ARTICLE
Journal of Cellular Biochemistry 110:531–544 (2010)

Journal of Cellular Biochemistry

The Human *IGF1R* IRES Likely Operates Through a Shine–Dalgarno-Like Interaction With the G961 Loop (E-Site) of the 18S rRNA and Is Kinetically Modulated by a Naturally Polymorphic polyU Loop

Zheng Meng,[1] Nateka L. Jackson,[2] Oleg D. Shcherbakov,[1] Hyoungsoo Choi,[2] and Scott W. Blume[1,2,3*]

[1] *Department of Biochemistry and Molecular Genetics, University of Alabama at Birmingham, Birmingham, Alabama 35294*

[2] *Department of Medicine, University of Alabama at Birmingham, Birmingham, Alabama 35294*

[3] *Comprehensive Cancer Center, University of Alabama at Birmingham, Birmingham, Alabama 35294*

ABSTRACT

IGF1R is a proto-oncogene with potent mitogenic and antiapoptotic activities, and its expression must be tightly regulated to maintain normal cellular and tissue homeostasis. We previously demonstrated that translation of the human *IGF1R* mRNA is controlled by an internal ribosome entry site (IRES), and delimited the core functional IRES to a 90-nucleotide segment of the 5′-untranslated region positioned immediately upstream of the initiation codon. Here we have analyzed the sequence elements that contribute to the function of the core IRES. The Stem2/Loop2 sequence of the IRES exhibits near-perfect Watson–Crick complementarity to the G961 loop (helix 23b) of the 18S rRNA, which is positioned within the E-site on the platform of the 40S ribosomal subunit. Mutations that disrupt this complementarity have a negative impact on regulatory protein binding and dramatically decrease IRES activity, suggesting that the *IGF1R* IRES may recruit the 40S ribosome by a eukaryotic equivalent of the Shine–Dalgarno (mRNA–rRNA base-pairing) interaction. The homopolymeric Loop3 sequence of the IRES modulates accessibility and limits the rate of translation initiation mediated through the IRES. Two functionally distinct allelic forms of the Loop3 poly(U)-tract are prevalent in the human population, and it is conceivable that germ-line or somatic variations in this sequence could predispose individuals to development of malignancy, or provide a selectable growth advantage for tumor cells. J. Cell. Biochem. 110: 531–544, 2010.

KEY WORDS: IGF1R; REGULATION OF GENE EXPRESSION AT THE TRANSLATIONAL LEVEL; 5′-UNTRANSLATED SEQUENCE; ALLELIC VARIATION

The type 1 insulin-like growth factor receptor (*IGF1R*) is a proto-oncogene with potent mitogenic, antiapoptotic, and transforming capabilities [Gooch et al., 1999; Pandini et al., 1999; LeRoith and Roberts, 2003; Kurmasheva and Houghton, 2006; Yanochko and Eckhart, 2006; Kim et al., 2007]. Precise control of *IGF1R* expression is critical for maintenance of normal cellular and tissue homeostasis and avoidance of malignant transformation, and this requires specific regulation at both the transcriptional and translational levels [Cooke and Casella, 1994]. The *IGF1R* mRNA contains an extraordinarily long 5′-untranslated region (1,040 nucleotides) [Cooke et al., 1991], projected to adopt a highly internally base-paired structure ($dG > -500$ kcal/mol) which represents a substantial impediment to ribosomal scanning. We discovered an internal ribosome entry site (IRES) within the human *IGF1R* 5′-UTR which provides an alternative mechanism for *IGF1R* translation initiation, allowing the 40S ribosome to bypass the obstacles presented by the highly structured 5′-UTR [Meng et al., 2005]. The *IGF1R* IRES was authenticated by its sensitivity to

Grant sponsor: National Institutes of Health, National Cancer Institute; Grant number: R01 CA108886.

Zheng Meng's present address is Division of Biology, California Institute of Technology, Pasadena, CA 91106.

Hyoungsoo Choi's present address is Department of Pediatrics, Seoul National University Bundang Hospital, Gyeonggi-do 463-707, Korea

*Correspondence to: Dr. Scott W. Blume, University of Alabama at Birmingham, 845 19th Street South, BBRB 765, Birmingham, AL 35294. E-mail: scott.blume@ccc.uab.edu

Received 2 January 2010; Accepted 3 February 2010 • DOI 10.1002/jcb.22569 • © 2010 Wiley-Liss, Inc.
Published online 26 March 2010 in Wiley InterScience (www.interscience.wiley.com).

deletion of the promoter from a bicistronic construct, its resistance in a monocistronic context to co-expression of a viral 2A protease, and its ability to function in vitro under conditions that block cap-dependent translation initiation [Meng et al., 2008]. The core functional IRES was delimited to the 3′-terminal 90 nucleotides of the 5′-UTR, positioned immediately upstream of the initiation codon. Two sequence-specific RNA-binding proteins, HuR and hnRNP C, which compete for interaction with the IRES and differentially regulate its activity, were identified and characterized, and evidence for pathological dysregulation of the *IGF1R* IRES in human breast tumor cells relative to non-transformed breast epithelial cells through changes in activities of RNA-binding IRES-regulatory proteins was presented [Meng et al., 2005, 2008].

Here we have performed a detailed examination of the core functional IRES by site-directed mutagenesis within the context of the full-length 5′-UTR. We report two sequence elements, which appear to be very important for translation initiation mediated by the *IGF1R* IRES. The Stem2/Loop2 sequence appears to be a focal point for operation and regulation of the IRES. It serves as the recognition site for a group of IRES-regulatory proteins and may directly facilitate ribosome recruitment via direct Shine–Dalgarno-like base-pairing interaction with the 18S rRNA. The Loop3 sequence, which exists in two distinct allelic size variants in the human population, appears to govern the maximal rate of translation initiation through the *IGF1R* IRES by limiting accessibility of the core functional IRES to positive IRES *trans*-acting factors (ITAFs).

MATERIALS AND METHODS

CAGE TAG ANALYSIS

The human CAGE tag database was accessed at http://gerg01.gsc.riken.jp/cage/hg17 [Carninci et al., 2006]. The CAGE Basic Viewer for *Homo sapiens* was used to search chromosome 15 sorted on gene symbol for *IGF1R*, and all of the CAGE tags associated with the *IGF1R* locus were tabulated. The positions of the tags relative to the ATG initiation codon were calculated, and the number of tags mapping to each site was plotted.

mfold ANALYSIS OF RNA STRUCTURE

The mfold 3.2 algorithm [Mathews et al., 1999; Zuker, 2003] was used to analyze projected secondary structures for the *wild-type* and mutant *IGF1R* 5′-UTR/IRES sequences.

NUCLEASE SENSITIVITY ASSAYS FOR RNA STRUCTURE MAPPING

The full-length *IGF1R* 5′-UTR (1,040 nucleotides) and isolated IRES RNA (90 nucleotides) were synthesized in vitro using T7 RNA polymerase (RiboMax, Promega) and 3′-end-labeled using ^{32}P-pCp and T4 RNA ligase. The RNA was then subjected to limited digestion with RNase T1 (cleaves after single-stranded G residues) or RNase V1 (cleaves within double-stranded RNA). Alkaline hydrolysis and RNase T1 digestion performed following heat denaturation of the RNA were used to generate appropriate reference landmarks.

BICISTRONIC REPORTER CONSTRUCTS

The parent bicistronic construct containing the full-length *IGF1R* 5′-UTR cloned between the coding sequences for Renilla and firefly luciferases [pDual IGF1R (1–1038)) has been described, Meng et al., 2005]. The first series of site-directed mutations was designed using mfold and introduced into the parent construct using the QuikChange kit (Stratagene). Inserts for the second series of site-directed mutations were designed using mfold and generated by PCR amplification from T47D (biallelic) genomic DNA, capturing the natural short (U_{13}) and long (U_{21}) variations of the Loop3 sequence, and utilizing mutant primers to introduce compensatory alterations into Stem2.

Base substitutions are as follows: mutLoop2: replace residues 1022–1023 (UU) with GC; mutStem2: replace residues 1017–1018 (GG) with CC; mutStem2_comp: replace residues 1017–1018 (GG) with CC and residues 1027–1028 (CC) with GG; minLoop3: reduce 988–1008 $(U)_{21}$ tract to $(U)_6$ (four residues remaining in Loop3, two residues in Stem3); IRESmutStem2a: replace residues 67–69 (GGA) with CCU and replace residues 76–78 (UCC) with AGG; IRESmut-Stem2b: replace residues 66–68 (GGG) with UAU and replace residues 77–79 (CCC) with AUA; IRESdimLoop3: capture natural $(U)_{13}$ allele from T47D (biallelic) genomic DNA using *wild-type* primers.

CELL LINES

The T47D human breast carcinoma cell line was obtained from ATCC and grown in RPMI 1640 medium with 10% fetal calf serum, 10 μg/ml insulin, and penicillin/streptomycin, and maintained in a humidified environment at 37°C with 5% CO_2. The T98G human glioblastoma, Saos-2 human osteosarcoma, and MCF-10A non-transformed human mammary epithelial cell lines were propagated according to ATCC recommendations.

IRES ACTIVITY ASSAYS

T47D cells were seeded at 50% confluence (~100,000 cells per well) into 24-well plates and allowed 36 h for attachment and recovery prior to transfection using Lipofectamine 2000 (0.67 μl/well) and bicistronic reporter plasmid (100 ng/well). Forty-eight hours following transfection, cell were harvested, lysates prepared, and firefly and *Renilla* luciferase activities measured using the Dual Luciferase assay kit (Promega). All assays were performed in triplicate or quadruplicate and repeated at least 2–4 times. IRES activity was calculated as the ratio of firefly to Renilla luciferase activities, and in some figures, presented relative to f/R of the full-length *IGF1R* 5′-UTR (set = 100%), and in other figures, related to the f/R ratio obtained with the control bicistronic construct pDual (with no *IGF1R* insert and no IRES, set = 1). To measure the effects of HuR or hnRNP C on activity of the IRES mutants, expression constructs for each of these RNA-binding proteins were cotransfected along with the bicistronic reporter constructs (as previously described) [Meng et al., 2005, 2008].

EXAMINATION OF HIGH-RESOLUTION CRYSTAL STRUCTURE OF *T. THERMOPHILUS* 30S RIBOSOMAL SUBUNIT WITH mRNA

The three-dimensional structure for the *T. thermophilus* 30S ribosomal subunit with mRNA (PDB ID: 1JGO) [Yusupova et al.,

2001] was accessed from the Structure database of the NCBI. The image was rotated to achieve optimal simultaneous visualization of the G691 loop of the rRNA, Shine–Dalgarno region of mRNA, and anti-Shine–Dalgarno region of rRNA (roughly the intersubunit face view). Segments of interest were highlighted using the interactive viewing program Cn3D.

