Regulation of RARβ2 mRNA Expression: Evidence for an Inhibitory Peptide Encoded in the 5′-untranslated Region

Kay Reynolds, Anne M. Zimmer, and Andreas Zimmer

Unit of Developmental Biology, National Institute of Mental Health, Bethesda, Maryland 20892-4090

Abstract. Regulation of mRNA translation and stability plays an important role in the control of gene expression during embryonic development. We have recently shown that the tissue-specific expression of the RARβ2 gene in mouse embryos is regulated at the translational level by short upstream open reading frames (uORFs) In the 5′-untranslated region (Zimmer, A., A.M. Zimmer, and K. Reynolds. 1994. *J. Cell Biol.* 127:1111–1119). To gain insight into the molecular mechanism, we have performed a systematic mutational analysis of the uORFs. Two series of constructs were tested: in one series, each uORF was individually inactivated by introducing a point mutation in its start codon; in the second series, all but one ORF were inactivated. Our results indicate that individual uORFs may have different functions. uORF4 seems to inhibit translation of the major ORF in heart and brain, while uORFs 2 and 5 appear to be important for efficient translation in all tissues. To determine whether the polypeptide encoded by uORF4 or the act of translating it, is the significant event, we introduced point mutations to create silent mutations or amino acid substitutions in uORF4. Our results indicate that the uORF4 amino acid coding sequence is important for the inhibitory effect on translation of the downstream major ORF.

THE scanning model for translational initiation predicts that the translational initiation site is dictated by the location and sequence context of the initiation codon (Kozak, 1989*a*). The 40S preinitiation complex binds to the 5′ cap-proximal region of the mRNA and migrates downstream until it encounters an AUG in a favorable context. At this point, the 60S ribosomal subunit binds to the complex and protein synthesis commences. Most eukaryotic mRNAs have a short 5′-untranslated region (UTR)[1] without stable secondary structures and no upstream AUGs. These features enable the preinitiation complex to efficiently scan the 5′-region for the translation start site. Complex 5′-UTRs with upstream AUGs and/or the ability to form stable secondary structures are found only in a small minority of transcripts (Kozak, 1989*a*). Strinkingly, most of these genes, which include homeobox genes, proto-oncogenes, transcription factors, and signal transduction components are thought to play an important role in the regulation of embryonic development. It has been proposed that expression of many of these genes is regulated at the translational level (Kozak, 1989*a*).

Please address all correspondence to A. Zimmer, Unit of Developmental Biology, Laboratory of Cell Biology, National Institutes of Mental Health, Building 36/3A17, Bethesda, MD 20892-4090. Tel.: (301) 496-7377. Fax: (301) 402-1748. E-mail: zimmer@codon.nih.gov.

1. *Abbreviations used in this paper*: lacZ, β-galactosidase gene; RA, retinoic acid; RARβ2, retinoic acid receptor-β2; uORF, open reading frame; UTR, untranslated region.

The retinoic acid receptor-β2 (RARβ2) mRNA has a complex 5′-UTR. RARβ2 is a retinoic acid (RA)–dependent transcription factor that belongs to the superfamily of steroid/thyroid hormone receptors (Evans, 1988). RA plays an important role as a signaling molecule during embryoic development and homeostasis. It is a very potent teratogen and can induce a large variety of severe congenital defects, depending on the dose and time of exposure. Administration during early gastrulation, for example, results in craniofacial and heart defects (Kochhar, 1967) as well as skeletal abnormalities (Kessel and Gruss, 1991; Kessel, 1992).

The family of RAR genes consists of three members designated as RARα, RARβ, and RARγ. Several different transcripts are generated by each of these genes through differential splicing and promoter usage. The transcripts have different 5′-UTRs and encode for proteins with distinct NH_2 termini. All isoforms share a common DNA-binding and COOH-terminal domain. Expression of the RARβ2 isoform is strongly induced upon RA treatment through an RA responsive element. The expression pattern in gastrulating embryos is consistent with a potential role in teratogenesis (Zimmer and Zimmer, 1992).

We have previously shown that expression of the RARβ2 mRNA is regulated by a translational mechanism. The 5′-UTR of the RARβ2 mRNA, but not of the other isoforms, contains five short upstream open reading frames (uORFs) which precede the major ORF (Reynolds

et al., 1991; Zelent et al., 1991; Zimmer et al., 1994). These uORFs encode small peptides of 16–21 amino acids. Three of these five uORFs have staggered organization in which the stop codon of one uORF overlaps with the start codon of the following uORF. We studied the role of the 5′-UTR in the regulation of the RARβ2 gene by fusing the RARβ2 5′-region including the RARβ2 promoter, the 5′-UTR, and the first 29 codons of the major open reading frame, to the coding region of the bacterial β-galactosidase gene (lacZ). Expression of the lacZ gene in these RARβ2-lacZ reporter constructs was subject to regulatory mechanisms mediated by the 5′-UTR (Reynolds et al., 1991; Zimmer and Zimmer, 1992). LacZ expression was surprisingly efficient in transgenic embryos, considering the obstacles that the uORFs are thought to create for scanning ribosomes. Strikingly, mutations in the overlapping start/stop codons or deletion of all uORFs led to deregulated translation of RARβ2-lacZ reporter constructs in heart and brain, suggesting that the uORFs control RARβ2 translation in a tissue-specific manner (Zimmer et al., 1994).

To gain insight into the molecular mechanism underlying the translational regulation, we have now examined the role of individual uORFs. Two series of constructs were generated. One series contained mutations in only one uORF, and in the other series, all except one uORF were mutated. Analysis of the lacZ expression pattern in transgenic embryos derived with these constructs indicated that the 5′-UTR is composed of inhibitory and stimulatory uORFs. uORF4, in particular, appeared to play the most important role for the inhibition of translation in heart and brain. We present evidence that the peptide encoded by uORF4 may be responsible for the inhibition of translation of the downstream major ORF in heart and brain tissues.

Materials and Methods

M1 to M5 Constructs

Individual uORFs in the RARβ2-5′-UTR were inactivated by introducing point mutations into the start codons using an oligonucleotide-directed mutagenesis system (Amersham Corp., Arlington Heights, IL) according to the manufacturer's recommendations. The template was a PstI fragment from the wild-type reporter construct RLZ79 (Reynolds et al., 1991) that was inserted into M13mp18.

The following oligonucleotides were used for site-directed mutagenesis: M1-GTCCT CTGCA ACCCA GTCTT; M2-TCATT TACCA ATTTC GAGGC; M3-CCAAA TGATC AATTA CCATT; M4-CCCGA ATCAA GAATT ATGAC; M5-ACGAA CTCCT TCAAA CTCTC. The mutated PstI fragment was reinserted into RLZ79 after partial PstI digestion. All mutations were verified by di-deoxy sequencing.

M01 to M05 Constructs

Constructs were made that had all but one of the uORFs inactivated by point mutations. HindIII–BamHI fragments that contained the entire RARβ2-5′-UTR from M2, M3, M4, M5, and M0 (Zimmer et al., 1994) were subcloned into Bluscript KS+. The resulting plasmids were named RM2, RM3, RM4, RM5, and RM0, respectively.

RM01: Mutations in the uORF4 and uORF5 start sites in constructs RM4 and RM5 were combined into one construct using a unique BsaAI site between uORF4 and uORF5 (Zimmer et al., 1994) to yield plasmid RM4+5. The HindIII–BamHI fragment was subsequently inserted into M13mp18 and start codons of uORF2 and uORF3 were mutated using the oligonucleotide: CATCA ATTAC CAATT TCCAGG.

