins detected by mass spectrometry in the tandem CAR-E and CAR-E$_{m2}$ RNPs is shown. The number of total and unique peptides and molecular weight of the proteins are indicated.

We previously showed that the antisense of the CARs are non-functional in export (10). Significantly, no CAR-Es are present in the antisense of the CARs at the same *P*-value output threshold ($1 \times 10^{-4}$) as used to identify the CAR-Es in the sense strand of the CARs. Moreover, no CAR-Es were detected in size-matched regions outside of the CARs in each naturally intronless mRNA at this threshold. Based on these observations, we pursued the CAR-E further.

To examine the role of the CAR-E in naturally intronless mRNA export, we mutated either one or more of the CAR-Es in HSPB3 mRNA. Although we made complete substitutions of all of these elements, the mutations had no apparent effect on cytoplasmic accumulation of HSPB3 mRNA (data not shown). In addition, we made a construct in which all three of the CAR-Es were mutated and five additional sequences that resembled the CAR-Es were mutated, but no effect of these mutations was observed (see 'Discussion' section). Thus, as an alternative approach to determine whether the CAR-E plays a role in the cytoplasmic accumulation of intronless mRNAs, we inserted 16 tandem copies of the most conserved CAR-E (CCAGTTCCTG from c-Jun) or its antisense counterpart (CAR-E$_{AS}$) upstream of β-globin cDNA (Figure 2A). These constructs were transfected into HeLa cells and FISH was carried out to determine the nucleocytoplasmic distribution of β-globin RNA. Constructs encoding β-globin cDNA alone or WT β-globin containing its two natural introns were analyzed as controls. Consistent with previous work (8–11), WT β-globin mRNA, but not the cDNA transcript, efficiently accumulated in the cytoplasm (Figure 2B). Significantly, insertion of the tandem CAR-E
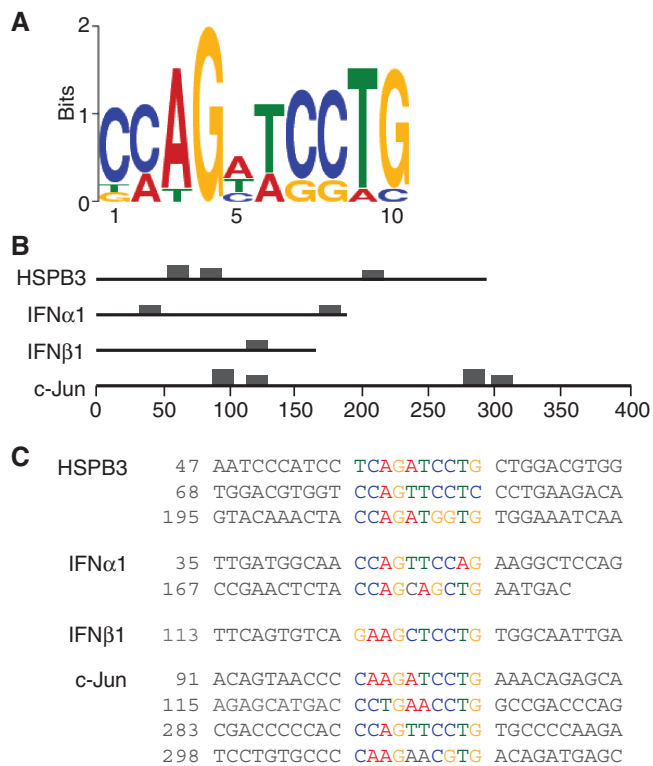
**A**



**B**



**C**

```
HSPB3    47  AATCCCATCC TCAGATCCTG CTGGACGTGG
         68  TGGACGTGGT CCAGTTCCTC CCTGAAGACA
        195  GTACAAACTA CCAGATGGTG TGGAAATCAA

IFNα1    35  TTGATGGCAA CCAGTTCCAG AAGGCTCCAG
        167  CCGAACTCTA CCAGCAGCTG AATGAC

IFNβ1   113  TTCAGTGTCA GAAGCTCCTG TGGCAATTGA

c-Jun    91  ACAGTAACCC CAAGATCCTG AAACAGAGCA
        115  AGAGCATGAC CCTGAACCTG GCCGCACCCAG
        283  CGACCCCCAC CCAGTTCCTG TGCCCCAAGA
        298  TCCTGTGCCC CAAGAACGTG ACAGATGAGC
```

**Figure 1.** CARs in naturally intronless mRNAs contain a conserved motif. (**A**) The 10-nt consensus element (CAR-E) identified in the CARs from HSPB3, c-Jun, IFNα1 or IFNβ1 is shown. (**B**) The locations of the CAR-Es in each CAR are indicated by the boxes. The height of box is proportional to $-\log(P\text{-value})$, truncated at the height for a motif with a *P*-value of $1 \times 10^{-10}$ (34). The numbering is relative to the first nucleotide of each CAR sequence. (**C**) The sequences of each of the CAR-Es (colored bases) surrounded by flanking sequence (black bases) are shown.

