Biochimica et Biophysica Acta 1789 (2009) 542–557

Contents lists available at ScienceDirect

Biochimica et Biophysica Acta

journal homepage: www.elsevier.com/locate/bbagrm

Review

# Structural and functional diversity of viral IRESes

Laurent Balvay, Ricardo Soto Rifo, Emiliano P. Ricci, Didier Decimo, Théophile Ohlmann *

*Unité de Virologie Humaine, Ecole Normale Supérieure de Lyon, Lyon F-693643, France*
*INSERM, U758, Lyon F-69364, France*
*IFR 128, Lyon F-693643, France*

ARTICLE INFO

*Article history:*
Received 20 March 2009
Received in revised form 17 July 2009
Accepted 19 July 2009
Available online 24 July 2009

*Keywords:*
Virus
Picornavirus
IRES
eIF4G
Translation initiation
eIF
HIV
Retrovirus
Flavivirus

ABSTRACT

Some 20 years ago, the study of picornaviral RNA translation led to the characterization of an alternative mechanism of initiation by direct ribosome binding to the 5′ UTR. By using a bicistronic vector, it was shown that the 5′ UTR of the poliovirus (PV) or the Encephalomyelitis virus (EMCV) had the ability to bind the 43S preinitiation complex in a 5′ and cap-independent manner. This is rendered possible by an RNA domain called IRES for Internal Ribosome Entry Site which enables efficient translation of an mRNA lacking a 5′ cap structure. IRES elements have now been found in many different viral families where they often confer a selective advantage to allow ribosome recruitment under conditions where cap-dependent protein synthesis is severely repressed. In this review, we compare and contrast the structure and function of IRESes that are found within 4 distinct family of RNA positive stranded viruses which are the (i) Picornaviruses; (ii) Flaviviruses; (iii) Dicistroviruses; and (iv) Lentiviruses.



## 1. Introduction

Viruses are intracellular parasites that rely on the components of the host cell for gene expression and replication. Although viral replication greatly varies from one virus to another, all known eukaryotic viruses have to produce messenger RNAs (mRNA) that can be translated by the cellular ribosomes. The Baltimore classification system of all viruses in six classes, based on the strategies adopted to produce mRNA from their genomes, emphasises this dependency on the cell's translational machinery [1]. Soon after infection, the host cell often tends to limit viral production and replication by shutting-off global translation [2]. This regulatory mechanism generally targets events from the initiation phase because it is the rate limiting step that determines overall protein production for most mRNAs. Many viral genomes have evolved to bypass this general inhibition of translation by developing mechanisms of initiation independent of the classical recognition of an $m^7G$ cap structure at the 5′ end of the mRNA. These mechanisms imply the utilization of internal ribosome entry sites (IRESes) which can promote 5′ end independent initiation.

### 1.1. The cap-dependent mechanism of translation initiation

In eukaryotes, protein synthesis is a complex process by which polypeptides are produced from mRNA templates read by ribosomes in a 5′ to 3′ direction. Translation can be divided in three distinct stages: initiation, elongation and termination. Initiation is the process that allows assembly of a constituted 80S ribosome on an mRNA in which the first translated codon is based paired with the anticodon of the aminoacylated initiator methionyl transfer RNA ($tRNA_i^{Met}$). Elongation is the phase during which the ribosome selects an aminoacylated tRNA and catalyzes the formation of a peptide bond between the polypeptide chain already synthesized and the incoming amino acid. Termination takes place when a stop codon enters the A site of the 80S ribosome, this triggers the release of the neo synthesized protein and ribosome dissociation into both 40S and 60S subunits. In eukaryotes the initiation phase of protein synthesis is generally the most controlled step and determines overall protein production, for recent reviews see [3,4].

