MOLECULAR AND CELLULAR BIOLOGY, Sept. 2004, p. 7622–7635 Vol. 24, No. 17
0270-7306/04/$08.00+0 DOI: 10.1128/MCB.24.17.7622–7635.2004

Internal Ribosome Entry Site Structural Motifs Conserved among Mammalian Fibroblast Growth Factor 1 Alternatively Spliced mRNAs

Yvan Martineau,[1] Christine Le Bec,[2] Laurent Monbrun,[1] Valérie Allo,[2] Ing-Ming Chiu,[3] Olivier Danos,[2] Hervé Moine,[4] Hervé Prats,[1] and Anne-Catherine Prats[1]*

Institut National de la Santé et de la Recherche Médicale U589, Hormones, Facteurs de Croissance et Physiopathologie Vasculaire, Institut Louis Bugnard, IFR31, CHU Rangueil, Toulouse,[1] Genethon, CNRS UMR 8115, Evry,[2] and UPR 9002 CNRS, Institut de Biologie Moléculaire et Cellulaire, Strasbourg,[4] France, and Department of Internal Medicine, The Ohio State University, Columbus, Ohio[3]

Received 8 December 2003/Returned for modification 5 January 2004/Accepted 25 May 2004

Fibroblast growth factor 1 (FGF-1) is a powerful angiogenic factor whose gene structure contains four promoters, giving rise to a process of alternative splicing resulting in four mRNAs with alternative 5′ untranslated regions (5′ UTRs). Here we have identified, by using double luciferase bicistronic vectors, the presence of internal ribosome entry sites (IRESs) in the human FGF-1 5′ UTRs, particularly in leaders A and C, with distinct activities in mammalian cells. DNA electrotransfer in mouse muscle revealed that the IRES present in the FGF-1 leader A has a high activity in vivo. We have developed a new regulatable TET OFF bicistronic system, which allowed us to rule out the possibility of any cryptic promoter in the FGF-1 leaders. FGF-1 IRESs A and C, which were mapped in fragments of 118 and 103 nucleotides, respectively, are flexible in regard to the position of the initiation codon, making them interesting from a biotechnological point of view. Furthermore, we show that FGF-1 IRESs A of murine and human origins show similar IRES activity profiles. Enzymatic and chemical probing of the FGF-1 IRES A RNA revealed a structural domain conserved among mammals at both the nucleotide sequence and RNA structure levels. The functional role of this structural motif has been demonstrated by point mutagenesis, including compensatory mutations. These data favor an important role of IRESs in the control of FGF-1 expression and provide a new IRES structural motif that could help IRES prediction in 5′ UTR databases.

During the last decade, the discovery of posttranscriptional gene complexity and of the coupling of transcriptional with posttranscriptional processes has revealed the importance of different levels of regulation in the control of gene expression. Furthermore, in regard to the level of mRNA translation, two main mechanisms are used for translation initiation (for a review, see reference 21). Most eukaryotic mRNAs are translated according to the cap-dependent ribosome scanning mechanism, whereas a smaller group of mRNAs, mostly coding for control proteins (such as growth factors, transcription factors, etc.), can be translated by internal ribosome entry, a cap-independent mechanism mediated by an RNA element present in the 5′ untranslated region (5′ UTR) of the mRNA, the internal ribosome entry site (IRES). In a few cases the mRNA coding region is involved (25, 26, 47). The discovery of IRESs in cellular capped mRNAs raised an interesting question that has been addressed by scientists for a few years: why should a cellular capped mRNA be translated by a cap-independent mechanism? Although the function of such RNA elements is far from being elucidated, several studies have reported that IRESs allow translation of specific cellular mRNAs to occur under conditions where cap-dependent translation is blocked, such as stress conditions (hypoxia, apoptosis, or amino acid deprivation) or during the M phase of the cell cycle (27). In addition, IRESs are highly tissue specific and developmentally regulated, at least in the cases of fibroblast growth factor 2 (FGF-2) and c-*myc*, whereas abnormal IRES activation is, for both genes, related to cell transformation (8, 13, 14, 17, 53).

Another function of IRESs is to participate in the control of alternative initiation of translation. This is obvious in the case of vascular endothelial growth factor (VEGF), where IRESs A and B control the use of a CUG and an AUG start codon, respectively, as well as in the cases of FGF-2 and Moloney murine leukemia virus mRNAs, where the IRESs are located between two alternative start codons (2, 23, 49, 50). In these three cases, IRES-dependent translation has, indirectly, important consequences for gene function, because it modulates the relative expression of protein isoforms having distinct intra- or extracellular localization and functions (6, 12, 45, 46). Recently it has been shown that a mutation in VEGF IRES A is correlated with a higher risk of amyotrophic lateral sclerosis, providing a first link between IRES deregulation and neurodegenerative disease in humans (32).

IRESs have been described mainly as RNA elements with secondary or tertiary structures allowing internal ribosome recruitment. The first RNA crystal structure has been obtained for the hepatitis C virus IRES (30). In other studies, IRES structures have been analyzed by chemical and enzymatic protection, revealing defined structural domains, including pseudoknots in picornaviruses, cricket paralysis-like virus, hepatitis C virus, and c-*myc* IRESs, whereas a G quartet structure has

* Corresponding author. Mailing address: Institut National de la Santé et de la Recherche Médicale U589, Hormones, Facteurs de Croissance et Physiopathologie Vasculaire, Institut Louis Bugnard, IFR31, CHU Rangueil, 31059 Toulouse Cedex 09, France. Phone: 33 (5) 61 32 21 42. Fax: 33 (5) 61 32 21 41. E-mail: pratsac@toulouse.inserm.fr.

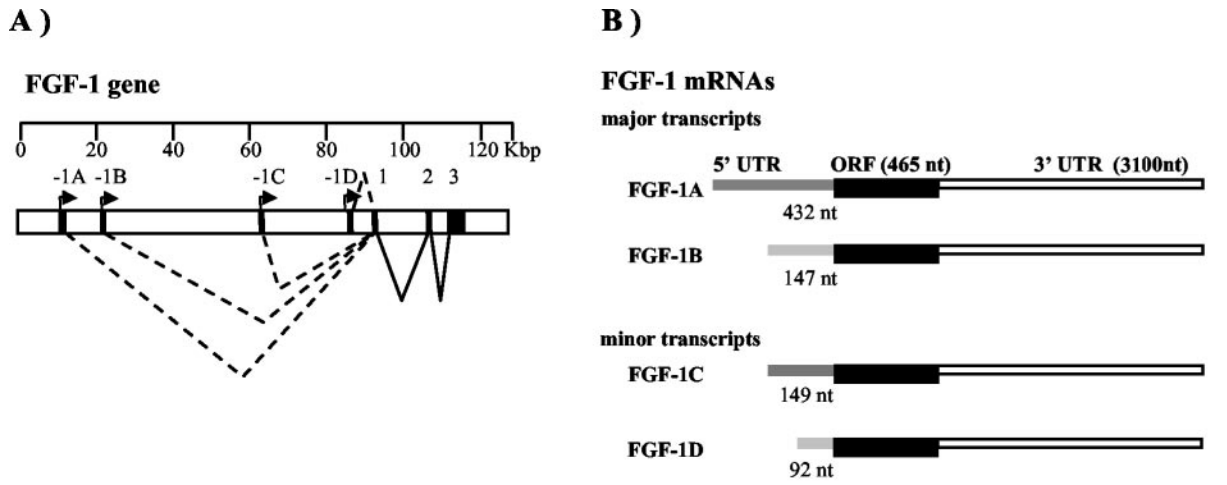


FIG. 1. Human FGF-1 gene and transcripts. (A) Organization of the human FGF-1 gene. Arrows represent the promoters 1A, 1B, 1C, and 1D. Alternative splicing of untranslated exons 1A, 1B, 1C, or 1D to exon 1 (depending on the promoter used) will generate mRNAs 1A, 1B, 1C, and 1D. (B) Structures of the FGF-1 mRNAs. mRNAs 1A and 1B are the most common and are expressed mainly in kidney brain and retina. mRNAs 1C and 1D are expressed at a very low level and can be induced by serum or transforming growth factor β (10).

been recently described for the FGF-2 IRES (4, 29, 33, 35). Several IRESs are able to recruit ribosome without the involvement of specific IRES *trans*-acting factors (ITAFs), whereas others need ITAF binding to be active (24, 43). ITAFs act as chaperones whose binding to the mRNA stabilizes the IRES in the correct conformation to be functional (37).

Although the existence of structural motifs in IRESs has been clearly demonstrated, up to now it has remained impossible to predict an IRES element by searching in the sequence data banks, except in the case of short RNA sequences, such as the Gtx 9-nucleotide (nt) motif, which can function as IRESs by allowing rRNA base pairing (7). In contrast to these RNA sequence motifs that can be easily searched for in the sequence libraries, IRES RNA structural motifs are difficult to predict. Furthermore, they show a large heterogeneity from one mRNA to another. In the present study, our objective was to identify, by a functional approach, IRES motifs conserved in a gene family and/or between mammalian species, with the purpose of using them for further IRES searching in silico, possibly leading to the discovery of mRNA families with similar IRES motifs which would be regulated in a coordinated manner.

FGFs make up a large family of 22 members that are conserved in vertebrates and are involved in the control of cell proliferation and differentiation, in development, and, for several of them, in angiogenesis (40). FGF-1 and FGF-2 are prototypes of this heparin binding growth factor family and share, in addition to their coding sequence homology, the feature of being major angiogenesis factors. However, the FGF-1 and FGF-2 gene structures are completely different: the FGF-2 gene is composed of a single transcription unit expressing five protein isoforms from alternative translation initiation codons, whereas the FGF-1 gene, expressing a single protein isoform, has four alternative tissue-specific promoters and is subjected to a process of alternative splicing that is conserved among mammals (2, 10, 36, 39, 46) (Fig. 1). Thus, each promoter is coupled with its 5′ untranslated exon, giving rise to four mRNAs with distinct UTRs, containing exons 1A, 1B, 1C, and 1D, respectively. The cell-specific regulation of these FGF-1 transcripts and subsequent protein synthesis is unique for each of them. Transcripts 1A and 1B are specific to maintenance and survival of cells, particularly cardiac and neuronal cells, whereas transcripts 1C and 1D are expressed only under specific conditions, such as serum, phorbol ester, or hormone stimulation. This suggests that each transcript may have a distinct role in development, normal cellular processes, and, upon aberrant regulation, disease (11, 41).

The presence in FGF-1 mRNAs of the four different UTRs A, B, C, and D, measuring 432, 147, 149, and 92 nt, respectively, and being strikingly conserved among mammals, suggests that translational regulation of FGF-1 expression should occur in addition to, and may be coupled with, the transcriptional regulation. Furthermore, IRES-dependent regulation of translation initiation has been described for the other main angiogenic factors, i.e., FGF-2, VEGF, and platelet-derived growth factor (PDGF), prompting us to search for IRESs in the four FGF-1 leaders (3, 23, 46, 49). In addition, the unusual conservation of such UTRs among mammals makes the FGF-1 gene a system of choice to identify conserved RNA structural motifs involved in translational control.

In this study, we identified IRESs in the alternative FGF-1 leaders, particularly in leaders A and C, by using a double luciferase bicistronic vector strategy. A new bicistronic regulatable TET OFF system was developed, allowing us to rule out the possibility of any cryptic promoter in the alternatives exons. In addition, the FGF-1 IRESs A of murine and human origins showed similar IRES activity profiles. We focused on IRES A to detect evolutionarily conserved IRES structural motifs. Enzymatic and chemical probing of the FGF-1 IRES A structure, as well as site-directed mutagenesis, revealed a new IRES structural domain conserved among mammals.

MATERIALS AND METHODS

Cell culture. Cell lines were obtained from the American Type Culture Collection (ATCC) and the European Collection of Cell Cultures (ECACC) and were cultivated according to the instructions of the suppliers.