resulted in accumulation of β-globin cDNA transcripts in the cytoplasm (Figure 2B). In contrast, β-globin cDNA transcripts containing the CAR-E$_{AS}$ had a nucleo-cytoplasmic distribution that was similar to β-globin cDNA transcripts alone (Figure 2B). Specifically, as readily seen in the low magnification fields, the CAR-E$_{AS}$ cDNA and cDNA transcripts alone were largely retained/degraded in the nucleus (Figure 2C). We note that the low level of FISH signal observed from the CAR-E$_{AS}$ cDNA and cDNA transcripts is not due to differences in transfection efficiency, as shown by western blot of the transfection control, EGFP (Figure 2D). In addition, our previous work showed that the transcript levels of different CMV constructs are the same at 5 min after microinjection, regardless of the transcribed sequence (9,10).

Consistent with our results showing enhanced export of β-globin cDNA containing the CAR-E, western analysis 1 or 2 days after transfection showed that high levels of protein were expressed from both WT and the CAR-E constructs, whereas little protein was detected from the cDNA or CAR-E$_{AS}$ construct (Figure 2D). These results indicate that the consensus element is functional in promoting cytoplasmic accumulation of mRNA and

production of protein. We note that insertion of 6 or 10 copies of the CAR-E promoted expression of β-globin cDNA, but the levels were significantly higher with 16 copies. Thus, we used the 16-copy CAR-E construct for further studies. To determine whether the effect of the tandem CAR-E element might be general, we inserted it upstream of two other cDNAs, Slu7 and DDX3x (Supplementary Figure S1). In both cases, the CAR-E resulted in increased mRNA export and greater levels of protein than observed with the corresponding cDNA alone (Supplementary Figure S1), suggesting that the function of the CAR-E may be general.

To further analyze the tandem CAR-E, we introduced dinucleotide mutations across each copy of the CAR-E and inserted this sequence upstream of β-globin cDNA (Figure 3A). These constructs were transfected into HeLa cells and western analysis was used to assay the levels of β-globin protein. After 1 day of transfection, no protein was detected from CAR-E mutants m1-m4 and less protein was detected from CAR-E mutant m5 versus either WT β-globin or the CAR-E cDNA (Figure 3B). Similar results were obtained 2 days after transfection (Figure 3B). Consistent with these data, FISH analysis showed that β-globin RNA containing the CAR-E mutants 1–5 either did not accumulate (mutant 1–4) or partially accumulated (mutant 5) in the cytoplasm, whereas β-globin RNA containing the CAR-E was mostly detected in the cytoplasm (Figure 3C). These data indicate that the CAR-E promotes cytoplasmic accumulation of RNA.

