The EMBO Journal vol.14 no.23 pp.6010–6020, 1995

Unique features of internal initiation of hepatitis C virus RNA translation

Joanna E.Reynolds, Ann Kaminski, Hilkka J.Kettinen, Ken Grace[1], Berwyn E.Clarke[1], Anthony R.Carroll[1], David J.Rowlands[1] and Richard J.Jackson[2]

Department of Biochemistry, Tennis Court Road, University of Cambridge, Cambridge CB2 1QW and [1]Biology Division, Wellcome Research Laboratories, Langley Court, Beckenham BR3 3BS, UK

[2]Corresponding author

The question of whether hepatitis C virus (HCV) RNA is translated by a mechanism of internal ribosome entry has been examined by testing whether insertion of HCV sequences between the two cistrons of a dicistronic mRNA promotes translation of the downstream cistron in rabbit reticulocyte lysates. Deletion analysis showed that efficient internal initiation required a segment of the HCV genome extending from about nucleotides 40–370 and that deletions from the 3′-end of this element were highly deleterious. As the authentic initiation codon for HCV polyprotein synthesis is at nucleotide 342, this demonstrates that, besides 5′-UTR sequences, a short length of HCV coding sequences is required for internal initiation. This finding was confirmed in transfection assays of BT7-H cells and was shown to be independent of the nature of the downstream reporter cistron. The strong requirement for coding sequences is in sharp contrast to internal initiation of picornavirus RNA translation. As a probable correlate with this, it was also found that the efficiency of internal initiation was only marginally compromised when the authentic initiation codon was mutated to a non-AUG codon, again in sharp contrast with the picornaviruses. The finding that coding sequences are required for internal initiation has important implications for the design of experiments to test for internal initiation of translation of cellular mRNAs.

Keywords: hepatitis C virus/internal ribosome entry/translation initiation

Introduction

Hepatitis C virus (HCV) is the major causative agent of non-A, non-B hepatitis worldwide, being transmitted principally by blood and blood-derived products (Choo *et al.*, 1989). Infection can frequently lead to chronic hepatitis or cirrhosis and is associated with the development of hepatocellular carcinoma (Saito *et al.*, 1990). HCV has a single-stranded RNA genome of positive polarity, ~9450 nucleotides in length, with a single large open reading frame (ORF). This encodes a polyprotein of ~3010 amino acid residues (variations depending on the isolate) which is cleaved by virus-encoded and cellular proteases to yield the various viral capsid and non-structural proteins. Although the order of the viral structural and non-structural proteins within the polyprotein resembles that found in flaviviruses and pestiviruses (Choo *et al.*, 1991; Takamizawa *et al.*, 1991), the 5′-untranslated region (5′-UTR) of HCV is much longer than that of the flaviviruses and more closely resembles the 5′-UTR of pestiviruses in terms of length and putative secondary structure (Han *et al.*, 1991; Brown *et al.*, 1992). The sequence of the HCV 5′-UTR is strongly conserved amongst different strains and isolates (Bukh *et al.*, 1992), much more conserved than any other region of the viral genome, a fact which has been exploited in diagnostic methods based on PCR techniques. It is 341 nucleotides long, containing three to five AUG triplets (depending on the individual strain or isolate) and has the potential to form extensive secondary structure (Brown *et al.*, 1992; Tsukiyama-Kohara *et al.*, 1992). In these respects the HCV 5′-UTR shares many properties with picornavirus 5′-UTRs.

It is now well established that initiation of translation of picornavirus RNAs is by a mechanism of internal ribosome entry, rather than by ribosome scanning from the 5′-end of the RNA (Jackson *et al.*, 1990, 1994). Internal initiation requires a defined segment of the picornavirus 5′-UTR which is generally known as the IRES (for 'internal ribosome entry segment') and is ~450 nucleotides long, but does not include the extreme 5′-proximal sequences of the viral genome. Insertion of a picornavirus IRES between the two cistrons of a laboratory-generated dicistronic construct results in a very significant stimulation of translation of the downstream cistron both *in vitro* and *in vivo*, to the extent that the yield of downstream cistron translation product may exceed that from the upstream cistron (Jang *et al.*, 1988; Pelletier and Sonenberg, 1988). This stimulation is independent of whether the dicistronic mRNA is capped or not and controls show that it cannot be explained by cleavage of the dicistronic RNA into two monocistronic mRNAs. Thus the unambiguous conclusion is that the picornavirus RNA segment promotes direct internal ribosome entry to initiate translation of the downstream cistron.

The common features of the 5′-UTRs of HCV and the picornaviruses have prompted the question of whether HCV RNA is also translated by internal initiation, but the results reported so far have been controversial. Using dicistronic constructs in which the downstream cistron was the first part of the HCV polyprotein coding sequence located in the same relative position to the HCV 5′-UTR as in the viral genome itself, Tsukiyama-Kohara *et al.* (1992) showed by *in vitro* translation in rabbit reticulocyte lysates or HeLa cell extracts that the HCV genome has

a quite efficient IRES. On the other hand, in a very comprehensive set of experiments involving both *in vitro* and *in vivo* assays, Yoo *et al.* (1992) could find no evidence for internal initiation, though in this case the dicistronic constructs used standard reporter cistrons and a linker was placed between the 3′-end of the HCV 5′-UTR segment and the initiation codon of the downstream cistron. More recently, several groups have confirmed the existence of an HCV IRES (Wang *et al.*, 1993; Fukushi *et al.*, 1994; Rijnbrand *et al.*, 1995) and are all agreed that deletions in the 3′-end of the HCV 5′-UTR abrogate IRES function. However, there is still some controversy as to the exact boundaries of the IRES and whether the IRES lies entirely within the 5′-UTR. While some groups have used standard reporters with no HCV coding sequences retained (Wang *et al.*, 1993), others have retained a short length of viral coding sequences (Rijnbrand *et al.*, 1995) and yet others have used HCV core protein coding sequences as the downstream reporter cistron of the dicistronic constructs (Tsukiyama-Kohara *et al.*, 1992; Fukushi *et al.*, 1994).

We have examined these questions and have paid particular attention to the possible requirement for HCV coding sequences for efficient internal initiation of translation. We show here that the HCV genome does indeed have an efficient IRES, but that quite unlike any picornavirus IRES, the 5′-proximal HCV polyprotein coding sequences are required for internal ribosome entry. These findings have important implications for the design of experiments to test for IRESs in cellular mRNAs and other species of positive strand RNA viruses.

Results

Strategy for testing for an HCV IRES and for mapping the 3′ boundary

In order to test the ability of HCV sequences to direct internal initiation, a standard dicistronic mRNA translation assay was used. The constructions followed the design previously used in this laboratory to study the IRESs of human rhinovirus type-2 and Theiler's murine encephalomyelitis virus (Borman and Jackson, 1992; Hunt *et al.*, 1993). The upstream cistron codes for *Xenopus laevis* cyclin B2 and the downstream cistron is a slightly truncated form of the influenza virus non-structural protein (NS) coding region as described previously (Borman and Jackson, 1992). The truncated form of this viral protein is designated as NS′. The control dicistronic plasmid, previously named pXLJ Con (Borman and Jackson, 1992), but renamed pXLJ0 for this work, has a short length of polylinker sequences between the two cistrons, which, together with the *Bam*HI site immediately after the AUG initiation codon of the NS′ cistron, facilitate the insertion of putative IRES sequences (Figure 1).

The starting point for testing whether there is an IRES within the 5′-proximal part of the HCV genome was a cDNA clone spanning nucleotides 40–405, according to the sequence of Han *et al.* (1991), which is now considered to be the complete sequence starting from the very 5′-end of the viral RNA. Reconstruction of the 5′-proximal 40 nucleotides of the HCV genome sequence was found to have no influence on internal initiation (data not shown), in agreement with the conclusions of Wang *et al.* (1993) regarding the 5′-boundary of the IRES, and thus most of the constructs used in this work started at nucleotide 40, except those for the experiment shown in Figure 5. As the authentic initiation codon is at nucleotide 342 in this sequence, our starting clone included 64 nucleotides of HCV coding sequences. Deletions were made from the 3′-end of these HCV sequences and the truncated segment incorporated first into a monocistronic construct with a downstream in-frame NS′ cistron, and thence into the dicistronic test plasmid (Figure 1). For the majority of these constructs, the sequence immediately downstream of the HCV sequences is ..CC**AUG**GAUCC.., where the AUG shown in bold is the initiation codon of the actual NS′ ORF, and the two C residues function as a linker between the HCV sequences and the NS′ coding sequences. In the case of pXL40-395.NS′, the fusion between HCV and NS′ sequences had to be made at the in-filled *Bam*HI site in order to maintain the correct reading frame and thus the two C residues and the AUG codon are missing (Figure 1). As the HCV initiation codon is in-frame with the NS′ ORF in all constructs, initiation at this HCV initiation codon should lead to the synthesis of an NS′ protein extended at the N-terminus by 3, 4, 10 and 17 amino acid residues of HCV core protein sequences respectively for XL40-351.NS′, XL40-354.NS′, XL40-372.NS′ and XL40-395.NS′ mRNAs (Figure 1), under the assumption that the initiating methionine residue will be cleaved off as a consequence of the neighbouring serine (Sherman *et al.*, 1985, 1993). In construct pXL40-339.NS′, the 3′ end-point of the HCV segment was such that the NS′ initiation codon was in the equivalent position to the authentic viral initiation codon (Figure 1) and so the product of translation of this transcript would be full length NS′ with no N-terminal extension. It should also be noted that this construct in fact has precisely the same sequence upstream of the initiation codon as is found in the HCV genome: the linker of two C residues brought in with the NS′ cDNA sequences fortuitously making an exact replacement of two C residues deleted from the HCV sequences by *Bal*31 treatment (Figure 1).

