New insights into internal ribosome entry site elements relevant for viral gene expression

Encarnación Martínez-Salas, Almudena Pacheco, Paula Serrano and Noemi Fernandez

Centro de Biología Molecular Severo Ochoa, Consejo Superior de Investigaciones Científicas - Universidad Autónoma de Madrid, Cantoblanco 28049 Madrid, Spain

A distinctive feature of positive-strand RNA viruses is the presence of high-order structural elements at the untranslated regions (UTR) of the genome that are essential for viral RNA replication. The RNA of all members of the family Picornaviridae initiate translation internally, via an internal ribosome entry site (IRES) element present in the 5' UTR. IRES elements consist of cis-acting RNA structures that usually require specific RNA-binding proteins for translational machinery recruitment. This specialized mechanism of translation initiation is shared with other viral RNAs, e.g. from hepatitis C virus and pestivirus, and represents an alternative to the cap-dependent mechanism. In cells infected with many picornaviruses, proteolysis or changes in phosphorylation of key host factors induces shut off of cellular protein synthesis. This event occurs simultaneously with the synthesis of viral gene products since IRES activity is resistant to the modifications of the host factors. Viral gene expression and RNA replication in positive-strand viruses is further stimulated by viral RNA circularization, involving direct RNA–RNA contacts between the 5' and 3' ends as well as RNA-binding protein bridges. In this review, we discuss novel insights into the mechanisms that control picornavirus gene expression and compare them to those operating in other positive-strand RNA viruses.

Initiation of protein synthesis in picornavirus RNAs

Picornavirus infectious particles consist of a single-stranded RNA molecule of positive polarity surrounded by a non-enveloped icosahedral capsid. The genome is a long RNA molecule (7000–8500 nt long, depending on the genera) in which 5' and 3' untranslated regions (UTRs) flank a single open reading frame. The 5' end of the viral RNA is covalently linked to the viral protein VPg, and the 3' end is polyadenylated (Fig. 1). The genome encodes a single polyprotein that is rapidly processed by viral encoded proteases yielding the mature viral proteins needed to accomplish the different steps of the viral replication cycle.

Picornavirus protein synthesis, which is the first step of viral gene expression, is controlled by the internal ribosome entry site (IRES) element using a cap-independent mechanism (Pelletier & Sonenberg, 1988; Jang et al., 1988). The viral RNA of different genera belonging to the family Picornaviridae initiate translation via an IRES element, an internal region of the 5' UTR that also drives protein synthesis of a second cistron when placed outside of its genetic context (Hellen & Sarnow, 2001; Martinez-Salas, 1999). Typically, picornavirus IRES elements span about 450 nt of the 5' UTR viral RNA. According to their secondary structure, IRES elements were initially classified into two groups (Jackson & Kaminski, 1995). Type I includes the IRES elements of enteroviruses [such as poliovirus (PV) or coxsackie B virus (CBV)] and human rhinoviruses (HRV) (Fig. 1, upper panel). IRES elements from cardioviruses [such as encephalomyocarditis virus (EMCV) and Theiler's murine encephalitis virus (TMEV)] and aphthoviruses (foot-and-mouth disease virus, FMDV) belong to type II (Fig. 1, lower panel), together with erbo- and parechovirus IRES elements (Hinton & Crabb, 2001; Nateri et al., 2000). Subsequently, a new class of IRES elements was found in teschovirus RNA (Pisarev et al., 2004; Chard et al., 2006a) that shares similarities with the IRES element of hepatitis C virus (HCV), a member of the family Flaviviridae. In agreement with this finding, comparison of 5' UTR sequences identified several picornaviruses with similarities in their RNA structure to HCV IRES (see Fig. 2), including the pseudoknot (Pk), domains IIId and IIle. The differences in IRES structural organization found in the RNA genomes of distinct picornavirus genera might have arisen by horizontal transmission, presumably generated by recombination between ancestors (Hellen & de Breyne, 2007).

IRES-dependent initiation of translation represents an alternative to the cap-dependent initiation mechanism operating on the majority of cellular mRNAs (Merrick, 2008).
Most eukaryotic mRNAs have a cap (m\(^7\)Gppp) structure at the 5’ end that is recognized by the initiation factor eIF4F. This key initiation factor comprises the cap-binding protein, eIF4E, the RNA helicase, eIF4A, and the scaffold protein, eIF4G (Sonenberg & Dever, 2003). To initiate translation, the 40S ribosomal subunit interacts with the eIF4F complex bound at the 5’ cap and scans in the 5’→3’ direction until an appropriate AUG triplet is found (Kozak, 1989). During the scanning process, recognition of the correct initiation codon is inhibited by stable RNA structures, by proteins bound to specific RNA sequences, as well as by upstream AUG codons (Hernandez-Sanchez et al., 2003; Svitkin et al., 2005). In contrast, initiation of translation mediated by IRES elements involves the direct recruitment of the translational machinery to an internal position in the mRNA. Moreover, the triplet recognized as the initiator codon in IRES-dependent protein synthesis does not need to conform to a conserved universal sequence context, in marked difference from the start AUG in mRNAs translated in a cap-dependent manner.

A characteristic of all picornavirus RNAs is the presence of a large number of AUG triplets located all along the 5’ UTR sequence that are not recognized as initiator codons (Hellen & Sarnow, 2001; Martinez-Salas et al., 2001). However, since the parameters affecting initiation codon selection differ between the RNA genome of members of the family Picornaviridae, the mechanism operating to discriminate the authentic initiation codon from other AUG triplets is still under discussion. There is no evidence to suggest conservation of primary sequence or secondary structure around the initiation triplet of all picornavirus RNAs. In PV RNA, a silent AUG (AUG586) is located within the ribosome-binding site but initiation only occurs at a downstream AUG triplet separated by 154 nt (AUG743) (Pilipenko et al., 1992). Different studies have proposed that the second PV AUG is reached after scanning (Kuge et al., 1989) or shunting (Hellen et al., 1994). The situation is different in EMCV where initiation occurs at AUG11, overlapping the ribosome-binding site (Kaminski et al., 1994). In this RNA, three different AUG codons (numbers 10, 11 and 12) are closely located, but AUG10 (8 nt upstream) and AUG12 (12 nt downstream) are not recognized as initiation triplets by the translation machinery. In other cardioviruses, such as TMEV, an out-of-phase initiation triplet (AUG*) within the initiation zone affects the expression of viral capsid proteins (Yamasaki et al., 1999). Translation initiation at the polyprotein AUG impairs initiation at the AUG* triplet,
suggesting that ribosomes land at the authentic initiation codon before scanning downstream to reach AUG*. This model was suggested because of the observation that infectious revertants, recovered from cDNAs with mutations at the authentic initiation triplet or a stop codon inserted upstream of the AUG*, showed a second site mutation that restored the AUG* frame into the polyprotein reading frame.