PCR AMPLIFICATION OF CORE *IGF1R* IRES SEQUENCE FROM HUMAN GENOMIC DNA

Human genomic DNA was obtained from commercial sources or prepared from cultured cell lines. The core functional IRES sequence was amplified using forward primer Apa-Seq: 5′-TCGGAG-TATTGTTTCCTTCGCCCTTG-3′ (nt 892–917) and reverse primer f: 5′-CGACGGTGGCAACTCGGGTTTCGC-3′ (nt 1201–1224); yielding a 333-bp product surrounding the core functional IRES (nt 951–1040). We tested several different PCR conditions in attempt to eliminate PCR error (most commonly +1, +2, or −1 U residues) through the long homopolymeric Loop3 sequence, but none was completely effective. We found that AmpliTaq combined with a buffer containing $(NH_4)_2SO_4$ in place of KCl and 10% glycerol as cosolvent enhanced the specificity and fidelity of this product more so than other polymerase/buffer combinations. PCR products were submitted for direct sequence analysis and also cloned into pSTblue1 and individual clones sequenced.

NORTHWESTERN ASSAY

The northwestern assay was performed essentially as described [Meng et al., 2008] except that a 10% gel was used and the magnesium concentration during incubation with probe was reduced to 2 mM. Briefly, cytoplasmic, nuclear, or whole cell extracts were prepared from T47D or MCF-10A cells, separated by SDS/PAGE, transferred to nitrocellulose, and then renatured sequentially in Tris-buffered saline followed by a detergent-free physiological buffer. Radiolabeled RNA probes (sense orientation) representing the *wild-type* or mutant full-length *IGF1R* 5′-UTR/IRES were synthesized using linearized monocistronic constructs as template and T7 RNA polymerase, with internal incorporation of α^{32}P-UTP. The probes were incubated with the blot in renaturation buffer for 2 h at room temperature, with gentle agitation, then washed three times (the first wash containing 1 mg/ml heparin) prior to autoradiographic exposure.

RESULTS

THE IRES WITHIN THE COMPLEX *IGF1R* 5′-UTR

In their original characterization of the human *IGF1R* mRNA, Cooke et al. [1991] determined by primer extension and RNase protection methods that the transcription start site was 1,038 bp upstream of the coding sequence. Data generated more recently by high-throughput genome-scale methodologies such as CAGE tag analysis allow us to re-examine this question (Fig. 1A). Although there is evidence for a minor population of *IGF1R* mRNA molecules that begin ~700 or ~900 nucleotides upstream of the initiation codon, and an additional small cluster of *IGF1R* transcription start sites ~50,000 bp further upstream (not shown), these new data confirm that the predominantly utilized transcription start site imparts an ~1,043-nucleotide long 5′-untranslated region to the *IGF1R* mRNA.

The long, G–C-rich 5′-untranslated region of the *IGF1R* mRNA is projected to adopt a secondary structure with extensive internal base-pairing (Fig. 1B) and high thermodynamic stability (dG = −515.2 kcal/mol). The core functional IRES, which had been localized to the 3′-terminal 90 nucleotides (951–1040) of the 5′-UTR [Meng et al., 2008], is projected to adopt a relatively simple structure composed of one central multi-loop and two terminal stem-loops (see bracket in Fig. 1B). Although the mfold algorithm suggested multiple base-pairing alternatives with similar thermodynamic stabilities available to the 5′-UTR as a whole, the core functional IRES itself appears to represent an independently folding domain, highly favored within the context of the full-length 5′-UTR. In fact, within the 20 highest scoring structures generated by mfold 3.2 for the 1,040-nucleotide 5′-UTR, 19 exhibited precisely the same 3-Loop folding pattern for the core functional IRES sequence. We were concerned, however, that the projected structures for the IRES could be influenced by its position so near the end of the sequence entered into the program. Therefore, we repeated the mfold analysis after adding the natural exon 1 and 2 sequences to the 5′-UTR (Fig. 1C). Although the structure projected for the 5′-UTR was dramatically altered by the presence of exons 1 and 2, the functional core IRES, now positioned near the middle of the sequence being analyzed, still preferentially adopted the same 3-Loop structure that had been highly favored in the original analysis.

A magnified image of the 3-Loop structure anticipated for the core functional IRES is presented as it might exist in the context of the full-length 5′-UTR (Fig. 1D), or as an isolated 90-nucleotide fragment (Fig. 1E) which displays even higher IRES activity than the full-length 5′-UTR [Meng et al., 2008]. We have designated the asymmetrical central multi-loop as Loop1, the six-nucleotide hairpin loop appended to a 4-bp helix as Loop2, and the large poly(U) sequence appended on a 6-bp stem as Loop3. RNase T1/V1 sensitivity data provide experimental support for the natural existence of this structure both in the intact 5'-UTR (Fig. 1F) as well as the isolated IRES (Fig. 1G). A symmetrical pattern of susceptibility to RNase V1 (which cleaves within double-stranded RNA) is observed for residues expected to lie within Stem2. G-residues within Stem2 (3 out of 4 base pairs are G·C) were partially or completely protected from RNase T1 digestion (which cleaves after single-stranded G residues) even in the absence of added magnesium, while G-residues within Stem3 (of which only 2 out of 6 base pairs are G·C) were more dependent on magnesium for base-pairing and protection from RNase T1 digestion.

We had previously noted that progressive 5′-deletions of the *IGF1R* 5′-UTR sequence were associated with progressive loss and then return of IRES activity [Meng et al., 2008]. We suspected that the loss of IRES activity observed with the intermediate deletions could be attributed to altered base-pairing patterns negatively impacting the core functional IRES. Indeed, we observe a strong correlation between measured IRES activity and the potential of the RNA to adopt the 3-Loop structure (Table I).

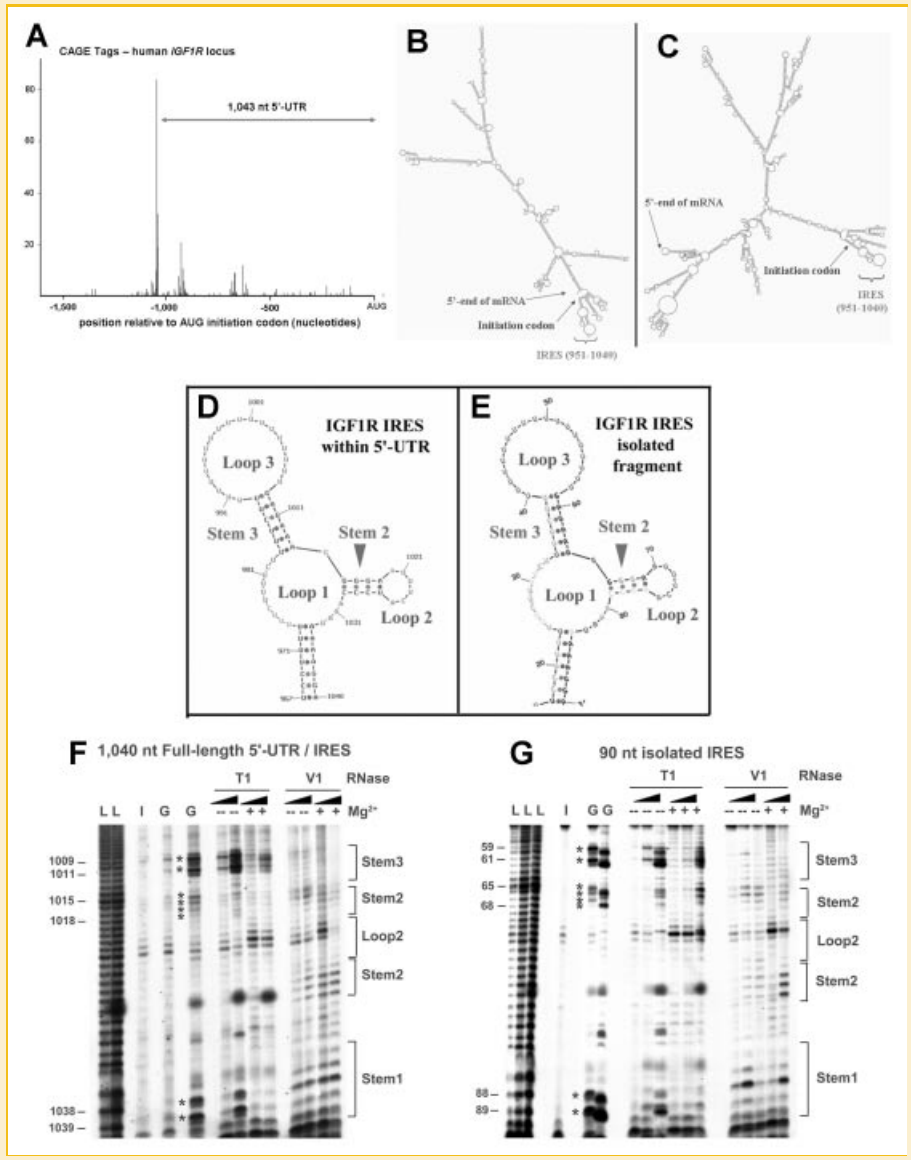

Fig. 1. The core functional IRES within the complex *IGF1R* 5′-UTR. A: The frequencies of CAGE tags (Cap Analysis of Gene Expression, derived from the publicly accessible database) [Carninci et al., 2006] associated with the human *IGF1R* locus, are plotted. The data confirm that the *IGF1R* mRNA contains an extraordinarily long 5′-untranslated region. B: Secondary structure projected for the human *IGF1R* 5′-UTR by the mfold 3.2 algorithm of Zuker et al. (dG = −515.2 kcal/mol). The IRES (marked by bracket) is projected to exist as an independently folding domain, highly favored within the 1,040-nucleotide 5′-UTR. C: Repeat mfold analysis of the *IGF1R* 5′-UTR including exon 1 and 2 sequences. There is no change in projected structure for the core functional IRES domain (bracket). D: Magnified view of structure projected for the core functional IRES domain within the context of the full-length 1,040-nucleotide 5′-UTR. E: Structure projected for the isolated core functional IRES (90 nucleotides). Note that the structural representations in this figure were produced using the Untangle mode of the mfold algorithm (helices widely separated for optimal visualization). F,G: Structural mapping by RNase sensitivity for the core functional IRES sequence, either within the context of the full-length *IGF1R* 5′-UTR (F) or alone (G). The full-length *IGF1R* 5′-UTR or the isolated IRES RNA was transcribed in vitro and 3′-end-labeled using ^{32}P-pCp and T4 RNA ligase, prior to treatment as follows: Lanes L: alkaline hydrolysis ladders. Lane I: intact RNA. Lanes G: RNase T1 digestion (0.01–1.0 U × 15 min at room temperature) performed following heat denaturation of RNA. The remaining lanes illustrate the pattern obtained following either RNase T1 digestion (cleaves single-stranded G residues) or RNase V1 digestion (preferentially cleaves double-stranded residues, 0.01–1.0 U × 15 min at room temperature) of the RNA, with or without added magnesium chloride (5 mM) as indicated. Products of digestion were separated on a 10% denaturing (8 M urea) polyacrylamide gel and exposed to autoradiographic film. Sequence coordinates are shown to the left of the gel, while the positions of projected structural features are indicated on the right. Asterisks on gel mark the positions of G residues.

