5' UTR of hepatitis A virus RNA: mutations in the 5'-most pyrimidine-rich tract reduce its ability to direct internal initiation of translation

Jane Silveira Carneiro, M. Equestre, P. Pagnotti, Alessandra Gradi, N. Sonenberg and R. Perez Bercoff

1 Department of Cellular & Developmental Biology, University of Rome, Viale di Porta Tiburtina 28, 00185-Rome, Italy
2 Laboratory of Virology, Istituto Superiore di Sanita, 00161-Rome, Italy and 3 Department of Biochemistry, McGill University, 3566 Drummond, Montreal PQ H3G 1Y6, Canada

The 5'-terminal untranslated region (5' UTR) of the uncapped hepatitis A virus (HAV) RNA contains two pyrimidine-rich sequences; one about 20 nucleotides (nt) in length in the vicinity of the AUG initiation codon (nt 706-726), and a longer one (about 40 nt) encompassing nt 100 to 140. The latter includes a 13 nt ‘core’ sequence (positions 126-138 in the HM175 strain) which is 80% identical to the pyrimidine-rich tract of poliovirus type 1 RNA (Mahoney strain). Representative cDNAs of the entire 5' UTR of HAV RNA were inserted in the intercistronic region of the bi-cistronic plasmid pSV-GH/CAT between the genes coding for the human growth hormone (GH) and bacterial chloramphenicol acetyltransferase (CAT). When COS-7 cells were transfected with these constructs they transiently expressed CAT indicating that the 5' UTR of HAV was efficiently directing internal initiation of translation of the reporter gene. Under similar conditions the 5' UTR of poliovirus type 2 (Lansing strain) was 30% more efficient in directing the expression of the CAT gene. Removal of the ‘core’ sequence from the 5'-distal pyrimidine-rich stretch extending between nt 117 and 131 in the HAV 5' UTR reduced the CAT activity in the lysates of transfected cells by 40%, whereas point mutations engineered in this segment strongly decreased (80% inhibition) the HAV-driven expression of the reporter gene. Limited mutations systematically introduced in the reiterated (U)UUUCCC motifs of the 5'-distal pyrimidine-rich tract identified two major functional domains extending between nt 100-106 and 113-119. Substitutions in these hexanucleotides abrogated internal initiation of translation, whereas similar changes in the neighbouring domains (nt 107-112 and 120-126) had no effect on the expression of the reporter gene, suggesting that the 5'-most pyrimidine-rich tract is indeed part of the structure(s) recognized by ribosomes and associated factors at initiation of translation and that the hexanucleotides 100-106 and 113-119 constitute an important part of it. Although HAV replicates better at 33 °C than at 37 °C, incubation of transfected cultures at 33 °C delayed the expression and slightly reduced the level of CAT activity in the cell lysates, but the overall effect of the mutations remained unchanged.

Introduction

The causative agent of human hepatitis A is a picornavirus recently assigned to the hepatovirus group. In contrast to other members of this family (entero-, rhino-, cardio- and aphthoviruses) Hepatitis A virus (HAV) replicates very slowly (several days compared with a few hours) in all tissue culture systems so far studied and has little noticeable effect on host cell macromolecular synthesis. There is no definitive explanation for this unusual behaviour.

The genome of HAV is a single-stranded RNA molecule, 7478 nucleotides (nt) in length, of positive polarity, the uncapped 5' end of which is covalently blocked by a small protein (VPg). The general organization of HAV genomic RNA is similar to that of other members of the family Picornaviridae: a single open reading frame (ORF) extends from nucleotide 734 to residue 7410, followed by a short extracistronic region (68 nt), and a 3'-terminal poly(A) tract. Two in-frame AUG codons initiate the ORF at positions 734–736 and 740–742 (nucleotide numbering according to the sequence published by Cohen et al., 1987; GeneBank accession number M14707). The second AUG seems to be preferred in vitro and used almost exclusively in vivo (Tesar et al., 1992).

