A conserved family of *elav*-like genes in vertebrates

PETER J. GOOD*

Laboratory of Molecular Genetics, National Institute of Child Health and Human Development, National Institutes of Health, Bethesda, MD 20892

ABSTRACT   A large family of genes encodes proteins with RNA recognition motifs that are presumed to bind RNA and to function in posttranscriptional regulation. Neural-specific members of this family include *elav*, a gene required for correct differentiation and maintenance of neurons in *Drosophila melanogaster*, and a related gene, HuD, which is expressed in human neuronal cells. I have identified genes related to *elav* and HuD in *Xenopus laevis*, zebrafish, and mouse that define a family of four closely related vertebrate *elav*-like genes (*elrA*, *elrB*, *elrC*, and *elrD*) in fish, frogs, and mammals. In addition to protein sequence conservation, a segment of the 3′-untranslated sequence of *elrD* is also conserved, implying a functional role in *elrD* expression. In adult frogs, *elrC* and *elrD* are exclusively expressed in the brain, whereas *elrB* is expressed in brain, testis, and ovary. During *Xenopus* development, *elrC* and *elrD* RNAs are detected by late gastrula and late neurula stages, respectively, whereas a nervous system-specific *elrB* RNA species is expressed by early tadpole stage. Additional *elrB* transcripts are detected in the ovary and early embryo, demonstrating a maternal supply of mRNA and possibly of protein. These expression patterns suggest a role for different *elav*-like genes in early development and neuronal differentiation. Surprisingly, *elrA* is expressed in all adult tissues tested and at all times during development. Thus, the widely expressed *elrA* is expected to have a related function in all cells.

The isolation of mutations in genes involved in development has identified many possible pathways of gene regulation in model organisms like *Drosophila melanogaster* and *Caenorhabditis elegans*. The subsequent identification of the corresponding gene products and a molecular analysis of their function has provided valuable clues to mechanisms underlying developmental gene regulation. A subset of developmentally important genes encodes RNA-binding proteins (RBPs) that contain a conserved domain called the RNA recognition motif (RRM). In *Drosophila*, examples of these genes include *cpo*, *nonA*, *squid*, *sxl*, *tra2*, and *orb* (1–3). An elegant example of the importance of RBPs is the genetic hierarchy of genes involved in somatic sex determination in *Drosophila*. The master regulator of this pathway is *sxl*, a gene that encodes an RBP that has been shown to bind to target RNAs such as its own pre-mRNA and the pre-mRNA of a downstream gene, *tra*, and determine sex-specific splice-site selection (4, 5). The identification of genes encoding RBPs important in invertebrate development allows the isolation of related genes in vertebrates that may have homologous functions.

Proteins that contain RRMs represent a large family of genes. The RRM motif is ≈80 aa in length and has been reported in >100 proteins (6, 7). These proteins have a modular domain structure where one to four RRMs are present in addition to auxiliary domains that may mediate either nucleic acid binding or protein–protein interactions (1, 8). RRMs differ in RNA-binding specificity: some RRMs bind to distinct RNA recognition sites, and others appear to bind RNA nonspecifically (1, 2, 8). RRM-containing proteins have

been demonstrated to be involved in many different posttranscriptional events including RNA processing, RNA transport, and translation (1, 2).

One example of a gene encoding an RBP that is required for proper development of the nervous system is *elav*. In *Drosophila*, *elav* was identified by mutants with an embryonic lethal, abnormal visual system phenotype (9, 10). Hypomorphic alleles of *elav* were also isolated in screens for behavioral phenotypes (11). A genetic analysis of the function of *elav* suggests a role in the development and maintenance of the nervous system (9, 10, 12, 13). Consistent with this broad function, the Elav protein is detected in the nervous system when neurons become postmitotic and is present in most, if not all, neurons (14–16). The sequence of the Elav protein predicts an RBP containing three RRMs along with auxiliary motifs (17).

Additional members of the *elav*-like gene (ELG) family have been identified in *Drosophila* and in humans. In *Drosophila*, *rbp9* was identified in a screen for cDNAs that encode RRMs (18, 19). The *rbp9* gene is expressed later than *elav* and is present throughout the central nervous system in the adult (19). In humans, autoantibodies in patients with small cell lung carcinoma-associated paraneoplastic neurologic syndrome recognize a family of proteins, HuC/PLE21, HuD, and Hel-N1, with substantial homology to Elav that have three RRMs in a similar arrangement but lacking the auxiliary motif at the amino terminus (20–22). These proteins are expressed in neurons and cell lines with neural characteristics (20, 22–25). Nothing is known about the developmental expression of the corresponding genes.

With the aim to analyze the function of ELGs during early neural development, I have isolated cDNAs encoded by ELGs in *Xenopus*, zebrafish, and mouse.† By comparing the predicted protein sequence for these genes to human ELGs, I conclude that there are at least four vertebrate ELGs. Analysis of tissue expression in adults demonstrates that two of these genes are strictly neural specific, one has a neural-specific RNA species and different transcripts in the ovary and testis, while the fourth gene is expressed widely. During development, each gene has a unique pattern of expression. Thus, two classes of ELGs are present, one that functions in the nervous system and early embryo and another that functions in most, if not all, cells.