CHO-K1 (ATCC CCL-61) is a Chinese hamster ovary carcinoma. COS-7

(ATCC CRL 1654) is a monkey kidney cell line transformed by simian virus 40 large T antigen. SK-Hep-1 (ATCC HTB52) is a human liver adenocarcinoma of endothelial origin. MEF-3T3 Tet-off (BD Clontech catalog no. C3018-1) is a mouse immortalized fibroblast cell line stably expressing the tetracycline-controlled transactivator (tTA-VP16). The 911 cell line (ECACC 95062101) is derived from human embryonic retinoblasts transformed by adenovirus type 5 DNA. doxycycline (Sigma catalog no. D9891) was added to culture medium to a final concentration of 10 nM 1 or 2 h before transfection.

Cell transfection. The different cell types were transfected with 1 μg of plasmid and JetPEI (Q-BIOgene) in 12-well tissue culture dishes. At 48 h after transfection, cell lysates were prepared for luminescence activity as previously described (23). LucR and LucF activities were measured by using the dual luciferase kit from Promega and a Berthold LB96V luminometer.

Cellular RNA purification. Total cellular RNA was prepared by the RNAble method (Eurobio), which is derived from the guanidinium thiocyanate procedure. A total of 5×10^6 cells were lysed on the dish in 1 ml of RNAble. RNA was extracted after addition of 250 μl of chloroform and precipitated with isopropanol. After ethanol washing and precipitation, RNA was quantified by measuring the absorbance at 260 nm and checked for integrity by electrophoresis on an agarose gel and ethidium bromide staining.

Reverse transcription and PCR analysis. The cDNAs were synthesized with a Superscript II kit (Invitrogen) by using 3 μg of total RNA and random hexamer nucleotides for human mRNA and FGF-1 3′ UTR-specific oligonucleotides for mouse mRNA. The PCRs were carried out with 0.5 U of *Taq* Platinum DNA polymerase (Invitrogen) in a final volume of 50 μl with 3 μl of cDNA (oligonucleotide sequences are available upon request). The reaction was performed on a Perkin-Elmer apparatus under the following conditions: 94°C for 5 min, then three cycles of 94°C for 30 s and 74°C for 2 min, then 25 cycles of 94°C for 30 s and 72°C for 2 min, and finally 72°C for 5 min. Amplification products were observed in SK-HEP-1 cells for all human FGF-1 5′ UTRs and in C2C12 cells for the mouse FGF-1A 5′ UTR.

Plasmid construction. Amplification products were cut with SpeI and NcoI and inserted in double luciferase plasmids by using a trimolecular strategy. The LucF fragment (NcoI-BglII) and the cytomegalovirus (CMV)-LucR fragment (SpeI-BglII) were purified from pCREL and pCRFL, respectively. A hairpin structure has been introduced after the LucR stop codon by replacing the CMV-LucR fragment by the CMV-LucR-Hairpin fragment (SpeI-BglII) from the pCRHL vector. All vectors have been described previously (14).

The TRE-CMV promoter fragment is derived from plasmid pUHD 10-3 described by Gossen and Bujard (18). XhoI and MluI sites were added by PCR to 5′ and 3′ extremities, respectively. A bimolecular strategy was used to insert this fragment into the double luciferase vector (SalI-MluI). All mutants described in Fig. 5 were created by PCR (oligonucleotides sequences are available upon request) and by the trimolecular strategy described above. The double luciferase vectors containing PDGF 5′ UTR sequence were derived from pCPL, described by Bernstein et al. (3). The EF-1α promoter was subcloned from pTracer-EF/Bsd/lacZ (Invitrogen) into pCRHL by removing the hairpin structure. In the final construct, EF-1α forward priming sites are located between 119 and 137 nt upstream from the NcoI site (ATG of LucF). All constructs were sequenced by the use of forward and reverse oligonucleotides located in LucR and LucF, respectively. The five point mutations in the DII domain (see Fig. 9) were created by using the QuikChange XL site-directed mutagenesis kit (Stratagene) with oligonucleotides designed according to the manufacturer's instructions.

For RNA probing and Northern blotting experiments, pKSA5 and pKSA9 were constructed. Double luciferase A5 and A9 deletion mutants were opened with SpeI, and filled with Klenow, purified, and then cut again with EcoRI, located in LucF, 732 bp downstream from the NcoI site (ATG). pKS+ DNA was cut by Ecl136II and EcoRI, purified, and ligated with the SpeI-Klenow/EcoRI fragments mentioned above, resulting in pKSA5 and pKSA9.

DNA plasmid electrotransfer in mouse skeletal muscle. For in vivo experiments, DNA plasmids were purified on Qiagen endotoxin-free columns. Naked DNA (50 μg in 30 μl in 0.9% NaCl) was injected and electrotransferred into the tibialis anterior muscle of mice (7-week-old female BALB/c). A group of three mice per plasmid was used, and each experiment was repeated at least three times. One week later, mice were sacrificed and muscle extracts were prepared for determination of luminescence activity. Conditions for intramuscular injections, electrotransfers, and preparation of muscle extracts have been previously reported (42). Luciferase activities were assayed in 20 μl of muscle lysate with 50 μl of reagent (dual luciferase kit; Promega) and a luminometer (Victor2; Perkin-Elmer).

Northern blotting. Northern blotting was performed with a NorthernMax kit (Ambion catalog no. 1940) according to manufacturer's instructions. Fifteen micrograms of total RNA per well was loaded on 0.8% agarose denaturing gels. Passive transfer was performed on Hybond N nitrocellulose membranes.

A complementary-strand LucF RNA probe was generated in vitro with MAXIscript T3 RNA polymerase (Ambion catalog no. 1324) according to the manufacturer's instructions in the presence of 50 μCi of [α-^{32}P]CTP with 1 μg of NcoI-linearized pKSA5 or pKSA9.

Structural probing reactions on in vitro-synthesized RNA. Large amounts of RNAs were generated with MEGAscript T7 (Ambion catalog no. 1334) according to the manufacturer's instructions with NcoI-linearized pKSA5 or pKSA9. In vitro-transcribed RNA (4 pmol) was denatured for 1 min at 95°C and chilled on ice for 5 min. Incubation was continued for 10 min at 25°C after addition of the various reaction buffers (10 mM HEPES-Na [pH 7.5], 50 mM KCl, and 2.5 mM $MgCl_2$ for dimethyl sulfate [DMS]; RNases T_1 [Invitrogen], T_2 [Invitrogen], and V_1 [Pierce]; 50 mM sodium borate [pH 8.0], 50 mM KCl, and 2.5 mM $MgCl_2$ for CMCT). Chemical modifications with DMS and CMCT were done for 5 and 20 min, respectively, in a 25-μl final volume as described previously (5). Enzymatic digestions were done with 0.5 Units of RNase T_1, with 0.05 U of RNase T_2, or with 0.005 U of RNase V_1 for 10 min at 25°C. For primer extensions, three different primers (RTA5, 5′-TTCTACCAGGACAGCTGCAGGC-3′; RTA9, 5′-TCTGGGGTGCTTTCTTTCTGGG-3′; and RTEX1, 5′-CAGCAGCTGCT GCTTGTGGCGC-3′) were labeled with [γ-^{32}P]ATP and combined with 2 pmol of synthetic RNA. Hybridization and reverse transcription with the murine leukemia virus reverse transcriptase were carried out as described previously (5). The samples were analyzed on 8% polyacrylamide sequencing gels containing 8 M urea.

RESULTS

IRESs are present in the human FGF-1 mRNA leaders. In order to identify putative IRESs in one or more of the four FGF-1 mRNA leader sequences, we used the bicistronic vector strategy that was previously successful in demonstrating the existence of the two VEGF IRESs (23). The bicistronic vector expresses two highly sensitive luciferases, LucR and LucF, encoded by the same mRNA under the control of the CMV promoter. The human FGF-1 mRNA leader A, B, C, and D sequences, as well as the exon 1 36-nt-long sequence common to the four leaders (Fig. 1), were introduced between the LucR and LucF genes (Fig. 2A). LucR, the first cistron, provides the level of cap-dependent translation. It is expected to reflect the activity of CMV sequences in the tissues or cells and thus to be proportional to the amount of bicistronic mRNA. LucF, the second cistron, is expressed proportionally to the IRES activity. In order to rule out the phenomenon of reinitiation, a series of vectors with a hairpin upstream from the 5′ UTR sequence, at the end of the first cistron, was constructed. Our negative (IRES-minus) control was a plasmid with only a hairpin between a couple of cistrons (pCRHL) (23). The use of such an assay, based on two highly sensitive reporters, allows rapid and concomitant quantitation of both reporters in transfected cells or tissue extracts.

Bicistronic constructs containing either one of the FGF-1 leaders or exon 1 or the FGF-2 IRES were used for transient transfection of MEF-3T3 cells, mouse immortalized fibroblast cells expressing the tetracycline transactivator tTA (Fig. 2B). This cell type was chosen because it also allowed us to develop a new system to detect cryptic promoters (Fig. 2). The IRES activity of each intercistronic region was measured by the LucF/LucR ratio and was calibrated by using the negative control pCRHL.

In MEF-3T3 cells, the four leaders had a significant LucF/LucR ratio (>6 arbitrary units [AU]) in the absence of the hairpin. Although the ratio decreased in the presence of the hairpin, the activity remained greater than background, at

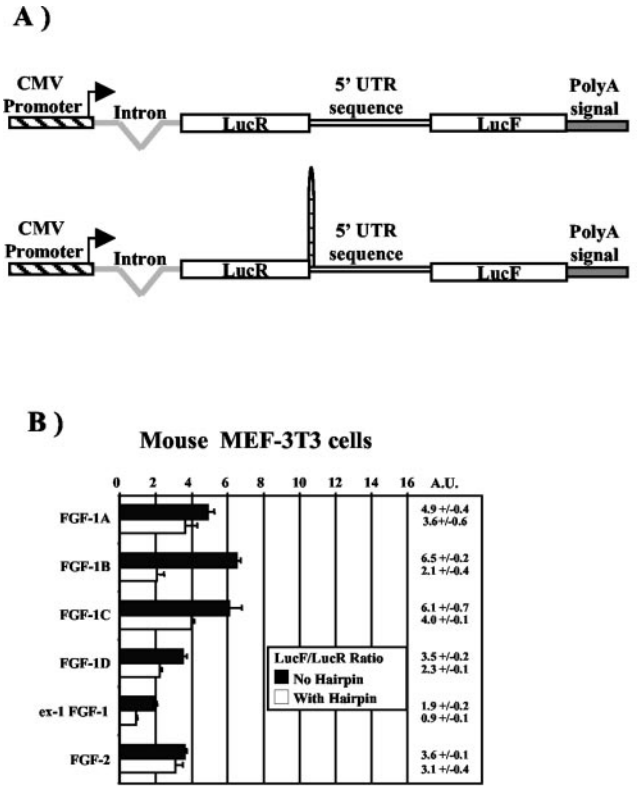


FIG. 2. Expression of FGF-1 leader-containing bicistronic vectors in transfected mammalian cell types. (A) Schematic representation of the bicistronic LucR-I-LucF vectors containing 1A, 1B, 1C, or 1D FGF-1 mRNA 5′ UTRs. Vector construction is described in Materials and Methods. Constructs contain the CMV promoter controlling the expression of a bicistronic LucR-I-LucF mRNA. A synthetic intron is present upstream from LucR, and a poly(A) site is present downstream from LucF (23). In the bottom construct, a hairpin has been introduced at the end of the luciferase renilla gene in order to prevent reinitiation and interaction between the FGF-1 5′ UTRs and the beginning of the bicistronic mRNA. (B) Transient transfection of murine embryo fibroblast MEF-3T3 cells with bicistronic LucR-I-LucF vectors containing either the 1A, 1B, 1C, or 1D FGF-1 5′ UTR in the intercistronic region, with or without a hairpin at the 3′ end of the LucR gene (see Materials and Methods). The bicistronic vectors pCRHL and pCRFL, containing either a hairpin or the FGF-2 IRES, were used as negative and positive controls, respectively (14). Cells were harvested 48 h after transfection, and luciferase activities present in cell extracts were measured. IRES activities were determined by calculating the LucF/LucR ratio. Data were calibrated by dividing the LucF/LucR ratio by the ratio of the pCRHL negative control, providing a relative IRES activity expressed in AU (23). The exact values are presented on the right. Experiments were repeated at least five times. Results represent means ± standard errors (calculated with Microsoft Excel software) from a representative experiment done in triplicate.

about 4 AU for leaders A and C and very low, but greater than 2 AU, for leaders B and D (Fig. 2B). In contrast, exon 1 was at the background level. These data suggested the presence of IRESs in the FGF-1 leaders, more clearly in leaders A and C. The leader A and C activities, although not very strong, were significant and at the same level as that of the FGF-2 IRES in this cell type.

