Splicing mediates the activity of four putative cellular internal ribosome entry sites

Brian T. Baranick*†, Nathan A. Lemp†‡, Jill Nagashima†, Kei Hiraoka†, Noriyuki Kasahara*†‡, and Christopher R. Logg†§

*Interdepartmental Program in Molecular Biology, ‡Department of Molecular and Medical Pharmacology, and †Department of Medicine, University of California, 675 Charles E. Young Drive South, MRL 1551, Los Angeles, CA 90095-7019

A growing number of cellular mRNAs are thought to possess internal ribosome entry sites (IRESs), sequences that permit translation of a transcript independent of its 5′ end and cap structure. Although dicistronic assays are the canonical method of testing sequences for IRES activity, they may produce false-positive results if unanticipated monocistronic RNAs arise from the dicistronic construct used. Using a dicistronic reporter system and a green fluorescent protein-tagged retrovirus to evaluate six previously reported cellular IRESs, we found that four contain 3′ splice sites whose activity was required for apparent IRES function and which resulted in formation of monocistronic transcripts by splicing. Bioinformatic analysis revealed that the 3′ splice sites identified in three of these putative IRESs are used in their native mRNAs and that the fourth is likely an artifactual sequence created during cDNA cloning. Our findings demonstrate a need for reexamination of other reported cellular IRESs by using careful RNA structural analysis to rule out splicing as the source of perceived IRES activity.

bicistronic | dicistronic | RNA splicing | translation

The vast majority of eukaryotic mRNAs are translated via a mechanism in which the 40S ribosomal subunit engages the mRNA at its methylguanosine-capped 5′ end (1). Upon associating with the transcript, these subunits are believed then to scan in the 5′ to 3′ direction for an appropriately situated AUG at which to begin translation (2, 3). A smaller number of mRNAs are translated by a 5′ end- and cap-independent mechanism wherein ribosomes are recruited to the transcript at an interior location through an internal ribosome entry site (IRES).

IRESs were first discovered in the picornaviruses encephalomyocarditis virus (EMCV) and poliovirus (4, 5). The RNA of these viruses possesses very long 5′ UTRs bearing many unutilized upstream AUGs (uAUGs) and, unlike cellular mRNA, is uncapped (6). Soon after the identification of picornaviral IRESs, a number of cellular mRNAs were also reported to contain IRESs. To date, at least 85 cellular IRESs have been described (7). The experimental grounds on which proof of most cellular IRESs rest, however, has been the subject of dispute (8–10).

A primary criticism of the data presented as establishing the existence of cellular IRESs concerns the plasmid-based dicistronic assay, the standard method of ascertaining IRES activity. In this assay, the candidate sequence is inserted between two reporter genes (5) so that both the upstream and downstream cistron are transcribed on the same RNA. If the test insert causes increased expression of the downstream cistron relative to the upstream cistron, the result is considered evidence for internal ribosome entry. However, the generation of even low levels of monocistronic RNAs from dicistronic constructs has the potential to falsely indicate IRES activity (8, 9, 11). One way that such RNAs could arise is through splicing of the dicistronic transcript due to the presence of a 3′ splice site (ss) in the test sequence [see supporting information (SI) Fig. 5]. A previous study on the putative XIAP IRES found that it contained a 3′ ss that was active in the context of a dicistronic luciferase reporter RNA and

that contributed to apparent IRES activity (11). In this study, we evaluated six arbitrarily chosen putative cellular IRESs and found that splicing was necessary for the apparent activity of four. We also show that the splice sites identified in three of these four putative IRESs are used in the biogenesis of the corresponding cellular mRNAs and that one IRES sequence is an artifact that incidentally functions as a 3′ ss.

Results

Six Putative Cellular IRESs Exhibit Little Activity in a Dicistronic Construct Depleted of Upstream 5′ ss. To allow the evaluation of IRESs in the absence of splicing, we constructed a dicistronic reporter plasmid system in which all transcribed sequence upstream of the IRES insert was extensively modified by silent mutagenesis to eliminate potential 5′ ss (SI Fig. 6). The upstream cistron in this construct is a *Gaussia* luciferase (GLuc) gene depleted of potential 5′ ss, and the downstream cistron encodes GFP (Fig. 1A). We evaluated six putative cellular IRES sequences in this system. These are derived from the 5′ UTRs of the BiP (12), eIF4G (13), NRF (14), Rbm3 (15), VEGF (16), and XIAP (17) mRNAs. The VEGF sequence used is a variant, termed SP163, reported to exhibit greater activity than the native 5′ UTR (16). Several non-IRES sequences, including the 5′ UTRs of the cap-dependent actin (18), β-globin (19), and JunD (20) mRNAs, were evaluated in parallel as controls.

Surprisingly, we found that only the EMCV IRES mediated levels of GFP expression clearly above those of the non-IRES sequences in this construct (Fig. 1B). This stands in particularly strong contrast to the original reports on the eIF4G, NRF, Rbm3, and XIAP IRESs, in which IRES activities appeared to be far higher. In those studies, IRES activities from dicistronic constructs containing these sequences were reported to be ≈40 or more times greater than from controls containing no IRES insert (13, 17) or 12 or more times greater than from those containing the EMCV or poliovirus IRES (14, 21). This discrepancy cannot be attributed to cell line-specific differences because in each of these studies at least one of the lines we used was also used. Also unexpectedly, we found that, despite our efforts to eliminate potential 5′ ss in the reporter construct, 3′ ss in the putative eIF4G and XIAP IRESs had caused splicing of the dicistronic transcript (SI Fig. 7). In both cases, a single 3′ ss in the IRES had spliced with multiple cryptic 5′ ss located within the GLuc sequence or, remarkably, within the IRES itself.