MATERIALS AND METHODS

Nomenclature. The genes corresponding to the cDNAs in this study are named *elav*-like *ribonucleoprotein A, B, C,* or *D* (*elrA*, *elrB*, *elrC*, or *elrD*). The previously identified human genes corresponding to this nomenclature are Hel-N1 = *elrB* (20), HuC/ple21 = *elrC* (21), and HuD = *elrD* (22).

Isolation of PCR Products Encoding Elav-Like Proteins. Degenerate primers corresponding to conserved amino acids between the Elav and HuD proteins, one sense (elav5S, PKTMTQ) and two antisense (elav6A, MTNYDE; and elav7A, FKTNKA), were synthesized. PCR amplification with these degenerate primers was performed as described (26, 27). For *Xenopus*, cDNA from a brain library (28) was used as template with primers elav5S and elav7A. For zebrafish and mouse, products were amplified by nested PCR amplification, first with the elav5S and elav7A primers, then with the elav5S and elav6A primers, from post-somitogenesis library DNA (a gift of David Grunwald, University of Utah) or mouse brain cDNA (Clontech), respectively. The relevant amplified bands were gel purified and inserted into the *Eco*RI site of pBluescript KS(+).

Isolation and Sequencing of *Xenopus* cDNA Inserts. Tadpole brain (a gift of Reiko Toyama, National Institutes of Health) and embryonic head (29) cDNA libraries in λZAPII were screened at low and high stringency with the cloned PCR products. This screen identified cDNAs with complete coding regions for *elrD*, *elrC*, and *elrB*. The same filters were reprobed at low stringency with a zebrafish *elrA* cDNA to identify a full-length cDNA for *elrA*. Two distinct types of cDNAs were identified for each gene, consistent with the pseudotetraploid nature of *Xenopus laevis* (30). Only one representative cDNA for each gene was sequenced on both strands using synthetic oligonucleotides. All sequences were analyzed with the Genetics Computer Group package of computer programs (31).

Isolation and Sequencing of Zebrafish cDNA Inserts. The zebrafish postsomitogenesis cDNA library in λZAPII was screened at low stringency with a *Xenopus elrD* cDNA probe. This screen identified a cDNA containing the complete coding region of the *elrD* gene and partial cDNAs for *elrA* and *elrC*.

RNA Analyses. Northern blots of *Xenopus* tissue and embryonic RNAs were performed as described (32). Blots were stripped of probe by washing twice with 10 mM Tris Cl (pH 8.0), 1 mM EDTA, and 0.1% SDS at 95°C for 10 min. The blots were exposed to film before reprobing to verify sufficient probe removal before reuse. After hybridization with cDNA probes, the filters were hybridized with an oligonucleotide to the 18S ribosomal RNA that was labeled with ^{33}P to ensure equivalent RNA amounts in each lane.

Reverse transcription–PCR reactions were performed as described (33), using 20 cycles of 94°C for 30 sec, 55°C for 15 sec, and 72°C for 60 sec. The primers corresponded to the 5′ end (5′-ATGGCAGTCAGACTGTGTGA-3′) and 3′ end (5′-ACTTTGGTGTAACACTTT-3′) of the *Xenopus elrB* coding region. The products were ethanol precipitated and then analyzed by a Southern blot using an *elrB* coding-region probe.

RESULTS AND DISCUSSION

Four ELGs in *Xenopus*. Degenerate primer PCR and low-stringency cDNA library screening led to the isolation of cDNAs of four genes, named *elrA*, *elrB*, *elrC*, and *elrD*, that are related to the *Drosophila elav* and human HuD genes. The predicted molecular masses of these proteins (ElrA = 36.0 kDa, ElrB = 42.6 kDa, ElrC = 38.1 kDa, and ElrD = 40.3 kDa) are consistent with the observed sizes of proteins that react with human HuD autoimmune sera (23, 25). These predicted proteins had identical domain arrangements: a short amino-terminal domain followed by two consecutive RRMs, a tether (or linker) region of 57–70 aa, and a carboxyl-terminal RRM (Fig. 1). All of the RRMs are highly conserved between three of these proteins, ElrB, ElrC, and ElrD; the RRMs for the ElrA protein are less well conserved. The tether regions are more divergent, although clearly related in sequence, whereas the amino termini of ElrB, ElrC, and ElrD share limited sequence similarity.