RM02, RM03, and RM04: To obtain these plasmids we took advantage of an AluI site between uORF3 and uORF4. A 141-bp PpuMI × AluI fragment contained the start codons of uORF2 and uORF3, and a 268-bp AluI × BamHI fragment contained the start sites of uORF4 and uORF5. RM02 was made by inserting a PpuMI × AluI fragment from RM3 and a AluI × BamHI fragment from RM0 into RM0 cut with PpuMI and BamHI. RM03 contained the PpuMI × AluI fragment from RM2 and the AluI × BamHI fragment from RM0, and RM04 the PpuMI × AluI fragment from RM0 and the AluI × BamHI fragment from RM5.

RM05: A BsaAI fragment that covers the uORF5 start codon from RM0 was replaced by the corresponding fragment from RM5. All uORFs, except uORF5, were deleted in the resulting plasmid RM05. All constructs were verified by sequencing.

The final MOX constructs were obtained by inserting a BamHI fragment from RLZ79 containing the lacZ gene into the BamHI site of the RM0X plasmids.

M4*, M04*, and M3* Constructs

Three constructs were made to test the effects of mutations in the amino acid sequence of uORF3 and uORF4.

M4* and M04* constructs were derived from RLZ79 and RMO, respectively, by introducing point mutations into the first or second bases of codons 5, 6, 12, and 13 of the uORF4 inferred peptide. The oligonucleotide used to obtain M4* was as follows: GCC CTT TTT TTG GCT CCT TGG CTG TTG GTC TTT TTT CCC GCC CCG. The oligonucleotide used to obtain M04* was similar, but with the addition of bases to reintroduce the ATG into its original context and was as follows: CCC TTT TTT GGC TCC TTG GCT GTT GGT CTT TTT TCC CGC CCC GAA TCA TG.

M3* was derived from RLZ79 by introducing point mutations that did not alter the uORF4 amino acid sequence. The oligonucleotide used to obtain M3* was as follows: CGG CCT TTT TTG ACA GGT AGG CTG TTG GTC TTT CTC CCA CCC CCG AATC.

Generation and Analysis of Transgenic Mice

Transgenic mouse lines and transgenic embryos were established and stained for β-galactosidase (β-gal) activity as previously described (Reynolds et al., 1991; Eisel et al., 1993).

β-Gal activity was determined by enzymatic assays in extracts made from E 12.5 embryos and adult tissues. Embryos were dissected free from extraembryonic tissues and stored at −80°C. DNA was isolated from the extraembryonic material and analyzed by PCR to identify transgenic embryos. Adult organs were dissected, divided into equal halves, and stored at −80°C. Frozen tissues were homogenized in 500 μl PM2 buffer (100 mM Na-phosphate, 2 mM $MgSO_4$, 0.1 mM $MnCl_2$, 0.5% [vol/vol] β-mercaptoethanol, pH 8.0). Cell debris was pelleted by centrifugation in a microcentrifuge. β-Gal activity was measured at 570 nm in a 96-well microtiter plate after adding 25 μl chlorophenolred-β-D-galactopyranoside to 100 μl embryo extract. At least three embryos or tissue extracts were analyzed from each transgenic line.

RNA was isolated using RNA Stat-60 (Tel-Test "B" Inc., Friendswood, TX) according to the manufacturer's recommendations. RNA from embryos was analyzed by RNase protection assays. RNA from adult tissues was analyzed by Northern blotting.

To obtain a template for in vitro translation, a Bsu36I × EcoRI fragment that contained most of the lacZ gene was deleted from RLZ79. A radiolabeled RNA probe was made by in vitro translation of this plasmid with SP6 polymerase in the presence of ^{32}P-α-GTP, after digestion with BsaAI.

The endogenous RARβ2 mRNA protects a 187-nt fragment, while the transgene protects a 423-nt fragment. 10 μg of total RNA from embryos were hybridized with 1.5×10^5 cpm of this probe in hybridization buffer (40 mM Pipes, pH 6.4; 1 mM EDTA; 0.4 M NaCl; 80% [vol/vol] formamide) at 50°C overnight. 350 μl ice-cold RNase digestion buffer (10 mM Tris, pH 7.5; 5 mM EDTA; 300 mM NaCl) and 4 μl RNase stock (200 μg/ml RNase T1; 4 mg/ml RNase A; 10 mM Tris, pH 7.5; 15 mM NaCl) was added to this mixture and incubated at 37°C for 60 min. Subsequently, 20 μl of 10% SDS and 10 μl proteinase K (10 mg/ml) were added and incubated for an additional 15 min. The mixture was extracted with 400 μl phenol/chloroform/isoamylalcohol (25:24:1) and the protected RNA was precipitated with ethanol after adding 1 μl tRNA (25 mg/ml) as a carrier. The pellet was dissolved in loading buffer (80% formamide; 10 mM EDTA; 1 mg/ml xylene cyanol FF; 1 mg/ml bromophenol blue) and separated on a 6% sequencing gel. Radiolabeled Bluescript SK+ DNA di-

gested with HpaII was used as a size marker. Subsequently, the gel was dried and analyzed using a Bio-Imaging Analyzer (Fuji Medical Systems Inc., Stanford, CT).

Total RNA from adult tissues (10 μg) was separated electrophoretically on a formaldehyde gel, blotted onto Nytran (S&S) membrane, and hybridized to a radiolabeled RNA probe. The blot was washed 3 × in 1 × SSPE/1% SDS at 65°C and 1 × in 0.1 × SSPE/1% SDS at 60°C and analyzed using a Bio-Imaging Analyzer.

RNA expression levels were quantitated using the NIH Image Program. The endogenous RARβ2 RNA or 28S RNA served as an endogenous standard in RNase protection assays or Northern blots, respectively. Translational efficiency was estimated by determining the ratio of β-gal activity (average of four samples) to RNA expression levels.

Results

Deletion of Single uORFs: Enhanced or Reduced Expression Levels

We have previously shown that expression of the RARβ2 gene in heart and brain tissues is regulated during embryonic development at the translational level by five uORFs in the 5′-UTR (Zimmer et al., 1994). To analyze the role of individual uORFs, we generated a series of constructs with point mutations in one of the uORF start codons. These mutations were designed to delete one of the uORFs while leaving the others intact (Fig. 1, M1 to M5; uORF1 is mutated in M1, uORF2 in M2, etc). With this series of constructs, we asked whether the deletion of only one uORF would be sufficient to alter expression in heart and brain. Several transgenic mouse lines and transgenic founder embryos were generated with these constructs. To analyze the lacZ expression pattern, embryos were isolated between embryonic days E 9.5 and E 12.5, fixed, and stained for β-gal activity (Fig. 2 and Table I).

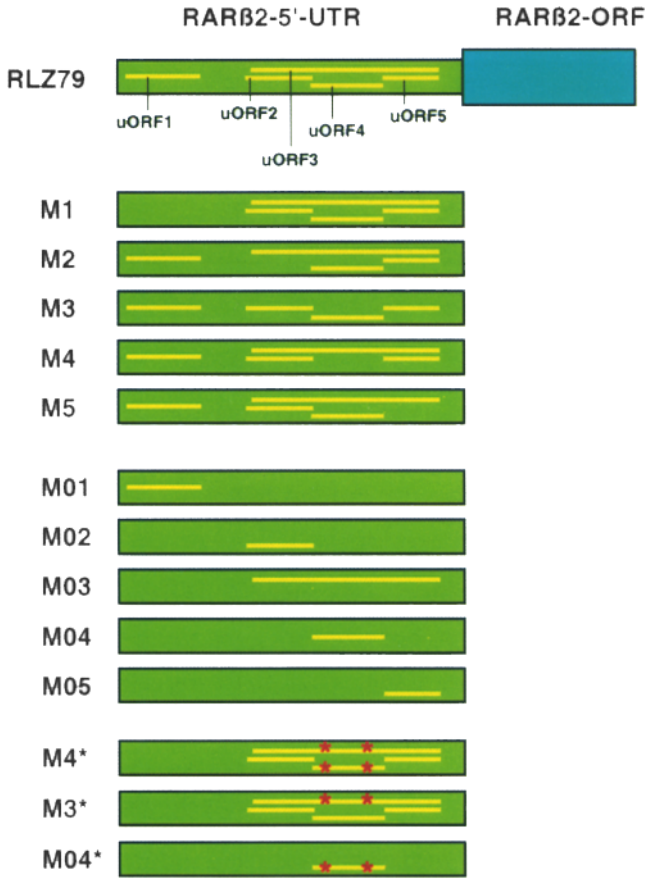

Figure 1. Schematic representation of the RARβ2-5′-UTR in the wild-type (RLZ79) and mutant constructs. In the M1–M5 series of constructs, one uORF was inactivated by introducing a point mutation into the start codon. (M1–M4) AUG is changed to UUG; (M5) AUG was changed to AAG. In the M01 to M05 series, all uORFs except one were mutated. Constructs M4*, M3*, and M04* contained mutations that altered the amino acid sequence of the uORF3 and/or uORF4 encoded peptides. See text for details.