We hypothesized that the low activity of the IRESs in the

Fig. 1. Evaluation of six putative cellular IRESs in a dicistronic reporter plasmid depleted of upstream 5′ splice sites. The ability of the six sequences to mediate second-cistron translation was compared with that of the EMCV IRES and various non-IRES control sequences. The control sequences include the 5′ UTRs of the β-globin, γ-actin, and JunD mRNAs and the putative BiP and XIAP IRESs in reverse-complement (r.c.) orientation. (*A*) Schematic of the reporter construct. As indicated, the 5′ UTR and GLuc coding sequence were fully reengineered to eliminate potential cryptic 5′ ss. The bent arrow denotes the start site of transcription from the CMV promoter. pA, polyadenylation signal. (*B*) Results of transfection of HeLa and NIH 3T3 cells with the dicistronic reporter plasmids. GLuc luminescence (open bars, left axis) and GFP fluorescence (filled bars, right axis) values are the mean ± SD of three experiments. GFP fluorescence values are relative, and that for the EMCV IRES was set to 100. RLU, relative light units.

dicistronic reporter RNA was due to the absence of upstream sequences that could efficiently function as 5′ ss and thereby permit formation of monocistronic transcripts. Although some of these IRESs have been tested for splicing by Northern analysis of RNA from cells transfected with dicistronic plasmids, this method may not be sufficiently sensitive to rule out low levels of spliced RNA (11). The use of a sensitive detection method is imperative because cap-mediated translation is generally more efficient than IRES-mediated translation and spliced transcripts can exhibit greater translational yield than otherwise identical unspliced transcripts (22). Additionally, Northern analysis may not provide adequate size resolution, because splicing events that remove only short segments of a dicistronic RNA can create fusions between the two cistrons, as we observed with transcripts containing the eIF4G and XIAP sequences (SI Fig. 7).

Fig. 2. Four of the six putative IRESs allow high transgene expression from a retrovirus genome but inhibit viral replication. (*A*) Schematic of the MLV genome and site of insertion of IRES-GFP cassettes. A fraction of the MLV RNA undergoes splicing at the indicated sites to allow expression of the *env* gene. Ψ, RNA packaging signal. (*B*) GFP expression in cells transfected with the proviral plasmids. Values are the mean ± SD of three experiments. Similar results were obtained by using NIH 3T3 cells (data not shown). (*C*) The presence of the eIF4G, NRF, Rbm3, and XIAP sequences in the viral genome inhibits replication. PC-3 cells were transfected with the proviral plasmids, and replication was monitored by quantitation of RT activity in the culture supernatant at 2, 4, and 6 days after transfection. Values are the mean of three replicate transfections.

Four Putative IRESs Cause Aberrant Splicing of Retroviral RNA. To screen the putative IRESs for 3′ ss using a more sensitive approach, we introduced the IRES-GFP cassettes from the dicistronic plasmids into a murine leukemia virus (MLV) proviral genome (Fig. 2*A*). MLV must maintain a balance between its full-length genomic transcript and spliced *env* transcript to undergo productive replication. We previously found that MLV is highly sensitive to modifications to its genome that perturb normal splicing of the viral transcript and that attenuation of replication can serve as a useful indicator of aberrant splicing (23). In stark contrast to the results obtained with the 5′ ss-depleted dicistronic constructs, the proviruses containing the eIF4G, NRF, Rbm3, and XIAP sequences produced exceedingly high levels of GFP upon transfection (Fig. 2*B*). Notably, the proviruses containing the eIF4G, NRF, Rbm3, and XIAP sequences also appeared to produce little replicating virus, as judged by reverse transcriptase (RT) activities in transfected cultures (Fig. 2*C*), the low amounts of virion RNA in the transfection supernatant (SI Fig. 8*A*), and the lack of progression in GFP-positivity in cells exposed to the supernatant (SI Fig. 8*B*). The proviral constructs containing the EMCV, BiP, or SP163 IRES, or no IRES, each produced efficiently replicating virus.