At Least Four ELGs in Vertebrates. A similar strategy was used to identify ELGs from zebrafish and mouse. Three types of zebrafish cDNAs were isolated, while a single PCR fragment from mouse cDNA was obtained. Fig. 2 presents a comparison of the tether and RRM3 domains for all vertebrate Elav-like proteins that are available. This comparison demonstrates that the zebrafish cDNAs are homologs of the *Xenopus elrA*, *elrC*, and *elrD* genes, whereas the mouse cDNA is a homolog of the *Xenopus elrA*. The relationship between these proteins is demonstrated by the conserved amino acid substitutions within the same gene in the different species. The *Xenopus*, zebrafish, and mouse proteins, along with the published human Elav-like proteins, suggest a family of four distinct vertebrate ELGs. The human genes Hel-N1, HuC/ple21, and HuD correspond to the *elrB*, *elrC*, *elrD* genes, respectively, whereas

N-term	RRM 1	RRM 2	Tether	RRM 3		% HuD

					% HuD
1 D	MVMIISTMEPQVSNGPTSNTSNGPSSNSRNCPSPMQTGAVTDDST	95
C	M.T......T.AN...GC********VGILNG.NGEA.......	50
N-term B	MAVRLCDVASLLRSGSWAAEPWTGQV.AAM.T.L.....C.*NTANCP.TI..S..VESNN*.E...	48
A	M.NGYEDH.DDVCRD.IGR	17
68 D	KTNLIVNYLPQNMQEEFRSLFGSIGEIESCKLVRDKIT*GQSLGYGFVNYIDPKDAEKAINTLNGLRLQTKTIKVSYARPSSASIR	100
C	................K.....................................V..N..D.........K................	94
RRM1 B	................LK.....................E...	97
A	...........D.L....S....V..A..I...VA..H.........LNA....R............S......F.....E..K	79
154 D	DANLYVSGLPKTMTQKELEQLFSQYGRIITSRILVDQVTGVSRGVGFIRFDKRIEAEEAIKGLNGQKPSGAAEPITVKFAN	99
C	.......S.....N...M...L..S.........	94
RRM2 B	....P..T.........	96
A	.....I....R.....DV.DM.LPF.H..N..V....A..L....A.......S......ASF..H..P.SS........A	70
235 D	NPSQKTSQALLSQLYQSPNRRYPGPLHHQAQRFRLDNLLNMAYGV**KRFSPITIDGMTSLVGMNIPGHTGT	97
C	......G....TH...TTA...T......T....SPLSI.********P........SV.N.A.VSLT.P.TA	61
Tether B	.....VNHTI................AQ................GIKS....MA.......A.I.F...A..	79
A	..N.NKNV.....ICH..A..FG..V.......***************....MGV.H.S.ISSV.VASSATS	51
307 D	GWCIFVYNLSPDSDESVLWQLFGPFGAVNNVKVIRDFNTNKCKGFGFVTMTNYDEAAMAIASLNGYRLGDRVLQVSFKTNKTHKS	99
C	..........EA.........T.......T...S.Q..A	92
RRM3 B	.........A..A...I...M.......T...S....A	92
A	.....I...GQ.A..GI...M.......T...................E.................KT.......S.S..*	86

Fig. 1.   Alignment of four Elav-like proteins of *Xenopus*. The top line is a diagram of a prototypical vertebrate Elav-like protein. The different domains are indicated by the boxes and are labeled as described in the text. The predicted protein sequences encoded by cDNAs representing four separate ELGs were aligned by the Genetics Computer Group PILEUP program (31) and are identified at the left with the letter of each gene. Periods and asterisks indicate amino acid identity and gaps, respectively, relative to the ElrD sequence. Each set of rows is labeled at the left with the domain name. The numbers at the right are the percent identity to the same region of the human HuD (ElrD) protein. N-term, amino-terminal domain.
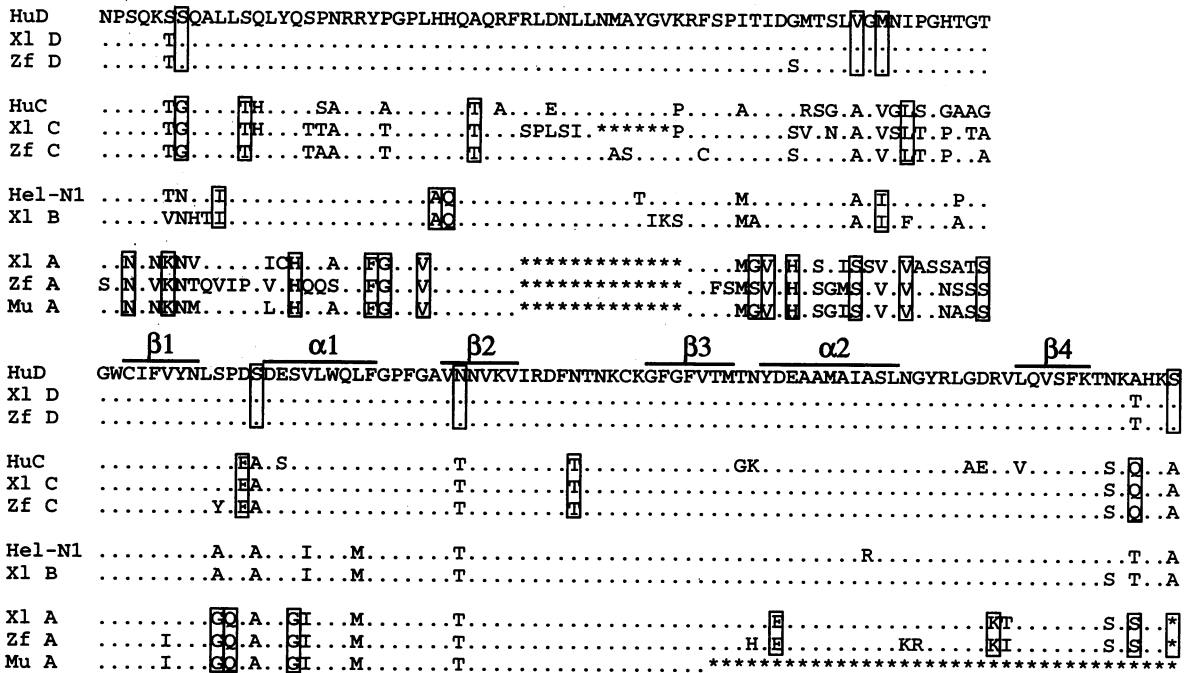
Biochemistry: Good

