

Tissue Specific Expression of the Retinoic Acid Receptor- β 2: Regulation by Short Open Reading Frames in the 5'-Noncoding Region

Andreas Zimmer, Anne M. Zimmer, and Kay Reynolds

Unit on Developmental Biology, Laboratory of Cell Biology, National Institute of Mental Health, Bethesda, Maryland 20892

Abstract. The 40-S subunit of eukaryotic ribosomes binds to the capped 5'-end of mRNA and scans for the first AUG in a favorable sequence context to initiate translation. Most eukaryotic mRNAs therefore have a short 5'-untranslated region (5'-UTR) and no AUGs upstream of the translational start site; features that seem to assure efficient translation. However, ~5–10% of all eukaryotic mRNAs, particularly those encoding for regulatory proteins, have complex leader sequences that seem to compromise translational initiation. The retinoic-acid-receptor- β 2 (RAR β 2) mRNA is such a transcript with a long (461 nucleotides) 5'-UTR that contains five, partially overlapping, upstream open reading frames (uORFs) that precede the major ORF. We have begun to investigate the function of this complex 5'-UTR in transgenic mice, by introducing muta-

tions in the start/stop codons of the uORFs in RAR β 2-lacZ reporter constructs. When we compared the expression patterns of mutant and wild-type constructs we found that these mutations affected expression of the downstream RAR β 2-ORF, resulting in an altered regulation of RAR β 2-lacZ expression in heart and brain. Other tissues were unaffected. RNA analysis of adult tissues demonstrated that the uORFs act at the level of translation; adult brains and hearts of transgenic mice carrying a construct with either the wild-type or a mutant UTR, had the same levels of mRNA, but only the mutant produced protein. Our study outlines an unexpected role for uORFs: control of tissue-specific and developmentally regulated gene expression.

IT is becoming increasingly clear that the 5' and 3' untranslated regions (UTRs)¹ of many eukaryotic mRNAs play an important role in the regulation of gene expression by modulating mRNA stability and translational efficiency. In fact, the expression of many genes is primarily regulated at the posttranscriptional level (for review see Jackson, 1993).

The initiation of protein synthesis is often the rate limiting step and determines how much protein is synthesized from one mRNA. According to the scanning model for translational initiation, the 40 S preinitiation complex binds to the 5'-cap site of the mRNA, and then migrates until it encounters the first AUG in a favorable context to initiate translation. However, in many eukaryotic genes, the first AUG is not the translational start site of the major open reading frame. In particular mRNAs that encode for proteins with a proposed function in cell growth and differentiation, often have one or more AUGs or small upstream open reading

frames (uORFs) that precede the major ORF (Kozak, 1991). These AUGs or uORFs usually inhibit translation from downstream, although some cases have been described where uORFs stimulated translation of the major ORF.

The retinoic acid receptor- β (RAR β) gene is one of those genes. It encodes a ligand-activated transcription factor and belongs to the superfamily of nuclear hormone receptors (Evans, 1988). Retinoic acid (RA) plays a critical role during embryonic development, homeostasis, and cell differentiation (for review see Eichele, 1993). Retinoic acid receptors act as molecular switches to transform the RA signal into a transcriptional response.

The RAR β gene encodes at least four different transcripts which are derived by differential splicing and promoter usage (Zelent et al., 1991; van der Leede et al., 1992; Nagpal et al., 1992). These transcripts contain common DNA and hormone-binding domains, but differ in the 5'-UTR and the NH₂-terminal part of the receptor. The RAR β 2 isoform has an unusually long 5'-UTR (461 nucleotides) that can form stable secondary structures and contains five small uORFs preceding the major ORF (de The et al., 1987; Reynolds et al., 1991; Zelent et al., 1991). Interestingly, uORF2, uORF4, and uORF5 have a staggered organization of overlapping start and stop codons. Similarly overlapping uORFs can be found in other regulatory genes, like homeobox genes. It has been suggested that such an arrangement might be par-

Address all correspondence to Dr. Andreas Zimmer, Unit on Developmental Biology, Laboratory of Cell Biology, National Institute of Mental Health, Bldg. 36/3A17, Bethesda, MD 20892. Telephone: (301) 496-7377; FAX: (301) 402-1748; e-mail: andreas@helix.nih.gov.

1. *Abbreviations used in this paper:* eIF-2, eukaryotic initiation factor 2; ORF, open reading frame; RA, retinoic acid; RAR β , retinoic acid receptor- β ; uORF, upstream open reading frame; UTR, untranslated region.

ticularly suitable for translational regulation (Kessel and Gruss, 1988).

We (Reynolds et al., 1991; Zimmer and Zimmer, 1992) and others (Mendelsohn et al., 1991; Shen et al., 1992) have used RAR β 2-lacZ reporter constructs in transgenic mice, to study the regulation of the RAR β 2 gene in normal embryos and after RA treatment. Our constructs were made by fusing the lacZ coding region to the NH₂-terminus of the RAR β 2 gene, leaving the 5' UTR intact. The expression pattern of the reporter gene constructs was tightly regulated during embryogenesis and closely followed that of the endogenous gene in most tissues. This was surprising, because it indicated that the 5'-UTR did not inhibit translation, as one might have expected. However, close examination of the published reports revealed some striking differences. For example, Mendelsohn et al. (1991) and Shen et al. (1992) both noted lacZ expression in the heart wall, which we never observed in our transgenics. Mendelsohn et al. (1991) and Shen et al. (1992) used constructs with a truncated 5'-UTR, thus indicating that this region might have indeed an important regulatory function that is not evident in all tissues.

To address the role of the uORFs, we generated RAR β 2-lacZ reporter constructs with mutations in all upstream AUGs or with mutations in overlapping start and stop codons. The expression patterns in transgenic mice of the mutant constructs were compared with a construct carrying the wild-type 5'-UTR (Reynolds et al., 1991; Zimmer and Zimmer, 1992). We found that expression of the mutant constructs was differently regulated in heart and brain, suggesting that the uORFs play an important role in regulating tissue specific expression of the RAR β 2 mRNA.

Materials and Methods

Constructs

Mutations in the RAR β 2-UTR were introduced using an oligonucleotide directed mutagenesis system (Amersham Corp., Arlington Heights, IL) according to the manufacturer's recommendations. The template for mutagenesis was a PstI fragment from the wild-type reporter construct RLZ79 (Reynolds et al., 1991) inserted into M13mpl8. The following oligonucleotides were used for site directed mutagenesis: M24-CCCGA ATCTA GAATT ATGAC; M45-ACGAA CTCCA ACAA CTCTC. M0: uORF1-GTCCT CTGCA ACCCA GTCTT; uORF2-TCATT TACCA ATTC CAGGC; uORF3-CCAAA TGATC AATTA CCATT; uORF4-CCCGA ATCAA GAATT ATGAC; uORF5-ACGAA CTCCT TCAAA CTCTC. Mutations were verified by di-deoxy sequencing and the mutated PstI fragment was reinserted into RLZ79 after partial PstI digestion.

Generation and Analysis of Transgenic Mice

Transgenic mice were established and embryos were analyzed for β -gal expression as previously described (Reynolds et al., 1991; Eisel et al., 1993). Briefly, transgenic embryos were dissected from euthanized mothers and fixed for 30–120 min at 4°C in fixative (1% formaldehyde, 0.1% glutaraldehyde, 0.02% NP-40 in PBS). After fixation embryos were washed twice in PBS for 30 min and incubated in staining solution (1 mg/ml 5-Bromo-4-Chloro-3-Indolyl- β -D-Galactopyranoside, 5 mM K₃Fe(CN)₆, 5 mM K₄Fe(CN)₆, 2 mM MgCl₂) at 30°C overnight. For whole mount stainings of organs from adult animals all solutions were supplemented with NP-40 to a final concentration of 0.2%. For the preparation of cryosections, the embryos were soaked in a cryoprotectant solution (3% polyethyleneglycol, 20% sucrose in 0.1 M PBS pH 7.4) overnight. Subsequently, they were quickly frozen on dry ice and 20- μ m thick sections were cut in a Reichert 2800 Frigocut cryostat. All sections were counterstained with 0.1% neutral red. To analyze lacZ staining in deeper brain structures, adult brains were fresh frozen, sectioned, fixed for 15 min at 4°C in fixative, washed twice in PBS, and incubated in staining solution at 30°C overnight.

