Human insulin-like growth factor II leader 2 mediates internal initiation of translation

Susanne K. PEDERSEN*, Jan CHRISTIANSEN†, Thomas v.O. HANSEN*, Martin R. LARSEN‡ and Finn C. NIELSEN*¹

*Department of Clinical Biochemistry, Rigshospitalet, Copenhagen, Denmark, †RNA Regulation Centre, Institute of Molecular Biology, University of Copenhagen, Denmark, and ‡Institute of Molecular Biology and Biochemistry, SDU, Odense University, Denmark

Insulin-like growth factor II (IGF-II) is a fetal growth factor, which belongs to the family of insulin-like peptides. During fetal life, the IGF-II gene generates three mRNAs with different 5′ untranslated regions (UTRs), but identical coding regions and 3′ UTRs. We have shown previously that IGF-II leader 3 mRNA translation is regulated by a rapamycin-sensitive pathway, whereas leader 4 mRNA is constitutively translated, but so far the significance of leader 2 mRNA has been unclear. Here, we show that leader 2 mRNA is translated efficiently in an eIF4E-independent manner. In a bicistronic vector system, the 411 nt leader 2 was capable of internal initiation via a phylogenetically conserved internal ribosome entry site (IRES), located in the 3′ half of the leader. The IRES is composed of an approx. 120 nt ribosome recruitment element, followed by an 80 nt spacer region, which is scanned by the ribosomal pre-initiation complex. Since cap-dependent translation is down-regulated during cell division, leader 2 might facilitate a continuous IGF-II production in rapidly dividing cells during development.

Key words: insulin-like growth factors, internal ribosome entry site, translation initiation, 5′ untranslated region.

INTRODUCTION

Insulin-like growth factor II (IGF-II) is a mitogenic peptide, structurally related to insulin and IGF-I, that regulates fetal development and growth (for a review, see [1]). The human IGF-II gene generates four mRNAs with different 5′ untranslated regions (UTRs), but identical coding regions and 3′ UTRs, by utilizing four different promoters: P1, P2, P3 and P4 (Figure 1). Promoters P2, P3 and P4 are active in fetal tissues and transformed cells, where they generate mature transcripts of 5.0, 6.0 and 4.8 kb respectively, whereas promoter P1 is used only in adult liver, giving rise to a transcript of 5.3 kb. The 5′ UTRs are structurally distinct and provide separate functions to the transcripts. The abundant 6.0 kb leader 3 mRNA is selectively mobilized in growing cells by a rapamycin-sensitive mechanism [2], and binds multiple copies of the family of IGF-II mRNA binding proteins ('IMPs'), which mediate sub-cytoplasmic mRNA localization [3]. Leader 4 mRNA is translated constitutively, whereas the adult liver-specific leader 1 may possess the ability to be translated in a cap-independent manner [4]. In contrast, the significance of leader 2 mRNA has been unclear.

Translation is a key stage in the regulation of gene expression and cell growth (for reviews, see [4,5]). Translation initiation is generally considered to occur by a cap-dependent scanning mechanism, where the 43 S ribosomal pre-initiation complex is brought into contact with the 5′ terminus of the mRNA via binding to eukaryotic initiation factor (eIF)4G in the eIF4F cap-binding complex, which also contains the cap-binding protein eIF4E, and eIF4A, which exhibits an ATP-dependent RNA helicase activity. Infection by a variety of RNA viruses results in the selective inhibition of the host translation, while the viral RNA is efficiently translated. Viral proteases cleave eIF4G-I

and -II, thereby disrupting the coupling between eIF4E and the 43 S pre-initiation complex in cap-dependent translation. The viral RNA, in contrast, initiates translation by a cap-independent mechanism via an internal ribosome entry site (IRES) in the 5′ UTR, which is capable of recruiting the 43 S ribosomal pre-initiation complex independently of an intact eIF4F complex (for a review, see [6]). Recruitment of the 43 S complex may be accomplished via at least three different mechanisms. In encephalo-myocarditis virus (EMCV) RNA, recruitment is mediated by binding of eIF4G/eIF4A and canonical eIFs 2, 3 and 4B, whereas the hepatitis C virus IRES associates directly with the 43 S pre-initiation complex via an apical loop and a pseudoknot. Finally, the cricket paralysis-like viruses appear to occupy the ribosomal P-site directly, and can assemble 80 S ribosomes without any eIF or the initiator tRNA. A decade ago, cap-independent translation was mainly regarded as a viral speciality that allowed the virus to gain control of the host translation machinery. After the demonstration of cap-independent translation of BiP (immunoglobulin heavy-chain binding protein) mRNA [7], however, it has become evident that an increasing number of cellular mRNAs are capable of internal initiation in situations where cap-dependent translation is compromised (reviewed in [6]). The cellular IRESs frequently comprise multiple non-contiguous sequences, the structural features of which have only been resolved in a few cases [6,8]. The mechanism by which the cellular IRESs function is still incompletely understood, and might involve RNA structural features, general initiation factors and other *trans*-acting factors. The Gtx homeodomain protein mRNA IRES recognizes the 40 S subunit via a complementary 9 nt sequence in the 5′ UTR [9], whereas internal initiation of ornithine decarboxylase mRNA is regulated by eIF4G [10]. Finally, factors such as heterogeneous nuclear ribonucleopro-
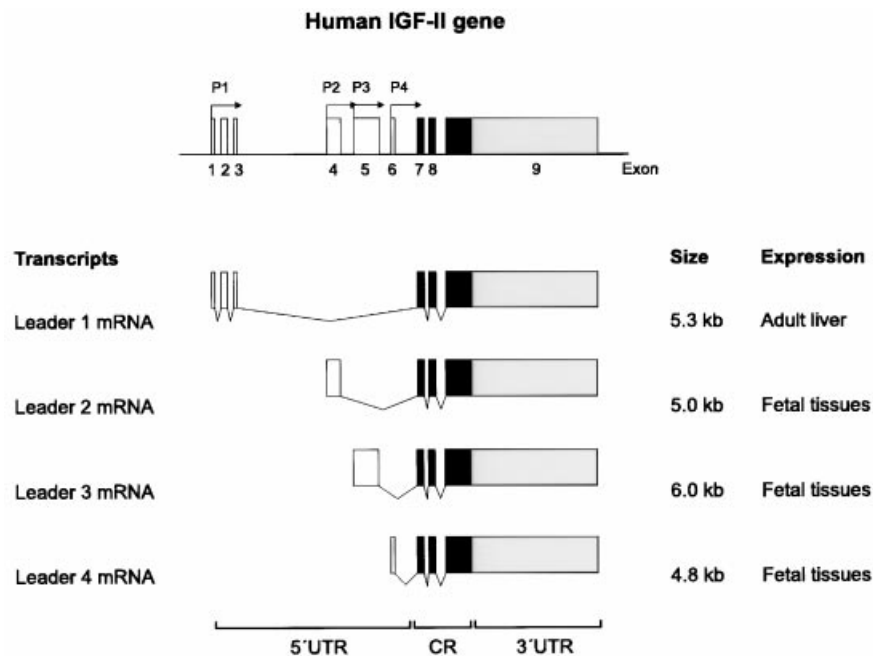
Figure 1   Schematic representation of the human IGF-II gene

The human IGF-II gene consists of nine exons, where the 180 amino acids of pre-proIGF-II are encoded by exon 7, exon 8 and the first 234 nt of exon 9 (black boxes). The gene generates four mRNAs designated leader 1, 2, 3 or 4 mRNA, with different 5′ UTRs but identical coding region and 3′ UTRs from four different promoters (P1, P2, P3 and P4). The sizes of the mRNAs and their expression in adult or fetal tissue are shown to the right.

tein C ('hnRNP C'), La and polypyrimidine tract binding protein ('PTB') have been implicated in the regulation of IRES-mediated translation of platelet-derived growth factor 2 ('PDGF'), X-linked inhibitor of apoptosis and apoptotic protease-activating factor 1 mRNA respectively [11–13].

