1474–1486 *Nucleic Acids Research, 2005, Vol. 33, No. 5*
doi:10.1093/nar/gki292

Translation of the F protein of hepatitis C virus is initiated at a non-AUG codon in a +1 reading frame relative to the polyprotein

Martin Baril and Léa Brakier-Gingras*

Département de Biochimie, Université de Montréal, 2900 Blvd Édouard-Monpetit, Montréal, Québec, Canada H3T 1J4

Received January 21, 2005; Revised and Accepted February 20, 2005

ABSTRACT

The hepatitis C virus (HCV) genome contains an internal ribosome entry site (IRES) followed by a large open reading frame coding for a polyprotein that is cleaved into 10 proteins. An additional HCV protein, the F protein, was recently suggested to result from a +1 frameshift by a minority of ribosomes that initiated translation at the HCV AUG initiator codon of the polyprotein. In the present study, we reassessed the mechanism accounting for the synthesis of the F protein by measuring the expression in cultured cells of a luciferase reporter gene with an insertion encompassing the IRES plus the beginning of the HCV-coding region preceding the luciferase-coding sequence. The insertion was such that luciferase expression was either in the +1 reading frame relative to the HCV AUG initiator codon, mimicking the expression of the F protein, or in-frame with this AUG, mimicking the expression of the polyprotein. Introduction of a stop codon at various positions in-frame with the AUG initiator codon and substitution of this AUG with UAC inhibited luciferase expression in the 0 reading frame but not in the +1 reading frame, ruling out that the synthesis of the F protein results from a +1 frameshift. Introduction of a stop codon at various positions in the +1 reading frame identified the codon overlapping codon 26 of the polyprotein in the +1 reading frame as the translation start site for the F protein. This codon 26(+1) is either GUG or GCG in the viral variants. Expression of the F protein strongly increased when codon 26(+1) was replaced with AUG, or when its context was mutated into an optimal Kozak context, but was severely decreased in the presence of low concentrations of edeine. These observations are consistent with a Met-tRNA$_i$-dependent initiation of translation at a non-AUG codon for the synthesis of the F protein.

INTRODUCTION

Hepatitis C virus (HCV) chronically infects around 200 million people worldwide and frequently causes liver cirrhosis and hepatocellular carcinoma (1–3). HCV is a member of the *Flaviviridae* family (4) and has a positive-stranded RNA genome of ~9.6 kb. This RNA contains an internal ribosome entry site (IRES) (5–7), which controls the initiation of translation of a large open reading frame encoding a polyprotein of ~3000 amino acids. Proteolytic cleavage of this polyprotein by host signal peptidases and two viral proteases produces four structural proteins (core, E1, E2 and P7) and six nonstructural proteins (NS2, NS3, NS4A, NS4B, NS5A and NS5B) [reviewed in (8–10)].

It was also observed that the core-encoding region of HCV expresses an additional protein of about 16–17 kDa at a low level (11–13). A conserved open reading frame in the +1 reading frame relative to the polyprotein was subsequently observed in the core-encoding region and it was shown that the sera of HCV-infected patients reacted against peptides corresponding to this +1 reading frame (14). Xu *et al*. (15) showed that this protein, which was named the F protein, was produced by a +1 translational frameshift by ribosomes that initiated translation at the AUG codon used for the synthesis of the HCV polyprotein and shifted the reading frame between codons 9 and 11 of this polyprotein. This mechanism was deduced from the N-terminal sequencing of the translation product generated when translating the beginning of the HCV-core region in a rabbit reticulocyte lysate (RRL). This +1 frameshift was observed with a HCV sequence that contains a stretch of 10 adenines (10A) encompassing the proposed frameshift site, a sequence that is underrepresented among the viral variants (2 out of 721 sequences available at the hepatitis virus database: http://s2as02.genes.nig.ac.jp/).

*To whom correspondence should be addressed. Tel: +1 514 343 6316; Fax: +1 514 343 2210; Email: lea.brakier.gingras@umontreal.ca

However, a stretch of A in a messenger RNA is known to promote frameshifting at a high efficiency *in vitro* (16), but much less in cultured mammalian cells where the control of the translation accuracy and of the maintenance of the reading frame is much more stringent (17). Also, antibodies against the F protein can be detected in patients infected with any HCV genotype, whether the 10A stretch is present or not (18,19). In an independent study, Boulant *et al*. (20) sequenced translational products generated by expressing in *Escherichia coli* a HCV fragment coding for the core protein and they also concluded that the F protein results from a +1 ribosomal frameshift, but at codon 42 of the polyprotein, an AGG codon. This codon is a rare codon in bacteria but not in humans (codon usage database: http://www.kazusa.or.jp/codon/). It could therefore promote a +1 frameshift by slowing down the elongating ribosomes in bacteria (17), but this does not hold for humans. Lastly, Vassilaki and Mavromara (21), when expressing a fusion of the HCV core-coding sequence to a luciferase reporter gene in cultured cells, detected another HCV protein, smaller than the protein generated by a +1 frameshift and showed that this protein results from the initiation of translation in a +1 reading frame relative to the polyprotein, at an AUG codon overlapping codon 86 or 88 of this polyprotein. All these observations raise questions on the mechanism that account for the synthesis of the F protein and led us to reassess this mechanism by investigating the expression of a luciferase reporter with an N-terminal insertion encompassing the beginning of the HCV-coding sequence *in vitro* and in cultured cells. In the site that was proposed by Xu *et al*. (15) to promote a +1 ribosomal frameshift, we introduced either the consensus sequence (AAAGAAAAAC) found in most HCV variants or the previously used stretch of 10A. We found that the previously reported +1 frameshift was artificially caused by the presence of the 10A stretch and that the synthesis of the F protein results from a direct initiation in the +1 reading frame at a non-AUG codon (GUG or GCG) overlapping codon 26 of the polyprotein.

MATERIALS AND METHODS

Construction of plasmids

All the plasmids used in this study were derived from pcDNA3.1-LUC, originating from pcDNA3.1/Hygro(+) (Invitrogen), where the firefly luciferase gene is inserted under the control of a CMV and a T7 promoter (22). Plasmid pCore1a-33, which contains the 5′-UTR plus the beginning of the coding sequence of HCV, was a gift from Dr H. Soudeyns (Hôpital Ste-Justine, Montréal) and was created from a viral sequence obtained from the plasma of a patient infected with HCV genotype 1a. This plasmid provided the HCV sequence used in this study. Portions of different length were amplified by PCR from this plasmid and inserted between the KpnI and BamHI sites of pcDNA3.1-LUC, creating the pHCV-LUC (0) series (Figure 2). These (0) constructs are such that the luciferase-coding sequence is in-frame with the HCV AUG initiation codon, so that only the ribosomes that translate in the reading frame of this AUG produce luciferase. For each pHCV-LUC (0) construct, a (+1) construct was created by inserting an adenine just after the BamHI site of the (0) construct. These (+1) constructs are such that the luciferase-coding sequence is in the +1 reading frame relative to the HCV AUG initiation codon, so that only ribosomes that translate in this +1 reading frame produce luciferase. Derivatives of pHCV-447-LUC (0) and (+1) constructs (the parental constructs that were used to generate deletion and substitution mutants) were created by PCR, by amplifying mutated DNA fragments with two primers for deletion mutants and with four primers for substitution mutants, according to the procedure of Ho *et al*. (23). The amplified DNA fragments were subcloned between the KpnI and BamHI sites of pcDNA3.1-LUC and all the constructs were verified by sequencing the entire insert.