In HAV RNA the long extracistronic sequence
extending between the uncapped 5' end and the beginning of the ORF contains nine (apparently unused) AUG initiation codons. As in other picornavirus genomes, the 5' UTR of HAV RNA presents a high degree of secondary structure: computer-generated models of putative RNA folding and analysis of the products of partial enzymic or chemical degradation suggested that it folds into six stem–loop domains with a long single-stranded region extending between nucleotides 94 and 154 (Brown et al., 1991; 1992). In none of these studies, however, was the putative secondary structure of the first 263 nt tested by chemical or enzymic analysis.

Nucleotide sequence analysis of the 5' UTR of several isolates of HAV revealed that besides a canonically located pyrimidine-rich stretch (nt 711–722) present in all picornavirus RNAs in the vicinity of the initiation codon (Nicholson et al., 1991), HAV RNA contains a second, in what appeared to be an abnormal location (nt 100–139) (Fig. 1). The length of the stretch varies in different HAV isolates, but its structure does not: a basic motif (U)UUUCCC, is reiterated three to five times with a few intercalated purines. Interestingly, the core of this sequence (nt 126–138) has 80% sequence identity when compared with the pyrimidine-rich tract (PRT) of poliovirus type 1 (Mahoney strain) (Fig. 1).

The presence of a second PRT 5'-distal to the translation initiation site was in many ways intriguing. In fact, this sequence is one of the most conserved regions of picornavirus RNA. Its length, and the distance between the tract and the following AUG codon are virtually constant in each group within the family Picornaviridae and its location relative to the translation initiation site divides picornaviruses into two groups: in cardio- and aphthoviruses the PRT precedes the initiator AUG by 16–18 nt, in entero- and rhinoviruses the polypyrimidine stretch lies 8–20 nt upstream from a silent, unused AUG codon and about 150 nt upstream of the beginning of the ORF. Mutational analysis of the 5' UTR of poliovirus type 2 (Lansing strain) revealed that substitutions in the PRT dramatically reduced the translation of a downstream reporter gene. Moreover, the pyrimidines making up the 5'-distal half of the stretch seemed to play a critical role in securing translation (Meerovitch et al., 1991a). All the available evidence indicates that the pyrimidine-rich sequence constitutes the 3' border of the ribosome landing pad, the structure responsible for internal initiation of translation in picornavirus RNA (Nicholson et al., 1991; Sonenberg & Meerovitch, 1990; Meerovitch et al., 1991b).

Thus, the presence of a second PRT transposed to a location removed from the translation initiation codon(s) raised the question as to whether this peculiar genome organization might account for the notoriously poor in vitro translation ability of HAV RNA and, possibly, for its lengthy replication cycle. In fact, there is considerable evidence suggesting that a precisely located pyrimidine-
rich segment of proper length may be one of the signals required to determine the AUG used at initiation of translation of picornavirus RNAs (Sonenberg & Meerovitch, 1990; Meerovitch et al., 1991b). Moreover, it was not clear whether the peculiar configuration of other picornavirus RNAs.

In vitro translation of a nested series of 5'-terminally truncated RNA transcripts of HAV cDNAs has shown that the region encompassing nt 345–734 was stringently required for translation in rabbit reticulocyte lysates (Brown et al., 1991, 1992). While these experiments clearly defined the minimal length of the extra-cistronic region of HAV RNA required for in vitro translation, they did not prove formally that translation of mRNAs with an intact HAV 5' UTR occurred via binding of ribosomes to internal (rather than 5'-terminal) sequences. They also suggested that the 5'-terminal half of HAV 5' UTR (including the first PRT) participated marginally (if at all) in translation, at least in the reticulocyte system. The issue was addressed more recently by Glass et al. (1993), who introduced a series of large deletions (ranging from 105 to 387 nt), or short insertions (4–6 nt) in the 734 nt HAV 5' UTR. Subsequently, the effect that these changes had on the ability of the remaining sequences of the 5' UTR to direct protein synthesis both in vitro and in vivo was investigated. However, none of the above-mentioned studies investigated the role of the 5'-distal PRT.

In an attempt to ascertain whether the 5'-most PRT was involved in the process of internal initiation of translation, mutations and deletions were engineered into a cDNA representation of HAV RNA 5' UTR. The native and mutated sequences were subsequently inserted in the intergenic region of the bi-cistronic plasmid pSV-GH/CAT (Nicholson et al., 1991), and the constructs were transfected into COS-7 cells where the transient expression of both genes was quantified. Here we report that the 5'-most PRT of HAV RNA is indeed required to allow translation in vivo, and contains two major functional domains encompassing nucleotides 100–106 and 113–119. Base substitutions in these sequences are incompatible with internal initiation of translation in COS-7 cells.