**The tandem CAR-E RNA associates with mRNA export factors**

We next sought to identify proteins that associate with the tandem CAR-E, using CAR-E$_{m2}$ as a negative control. To do this, we biotinylated the RNAs and isolated the RNPs by binding to streptavidin agarose (42). Total proteins from equivalent amounts of these RNPs were then analyzed by mass spectrometry. Not unexpectedly, both types of RNPs contained a large number of RNA-related proteins, including those associated with the spliceosome, the polyadenylation machinery as well as hnRNP proteins (Supplementary Table S1). Of particular relevance to this study, we found that TREX complex components, including THOC2, THOC5, THOC6, THOC3, THOC1 and UAP56, were specifically enriched in the CAR-E RNP compared with the CAR-E$_{m2}$ RNP (Table 1). Interestingly, we also observed a strong enrichment in the CAR-E RNP for components of the Prp19 complex and U2AF2 (Table 1), both of which function in mRNA export (18,32). Both the CAR-E and CAR-E$_{m2}$ RNPs contained similar numbers of peptides derived from the CAP-binding complex proteins (Table 1), consistent with the presence of the cap on both the CAR-E and CAR-E$_{m2}$ RNAs. However, in contrast to the CAR-E, the CAR-E$_{m2}$ RNP was not enriched in mRNA export factors, but was instead enriched in proteins such as PTB, ADAR, KHSRP and other RNA-binding proteins (Table 1 and Supplementary Table S1). Together,
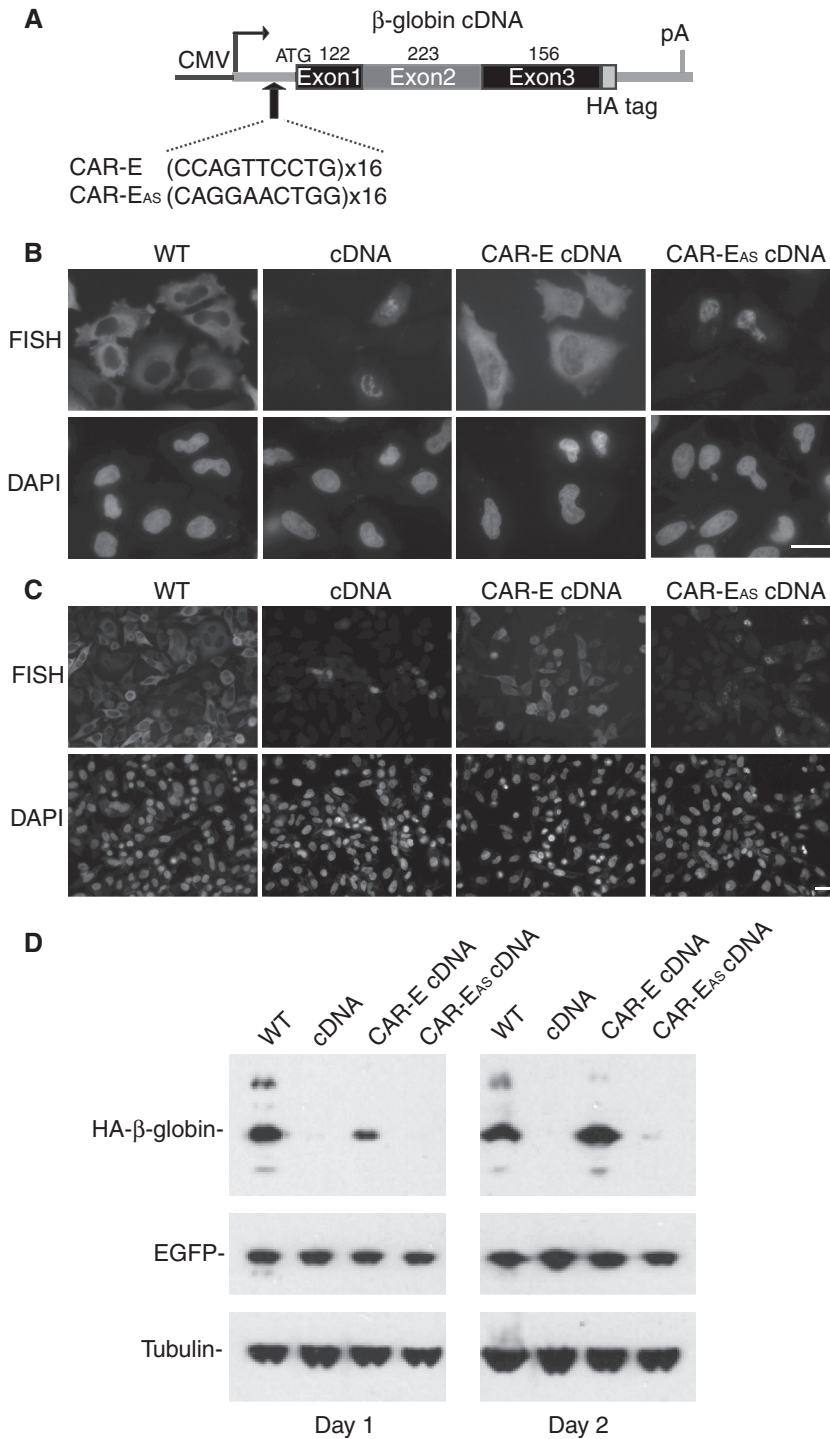
**A**

β-globin cDNA

CMV

ATG 122      223      156     pA

Exon1   Exon2   Exon3

HA tag

CAR-E  (CCAGTTCCTG)x16
CAR-E$_{AS}$ (CAGGAACTGG)x16

**B**



**C**



**D**



Day 1     Day 2

**Figure 2.** Tandem CAR-E promotes cytoplasmic accumulation of β-globin cDNA and β-globin protein expression. (**A**) Schematic of CMV-β-globin cDNA construct indicating the position where the CAR-E and CAR-E$_{AS}$ were inserted. The start codon (ATG), HA tag and BGH polyA signal (pA) are indicated. The sizes of each of the exons (in nucleotides) are shown. (**B**) FISH was used to determine the nucleocytoplasmic distribution of the indicated transcripts 24 h after transient transfection of the respective constructs into HeLa cells. DAPI staining was used to identify the nucleus. Scale bar: 10 μm. (**C**) Low magnification of the data shown in panel B. (**D**) Western analysis at 24 h (Day 1) or 45 h (Day 2) after transient transfection of the indicated constructs into HeLa cells. HA-β-globin, co-transfection control EGFP and tubulin (loading control) are indicated.

these data suggested that the reason that the tandem CAR-E, but not the tandem CAR-E$_{m2}$, facilitates cytoplasmic accumulation of β-globin cDNA is because the CAR-E recruits mRNA export factors.

