Translational efficiency of EMCV IRES in bicistronic vectors is dependent upon IRES sequence and gene location

Yury A. Bochkov and Ann C. Palmenberg
University of Wisconsin-Madison, Madison, WI, USA

BioTechniques 41:283-292 (September 2006)
doi 10.2144/000112243

INTRODUCTION

The internal ribosomal entry site (IRES) from encephalomyocarditis virus (EMCV) is a popular RNA element used widely in experimental and pharmaceutical applications to express proteins in eukaryotic cells or cell-free extracts. Inclusion of the wild-type element in monocistronic or bicistronic messenger RNAs (mRNAs) confers a high level of cap-independent translation activity to appropriately configured cistrons. The history of this element and the experimental consequences of sequence derivations inherent to commercial IRES vectors are less well known. Compared head-to-head with dual-luciferase reporter constructs, a native EMCV IRES in a bicistronic configuration directed 8- to 10-fold more protein than a similarly configured pIRES vector. It also produced nearly twice as much protein as pCITE®-1, an early monocistronic iteration, harboring a suboptimal A7 sequence in a crucial structural motif. The results indicate that investigators should be aware of and carefully report the sequence of their IRES in any comparative study. The preferred IRES (viral bases 273–845) and the minimum IRES (viral bases 400–836) for optimum activity are illustrated.

The internal ribosomal entry site (IRES) from encephalomyocarditis virus (EMCV) is a noncoding RNA fragment noted for its ability to initiate high levels of cap-independent protein synthesis in mammalian cells and cell-free extracts. In 1986, shortly after EMCV-R was sequenced (1), the region responsible for this activity was localized to a genome fragment about 430 bases long, immediately 5’ to the AUG, which begins the viral polyprotein open reading frame (ORF) (Figure 1A) (2). When the enhancer element was excised and linked to other portions of the virus ORF, the resulting T7 transcripts were highly active messenger RNAs (mRNAs) in the absence of 5’ capping reactions.

This useful discovery was commercialized in 1990 by Novagen (Madison, WI, USA) in the form of pCITE®-1, a vector that allowed easy linkage of exogenous cistrons onto the cap-independent translation enhancer for transcription of hybrid mRNAs and facile protein expression in cell-free extracts. pCITE-1 enjoyed a wide distribution because of the high level of enhancer-confferred protein synthesis (3). In 1992, that popularity led to the release of pCITE-2, a derivative vector, differing in several important features.

First, an expanded multiple cloning site (MCS) was engineered 3’ of the natural virus AUG, to allow ligation of cistrons from multiple reading frames. The original plasmid, based on viral pE5LVP0 (2) had utilized only native BalI (MscI) and NcoI sequences (Figure 1B). Second, an ACC triplet was inserted upstream of the AUG, to create a new NcoI site and to bring this codon into better accord with canonical (at that time) eukaryotic initiation sequences. The third difference corrected an isolate-specific discrepancy within the enhancer element, whereby the published EMCV-R sequence, 5’-GGTTAAAAACGTC-3’ (the A6 sequence), had by chance, been recloned by Novagen as 5’-GGTAAAAAACGTC-3’ (the A7 sequence).

The questionable oligo(A) was in a highly conserved cardioivirus and aphthovirus bifurcation loop at the junction of stems J and K (Figure 1A), and pCITE-2 was adjusted relative to pCITE-1 to conform with the native sequence (4). The consequence of this change was not immediately appreciated. It was only later reported that A6 versus A7 segments actually required somewhat different cohorts of translation factors for optimal activity (5) and therefore had the potential to respond differently in a range of cell types. It is now recognized that this particular loop is crucial to translational function, because it interacts with translation initiation factors (6) and, in its genome context, with the EMCV Leader protein (7).