We then performed RT-PCR on the viral RNA with primer sets designed to identify splicing between potential 3′ ss in the IRESs and either the native MLV 5′ ss or potential cryptic 5′ ss in the *env* gene (Fig. 3*A*). Whereas none of the proviruses that
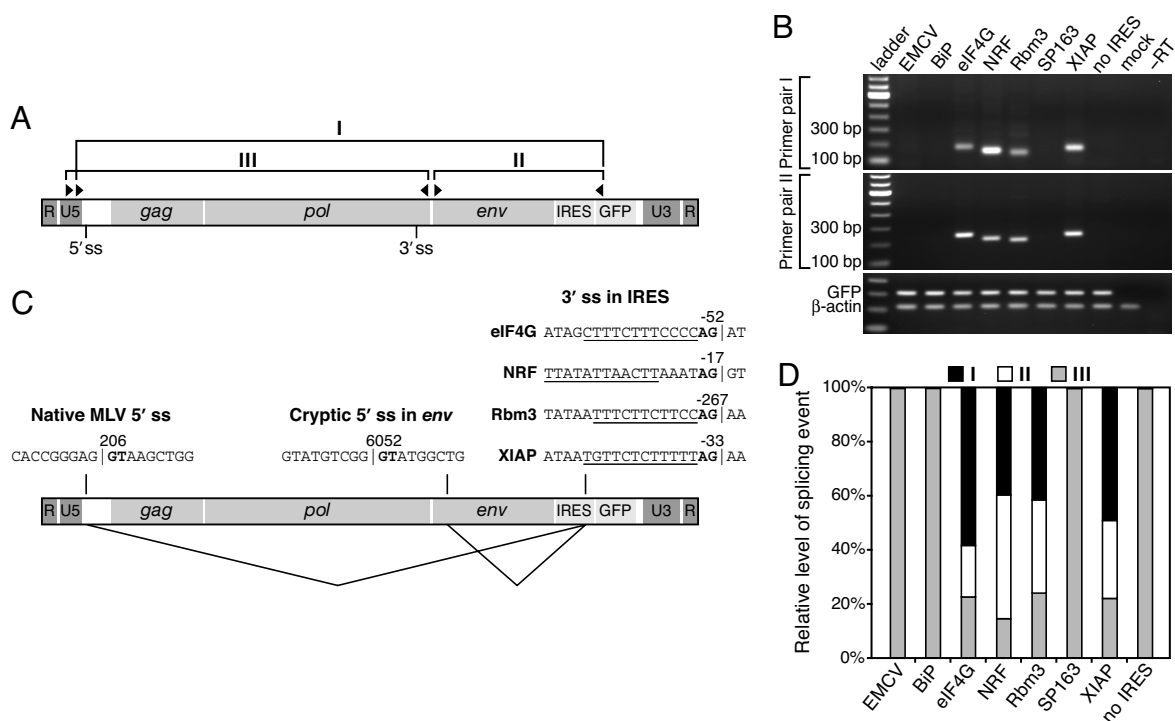
Baranick *et al.*

Fig. 3. The eIF4G, NRF, Rbm3, and XIAP IRESs contain 3′ splice sites. (*A*) Schematic of the unspliced viral RNA showing the locations of the primers used to amplify spliced isoforms by RT-PCR. Primer set I detects splicing between the native MLV 5′ ss and a 3′ ss in the IRES; set II detects splicing between a cryptic 5′ ss within *env* and a 3′ ss in the IRES; and set III detects splicing between the native 5′ and 3′ MLV splice sites. (*B*) Results of RT-PCR with primer sets I and II. (*C*) Splicing events identified by sequencing of the amplified products in *B*. Vertical lines indicate exon–intron and intron–exon borders, and the GT and AG dinucleotides of the 5′ and 3′ ss, respectively, are in bold. The numbering in the MLV sequence is in accordance with GenBank accession no. J02255. The numbering in each IRES sequence represents the position relative to the start codon in the respective cellular mRNA, such that the first base of this codon is +1. Underlined sequences represent pyrimidine-rich tracts likely to be involved in splicing. (*D*) Quantitation of the spliced RNAs by real-time RT-PCR. The 0–100% scale assumes that only the three quantitated splice isoforms are present, although others may occur.

produced efficiently replicating virus yielded product with these primers, those containing the eIF4G, NRF, Rbm3, and XIAP sequences did with both sets (Fig. 3*B*). Sequencing of the amplicons showed that these four proposed IRESs all contained 3′ ss that had spliced with both the native MLV 5′ ss and a cryptic 5′ ss in *env* (Fig. 3*C*) and that the eIF4G and XIAP 3′ ss were the same as those used in the GLuc-GFP dicistronic RNAs. The aberrant splicing events had occurred in each case at levels higher even than that between the native 5′ and 3′ MLV ss (Fig. 3*D*). These results implied that the IRES 3′ ss were responsible for the strong GFP expression in transfected cells, because splicing between the native MLV 5′ ss and the IRES 3′ ss positions the GFP start codon as the first AUG in the viral RNA, and, like cellular mRNA, retroviral RNA is capped (24).

3′ Splice Site Activity Corresponds with Apparent IRES Activity. To ascertain the ability of the eIF4G, NRF, Rbm3, and XIAP sequences to mediate GFP expression in this system in the absence of splicing, we sought to generate point mutations that inactivated the identified 3′ ss. As an approach to producing such mutations, we serially passaged the viruses containing these IRESs to evolve mutants in which the splice sites had been inactivated through natural selection. We previously used this method to isolate rapidly replicating mutants of a retrovirus that had been crippled by an oversplicing defect (23). From five passaged isolates of the NRF and Rbm3 viruses, we identified three unique mutants, each of which possessed a single or double point mutation at or adjacent to the IRES 3′ ss (SI Fig. 9 *A* and *B*). From the eIF4G and XIAP viruses, however, we were able to identify only mutants that had acquired deletions involving sizable segments of the IRES sequence, all of which resulted in

loss of the 3′ ss (SI Fig. 9 *A* and *C*). We therefore introduced subtler point mutations into the eIF4G and XIAP 3′ ss by site-directed mutagenesis (SI Fig. 9*D*).

