Molecular Biology Reports **19**: 147–159, 1994.
© 1994 *Kluwer Academic Publishers. Printed in Belgium.*

Internal initiation of translation of picornavirus RNAs

Richard J. Jackson, Sarah L. Hunt, Catherine L. Gibbs & Ann Kaminski
Department of Biochemistry, University of Cambridge, Tennis Court Road, Cambridge CB2 1QW, UK

Received 13 December 1993; accepted 17 December 1993

Key words: translation initiation, internal ribosome entry, pyrimidine tract binding protein, oligopyrimidine tract

Introduction

The animal picornavirus family comprises many different species which cause a wide range of diseases, e.g., acute paralysis (poliovirus and enterovirus-71), hemorrhagic conjunctivitis (enterovirus-70), viral myocarditis (coxsackie B viruses), foot and mouth disease, hepatitis, common colds [1]. In spite of this diversity of diseases and target tissues, all picornaviruses are quite similar in structure and genetic information. The genome is a positive strand RNA molecule about 7,500 nt. long (depending on the species), with a single long open-reading frame coding for a polyprotein which includes either one or two protease domains (likewise depending on the species) that cleave the polyprotein into the various structural and non-structural proteins. The virion RNA is unusual in that it lacks the 5′-cap structure universal to all eukaryotic cellular mRNAs, and instead has a covalently linked small peptide of about 25 amino acid residues (VPg). Following entry of the virus into the infected cell and the subsequent uncoating, the VPg is rapidly destroyed and thus the RNA is actually translated as uncapped RNA with a 5′ pU... end.

The 5′-untranslated regions (5′-UTR) of picornavirus RNAs are unusually long (ranging from 610 to more than 1000 nt. depending on the virus species) in comparison with the average cellular mRNA, and are thought to have regions of complex secondary structure [2]. In addition, the picornavirus 5′-UTRs have many AUG codons (up to a maximum of 15) which do not appear to be used as translation initiation sites, despite the fact that some of them are located in a local context background which would be regarded as very favourable to initiation according to the rules of the scanning ribosome model. What is particularly remarkable about these upstream AUG codons is that the majority are very poorly conserved between closely related virus species or between different serotypes of the same species, or even between different isolates of any one serotype [3]. This argues that these AUG triplets are not functional translation initiation sites. All these features of the 5′-UTR suggest that picornaviral RNAs could not be translated efficiently by the 5′-end dependent ribosome scanning mechanism [4, 5] thought to be operative on all cellular mRNAs and capped viral mRNAs.

Evidence for internal initiation of translation of picornavirus RNAs

Direct evidence for an alternative mechanism of internal ribosome entry was obtained by insertion of the picornavirus 5′-UTR as the intercistronic spacer element in laboratory-generated dicistronic constructs [6, 7]. The outcome was a very large increase in the yield of product from the downstream cistron, which, in the absence of the picornavirus element, is translated very inefficiently as very few ribosomes scan through the intercistronic region after translation of the upstream cistron. However, the inserted picornavirus sequences clearly do not promote downstream cistron translation simply because they increase the probability that ribosomes which have translated the upstream cistron will resume scanning through the intercistronic spacer, since the insertion often results in a yield of downstream cistron translation product far in excess of the yield from the 5′-proximal cistron. Other controls show that the downstream cistron real-

ly is expressed from dicistronic mRNAs and not from monocistronic mRNAs generated by nuclease cleavage in the intercistronic spacer. Thus the unambiguous conclusion is that the picornavirus 5′-UTR segment is allowing direct internal ribosome entry to the downstream cistron.

Deletion analysis allows the minimal segment of the picornavirus 5′-UTR required for internal initiation to be defined, and this is generally known as the IRES (for 'internal ribosome entry segment'). Such deletion analyses are best carried out using the dicistronic mRNA assay system, since extensive deletions of the 5′-UTR which abolish IRES function in the dicistronic mRNA assay may allow translation of a monocistronic version via a 5′-end dependent scanning mechanism. However, it is now clear that if the whole viral 5′-UTR, or even just the minimal IRES, is present in a monocistronic construct, ribosome scanning from the 5′-end does not penetrate through the IRES (presumably because of the secondary structure and/or the upstream AUG triplets); virtually all initiation at the authentic site is by internal ribosome entry and none by scanning [8, 9].

The minimal IRES element of several picornavirus species has been mapped (see [10] for primary references to the sequences and IRES delineation results). In all cases it is ~ 450 nt. long, but the surprising result is that there is no single consensus IRES sequence common to all viruses. Instead, the viruses fall into three distinct groups (Table 1) based on IRES sequence conservation and, more markedly, secondary structure [11]: (i) the enteroviruses (e.g. poliovirus) and rhinoviruses; (ii) the cardioviruses and aphthoviruses (foot and mouth disease virus), as well as the recently characterised echovirus-22 [12]; and (iii) hepatitis A virus. Although there is clear but somewhat patchy conservation of primary sequence within each group, there are no extended primary sequence motifs shared between the groups apart from an oligopyrimidine tract near the 3′-end of the IRES.

Despite this divergence of IRES structure between the different groups of picornaviruses, a common model for internal initiation is beginning to emerge. This model is shown schematically in Fig. 1, and has the following features, which will be discussed in subsequent sections. For a more detailed examination of these questions the reader is referred to two recent reviews [10, 13].

1) The ribosome binding site, defined as the most 5′-proximal point at which initiation can occur, is located at the 3′-end of the IRES element, at or very close to the authentic initiation site on the cardiovirus RNAs, or the upstream of the two initiation sites of FMDV [11]. In contrast, in enteroviruses and rhinoviruses, initial ribosome binding is at or near the most 5′-distal AUG triplet within the 5′-UTR, an AUG triplet which is not used significantly as a translation initiation site but seems to be just an entry site from which the ribosomes are transferred, most probably by scanning, to the authentic initiation codon, the next AUG codon downstream.
2) The selection of the ribosome entry/internal initiation site is determined in part on its distance and spacing from upstream motifs within the IRES element, of which the oligopyrimidine tract may be one significant motif in some but not all species of picornaviruses.
3) The secondary and tertiary structure of the IRES element is critical for internal initiation, whilst the essential primary sequence motifs are quite short and are located mainly in unpaired segments. The function of these essential primary sequence motifs is thought to be either to provide RNA tertiary structure interactions, or to be part of the ribosome binding site, or to be the binding site for trans-acting protein factors, which in turn facilitate internal ribosome binding.