Methods

Cells and bacteria. COS-7 cells were routinely maintained as monolayers in Dulbecco's modified Eagle's medium (DMEM) with 2% synthetic serum substitute Ultroser G (Sepracor, France) and antibiotics. Cultures of Escherichia coli TG-1 were transformed following treatment with calcium chloride and selected for ampicillin resistance following standard procedures.

Plasmids construction. The molecular cloning of the genome of the fast-growing strain of HAV has been described previously (Venuti et al., 1985; Divizia et al., 1986). Digestion of plasmid DNA containing poliovirus or HAV cDNA sequences, dephosphorylation by alkaline phosphatase treatment, purification of DNA fragments by agarose gel electrophoresis and QN+-butanol extraction (Landridge et al., 1980), ligation of proper DNA fragments, and propagation of the constructs in competent bacteria were performed by standard procedures (Sambrook et al., 1989). DNA-modifying enzymes were used according to the manufacturers' instructions.

Oligonucleotide-directed and PCR mutagenesis. DNA fragments containing the sequences to be mutated were subcloned in the multiple cloning site (MCS) of the phagemid BlueScript KS+ (Stratagene), and propagated in E. coli strain TG-1. Single-stranded recombinant phagemid DNA was rescued by superinfection with the defective phage M13K07 as described in Sambrook et al. (1989) with minimal modifications. Three to five µg of ssDNA were used as template in oligonucleotide-directed mutagenesis according to the procedure of Taylor et al. (1985) (Amersham). Alternatively, mutated cDNA sequences were generated in PCR driven by pairs of complementary oligonucleotides containing the desired mutations. This was done as described by Jones & Winstonfor (1991) except that linear cDNA fragments were generated in separate reactions and following purification they were allowed to hybridize to the overlapping sequences, followed by extension and amplification by PCR. The mutated linear cDNA fragments were gel-purified, digested with the enzyme AvrII (which cuts HAV cDNA at bases 82 and 248), and introduced by 'cassette exchange' into the AvrII-digested, phosphatase-treated and gel-purified plasmid pSV-GH-HEP-CAT, where the deleted or mutated sequences replaced the portion extending between the two AvrII sites of the HAV 5' UTR. The orientation and correct insertion of the mutated fragments were ascertained by direct DNA sequencing.

Transfection. COS-7 cells (4 × 10^5 to 6 × 10^5) in 10 ml DMEM plus 2% Ultroser, were seeded in 10 cm diameter plastic dishes (NUNC clone). Two days later, the medium was removed and the cultures were transfected with 10 to 15 µg of the bi-cistronic plasmid DNA according to the DEAE-dextran/chloroquine procedure (Gorman, 1985). Following incubation at the indicated temperature, the amount of human growth hormone (GH) present in the supernatants was determined by radioimmunounaasay (RIA) in triplicate 100 µl aliquots, and this value was used as an indication of the efficiency of transfection (Nicholson et al., 1991). The cells were then detached from the plates, washed in ice-cold PBS, suspended in 300 µl 250 mM-Tris–HCl pH 8.3, frozen and thawed three times, and the lysates processed as described (Sambrook et al., 1989). The protein concentration was determined in duplicate 10 µl samples (Bio-Rad), and the chloramphenicol acetyltransferase (CAT) activity was measured by standard procedures after correcting for the protein content. The acetylated forms of [14C]chloramphenicol were resolved from the unmodified ones by ascending TLC in silica gel plates. Following autoradiography, the relative activities were determined either by cutting out the corresponding areas of the TLC plates and counting the radioactivity by liquid scintillation, or this was estimated in a photoimaging device (FUJI 2000).

Computer-assisted sequence analyses. Sequence analysis was performed using the GCG Software package, University of Wisconsin (Devereux et al., 1984).