A new TET OFF bicistronic system to check for the absence of internal promoters. The FGF-1 transcription start sites have been extensively studied previously, rendering very improbable the existence of a cryptic promoter in any of the FGF-1 leaders (38). To check such an eventuality, we developed a TET OFF regulatable system in which the bicistronic cassette is under the control of a doxycycline-repressible CMV promoter (see Fig. 3A). In such a construct, the LucF/LucR ratio is expected to be stable independently of the CMV expression level if expression of the second cistron is IRES dependent. In contrast, if there is an intercistronic cryptic promoter, LucF expression will be independent of the CMV promoter, and thus the LucF/LucR ratio will increase proportionally to the repression of CMV promoter by doxycycline. As positive controls, the EF-1α promoter or the PDGFb 5′ UTR (containing a cryptic promoter) was introduced in the intercistronic region.

The different bicistronic cassettes used for Fig. 2 were transferred into the TET OFF construct, and the resulting plasmids were used to transfect MEF-3T3 cells expressing the tetracycline transactivator tTA. Luciferase activities were measured following treatment with 10 nM doxycycline, and we showed that LucR activity, reflecting the mRNA level, was repressed by 5- to 10-fold under such conditions (Fig. 3B). The lower LucR activity observed with the construct containing the EF-1α promoter could result from interference between the two promoters contained in this construct.

The LucF/LucR ratio increased by 20-fold after doxycycline treatment when the EF-1α promoter was present between the two cistrons (Fig. 3C). In contrast, this ratio remained unchanged in the presence of any FGF-1 leader in the intercistronic region (about 5 AU for leaders A, B, and C) (Fig. 3C). To check the sensitivity of the TET OFF bicistronic system, we introduced into the intercistronic region the PDGFb leader, which was recently shown to contain a cryptic promoter (much weaker than EF-1α); in this case the LucF/LucR ratio increased by 34-fold in response to doxycycline, showing that the system is able to detect a weak promoter.

As a control, the integrity of the bicistronic mRNAs was also analyzed by Northern blotting with a LucF riboprobe on total RNAs purified from transfected MEF-3T3 cells in the absence of doxycycline treatment (Fig. 3D). Figure 3D clearly shows the presence of monocistronic mRNA when the intercistronic region contains the EF-1α promoter; in addition, the bicistronic mRNA was hardly detectable with this construct, confirming that the LucR activity was indeed reflecting the bicistronic mRNA level (Fig. 3B) and confirming our hypothesis of promoter interference. In contrast, the LucF riboprobe detected the mRNA at the expected bicistronic size for constructs containing either the PDGFb leader or any of the four FGF-1 leaders in the intercistronic region in all of the cell types tested (Fig. 3D and data not shown). Thus, Northern blotting did not allow detection of the cryptic PDGF promoter under such conditions, whereas the TET OFF bicistronic system clearly detected it. This confirmed that the Northern blot technique is much less sensitive to detect cryptic promoters than this system based on the measurements of enzymatic activities rather than mRNA levels. Furthermore, the TET OFF bicistronic system is more reliable than the promoter-less vector approach: it allows control of the first-cistron LucR level and measurement of LucF/LucR ratio, while the promoterless system does not. This approach could thus be used to check the absence of a promoter in any putative IRES element, with the only limitation of

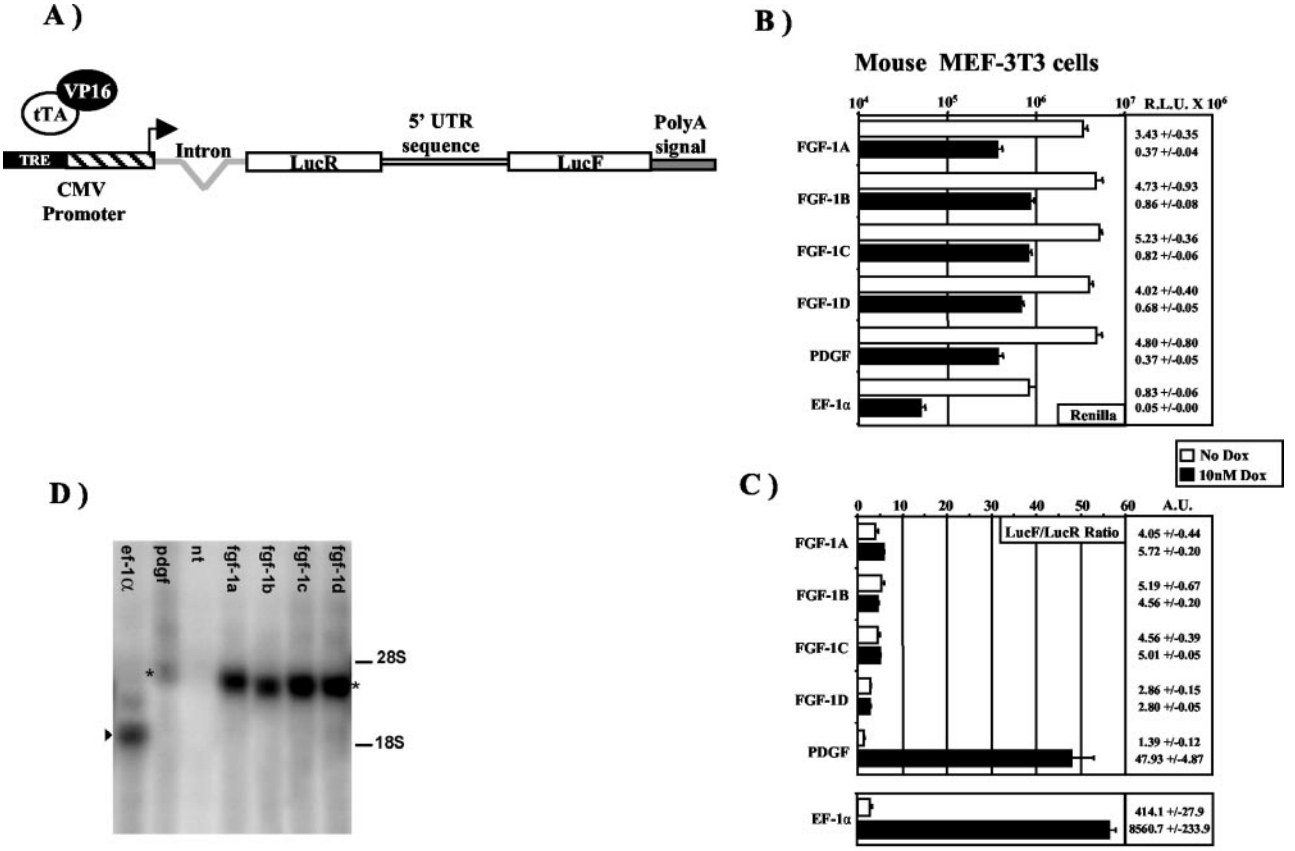


FIG. 3. Development of a TET OFF regulatable system to check for the absence of cryptic promoters in the bicistronic Luc-I-Luc mRNAs. (A) Schematic representation of the bicistronic TET OFF LucR-I-LucF vector. We constructed a bicistronic vector with a promoter repressible by tetracycline, the TET OFF system. A tetracycline-responsive element (TRE) was placed in front of the CMV promoter (devoid of its enhancer). The principle is that in the absence of tetracycline, the chimeric tTA-VP16 protein is bound to the TRE element and thus the CMV promoter is active. In the presence of tetracycline or tetracycline analogs such as doxycycline, tTA will bind the antibiotic and will be unable to bind the TRE, thus preventing CMV promoter transactivation. Bicistronic TRE-CMV vectors were constructed with the different FGF-1 5′ UTRs or a hairpin in the intercistronic region. As positive promoter controls, the EF-1α promoter or the PDGFb 5′ UTR (containing a cryptic promoter) was also introduced in the intercistronic region. (B and C) Transfection of MEF-3T3 cells expressing tTA-VP16 with the TET regulatable bicistronic constructs described for panel A. (B) At 2 h prior to transfection, cells were treated or not with 10 nM doxycycline (Dox). Luciferase activities were measured as described for Fig. 2. LucR activities, expected to reflect the bicistronic mRNA levels, are represented for each construct. (C) The LucF/LucR ratio is expressed in AU measured as described for Fig. 2. A different scale has been used to present the values for the EF-1α-containing vector, whose LucF/LucR ratio is a higher order of magnitude. (D) Analysis of bicistronic mRNAs by Northern blotting. Northern blotting was performed with a LucF ^{32}P-labeled riboprobe with total RNAs purified from MEF-3T3 cells transfected with the bicistronic vectors used for Fig. 2 C, containing either the FGF-1A, -1B, -1C, or -1D or PDGFb 5′ UTR or the EF-1α promoter (see Materials and Methods). For each transfection, the fragment present in the intercistronic region is indicated at the top of the lane. Migrations of 28S (4.7 kb) and 18S (1.2 kb) rRNAs are indicated. Migration of the monocistronic LucF mRNA is indicated by an arrowhead. Migration of the bicistronic mRNAs is indicated by an asterisk.

the system being the requirement for cells expressing the tetracycline transactivator tTA.

In addition, reverse transcription-PCR experiments were performed with the bicistronic mRNAs in order to detect possible splicing events, as recently described for another 5′ UTR (51); the results showed that one band is amplified with the bicistronic mRNAs containing the FGF-1 5′ UTRs, ruling out the hypothesis of splicing (data not shown).

All these data give evidence of bicistronic mRNA integrity, allowing us to conclude that the second-cistron expression of the bicistronic mRNAs is due to IRES activity.

The human FGF-1 mRNA IRESs are selectively active in different cell types. The bicistronic constructs used for Fig. 2 were used for transient transfection of SK-Hep-1 cells (human liver adenocarcinoma), 911 cells (human embryonic retinoblasts transformed by adenovirus type 5 DNA), COS-7 cells (monkey kidney cells transformed by simian virus 40 large T antigen), and CHO-K1 cells (Chinese hamster ovary carcinoma) (Fig. 4). The IRES activity of each intercistronic region was measured as for Fig. 2. The bicistronic mRNA status was checked by Northern blotting as for Fig. 3D (data not shown).