Pictures of histological sections and transgenic embryos were taken with a Zeiss Axiophot or Nikon stereomicroscope. Color images were scanned on Kodak Photo CD, processed with Adobe Photoshop, and printed on a Tektronics Phaser IISDX.

LacZ activity was measured in embryo or tissue extracts essentially as described (König et al., 1991). Briefly, tissues were homogenized in PM2 buffer (100 mM Na-phosphate, 2 mM MgSO₄, 0.1 mM MnCl₂, 0.5% (vol/vol) β -mercaptoethanol, pH 8.0). We used 500 μ l each embryo. Cell debris was pelleted by centrifugation in a microcentrifuge and 100 μ l of the supernatant was transferred either directly, or after dilution into a 96-well microtiter plate. The reaction was started by adding 25 μ l of Chlorophenol red- β -D-galactopyranoside solution (5 mg/ml in PM2 buffer) and measured in a microplate reader (OD at 570 nm).

RNAse Protection Assays

RNA from adult organs was purified by homogenization in guanidinium thiocyanate and centrifugation through a CsTFA gradient (Okayama et al., 1987). Quantitative RNA protection assays were performed essentially as described (Sambrook et al., 1989). 10 μ g of total RNA were hybridized with 1.5×10^5 cpm of probe in hybridization buffer (40 mM Pipes, pH 6.4; 1 mM EDTA; 0.4 M NaCl; 80% formamide) at 50°C for 16 h. Subsequently, 300 μ l RNase digestion mixture (10 mM Tris, pH 7.4; 300 mM NaCl; 5 mM EDTA; 2 μ g/ml RNase T1; 40 μ g/ml RNase A), was added and incubated at 37°C for 60 min. The protected RNA was precipitated after inactivation of RNAases by proteinase K digestion and phenol/chloroform extractions, dissolved in formamide loading buffer and separated on a 6% sequencing gel (Sequagel). Subsequently, the gels were dried and analyzed using a Bio-Imaging Analyzer (Fuji BAS 2000). Expression levels were quantitated using the "Radioaktive" function and also independently after importing the raw data into a Macintosh Quadra 800 computer using the public domain NIH Image program (written by Wayne Rasband at the U.S. National Institutes of Health and available from the Internet by anonymous ftp from zippy.nimh.nih.gov or on floppy disk from National Technical Information Service, 5285 Port Royal Rd., Springfield, VA 22161, part number PB93-504868).

Results

The 5'-UTR Is Highly Conserved between Mouse and Human

The short upstream open reading frames in the RAR β 2 5'-UTR are highly conserved between mouse and human (Fig. 1; de The et al., 1987; Reynolds et al., 1991; Zelent et al., 1991), which indicates that this region has an important function. Particularly striking is the overlapping arrangement of uORF2, uORF4, and uORF5. The stop codon of uORF2 overlaps with the uORF4 start codon, and the uORF4 stop codon overlaps with the AUG of uORF5 (Fig. 1). The organization of these uORFs is almost identical in mouse and humans. However, because of a frameshift mutation in uORF3/uORF5, the amino acid sequence of the human and mouse uORF5 peptides and the COOH-terminal half of the uORF3 peptides would be different.

Mutations in Overlapping Start and Stop Codons Results in Altered Regulation of LacZ Expression in Heart and Brain

The reporter constructs used in this study were derived from the RAR β 2-lacZ expression vector RLZ79 (Reynolds et al., 1991). This construct contained about 1.25 kb of RAR β 2 genomic sequences, including the RAR β 2 promoter, the RAR β 2 5'-UTR and the first 29 amino acids of the RAR β 2-ORF, to which the lacZ coding region was fused (Fig. 2). Consequently, translation of the lacZ gene was subject to potential control mechanisms located within the 5'-UTR. We analyzed the expression pattern in transgenic founder embryos and transgenic lines. Transgenic founder embryos

human gtcacagaagtagtaggaagtgtgact
mouse GTGACAGAAGTGGTAGGAAGTGAGCT

gttcagaggcaggagggtctattctttgcaaaa.ggggggaccagaattccccatgctgag
 GCTCCGAGGCAGGAGGGTCTATTCTTTGCCAAAAGGGGGGACCAGAGTCCCGTGCGCC

uORF1
 M P R T R A I R A G F V W A P
 ctgtttgaggactgggatgcccagagaacgcgagcgcgatccgagcagggttctctgggcacc
 CGCCACAAAGACTGGGATGCAGAGGACGCGAGCCACCCGGGCAGGGAGCGCTGGGCACC
 M Q R T R A T R A G S V W A P
 S G *
 gtcgggtaggtagccggaacccattcgggaaggctttttgcaagcatttacttgaaggag
 GCGGGGTAGGA.CCCGCGCTCCCGAGGCCGTCGCG...GCGTGCCTGGAAGGGAG
 . A G *

aaacttgggatctttctgggaacccccgccccggctggttggccgagcaagcctggaaa
 AACTTGGGATCGTGCGGGAA.CCCCGCCCTGCGTGGATCGGCCGAGCGAGCCTGGAAA

uORF2 uORF3
 M I I W I N Y R L L A G L S V I I H
 M V N D H L D Q L Q A F S W L V C H N S
 atggtaaatgatcatttggatcaattacaggcttttagctggctgtctgtcataatca
 ATGGTAAATGATCATTGGATCAATTACAGGCTTTAGCTGGCTTGTCTGTCAATAATCA
 M V N D H L D Q L Q A F S W L V C H N S
 M I I W I N Y R L L A G L S V I I H

uORF4 **uORF5**
 M I R G W E K D Q Q P T C Q K R G R V *
 D S G L G K R P T A Y V P K K G Q S L M
 *
 tgcctggggctgggaaaaagacacacagcctacgtgccccaaaaagggcagagtttgat
 TGATTCGGGGCTGGGAAAAAGACCAACAGCCTACGTGCCAAAAAAGGGCAGAGTTTGAT
 *
 D S G L G K R P T A Y V P K K G Q S L M
 M I R G W E K D Q Q P T C Q K R G R V *

E L G G L F Y A I C L H T *
 ggagttgggtggactttctatgcccatttgcctccacacctagaggataagcacttttgc
 GGAGTT.CGTGGACTTTTCTGTGGCGCTCGCCTCCACACCTAGAGGATAAGCACTTTTGC
E F V D F S V R L A S T P R G *

agacattcagtgcaagggagatcatgtttgactgtatg
 AGA.GCGCGGTGC.GGAGAGATCATGTTTACTGTATG

Figure 1. Comparison of the human and mouse RAR β 5'-UTRs. Shown are the sequences from the transcriptional start site (de The et al., 1990) to the RAR β -AUG and the conceptual translation of all five uORFs. The amino acid sequence of uORF5 is indicated in bold letters, because uORF3 and uORF5 are in the same reading frame.

were particularly informative. We routinely obtained between 10% and 30% transgenic embryos from every litter, enabling us to examine several independently derived founder animals within a single experiment. Transgenic lines were most useful for following the expression patterns during development.