In 1988, the availability of vectors to enhance translation, in any format, sparked lively interest in new eukaryotic expression techniques. One particularly innovative experimental series linked the (pE5LVP0) segment downstream of a cap-dependent reporter gene in a bicistronic configuration. Translation of the EMCV-driven cistron proved independent of the upstream gene, an activity that could not be attributed to leaky scanning or ribosome read-through mechanisms (8). Dubbed an internal ribosome entry site, or IRES, the moniker has stuck with this element and has since been applied to many other virus or cellular RNA fragments that also direct eukaryotic translation in a cap-independent manner. To be sure, other viral noncoding segments, most notably from poliovirus (9), were tested at about this time in analogous bicistronic constructions and also shown to be IRESs. The superior activity of the EMCV IRES in multiple cell and cell-free systems, however, popularized its use.

In 1995, Clontech capitalized on the bicistronic potential and released a new vector (pIRES) that could be delivered directly into mammalian cells as cDNA. Transfection with pIRES induced nuclear transcription of capped bicistronic mRNAs, driven by a cytomegalovirus (CMV) promoter. Translation of any upstream gene “A” inserted into the 5’-most MCS-A was cap-dependent. Linked in tandem was a modified EMCV IRES, a second MCS (B), and a poly(A) signal sequence. Protein synthesis for gene “B” inserted at MCS-B was IRES-dependent. Since the native EMCV IRES usually overexpressed protein relative to an equivalent capped mRNA, Clontech chose to partially disable the bicistronic...
EMCV segment to attenuate B cistron translation down to the more moderate level of the A cistron (10). Accordingly, pIRES vectors have A7 rather than A6 bifurcation sequences, and the MCS-B does not include a native EMCV AUG (Figure 1B).

Today’s literature describes thousands of cDNA constructions utilizing IRES technology. Mono- and bicistronic vectors using the poliovirus or EMCV IRES are in common experimental and commercial use for heterologous protein expression, gene therapy, and gene transfer experiments. Compared head-to-head, an EMCV IRES generally directs higher translation levels than any other IRES or capped mRNA (11). But there are several reports from a variety of cell cultures, using various reporters genes, suggesting EMCV activity, particularly in bicistronic constructions, is sometimes unexpectedly low (12–14). Conflicting reports document 6% to 100% activity, relative to A cistron expression (12,13,15). The differences among pCITE-1, pCITE-2, pCITE-4 (Novagen’s current vector), pIRES, and the original pESLVP0 segments are acknowledged to link directly to the relative IRES expression activities. Nonetheless, the ubiquitous practice of sharing of plasmids among colleagues usually means that few laboratories accurately trace or correctly report the parentage of their preferred IRES segments. While the nature of the A and B ORFs chosen for these tests can influence the outcome (16), several low activity IRESs have their origins in pIRES, pCITE-1, or pCITE-2 vectors constructed for other purposes and perhaps erroneously assumed to be of native EMCV sequence (14). Other low activity sequences have suboptimal configurations in their IRES-cistron linkages (13), but the authors were seemingly unaware their choice of sequence, context, and IRES integrity were integral to translational activity.

A functional IRES is a dynamic RNA structure. To be optimal, the EMCV element, or any IRES for that matter, must be introduced in an orientation that allows it to fold properly without structural or heterologous sequence interference (4). The orientation of the IRES AUG at the 5’ border of the cistron, is particularly important (17,18). If the requirements are violated, an improperly configured IRES will preclude 40S ribosomal subunit recruitment. To clarify exactly how the IRES origin influences expression, we constructed a matched series of bicistronic, dual-luciferase reporter plasmids reproducing common commercial arrangements of the EMCV IRES. The translational efficiencies of the B cistron reporters clearly depended upon the specific IRES sequence and linkage to its gene.