Transient transfections and luciferase assays

Transfections of the plasmids of the pHCV-LUC series into HEK 293FT cells (Human Embryonic Kidney fibroblast-like cells; Invitrogen) were carried out as described previously (22), using a standard calcium phosphate precipitation method (24) with 3 μg of pHCV-LUC (0) or (+1) and 1 μg of pcDNA3.1/Hygro(+)/lacZ, which codes for β-galactosidase. For luciferase assays, 1 μl of a 600 μl cell extract was added to 25 μl of the Luciferase Assay Reagent (Promega) and the amount of light emitted was measured with a Berthold Lumat LB 9507 luminometer. The synthesis of the F protein relative to the synthesis of the polyprotein was calculated by dividing the luciferase activity of the (+1) construct by the luciferase activity of the corresponding (0) construct, unless otherwise indicated. The β-galactosidase activity was measured with the chlorophenolred-β-galactopyranoside substrate (Calbiochem) (25) with aliquots of 10 μl of cell extracts and it was used to normalize luciferase activities for variations in transfection efficiency.

In vitro transcription and translation

In vitro transcriptions were carried out essentially as described previously (22), using FbaI-linearized pHCV-LUC (0) and (+1) constructs. The RNA transcripts (0.2 μg) were translated in 25 μl of RRL (Promega) at 30°C for 20 min, a reaction time for which the translation system functions at its maximal rate. When the translation assays were made in the presence of edeine, the RRL was preincubated with the indicated concentration of this compound for 5 min at 37°C before the addition of the RNA transcripts. The reaction was stopped by the addition of EDTA to a final concentration of 6 mM. For luciferase assays, 2.5 μl of the translation mixture was added to 25 μl of the Luciferase Assay Reagent and the amount of light emitted was measured with a Berthold Lumat LB 9507 luminometer. The synthesis of the F protein relative to the synthesis of the polyprotein was calculated as described above.

RESULTS

Mapping of the HCV-coding region required to produce the F protein

To reassess the mechanism controlling the synthesis of the F protein from HCV RNA, we introduced fragments of different length of the HCV genotype 1a genome (Figure 1) before the coding sequence for the firefly luciferase gene (*luc*) harbored by a pcDNA3.1/Hygro(+) plasmid, under the control of a CMV and a T7 promoter. This generated a series of

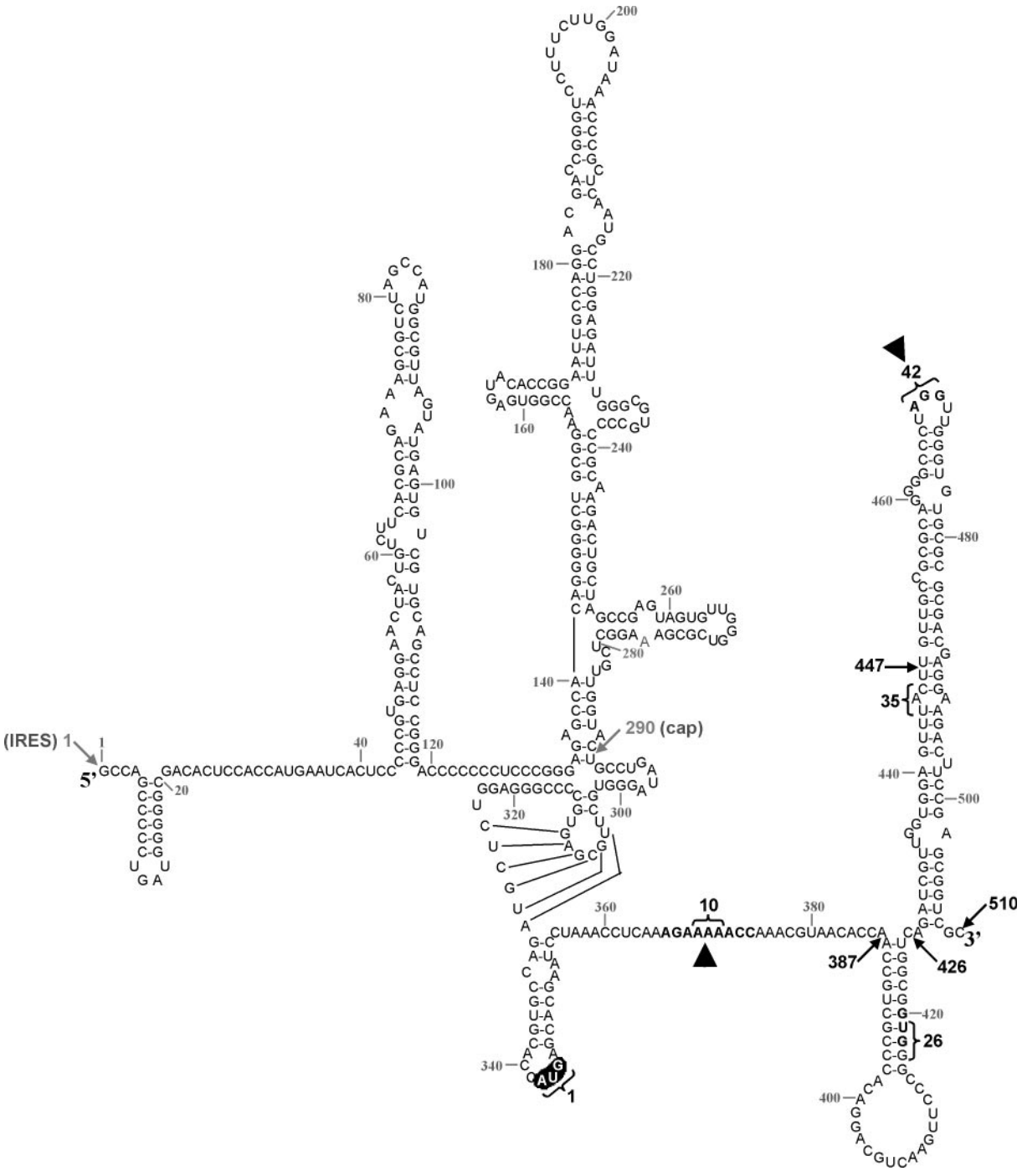

Figure 1. Sequence and secondary structure of the HCV IRES [based on reference (60)] and beginning of the HCV-coding sequence (nt 1–510 of genotype 1a). The initiation codon of the HCV polyprotein (codon 1, nt 342–344) is highlighted. The gray arrows point to the 5′ end (nt 1) of the IRES and to the 5′end (nt 290) of the HCV insertion in the constructs without IRES (see Figure 2). The black arrows point to the 3′ end of segments of different length of the HCV-coding sequence that were used in this study. The number preceding the arrows corresponds to the last nucleotide of these segments. The sites where a +1 ribosomal frameshift was previously proposed to occur (see the text) are in bold and marked by arrowheads. The brackets indicate codons of the HCV polyprotein with their corresponding number, which are referred to in the text. The codon overlapping codon 26 in the +1 reading frame to which we also refer in the text is in bold. It was verified, using the *mfold* program (61), that the predicted secondary structure of the HCV sequence was not altered when it was fused to the luciferase-coding sequence.

pHCV-LUC (0) constructs, an example of which, pHCV-447-LUC (0), is shown in Figure 2A. In these constructs, the insertions are such that ribosomes initiate translation at the HCV AUG initiation codon (nt 342–344 of the HCV RNA) and translate the portion of HCV viral RNA and the fused in-frame *luc* sequence. Synthesis of luciferase in these (0) constructs mimics the expression of the HCV polyprotein. These constructs contained either the complete IRES, starting at nt 1 of HCV RNA (thereafter referred to as IRES constructs) or a small fragment of the 5′-UTR of HCV RNA, starting at nt 290

