Brief communication

A mini-IRES sequence for stringent selection of high producers

JUN YAN[1], HAILIN YANG[2], GUOHUA YUE[2] and WENDA GAO[1,3,*]

[1]*Department 1, Institute of Surgery Research, State Key Laboratory of Trauma, Burns and Combined Injury, Research Institute for Traffic Medicine, Daping Hospital, Third Military Medical University, Chongqing 400042, China*

[2]*Antagen Pharmaceuticals, Inc., 650 Albany Street, Boston, MA 02118, USA*

[3]*Antagen Institute for Biomedical Research, 650 Albany Street, Boston, MA 02118, USA*

**Corresponding author (Fax, 01-617-401-8344; Email, wendagao01@gmail.com)*

Internal Ribosome Entry Site (IRES) sequences have been widely used to link the expression of two independent proteins on the same mRNA transcript. Genes encoding fluorescent proteins or drug-resistance enzymes are usually placed downstream of IRES, serving as expression indicators or selection markers. In biological applications where the upstream gene-of-interest is to be expressed at extremely high levels, it is often desirable to purposely reduce IRES downstream gene expression to economize the cellular resources and/or to generate more stringent selection pressure. Here we describe a miniature IRES mutant sequence (IRESmut3) with dramatically diminished co-translational efficiency to fulfill these purposes.

[Yan J, Yang H, Yue G and Gao W 2013 A mini-IRES sequence for stringent selection of high producers. *J. Biosci.* **38** 245–249] **DOI** 10.1007/s12038-013-9323-x

1. Introduction

The ability to express two or more independent proteins under the same promoter in multicistronic expression vectors is of great importance in molecular and cellular biology. Before the discovery and application of Picornavirus-derived 2A peptides and 2A peptide-like sequences (also called as *cis*-acting-hydrolase elements [CHYSELs]) (Szymczak and Vignali 2005), researchers often use Internal Ribosome Entry Site (IRES) sequences from Encephalomyocarditis virus (EMCV) to link two genes transcribed from a single promoter within the same recombinant vectors (Ngoi *et al.* 2004). 2A-based multicistronic vectors are particularly useful for expressing multiple subunits of a complex protein or multiple independent proteins that cooperate functionally, e.g. subunits of enzymes or immunoglobulin heavy and light chains that should be produced at an equal molar ratio to facilitate their assembly into functional proteins (Fang *et al.* 2005). The obvious advantages of using 2A sequences are their smaller sizes (about 60–70 bp) and 1:1 ratio autocleavage of upstream and downstream proteins (Ryan *et al.* 1991; de Felipe *et al.* 2003). In this sense, using IRES is not ideal, as IRES sequences have a number of limitations, including their large size (about 600 bp) and variability in expression of downstream genes. In many cases, it has been reported that a gene transcribed upstream of an IRES is expressed strongly, whereas a gene placed downstream is expressed at lower levels (Zhou *et al.* 1998; Mizuguchi *et al.* 2000). Nevertheless, such disadvantages of IRES could be of use in some other situations when the genes downstream of IRES, usually fluorescent protein indicators and drug selection markers, need to be expressed at much lower levels, while the genes upstream of IRES, usually therapeutic proteins, are expressed at very high levels in mammalian cell lines.

Traditionally, to obtain highly productive mammalian cell lines, it is necessary to screen large numbers of clones after DNA transfection. This is mainly due to the low selection stringencies, creating many, but low protein-producing clones. Raising the stringency of selection often helps to reduce the number of generated clones and thereby increases the chance to identify a high protein-producing clone (Wurm 2004). It has been reported that mutations in the wild type (WT) IRES sequences could significantly affect its co-

Keywords. Co-translation; EGFP; IRES; mutation

Published online: 25 April 2013

translational efficiency. For instance, insertion of a *Hind*III site, in place of the initiating ATG codon of the wild-type EMCV IRES, causes about 10-fold loss of expression from the second cistron, whereas expression from the first cistron remains unaffected (Martin *et al.* 2006). Therefore, using IRES mutants with much attenuated co-translational activity for downstream selection markers will increase the selection stringency and facilitate the screening for those high protein-producing clones.