In the present study, we examine the translation of IGF-II leader 2 mRNA and show that the leader is capable of mediating cap-independent translation initiation via an IRES located in the 3′ half of the leader. Since cap-dependent translation might be down-regulated during cell division, leader 2 mRNA might ensure a steady production of IGF-II during fetal stress and in rapidly dividing cells during development.

EXPERIMENTAL

Plasmid constructs

The bicistronic vector pBicFire was generated by inserting the *Firefly* and *Renilla* luciferase open reading frames with a small 30 bp intercistronic spacer containing the unique restriction sites *Eco*RI, *Bam*HI, *Eco*RV and *Xho* downstream of the cytomegalovirus (CMV) promoter of pcDNA$_{3.1}$($+zeo$) (Invitrogen, Carlsbad, CA, U.S.A.). The *Firefly* and *Renilla* luciferase open reading frames were derived from the pGL3-Basic and pRL-SV40 vectors (Promega, Madison, WI, U.S.A.) by PCR and contained no extraneous sequences, except for attached polylinkers. The 5′ UTRs of gastrin, EMCV (segment 303–859; Genbank® X74312), poliovirus (segment 50–749; kindly given by P. Sarnow), BiP and leaders 2, 3 and 4 of human IGF-II mRNA were cloned into the *Bam*HI and *Eco*RV sites of the intercistronic spacer. To introduce an upstream stable RNA hairpin in pBicFire, a 30 bp *Hin*dIII–*Eco*RI fragment from pGEM7Z (Promega) [13] was cloned as an inverted tandem repeat into the unique *Hin*dIII site immediately

upstream of the *Firefly* luciferase open reading frame of pBicFire. *In vitro* site-directed mutagenesis was performed with the QuikChange™ site-directed mutagenesis kit, according to the manufacturer's instructions (Stratagene, La Jolla, CA, U.S.A.). AUG mutations were generated with the following pairs of primers: 5′-GCC TCT CTG TCT CCT ACG AAG TC<u>A CCA TGG</u> CAA CTC GGA TTT GGG AAA TTT CTC TCT AGC (sense) and 5′-GCT AGA GAG AAA TTT CCC AAA TCC GAG TTG CCA TGG TGA CTT CGT (antisense) (position 92); and 5′-CGC CTG CCA CAG AGC GT<u>A CCA TGG</u> CTC GCT GCC TGA GCT CCT GGT GC (sense) and 5′-GCA CCA GGA GCT CAG GCA GCG AGC CAT GGT ACG CTC TGT GGC AGG CG (antisense) (position 348). The resulting AUG codons embedded in a minimal Kozak consensus sequence are shown underlined in the sense primers. The coding region of eIF4E was inserted downstream of the CMV promoter in pcDNA$_{3.1}$($+zeo$) (Invitrogen).

Cell culture and transient transfection

Human RD rhabdomyosarcoma cells (A.T.C.C. no. CCL-136) were maintained in RPMI 1640 medium supplemented with 10 % (v/v) fetal bovine serum, and human SK-N-MC neuroblastoma cells (A.T.C.C. no. HTB-10) were cultured 1:1 in Dulbecco's modified Eagle's medium [1000 mg/l glucose/200 mM glutamax-1 and Nutrient mix F-12 with glutamax-1 supplemented with 15 % (v/v) fetal bovine serum and 1 % (w/v) non-essential amino acids]. Cells were transiently transfected with LIPOFECT-AMINE™ according to the manufacturer's instructions (Life Technologies, Tåstrup, Denmark). Briefly, 30 000 cells/cm² were seeded in 24-well plates 24 h before transfection. The cells were transfected with 400 ng of reporter plasmid DNA and 50 ng of pRL-SV40 plasmid (Promega), which was used for normaliza-
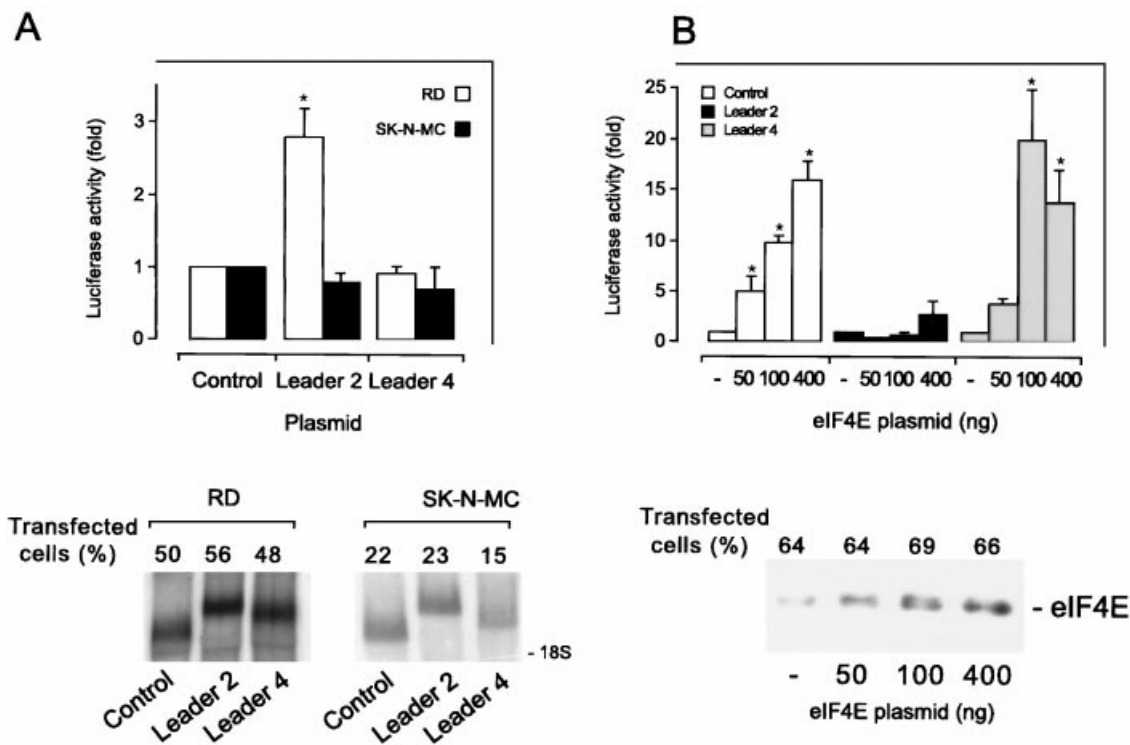
Figure 2   Translation of IGF-II leader 2, 3 and 4 chimaeric mRNAs

