Published in final edited form as:

Gene. 2012 December 15; 511(2): 161–168. doi:10.1016/j.gene.2012.09.088.

An RNA electrophoretic mobility shift and mutational analysis of *rnp-4f* 5'-UTR intron splicing regulatory proteins in *Drosophila* reveals a novel new role for a dADAR protein isoform

G. Girija Lakshmi, Sushmita Ghosh, Gabriel P. Jones, Roshni Parikh, Bridgette A. Rawlins, and Jack C. Vaughn*

Department of Zoology, Cell Molecular and Structural Biology Program, Miami University, Oxford, OH 45056, United States

Abstract

Alternative splicing greatly enhances the diversity of proteins encoded by eukaryotic genomes, and is also important in gene expression control. In contrast to the great depth of knowledge as to molecular mechanisms in the splicing pathway itself, relatively little is known about the regulatory events behind this process. The 5'-UTR and 3'-UTR in pre-mRNAs play a variety of roles in controlling eukaryotic gene expression, including translational modulation, and nearly 4,000 of the roughly 14,000 protein coding genes in *Drosophila* contain introns of unknown functional significance in their 5'-UTR. Here we report the results of an RNA electrophoretic mobility shift analysis of *Drosophila rnp-4f* 5'-UTR intron 0 splicing regulatory proteins. The pre-mRNA potential regulatory element consists of an evolutionarily-conserved 177-nt stem-loop arising from pairing of intron 0 with part of adjacent exon 2. Incubation of *in vitro* transcribed probe with embryo protein extract is shown to result in two shifted RNA-protein bands, and protein extract from a dADAR null mutant fly line results in only one shifted band. A mutated stem-loop in which the conserved exon 2 primary sequence is changed but secondary structure maintained by introducing compensatory base changes results in diminished band shifts. To test the hypothesis that dADAR plays a role in intron splicing regulation *in vivo*, levels of unspliced *rnp-4f* mRNA in dADAR mutant were compared to wild-type *via* real-time qRT-PCR. The results show that during embryogenesis unspliced *rnp-4f* mRNA levels fall by up to 85% in the mutant, in support of the hypothesis. Taken together, these results demonstrate a novel role for dADAR protein in *rnp-4f* 5'-UTR alternative intron splicing regulation which is consistent with a previously proposed model.

Keywords

Alternative splicing regulation; *cis*-regulatory elements; *trans*-regulatory elements; dADAR function; *rnp-4f* gene

© 2012 Elsevier B.V. All rights reserved.

*Corresponding author. Tel: +1 513 529 3150; fax: +1 513 529 6900. vaughnjc@muohio.edu.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

1. Introduction

Drosophila rnp-4f is a nuclear gene (Petschek et al., 1997; Hess et al., 2006) which is represented by orthologues in many other eukaryotic species, including yeast *Prp24* and human *p110* (Rader and Guthrie, 2002; Bell et al., 2002). The encoded protein RNP-4F is a splicing assembly factor and functions as a chaperone to unfold U6-snRNA prior to its association with U4-snRNA subsequent to forming a heterotrimer with U5-snRNA to enter the assembling spliceosome. Temperature-sensitive mutations in pre-mRNA processing (PRP) genes were early shown to result in the accumulation of unspliced pre-mRNAs in yeast (Rosbash et al., 1981), revealing the existence of at least 100 different genes which encode proteins important for splicing. More recent studies place their number closer to 300 (reviewed in Jurica and Moore, 2003). In contrast to our detailed understanding of the spliceosomal molecular mechanism (reviewed in Black, 2003; Wahl and Luhrmann, 2009), we know very little about the regulation of the PRP proteins. The fruit-fly *Drosophila* is an important model organism, having a completely sequenced and largely annotated genome containing about 14,000 genes (Adams et al., 2000) and offers many advantages for the understanding of higher organism molecular processes. Many evolutionarily conserved molecular pathways exist in flies and about 70% of human genetic disease genes have clear orthologues in *Drosophila* (Banfi et al., 1996; Fortini et al., 2000). Both inherited and acquired defects in pre-mRNA processing are being increasingly recognized as causes of human diseases (reviewed in Nissim-Rafinia and Kerem, 2002; Cooper et al., 2009), and nearly all pre-mRNA processing events are controlled by a combination of protein factors, including especially the PRP proteins. The elucidation of genetic mechanisms regulating pre-mRNA processing may lead to future diagnostic and therapeutic breakthroughs.

The *Drosophila rnp-4f* gene encodes two different major mRNA isoform classes. These isoforms are identical in their coding potential but differ primarily by 177-nt due to the alternative splicing pattern in the 5'-UTR around intron 0. These isoforms are developmentally regulated, the 5'-UTR unspliced variant being abundant during mid-embryo stages in comparison to the shorter isoform (Fetherson et al., 2006; Chen et al., 2007). These observations raise interesting questions as to the control mechanism and functional significance of the 5'-UTR alternative splicing decision. Insofar as the major difference between the two mRNA isoforms is found in their differing 5'-UTRs, our attention in this report is focused on this region to understand the regulation of the splicing event. The longer isoform contains a highly conserved secondary structure in which intron 0 pairs with part of adjacent exon 2 (Chen et al., 2009), itself highly conserved, to form a long stable stem-loop (Fig. 1). Evolutionary conservation shows the operation of natural selection and implies an important function for these features. *In situ* hybridization studies have shown that the longer *rnp-4f* mRNA isoform is localized in the developing fly central nervous system (Chen et al., 2007), as is *dADAR* mRNA (Palladino et al., 2000). *dADAR* encodes a protein which utilizes double-stranded RNAs for substrate in the A-to-I editing of some pre-mRNAs (reviewed in Bass, 2001; Gott, 2007). The single-copy *dADAR* gene encodes two different major mRNA isoforms, designated full-length (Palladino et al., 2000) and truncated (Ma et al., 2002; Chen et al., 2009), of which both contain a double-stranded RNA binding motif but only the longer variant contains a catalytic domain (Fig. 2).