Proc. Natl. Acad. Sci. USA 92 (1995) 4559

```
HuD   NPSQKSSQALLSQLYQSPNRRYPGPLHHQAQRFRLDNLLNMAYGVKRFSPITIDGMTSLVGMNIPGHTGT
Xl D  .....T.....................................................
Zf D  .....T..............................................S....

HuC   .....TG....TH....SA...A......T A...E.........P....A....RSG.A.VGIS.GAAG
Xl C  .....TG....TH....TTA...T.....T...SPLSI.******P........SV.N.A.VSLT.P.TA
Zf C  .....TG....T.....TAA...T.....T..........AS.....C......S....A.V.LT.P..A

Hel-N1  .....TN..T...........AQ.............T.......M........A.T.....P..
Xl B    .....VNHTT...........AQ.................IKS....MA.......A.T.F...A..

Xl A  ..N.NKNV.....IGH..A..FG..V.......************....MGV.H.S.ISSV.VASSATS
Zf A  S.N.VKNTQVIP.V.HQQS..FG..V.......*************..FSMSV.H.SGMS.V.M..NSSS
Mu A  ..N.NKNM.....L.H..A..FG..V.......*************....MGV.H.SGIS.V.M..NASS
```

```
          β1        α1         β2          β3       α2           β4
HuD   GWCIFVYNLSPDSDESVLWQLFGPFGAVNNVKVIRDFNTNKCKGFGFVTMTNYDEAAMAIASLNGYRLGDRVLQVSFKTNKAHKS
Xl D  .............................................................................T..
Zf D  .............................................................................T..

HuC   ..........EA.S.............T.....T.............GK.................AE..V......S.Q..A
Xl C  ..........EA..............T.....T.........................................S.Q..A
Zf C  ........Y.EA..............T.....T.........................................S.Q..A

Hel-N1  .........A..A...I...M.......T.......................R................T..A
Xl B    .........A..A...I...M.......T.....................................S T..A

Xl A  ........GQ.A..GI...M.......T..............................E............KT.......S.S..*
Zf A  .....I..GQ.A..GI...M.......T...............H.E..........KR.....KI.......S.S..*
Mu A  .....I..GQ.A..GI...M.......T....************************************************
```

FIG. 2. Alignment of the tether region (*Upper*) and third RRM (*Lower*) of different Elav-like proteins. The predicted protein sequences for different cDNAs were aligned and displayed as in Fig. 1. Xl, Zf, and Mu correspond to *Xenopus*, zebrafish, and mouse sequences, respectively, with the letter indicating the gene name; previously identified proteins are labeled with their original names. The mouse sequence is from a partial cDNA fragment lacking the carboxyl terminus. The locations of putative α-helix and β-sheet motifs within the third RRM (7) are indicated by the overlined sequences and are labeled. Boxed amino acid residues are unique to a particular gene in all species examined.

elrA is an additional member of this family. The *elrA* gene is present in *Xenopus*, zebrafish, and mouse, demonstrating its presence in many vertebrate species.

The protein sequences of Elav and Rbp9, the two *Drosophila* ELGs, are about equally similar to the vertebrate ELGs. In terms of domain structure, the *Drosophila* proteins have a large amino-terminal domain characterized by stretches of alanine and glutamine for Elav and asparagine and glutamine for Rbp9 (19). For the *elav*-encoded protein, this amino-terminal domain is needed for complete function (34). The relative developmental expression patterns of *elav* and *rbp9* are similar to *elrC* and *elrD*; *elav* and *elrC* are expressed very early in nervous system development, whereas *rbp9* and *elrD* are expressed later (see below; refs. 19 and 34). The presence of two nervous system-specific ELGs expressed in a similar developmental pattern implies that their function in neurons may have been conserved in evolution from insects to vertebrates.