The internal ribosome entry site is at the 3′-end of the IRES

The idea that ribosome binding occurs at the 3′-end of the IRES originates from work on the cardio-/aphthoviruses, particularly EMCV strain-R [9]. In this strain the viral polyprotein open-reading frame starts at the eleventh AUG triplet at nt. 834 [14, 15], though some initiation may also occur at AUG-12 (nt. 846) and it is technically difficult to quantitate the relative frequency of initiation at these two sites. In this particular strain of EMCV [15], but not in any other cardiovirus and in only two out of nine strains of FMDV [16, 17], there is another AUG triplet (AUG-10) located at nt. 826, upstream of the authentic initiation site(s) and out of frame with respect to the viral polyprotein coding sequence (Fig. 2). Since the context of AUG-10 is favourable for initiation according to the rules of the scanning ribosome model, one would expect that AUG-10 would be used in preference to AUG-11 if ribosomes approached AUG-11/12 by scanning from an upstream entry site, and this preference for AUG-10

Table 1. Classification of picornaviruses according to IRES sequence homologies

Entero-/rhinoviruses		
Bovine enterovirus		BEV
Poliovirus group	Polioviruses (3 serotypes)	PV
	Enterovirus-70	
	Coxsackie A21 virus	
Coxsackie B virus group	Coxsackie B viruses	CBV
	Coxsackie A9, A16 viruses[1]	
	(some) Echoviruses[2]	
Human rhinoviruses		HRV
Cardio-/aphthoviruses		
Echovirus-22[3]		
Cardioviruses	Encephalomyocarditis virus and mengovirus	EMCV
	Theiler's murine encephalomyelitis virus	TMEV
Aphthoviruses	Foot and mouth disease virus	FMDV
Hepatoviruses	Hepatitis A virus	HAV

The table gives the standard abbreviations used throughout this article for each virus species.
[1] Some viruses designated as coxsackie A viruses by classical criteria have IRES sequences much closer to those of the coxsackie B viruses than to coxsackie A21 virus.
[2] Partial sequences of the 5′-UTR of some echoviruses show close homology with the IRESes of coxsackie B viruses (L. Kinnunen and T. Poyry, personal communication).
[3] The coding sequence of echovirus-22 shows no homology with any other picornavirus, but the putative IRES is quite similar to the cardio- and aphthovirus IRES sequences [12].

was in fact seen in translation assays with monocistronic RNAs from which most of the IRES sequences had been deleted in order to convert the RNA into one translated by the conventional scanning mechanism. However, when the intact IRES was present and thus internal initiation was operative, translation was initiated almost exclusively at AUG-11/12, with less than 1% of initiation events occurring at AUG-10, which demonstrates that the ribosomes must enter directly at AUG-11 (or at least at a site between AUG-10 and AUG-11), and do not select AUG-11 by scanning from an upstream entry site [9].

The polyprotein reading frame of hepatitis A virus starts with two in-frame AUG codons 6 nt. apart. If most of the IRES is deleted to convert the RNA into one translated by the scanning ribosome mechanism, the two sites are used with roughly equal frequency, but if the intact IRES is present and internal initiation is operative, the downstream of the two AUG codons is the preferred, but not the exclusive, initiation site [18]. This again argues for direct ribosome entry at these sites rather than scanning from an upstream entry site.

In contrast to the cardioviruses and hepatitis A virus, in which the IRES element extends virtually right up to the authentic initiation site, in the entero-/rhinoviruses the 3′-boundary of the IRES element, as defined by the dicistronic mRNA assay, is located some 25 nt. downstream of the oligopyrimidine tract (as in cardioviruses and hepatitis A virus) but some 40 nt. or 160 nt. in rhinoviruses and enteroviruses, respectively (Fig. 3), upstream of the authentic initiation site [19, 20]. Nevertheless, there are mutants of poliovirus and coxsackie B1 virus bearing deletions of almost all of this $\sim$ 160 nt. segment which are viable provided that (a) the 5′-proximal UUUC portion of the oligopyrimidine tract is retained, and (b) the authentic initiation codon is located 20–25 nt. (ideally 22 nt.) downstream of the C residue in this vestigial oligopyrimidine motif. Such mutants obviously resemble cardioviruses in the location of the authentic initiation codon relative to upstream IRES sequences including the oligopyrimidine tract [21–25].

In the wild type genomes of the entero-/rhinoviruses there is an AUG triplet which is located