β-Gal staining in these embryos was compared to that of embryos generated with the wild-type RLZ79 construct (Reynolds et al., 1991; Zimmer and Zimmer, 1992; Zimmer et al., 1994). E 12.5 embryos from each construct were sectioned to evaluate the β-gal staining in internal organs. We found that transgenic animals carrying the M1 construct expressed lacZ in the same pattern as mice with the wild-type RLZ79 construct. In particular, there was no detectable β-gal staining in the hearts and brains of TG-M1 embryos (Fig. 2 *A*, Table I). This indicates that the deletion of uORF1 does not affect the translation of the downstream major ORF within the limits of detection. Hence, we have used the β-gal staining pattern of TG-M1 embryos as a reference for comparison of other mutant constructs.

Interestingly, TG-M4 embryos expressed lacZ in the heart and brain, while TG-M1, TG-M2, TG-M3, and TG-M5 embryos did not express lacZ in these tissues (Fig. 2). At E 9.5, TG-M4 showed β-gal staining in the brain. E 10.5 and older embryos also expressed lacZ in the heart. β-Gal staining in other tissues of TG-M4 embryos was indistinguishable from TG-M1 orTG-RLZ79 embryos. These results indicate that the deletion of uORF4 leads to translation of the RARβ2 mRNA in heart and brain and points to an important role for uORF4 in the inhibition of translation in these tissues.

Strikingly, the mutation of either uORF2, uORF3, or uORF5 in constructs M2, M3, or M5, respectively, seemed to reduce lacZ expression levels (Fig. 2). We consistently found relatively weak β-gal staining in all embryos generated from these constructs, particularly in the thoracic region of the spinal cord at E 12.5. This was never seen with the wild-type RLZ79, M1, or M4 constructs. Histological sections of TG-M2 and TG-M5 embryos revealed that lacZ staining in other tissues, i.e., head mesenchyme, trachea, oesophagus, umbilical cord, and genital eminence, was very weak or not detectable. TG-M2 embryos also failed to express in nerve sheets (Fig. 2 *B*). At E 9.5, lacZ staining was weaker in the most rostral and caudal regions of TG-M5 embryos and very weak in TG-M2 embryos.

To get a better estimate of the amount of protein synthesized, we quantitated RARβ2lacZ mRNA and protein levels in E 12.5 embryos. We found that mRNA levels in M2 and M5 embryos were 12.5% and 45% of that of RLZ79 embryos, respectively. However, protein levels were only ~3% and 5% of the RLZ79 embryos (Fig. 3), indicating a very inefficient translation of these constructs. Translation of M1 and M4 were also about 25% lower than RLZ79. These results suggest that the precise organization of uORFs is important to ensure efficient translation of the major ORF.

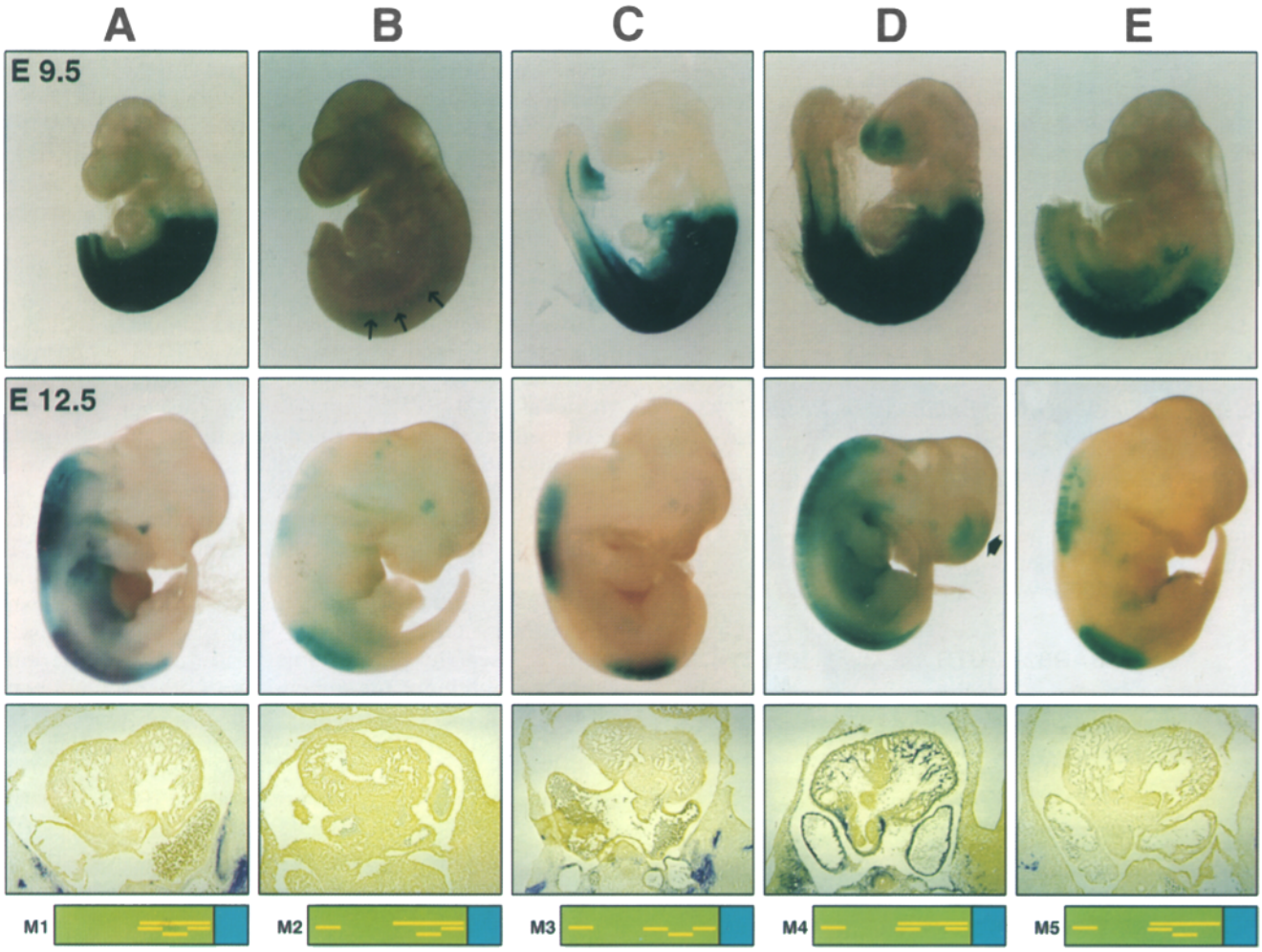

Figure 2. Comparison of lacZ expression patterns of constructs with a mutation in a single uORF. A schematic representation of the 5′-UTR of the mutant constructs is given at the bottom. M1 had a mutation in the AUG of uORF1, M2 in uORF2, etc. The major ORF is indicated by a blue box. Embryos were removed and stained for β-gal activity under identical conditions. Whole mount staining of E 9.5 embryos from each construct is shown in the upper panel; E 12.5 embryos are shown in the middle panel; and transverse sections through the hearts of E 12.5 embryos are shown in the lower panel. At E 9.5 most constructs expressed lacZ at similar levels except M2 which expressed only in the lateral mesenchyme (*small arrows*). At E 12.5 β-gal staining was consistently weaker in M2 and M5 embryos. The large arrow indicates β-gal staining in the telencephalon of M4 embryos. Note the β-gal staining in the heart wall and papillary muscles of M4 mice.