A unique feature of the aphthovirus viral RNA is the presence of two initiation codons, conserved in all FMDV isolates (Carrillo et al., 2005), which are used to initiate translation of the leader (L) protein in two forms, Lab and Lb (Fig. 1, lower panel). Thus, two in-frame AUG triplets, 84 nt apart, are used as translation initiator codons (Belsham, 1992), with the peculiarity that the second one is used more frequently (80–90%) in infected cells as well as in cells transfected with chimeric RNAs (Lopez de Quinto & Martinez-Salas, 1999). The efficiency of recognition of the second AUG is independent of AUG1 and modification of the AUG1 context to optimize recognition does not inhibit initiation at the second AUG. Neither is initiation at AUG2 abrogated by the presence of a stable RNA structure in the spacer sequence or by antisense molecules bound to AUG1 (Lopez de Quinto & Martinez-Salas, 1999). However, while initiation at each of these AUGs is IRES-dependent, the subset of factors required to reconstitute 48S initiation complexes is different. In addition to eIF4G, eIF4A, eIF3 and eIF2, recognition of AUG1 is dependent on eIF1A, but initiation at AUG2 depends on eIF1 (Andreev et al., 2007). Stabilization of a stem–loop between the initiation triplets induces a small decrease in assembly of the 48S complexes at AUG2, accompanied by an increase in the AUG1/AUG2 initiation ratio, as well as a moderate reduction of protein synthesis initiated at AUG2 in transfected cells. This differential requirement of factors suggests that different mechanisms operate to select the translation initiation codon in FMDV RNA.

Unusual features also characterize the initiation codon region of other viral IRES. The functional AUG codon is internal to the HCV IRES sequence (Reynolds et al., 1995), and forms part of a tertiary motif that includes a stem–loop of domains IV and IIIf (Wang et al., 1995). One of the most striking differences in the initiation codon region with the mRNAs translated using a cap-dependent manner is present in the dicistrovirus intergenic region (IGR). In this RNA, triplets GCC, GCU, GCA, CAA or CUG (encoding alanine or glutamine) are used to initiate translation internally in the context of a high-order RNA structure (Sasaki & Nakashima, 1999, reviewed by Jan, 2006) (see Fig. 3).

### Requirement of host factors for internal translation initiation in RNA viruses

**Picornavirus IRES-driven protein synthesis**

Early during the study of picornavirus IRES-driven translation it was found that translation initiation was dependent on host proteins (reviewed by Belsham & Sonenberg, 2000; Pestova et al., 2001). The initiation factors eIF4G, eIF4A, eIF2 and eIF3 are required for 48S complex formation in a reconstituted 40S ribosome assay with representative members of type II IRES elements (Table 1) EMCV, TMEV and FMDV (Pestova et al., 1996; Kolupaeva et al., 1998; Pilipenko et al., 2000, 2001). Accordingly, translation initiation promoted by the EMCV IRES, and to lesser extent PV IRES elements, is sensitive to a dominant-negative mutant of eIF4A (Pause et al., 1994; Table 1.

<table>
<thead>
<tr>
<th>IRES type (virus)</th>
<th>eIF4E</th>
<th>eIF4G</th>
<th>eIF4A</th>
<th>eIF3</th>
<th>eIF2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Picornavirus</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Type I (PV, HRV, CVB)</td>
<td>−</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Type II (EMCV, TMEV, FMDV)</td>
<td>−</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>HCV-like (teschovirus)</td>
<td>−</td>
<td>−</td>
<td>−</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Flavivirus</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HCV</td>
<td>−</td>
<td>−</td>
<td>−</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>CSFV, BVDV (pestivirus)</td>
<td>−</td>
<td>−</td>
<td>−</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Dicistrovirus</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5' UTR (RhPV, PSIV)</td>
<td>−</td>
<td>−</td>
<td>−</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>IGR (PSIV, CrPV)</td>
<td>−</td>
<td>−</td>
<td>−</td>
<td>−</td>
<td>−</td>
</tr>
</tbody>
</table>

**Fig. 3.** Dicistrovirus genome schematic representation. The 5' UTR IRES and the IGR IRES precede two open reading frames encoding the non-structural and structural proteins, respectively. A non-AUG codon is used to start translation dependent on the IGR.
proteases (Martinez-Salas, 2000). This study demonstrated that eIF4G that is produced from its cleavage by the 2A protease of entero- and rhinoviruses, or the aphthovirus Lb protease (Glaser et al., 2001; Gradi et al., 2004; Lampehr et al., 1993; Ziegler et al., 1995), contains the binding sites for eIF3 and eIF4A, and it is sufficient to promote 48S complex initiation with the type II picornavirus IRES elements (Lomakin et al., 2000). Moreover, the Lb protease processes C-terminal products of eIF4GI and eIF4GII bind to the FMDV IRES with the same efficiency as the unprocessed proteins (Lopez de Quinto et al., 2001). In agreement with this, picornavirus IRES-dependent protein synthesis is resistant to inhibitory conditions for cap-dependent translation initiation, which occur following eIF4G cleavage by the aphthovirus Lb or the enterovirus 2A proteases (Martinez-Salas et al., 2001), or by 4E-BP1 dephosphorylation during cardiovirus infection (Gingras et al., 1996; Svitkin et al., 1998). In addition, a novel role of the EMCV 2A protein, which does not have proteolytic activity, can contribute to explain host mRNA translation inhibition in cardiovirus-infected cells. EMCV 2A, which has been found to be associated with 40S ribosomal subunits (Groppo & Palmenberg, 2007), might confer to these modified 40S subunits a preferential capacity to associate with the viral RNA, preventing host mRNA translation.