**MATERIALS AND METHODS**

**Eukaryotic Expression Vectors**

Plasmid pRL-CMV (Promega, Madison WI, USA) encodes a Renilla luciferase gene (Rluc) whose transcription is driven by a CMV promoter. Translation after cDNA transfection into cells is cap-dependent. Plasmid pIRES (Clontech, Mountain View, CA, USA) has a similar promoter, 2 MCS (-A and -B) located on either side of an attenuated EMCV IRES (att-...
IRES), and a simian virus 40 (SV40) poly(A) signal downstream of MCS-B. Plasmid pF/- is a pIRES derivative originally called pIRES-luc (19). It links a firefly luciferase gene (Fluc) from pGEM-Dirit (Promega) into MCS-A through the Nhel-MulI restriction sites. MCS-B is empty. In plasmid pFLR-att, an Rluc gene, amplified by PCR from pRL-CMV (PR1 and PR2; see Table 1 for cloning primers) was inserted into MCS-B of pF/- using SalI-NolI sites. This bicistronic sequence has the A7 Clontech (attenuated) EMCV IRES (Figure 1B). Plasmid pF/- is the reciprocal of pF-. An Rluc gene was amplified by PCR from pFL-CMV (PR3 and PR4) and then inserted into pIRES MCS-A using Nhel-MulI sites. To make pFL/R-7A, a segment of pIRES was amplified by PCR (PR5 and PR6), combined with an Rluc product (PR7 and PR2), and then amplified with flanking primers. The resulting fragment was digested with PstMI (within the IRES) and NolI and cloned into pF/- that had been similarly digested. This plasmid differs from pF/R-wt by a single nucleotide insertion. It is A7 rather than A6 in the JK bifurcation loop. Plasmid pF/R-wt is similar to pF/R-att, except the 3’ half of the IRES sequence was replaced with an equivalent native EMCV segment (PR5 and PR6) derived from plasmid pECo (20). This IRES is an A6 variety, and the AUG of the Rluc ORF is in identical orientation to the AUG of the EMCV genome polyprotein ORF (Figure 1B). The identities of all cDNAs were confirmed by diagnostic restriction analyses and sequencing.

Cell-Free Translation

Plasmid DNA (1 μg) was linearized by digestion with HpaI, then transcribed into RNA with T7 RNA polymerase (Invitrogen, Carlsbad, CA, USA; 50 μL reaction for 90 min at 37°C), in the presence of Ribo mG cap Analog (Promega). The RNA was isolated with the presence of Ribo mG Cap Analog (Promega). The RNA was isolated using RNeasy® Mini kits (Qiagen, Valencia, CA, USA), eluted with water, then concentrated by size, integrity, and concentration (by staining and spectrophoto metrically) of the full-length products after agarose gel fractionation (gel patterns and Northern analyses not shown). Cell-free translation reactions in 15 μL rabbit reticulocyte extracts using [35S]Met label and 1 μg RNA transcripts were as described previously (4). After the reactions were complete (1 h at 30°C), two or four aliquots (1 μL each) from each sample were assayed for acid insoluble [35S]Met incorporation. Additional samples were fractionated by sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE; 12% Laemmli gels), the gels were dried, and the bands were visualized by phosphorimaging on an Amersham Biosciences’ Typhoon™ 9200 scanner (GE Healthcare, Piscataway, NJ, USA). The relative band intensities were captured with Molecular Dynamics ImageQuant™ software (GE Healthcare).

RESULTS AND DISCUSSION

Clontech pIRES was the founder plasmid for five mono- and bicistronic reporter cDNAs encoding functional genes for Renilla (Rluc) and/or firefly luciferase (Fluc) (Figure A). Our plasmid names reflect the order of their reporters (i.e., 5’ F/R 3’), and whether the IRES and adjoining MCS-B were the same as pIRES (e.g., pF/R-att), or pCITE-1 (e.g., pF/R-7A), or the native EMCV sequence from pEC9 (e.g., pF/R-wt). Plasmids pF/- and pR/- had alternate reporters in the upstream (MCS-A) cistron, with the second (MCS-B) cistron left blank. Plasmid pRL-CMV (Promega) is monocistronic and differs from pIRES in the specific sequence between the cap and MCS-A.

A T7 promoter is included in all plasmids to achieve facile in vitro transcription. Inclusion of mG in the reactions gives capped mRNAs that readily direct protein synthesis in rabbit reticulocyte extracts. Protein synthesis was monitored by [35S]Met incorporation into Fluc and Rluc bands after SDS-PAGE fractionation (Figure B). The relative protein content (RPC) directed by each sequence is reported as the observed band intensities (pixels) divided by the methionine per protein (Fluc, 15 Met; Rluc, 9 Met; Figure C).