2. Materials and methods

Using overlapping PCR, we generated a series of mammalian expression vectors containing WT IRES, mini-IRESmut1, mini-IRESmut2 and mini-IRESmut3, respectively (figure 1). Sequencing results confirmed the desired mutations. These vectors all expressed DsRed-E2 (a generous gift from Dr. Robert Keenan of the University of Chicago) (Strack *et al.* 2008) at upstream, and d1EGFP (a generous gift from Dr. Stephen Elledge of Harvard Medical School) (Yen *et al.* 2008) at downstream, linked by the various IRES elements in between. The human CMV immediate early promoter and the SV40 late poly(A) sequences were used throughout in these vectors (not shown). We then transfected each individual vectors into HEK293 cells using FuGene 6 reagent (Roche). Expression of the fluorescent signals was analysed by flow cytometry (FACS) at 72 h post transfection.

3. Results

We found that WT IRES vector expressed the strongest d1EGFP signal, as expected, whereas the signal strength of d1EGFP in all the IRES mutants was dramatically diminished (figure 2a). In all these cases, DsRed-E2 expression was not affected by various IRES constructs (data not shown).

Because the shortest mini-IRESmut3 directed the least d1EGFP expression, we determined whether it could be used as a faithful marker for high-producing cells. This is indeed the case. In cells transfected with WT IRES or mini-IRESmut3 vectors, the top 0.3% of cells expressing the most potent d1EGFP signals expressed DsRed-E2 at equivalent

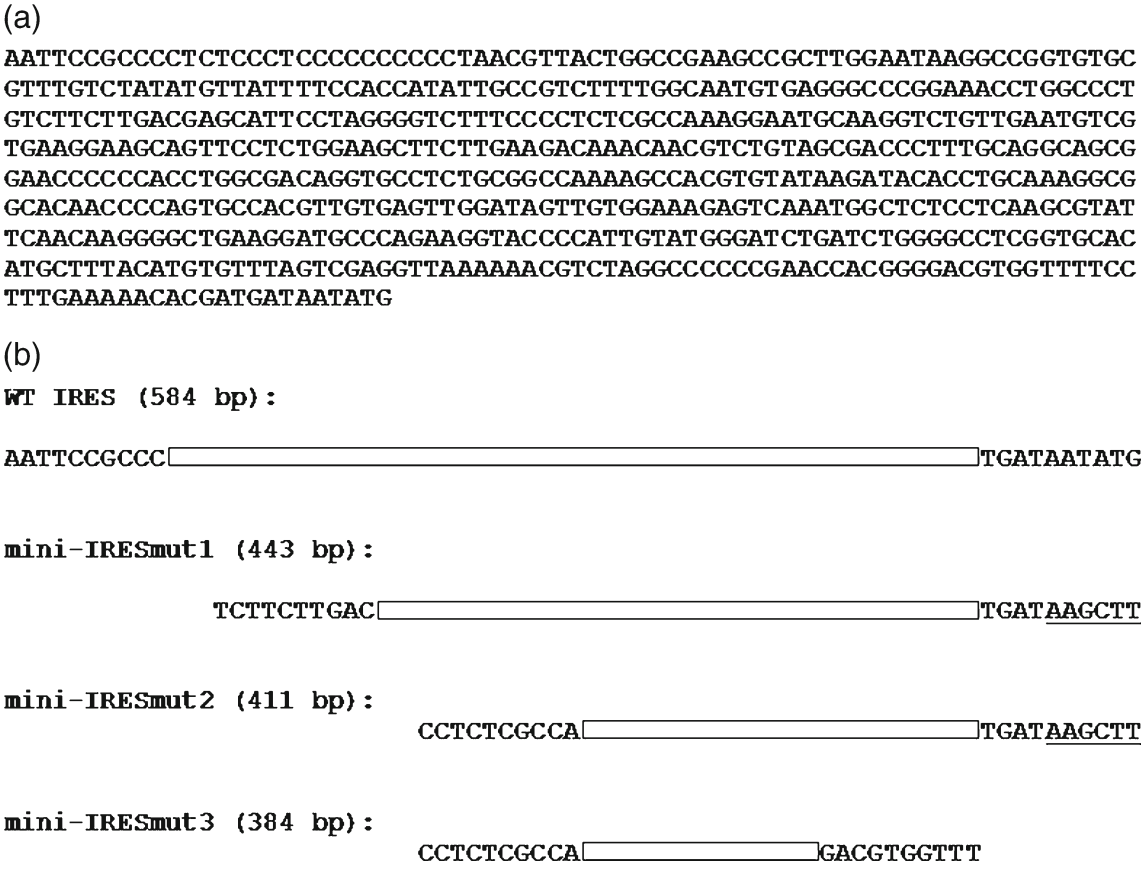

Figure 1. Sequences of wild type and mutant IRES. The full-length wild type IRES sequence was denoted as a reference in (**a**), and the starting and ending nucleotide sequences of IRES mutants were denoted in (**b**). The bars were not in scale. The 3′ mutation to HindIII site was underlined.

high levels in both cases (figure 2b). Yet, d1EGFP signals of the latter were much weaker than those of the former, allowing a much larger window in FL1 channel for the selection of high producers by gating on d1EGFP positive cells.