SITE-DIRECTED MUTAGENESIS OF CRITICAL RESIDUES WITHIN THE *IGF1R* IRES

To further dissect the IRES, we introduced site-directed mutations within the core *IGF1R* IRES sequence, to identify individual sequence elements important for IRES operation and regulation. The initial series of mutations were introduced and tested within the context of the full-length 1,040-nucleotide 5′-UTR (Fig. 2A,B). We focused our attention on two features: Stem2/Loop2 and the homopolymeric Loop3.

The Stem2/Loop2 sequence was of particular interest because the hexaloop (AUUUCA) is very similar to a sequence element (UUUCC)

TABLE I. Correlation Between IRES Function and Capacity to Adopt the 3-Loop IRES Structure

5′-UTR sequence included	Relative % IRES activity[a]	Propensity to adopt IRES structure[b]
1–1040	100	95% (19/20)
205–1040	38	65% (13/20)
426–1040	2	0% (0/20)
677–1040	7	8% (1/12)
889–1040	34	20% (1/5)
951–1040	216	25% (1/4)

[a]Relative IRES activity is measured using the bicistronic construct in which the full-length *IGF1R* 5′-UTR or fragments thereof are positioned between the *Renilla* and firefly luciferase coding sequences. IRES activity is measured by the ratio of firefly to *Renilla* luminescence, and expressed relative to the f/R ratio obtained with the full-length 5′-UTR (nucleotides 1–1040, set to 100%), and also normalized to the f/R ratio obtained with the control construct (no *IGF1R* sequence insert and no IRES, set to 0%).
[b]Propensity to adopt the 3-loop structure illustrated in the figures is assessed by mfold 3.2 projections of up to 20 of the highest scoring secondary structures for each 5′-UTR fragment. Note that the two smallest fragments of the 5′-UTR, in which IRES activity returns, are capable of adopting the 3-Loop structure, but show less structural constraint than the full-length 5′-UTR.

thought to contribute to the function of several viral IRESs [Pilipenko et al., 1992; Scheper et al., 1994], and also very similar to the consensus (AUUUA) for the A–U-rich element binding protein HuR [Antic and Keene, 1997], which we had already determined binds to the IRES and potently regulates its activity [Meng et al., 2005]. We found that mutation of two adjacent nucleotides within either Loop2 or Stem2 resulted in ∼50% loss of IRES activity. Five additional mutants altering from one to four residues within Loop2 were also associated with ∼25–50% decrease in IRES activity (data not shown). A compensatory mutation in Stem2, modifying 4 out of 1,040 nucleotides and with no anticipated change in secondary structure, was associated with >80% loss of IRES activity, suggesting that the primary sequence of this region may be critical for internal ribosome entry.

To determine what effect the large poly(U)-tract comprising Loop3 might exert on IRES function, the (U_{19}) sequence was minimized to a tetraloop. Based on mfold analysis of the mutated sequence, no additional changes to the core IRES structure were anticipated (Fig. 2A). (In fact, the identical 3-Loop structure was projected for each of the top 20 scoring structures generated for both the Stem2 compensatory mutation and the Loop3 minimization mutation.) Minimization of the Loop3 poly(U)-tract was associated with a dramatic (∼2.5×) increase in IRES activity (Fig. 2B), suggesting that, when intact, this long homopolymeric sequence serves as a negative regulator of *IGF1R* IRES-mediated translation initiation.

To further establish the importance of Stem2/Loop2 and Loop3 to IRES function, we tested the mutant constructs in a series of representative human tumor cell lines in which the *IGF1R* IRES is active. Very similar results were obtained in T47D (breast carcinoma in which the *IGF1R* IRES was originally characterized, very high IRES activity), T98G (glioblastoma, moderately high IRES activity), and Saos-2 (osteosarcoma, modest but readily detectable IRES activity) (Fig. 2C). Note that the compensatory Stem2 mutation was severely debilitating to IRES activity, while the Loop3 minimization mutant exhibited a marked increase in IRES activity, in all three cell lines.

To confirm these results, a second series of mutations was designed within the context of the 90-nucleotide core functional IRES. A compensatory substitution within Stem2 retaining the natural G–C-rich sequence composition resulted in near 50% decrease in IRES activity, while a similar mutation which replaced the three G–C base pairs with A–U base pairs exhibited >80% loss of IRES activity (Fig. 2D,E). This result suggests that, in addition to the changes in primary sequence, weakening of this stem may also have severe consequences for IRES function. To further assess the effect of the size of the homopolymeric Loop3 sequence on IRES activity, a more modest diminution of the polyU-tract (from U_{21} to U_{13}, representing the naturally occurring allelic variant of this sequence, derived by PCR amplification from T47D genomic DNA, see below) was tested. The decrease in size of Loop3 was again associated with a modest but significant (∼20%) increase in IRES activity.

Together, these findings suggest that Stem2/Loop2 may be the key sequence element facilitating internal ribosomal entry, while Loop3 may serve to modulate the activity of the IRES.

KINETIC ANALYSIS OF THE *IGF1R* IRES

We extended our assays to further assess the effects of the Stem2 and Loop3 mutations on function of the *IGF1R* IRES over time (Fig. 3). The full-length 1,040-nucleotide *IGF1R* 5′-UTR and the 90-nucleotide isolated core IRES exhibit nearly identical kinetic profiles, with the isolated IRES exhibiting ∼25% higher activity. By the 48- and 72-h time points, the activity of the *IGF1R* IRES exceeds that of the well-characterized encephalomyocarditis virus (EMCV) IRES by a considerable margin, although the viral IRES is activated much more rapidly in the cells (peaking at the 4–8 h time points). The Loop3 minimization mutant has lost this intrinsic dampening effect, allowing the *IGF1R* IRES to accelerate at a level that approaches that of the viral IRES, and ultimately displaying a higher activity than any of the other IRES constructs. The near complete loss of IRES function of the Stem2 compensatory mutant is obvious.

These results further support the conclusion that Stem2/Loop2 contains the sequence most critical for the operation of the IRES (recruitment of the 40S ribosomal subunit), while the poly(U) tract of Loop3 may serve to limit the maximal rate of translation initiation mediated through the IRES.

POSSIBLE SHINE–DALGARNO-LIKE (mRNA–rRNA BASE-PAIRING) INTERACTION BETWEEN THE *IGF1R* IRES AND THE HUMAN 18S rRNA

The fundamental purpose of the IRES is to recruit the 40S ribosomal subunit. A direct mRNA–rRNA base-pairing interaction between the