Although none of the mutations completely eliminated use of the 3′ ss, we observed a clear relationship between the level of residual splicing and GFP expression (Fig. 4), and those mutants exhibiting the least splicing at the mutated site also expressed GFP at levels approaching that of the provirus containing no IRES. Consistent with these results, mutations in elements of the 3′ ss (either the polypyrimidine tract or AG dinucleotide) in the eIF4G, Rbm3, and XIAP sequences were previously found to greatly reduce apparent IRES activity (13, 15, 17). We additionally observed a strong correspondence between residual splicing and inhibition of viral replication (SI Fig. 10), demonstrating that the replication impairment in the four parental viruses was indeed a consequence of anomalous splicing.

One Putative IRES Sequence Appears to Be a Cloning Artifact. To determine whether the 3′ ss are active in their native transcripts, we examined the evidence among ESTs. Unexpectedly, while searching for ESTs corresponding to the proposed Rbm3 IRES, we found that most ESTs matching the 5′ UTR of the Rbm3 sequence (GenBank accession no. AY052560) on which the earlier studies were based map to mouse chromosome 18, rather than chromosome X, where the Rbm3 gene is located (data not shown). Further analysis of this sequence indicated that it represents a cloning artifact in which a cDNA for the Thoc1 gene on chromosome 18 had recombined with an Rbm3 cDNA (SI Fig. 11). The region of the 5′ UTR of AY052560 containing the 22-nt reported IRES sequence, and the 3′ ss we identified, in fact corresponds to the reverse complement of partially spliced
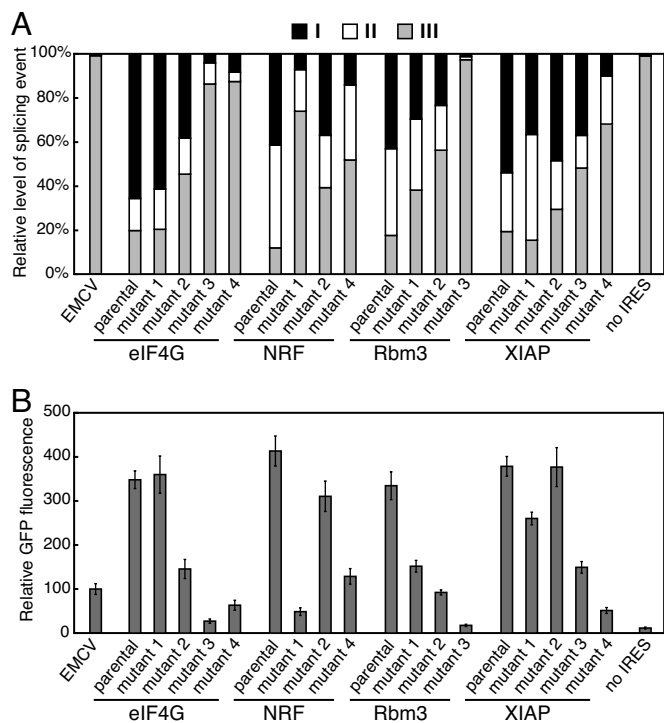
Fig. 4. Activity of the putative eIF4G, NRF, Rbm3, and XIAP IRES 3′ ss correlates with GFP expression level. (A) Real-time RT-PCR quantitation of spliced RNAs produced in cells transfected with proviral constructs containing point mutations in the identified 3′ ss. The primer sets used are as shown in Fig. 3A, and the 3′ ss mutations are detailed in SI Fig. 9 B and D. (B) GFP expression in the transfected cells. Values are the mean ± SD from three experiments.

Thoc1 mRNA and is thus not an authentic part of the Rbm3 transcript.

The 3′ Splice Sites in Three Putative IRESs Are Used in Their Native mRNAs. Inspection of the ESTs of the human and mouse Uni-Gene clusters for eIF4G, NRF, and XIAP demonstrated that the identified splice sites are all used in their native cellular transcripts (Table 1). The 3′ ss in the putative eIF4G IRES is known to be used in formation of several of the eIF4G transcript variants (SI Fig. 12A) (25). In transcript variant 5, which contains the proposed IRES and is transcribed from its own promoter, the 3′ ss is not thought to be used. However, the majority of ESTs spanning this site showed splicing (SI Fig. 12B). In each of the NRF and XIAP clusters, several or more ESTs had been spliced at the identified site whereas either none or only one contained

Table 1. The 3′ ss identified in the putative eIF4G, NRF, and XIAP IRESs are utilized in their native mRNAs

	Human		Mouse	
mRNA	Spliced	Unspliced	Spliced	Unspliced
eIF4G (all variants)	69	11	37	1
eIF4G (variant 5)	19	ND	1	ND
NRF	7	1	9	0
XIAP	12	1	16	1

Shown are the number of ESTs in the eIF4G, NRF, and XIAP UniGene clusters that were spliced at the identified site or that contain the site unspliced. ND, not determinable, because it is not possible to distinguish between an EST from an unspliced eIF4G variant 5 transcript and one from an incompletely processed transcript of another variant. GenBank accession numbers for all ESTs are given in SI Table 2.

the 3′ ss in unspliced form. Structural schematics of these ESTs are shown in SI Figs. 13 and 14, respectively.