The functional contribution of eIF4G to IRES-dependent translation was further supported by the strong correlation found between eIF4G–IRES interaction and IRES activity in transfected cells using an extensive set of FMDV IRES mutants bearing nucleotide substitution in a conserved structural motif of domain 4 (Lopez de Quinto & Martinez-Salas, 2000). This study demonstrated that eIF4G binding to the IRES element is an essential step in the recruitment of the translational machinery in vivo. In contrast to IRES mutants defective in eIF4G binding, those impaired in eIF4B interaction barely reduced IRES activity (Lopez de Quinto et al., 2001), consistent with the observation that eIF4B stimulated 48S complex formation on the EMCV IRES by only twofold (Pestova et al., 1996).

Early studies of the PV IRES, which was inactive in rabbit reticulocyte lysate (RRL) translation system, showed that its activity was enhanced upon supplementation with HeLa cell extracts (Dorner et al., 1984). This observation allowed the discovery of IRES trans-acting factors (ITAFs). To list a few, the polypyrimidine tract-binding protein (PTB), the La human autoantigen, the poly(rC)-binding protein (PCBP) and the upstream-of-N-ras protein (UNR) act as auxiliary factors by interacting with the picornavirus IRES elements (Hunt et al., 1999; Luz & Beck, 1991; Meerovitch et al., 1993; Walter et al., 1999; reviewed by Belsham & Sonenberg, 2000). Analysis of the different IRES auxiliary factors has proposed that their role is to facilitate the structural organization of IRES elements by acting as RNA chaperones. Nevertheless, their mode of action is still unknown. IRES elements that do not share overall structural similarity interact with common proteins, probably due to the presence of short primary sequence motifs that provide the binding sites for such factors. In this regard, it has to be taken into consideration that RNA–protein binding per se does not imply functional requirement.

In turn, it also seems that closely related IRES elements that share secondary structure behave differently in terms of functional RNA–protein association. One representative example is PTB that together with the proliferation-associated factor ITAF45, is specifically required for 48S complex formation in vitro with FMDV IRES but not with cardiovascular IRES (Pilipenko et al., 2000, 2001), and it is also consistent with the differential effect of ITAF45 depletion on IRES activity (Monie et al., 2007). The presence of PTB and ITAF45 exerts an additive effect on the 48S complex formation at both FMDV start codons, although reconstitution complexes on both AUGs were also detectable in the absence of either of these proteins (Andreev et al., 2007).

Novel picornavirus IRES-interacting factors have been recently identified by RNA chromatography methods and mass spectrometry analysis (Bedard et al., 2007; Merrill et al., 2006; Kim et al., 2004; Choi et al., 2004). Some of these proteins have been shown to act as activators or repressors of IRES activity. One example of the activators recently reported is the splicing factor SRp20, which functions in PV IRES-mediated translation via its interaction with PCBP2. Targeting SRp20 in HeLa cells with short interfering RNAs resulted in a decrease in PV RNA translation (Bedard et al., 2007). Conversely, the double-stranded RNA-binding protein 76 (DRBP76) is an example of a repressor; this factor binds to the HRV2 IRES in neuronal cells and inhibits translation of the PV–HRV chimeras, and thereby viral propagation (Merrill & Gromeier, 2006). It is likely that the tissue-specific expression of some of these proteins plays a specific role during viral pathogenesis, as shown before for TMEV and HRV (Merrill et al., 2006; Pilipenko et al., 2001), although in the case of PV infection, tissue tropism does not appear to be mediated by the IRES element (Kauder & Racaniello, 2004). The characterization of novel trans-acting factors may also help to explain the functional competition observed between different IRES elements (Reigadas et al., 2005). In any case, deciphering the role played by the IRES auxiliary factors will help to understand why there are so many different strategies, at least apparently, employed by diverse RNAs during the process of internal initiation.
**Hepatitis C and dicistrovirus IRES-driven protein synthesis**

Protein synthesis initiation on HCV and pestivirus RNAs, which belong to the family *Flaviviridae*, is also driven by IRES elements (Fletcher & Jackson, 2002; Tsukiyama-Kohara et al., 1992). The IRES region of these viral RNAs encompasses about 300 nt, and differs from the picornavirus IRES elements in RNA structural organization (Fig. 2) and protein requirement (Table 1) (Fraser & Doudna, 2007). To assemble 48S initiation complexes in vitro, the IRES of HCV as well as classical swine fever virus (CSFV) requires the presence of the 40S ribosomal subunit, eIF3 and the eIF2/GTP/Met-tRNA(i) ternary complex (Ji et al., 2004; Pestova et al., 1998).

Binding of the HCV IRES to the 40S ribosomal subunit is the first step of translation initiation on the viral RNA. Interaction with the 40S subunit requires hairpins IIId and IIIe of the IRES (Fig. 2) (Kieft et al., 2001; Lukavsky et al., 2000), but formation of both 48S and 80S complexes is dependent on the HCV IRES structure (Otto & Puglisi, 2004). The role of domain II in HCV IRES activity has remained elusive for many years; site-directed mutations affecting this IRES region impaired translation initiation (Otto & Puglisi, 2004), but its presence was not needed to assemble binary complexes (Spahn et al., 2001). Recently, it has been found that domain II of the HCV IRES plays a functional role in later translation steps, promoting eIF5-induced GTP hydrolysis during 80S ribosome assembly and eIF2/GDP release from the initiation complex. This function depends both on the bent conformation of domain II and on an apical hairpin loop and loop E (Locker et al., 2007).