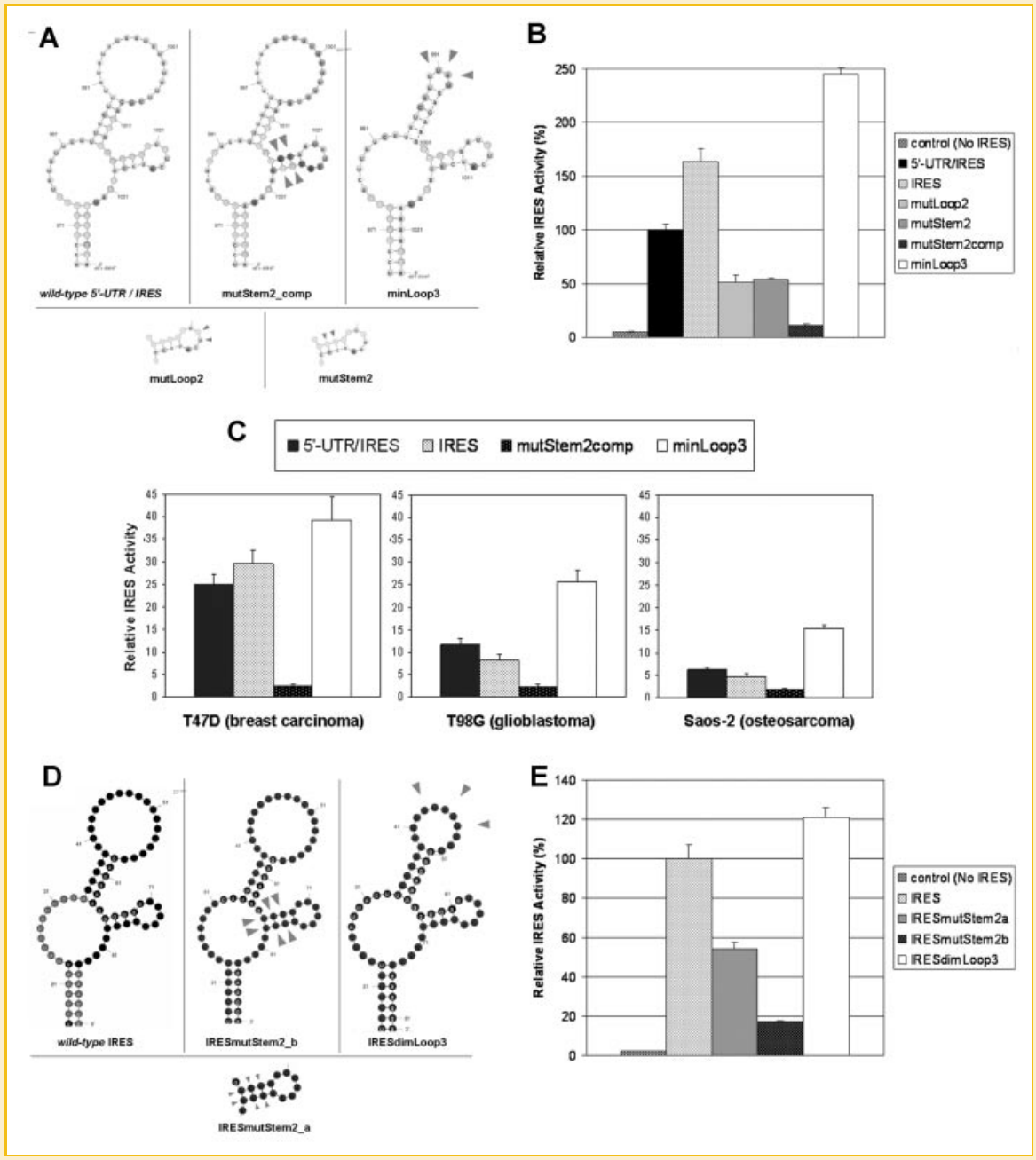

Fig. 2. Site-directed mutagenesis implicates Stem2/Loop2 and Loop3 as critical sequence elements for operation and regulation of the *IGF1R* IRES. A: mfold 3.2 projected structures of the core functional IRES domain within the context of the full-length 1,040-nucleotide 5′-UTR for the *wild-type* sequence, the compensatory mutation to Stem2, and minimization of Loop3. The positions of altered residues are marked by triangles. The structural representations in this figure were produced using the Natural mode of the mfold algorithm (maintains uniform spacing between residues). The positions of altered residues for the mutLoop2 and mutStem2 (not compensatory) constructs are shown in the lower panels. B: Relative IRES activity associated with the *wild-type* full-length 1,040-nucleotide *IGF1R* 5′-UTR (including the IRES), the isolated 90-nucleotide core functional IRES, mutLoop2 and mutStem2 (inactivating mutations), compensatory mutation to Stem2 (severely debilitating mutation), and minimization of Loop3 (activating mutation). Each of the mutants in this figure was tested in the context of the full-length 1,040-nucleotide 5′-UTR. T47D human breast carcinoma cells were transiently transfected with the appropriate bicistronic reporter construct, and firefly and *Renilla* luciferase levels measured in extracts prepared from cells 48 h posttransfection. IRES activity is calculated as the ratio of firefly to *Renilla* luciferase activity (controlling for transfection efficiency, cell density, pipetting error), and expressed as a percentage relative to that observed with the *wild-type* full-length 5′-UTR. Values obtained for the control construct (with no *IGF1R* insert and no IRES) are shown for comparison. C: Evaluation of activity of representative IRES mutants in three different human tumor cell lines. Transient transfections of the indicated wild-type and mutant bicistronic reporter constructs were performed as described above using the T47D human breast carcinoma cell line, T98G human glioblastoma cell line, or Saos-2 human osteosarcoma cell line. In these three graphs, IRES activity is expressed relative to the firefly/*Renilla* ratio obtained with the control construct (no *IGF1R* insert and no IRES, set = 1) and shown on the same scale to facilitate direct comparisons across the different cell types. A normalized f/R ratio of 2 relative to the control is usually considered the minimum threshold for detectable IRES activity, while a score of 8–10 is typical for the well-characterized EMCV IRES. The *IGF1R* IRES is active in all three of these tumor cell lines, however, the degree of activation varies T47D > T98G > Saos-2. The characteristic loss or gain in IRES activity associated with the compensatory Stem2 mutant and the minimization of Loop3, respectively, is observed in all three cell lines. D: mfold 3.2 projected structures of the *wild-type* isolated core IRES domain, IRESmutStem2b (compensatory mutation replacing three G·C base pairs with A·U base pairs), and IRESdimLoop3 (diminution of the Loop3 poly(U)-tract to match the naturally occurring small Loop3 allele (11 residues compared to 19 in the standard *wild-type* construct)). The positions of residues altered in MutStem2a, another compensatory mutation of Stem2 that retains the G–C rich composition of the native sequence, are shown in the lower panel. E: Relative IRES activity associated with *wild-type* core functional IRES, IRESmutStem2a (inactivating mutation), IRESmutStem2b (severely debilitating mutation), and IRESdimLoop3 (activating mutation) in T47D cells.

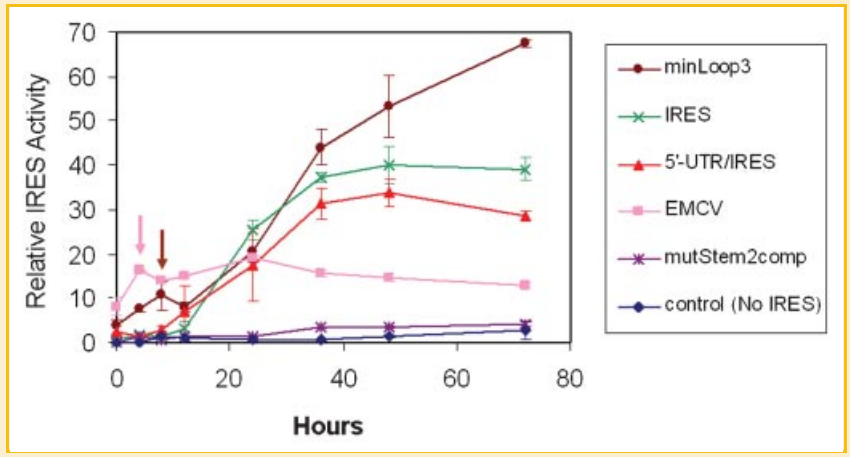

Fig. 3. Kinetic analysis of *IGF1R* IRES activity. Transient transfections of reporter constructs into T47D human breast carcinoma cells were performed as described above for *wild-type* and selected IRES mutants. Cells were harvested at varying times following transfection, and firefly and *Renilla* luciferase activities of cell lysates assayed. IRES activity is calculated as the ratio of firefly to *Renilla* luciferase activity, normalizing to the average f/R ratio observed with the control construct (set = 1). Control, parent bicistronic reporter construct with no *IGF1R* sequence (and no IRES); EMCV, bicistronic reporter with encephalomyocarditis virus (EMCV) IRES positioned between *Renilla* and firefly luciferase coding sequences; 5′-UTR/IRES, bicistronic reporter in which full-length (1,040 nucleotides) *wild-type* human *IGF1R* 5′-UTR (including the IRES) is positioned between *Renilla* and firefly luciferase coding sequences; IRES, bicistronic reporter containing the isolated core *IGF1R* IRES sequence (90 nucleotides); mutStem2comp, bicistronic reporter containing full-length *IGF1R* 5′-UTR with 4-base compensatory mutation to Stem2 of the IRES; minLoop3, bicistronic reporter containing full-length *IGF1R* 5′-UTR with 15-base deletion within the Loop3 poly(U)-tract of the IRES. The pink arrow marks the early acceleration of the viral (EMCV) IRES which peaks at ~4–8 h posttransfection. The brown arrow marks the enhanced acceleration of *IGF1R* IRES activity observed in association with the Loop3 minimization mutant.

IGF1R IRES and the 18S rRNA could conceivably facilitate ribosome recruitment. Upon careful examination, we noted that there exists a near-perfect Watson–Crick complementarity between the Stem2/Loop2 sequence of the *IGF1R* IRES and the G961 loop/helix 23b of the human 18S rRNA (Fig. 4A,B). The 959–964 region of the human 18S rRNA is known to be accessible to chemical probes, and therefore potentially available for direct base-pairing interactions with the mRNA [Demeshkina et al., 2000, 2003]. Our mutational analyses indicate that alterations to the Stem2Loop2 sequence which disrupt this complementarity are associated with a substantial loss of IRES activity. Furthermore, the G961 loop is projected to occupy an ideal position in three-dimensional space, on the platform of the 40S ribosomal subunit, from which to interact with the mRNA and facilitate translation initiation (Fig. 4C). Based on the relative positions of the mRNA and rRNA deduced from the crystal structure of the *T. thermophilus* 30S ribosomal subunit [Yusupova et al., 2001], it appears that the G693 loop of the prokaryotic 16S rRNA (homologous to the eukaryotic G961 loop of the 18S rRNA) lies in the immediate vicinity of both the anti-Shine–Dalgarno sequence (which is not present in eukaryotes) as well as the Shine–Dalgarno segment of the proximal 5′-untranslated region of the mRNA at the E-site. Together these findings suggest that the *IGF1R* IRES may recruit the 40S ribosome by a eukaryotic equivalent of the Shine–Dalgarno interaction, with the G961 loop of the 18S rRNA substituting for the anti-Shine–Dalgarno segment and base-pairing directly with the Stem2/Loop2 sequence of the IRES. The parallels between the mRNA–rRNA base-pairing interaction potentially involved in the operation of the *IGF1R* IRES in humans and the classical Shine–Dalgarno interaction of prokaryotes are summarized in Table II.

ALLELIC VARIATION OF THE Loop3 SEQUENCE OF THE *IGF1R* IRES

Examination of multiple GenBank entries containing the human *IGF1R* 5′-untranslated sequence reveals evidence of natural variation in the size of the Loop3 poly(U)-tract, ranging from 14 to 24 contiguous U residues, with a bimodal distribution centered around U_{16} and U_{24} (Fig. 5A). To explore this issue further, we PCR amplified the core *IGF1R* IRES sequence from several different sources of human genomic DNA (Fig. 5B). In most cases, a doublet (two distinct bands) was observed on agarose gel electrophoresis of the PCR product. Direct sequencing of these PCR products generated results that became challenging to interpret upon reaching the Loop3 sequence from either direction; however, we were able to definitively establish that the results were a composite of two distinct alleles (most commonly U_{21} and U_{13}) differing by eight in the number of residues in the Loop3 poly(U)-tract, with additional minor variants generated by polymerase error (most commonly +1, +2, or −1 U residues). However, for several of the human tumor cell lines tested, the PCR-amplified core *IGF1R* IRES yielded only one band on agarose gel, apparently representing only the smaller of the two alleles. To confirm these findings, PCR products generated from two representative cell lines: T47D (biallelic) and MDA-MB-231 (monoallelic) were cloned and multiple individual clones sequenced. The results were indicative of the existence of two distinct alleles, clustered around U_{13} and U_{21}, in T47D cells, but only one allele, clustered around U_{13}, in MDA-MB-231 (Fig. 5C).