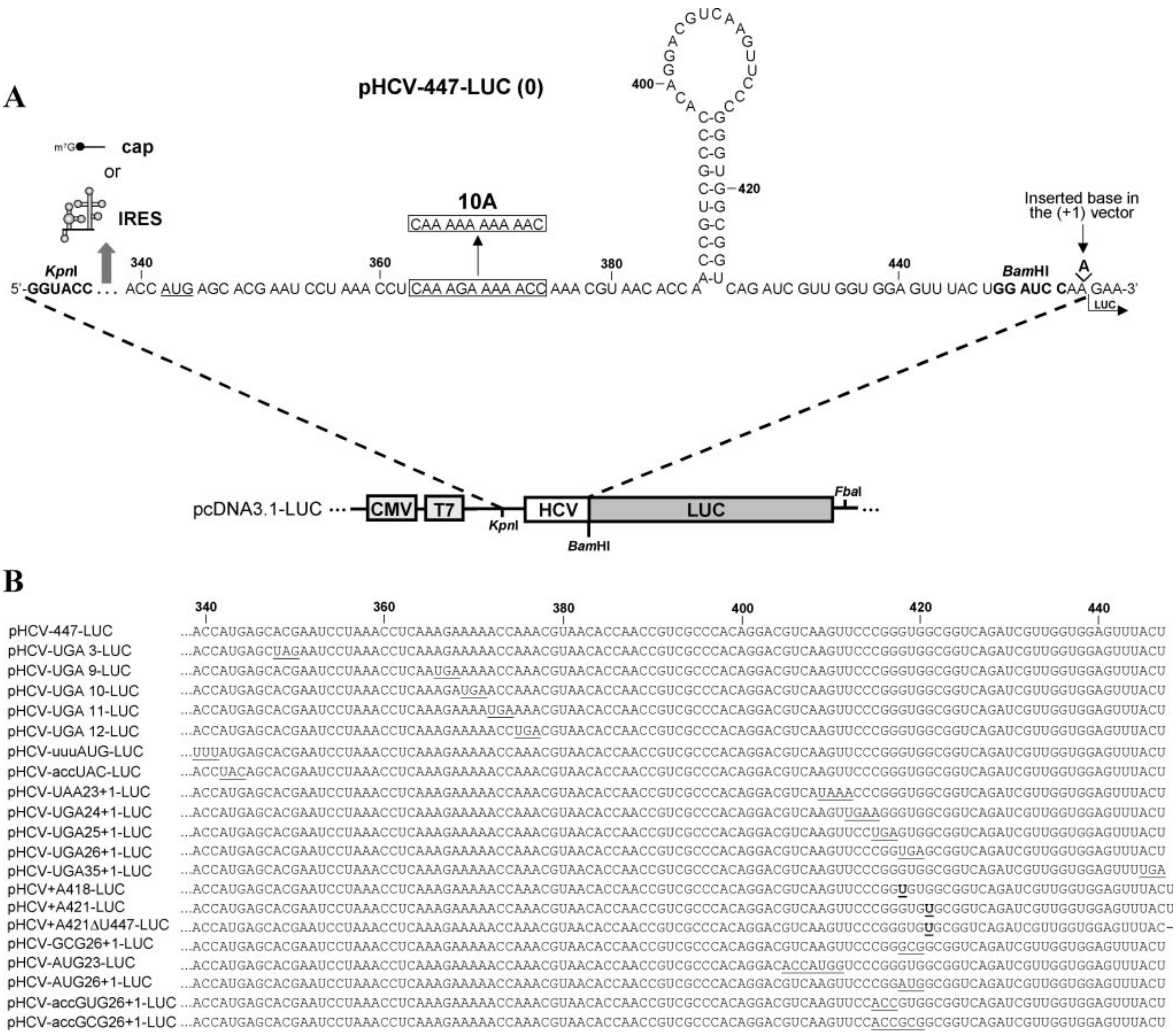

Figure 2. Description of the luciferase vectors used for the study of the expression of HCV F protein *in vitro* and in cultured cells. Nucleotides are numbered according to Figure 1. (**A**) The construct presented in this figure is pHCV-447-LUC, where a portion of the HCV-coding sequence extending to nt 447 is inserted upstream the coding sequence of the firefly luciferase reporter gene. Inserts of different length are: pHCV-387-LUC, pHCV-426-LUC, pHCV-447-LUC and pHCV-510-LUC, where the number indicates the last nucleotide of the insertion (see Figure 1). The AUG initiation codon (nt 342–344) of the HCV polyprotein is underlined. For the (0) constructs, the luciferase sequence is in-frame with this AUG initiation codon. For these (0) constructs, the expression of luciferase mimics the synthesis of the polyprotein. For the (+1) constructs, an adenine was added immediately after the BamHI restriction site (at position 455), so that only ribosomes translating the HCV sequence in the +1 reading frame relative to the HCV AUG initiator codon synthesize luciferase. With these (+1) constructs the expression of luciferase mimics the synthesis of the F protein. All the constructs were cloned by inserting a PCR product containing the investigated HCV sequence between the KpnI–BamHI sites of the pcDNA3.1-LUC vector. This HCV sequence was obtained from a plasmid that contains the 5′-UTR plus the beginning of the coding sequence of HCV genotype 1a (see Materials and Methods). In all the pHCV-LUC (0) and (+1) constructs, the HCV-coding sequence is preceded either by the complete HCV 5′-UTR (IRES constructs) or by a small segment (nt 290–341) of the HCV 5′-UTR region (cap constructs). The boxed sequence, which corresponds to a HCV consensus sequence, was mutated in pHCV-387-LUC, pHCV-426-LUC, pHCV-447-LUC and pHCV-510-LUC, creating derivatives with a stretch of 10A (10A constructs). (**B**) Sequences of the HCV-coding region of the derivatives of pHCV-447-LUC constructed in this study. Nucleotides that are mutated are underlined and nucleotides that are inserted and deleted are in bold and indicated by a dashed line, respectively.

(thereafter referred to as cap constructs). In the cap constructs, initiation of translation is mediated through recruitment of the 40S ribosomal subunits to the 5′ end of the mRNA and scanning of the 5′-UTR until the 40S subunits encounter the polyprotein AUG initiation codon [see (26–29) for reviews of cap-dependent initiation]. In the IRES constructs, translation initiation occurs through the HCV IRES-mediated binding of the 40S subunits directly to the polyprotein AUG initiation codon [see (6,30) for reviews of HCV IRES-dependent initiation]. For each (0) construct, a (+1) construct was made by adding an adenine before the *luc* coding sequence (Figure 2A), so that the luciferase-coding sequence is in the +1 reading frame relative to the HCV AUG initiation codon and that only ribosomes that translate the HCV viral RNA in the