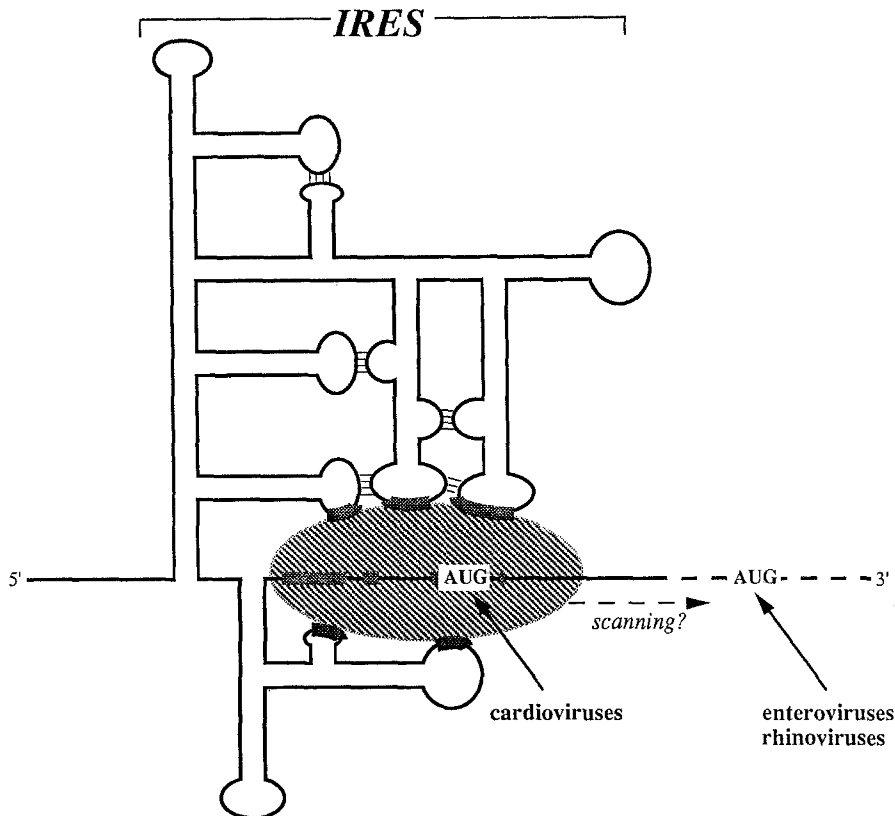


Fig. 1. Schematic model for internal initiation of picornavirus RNA translation. The IRES is viewed as a complex secondary and tertiary structure element, the main function of which is to present a number of short primary sequence motifs, mainly in unpaired regions and depicted as thick shaded lines, in the correct spatial organisation. These motifs are presumed to be critical binding sites for the ribosome itself (shaded oval), or for protein factors which promote internal ribosome entry. The ribosome binding site is at the 3′-end of the IRES at an AUG triplet. This AUG codon is the authentic initiation site for viral polyprotein synthesis in the case of the cardioviruses. In the entero-/rhinoviruses it serves only as a ribosome entry site, not as a functional initiation site, and the ribosomes are transferred, probably by a scanning mechanism, to the polyprotein initiation site which is the next AUG codon downstream.

about 20–25 nt. downstream of the oligopyrimidine tract (i.e. in the same relative position as the authentic initiation site in cardioviruses and HAV) and is absolutely conserved amongst all viruses of this group [2, 11]. No product from initiation at this site has been detected *in vivo*, and in cell-free translation assays initiation at this site occurs with very low frequency unless the Mg^{2+} concentration is raised to far above the optimum for translation of all other mRNAs, including both cellular mRNAs (dependent on scanning ribosomes) and IRES-dependent RNAs. Moreover, as the length (as well as the sequence) of the reading frame following this AUG triplet varies so widely between closely related viruses, ranging from 4 codons in enterovirus-70 to some 70 codons in poliovirus type-1 [26, 27], it seems stretching credibility beyond its limit to imagine that any translation product initiated at this site could be biologically relevant. Nevertheless, when the AUG triplets in the 5′-UTR of poliovirus type-2 were mutated individually, the infectivity of these mutants was compromised only in the case of the mutation of this particular AUG, which resulted in a small plaque phenotype and a reduction of *in vitro* translation efficiency of some 60–70% [28, 29]. Thus the evidence suggests that although this AUG is not used to any significant extent as a functional initiation site, it must be quite important, but not absolutely essential, for internal initiation, possibly as a proximal recognition element of the ribosome entry site.

Fig. 2. The IRES elements of cardio-/aphthoviruses: their position within the 5′-untranslated region, their secondary structure and the sequences at the ribosome binding site. The upper part of the figure depicts the 5′-UTRs of these viruses drawn to scale, with the position of the authentic initiation site for viral polyprotein synthesis shown by the arrow, and the IRES element, either proven by deletion analysis or assumed on the basis of sequence homologies, shown as a black bar. In the lower left is shown the consensus secondary structure model of the IRESes of this group of viruses; the asterisk denotes the position of the only AUG triplet conserved in all strains and species. The sequences at the 3′-end of the IRES, at the actual ribosome entry site, are shown lower right: the consensus sequence for each species is given in upper case, and the strain-dependent variations in lower case. The authentic initiation site is shown in bold, and the positions of the three AUG codons of EMCV strain-R discussed in the text (AUG-10, 11 & 12) are indicated. These sequences at the ribosome binding site are joined directly to the position of the arrowhead in the secondary structure cartoon.

Events following internal ribosome entry

If the ribosome entry site in entero-/rhinoviruses is at or very near this AUG triplet located ~ 25 nt. downstream of the oligopyrimidine tract, this raises the question of how the ribosome subsequently reaches the authentic initiation codon, which is invariably the next AUG codon downstream, at a distance of some 40 nt. in rhinoviruses and ~ 160 nt. in enteroviruses. There is considerable circumstantial evidence that this transfer is by a scanning mechanism, at least in polioviruses. The ~ 160 nt. segment between the entry site and the initiation codon includes a segment of 100 nt. which is hypervariable between different serotypes and even between different isolates of the same serotype, yet never includes an AUG triplet [3]. This, together with the fact that deletions in this region and most types of insertion have little influence on infectivity, argues that the transfer of the ribosome from the entry site to the initiation codon is indifferent to the actual intervening sequences. The only type of insertions which are inhibitory to *in vitro* translation initiation at the authentic site or reduce infectivity are those that would be predicted to prevent a scanning ribosome from reaching the authentic initiation codon: introduction of a predicted hairpin loop structure or an AUG codon [6, 30]. Significantly, when an AUG codon with a poor local sequence context for recognition by a scanning ribosome was inserted in this intervening region, all of the phenotypic revertant viruses which arose were found