Preliminary studies have shown that the abundance of the longer *mnp-4f* mRNA isoform in *dADAR* null mutant is about 30% less than that seen in wild-type during embryogenesis (Chen et al., 2009), suggesting a role for dADAR protein in the intron 0 alternative splicing decision. Taken together, these observations led us to hypothesize that dADAR protein may target *mnp-4f* transcripts within the 5'-UTR and regulate alternative splicing. Here we report the results of a combined RNA electrophoretic mobility shift (REMSA) and mutational analysis to test this hypothesis.

2. Materials and methods

2.1. *Drosophila* propagation and embryo preparation

D. melanogaster strain Oregon R lab stock was obtained from the Bloomington *Drosophila* Stock Center at Indiana University. A *dADAR* null mutant fly line was the generous gift of Enbo Ma (Ma et al., 2001). This viable homozygous mutant has previously been shown by Western analysis not to encode a protein (Chen et al., 2009). Fly propagation and embryo collection methods employed were as previously described (Peters et al., 2003). Staged embryos (Campos-Ortega and Hartenstein, 1997) were stored in aliquots at -80°C until used for RNA or protein isolations.

2.2. Preparation of RNA and protein extracts from *Drosophila* developmental stages

Total cell RNAs for use in REMSA and real-time qRT-PCR experiments were isolated from about 50 ug of tissue using the Maxwell 16 IVD instrument as specified by the supplier (Promega) for -80°C storage. RNA and protein concentrations were estimated by OD₂₆₀ and OD₂₈₀ spectrophotometry, respectively (NanoDrop Technologies). Proteins were isolated from 100 mg staged embryos (0–2 h, 2–4 h, 4–8 h, 8–12 h, and 12–16 h) as previously described (Chen et al., 2009), and stored at -80°C .

2.3. Preparation of 32-P labeled RNA stem-loop probes for REMSA

A DNA fragment containing the wild-type 177-nt 5'-UTR intron 0 regulatory region (Fig. 3A) was prepared by PCR amplification from genomic DNA, using primers whose design was based on the complete *mnp-4f* gene sequence (GenBank no. **NM 078492**). The forward primer sequence was 5'-TAT[**GGGCCC**]GGAAGTCTTGAAAGTTTTT-3' and contained an adapter (bracketed) for subsequent cutting with endonuclease *ApaI* prior to plasmid insertion, in addition to the two mutated nucleotides U→G and G→C (italicized) to prevent association with U1-snRNP (Lewin, 1994) which may have resulted in confusing bands during REMSA. This potential result could have developed from association of this snRNP with the RNA probe utilized, insofar as that probe contained an intron. The reverse primer sequence also contained adapters (bracketed) for subsequent cutting with endonuclease *SphI*: 5'-TAT[**GCATGC**]TTCCGTGGAAGTCAATTTTT-3' prior to plasmid insertion. An extensively mutated stem-loop probe was commercially synthesized (GenScript), with adapters at the ends. Here, virtually every nucleotide in the evolutionarily conserved exon 2 sequence was mutated (Fig. 3B), while changing the intron 0 sequence pairing with this region by introducing compensatory base changes, thereby retaining a secondary structure identical to that in wild-type.

Following cutting with restriction endonucleases *ApaI* and *SphI*, DNA fragments were positionally cloned into “pGEM-T Easy” plasmid vector (Promega). Orientation of fragment inserts within the vector with respect to the T7 and SP6 RNA polymerase promoters was verified by DNA sequencing. *In vitro* transcriptions of 32P-labeled sense (corresponds to mRNA sequence) and antisense strand (control and complementary to mRNA transcript) RNA probes were carried out using a commercially-available “Riboprobe” kit (Promega), following the directions of the supplier. The antisense control RNA molecule is not predicted to contain any significant stem structure (Zuker, 2003). The RNA probes were precipitated, washed, then dissolved in 60 uL binding buffer (Chabanon et al., 2005), consisting of 40 mM NaCl, 5 mM MgCl₂, 2 mM DTT, 30 mM Tris-HCl, pH 7.6, 40 U RNasin (Promega), heated to 90 °C for 5 min, and slowly cooled to room temperature for annealing. RNA probe preparations were purified by non-denaturing 5% polyacrylamide gel electrophoresis containing RNA sizers in 0.5× TBE (89 mM Tris-HCl, 0.5 mM disodium EDTA, 0.1% Tween-20, pH 8.3), stained with ethidium bromide, and excision of the RNA band. The purified RNA probe band was then recovered using a commercially available “ZR small RNA PAGE recovery” kit (Zymo Research). Specific activities of RNA probes were typically about 1×10^8 cpm/ug.

2.4. RNA electrophoretic mobility shift assay

The synthesized probes (1–10 ng/uL) and whole embryo protein extract (10–50 ug/mL) were combined and incubated at 25 °C for 30–60 min in binding buffer (Chabanon et al., 2005). Trials were run to determine the optimal amounts of probe, protein, incubation temperature, and incubation time (Thomson et al., 1999). The RNA-protein complex was then loaded onto a 5% non-denaturing polyacrylamide gel (acrylamide:bis-acrylamide in the ratio of 35:1 dissolved in 0.5× TBE) and run for 600 V-h at 4 °C. The probe alone and control RNA incubated with protein extract were also run. An RNA sizer (Ambion) was included in each gel. RNA bands were visualized by scanning into a STORM phosphorimager (Molecular Dynamics). The images shown in the REMSA figures were obtained over a period of two years during the course of this study. Differences in radioactivity levels between the various experiments make it impossible to compare band intensities between figures, although those within any given figure are readily comparable. The backgrounds encountered in the different gel images and revealed in the phosphorimager scans differ appreciably. The images shown are a true reflection of results typically obtained for each experiment, after many repeats.