Cross-linking studies of the HCV IRES binary complex with the human 40S ribosomal subunit identified ribosomal proteins S5, S16, p40 and S3a bound to hairpin IIIe (Laletina et al., 2006). Many ribosomal proteins were identified by mass spectrometry in the HCV IRES complexes assembled with native 40S subunits, in addition to RACK1, eIF3 proteins and nucleolin (Yu et al., 2005). Similar methodology applied to the binary HCV IRES–40S complex revealed that ribosomal proteins S25 and S29 contain different modifications than those present in the native 40S subunit. In addition to PTB, La autoantigen and UNR, which also interact with the picornavirus IRES (Belsham & Sonenberg, 2000), other HCV IRES-binding proteins include UNR-interacting protein, plasminogen activator inhibitor-1 protein, Ewing sarcoma breakpoint 1 region protein and several cytoskeleton proteins (Lu et al., 2004). The functional relevance of these factors in IRES activity remains to be determined.

A fully divergent class of IRES elements are present in the genome of the family *Dicistroviridae* (Wilson et al., 2000b). This family of RNA viruses possesses a genome that is naturally dicistronic (Fig. 3); it contains two IRES elements, one in the 5′ UTR responsible for translation of the first cistron and another, the IGR element, that drives protein synthesis of the second cistron from a non-AUG triplet (Jan, 2006). Thus, translation of each open reading frame depends on an active IRES element with functional and structural features that differ from any of the IRES elements described above. Indeed, the 5′ UTR of Rhopalsosiphum padi virus (RhPV) has been reported to contain an IRES that is active in plant and mammalian cell-free translation systems (Groppelli et al., 2007; Woolaway et al., 2001), and can assemble 48S initiation complexes in vitro with eIF2, eIF3 and eIF1 (Terenin et al., 2005) (Table 1). Two separate unstructured regions within the 5′ UTR seem to be critical for IRES function, but sequences of the 5′ UTR can be deleted without affecting initiation-complex formation. The RNA genome of another dicistrovirus, *Plautia stali* intestine virus (PSIV) also contains a 5′ IRES that was active in an insect cell lysate, but in contrast to RhPV, it did not function in RRL or wheatgerm (Shibuya & Nakashima, 2006). RNA probing of the 5′ IRES element of PSIV indicates that its structure is distinct from that of RhPV, suggesting differences in the mechanism of translation initiation mediated by the 5′ IRES elements of dicistroviruses.

Among the IRES elements that have fewer requirements for activity are the IGR elements present in the intercistronic space of the dicistrovirus genomes. In marked difference to any other IRES described to date, the IGRs of dicistroviruses assemble initiation complexes in the absence of any eIFs (Table 1) (Spahn et al., 2004; Wilson et al., 2000a). This RNA sequence establishes direct contacts with ribosomal proteins (Schuler et al., 2006). Mapping of the PSIV IGR interactions with ribosomal proteins revealed that the strongest signal was S25 (Nishiyama et al., 2007), a protein located near S5 that has been shown to interact with this IGR (Pfingsten et al., 2006).

**Relevance of RNA structure for IRES function**

Viral IRES elements are distributed into groups that are well differentiated in terms of RNA structure and trans-acting factor requirements (Balvay et al., 2007; Fraser & Doudna, 2007). IRES elements are also employed for the translation of some cellular mRNAs during conditions that inhibit cap-dependent translation initiation (Baird et al., 2006; Stoneley & Willis, 2004). No apparent conservation of significant length of primary sequence is detected between viral and cellular IRES elements, with the exception of a polypyrimidine tract, which is commonly found in both viral and cellular IRES elements (Bushell et al., 2006; Honda et al., 1996; Kolupaeva et al., 1996; Mitchell et al., 2003; Witherell & Wimmer, 1994). Additionally, cellular IRES elements do not share overall structural similarity (Baird et al., 2007), indicating that short motifs may control the interaction with trans-acting factors needed for IRES activity. In contrast, viral IRES are organized into high-order structures that differ between distant families, but it is common to find conserved domains between the RNA structure of related viruses.
Functional and structural studies of IRES elements have shown a close relationship between RNA folding and IRES activity (Martinez-Salas & Fernandez-Miragall, 2004). This tight relationship suggests that the RNA structure of IRES elements regulates translation efficiency. On the other hand, structural studies performed on the HCV IRES and the IGR of members of the family Dicistroviridae have shown the capacity of these IRES elements to be accommodated in the interface of the ribosomal subunits (Spahn et al., 2001; Boehringer et al., 2005; Schuler et al., 2006; Pfingsten et al., 2006). Although the IGR and the HCV IRES elements exhibit different structural organization (Kief et al., 2002; Rijnbrand et al., 2004; Jan & Sarnow, 2002; Nishiyama et al., 2003) and their binding sites in the ribosomal subunit are different, similar conformational changes are induced in the 40S ribosomal subunit (Spahn et al., 2004). This finding opens the possibility that IRES elements could share the property of having a universal structural IRES motif (USIM) that could mediate its direct interaction with the 40S subunit. However, to date this putative universal RNA motif has remained elusive.

RNA structural elements in IRES elements

Regarding the RNA organization of IRES elements belonging to viruses addressed in this review, they are grouped in clearly different categories. The first category includes picornavirus IRES elements that occupy an internal position in the genomic RNA without interference of the upstream structural elements. Computer prediction studies of picornavirus RNAs in conjunction with mutational analysis evidenced the presence of several stem–loops organized in stable structural domains all along the 5′ UTR (Pilipenko et al., 1989; reviewed by Hellen & Sarnow, 2001). Functional analysis of the 5′ UTR sequences in enteroviruses (Fig. 1, top panel) revealed that stem–loops II to VI are essential for IRES function (Nicholson et al., 1991; Pelletier et al., 1988; Pilipenko et al., 1992).

In comparison to entero- and rhinovirus RNA, the 5′ UTR of cardio- and aphthovirus is longer and contains more structural elements. In FMDV RNA, the S region, a poly(C) tract of about 200 residues, two to four Pk, the cis-acting replication element (cre) and the IRES structural elements are arranged consecutively (Fig. 1, bottom panel). It is well established that in type II IRES the distal stem–loops, termed 2 and 4-5 (or H and J-K-L), are involved in interactions with host factors (Clark et al., 2003; Kolupaeva et al., 1998; Lopez de Quinto & Martinez-Salas, 2000; Pilipenko et al., 2000). However, the role performed by the central domain (termed 3 or I) that mediates long-range RNA–RNA interactions (Ramos & Martinez-Salas, 1999) is still unknown. Its apical region contains a conserved GNRA motif (N stands for any nucleotide and R for purine) that is also present in stem–loop IV of PV IRES elements (Jackson & Kaminski, 1995); this motif does not tolerate nucleotide substitutions, deletions or insertions (Lopez de Quinto & Martinez-Salas, 1997; Robertson et al., 1999), unlike a different GNRA motif present in a lateral bulge of stem–loop V of PV IRES (Malnou et al., 2002). On the other hand, substitutions in the conserved C-rich loop of FMDV do not impair IRES activity (Martinez-Salas et al., 2002).