The first mutation tested, M24, had a two nucleotide exchange in the uORF2 stop codon and the uORF4 start codon (Fig. 2), thus extending uORF2 (162 bp vs 63 bp) and deleting uORF4. Four transgenic lines that expressed the transgene were derived with this construct. The expression pattern was analyzed in embryos between day 9.5 and 12.5 post coitus (p.c.), and in adults. The expression patterns of the wild-type construct (RLZ79) and M24 were very similar in most tissues throughout development (Fig. 3). However, all M24 lines expressed lacZ at high levels in heart and three lines also expressed the transgene in the telencephalon. Of two transgenic lines and 22 founder embryos derived with RLZ79, none expressed in the heart, and only one expressed at detectable levels in the brain. LacZ expression in the telencephalon of this embryo was much weaker than typi-

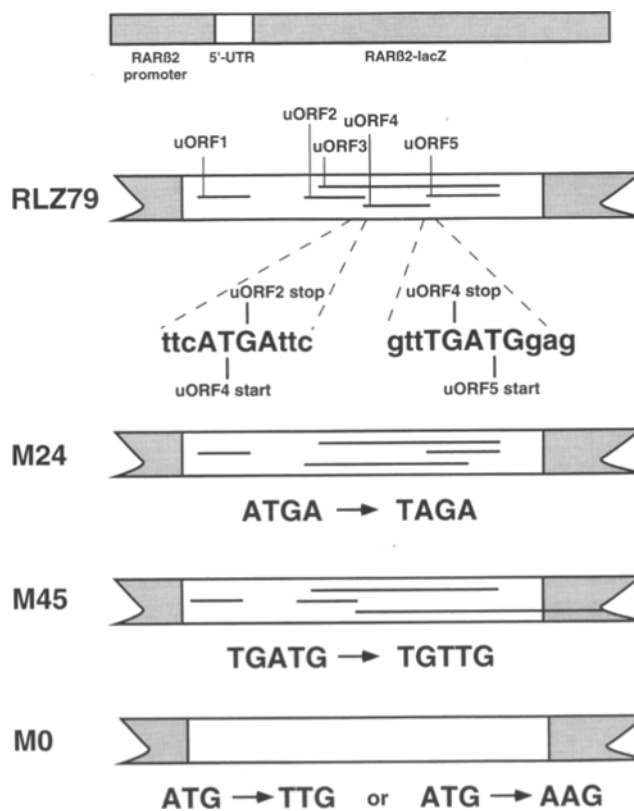


Figure 2. Schematic representation of reporter plasmids employed in this study. The RAR β 2-lacZ expression vector RLZ79 (Reynolds et al., 1991) contained \sim 1.25 kb of RAR β genomic sequences, including the RAR β promoter, the RAR β 5'-UTR and the first 29 amino acids of the RAR β -ORF, to which the lacZ coding region was fused. Translation of the lacZ gene was subject to potential control mechanisms located within the 5'-UTR. Open reading frames in the 5'-UTR are indicated as lines. The exchange of two nucleotides in mutation M24 eliminates the uORF2 stop codon and the uORF4 start codon. As a consequence, uORF2 terminates further downstream, and uORF5 is deleted. The single nucleotide exchange in M45, mutates the uORF4 stop codon and the uORF5 start codon. As uORF4 and the major ORF are in the same reading frame, the possibility exists that a larger fusion protein of uORF4 and RAR β 2-lacZ is made. In construct M0, the AUGs of uORF1 to uORF4 are mutated by changing the adenine to a thymidine, while in uORF5 the thymidine was mutated to an adenine.

cally observed in M24 embryos. Thus, the mutation in M24 alters the regulation of expression of the RAR β 2-lacZ gene in heart and telencephalon.

Two of the four lines also expressed at high levels in the peritoneum of day 12.5 embryos (Fig. 3 E). However, we believe that this high level of expression can not be attributed to the mutation, because we also found lacZ expression in the peritoneum in \sim 35% of all RLZ79 embryos analyzed.

The second construct, M45, had a single nucleotide exchange to mutate the uORF4 stop codon and the overlapping uORF5 start codon (Fig. 2). Because uORF4 and the major RAR β 2-lacZ ORF are in the same reading frame, this mutation could result in the production of a larger read-through fusion protein. With this construct we analyzed one transgenic line and the litters from three independently derived transgenic founder mothers. M45 embryos expressed the reporter gene at very high levels in the heart at all develop-