2.5. Real-time qRT-PCR of *rnp-4f* 5'-UTR unspliced mRNA levels during development

A quality control step was utilized to assure that RNA samples were free from DNA contamination, using conventional RT-PCR. This was done by using PCR primers bracketing the 59-bp intron in the *rp49* gene (O'Connell and Rosbash, 1984), followed by 2% agarose gel electrophoresis to verify that only the expected, spliced, 480-bp cDNA species was present. Reverse transcription of 250 ng total cell RNAs in 20 uL reaction volume from wild-type and *dADAR* mutant and the associated negative controls were carried out using the GoScript Reverse Transcription System (Promega) from developmental stages 0–2 h, 2–4 h, 4–8 h, 8–12 h, and 12–16 h. The forward primer used was 5'-GCTTCAAGATGACCATCC-3' and the reverse primer was 5'-

AAGTCCGGTATATTAACGTT-3'. The cycling protocol used was: 95 °C 3 min, followed by 30 cycles of 95 °C 30 s, 55 °C 30 s, 72 °C 30 s.

A computer-assisted program (<http://biotools.umassmed.edu>) was utilized to design primers for realtime PCR which would result in products of about 100-bp length within the *rnp-4f* 5'-UTR region of interest as well as in *rp49*, which was used as the reference. The forward primer used for *rp49* was 5'-ATGCTAAGCTGTCGCACAAA-3' and the reverse primer was 5'-GTTCGATCCGTAACCGATGT-3'. The forward primer for *rnp-4f* located within intron 0 was 5'-TCCAAGAAAAAGTGCCCAAT-3' and the reverse primer located within exon 2 was 5'-CAGGGCTGGCCAATCTAATA-3'. A typical one-step PCR reaction mixture contained 1 uL RT product, with final primer concentrations of 0.2 uM in 25 uL volume. Reactions were carried out in triplicate for each sample using the iCycler instrument (BioRad) and GoTaq qPCR master mix (Promega), according to the manufacturer's directions. The quantification of mRNA levels was based on the mathematical model developed by Pfaffl (2001). Results were expressed as ratio intensity of *rnp-4f* unspliced mRNA isoform relative to the *rp49* reference for both wild-type and *dADAR* mutant at each developmental stage.

3. Results

3.1. Incubation of stem-loop RNA probe with embryo protein extract results in two shifted RNA-protein bands in wild-type

A modified wild-type *rnp-4f* 5'-UTR stem-loop DNA fragment was amplified from genomic DNA (Fig. 3A) in which two nucleotides were mutated at the 5'-splice acceptor site to prevent association with U1-snRNP (Lewin, 1994). This fragment was cloned into plasmid vector, linearized with either *ApaI* or *SphI* endonuclease according to whether the sense or antisense transcript was required, and 32P-labeled probe was produced by *in vitro* transcription. Following purification and annealing, 1 ng of RNA probe was incubated for 1 h at 25 °C with 50 ug of 8–16 h embryo protein extract, combining 8–12 and 12–16 h stage extracts. These developmental stages were utilized since it has previously been shown that a peak in unspliced *rnp-4f* mRNA isoform (Chen et al., 2007) and also *dADAR* mRNA (Chen et al., 2009) occurs at this period. The control RNA was also incubated in this way.

Following non-denaturing polyacrylamide gel electrophoresis, it was found (Fig. 4A, lanes 1 and 5) that two shifted bands, designated "S" and "L", were present. No shifted RNA band was observed for the control (lane 3). The more rapidly migrating "S" band is relatively broad, and may consist of two different RNA-protein complexes.

3.2. Protein binding to the stem-loop is specific and in the absence of dADAR protein the "S" band shift is not present

Prior to gel electrophoresis, some samples were incubated with proteinase K to verify that protein is responsible for the observed band shift, and this was the expected result (Fig. 4B, lane 1). Specificity of RNA-protein complex formation is commonly verified by competition experiments (Thomson et al., 1999; Murani et al., 2009). It was found that using increasing amounts of non-radioactive RNA probe (lanes 2 and 3) results in inhibition of band shift, showing that the binding is specific. Previous work (Chen et al., 2009) had suggested that

dADAR protein may bind to the *rnp-4f* 5'-UTR stem-loop structure. This idea was tested by utilizing protein extract from *dADAR* null mutant embryos, and it was found that no band shift occurs for the "S" RNA-protein complex (Fig. 4B, lanes 4 and 5), whereas the "L" band is only partially affected. This result is interpreted to show that dADAR protein binds to the RNA stem-loop. It was also found that recombinant full-length dADAR protein (Chen et al., 2009) results in a band shift, although the position of the band varied depending on the amount of purified protein used (results not shown).

3.3. Utilization of mutated *rnp-4f* long stem-loop RNA probes to localize cis-acting regions involved in dADAR protein binding

It has been shown (Fetherson et al., 2006; Chen et al., 2009) that the putative *rnp-4f* 5'-UTR splicing regulatory region contains two evolutionarily-conserved features. One is the potential to fold into a long stable stem-loop secondary structure and the other is a highly conserved bipartite nucleotide sequence in the region of exon 2 adjacent to intron 0. A mutated RNA probe was constructed in which virtually every nucleotide within the conserved exon 2 tract was changed in such a way as to conserve the secondary structure. Utilization of increasing amounts of wild-type embryo protein extract for REMSA while holding the probe level to 1 ng consistently showed that band shift is diminished for the "S" but not the "L" RNA-protein complex. This can be most clearly seen by viewing the RNA probe intensities at the base of each gel lane (Fig. 5A, lanes 2–4).