To ascertain the possible frequency of 3′ splice sites in native mRNAs of other reported cellular IRESs, we performed a survey of ESTs from ≈80 other such elements. We identified 12 other putative IRESs for which ESTs indicate the presence of 3′ ss (SI Table 3). Whether or not splicing at these sites contributed to apparent IRES function in previous studies merits investigation.

Discussion

Our findings strongly support an alternative explanation, splicing, for the seeming ability of the putative eIF4G, NRF, Rbm3, and XIAP IRESs to mediate internal ribosome entry. They moreover undermine much of the original rationale for postulation of IRES-mediated translation of the corresponding cellular mRNAs. Like picornaviral RNAs, these cellular mRNAs appeared to possess unusually long 5′ UTRs burdened with a large number of uAUGs that scanning ribosomes would somehow need to traverse before reaching the start codon. Internal ribosome entry was proposed as the mechanism whereby these obstacles to translation could be circumvented (14, 17, 21, 26). However, the natural 5′ UTR of Rbm3 is much shorter than described and contains no uAUGs (SI Fig. 11). The putative IRES-containing eIF4G 5′ UTR is also shorter than reported and contains no more than a single uAUG, because of both splicing within the putative IRES sequence itself and the use of an alternative transcription start site (SI Fig. 12B). In the case of the NRF and XIAP mRNAs, most of the putative IRES sequence is located within an intron whose excision results in a UTR that is completely free of uAUGs and is of typical length. The mature structures of these mRNAs are fully consistent with translation via the 5′ end- and cap-dependent scanning mechanism and thus do not require invocation of a nonconventional mode of translation initiation.

Previous findings also support our contention that the putative eIF4G, NRF, and XIAP IRESs function via splicing. Whereas studies on the eIF4G IRES that used conventional plasmid-based dicistronic reporter assays consistently found very high apparent IRES activities (13, 26, 27), testing of this sequence under conditions in which the reporter RNA was not subject to nuclear processing—by use of a bacteriophage T7 RNA polymerase-driven cytoplasmic expression plasmid (28), direct transfection of dicistronic RNA into the cytoplasm, or in vitro translation (29)—yielded activity comparable to that of controls containing no IRES. In each case, however, parallel testing with controls containing picornaviral IRESs produced clearly positive results for IRES function. Similarly, the NRF IRES was reported to be inactive when tested by in vitro translation (30), and the XIAP IRES was found to be barely active or inactive in direct RNA transfections (11) and in the T7 vaccinia virus expression system (31) wherein the RNA is transcribed in the cytoplasm. The prototypical EMCV IRES, by contrast, is functional in all of the aforementioned systems (4, 28, 32, 33).

It has been argued (34) that the previously observed activity of the 3′ splice site in the XIAP IRES (11) was a result of the presence of strong 5′ ss in the particular dicistronic construct in which it had been tested. However, even in the context of the GLuc-GFP dicistronic reporter from which we attempted to eliminate all likely 5′ ss upstream of the IRES, the 3′ ss in the XIAP and eIF4G sequences caused activation of multiple cryptic 5′ ss. These included 5′ ss within the IRES sequences themselves, as well as one in the GLuc gene bearing an unconventional GC donor dinucleotide (SI Fig. 7). It is clear that this promiscuous splicing was not due to a unique characteristic of our dicistronic construct, because these two putative IRESs, as well as the NRF and Rbm3 sequences, also potently activated at least one cryptic 5′ ss in the MLV RNA. More importantly, the identified eIF4G,

NRF, and XIAP 3′ ss are all used in the maturation of their respective natural mRNAs.

It is notable that, in both of our reporter systems, the BiP and SP163 sequences yielded activity comparable to the non-IRES control sequences. By comparison, the original studies on these elements described activities 15- to 20-fold greater than negative controls (16, 35). Subsequent studies also reported apparent IRES activities far lower than those originally described. For the BiP IRES, these ranged between 1- and 4-fold over non-IRES controls (16, 36–38), and, for the SP163 IRES, the one study we identified reported no IRES function, in agreement with our results (39). Given that we observed 3-fold differences in apparent "IRES" activity among the negative control sequences we tested, we speculate that the originally observed activities of the putative BiP and SP163 IRESs might have been due to the particular reporter systems in which they were evaluated. Further investigation is needed to resolve this question.

Because most studies relying on the dicistronic test for IRES activity have not included RNA structural analysis beyond Northern blotting, and because many others included no RNA analysis at all (8, 9, 40), the question of whether or not the dicistronic constructs used yield any monocistronic RNA remains open. Perhaps significantly, it has been noted that a number of other cellular IRESs have also been found to function in conventional plasmid-based dicistronic assays but not in assays that preclude nuclear processing of the reporter RNA, implying the need for a "nuclear experience" for the function of these putative IRESs (41). Our data indicate that splicing may in at least some cases represent this essential nuclear experience. Nevertheless, the high incidence of active 3′ ss in the putative cellular IRESs we tested suggests that conclusive demonstration of internal ribosome entry by other reported cellular IRESs will require rigorous reexamination with sensitive methods to rule out splicing.