We also determined translational efficiency of the mutant constructs in hearts and brains from adult animals. β-Gal enzymatic activity in these tissues was very low in TG-M1, TG-M2, TG-M5, and TG-RLZ79 animals, but TG-M4 mice exhibited significant activity. Comparison of protein and mRNA levels revealed that the translational efficiency of M4-RNA was 7 (brain)–12 (heart) times higher than RLZ79-RNA (Fig. 3).

Constructs with Single uORFs: Inhibitory and Stimulatory uORFs

The above results suggest that uORF4 is involved in the inhibition of translation of the downstream RARβ2-AUG in heart and brain, while other uORFs seemed to be required for efficient translation. To determine the effect of a single uORF on lacZ expression, we generated a series of constructs in which all but one uORF were mutated (Fig. 1; constructs M01 to M05; uORF1 is intact in M01, uORF2 in M02, etc.). With this series we asked whether a single uORF would be sufficient to inhibit expression in heart and brain. Transgenic lines were established with M01 only. The other constructs were analyzed in several independently derived transgenic founder embryos (see Table I).

TG-M04 and TG-M03 embryos did not express lacZ at any developmental stage in heart or brain (Fig. 4), thus supporting the inhibitory role of uORF4 on brain and heart expression of lacZ. It was somewhat surprising that uORF3 as a solitary uORF also abolished heart and brain expression in TG-M03 embryos, because its deletion in construct M3 had no effect. It is possible that uORF3 is a weaker repressor of translation or it may be related to the fact that uORF3 is the longest of all uORFs.

Interestingly, TG-M01, TG-M02, and TG-M05 embryos expressed lacZ at different levels in heart and brain. In the

Table I. Transgenic Lines and Founder Embryos

Construct	TG lines	F0 embryos	9.5 H	9.5 B	10.5 H	10.5 B	11.5 H	11.5 B	12.5 H	12.5 B
RLZ79	2	22	−	−	−	−	−	−	−	a
M1	4	0	−	−	−	−	−	−	−	−
M2	1	4	−	−	−	−	−	−	−	−
M3	0	9	−	−	nd	nd	nd	nd	−	−
M4	5	16	−	++	+	++	+	++	++	+
M5	2	6	−	−	−	−	−	−	−	−
M01	3	0	++	+	++	+	+	−	+	−
M02	0	4	+	+	nd	nd	nd	nd	+	−
M03	0	7	−	−	nd	nd	nd	nd	−	−
M04	0	6	−	−	nd	nd	nd	nd	−	−
M05	0	8	−	+	nd	nd	nd	nd	+	++
M4*	0	16	−	++	nd	nd	nd	nd	++	b
M3*	0	10	−	−	nd	nd	nd	nd	−	c
M04*	0	15	−	+	nd	nd	nd	nd	+	+

Summary of the expression of constructs in the heart (H) or the brain (B) of transgenic embryos. Several transgenic lines (TG lines) or transgenic founder embryos (F0 embryos) were analyzed for each construct.
−, no expression; +, weak expression; ++, strong expression; nd, not done; a , only one line expressed in the telencephalon (Zimmer et al., 1994); b = one out of twelve embryos expressed in the brain; c, one embryo expressed in all regions of the brain.

brain, all embryos expressed a relatively low level at E 9.5, although β-gal staining in TG-M05 embryos was somewhat stronger than in TG-M01 or TG-M02 embryos. At E 12.5, only TG-M05 embryos expressed strongly in the brain. We compared the amount of β-gal staining in E 12.5 embryos generated with M0 (Zimmer et al., 1994) or M05 and found much stronger staining in TG-M05 embryos (data not shown). Hence, it appears that uORF5 stimulates translation in the brain.

In the heart, TG-M01 exhibited more β-gal staining at E 9.5 when compared with other embryos of this series or with TG-M0 embryos, thus suggesting a stimulatory role for uORF1. At E 12.5, TG-M01, TG-M02, TG-M05, and TG-M0 expressed at similar levels.

uORF4 Encodes an Inhibitory Peptide

To determine if the inhibitory effect of uORF4 is dependent on its amino acid coding sequence, we have changed the amino acid sequence by introducing point mutations into the first or second bases of codons 5, 6, 12, and 13 (Fig. 5 *A*). Because uORF3 and uORF4 overlap, these mutations also changed the amino acid sequence of the uORF3 encoded peptide. The resulting plasmid was called M4*. Another construct, M3*, contained point mutations in nucleotides adjacent to those that were altered in M4* (Fig. 5 *B*). These mutations, however, preserved the amino acid sequence of the peptide encoded by uORF4 but changed the uORF3 peptide. The general organization of uORFs in these constructs was not changed.

Four E 9.5 and twelve E 12.5 transgenic founder embryos were derived with this construct and stained for β-gal expression. Representative embryos are shown in Fig. 5 A. At E 9.5, all TG-M4* embryos expressed in the brain, but not in the heart. At E 12.5, eight out of twelve TG-M4*