Results

Bi-cistronic constructs

A cDNA representation of the entire 5' UTR of HAV RNA (bases 1–744) was obtained (Fig. 2a) by digestion
of plasmid pHepA (Venuti et al., 1985) with the enzymes XhoI and XbaI (which cut in the MCS of the BlueScript vector) and at base 744 of the HAV insert. The fragment was blunt-ended, gel-purified and ligated into the intergenic region (EcoRV site) of the bi-cistronic expression vector pSV-GH/CAT. This resulted in a construct where the two legitimate AUG initiation codons of HAV RNA (735–737 and 741–743) preceded in-phase the initiation triplet of the CAT gene. In a series of preliminary experiments we observed that lysates of COS-7 cells transfected with this construct were devoid of (or had very little) CAT activity. In principle, this could result from either down regulation of translation induced by the HAV sequences, or else by internal initiation of translation at any of the HAV AUG codons present in this construct. Were the latter the case, a fusion protein would be synthesized with 13 to 15 extra N-terminal amino acids and little (or no) enzymic activity. The latter
interpretation seems to be correct, since mutation of the HAV AUG triplets into UUU endowed the constructs with the ability to consistently express CAT activity (see below).

**Mutational analysis**

Two kinds of specific modifications were introduced in the 5'-distal PRT. In a first series of experiments, the 'core' sequence that shares 80% identity with the pyrimidine-rich tract of poliovirus was removed by oligonucleotide-directed mutagenesis. An oligonucleotide of the form: 5'TTCCCTTCTCTTGCTC[GCTT-GTAAATATTAATTCC 3' was synthesized to remove the 'core' sequence of the PRT. The 3'-terminal 19 nt of this primer hybridize to HAV cDNA from base 139 to 157 (nucleotide numbering according to Cohen et al., 1987), whereas the remaining 15 bases at the 5' terminus reconstruct a consensus sequence encompassing nt 101 to 117 (note the purine, G, at the end of the pyrimidine stretch).

Subsequently, point mutations were engineered by means of a synthetic 51-mer of the form: TTTCCTTTGCAAGGGACCCCCTATAGCCCTTTCT AAAGCTTTGTAATAT (the bases underlined indicate the positions at which the original nucleotide has been replaced by the complementary one). These substitutions were designed to disrupt the potential secondary folding of the PRT. All the changes engineered were checked by direct sequence analysis.

The cDNA representation of the 5' UTR of HAV RNA, with or without the modifications described, were then removed from the phagemid vector by restriction enzyme digestion (Fig. 2a) and inserted in the intergenic region of the bi-cistronic construction pSV-GH/CAT. COS-7 cells were transfected with 10 to 15 ~g of the plasmid DNA according to the DEAE-dextran/chloroquine procedure. Following incubation at the indicated temperature, the presence of GH in the supernatant, and that of CAT in the cellular lysate were assayed as described by Nicholson et al. (1991).

Transfection of COS-7 cells with bi-cistronic constructs carrying either wild-type or mutated 5' UTR of HAV cDNA was equally efficient, and comparable amounts of human GH were consistently recovered in the supernatants of the cultures 48 h after transfection (Fig. 3a).

Under these conditions, the expression of the CAT gene directed by the 5' UTR of HAV RNA or by its mutated forms differed considerably. The deletion that removed a short run (116–132) from the first pyrimidine-rich tract, reduced the translation of the CAT gene by over 40%. This value is the average of triplicate samples in four different experiments, the range of variation of the observed inhibition extending between 38% and 45%. More significantly, the long string of point mutations engineered in the 5'-most PRT resulted in a 70 to 80% reduction of the expression of the reporter gene (Fig. 3b).

These reductions could not be accounted for by the reduced ability of the constructs to transfect the target cells, since the expression of the GH gene was as effective as in cultures transfected with the parental wild-type construct. Nor could this be explained by a mechanism of intramolecular competition for translation between the GH and the CAT genes, the flanking sequences preceding the former being the same in both constructs. The observed 40 and 80% reduction in the expression of CAT appeared as a result of the diminished ability of the mutated 5' UTRs to mediate the translation of the downstream reporter gene, suggesting that the first PRT of HAV RNA (nt 100–140) is indeed part of the structure required to enable the internal initiation of translation. Although the fast-growing strain of HAV replicates better at 33 °C than at 37 °C (Venuti et al., 1985), incubation of transfected cultures at the former temperature did not modify these results (data not shown).