The aphthovirus IRES GNRA motif adopts a tetraloop conformation (Fernandez-Miragall & Martinez-Salas, 2003) that is responsible for the organization of the adjacent stem–loops. Subsequently, nuclear magnetic resonance analysis confirmed the presence of a GNRA tetraloop in EMCV and PV IRES elements (Phelan et al., 2004; Du et al., 2004). Site-directed substitutions in the FMDV GNRA motif led to a local reorganization of the apical region that allowed the identification of a second motif (Fernandez-Miragall et al., 2006), presumptively encompassing the GNRA receptor. It is noteworthy that this motif is located in an invariant region of about 100 FMDV sequences (Carrillo et al., 2005) that can potentially form Watson–Crick pairs with nearby residues according to computer-folding programs. In spite of the high genetic variability of FMDV viral RNA (Domingo et al., 1992), sequence variability from field isolates shows rare substitutions in the GUAA sequence to GCAA or GCAG, always compatible with GNRA motifs. Conversely, nucleotides engaged in base pairs often show covariation (Martinez-Salas & Fernandez-Miragall, 2004), strongly supporting the need to preserve IRES structure for internal initiation.

Picornavirus IRES elements appear to have a modular organization; each module occupies a domain or structural element that accomplishes a distinct function during internal initiation. In this way every domain is necessary but not sufficient to achieve protein synthesis initiation. For example, RNA sequences encompassing domains 4-5 of the IRES, which contain the capacity to interact with initiation factors eIF4G, eIF4B, eIF3 and PTB (Pilipenko et al., 2000; Lopez de Quinto et al., 2001), are not sufficient to drive internal initiation. In full agreement with this observation, domain 3 of the aphtho- and cardiovirus IRES is an integral part of the IRES element, as mutations within this region impair IRES activity (Lopez de Quinto & Martinez-Salas, 1997; Robertson et al., 1999). This IRES region also contains a structural element where the virus has evolved a tRNA structural mimicry that renders it a substrate for RNase P ribozyme reaction in vitro (Serrano et al., 2007), a property shared with other IRES elements and the NS2 coding region of the HCV genome (Nadal et al., 2002; Sabariego et al., 2004). RNase P is a structure-dependent endonuclease involved in the processing of the tRNA precursor within the nucleus (Evans et al., 2006) that also recognizes, as substrate, the viral RNA of viruses containing tRNA-like structures at the 3′ end of the genome (Guerrier-Takada et al., 1988) and other natural RNAs (Li & Altman, 2003; Altman et al., 2005). On the basis of its recognition by RNase P, it was inferred that the HCV and pestivirus IRES contained a structural element
that mimics the tRNA-like structure (Lyons & Robertson, 2003). Recognition of a secondary response of defective FMDV IRES mutants with modified RNA structures to ribozyme cleavage, as also occurs in a variant RNA molecule present in the HCV quasispecies spectrum (Piron et al., 2005). The significance of the RNase P recognition motif in IRES elements is still unknown, as there is no definitive proof for its involvement in the translation process. Nevertheless, it does not constitute an RNA processing motif in transfected cells (Piron et al., 2005). This observation is consistent with the fact that the entire picornavirus infection cycle, as well as that of HCV and pestivirus, occurs in the cytoplasm of infected cells; therefore, the viral RNA has no access to RNase P.

The 5′ UTRs of the RNA genome of HCV and pestiviruses contain IRES elements whose RNA structural organization represents a different category other than the picornavirus type I and II IRES (see Fig. 2) (Rijnbrand et al., 2004). Site-directed mutagenesis and RNA probing analysis have shown that the HCV IRES consists of three main structural domains, II, III and IV (Honda et al., 1996; Rijnbrand et al., 2000). Under physiological ionic concentration, the HCV IRES adopts a tertiary fold that is essential to promote internal initiation (Kieft et al., 1999). Specific stem–loops, namely IIIa, IIId and IIIc, converge in a four-way junction that participates in the interaction with eIF3 (Kieft et al., 2001). On the other hand, subdomains IIId and IIIe conform short stem–loops involved in the interaction with the 40S subunit (Kolupaeva et al., 2000). A peculiarity of the HCV IRES structure is the formation of a Pk that involves residues of the subdomain IIIIF (Wang et al., 1995; this feature is conserved with the pestivirus IRES (Rijnbrand et al., 1997).

Another category of IRES is represented by the IGR elements, which adopt a tertiary folding that includes three Pk (Fig. 3) (Jan & Sarnow, 2002; Kanamori & Nakashima, 2001). As a result of these unusual RNA-folding properties, the IGR mimics the initiator tRNA during internal initiation (Wilson et al., 2000a; Jan et al., 2003; Pestova & Hellen, 2003). The three-dimensional architecture of the ribosome-binding domain from the IGR IRES elements is organized around a core helical scaffold, around which the rest of the RNA molecule folds (Costantino & Kieft, 2005). However, subtle changes in the folding pattern of IGRs corresponding to different discistrovirus genomes together with the presence of an additional secondary structure element suggest differences in the interaction with the large ribosomal subunit (Pfingsten et al., 2007).

**IRES activity is resistant to specific host protein cleavage and phosphorylation changes occurring in infected cells**

Picornavirus infection induces a large number of modifications in the host cell that lead to the shut off of cellular protein synthesis. This effect, which is mainly due to delocalization, proteolysis and changes in phosphorylation of host factors, favours viral expression as the cellular machinery is subverted from its normal role (Sarnow, 2003). Efficiency of protein synthesis is mainly controlled at the initiation step; in this regard, RNA viruses have developed specialized strategies to impair host factors responsible for translation initiation control (Sonenberg & Dever, 2003). In turn, the cell has evolved responses to recognize and fight the incoming infectious agent (Cole, 2007; He et al., 2003).