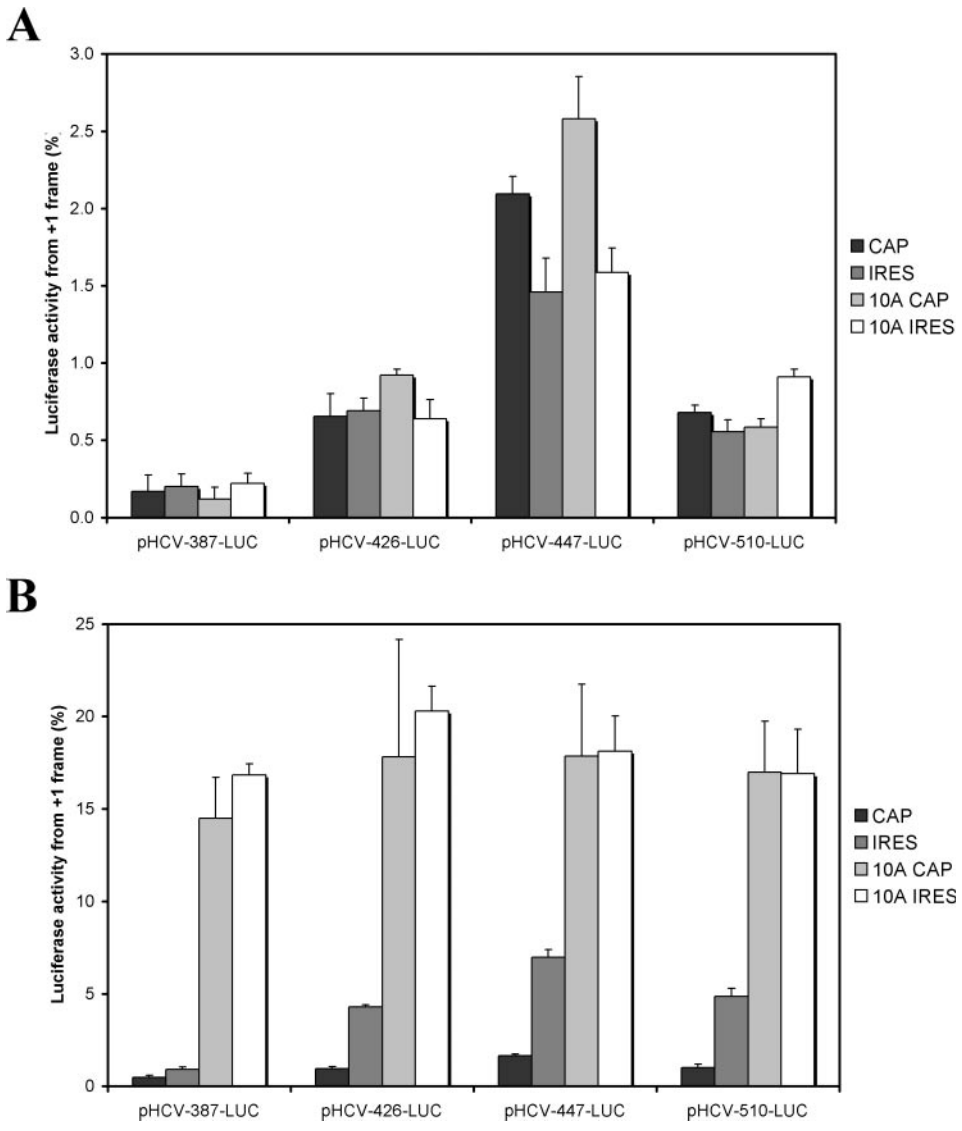

Figure 3. Synthesis of the F protein with constructs containing insertions of different length of the HCV-coding sequence. The synthesis of the F protein was measured with constructs containing HCV segments of different length described in Figure 2. In the pHCV-387-LUC, pHCV-426-LUC, pHCV-447-LUC and pHCV-510-LUC cap and IRES constructs, the consensus HCV sequence (AAAGAAAAAC, nt 364–373) occupies the site where a +1 frameshift was proposed to occur. In the 10A corresponding constructs, the consensus HCV sequence is replaced with a 10A stretch. (**A**) Synthesis of the F protein in cultured cells. Synthesis of the F protein was measured after co-transfection of 293FT cells with 3 μg of a pHCV-LUC (0) or (+1) construct and 1 μg of pcDNA3.1/Hygro(+)/lacZ, which is used to normalize for variations in transfection efficiency. (**B**) Synthesis of the F protein *in vitro*. *In vitro* translation experiments were carried out in 25 μl of RRL with 0.2 μg of mRNAs transcribed from the FbaI-digested pHCV-LUC constructs. Results are reported as the amount of F protein synthesized (assessed by the activity of luciferase in the +1 reading frame) relative to the amount of the polyprotein synthesized (assessed by the activity of luciferase in the 0 reading frame) and were calculated as described in the text. Each value represents the mean ± standard error of four to six independent experiments.

+1 reading frame synthesize luciferase. Synthesis of luciferase in these (+1) constructs mimics the expression of the HCV F protein. The synthesis of luciferase in the pHCV-LUC (0) and (+1) constructs was assessed in cultured cells and *in vitro* in an RRL.

We started by defining the HCV-coding region required to synthesize the maximal amount of the F protein. To this end, four series of (0) and (+1) cap and IRES constructs of different length were produced: pHCV-387-LUC, pHCV-426-LUC, pHCV-447-LUC and pHCV-510-LUC, where the number in the name of each construct indicates the nucleotide at the 3′ end of the HCV insertion (see Figure 1). In all these constructs the sequence encompassing codons 9–11 of the polyprotein, corresponding to nt 366–374 of the HCV viral RNA, is the consensus HCV sequence. Derivatives of all these constructs were made by mutagenesis, which contained the stretch of 10A (thereafter referred to as 10A constructs) that was previously proposed by Xu *et al*. (15) to mediate a +1 ribosomal frameshift (see Figure 2A).

Figure 3 shows the amount of the F protein (assessed by expression of luciferase in the +1 frame) synthesized in cultured cells (Figure 3A) or *in vitro* (Figure 3B) with the four series of constructs described above and with the corresponding 10A constructs. The synthesis of the F protein is expressed

relative to the amount of polyprotein. The background of expression, which is ~0.4% in cultured cells and ~1% *in vitro*, was subtracted from the results presented in the figures. This background was determined with a control construct where a UGA stop codon was inserted immediately preceding the *luc* coding sequence. In cultured cells with pHCV-387-LUC, the F protein was produced at 0.1–0.2% of the polyprotein for all members of this series (IRES or cap constructs with the consensus or 10A sequence), which is not significantly different from the background. Lengthening the HCV insertion in pHCV-426-LUC increased the relative synthesis of the F protein to 0.6–0.9% and an additional lengthening of the HCV insertion with nt 427–447 in pHCV-447-LUC further increased the synthesis of the F protein to 1.5–2.6% of the polyprotein. However, a longer HCV insertion in pHCV-510-LUC decreased the relative synthesis of the F protein to a level similar to that obtained with pHCV-426-LUC.