All constructs were made in a pGEM-2 vector, thus allowing capped and uncapped transcripts to be prepared by transcription with bacteriophage T7 RNA polymerase. These transcripts were translated in rabbit reticulocyte lysates in the presence of [^{35}S]methionine over a range of RNA concentrations and the relative yields of cyclin from the upstream cistron and NS′ from the downstream cistron were assessed by autoradiography following gel electrophoresis of the translation assay samples.

It was found that for uncapped RNAs, the yield of translation product from the upstream cistron (*Xenopus laevis* cyclin B2) was relatively constant for all constructs at any given RNA concentration (Figure 2). The yield of cyclin was generally stimulated by capping the transcript (Figure 2), although the degree of stimulation varied between the different transcripts, possibly reflecting differences in the degree of carry-over of m^{7}GpppG cap analogue from the transcription reaction.

In the case of the two transcripts with the longest lengths of HCV coding sequences (pXL40-395.NS′ and pXL40-372.NS′), there was efficient translation of the downstream cistron, indicative of internal initiation of translation. Consistent with this interpretation, the yield

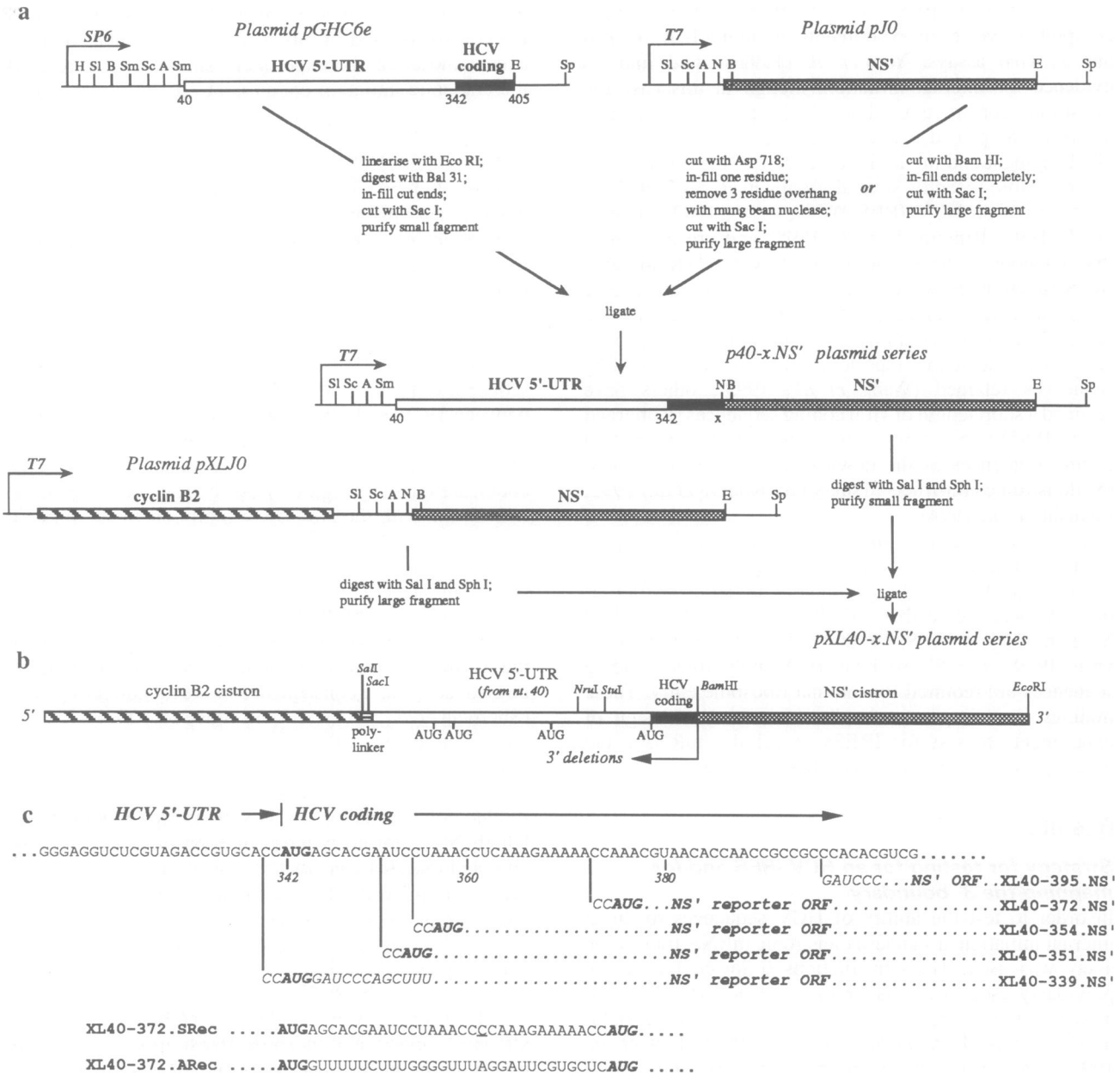

Fig. 1. Schematic diagram, not to scale, of the constructs used to test for an HCV IRES and to map its boundaries. In (**a**) the cloning strategy is depicted: the HCV 5′-UTR segment shown as an open bar; the start of the HCV coding sequences as a black bar; and linker sequences depicted by thin lines. Restriction enzyme sites are represented as follows: *Asp*718 (A), *Bam*HI (B), *Eco*RI (E), *Hin*dIII (H), *Nco*I (N), *Sac*I (Sc), *Sal*I (Sl), *Sma*I (Sm), *Sph*I (Sp). Plasmids pJ0 and pXLJ0, previously named pXLJ Con, are described in Borman and Jackson (1992). The nomenclature of the plasmids takes the form p40-x.NS′ for monocistronic constructs with HCV sequences extending from nucleotide 40 to residue x numbered according to the sequence of Han *et al.* (1991), and pXL40-x.NS′ for the corresponding dicistronic construct. In (**b**) a more detailed schematic diagram of the dicistronic constructs is shown, with the approximate locations of the three AUG triplets in the HCV 5′-UTR segment from nucleotide 40, in addition to the authentic HCV initiation codon. The *Nru*I, *Stu*I and *Bam*HI sites used for the subcloning steps described in the text are shown. In (**c**) is shown the sequence at the junction of the HCV and NS′ sequences in the various dicistronic constructs bearing 3′ deletions of the HCV segment. The authentic HCV initiation codon is shown in bold, as is the AUG codon at the start of the NS′ ORF. The sequences of the 'sense' and 'antisense' reconstructions of pXL40-372.NS′ (pXL40-372.SRec and pXL40-372.ARec, respectively) are given between the authentic initiation codon and the start of the NS′ coding sequences, and the single site mutation necessary to eliminate an in-frame stop codon in the antisense orientation is underlined in the pXL40-372.SRec sequence.

of NS′ polypeptide from these two RNAs was not significantly affected by capping the transcripts, although the yield of cyclin from these two mRNAs was quite strongly influenced by capping (Figure 2). The efficiency of the internal initiation is quite high, given that cyclin B2 has 17 methionine residues and the downstream cistron translation product would have nine in the case of pXL40-372.NS′ or 8 for pXL40-395.NS′, again assuming that the initiating methionine residue would be cleaved off as a consequence of the neighbouring serine (Sherman *et al.*, 1985, 1993).

The most striking point, however, is that translation from the downstream cistron was very much less efficient with the transcripts of the other dicistronic constructs.

