Nucleic Acids Research, 2003, Vol. 31, No. 1 427–428
DOI: 10.1093/nar/gkg003

IRESdb: the Internal Ribosome Entry Site database

Sophie Bonnal, Christel Boutonnet, Leonel Prado-Lourenço and Stéphan Vagner*

INSERM U397, Institut Louis Bugnard, CHU Rangueil, 31043 Toulouse, France

Received August 14, 2002; Accepted September 11, 2002

ABSTRACT

Internal Ribosome Entry Sites (IRES) are *cis*-acting RNA sequences able to mediate internal entry of the 40S ribosomal subunit on some eukaryotic and viral messenger RNAs upstream of a translation initiation codon. These sequences are very diverse and are present in a growing list of mRNAs. Novel IRES sequences continue to be added to public databases every year and the list of unknown IRESes is certainly still very large. The IRES database is a comprehensive WWW resource for internal ribosome entry sites and presents currently available general information as well as detailed data for each IRES. It is a searchable, periodically updated collection of IRES RNA sequences. Sequences are presented in FASTA form and hotlinked to NCBI GenBank files. Several subsets of data are classified according to the viral taxon (for viral IRESes), to the gene product function (for cellular IRESes), to the possible cellular regulation or to the *trans*-acting factor that mediates IRES function. This database is accessible at http://ifr31w3.toulouse.inserm.fr/IRESdatabase/.

INTRODUCTION

The first step in eukaryotic translation initiation, which consists in the recruitment of the 40S small ribosomal subunit to the eukaryotic messenger RNA (mRNA), requires several features of RNA polymerase II-transcribed mRNAs: (i) the ^{7}methyl guanosine capped structure at the 5′ end, (ii) the poly(A) tail at the 3′ end and (iii) the internal ribosome entry site located upstream of the translation initiation codon (1–3). These three features can act either together or independently, depending on the mRNA or the cell status, to recruit the small subunit of the ribosome. While the role of the 5′ cap structure and the 3′ poly(A) tail in translation initiation were directly envisaged to be of general importance, the role of an Internal Ribosome Entry Sites (IRES) sequence was initially thought to be restricted to the picornavirus group. However, the process of internal entry of ribosomes appears to be far more extended than previously described. Indeed, IRESes not only exist in picornaviruses but are also present in flaviviruses, plant viruses, retroviruses and even DNA viruses. There is also a continuously growing list of IRES-containing cellular mRNAs. The important feature is that cellular IRESes are found in crucial messenger RNAs encoding key regulatory proteins (transcription factors, growth factors and kinases). More importantly, evidence that internal entry of ribosomes is controlled during the cell cycle or apoptosis demonstrates the importance of this new mechanism of translation initiation and suggests that internal entry of ribosomes could be a key player of cellular proliferation and/or differentiation (4,5). The mechanism of internal entry of ribosomes is not fully elucidated because IRES sequences are very diverse. For instance, an IRES can be a short 9 nt-long sequence (6) and can also be a large highly structured 600 nt-long sequence (7,8). It is generally admitted that the primary RNA sequence determines two main requirements of IRES function that can act either together or independently: (i) the RNA structure, (ii) the binding sites for *trans*-acting factors that will contribute to the recruitment of the ribosome (2,3).

As the understanding of the internal ribosome entry mechanism goes forward, IRESes begin to be used as biotechnological tools, more particularly for gene therapy. A direct application is the synthesis of several proteins of interest from one multicistronic mRNA. One advantage of viral IRESes is their capability to sustain protein synthesis in a broad range of cellular types and tissues, albeit with different efficiencies. On the other hand, the tissular tropism displayed by some cellular IRESes could be useful to target specific organs (9,10). Similarly, the fact that IRES activity can be modulated in response to mitotic stimuli, hypoxia and other stresses would add further control over expression of a given transgene.

DATABASE CONTENT AND ORGANISATION

As of August 2002, the IRES database contains information for the 30 virally-encoded IRESes and the 50 IRESes present on eukaryotic mRNAs. The data are derived from primary sequence databases (NCBI GenBank, EMBL) completed with bibliographic references retrieved from the PubMed database. An IRES can be found either with a general retrieval system or with one of the four classifications available.