Our results with cultured cells show that the synthesis of the F protein occurs whether initiation of translation is cap-dependent or IRES-dependent and that the efficiency of this synthesis is maximal for pHCV-447-LUC, which contains the first 35 codons from the HCV-coding sequence. The level of expression was always low, even for pHCV-447-LUC, for which it is about 5-fold the background value, but it was reproducibly observed with independent constructs and DNA preparations. Our results also show that the presence of the 10A sequence used by Xu *et al*. (15) does not increase the amount of the F protein synthesized compared with the consensus sequence. This observation is in agreement with the fact that, as mentioned above, antibodies against the F protein were detected in patients infected with any HCV genotype regardless of the presence or the absence of the 10A stretch (18,19). We also assayed our constructs *in vitro* by translating in an RRL the various corresponding HCV-luciferase mRNAs obtained by transcription of the linearized plasmids. It can be seen (Figure 3B) that, with the consensus HCV sequence, the pattern of synthesis of the F protein is similar to that obtained in cultured cells, i.e. it does not differ significantly from the background for pHCV-387-LUC, increases with the pHCV-426-LUC and pHCV-447-LUC constructs, reaching a maximum level of 1.7 and 7.0% for the cap and IRES constructs, respectively, but decreases with pHCV-510-LUC to the levels observed with pHCV-426-LUC. However, with the constructs containing the 10A stretch, synthesis of the F protein is around 15–20%, independent of the length of the HCV insertion. *In vitro* translational systems are less accurate than cultured cells (17), as recalled in the Introduction. From comparing the results in cultured cells and *in vitro*, we conclude that the presence of the 10A stretch, a slippery sequence that is known to promote frameshifting *in vitro*, favors a +1 frameshift allowing the synthesis of the F protein in the +1 reading frame relative to the polyprotein. This efficient frameshift likely masks any other event occurring at a lower efficiency that could contribute to the synthesis of the F protein. The observation that the F protein was produced by a +1 ribosomal frameshift (15) appears to be linked to the use of a peculiar viral sequence and this prompted us to reassess the mechanism accounting for the synthesis of the F protein, using the consensus viral sequence and not the 10A stretch at the site previously proposed for frameshifting.

HCV F protein is not produced by a +1 ribosomal frameshift

To investigate the mechanism responsible for the synthesis of HCV F protein, we used pHCV-447-LUC, the construct with which the amount of the F protein synthesized is maximal. We first introduced mutations in pHCV-447-LUC (cap or IRES construct, with the consensus sequence for codons 9–11), by inserting a UGA stop codon in the HCV polyprotein reading frame. This stop codon was introduced at codon 3, 9, 10, 11 or 12, generating pHCV-UGA3-LUC, pHCV-UGA9-LUC, pHCV-UGA10-LUC, pHCV-UGA11-LUC and pHCV-UGA12-LUC (see Figure 2B for the description of these mutants and the other mutants used in the rest of this study). We reasoned that if the F protein was produced by a +1 ribosomal frameshift, inserting a stop codon in the HCV polyprotein reading frame should abolish translation in the +1 reading frame when it is inserted before the frameshift site, since no elongating ribosomes would then reach this frameshift site. However, assays in cultured cells showed that none of the inserted stop codon decreased the synthesis of the F protein (Figure 4), whereas all the inserted stop codons severely impaired the synthesis of the polyprotein (data not shown). The same observation was made with translation assays in an RRL (data not shown). These results therefore indicate that the F protein is not produced by elongating ribosomes that initiated translation at the polyprotein AUG initiation codon and shifted the reading frame at a specific site.

A decrease in the efficiency of the initiation of translation of HCV polyprotein favors the synthesis of the F protein

If the F protein is not synthesized by a +1 ribosomal frameshift, a likely hypothesis is that it results from the initiation of translation in the +1 reading frame relative to the reading frame of the HCV polyprotein. If this is the case, the translational start site for the polyprotein and the one for the F protein must compete for the amount of available 40S ribosomal subunits. To get further insight into the mechanism of synthesis of the F protein, we made two derivatives of pHCV-447-LUC. The first one is pHCV-uuuAUG-LUC, for which the 3 nt upstream of the AUG initiator codon were mutated so as to change its context from an optimal context for initiation of translation, according to Kozak (31), to a weak context. The second one is pHCV-accUAC-LUC, for which the initiator codon for the polyprotein was mutated to UAC so as to impair recognition by Met-tRNA$_i$. It can be seen in Figure 5 that the translation efficiency of the polyprotein was decreased by 5- to 10-fold with pHCV-uuuAUG-LUC and by 15- to 50-fold with pHCV-accUAC-LUC in the IRES and cap constructs, respectively. However, the translation of the F protein increased by 2- to 3-fold in pHCV-uuuAUG-LUC and pHCV-accUAC-LUC, whether the initiation of translation was cap or IRES dependent. These results confirm that the synthesis of the F protein does not result from a +1 frameshift by elongating ribosomes, since reducing the synthesis of the polyprotein did not decrease the amount of the F protein produced but rather increased it. They support the hypothesis that the synthesis of the F protein results from an initiation in the +1 reading frame relative to the AUG initiator codon used for the polyprotein.

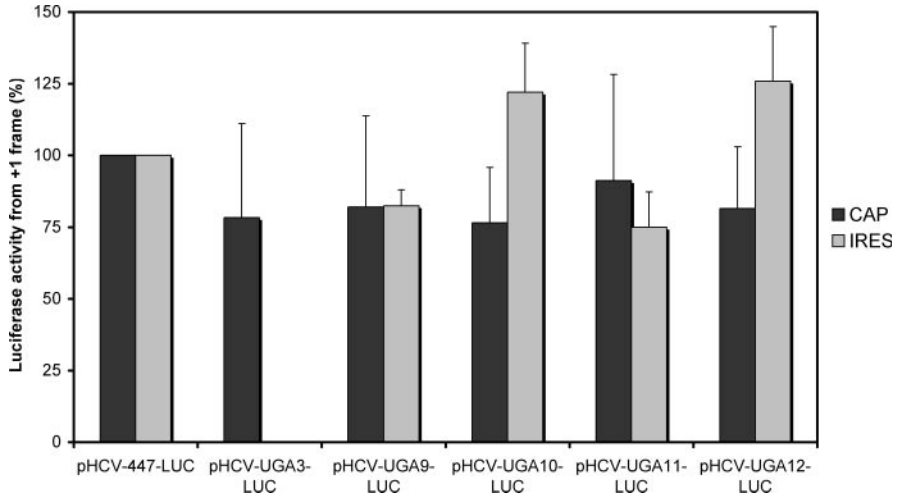

Figure 4. Synthesis of the F protein when a stop codon is introduced in the reading frame of the polyprotein. A UGA stop codon was introduced in pHCV-447-LUC, in the reading frame of the polyprotein, upstream or across the previously proposed +1 ribosomal frameshift site (15) at codon 3, 9, 10, 11 or 12. This generated pHCV-UGA3-LUC, pHCV-UGA9-LUC, pHCV-UGA10-LUC, pHCV-UGA11-LUC and pHCV-UGA12-LUC. All these constructs either contained or did not contain the HCV IRES (IRES and cap construct, respectively), except pHCV-UGA3-LUC for which the HCV IRES was not constructed since mutating nt 348–350 of the HCV RNA destabilizes a hairpin structure that is important for the IRES function (62). Assays were made in cultured cells as described in the legend to Figure 3. The synthesis of the F protein in the mutant constructs is expressed relative to that obtained with pHCV-447-LUC (+1), which is arbitrarily set at 100%. Each value represents the mean ± standard error of four to six independent experiments.