Materials and Methods

Plasmid Construction. The dicistronic reporter plasmids were constructed by using pGL4.75[hRluc/CMV] (Promega) as the backbone. First, the HindIII-XbaI region of pGL4.75[hRluc/CMV] containing the 5′ UTR and *Renilla* luciferase gene was replaced with a fragment containing a 5′ UTR and GLuc (42) coding sequence reengineered to eliminate potential 5′ ss. A multiple cloning site (MCS) was included after the GLuc stop codon for insertion of IRES and GFP sequences. The 5′ UTR-GLuc-MCS cassette was synthesized in its entirety to our specifications by Bio Basic. Before synthesis, we identified potential 5′ ss in the cassette using the NetGene2 (43) (www.cbs.dtu.dk/services/NetGene2) and NNSplice (44) (www.fruitfly.org/seq_tools/splice.html) ss prediction programs. Wherever possible, potential ss were eliminated by silent mutation. To produce the GLuc-GFP reporter plasmids, we inserted the IRES and GFP sequences into the MCS of the resulting plasmid (pCGLuc) in three-fragment ligations. The cellular IRES sequences (SI Fig. 15) and non-IRES control sequences were obtained by custom synthesis, annealing synthetic polynucleotides, or PCR amplification from human genomic DNA. The dicistronic control plasmid containing no IRES was produced by inserting the GFP transgene alone into pCGLuc. The dicistronic plasmid containing the EMCV IRES was constructed by ligation of the IRES amplified from plasmid pACE-GFP (45) and the GFP gene into pCGLuc.

All provirus plasmids were derived from pACE-GFP, which contains a full-length amphotropic MLV genome encoding a replication-competent virus bearing an EMCV IRES-GFP cassette inserted immediately downstream of the *env* gene (45). The IRES-GFP cassette in this plasmid was replaced with the IRES-GFP cassettes from the GLuc-GFP dicistronic constructs to generate proviruses containing the cellular IRESs. The provirus plasmids containing 3′ ss mutations were generated with the QuikChange II XL site-directed mutagenesis kit (Stratagene). All constructs used in this study were sequenced to verify their integrity.

Cell Culture and Transfections. 293T, HeLa, and NIH 3T3 cells were cultivated in DMEM, and PC-3 cells were cultivated in RPMI medium 1640. All media were supplemented with 10% FBS. Transfections were carried out by using FuGENE 6 (Roche). β-Galactosidase expression plasmid pCH110 was cotransfected with each provirus plasmid, and the β-Galactosidase Enzyme Assay System (Promega) was

used to quantitate transfection efficiency. Luciferase and β-galactosidase activity and GFP fluorescence were determined at 48 h after transfection.

GFP Quantitation by Flow Cytometry. Cells were analyzed by using a Beckman Coulter Epics XL flow cytometer and EXPO32 ADC software. A 488-nm laser was used for excitation, and emission by GFP was measured through a 510/20-nm bandpass filter. GFP fluorescence is calculated as the percentage of GFP-positive cells multiplied by the mean fluorescence intensity of these cells. For transfections with proviral constructs, the β-galactosidase activity from cotransfected pCH110 was quantitated as described above and used to normalize GFP fluorescence values according to transfection efficiency.

Luciferase Assays. GLuc activity in cell culture supernatants was determined by using the Gaussia Luciferase Assay Kit (New England BioLabs) and a Sirius luminometer (Berthold Instruments). All luminescence readings were corrected for cell number and for background luminescence from mock-transfected cells exposed to substrate.

Virion RNA Dot Blotting. Measurement of virion RNA in the supernatant of transfected 293T cultures was performed as previously described (46). Briefly, 300 μl of filtered (0.45 μM) supernatant was combined with denaturing buffer and transferred to a nylon membrane using a vacuum manifold. The RNA was probed with a random-primed, [α-^{32}P]dCTP-labeled probe specific for the MLV *env* region. Hybridization was carried out by using MiracleHyb buffer (Stratagene), and signal quantitation was performed with an FX Pro Plus MultiImager (Bio-Rad).

Reverse Transcriptase Assays. Assays were performed on supernatant of transfected PC-3 cells collected every 2 days after transfection. RT activity was measured by using the Quan-T-RT kit (GE Healthcare) per the manufacturer's instructions except for the following modification: The 1× assay buffer was altered to contain 50 mM Tris (pH 8.0), 40 mM NaCl, 4 mM MnCl$_2$, 3 mM 2-mercaptoethanol, 5 mM spermidine, and 0.05% Nonidet P-40. The reactions products were quantitated by using a Packard TopCount microplate scintillation counter.

cDNA Synthesis. Total RNA was isolated from transfected cells using an RNeasy Mini kit with optional DNase I treatment (Qiagen). Cell homogenization before RNA isolation was carried out with QIAshredder columns. Two micrograms of RNA was reverse-transcribed by using SuperScript III reverse transcriptase and random hexamer primers (Invitrogen). Reaction products were treated with RNase H (Invitrogen) after each cDNA synthesis reaction.

RT-PCR. Standard endpoint RT-PCR was performed with Phusion DNA polymerase (Finnzymes). For amplification of spliced viral transcripts, each cycle consisted of 8 seconds at 98°C, 22 seconds at 58°C, and 8 seconds at 72°C. For luciferase-GFP dicistronic reporter transcripts, each cycle consisted of 8 seconds at 98°C, 22 seconds at 55°C, and 15 seconds at 72°C. Real-time RT-PCR was carried out with TaqMan Universal PCR MasterMix and an ABI PRISM 7700 Sequence Detection System (Applied Biosystems). Spliced transcript values were normalized to GFP values to control for transfection efficiency, and GFP values were normalized to β-actin values. The amount of each spliced transcript was calculated as a percentage of the total of all spliced transcripts for each sample. Primer and probe sequences are shown in SI Table 4.