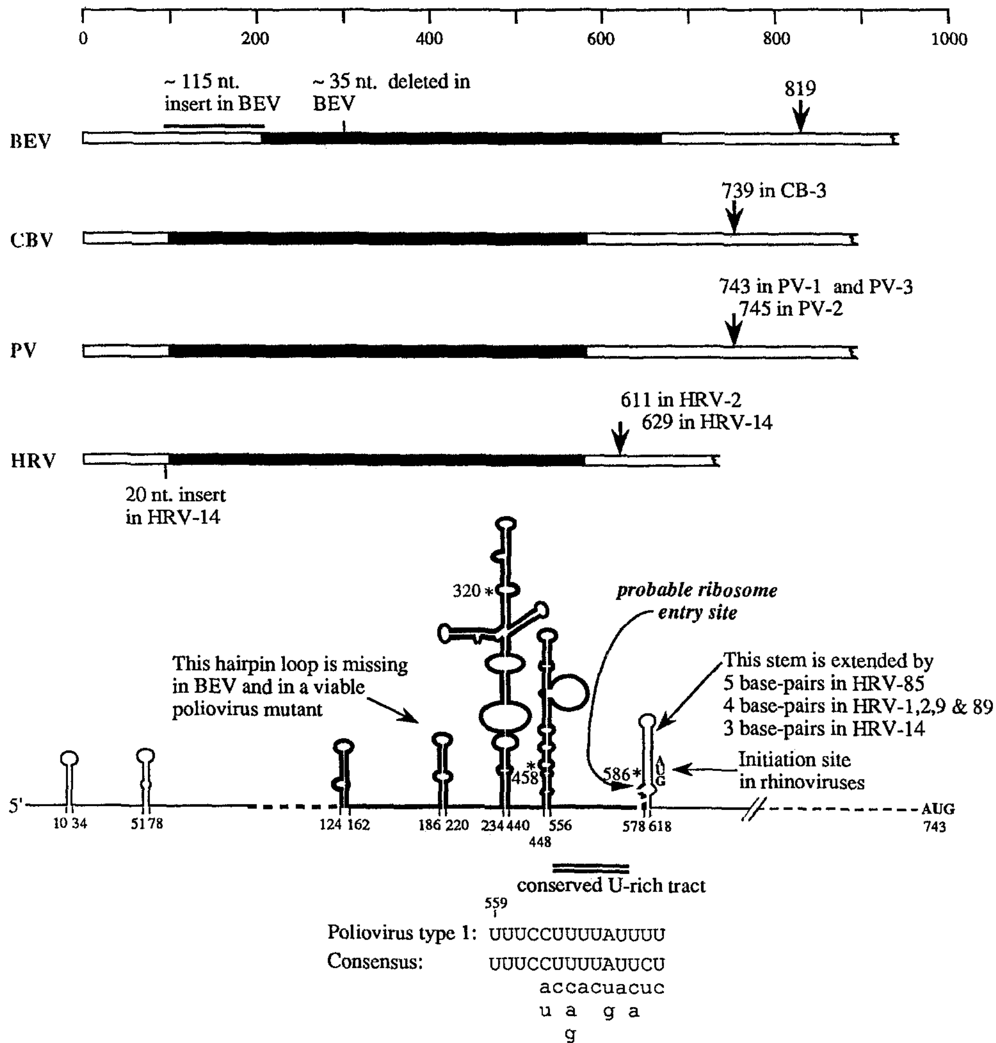


Fig. 3. The IRES elements of entero-/rhinoviruses: position within the 5′-untranslated region, secondary structure and the position of conserved AUG triplets. The upper part of the figure depicts the 5′-UTRs of these viruses drawn to scale, with the position of the authentic initiation site for viral polyprotein synthesis shown by the arrow, and the IRES element, either proven by deletion analysis or assumed on the basis of sequence homologies, shown as a black bar. Below is shown the consensus secondary structure model of the 5′-UTR of this group of viruses, using the nucleotide numbering of poliovirus type-1. The first and last base-paired nucleotides in each stem are numbered, and asterisks show the positions of the AUG triplets in the 5′-UTR that are absolutely conserved in all viruses of this group. The IRES element is depicted by the thickened line; uncertainties with regard to the exact boundaries at each end are denoted by the broken thickened line. The position and sequence of the start of the conserved oligopyrimidine tract is shown, with the consensus in upper case letters and the variations found in different species or strains in lower case.

to have lost the inserted AUG codon, the majority by point mutation and a few by deletion [30, 31].

Thus the internal initiation directed by both the entero-/rhinovirus IRES and the cardiovirus IRES share the common features of internal ribosome entry at an AUG codon located at the 3′-end of the IRES ~ 25 nt. downstream of the oligopyrimidine tract, but there are differences as to events which follow initial ribosome entry at this site. In the cardioviruses, virtually all the entering ribosomes initiate translation at this site (AUG-11 in EMCV strain-R), though a few may use the next AUG codon which in EMCV is located in-frame a short distance downstream. In contrast, in the entero-/rhinoviruses the AUG motif at the 3′-end of the IRES does not seem to be used at all as a functional initiation site, but appears to be a determinant of

the ribosome entry site from which the ribosomes are transferred by a scanning mechanism to the authentic initiation site (Figs. 1 & 3). Foot and mouth disease virus which has an IRES structure resembling that of the cardioviruses represents an intermediate between these two extremes: a minority of the entering ribosomes initiate translation at the AUG codon at the 3′-end of the IRES (equivalent to AUG-11 of EMCV strain-R), and the majority appear to scan to the next AUG codon located some 84 nt. further downstream and initiate translation at this site [32, 33].

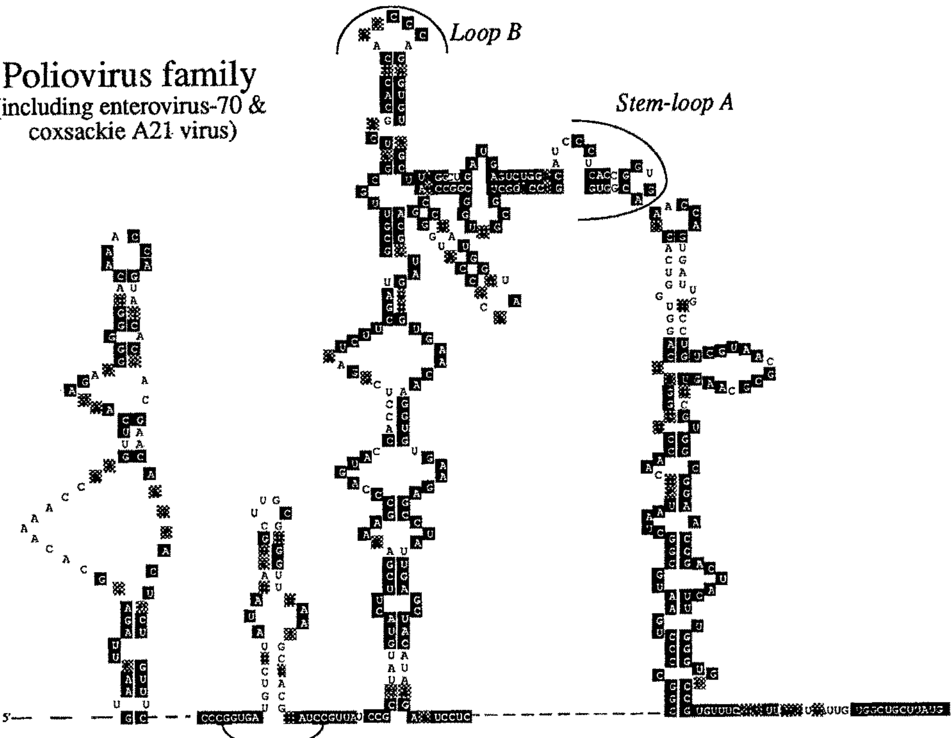
The internal ribosome entry site is determined by its spacing from upstream IRES elements