The LucF/LucR ratios were quite similar in SK-Hep-1 and 911 cells: in the absence of the intercistronic hairpin, they were around 9 AU for leaders A and C in the three cell types, similar to that observed for the FGF-2 IRES (Fig. 4A and B). The LucF/LucR ratios for leaders B and D and exon 1 were always between 3 and 5 AU. In the presence of the intercistronic hairpin, leader A had an increased LucF/LucR ratio of 15.2

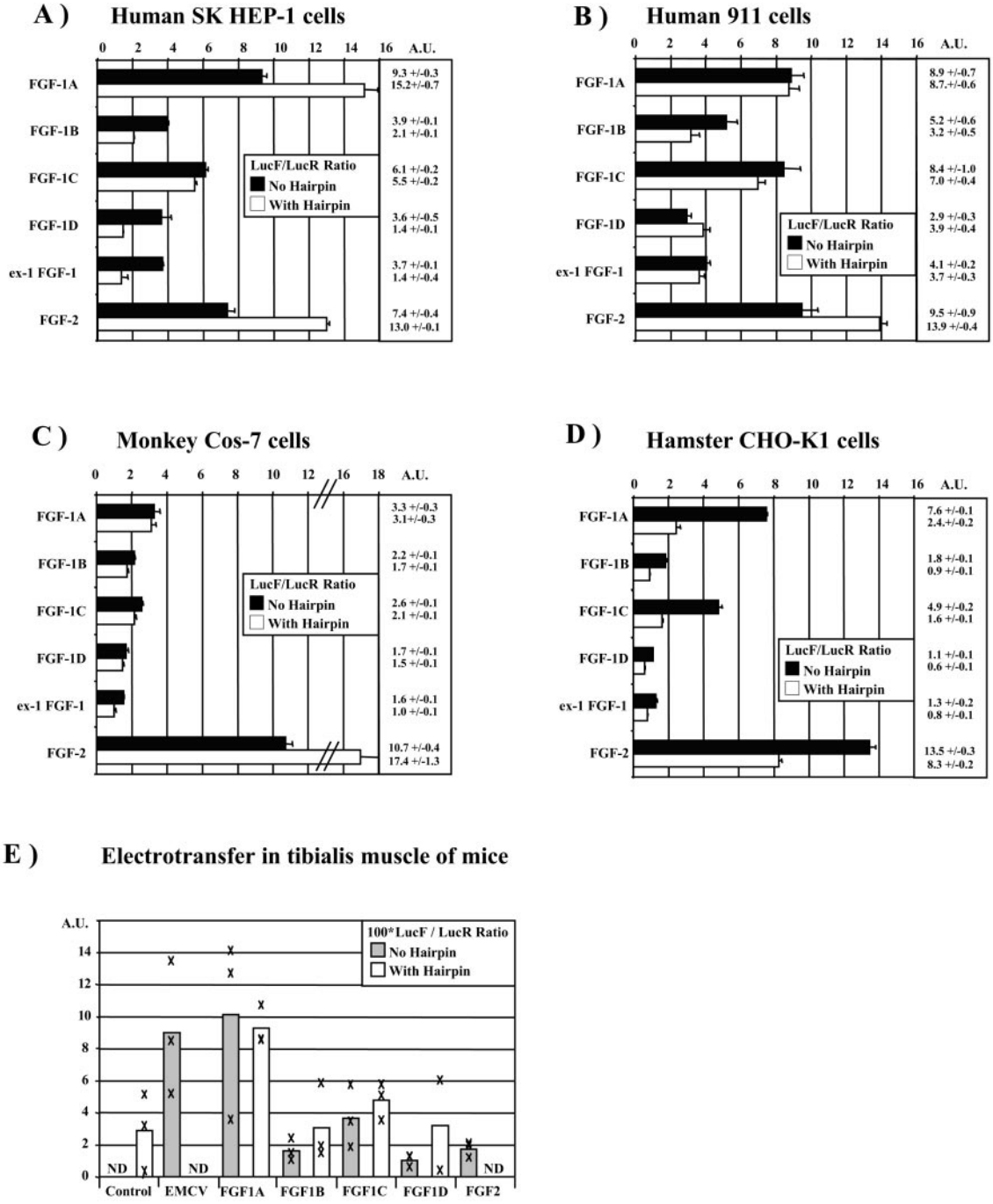


FIG. 4. Analysis of the FGF-1 5′ UTR IRES activities in cells from several mammalian species and in the tibialis anterior muscle of mice. (A to D) Human SK-Hep-1 adenocarcinoma cells (A), human 911 retinoblasts (B), simian kidney COS-7 cells (C), and Chinese hamster ovary CHO-K1 cells (D) were transiently transfected with bicistronic LucR-I-LucF vectors devoid of an intercistronic hairpin, and IRES activities were measured as described for Fig. 2B. Experiments were repeated three to five times. Results represent means ± standard errors from a representative experiment done in triplicate. (E) Expression of FGF-1 leader-containing bicistronic vectors in electrotransferred mouse muscle in vivo. FGF-1 IRES activities were analyzed in vivo by DNA electrotransfer in mouse muscle. The bicistronic LucR-I-LucF vectors used for Fig. 2 were electrotransferred in mouse muscle as described in Materials and Methods. Muscles were extracted, and luciferase activities were measured in muscle lysates. The LucF/LucR ratio is indicated by a cross for each mouse, and the histogram represents the average for each construct. These results are representative of those from experiments that were repeated at least three times for each construct. The negative control corresponds to the construct with an intercistronic hairpin. ND, not determined.

AU in SK-Hep-1 cells but the LucF/LucR ratio remained at about 9 AU in 911 and CHO cells, while the leader C ratio remained stable in the three cell types (Fig. 4A, B, and D). In the presence of the hairpin, FGF-2 IRES activity increased in both cell types. In contrast, leaders B and D and exon 1 had distinct behaviors: in SK-Hep-1 cells, leader B activity decreased to 2 AU, whereas the ratios for leader D and exon 1 fell to 1 AU (like that of the negative control pCRHL). In 911 cells, the ratio remained at 3 to 4 AU for the three intercistronic sequences.

In COS-7 cells, LucF/LucR ratios were low, independent of the presence of the hairpin, for the four leaders as well as exon 1; IRES activity was significant (>3 AU) for leader A only, showing a drastic difference from the FGF-2 IRES activity, which is greater than 10 AU in this cell type (Fig. 4C). Low activities were observed in CHO cells, where the presence of the hairpin led to a striking decrease in the ratio, indicating a greater contribution of reinitiation in this cell type than in the others (Fig. 4D).

These data clearly suggested the presence of IRESs in leaders A and C, which have significant activity in SK-Hep-1, 911, and MEF-3T3 cells (Fig. 2 and 4). The enhancing effect of the hairpin on IRES activity that was sometimes observed for the FGF-1A leader and the FGF-2 IRES will be discussed below. IRES activity was either nonexistent (SK-Hep, COS-7, and CHO cells) or weak in regard to leaders B and D (MEF and 911 cells). In 911 cells, leader B and D IRES activity followed that of the exon 1 common sequence.

These data also indicated that the FGF-1 IRES activity profiles are completely different from that of the FGF-2 IRES, which is more active than FGF-1 IRESs in COS-7 and CHO-K1 cells but not in other cell types (Fig. 4C and D).

The human FGF-1A IRES is high specifically in mouse muscle in vivo. Based on previous data showing that IRES tissue specificities appear more strikingly in vivo than ex vivo (13, 14), we developed a new approach to test IRES activities in vivo in mouse skeletal muscle (Fig. 4E). The different DNA constructs, containing various mRNA leaders (FGF-1, FGF-2, and encephalomyocarditis virus [EMCV]) with or without a hairpin downstream from LucR (Fig. 2A), were electrotransferred into mouse muscle after intramuscular injection, and luciferase activities were measured from muscle extracts (see Materials and Methods); strikingly, FGF-2 IRES activity was hardly greater than the background, as observed for FGF-1 leaders B and D (Fig. 4). In contrast, FGF-1 IRES C showed detectable activity, whereas IRES A had high IRES activity, equivalent to that of the EMCV IRES (whereas it is 10-fold less efficient than EMCV in transiently transfected cells [data not shown]).

These data show that in vivo, the IRES activity profile is distinct from that observed in transfected cells. In particular, FGF-1 IRES A and C activities are drastically different in vivo, whereas they are similar in cell culture. In fact, the FGF-1 A IRES activity is much stronger in vivo than in cell culture, compared to that of the EMCV IRES, suggesting the involvement of one or several ITAFs.

Minimal IRESs A and C are contained in fragments of 118 and 103 nt, respectively. The mapping of FGF-1A IRES was addressed by deletion mutagenesis, using the hairpin-containing constructs to avoid possible reinitiation (Fig. 5A).

On the one hand, constructs with a deletion of the exon 1 sequence (nt 387 to 432, common to the four leaders) maintained 74% of the activity of the non-5′-deleted leader A, suggesting that exon 1 is involved but dispensable (Fig. 5B, compare bar A with bar A2). Further deletions of the 3′ leader (up to nt 342 or 323) maintained activities of 66 and 51% compared to that of the undeleted leader, respectively (bars A19 and A20). In contrast 3′ deletions up to nt 259 and 209 decreased IRES activity to 17 and 18%, respectively.

On the other hand, deletion of the 5′ part of the leader up to position 214 affected the IRES activity very slightly (Fig. 5B, bar A5). 5′ deletion up to position 269 maintained an IRES activity of 34% in the presence or absence of the exon 1 sequence (Fig. 5B, bars A8 and A9), whereas 5′ deletion up to nt 303 resulted in IRES activity of 15% (bars A10 and A11).

Our conclusion was that the minimal FGF-1 IRES A is contained in a 118-nt fragment between nt 269 and 387 (A9). An indispensable element is present between nt 269 and 303, whereas two elements, located between nt 214 and 269 and nt 387 to 432, respectively, show an enhancing effect when present simultaneously (A7). The segment 387 to 432, containing the 36 nt of exon 1 sequence common to the four FGF-1 leaders, has no IRES activity by itself in this cell type (Fig. 5B, Ex-1). The data in Fig. 5B also show that FGF-1 IRES A has a wide starting window rather than a precise translation starting point, as the distance from the IRES to the AUG start codon seems to be flexible.

We also looked for a minimal IRES for IRES C (Fig. 5C). Deletion of the 36-nt segment corresponding to exon 1 increased IRES C activity by 38% (Fig. 5C, bars C and C2). Thus, the 36-nt segment of exon 1 may have enhancing or silencing effects on IRES activity when present in IRES A or C, respectively. Such data probably reflect the involvement of this segment in RNA structure, resulting in IRES stabilization or destabilization. Further deletions revealed that an element located between nt 1 and 59 of leader C is required for IRES C activity (Fig. 5C, bars C3 and C4). In conclusion, IRES C also show flexibility in regard to the initiation codon position, and both IRESs A and C need elements located in the 5′ part rather in the 3′ part of the leader.

The murine and human FGF-1 leaders A share similar IRES activities. To define a putative RNA structural motif responsible for FGF-1 IRES A activity, our first approach was to analyze the IRES activity of the mouse FGF-1 leader A.

The ability of murine leader A to function as an IRES was assessed in three mammalian cell types (SK-Hep-1, COS-7, and HeLa cells) (Fig. 6). The data clearly show that the murine leader A has an IRES activity (slightly greater than that of the human IRES) with a cell specificity profile similar to that of the human leader A but radically different from that of the FGF-2 IRES (Fig. 6).

Alignment of the complete human and murine leaders A revealed 70% homology, which is high for a noncoding region (coding sequence homology is 90%). In addition, several segments are 100% conserved. Interestingly, most of these conserved blocks are located in the A9 minimal IRES (between nt 269 and 387) (see Fig. 8A).