(**A**) Monocistronic IGF-II leader 2 or 4 luciferase reporter constructs and a pcDNA$_{3.1}$($+$ zeo) luciferase control were expressed in RD rhabdomyosarcoma or SK-N-MC neuroepithelial cells. The luciferase activities are expressed relative to the pcDNA$_{3.1}$($+$ zeo) control vector. All data were corrected for differences in transfection efficiency by normalization to the activity of the included pRL-SV40 construct. The results are shown as the means $\pm$ S.E.M. for three independent experiments. The lower panels show a representative Northern blot of the luciferase control, leader 2–luciferase and leader 4–luciferase mRNAs in RD (left) and SK-N-MC (right) cells. The percentages of transfected cells are indicated on top of the lower panels. (**B**) Co-expression of monocistronic luciferase control (open bars), leader 2 (black bars) or leader 4 (grey bars) luciferase constructs with increasing amounts of an eIF4E-encoding plasmid in RD cells. Results are shown as means $\pm$ S.E.M. for three experiments. The asterisks in (**A**) and (**B**) show statistical significance ($P < 0.05$). The lower panel shows a Western blot analysis of the dose-dependent expression of eIF-4E, and the percentages of transfected cells are given above the gel.

tion. After 48 h, luciferase activity was measured with the Dual Luciferase Assay kit (Promega). To determine the percentage of cells that were transfected, 50 ng of pEGFP were included in the transfection reaction, and the numbers of fluorescent and non-fluorescent cells were counted.

Western blot analysis

Cellular proteins were separated on an SDS/15 % polyacrylamide gel and transferred to PVDF Immobilon-P membranes (Millipore, Bedford, MA, U.S.A.). After blocking, filters were incubated overnight at 4 °C with anti-(eIF-4E) antibody (Transduction Laboratories, Lexington, KY, U.S.A.) in blocking solution, and with horseradish-peroxidase-conjugated anti-mouse IgG in blocking solution for 1 h at room temperature (25 °C). Bound antibody was detected with enhanced chemiluminescence reagents, in accordance with the manufacturer's instructions (Pierce, Rockford, IL, U.S.A.).

Northern blot analysis

Total RNA was isolated from transfected RD cells cultured in six-well plates by the guanidinium thiocyanate method [14]. RNA was denatured in glyoxal/DMSO, separated on a 1 % agarose gel, transferred to a Hybond-N membrane (Amersham Biosciences, Piscataway, NJ, U.S.A.), and hybridized with a ^{32}P-labelled full-length *Renilla* luciferase cDNA probe. Auto-

radiography was performed over a 16 h period, and the hybridization signals were visualized with a Fuji BAS 2000 Bioimager.

Promoter activity of IGF-II leader 2 segment comprising nt 200–400

Leader 2 segment comprising nt 200–400 or the EMCV IRES was cloned into the *Mlu*I/*Xho*I sites of pGL3-Basic (Promega). pGL3-SV40, pGL3-Basic or pGL3-Basic with the IRES segments were transfected into RD cells, as described above, together with pRL-0 (Promega), that was included for normalization. Cells were harvested after 48 h and analysed for luciferase activity.

Chemical probing of the leader 2 IRES

Structural probing of IGF-II leader 2 was performed essentially as described previously [15]. Briefly, RNA was generated by T7 RNA polymerase-directed *in vitro* transcription (Promega) and purified by gel filtration in Microspin G50 spin columns (Pharmacia), followed by precipitation with ethanol using 1.5 µg of tRNA as the carrier. Of the leader 2 RNA or the leader 2 segment 200–321 RNA, 4 µg was renatured by heating at 56 °C for 10 min in 70 mM Hepes/KOH, pH 7.8/10 mM MgCl$_2$/270 mM KCl followed by slow cooling to approx. 20 °C. The RNAs were modified with 1-cyclohexyl-3-(2-morpholinoethyl)-carbodi-imide metho-p-toluene sulphonate (CMCT; Sigma, St

Louis, MO, U.S.A.), dimethylsulphate (DMS; Merck, Gibbstown, NJ, U.S.A.) and kethoxal (Nutritional Biochemicals, Cleveland, OH, U.S.A.). The modified bases were analysed by reverse transcription [15], and were run on 6 % (w/v) polyacrylamide/urea sequencing gels, alongside sequencing reactions. Unmodified RNA was sequenced in a reaction mix containing 1 μl of a 0.2 mM solution of one of the ddNTPs and reverse transcriptase.

RESULTS

Translation of IGF-II leader 2 mRNA

To examine the effect of IGF-II mRNA leader 2 on translation *in vivo*, leader 2 mRNA was fused to a luciferase reporter, and the construct was transiently expressed in human RD rhabdomyosarcoma or SK-N-MC neuroblastoma cells (Figure 2A, upper panel). IGF-II leader 4 constructs and a luciferase control were also included for comparison. Whereas translation of leader 4–luciferase mRNA was comparable with that of the basic pcDNA 3.1 luciferase mRNA in both RD and SK-N-MC cells, translation of the leader 2 mRNA was approx. three times more efficient than that of the luciferase control RNA in RD cells, and translation of the two transcripts was approximately the same in SK-N-MC cells. In control experiments, the levels of the control, leader 2 and 4 mRNAs were examined by Northern blot analysis, and this showed that similar amounts of mRNA were expressed from each of the constructs (Figure 2A, lower panel). To determine whether translation initiation of leader 2 mRNA was regulated by eIF4E, leader 2 and leader 4 luciferase constructs and a luciferase control were co-transfected with increasing amounts of a plasmid encoding eIF4E (Figure 2B). Whereas the luciferase control and leader 4 mRNA responded with a dose-dependent increase in translation, eIF4E had no significant effect on the translation of leader 2-containing luciferase mRNA.