In view of such precise discrimination between initiation at AUG-10 and AUG-11 of EMCV strain-R, which are located only 8 nt. apart at a site over 800 nt. from the 5′-end of the virion RNA, it would be reasonable to suppose that there must be local sequence determinants around AUG-11 directing ribosomes to bind at this site rather than at AUG-10. However, the sequences between the oligopyrimidine tract and the AUG codon are not particularly well conserved in any group of viruses. In TMEV and FMDV and in all entero-/rhinoviruses the sequences immediately following the oligopyrimidine tract seem to be a hot spot for strain-dependent variations [16, 17], a variability which extends right up to the initiation codon in FMDV [16]. Even though this variable sequence in polioviruses and coxsackie B viruses is followed by 9 residues immediately preceding the AUG triplet that are very strongly conserved, this conservation does not extend to the closely related rhinoviruses and bovine enterovirus. Overall, the most conserved feature of the segment between the pyrimidine-rich tract and the AUG triplet is its length and the relative paucity of G-residues, which taken together suggest that it may play the role of an unstructured spacer rather than comprising important primary sequence determinants directing the ribosome to the authentic initiation site. This is supported by the fact that deletion mutants and linker substitution mutants of poliovirus in which the sequence between the pyrimidine-rich tract and the AUG triplet is radically altered retain full infectivity [21, 22, 24, 25].

The concept of this segment as a spacer suggests that the critical parameter determining the selection of the initiation site might be the distance of the AUG codon from the oligopyrimidine tract or from other upstream motifs in the IRES. When the distance was shortened by deletion, initiation on monocistronic constructs with the EMCV (strain-R) IRES was exclusively at AUG-12 rather than the usual AUG-11, though overall efficiency of translation was decreased. When the distance was increased by insertions, there was a 15-fold increase in initiation at AUG-10 coupled with a decrease at AUG-11, but a complete switch in the preference between these sites did not occur [34]. These results show that the spacing between the AUG codon(s) and upstream elements such as the oligopyrimidine tract are indeed an important determinant of the selection of the initiation site, but there must clearly be other signals, as yet unidentified, which cause AUG-11 to be preferred even when it is located at what would seem to be greater than the optimum spacing.

Spacing from the oligopyrimidine tract and other upstream IRES elements also offers an explanation why the second of the two in-frame AUG codons in HAV is the preferred initiation site when internal initiation is operative [18], as mentioned above. In addition this spacing also seems important for internal ribosome entry in the entero-/rhinoviruses, even though in this case the relevant AUG triplet is not a functional initiation site but is apparently an important element of the ribosome entry site. An increase in this spacing in poliovirus type-1 resulted in phenotypes ranging from small plaque to quasi-lethal. Significantly, the phenotypic revertants were mostly of two types: (i) deletions which restored a spacing between the oligopyrimidine tract and the cryptic AUG codon similar to that found in wild type genomes; (ii) point mutations which introduced a new AUG triplet some 25 nt. downstream from the oligopyrimidine tract [25].

Although the influence of the position of the AUG initiation codon/ribosome entry site is most conveniently measured in terms of the spacing between this AUG codon and the oligopyrimidine tract located a short distance upstream, it is likely that the critical parameter is the spacing of the AUG codon from several upstream IRES motifs, amongst which the oligopyrimidine tract may be of relatively minor importance and might more properly be considered as the start of the unstructured spacer element immediately preceding the initiation site. Mutation of all but two of the pyrimidine residues of the EMCV (strain R) IRES to purines (mainly A residues) reduced the efficiency of translation initiation *in vivo* and *in vitro* by only 10–15%, and mutation of the whole tract resulted in only a 30–35% reduction [34]. In contrast, mutation (to G residues) of some of the pyrimidine residues of the



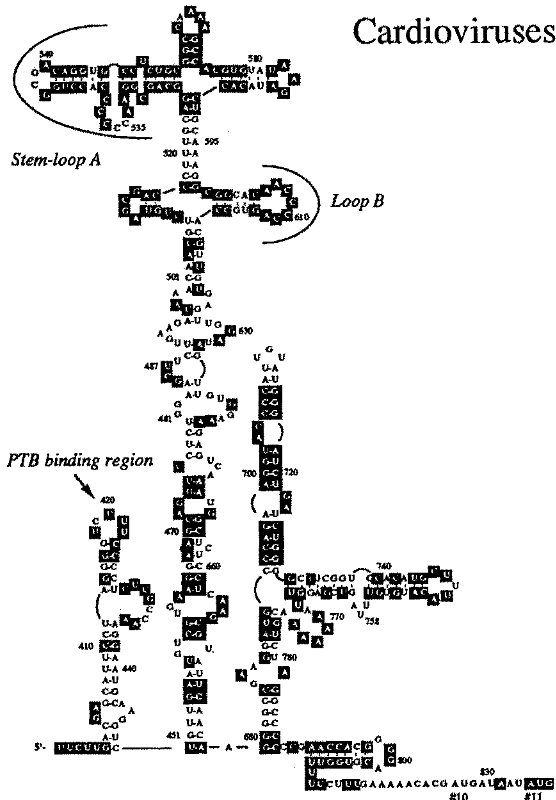


Fig. 4. Patterns of sequence conservation in picornavirus IRESes. The primary and secondary structure of poliovirus type-1 is depicted top left: blacked out residues denote positions which are absolutely conserved amongst all known polioviruses, enterovirus-70 and coxsackie A21 virus, and shaded residues are positions where conservation is high (in > 85% of sequences under consideration) but not absolute. Below is shown the same sequence, but depicting, with the same symbols, the pattern of sequence conservation amongst all known polioviruses, coxsackie B viruses, bovine enteroviruses and human rhinoviruses. At the top right is shown the primary and secondary structure of the EMCV (strain-R) IRES, with the residues that are absolutely conserved amongst all cardioviruses blacked out. Below is shown the pattern of absolutely conserved positions in both cardioviruses and aphthoviruses. The significance of stem-loop A and loop B is discussed in the text, as is the binding site in the cardio-/aphthovirus IRES for PTB, polypyrimidine tract binding protein.

Poliovirus family
plus Coxsackie B viruses
plus Human rhinoviruses
plus Bovine enteroviruses

Loop B

Stem-loop A

Cardioviruses
plus Aphthoviruses

Stem-loop A

Loop B

PTB binding region

Fig. 4. Continued.

oligopyrimidine tract of the IRES of the closely related FMDV had a more serious effect on internal initiation efficiency [35].