To study DsRed-E2 expression in d1EGFP-positive cells versus d1EGFP-negative cells, we re-plotted the above FACS data in figure 2c, which clearly shows that d1EGFP-negative cells are mainly DsRed-E2 negative in both constructs (77.1% and 73.8%, respectively). Moreover, when we put a narrow gate on DsRed-E2 positive cells (signal intensity at around 10^3), we found that both WT-IRES and mini-IRESmut3 constructs have similar percentages of DsRed-E2 positive cells in this gate (1.6% vs 1.2%). However, for cells with the similar DsRed-E2 expression, d1EGFP signals vary dramatically in these two constructs. Compared to mini-IRESmut3, WT-IRES gives a much wider spectrum of d1EGFP expression. The much larger amounts of d1EGFP protein translated by WT-IRES, together with the fast turnover rate of d1EGFP, could render high heterogeneity in cells emitting d1EGFP signal. The WT-IRES-d1EGFP construct will thus cause uncertainty for users to select high producers based on d1EGFP signal. In contrast, the mini-IRESmut3 construct gives almost linear correlation between d1EGFP signal and DsRed-E2 expression, especially when DsRed-E2 signal intensity is above 10^2. Therefore, we think mini-IRESmut3 combined with d1EGFP will have better use in selection of high-producing clones.

To prove this on single clone level, we transfected CHO cells with DsRed-mini-IRESmut3-d1EGFP vector and re-