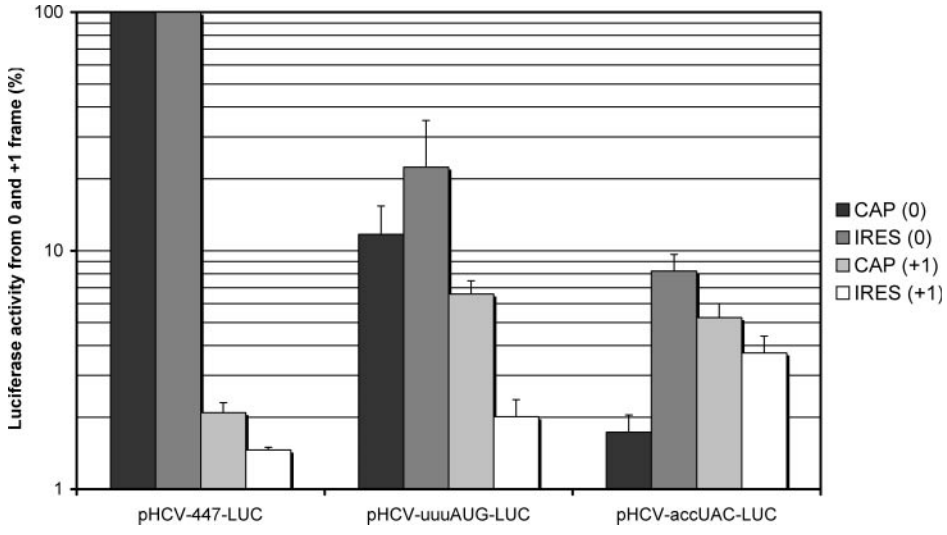

Figure 5. Synthesis of the polyprotein and the F protein when the AUG start codon of the polyprotein or its context is mutated. Derivatives of pHCV-447-LUC were constructed by mutating the AUG (nt 342–344) polyprotein start codon to UAC or its upstream context from ACC to UUU, generating pHCV-accUAC-LUC and pHCV-uuuAUG-LUC, respectively. Translation efficiencies are indicated on a logarithmic scale for the polyprotein and the F protein. The synthesis of the polyprotein measured with the mutant (0) constructs and that of the F protein measured with the (+1) constructs are expressed relative to the synthesis of the polyprotein with pHCV-447-LUC (0), which is arbitrarily set at 100%. Assays were made in cultured cells as described in the legend to Figure 3. Each value represents the mean ± standard error of four to six independent experiments.

Location of the F protein translation start site

The next step consists in determining where the translation initiation site for the F protein is located. There is no AUG codon in the +1 reading frame in the HCV-coding sequence inserted in pHCV-447-LUC, which suggests that the translation initiation for the F protein takes place at a non-AUG initiation codon. Initiation at a non-AUG codon is now well-documented in mammalian cells and in viruses [reviewed in (32–34)]. To locate the initiation codon for the F protein, we introduced a stop codon at various positions in the +1 reading frame of the HCV-coding sequence in pHCV-447-LUC, knowing that insertion of a stop codon in the +1 reading frame should impair translation of the F protein if it is positioned at the translational start site or downstream of it. We first narrowed the region encompassing the translation initiation site for the F protein by inserting a stop codon every tenth codon. We found that the synthesis of the F protein was not impaired when a stop codon was inserted at positions overlapping codon 10 or 20 of the polyprotein sequence in the +1 reading frame [codon 10(+1) or 20(+1)],

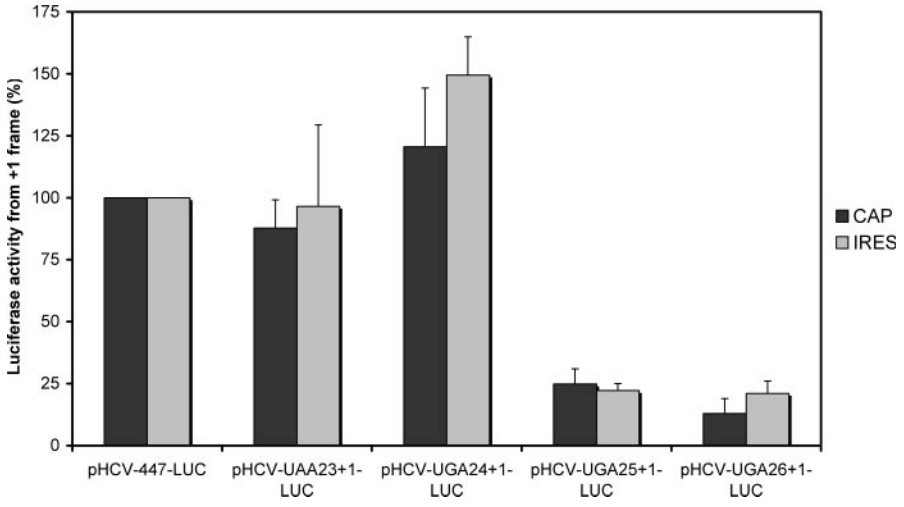

Figure 6. Identification of the translation start site of the F protein. A stop codon was introduced in pHCV-447-LUC at codons overlapping codon 23, 24, 25 or 26 of the polyprotein in the +1 reading frame, generating pHCV-UAA23+1-LUC, pHCV-UGA24+1-LUC, pHCV-UGA25+1-LUC and pHCV-UGA26+1-LUC. Assays were made in cultured cells, as described in the legend to Figure 3. The efficiency of the synthesis of the F protein from construct pHCV-447-LUC (+1) is arbitrarily set at 100%. Each value represents the mean ± standard error of four to six independent experiments.

but this synthesis was strongly decreased when the stop codon was inserted at codon 30(+1) (data not shown). We then introduced a stop codon in the +1 reading frame overlapping codon 23, 24, 25 or 26 of the polyprotein, generating pHCV-UAA23+1-LUC, pHCV-UGA24+1-LUC, pHCV-UGA25+1-LUC and pHCV-UGA26+1-LUC, respectively. In pHCV-UAA23+1-LUC and pHCV-UGA24+1-LUC, the nucleotide downstream of the inserted stop codon was mutated to an adenine to ensure efficient termination of translation at this codon (35). It can be seen (Figure 6) that, in cultured cells the synthesis of the F protein is not decreased with pHCV-UAA23+1-LUC and pHCV-UGA24+1-LUC, but, with pHCV-UGA25+1-LUC and pHCV- UGA26+1-LUC, it is reduced to about one-fourth or one-fifth, to a value near the background. This decrease was also observed with the same two constructs in an RRL (data not shown). An immediate conclusion would be that initiation of translation for the F protein occurs at the CGG codon overlapping codon 25 of the polyprotein. However, we considered the possibility that the GUG codon 26(+1) could also be the translation start site and that mutating codon 25(+1) could affect the F protein synthesis by altering the context of codon 26(+1).

To discriminate between codons 25(+1) and 26(+1) as candidates for the initiation of translation of the F protein, we made another series of mutants derived from pHCV-447-LUC (Figure 7A): an additional nucleotide (U) was added immediately after codon 25(+1) (pHCV+U418-LUC) or after codon 26(+1) (pHCV+U421-LUC). These insertions change the reading frame of the coding sequence of the reporter gene and should impair luciferase expression, which monitors the synthesis of the F protein if translation initiation of the F protein takes place before the inserted nucleotide. The synthesis of the F protein was not impaired with pHCV+U418-LUC, but was strongly reduced with pHCV+U421-LUC (Figure 7B), which indicates that the translational start site for the F protein is located at codon 26(+1). We also deleted nt 447 (U) in pHCV+U421-LUC, generating pHCV+U421ΔU447-LUC in which the luciferase sequence was replaced in the reading frame corresponding to codon 26(+1) (Figure 7B). In this mutant, the synthesis of the F protein was restored to its value in pHCV-447-LUC, confirming that the decrease in the synthesis of the F protein observed with pHCV+U421-LUC resulted from a change in the reading frame and that codon 26(+1) is used to initiate translation of the F protein.