Figure 2. Cell-free expression. (A) Plasmid maps illustrate the internal ribosomal entry site (IRES) configurations described in the Materials and Methods section. (B) [35S]Met-labeled proteins produced in reticulocyte extracts were fractionated by sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) and visualized by phosphorimaging. The Fluc and Rluc bands are indicated. (C) The band intensities (pixels x 10^4) in the gel depicted in panel B were digitized using ImageQuant and then divided by methionine per protein (Fluc, 15; Rluc, 9), wt, wild-type.
Short Technical Reports

In repeated experiments with matched samples of full-length RNAs, Rluc synthesis was maximum (RPC of 4295) when the reporter was configured in the B cistron of pF/R-wt. Synthesis of Fluc from the capped A cistron (RPC of 423) of the same mRNA was usually about 8- to 10-fold lower, a ratio observed with many samples and over different translation incubation times (all data not shown). Fluc expression increased (RPC of 848) only when the B cistron was left empty (pF/-), a finding consistent with results from other cell-free systems that propose a direct competition between a functional IRES and cap-dependent translation modes (21). Despite this competition, Rluc synthesis from pF/R-wt was about 40%–50% higher than when the same gene was tested in pIRES (pF/-) or pRL-CMV monocistronic (capped) configurations (RPC of 3077 and 2848, respectively). Therefore, reticulocyte ribosomes tend to initiate more effectively with the IRES than with a cap, and typically, the B cistron under direction of this IRES synthesized more protein per mRNA than an equivalent monocistronic capped transcript or bicistronic upstream A gene.

For this to happen, the IRES must be native. A pcITE-1-like sequence with A7 instead of A6 in the JK bifurcation loop (pF/R-7A) produced a similar amount of Fluc (RPC of 393 versus 432), but only 56% as much Rluc (RPC of 2407 versus 4295) as the wild-type sequence. The pIRES-like sequence of pF/R-att (RPC of 380) was about 11-fold lower than with wild-type or only about 8% as effective. Designed to produce nearly equivalent amounts of protein from both cistrons, Fluc synthesis (RPC of 402) and Rluc synthesis (RPC of 380) from this configuration were indeed quite similar. The pIRES sequence not only includes the suboptimal A7 bifurcation loop, but the initiating AUG for the MCS-B sequence, added by the user during cloning, is also displaced significantly (32 bases in pF/R-att) relative to the natural codon location. Clearly, this attenuated IRES is not an ideal selection for conditions where maximum protein synthesis is desired.

To test whether the relative IRES expression was unique to cell-free extracts, the panel of cDNAs was transfected into HeLa cells. Fluc activity from monocistronic pF/- was only marginally higher than from any bicistronic A cistron (Figure 3A), and as expected, Rluc activity from the B cistron was more responsive to the mode of expression (Figure 3B). The wt-IRES exceeded the A7 IRES by 15%–20% in Rluc relative luminescence units (RLU), and both sequences produced an average of 5–6 times more Rluc activity at all time points than the attenuated IRES. When mRNA concentrations from pF/R-wt and pF/R-att were directly compared by quantitative reverse transcription PCR (RT-PCR), the levels per cell were identical (data not shown). Therefore, any differences in RLU were due to the efficiency of protein translation. Since the specific activity of Fluc is roughly 8–10 times that of Rluc, the total comparative data are consistent with the idea that pIRES cDNA (pF/-att) produced a molar ratio of roughly 1:1 for the A:B cistron proteins, while the wt-IRES was at about 1:7, and the A7-IRES was at about 1:6. Interestingly, when the B cistron was empty (pR/-), Rluc activity from the capped cistron slightly exceeded that from the wt-IRES. As with pF/-, the relative edge for this expression mode may reflect the lack of a competing IRES. It is also possible any monocistronic RNA would have a slight advantage in transcription rate over a longer bicistronic or even a slightly preferred transfection efficiency because of the smaller cDNAs (Figure 3A).