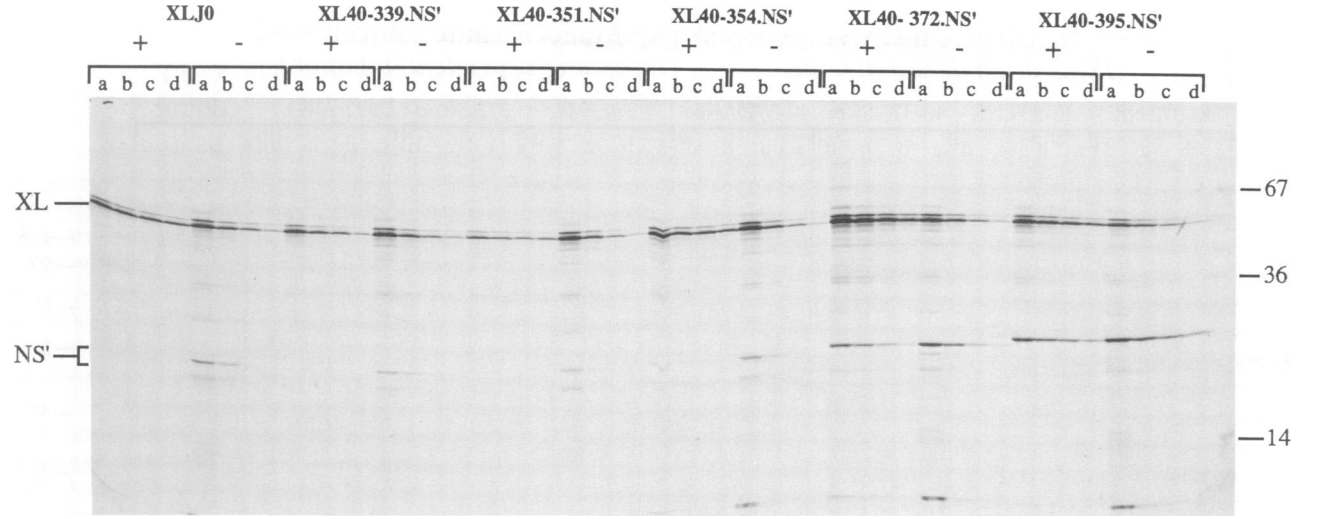

Fig. 2. Mapping the 3′ boundary of the HCV IRES by *in vitro* translation of 3′ deletion mutants. Capped (+) and uncapped (–) transcripts of pXLJ0, pXL40-339.NS′, pXL40-351.NS′, pXL40-354.NS′, pXL40-372.NS′ and pXL40-395.NS′ were translated in rabbit reticulocyte lysate at final concentrations of: (a) 50; (b) 25; (c) 12.5 and (d) 6.25 μg/ml. Translation products were resolved by gel electrophoresis using a 20% polyacrylamide gel and the resulting autoradiograph is shown. XL indicates the cyclin B2 cistron translation product and NS′ indicates products derived from the downstream cistron. The size of molecular weight markers is shown on the right in kDa.

With XL40-339.NS′ and XL40-351.NS′ mRNAs the yield of the NS′ derived product was no greater than the negative control, XLJ0 (Figure 2), whereas for XL40-354.NS′ it was very slightly above this background level. Thus the 3′ boundary of the HCV IRES lies between nucleotides 354 and 372, which means that, given that the authentic initiation codon is at nucleotide 342, a short length of the HCV polyprotein coding sequence is required for internal ribosome entry, unlike the case with any picornavirus IRES investigated to date.

In vivo expression studies

In view of the unprecedented nature of these results, we sought to confirm them by *in vivo* transfection assays, even though it would be very surprising and quite unparalleled if the requirements for internal initiation *in vivo* were less stringent than *in vitro*. The same plasmid constructs were used as for the *in vitro* translations and these were transfected into BT7-H cells (Whetter *et al.*, 1994) which had been infected with a recombinant vaccinia virus that expresses bacteriophage T7 RNA polymerase (Fuerst *et al.*, 1986). Subsequently, the cells were incubated with [^{35}S]methionine and the labelled proteins were analysed by one-dimensional gel electrophoresis (Figure 3). With the control plasmid, pXLJ0, the labelled cyclin product was clearly visible against the background of labelled vaccinia encoded proteins (compare lanes 6 and 7 of Figure 3) and transfections with various plasmids of the pXL40-x.NS′ series all gave a very similar yield of cyclin, confirming that transfection efficiencies were consistent and reproducible. Labelled NS′ resulting from translation of the downstream cistron could clearly be seen in the case of transfections with pXL40-395.NS′ and pXL40-372.NS′ (lanes 1 and 2 of Figure 3). On the other hand it was barely visible in the case of pXL40-339.NS′ (lane 5) and any labelled NS′ in lanes 3 and 4 was at such low yield that it is undetectable against the background of vaccinia encoded proteins. Immunoprecipitation assays verified the identity of the cyclin and NS′ products, and

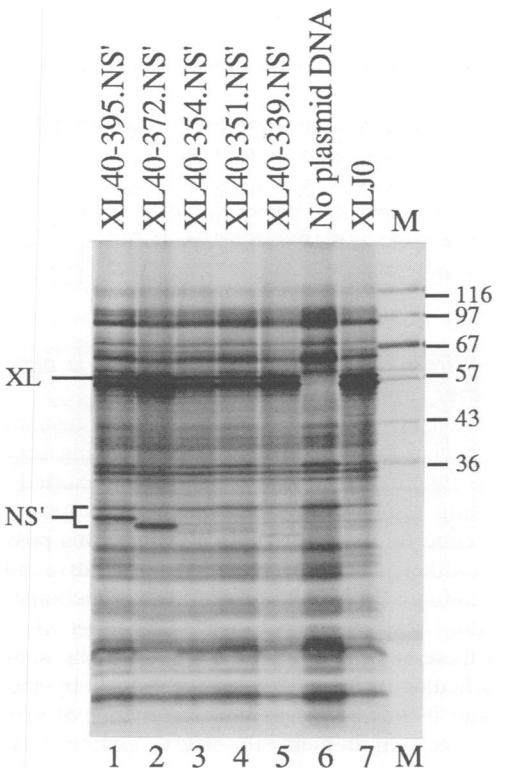

Fig. 3. Mapping the 3′ boundary of the HCV IRES by *in vivo* transfection assays. BT7-H cells were infected with recombinant vaccinia virus vTF7-3 and were transfected 60 min later with 5 μg of the designated plasmid DNAs. Metabolic labelling with [^{35}S]methionine was carried out for a 2 h period starting 18 h post-transfection and then the cells were processed for gel electrophoresis using a 20% polyacrylamide gel. The resulting autoradiograph is shown. The positions of the molecular weight standards (in kDa) are shown in the right hand margin and the positions of the cyclin (XL) and NS′ translation products in the left margin.

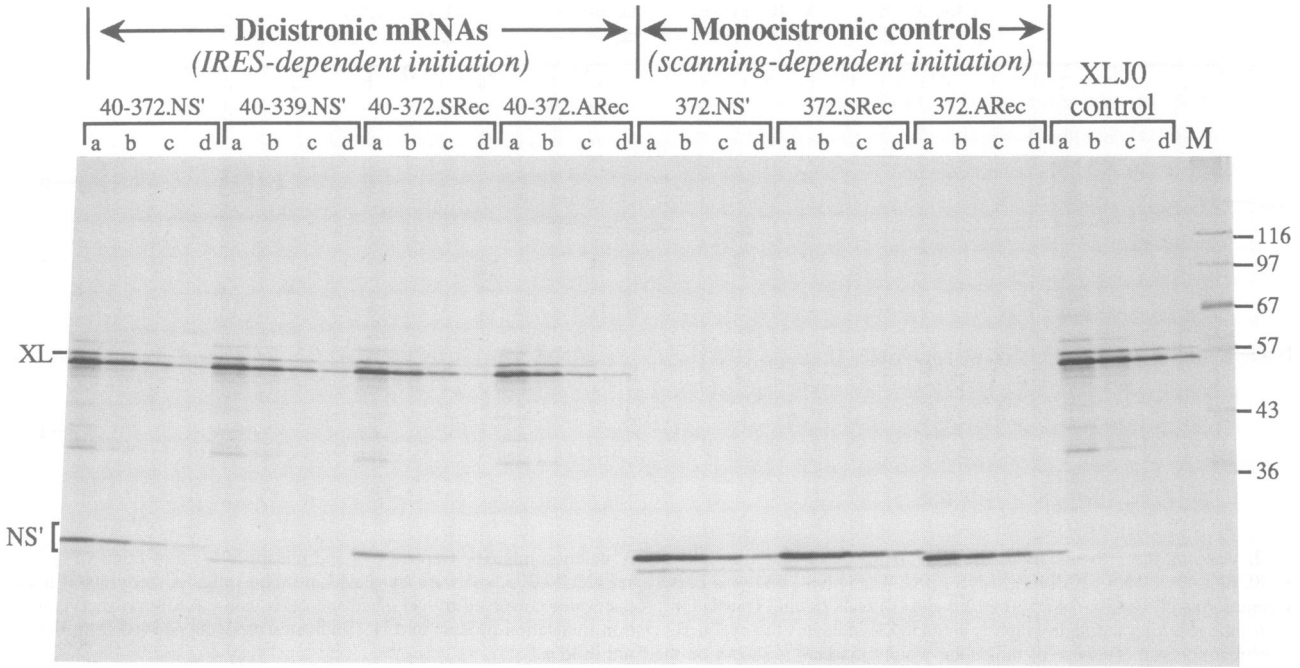

Fig. 4. Internal initiation requires HCV coding sequences *per se* and not just as a spacer. Capped transcripts of the designated constructs were translated in rabbit reticulocyte lysate at final concentrations of: (a) 50; (b) 25; (c) 12.5 and (d) 6.25 μg/ml. Translation products were resolved by gel electrophoresis using a 20% polyacrylamide gel and the resulting autoradiograph is shown. XL indicates the cyclin B2 cistron translation product and NS′ indicates products derived from the downstream cistron. The size of molecular weight markers is shown on the right in kDa.

confirmed that only pXL40-395.NS′ and pXL40-372.NS′ gave a significant yield of NS′ (data not shown). Thus the results of the transfection assays precisely parallel the *in vitro* data: efficient IRES function is observed with pXL40-395.NS′ and pXL40-372.NS′, but not with any construct that retained fewer codons of HCV coding sequence.

The requirement for coding sequences is also seen with other reporter cistrons

We next considered the possibility that these results reflect not so much a positive effect of HCV coding sequences on IRES function as a negative influence exerted by the NS′ coding sequences if they are brought too close to the authentic initiation site. Nevertheless, this possibility seemed unlikely since we have found no negative influence of NS′ coding sequences adjacent to the functional initiation codon of picornavirus IRESs (Hunt *et al.*, 1993). Thus if there is an inhibitory effect of the NS′ sequences it is particular to internal initiation driven by the HCV IRES, and it thus becomes almost a matter of semantics whether the requirement for HCV coding sequences reflects a positive contribution of such sequences or the absence of a negative influence.