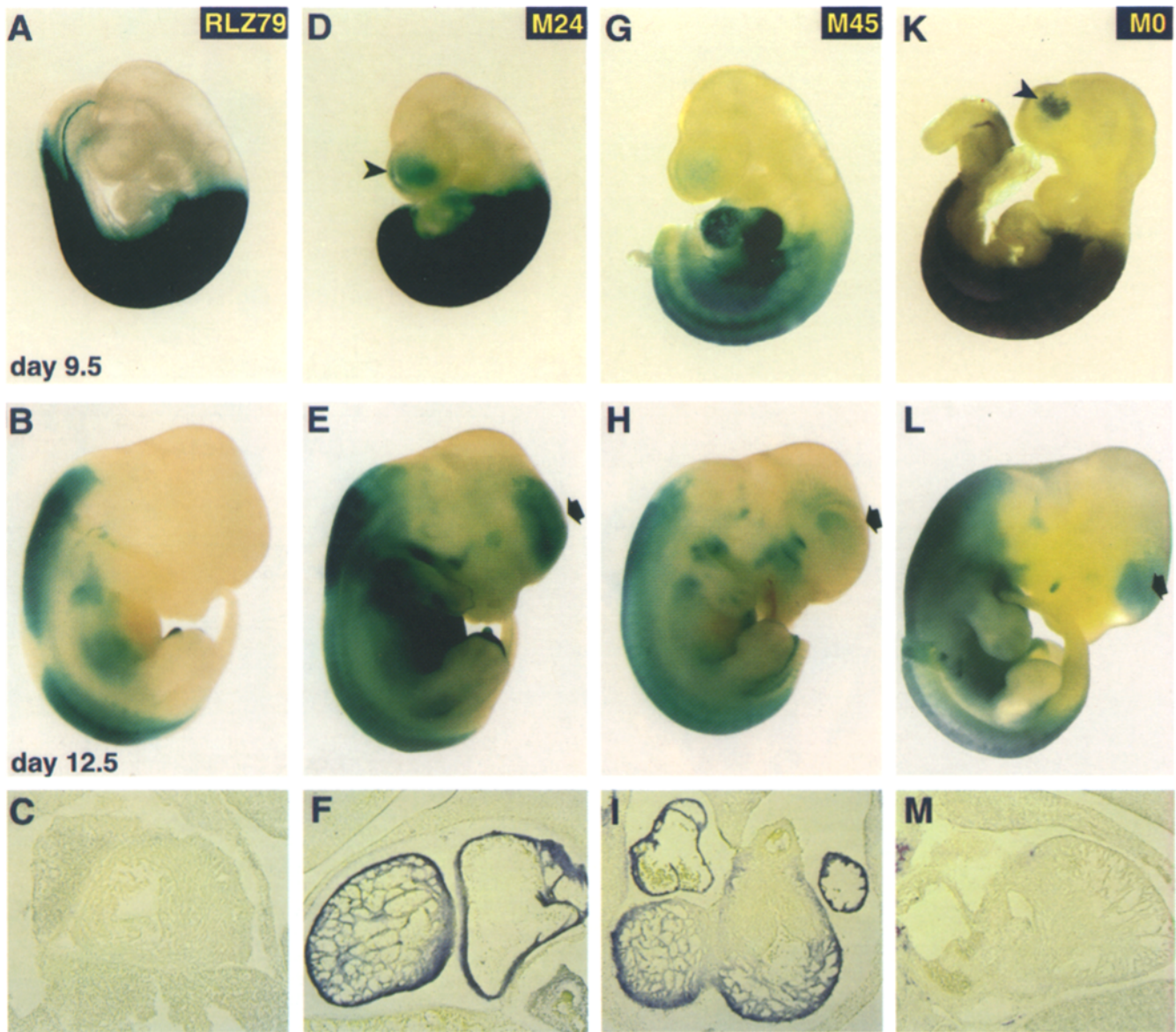


Figure 3. Mutations in the uORFs result in altered regulation of lacZ expression in the heart and brain of transgenic embryos. Whole mount staining of representative embryos at day 9.5 p.c. and 12.5 p.c. are shown in the upper panel (A, D, G, and K) and the middle panel (B, E, H, and L), respectively. Note lacZ expression in the telencephalon (arrows) of embryos derived with M24, M45, or M0. M45 embryos also expressed at high levels in the hearts of day 9.5 embryos (asterisk) while M24 expressed only weakly at this stage. Sagittal (C, F, and M) and transverse (I) sections show high levels of lacZ expression in the heart wall and also in papillary muscles of day 12.5 p.c. M24 and M45 embryos. M0 expressed only in a few cells (M).

mental stages and also at low levels in the brain. The M45 expression pattern in other tissues was similar to the wild-type construct, although staining in the spinal cord of some M45 embryos was weaker (Fig. 3). These results suggest that the uORFs inhibit expression of the RAR β 2-ORF in a tissue specific manner in heart and brain.

β -gal staining in the heart of M24 and M45 embryos was particularly strong in the heart wall and weaker in papillary muscle fibers (Fig. 3, F and I). This distribution is not an artifact caused by incomplete penetration of the staining solution, as we found a similar pattern when the staining was performed with sections (data not shown).

Both mutations exhibited a distinct temporal expression pattern in heart and brain during development and also in

adults (see Table I). For example, M24 embryos expressed lacZ in the heart only at moderate levels at day 9.5 p.c., when compared to M45 embryos. However, at day 12.5 p.c. and in adults, lacZ staining in the heart was very similar with both constructs. In the brain, β -gal activity of M24 embryos and adult mice was very high throughout all developmental stages, while M45 transgenics showed much weaker lacZ staining.

Deletion of All uORFs

The mutations in M24 and M45 suggest that the altered regulation of lacZ expression is tissue specific. However, the possibility exists that the presence of uORFs that were not af-

Table 1. LacZ Expression in Heart and Brain of Transgenic Lines and Founder Embryos

Construct	TG lines	FO embs	9.5		10.5		11.5		12.5	
			H	B	H	B	H	B	H	B
RLZ79	2	22	-	-	-	-	-	-	-	*
M0	1	10	+	++	+	++	nd	nd	+	++
M24	4†	0	+	++	++	+	++	++	+++	+++
M45	4	0	+++	+	+++	-	+++	+	+++	+

Expression of 5'-UTR mutants in the heart (H) or the brain (B) of transgenic mice. The number of transgenic lines or transgenic founder embryos that were analyzed for each construct is indicated. -, no expression; +, weak expression; ++, strong expression; +++, very strong expression; nd, not done.

* Of 22 founder embryos and two transgenic lines analyzed, only one expressed in the telencephalon.

† We examined one transgenic line throughout embryonic development and in adults, and the litters from three independently derived transgenic founder females on days 10.5 or 12.5 p.c., respectively.

ected by these mutations may prevent translation of the major ORF in other tissues. It was therefore important to investigate the expression pattern of a reporter construct in which all uORFs were mutated. This was achieved by introducing single nucleotide exchanges into all upstream AUGs.

One transgenic line and 10 founder embryos were gener-

ated with M0 and analyzed at different developmental stages. Representative day 9.5 and day 12.5 embryos are shown in Fig. 3, K and L. The overall pattern of expression was very similar in RLZ79 and M0 embryos. However, M0 founders exhibited, as expected, lacZ activity in the heart and/or brain (Fig. 3, K-M), although expression in the heart was weaker

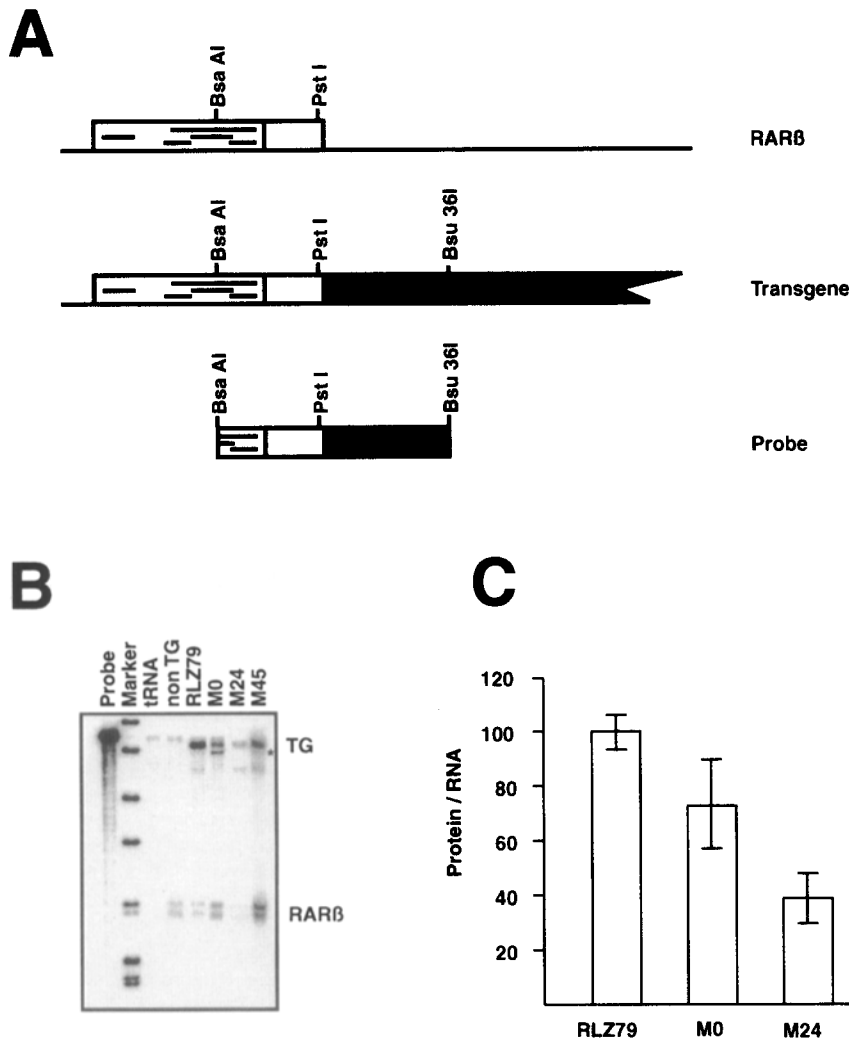


Figure 4. Determination of RARβ2-lacZ RNA and protein levels in day 12.5 embryos. (A) Organization of the endogenous RARβ2 and the transgene. A BsaAI-Bsu36I fragment of the transgene was used to synthesize an antisense RNA probe. The endogenous RARβ mRNAs protect a 187-nt fragment (BsaAI to PstI). The transgene protects a 423-nt fragment (BsaAI to Bsu36I) of the RNA probe. (B) Total RNA from various transgenic embryos was hybridized to the radiolabeled RNA probe and digested with RNase as described in Materials and Methods. The protected fragments were separated on a sequencing gel (Sequagel) and analyzed using a Bio-Imaging Analyzer. Radiolabeled bluescriptSK+ DNA digested with HpaII was used as size-marker. Note that the probe overlaps with the AUG of uORF5. This AUG is mutated by a single nucleotide exchange in constructs M0. As a consequence RNA from M0 and M45 embryos is partially cleaved at this position to yield a smaller fragment (indicated by asterisk). (C) Quantitation of RARβ2-lacZ RNA and protein levels. Protein levels were determined enzymatically in at least three embryos (see Materials and Meth-

ods). RNA levels were estimated using a Bio-Imaging Analyzer (Fuji BAS 2000) and also independently after importing the raw data into a Macintosh Quadra 800 computer using the public domain NIH Image program.

was weaker compared to M24 and M45 embryos. These data support our hypothesis that the translational regulation by the uORFs is tissue specific.