3.4. Developmental REMSA shows that the protein responsible for the "S" band shift is developmentally regulated

Wild-type RNA probes (1 ng) were incubated with embryo protein extract (50 ug) from wild-type developmental stages for REMSA. The results (Fig. 5B) show that the band intensity of the slower migrating "L" RNA-protein complex is similar at all developmental stages, so that the protein responsible for this band shift does not appear to be developmentally regulated. The diminished band intensity shown for the 2–4 h stage was not reproducibly observed. In contrast, the band intensity of the faster migrating "S" RNA-protein complex is consistently found to be diminished in 0–2 and 2–4 h stage embryos, but then is greater in later developmental stages. This can be most clearly seen by viewing the RNA probe intensities at the base of each gel lane. The protein responsible for these "S" band shifts is interpreted to be developmentally regulated.

3.5. Developmental real-time qRT-PCR comparing levels of 5'-UTR unspliced *rnp-4f* mRNA in dADAR mutant to wild-type

The REMSA results (Fig. 4B) show that in the absence of dADAR protein, levels of *rnp-4f* stem-loop RNA-protein binding are nearly undetectable in the "S" band, while mobility of the "L" band is unaffected, suggesting that dADAR protein may be directly binding to the stem-loop *in vitro*. To test the hypothesis that this effect also occurs *in vivo*, real-time qRT-PCR was utilized in which levels of unspliced *rnp-4f* 5'-UTR mRNA were compared between wild-type and *dADAR* mutant. The results for wild-type (Fig. 6) show that mRNA levels are elevated at 0–2 h, decline during later stages, rise to a peak at 8–12 h, then decline. In the *dADAR* mutant the profile is qualitatively similar, but levels of unspliced

rnp-4f mRNA are significantly lower than in wild-type, being 85% lower at 2–4 h, 50% lower at 4–8 h, and 30% lower at 8–12 h of development.

4. Discussion

4.1. The REMSA method for detection and analysis of RNA-binding regulatory proteins

The REMSA method has often been successfully used to detect regulatory proteins that bind to RNAs (Leedman et al., 1995; Thomson et al., 1999; Chabanon et al., 2005; Kramarova et al., 2008). In using this method, it is essential that appropriate controls are employed to show that the detected binding is specific. We have used three different methods of confirmation. In one, the complement to the RNA probe molecule was utilized and did not show any RNA band shift after incubation with embryo protein extract (Fig. 4A, lane 3). However, this control molecule did not contain duplex regions longer than eight nucleotides, as shown by computer-assisted folding analysis (Zuker, 2003), so that proteins binding nonspecifically to a probe containing a long RNA duplex could not be ruled out. In a second specificity control, competition experiments were employed. In this method (Thomson et al., 1999) unlabeled RNA probe was added to the binding mixture in increasing amounts compared to the labeled probe. The results show that when unlabeled probe was increased from 10 to 100 ng, mobility shift of both the “S” and “L” bands was abolished (Fig. 4B). This method has previously been successfully used to demonstrate binding specificity. For example, in a study of binding *Secreted Phosphoprotein 1* (SPP1) mRNA to the nuclear protein C/EBP β , association with labeled probe was completely eliminated in the competition experiment (Murani et al., 2009). Taken by itself, this second evidence for binding specificity is considered to be weak and could have been improved by utilization of an unrelated RNA structure.

A third demonstration of binding specificity, in our study, utilized embryo protein extract obtained from *dADAR* null mutant flies, which showed that in the absence of dADAR protein the more rapidly migrating shifted RNA-protein “S” band was absent (Fig. 4B, lanes 4 and 5). Insofar as it is well established that dADAR binds to long double-stranded RNA substrates which need not be perfectly base-paired (reviewed in Bass, 2002; Gott, 2007), this result is interpreted to show that the band shift observed for the more rapidly migrating RNA-protein duplex is due to dADAR binding. Utilization of recombinant full-length dADAR protein in our study showed that a band shift does occur, verifying that a dADAR-*rnp-4f* stem-loop complex can be formed *in vitro* (not shown). We cannot eliminate the possibility, however, that the observed elimination of band shifting using a dADAR mutant protein extract was indirect, wherein an unknown protein forms a complex with dADAR, since it has been shown that dADAR has the capacity to interact with other proteins and regulate their activity (Bhogal et al., 2011).

4.2. Combined REMSA and mutational analysis for identification of cis-acting nucleotide tracts

Examination of the *rnp-4f* 5'-UTR potential intron splicing regulatory region shows that the exon 2 component of the long stem-loop has a highly conserved bipartite sequence which is notably more conserved than the adjacent intron 0 (Fig. 1B). Such a conserved feature

implies the operation of natural selection and thereby a function, and we tested the hypothesis that this sequence may in part serve as a recognition site for binding of regulatory proteins using the REMSA method. This was studied by constructing RNA probe molecules in which virtually every nucleotide in the conserved exon 2 tract was changed in such a way that the conserved secondary structure would be retained, by introducing compensatory base changes in the intron 0 pairing tract. Utilization of increasing amounts of wild-type embryo protein extract for REMSA showed that band shift is diminished for the “S” but not the “L” RNA-protein complex (Fig. 5A, lanes 2–4). This was an unexpected result, since previous studies which utilized perfectly paired RNA duplexes did not show any strong sequence conservation feature for binding of ADAR proteins to duplex RNAs (Polson and Bass, 1994; Petschek et al., 1997). However in a more recent study, a sequence-specific requirement was demonstrated in the case of mammalian *GluR-2* R/G RNA editing (Stefl et al., 2010). Here, two edited stem-loops are contained in a 71-nt region. When mutated bases in the binding sites were studied for their affinity to ADAR binding using fluorescence anisotropy, it was found that binding affinity is considerably reduced compared to wild-type, showing that both primary and secondary structure of the binding site are important.