**Knockdown of Prp19 components or U2AF2 results in nuclear retention of naturally intronless mRNAs**

Previously, we showed that TREX components (UAP56 and THOC2) and the mRNA export receptor Nxf1 were
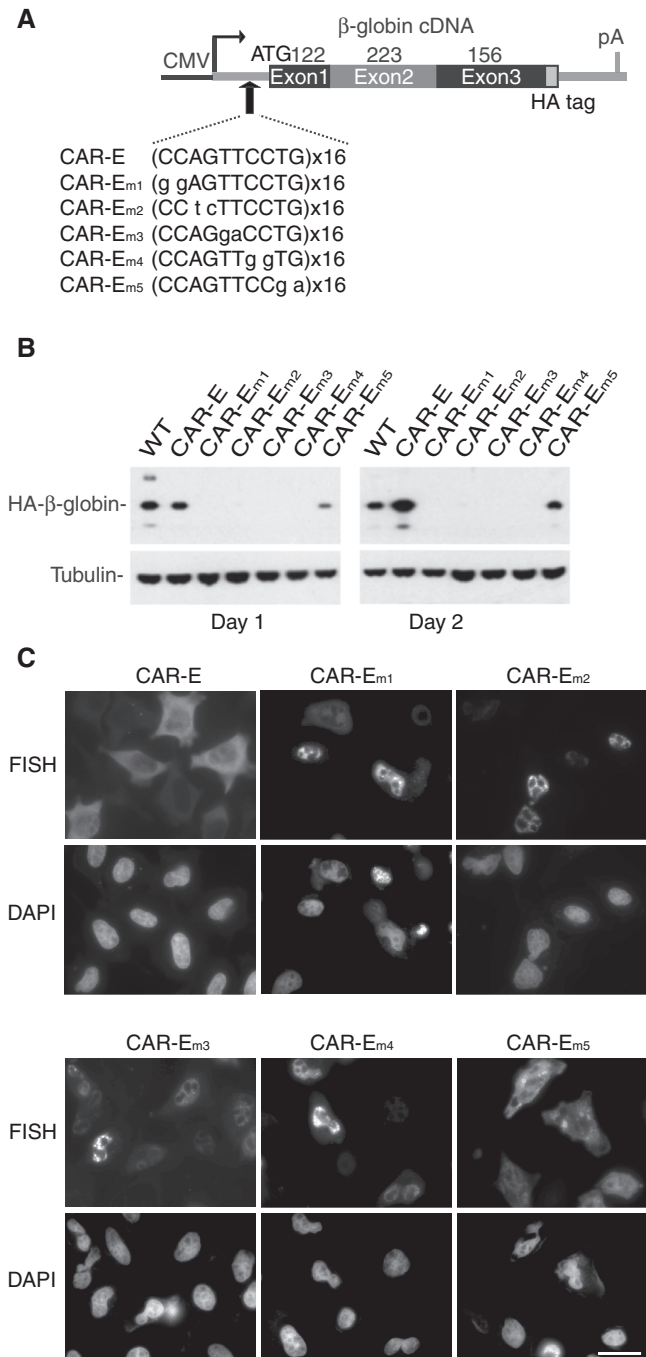
**Figure 3.** Tandem CAR-E containing dinucleotide substitutions is non-functional in RNA export. (**A**) Schematic of CMV-β-globin cDNA construct showing sequences of the CAR-E and each of the mutants that were inserted into the cDNA. Mutated nucleotides are shown in lower case. (**B**) Western analysis at 24 h (Day 1) or 45 h (Day 2) after transient transfection of the indicated constructs into HeLa cells. HA-β-globin and tubulin (loading control) are indicated. (**C**) Nucleocytoplasmic distribution of CAR-E$_{m1-5}$ transcripts was determined by FISH. DAPI staining was used to identify the nucleus. Scale bar: 10 μm.

required for cytoplasmic accumulation of HSPB3, IFNα1 and IFNβ1 (10). On the basis of our mass spectrometry data of the CAR-E RNP, we next sought to determine whether Prp19 complex components or U2AF2 also

functioned in cytoplasmic accumulation of naturally intronless mRNAs. We were especially interested in these factors because we recently found that several of the Prp19 complex components (e.g. XAB2, CRNKL1, CDC5L and PPIE) specifically associate with UAP56 (13). Moreover, recent work in *Saccharomyces cerevisiae* showed that the Prp19 complex is required for TREX occupancy at intronless genes and that components of this complex interact both genetically and biochemically with the TREX complex (12). Finally, it is known that U2AF2 associates with UAP56 (43) and functions in intronless mRNA export in *Drosophila* (32). To determine whether Prp19 components or U2AF2 function in cytoplasmic accumulation of naturally intronless mRNAs, we microinjected the IFNβ1 construct driven by the CMV promoter (Figure 4A) into the nuclei of cells knocked down using siRNAs against the Prp19 components Xab2, Aqr, Isy1, Crnkl1 (Figure 4B–I) or U2AF2 (Figure 4J and K). Cells transfected with non-targeting siRNA were used as a negative control. We note that we used microinjection rather than transfection of the CMV constructs for this analysis because we were examining knockdown cells. However, in a great deal of previous work [e.g. (10,11)] and in our present study, we found that expression of CMV constructs is the same for both microinjection and transfection. Significantly, this analysis revealed that IFNβ1 was retained in the nucleus when each of the Prp19 components or U2AF2 was knocked down (Figure 4C, E, G, I and K). These data indicate that Prp19 components and U2AF2 function in export of IFNβ1 mRNA. We obtained the same results when HSPB3 or IFNα1 driven by the CMV promoter were microinjected into the Prp19 component or U2AF2 knockdown cells (Supplementary Figure S2). Knockdown of other Prp19 components, including Prp19 itself, CCDC16, PPIE and PLRG1, resulted in partial or no nuclear retention of intronless mRNA (data not shown). However, the observation that knockdown of several Prp19 complex components and U2AF2, which associates with the Prp19 complex, results in nuclear retention of each of the naturally intronless mRNAs that we tested indicates a role for these factors in intronless mRNA export.