Identification of an evolutionarily conserved IRES structural motif. The RNA structure of the human FGF-1 IRES A was determined by using enzymatic and chemical probes. In

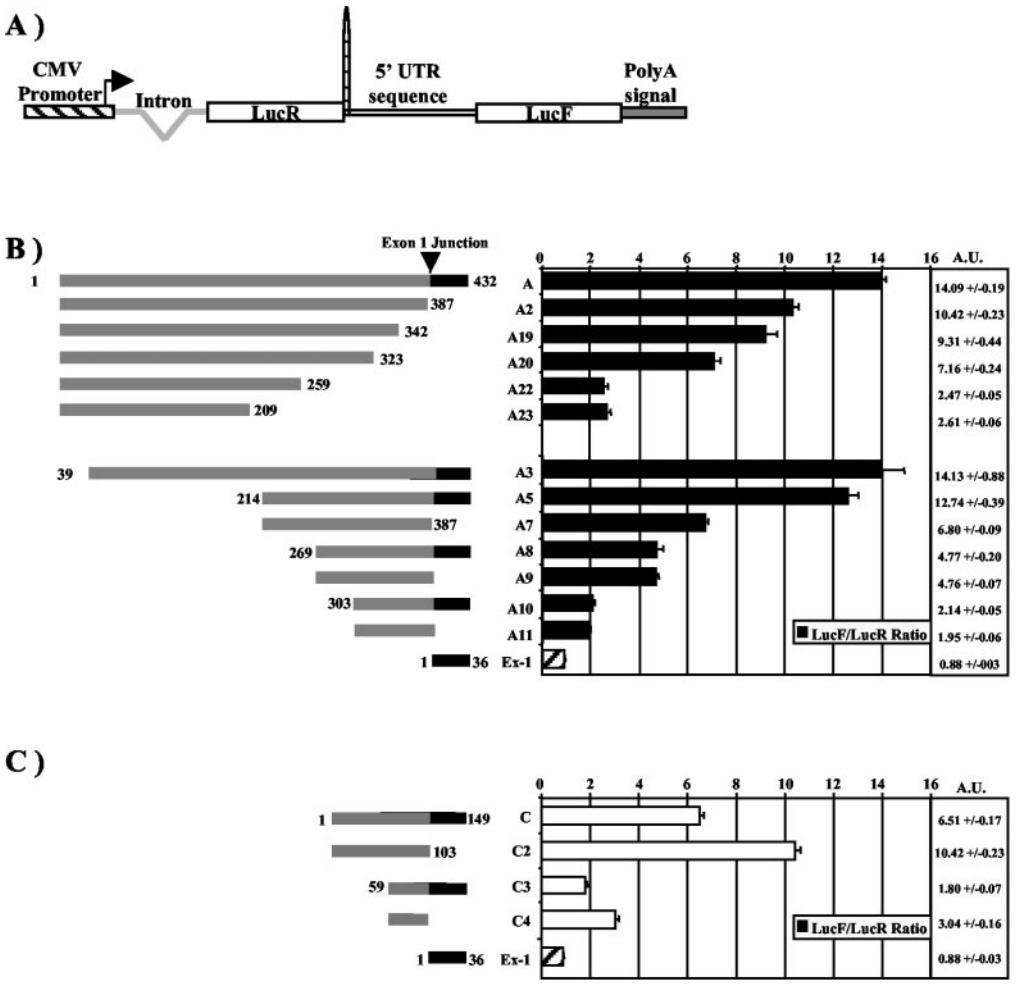


FIG. 5. Mapping of FGF-1 IRESs A and C. (A) Hairpin-containing vector used for mapping. Different portions of FGF-1A or -1C 5′ UTRs were introduced in the intercistronic region, downstream from the hairpin. (B and C) SK-Hep-1 cells were transfected with the different bicistronic constructs, and IRES activities were calculated from the LucF/LucR ratio and expressed in AU as described for Fig. 2. For each leader, nucleotide numbering is from the 5′ end. The different deletions in the A or C 5′ UTRs are shown and are named A*n* or C*n*, respectively. (B) Leader A and exon 1 (Ex-1); (C) leader C and exon 1. The precise IRES activity values are also indicated. The black boxes in the 3′ ends of the leaders correspond to the exon 1 36 nt shared by the four FGF-1 5′ UTRs.

order to detect possible structural variations between the minimal IRES and longer fragments retaining 100% of the IRES activity, probing experiments were performed with both the A9 fragment (minimal IRES) and the A5 fragment (retaining almost 100% activity) (Fig. 7A and data not shown). Three distinct stem-loop domains (DI, DII, and DIII) appeared in the A5 fragment, whereas the Y structure formed by the two asymmetrical stem-loops DII and DIII was identically present in the shorter A9 fragment (Fig. 7B). It is remarkable that the secondary structure obtained from the probing was similar to the structure prediction obtained by RNA folding in silico (Zucker Mfold software) (56) (see Fig. 8A). Furthermore, the IRES activities of RNA fragments with stem-loop DII deleted were abolished (Fig. 5B, bars A10 and A11). In contrast, deletion in the stem-loop DI or DIII only generated a decrease of IRES activity (Fig. 5B, compare bars A9 with A7 and A19 and bar A20 with A2) without abolishing it. Thus, DII appears to be the main IRES structural motif, corresponding to nt 276 to 299 of the FGF-1A leader, whereas DI and DIII, on each side of DII, have a helper effect, possibly by favoring DII structural stabilization.

Following the hypothesis that the DII IRES structural motif, if important for IRES function, should be conserved in mammals, we compared the nucleotide sequence and RNA folding predictions for the segment corresponding to the human minimal IRES A (A9) in six mammalian species (Fig. 8). The mammalian sequences compared to the human sequence were from small-eared galago monkey (a kind of lemur), baboon, cow, mouse, and rat. Interestingly, the nucleotide sequence alignment of the six FGF-1A leaders (Fig. 8A), showing a minimal homology of 60% (versus 86% for the coding sequence [not shown]), revealed a high conservation in the region corresponding to A9, with several 100% conserved nucleotide blocks (Fig. 8A). Furthermore, the in silico RNA folding of the segments corresponding to A9 (Fig. 8B) revealed for all mammals a similar Y structure with two asymmetrical stems-loops, corresponding to the DII and DIII domains identified in Fig. 7. The DIII-like domains revealed some divergence but

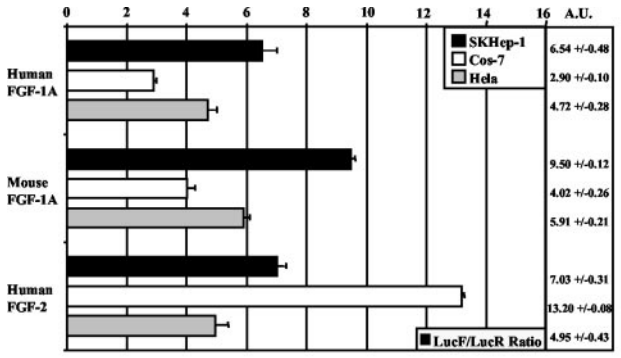


FIG. 6. Comparison of human and mouse FGF-1A 5′ UTRs for sequence homology and IRES activity in mammalian cell types. The entire mouse FGF-1A 5′ UTR was introduced in the Luc-R-I-LucF bicistronic vector described for Fig. 2A. SK-Hep-1, COS-7, and HeLa cells were transfected with the constructs containing the human or murine sequences, and IRES activities were determined as described for Fig. 2 by measurement of the LucF/LucR ratio calibrated to the negative control pCRHL devoid of IRES. Human FGF-2 was used as a positive control. Precise values are indicated on the right. Results represent means ± standard errors from a representative experiment done in triplicate.

also interesting similarities (CUG or CUC bulge, CAGU motif in or close to the loop, etc.). The DII domain was highly conserved in the six mammalian species, with a U bulge and an AACU tetranucleotide loop, supporting the role of DII as the essential IRES structural motif in concordance with the functional assay shown Fig. 5.

The conserved structural motif DII is functionally important for FGF-1 IRES A activity. In addition to use of the deletions presented in Fig. 5 that highlight the function of domain DII in IRES activity, we addressed the functional involvement of domains DI, DII, and DIII by two approaches of domain deletion-reinsertion and mutagenesis in the stem and loop of DII (Fig. 9). As shown in Fig. 9A, deletion of the three domains in leader A resulted in an IRES activity of only 37% (Fig. 9A, panel b). Each of the structural domains were then reintroduced in the leader. IRES activity was thus restored to 54 and 87% by reinsertion of domains DI and DII, respectively (Fig. 9A, panels c and d). In contrast the reinsertion of DIII did not restore any additional IRES activity (Fig. 9A, panel e). These data indicated that the presence of the DII domain in the leader seems to be responsible by itself for about 50% of the IRES activity.

To determine whether the DII structure is important, we generated mutations in the DII stem (Fig. 9B). According to enzymatic and chemical sensitivities (Fig. 7B), the G/C stem seems to be the more stable element of the DII structure. The three Cs that are base paired to three Gs in the wild type were changed to three Gs. The RNA folding of this mutated DII is radically different from that of the wild type (due to the disappearance of the U bulge and the AACU loop) (Fig. 9B, compare left and middle structures). The IRES activity of the complete FGF-1A leader with this altered structure was reduced to 45%. In addition, we introduced compensatory mutations by changing the three Gs to three Cs. According to the RNA folding prediction, the DII stem-loop is thus recreated (Fig. 9B, right structure). This compensatory mutation results in IRES activity recovery of 69%.

The disappearance of the bulge and loop in the stem mutant (Fig. 9B, middle panel) led us to mutate these five nucleotides, which are also the most conserved nucleotides of DII. The nucleotides on the top of the DII structure are the most accessible for putative ITAF binding and are also very sensitive to chemical agents (Fig. 7B). These five nucleotides were randomly changed by the QuikChange approach (see Materials and Methods). Twelve different leader mutants were analyzed for their IRES activities, revealing a 20 to 50% decrease (Fig. 9C). The weakest activities were observed when the U bulge was changed to C or A and the AACU was changed to GCAA (Fig. 9C, mutants 11 and 12).

These data demonstrated the importance of the DII stem, bulge, and loop in the IRES activity of the 432-nt-long sequence of FGF-1 leader A, allowing us to define DII as a new IRES structural motif.

DISCUSSION

The main data of this report are the identification of new IRESs with distinct activity features in the gene of FGF-1, a prototype member of the FGF family. These IRESs, present in different FGF-1 mRNAs resulting from the use of alternative promoters and splicing sites, provide a striking example of coupling of translation with transcription and splicing. Furthermore, we have shown that IRES A activity is conserved in human and mouse, and we have identified a 27-nt structural motif in IRES A, fully conserved among mammals at both the nucleotide and structural levels, which has a crucial role in IRES A activity. This provides new data for IRES prediction in 5′ UTR databases. We have also developed an original system to detect cryptic promoters, which is more reliable than Northern blotting to rule out the hypothesis of second-cistron expression resulting from transcription from an internal promoter.

In a previous report, the existence of several cellular IRESs had been questioned due to the absence of convincing data ruling out the presence of a cryptic promoter or aberrant splicing events (31). Furthermore, aberrant splicing events and promoters have been detected in the 5′ UTRs of several mRNAs described as containing IRESs (1, 15, 19, 20, 51). Such observations are not incompatible with the presence of IRESs but necessitate separation of transcription and translation events (in vitro translation, cell transfection with mRNA, etc.) to prove the presence of an IRES. The bicistronic vector developed in the present study, with a CMV tetracycline-regulatable promoter, provides a reliable tool to demonstrate both the presence of an IRES and the absence of an internal promoter. The only limitation of the system is the requirement for cells expressing the TET transactivator tTA.

The introduction of a promoter in the intercistronic region of the bicistronic construct allows us to demonstrate a strong interference phenomenon, as the internal EF-1α promoter almost abolishes the upstream CMV promoter activity (Fig. 4A). This reminds us of the interest of IRESs from a biotechnological point of view, allowing us to avoid the coexistence of several promoters on the same vector expressing several genes.