4.3. Developmental regulation of the protein (s) responsible for the RNA-dADAR REMSA band shift

The developmental REMSA results (Fig. 5B) show that the protein(s) responsible for the RNA-protein band shift in the “S” band is developmentally regulated. This protein is less abundant in the 0–2 and 2–4 h embryo stages, then increases in abundance during later developmental stages. This pattern is similar to that previously observed using semi-quantitative RT-PCR to study *dADAR* isoform mRNA levels during development (Chen et al., 2009). There, the full-length mRNA isoform was found to be present at similar levels during early embryo stages, then to rise to a peak by the 8–12 h stage. In contrast, the truncated *dADAR* mRNA isoform was shown to be highly abundant in 0–2 h embryos, followed by a decrease to basal levels at 2–4 and 4–8 h stages, after which the abundance rose to a peak at 8–12 h as for the full-length isoform. In comparing these developmental RT-PCR results to those obtained by REMSA (Fig. 5B), it is apparent that the full-length *dADAR* mRNA isoform profile is closer to that observed by REMSA, from which we infer that the full-length and not the truncated dADAR protein isoform is most likely the protein responsible for the RNA-dADAR REMSA band shift in the “S” band.

4.4. Involvement of dADAR protein in *rnp-4f* 5'-UTR intron splicing regulation as shown by real-time qRT-PCR

In previously reported work, semi-quantitative RT-PCR results showed that during embryogenesis unspliced *rnp-4f* mRNA levels are 30% less in the *dADAR* mutant at the 8–12 h developmental stage *in vivo* (Chen et al., 2009). This effect was attributed to an inhibitory action of dADAR protein on *rnp-4f* 5'-UTR intron splicing. We were suspicious as to the accuracy of this value insofar as our REMSA results, obtained *in vitro* in the present study, show that nearly all of the RNA-protein complex in band “S” disappears in the absence of dADAR protein. To further test the hypothesis that dADAR plays a role in *rnp-4f* 5'-UTR intron splicing regulation, levels of unspliced *rnp-4f* mRNA in *dADAR* mutant were compared to wild-type *via* real-time qRT-PCR, a more sensitive and precise

method to make the comparison. The results (Fig. 6) show that levels of unspliced *rnp-4f* mRNA are significantly lower than in wild-type, being 85% lower at 2–4 h, 50% lower at 4–8 h, and 30% lower at 8–12 h of development. Taken together, these results support the hypothesis that dADAR plays a role in *rnp-4f* 5'-UTR intron splicing regulation. It is however not yet clear as to whether the full-length or the truncated dADAR protein isoform (Fig. 2) is responsible for this function *in vivo*. Comparisons between REMSA and previously reported RT-PCR results, discussed above, favor the conclusion that the full-length dADAR protein isoform is responsible.

4.5. Involvement of ADARs during intron splicing decisions

Several studies have shown that the editase function of ADAR proteins can regulate intron splicing decisions (reviewed in Valente and Nishikura, 2005). For example, an AG-to-GG edit at the 3'-splice junction in *Drosophila* brain *rnp-4f* mRNA intron IV was early shown to inhibit intron splicing (Petschek et al., 1997). In mammalian brain ADAR2, autoediting of a nucleotide in an intron generates a new 3'-splice site, resulting in a 47-nt insertion leading to a frameshift creating a shorter and inactive protein lacking the carboxyl half of the protein molecule, which appears to be a negative feedback mechanism to regulate endogenous levels of ADAR2 (Rueter et al., 1999). In contrast, no RNA editing has been observed in the *Drosophila rnp-4f* long stem-loop during embryogenesis (Chen et al., 2009). This however does not in itself disprove the interpretation that full-length dADAR protein binds to the stem-loop *in vivo*, as virtually no significant RNA editing has been shown to exist in any *Drosophila* mRNAs during embryogenesis (Hanrahan et al., 2000; Ma et al., 2002), despite presence of the full-length dADAR protein itself (Chen et al., 2009).

The 5'-UTR and 3'-UTR in pre-mRNAs play a variety of roles in controlling eukaryotic gene expression, including translational modulation (reviewed in Mignone et al., 2002; Pickering and Willis, 2005). Our results demonstrate what is interpreted to be a novel role for an ADAR, in which *rnp-4f* 5'-UTR intron splicing is inhibited by steric hindrance and not by nucleotide editing, following protein binding to a long stem-loop generated by intron-exon pairing. In a highly comprehensive recent study, it was determined that nearly 4,000 of the roughly 14,000 protein coding genes in *Drosophila* contain introns of unknown functional significance in their 5'-UTR (J. Carlson, personal communication), of which many are present in the developing central nervous system. It will be interesting to determine whether dADAR plays a role during alternative intron splicing in other genes.

4.6. The functional significance of *rnp-4f* 5'-UTR regulatory region retention during embryogenesis

A working model has previously been proposed for the role of nuclear dADAR protein in *D. melanogaster rnp-4f* 5'-UTR intron alternative splicing regulation (Chen et al., 2009). According to this model, within the developing embryo dADAR and *rnp-4f* 5'-UTR unspliced mRNA localizations and presumably the proteins they encode are identical, both being primarily found in the germ band and subsequently the central nervous system. Binding of nuclear dADAR to the *rnp-4f* pre-mRNA 5'-UTR long stem-loop sterically hinders subsequent association with snRNP splicing factors to the intron splice junctions and branch point, and thus inhibits intron splicing. Following embryogenesis and also in most

adult tissues, excluding the central nervous system, dADAR protein is at relatively low levels and is not available for binding, enabling alternative splicing. This model predicts that when *dADAR* null mutant embryos are utilized, alternative splicing around intron 0 would be enhanced, as shown in the present study by the REMSA results (Fig. 4B). A second non-dADAR protein designated “X” is also observed to bind the stem-loop *in vitro* using REMSA and has not yet been identified. It is possible that this arises due to non-specific binding to a protein capable of association with any duplex RNA of sufficient length. In future research, it should be possible to identify the regulatory protein(s) which bind to the stem-loop, perhaps by using MALDI-TOF or a related technology (reviewed in Veraksa, 2010). The functional significance of *rnp-4f* 5'-UTR regulatory region retention during embryogenesis is the subject of our current research, and those results will be soon reported in a companion paper.