As a first step towards examining this question, we took pXL40-339.NS′, which has its initiation codon in the normal position as in the HCV genome but lacks any HCV coding sequences, and reconstructed the active pXL40-372.NS′ from it using synthetic oligonucleotides. The reconstruction strategy was such that the oligonucleotides could be inserted in either orientation. In one orientation the authentic HCV coding sequences would be regenerated (with a single site mutation) and in the other a quite different sequence would be generated (Figure 1). But in both orientations the NS′ coding sequences would be at the same distance from the authentic initiation site. The single mutation was necessary to eliminate an in-frame stop codon in the antisense orientation. These two reconstructions are referred to as pXL40-372.SRec and pXL40-372.ARec (for sense and antisense reconstruction, respectively).

The results of *in vitro* translation assays of capped RNAs show that the translation of the IRES-dependent downstream cistron of XL40-372.SRec was almost as efficient as that of XL.40-372.NS′ itself (Figure 4). The small difference in efficiency may possibly be due to the single point mutation by which pXL40-372.SRec differs from pXL40-372.NS′. What is most striking, however, is the fact that translation of XL40-372.ARec gave virtually no NS′ product whatsoever and IRES-dependent translation with this dicistronic mRNA was, if anything, even less efficient than with XL40-339.NS′, the parent for the two reconstructions (Figure 4).

As a control to show that the translation of the NS′ cistron of XL40-372.ARec was defective only in the context of IRES-dependent initiation, we deleted the upstream cyclin cistron as well as most of the HCV 5′-UTR up to the unique *Stu*I site (Figure 1) from each of these constructs, to generate monocistronic constructs which would be translated by the normal 5′-end dependent scanning mechanism. *In vitro* transcription of these monocistronic constructs would generate an mRNA with a 5′-UTR of 80 nucleotides length, the first 17 nucleotides being polylinker sequences followed by nucleotides 279–341 of the HCV genome, numbered according to the sequence of Han *et al.* (1991). When capped transcripts of these constructs were translated *in vitro* the yield of NS′ product from the monocistronic derivative 372.ARec was as high as from the equivalent derivatives of 372.SRec and 372.NS′ (Figure 4). (The greater gel mobility of the

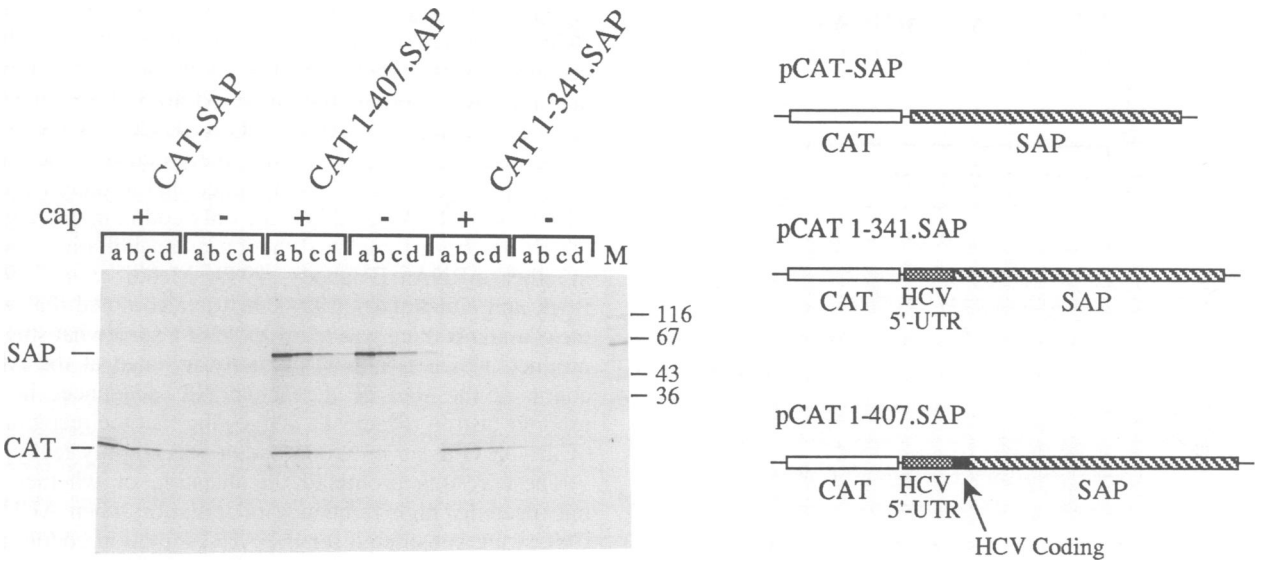

Fig. 5. The requirement for HCV coding sequences for internal initiation is seen with other reporter cistrons. Capped (+) and uncapped (–) transcripts, as indicated, of the designated constructs were translated in rabbit reticulocyte lysate at final concentrations of: (a) 50; (b) 25; (c) 12.5; and (d) 6.25 μg/ml. Translation products were resolved by gel electrophoresis using a 20% polyacrylamide gel, and the resulting autoradiograph is shown. The size of molecular weight markers is shown on the right in kDa. CAT indicates the chloramphenicol acetyltransferase product from the upstream cistron and SAP the secreted alkaline phosphatase product from the downstream cistron, as illustrated by the schematic diagram, approximately to scale, of the constructs: polylinker and UTR sequences are denoted by thin lines; the CAT cistron by an open rectangle; the SAP cistron by the cross-hatched rectangle; the complete HCV 5′-UTR by the stippled rectangle; and the first 22 codons of HCV core protein coding sequences by the filled black rectangle.

product from 372.ARec, which is also found with the very small amount of NS′ product from the dicistronic XL40-372.ARec, is probably due to the unusual amino acid composition at the N-terminus, since no abnormalities or unplanned additional mutations could be found on extensive sequencing of this construct.) We can therefore conclude that the translation of the NS′ cistron of XL40-372.ARec is defective only when such translation is by internal initiation. This allows the further conclusion that IRES function requires HCV coding sequences *per se* and not just as a spacer to separate the NS′ coding sequences by a sufficient length from the functional initiation codon.

In a completely different approach to this problem, a dicistronic construct was made with two other reporter cistrons: chloramphenicol acetyl transferase (CAT) was used as the upstream cistron and secreted alkaline phosphatase (SAP) in the downstream position. The constructs were made with the complete HCV 5′-UTR as the intercistronic spacer and included either no HCV coding sequences (pCAT 1-341.SAP) or the 5′-proximal 22 codons (pCAT 1-407.SAP). These were translated *in vitro* as either capped or uncapped mRNAs (Figure 5). As expected, the yield of CAT translated from the upstream cistron was significantly increased by capping the mRNAs, but was largely independent of whether the intercistronic spacer was polylinker sequences or the HCV 5′-UTR, or whether HCV coding sequences were retained. Strikingly, synthesis of SAP from the downstream cistron was observed only in the case of the construct which included HCV coding sequences and the yield was uninfluenced by capping the transcripts (Figure 5), a classical indicator of internal initiation. If no HCV coding sequences were present, the yield of SAP was equally low regardless of whether the complete HCV 5′-UTR or just polylinker sequences were present as intercistronic spacer.

The results of both the reconstruction experiment and the use of alternative reporter cistrons strongly indicate that HCV coding sequences are an essential element of the HCV IRES, and that the role of the coding sequences is not just to separate possible inhibitory sequences from close proximity to the functional initiation site.

The HCV IRES can promote internal initiation even at non-AUG codons

In a series of experiments to investigate the consequences of mutation of the AUG triplets in the HCV IRES, we came across another property unique to the HCV IRES. When the authentic initiation codon at nucleotide 342 of XL40-372.NS′ was mutated to AUU, we were surprised to find that the efficiency of internal initiation of downstream (NS′) cistron translation was only slightly compromised (Figure 6a). Under standard conditions we estimate the relative efficiency to be ~90%, but under extreme conditions of sub-optimal Mg^{2+}, or supra-optimal K^+ concentrations, the frequency of initiation was more severely reduced at AUU than at AUG (data not shown).