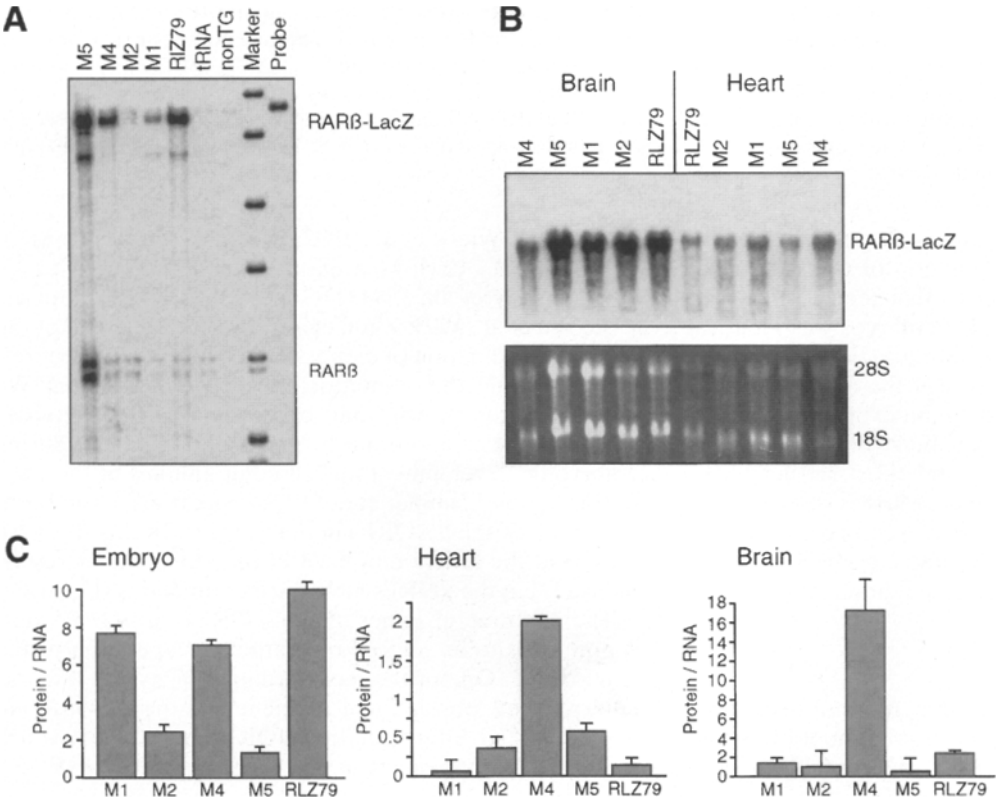

Figure 3. Determination of RARβ2-lacZ mRNA and protein levels. (*A*) RNA from E 12.5 embryos was analyzed by ribonuclease protection assays. The top bands represent the transgenic transcripts and the lower bands, the endogenous RARβ2 mRNAs. Comparison of the endogenous RARβ2 expression levels with the transgene mRNA showed that expression levels varied between two and eightfold between individual lines. (*B*) Northern blot analysis of heart and brain tissues from adult animals. The top panel shows the autoradiograph. The lower panel shows the ethidium bromside-stained RNA gel. (*C*) Quantitation of RARβ2-lacZ RNA and protein levels. Protein levels were determined enzymatically (see Materials and Methods). Translational efficiency is expressed in arbitrary values as the ratio of protein activity to RNA level. Averages and standard errors were compiled from at least three independent β-gal assays. Note that brain tissues exhibited significantly more β-gal activity than heart tissues.

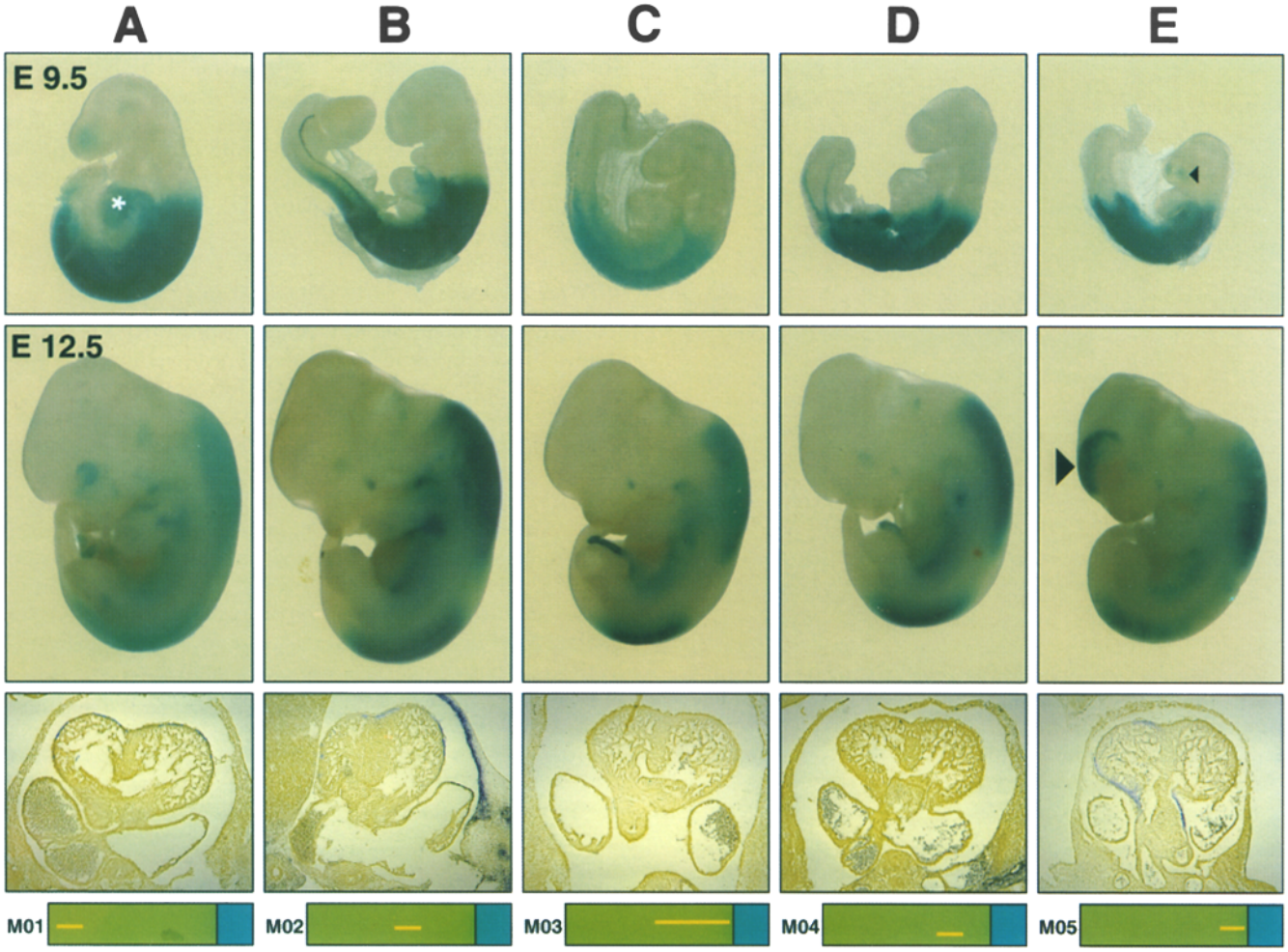

Figure 4. Comparison of lacZ expression patterns of constructs with solitary uORFs. The bottom panel shows a schematic representation of the 5′-UTR of the constructs used in this experimental series. In M01, all uORFs were mutated except uORF1, in M02 all except uORF2, etc. Whole mount staining of E 9.5 embryos is shown in the upper panel and E 12.5 embryos, in the middle panel. Transverse sections through the hearts of E 12.5 embryos are shown in the lower panel. Note β-gal staining in the heart of TG-M01 embryos at E 9.5 (*asterisk*) and in the telecephalon of TG-M05 embryos (*arrowheads*). TG-M01 and TG-M05 embryos expressed in more cells of the heart wall compared to embryos generated with M0, a construct that lacks all uORFs (Zimmer et al., 1994). No expression was seen in papillary muscle (compare to M4, Fig. 5). LacZ expression in TG-M02 embryos was similar to that of TG-M0 embryos. TG-M03 and TG-M04 embryos did not express in the heart.

embryos expressed in the heart, but only one expressed in the brain. Thus, the altered regulation of translation observed with M4* is very similar to that seen with M4. In contrast, β-gal activity in TG-M3* embryos was identical to that of wild-type TG-RLZ79 embryos (Fig. 5 *B*).

Finally, we wanted to determine if the altered uORF4 coding sequence (in M4*) could inhibit expression in heart and brain when introduced as a solitary uORF in the 5′-UTR, as the wild-type sequence did (in construct M04). The corresponding mutant construct M04* is shown in Fig. 5 *C*. A total of nine E 9.5 and six E 12.5 embryos were analyzed. As shown in Fig. 5 *C*, the mutated uORF4 sequence failed to inhibit expression in heart and brain.

Discussion

There is accumulating evidence that retinoid-dependent pathways play a crucial role in the regulation of brain and heart development and myocardial differentiation (Hart et al., 1992; Wiens et al., 1992; Kastner et al., 1994; Mendelsohn et al., 1994; Moklentin et al., 1994; Sucov et al., 1994; Yutzey et al., 1994; Dyson et al., 1995; Twal et al., 1995; Wu et al., 1995; Zhou et al., 1995). It is likely that the precise coordination of expression of the mediators of retinoid-signaling, the retinoic acid receptors, is essential. We have recently shown that expression of the RARβ2 mRNA in heart and brain is specifically regulated during embryonic development and in adult animals at the level of translation (Zimmer et al., 1994). The translational control involves small uORFs in the 5′-UTR. In this study we looked at the function of each of the uORFs. Our results indicate that the uORFs have distinct functions (Fig. 6).