Although the tract in entero-/rhinovirus IRESes is quite long (13–17 pyrimidine residues interrupted by 0–2 purines, depending on the species), deletion mutants which lack most of this tract are viable provided the extreme 5′-proximal UUUC motif is retained [21–25]. Point mutations (usually substitution by G residues) in this UUUC motif abrogated IRES function *in vitro* and infectivity in tissue culture cells, but mutations further downstream in the tract had little effect in most [19, 23, 36] , but not all [29], investigations. Thus, although entero-/rhinovirus IRES function seems to require the retention of more of the oligopyrimidine tract than is the case with the EMCV IRES, the minimal tract is very short. Therefore the important spacing parameter may be not so much the distance of the AUG triplet from the oligopyrimidine tract as from other motifs located further upstream in the IRES.

The importance of IRES secondary and tertiary structure

Thus far this account has focussed almost exclusively on the extreme 3′-end of the IRES element, the ~ 25 nt. between the oligopyrimidine tract and the AUG codon that constitutes the ribosome entry site/translation initiation site, without directly addressing the fact that upstream of the oligopyrimidine tract there are ~ 425 nt. of the IRES element all of which is just as important for internal initiation as the 3′-proximal 25 nt. There is much evidence that the secondary and tertiary structure of this ~ 425 nt. segment is essential for IRES function. Thus, small deletions or insertions or linker substitutions at certain sites have been shown to abrogate IRES function, although there are some sites where such mutations are tolerated [24, 37–39]. More direct evidence for the importance of secondary structure comes from examples of point mutations compromising IRES function which can be suppressed by second site mutations which would restore base-pairing according to the secondary structure model [21, 24].

The secondary structures of picornavirus IRESes are quite well established, partly through biochemical probing experiments, but mainly through phylogenetic comparisons [17, 40–44]. The high error frequency of RNA replication has resulted in considerable divergence of sequences between different serotypes of any one species, and between different species, and this provides a great deal of evidence for secondary structure predictions. When closely related viruses are compared, for example the poliovirus family (which includes enterovirus-70 and coxsackie A 21 virus), there is seen to be quite strong conservation of primary sequence except in some regions, mainly regions which seem to be base-paired according to the pattern of co-variances between species (Fig. 4). A similar pattern is seen if a comparison is made between all the coxsackie B viruses, or all the rhinoviruses. However, when a comparison is made between any two of these 'families' of the entero-/rhinoviruses, or between all four such families, the conservation drops quite markedly and is seen to be focussed mainly in unpaired residues (Fig. 4). A similar pattern emerges if the same exercise is carried out with the cardio-/aphthoviruses (Fig. 4).

If we now compare what is conserved amongst all entero-/rhinoviruses and amongst all cardio-/aphthoviruses, it is seen that even though a linear alignment fails to reveal any extended primary sequence homologies between the IRESes of the two groups, there are some motifs which appear very similar in the secondary structure conservation maps, e.g. stem-loop A and loop B (Fig. 4). Although such similarity may be fortuitous, it is intriguing that mutation of the two conserved unpaired residues at the end of stem-loop A has precisely the same outcome on poliovirus IRES function as on EMCV IRES function: mutation of the 5′-proximal G residue causes only a small inhibition, but mutation of the 3′-proximal A residue inhibits internal initiation by 95% in both cases (A. Kaminski, K. Kean and R.J. Jackson, unpublished results).

These patterns of conservation of IRES sequences suggest that what is required for internal initiation is certain quite short primary sequence motifs, mainly unpaired regions, held in the appropriate three dimensional spatial arrangement by the correct secondary and tertiary structure. Moreover, even though the IRESes of the two major groups have quite different primary sequences, and secondary structures which are clearly distinct, it seems quite possible that the three-dimensional array of the essential IRES motifs may be quite similar between the two classes. Nevertheless, we may anticipate that there should be some subtle differences, since the trans-acting factor requirements for internal initiation differ between entero-/rhinovirus IRESes and cardio-/aphthovirus IRESes, as discussed below.

Trans-acting protein factors necessary for internal initiation

It is generally believed that the same set of canonical initiation factors is needed for internal initiation as for initiation by the scanning ribosome mechanism, and indeed there is one long-standing report to this effect [45]. This report predated the discovery of eIF-4F the cap binding factor, the one factor over which there is some doubt as to its requirement for internal initiation. However, as eIF-4F has RNA helicase activity working in the 3′–5′ direction independent of the 5′-cap (as well as cap-dependent 5′–3′ helicase activity) it is by no means certain that it is not required for internal initiation [46, 47].

In addition to the canonical initiation factors internal initiation almost certainly needs additional trans-acting protein factors, since no picornavirus IRES has been reported to function in the wheat germ system, and in the reticulocyte lysate only the cardio-/aphthovirus IRESes function efficiently, whilst poliovirus and rhinovirus RNAs are inefficient messages unless HeLa cell factors (or factors from L-cells or Krebs II ascites cells) are added [48–52]. As both cell free systems translate scanning-dependent mRNAs efficiently, the implication is that the initiation on the cardio-/aphthovirus IRES (and perhaps also on the entero-/rhinovirus IRESes) requires factors present in reticulocytes but not in wheat germ, and entero-/rhinovirus IRES function also requires factors present in HeLa cells in higher abundance than in reticulocytes. HAV IRES function seems to need factors present in mammalian liver at higher abundance than in reticulocyte lysates [53].

One very commonly used approach to identify these factors is to carry out UV cross-linking assays using the whole IRES element as labelled probe, or gel-retardation assays with sub-domains of the IRES, an approach so far applied only to the poliovirus IRES. The latter approach has identified a 50 kDa HeLa cell protein, which is likely to be associated with the endoplasmic reticulum and binds to an internal stem-loop in the poliovirus IRES [54]. However, as this loop is missing in bovine enterovirus and can be deleted from poliovirus with no loss of viability, it seems unlikely that this protein-RNA interaction is essential for internal initiation [39, 55, 56]. The other specific binding protein identified by gel-retardation assays binds in the region of the putative ribosome entry site, and has been identified as the auto-antigen La [57, 58]. Addition of recombinant La to reticulocyte lysates stimulated poliovirus RNA translation, but only at unphysiological concentrations of La in excess of the levels present in HeLa cell extracts [58].