It has previously been shown that uORFs can inhibit the translation of transfected constructs in cultured cells. To determine if the translation of the RLZ79 construct is also inhibited in our transgenic lines, we wanted to compare the amount of protein made from the RLZ79 and the mutant constructs quantitatively. Therefore, day 12.5 embryos were isolated and the lacZ-mRNA and protein levels were determined by RNase protection and enzymatic assays, respectively. To facilitate the quantitation of RNA levels, we used a probe that simultaneously protected the endogenous RAR β mRNA and the transgene (Fig. 4). Thus, the endogenous RAR β expression level served as a reference. The comparison of RNA and protein levels revealed similar protein/RNA ratios in RLZ79, M0, and M45 embryos. However, translation of M24 was overall reduced, despite its strong expression in heart and telencephalon. These data indicate that the uORFs in the wild-type construct do not inhibit translation of the major open reading frame in most tissues.

mRNA, but Not Protein Expression, Is Similar in Mutant and Wild-Type Mice

To investigate whether the low lacZ levels in heart and brain

of the wild-type construct could be accounted for by reduced mRNA levels, we analyzed RAR β -lacZ RNA and protein in hearts and brains of adult RLZ79 and M24 transgenic mice. Preliminary experiments showed that the lacZ activity was readily detectable in organ extracts from M24 transgenic mice, while we failed to detect any lacZ activity above background in the hearts and brains of RLZ79 transgenic mice, when measured in enzymatic assays. Therefore, lacZ activity was estimated by staining of organs from several mice of similar age (3–6 mo), either as whole mounts or after sectioning. RNAs were analyzed in parallel.

We found that RAR β -lacZ RNA levels were very similar in M24 and RLZ79 tissues (Fig. 5 B). However, β -gal staining patterns in both transgenic lines varied dramatically. In the brain, RLZ79 expressed β -gal activity only in a very small number of neurons, while M24 stained for β -gal in a large number of cells in the cortex (Fig. 5 A), the medial habenular nucleus, the striatum, and the telencephalic ependyma (Fig. 6). The β -gal activity in those few cells in the brains of RLZ79 mice that stained positive seemed to be similar to M24 mice. β -gal staining in the hearts was very strong in M24 transgenics, while RLZ79 transgenics were completely negative (Fig. 7). These data indicate that the uORFs inhibit translation of the RAR β -lacZ mRNA in heart and brain.

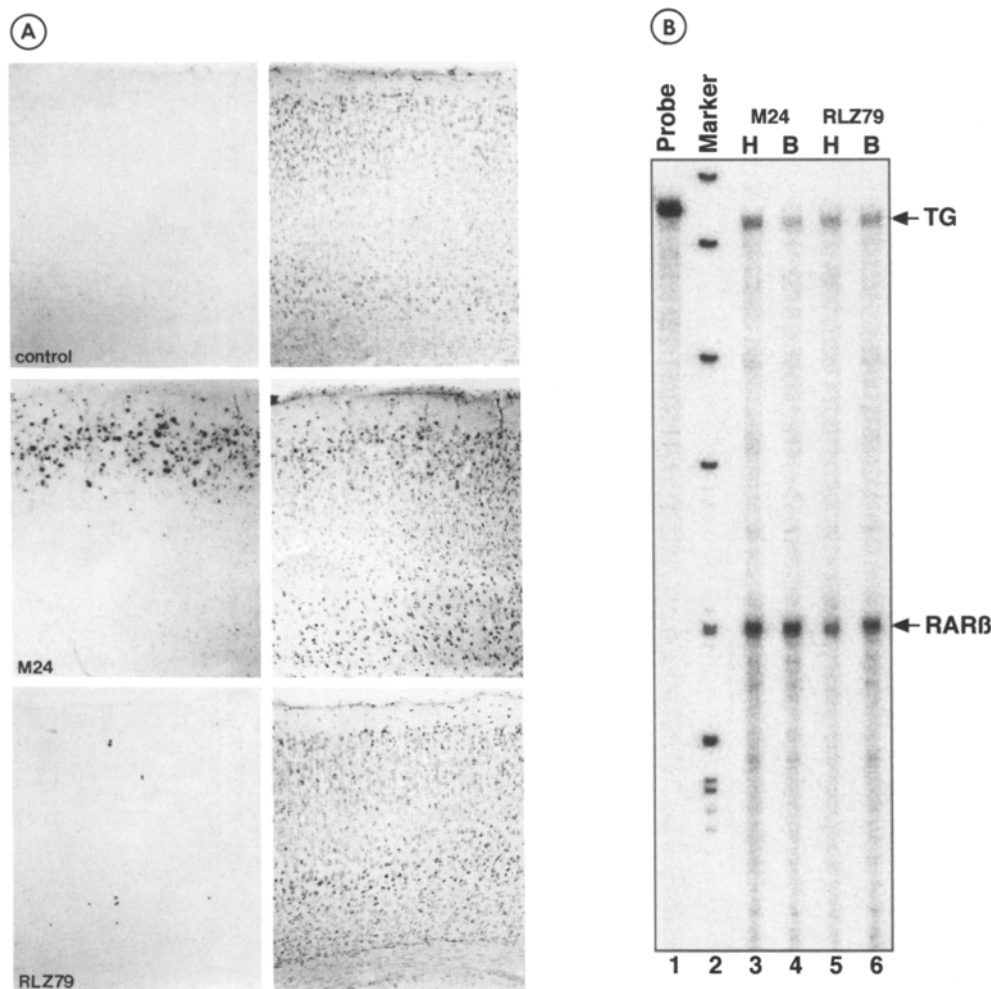


Figure 5. LacZ expression in brain and heart of adult mice. (A) Brain sections were cut after β -gal staining and the tissue was counterstained with neutral red as described in Materials and Methods. To visualize the β -gal staining, sections were photographed with a red filter (*left panel*). The right panel shows the same sections photographed under full light conditions. Non-transgenic controls showed no β -gal activity (*top row*). RLZ79 expressed only in very few cells in the cortex (*middle row*), while M24 (*bottom row*) expressed in a very large number of cells. (B) RNase protection assays of hearts and brains. Analysis was performed as described in Fig. 3 and Materials and Methods. The wild-type RLZ79 mRNA is expressed in the brain at slightly higher levels than M24 although protein staining is detected in fewer cells (see also Fig. 6). Expression in the heart is comparable with both constructs. Note however, that we found no β -gal staining in RLZ79 hearts (Fig. 7).

