Regulation of the ELAV target ewg: insights from an evolutionary perspective

Matthias Soller1, Min Li and Irmgard U. Haussmann
School of Biosciences, University of Birmingham, Edgbaston, Birmingham B15 2TT, U.K.

Abstract
ELAV (embryonic lethal abnormal visual system)/Hu family proteins are prototype RNA-binding proteins with binding preferences for AU-rich regions. Due to frequent occurrence of AU-rich motifs in introns and untranslated regions, it is poorly understood how gene-specific RNA-binding proteins, such as ELAV/Hu family members, recognize their complement of target RNAs in a complex cellular environment. The powerful genetic tools of Drosophila make the fruitfly an excellent model to study alternative mRNA processing in vivo in a developing organism. Recent sequencing of 12 Drosophila genomes will provide a novel resource to enhance our understanding of how gene-specific regulation of mRNA processing is achieved by ELAV/Hu family proteins.

Introduction
ELAV (embryonic lethal abnormal visual system) from Drosophila is the founding member of a family of RNA-binding proteins including human Hu. In Drosophila, there are three members [ELAV, Fne (found in neurons) and RBP9 (RNA-binding protein 9)] and in humans, there are four (HuB–HuD and HuR). ELAV/Hu family proteins are prototype RNA-binding proteins containing three RRMs (RNA recognition motifs). ELAV, like its human homologue HuD, is expressed in neurons and is one of the first signs of neuronal differentiation [1].

Historically, human Hu proteins have first been implicated in cytoplasmic post-transcriptional regulation, such as mRNA stability, localization and translatability [2]. Due to ELAV’s nuclear localization in Drosophila, ELAV has mainly been associated with alternative splicing regulation [3]. Recently, a nuclear function for HuD has also been identified in regulating the alternative splicing of the calcitonin/calcitonin related gene by inhibiting 3’-end processing [4], a mechanism that has first been described for ELAV in the Drosophila ewg (erect wing) gene [5]. Therefore ELAV/Hu family proteins are very versatile and might be involved in any type of alternative mRNA processing.

ELAV is a gene-specific regulator of pre-mRNA processing
Only a few neuronal genes that are alternatively spliced are also regulated by ELAV [3,6]. Hence, ELAV is a gene-specific regulator of alternative RNA processing. Currently known targets are ewg, a transcriptional regulator homologous with human NRF-1 (nuclear respiratory factor 1), nrg (neuroglian), a cell adhesion molecule homologous with human NCAM L1 and arm (armadillo), a cell adhesion molecule and transcriptional regulator of Wnt/Wingless signalling, ELAV also autoregulates its own expression [7]. In addition, ELAV has been implicated in regulating the expression of comm (commissureless) required for midline crossing of extending axons [8]. In ewg and nrg genes, ELAV regulates alternative splicing of terminal exons (Figure 1A) [5,9]. ELAV is both necessary in photoreceptor neurons and sufficient in non-neuronal wing imaginal discs for neuron-specific alternative splicing of ewg, nrg and arm (Figures 1B–1E) [3,6].

EWG protein expression requires an ELAV-regulated splicing mechanism
So far, the ELAV target studied in most detail is ewg, where ELAV is required for splicing of the last intron (Figure 1A) [5,6,10,11]. In the absence of ELAV, all ewg transcripts are polyadenylated in intron 6 and no EWG protein is expressed in photoreceptor neurons (Figures 1B–1F) [5]. A combination of in vitro binding assays and in vivo transgene analysis identified an ELAV-binding site of ~130 nt in the distal proximity of a poly(A) site (pA2) in the last ewg intron 6 (Figure 1A) [5,11]. In vitro, ELAV inhibits cleavage of ewg pA2 in a sequence- and dose-dependent manner. In vivo, mutating the ELAV-binding site abolishes splicing of ewg intron 6. Interestingly, ELAV seems not to inhibit recognition of the regulated poly(A) site by CstF (cleavage stimulatory factor), but rather interferes with later steps in 3’-end processing. Binding of ELAV together with a partial poly(A) complex consisting of at least CPSF (cleavage and polyadenylation specificity factor) and CstF is further required to inhibit inclusion of microexon I. In the presence of ELAV, splicing of ewg intron 6 merely occurs by default. Intriguingly, splicing of the last ewg intron is also required for EWG protein expression. The mechanism that leads to expression of EWG protein only from intron 6 spliced RNAs, however, is not known.
ELAV-regulated splicing of ewg intron 6 is required for EWG protein expression