Conserved Sequence Elements in the 3′-Untranslated Region (3′-UTR) of *elrD* Genes. The 3′-UTR of the *Xenopus* and zebrafish *elrD* genes also shares similarity with the human HuD 3′-UTR (Fig. 3). The conserved sequences are 71–234 nt downstream of the termination codon and are unlikely to encode a protein, since out-of-frame insertion/deletions are needed to efficiently align the sequences. A conserved 50-nt element is 76% identical between the three species. For the *Xenopus* and human UTR, this conservation is expanded to 100 nt with 90% identity. The 50-nt element can be folded into

a conserved 8-bp stem with a 15-bp loop followed by a 16-nt element that is perfectly conserved between all three species.

The significance of this conserved sequence is unknown. A similar conservation in the 3′-UTR can be seen in other RBP genes such as the poly(A)-binding protein, heterogeneous nuclear ribonucleoprotein A2, nucleolin, and Y-box binding proteins (27, 35) as well as in many other genes (35). This element might be involved in the regulation of expression of the ElrD protein by the binding of RBPs. For other mRNAs, conserved sequences in the 3′-UTR regulate both translation and stability of the corresponding RNA (36, 37). An attractive possibility, given the conservation of this element and the corresponding ElrD protein, is that ElrD binds to its own mRNA and autoregulates the level of ElrD protein. This type of autoregulation by RBPs is not uncommon (5).

Expression of ELGs in Adult Tissues. I examined RNA levels for *Xenopus* ELGs by Northern blots to look for tissue-specific expression in adult frogs (Fig. 4). RNA species hybridizing to *Xenopus elrC* (5.0 and 4.5 kb) and *Xenopus elrD* (3.5 kb) probes were found only in brain. *Xenopus elrB* probes detected a brain-specific transcript of 4.5 kb, an RNA species of 4.0 kb in testes and ovary, and several transcripts <4.0 kb in ovary. Surprisingly, *Xenopus elrA* encodes RNA species of 3.0 and 1.8 kb that were present in all tissues tested. Thus, while three of the *Xenopus* genes (*elrB*, *elrC*, and *elrD*) have a brain-specific expression pattern, *elrA* is expressed throughout the adult frog.
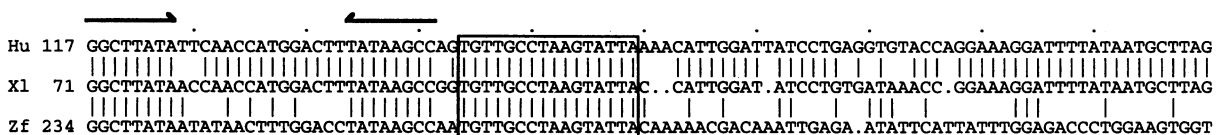
```
Hu 117 GGCTTATATTCAACCATGGACTTTATAAGCCAGTGTTGCCTAAGTATTAAAACATTGGATTATCCTGAGGTGTACCAGGAAAGGATTTTATAATGCTTAG
       |||||||| ||||||||||||||||||||| |||||||||||||||| ||||||| |||||| | ||| |||||||||||||||||||||||||
Xl 71  GGCTTATAACCAACCATGGACTTTATAAGCCGCTGTTGCCTAAGTATTAC..CATTGGAT.ATCCTGTGATAAACC.GGAAAGGATTTTATAATGCTTAG
       |||||||| | |||| |||||||||||||||| ||||||||||||||| | | || | | ||| | ||| | || | |
Zf 234 GGCTTATAATATAACTTTGGACCTATAAGCCAATGTTGCCTAAGTATTACAAAAACGACAAATTGAGA.ATATTCATTATTTGGAGACCCTGGAAGTGGT
```

FIG. 3. Lineup of a conserved element in the 3′-UTR of the vertebrate *elrD* genes. The number at the left refers to the distance from the termination codon while Hu, Xl, and Zf indicate human, *Xenopus*, and zebrafish genes, respectively. A conserved potential stem-loop is indicated by the arrows above the sequence, and a conserved sequence is boxed.
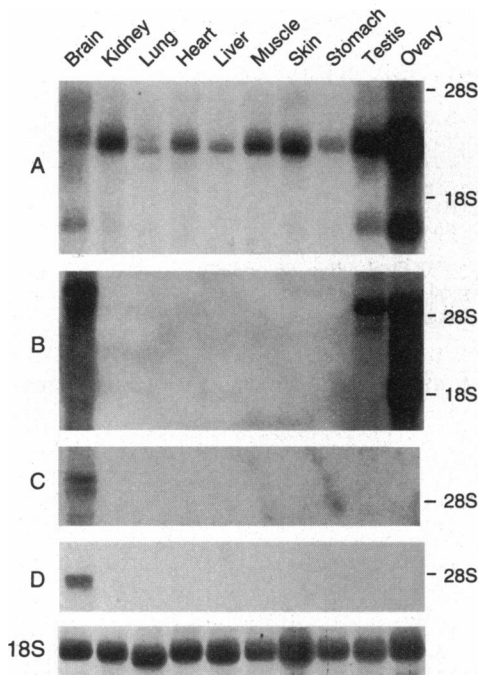
FIG. 4. Expression of *Xenopus* ELGs in different adult tissues. Northern blots with 7.5 μg of RNA from different adult tissues were hybridized with *Xenopus elav*-like cDNA probes. The probe used is indicated on the left, and the migration of 28S and 18S rRNA is indicated at the right. The bottom panel is a filter rehybridized with an oligonucleotide that detects 18S rRNA.