Identification of an IRES in IGF-II leader 2

Since eIF4E had no effect on leader 2 translation, it was conceivable that the leader was able to initiate translation via an IRES. To examine this possibility, leader 2 was inserted between the *Firefly* and *Renilla* luciferase-coding regions of the bicistronic pBicFire vector (Figure 3A), and transfected into RD cells. The IGF-II leader 4, gastrin and BiP mRNA leaders, and the EMCV and poliovirus IRESs, were included for comparison (Figure 3B). Whereas insertion of IGF-II leader 4 or the gastrin leader had no significant effect on translation of the downstream cistron, leader 2 and the BiP leader increased translation of *Renilla* luciferase approx. 4-fold (Figure 3B). The EMCV and poliovirus IRESs were moderately active in RD cells and increased translation of the second cistron 4- and 9-fold respectively. To locate the putative IRES, leader 2 was divided into the segments depicted in Figure 3(C) and examined as above. Whereas segments comprising nt 1–300, 100–300, 200–300 and 300–400 exhibited a reduced activity compared with the full-length leader, the activity of segment 100–400 was similar to that of the full-length construct, and segment 200–400 increased translation 11-fold (Figure 3C). Insertion of two copies of the 100–400 segment doubled its activity from 3.5-fold to approx. 8-fold over the basal level. To exclude the possibility that translation of the downstream cistron was caused by re-initiation of ribosomes following translation of the first cistron, a stable hairpin was inserted in front of the *Firefly* luciferase-coding region in pBicFire with segment 200–400 in the intercistronic spacer. The hairpin reduced translation of the first cistron by > 90 %, whereas translation of the second cistron remained unaffected (Figure 3C;
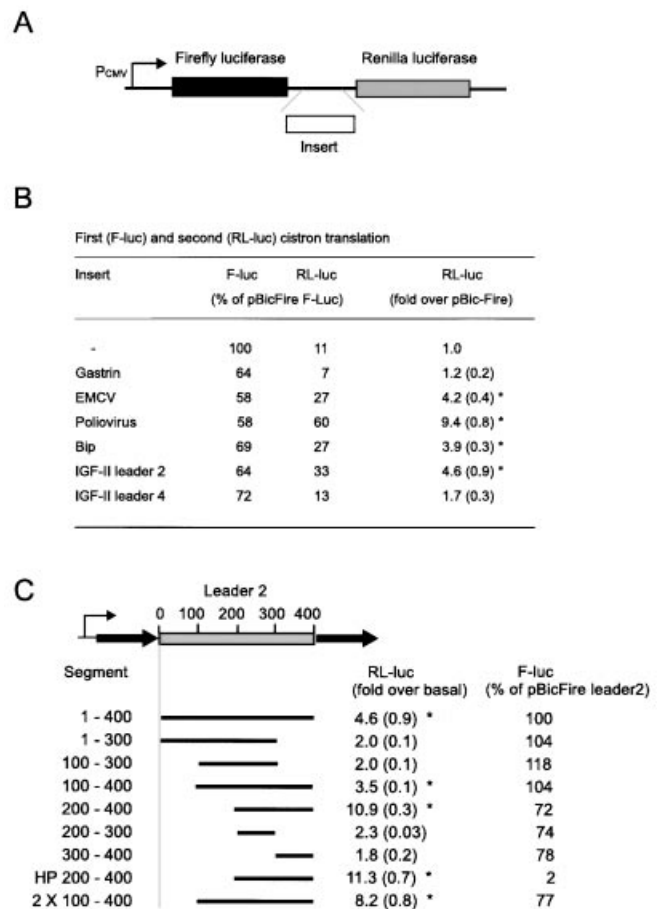

Figure 3   Identification of an IRES in IGF-II leader 2

(**A**) Schematic representation of the bicistronic pBicFire. The coding regions of *Firefly* and *Renilla* luciferase were inserted downstream of a CMV promoter, and the 5′ UTRs were placed in the intercistronic spacer. (**B**) pBicFire constructs with gastrin, Bip or IGF-II leaders and viral IRESs in the intercistronic spacer were transiently expressed in RD cells. The leaders are indicated at the left, and the relative activities of first (F-luc) and second (RL-Luc) cistrons compared with pBicFire first cistron (F-Luc) are shown in the left and middle columns respectively. The right column shows the calculated fold increase. The results are shown as means ± S.E.M. for three experiments. (**C**) Deletion analysis of IGF-II leader 2. The indicated fragments were inserted into pBicFire and transfected into RD cells, as described above. 'HP 200–400' designates that a stable hairpin was inserted upstream of the reading frame of *Firefly* luciferase in pBicFire containing nt segment 200–400, whereas '2 × 100–400' indicates that the vector contained two copies of the segment. (*$P < 0.05$).

HP 200–400). Moreover, to examine whether segment 200–400 disrupted the integrity of the bicistronic mRNA or functioned as a separate promoter, we performed a Northern blot analysis on RD cells transfected with pBicFire, or pBicFire with the EMCV leader, leader 2 segment 200–400 and leader 2 segment 100–200 bicistronic constructs and hybridized with a *Renilla* probe. In addition, fragments were inserted into pGL3-Basic to reveal a cryptic promoter activity (Figures 4A and 4B). Results from the Northern blot analysis did not disclose the presence of aberrant transcripts corresponding to cleaved bicistronic full-length mRNA containing segment 200–400, but a weak cross-reactivity to 18 S RNAs and a high-molecular-mass extended transcript were apparent. The activity of pGL3-Basic with segment 200–400 was < 0.5 % of pGL3-SV40 and identical with that of pGL3-Basic. Taken together, these data indicate that leader 2 directs IRES-mediated translation from a *cis*-acting element residing between nt 200 and 400.
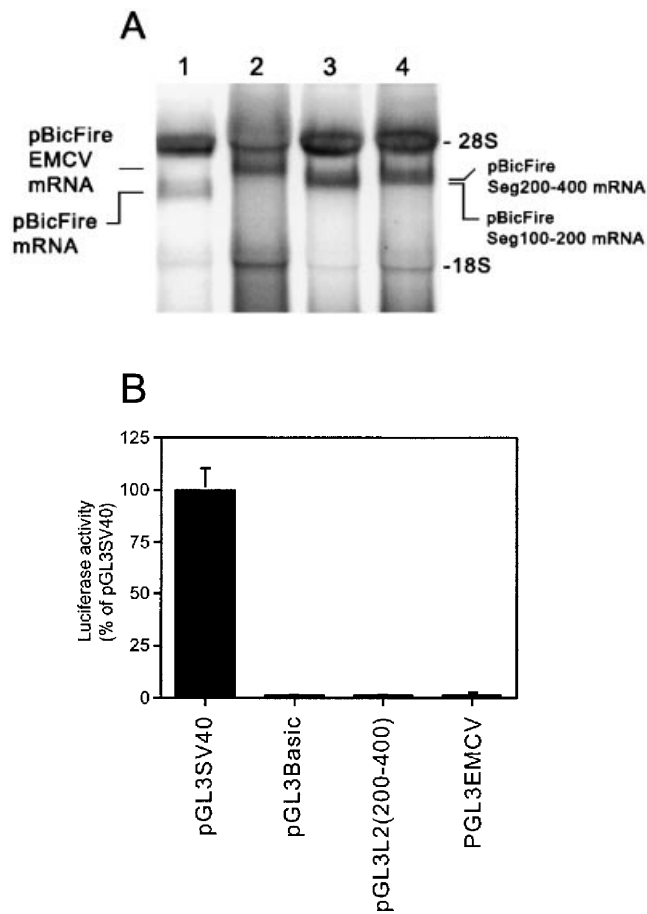
A

B

Figure 4   IGF-II leader 2 segment 200–400 mRNA expression and promoter activity

(**A**) RD cells were transfected with pBicFire (lane 1), pBicFire EMCV (lane 2), pBicFire segment 200–400 (lane 3) and pBicFire segment 100–200 (lane 4), before the total RNA was isolated and analysed by Northern blot analysis. The blot was hybridized with a full-length *Renilla* cDNA probe. (**B**) RD cells were transfected with pGL3-SV40, pGL3-Basic, pGL3-Basic with leader 2 segment 200–400 or pGL3-Basic with the EMCV leader, as indicated, and the luciferase activity was determined. The data are expressed as percentages of pGL3-SV40 activity and are shown as means ± S.E.M.