Acknowledgments

We especially wish to thank Enbo Ma for his generous gift of the *dADAR* null mutant fly stock, and Gary Janssen for helpful discussions on the REMSA technique. We also thank John Hawes and Xiaoyun Wu for their help in running the ABI 3100 DNA sequencings. We are grateful to Jing Chen for working out the real-time qRT-PCR protocol used in this work. This work was primarily supported by National Institutes of Health (NIH) Grant 1-R15-GM093895-01 to J. Vaughn. Support was also provided by award of a Miami University Graduate Assistantship to G. Lakshmi. Finally, support was provided by award of a Howard Hughes Summer Research Internship to R. Parikh, and by awards of Miami University USS Summer Research Internships to G. Jones and B. Rawlins. A preliminary report of these results was given at the 53rd Annual *Drosophila* Research Conference, Chicago, IL in March 2012.

Abbreviations

bp	base pair
dADAR	<i>Drosophila</i> adenosine deaminase acting on RNA
cDNA	DNA complementary to RNA
DEAM	deaminase domain in dADAR protein
DRBM/dsRBD	double-stranded RNA binding domain
mRNA	messenger RNA
NLS	nuclear localization signal
nt	nucleotide
PCR	polymerase chain reaction
PRP proteins	pre-mRNA processing proteins
REMSA	RNA electrophoretic mobility shift assay
RNP-4F	ribonucleoprotein encoded by <i>rnp-4f</i> gene in polytene band 4F
RT-PCR	reverse transcriptase-polymerase chain reaction
U-snRNP	U-series small nuclear ribonucleoprotein
3'-UTR	3'-untranslated mRNA region

5'-UTR

5'-untranslated mRNA region

References

- Adams MD, Celniker SE, Holt RA, Evans CA, et al. The genome sequence of *Drosophila melanogaster*. Science. 2000; 287:2185–2196. [PubMed: 10731132]
- Banfi S, Borsani G, Rossi E, Bernard L, Guffanti A, Rubboli F, Marchitelli A, Giglio S, Coluccia E, Zollo M. Identification and mapping of human cDNAs homologous to *Drosophila* mutant genes through EST database searching. Nat Genet. 1996; 13:167–174. [PubMed: 8640222]
- Bass, BL., editor. RNA Editing. Oxford University Press; New York, NY: 2001.
- Bell M, Schreiner S, Damianov A, Reddy R, Bindereif A. p110, a novel human U6 snRNP protein and U4/U6 snRNP recycling factor. EMBO J. 2002; 21:2724–2735. [PubMed: 12032085]
- Bhogal B, Jepson JE, Savva YA, Pepper AS, Reenan RA, Thomas A. Modulation of dADAR-dependent RNA editing by the *Drosophila* fragile X mental retardation protein. Nature Neurosci. 2011; 14:1517–1524. [PubMed: 22037499]
- Black DL. Mechanisms of alternative pre-messenger RNA splicing. Annu Rev Biochem. 2003; 72:291–336. [PubMed: 12626338]
- Campos-Ortega, JA.; Hartenstein, V. The Embryonic Development of *Drosophila melanogaster*. Springer-Verlag; Berlin: 1997.
- Chabanon H, Mickleburgh I, Burtle B, Pedder C, Hesketh J. An AU-rich stem-loop structure is a critical feature of the perinuclear localization signal of *c-myc* mRNA. Biochem J. 2005; 392:475–483. [PubMed: 16042622]
- Chen J, Concel VJ, Bhatla S, Rajeshwaran R, Smith DLH, Varadarajan M, Backscheider KL, Bockrath RA, Petschek JP, Vaughn JC. Alternative splicing of an *mnp-4f* mRNA isoform retaining an evolutionarily-conserved 5'-UTR intronic element is developmentally regulated and shown *via* RNAi to be essential for normal central nervous system development in *Drosophila melanogaster*. Gene. 2007; 399:91–104. [PubMed: 17582706]
- Chen J, Lakshmi GG, Hays DL, McDowell KM, Ma E, Vaughn JC. Spatial and temporal expression of *dADAR* mRNA and protein isoforms during embryogenesis in *Drosophila melanogaster*. Differentiation. 2009; 78:312–320. [PubMed: 19720447]
- Cooper TA, Wan L, Dreyfuss G. RNA and disease. Cell. 2009; 136:777–793. [PubMed: 19239895]
- Fetherson RA, Strock SB, White KN, Vaughn JC. Alternative pre-mRNA splicing in *Drosophila* spliceosomal assembly factor RNP-4F during development. Gene. 2006; 371:234–245. [PubMed: 16497447]
- Fortini ME, Skupski MP, Boguski MS, Hariharan IK. A survey of human disease gene counterparts in the *Drosophila* genome. J Cell Biol. 2000; 150:F23–F30. [PubMed: 10908582]
- Gott JM. RNA Editing. Meth Enzymol. 2007; 424:1–518.
- Hanrahan CJ, Palladino MJ, Ganetzky B, Reenan RA. RNA editing of the *Drosophila para* Na⁺ channel transcript: evolutionary conservation and developmental regulation. Genetics. 2000; 155:1149–1160. [PubMed: 10880477]
- Hess KA, Simone AA, Petschek JP. Spatial and temporal expression of *4f-rnp* gene in *Drosophila melanogaster*. Differentiation. 1996; 61:103–111. [PubMed: 8983176]
- Jurica MS, Moore MJ. Pre-mRNA splicing: awash in a sea of proteins. Mol Cell. 2002; 12:5–14. [PubMed: 12887888]
- Kramarova TV, Antonicka H, Houstek J, Cannon B, Nedergaard J. A sequence predicted to form a stem-loop is proposed to be required for formation of an RNA-protein complex involving the 3'-UTR of β -subunit F₀F₁-ATPase mRNA. Biochim Biophys Acta. 2008; 1777:747–757. [PubMed: 18538128]
- Leedman PJ, Stein AR, Chin WW. Regulated specific protein binding to a conserved region of the 3'-untranslated region of thyrotropin β -subunit mRNA. Mol Endo. 1995; 9:375–387.
- Lewin, B. Genes V. Oxford University Press; New York, NY: 1994.