DIFFERENTIAL MODULATION OF MUTANT IRES CONSTRUCTS BY RNA-BINDING IRES-REGULATORY PROTEINS HuR AND hnRNP C

We have previously characterized the RNA-binding protein HuR as a potent repressor of the *IGF1R* IRES [Meng et al., 2005], while hnRNP

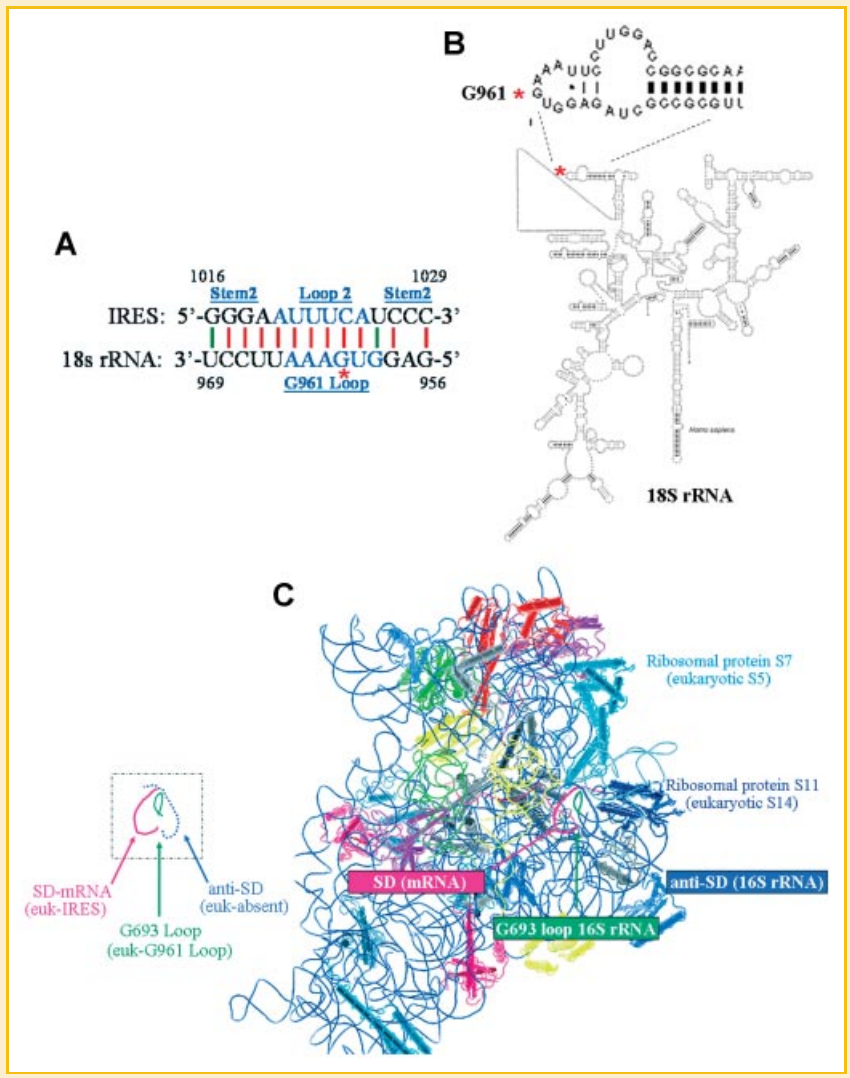

Fig. 4. Potential Watson–Crick base-pairing between the *IGF1R* IRES and the 18S rRNA. A: Sequence alignment demonstrating potential Watson–Crick complementarity between Stem2/Loop2 of the *IGF1R* IRES and the G961 loop/helix 23b of the 18S rRNA (accession K03432). Residues within Loop2 of the *IGF1R* IRES and the G961 loop of the 18S rRNA which are unpaired in their native conformation are highlighted in blue text. Note that the two loops are already in nearly perfect register for optimal Watson–Crick alignment. Potential intermolecular G·C and A·U base pairs are designated by red lines; G·U base pairs by green lines. B: Position of the G961 Loop on a two-dimensional structural representation of the human 18S rRNA [Cannone et al., 2002]. G961 is marked with a red asterisk. C: Crystal structure of the *T. thermophilus* 30S ribosomal subunit (PDB accession 1JGO) revealing the relative positions of the Shine–Dalgarno sequence within the 5′-untranslated region of the mRNA (highlighted red), the G693 loop of the 16S rRNA (homologous to the G961 Loop of eukaryotic 18S rRNA, highlighted green), and the anti-Shine–Dalgarno segment of the 16S rRNA (highlighted blue, not present in eukaryotes). The view is from the intersubunit face. The remainder of the 16S rRNA is represented by the thin dark blue line, and the associated mRNA is represented by a thin red line. The close proximal relationship of the three highlighted features can be appreciated from the simplified inset diagram on the left. The *IGF1R* IRES may recruit the 40S ribosome by direct base-pairing with the G961 loop of the 18S rRNA, in what would represent a eukaryotic equivalent of the Shine–Dalgarno interaction.

C appears to activate the IRES [Meng et al., 2008]. We were interested to determine what consequences the representative Stem2 compensatory (decrease in function) and Loop3 minimization (gain in function) IRES mutations would have on the IRES-regulatory capabilities of these RNA-binding proteins. Experiments were performed in which the various IRES reporter constructs were cotransfected with expression vectors for HuR or hnRNP C, as had been done originally to establish the IRES-regulatory activities of these RNA-binding proteins (Fig. 6). We found that each mutant IRES construct remained sensitive to the positive or negative modulatory effects of each of the RNA-binding proteins, with one exception: the severely debilitated Stem2 compensatory mutant IRES could not be activated by hnRNP C.

MUTATIONS IN THE CORE IRES SEQUENCE ALTER PROTEIN BINDING TO THE *IGF1R* 5′-UTR

We have developed a high-resolution northwestern protocol to analyze sequence-specific interactions between putative ITAFs and the *IGF1R* 5′-UTR. The northwestern is a functional assay of RNA-binding activity performed under highly stringent conditions. Utilizing the *wild-type* sequence as a probe, a reproducible pattern of bands representing a series of proteins capable of interacting

TABLE II. mRNA–rRNA Interactions Facilitating Translation Initiation[a]

Prokaryotic	Eukaryotic
Shine–Dalgarno interaction	IRES–18S Interaction
4–9 bp	9 bp (12 with 2 G–U)
Centered at −10 from AUG	Centered at −18 from AUG
Involves anti-SD sequence at 3′-terminus of 16S rRNA (E-site)	Involves G961 loop (helix 23b) in middle of 18S rRNA (E-site)

[a]Comparison of the prokaryotic Shine–Dalgarno interaction to the Stem2/Loop2 (mRNA)–18S rRNA interaction proposed as a mechanism for ribosome recruitment by the *IGF1R* IRES.

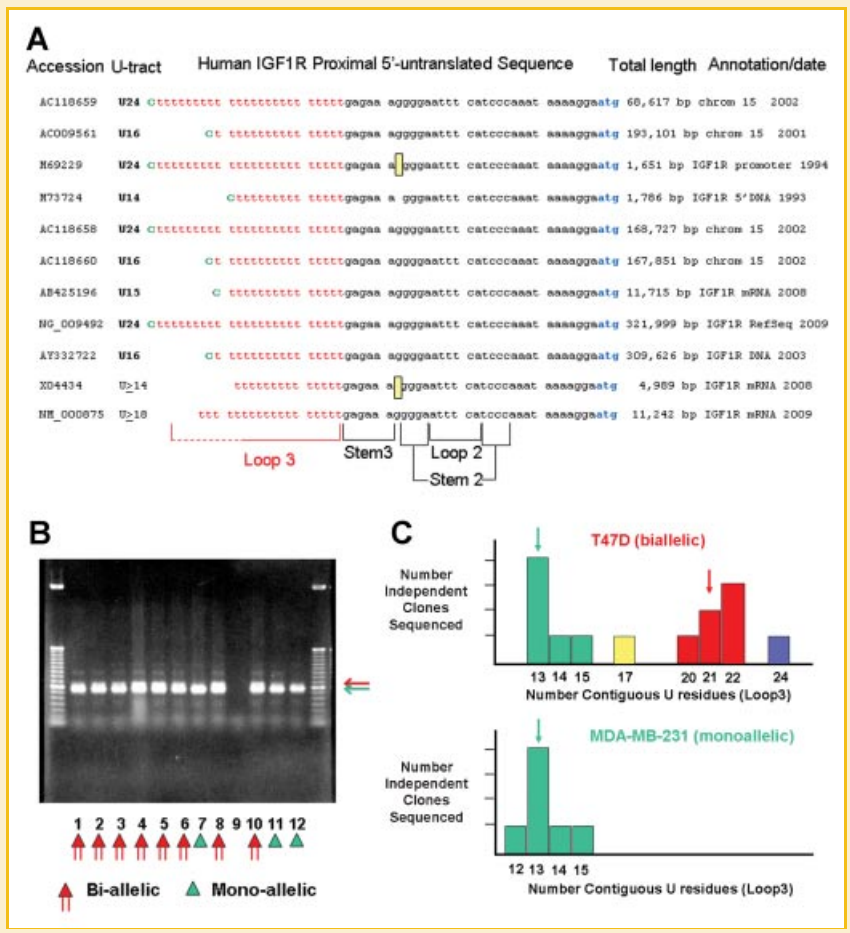

Fig. 5. Evidence for natural allelic variation of the Loop3 poly(U) sequence in human genomic DNA. A: Multiple individual GenBank entries containing the human *IGF1R* 5′-untranslated sequence are aligned at the authentic *IGF1R* initiation codon (AUG, blue text). The sequence surrounding Loop2 of the IRES is in black text. The long U-tract comprising Loop3 of the IRES (and two U residues within Stem3) is in red text. The C residue marking the end of the variable U-tract is in green text. The yellow rectangle indicates a missing G residue in two of the sequence entries. There are no other variations within the core IRES sequence amongst these GenBank entries. Note that these sequences range from primary entries submitted by individual labs over 15 years ago to recently generated high-throughput sequencing data. There appear to be two distinct naturally occurring alleles, differing in the size of the Loop3 U-tract by ~8 residues. B: Human genomic DNA was obtained from commercial sources or continuous cell lines as indicated and used as template for PCR amplification of the core *IGF1R* IRES sequence. Products were analyzed on a 1.75% agarose gel and stained with ethidium bromide. Lane 1: placenta (ClonTech); Lane 2: blood (Promega); Lane 3: blood (Promega, 2nd lot); Lane 4: T47D breast tumor cell line; Lane 5: MCF-7 breast tumor cell line; Lane 6: BT-20 breast tumor cell line; Lane 7: MDA-MB-231 breast tumor cell line; Lane 8: MCF-10A mammary epithelial cell line; Lane 9: negative control (no template); Lane 10: H146 small cell lung cancer cell line; Lane 11: Saos-2 osteosarcoma cell line; Lane 12: T98G glioblastoma cell line. In most instances, the product is a clear doublet, and direct sequence analysis (in either direction) is consistent with superimposition of two distinct alleles which differ only in size of the poly(U)-tract which comprises Loop3 of the projected IRES structure. In total, we have examined six sources of normal human genomic DNA, all of which appear to be biallelic (four confirmed by direct sequencing), and seven different human tumor cell lines, three of which (lanes 7, 11, 12) appear to be monoallelic (one confirmed by direct sequencing). C: Size distribution of the Loop3 poly(U)-tract assessed in individual clones amplified from human genomic DNA. The *IGF1R* core IRES sequence was PCR-amplified from genomic DNA obtained from two representative tumor cell lines: T47D and MDA-MB-231, which appeared to be biallelic and monoallelic, respectively, based on direct sequence analysis of PCR product. The PCR product from each of these cell lines was cloned into pSTblue1 and sequence analysis performed on plasmid DNA obtained from multiple individual clones. The results are presented graphically. The arrows at 13 and 21 are indicative of the apparent sizes of the U-tracts as determined by direct sequence analysis of the PCR products prior to cloning. For T47D (biallelic), out of 14 clones sequenced, 6 exhibited 13 ± 2 U residues (short allele) while 6 exhibited 21 ± 2 U residues (long allele). For MDA-MB-231 (monoallelic), all seven clones sequenced exhibited 13 ± 2 U residues (short allele).