## DISCUSSION

In this study, we identified a 10-nt consensus sequence, the CAR-E, which is common to the CARs that we previously identified in three naturally intronless mRNAs (10) and is also present in the CAR found in c-Jun mRNA (33). The CAR-E is not present in sequences that do not promote export, such as the antisense of the CARs or regions surrounding the CARs. Evidence that the CAR-E is functional was provided by the observation that insertion of tandem copies of the CAR-E upstream of β-globin cDNA rescued the cytoplasmic accumulation of this cDNA transcript, which is normally retained/degraded in the nucleus. In contrast, the antisense of the tandem CAR-E was not functional in this assay. In addition, several dinucleotide mutations abolished the function of the tandem CAR-E,
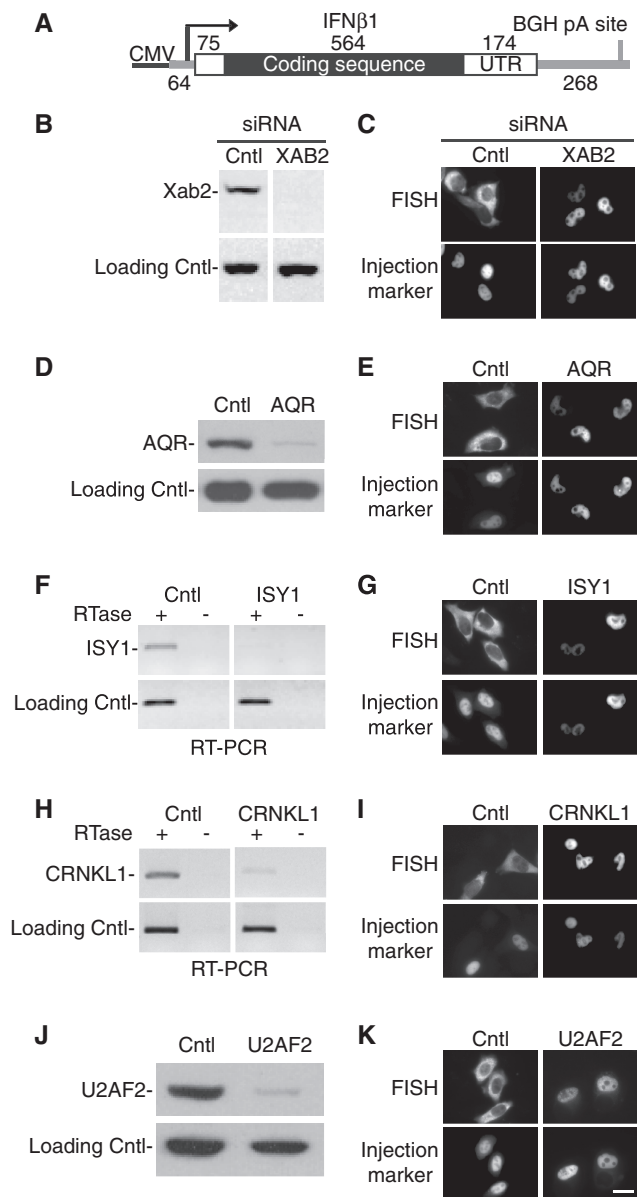
**Figure 4.** RNAi of Prp19 complex components or U2AF2 blocks export of naturally intronless IFNβ1 mRNAs. (**A**) Schematic of CMV-IFNβ1 construct used for microinjection. The length of the 5′-UTR, coding region and 3′-UTR are indicated. Vector sequences are shown as gray lines. BGH pA: bovine growth hormone polyA site. (**B**, **D**, **J**) Western blots showing levels of XAB2 (B), AQR (D) or U2AF2 (J) after transfecting HeLa cells with targeting or non-targeting (Cntl) siRNAs as indicated. Tubulin was used as a loading control. (**F** and **H**) RT–PCR showing levels of ISY1 or CRNKL1 mRNAs after transfecting HeLa cells with targeting or non-targeting (Cntl) siRNAs as indicated. Endogenous DNAJB1 was amplified by PCR as a loading control. Reverse transcription was carried out in reactions that contained (+) or lacked (−) reverse transcriptase (RTase) followed by PCR. (**C**, **E**, **G**, **I**, **K**) CMV-IFNβ1 naturally intronless construct was microinjected into nuclei of the indicated knockdown or control cells. FISH was carried out to determine the nucleocytoplasmic distribution of the RNAs. Dextran 70 kDa was used an injection marker. Scale bar: 10 μm.

indicating that the effect we observed is not due to a random sequence but is instead specific to the CAR-E sequence. The tandem CAR-E is similar to CARs in that they both promote cytoplasmic accumulation of

RNA. This effect is distinguished from elements that enhance RNA stability, such as the ENE in the Kaposi's sarcoma-associated herpesvirus, which stabilizes the RNA but does not affect RNA localization or protein production (44,45).