***In Vitro* Adaptation and Cloning of Virus Mutants.** PC-3 cells were transfected with provirus plasmid and cultivated for 3 weeks to allow the outgrowth of efficiently replicating virus mutants. At 3 weeks after transfection, the culture supernatant was used to infect fresh cells. These cells were propagated for 1 week, after which point a final infection was carried out. Genomic DNA was isolated from these cells after 1 week, and PCR was used to amplify the IRES-GFP cassettes from the adapted provirus using primers ACTGATCT-TACTCTTTGGACCTTG and CCCTTTTTCTGGAGACTAAATAA. The resulting products were reintroduced into the parental plasmid, replacing the entire IRES-GFP cassette.

EST Analysis. The human and mouse EST sequences of the eIF4G (EIF4G1), NRF (NKRF), and XIAP (BIRC4) UniGene clusters (build no. 203 for human and no. 164 for mouse) were obtained from the National Center for Biotechnology Information web site. These ESTs were aligned to their respective genomic loci by using Spidey (www.ncbi.nlm.nih.gov/spidey/index.html) to identify those that span the 3′ ss identified by RT-PCR and determine whether or not they exhibited splicing at these sites.

CELL BIOLOGY

1. Gray NK, Wickens M (1998) Control of translation initiation in animals. *Annu Rev Cell Dev Biol* 14:399–458.
2. Kozak M (1989) The scanning model for translation: An update. *J Cell Biol* 108:229–241.
3. Kozak M (2002) Pushing the limits of the scanning mechanism for initiation of translation. *Gene* 299:1–34.
4. Jang SK, et al. (1988) A segment of the 5′ nontranslated region of encephalomyocarditis virus RNA directs internal entry of ribosomes during in vitro translation. *J Virol* 62:2636–2643.
5. Pelletier J, Sonenberg N (1988) Internal initiation of translation of eukaryotic mRNA directed by a sequence derived from poliovirus RNA. *Nature* 334:320–325.
6. Belsham GJ, Sonenberg N (2000) Picornavirus RNA translation: Roles for cellular proteins. *Trends Microbiol* 8:330–335.
7. Baird SD, Turcotte M, Korneluk RG, Holcik M (2006) Searching for IRES. *RNA* 12:1755–1785.
8. Kozak M (2001) New ways of initiating translation in eukaryotes? *Mol Cell Biol* 21:1899–1907.
9. Kozak M (2005) A second look at cellular mRNA sequences said to function as internal ribosome entry sites. *Nucleic Acids Res* 33:6593–6602.
10. Schneider R, et al. (2001) New ways of initiating translation in eukaryotes. *Mol Cell Biol* 21:8238–8246.
11. Van Eden ME, Byrd MP, Sherrill KW, Lloyd RE (2004) Demonstrating internal ribosome entry sites in eukaryotic mRNAs using stringent RNA test procedures. *RNA* 10:720–730.
12. Yang Q, Sarnow P (1997) Location of the internal ribosome entry site in the 5′ non-coding region of the immunoglobulin heavy-chain binding protein (BiP) mRNA: Evidence for specific RNA-protein interactions. *Nucleic Acids Res* 25:2800–2807.
13. Gan W, LaCelle M, Rhoads RE (1998) Functional characterization of the internal ribosome entry site of eIF4G mRNA. *J Biol Chem* 273:5006–5012.
14. Oumard A, Hennecke M, Hauser H, Nourbakhsh M (2000) Translation of NRF mRNA is mediated by highly efficient internal ribosome entry. *Mol Cell Biol* 20:2755–2759.
15. Chappell SA, Mauro VP (2003) The internal ribosome entry site (IRES) contained within the RNA-binding motif protein 3 (Rbm3) mRNA is composed of functionally distinct elements. *J Biol Chem* 278:33793–33800.
16. Stein I, et al. (1998) Translation of vascular endothelial growth factor mRNA by internal ribosome entry: Implications for translation under hypoxia. *Mol Cell Biol* 18:3112–3119.
17. Holcik M, Lefebvre C, Yeh C, Chow T, Korneluk RG (1999) A new internal-ribosome-entry-site motif potentiates XIAP-mediated cytoprotection. *Nat Cell Biol* 1:190–192.
18. Johannes G, Sarnow P (1998) Cap-independent polysomal association of natural mRNAs encoding c-myc, BiP, and eIF4G conferred by internal ribosome entry sites. *RNA* 4:1500–1513.
19. Grifo JA, Tahara SM, Morgan MA, Shatkin AJ, Merrick WC (1983) New initiation factor activity required for globin mRNA translation. *J Biol Chem* 258:5804–5810.
20. Short JD, Pfarr CM (2002) Translational regulation of the JunD messenger RNA. *J Biol Chem* 277:32697–32705.
21. Chappell SA, Owens GC, Mauro VP (2001) A 5′ leader of Rbm3, a cold stress-induced mRNA, mediates internal initiation of translation with increased efficiency under conditions of mild hypothermia. *J Biol Chem* 276:36917–36922.
22. Nott A, Le Hir H, Moore MJ (2004) Splicing enhances translation in mammalian cells: An additional function of the exon junction complex. *Genes Dev* 18:210–222.