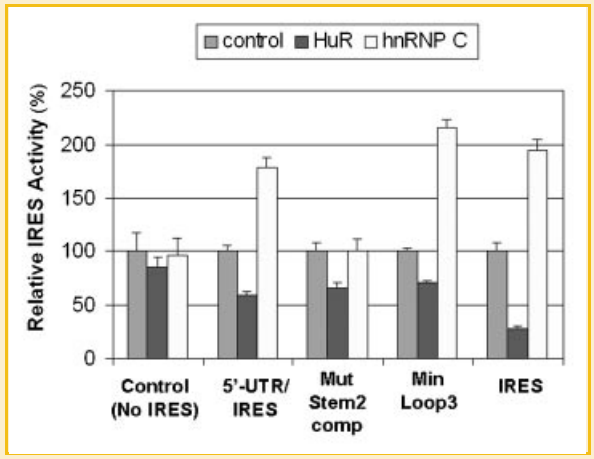

Fig. 6. Effect of IRES-regulatory proteins HuR and hnRNP C on activity of mutant IRES constructs. Expression vectors for RNA-binding proteins HuR or hnRNP C were transiently cotransfected into T47D cells along with bicistronic reporter constructs containing the *wild-type* or mutant *IGF1R* 5′-UTR/IRES sequence. After 48 h, cells were harvested and IRES activity determined from the ratio of firefly to *Renilla* luciferase levels as described above. IRES activity for each construct is expressed relative to that measured in absence of any ectopically expressed RNA-binding protein (i.e., cotransfection with pcDNA3.1 without insert). HuR, already characterized as a potent IRES repressor, maintained its negative influence on each of the mutant constructs. hnRNP C, previously shown to activate the *IGF1R* IRES, retained its activation potential for each of the mutant IRES constructs except the complementary mutation to Stem2. The control construct with no *IGF1R* sequence and no IRES is not significantly affected by either of these RNA-binding proteins.

specifically with the *IGF1R* 5′-UTR is observed (Fig. 7). To test whether the Stem2 or Loop3 IRES mutations might alter the pattern of protein binding to the *IGF1R* 5′-UTR, the northwestern assay was repeated using the mutant RNA sequences as probes. A marked change in affinity of multiple individual proteins for binding the mutant 5′-UTR RNA sequences was observed. The compensatory mutation to Stem2, which dramatically decreases IRES activity, was associated with a near complete loss of binding of a number of the putative ITAFs (particularly bands 2, 3, 4, 4b, 5, 8a, and 8d) detected by the northwestern assay. A distinct subset of bands (7, 9, and 10) remained relatively unchanged or even increased in intensity, confirming that the effect of the Stem2 mutation on protein binding to the *IGF1R* 5′-UTR was selective. Interestingly, the minimization of Loop3, which dramatically increases IRES activity, was associated with an equally dramatic increase in affinity for the same ITAFs which had been negatively impacted by the Stem2 mutation. Importantly, we have definitively characterized bands 2, 4, and 5 as "internal" ITAFs, binding within the 90-nucleotide core functional IRES sequence, thus their altered affinity for the mutant IRES probes can be rationalized on this basis. Band 8d, which we have determined to be hnRNP C (also an internal ITAF), exhibits the same trend as the other internal ITAFs (decreased affinity with the Stem2 mutation, increased affinity with minimization of Loop3). In contrast, we have determined that bands 7, 9, and 10 are "external" ITAFs, binding primarily outside of the core IRES sequence, and indeed these bands show very little variation in intensity accompanying mutation of the core IRES sequence.

DISCUSSION

It has been widely held that the distance from the beginning of the mRNA to the initiation codon is usually <200 nucleotides, that only a relatively small proportion of eukaryotic mRNAs have 5′-UTRs of substantially greater length, and that these long 5′-UTRs tend to be associated with proto-oncogenes and other factors integrally involved in the control of cell proliferation and survival. A longer 5′-untranslated region provides a larger space for RNA structural features and protein binding sites, and therefore greater opportunities for specific regulation of gene expression at the translational level. Indeed, we have begun to discover many intricate regulatory mechanisms provided by these complex 5′-UTRs which allow the translational efficiency of individual mRNAs to be selectively modulated [Fernandez et al., 2005; Galban et al., 2008; Jo et al., 2008; Spriggs et al., 2008; Coleman and Miskimins, 2009; Conte et al., 2009]. However, emerging results of the ENCODE project [Birney et al., 2007], which involves an unbiased high-throughput mapping of transcriptional activity, have shown that individual annotated genes are associated with an average of five distinct transcription start sites, many of which are positioned thousands of base pairs further upstream from the coding sequences [Denoeud et al., 2007], suggesting that complex 5′-untranslated sequences such as that associated with the *IGF1R* mRNA are likely to be much more common than previously recognized. Such an expansion of the transcriptional landscape potentially provides an even greater opportunity for gene-specific translational regulation. However, if indeed these additional mRNA species with massively extended 5′-untranslated sequences are to function as translatable mRNAs, it is also reasonable to anticipate that IRESs may also be much more commonly utilized than previously suspected [Baird et al., 2006]. Consequently, it is very important that we determine how IRESs operate and how they are regulated in order to understand how these newly discovered transcripts can be effectively translated.

COMPOSITE MODEL FOR OPERATION AND REGULATION OF THE *IGF1R* IRES

Translation initiation in bacteria inherently requires "internal ribosomal entry" due to the polycistronic nature of the prokaryotic mRNAs [Shine and Dalgarno, 1974; Steitz and Jakes, 1975]. Our findings suggest that the *IGF1R* IRES may recruit the 40S ribosome by a eukaryotic equivalent of the prokaryotic Shine–Dalgarno interaction. The first point of contact between the *IGF1R* IRES and the 18S rRNA would likely involve Loop2 of the IRES and the G961 loop of the 18S rRNA. The Stem2 structure may serve an important purpose in presenting the Loop2 sequence in optimal geometrical orientation for the initial "kissing" interaction with the unpaired residues of the G961 loop. Subsequent melting of the Stem2 structure would be required to expose the full Watson–Crick mRNA–rRNA base-pairing potential, a dynamic process likely regulated by sequence-specific RNA-binding proteins. We have observed two different complementary mutations to the Stem2 sequence which result in >80% loss of IRES activity. These data support the concept that Stem2/Loop2 may be the key determinant of *IGF1R* IRES function, and that a direct base-pairing interaction with the G961

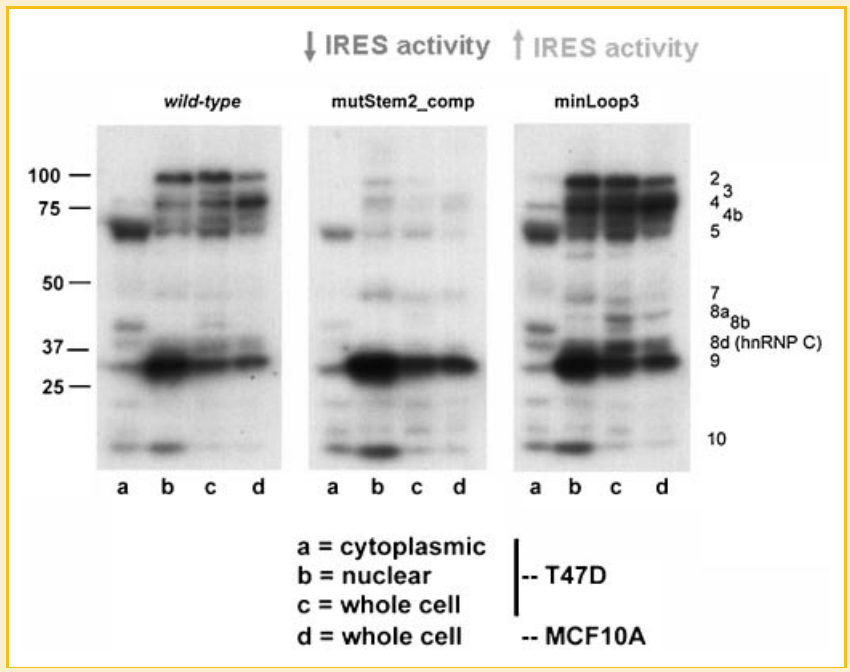

Fig. 7. Mutations to critical sequence elements of the *IGF1R* IRES alter the pattern of protein binding to the *IGF1R* 5′-UTR. A high-resolution northwestern assay was performed to assess sequence-specific RNA–protein interactions involving the *IGF1R* 5′-UTR. Protein extracts from T47D or MCF-10A cells were separated by SDS/PAGE, transferred to nitrocellulose, and renatured on blot in a physiological buffer prior to incubation with radiolabeled sense RNA probes representing the *wild-type* or mutant *IGF1R* 5′-UTR/IRES sequence, using excess tRNA as a non-specific competitor. Following incubation, the blots were washed in the presence of heparin (to eliminate weak or unstable RNA–protein interactions) then exposed to autoradiographic film. The northwestern is a functional assay of RNA-binding activity, and a reproducible pattern of proteins specifically recognizing the *IGF1R* 5′-UTR sequence (putative IRES-regulatory proteins or ITAFs) is obtained. The IRES mutations are associated with substantial alterations in affinity for individual IRES-regulatory proteins. The results implicate Stem2/Loop2 as the most important binding site for the internal ITAFs regulating ribosome recruitment, while Loop3 appears to limit accessibility of the critical Stem2/Loop2 sequence to these translation-regulatory proteins.

loop (h23b) of the 18S rRNA at the E-site of the 40S ribosomal subunit may provide the fundamental basis for ribosome recruitment by the IRES.