The four alternative leaders of the FGF-1 mRNAs have a

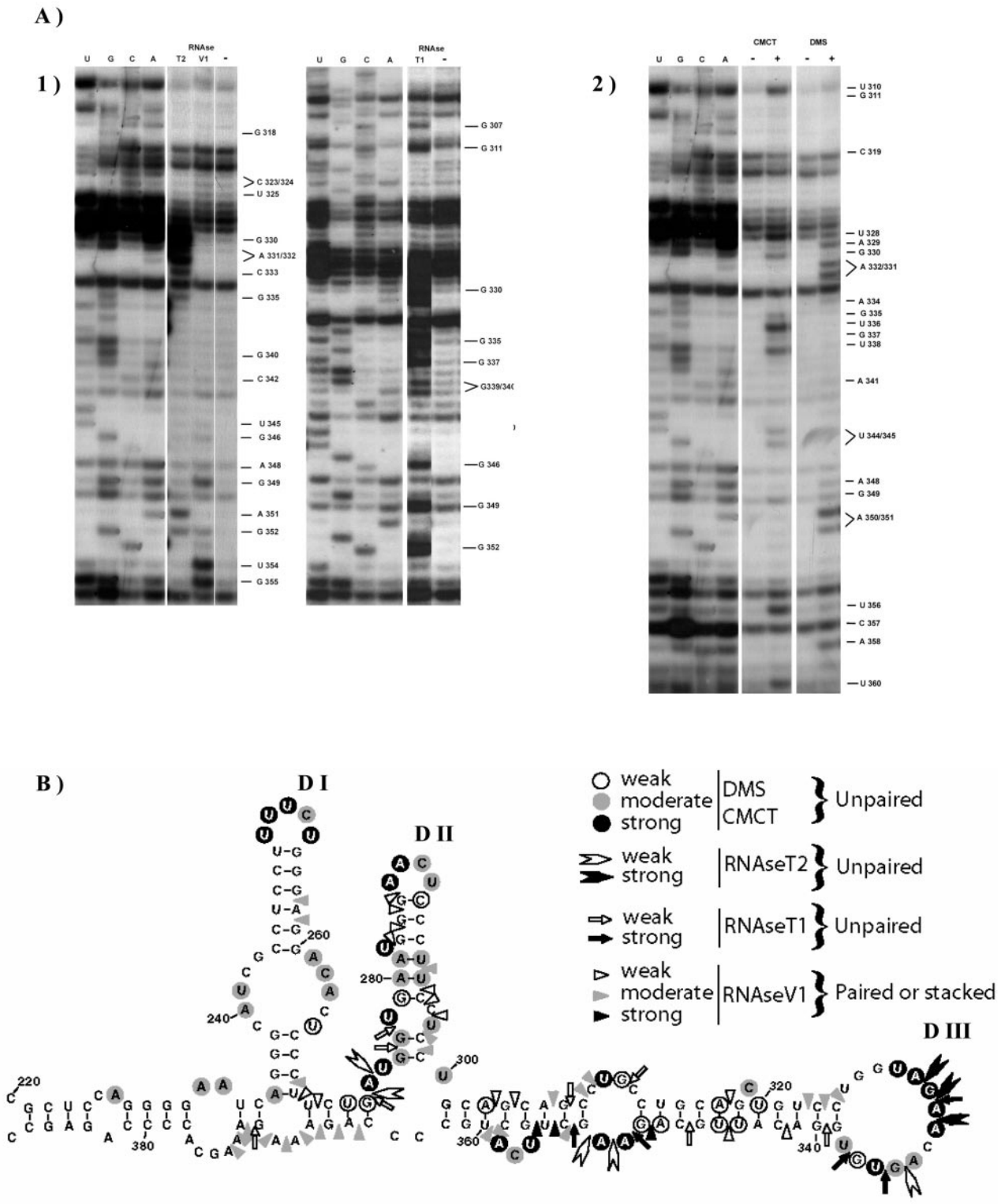


FIG. 7. RNA secondary structure of human FGF-1 IRES A. (A)Enzymatic (panel 1) and chemical (panel 2) probing of in vitro-transcribed A9 fragment. Cleavage and modification sites were detected by primer extension with the 5′-^{32}P-end-labeled RTA9 oligonucleotide, which hybridizes at positions 363 to 385 of the FGF-1A RNA leader. The resulting cDNA was separated on an 8% polyacrylamide–8 M urea sequencing gel and analyzed by autoradiography. RNA sequencing reactions were run in parallel. The nature and positions of cleaved (panel 1) and modified (panel 2) bases are indicated at the right. RNase V_1, RNase T_1, RNase T_2, or the two chemical agents (DMS or CMCT) were added (+) or not (−) before the reverse transcription step. Identical results were found for in vitro-transcribed A5 fragment (data not shown). (B) RNA secondary structure model of the fragment from nt 220 to 386, showing results from enzymatic cleavage and chemical modification experiments. The A9 region is from nt 269 to 387. Sites of RNase cleavage and chemical modification are indicated.

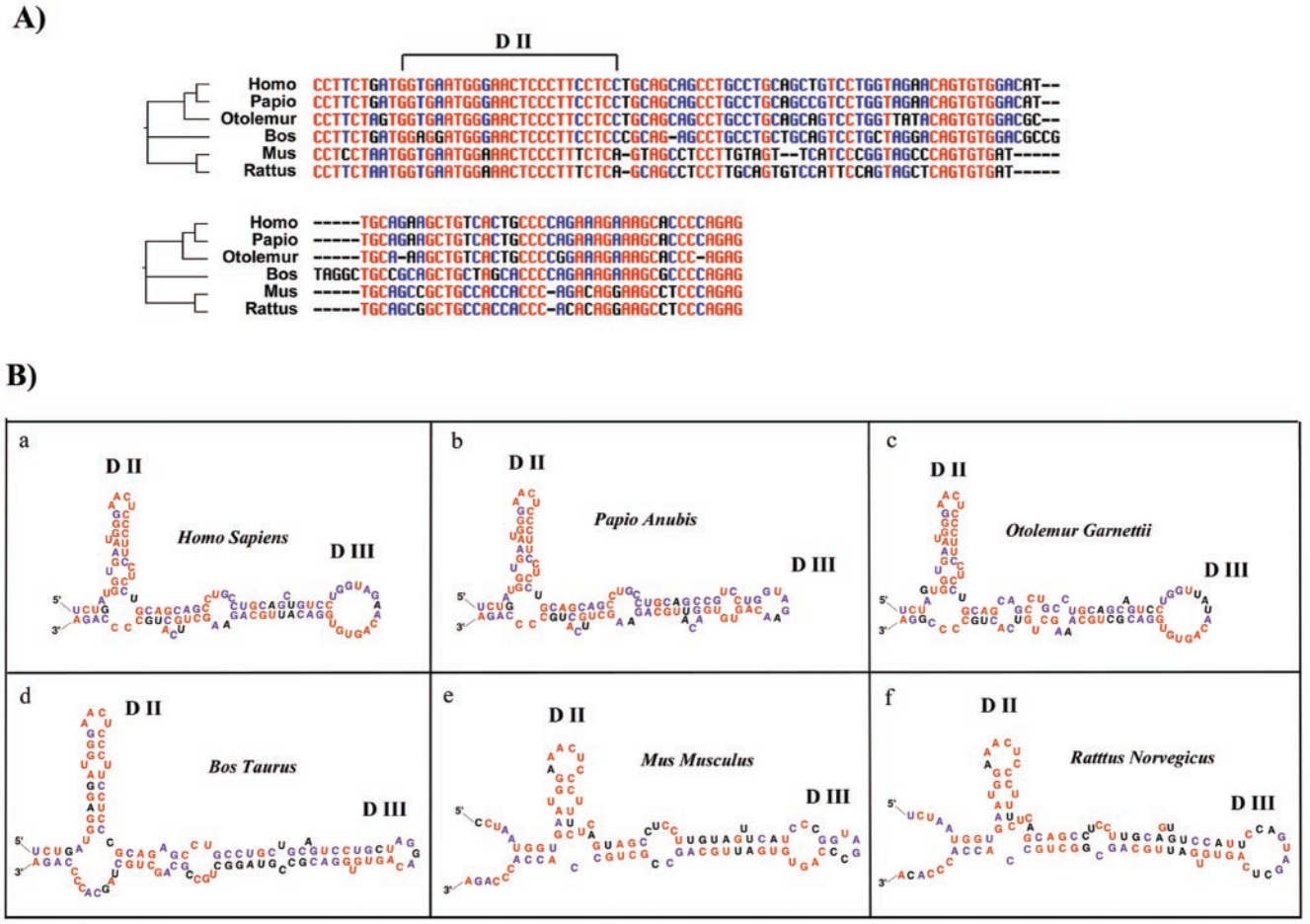
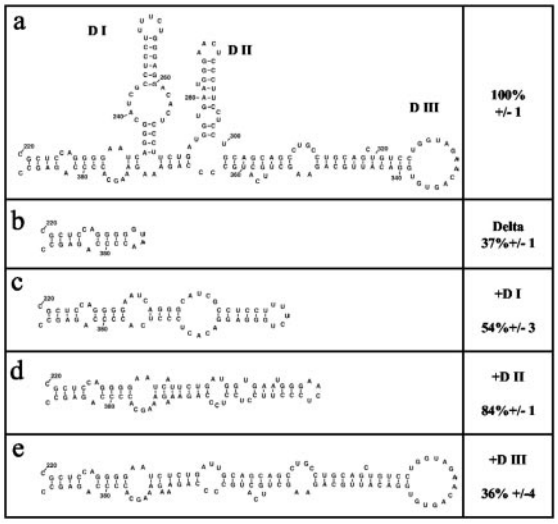


FIG. 8. Alignment and folding of six RNA orthologous sequences corresponding to FGF-1 IRES A. (A) A phylogenic tree was obtained with 5′ UTR orthologous sequences from human (*Homo sapiens*), baboon (*Papio anubis*), small-eared galago monkey (*Otolemur garnettii*), cow (*Bos taurus*), mouse (*Mus musculus*), and rat (*Rattus norvegicus*) by using ClustalW and Phylip software. Nucleotide blocks corresponding to the human minimal IRES (A9 fragment, nt 269 to 387) were compared for the six mammalian species. Alignment was obtained by using Multalin software. Red, blue, and black nucleotides are 100% conserved, more than 50% conserved, and not conserved, respectively. The position of domain DII is indicated. (B) The region corresponding to the DII and DIII domains (nt 270 to 368) of the human FGF-1 IRES A was folded by using Zuker-derived folding software for the five other mammalian species shown in panel A. Domains DII and DIII, conserved in the six species, are indicated.

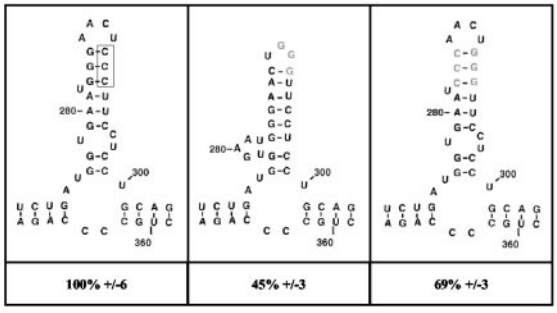
common 3′ segment: 36 nt of the 5′ part of exon 1 (Fig. 1). This segment has no IRES activity by itself in the different cell types tested, except for the 911 cell line. Unexpectedly, this 36-nt segment has an enhancer effect on IRES A when combined with the segment from nt 214 to 269 but has a silencer effect on IRES C (Fig. 5). 3′ deletions in the leader sequences indicate that the FGF-1 IRESs are flexible in regard to the position of the start codon, in contrast to the mechanism of ribosome recruitment described for the IRESs of cardioviruses such as EMCV (type I) (28). The FGF-1 IRESs more resemble type II IRESs (enterovirus), where ribosome recruitment is followed by scanning to reach the initiation codon. Another possibility is the existence of a large starting window, allowing ribosome landing at different start codons within the window, without scanning (44). It has been shown for the Gtx mRNA that a 9-nt fragment complementary to the 18S rRNA is involved in ribosome attachment to mRNA (7). Such a mechanism is not valid for FGF-1, as the three IRESs do not contain the 9-nt segment. In addition, we have shown that this element is present in the mRNA leader of another FGF family member, FGF-2, but is not required for FGF-2 IRES activity (4). Thus, rRNA complementarity is a crucial parameter for IRES activity in a few but not all cases. It must also be mentioned that 18S rRNA base pairing has also been described to be inhibitory of translation initiation by preventing ribosome progression on the mRNA (22, 52).