Functional characterization of the IGF-II leader 2 IRES

To determine the relative significance of cap-dependent and IRES-mediated leader 2 translation, and whether IRES-mediated translation initiation involved scanning by the 43 S complex from the recruitment element to the AUG codon, artificial optimized AUG start codons were inserted at positions 92 and 348 in mono- and bi-cistronic leader 2 constructs (Figure 5A). The artificial AUGs in the bicistronic vector did not affect translation of the upstream cistron, and they resulted in small out-of-frame upstream reading frames of 10 and 37 amino acids respectively. Whereas the upstream AUG_{92} did not reduce translation in either the monocistronic or the bicistronic construct, insertion of AUG_{348} essentially abolished translation of the downstream cistron in both mono- and bi-cistronic contexts (Figure 5B). Since this indicated that the internal initiation involved scanning of the 43 S ribosomal pre-initiation complex through the terminal part of the leader, we fused the conserved segment 200–300 of leader 2 (see below) to the 78-nt-long gastrin

A

B

C
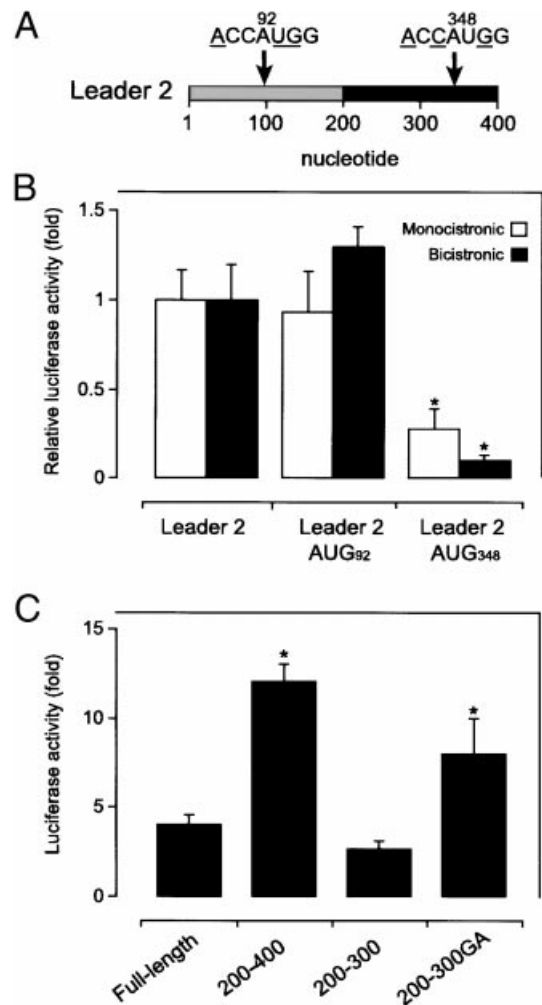
Figure 5   Insertion of artificial AUG start codons in leader 2

(**A**) Artificial optimized AUG start codons at positions 92 and 348 were inserted into mono- and bi-cistronic leader 2 constructs, as indicated. Substituted nucleotides are underlined, and the resulting Kozak consensus sequences are shown. (**B**) Expression of mono- and bi-cistronic leader 2 constructs containing artificial AUGs in position 92 or 348 in RD cells. Results are expressed relative to the mono- and bi-cistronic leader 2 constructs, respectively, and are shown as means ± S.E.M. values for three independent experiments. (**C**) Activity of leader 2 segment 200–300 after fusion to the 78-nt-long gastrin 5′ UTR ('200–300GA'). *Renilla* luciferase activities originating from the downstream coding region were normalized to the activity of the upstream firefly luciferase activity. The results are given as means ± S.E.M. for three experiments. (*$P < 0.05$).

mRNA leader (Figure 5C). Segment 200–300 alone was incapable of promoting internal initiation, but IRES activity was partially restored after fusion to the gastrin leader. Taken together, these results indicate that the leader 2 IRES can be separated into a putative recruitment element and a spacer region, which is scanned by the 43 S ribosomal pre-initiation complex to the translation start codon.

Structure of IGF-II leader 2 IRES

Alignment of leader 2 from five different species shows that the segment comprising nt 200–321 is highly conserved (84 % identity), compared with the segments comprising nt 1–200 and 322–400, which exhibit 40 % and 49 % identity respectively (Figure 6A). The most striking feature of the element containing