Developmental Pattern of ELG Expression. Northern blots with RNAs from different *Xenopus* embryonic stages were probed to examine developmental regulation of these genes (Fig. 5*A*). As expected from the ubiquitous tissue distribution of *Xenopus elrA*, this gene is expressed at all developmental stages tested; a similar expression is seen during zebrafish development (data not shown). The high RNA level in ovary relative to that in embryos is likely the consequence of the higher amount of mRNA in ovary that results from an abundance of immature oocytes. The neural-specific genes *elrC* and *elrD* are expressed after stage 11 and stage 19, respectively, corresponding to late gastrula and late neurula embryos. The homologous genes in the zebrafish are expressed at equivalent stages (data not shown).

The *elrB* gene exhibited a complex pattern of expression. The neural-specific 4.5-kb transcript is expressed late in development, by early tadpole stage (stage 44). The 4.0-kb ovary-specific transcript is not present in early embryos; this RNA species is specific to follicle cells surrounding the oocytes (data not shown). Multiple RNA species <4.0 kb are present in the ovary and persist in the early embryo at least up to gastrulation (stage 10). These RNA species are not due to degradation as judged by rehybridizing the same filter with probes either for 18S rRNA (Fig. 5*A*) or for a ribosomal protein (data not shown). To determine whether these oocyte and embryonic RNA species encode the ElrB protein, these same RNA samples were analyzed by reverse transcription–PCR with primers that flank the coding region (Fig. 5*B*). These reactions showed that at least some of the maternal transcripts can encode the full-length protein.

Possible Functions of the ELGs. The nervous system-specific ELGs are probably involved in the regulation of neuronal gene expression. In *Drosophila*, genetic experiments suggest that *elav* is required for correct differentiation and maintenance of the neuronal phenotype. Both Elav and Rbp9 are expressed exclusively in neurons, although at different developmental stages. The vertebrate nervous system-specific
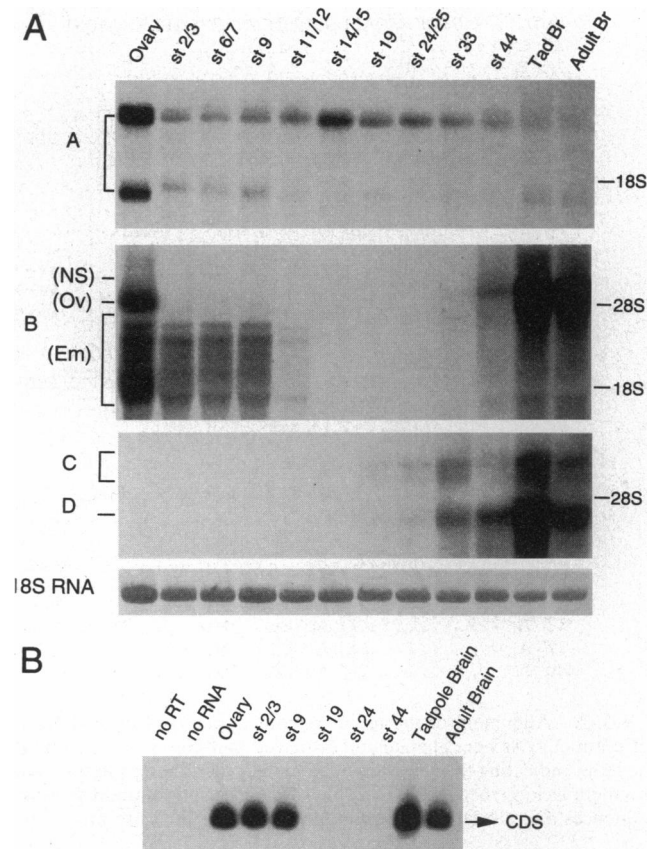

FIG. 5. Expression of *Xenopus* ELGs during embryonic development. (*A*) Total RNA (7.5 μg) from staged embryos was analyzed by Northern blots probed with *Xenopus elav*-like cDNA probes. The lanes are labeled with the embryo stage (st) or tissue. The probe used is identified on the left, and the migration of 28S and 18S rRNA is indicated on the right. The different RNA species that hybridized to the *elrB* probe are labeled NS (nervous system specific), Ov (ovary specific), and Em (embryo specific). The third panel is a blot probed with both *elrC* and *elrD* probes. The migration of the respective RNA species is indicated. The bottom panel is a filter rehybridized with an oligonucleotide that detects 18S rRNA. (*B*) The coding region of *elrB* was amplified by reverse transcription–PCR from the indicated RNA samples and detected on a Southern blot probed with an *elrB* probe. The lanes are labeled with the source RNA; no RT and no RNA are reactions containing brain RNA with no reverse transcriptase and no input RNA, respectively. The arrow labeled CDS indicates the migration of full-length coding region amplified with the same primers from plasmid DNA. Tad Br, tadpole brain; Adult Br, adult brain.