In UV cross-linking assays it is a common observation that many more proteins can be cross-linked to the polio- or rhinovirus IRES than to the cardio-/aphthovirus IRES, and that these proteins seem to be very much more abundant in HeLa cell extracts than in reticulocyte lysates [36, 59–65]. Most of these proteins remain unidentified, except for one which is cross-linkable to all picornavirus IRESes and appears to have higher affinity for the cardio-/aphthovirus IRES (it is actually the only protein cross-linkable to the TMEV and FMDV IRESes) than for entero-/rhinovirus IRESes [65]. This protein doublet of 56–60 kDa has been identified as polypyrimidine tract binding protein (PTB), previously implicated as having a role in pre-mRNA splicing [65–68], though it now appears that the active entity for splicing is a complex between PTB and another protein, PSF [69]. Surprisingly, the strong binding site for PTB in the cardio-/aphthovirus IRES is not the oligopyrimidine tract at the 3′-end of the IRES, but a stem-loop with a U-rich loop, located near the 5′-end [59, 61–63]. An essential role for PTB in cardiovirus IRES function is suggested by the fact that mutations of one side of the stem reduced PTB binding and compromised internal initiation, and compensating mutations designed to restore the base-pairing of the stem rescued both functions [59, 66]. However, when HeLa cell extracts were immunodepleted of PTB, the translation of IRES-driven mRNAs but not of globin mRNA was inhibited, yet IRES function could not be recovered by addition of recombinant PTB [66].

Using a dicistronic mRNA with the rhinovirus 5′-UTR in the intercistronic spacer and carrying out translation assays in the reticulocyte lysate, we find that the HeLa cell activity which specifically stimulates downstream cistron translation fractionates into two components [65]. Each is active on its own and, when tested in combination, their effects are additive rather than synergistic. One component co-purifies precisely with PTB, but it is not yet clear whether the active entity is PTB itself or a PTB/PSF complex. The other co-purifies precisely with a 97 kDa protein cross-linkable to the HRV and poliovirus IRES, but not to the EMCV IRES, and present in very much greater abundance in HeLa cell extracts than in reticulocyte lysates [65]. p97 seems to be present in a high molecular weight complex which possibly includes a minor fraction of the PTB present in HeLa cell extracts [65].

Obviously there is still considerable confusion over what trans-acting proteins are needed for internal initiation of translation of picornavirus RNAs, but the work is at a position where one can expect a clearer picture to emerge soon. There is a strong suspicion that PTB is involved in some form, but as recombinant PTB couldn't rescue the ability of immunodepleted extracts to translate IRES-driven mRNAs, the active entity may be a modified form of PTB, or one particular isoform of this protein (there are at least four such isoforms generated by alternative splicing [67, 68]), or a complex of PTB with other proteins.

Coda

Although internal initiation of translation of all picornavirus genomes can be described by a single model with only slight variations between different species, it is far from obvious that this is applicable to other cases of internal initiation, of which the best characterised examples are: GRP 78 mRNA (the 78 kDa glucose regulated protein, also known as immunoglobulin heavy chain binding protein), the *antennapedia* mRNA of *Drosophila*, a coronavirus mRNA, and hepatitis C virus RNA [70–73]. None of these mRNAs has IRES sequences showing any resemblance to any picornavirus IRES. However, perhaps this should not be considered so surprising, since the IRESes of the three groups of picornaviruses show little resemblance to each other at first sight, and the model shown in Fig. 1 carries the implication that a functional IRES could be generated from many quite different primary sequences. It is extraordinary enough that two different modes of initiation should have evolved (internal initiation and the scanning ribosome mechanism), and it would seem beyond credibility that there should be two or more entirely different mechanisms of internal ribosome entry. The challenge for the future is to identify the features common to internal initiation on both picornaviral and non-picornaviral examples, and also to identify the way in which internal initiation undoubtedly shares features in common with the scanning ribosome mechanism.

Acknowledgements

We would like to thank former members of our group and our collaborators for all their numerous inputs which have been instrumental in developing the ideas put forward here: Andy Borman, John Hershey, Michael Howell, Kathie Kean, Carola Lempke, Anne MacBride, Sue Milburn and Bert Semler. We also thank Elizabeth Hoey, Leena Kinnunen, Nick Knowles, Tuija Poyry and Glyn Stanway for supplying unpublished picornavirus 5′-UTR sequences. Work carried out in the authors' laboratory was supported by a grant from the Wellcome Trust.

References

1. Rotbart HA (1989) In: Semler BL & Ehrenfeld E (Eds) Molecular Aspects of Picornavirus Infection and Detection (pp 243–264) American Society for Microbiology, Washington, DC
2. Agol VI (1991) Adv. Virus Res. 40: 103–180
3. Poyry T, Kinnunen L & Hovi T (1992) J. Virol. 66: 5313–5319
4. Kozak M (1989) J. Cell Biol. 108: 229–241
5. Kozak M (1992) Crit. Rev. Biochem. Mol. Biol. 27: 385–402
6. Pelletier J & Sonenberg N (1988) Nature 334: 320–325
7. Jang SK, Krausslich H-G, Nicklin MJH, Duke GM, Palmenberg AC & Wimmer E (1988) J. Virol. 62: 2636–2643
8. Howell MT, Kaminski A & Jackson RJ (1990) In: Brinton M & Heinz FX (Eds) New Aspects of Positive Strand RNA Viruses (pp 144–151) American Society for Microbiology, Washington, DC
9. Kaminski A, Howell MT & Jackson RJ (1990) EMBO J. 9: 3753–3759
10. Kaminski A, Hunt SL, Gibbs CL & Jackson RJ (1994) In: Setlow J. (Ed) Genetic Engineering, Principles and Methods (Vol. 16, in press) Plenum Press, New York
11. Jackson RJ, Howell MT & Kaminski A (1990) Trends Biochem. Sciences 15: 477–483
12. Hyypia T, Horsnell C, Maaronen M, Khan M, Kalkkinen N, Auvinen P, Kinnunen L & Stanway G (1992) Proc. Natl. Acad. Sci. USA 89: 8847–8851
13. Meerovitch K & Sonenberg N (1993) Seminars in Virology 4: 217–227
14. Palmenberg AC, Kirby EM, Janda MR, Drake NL, Duke GM, Potratz KF & Collett MS (1984) Nucleic Acids Res. 12: 2969–2985
15. Duke GM, Hoffman MA & Palmenberg AC (1992) J. Virol. 66: 1602–1609
16. Sangar DV, Newton SE, Rowlands DJ & Clarke BE (1987) Nucleic Acids Res. 15: 3305–3315
17. Pritchard AE, Calenoff MA, Simpson S, Jensen K & Lipton HL (1992) J. Virol. 66: 1951–1958
18. Tesar M, Harmon SA, Summers DF & Ehrenfeld E (1992) Virology 186: 609–618
19. Nicholson R, Pelletier J, Le SY & Sonenberg N (1991) J. Virol. 65: 5886–5894
20. Borman A & Jackson RJ (1993) Virology 188: 685–696
21. Kuge S & Nomoto A (1987) J. Virol. 61: 1478–1487
22. Iizuka N, Kohara M, Hagino-Yamagishi K, Abe S,