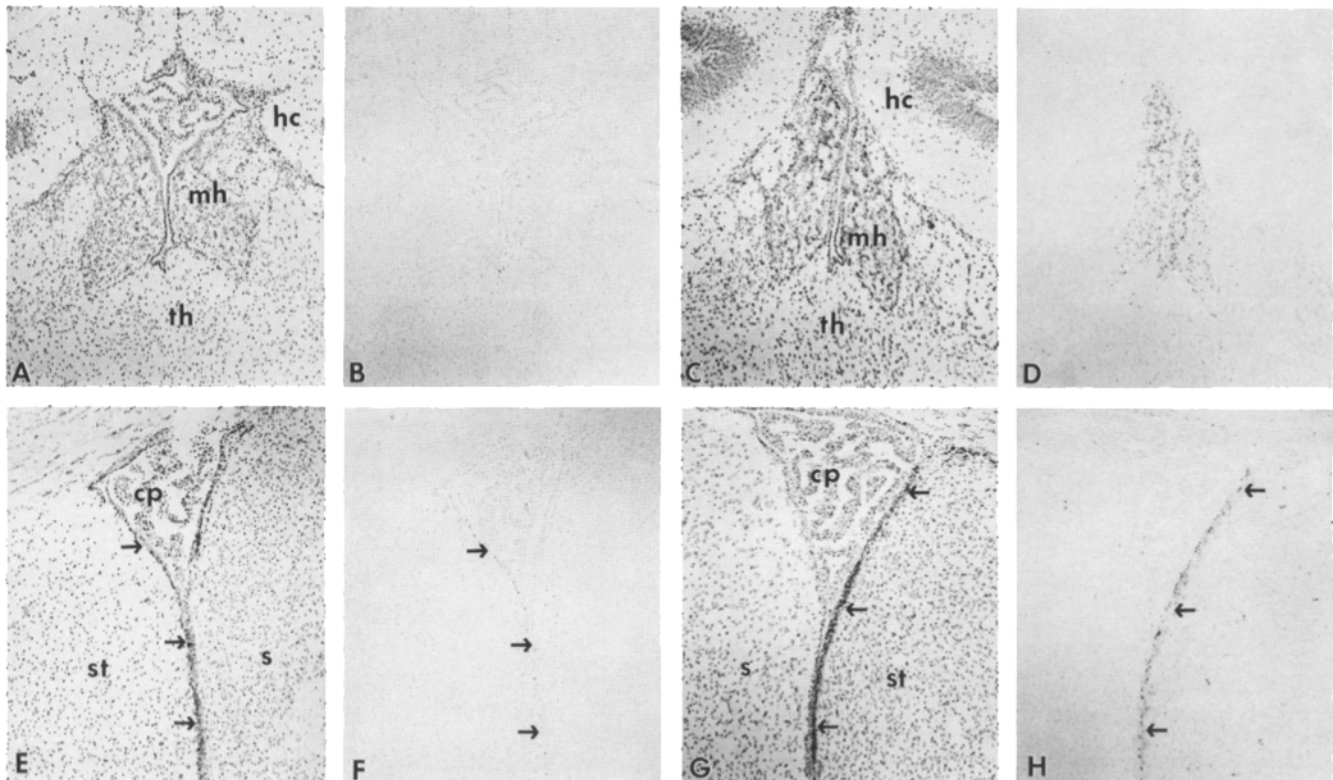


Figure 6. *LacZ* expression in the medial habenular nucleus, striatum, and ependyma. To compare *lacZ* expression in deeper brain structures between RLZ79 and M24, frozen sections from adult brains were stained for β -gal activity. (A and B) Sections through the medial habenular nucleus of RLZ79 transgenic mice showed β -gal staining only in very few cells while M24 (C and D) expressed in almost all cells. A similar pattern was seen in the ependyma (arrows) and striatum (st), where RLZ79 (E and F) expressed in very few and M24 (G and H) in many cells. Note that RNase protection assays revealed higher RNA levels from RLZ79 than from M24 (Fig. 5). *th*, thalamus; *mh*, medial habenular nucleus; *hc*, hippocampus; *cp* choroid plexus; *s*, septum.

Discussion

The complex 5'-UTR of the RAR β 2 mRNA is highly conserved between mice and humans, indicating an important function. Such a complex leader is not unique for the RAR β 2 mRNA, but can be found in ~5% of all vertebrate mRNAs, in particular those that encode for proteins with a proposed function in the regulation of cell growth and differentiation (Kozak, 1991). However, the role of these 5'-UTRs is unknown, although evidence suggests that they negatively interfere with translation (Marth et al., 1988; Rao et al., 1988; Arrick et al., 1991; Darveau et al., 1985). It was therefore very surprising to see that the mutation of RAR β 2-uORFs yielded no increase in the overall protein synthesis.

Nevertheless, mutations in uORFs in the 5'-UTR of RAR β 2-lacZ reporter gene constructs lead to differently regulated lacZ expression in hearts and brains of transgenic mice, while the expression in other tissues was very similar. Interestingly, every mutation tested effected the level and time of expression during embryogenesis differently. M45, for example, already expressed at high levels in the hearts of day 9.5 embryos, in contrast to M24 and M0. However, expression in the brain was higher with M24 and M0 than with M45, at all developmental stages. These results show that uORFs in the RAR β 2-5'-UTR can modulate gene expression in a tissue specific and developmentally regulated manner.

The best studied example of translational regulation by uORFs is the control of *GCN4* gene expression in *Saccharomyces cerevisiae* (Mueller and Hinnebusch, 1986; Miller and Hinnebusch, 1989; Abastado et al., 1991; Dever et al., 1992). The *GCN4*-open reading frame (ORF) is preceded by four small uORFs. Ribosomes initially translate the first uORF and subsequently reinitiate at a downstream AUG. Under conditions of amino acid sufficiency, ribosomes reinitiate at AUG of uORF2-4, instead of the *GCN4*-AUG. The frequency of reinitiation at uORF2-4 is reduced in amino acid starved cells and thus, some ribosomes resume translation at the *GCN4*-AUG (Mueller and Hinnebusch, 1986; Abastado et al., 1991). The reinitiation efficiency is regulated by the protein kinase GCN2, which is activated by uncharged tRNAs and phosphorylates the eukaryotic initiation factor 2 (eIF-2). When phosphorylated, eIF-2 sequesters the guanine nucleotide exchange factor eIF-2B, and thus reduces the level of active ternary complexes (Dever et al., 1992). Phosphorylation of eIF-2 by the heme regulated kinase (HRI) or dsI has initially been described as a mechanism regulating the general protein synthesis in mammalian cells (Farrell et al., 1977; Cooper and Farrell, 1977; De Benedetti and Baglioni, 1984; Chen et al., 1991). Reinitiation is thought to be more sensitive to reduced levels of ternary complexes than cap-dependent initiation events, because it is less efficient (Kozak, 1989). Thus, a molecular