ELGs also are expressed in most neurons (22, 23, 25). In an early embryo, a monoclonal antibody that recognizes both the ElrC and ElrD proteins (16A11) is a very early marker of the neuronal phenotype (24, 25). This early neuronal expression of ELGs is consistent with the expression of *elrC* in *Xenopus* where RNA is detected after the late gastrula stage (see Fig. 5), when the first neurons are differentiating (38).

The presence of the RRMs in the ELGs suggests that the corresponding proteins bind target RNAs to form ribonucleoprotein particles. The arrangement of RRMs suggests the presence of two RNA-binding domains, where the first and second RRMs combine to form a single binding domain, with the third RRM being the second binding domain. Using a random RNA selection technique, Levine *et al.* (39) identified a putative RNA target sequence that binds preferentially to the third RRM of the human ElrB protein (Hel-N1). Given the high conservation of the third RRM between members of this family, other Elav-like proteins would also be expected to bind this RNA target. This target sequence is present in the 3'-UTR of several mRNAs involved in regulating growth, including

mRNAs encoding c-Myc, c-Fos, and the Id repressor (20, 39), and has been implicated in regulating mRNA stability (37). On this basis, Levine *et al.* (39) proposed that the Elav-like proteins bind to the 3'-UTR of these growth-regulating genes, affect the stability or translation of the corresponding mRNAs, and thereby regulate growth. Similar models for the function of Elav and the ELGs in vertebrates have been proposed by others (12, 25).

Identifying the RNA processing event that may be regulated by Elav-like proteins will require determination of the localization of the proteins in the cell. Antisera directed against both Elav and Rbp9 stain the nucleus (14, 16); thus, these proteins may regulate events in the nucleus such as alternative splicing, polyadenylylation, and nuclear transport. In contrast to these results, the vertebrate Elav-like proteins recognized by autoimmune sera and monoclonal antibody 16A11 are seen in both the nucleus and the cytoplasm (21–25). These results could signify the presence of different cross-reacting isoforms in the nucleus and cytoplasm, each of which could have a distinct function. Alternatively, individual Elav-like proteins could be present in, and possibly shuttle between, nuclear and cytoplasmic compartments.

Multiple Functions of the *elrB* Gene. *elrB* is expressed in a complex temporal pattern of RNAs of different sizes in the nervous system, reproductive organs, and early embryo. Thus, ElrB protein may have a function in germ cell formation and early development in addition to its neuronal function. During early development, much of the regulation of gene expression is posttranscriptional (40, 41). Given that the human homolog of *elrB*, Hel-N1, encodes a protein that binds to uridylate-rich sequences (20, 39), it is possible that the ElrB protein may be involved in binding to uridylate-rich sequences that have been implicated in regulating polyadenylylation and translation in *Xenopus* oocytes and early embryos (reviewed in ref. 42).

A Widely Expressed Member of the ELG Family. The *elrA* gene appears to be unique to vertebrates, since no ubiquitously expressed ELG has been described in *Drosophila* despite an extensive search for RRM-containing genes (18). The ElrA protein is the most divergent family member, possibly implying that during evolution this gene arose as a duplication of one of the nervous system-specific ELGs and later evolved a new function that became required in all cells. The divergence in sequence suggests that the *elrA* product and the other Elav-like proteins bind to different target RNA sequences and thus regulate different genes. In *Xenopus*, a similar relationship is seen for another RBP family, the nrp family. *nrp1* encodes a neural-specific RBP that is expressed in proliferating regions of the nervous system (33). In a search for related genes, Good *et al.* (27) identified the *xrp1* gene, which shares a high degree of sequence similarity yet is expressed in all adult tissues and developmental stages. The evolution of ubiquitously expressed and nervous system-specific RBPs may reflect the inherent complexity of the nervous system and the requirement for alternative means of gene regulation.