The mutation of either uORF2, ORF4, or uORF5 start codon leads to an altered regulation of expression of the major ORF. One of the most striking effects was a markedly reduced translational efficiency of the downstream RARβ2-AUG after mutating uORF2 or uORF5, which was particularly evident in the thoracic region of E 12.5

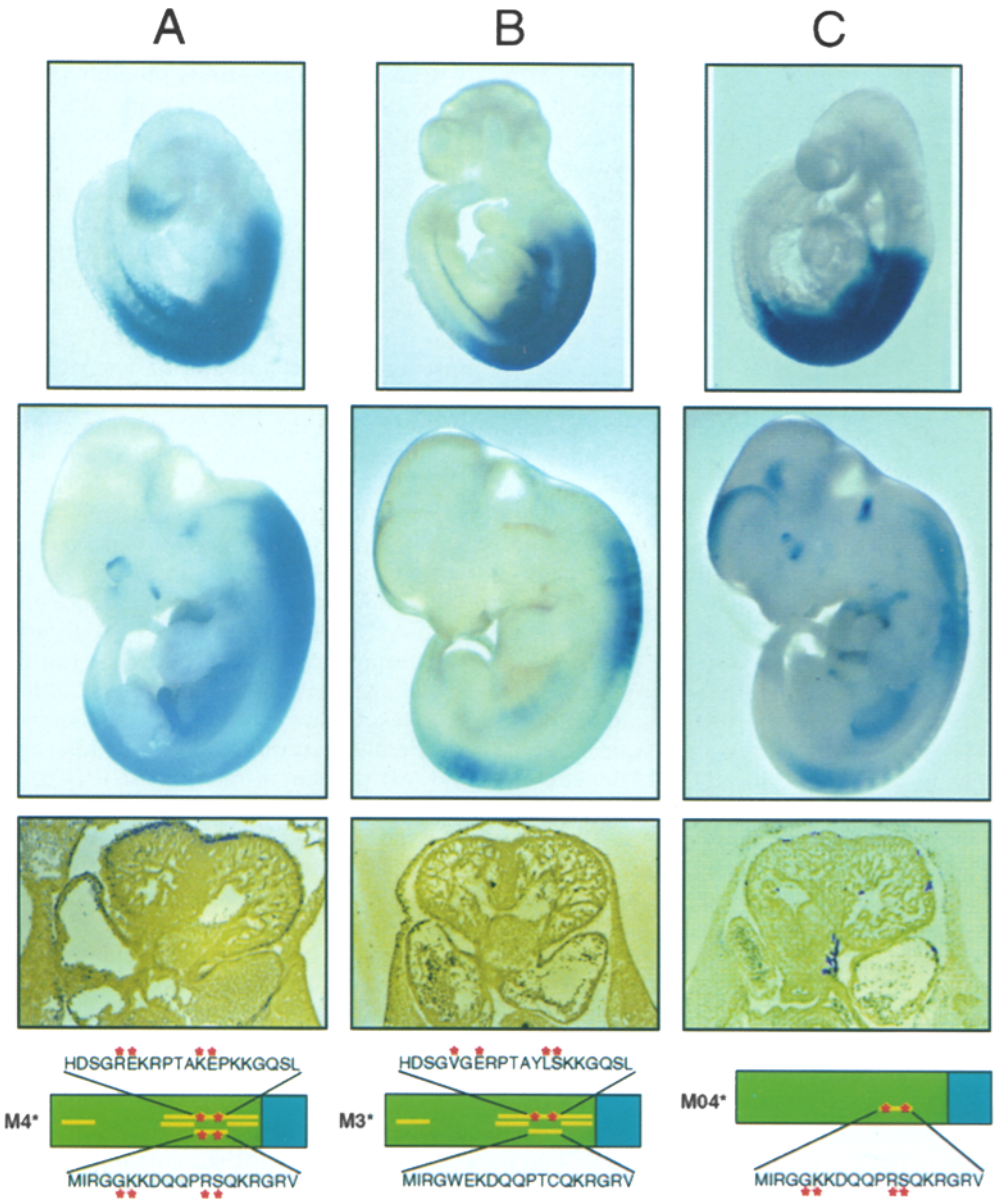

Figure 5. (*A*) Changes to the uORF4 amino acid sequence. The amino acid sequence of uORF4 was altered in construct M4* by introducing point mutations into the first or second bases of codons 5, 6, 12, and 13. These mutations also altered the amino acid sequence of the uORF3 encoded peptide, but did not change the overall organization of uORFs. Transgenic embryos were isolated on E 9.5 (*top panel*) or E 12.5 (*middle panel*) and stained for β-gal activity. Note lacZ expression in the brain at E 9.5 and heart at E 12.5 (*lower panel*). (*B*) Point mutations that altered the nucleotide sequence in the uORF4 region and the amino acid sequence of the uORF3 encoded peptide, but not the uORF4 encoded peptide. The mutated nucleotides were in neighboring positions of those changed in M4*. LacZ expression in TG-M3* embryos was identical to wild-type TG-RLZ79 embryos and not found in hearts and brains. (*C*) the mutated uORF4 as a solitary uORF. The M04* construct was similar to the M04 construct, but contained mutations in the amino acid sequence of uORF4. Note that the mutated uORF4 sequence did not inhibit expression in heart and brain, as did the wild-type sequence.

embryos. These results indicate that both uORFs are important for efficient expression of the major ORF in all tissues. We also found relatively low mRNA levels in TG-M2 and TG-M5 mice. This could be due to transcriptional variation, which is commonly observed between individual transgenic lines, or it may be related to the fact that mRNA stability is often correlated with translation (Atwater et al., 1990). Although mutating the AUG or uORF1 had no apparent effect on the expression pattern of the reporter gene construct, uORF1 seemed to enhance translation in the heart when present as a solitary uORF. These findings suggest that all of the uORFs in the RARβ2-5′-UTR inhibit or enhance translation of the downstream major ORF in heart and/or brain tissues. It appears that the complex arrangement of uORFs is important to prevent translation in heart and brain, while permitting expression in other tissues.

Is it possible that some of the general translational effects observed with constructs M2 and M5, or the tissue-specific effects seen with constructs M4 and M4*, or M24 and M45 (Zimmer et al., 1994), are due to alterations in mRNA secondary structure of the mutant transcripts? To explore this possibility we have analyzed the mRNA secondary structure of the human and mouse 5′-UTRs and those of the mutant constructs by the method of Zuker (1989). This analysis revealed that the RNA secondary structure of mouse and human 5′-UTRs is very different, despite the fact that the sequence, in particular the region spanning uORF2 to uORF5, is highly conserved (Zimmer et al., 1994). The 5′-UTRs from both species are predicted to form unrelated stem-loop structures and have different free energies (nt 1–592; human, −186.3 kcal; mouse, −205.0 kcal/mol). Furthermore, potential alterations in mRNA secondary structure of the mutant constructs could not be reconciled with their expression pattern.