(A) Schematic diagram of the ewg gene. The open reading frame of the only EWG protein detected is shown in yellow [10]. The ELAV-regulated poly(A) site (pA2) in intron 6 is indicated in red. (B-E) Eye imaginal discs with ELAV-null clones with photoreceptors stained with anti-ELAV (B), anti-EWG (C) and CD8GFP (D, driven by elavGAL4 in the clone). A merged picture of (B-D) is shown in (E). Note that the anti-EWG polyclonal antibody recognizes epitopes in all parts of EWG [10]. Clones were induced in female animals of the following genotype: elav e5 UASCD8GFP FRT19A/tubGAL80 hsflp FRT19A; elavGAL4 (3rd). Scale bar (E), 25 μm. (F) Schematic diagram of ELAV-regulated ewg intron 6 processing. In the presence of ELAV, as in photoreceptor neurons, ewg intron 6 is spliced. In the absence of ELAV, all transcripts are polyadenylated in intron 6 and no EWG protein is expressed. (G) Model of ELAV-regulated splicing of ewg intron 6. ELAV binds together with a partial polyadenylation complex consisting of at least CPSF and CstF to the regulated pA2 site and inhibits 3′-end processing. ELAV together with CPSF and CstF is further required to inhibit inclusion of microexon I. Splicing of exons H–J then occurs by default.

A phylogenetic approach to define the ELAV ewg binding site

Rather unexpectedly, the binding site of ELAV in the ewg gene extends over ~130 nt and harbours several AU4−6 motifs. To identify elements that are important for ELAV binding, we used a phylogenetic analysis of the ewg binding site in very closely related Drosophila species of the melano-gaster group that separated approx. 8 million years ago. The RNA-binding part of ELAV is highly conserved even in distantly related species that separated approx. 40 million years ago (see below, Figure 2). In contrast, intronic sequences start to diverge much earlier. Through purifying selection, nucleotides important for ELAV binding change with a lower rate than those unimportant for binding. This led to the identification of six conserved AU4−6 motifs. Consistent with such an extended binding site, we have shown that ELAV forms a defined multimeric complex on ewg RNA in vitro [11]. Results from a number of biochemical experiments indicate that ELAV forms a dodecameric complex on ewg RNA in vitro. Evolutionary conservation of six AU4−6 motifs mirrors the stoichiometry of the ELAV complex. Binding of a multimeric complex in vivo is supported by mutational analysis with transgenes. Mutations in multiple AU4−6 motifs are required to effectively inhibit splicing of ewg intron 6. Furthermore, introducing spacer sequences did not inhibit ELAV function, consistent with a proposed model whereby the spacer part is extruded from a multimeric complex [12]. Flexibility in the positioning of AU4−6 motifs is also indicated by deletions in the ewg ELAV-binding site in closely related species. Although we have shown that multiple and spaced AU4−6 motifs are important for ELAV binding, we have been unable to predict further ELAV target genes by bioinformatics approaches due the frequent occurrence of AU motifs in introns and untranslated regions.

Further insights into ELAV gene-specific regulation

Recent sequencing of 12 Drosophila genomes provides a wealth of information for structure–function analysis of ELAV binding to target RNA [13]. Alignment of ELAV proteins of the 12 Drosophila species reveals an astonishing conservation of ELAV protein in the RNA-binding part. Essentially, 100% of the three RRMs are conserved and this part, termed RBD60, is sufficient for full RNA-binding activity, and
rescues lethality and visual system defects of elav mutants (Figure 2) [11,14]. The only difference in the RNA-binding part found in some species is the absence of a 13-amino-acid extension in the first RRM, which is also not conserved in the mosquitoes Culex pipiens and Aedes aegypti, the moth Bombyx mori and human HuD (Figure 2). The 13-amino-acid extension is not required for viability of elav mutants in Drosophila melanogaster [15] and seems to have been lost independently during evolution in closely related Drosophila simulans and Drosophila sechellia, and in Drosophila erecta. The A/Q N-terminal domain is specific to the Drosophila lineage and is not involved in RNA binding (Figure 2) [11].

The secrets of ELAV’s gene-specific binding probably lie, to a large part, in the nucleic acid sequence of its targets [12]. The availability of the genome sequences of the 12 Drosophila species will therefore be instrumental in deciphering the combinatorial code necessary for ELAV’s gene-specific binding.

This work is supported by the U.K. BBSRC (Biotechnology and Biological Sciences Research Council).

References
5 Lisbin, M.J., Qiu, J. and White, K. (2001) The neuron-specific RNA-binding protein ELAV regulates neuroglian alternative splicing in neurons and binds directly to its pre-mRNA. Genes Dev. 15, 2526–2538
9 Lisbin, M.J., Qiu, J. and White, K. (2001) The neuron-specific RNA-binding protein ELAV regulates neuroglian alternative splicing in neurons and binds directly to its pre-mRNA. Genes Dev. 15, 2546–2561

Received 18 February 2008
doi:10.1042/BST0360502