The surprisingly high efficiency of internal initiation at an AUU codon prompted an investigation of the properties of other mutant alleles of the authentic initiation codon. The substitution of a CUG was found to have only a minimal influence on efficiency of translation of the downstream cistron by internal initiation (Figure 6a). The efficiency of utilization of a CUG at this site was similar to that of an AUU. Mutation of the authentic initiation codon to AAG, GAG or GCG caused a significant decrease in translation efficiency, to ~40–50% with respect to the AUG parent (Figure 6a, lanes 5–7). However, on closer inspection of the autoradiographs it can be seen that the NS′ product translated from these mRNAs is marginally smaller than that obtained on translation of constructs

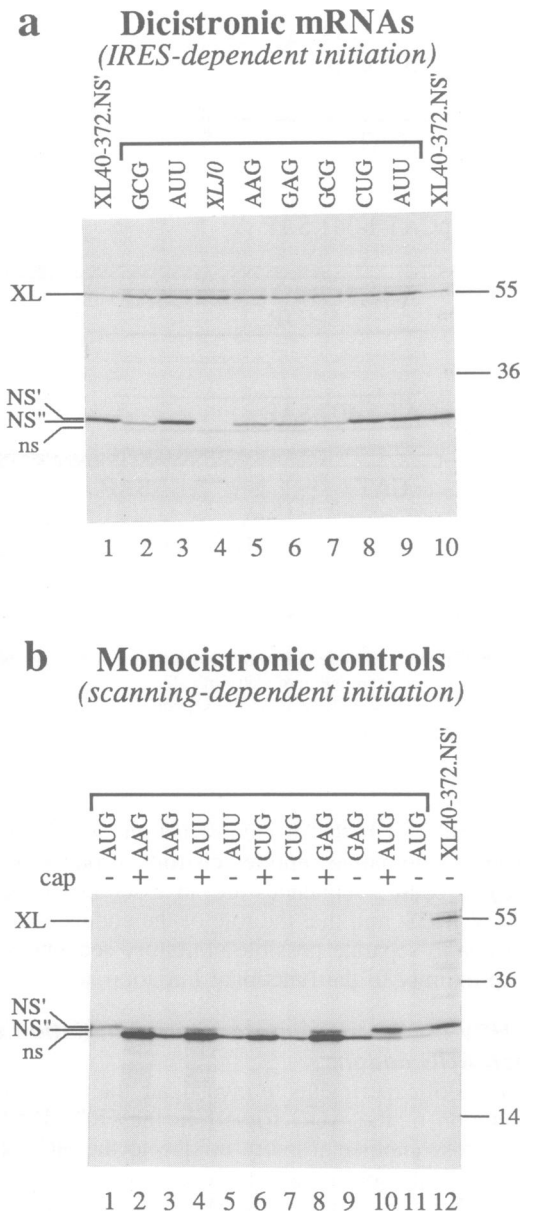

Fig. 6. Translation of (**a**) dicistronic mRNAs with different mutations of the authentic HCV initiation codon and (**b**) monocistronic derivatives of the same constructs with most of the HCV 5′-UTR sequences deleted. (a) Uncapped transcripts of dicistronic constructs pXL40-372.NS′, pXLAUU.NS′, pXLCUG.NS′, pXLAAG.NS′, pXLGAG.NS′, pXLGCG.NS′ and the control pXLJ0, were translated for 90 min in rabbit reticulocyte lysate at a final concentration of 50 μg/ml. (b) Capped (+) or uncapped (–) transcripts, as indicated, of the monocistronic (IRES-deleted) derivatives of these same mutants were translated for 90 min in rabbit reticulocyte lysate at a final concentration of 50 μg/ml and an uncapped transcript of the dicistronic pXL40-372.NS′ was also translated as a reference control. Translation products were resolved by gel electrophoresis using (a) a 10–30% polyacrylamide gradient gel or (b) a 20% polyacrylamide gel, and the resulting autoradiographs are shown. XL indicates the cyclin B2 cistron translation product, NS′ indicates products derived from the downstream cistron, NS″ the product thought to be initiated at the ACG codon (see text) and ns the product initiated at the AUG codon at the start of the NS′ coding sequences proper (Figure 1). The size of molecular weight markers is shown on the right in kDa.

with AUG, CUG or AUU at the position of the authentic initiation site (Figure 6a, compare lanes 2 and 3, and lanes 5–7 with lanes 8–10). Thus it would appear that mutation of the AUG to AAG, GAG or GCG has resulted in a switch in initiation site selection to another codon a short distance downstream in the HCV coding sequences. Given that the HCV coding region in XL40-372.NS′ mRNA has the sequence AUG.AGC.ACG.AAU.CCU.AAA.CCU.CAA.AGA.AAA it seems most likely that it is the ACG codon that is being used as the functional initiation codon with these mRNAs, as this is the only codon in this region which has been found to function as an initiation codon in other mRNAs (Peabody, 1989; Mehdi *et al.*, 1990; Böck and Kolakofsky, 1994). It was also noted that with these mutants there was a low yield of a somewhat smaller product, which is almost certainly initiated at the AUG codon at the start of the actual NS′ sequences in the reporter cistron (Figure 1), suggesting that the recognition of the ACG as an initiation codon was slightly leaky.

These results prompted the question of whether the unexpectedly high frequency of initiation when AUU or CUG were substituted for the AUG initiation codon, and the fairly high efficiency of utilization of the downstream ACG in the case of the other substitution mutants, was a characteristic of the internal initiation process promoted by the HCV IRES, or whether some local context features would allow efficient initiation at these sites even by scanning ribosomes. Accordingly, the upstream cyclin cistron and most of the HCV 5′-UTR were deleted to generate monocistronic constructs that would be translated by the scanning mechanism: the HCV 5′-UTR sequences were cut at a unique *Stu*I site and the downstream *Bam*HI site (Figure 1) and this short fragment was inserted into *Hin*cII- and *Bam*HI-cut monocistronic p40-372.NS′. When transcripts of these constructs were translated *in vitro*, it was found that capping the transcripts significantly stimulated translation, as would indeed be expected if these monocistronic mRNAs were translated by a scanning mechanism (Figure 6b). With the wild-type (AUG) sequence, most of the initiation occurred at the HCV polyprotein initiation codon, but the scanning was somewhat leaky and thus there was initiation at the next AUG codon, located at the start of the NS′ ORF proper (Figure 6b). This leakiness is not entirely unexpected given that the context of the HCV initiation codon is ACCAUGAGC, for although the upstream context is ideal (Kozak, 1986, 1989), the downstream AGC is unfavourable for initiation site recognition by scanning ribosomes (Grünert and Jackson, 1994).

In the case of the mRNAs in which the HCV initiation codon had been mutated, virtually all initiation was from the downstream AUG at the start of the NS′ coding sequences, but there was a low yield of a somewhat larger product, whose size suggests that it may have been initiated at the ACG codon rather than at the AUU, CUG, AAG or GAG codons substituting the normal HCV translation initiation codon (Figure 6b). According to the recently revised context rules (Grünert and Jackson, 1994), the context of the ACG is somewhat more favourable to recognition by scanning ribosomes than is the context of the authentic initiation site and thus the preference for initiation by scanning ribosomes at the ACG rather than the AUU, CUG, AAG or GAG is not wholly surprising. However, the important conclusion is that the high efficiency of utilization of the AUU or CUG codon, and the somewhat less frequent initiation at the ACG codon

(Figure 6), of the various mutant dicistronic mRNAs is a unique feature of IRES-driven internal initiation, which is not seen when initiation is by the scanning ribosome mechanism.

Discussion

The results presented here confirm that there is an IRES in the HCV genome capable of conferring internal initiation of translation of a downstream heterologous reporter cistron, in support of the conclusions of Tsukiyama-Kohara *et al.* (1992) and Wang *et al.* (1993), and in contradiction with Yoo *et al.* (1992). In retrospect, the failure of Yoo *et al.* (1992) to find evidence for an IRES can almost certainly be explained by the design of their constructs, which included an arbitrary spacer of ~30 nucleotides between the 3′-end of the HCV 5′-UTR and the initiation codon of the downstream reporter cistron. This construct design would certainly compromise the efficiency of internal initiation dependent on cardiovirus IRESs (Kaminski *et al.*, 1994a) and would seem to be even more deleterious in the case of HCV.

The most surprising and noteworthy of the results reported here is the finding that HCV coding sequences are required for internal initiation; in other words, the IRES extends into the coding region and its 3′ boundary lies some 12–30 nucleotides downstream of the authentic initiation codon. As the same 3′ boundary was observed in the *in vivo* transfection experiments, this cannot be an artefact of the rabbit reticulocyte lysate *in vitro* system and, indeed, it would have been almost unprecedented for the *in vitro* system to have proven to be more stringent in requirements for IRES activity than *in vivo*. Moreover, the requirement for HCV coding sequences is independent of the nature of the reporter cistron fused to the viral coding sequences, which strongly suggests that the HCV coding region is required *per se* and not just as a neutral spacer to separate the initiation codon from particular reporter coding sequences that might possibly be inhibitory if located immediately downstream of the AUG codon. In further support of this conclusion, it has been found that when the endogenous IRES in the poliovirus genome was replaced by the HCV IRES, a viable chimeric virus was only obtained if HCV core protein coding sequences were included in addition to the HCV 5′-UTR (H-h.Lu and E.Wimmer, submitted). Finally, in a search for a potential IRES in the genome of the related pestivirus, classical swine fever virus, we have noted a similar though slightly less stringent requirement for coding sequences: constructs which retain less than five codons of viral coding sequence allow an efficiency of internal initiation only 10–15% of that observed if 17 or more codons are retained (S.P.Fletcher and R.J.Jackson, unpublished observations).