Pioneering studies of picornavirus-infected cells showed the cleavage of eIF4G (Etchison et al., 1982), but proteolysis of other host factors also contribute to virus pathogenesis (reviewed by Lloyd, 2006). Transcription factors and other proteins involved in gene expression and cell signalling are cleaved during picornavirus infection (Clark & Dasgupta, 1990; Falk et al., 1990; Yalamanchili et al., 1996). The number of host factors identified as targets of viral proteases has increased with the advances of proteomic analysis; characterization of the protein patterns in CVB3-infected cells revealed modification of UNR, nucleophosmin, lamin and the p38 mitogen-activated protein kinase (Rassmann et al., 2006). Specific cleavage of Gemin 3, a nuclear protein, was recently found during PV infection, together with a redistribution of the target proteins from the cell nucleus to the nuclear periphery (Almstead & Sarnow, 2007). Gemin 3 is a component of the macromolecular complex that mediates U snRNP assembly (Battle et al., 2006), and therefore its processing has a profound impact on mRNA splicing. A close location of viral protease targets and the virus-encoded proteases is consistent with the perinuclear location of viral proteins, including protease precursors in EMCV-, FMDV-, PV- and HRV-infected cells (Garcia-Briones et al., 2006; Aminve et al., 2003; Aminve et al., 2004; Sharma et al., 2004).

Coincident with host factor cleavage, efficient picornavirus IRES performance ensures the viral protein accumulation needed to accomplish virus production. Two aphthovirus gene products, L and 3C, are efficient proteases that recognize, as substrates, several host components, in addition to the viral polyprotein (Fig. 1). As already mentioned, the L protease cleaves the translation initiation factor eIF4G (Lopez de Quinto & Martinez-Salas, 2000; Medina et al., 1993), and in the late steps of infection the 3C protease also processes eIF4G (Strong & Belsham, 2004). Enterovirus and rhinovirus genomes also encode two proteases, 2A and 3C (Fig. 1). The protease 2A, located in the coding region preceding the non-structural proteins (Cuconati et al., 1998), is responsible for cleavage of eIF4G (Lampheer et al., 1993).

Partial processing of other RNA-binding proteins, PABP and PCBP, was observed in PV and CBV infection (Perera et al., 2007; Kuyumcu-Martinez et al., 2002, 2004; Lerner & Nicchitta, 2006). Subsequently, cleavage of eIF3a and b, PABP and PTB was reported in FMDV-infected cells (Rodriguez Pulido et al., 2007). Proteolysis of PABP and
PTB correlated with the extent of cytopathic effect in infected cells, whereas eIF3a, eIF3b and eIF4G were cleaved at early times post-infection, but presumably after the first round of viral RNA translation. Cleavage of eIF3a–b in FMDV-infected cells, in addition to eIF4G, may contribute to host translation shutdown (Table 2). These polypeptides belong to the 13 subunits of eIF3, a translation factor essential for the assembly of the 48S initiation complexes on the host mRNAs via eIF3–eIF4G interaction (Hinnebusch, 2006; LeFebvre et al., 2006). Cleavage of PTB is not expected to alter cap-dependent translation initiation; although recombinant PTB fragments similar to those generated during PV infection inhibit IRES-dependent translation (Back et al., 2002), the newly replicated viral RNA is presumably encapsidated at the late stages of the FMDV infection cycle when PTB is cleaved, and thereby not used as template for translation.

Proteolysis of eIF4G and PABP in picornavirus-infected cells disrupts the functional bridge that connects the 3′ end of the host polyadenylated mRNAs with the 5′ cap via the trimeric eIF4F complex. During FMDV infection, cleavage of PABP was at least partially due to the Lb protease activity (Rodriguez Pulido et al., 2007). However, since FMDV IRES stimulation in transfected cells expressing the Lb protease is achieved by the 3′ UTR heteropolymeric region devoid of a poly(A) tail (Lopez de Quinto et al., 2002), it is expected that cleavage of PABP might inhibit cap-dependent initiation without a significant effect on internal initiation (Table 2).

Cells infected by RNA viruses accumulate double-stranded RNA (dsRNA). This molecule leads to the activation of the cellular kinase PKR, a protein that has a key role in the innate immune response to viral infection (Cole, 2007). The RNA-dependent protein kinase (PKR) is responsible for phosphorylation of eIF2α, causing a reduction of eIF2/GTP/Met-tRNA(i)(Met) ternary complexes, and thus inhibiting mRNA translation (Sonenberg & Dever, 2003).

To counteract this response, RNA viruses have evolved distinct strategies (Katze et al., 2002). In HCV-infected patients the virus establishes persistent infections, reflecting the evasion of host immunity and interference with interferon (IFN) innate immune defences. Cellular antiviral immunity is based on the host recognition of virus infection-associated dsRNA. This viral replication product is sensed by Toll-like receptor 3 (TLR3) and the RNA helicases, retinoic acid inducible gene I (RIG-I). Detection of dsRNA, assisted by the protein-interacting CARD domains, results in the activation of the transcription factors IFN regulatory factor 3 (IRF3) and NF-κB. Recent data have shown that cardif, a CARD-adaptor protein, is cleaved by the HCV NS5-4A gene product (Meylan et al., 2005), a viral encoded serine protease that blocks IFN-β production. CARD-adaptor proteins interact with RIG-I and recruit IKK kinases, leading to the activation of NF-κB and IRF3. Thus, the NS3-4A protease (Fig. 2) induces specific proteolysis of TLR that bind viral derived products early in infection (Meylan & Tschopp, 2006). This proteolysis event prevents signalling to kinases responsible for the activation of IFN-β promoter (Foy et al., 2005; Li et al., 2005) and thereby interfering with host innate immune defences.