Despite the functionality of the multimerized CAR-E in promoting export of cDNA transcripts, we were unable to observe an export defect when the elements were mutated in their natural context in the CAR. However, our previous work showed that the antisense of the CARs have no activity in mRNA export (10). Thus, the sense strand of the CARs must contain elements or structures that promote cytoplasmic accumulation of the intronless mRNAs. One of these elements may be the CAR-E that we identified here, but additional sequences remain to be identified. Recently, a new, curated resource called the Intronless Gene Database was reported (46). It contains 687 human intronless genes, which represents ~3% of the genome (46,47). In a global *in silico* analysis of these genes, we identified six 10-nt motifs that were present in >90% of the intronless genes with one highly similar to CAR-E. It is possible that one or more of these motifs function in intronless mRNA export and/or that a structure(s) is involved. These possibilities require further investigation.

Consistent with the possibility that the CAR-Es function in intronless mRNA export, we found that the tandem CAR-E, but not the mutant CAR-E$_{m2}$, assembles into an RNP that is enriched in the TREX mRNA export machinery as well as Prp19 complex components and U2AF2. Interestingly, both the Prp19 complex and U2AF2 were previously found to function in the intronless mRNA export pathway in yeast and *Drosophila*, respectively (18,32). Moreover, knockdown of TREX components, Prp19 complex components (Xab2, Aqr, Crnkl1 or Isy1), or U2AF2 resulted in nuclear retention of the naturally intronless mRNAs [(10), this study]. Knockdown of other Prp19 complex components, such as Prp19, CCDC16, PPIE and PLRG1, had partial or no effect on export of intronless mRNAs (data not shown). These results may be explained by functional redundancy as observed for UAP56 and URH49 (9,13,48); insufficient knockdown to observe a robust phenotype or some of the Prp19 components may form different complexes with distinct functions.

In addition to the enrichment of TREX and Prp19 complex components, we observed that many other proteins, including splicing factors, were enriched in the CAR-E RNP (Supplementary Table S1). As splicing promotes export of mRNAs generated by splicing, an interesting possibility is that splicing factors also play a role in export of naturally intronless mRNAs. This possibility is supported by previous work showing a role for SR proteins in histone mRNA export (49) and the studies showing a role for UAP56 (43,50), U2AF2 (32) and the Prp19 complex in intronless mRNA export (18). Another possible explanation for the association of splicing factors with the CAR-E mRNP is that these factors participate in protein–protein interactions with export factors. We note that the TREX components Aly and CIP29 were not detected in the CAR-E mRNP. Extensive additional

studies are needed to determine whether other factors enriched in the CAR-E RNP are involved in export of naturally intronless mRNAs.

Our data, together with previous work, suggest that the mRNA export machinery can be recruited either in a sequence-dependent manner (this study) or in a splicing-dependent manner (7). Thus, the CAR-E may have general utility for enhancing the expression of cDNAs without the need for insertion of an intron.

## SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online: Supplementary Table 1 and Supplementary Figures 1–2.

## ACKNOWLEDGEMENTS

The authors thank Marlene Winkelbauer-Hurt, Eric Folco, Yong Yu and Tomohiro Yamazaki for useful discussions and comments on the manuscript. They also thank the Nikon Imaging Center at Harvard Medical School for assistance with microscopy.