1. Virally-encoded IRESes are classified according to the NCBI viral taxonomy.

*To whom correspondence should be addressed. Tel: +33 561 32 31 28; Fax: +33 561 32 21 41; Email: vagner@toulouse.inserm.fr

The authors wish it to be known that, in their opinion, the first three authors should be regarded as joint First Authors

2. Cellular mRNAs-encoded IRESes are classified according to the gene product function.
3. IRESes can be found in a classification of cellular conditions that modulate IRES function (development, stress response, differentiation, cell cycle and cell-type specific).
4. IRESes can be retrieved in a classification of *trans*-acting factors that interact with IRESes (canonical translation initiation proteins, ribosomal factors and RNA binding proteins). Hyperlinks to NCBI accession numbers for these factors are also available.

Information for each IRES (including links to PubMed and NCBI GenBank databases) is presented in independent sections as follows.

1. The abbreviated and full name of the virus or the gene that contain the IRES, followed by a link to a sequence data file.
2. The possible or proven cellular condition in which a modulation of the IRES activity was observed.
3. The proteins that were shown to interact physically with the IRES RNA sequence but were not proven to affect the IRES-dependent translation.
4. The proteins that interact with and modulate translation of the IRES-containing mRNA.

The sequence data file contains the sequence of the IRES in FASTA form as well as hyperlinks to GenBank. Sequence and/or RNA structure information for domains of the IRES that were shown to interact with *trans*-acting factors is also included in this file.

Finally, a 'biotechnology' section presents data concerning the use of IRESes (i) for expression of a selection gene, (ii) for co-expression of proteins and (iii) for a gene therapy purpose. Information on the type of the IRES, the nature of the proteins produced, the type of vector and the cellular model used is provided.

DATABASE AVAILABILITY AND CITATION

Access to the IRES database is possible through the World Wide Web at http://ifr31w3.toulouse.inserm.fr/IRESdb/. Users of the database should cite the present publication as reference. Comments, corrections and new entries are welcome. Submission forms for new information on IRESes are provided in the web site.

ACKNOWLEDGEMENTS

We thank Stefania Millevoi for helpful comments on the manuscript. This work was supported by INSERM, ARC (Association pour la Recherche sur le Cancer) and the French Ministry of Research (ACI 'Jeunes chercheurs').

REFERENCES

1. Gray,N.K. and Wickens,M. (1998) Control of translation initiation in animals. *Annu. Rev. Cell. Dev. Biol.*, **14**, 399–458.
2. Hellen,C.U. and Sarnow,P. (2001) Internal ribosome entry sites in eukaryotic mRNA molecules. *Genes Dev.*, **15**, 1593–1612.
3. Vagner,S., Galy,B. and Pyronnet,S. (2001) Irresistible IRES. Attracting the translation machinery to internal ribosome entry sites. *EMBO Rep.*, **2**, 893–898.
4. Holcik,M., Sonenberg,N. and Korneluk,R.G. (2000) Internal ribosome initiation of translation and the control of cell death. *Trends Genet.*, **16**, 469–473.
5. Sachs,A.B. (2000) Cell cycle-dependent translation initiation: IRES elements prevail. *Cell*, **101**, 243–245.
6. Chappell,S.A., Edelman,G.M. and Mauro,V.P. (2000) A 9-nt segment of a cellular mRNA can function as an internal ribosome entry site (IRES) and when present in linked multiple copies greatly enhances IRES activity. *Proc. Natl Acad. Sci. USA*, **97**, 1536–1541.
7. Kieft,J.S., Zhou,K., Jubin,R., Murray,M.G., Lau,J.Y. and Doudna,J.A. (1999) The hepatitis C virus internal ribosome entry site adopts an ion-dependent tertiary fold. *J. Mol. Biol.*, **292**, 513–529.
8. Kieft,J.S., Zhou,K., Grech,A., Jubin,R. and Doudna,J.A. (2002) Crystal structure of an RNA tertiary domain essential to HCV IRES-mediated translation initiation. *Nature Struct. Biol.*, **9**, 370–374.
9. Creancier,L., Morello,D., Mercier,P. and Prats,A.C. (2000) Fibroblast growth factor 2 internal ribosome entry site (IRES) activity ex vivo and in transgenic mice reveals a stringent tissue-specific regulation. *J. Cell. Biol.*, **150**, 275–281.
10. Creancier,L., Mercier,P., Prats,A.C. and Morello,D. (2001) c-myc Internal ribosome entry site activity is developmentally controlled and subjected to a strong translational repression in adult transgenic mice. *Mol. Cell. Biol.*, **21**, 1833–1840.