In view of this overwhelming evidence that the HCV IRES extends into the coding region, it is pertinent to ask why Wang *et al.* (1993) obtained what appeared to be quite efficient internal initiation using only HCV 5′-UTR sequences. However, we note that there is considerable fortuitous homology between the nucleotide sequences at the 5′-end of the luciferase ORF and the HCV polyprotein ORF (Figure 7). If the requirement for HCV coding sequences in our hands is due to a need for base-pairing

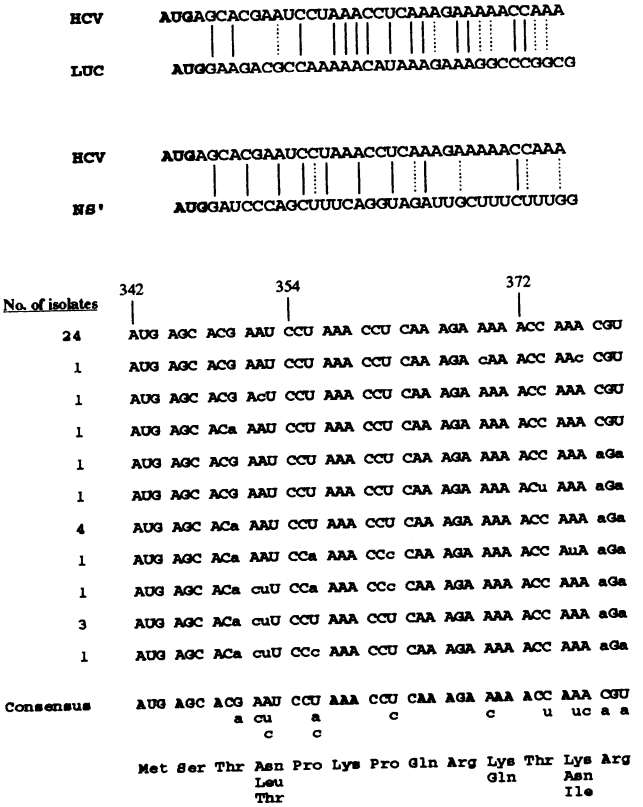

Fig. 7. Features of the start of the HCV coding sequences. In the top part of the figure the first 36 nucleotides of HCV coding sequences are separately aligned with the corresponding segment of the luciferase (LUC) and NS′ ORFs. Five different alignments were examined: the direct comparison and alignments displaced by up to two residues in either direction. The alignment shown, displaced by one residue, gave the best match between HCV and NS′ sequences of all five alignments, but only the second best match between HCV and luciferase coding sequences (the best match being achieved by a displacement of two residues). A solid line denotes a perfect match between the two sequences under comparison. A dashed line denotes a match that could conserve any base-pairing which occurs in the case of the HCV sequences; i.e. an A or a C in the HCV sequence (which might be paired with a U or a G, respectively) could be substituted by a G or a U, respectively, in the luciferase or NS′ coding sequences with only slight perturbation of base-pairing. The lower part shows the variations in this sequence amongst 39 different worldwide isolates examined by Bukh *et al.* (1992) for variations in the 5′-UTR sequences. The coding sequences, which are not actually published in Bukh *et al.* (1992) were obtained via the GenBank accession numbers in that publication. The top line of this table is considered the prototypic sequence (24 isolates); the sequence variations in the other isolates are shown in lower case. At the bottom, the consensus sequence is given in upper case letters, with invariant residues in bold and all the variations shown in lower case.

between these coding sequences and some motifs in the HCV 5′-UTR, then we conclude that similar base-pairing would be much more likely if the downstream cistron is the luciferase ORF rather than the NS′ ORF (Figure 7). Wang *et al.* (1993) did not actually test the influence of HCV coding sequences and thus we do not know whether they might have observed even higher internal initiation efficiency if the viral coding sequences had been retained.

The high mutation rate on RNA viruses (Holland *et al.*, 1982; Domingo *et al.*, 1985) generally leads to considerable sequence variation in the wobble positions of the coding sequences of different strains and different isolates. For example, in a study of a 60 codon segment

of the coding region of 62 different wild-type 1 poliovirus isolates from five different continents, all the wobble positions showed variations and any one isolate deviated from the prototypic sequence in ~40–60% of the wobble positions (Rico-Hesse *et al.*, 1987). In contrast, in an analysis of 39 different worldwide and thus almost certainly independent HCV isolates (Bukh *et al.*, 1992), no fewer than 24 were absolutely identical in the first 11 codons immediately downstream of the initiation codon and six of these codons were invariant in the wobble position over all 39 sequences (Figure 7). Of the 15 isolates that differ from the prototypic sequence, two show variation in ~30% of the wobble positions in these 11 codons, one in ~20%, 10 in 10% of the wobble positions and in two cases the variations are confined to non-wobble positions (Figure 7). An analysis which extended the comparison further into the coding sequences, albeit with a smaller number of different strains and isolates, showed that the typical RNA virus pattern of strain-dependent variation in at least 80% of the wobble positions occurs only after about the 50th codon (Simmonds *et al.*, 1993). Clearly, there seems to be strong selective pressure for particular nucleotide sequences at the start of the HCV ORF which cannot be explained solely as constraints dictated by the need for a particular amino acid sequence. An involvement of the nucleotide sequence of the start of the coding region in the function of the IRES could provide an explanation for this unusually high sequence conservation.

According to the picornavirus paradigm for internal initiation, a requirement for coding sequences (or more strictly, for viral sequences immediately downstream of the actual ribosome entry site) is quite unexpected. In the case of the cardioviruses, encephalomyocarditis virus (EMCV) and Theiler's murine encephalomyelitis virus (TMEV), in which the actual ribosome entry site is at the authentic initiation codon, internal initiation shows no requirement specifically for immediate downstream viral coding sequences, although certain motifs, especially G-rich sequences, can be inhibitory if located just after the AUG initiation codon (Hunt *et al.*, 1993). This conclusion is consistent with the numerous examples in the literature of successful expression of various reporters linked to the EMCV IRES without the inclusion of any viral coding sequences. In the case of poliovirus, in which the actual ribosome entry site is thought to be at a silent AUG triplet some 160 nucleotides upstream of the functional initiation codon, quite extensive mutation of the sequences immediately downstream of the silent AUG triplet had a rather small effect on internal initiation efficiency, at least *in vitro* (Meerovitch *et al.*, 1991).

Another unprecedented feature of internal initiation driven by the HCV IRES is the high efficiency even when AUU or CUG is substituted for the normal AUG initiation codon. Although CUG and to a lesser extent AUU have been found to function as initiation codons in the background of other mRNAs which are almost certainly translated by the scanning ribosome mechanism (Peabody, 1989; Mehdi *et al.*, 1990), these results with the HCV IRES are remarkable for the almost unprecedented high efficiency with which these non-AUG codons are used as initiation sites. Only in the case of a GUG codon in the human parainfluenza-1 virus P/C mRNA has such a high initiation efficiency been observed relative to that with an AUG codon in the same position (Böck *et al.*, 1992).

In the case of EMCV IRES, in which the AUG codon at the 3′-end of the IRES is both the ribosome entry site and the authentic initiation codon (Kaminski *et al.*, 1990), mutation of this AUG to AUU results in a 75% decrease in the overall yield of translation product and a switch in initiation site to the next AUG codon downstream of the site of mutation (unpublished observations). In polioviruses the AUG triplet at the 3′-end of the IRES is not used to any detectable extent as a functional initiation site, but appears to be an element of the ribosome entry site from which the ribosomes are transferred to the downstream authentic initiation codon by a process akin to scanning (Jackson *et al.*, 1990, 1994; Pilipenko *et al.*, 1992). However, even though this AUG at the 3′-end of the IRES is not a functional initiation site, its mutation to AUU, UUG or AAG results in a 3-fold decrease in the yield of reporter gene product initiated at the next AUG codon downstream in monocistronic constructs translated in HeLa cell extracts (Meerovitch *et al.*, 1991) and mutation to UUG, GUG or ACG results in a viral clone with a small plaque phenotype *in vivo* (Pelletier *et al.*, 1988; Pilipenko *et al.*, 1992).

This contrast between the HCV IRES and the picornavirus IRESs in respect of the outcome of substituting a non-AUG codon may again be explicable by the fact that coding sequences, or sequences downstream of the relevant AUG triplet, contribute to IRES function in the case of HCV but not in the picornaviruses (Hunt *et al.*, 1993). In the picornavirus IRESs the AUG entry site codon is an important determinant of internal initiation, as is its distance from upstream elements such as the conserved oligopyrimidine tract (Pilipenko *et al.*, 1992; Kaminski *et al.*, 1994a), whereas sequences downstream of the AUG seem relatively unimportant. In contrast, in HCV the authentic initiation site lies between strong determinants of internal initiation located both upstream and downstream, and thus it is easier to imagine a mutation to a non-AUG codon being tolerated than in the case of the picornavirus IRESs where all the determinants for internal initiation lie upstream of the AUG codon, and downstream coding sequences seem to play no direct role (Hunt *et al.*, 1993; Jackson *et al.*, 1994).

It is clear that internal initiation driven by the HCV IRES differs from the picornavirus paradigm in a great many respects, but most notably in the requirement for a short length of coding sequences. Apart from its intrinsic interest, the most important general implication of these results is that in the search for IRESs in cellular and other viral mRNAs, the experimental design should allow for the possibility that coding sequences could be important for IRES function. Any negative results which have been obtained in experiments following the usual practice of inserting just the 5′-UTR into the dicistronic test construct must now be considered as inconclusive until the test has been repeated with coding sequences included in the construction. We have set out elsewhere a revised strategy for testing for IRESs which covers the possibility that coding sequences or even 3′-UTR sequences may be an integral part of the IRES (Kaminski *et al.*, 1994b).