## FUNDING

## REFERENCES

1. Cramer,P., Srebrow,A., Kadener,S., Werbajh,S., de la Mata,M., Melen,G., Nogues,G. and Kornblihtt,A.R. (2001) Coordination between transcription and pre-mRNA processing. *FEBS Lett.*, **498**, 179–182.
2. Maniatis,T. and Reed,R. (2002) An extensive network of coupling among gene expression machines. *Nature*, **416**, 499–506.
3. Moore,M.J. and Proudfoot,N.J. (2009) Pre-mRNA processing reaches back to transcription and ahead to translation. *Cell*, **136**, 688–700.
4. Perales,R. and Bentley,D. (2009) "Cotranscriptionality": the transcription elongation complex as a nexus for nuclear transactions. *Mol. Cell*, **36**, 178–191.
5. 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.
6. Reed,R. and Hurt,E. (2002) A conserved mRNA export machinery coupled to pre-mRNA splicing. *Cell*, **108**, 523–531.
7. Kohler,A. and Hurt,E. (2007) Exporting RNA from the nucleus to the cytoplasm. *Nat. Rev. Mol. Cell. Biol.*, **8**, 761–773.
8. Lu,S. and Cullen,B.R. (2003) Analysis of the stimulatory effect of splicing on mRNA production and utilization in mammalian cells. *RNA*, **9**, 618–630.
9. Dias,A.P., Dufu,K., Lei,H. and Reed,R. (2010) A role for TREX components in the release of spliced mRNA from nuclear speckle domains. *Nat. Commun.*, **1**, 97.
10. Lei,H., Dias,A.P. and Reed,R. (2011) Export and stability of naturally intronless mRNAs require specific coding region sequences and the TREX mRNA export complex. *Proc. Natl Acad. Sci. USA*, **108**, 17985–17990.
11. Valencia,P., Dias,A.P. and Reed,R. (2008) Splicing promotes rapid and efficient mRNA export in mammalian cells. *Proc. Natl Acad. Sci. USA*, **105**, 3386–3391.
12. Strasser,K., Masuda,S., Mason,P., Pfannstiel,J., Oppizzi,M., Rodriguez-Navarro,S., Rondon,A.G., Aguilera,A., Struhl,K., Reed,R. *et al.* (2002) TREX is a conserved complex coupling transcription with messenger RNA export. *Nature*, **28**, 28.
13. 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.
14. Cheng,H., Dufu,K., Lee,C.-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.
15. Abruzzi,K.C., Lacadie,S. and Rosbash,M. (2004) Biochemical analysis of TREX complex recruitment to intronless and intron-containing yeast genes. *EMBO J.*, **23**, 2620–2631.
16. 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 Pcf11. *Mol. Cell*, **33**, 215–226.
17. Zenklusen,D., Vinciguerra,P., Wyss,J.C. and Stutz,F. (2002) Stable mRNP formation and export require cotranscriptional recruitment of the mRNA export factors Yra1p and Sub2p by Hpr1p. *Mol. Cell. Biol.*, **22**, 8241–8253.
18. Chanarat,S., Seizl,M. and Straber,K. (2011) The Prp19 complex is a novel transcription elongation factor required for TREX occupancy at transcribed genes. *Genes Dev.*, **25**, 1147–1158.
19. David,C.J., Boyne,A.R., Millhouse,S.R. and Manley,J.L. (2011) The RNA polymerase II C-terminal domain promotes splicing activation through recruitment of a U2AF65-Prp19 complex. *Genes Dev.*, **25**, 972–983.
20. Kuraoka,I., Ito,S., Wada,T., Hayashida,M., Lee,L., Saijo,M., Nakatsu,Y., Matsumoto,M., Matsunaga,T., Handa,H. *et al.* (2008) Isolation of XAB2 complex involved in pre-mRNA splicing, transcription, and transcription-coupled repair. *J. Biol. Chem.*, **283**, 940–950.
21. Chan,S.P., Kao,D.I., Tsai,W.Y. and Cheng,S.C. (2003) The Prp19p-associated complex in spliceosome activation. *Science*, **302**, 279–282.
22. Boyne,J.R., Colgan,K.J. and Whitehouse,A. (2008) Herpesvirus saimiri ORF57: a post-transcriptional regulatory protein. *Front Biosci.*, **13**, 2928–2938.
23. Chen,I.H., Sciabica,K.S. and Sandri-Goldin,R.M. (2002) ICP27 interacts with the RNA export factor Aly/REF to direct herpes simplex virus type 1 intronless mRNAs to the TAP export pathway. *J. Virol.*, **76**, 12877–12889.
24. Guang,S., Felthauser,A.M. and Mertz,J.E. (2005) Binding of hnRNP L to the pre-mRNA processing enhancer of the herpes simplex virus thymidine kinase gene enhances both polyadenylation and nucleocytoplasmic export of intronless mRNAs. *Mol. Cell. Biol.*, **25**, 6303–6313.
25. Toth,Z. and Stamminger,T. (2008) The human cytomegalovirus regulatory protein UL69 and its effect on mRNA export. *Front Biosci.*, **13**, 2939–2949.
26. Wang,W., Cui,Z.Q., Han,H., Zhang,Z.P., Wei,H.P., Zhou,Y.F., Chen,Z. and Zhang,X.E. (2008) Imaging and characterizing influenza A virus mRNA transport in living cells. *Nucleic Acids Res.*, **36**, 4913–4928.
27. Huang,Y. and Carmichael,G.G. (1997) The mouse histone H2a gene contains a small element that facilitates cytoplasmic accumulation of intronless gene transcripts and of unspliced HIV-1-related mRNAs. *Proc. Natl Acad. Sci. USA*, **94**, 10104–10109.
28. Huang,Y., Wimler,K.M. and Carmichael,G.G. (1999) Intronless mRNA transport elements may affect multiple steps of pre-mRNA processing. *EMBO J.*, **18**, 1642–1652.
29. Huang,Y., Gattoni,R., Stevenin,J. and Steitz,J.A. (2003) SR splicing factors serve as adapter proteins for TAP-dependent mRNA export. *Mol. Cell*, **11**, 837–843.
30. Gilbert,W. and Guthrie,C. (2004) The Glc7p nuclear phosphatase promotes mRNA export by facilitating association of Mex67p with mRNA. *Mol. Cell*, **13**, 201–212.
31. Hargous,Y., Hautbergue,G.M., Tintaru,A.M., Skrisovska,L., Golovanov,A.P., Stevenin,J., Lian,L.Y., Wilson,S.A. and Allain,F.H. (2006) Molecular basis of RNA recognition and TAP