- Ma E, Gu XQ, Wu X, Xu T, Haddad GG. Mutation in pre-mRNA adenosine deaminase markedly attenuates neuronal tolerance to O₂ deprivation in *Drosophila melanogaster*. *J Clin Invest*. 2001; 107:685–693. [PubMed: 11254668]
- Ma E, Tucker MC, Chen Q, Haddad GG. Developmental expression and enzymatic activity of pre-mRNA deaminase in *Drosophila melanogaster*. *Mol Brain Res*. 2002; 102:100–104. [PubMed: 12191498]
- Mignone F, Gissi C, Liuni S, Pesole G. Untranslated regions of mRNAs. *Genome Biol*. 2002; 3 reviews 0004.1–0004.10.
- Mount SM, Burks C, Hertz G, Stormo GD, White O, Fields C. Splicing signals in *Drosophila*: Intron size, information content, and consensus sequences. *Nucleic Acids Res*. 1992; 20:4255–4262. [PubMed: 1508718]
- Murani E, Ponsuksili S, Seyfert H, Shi X, Wimmers K. Dual effect of a single nucleotide polymorphism in the first intron of the porcine *Secreted phosphoprotein 1* gene: allele-specific binding of C/EBP β and activation of aberrant splicing. *BMC Mol Biol*. 2009; 10:96–106. [PubMed: 19845957]
- Nissim-Rafinia M, Kerem B. Splicing regulation as a potential genetic modifier. *Trends Genet*. 2002; 18:123–127. [PubMed: 11858835]
- O'Connell P, Rosbash M. Sequence, structure, and codon preference of the *Drosophila* ribosomal protein 49 gene. *Nucleic Acids Res*. 1984; 12:5495–5513. [PubMed: 6087289]
- Palladino MJ, Keegan LP, O'Connell MA, Reenan RA. *dADAR*, a *Drosophila* double-stranded RNA-specific adenosine deaminase is highly developmentally regulated and is itself a target for RNA editing. *RNA*. 2000; 6:1004–1018. [PubMed: 10917596]
- Peters NT, Rohrbach JA, Zalewski BA, Byrket CM, Vaughn JC. RNA editing and regulation of *Drosophila 4f-rnp* expression by *sas-10* antisense readthrough mRNA transcripts. *RNA*. 2003; 9:698–710. [PubMed: 12756328]
- Petschek JP, Scheckelhoff MR, Mermer MJ, Vaughn JC. RNA editing and alternative splicing generate mRNA transcript diversity from the *Drosophila 4f-rnp* locus. *Gene*. 1997; 204:267–276. [PubMed: 9434193]
- Pfaffl MW. A new mathematical model for relative quantification in real-time RT-PCR. *Nucleic Acids Res*. 2001; 29:2002–2007.
- Pickering BM, Willis AE. The implications of structured 5'-untranslated regions on translation and disease. *Semin Cell Dev Biol*. 2005; 16:39–47. [PubMed: 15659338]
- Polson AG, Bass BL. Preferential selection of adenosines for modification by double-stranded RNA adenosine deaminase. *EMBO J*. 1994; 13:5701–5711. [PubMed: 7527340]
- Rader SD, Guthrie C. A conserved Lsm-interaction motif in Prp24 required for efficient U4/U6 di-snRNP formation. *RNA*. 2002; 8:1378–1392. [PubMed: 12458792]
- Rosbash M, Harris PK, Woolford JL, Teem JL. The effect of temperature-sensitive RNA mutants on the transcription products from cloned ribosomal protein genes of yeast. *Cell*. 1981; 24:679–686. [PubMed: 7018694]
- Rueter SM, Dawson RB, Emeson RB. Regulation of alternative splicing by RNA editing. *Nature*. 1999; 399:75–80. [PubMed: 10331393]
- Stefl R, Oberstrass FC, Hood JL, Jourdan M, Zimmermann M, Skrisovska L, Maris C, Peng L, Hofr C, Emeson RB, Allain FH. The solution structure of the ADAR2 dsRBM-RNA complex reveals a sequence-specific readout of the minor groove. *Cell*. 2010; 143:225–237. [PubMed: 20946981]
- Thomson AM, Rogers JT, Walker CE, Staton JM, Leedman PJ. Optimized RNA gel-shift and UV cross-linking assays for characterization of cytoplasmic RNA-protein interactions. *BioTechniques*. 1999; 27:1032–1040. [PubMed: 10572651]
- Valente L, Nishikura K. ADAR gene family and A-to-I RNA editing: diverse roles in post-transcriptional gene regulation. *Prog Nucleic Acid Res Mol Biol*. 2005; 79:299–338. [PubMed: 16096031]
- Veraksa A. When peptides fly: advances in *Drosophila* proteomics. *J Proteomics*. 2010; 73:2158–2170. [PubMed: 20580952]
- Wahl MC, Luhrmann R. The spliceosome: design principles of a dynamic RNP machine. *Cell*. 2009; 136:701–718. [PubMed: 19239890]

Zuker M. Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Res.* 2003; 31:3406–3415. [PubMed: 12824337]

Highlights

- A conserved primary and secondary structure exists in the *rnp-4f* pre-mRNA 5'-UTR
- The REMSA method shows that two proteins bind to the *rnp-4f* 5'-UTR stem-loop
- Mutated 5'-UTR study shows both primary and secondary structure are essential
- Mutated *dADAR* gene study shows that its encoded protein binds to the stem-loop
- A novel role for dADAR protein is shown for *rnp-4f* alternative splicing regulation

Fig. 1.