Another important parameter which can control IRES activity is the presence of small upstream open reading frames (uORFs). A recent report has shown that activation of the arginine-lysine transporter Cat-1 mRNA IRES in response to amino acid deficiency is mediated by translation of a uORF that is responsible for mRNA leader remodeling (16, 54). Alternatively, translation of a uORF can decrease IRES activity if RNA unwinding destabilizes the IRES structure, as shown for the FGF-2 IRES in a previous report (2). The involvement of a uORF in IRES activity modulation remains probable in the case of FGF-1 IRES A, as several conserved AUGs in favorable contexts are present in the leader. One of these ORFs overlaps the minimum IRES A9 (Fig. 5). In contrast, no AUG codon is present in IRES C; however, there are several

A)



B)



C)

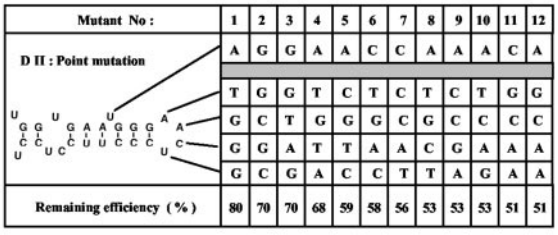


FIG. 9. Domain and nucleotides mutations of human FGF-1 IRES A. The IRES activities of the different constructs, analyzed as described for Fig. 2 by using the double luciferase vector in SH-Hep-1-transfected cells, are indicated as percentages of the activity of the wild-type complete leader 1A IRES. (A) Domain deletion analysis of human FGF-1 IRES A. DI, DII, and DIII domains were deleted from the complete sequence (a) to create mutant Delta (b), and then domain DI (c), DII (d), or DIII (e) was added to Delta in order to estimate their role in IRES A function. (B) Point mutations affecting domain DII structural stability in the complete FGF-1 leader A. The original DII domain contains a three-G/C pairing stem (RNA structure on the left). The three Cs were mutated to three Gs (in gray, structure in the middle) to suppress interaction and to disrupt the structure. Finally, the three original Gs were mutated to three Cs (in gray, structure on the right) in order to restore the G/C pairing and to rebuild the starting structure. (C) Point mutations of the five most accessible and conserved nucleotides in domain DII (U bulge and AACU loop). Using the QuikChange method, nucleotides were randomly and simultaneously mutated in the context of the complete human FGF-1A leader.

possible uORFs whose translation could occur by initiation at CUG and GUG codons (data not shown).

Structural motifs found after chemical and enzymatic probing fully correspond to the structure prediction obtained in silico (Fig. 7). DII and DIII of the human RNA structure are present in both the minimum IRES, A9 (Fig. 5, nt 269 to 387), and the longest segment, A5 (nt 214 to 435); in addition, these motifs are conserved in the structure prediction obtained for primates, cow, mouse, and rat (Fig. 8). DII is conserved both as a structure and at the nucleotide sequence level; one can notice in the stem a purine-rich stretch hybridized to a pyrimidine stretch close to a U bulge and in the loop the AACU tetranucleotide present in the six mammalian species. We show that DII has a crucial role in IRES activity. Furthermore, mutation of the bulged U and of the tetranucleotide loop or of the pyrimidine stretch is able to decrease IRES activity by 50%, whereas compensatory mutations in the stem-loop restores IRES activity, demonstrating the importance of the DII RNA structure (Fig. 9). The conserved nucleotide sequence of the DII domain may be used to search for new IRESs in silico in 5′ UTR databases. Remarkably, the bottom of the DII stem is metastable as indicated by the presence in this region of reactivities towards both single-strand-specific (DMS, CMCT, and T_1) and structure-specific (V_1) probes. As this indicates the simultaneous presence of open and closed conformations within the stem of DII, it would be interesting to determine which is the active form (e.g., by performing mutations stabilizing or destabilizing the bottom part of stem II). In regard to DI and DIII, we suggest that they could act as auxiliary structures that would stabilize the DII structure, thus increasing the IRES activity.

We must mention the limitations of our approach to determine RNA secondary structure (although it is widely used in the field). We used in vitro-transcribed RNA that was heat denatured and cooled prior to probing for structure; such RNA structures might bear no resemblance to those existing in living cells, where mRNA is folded when transcribed in a 5′-to-3′ direction. Furthermore, RNA in cells is always associated with proteins, which can stabilize a given structure that is not always the most thermodynamically stable. However, the compensatory mutagenesis presented in Fig. 9B remains a strong argument to confirm that the DII structure obtained by probing corresponds to an existing mRNA secondary structure in the cell.

Enhanced IRES activity has also been observed for FGF-1 IRES A and for the FGF-2 IRES when a hairpin is present at the end of the first cistron (Fig. 2B). This phenomenon has previously been reported for antennapedia IRES (55). Here again we can hypothesize that the hairpin might act as a "clip" structure, stabilizing the IRES structure by preventing interactions with upstream sequences located in the first cistron.

The presence of such an evolutionarily conserved IRES structure suggests that the structural motifs provide binding sites for ITAFs, especially the AACU loop of DII. Interestingly, DI is pyrimidine rich and could bind pyrimidine tract binding protein (PTB), whereas DII and DIII contain purine-rich stretches resembling the Unr consensus (48). The requirement of both PTB (or neuronal PTB) and Unr has been described recently for the APAF-1 IRES, and both proteins are

thought to act as chaperone proteins changing the IRES structure into one that permits translation (37).

The IRES structural motifs identified in IRES A are completely absent from IRES C, indicating that the FGF-1 IRESs presumably bind different ITAFs (if any) and consequently will be regulated differently, in agreement with our data showing different ex vivo and in vivo behaviors for each of them (Fig. 2 and 4). A bioinformatic approach using the FGF-1 IRES A motif can now be used for searching similar motifs in other members of the FGF family and in all 5′ UTRs in general. A previous study described a putative common Y-type structural motif putatively involved in the internal initiation of translation of cellular mRNAs, in particular the FGF-2 mRNA (34). However, this predicted IRES motif, just upstream of the translation initiation codons, does not correspond to the localization of the recently characterized functional FGF-2 IRES (4). This strong discordance clearly shows that an in silico approach cannot be sufficient by itself and that both functional and structural approaches are necessary to characterize an IRES motif.

A last point to discuss is the specificity and physiological relevance of FGF-1 IRESs. Examination of IRES activities in different cell types as well as in vivo in mouse muscle has shown that the three FGF-1 IRESs behave differently from the FGF-2 IRES. This indicates that although FGF-1 and FGF-2 are both involved in angiogenesis, cancer, cardiovascular diseases, and neural development, their translational regulation is not coordinated and is probably regulated by distinct factors. An interesting possibility would be that the IRES tissue specificity would be coupled to that of the corresponding promoter. Indeed, the FGF-1 promoter A is active in the kidney, whereas promoter B is active in the brain. Promoters C and D, characterized in a variety of cultured cells, including vascular smooth muscle cells, can be induced by different biological response modifiers, including serum and transforming growth factor β (9). The coupling of FGF-1 transcriptional and translational regulation cannot be reliably studied in transformed cultured cells; we have previously shown that regulation is not physiological in such cell systems and that experiments must rather be done in vivo or ex vivo with primary cells (13, 14). In addition, transient transfection must be avoided, as it represents a cellular stress by itself.

The question of the significance of several IRESs for expression of a single growth factor, FGF-1, remains the same as that asked about the presence of four alternative promoters. At the moment, we can propose that multiple controlling mechanisms for expression of a single gene provide the possibility for this gene to be regulated by different signaling pathways. In particular, an IRES-dependent translational stimulation of FGF-1 synthesis may be crucial for the cell to provide an adapted and immediate response to different stimulations such as cell migration, cell survival, angiogenesis, and wound healing, all of which involve FGF-1. IRES-dependent regulation of translation, described in other cases to be stimulated by stress, is a subtle way to complete the transcriptional regulation depending on the alternative promoters. When the promoters are constitutive (as is the case for 1A and 1B), IRESs will also provide an inducibility mechanism allowing the cell to respond rapidly to an exogenous stimulus.

ACKNOWLEDGMENTS

We thank J. Iacovoni for help with bioinformatics. We also thank B. Ehresmann, C. Ehresmann, and P. Romby for hosting Y. Martineau in their laboratory in Strasbourg, France, and for helpful discussions.

This work was supported by grants from INSERM APEX, Université Paul Sabatier, Association pour la Recherche sur le Cancer (ARC), Ligue pour la Recherche contre le Cancer, Association Française contre les Myopathies (AFM), Conseil Régional Midi-Pyrénées, European Commission FP5 (QOL-2000-3.1.2, consortium CONTEXTH contract QLRT-2000-00721), and the French Ministry of Research (decision no. 01H0387). Y. Martineau was financed by doctoral fellowships from the French Minister of Research and from ARC.

REFERENCES

1. **Akiri, G., D. Nahari, Y. Finkelstein, S. Y. Le, O. Elroy-Stein, and B. Z. Levi.** 1998. Regulation of vascular endothelial growth factor (VEGF) expression is mediated by internal initiation of translation and alternative initiation of transcription. Oncogene **17:**227–236.
2. **Arnaud, E., C. Touriol, C. Boutonnet, M. C. Gensac, S. Vagner, H. Prats, and A. C. Prats.** 1999. A new 34-kilodalton isoform of human fibroblast growth factor 2 is cap dependently synthesized by using a non-AUG start codon and behaves as a survival factor. Mol. Cell. Biol. **19:**505–514.
3. **Bernstein, J., O. Sella, S. Y. Le, and O. Elroy-Stein.** 1997. PDGF2/c-sis mRNA leader contains a differentiation-linked internal ribosomal entry site (D-IRES). J. Biol. Chem. **272:**9356–9362.
4. **Bonnal, S., C. Schaeffer, L. Creancier, S. Clamens, H. Moine, A. C. Prats, and S. Vagner.** 2003. A single internal ribosome entry site containing a G quartet RNA structure drives fibroblast growth factor 2 gene expression at four alternative translation initiation codons. J. Biol. Chem. **278:**39330–39336.
5. **Brunel, C., and P. Romby.** 2000. Probing RNA structure and RNA-ligand complexes with chemical probes. Methods Enzymol. **318:**3–21.
6. **Bugler, B., F. Amalric, and H. Prats.** 1991. Alternative initiation of translation determines cytoplasmic or nuclear localization of basic fibroblast growth factor. Mol. Cell. Biol. **11:**573–577.
7. **Chappell, S. A., G. M. Edelman, and V. P. Mauro.** 2000. A 9-nt segment of a cellular mRNA can function as an internal ribosome entry site (IRES) and when present in linked multiple copies greatly enhances IRES activity. Proc. Natl. Acad. Sci. USA **97:**1536–1541.
8. **Chappell, S. A., J. P. LeQuesne, F. E. Paulin, M. L. deSchoolmeester, M. Stoneley, R. L. Soutar, S. H. Ralston, M. H. Helfrich, and A. E. Willis.** 2000. A mutation in the c-myc-IRES leads to enhanced internal ribosome entry in multiple myeloma: a novel mechanism of oncogene de-regulation. Oncogene **19:**4437–4440.
9. **Chiu, I. M., K. Touhalisky, and C. Baran.** 2001. Multiple controlling mechanisms of FGF1 gene expression through multiple tissue-specific promoters. Prog. Nucleic Acid Res. Mol. Biol. **70:**155–174.
10. **Chotani, M. A., and I. M. Chiu.** 1997. Differential regulation of human fibroblast growth factor 1 transcripts provides a distinct mechanism of cell-specific growth factor expression. Cell Growth Differ. **8:**999–1013.
11. **Chotani, M. A., R. A. Payson, J. A. Winkles, and I. M. Chiu.** 1995. Human fibroblast growth factor 1 gene expression in vascular smooth muscle cells is modulated via an alternate promoter in response to serum and phorbol ester. Nucleic Acids Res. **23:**434–441.
12. **Couderc, B., H. Prats, F. Bayard, and F. Amalric.** 1991. Potential oncogenic effects of basic fibroblast growth factor requires cooperation between CUG and AUG-initiated forms. Cell Regul. **2:**709–718.
13. **Creancier, L., P. Mercier, A. C. Prats, and D. Morello.** 2001. c-myc internal ribosome entry site activity is developmentally controlled and subjected to a strong translational repression in adult transgenic mice. Mol. Cell. Biol. **21:**1833–1840.
14. **Creancier, L., D. Morello, P. Mercier, and A. C. Prats.** 2000. Fibroblast growth factor 2 internal ribosome entry site (IRES) activity ex vivo and in transgenic mice reveals a stringent tissue-specific regulation. J. Cell Biol. **150:**275–281.
15. **Dumas, E., C. Staedel, M. Colombat, S. Reigadas, S. Chabas, T. Astier-Gin, A. Cahour, S. Litvak, and M. Ventura.** 2003. A promoter activity is present in the DNA sequence corresponding to the hepatitis C virus 5′ UTR. Nucleic Acids Res. **31:**1275–1281.
16. **Fernandez, J., I. Yaman, R. Mishra, W. C. Merrick, M. D. Snider, W. H. Lamers, and M. Hatzoglou.** 2001. Internal ribosome entry site-mediated translation of a mammalian mRNA is regulated by amino acid availability. J. Biol. Chem. **276:**12285–12291.
17. **Galy, B., A. Maret, A. C. Prats, and H. Prats.** 1999. Cell transformation results in the loss of the density-dependent translational regulation of the expression of fibroblast growth factor 2 isoforms. Cancer Res. **59:**165–171.
18. **Gossen, M., and H. Bujard.** 1992. Tight control of gene expression in mammalian cells by tetracycline-responsive promoters. Proc. Natl. Acad. Sci. USA **89:**5547–5551.
19. **Han, B., Z. Dong, and J. T. Zhang.** 2003. Tight control of platelet-derived