23. Logg CR, Baranick BT, Lemp NA, Kasahara N (2007) Adaptive evolution of a tagged chimeric gammaretrovirus: Identification of novel cis-acting elements that modulate splicing. *J Mol Biol* 369:1214–1229.
24. Rabson AB, Graves BJ (1997) in *Retroviruses*, eds Coffin JM, Hughes SH, Varmus H (Cold Spring Harbor Lab Press, Plainview, NY), pp 205–261.
25. Byrd MP, Zamora M, Lloyd RE (2005) Translation of eukaryotic translation initiation factor 4GI (eIF4GI) proceeds from multiple mRNAs containing a novel cap-dependent internal ribosome entry site (IRES) that is active during poliovirus infection. *J Biol Chem* 280:18610–18622.
26. Gan W, Rhoads RE (1996) Internal initiation of translation directed by the 5′-untranslated region of the mRNA for eIF4G, a factor involved in the picornavirus-induced switch from cap-dependent to internal initiation. *J Biol Chem* 271:623–626.
27. Wong ET, Ngoi SM, Lee CG (2002) Improved co-expression of multiple genes in vectors containing internal ribosome entry sites (IRESes) from human genes. *Gene Ther* 9:337–344.
28. Finn J, MacLachlan I, Cullis P (2005) Factors limiting autogene-based cytoplasmic expression systems. *FASEB J* 19:608–610.
29. Han B, Zhang JT (2002) Regulation of gene expression by internal ribosome entry sites or cryptic promoters: The eIF4G story. *Mol Cell Biol* 22:7372–7384.
30. Reboll MR, et al. (2007) NRF IRES activity is mediated by RNA binding protein JKTBP1 and a 14-nt RNA element. *RNA* 13:1328–1340.
31. Holcik M, Gordon BW, Korneluk RG (2003) The internal ribosome entry site-mediated translation of antiapoptotic protein XIAP is modulated by the heterogeneous nuclear ribonucleoproteins C1 and C2. *Mol Cell Biol* 23:280–288.
32. Elroy-Stein O, Moss B (1990) Cytoplasmic expression system based on constitutive synthesis of bacteriophage T7 RNA polymerase in mammalian cells. *Proc Natl Acad Sci USA* 87:6743–6747.
33. Wang Z, Weaver M, Magnuson NS (2005) Cryptic promoter activity in the DNA sequence corresponding to the pim-1 5′-UTR. *Nucleic Acids Res* 33:2248–2258.
34. Holcik M, et al. (2005) Spurious splicing within the XIAP 5′ UTR occurs in the Rluc/Fluc but not the betagal/CAT bicistronic reporter system. *RNA* 11:1605–1609.
35. Macejak DG, Sarnow P (1991) Internal initiation of translation mediated by the 5′ leader of a cellular mRNA. *Nature* 353:90–94.
36. Kim JG, Armstrong RC, Berndt JA, Kim NW, Hudson LD (1998) A secreted DNA-binding protein that is translated through an internal ribosome entry site (IRES) and distributed in a discrete pattern in the central nervous system. *Mol Cell Neurosci* 12:119–140.
37. Kim YK, Hahm B, Jang SK (2000) Polypyrimidine tract-binding protein inhibits translation of bip mRNA. *J Mol Biol* 304:119–133.
38. Nevins TA, Harder ZM, Korneluk RG, Holcik M (2003) Distinct regulation of internal ribosome entry site-mediated translation following cellular stress is mediated by apoptotic fragments of eIF4G translation initiation factor family members eIF4GI and p97/DAP5/NAT1. *J Biol Chem* 278:3572–3579.
39. Yu X, et al. (2003) Lentiviral vectors with two independent internal promoters transfer high-level expression of multiple transgenes to human hematopoietic stem-progenitor cells. *Mol Ther* 7:827–838.
40. Kozak M (2003) Alternative ways to think about mRNA sequences and proteins that appear to promote internal initiation of translation. *Gene* 318:1–23.
41. Spriggs KA, Bushell M, Mitchell SA, Willis AE (2005) Internal ribosome entry segment-mediated translation during apoptosis: The role of IRES-trans-acting factors. *Cell Death Differ* 12:585–591.
42. Tannous BA, Kim DE, Fernandez JL, Weissleder R, Breakefield XO (2005) Codon-optimized Gaussia luciferase cDNA for mammalian gene expression in culture and in vivo. *Mol Ther* 11:435–443.
43. Hebsgaard SM, et al. (1996) Splice site prediction in Arabidopsis thaliana pre-mRNA by combining local and global sequence information. *Nucleic Acids Res* 24:3439–3452.
44. Reese MG, Eeckman FH, Kulp D, Haussler D (1997) Improved splice site detection in Genie. *J Comput Biol* 4:311–323.
45. Logg CR, Logg A, Matusik RJ, Bochner BH, Kasahara N (2002) Tissue-specific transcriptional targeting of a replication-competent retroviral vector. *J Virol* 76:12783–12791.
46. Onodera M, et al. (1997) A simple and reliable method for screening retroviral producer clones without selectable markers. *Hum Gene Ther* 8:1189–1194.