1. Kenan, D. J., Query, C. C. & Keene, J. D. (1991) *Trends Biochem. Sci.* **16**, 214–220.
2. Burd, C. G. & Dreyfuss, G. (1994) *Science* **265**, 615–621.
3. Bellen, H. J., Kooyer, S., D'Evelyn, D. & Pearlman, J. (1992) *Genes Dev.* **6**, 2125–2136.
4. Wolfner, M. F. (1988) *Trends Genet.* **4**, 333–337.
5. Mattox, W., Ryner, L. & Baker, B. S. (1992) *J. Biol. Chem.* **267**, 19023–19026.
6. Landsman, D. (1992) *Nucleic Acids Res.* **20**, 2861–2864.
7. Birney, E., Kumar, S. & Krainer, A. R. (1993) *Nucleic Acids Res.* **21**, 5803–5816.
8. Bandziulis, R. J., Swanson, M. S. & Dreyfuss, G. (1989) *Genes Dev.* **3**, 431–437.
9. Campos, A. R., Grossman, D. & White, K. (1985) *J. Neurogenet.* **2**, 197–218.
10. Jimenez, F. & Campos-Ortega, J. A. (1987) *J. Neurogenet.* **4**, 179–200.
11. Homyk, T., Jr., Szidonya, J. & Suzuki, D. T. (1980) *Mol. Gen. Genet.* **177**, 553–565.
12. Yao, K.-M., Samson, M.-L., Reeves, R. & White, K. (1993) *J. Neurobiol.* **24**, 723–739.
13. Homyk, T., Jr., Isono, K. & Pak, W. L. (1985) *J. Neurogenet.* **2**, 309–324.
14. Robinow, S. & White, K. (1991) *J. Neurobiol.* **22**, 443–461.
15. Robinow, S. & White, K. (1988) *Dev. Biol.* **126**, 294–303.
16. Bier, E., Ackerman, L., Barbel, S., Jan, L. & Jan, Y. N. (1988) *Science* **240**, 913–916.
17. Robinow, S., Campos, A. R., Yao, K.-M. & White, K. (1988) *Science* **242**, 1570–1572.
18. Kim, Y.-J. & Baker, B. S. (1993) *Mol. Cell. Biol.* **13**, 174–183.
19. Kim, Y.-J. & Baker, B. S. (1993) *J. Neurosci.* **13**, 1045–1056.
20. King, P. H., Levine, T. D., Fremeau, R. T. & Keene, J. D. (1994) *J. Neurosci.* **14**, 1943–1952.
21. Sakai, K., Gofuku, M., Kitagawa, Y., Ogasawara, T., Hirose, G., Yamazaki, M., Koh, C.-S., Yanagisawa, N. & Steinman, L. (1994) *Biochem. Biophys. Res. Commun.* **199**, 1200–1208.
22. Szabo, A., Dalmau, J., Manley, G., Rosenfeld, M., Wong, E., Henson, J., Posner, J. B. & Furneaux, H. M. (1991) *Cell* **67**, 325–333.
23. Dalmau, J., Furneaux, H. M., Cordon-Cardo, C. & Posner, J. B. (1992) *Am. J. Pathol.* **141**, 881–886.
24. Marusich, M. F. & Weston, J. A. (1992) *Dev. Biol.* **149**, 295–306.
25. Marusich, M. F., Furneaux, H. M., Henion, P. D. & Weston, J. A. (1994) *J. Neurobiol.* **25**, 143–155.
26. Taira, M., Jamrich, M., Good, P. J. & Dawid, I. B. (1992) *Genes Dev.* **6**, 356–366.
27. Good, P. J., Rebbert, M. L. & Dawid, I. B. (1993) *Nucleic Acids Res.* **21**, 999–1006.
28. Richter, K., Grunz, H. & Dawid, I. B. (1988) *Proc. Natl. Acad. Sci. USA* **85**, 8086–8090.
29. Hemmati-Brivanlou, A., de la Torré, J. R., Holt, C. & Harland, R. M. (1991) *Development (Cambridge, U.K.)* **111**, 715–724.
30. Graf, J.-D. & Kobel, H. R. (1991) *Methods Cell Biol.* **36**, 19–34.
31. Devereux, J., Haeberli, P. & Smithies, O. (1984) *Nucleic Acids Res.* **12**, 387–395.
32. Good, P. J., Richter, K. & Dawid, I. B. (1990) *Proc. Natl. Acad. Sci. USA* **87**, 9088–9092.
33. Richter, K., Good, P. J. & Dawid, I. B. (1990) *New Biol.* **2**, 556–565.
34. Yao, K.-M., Samson, M. L., Reeves, R. & White, K. (1993) *J. Neurobiol.* **24**, 723–739.
35. Duret, L., Dorkeld, F. & Gautier, C. (1993) *Nucleic Acids Res.* **21**, 2315–2322.
36. Sonenberg, N. (1994) *Curr. Opin. Genet. Dev.* **4**, 310–315.
37. Sachs, A. B. (1993) *Cell* **74**, 413–421.
38. Lamborghini, J. E. (1980) *J. Comp. Neurol.* **189**, 323–333.
39. Levine, T. D., Gao, F., King, P. H., Andrews, L. G. & Keene, J. D. (1993) *Mol. Cell. Biol.* **13**, 3494–3504.
40. Dworkin, M. B. & Dworkin-Rastl, E. (1990) *Mol. Reprod. Dev.* **26**, 261–297.
41. Standart, N. (1992) *Semin. Dev. Biol.* **3**, 367–379.
42. Wickens, M. (1992) *Semin. Dev. Biol.* **3**, 399–412.