Further information on the initiation of translation of the F protein

Codon 26(+1), which we identified as the start codon for the synthesis of the F protein, is GUG in pHCV-447-LUC. However, this codon is GUG in only 18% of the HCV sequences available at the hepatitis virus database, but it is GCG in 81% of the sequences. We therefore decided to mutate this GUG codon to GCG, generating pHCV-GCG26+1-LUC, and to investigate the synthesis of the F protein in the corresponding cap and IRES constructs. The synthesis of the F protein was not decreased in the IRES construct (Figure 8), revealing that translation initiation of the F protein can take place at codon 26(+1) of the polyprotein, whether it is GUG or GCG. However, with the corresponding cap construct, the synthesis of the F protein was strongly decreased.

It was previously shown that the 40S subunits that initiate translation with the HCV IRES cannot use a leaky scanning to initiate downstream from the polyprotein AUG initiator codon (36). This suggests that initiation of translation at codon 26(+1) should result from direct positioning of the 40S subunits at this codon when the initiation is IRES-mediated. To verify this suggestion, we introduced, in the 0 reading frame, an AUG with an optimal Kozak context (accAUGg) at codon 23 of the HCV polyprotein coding sequence for both the cap and IRES constructs of pHCV-447-LUC. With the pHCV-AUG23-LUC cap construct, the 40S subunits that scan the viral RNA after skipping the first initiator start site should initiate translation at the inserted AUG at codon 23, reducing the amount of 40S subunits available for an initiation at codon

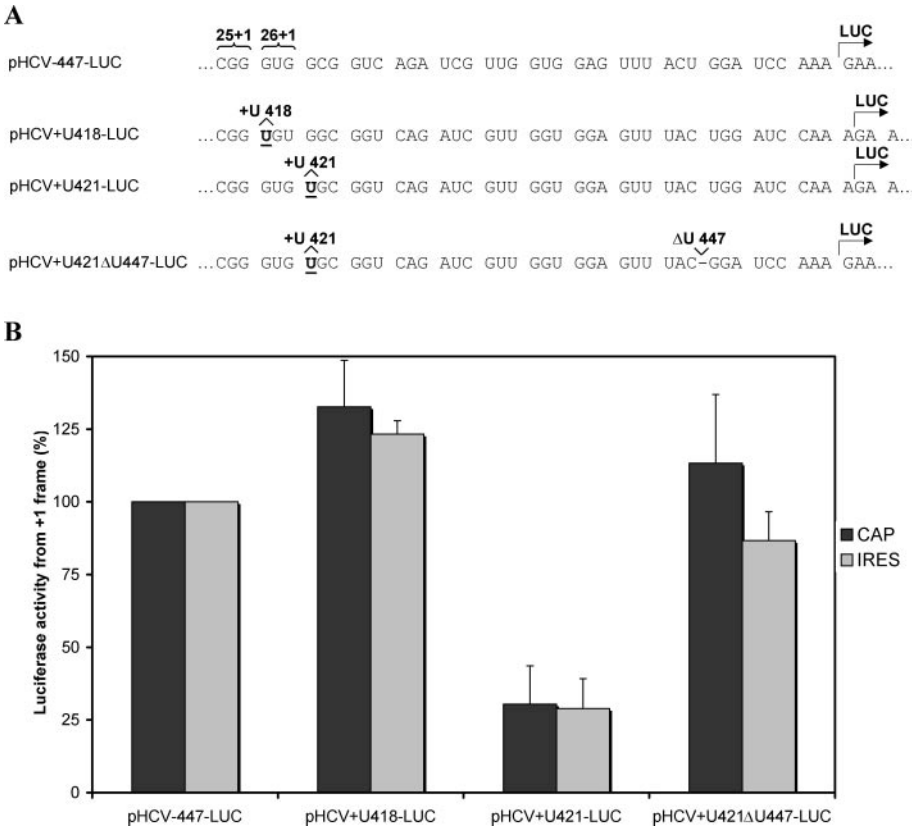

Figure 7. Confirmation of the location of the translation start site of the F protein. (**A**) Description of the mutants used for this confirmation. A nucleotide was added in pHCV-447-LUC (+1) immediately after codon 25(+1) or 26(+1), generating pHCV+U418-LUC and pHCV+U421-LUC, respectively, where the reading frame of luciferase that monitors the synthesis of the F protein is shifted by 1 nt relative to its reading frame in pHCV-447-LUC (+1). Nt 447 was deleted (dashed line) in pHCV+U421-LUC, generating pHCV+U421ΔU447-LUC, in which the luciferase reading frame is restored. (**B**) Assays were made in cultured cells as described in the legend to Figure 3. The efficiency of synthesis of the F protein from construct pHCV-447-LUC (+1) is arbitrarily set at 100%. Each value represents the mean ± standard error of four to six independent experiments.

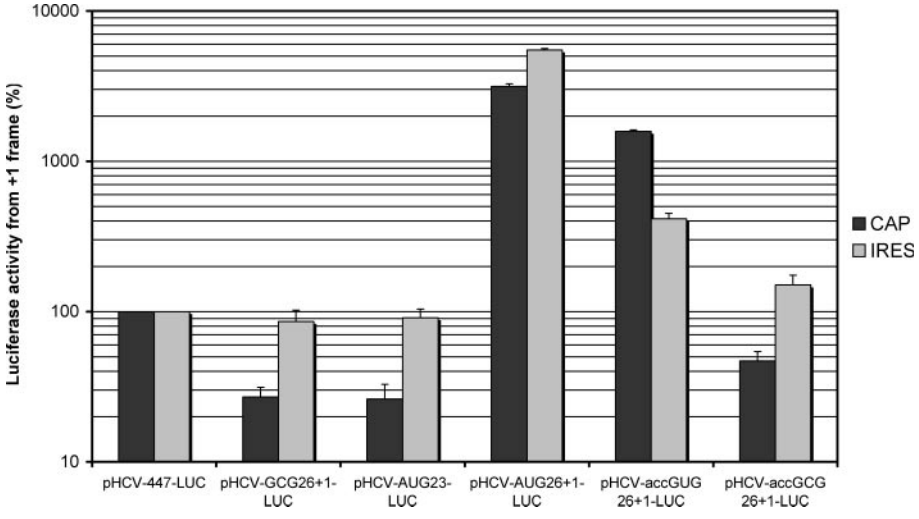

Figure 8. Further details on the initiation of translation of the F protein. In pHCV-GCG26+1-LUC, the initiation codon of the F protein, which is GUG in pHCV-447-LUC, was mutated to GCG. In pHCV-AUG23-LUC, an AUG codon with an optimal Kozak context for initiation was introduced in the 0 frame at codon 23 of pHCV-447-LUC. In pHCV-AUG26+1-LUC, the GUG initiation codon of the F protein was mutated to AUG. In pHCV-accGUG26+1-LUC and pHCV-accGCG26+1-LUC, the Kozak context of codon 26(+1) was mutated to ACC, the initiator codon being GUG and GCG, respectively. The synthesis of the F protein is indicated on a logarithmic scale. Assays were made in cultured cells, as described in the legend to Figure 3. The synthesis of the F protein from pHCV-447-LUC (+1) is arbitrarily set at 100%. Each value represents the mean ± standard error of four to six independent experiments.

26(+1). This is indeed the case for this cap construct (Figure 8), with which synthesis of the F protein was strongly reduced. However, the synthesis of the F protein remained unchanged for the pHCV-AUG23-LUC IRES construct, demonstrating that the initiation of translation occurs after direct positioning of the 40S subunits around codon 26(+1), as expected.