The vast majority of eukaryotic mRNAs (except those found in mitochondria) begins with a 5′ terminal 7-methyl-guanosine ($m^7G$) cap structure linked to the second nucleotide by a 5′–5′ phosphodiester bond. The cap is followed by a 5′ untranslated region (5′ UTR) which may vary in length and nucleotide composition but is typically comprised between 50 and 120 nucleotides [5]. The open reading frame encodes the protein and is ended by a 3′ untranslated region (3′ UTR) of variable length followed by a stretch of adenylate residues

* Corresponding author. INSERM, U758, Lyon F-69364, France. Tel.: +33 4 72 72 89 53; fax: +33 4 72 72 81 37.
*E-mail address:* tohlmann@ens-lyon.fr (T. Ohlmann).

1874-9399/$ – see front matter © 2009 Published by Elsevier B.V.
doi:10.1016/j.bbagrm.2009.07.005

known as the poly(A) tail. In the canonical mechanism for translation initiation or "5′ scanning model" initially described by Marilyn Kozak [6,7], the 40S ribosomal small subunit binds to the 5′ terminal cap moiety and linearly scans the 5′ UTR in a 5′ to 3′ direction until it locates and recognizes the initiation codon. This process requires at least 12 different eukaryotic initiation factors (eIFs). The first event is the recruitment of the eukaryotic translation initiation factor 2 (eIF2)–GTP–$tRNA_i^{Met}$ ternary complex to the 40S ribosome to form a 43S preinitiation complex; this is also catalyzed by eIF5 which interacts with both eIF2 and eIF3 [8]. This complex formation is stabilized by eIF3, the largest initiation factor composed of at least 11 subunits [9–11], eIF1 and eIF1A [12–15]. The 43S preinitiation complex can be recruited to the 5′ capped end of the mRNA by its interaction with the heterotrimeric eIF4F factor which is constituted by eIF4E, the cap binding protein [16,17] associated with eIF4G [18], and eIF4A an ATPase RNA helicase which unwinds local RNA secondary structures [19]. This process is stimulated by a physical and functional interaction between the poly (A) binding protein (Pab1P) at the 3′ end and the initiation factor eIF4G located at the 5′ end [20–23] resulting in the pseudo circularization of the mRNA. Translation is also stimulated by the homodimer eIF4B which enhances the RNA helicase enzyme activity of eIF4A [24,25]. After initial binding to the 5′ end of the mRNA and scanning [6,26], the ribosome will then locate an AUG start codon surrounded by a good nucleotide consensus which is defined by a purine at the −3 position and a G at the +1 position [27]. Computer based analysis of different transcripts has shown that the first AUG codon encountered is usually selected (in a percentage varying between 90% and 60% depending on the studies and the surrounding context [28,29]). Accuracy of initiation codon selection is promoted by both eIF1 and eIF1A, which trigger the formation of a "closed" conformation when a proper codon–anticodon base pairing between the mRNA and the $tRNA_i^{Met}$ is found [30]. At this stage, eIF5 induces hydrolysis of eIF2-bound GTP to release some eIF2-GDP. Then the dissociation of initiation factors is promoted by the GTPase factor eIF5B and allows the association of the 60S subunit to form the 80S ribosome with the $tRNA_i^{Met}$ in the P site marking the end of the initiation phase.

### 1.2. Internal ribosomes entry sites

The study of poliovirus (PV) replication in the early 70's revealed that the viral mRNA was structurally distinct from cellular mRNAs as it does not contain a 5′ terminal cap structure but a VPg protein covalently linked to the 5′ end [31]. In addition, the 5′ UTR of PV is unusually long, rich in secondary RNA structures [32,33] and contains multiple upstream AUG codons. Finally, the finding that PV entry into the host cell is accompanied by a severe shut-off of host cell translation [34] with no impact on viral translation, suggested the use of a novel mechanism for translation initiation. Thus, it was hypothesized that the ribosome could bind to the viral mRNA in an internal position rather than at the mRNA 5′ end. By using bicistronic mRNA assays it was shown that the highly structured 5′ UTR of poliovirus was able to promote internal entry of the 43S ribosomal subunit on its mRNA, characterizing the first viral IRES [35]. This discovery was almost immediately extended to EMCV [36], and to other classes of RNA viruses such as the Hepatitis C Virus (HCV) [37], and to some DNA viruses [38,39].