Is it possible that a minor form of mRNA is generated through tissue-specific splicing and promoter usage from some mutant constructs tested in this study, and could this

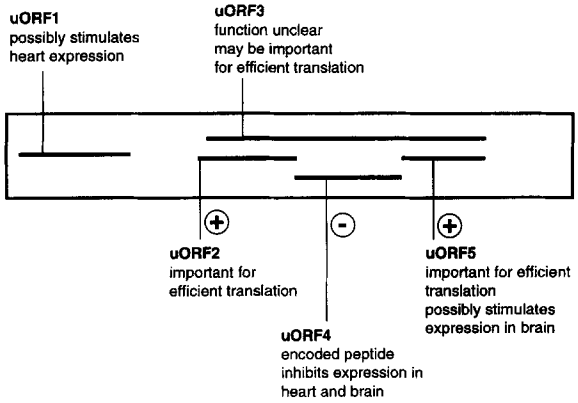

Figure 6. Summary of the functional analysis of uORFs.

minor form be the major functional mRNA? It is conceivable that the β-gal assay is sensitive enough to detect such a minor mRNA form. We cannot address this question directly in transgenic embryos, due to the limited availability and heterogeneity of cell populations in embryonic brain and heart tissues. However, we have now identified a tissue culture system, in which we can reproduce the differential translation that we find in vivo (Reynolds, K., A.M. Zimmer, unpubished results). This system will help us to address these and other questions in future experiments. While we cannot exclude the possibility that subtle changes in RNA secondary structure, or the production of minority transcripts, contribute to the translational regulation, we feel that, in summary, our results provide overwhelming evidence that the organization and sequence of uORFs is most critical for translational regulation of the RARβ2 mRNA.

The AUG of uORF2 is in the best sequence context of all uORFs (UGGAAAAUGG). It is the only one that contains the preferred bases in the critical positions −3 (A) and +4 (G). The AUG of uORF1, on the other hand, is in a relatively poor sequence context (ACUGGGAUGC). Therefore, the AUG of uORF2 is the most 5′-proximal AUG with a good sequence context and should be translated efficiently according to the scanning model for translational initiation. Indeed our results show that the deletion of uORF2 in the context of the wild-type 5′-UTR results in very low β-gal activities indicating an important function for uORF2. However, uORF2 is not required when the other uORFs are mutated as in construct M0 (Zimmer et al., 1994), or in the M0x series of constructs (Fig. 4). These findings suggest that uORF2 is important to alleviate the inhibitory effects of other uORFs. If one assumes that uORF2 is translated, it is somewhat surprising that uORF4, the start codon of which overlaps the uORF2 stop codon, has such a strong inhibitory effect on translation in heart and brain. In eukaryotic cells, the reinitiation frequency increases with the distance between the 5′- and 3′-cistrons (Kozak, 1991*a*) which might suggest that initiation at uORF4 should be meager after translation of uORF2. However, reinitiation is very efficient in bacteria when the 5′ and 3′ ORFs overlap.

The molecular basis for the positive effects of uORF1 and uORF5 remains to be determined, but it may be related to the fact that the RARβ2-5′-UTR can form complex secondary structures (Zelent et al., 1991). Stable secondary structures (−50 or −60 kcal/mol) constitute a strong barrier for scanning 40S ribosomal subunits (Pelletier and Sonenberg, 1985; Kozak, 1986), while 80S ribosomes can penetrate these structures to some extent (Kozak, 1989*b*). It has therefore been proposed that translation of uORFs by 80S ribosomes might alleviate the inhibitory effect (Kozak, 1991*b*).

uORF4 was important for inhibiting the translation of the major RARβ2-ORF. Mutation of uORF4 resulted in much more efficient lacZ expression in heart and brain tissues, but had little or no effect on other tissues (Figs. 2 *D* and 3 *C*). The presence of uORF4 as a solitary uORF was sufficient to block the expression of the reporter gene in hearts and brains. uORFs in the 5′-UTR as translational inhibitors have been reported for other cellular and viral genes such as TGF-β3 (Arrick et al., 1991), GCN4 (Mueller and Hinnebusch, 1986), and the Rous sarcoma virus (Donzé and Spahr, 1992). Several structural features have been determined that modulate the inhibitory effects of uORFs on translation downstream: (*1*) The sequence context of the upstream AUG (Kozak, 1984); (*2*) the distance between the uORF termination site and the AUG of the major ORF (Peabody and Berg, 1986; Kozak, 1987; Miller and Hinnebusch, 1989; Grant et al., 1994); (*3*) the sequence context of the uORF termination site (Miller and Hinnebusch, 1989); (*4*) the length of the uORF (Kozak, 1987; Werner et al., 1987; Abastado et al., 1991); and (*5*) the sequence of the uORF encoded peptide (Geballe and Morris, 1994). To explain the effects of uORFs on translation of the major ORF, it has been proposed that ribosomes (or the 40S subunit) remain associated with the mRNA after translating the uORF and reinitiate translation at a downstream AUG. Hence, the act of translating the uORFs has been postulated to be critical in a regulatory mechanism involving reinitiation while the sequence of the uORFs has little or no importance (Williams et al., 1988).

In contrast, our data suggest that the inhibitory effect of uORF4 on translation of the downstream major uORF depended on the uORF4 coding sequence. Changes in the nucleotide sequence which preserve the uORF4 amino acid coding sequence (in M3*) did not alter the lacZ expression pattern, while changes in the uORF4 amino acid coding information (in M4*) or mutation of the uORF4 AUG (in M4) resulted in a similarly altered regulation of lacZ expression in hearts and brains. This finding indicates that the translation of uORF4 is not the significant event. Our data are most easily explained by the assumption that the peptide encoded by uORF4 has an inhibitory function. Interestingly, we have previously observed that extremely high levels of expression in the heart can be achieved by mutating the stop codon of uORF4 (construct M45; Zimmer et al., 1994). Since uORF4 and the major ORF are in the same reading frame, this mutation may yield a larger fusion protein that initiates at the uORF4 AUG. It is conceivable that this fusion protein accounts for the high levels of lacZ activity in the hearts of TG-M45 embryos which would, in turn, indicate that uORF4 is efficiently translated in the heart. Alternatively, the extension of uORF4 might inactivate it and thus lead to a higher translation of the downstream major uORF.

Interestingly, it has been shown that the translation of the S-Adenosylmethionine Decarboxylase (AdoMetDC) mRNA, the human cytomegolovirus gpUL48 mRNA, and the yeast carbamoyl-phosphate (CPA1) mRNA also depend on the amino acid coding sequence of small uORFs in the 5′-leader sequence of those mRNA (Werner et al., 1987; Schleiss et al., 1991; Hill and Morris, 1992; Degnin et al., 1993; Hill and Morris, 1993). Mis-sense mutations in these uORFs, as well as their deletion, resulted in altered translation of the downstream major ORF. The molecular basis of the translational regulation, however, remains to be determined.

There is growing evidence that translation of mRNA is an important step in the regulation of many specific eukaryotic genes (Geballe and Morris, 1994). In particular, mRNAs that encode for retinoic acid receptors, homeobox genes, proto-oncogenes, growth factors, and other proteins with an important role in embryonic development, often contain small uORFs (Kessel and Gruss, 1988; Kozak, 1991*b*). This distinguishing feature suggests that translational regulation of gene expression may be common among those genes.

We thank Dr. Mike Brownstein for support and oligonucleotide synthesis, Monika König for help with the Northern blot, Rick Dreyfuss for excellent microphotography, and our colleagues for their valuable comments on the manuscript.

Received for publication 8 April 1996 and in revised form 31 May 1996.

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. Derynick. 1991. Inhibition of translation of transforming growth factor-β 3 mRNA by its 5′ untranslated region. *Mol. Cell. Biol.* 11:4306–4313.

Atwater, J.A., R. Wisdom, and I.M. Verma. 1990. Regulated mRNA stability. *Annu. Rev. Genet.* 24:519–541.

Degnin, C.R., M.R. Schleiss, J. Cao, and A.P. Geballe. 1993. Translational inhibition mediated by a short upstream open reading frame in the human cytomegalovirus gpUL4 (gp48) transcript. *J. Virol.* 67:5514–5521.

Donzé, O., and P.-F. Spahr. 1992. Role of the open reading frames of Rous sarcoma virus leader RNA in translation and genome packaging. *EMBO (Eur. Mol. Biol. Organ.) J.* 11:3747–3757.

Dyson, E., H.M. Sucov, S.W. Kubalak, G.W. Schmid-Schönbein, F.A. DeLano, R.M. Evans, J.J. Ross, and K.R. Chien. 1995. Atrial-like phenotype is associated with embryonic ventricular failure in retinoid X receptor α −/− mice. *Proc. Natl. Acad. Sci. USA* 92:7386–7390.