A

```
           1            20            40                    60            80            100
Human  ----GUUCCGGCUCGCCGGCCACCCACUGCAGCUUCCCCAACCCCGCGCACAGCG---GGCACUG---GUUUCGGGCCUCUCU--GUCUCCUACGAAGUCCCCAGAGCAACUCGGAUUUG
Sheep  AUUGCUCCCGGCCCCAGCGCGGUCU-CUGCAGCUUCCGCGGGCCC-GCACCCCGCG---GGCACCG---GUUCUGGGCCUCUCCGCGUCUGCUGCGAAGUCCCCGGAGCUCCUUGGAUGCG
Horse  ---ACUUGCGUCCCC--CGGUAUCU-GCGCGUCGUUCCCGGCCC-CAACCCCGCGCACAGCACCACUAGUUUCGGGCCUCUCCGAGUUUCGUACGGAGUCCCCCGAGCUCUCAGAUUUG
Mouse  --------------------------GGCAGCUUCC--------CGUCCUGCGCAUAACAGUG---GUAUCGG-CCUCGUCACUUCUCCUACGGUGUCCCAGCGGCAGCACAGAUUUU
Rat    --------------------------GGCAGCUUCC--------CGUCCUGCGCAUAACAUCG--AUAUCGG-CCUCGUCACUUCUCCUACGGUGUCCCAGCGGCAGCAGAUUUU
                                 *  *  *        *  ***    **     *   ** ****     * *  * **   *****   ** *  ***

          120           140           160           180           200           220
Human  GGAAAUUUCUCUCUAGCGUUGCCCAAACACAC--UUGGGUCGGCCGCGCGCCC-UCAGGACGUGGACAGGGAGGGCUUC-CCCGUGUCCAGGAAAGCGACCGGGCAUUGCCCCCAGUCUCC
Sheep  GGAAGUUUCUCUC--GUCCUCCUCACACACAU-UCGGUCGGGCAGCUUGUCCGCCAGGCAGUAGGCAGGG----CUUCUCCCCGCGUCCCAG---AACGACUGGGCAUUGCCCCCAGUUUCC
Horse  GGAAAUUUAUCUCU-GUACUCCCCAAACACAC-UUGGUUCGGCAGUGCGUCCGCCAGGACGUGGGCAGU-----CUUCUCUCGAGUCUUAGGAAAACGACUGGGGCAUUGCCCCCAGUUUCC
Mouse  GGAAAAGUGU-----GUGUGCCCCAAACACGCACUGGAC--ACUGCGCGGCGGCCAGGACGUGGGCAGUG----UUGCUCCCGUGUCCAGGAAAACGACUGG-CAUUGCCCCCAGUUUCC
Rat    GGAAAAGUGU-----GUGUGCCCCAAACACGCACCGGAC--ACUGCGCGGCGGCCAGGACGUGGGCAGUG----CUGCUCCCGUGUCCAGGAAAACCACUGGGGCAUUGCCCCCAGUUUCC
       ****   * *      *  ** ****    **    *  **    *  *   **** ** * ***   * *  **  *** **   *  * ** **  ************* ***

          240           260           280           300           320
Human  CCCAAAUUUGGGCAUUGUCCCGGGUCUUCCAACGGACUGGGCGUUGCUCCCGGACACUGAGGACUGGCCCCGGGGUCUCGCUCACCUUCAGCAGCGUCCAC--------CGCCUGCC--
Sheep  CCCAAAUUUGGGCAUUGUCCCGGGUCUUCCAACGGACUGGGCGUUGCCCCGG-ACGCUGGGGACUGCCCC-GGGGUCUCGCUCACCUUCAGC-GCGCCUAC-------CGCCCGCC--
Horse  CCCAAAUUAUGGGCAUUGUCCCCAAACACAC-UUGGUUCGGCAGUGCGUCCGCCAGGACGUGGGCAGU-----CUUCUCUCGAGUCUUAGGAAAACGACUGGGGCAUUGCCCCCAGUUUCC
Mouse  CCCAAAUUUGGGCAUUGUCCCGGGUCUUCCAACGGACUGG-CGUUGCCCCGGUUCACUGGGGACUGCCCUUGGGGUCUCGCUCACCUUCAGCACCGUCCACUUCAGCAGCUCCCCACUUC
Rat    CCCAAAUUUGGGCAUUGUCCCGGGUCUUCCAACGGACUGGGCGUUGCCCGGGGACACUGGGGACUGCCUUUGGGGUCUCGCUCACCUUCAGCAACGUCCACUUCAGCAGCUCCCACUUC
       *********************** ****** ******   *  *    *  * *   ** *****  *  *   ***** ************* **  *        * **  *

          340           360           380           400
Human  -ACAGAGCGUUCGAUCGCUCGCUGCCUGAGCUCCUGGUGCGCCCGCGGGACGCAGCCUCC-AGCU-UCGCG-------
Sheep  -GCAGAUCGCUCG----CUCGCUGCCUCGACUCCUC-CGCGCCCGC-GACGCAGCCUUC-GGCC-UCCAGCCUCUGCG
Horse  -GCAGAGCGCUCG----CUCGCUGCUUCGGCUCCUGGCGGCCGCCCGCAGACGCAGCCUCCCAGCAG-UCCAG------
Mouse  AUCCGCUCUCCCCG---CCCUUCUCCUCCCGA-UCCUCCUGCGCCAGCGGACCCGACCUUC-GGCCCUUGUG-------
Rat    UGCAGCUCUCCCCG---CCCUUCU----CCGA-UCCUCCUGCGCCA-CGGACCCGACCUUC-GGCC-UUGCG-------
        *  *    **     *   **     ****  *****  * ***  * *** *  **  *      *
```
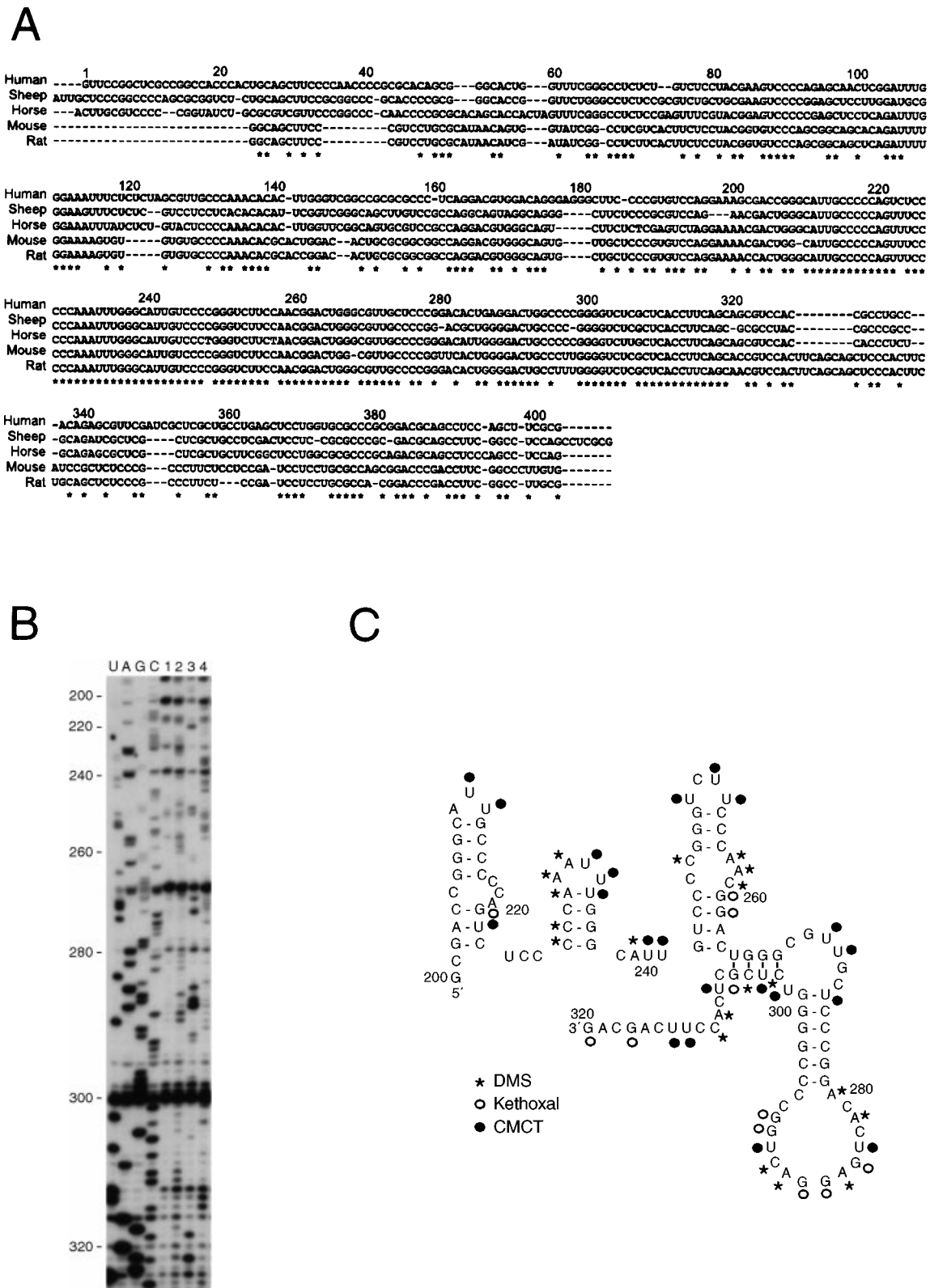
B