binding by the SR proteins SRp20 and 9G8. *EMBO J*, **25**, 5126–5137.

32. Blanchette,M., Labourier,E., Green,R.E., Brenner,S.E. and Rio,D.C. (2004) Genome-wide analysis reveals an unexpected function for the *Drosophila* splicing factor U2AF50 in the nuclear export of intronless mRNAs. *Mol. Cell*, **14**, 775–786.

33. Guang,S. and Mertz,J.E. (2005) Pre-mRNA processing enhancer (PPE) elements from intronless genes play additional roles in mRNA biogenesis than do ones from intron-containing genes. *Nucleic Acids Res.*, **33**, 2215–2226.

34. Bailey,T.L., Boden,M., Buske,F.A., Frith,M., Grant,C.E., Clementi,L., Ren,J., Li,W.W. and Noble,W.S. (2009) MEME SUITE: tools for motif discovery and searching. *Nucleic Acids Res.*, **37**, W202–W208.

35. Grant,C.E., Bailey,T.L. and Noble,W.S. (2011) FIMO: scanning for occurrences of a given motif. *Bioinformatics*, **27**, 1017–1018.

36. Chua,K. and Reed,R. (1999) Human step II splicing factor hSlu7 functions in restructuring the spliceosome between the catalytic steps of splicing. *Genes Dev.*, **13**, 841–850.

37. Lee,C.S., Dias,A.P., Jedrychowski,M., Patel,A.H., Hsu,J.L. and Reed,R. (2008) Human DDX3 functions in translation and interacts with the translation initiation factor eIF3. *Nucleic Acids Res.*, **36**, 4708–4718.

38. Villen,J., Beausoleil,S.A. and Gygi,S.P. (2008) Evaluation of the utility of neutral-loss-dependent MS3 strategies in large-scale phosphorylation analysis. *Proteomics*, **8**, 4444–4452.

39. Eng,J.K., McCormack,A.L. and Yates,J.R. (1994) An approach to correlate tandem mass spectral data of peptides with amino acid sequences in a protein database. *J. Am. Soc.Mass Spectr.*, **5**, 976–989.

40. Elias,J.E. and Gygi,S.P. (2007) Target-decoy search strategy for increased confidence in large-scale protein identifications by mass spectrometry. *Nat. Methods*, **4**, 207–214.

41. Huttlin,E.L., Jedrychowski,M.P., Elias,J.E., Goswami,T., Rad,R., Beausoleil,S.A., Villen,J., Haas,W., Sowa,M.E. and Gygi,S.P. (2010) A tissue-specific atlas of mouse protein phosphorylation and expression. *Cell*, **143**, 1174–1189.

42. Zhou,Z., Licklider,L.J., Gygi,S.P. and Reed,R. (2002) Comprehensive proteomic analysis of the human spliceosome. *Nature*, **419**, 182–185.

43. 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.

44. Conrad,N.K., Mili,S., Marshall,E.L., Shu,M.D. and Steitz,J.A. (2006) Identification of a rapid mammalian deadenylation-dependent decay pathway and its inhibition by a viral RNA element. *Mol. Cell*, **24**, 943–953.

45. Conrad,N.K. and Steitz,J.A. (2005) A Kaposi's sarcoma virus RNA element that increases the nuclear abundance of intronless transcripts. *EMBO J*, **24**, 1831–1841.

46. Louhichi,A., Fourati,A. and Rebai,A. (2011) IGD: a resource for intronless genes in the human genome. *Gene*, **488**, 35–40.

47. Grzybowska,E.A. (2012) Human intronless genes: functional groups, associated diseases, evolution, and mRNA processing in absence of splicing. *Biochem. Biophys. Res. Commun.*, **424**, 1–6.

48. Pryor,A., Tung,L., Yang,Z., Kapadia,F., Chang,T.H. and Johnson,L.F. (2004) Growth-regulated expression and G0-specific turnover of the mRNA that encodes URH49, a mammalian DExH/D box protein that is highly related to the mRNA export protein UAP56. *Nucleic Acids Res.*, **32**, 1857–1865.

49. Huang,Y. and Steitz,J.A. (2001) Splicing factors SRp20 and 9G8 promote the nucleocytoplasmic export of mRNA. *Mol. Cell*, **7**, 899–905.

50. Luo,M.J., 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.