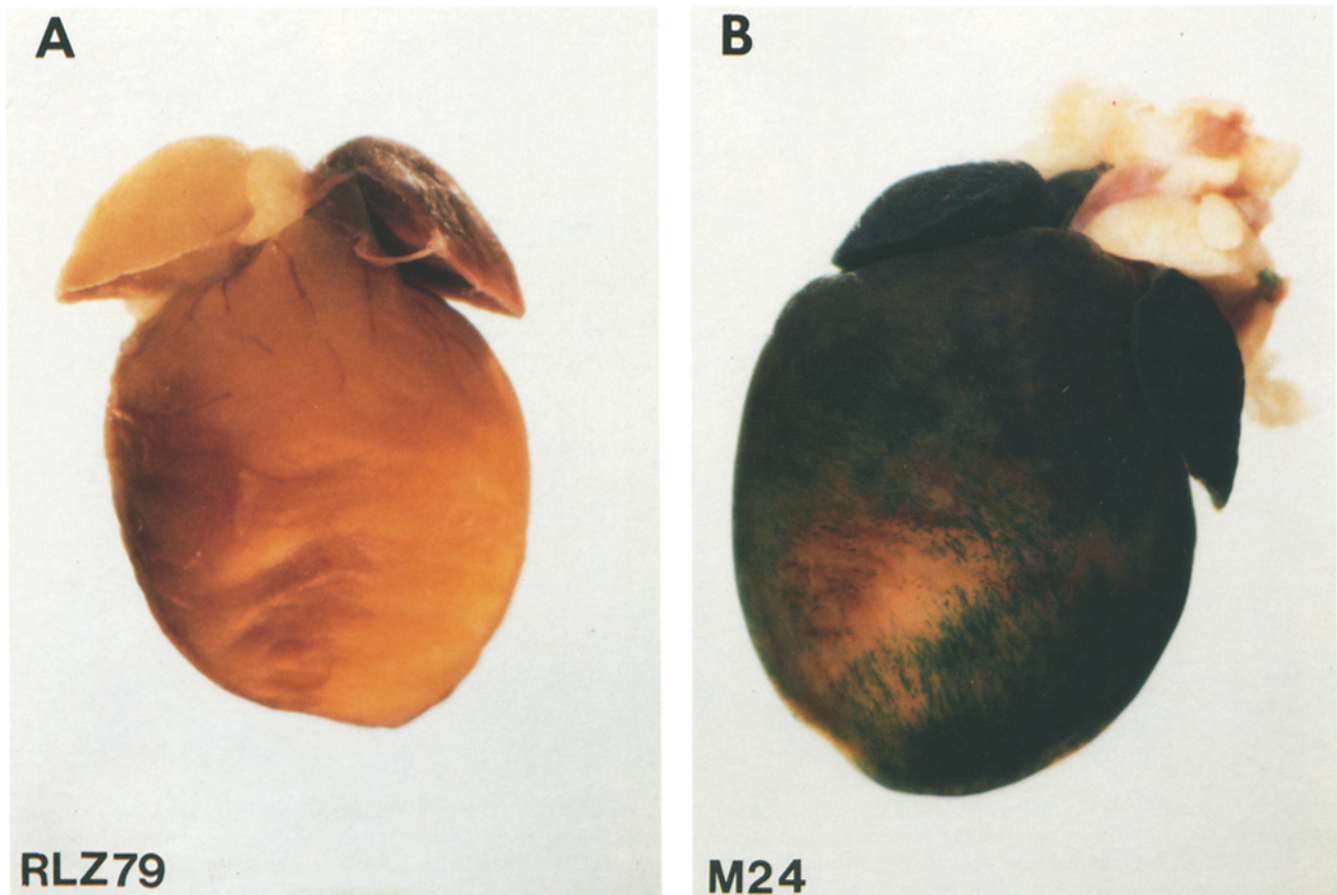


Figure 7. *LacZ* expression in adult hearts. Whole mount β -gal staining of adult hearts shows that RLZ79 (A) does not express β -gal, in contrast to M24 (B). Note that RLZ79 and M24 mRNA levels were similar (Fig. 4).

mechanism involved in the regulation of the general protein synthesis can be used to control gene specific translation as well.

Surprisingly, very little is known about the role of uORFs in complex leaders of vertebrate genes. Experiments involving expression constructs in which all or parts of the 5'-UTR were deleted showed that these leaders impaired translation (Marth et al., 1988; Ra et al., 1988; Arrick et al., 1991; Darveau et al., 1985). Therefore, it has been speculated that these 5'-UTRs might serve to reduce translation levels and that alterations in the translational efficiency of some proto-oncogenes may contribute to the development of the neoplastic phenotype (Darveau et al., 1985; Marth et al., 1988; Rao et al., 1988). This idea was supported by the finding that overexpression of eIF-4F (or a 24-kD subunit eIF-4E; Lazaris et al., 1990; Koromilas et al., 1992) or the IFN-inducible dsRNA-dependent protein kinase (dsRNA-PK; Koromilas et al., 1992) results in malignant transformation, possibly by increasing the translation of normally poorly translated mRNAs.

Our results show that a translational mechanism involving uORFs can be a very effective means to control tissue- or cell-specific gene expression patterns. This suggests that the translational control involving complex 5'-UTRs may be used to fine tune tissue specific gene expression patterns, rather than modulating the general level of protein synthesis. The molecular basis for tissue specific translational control

of the *RAR β 2* gene remains to be determined, but the distinct expression patterns of the mutations indicate that several factors may be involved. It is conceivable that specific differences in the composition of initiation and/or elongation factors permit the translational machinery to traverse the uORFs in some cells, but not in others. A tissue specific expression pattern has been reported for different eIF-4A genes (Nielsen and Trachsel, 1988) and eIF-4E (Jaramillo et al., 1991), but a comprehensive analysis for other factors is lacking. The activity of the translational machinery is modulated by posttranslational modifications, in particular through phosphorylation by several protein kinases (Tuazon et al., 1989; Hershey, 1991; Merrick, 1992; Samuel, 1992). It is conceivable that the tissue specific and developmentally coordinated activation or expression of these protein kinases may account for tissue specific differences in the ability of ribosomes to translate the *RAR β 2*-ORF.

The high degree of evolutionary conservation indicates that the *RAR β 2* isoform has a unique and important function. Furthermore, the very good conservation of the regulatory 5'-UTR suggests that the precise control of *RAR β 2* translation in heart and brain is important. Interestingly, a relatively small number of cells in the brain expressed the wild-type construct at high levels, while most cells were negative. As judged by the staining intensity, β -gal levels in those cells were similar to that seen in the brain cells of mutant mice. Thus it appears that the regulation of *RAR β 2*

translation in the brain is cell specific. A unique feature of the brain is the enormous number of specific cell-cell interactions, totaling as many as 10^{15} (Kandel et al., 1991). It is evident that this complexity can be specified only through combinatorial effects of many genes in conjunction with epigenetic processes. Conceivably, synaptic inputs might influence the translation of regulatory proteins by modulating the translational machinery through the activity of G-coupled receptors. Such a mechanism could play an important role in fine tuning the expression of many genes in neurons and in the specification of neuronal identity.

We thank Dr. Mike Brownstein for support and oligonucleotide synthesis, Rick Dreyfuss for excellent photographic work, and Drs. Mike Brownstein, Helen Lee Helmich, Marilyn Kozak, Andreas Kurtz, and Howard Nash for valuable comments on different versions of this manuscript.

Received for publication 6 June 1994 and in revised form 27 July 1994.