### Implications of viral RNA 5′–3′ end interactions in stimulation of IRES activity and the infection cycle

The UTRs of the picornavirus genomes, as well as of other positive-strand viruses, contain multiple functional elements that play specific roles during the viral cycle. In this regard, specialized structural motifs in the viral UTRs are recognized by viral and host RNA-binding proteins that control translation, replication and interaction of viral replication complexes within host cell macromolecular structures. This process has been studied in great detail in picornavirus RNAs. The cloverleaf (CL) structure in the 5′ end of the PV genome (Fig. 1) and the downstream C-rich spacer interact with the cellular protein PCBP2 and the viral protein 3CD (Gamarnik & Andino, 1997; Toyoda et al., 2007). This ternary complex is essential for RNA replication and, hence, virus proliferation. In addition, the efficiency of picornavirus RNA replication is greatly enhanced by protein bridges that interact with both ends of the viral RNA, mediating RNA circularization. This is illustrated by the functional bridge generated by PCBP2, a dual interactor of the PV 5′ UTR and the host factor PABP, which in turn recognizes the viral poly(A) tail (Herold & Andino, 2001).

Host and viral RNA-binding proteins have also been proposed to mediate 5′–3′ contacts in the genome of other positive-strand RNA viruses. A large variety of proteins recognize specific conformational motifs in the termini of the viral RNA from flaviviruses and rotaviruses (Isken et al., 2003, 2004; Vende et al., 2000; De Nova-Ocampo et al., 2002) and in some cases, mediate the interaction with third components. For example, translation of rotavirus mRNAs, which are capped but not polyadenylated, is accomplished through the action of the viral non-structural protein NSP3; this protein specifically binds the 3′ sequence of viral mRNAs and interacts with eIF4G and RoXAN, forming a ternary complex (Groft & Burley, 2002; Piron et al., 1998; Vitour et al., 2004). Specific motifs in the 5′ and 3′ ends of the RNA from a pestivirus, bovine viral

### Table 2. Effect of protein cleavage on translation

<table>
<thead>
<tr>
<th>Host factor</th>
<th>Viral protease</th>
<th>IRES-dep</th>
<th>Cap-dep</th>
</tr>
</thead>
<tbody>
<tr>
<td>eIF4G</td>
<td>L, 2A, 3C</td>
<td>Stimulation</td>
<td>Inhibition</td>
</tr>
<tr>
<td>eIF3a, b</td>
<td>3C</td>
<td>Inhibition</td>
<td>Inhibition</td>
</tr>
<tr>
<td>PABP</td>
<td>L, 2A, 3C</td>
<td>–</td>
<td>Inhibition</td>
</tr>
<tr>
<td>PTB</td>
<td>3C</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>PCBP</td>
<td>3C</td>
<td>–</td>
<td>–</td>
</tr>
</tbody>
</table>
diarrhea virus (BVDV), interact with the NFAR proteins (Isken et al., 2004). These host factors have in common a double-stranded RNA-binding motif (dsRBM) shared with PKR and the dsRNA-activated protein kinase that senses the presence of dsRNA in infected cells (Cole, 2007). Another example of the function carried out by RNA-binding proteins recognizing UTRs of viral RNAs is illustrated by the interaction between eIF1A and the 3’ stem–loop of flaviviruses, which facilitates minus-strand synthesis by promoting an interaction between the 3′ end of the genome and the replication complexes (Davis et al., 2007).

The role of viral 3′ UTR sequences in translation enhancement has been a controversial issue. Perhaps this is due to the different experimental approaches used to study this event that strongly influences the results depending on whether the study was carried out in vitro using RRL, or in vivo using tissue culture cells or mice (McCaffrey et al., 2002). The poly(A) tail present in the majority of eukaryotic mRNAs improves the efficiency of translation initiation through recruitment of PABP, enabling its interaction with eIF4F located at the mRNA 5′ end. RNA virus genomes that utilize IRES elements to promote cap-independent translation are differentially affected by PABP and the poly(A) status. Translation of polyadenylated CBV3 as well as capped-RNAs displayed increased sensitivity to the PABP inhibitor Paip2 compared with EMCV or HCV. Sucrose density gradient analyses suggested a stimulatory role for PABP and 3′ poly(A) in the CBV3 initiation phase of protein synthesis, while assembly of HCV and EMCV RNAs into ribosomal complexes was only slightly affected by either factor (Bradrick et al., 2007). Accordingly, the HCV IRES was not affected by the presence of PABP and a poly(A) tail (Michel et al., 2001), a result consistent with the fact that the HCV RNA is not polyadenylated. However, an independent study claims a translation enhancer function of the HCV 3′ UTR in the variable region, the poly(U:C) tract, and the 3′-terminal stem–loop 1 of the conserved 3′ X region (Song et al., 2006). The differential effects of PABP and poly(A) on translation imply mechanistic differences between viral IRES elements and suggest modulating roles for PABP and the poly(A) tail at multiple phases of translation.

A functional link between the aphthovirus IRES and the 3′ end of the viral RNA was suggested by the specific stimulation of IRES activity by the FMDV 3′ UTR, irrespective of Lb protease coexpression in transfected cells or the absence of poly(A) in the translated transcripts (Lopez de Quinto et al., 2002). This link could take place through direct RNA–RNA contacts, through protein bridges mediating RNA circularization or both. In other RNA viruses, e.g. Dengue, genome circularization promoted by direct RNA–RNA contacts involving inverted terminal repeats has been shown to occur in vitro, and to promote viral replication in vivo (Alvarez et al., 2005; Filomatori et al., 2006). In other virus models such as RNA plant viruses, direct RNA–RNA interaction between a small number of nucleotides in the 3′ and the 5′ UTR sequences of the genome controls cap-independent translation initiation (Miller & White, 2006; Rakotondrafarana et al., 2006; Shen & Miller, 2007).

The S region at the 5′ terminus of FMDV RNA is predicted to adopt a hairpin structure (Escarmis et al., 1992). The 3′ UTR of FMDV RNA is composed of two stem–loops and a poly(A) tract that are required for replication and infectivity (Saiz et al., 2001). In the viral genome of other picornaviruses, the 3′ UTR is organized as two stem–loops that adopt a quasi-globular organization (van Ooij et al., 2006) and constitute essential determinants of virus replication (Dobrikova et al., 2006). A long-range stimulation of type I picornavirus IRES by the 3′ UTR sequences was found in PV RNA (Dobrikova et al., 2003), where the IRES activity in neuronal cells was connected to the presence of particular 3′ UTR structural elements. This result suggests a functional role of 3′ end–IRES interaction in viral pathogenesis.