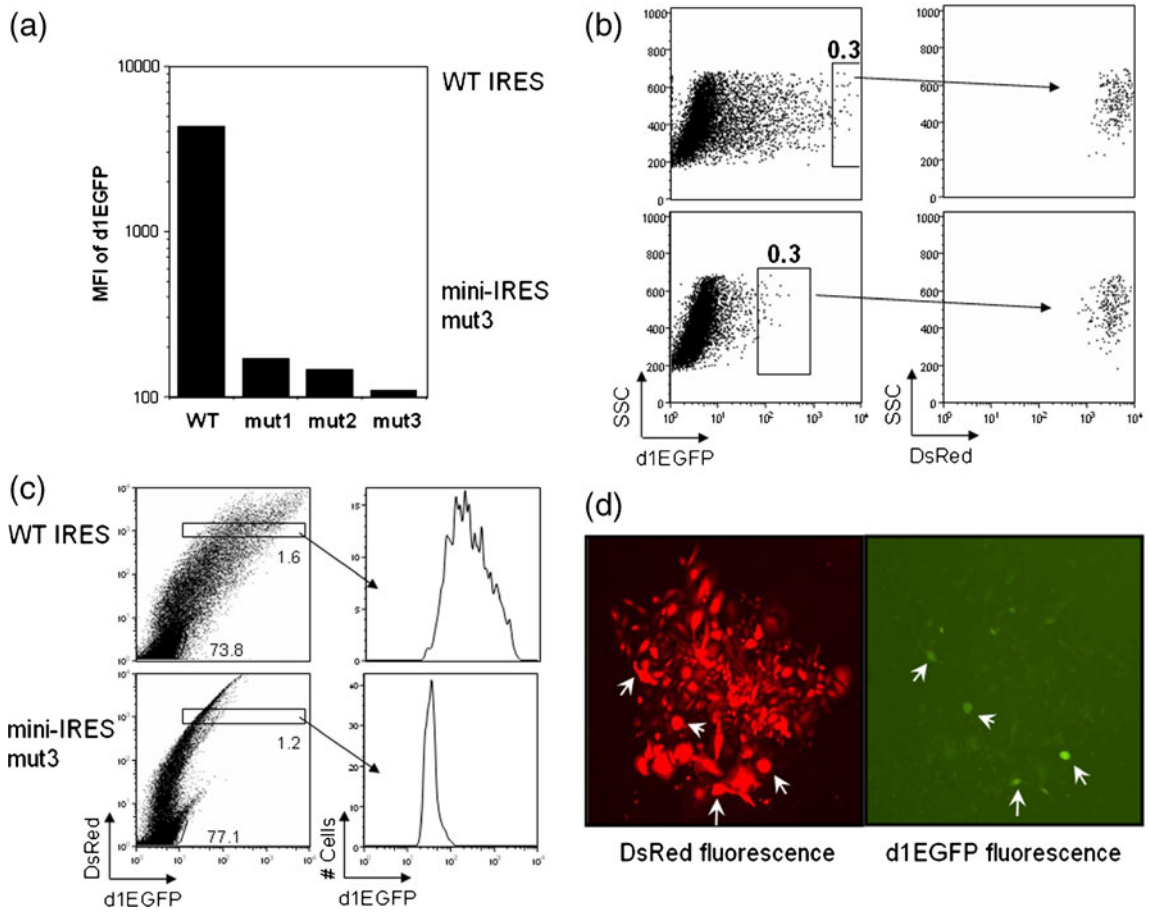

Figure 2. Validation of IRES mutants. (**a**) HEK293 cells were transfected with vectors containing WT IRES, mini-IRESmut1, mini-IRESmut2 and mini-IRESmut3, respectively, and the green fluorescence intensity of d1EGFP+ cell populations was compared after 3 days. (**b**) HEK293 cells were transfected with WT IRES or mini-IRESmut3 vectors, and the top 0.3% of cells expressing the most potent d1EGFP signals were gated to show the equivalent high level DsRed expression. (**c**) Same experiment in b, except that d1EGFP vs. DsRed-E2 signals were shown in dot plots. A narrow gate on cells with similar DsRed-E2 expression showed dramatic difference in their d1EGFP signals. (**d**) CHO cells were transfected with DsRed-mini-IRESmut3-d1EGFP vector, and subcloned in 96-well plate with Zeocin selection. Single clones were photographed with fluorescence microscopy. The exposure time for DsRed is 40 ms, whereas for d1EGFP is 4 s. Arrows point to cells of the same clone under red and green fluorescent channels.

plated all the cells into half of a 96-well plate with Zeocin selection (300 μg/mL, Invitrogen). After 10 days, 23 out of 48 wells contained Zeocin-resistant single clones. Nearly all the clones were DsRed-positive, but only a few were d1EGFP-positive, which were likely to be high-producing ones. This finding is compatible with FACS data obtained with HEK293 cells transfected with mini-IRESmut3 vector, where cells expressing DsRed-E2 at levels above the negative gate but below 10^2 are still d1EGFP-negative (figure 2c, lower panel). These cells are most likely DsRed-E2 low producers. In fact, in mini-IRESmut3 transfectants, only 20% cells highly positive for DsRed-E2 (signal intensity> 5×10^2) emit d1EGFP signals in FACS. Fluorescence microscopy of transfected CHO subclones confirmed that while DsRed-E2 upstream of mini-IRESmut3 was strongly expressed, d1EGFP downstream of mini-IRESmut3 could only be barely detected in a few clones (figure 2d). The much wider d1EGFP signal range is ideal for the selection of even higher producers after DHFR- or GS-mediated gene amplification steps.

4. Discussion

It has been appreciated that certain mRNA secondary structures of the wild-type IRES critically affect its co-translational efficiency. In mini-IRESmut1, 5' deletion of 141 nt from the wild-type IRES sequence alone reduces co-translational activity about 3-5 fold (Jang and Wimmer 1990), while 3' mutation to *Hind*III site alone knocks down 10-fold activity (Martin *et al.* 2006). It seems there is cooperative activity of the 5′ and 3′ secondary structures, as both mutations dramatically reduced IRES activity (figure 2a). Although in mini-IRESmut2, further 5′ deletion of 32 nt extends into the stem-loop E, which is crucial for internal ribosome entry, there was only small increment in reduction of IRES activity. Likewise, in mini-IRESmut3, 3′ deletion of additional 27 nt from mini-IRESmut2 only showed a small reduction of IRES activity (figure 2a), although over 100-fold reduction of IRES activity was noted when 5′ sequences were kept as wild type (Jang and Wimmer 1990). These results indicate that deletion/mutation at both 5′ and 3′ ends effectively destroyed critical secondary structures needed for efficient internal ribosome entry.