C

★ DMS
○ Kethoxal
● CMCT

Figure 6   Sequence and structural probing of human leader 2

(**A**) The sequence of human leader 2 (GenBank® X53038) was aligned with the corresponding sequences of sheep (GenBank® U00664), horse (GenBank® AF020598), mouse (GenBank® M36329) and rat (GenBank® X14833). Identical nucleotides are indicated with an asterisk. (**B**) Leader 2 RNA was incubated with buffer (track 1), or modified with CMCT (track 2), kethoxal (track 3) or DMS (track 4), and the products were analysed by primer extension. The dideoxynucleotide sequence tracks are shown at the left. A modified base gives rise to a reverse transcript, that is one nucleotide shorter than the corresponding band in the sequencing track. (**C**) Putative secondary structure of the 200–320 region in leader 2. Bases chemically modified by DMS, kethoxal or CMCT are as shown.

nt 200–321 is the five GGG repeats, but the leader does not share any obvious similarities to other viral or cellular IRESs or 5′ UTRs. On the basis of the above results, which indicate that the conserved nt 200–321 region is required for recruitment of the 43 S ribosomal pre-initiation complex, we determined a putative secondary structure for this region by probing chemically with CMCT, DMS and kethoxal (Figures 6B and 6C). The region comprising nt 200–321, in the context of both the full-length leader and the isolated 200–400 segment (results not shown), was subjected to modification. The probing pattern was identical in the two sets of experiments, showing that the nt 200–321 region exhibits a low Watson–Crick structuring potential, with four small stem–loop structures. The element does not exhibit any apparent structural similarities to the conserved GNRA loops or 'Y'-shape-like structures reported previously, which have been identified in some viral and cellular IRESs [11,16].

DISCUSSION

IGF-II leader 2 mRNA is expressed during fetal life and in transformed cells. Leader 2 is relatively long (411 nt), and its mRNA is translated efficiently, particularly in fetal liver [17,18]. In agreement with these data, we found that chimaeric leader 2–luciferase mRNA was efficiently translated, particularly in RD cells that express high levels of endogenous IGF-II mRNA [19]. Whereas translation of chimaeric leader 4 mRNA was increased dose-dependently by the overexpression of eIF4E, leader 2-mediated translation was refractory to this factor, suggesting that it is translated by a cap-independent mechanism. However, we cannot exclude the possibility that conventional scanning also plays a role, since leader 2 is efficiently translated by this process *in vitro* [17]. Moreover, the phylogenetic pressure on the absence of upstream AUG codons actually suggests that a scanning mechanism might be operational under physiological conditions. Bifunctional leaders have been identified in other cellular mRNAs, such as the c-Myc and basic fibroblast growth factor mRNAs [20,21], and ornithine decarboxylase mRNA is translated by conventional scanning during G_1 phase and by an IRES-mediated initiation during the G_2/M phase [22].

Insertion of leader 2 segments into a bicistronic vector identified a single IRES in the region comprising nt 200–400. The presence of an IRES in this region was supported further by the lack of translational repression following the insertion of an upstream AUG codon at position 92 in the monocistronic construct, and a hairpin before the first cistron in the bicistronic constructs. Moreover, there was no indication that the segment disrupted the integrity of the bicistronic mRNA, or that the element functioned as a cryptic promoter. The activity of the isolated element was significantly higher than that of the full-length leader, and insertion of two copies of the region doubled this enhancement, indicating that the IRES functions as a module independently of the remaining leader sequence. The sequence and secondary structure of the leader 2 IRES do not exhibit any obvious similarities to other IRESs. In particular, there was no evidence of either the 'Y'-shape structures that have been identified in some viral and cellular IRESs [11,23] or a pyrimidine-rich stretch that separates the *cis*-acting IRES domain from the initiation codon [24]. At the functional level, there is a resemblance with the family of rhino-/entero-virus IRESs [25], which consists of a putative recruitment element followed by a spacer that is scanned by the pre-initiation complex. The element resides in the nt 200–320 region, and it is fairly unstructured, at least in the context of Watson–Crick base-pairing, although highly conserved. The following 80 nt comprise the putative spacer region that could be replaced with the completely unrelated gastrin leader sequence, leading to the near-complete restoration of IRES activity. The depicted hairpins of the recruitment element are not stable under probing conditions, so it is feasible that recruitment of the 43 S pre-initiation complex or binding of *trans*-acting factors could involve the direct recognition of primary sequence elements in the IRES, as recently described for Gtx homeodomain mRNA [9,26].

Most IRES-containing mRNAs encode growth factors or proto-oncogenes with crucial roles in cell division and survival [27]. Typically, the IRES ensures that the transcripts are translated during cellular stress, apoptosis and at the G_2/M transition during mitosis, where the supply of eIF-4E is compromised. In this way, IRES-mediated translation of vascular endothelial growth factor and basic fibroblast growth factor becomes activated during hypoxia [28,29] and heat shock [30] respectively, whereas the IRES-containing X-linked inhibitor-of-apoptosis and c-Myc mRNAs are translated during the initial phases of programmed cell death [31,32]. Moreover, internal initiation of basic fibroblast growth factor mRNA was shown recently to be both tissue-specific and strongly enhanced during late embryogenesis and in the adult brain [33]. IGF-II is a fetal growth factor that promotes exponential growth of the embryo during late development [34]. At the genomic level, IGF-II expression is controlled by imprinting [35], but otherwise the production of fetal IGF-II relies heavily on post-transcriptional mechanisms, thus allowing a rapid control of gene expression during crucial developmental stages. Similar to the 5′ UTRs in other mRNAs involved in growth and cell-cycle control, IGF-II leaders 2 and 3 are relatively large and provide extensive platforms for the recognition of specific *trans*-acting factors, which may confer selective functions to the transcripts. Since IGF-II leader 2 exhibits an IRES, the transcript might provide a continuous production of IGF-II and growth of the embryo during fetal stress conditions and mitosis. Moreover, it rationalizes the presence of an IRES in the IGF-II/mannose 6-phosphate receptor [27], which mediates the rapid internalization and degradation of IGF-II, thereby maintaining a careful balance between the production and inactivation of IGF-II in the expanding embryo.