Compensatory mutation to Stem2 also markedly decreases affinity for a number of individual bands detected by northwestern analysis, indicating that the Stem2/Loop2 microdomain serves as a critical recognition sequence for as many as seven distinct RNA-binding proteins (including hnRNP C) which interact specifically with the core functional IRES sequence and may directly influence the efficiency of ribosome recruitment to the IRES. Taken together, these findings suggest that Stem2/Loop2 may be the focal point for ribosome recruitment and its regulation by internal ITAFs.

Minimization of the size of the Loop3 poly(U)-tract actually increases affinity for several putative IRES-regulatory proteins. Loop3 minimization also dramatically increases IRES activity and alters the kinetics of IRES activation, inducing a much more rapid acceleration. These results suggest that the physiological function of the intact Loop3 sequence may be to limit accessibility of the critical Stem2/Loop2 sequence to internal ITAFs and perhaps also the 40S ribosomal subunit itself, ultimately governing the maximal rate of translation initiation mediated through the IRES. Both HuR and hnRNP C are known to bind preferentially to U-rich sequences [Gorlach et al., 1994; Lopez de Silanes et al., 2004; Coleman and Miskimins, 2009]. However, the fact that the Loop3 minimization mutant remains responsive to both hnRNP C and HuR indicates that the long homopolymeric (U)-tract of Loop3 is not likely the primary binding site for these IRES-regulatory proteins. Rather, these results suggest that hnRNP C and HuR may be directly involved in facilitating or blocking ribosome recruitment, respectively, probably through interactions with Stem2/Loop2.

mRNA–rRNA INTERACTIONS FACILITATING TRANSLATION INITIATION

The proposed base-pairing interaction would position the proximal 5′-untranslated region of the *IGF1R* mRNA at the E-site on the platform of the 40S ribosome. This location on the small ribosomal subunit has been proposed to serve as a dedicated site for a variety of translation-regulatory phenomena involving 5′-untranslated sequences of mRNA [Marzi et al., 2007]. The mRNA exit site, so named because of its role during translation elongation, is bounded by helix 23, helix 26, rpS7 (eukaryotic S5), and rpS11 (eukaryotic S14) [Kaminishi et al., 2007], and actually serves as a docking and entry site during translation initiation. Viral IRESs have also been shown to contact the 40S ribosome in this region [Kieft, 2008].

Shine–Dalgarno-like interactions have been suspected of contributing to the function of several viral IRESs [Le et al., 1992; Pilipenko et al., 1992; Scheper et al., 1994; Yang et al., 2003], and may indeed be utilized by other cellular IRESs as well [Dresios et al., 2006]. In addition, two regulatory mechanisms distinct from internal ribosomal entry have been described which appear to

depend on mRNA–rRNA base-pairing to recruit or reposition the ribosome. Ribosomal shunting has been demonstrated for the adenovirus late mRNA, with evidence for Watson–Crick base-pairing of 5′-untranslated sequence to the 3′-terminal hairpin of the human 18S rRNA, located on solvent accessible surface of 40S ribosomal subunit [Yueh and Schneider, 2000]. A similar mechanism has been proposed for the heat shock protein 70, an endogenous cellular mRNA. A process of translation re-initiation is utilized by the feline calicivirus to facilitate expression of the VP2 open-reading frame from a naturally bicistronic viral RNA in eukaryotic cells [Luttermann and Meyers, 2009]. This mechanism requires an unpaired sequence element located between the two open-reading frames, which has the potential for base-pairing interaction (nine contiguous nucleotides of perfect Watson–Crick complementarity) with the loop of helix 26 of the 18S rRNA. Like helix 23 (which we propose may be involved in interaction with the *IGF1R* IRES), helix 26 of the 18S rRNA is also positioned near the E-site of the 40S ribosomal subunit. For each of these mechanisms (shunting, re-initiation, IRES), the complementarity to rRNA is not the only factor contributing to ribosome recruitment. Secondary structures formed by neighboring sequences within the mRNA appear to be important for optimal presentation of the complementary sequence to ribosomal RNA, and protein–protein interactions likely facilitate the interaction between the mRNA and the ribosome.

However, based on our BLAST analyses, it appears that complementarity to the G961 loop of the 18S rRNA may be unique to the *IGF1R* IRES. If our interpretation is correct, for such an interaction to be reserved solely for *IGF1R* would reflect just how biologically important the translational regulation of this critical growth control gene may be to the phenotypic integrity of the cell. Indeed, the significance of *IGF1R* within the hierarchy of genes controlling cell proliferation and survival is not without support, as it appears that many other growth factor/receptor systems such as EGFR, VEGF, and the estrogen receptor, and other growth control genes such as cyclin D1, may function through or be controlled by *IGF1R* [Stewart et al., 1990; Coppola et al., 1994; Jones et al., 2008]. *IGF1R* is considered by cellular physiologists to be the single most powerful potentiator of survival signaling in the mammalian cell [Peruzzi et al., 1999], and *IGF1R* appears to be essentially required for malignant transformation [Sell et al., 1994]. (This is not to suggest that other cellular IRESs may not utilize Shine–Dalgarno-like base-pairing with the 18S rRNA, but rather that interaction with the seemingly ideally positioned G961 loop may be reserved for *IGF1R*.)

POSSIBLE CANCER-RELEVANCE OF THE *IGF1R* IRES

IGF1R levels are low in normal differentiated adult cells [Werner et al., 1989], whereas *IGF1R* overexpression is implicated in the pathogenesis of a large proportion of human cancers [Zhang and Yee, 2000; LeRoith and Roberts, 2003; Kolb et al., 2008]. Our lab has begun to accumulate substantial evidence that pathological dysregulation of the *IGF1R* IRES could be responsible for *IGF1R* overexpression in a proportion of human breast tumors [Meng et al., 2008]. There is clear evidence from multiple GenBank entries and our own focused sequence analyses that the Loop3 U-tract sequence is naturally polymorphic in the human population, with the existence of two distinct alleles, differing by eight in the number of U residues. While each of six different sources of normal human genomic DNA we examined appear to contain both the long and short alleles, three of the seven tumor cell lines screened appear to contain only the small Loop3 allele. As little as 50% increase in IGF1R protein levels may allow cells to proliferate in an anchorage-independent manner (a surrogate assay for tumorigenicity) [Rubini et al., 1997]. Considering the functional relationship we have observed between the size of the Loop3 polyU-tract and IRES activity, it is conceivable that allelic variation of this sequence, whether present in the germ-line or somatically acquired, could be associated with a predisposition to malignant transformation or provide a selectable proliferative/survival advantage for the tumor cells. Studies are currently underway to further explore this question.

REFERENCES

Antic D, Keene JD. 1997. Embryonic lethal abnormal visual RNA-binding proteins involved in growth, differentiation, and posttranscriptional gene expression. Am J Hum Genet 61:273–278.

Baird SD, Turcotte M, Korneluk RG, Holcik M. 2006. Searching for IRES. RNA 12:1755–1785.

Birney E, Stamatoyannopoulos JA, Dutta A, Guigo R, Gingeras TR, Margulies EH, Weng Z, Snyder M, Dermitzakis ET, Thurman RE, Kuehn MS, Taylor CM, Neph S, Koch CM, Asthana S, Malhotra A, Adzhubei I, Greenbaum JA, Andrews RM, Flicek P, Boyle PJ, Cao H, Carter NP, Clelland GK, Davis S, Day N, Dhami P, Dillon SC, Dorschner MO, Fiegler H, Giresi PG, Goldy J, Hawrylycz M, Haydock A, Humbert R, James KD, Johnson BE, Johnson EM, Frum TT, Rosenzweig ER, Karnani N, Lee K, Lefebvre GC, Navas PA, Neri F, Parker SC, Sabo PJ, Sandstrom R, Shafer A, Vetrie D, Weaver M, Wilcox S, Yu M, Collins FS, Dekker J, Lieb JD, Tullius TD, Crawford GE, Sunyaev S, Noble WS, Dunham I, Denoeud F, Reymond A, Kapranov P, Rozowsky J, Zheng D, Castelo R, Frankish A, Harrow J, Ghosh S, Sandelin A, Hofacker IL, Baertsch R, Keefe D, Dike S, Cheng J, Hirsch HA, Sekinger EA, Lagarde J, Abril JF, Shahab A, Flamm C, Fried C, Hackermuller J, Hertel J, Lindemeyer M, Missal K, Tanzer A, Washietl S, Korbel J, Emanuelsson O, Pedersen JS, Holroyd N, Taylor R, Swarbreck D, Matthews N, Dickson MC, Thomas DJ, Weirauch MT, Gilbert J, et al. 2007. Identification and analysis of functional elements in 1% of the human genome by the ENCODE pilot project. Nature 447:799–816.

Cannone JJ, Subramanian S, Schnare MN, Collett JR, D'Souza LM, Du Y, Feng B, Lin N, Madabusi LV, Muller KM, Pande N, Shang Z, Yu N, Gutell RR. 2002. The comparative RNA web (CRW) site: An online database of comparative sequence and structure information for ribosomal, intron, and other RNAs. BMC Bioinformatics 3:2.

Carninci P, Sandelin A, Lenhard B, Katayama S, Shimokawa K, Ponjavic J, Semple CA, Taylor MS, Engstrom PG, Frith MC, Forrest AR, Alkema WB, Tan SL, Plessy C, Kodzius R, Ravasi T, Kasukawa T, Fukuda S, Kanamori-Katayama M, Kitazume Y, Kawaji H, Kai C, Nakamura M, Konno H, Nakano K, Mottagui-Tabar S, Arner P, Chesi A, Gustincich S, Persichetti F, Suzuki H, Grimmond SM, Wells CA, Orlando V, Wahlestedt C, Liu ET, Harbers M, Kawai J, Bajic VB, Hume DA, Hayashizaki Y. 2006. Genome-wide analysis of mammalian promoter architecture and evolution. Nat Genet 38:626–635.

Coleman J, Miskimins WK. 2009. Structure and activity of the internal ribosome entry site within the human p27 Kip1 5′-untranslated region. RNA Biol 6:84–89.

Conte C, Ainaoui N, Delluc-Clavieres A, Khoury MP, Azar R, Pujol F, Martineau Y, Pyronnet S, Prats AC. 2009. Fibroblast growth factor 1 induced during myogenesis by a transcription-translation coupling mechanism. Nucleic Acids Res 37:5267–5278.

Cooke DW, Casella SJ. 1994. The 5′-untranslated region of the IGF-I receptor gene modulates reporter gene expression by both pre- and post-transcriptional mechanisms. Mol Cell Endocrinol 101:77–84.

Cooke DW, Bankert LA, Roberts CT, Jr., LeRoith D, Casella SJ. 1991. Analysis of the human type I insulin-like growth factor receptor promoter region. Biochem Biophys Res Commun 177:1113–1120.