**Identification of functional domains in the 5'-distal PRT**

Since the above 18-point mutation introduced base substitutions distributed over most of the 5' PRT, it was not clear whether this stretch behaved as a single functional block or else defined areas contributed differently in permitting internal initiation of translation. To discriminate between these possibilities, short strings of mutations were systematically introduced in the PRT and the effect of these changes on cap-independent translation was evaluated. The cDNA fragments carrying the mutations depicted in Fig. 2(b) were generated by PCR mutagenesis and inserted into AvrII-digested pSV-GH-HEP-CAT vector. This generated a series of constructs carrying limited substitutions (Fig. 2(b)).

Upon transfection of COS-7 cells there was no significant variation in the cap-dependent expression of the upstream gene as revealed by the levels of GH present in the supernatants (Fig. 3a), indicating that these constructs were equally efficient in transfecting target cells. The expression of the downstream cistron, on the other hand, revealed considerable differences in the ability of the mutated 5' UTRs to drive internal initiation of translation. The mutation Hep100 (in which the 7 nt string UUUUCCCC (positions 100–106) had been replaced by AAAAGGG) totally abrogated translation of the reporter CAT gene (Fig. 3b). Similarly, there was no CAT activity in lysates of cells transfected with the mutant Hep113, in which a 6 nt substitution had been introduced in the string UUCCCC (positions 114–
Fig. 3. (a) COS-7 cells transfected with the DNA constructs described in the text were incubated for 48 to 72 h. The amounts of human growth hormone (GH) present in the supernatants were determined by radioimmunoassay in triplicate 100 µl samples and used as an indication of the efficiency of transfection. (b) Cells were lysed and the CAT activity was measured in vitro within the linear range of the reaction after correcting for the protein content of the lysates. The acetylated forms of [14C]chloramphenicol were resolved from the unmodified ones by ascending TLC in silica gel plates. Following autoradiography, the relative activities were determined either by cutting the corresponding areas of the TLC plate and counting the radioactivity by liquid scintillation, or this was estimated in a photoimaging device (FUJI 2000).

119). Mutations in the neighbouring domains 107–113 (UUUCCCCC) or 120–126 (UUCCUU), had no effect on translation.

These two short strings extending between nt 100–106 and 114–119 appear to constitute two major, independent, functional domains of the 5'-most PRT of HAV.
RNA, and are part of the cis-acting structures responsible for the internal initiation of translation in vivo.

Discussion

The genomic RNA of picornaviruses is exceptional among eukaryotic mRNAs for the lack of a 5'-terminal cap. In the absence of this structure, other elements must secure the proper initiation of translation. Ribosome-binding and eIF-2 protection experiments provided direct evidence in support of the contention that the cellular ribosomes and associated initiation factors recognize internal (rather than 5'-terminal) sequences of Mengovirus RNA, consistent with the notion that the genomic RNA of picornaviruses may have evolved (or conserved during evolution) a most efficient mechanism of initiation of translation that allows it to bypass the need for either a 5' end or a cap (Perez Bercoff, 1982; Perez Bercoff & Kaempfer, 1982; Degener et al., 1983).

In the case of poliovirus RNA (Lansing strain of type 2) the cis-acting elements required for internal initiation of translation map between nt 141 and 631 of the genome (Pelletier & Sonenberg, 1988).

The 5' UTR of HAV RNA can also direct internal initiation of translation of a downstream gene, albeit less efficiently (30%) than the 5' UTR of poliovirus RNA. The difference in efficiency cannot be accounted for by the difference in optimal growth temperature of these viruses (Venuti et al., 1985; Divizia et al., 1986), since the difference between poliovirus- and HAV-driven expression of the CAT gene remained unchanged when the cultures transfected with the bi-cistronic constructs were incubated at 33 °C instead of 37 °C (not shown).

Similar to other picornavirus RNAs, a pyrimidine-rich tract precedes the AUG initiation codon, but in HAV RNA the tract (12 bases) is the shortest so far described (Nicholson et al., 1991). The presence of a longer pyrimidine-rich sequence in an area far removed from the legitimate translation initiation codon is reminiscent of the situation found in aphtho- and cardioviruses, where a long poly(C) tract is close to (but not at) the 5' end of the genome (Perez Bercoff & Gander, 1977), and raised the question as to whether this sequence was one of the elements involved in the initiation of translation.