Structure and sequence alignment around the *D. melanogaster rnp-4f* 5'-UTR regulatory region. (A) Orientation diagram showing locations of exons 1 and 2 (boxed) interrupted by intron 0. The indicated 177-nt tract is predicted to fold into a long stable evolutionarily-conserved stem-loop in many *Drosophila* species (Chen et al., 2009). (B) Phylogenetic sequence alignment between *D. melanogaster* and related species: *D. sechellia*, *D. mauritiana*, *D. simulans*, *D. yakuba*, and *D. erecta* for the homologous 177-nt tract. Nucleotides conserved in at least five of the six species are boxed within intron 0 and the exon 2 pairing region. The putative branch point consensus sequence within intron 0 is underlined (Mount et al., 1992).

Fig. 2.

Structure of the *D. melanogaster* *dADAR* gene and its two major isoform classes. Introns are indicated by thin lines. (A) Splicing patterns for full-length isoform class present in both embryos and adults (Palladino et al., 2000). Positions of alternate promoters (P1, P2), alternative start codons (AUG), nuclear localization signal (NLS), double-stranded RNA binding domains (dsRBD-1, dsRBD-2), and deaminase domain (DEAM) are indicated. (B) Splicing patterns for truncated isoform present only in embryos (Ma et al., 2002; Chen et al., 2009). This polyadenylated isoform terminates within intron 6. (C) Comparison of protein domains between full-length and truncated *dADAR* isoform protein classes.

Fig. 3.

Potential secondary structures and sequences of *rnp-4f* wild-type and mutated 5'-UTR regulatory region RNA molecules used in this study. Structures are each split into two parts to enhance clarity of the nucleotides. (A) Structure of the modified wild-type long stem-loop RNA. Two nucleotides at the 5'-acceptor site are mutated as indicated, to diminish binding of U1-snRNP. Positions of the 5'-end and two alternative 3'-ends are marked by heavy arrows. The conserved exon 2 sequence region involved in pairing is bracketed. (B) Structure of mutated long stem-loop RNA in which the majority of nucleotides in the conserved exon 2 region were changed (limited to region between arrows) in such a way that the conserved secondary structure is retained. Positions of the 5'- and 3'-intron splice junctions are shown by heavy arrows.

Fig. 4.

REMSA to detect *rnp-4f* long stem-loop binding proteins in *Drosophila* embryos. (A) Incubation of wild-type stem-loop 32P-labeled RNA probe (Fig. 3A) with combined 8–12 and 12–16 h stage wildtype total protein embryo extract. Non-denaturing polyacrylamide gel electrophoresis detects two shifted bands (“S” and “L”), representing RNA-protein complexes (lanes 1 and 5), in comparison to mobility for the RNA probe alone (lanes 2 and 4) and the control (lane 3). (B) Incubation of 32P-labeled stem-loop RNA probe with (separately) 8–12 and 12–16 h stage wild-type total protein embryo extract, followed by proteinase K digestion shows no band shift (lane 1). RNA competition using increasing amounts of non-radioactive RNA probe (lanes 2 and 3) shows that the protein binding is specific to the stem-loop. Protein extract from *dADAR* mutant (lanes 4 and 5) shows that no band shift occurs for the more rapidly migrating RNA-protein complex “S” in the absence of dADAR protein, while the “L” band shift is unaffected.

Fig. 5.

Utilization of mutated *rnp-4f* long stem-loop RNA probes to localize *cis*-acting regions involved in protein binding. (A) Incubation of mutated stem-loop in which the majority of conserved exon 2 pairing region nucleotides have been changed (Fig. 3B) with different amounts of wild-type embryo protein extract. In comparison to migration of RNA probe in the absence of protein extract (lanes 1 and 5) and wild-type probe in the presence of protein extract (lanes 6–8), band shift in the “S” but not the “L” complex is diminished. (B) REMSA to study the developmental regulation of *rnp-4f* long stem-loop binding proteins. Incubation of wild-type stem-loop 32P-labeled RNA probe (Fig. 3A) with wild-type embryo total protein from developmental stages shows that the protein responsible for the “S” RNA-protein band shift is developmentally regulated but that the protein responsible for the “L” RNA-protein band shift is not. The band intensities are diminished at 0–2 and 0–4 h stages, then increase at later stages for the “S” complex. The diminished band intensity shown here for the “L” RNA-protein band at 2–4 h was not reproducibly observed.

Fig. 6.

Developmental real-time qRT-PCR comparing levels of 5'-UTR unspliced *rnp-4f* mRNA in *dADAR* mutant to wild-type *in vivo*, relative to constitutively expressed *rp49*. (A) The *rnp-4f* mRNA abundance in wild-type across development. (B) The *rnp-4f* mRNA abundance in *dADAR* mutant across development. In comparison to wild-type, mRNA levels in the mutant are significantly reduced at several stages (asterisks) during embryonic development. Student's t-test indicated significant difference between control and experimental embryos ($P < 0.01$), as indicated.