REFERENCES

1   Nielsen, F. C. (1992) The molecular and cellular biology of insulin-like growth factor II. Progr. Growth Factor Res. **4**, 257–290

2   Nielsen, F. C., Ostergaard, L., Nielsen, J. and Christiansen, J. (1995) Growth-dependent translation of IGF-II mRNA by a rapamycin-sensitive pathway. Nature (London) **377**, 358–362

3   Nielsen, J., Christiansen, J., Lykke-Andersen, J., Johnsen, A. H., Wewer, U. M. and Nielsen, F. C. (1999) A family of insulin-like growth factor II mRNA-binding proteins represses translation in late development. Mol. Cell. Biol. **19**, 1262–1270

4   Merrick, W. C. and Hershey, J. W. (1996) The pathway and mechanism of eukaryotic protein synthesis. In Translation Control (Hershey, J. W., Mathews, M. B. and Sonenberg, N., eds.), pp. 31–69, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY

5   Mathews, M. B., Sonenberg, N. and Hershey, J. W. (1996) Origins and targets of translational control. In Translation Control (Hershey, J. W., Mathews, M. B. and Sonenberg, N., eds.), pp. 1–29, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY

6   Hellen, C. U. T. and Sarnow, P. (2001) Internal ribosome entry sites in eukaryotic mRNA molecules. Genes Dev. **15**, 1593–1612

7   Macejak, D. G. and Sarnow, P. (1991) Internal initiation of translation mediated by the 5′ leader of a cellular mRNA. Nature (London) **353**, 90–94

8   Le Quesne, J. P., Stoneley, M., Fraser, G. A. and Willis, A. E. (2001) Derivation of a structural model for the c-myc IRES. J. Mol. Biol. **310**, 111–126

9   Chappell, S. A., Edelman, G. M. and Mauro, V. P. (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. U.S.A. **97**, 1536–1541

10 Hayashi, S., Nishimura, K., Fukuchi-Shimogori, T., Kashiwagi, K. and Igarashi, K. (2000) Increase in cap- and IRES-dependent protein synthesis by overproduction of translation initiation factor eIF4G. Biochem. Biophys. Res. Commun. **277**, 117–123

11 Sella, O., Gerlitz, G., Le, S. Y. and Elroy-Stein, O. (1999) Differentiation-induced internal translation of c-sis mRNA: analysis of the cis elements and their differentiation-linked binding to the hnRNP C protein. Mol. Cell. Biol. **19**, 5429–5440

12 Holcik, M. and Korneluk, R. G. (2000) Functional characterization of the X-linked inhibitor of apoptosis (XIAP) internal ribosome entry site element: role of La autoantigen in XIAP translation. Mol. Cell. Biol. **20**, 4648–4657

13 Mitchell, S. A., Brown, E. C., Coldwell, M. J., Jackson, R. J. and Willis, A. E. (2001) Protein factor requirements of the Apaf-1 internal ribosome entry segment: roles of polypyrimidine tract binding protein and upstream of N-ras. Mol. Cell. Biol. **21**, 3364–3374

14 Chomczynski, P. and Sacchi, N. (1987) Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. Anal. Biochem. **162**, 156–159

15 Lykke-Andersen, J. and Garrett, R. A. (1994) Structural characteristics of the stable RNA introns of archaeal hyperthermophiles and their splicing junctions. J. Mol. Biol. **243**, 846–855

16 Robertson, M. E., Seamons, R. A. and Belsham, G. J. (1999) A selection system for functional internal ribosome entry site (IRES) elements: analysis of the requirement for a conserved GNRA tetraloop in the encephalomyocarditis virus IRES. RNA **5**, 1167–1179

17 De Moor, C. H., Jansen, M., Bonte, E. J., Thomas, A. A., Sussenbach, J. S. and Van den Brande, J. L. (1994) Influence of the four leader sequences of the human insulin-like-growth-factor-2 mRNAs on the expression of reporter genes. Eur. J. Biochem. **226**, 1039–1047

18 De Moor, C. H., Jansen, M., Sussenbach, J. S. and Van den Brande, J. L. (1994) Differential polysomal localization of human insulin-like-growth-factor-2 mRNAs in cell lines and foetal liver. Eur. J. Biochem. **222**, 1017–1024

19 Nielsen, F. C., Gammeltoft, S. and Christiansen, J. (1990) Translational discrimination of mRNAs coding for human insulin-like growth factor II. J. Biol. Chem. **265**, 13431–13434

20 Stoneley, M., Paulin, F. E., Le Quesne, J. P., Chappell, S. A. and Willis, A. E. (1998) c-Myc 5′ untranslated region contains an internal ribosome entry segment. Oncogene **16**, 423–428

21 Arnaud, E., Touriol, C., Boutonnet, C., Gensac, M. C., Vagner, S., Prats, H. and Prats, A. C. (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

22 Pyronnet, S., Pradayrol, L. and Sonenberg, N. (2000) A cell cycle-dependent internal ribosome entry site. Mol. Cell **5**, 607–616

23 Le, S. Y. and Maizel, Jr, J. V. (1997) A common RNA structural motif involved in the internal initiation of translation of cellular mRNAs. Nucleic Acids Res. **25**, 362–369

24 Gan, W., LaCelle, M. and Rhoads, R. E. (1998) Functional characterization of the internal ribosome entry site of eIF4G mRNA. J. Biol. Chem. **273**, 5006–5012

25 Ehrenfeld, E. (1996) Initiation of translation by picornavirus RNAs. In Translation Control (Hershey, J. W., Mathews, M. B. and Sonenberg, N., eds.), pp. 549–553, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY

26 Hu, M. C., Tranque, P., Edelman, G. M. and Mauro, V. P. (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. U.S.A. **96**, 1339–1344

27 Johannes, G., Carter, M. S., Eisen, M. B., Brown, P. O. and Sarnow, P. (1999) Identification of eukaryotic mRNAs that are translated at reduced cap binding complex eIF4F concentrations using a cDNA microarray. Proc. Natl. Acad. Sci. U.S.A. **96**, 13118–13123

28 Stein, I., Itin, A., Einat, P., Skaliter, R., Grossman, Z. and Keshet, E. (1998) Translation of vascular endothelial growth factor mRNA by internal ribosome entry: implications for translation under hypoxia. Mol. Cell. Biol. **18**, 3112–3119

29 Huez, I., Creancier, L., Audigier, S., Gensac, M. C., Prats, A. C. and Prats, H. (1998) Two independent internal ribosome entry sites are involved in translation initiation of vascular endothelial growth factor mRNA. Mol. Cell. Biol. **18**, 6178–6190

30 Vagner, S., Touriol, C., Galy, B., Audigier, S., Gensac, M. C., Amalric, F., Bayard, F., Prats, H. and Prats, A. C. (1996) Translation of CUG- but not AUG-initiated forms of human fibroblast growth factor 2 is activated in transformed and stressed cells. J. Cell Biol. **135**, 1391–1402

31 Holcik, M., Lefebvre, C., Yeh, C., Chow, T. and Korneluk, R. G. (1999) A new internal-ribosome-entry-site motif potentiates XIAP-mediated cytoprotection. Nat. Cell Biol. **1**, 190–192

32 Stoneley, M., Chappell, S. A., Jopling, C. L., Dickens, M., MacFarlane, M. and Willis, A. E. (2000) c-Myc protein synthesis is initiated from the internal ribosome entry segment during apoptosis. Mol. Cell. Biol. **20**, 1162–1169

33 Créancier, L., Morello, D., Mercier, P. and Prats, A. C. (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

34 DeChiara, T. M., Efstratiadis, A. and Robertson, E. J. (1990) A growth-deficiency phenotype in heterozygous mice carrying an insulin-like growth factor II gene disrupted by targeting. Nature (London) **345**, 78–80

35 DeChiara, T. M., Robertson, E. J. and Efstratiadis, A. (1991) Parental imprinting of the mouse insulin-like growth factor II gene. Cell **64**, 849–859