Earlier studies based on the in vitro translation of a nested series of mRNAs carrying progressively shorter or partially deleted 5' UTRs led to the conclusion that the entire 5'-distal half (including the PRT) of HAV RNA participated only marginally (if at all) in translation (Brown et al., 1991, 1992, 1994). The results reported here contradict this view, and demonstrate instead that the structures required for internal initiation of translation in vivo extend at least up to (and include) the 5'-distal PRT. Moreover, two well-defined domains appear to play a major role in this process, namely the heptanucleotide UUUUCCC (positions 100–106) and the closely related motif UUCCCC (nt 113–119). The discrepancy can be accounted for by the intrinsic differences between the systems used in these two studies. It is not surprising that the results obtained by in vitro translation of a phenolized mRNA in cell-free lysate differ from those produced when the mRNAs are transcribed (and presumably processed), transported and translated by the cellular machinery in the physiological environment of the intact cell. Whetter et al. (1994) reported that long deletions in the 5' UTR of HAV RNA were compatible with efficient translation in vivo of monocistronic RNAs. Since these studies also showed that removal of almost the entire 5' UTR (Δnt 1–633) increased the expression of the downstream gene several fold it was obvious that the observed effect was simply due to 5'-terminal mRNA scanning of the monocistronic RNA, and could not be accounted for by internal initiation of translation.

The involvement of the 5'-distal PRT of HAV RNA in the process of translation raises the question of the role of this element, which is unique among picornavirus RNAs. In fact, whereas the poly(C) tract of aphtho- and cardiovirus RNAs is an uninterrupted homopolymeric sequence of variable length (80 to over 150 nt), the 5'-distal pyrimidine stretch of HAV RNA presents strings of a reiterated hexanucleotide of the form UUUCCC, interrupted by single purines. The number of copies (three to five) of the repeated hexanucleotide and the total length of the 5'-most PRT varies in different HAV isolates at different stages of their adaptation to cell culture conditions, as if the minimal length required varied under varying growing conditions. Moreover, such a peculiar structure of juxtaposed repeated blocks suggest that the PRT might have originated by reiterative transcription of the basic motif; a mechanism used to generate, for instance, the long 3'-terminal poly(A) tract of rhabdovirus mRNAs. This interpretation is consistent with the findings of Rieder et al. (1993), who recovered foot-and-mouth disease virus (FMDV) with a poly(C) tract 42 nt in length following transfection of BHK cells with a genetically engineered construct carrying two cytidine residues instead of a real poly(C) stretch. The location of the 5'-most PRT of HAV RNA, reminiscent of the poly(C) tract of cardio- and aphthoviruses, suggests that it may constitute an important part of the signal(s) recognized by, and interacting with cellular factor(s) at initiation of translation. In this context, it is noteworthy that the interaction between the cellular protein p57 and the 5' UTR of FMDV RNA involves the PRT close to the initiator AUG, and a second site 300 nt upstream, possibly brought into proximity by secondary or tertiary folding (Luz & Beck, 1991).
After this paper was completed, Shaffer et al. (1994) reported studies on the replication ability of HAV carrying extensive deletions (14 to 46 bases) of the 5′-most PRT. Apparently, none of these deletions (except those extending up to nt 140–144, which resulted in a temperature-sensitive phenotype) had noticeable effects those extending up to nt 140-144, which resulted in a translation, it would be interesting to investigate whether such putative factor(s) are present in similar amounts in COS-7 cells (used in our studies) and in BS-C-1 or FRHk-4 cells, where the replication experiments were conducted.

J. S. C. is on leave of absence from IBAMA, Brasilia, and is recipient of a fellowship from the Brazilian C.N.P.Q. This work was partly supported with funds of the Canadian Medical Research Council, NATO Collaborative Research Grant 890157, Italian National Research Council (CNR) (Progetti Finalizzati ‘Ingegneria Genetica’ and ‘Biotecnologie e Biostrumentazzione’), and Istituto Pasteur/Fondazione Cenci-Bolognetti of Rome.

References


(Received 16 November 1994; Accepted 10 January 1995)