Internal entry of ribosomes has also been evidenced in some eukaryotic cellular mRNAs that have the ability to remain associated with polysomes in poliovirus-infected cells when cap-dependent translation of host cell mRNAs was severely repressed ([40,41], review [42]).

Just about 20 years later, internal initiation has now emerged as an alternative way to bind a 43S ribosome at the correct initiation codon in a manner independent of the 5′cap structure of the viral or cellular mRNA considered. However, the exact order of events and the set of proteins required for efficient ribosomal binding is different between different IRESes.

## 2. Experimental demonstration of IRES activity

A peculiar characteristic of viral IRESes is the absence of a consensus or motif in their primary sequence, rendering a computer-based search very difficult [43,44]. Nevertheless, viral IRES elements often share several common structural characteristics such as long and structured 5′ UTRs, several upstream AUG triplets located before the authentic start codon, and, often but not always, the absence of a cap structure at the 5′ end of the RNA. These features are not compatible with an efficient cap-dependent translation and reflect the use of an alternative mechanism for translation initiation. However, only a thorough, functional assessment of IRES activity can validate a candidate sequence. Historically internal ribosome binding was first tested by measuring protein production from the second open reading frame (ORF) of mRNAs containing two ORFs called bicistronic mRNAs [35,36] or by looking at the translation of circular mRNAs [45].

### 2.1. Bicistronic constructs

This test needs the construction of a synthetic mRNA which contains two open reading frames (ORFs) separated by the candidate IRES, so that this single bicistronic RNA molecule codes for two reporter genes (see Fig. 1). The first cistron is translated in a 5′ cap-dependent manner whereas efficient production of the second one can only take place if ribosomes are recruited internally by the candidate IRES sequence located in the intercistronic spacer. The ratio of protein production measured from translation of both cistrons gives an indication of the relative strength of the IRES.

This bicistronic assay has rapidly been considered as the gold standard to define internal initiation and it is the most widely used method for testing putative IRES sequences. It can be utilised both *in vitro* in the reticulocyte lysate or cell extracts but also *ex vivo* in cultured cells. However, its utilization *ex vivo* should be thoroughly controlled as it has caused several drawbacks and misinterpretations. In fact, this test has been criticized [46,47] on the basis that even low amount of alternative monocistronic mRNA transcripts can result in a misinterpretation of an IRES activity. These constructs may be generated by different ways: i.e. the presence of a cryptic promoter in the sequence of interest, or by a splicing event that would skip completely the first ORF. These mRNAs, in very low concentration, could be poorly detectable but highly translatable and could distort the interpretation. This has been demonstrated in the case of the *Renilla* and *Firefly luciferase* reporter genes (widely used in bicistronic constructs) which have been shown to contain a splice-donor sequence (Renilla) or a cryptic promoter (Firefly) respectively [48]. Thus when testing IRES activity by transfection of bicistronic DNA test constructs in cells, it is essential to check for the presence of alternate transcripts by stringent RNA tests, such as that proposed by Van Eden et al. [49], which is based on RNAi combined to RT-PCR detection of aberrant RNAs. Nevertheless, one easy way to circumvent the presence of cryptic promoters or splicing sites, is to use direct transfection of *in vitro* transcribed bicistronic RNA [50–52]. This method rules out the intervention of any possible nuclear processing events but may also prevent proper association with nuclear ITAFs (IRES trans-acting factors) involved in RNP complex formation and full IRES activity, and it should be noted that the influence of a nuclear event for efficient translation of cellular IRESes was pointed out by Willis et al. [53].

### 2.2. Monocistronic RNAs

Virtually all eukaryotic genes, even those containing IRESes, are transcribed as monocistronic messenger RNAs. Thus, one should bear in mind that candidate IRESes must also be assessed as monocistronic entities (i.e. with no 5′ flanking sequence and a free 5′ end) as it is the relevant physiologically context. Although not useful when demon-