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.

GeBalle. A.P., and D.R. Morris. 1994. Initiation codons within 5′-leaders of mRNAs as regulators of translation. *Trends Biochem. Sci.* 19:159–164.

Grant, C.M., P.F. Miller, and A.G. Hinnebusch. 1994. Requirements for intercistronic distance and level of eukaryotic initiation factor 2 activity in reinitiation on GCN4 mRNA vary with the downstream cistron. *Mol. Cell. Biol.* 14:2616–2628.

Hart, R.C., K.J. Winn, and E.R. Unger. 1992. Avian model for 13-*cis*-retinoic acid embryopathy: morphological characterization of ventricular septal defects. *Teratology.* 46:533–539.

Hill, J.R., and D.R. Morris. 1992. Cell-specific translation of S-adenosylmethionine decarboxylase mRNA. Regulation by the 5′ transcript leader. *J. Biol. Chem.* 267:21886–21893.

Hill, J.R., and D.R. Morris. 1993. Cell-specific transational regulation of S-adenosylmethionine decarboxylase mRNA. Dependence on translation and coding capacity of the *cis*-acting upstream open reading frame. *J. Biol. Chem.* 268:726–731.

Kastner, P., J.M. Grondona, M. Mark, A. Gansmuller, M. LeMeur, d. Decimo, J.L. Vonesch, P. Dolle, and P. Chambon. 1994. Genetic analysis of RXR α developmental function: convergence of RXR and RAR signaling pathways in heart and eye morphogenesis. *Cell.* 78:987–1003.

Kessel, M. 1992. Respecification of vertebral identities by retinoic acid. *Development.* 115:487–501.

Kessel, M., and P. Gruss. 1988. Open reading frames and translational control. *Nature (Lond.).* 332:117–118.

Kessel, M., and P. Gruss. 1991. Homeotic transformation of murine vertebrae and concomitant alteration of Hox codes induced by retinoic acid. *Cell.* 67: 89–104.

Kochhar, D.M. 1967. Teratogenic activity of retinoic acid. *Acta path. et microbiol. scandinav.* 70:398–404.

Kozak, M. 1984. Selection of initiation sites by eucaryotic ribosomes: effect of inserting AUG triplets upstream from the coding sequence for preproinsulin. *Nucleic Acids Res.* 12:3873–3893.

Kozak, M. 1986. Influence of mRNA secondary structure on initiation by eukaryotic ribosomes. *Prc. Natl. Acad. Sci. USA.* 83:2850–2854.

Kozak, M. 1987. Effects of intercistronic length on the efficiency of reinitiation by eucaryotic ribosomes. *Mol. Cell. Biol.* 7:3438–3445.

Kozak, M. 1989*a*. The scanning model for translation: an update. *J. Cell Biol.* 108:229–241.

Kozak, M. 1989*b*. Circumstances and mechanisms of inhibition of translation by secondary structure in eucaryotic mRNAs. *Mol. Cell. Biol.* 9:5134–5142.

Kozak, M. 1991*a*. Structural features in eukaryotic mRNAs that modulate the initiation of translation. *J. Biol. Chem.* 266:19867–19870.

Kozak, M. 1991*b*. An analysis of vertebrate mRNA sequences: intimations of translational control. *J. Cell Biol.* 115:887–903.

Mendelsohn, C., D. Lohnes, D. Decimo, T. Lufkin, M. LeMeur, P. Chambon, and M. Mark. 1994. Function of the retinoic acid receptors (RARs) during development (II). Multiple abnormalities at various stages of organogenesis in RAR double mutants. *Development.* 120:2749–2771.

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.

Molkentin, J.D., D.V. Kalvakolanu, and B.E. Markham. 1994. Transcription factor GATA-4 regulates cardiac muscle-specific expression of the α-myosin heavy-chain gene. *Mol. Cell. Biol.* 14:4947/-4957.

Mueller, P.P., and A.G. Hinnebusch. 1986. Multiple upstream AUG codons mediate translational control of GCN4. *Cell.* 45:201–207.

Peabody, D.S., and P. Berg. 1986. Termination-reinitiation occurs in the translation of mammalian cell mRNAs. *Mol. Cell. Biol.* 6:2695–2703.

Pelletier, J., and N. Sonenberg. 1985. Insertion mutagenesis to increase secondary structure within the 5′ noncoding region of a eukaryotic mRNA reduces translational efficiency. *Cell.* 40:515–526.

Reynolds, K., E. Mezey, and A. Zimmer. 1991. Activity of the β-retinoic acid receptor promoter in transgenic mice. *Mech. Dev.* 36:15–29.

Schleiss, M.R., C.R. Degnin, and A.P. Geballe. 1991. Translational control of human cytomegalovirus gp48 expression. *J. Virol.* 65:6782–6789.

Sucov, H.M., E. Dyson, C.L. Gumeringer, J. Price, K.R. Chien, and R.M. Evans. 1994. RXR alpha mutant mice establish a genetic basis for vitamin A signaling in heart morphogenesis. *Genes Dev.* 8:1007–1018.

Twal, W., L. Roze, and M.H. Zile. 1995. Anti-retinoic acid monoclonal antibody localizes all-trans-retinoic acid in target cells and blocks normal development in early quail embryo. *Dev. Biol.* 168:225–234.

Werner, M., A. Feller, F. Messenguy, and A. Pierard. 1987. The leader peptide of yeast gene CPA1 is essential for the translational repression of its expression. *Cell.* 49:805–813.

Wiens, D.J., T.K. Mann, D.E. Fedderson, W.K. Rathmell, and B.H. Franck. 1992. Early heart development in the chick embryo: effects of isotretinoin on cell proliferation, α-actin synthesis, and development of contractions. *Differentiation.* 51:105–112.

Williams, N.P., P.P. Mueller, and A.G. Hinnebusch. 1988. The positive regulatory function of the 5′-proximal open reading frames in GCN4 mRNA can be mimicked by heterologous, short coding sequences. *Mol. Cell. Biol.* 8: 3827–3836.

Wu, J., M. Garami, L. Cao, Q. Li, and D.G. Gardner. 1995. 1,25(OH)2D3 suppresses expression and secretion of atrial natriuretic peptide from cardiac myocytes. *Am. J. Physiol.* 268:E1108–1113.

Yutzey, K.E., J.T. Rhee, and D. Bader. 1994. Expression of the atrial-specific myosin heavy chain AMHC1 and the establishment of anteroposterior polarity in the developing chicken heart. *Development.* 120:871–883.

Zelent, A., C. Mendelsohn, P. Kastner, A. Krust, J.M. Garnier, F. Ruffenbach, P. Leroy, and P. Chambon. 1991. Differentially expressed isoforms of the mouse retinoic acid receptor β generated by usage of two promoters and alternative splicing. *EMBO (Eur. Mol. Biol. Organ.) J.* 10:71–81.

Zhou, M.D., H.M. Sucov, R.M. Evans, and K.R. Chien. 1995. Retinoid-dependent pathways suppress myocardial cell hypertrophy. *Proc. Natl. Acad. Sci. USA.* 92:7391–7295.

Zimmer, A., and A. Zimmer. 1992. Induction of a RARβ2-lacZ transgene by retinoic acid reflects the neuromeric organization of the central nervous system. *Development.* 116:977–983.

Zimmer, A., A.M. Zimmer, and K. Reynolds. 1994. Tissue specific expression of the retinoic acid receptor-β2: regulation by short open reading frames in the 5′-noncoding region. *J. Cell Biol.* 127:1111–1119.

Zuker, M. 1989. Computer prediction of RNA structure. *Methods Enzymol.* 180:262–288.