The 3′ end of the FMDV genome establishes two distinct strand-specific long-range RNA–RNA interactions, one with the IRES element and the other with the S region (Fig. 4a, b) (Serrano et al., 2006). A high-order structure adopted by the entire IRES and the 3′ UTR was essential for RNA interaction, whereas the S region interacted with each of the stem–loops at the genome 3′ end. The possibility that proteins might stabilize RNA–RNA bridges in the viral genome was supported by the specific interaction of polypeptides with the viral UTRs (Rodriguez Pulido et al., 2007; Serrano et al., 2006). Two proteins with the same mobility interact with the S region and the 3′ UTR. One of them, presumably PCBP, competed for binding to 3′ UTR, suggesting that this protein interacts with both ends of the FMDV genome. Additionally, the long poly(C) tract located in the 5′ UTR between the S region and the IRES, is also a candidate to interact with PCBP. Whether its recognition by host factors displaces the balance of IRES–3′ UTR to S–3′ UTR interactions awaits further investigation.

Thus, bridging of 5′ and 3′ ends in the picornavirus genome involves both direct RNA–RNA contacts and RNA–protein interaction (Fig. 4a, b); this observation provides a mechanistic basis for translation stimulation and replication of the viral RNA. During PV infection, it is believed that translation and replication occur consecutively, but not simultaneously, on the same viral RNA molecule (Gamarnik & Andino, 1998; Novak & Kirkegaard, 1994). In FMDV RNA, the 3′ UTR–IRES interaction was not affected by the formation of the S–3′ UTR complexes, indicating that different sequences are involved in these contacts. On the other hand, these contacts do not occur simultaneously in the same RNA molecule since ternary complexes were not detected under different permissive binding conditions (Serrano et al., 2006). Thereby, a switch from translation to replication
may be governed by a transition from the 3’ UTR–IRES to the 3’ UTR–S complexes during early stages of infection.

As mentioned above, several cellular proteins are cleaved by viral-encoded proteases during picornavirus infection. Such cleavage events are likely to be involved in the successive switches from viral translation to viral RNA replication and, finally, to virion assembly. The host protein PCBP2 together with the viral 3CD were identified as candidates for regulating such a mechanism (Gamarnik & Andino, 1998). PCBP2 is required for translation initiation on picornavirus genomes with type I IRES elements (Blyn et al., 1997) and also for RNA replication. PCBP2 forms functional complexes with components of the viral translation and replication machinery. PCBP1 and 2 are cleaved during the mid-late phase of PV infection by the viral proteases 3C/3CD (Perera et al., 2007). The cleavage results in a truncated PCBP2 that is unable to function in translation but maintains its activity in viral RNA replication. This event may promote the switch from viral translation to RNA replication.

**Evolutionary constraints imposed by viral UTRs**

A large number of picornavirus chimeras have been generated by the exchange of UTRs. In many cases, the resulting cDNAs exhibited a compromised infectivity. Among others, chimeras exchanging IRES elements between entero- and cardioviruses delayed viral growth (Alexander et al., 1994; Gromeier et al., 1996), exchanging 3’ UTR sequences between PV and bovine enteroviruses diminished replication efficiency (Rohll et al., 1995), while substitution of swine vesicular disease virus 3’ UTR for that of FMDV abrogated viral infectivity (Saiz et al., 2001). The compromised activity of picornavirus genetic hybrids can be associated with specific properties of the UTRs; structural conformations adopted by sequences in the respective genomic RNAs might have evolved to interact with specific trans-acting factors. Additionally, intramolecular interactions involving specific 3’ UTR high-order structures would not be conserved in the hybrid RNA. Therefore, the evolution of the viral RNA might be strongly influenced by RNA–RNA intramolecular interactions that confer phenotypic properties to the viral RNA molecule.

In addition to host factors, a link between UTR sequences and viral proteins has been reported in CBV RNA carrying a heterologous IRES element together with a modified 3’ UTR. Genetic adaptation of these constructs conferring a cell type-specific propagation deficit in neuroblastoma cells resulted in the selection for mutations in the non-structural proteins 3A and 3C. Thus, a functional complex of 3A/3C or 3AB/3CD precursors with the IRES and 3’ UTR is important for viral propagation (Florez de Sessions et al., 2007).

Although the signals that control translation and replication in the viral UTRs are partially overlapping, they can differ significantly between related viruses. The viral RNA genome of the pestivirus BVDV that also contains an IRES (Fletcher & Jackson, 2002) shares, with the HCV IRES, a similar RNA structure. However, functional differences exist between the viral genome of HCV and BVDV (Grassmann et al., 2005; Rijnbrand et al., 2004). In the 5’ UTR of BVDV RNA, replication signals are restricted to the 5’-terminal domain I. Specific HCV replication signals reside in domain I of the viral RNA but also involve domains II and III that constitute the functional IRES. While domain I of the BVDV viral RNA supports IRES activity, domain I in HCV RNA seems to downregulate IRES function. These data suggest that the genomes of HCV and BVDV apply different mechanisms to coordinate viral protein translation and RNA synthesis.

**Concluding remarks**

During the past decade, the study of UTRs located at each end of positive-strand viral RNAs has revealed an enormous potential for the control of gene expression.
and viral replication. In the life cycle of positive-strand RNA viruses the genome serves as a dual template for translation and replication of the viral RNA. These highly dynamic processes must be properly balanced to ensure efficient viral proliferation. To achieve this, specific high-order RNA structures at the termini of virus genomes play a key role in regulating translation and viral RNA synthesis. Combination of structural and functional studies, together with novel proteomic advances, will help to decipher the mechanistic details of how these functions are controlled by viral-specific RNA structures working in close association with host and viral proteins.

Acknowledgements

We are grateful to C. Gutierrez for helpful suggestions on the manuscript. This work was supported by grants BFU-2005-00948 and by an Institutional grant from Fundación Ramón Areces.

References