Materials and methods

Plasmid constructs

The starting construct pDX149 was obtained from Wellcome Diagnostics. It consists of nucleotides 40–405 of the HCV genome sequence as defined by Han *et al.* (1991) and was derived from a British patient infected with virus of genotype 1b. This sequence was flanked on the 5′-side by residual polylinker sequences (*Eco*RI, *Sac*I, *Asp*718, *Sma*I) from previous subcloning steps and on the 3′-side by an *Eco*RI site. The *Eco*RI fragment was subcloned into the *Eco*RI site of pGEM-1 (Promega) and a clone (pGHC6) selected with the orientation that transcription by SP6 RNA polymerase would produce a sense transcript. pGHC6 was digested with *Sac*I (there being two such sites, one in the pGEM-1 polylinker sequences proper and the other from the residual polylinker sequences of pDX149) and religated, generating pGHC6e, which now has a unique *Eco*RI site downstream of the HCV sequences. To generate deletions from the 3′-end of the HCV sequences, pGHC6e was digested with *Eco*RI, and then with *Bal*31 nuclease for various times (Figure 1). The resected ends were in-filled, then the HCV sequence was released by digestion with *Sac*I and gel-purified. This fragment was then ligated into cut pJ0, a monocistronic NS′ derivative plasmid in pGEM-2 as vector which has been described previously (Borman and Jackson, 1992). Two different insertion strategies were used (Figure 1). For the majority of constructs, pJ0 was digested with *Asp*718 and incubated with Klenow fragment of DNA polymerase I and dGTP, to partially in-fill the ends, then the remaining overhang removed by digestion with mung bean nuclease. It was then digested with *Sac*I and the HCV segment inserted between the *Sac*I site and the partially in-filled and blunted *Asp*718 site, generating monocistronic constructs with the HCV sequences followed by the sequence ..CCATGGATCC..., where the AUG codon is the start of the NS′ ORF (Figure 1). Alternatively, for the construction of p40-395.NS′, pJ0 was first cut at the unique *Bam*HI site immediately downstream of this ATG triplet and the ends in-filled completely, before digestion with *Sac*I and ligation of the HCV sequences between the *Sac*I site and the blunted *Bam*HI site.

These monocistronic derivatives were converted to dicistronic constructs by a similar strategy as described by Borman and Jackson (1992), except that the insert was excised using *Sal*I to cut on the upstream side, but *Sph*I rather than *Eco*RI on the downstream side. The excised insert was gel-purified and ligated into the dicistronic control plasmid, pXLJ0, previously named pXLJ Con in Borman and Jackson (1992), which had been cut with *Sal*I and with *Sph*I.

Point mutations were made in pXL40-372.NS′ by site-directed mutagenesis. The smaller *Sal*I–*Bam*HI fragment from pXL40-372.NS′ was subcloned between the *Sal*I and *Bam*HI sites of the phagemid vector pKTO, to generate single-stranded dU-substituted template strand for site-directed mutagenesis by the use of mismatched oligodeoxynucleotides as described by Brierley *et al.* (1989) and based upon the method of Kunkel (1985). To generate a variety of mutant alleles of the authentic HCV initiation codon, an oligonucleotide was used with 4-fold redundancy in the first position (corresponding to the A) and 3-fold redundancy (A, C and T) in the middle position.

Monocistronic derivatives lacking most of the HCV 5′-UTR and designed to be translated by the scanning ribosome mechanism were generated by digestion of the appropriate dicistronic construct with *Bam*HI and *Stu*I (Figure 1), followed by insertion of this short fragment between the *Bam*HI site and the *Hin*cII site in the polylinker of the monocistronic p40-372.NS′.

The dicistronic reconstructions pXL40-372.SRec and pXL40-372.ARec were made from pXL40-339.NS′ by first sub-cloning the *Nru*I–*Bam*HI fragment from pXL40-339 NS′ (Figure 1) between the *Sma*I and *Bam*HI sites in the cloning cassette of pGEM-1. The latter was then opened at the unique *Nco*I site and ligated with a pair of annealed synthetic oligonucleotides designed to have *Nco*I 5′ protrusions at both ends, and which could thus be inserted in either orientation. The resulting plasmids were then digested with *Stu*I and *Bam*HI, and the smaller fragment ligated with the large fragment of pXL40-339.NS′ previously cut with *Stu*I and *Bam*HI (Figure 1).

The CAT–SAP series of constructs have upstream and downstream reporters of CAT and SAP, respectively. The SAP cDNA was amplified by PCR with primers which introduced *Eco*RI and *Aat*II sites at the 5′-end and *Xba*I and *Asp*718 sites at the 3′-end of the gene, and the product was cloned between the *Eco*RI and *Asp*718 sites of Bluescript KS+ (Stratagene) to produce pKSAP2. After verification of the sequence, the *Aat*II–*Asp*718 fragment from pKSAP2 was inserted between the *Aat*II and *Asp*718 sites of a pKS+ based plasmid (pKHC-341) which contains the complete 5′-UTR of HCV plus a portion of the virus ORF coding for the core protein, the HCV sequences being flanked by a *Sal*I site (introduced by linker mutagenesis) on the upstream side and an *Asp*718 site on the downstream side. Since the *Aat*II site in pKHC-341 is within the core coding region, cutting between nucleotides 406 and 407, the resulting monocistronic construct (pKUSAP2) has the complete 5′-UTR of HCV followed by the first 22 codons of core protein coding sequence fused directly to the SAP ORF. The CAT gene was amplified by PCR using primers which introduced an *Xba*I site at the 5′-end and a *Sal*I site followed by a *Bam*HI site at the 3′-end. The product was cloned between the *Xba*I and *Bam*HI sites of Bluescript KS+ to generate pKCAT2 and the sequence was verified. Then pCAT 1-407.SAP was constructed by sub-cloning the *Sal*I–*Asp*718 fragment from pKUSAP2 into pKCAT2 that had been digested with *Sal*I and *Asp*718.

The control construct pCAT-SAP was made by digesting pCAT 1-407.SAP with *Sal*I and *Aat*II, blunting the ends using T4 DNA polymerase and then religating. This eliminated all HCV sequences and reduced the intercistronic spacer to a length of five residues. To construct pCAT 1-341.HCV, plasmid pCAT 1-407.SAP was digested with *Nhe*I, which cuts in the middle of the HCV 5′-UTR, and with *Sph*I, which cuts after the sixth codon in the SAP coding sequence, and this fragment was replaced by the *Nhe*I–*Sph*I fragment from pHCVSAP, a monocistronic construct consisting of the entire HCV 5′-UTR with the SAP reading frame fused directly to the authentic HCV initiation codon.

All plasmids were propagated by standard methods in *Escherichia coli* TG1, using ampicillin selection (Sambrook *et al.*, 1989).

Transcription reactions and translation assays

All plasmids with the NS′ cistron were linearized by digestion with *Eco*RI prior to transcription (Figure 1), whereas plasmids of the pCAT-SAP series were linearized with *Asp*718. The production of capped or uncapped RNAs by transcription with bacteriophage T7 RNA polymerase was as described previously (Borman and Jackson, 1992; Grünert and Jackson, 1994; Kaminski *et al.*, 1994a). Capping efficiency was estimated at ~70% (Grünert and Jackson, 1994). Transcripts were translated in micrococcal nuclease treated rabbit reticulocyte lysate under the conditions described previously (Grünert and Jackson, 1994), with added KCl at 100 mM and Mg^{2+} at 0.5 mM. Aliquots of the translation assay samples were subjected to gel electrophoresis using 20% polyacrylamide gels as described previously (Jackson, 1986; Dasso and Jackson, 1989). Stained, dried gels were exposed to Hyperfilm β-Max (Amersham International) usually overnight.

Transfection assays

BT7-H cells (Whetter *et al.*, 1994) in 6-well Costar plates were infected with recombinant vaccinia virus vTF7-3 (Fuerst *et al.*, 1986) at a m.o.i. of 10. After 60 min at 37°C, the inoculum was removed and the cell sheet washed once with OPTIMEM (Gibco-BRL). DNA transfection was carried out using 5 μg plasmid DNA and 30 μg DOTAP (Boehringer Mannheim) as described in the manufacturer's protocol. At 17 h post-transfection the medium was replaced with methionine-free Eagle's medium. After a further 60 min, 5 μCi [^{35}S]methionine was added per well and the cells incubated for a further 2 h. Cells were scraped from the plates, pelleted by centrifugation and processed for gel electrophoresis.

Acknowledgements

We thank Stefan Grünert for his patience and help in the early stages of this work; Catherine Gibbs for technical support; Tim Hunt for the gift of reticulocyte lysates; B.Rodgers and D.Parker of the former Wellcome Diagnostics (now Murex Diagnostics) for the original gift of pDX149; and Eckard Wimmer for communicating results prior to publication. J.Reynolds gratefully acknowledges the award of a studentship from the former Science and Engineering Research Council (now the Biotechnology and Biological Sciences Research Council).

References

Böck,R. and Kolakofsky,D. (1994) Positions +5 and +6 can be major determinants of the efficiency of non-AUG initiation codons for protein synthesis. *EMBO J.*, **13**, 3608–3617.

Böck,R., Curran,J., Matsuoka,Y., Compans,R. and Kolakofsky,D. (1992) The parainfluenza virus type 1 P/C gene uses a very efficient GUG codon to start its C′ protein. *J. Virol.*, **66**, 1765–1768.

Borman,A. and Jackson,R.J. (1992) Initiation of translation of human

rhinovirus RNA: mapping the internal ribosomal entry site. *Virology*, **188**, 685–696.