The combined data reiterate the importance of the native EMCV IRES configuration for protein optimum expression in bicistronic contexts. Virus mapping experiments have suggested that ribosome recognition may actually be enhanced by including the first few codons of the polyprotein ORF (preferred IRES; Figure 3A) (22). However, since most IRES users prefer to avoid amino-terminal fusion sequences, very good expression (as shown here) can usually be achieved by linking the heterologous ORF directly to the native AUG in its native IRES context (5,17). A caveat to this linkage is that rare or unusual codons for bulky, charged amino acids (e.g., Pro, Trp, Cys, Arg, Lys) should probably be

**Table 1. Primers for cDNA Construction**

<table>
<thead>
<tr>
<th>Primer</th>
<th>Sequence</th>
<th>Restriction Site</th>
</tr>
</thead>
<tbody>
<tr>
<td>PR1 (+)</td>
<td>5’-TATAGGTCGACATGACTTCCAGT-3’</td>
<td>SalI</td>
</tr>
<tr>
<td>PR2 (-)</td>
<td>5’-CAGTGCGGCCGCTTATAGGTTCATTTTGGAG-3’</td>
<td>NotI</td>
</tr>
<tr>
<td>PR3 (+)</td>
<td>5’-TATAGGCTAGCATGACTTCCAGT-3’</td>
<td>NcoI</td>
</tr>
<tr>
<td>PR4 (+)</td>
<td>5’-CATTAAACGGTTATGTTCACTTTTGGAG-3’</td>
<td>MluI</td>
</tr>
<tr>
<td>PR5 (+)</td>
<td>5’-GAAGGTACCCTATGAGTGGATCTG-3’</td>
<td>PflMl</td>
</tr>
<tr>
<td>PR6 (-)</td>
<td>5’-ACTTTCCAGATCATATTACATCGTG-3’</td>
<td>—</td>
</tr>
<tr>
<td>PR7 (+)</td>
<td>5’-CAGCATGATAATGACTTCCAGT-3’</td>
<td>—</td>
</tr>
</tbody>
</table>

(+) Forward primer; (-) reverse primer.

**Figure 3. Luciferase activity.** Plasmid cDNAs were transfected into HeLa cells as described in the Materials and Methods section. At the indicated times, samples were assayed for (A) firefly or (B) Renilla activity. Duplicates were assayed, then averaged for each data point. Error bars indicate standard deviations from three separate experiments. RLU, relative luminescence units.

**Vol. 41 ı No. 3 ı 2006**

**BioTechniques** ı www.biotechniques.com

288
avoided within the first 2–3 codons of the cistron. The 5′ side of the minimum IRES needs to avoid disruption of stem H (4) defining a total viral fragment of about 430 bases, including the A6 bifurcation loop in the JK segment. All pCITE vectors, as well as the original pE5LVp0, extend 5′ for an additional 120–130 viral bases into an upstream oligo(C) tract (e.g., to base 273). The extra length is not required for ribosome recognition, but it can provide a useful spacer between the A and B cistrons, allowing the IRES to fold without undue structural constraints. If the IRES is placed too close to the A cistron or immediately adjacent to a 5′ cap (i.e., preceding the A cistron of a monocistronic DNA vector), it is likely to fail or interfere with translation from the upstream ORF. As a general rule, attention to the IRES origin and preferred sequence are required for optimum activity. Methods sections describing comparative IRES expression experiments should always be careful to cite this information.

ACKNOWLEDGMENTS

This work was supported by National Institutes of Health (NIH) grant no. AI-17331 to A.C.P. The authors thank Dr. Aleksey Aminev for the generous gift of plasmid pIRES-luc and Ms. Rachel Groppo for plasmid pF/R-wt.

COMPETING INTERESTS STATEMENT

A.C.P. is the senior author on the original, as well as subsequent papers describing the EMCV IRES technology, which is wholly owned by the Wisconsin Alumni Research Foundation (WARF) and who also administers the U.S. patent covering this technology. A.C.P. receives a fractional royalty. Y.A.B. declares no competing interests.

REFERENCES