Komatsu T, Tago K, Arita M & Nomoto A (1989) J. Virol. 63: 5354–5363
23. Iizuka N, Yonekawa H & Nomoto A (1991) J. Virol. 65: 4867–4873
24. Haller AA & Semler BL (1992) J. Virol. 66: 5075–5086
25. Pilipenko EV, Gmyl AP, Maslova SV, Svitkin YV, Sinyakov AN & Agol VI (1992) Cell 68: 119–131
26. Toyoda H, Kohara M, Kataoka Y, Suganuma T, Omata T, Imura N & Nomoto A (1984) J. Mol. Biol. 174: 561–585
27. Ryan MD, Jenkins O, Hughes PJ, Brown A, Knowles NJ, Booth D, Minor PD & Almond JW (1990) J. Gen. Virol. 71: 2291–2299
28. Pelletier J, Flynn ME, Kaplan G, Racaniello V & Sonenberg N (1988) J. Virol. 62: 4486–4492
29. Meerovitch K, Nicholson R & Sonenberg N (1991) J. Virol. 65: 5895–5901
30. Kuge S, Kawamura N & Nomoto A (1989) J. Virol. 63: 1069–1075
31. Kuge S, Kawamura N & Nomoto A (1989) J. Mol. Biol. 207: 175–182
32. Belsham GJ & Brangwyn JK (1990) J. Virol. 64: 5389–5395
33. Belsham GJ (1992) EMBO J. 11: 1105–1110
34. Kaminski A, Belsham GJ & Jackson RJ (1994) EMBO J. 13: 1673–1681
35. Kuhn R, Luz N & Beck E (1990) J. Virol. 64: 4625–4631
36. Pestova TV, Hellen CUT & Wimmer E (1991) J. Virol. 65: 6194–6204
37. Trono D, Andino R & Baltimore D (1988) J. Virol. 62: 2291–2299
38. Trono D, Pelletier J, Sonenberg N & Baltimore D (1988) Science 241: 445–448
39. Dildine SL & Semler BL (1989) J. Virol. 63: 847–862
40. Rivera VM, Welsh D & Maizel JV (1988) Virology 165: 42–50
41. Pilipenko EV, Blinov VM, Chernov BK, Dimitrieva TM & Agol VI (1989) Nucleic Acids Res. 17: 5701–5711
42. Pilipenko EV, Blinov VM, Romanova LI, Sinyakov AN, Maslova SV & Agol VI (1989) Virology 168: 201–209
43. Skinner MA, Racaniello VR, Dunn G, Cooper J, Minor PD & Almond JW (1989) J. Mol. Biol. 207: 379–392
44. Le SY & Zuker M (1990) J. Mol. Biol. 216: 729–741
45. Staehelin T, Trachsel H, Erni B, Boschetti A & Schreier MH (1975) Proceedings 10th FEBS Meeting 39: 309–323
46. Rozen F, Edery I, Meerovitch K, Dever TE, Merrick WC & Sonenberg N (1990) Mol. Cell. Biol. 10: 1134–1144
47. Jaramillo M, Dever TE, Merrick WC & Sonenberg N (1991) Mol. Cell. Biol. 11: 5992–5997
48. Brown B & Ehrenfeld E (1979) Virology 97: 396–405
49. Dorner AJ, Semler BL, Jackson RJ, Hanecak R, Duprey E & Wimmer E (1984) J. Virol. 50: 507–514
50. Phillips BA & Emmert A (1986) Virology 148: 255–267
51. Svitkin YV, Pestova TV, Maslova SV & Agol VI (1988) Virology 166: 394–404
52. Pelletier J, Kaplan G, Racaniello VR & Sonenberg N (1988) J. Virol. 62: 2219–2227
53. Glass MJ & Summers DF (1993) Virology 193: 1047–1050
54. Najita L & Sarnow P (1990) Proc. Natl. Acad. Sci. USA 87: 5846–5850
55. Earle JAP, Skuce RA, Flemming CS, Hoey EM & Martin SJ (1988) J. Gen. Virol. 69: 253–263
56. Dildine SL & Semler BL (1992) J. Virol. 66: 4364–4376
57. Meerovitch K, Pelletier J & Sonenberg N (1989) Genes Dev. 3: 1026–1034
58. Meerovitch K, Svitkin YV, Lee HS, Lejbkowicz F, Kenan DJ, Chan EKL, Agol VI, Keene JD & Sonenberg N (1993) J. Virol. 67: 3798–3807
59. Jang SK & Wimmer E (1990) Genes Dev. 4: 1560–1572
60. Luz N & Beck E (1990) FEBS Lett. 269: 311–314
61. Borovjagin AV, Evstafieva AG, Ugarova TY & Shatsky IN (1990) FEBS Lett. 261: 237–240
62. Borovjagin AV, Ezrokhi MV, Rostapshov VM, Ugarova TY, Bystrova TF & Shatsky IN (1991) Nucleic Acids Res. 19: 4999–5005
63. Luz N & Beck E (1991) J. Virol. 65: 6486–6494
64. Gebhard JR & Ehrenfeld E (1992) J. Virol. 66: 3101–3109
65. Borman A, Howell MT, Patton JG & Jackson RJ (1993) J. Gen. Virol. 74: 1775–1788
66. Hellen CUT, Witherell GW, Schmid M, Shin SH, Pestova TV, Gil A & Wimmer E (1993) Proc. Natl. Acad. Sci. USA 90: 7642–7646
67. Patton JG, Mayer SA, Tempst P & Nadal-Ginard B (1991) Genes Dev. 5: 1237–1251
68. Gil A, Sharp PA, Jamison SJ & Garcia-Blanco MA (1991) Genes Dev. 5: 1224–1236
69. Patton JG, Porro EB, Galceran J, Tempst P & Nadal-Ginard B (1993) Genes Dev. 7: 393–406
70. Macejak DG & Sarnow P (1991) Nature 353: 90–94
71. OH S-K, Scott MP & Sarnow P (1992) Genes Dev. 6: 1643–1653
72. Liu DX & Inglis SC (1992) J. Virol. 66: 6143–6154
73. Wang C, Sarnow P & Siddiqui A (1993) J. Virol. 67: 3338–3344