References

- Abastado, J. P., P. F. Miller, and A. G. Hinnebusch. 1991. A quantitative model for translational control of the GCN4 gene of *Saccharomyces cerevisiae*. *New Biol.* 3:511-524.
- Arrick, B. A., A. L. Lee, R. L. Grendell, and R. Derynck. 1991. Inhibition of translation of transforming growth factor-beta 3 mRNA by its 5' untranslated region. *Mol. Cell Biol.* 11:4306-4313.
- Chen, J. J., M. S. Throop, L. Gehrke, I. Kuo, J. K. Pal, M. Brodsky, and I. M. London. 1991. Cloning of the cDNA of the heme-regulated eukaryotic initiation factor 2 alpha (eIF-2 alpha) kinase of rabbit reticulocytes: homology to yeast GCN2 protein kinase and human double-stranded-RNA-dependent eIF-2 alpha kinase. *Proc. Natl. Acad. Sci. USA.* 88:7729-7733.
- Cooper, J. A., and P. J. Farrell. 1977. Extracts of interferon-treated cells can inhibit reticulocyte lysate protein synthesis. *Biochem. Biophys. Res. Commun.* 77:124-131.
- Darveau, A., J. Pelletier, and N. Sonenberg. 1985. Differential efficiencies of in vitro translation of mouse c-myc transcripts differing in the 5' untranslated region. *Proc. Natl. Acad. Sci. USA.* 82:2315-2319.
- De Benedetti, A., and C. Baglioni. 1984. Inhibition of mRNA binding to ribosomes by localized activation of dsRNA-dependent protein kinase. *Nature (Lond.)* 311:79-81.
- de The, H., A. Marchio, P. Tiollais, and A. Dejean. 1987. A novel steroid thyroid hormone receptor-related gene inappropriately expressed in human hepatocellular carcinoma. *Nature (Lond.)* 330:667-670.
- Dever, T. E., L. Feng, R. C. Wek, A. M. Cigan, T. F. Donahue, and A. G. Hinnebusch. 1992. Phosphorylation of initiation factor 2 alpha by protein kinase GCN2 mediates gene-specific translational control of GCN4 in yeast. *Cell.* 68:585-596.
- Eichele, G. 1993. Retinoids in embryonic development. *Ann. NY Acad. Sci.* 678:22-36.
- Eisel, U., K. Reynolds, M. Riddick, A. Zimmer, H. Niemann, and A. Zimmer. 1993. Tetanus toxin light chain expression in Sertoli cells of transgenic mice causes alterations of the actin cytoskeleton and disrupts spermatogenesis. *EMBO (Eur. Mol. Biol. Organ.) J.* 12:3365-3372.
- Evans, R. M. 1988. The steroid and thyroid hormone receptor superfamily. *Science (Wash. DC).* 240:889-895.
- Farrell, P. J., K. Balkow, T. Hunt, R. J. Jackson, and H. Trachsel. 1977. Phosphorylation of initiation factor eIF-2 and the control of reticulocyte protein synthesis. *Cell.* 11:187-200.
- Hershey, J. W. 1991. Translational control in mammalian cells. *Annu. Rev. Biochem.* 60:717-755.
- Jackson, R. J. 1993. Cytoplasmic regulation of mRNA function: the importance of the 3' untranslated region. *Cell.* 74:9-14.
- Jaramillo, M., J. Pelletier, I. Edery, P. J. Nielsen, and N. Sonenberg. 1991. Multiple mRNAs encode the murine translation initiation factor eIF-4E. *J. Biol. Chem.* 266:10446-10451.
- Kandel, E. R., J. H. Schwartz, and T. M. Jessell. 1991. Principles of Neural Science. 1135. Elsevier, New York. 1135.
- Kessel, M., and P. Gruss. 1988. Open reading frames and translational control. *Nature (Lond.)* 332:117-811.
- König, M., L. C. Mahan, J. W. Marsh, J. S. Fink, and M. J. Brownstein. 1991. Method for identifying ligands that bind to cloned G_i- or G_q-coupled receptors. *Mol. Cell. Neurosci.* 2:331-337.
- Koromilas, A. E., K. A. Lazaris, and N. Sonenberg. 1992a. mRNAs containing extensive secondary structure in their 5' non-coding region translate efficiently in cells overexpressing initiation factor eIF-4E [published erratum appears in *EMBO (Eur. Mol. Biol. Organ.) J.* 1992, Dec;11(13):5138] *EMBO (Eur. Mol. Biol. Organ.) J.* 11:4153-4158.
- Koromilas, A. E., S. Roy, G. N. Barber, M. G. Katze, and N. Sonenberg. 1992b. Malignant transformation by a mutant of the IFN-inducible dsRNA-dependent protein kinase. *Science (Wash. DC).* 257:1685-1689.
- Kozak, M. 1989. The scanning model for translation: an update. *J. Cell Biol.* 108:229-241.
- Kozak, M. 1991. An analysis of vertebrate mRNA sequences: intimations of translational control. *J. Cell Biol.* 115:887-903.
- Lazaris, K. A., K. S. Montine, and N. Sonenberg. 1990. Malignant transformation by a eukaryotic initiation factor subunit that binds to mRNA 5' cap. *Nature (Lond.)* 345:544-547.
- Marth, J. D., R. W. Overell, K. E. Meier, E. G. Krebs, and R. M. Perlmutter. 1988. Translational activation of the lck proto-oncogene. *Nature (Lond.)* 332:171-173.
- Mendelsohn, C., E. Ruberte, M. LeMeur, G. Morriss-Kay, and P. Chambon. 1991. Developmental analysis of the retinoic acid-inducible RAR-β2 promoter in transgenic animals. *Development.* 113:723-734.
- Merrick, W. C. 1992. Mechanism and regulation of eukaryotic protein synthesis. *Microbiol. Rev.* 56:291-315.
- Miller, P. F., and A. G. Hinnebusch. 1989. Sequences that surround the stop codons of upstream open reading frames in GCN4 mRNA determine their distinct functions in translational control. *Genes Dev.* 3:1217-1225.
- Mueller, P. P., and A. G. Hinnebusch. 1986. Multiple upstream AUG codons mediate translational control of GCN4. *Cell.* 45:201-207.
- Naggal, S., A. Zelent, and P. Chambon. 1992. RAR-beta 4, a retinoic acid receptor isoform is generated from RAR-beta 2 by alternative splicing and usage of a CUG initiator codon. *Proc. Natl. Acad. Sci. USA.* 89:2718-2722.
- Nielsen, P. J., and H. Trachsel. 1988. The mouse protein synthesis initiation factor 4A gene family includes two related functional genes which are differentially expressed. *EMBO (Eur. Mol. Biol. Organ.) J.* 7:2097-2105.
- Okayama, H., M. Kawauchi, M. Brownstein, F. Lee, T. Yokota, and K. Arai. 1987. High-efficiency cloning of full-length cDNA; construction and screening of cDNA expression libraries for mammalian cells. *Methods Enzymol.* 154:3-28.
- Rao, C. D., M. Pech, K. C. Robbins, and S. A. Aaronson. 1988. The 5' untranslated sequence of the c-sis/platelet-derived growth factor 2 transcript is a potent translational inhibitor. *Mol. Cell Biol.* 8:284-292.
- Reynolds, K., E. Mezey, and A. Zimmer. 1991. Activity of the beta-retinoic acid receptor promoter in transgenic mice. *Mech. Dev.* 36:15-29.
- Sambrook, J., E. F. Fritsch, and T. Maniatis. 1989. Molecular Cloning: A Laboratory Manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.
- Samuel, C. E. 1992. Role of the RNA-dependent protein-kinase in the regulated expression of genes in transfected cells. *Pharmacology & Therapeutics.* 54:307-317.
- Shen, S., C. E. Vandenbrink, W. Kruijer, and P. T. Vandersaag. 1992. Embryonic stem-cells stably transfected with mra-β-2-lacZ exhibit specific expression in chimeric embryos. *Int. J. Dev. Biol.* 36:465-476.
- Tuazon, P. T., W. C. Merrick, and J. A. Traugh. 1989. Comparative analysis of phosphorylation of translational initiation and elongation factors by seven protein kinases. *J. Biol. Chem.* 264:2773-2777.
- van der Leede, B., G. E. Folkers, F. A. Kruyt, and P. van der Saag. 1992. Genomic organization of the human retinoic acid receptor beta 2. *Biochem. Biophys. Res. Commun.* 188:695-702.
- Zelent, A., C. Mendelsohn, P. Kastner, A. Krust, J. M. Garnier, F. Ruffenach, P. Leroy, and P. Chambon. 1991. Differentially expressed isoforms of the mouse retinoic acid receptor beta generated by usage of two promoters and alternative splicing. *EMBO (Eur. Mol. Biol. Organ.) J.* 10:71-81.
- Zimmer, A., and A. Zimmer. 1992. Induction of a RAR beta 2-lacZ transgene by retinoic acid reflects the neuromeric organization of the central nervous system. *Development.* 116:977-983.