growth factor B/c-sis expression by interplay between the 5′-untranslated region sequence and the major upstream promoter. J. Biol. Chem. **278:** 46983–46993.
20. **Han, B., and J. T. Zhang.** 2002. Regulation of gene expression by internal ribosome entry sites or cryptic promoters: the eIF4G story. Mol. Cell. Biol. **22:**7372–7384.
21. **Hellen, C. U., and P. Sarnow.** 2001. Internal ribosome entry sites in eukaryotic mRNA molecules. Genes Dev. **15:**1593–1612.
22. **Hu, M. C., P. Tranque, G. M. Edelman, and V. P. Mauro.** 1999. rRNA-complementarity in the 5′ untranslated region of mRNA specifying the Gtx homeodomain protein: evidence that base-pairing to 18S rRNA affects translational efficiency. Proc. Natl. Acad. Sci. USA **96:**1339–1344.
23. **Huez, I., L. Creancier, S. Audigier, M. C. Gensac, A. C. Prats, and H. Prats.** 1998. Two independent internal ribosome entry sites are involved in translation initiation of vascular endothelial growth factor mRNA. Mol. Cell. Biol. **18:**6178–6190.
24. **Hunt, S. L., J. J. Hsuan, N. Totty, and R. J. Jackson.** 1999. unr, a cellular cytoplasmic RNA-binding protein with five cold-shock domains, is required for internal initiation of translation of human rhinovirus RNA. Genes Dev. **13:**437–448.
25. **Hwang, L. H., C. L. Hsieh, A. Yen, Y. L. Chung, and D. S. Chen.** 1998. Involvement of the 5′ proximal coding sequences of hepatitis C virus with internal initiation of viral translation. Biochem. Biophys. Res. Commun. **252:**455–460.
26. **Jaag, H. M., L. Kawchuk, W. Rohde, R. Fischer, N. Emans, and D. Prufer.** 2003. An unusual internal ribosomal entry site of inverted symmetry directs expression of a potato leafroll polerovirus replication-associated protein. Proc. Natl. Acad. Sci. USA **100:**8939–8944.
27. **Jackson, R. J.** 1991. mRNA translation. Initiation without an end. Nature **353:**14–15.
28. **Jackson, R. J., and A. Kaminski.** 1995. Internal initiation of translation in eukaryotes: the picornavirus paradigm and beyond. RNA **1:**985–1000.
29. **Kanamori, Y., and N. Nakashima.** 2001. A tertiary structure model of the internal ribosome entry site (IRES) for methionine-independent initiation of translation. RNA **7:**266–274.
30. **Kieft, J. S., K. Zhou, A. Grech, R. Jubin, and J. A. Doudna.** 2002. Crystal structure of an RNA tertiary domain essential to HCV IRES-mediated translation initiation. Nat. Struct. Biol. **9:**370–374.
31. **Kozak, M.** 2001. New ways of initiating translation in eukaryotes? Mol. Cell. Biol. **21:**1899–1907.
32. **Lambrechts, D., E. Storkebaum, M. Morimoto, J. Del-Favero, F. Desmet, S. L. Marklund, S. Wyns, V. Thijs, J. Andersson, I. van Marion, A. Al-Chalabi, S. Bornes, R. Musson, V. Hansen, L. Beckman, R. Adolfsson, H. S. Pall, H. Prats, S. Vermeire, P. Rutgeerts, S. Katayama, T. Awata, N. Leigh, L. Lang-Lazdunski, M. Dewerchin, C. Shaw, L. Moons, R. Vlietinck, K. E. Morrison, W. Robberecht, C. Van Broeckhoven, D. Collen, P. M. Andersen, and P. Carmeliet.** 2003. VEGF is a modifier of amyotrophic lateral sclerosis in mice and humans and protects motoneurons against ischemic death. Nat. Genet. **34:**383–394.
33. **Le, S. Y., J. H. Chen, N. Sonenberg, and J. V. Maizel.** 1992. Conserved tertiary structure elements in the 5′ untranslated region of human enteroviruses and rhinoviruses. Virology **191:**858–866.
34. **Le, S. Y., and J. V. Maizel, Jr.** 1997. A common RNA structural motif involved in the internal initiation of translation of cellular mRNAs. Nucleic Acids Res. **25:**362–369.
35. **Lyons, A. J., J. R. Lytle, J. Gomez, and H. D. Robertson.** 2001. Hepatitis C virus internal ribosome entry site RNA contains a tertiary structural element in a functional domain of stem-loop II. Nucleic Acids Res. **29:**2535–2541.
36. **Madiai, F., K. V. Hackshaw, and I. M. Chiu.** 1999. Characterization of the entire transcription unit of the mouse fibroblast growth factor 1 (FGF-1) gene. Tissue-specific expression of the FGF-1.A mRNA. J. Biol. Chem. **274:**11937–11944.
37. **Mitchell, S. A., K. A. Spriggs, M. J. Coldwell, R. J. Jackson, and A. E. Willis.** 2003. The Apaf-1 internal ribosome entry segment attains the correct structural conformation for function via interactions with PTB and unr. Mol. Cell **11:**757–771.
38. **Myers, R. L., R. A. Payson, M. A. Chotani, L. L. Deaven, and I. M. Chiu.** 1993. Gene structure and differential expression of acidic fibroblast growth factor mRNA: identification and distribution of four different transcripts. Oncogene **8:**341–349.
39. **Myers, R. L., S. K. Ray, R. Eldridge, M. A. Chotani, and I. M. Chiu.** 1995. Functional characterization of the brain-specific FGF-1 promoter, FGF-1.B. J. Biol. Chem. **270:**8257–8266.
40. **Nishimura, T., Y. Nakatake, M. Konishi, and N. Itoh.** 2000. Identification of a novel FGF, FGF-21, preferentially expressed in the liver. Biochim. Biophys. Acta **1492:**203–206.
41. **Payson, R. A., M. A. Chotani, and I. M. Chiu.** 1998. Regulation of a promoter of the fibroblast growth factor 1 gene in prostate and breast cancer cells. J. Steroid Biochem. Mol. Biol. **66:**93–103.
42. **Perez, N., P. Plence, V. Millet, D. Greuet, C. Minot, D. Noel, O. Danos, C. Jorgensen, and F. Apparailly.** 2002. Tetracycline transcriptional silencer tightly controls transgene expression after in vivo intramuscular electrotransfer: application to interleukin 10 therapy in experimental arthritis. Hum. Gene Ther. **13:**2161–2172.
43. **Pestova, T. V., I. N. Shatsky, S. P. Fletcher, R. J. Jackson, and C. U. Hellen.** 1998. A prokaryotic-like mode of cytoplasmic eukaryotic ribosome binding to the initiation codon during internal translation initiation of hepatitis C and classical swine fever virus RNAs. Genes Dev. **12:**67–83.
44. **Pilipenko, E. V., A. P. Gmyl, S. V. Maslova, G. A. Belov, A. N. Sinyakov, M. Huang, T. D. Brown, and V. I. Agol.** 1994. Starting window, a distinct element in the cap-independent internal initiation of translation on picornaviral RNA. J. Mol. Biol. **241:**398–414.
45. **Prats, A. C., G. De Billy, P. Wang, and J. L. Darlix.** 1989. CUG initiation codon used for the synthesis of a cell surface antigen coded by the murine leukemia virus. J. Mol. Biol. **205:**363–372.
46. **Prats, H., M. Kaghad, A. C. Prats, M. Klagsbrun, J. M. Lelias, P. Liauzun, P. Chalon, J. P. Tauber, F. Amalric, J. A. Smith, et al.** 1989. High molecular mass forms of basic fibroblast growth factor are initiated by alternative CUG codons. Proc. Natl. Acad. Sci. USA **86:**1836–1840.
47. **Reynolds, J. E., A. Kaminski, A. R. Carroll, B. E. Clarke, D. J. Rowlands, and R. J. Jackson.** 1996. Internal initiation of translation of hepatitis C virus RNA: the ribosome entry site is at the authentic initiation codon. RNA **2:**867–878.
48. **Triqueneaux, G., M. Velten, P. Franzon, F. Dautry, and H. Jacquemin-Sablon.** 1999. RNA binding specificity of Unr, a protein with five cold shock domains. Nucleic Acids Res. **27:**1926–1934.
49. **Vagner, S., M. C. Gensac, A. Maret, F. Bayard, F. Amalric, H. Prats, and A. C. Prats.** 1995. Alternative translation of human fibroblast growth factor 2 mRNA occurs by internal entry of ribosomes. Mol. Cell. Biol. **15:**35–44.
50. **Vagner, S., A. Waysbort, M. Marenda, M. C. Gensac, F. Amalric, and A. C. Prats.** 1995. Alternative translation initiation of the Moloney murine leukemia virus mRNA controlled by internal ribosome entry involving the p57/PTB splicing factor. J. Biol. Chem. **270:**20376–20383.
51. **Van Eden, M. E., M. P. Byrd, K. W. Sherrill, and R. E. Lloyd.** 2004. Demonstrating internal ribosome entry sites in eukaryotic mRNAs using stringent RNA test procedures. RNA **10:**720–730.
52. **Verrier, S. B., and O. Jean-Jean.** 2000. Complementarity between the mRNA 5′ untranslated region and 18S ribosomal RNA can inhibit translation. RNA **6:**584–597.
53. **West, M. J., N. F. Sullivan, and A. E. Willis.** 1995. Translational upregulation of the c-myc oncogene in Bloom's syndrome cell lines. Oncogene **11:**2515–2524.
54. **Yaman, I., J. Fernandez, H. Liu, M. Caprara, A. A. Komar, A. E. Koromilas, L. Zhou, M. D. Snider, D. Scheuner, R. J. Kaufman, and M. Hatzoglou.** 2003. The zipper model of translational control: a small upstream ORF is the switch that controls structural remodeling of an mRNA leader. Cell **113:** 519–531.
55. **Ye, X., P. Fong, N. Iizuka, D. Choate, and D. R. Cavener.** 1997. Ultrabithorax and Antennapedia 5′ untranslated regions promote developmentally regulated internal translation initiation. Mol. Cell. Biol. **17:**1714–1721.
56. **Zuker, M.** 2003. Mfold web server for nucleic acid folding and hybridization prediction. Nucleic Acids Res. **31:**3406–3415.