The residual IRES activity, as exemplified by mini-IRESmut3 (figure 2b–d), could be potentially useful when coupled with DHFR- or GS-mediated gene amplification strategy. It is conceivable that if wild-type IRES sequence were used, and after gene amplification with hundreds of copies, the green fluorescence signal of EGFP in the high-producing clones would easily exceed the detection window of FACS. Thus, using greatly attenuated mini-IRESmut3 would solve this problem. Also note that in our experiments, we used the unstable d1EGFP instead of the wild-type EGFP. According to Clontech, residues 422–461 of mouse ornithine decarboxylase (MODC) were fused to the C-terminus of wild-type EGFP to make the d2EGFP variant, which has a half-life of 2 h. This region of MODC contains a PEST amino acid sequence that targets the protein for degradation and results in rapid protein turnover. d1EGFP was derived from d2EGFP by a few amino acid substitutions in the MODC region (Glu-428 to Ala, Glu-430 to Ala, and Glu-431 to Ala). d1EGFP has a half-life of approximately 1 h, in sharp contrast to wild-type EGFP, which has a half-life of about 26 h. Accordingly, the green fluorescence intensity of d1EGFP is about 3 orders weaker than that of wild-type EGFP (Yen *et al.* 2008). It is also noteworthy that there might be certain synergy between the effect of IRES mutation and the use of destabilized d1EGFP, as a clear linear relation in detection signals could be found when both are used (Compare lower vs. upper panels of figure 2c). Thus, by using various IRES mutants and EGFP derivatives, high-producing clones can be easily selected by flow cytometry after gene amplification.

Acknowledgements

This work was supported by Juvenile Diabetes Research Foundation (1-2007-551 and 41-2009-760 to WG), and by the National Natural Science Foundation of China (81070624, 81273227), Natural Science Foundation Project of CQ CSTC (2011BB5034) and Foundation for Returned Scholars Sponsored by Department of Human Resources and Social Security of China (CQ-2011-235).

References

de Felipe P, Hughes LE, Ryan MD and Brown JD 2003 Co-translational, intraribosomal cleavage of polypeptides by the foot-and-mouth disease virus 2A peptide. *J. Biol. Chem.* **278** 11441–11448

Fang J, Qian JJ, Yi S, Harding TC, Tu GH, VanRoey M and Jooss K 2005 Stable antibody expression at therapeutic levels using the 2A peptide. *Nat. Biotechnol.* **23** 584–590

Jang SK and Wimmer E 1990 Cap-independent translation of encephalomyocarditis virus RNA: structural elements of the internal ribosomal entry site and involvement of a cellular 57-kD RNA-binding protein. *Genes Dev.* **4** 1560–1572

Martin P, Albagli O, Poggi MC, Boulukos KE and Pognonec P 2006 Development of a new bicistronic retroviral vector with strong IRES activity. *BMC Biotechnol.* **6** 4

Mizuguchi H, Xu Z, Ishii-Watabe A, Uchida E and Hayakawa T 2000 IRES-dependent second gene expression is significantly lower than cap-dependent first gene expression in a bicistronic vector. *Mol. Ther.* **1** 376–382

Ngoi SM, Chien AC and Lee CG 2004 Exploiting internal ribosome entry sites in gene therapy vector design. *Curr. Gene Ther.* **4** 15–31

Ryan MD, King AM and Thomas GP 1991 Cleavage of foot-and-mouth disease virus polyprotein is mediated by residues located within a 19 amino acid sequence. *J. Gen. Virol.* **72** 2727–2732

Strack RL, Strongin DE, Bhattacharyya D, Tao W, Berman A, Broxmeyer HE, Keenan RJ and Glick BS 2008 A noncytotoxic DsRed variant for whole-cell labeling. *Nat. Methods* **5** 955–957

Szymczak AL and Vignali DA 2005 Development of 2A peptide-based strategies in the design of multicistronic vectors. *Expert Opin. Biol. Ther*. **5** 627–638

Wurm FM 2004 Production of recombinant protein therapeutics in cultivated mammalian cells. *Nat. Biotechnol*. **22** 1393–1398

Yen HC, Xu Q, Chou DM, Zhao Z and Elledge SJ 2008 Global protein stability profiling in mammalian cells. *Science* **322** 918–923

Zhou Y, Aran J, Gottesman MM and Pastan I 1998 Co-expression of human adenosine deaminase and multidrug resistance using a bicistronic retroviral vector. *Hum. Gene Ther*. **9** 287–293

MS received 20 December 2012; accepted 26 March 2013

Corresponding editor: SEYED E HASNAIN