Brierley,I., Digard,P. and Inglis,S.C. (1989) Characterization of an efficient coronavirus ribosomal frameshifting signal: requirement for an RNA pseudoknot. *Cell*, **57**, 537–547.

Brown,E.A., Zhang,H., Ping,L. and Lemon,S.M. (1992) Secondary structure of the 5′ non translated regions of hepatitis C virus and pestivirus genomic RNAs. *Nucleic Acids Res.*, **19**, 5041–5045.

Bukh,J., Purcell,R.H. and Miller,R.H. (1992) Sequence analysis of the 5′ noncoding region of hepatitis C virus. *Proc. Natl Acad. Sci. USA*, **89**, 4942–4946.

Choo,Q.-L., Kuo,G., Weiner,A.J., Overby,L.R., Bradley,D.W. and Houghton,M. (1989) Isolation of a cDNA clone derived from a blood-borne non-A, non-B viral hepatitis genome. *Science*, **244**, 359–362.

Choo,Q.-L. *et al.* (1991) Genetic organization and diversity of the hepatitis C virus. *Proc. Natl Acad. Sci. USA*, **88**, 2451–2455.

Dasso,M.C. and Jackson,R.J. (1989) Efficient initiation of mammalian mRNA translation at a CUG codon. *Nucleic Acids Res.*, **17**, 6485–6497.

Domingo,E. *et al.* (1985) The quasispecies (extremely heterogeneous) nature of viral RNA genome populations: biological relevance—a review. *Gene*, **40**, 1–8.

Fuerst,T.R., Niles,E.G., Studier,F.W. and Moss,B. (1986) Eukaryotic transient expression system based on recombinant vaccinia virus that synthesises bacteriophage T7 RNA polymerase. *Proc. Natl Acad. Sci. USA*, **83**, 8122–8126.

Fukushi,S., Katayama,K., Kurihara,C., Ishiyama,N., Hoshino,F.B., Ando,T. and Oya,A. (1994) Complete 5′ noncoding region is necessary for the efficient internal initiation of hepatitis C virus RNA. *Biochem. Biophys. Res. Commun.*, **199**, 425–432.

Grünert,S. and Jackson,R.J. (1994) The immediate downstream codon strongly influences the efficiency of utilization of eukaryotic translation initiation sites. *EMBO J.*, **13**, 3618–3630.

Han,J.H. *et al.* (1991) Characterization of the terminal regions of hepatitis C viral RNA: identification of conserved sequences in the 5′ untranslated region and poly(A) tails at the 3′ end. *Proc. Natl Acad. Sci. USA*, **88**, 1711–1715.

Holland,J.J., Spindler,K., Horodyski,F., Grabau,E., Nichol,S. and VanDepol,S. (1982) Rapid evolution of RNA genomes. *Science*, **215**, 1577–1585.

Hunt,S.L., Kaminski,A. and Jackson,R.J. (1993) The influence of viral coding sequences on the efficiency of internal initiation of translation of cardiovirus RNAs. *Virology*, **197**, 801–807.

Jackson,R.J. (1986) A detailed kinetic analysis of the in vitro synthesis and processing of encephalomyocarditis virus products. *Virology*, **149**, 114–117.

Jackson,R.J., Howell,M.T. and Kaminski,A. (1990) The novel mechanism of initiation of picornavirus RNA translation. *Trends Biochem. Sci.*, **15**, 477–483.

Jackson,R.J., Hunt,S.L., Gibbs,C.L. and Kaminski,A. (1994) Internal initiation of translation of picornavirus RNAs. *Mol. Biol. Reports*, **19**, 147–159.

Jang,S.K., Krausslich,H.-G., Nicklin,M.J.H., Duke,G.M., Palmenberg, A.C. and Wimmer,E. (1988) A segment of the 5′ nontranslated region of encephalomyocarditis virus RNA directs internal entry of ribosomes during in vitro translation. *J. Virol.*, **62**, 2636–2643.

Kaminski,A., Howell,M.T. and Jackson,R.J. (1990) Initiation of encephalomyocarditis virus RNA translation: the authentic initiation site is not selected by a scanning mechanism. *EMBO J.*, **9**, 3753–3759.

Kaminski,A., Belsham,G.J. and Jackson,R.J. (1994a) Translation of encephalomyocarditis virus RNA; parameters influencing the selection of the internal initiation site. *EMBO J.*, **13**, 1673–1681.

Kaminski,A., Hunt,S.L., Gibbs,C.L. and Jackson,R.J. (1994b) Internal initiation of mRNA translation in eukaryotes. In Setlow,J. (ed.), *Genetic Engineering: Principles and Methods.* Vol. 16. Plenum Press, New York, pp. 115–155.

Kozak,M. (1986) Influences of mRNA secondary structure on initiation by eukaryotic ribosomes. *Proc. Natl Acad. Sci. USA*, **83**, 2850–2854.

Kozak,M. (1989) The scanning model for translation: an update. *J. Cell. Biol.*, **108**, 229–241.

Kunkel,T.A. (1985) Rapid and efficient site-directed mutagenesis without phenotypic selection. *Proc. Natl Acad. Sci. USA*, **82**, 488–492.

Meerovitch,K., Nicholson,R. and Sonenberg,N. (1991) In vitro mutational analysis of cis-acting RNA translational elements within the poliovirus type 2 5′ untranslated region. *J. Virol.*, **65**, 5895–5901.

Mehdi,H., Ono,E. and Gupta,K.C. (1990) Initiation of translation at CUG, GUG and ACG codons in mammalian cells. *Gene*, **91**, 173–178.

Peabody,D.S. (1989) Translation initiation at non-AUG triplets in mammalian cells. *J. Biol. Chem.*, **264**, 5031–5035.

Pelletier,J. and Sonenberg,N. (1988) Internal initiation of translation of eukaryotic mRNA directed by a sequence derived from poliovirus RNA. *Nature*, **334**, 320–325.

Pelletier,J., Flynn,M.E., Kaplan,G., Racaniello,V. and Sonenberg,N. (1988) Mutational analysis of upstream AUG codons of poliovirus RNA. *J. Virol.*, **62**, 4486–4492.

Pilipenko,E.V., Gmyl,A.P., Maslova,S.V., Svitkin,Y.V., Sinyakov,A.N. and Agol,V.I. (1992) Prokaryotic-like cis elements in the cap-independent internal initiation of translation on picornavirus RNA. *Cell*, **68**, 119–131.

Rico-Hesse,R., Pallansch,M.A., Nottay,B.K. and Kew,O.M. (1987) Geographic distribution of wild poliovirus type 1 genotypes. *Virology*, **160**, 311–322.

Rijnbrand,R., Bredenbeek,P., van der Straten,T., Whetter,L., Inchauspe,G., Lemon,S. and Spaan,W. (1995) Almost the entire 5′ non-translated region of hepatitis C virus is required for cap independent translation. *FEBS Lett.*, **365**, 115–119.

Saito,I. *et al.* (1990) Hepatitis C virus infection is associated with the development of hepatocellular carcinoma. *Proc. Natl Acad. Sci. USA*, **87**, 6547–6549.

Sambrook,J., Fritsch,E.F. and Maniatis,T. (1989) *Molecular Cloning. A Laboratory Manual.* Cold Spring Harbour Laboratory Press, Cold Spring Harbour, NY.

Sherman,F., Stewart,J.W. and Tsunasawa,S. (1985) Methionine or not methionine at the beginning of a protein. *Bioessays*, **3**, 27–31.

Sherman,F., Moerschell,R.P., Tsunasawa,S., Sternglanz,R. and Dumont,M.E. (1993) Co- and posttranslational processes and mitochondrial import of yeast cytochrome c. In Ilan,J. (ed.), *Translational Regulation of Gene Expression 2.* Plenum Press, New York, pp. 117–141.

Simmonds,P., McOmish,F., Yap,P.L., Chan,S.-W., Lin,C.K., Dushieko,G., Saeed,A.A. and Holmes,E.C. (1993) Sequence variability in the 5′ non-coding region of hepatitis C virus: identification of a new virus type and restrictions on sequence diversity. *J. Gen. Virol.*, **74**, 661–668.

Takamizawa,A. *et al.* (1991) Structure and organization of the hepatitis C virus genome isolated from human carriers. *J. Virol.*, **65**, 1105–1113.

Tsukiyama-Kohara,K., Iizuka,N., Kohara,M. and Nomoto,A. (1992) Internal ribosome entry site within hepatitis C virus RNA. *J. Virol.*, **66**, 1476–1483.

Wang,C., Sarnow,P. and Siddiqui,A. (1993) Translation of human hepatitis C virus RNA in cultured cells is mediated by an internal ribosome-binding mechanism. *J. Virol.*, **67**, 3338–3344.

Whetter,L.E., Day,S.P., Elroy-Stein,O., Brown,E.A. and Lemon,S.M. (1994) Low efficiency of the 5′ non-translated region of hepatitis A virus RNA in promoting CAP independent translation in permissive monkey kidney cells. *J. Virol.*, **68**, 5253–5263.

Yoo,B.J., Spaete,R., Geballe,A.P., Selby,M., Houghton,M. and Han,J.H. (1992) 5′ end-dependent translation of hepatitis C viral RNA and the presence of putative positive and negative translation control elements within the 5′ untranslated region. *Virology*, **191**, 889–899.

Received on August 2, 1995