Coppola D, Ferber A, Miura M, Sell C, D'Ambrosio C, Rubin R, Baserga R. 1994. A functional insulin-like growth factor I receptor is required for the mitogenic and transforming activities of the epidermal growth factor receptor. Mol Cell Biol 14:4588–4595.

Demeshkina N, Repkova M, Ven'yaminova A, Graifer D, Karpova G. 2000. Nucleotides of 18S rRNA surrounding mRNA codons at the human ribosomal A, P, and E sites: A crosslinking study with mRNA analogs carrying an aryl azide group at either the uracil or the guanine residue. RNA 6:1727–1736.

Demeshkina N, Laletina E, Meschaninova M, Ven'yaminova A, Graifer D, Karpova G. 2003. Positioning of mRNA codons with respect to 18S rRNA at the P and E sites of human ribosome. Biochim Biophys Acta 1627:39–46.

Denoeud F, Kapranov P, Ucla C, Frankish A, Castelo R, Drenkow J, Lagarde J, Alioto T, Manzano C, Chrast J, Dike S, Wyss C, Henrichsen CN, Holroyd N, Dickson MC, Taylor R, Hance Z, Foissac S, Myers RM, Rogers J, Hubbard T, Harrow J, Guigo R, Gingeras TR, Antonarakis SE, Reymond A. 2007. Prominent use of distal 5′ transcription start sites and discovery of a large number of additional exons in ENCODE regions. Genome Res 17:746–759.

Dresios J, Chappell SA, Zhou W, Mauro VP. 2006. An mRNA–rRNA base-pairing mechanism for translation initiation in eukaryotes. Nat Struct Mol Biol 13:30–34.

Fernandez J, Yaman I, Huang C, Liu H, Lopez AB, Komar AA, Caprara MG, Merrick WC, Snider MD, Kaufman RJ, Lamers WH, Hatzoglou M. 2005. Ribosome stalling regulates IRES-mediated translation in eukaryotes, a parallel to prokaryotic attenuation. Mol Cell 17:405–416.

Galban S, Kuwano Y, Pullmann R, Jr., Martindale JL, Kim HH, Lal A, Abdelmohsen K, Yang X, Dang Y, Liu JO, Lewis SM, Holcik M, Gorospe M. 2008. RNA-binding proteins HuR and PTB promote the translation of hypoxia-inducible factor 1alpha. Mol Cell Biol 28:93–107.

Gooch JL, Van Den Berg CL, Yee D. 1999. Insulin-like growth factor (IGF)-I rescues breast cancer cells from chemotherapy-induced cell death—Proliferative and anti-apoptotic effects. Breast Cancer Res Treat 56:1–10.

Gorlach M, Burd CG, Dreyfuss G. 1994. The determinants of RNA-binding specificity of the heterogeneous nuclear ribonucleoprotein C proteins. J Biol Chem 269:23074–23078.

Jo OD, Martin J, Bernath A, Masri J, Lichtenstein A, Gera J. 2008. Heterogeneous nuclear ribonucleoprotein A1 regulates cyclin D1 and c-myc internal ribosome entry site function through Akt signaling. J Biol Chem 283:23274–23287.

Jones RA, Campbell CI, Petrik JJ, Moorehead RA. 2008. Characterization of a novel primary mammary tumor cell line reveals that cyclin D1 is regulated by the type I insulin-like growth factor receptor. Mol Cancer Res 6:819–828.

Kaminishi T, Wilson DN, Takemoto C, Harms JM, Kawazoe M, Schluenzen F, Hanawa-Suetsugu K, Shirouzu M, Fucini P, Yokoyama S. 2007. A snapshot of the 30S ribosomal subunit capturing mRNA via the Shine–Dalgarno interaction. Structure 15:289–297.

Kieft JS. 2008. Viral IRES RNA structures and ribosome interactions. Trends Biochem Sci 33:274–283.

Kim HJ, Litzenburger BC, Cui X, Delgado DA, Grabiner BC, Lin X, Lewis MT, Gottardis MM, Wong TW, Attar RM, Carboni JM, Lee AV. 2007. Constitutively active type I insulin-like growth factor receptor causes transformation and xenograft growth of immortalized mammary epithelial cells and is accompanied by an epithelial-to-mesenchymal transition mediated by NF-kappaB and snail. Mol Cell Biol 27:3165–3175.

Kolb EA, Gorlick R, Houghton PJ, Morton CL, Lock R, Carol H, Reynolds CP, Maris JM, Keir ST, Billups CA, Smith MA. 2008. Initial testing (stage 1) of a monoclonal antibody (SCH 717454) against the IGF-1 receptor by the pediatric preclinical testing program. Pediatr Blood Cancer 50:1190–1197.

Kurmasheva RT, Houghton PJ. 2006. IGF-I mediated survival pathways in normal and malignant cells. Biochim Biophys Acta 1766:1–22.

Le SY, Chen JH, Sonenberg N, Maizel JV. 1992. Conserved tertiary structure elements in the 5′ untranslated region of human enteroviruses and rhinoviruses. Virology 191:858–866.

LeRoith D, Roberts CT, Jr. 2003. The insulin-like growth factor system and cancer. Cancer Lett 195:127–137.

Lopez de Silanes I, Zhan M, Lal A, Yang X, Gorospe M. 2004. Identification of a target RNA motif for RNA-binding protein HuR. Proc Natl Acad Sci USA 101:2987–2992.

Luttermann C, Meyers G. 2009. The importance of inter- and intramolecular base pairing for translation reinitiation on a eukaryotic bicistronic mRNA. Genes Dev 23:331–344.

Marzi S, Myasnikov AG, Serganov A, Ehresmann C, Romby P, Yusupov M, Klaholz BP. 2007. Structured mRNAs regulate translation initiation by binding to the platform of the ribosome. Cell 130:1019–1031.

Mathews DH, Sabina J, Zuker M, Turner DH. 1999. Expanded sequence dependence of thermodynamic parameters improves prediction of RNA secondary structure. J Mol Biol 288:911–940.

Meng Z, King PH, Nabors LB, Jackson NL, Chen CY, Emanuel PD, Blume SW. 2005. The ELAV RNA-stability factor HuR binds the 5′-untranslated region of the human IGF-IR transcript and differentially represses cap-dependent and IRES-mediated translation. Nucleic Acids Res 33:2962–2979.

Meng Z, Jackson NL, Choi H, King PH, Emanuel PD, Blume SW. 2008. Alterations in RNA-binding activities of IRES-regulatory proteins as a mechanism for physiological variability and pathological dysregulation of IGF-IR translational control in human breast tumor cells. J Cell Physiol 217:172–183.

Pandini G, Vigneri R, Costantino A, Frasca F, Ippolito A, Fujita-Yamaguchi Y, Siddle K, Goldfine ID, Belfiore A. 1999. Insulin and insulin-like growth factor-I (IGF-I) receptor overexpression in breast cancers leads to insulin/IGF-I hybrid receptor overexpression: Evidence for a second mechanism of IGF-I signaling. Clin Cancer Res 5:1935–1944.

Peruzzi F, Prisco M, Dews M, Salomoni P, Grassilli E, Romano G, Calabretta B, Baserga R. 1999. Multiple signaling pathways of the insulin-like growth factor 1 receptor in protection from apoptosis. Mol Cell Biol 19:7203–7215.

Pilipenko EV, Gmyl AP, Maslova SV, Svitkin YV, Sinyakov AN, Agol VI. 1992. Prokaryotic-like cis elements in the cap-independent internal initiation of translation on picornavirus RNA. Cell 68:119–131.

Rubini M, Hongo A, D'Ambrosio C, Baserga R. 1997. The IGF-I receptor in mitogenesis and transformation of mouse embryo cells: Role of receptor number. Exp Cell Res 230:284–292.

Scheper GC, Voorma HO, Thomas AA. 1994. Basepairing with 18S ribosomal RNA in internal initiation of translation. FEBS Lett 352:271–275.

Sell C, Dumenil G, Deveaud C, Miura M, Coppola D, DeAngelis T, Rubin R, Efstratiadis A, Baserga R. 1994. Effect of a null mutation of the insulin-like growth factor I receptor gene on growth and transformation of mouse embryo fibroblasts. Mol Cell Biol 14:3604–3612.

Shine J, Dalgarno L. 1974. The 3′-terminal sequence of *Escherichia coli* 16S ribosomal RNA: Complementarity to nonsense triplets and ribosome binding sites. Proc Natl Acad Sci USA 71:1342–1346.

Spriggs KA, Stoneley M, Bushell M, Willis AE. 2008. Re-programming of translation following cell stress allows IRES-mediated translation to predominate. Biol Cell 100:27–38.

Steitz JA, Jakes K. 1975. How ribosomes select initiator regions in mRNA: Base pair formation between the 3′ terminus of 16S rRNA and the mRNA during initiation of protein synthesis in *Escherichia coli*. Proc Natl Acad Sci USA 72:4734–4738.

Stewart AJ, Johnson MD, May FE, Westley BR. 1990. Role of insulin-like growth factors and the type I insulin-like growth factor receptor in the estrogen-stimulated proliferation of human breast cancer cells. J Biol Chem 265:21172–21178.

Werner H, Woloschak M, Adamo M, Shen-Orr Z, Roberts CT, Jr., LeRoith D. 1989. Developmental regulation of the rat insulin-like growth factor I receptor gene. Proc Natl Acad Sci USA 86:7451–7455.

Yang D, Cheung P, Sun Y, Yuan J, Zhang H, Carthy CM, Anderson DR, Bohunek L, Wilson JE, McManus BM. 2003. A Shine–Dalgarno-like sequence mediates in vitro ribosomal internal entry and subsequent scanning for translation initiation of coxsackievirus B3 RNA. Virology 305:31–43.

Yanochko GM, Eckhart W. 2006. Type I insulin-like growth factor receptor over-expression induces proliferation and anti-apoptotic signaling in a three-dimensional culture model of breast epithelial cells. Breast Cancer Res 8:R18.

Yueh A, Schneider RJ. 2000. Translation by ribosome shunting on adenovirus and hsp70 mRNAs facilitated by complementarity to 18S rRNA. Genes Dev 14:414–421.

Yusupova GZ, Yusupov MM, Cate JH, Noller HF. 2001. The path of messenger RNA through the ribosome. Cell 106:233–241.

Zhang X, Yee D. 2000. Tyrosine kinase signalling in breast cancer: Insulin-like growth factors and their receptors in breast cancer. Breast Cancer Res 2:170–175.

Zuker M. 2003. mfold web server for nucleic acid folding and hybridization prediction. Nucleic Acids Res 31:3406–3415.