Initiation of translation with the initiator tRNA, Met-tRNA$_i$, at a non-AUG codon in the ribosomal P site is well documented with codons that differ from AUG by only 1 nt (32–34), such as the GUG codon identified here as the initiation site of translation of the F protein. However, to our knowledge, this is the first time that a Met-tRNA$_i$-mediated initiation is observed with an initiator codon that differs from AUG by 2 nt, such as the GCG codon found in this study. An alternative possibility could be that the synthesis of the F protein results from an initiation of translation in the A site of the ribosome without requiring the initiator Met-tRNA$_i$. With this mode of initiation, which was demonstrated in insect viruses (37,38), translation is initiated by the binding of an aminoacyl-tRNA cognate to the codon in the A site whereas the P site is occupied by a structure involving an RNA–RNA interaction between the codon in the P site and a complementary sequence in the viral RNA. In this Met-tRNA$_i$-independent mechanism, the efficiency of the initiation of translation should be independent of the nature of the start codon. However, in our study, when codon 26(+1) was mutated to AUG generating pHCV-AUG26+1-LUC, the synthesis of the F protein dramatically increased with the cap construct as well as with the IRES construct (Figure 8). This suggests that the initiation of translation of the F protein does not occur in the A site, but takes place in the P site with Met-tRNA$_i$ recognizing a non-AUG codon at position 26(+1), although less efficiently than an AUG codon. Furthermore, when the 3 nt upstream codon 26(+1) were mutated to ACC (pHCV-accGUG26+1-LUC), so as to create an optimal Kozak context for this codon, the synthesis of the F protein increased by 4- and 15-fold in the IRES and cap constructs, respectively (Figure 8). The effect of an optimal Kozak context was also observed when codon 26(+1) is GCG (pHCV-accGCG26+1-LUC), although to a much lesser extent. Although it was observed in one case that a Kozak context could also enhance initiation at the A site (39), this context generally enhances the initiation with Met-tRNA$_i$ at the ribosomal P site, in cap- and IRES-mediated translation (40). This effect of the Kozak context on the synthesis of the F protein thus also supports a translation initiation with Met-tRNA$_i$ in the P site.

Effect of edeine on translation initiation of the F protein

To confirm that translation initiation of the F protein is carried out with Met-tRNA$_i$ in the P site, we performed *in vitro* translation experiments in the presence of edeine with transcripts generated from IRES constructs. This compound is a peptide antibiotic, which interferes with initiation by impairing AUG codon recognition by the Met-tRNA$_i$ in the P site at concentrations below 1 μM (41,42). At concentrations higher than 1 μM, it also interferes with binding of tRNA to the A site (43,44). A previous report by Sarnow and his collaborators (38) demonstrated that the presence of 0.25 μM edeine reduced IRES-mediated initiation at the P site by 90%, while initiation of translation at the A site for the cricket paralysis virus was only inhibited by 20%. We found here that translation of the polyprotein is inhibited up to 95% when the concentration in edeine increases from 0.1 to 0.3 μM (Figure 9). A similar response to edeine is observed for the synthesis of the F protein, whether the 26(+1) start codon is GUG, GCG or AUG. These results fully support a translation initiation for the F protein at the P site.

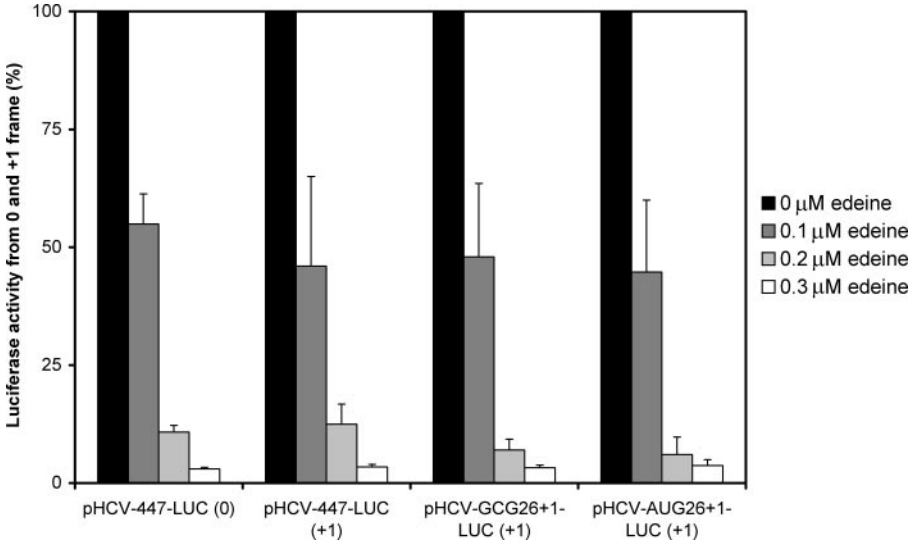

Figure 9. Inhibition of the synthesis of the polyprotein and the F protein by edeine. *In vitro* translation experiments were carried out in 25 μl of RRL containing 0, 0.1, 0.2 or 0.3 μM edeine, with 0.2 μg of mRNAs transcribed from the FbaI-digested pHCV-LUC IRES constructs. For each construct, the protein synthesis in the absence of edeine (0 μM) is arbitrarily set at 100%. Each value represents the mean ± standard error of three independent experiments.

DISCUSSION

In this study, we reassessed the mechanism accounting for the synthesis of the F protein of HCV, a protein whose role is still unknown and which is translated from the viral RNA in a +1 reading frame relative to the HCV polyprotein. It was previously proposed that the F protein results from a +1 ribosomal frameshift which takes place in the beginning of the core-coding region of the polyprotein and is made by a minority of ribosomes that initiated translation at the AUG start codon of the polyprotein (15). Our results demonstrate that this conclusion was inferred from *in vitro* translation assays using an underrepresented shifty sequence encompassing a stretch of 10A, which promotes this +1 ribosomal frameshift *in vitro*. Using the consensus HCV sequence (AAAGAAAAAC) instead of this shifty sequence, we showed that the synthesis of the F protein does not result from a ribosomal frameshift. We rather found that the initiation of the synthesis of the F protein takes place at a codon overlapping codon 26 of the polyprotein coding sequence in the +1 reading frame, codon 26(+1). Codon 26(+1) is GUG in 18% of the HCV sequences available and GCG in 81% of the sequences available. We found that a GUG codon at position 26(+1) is recognized as a translation start site by the ribosomes whether the initiation of translation is mediated by the HCV IRES or is cap-dependent, whereas initiation of translation at a GCG codon requires the presence of the IRES. When the initiation of translation is IRES-mediated, we showed that the 40S subunits that initiate translation at the start codon of the F protein do not reach this codon by scanning, but are directly positioned at this codon. We also observed that the synthesis of the F protein strongly increases when codon 26(+1) is mutated to AUG or when the context of this codon is mutated into an optimal Kozak context. Furthermore, at low concentrations of edeine that interfere with codon recognition by Met-tRNA$_i$ in the P site but do not impair binding of a tRNA to the A site, the synthesis of the F protein was severely inhibited. These observations are consistent with a Met-tRNA$_i$-dependent initiation of translation at the P site for the F protein.

When inserting segments of different length of the HCV-coding region before the coding sequence of the reporter gene, we found that the synthesis of the F protein requires the first 35 codons of the HCV polyprotein to reach its maximal efficiency. Under these conditions, the F protein is synthesized in cultured cells at an efficiency of ~2% relative to the synthesis of the polyprotein. Although the efficiency of translation of the F protein in the context of the complete viral genome is not known, we can assume that it is comparable with the value we observed. A low level of synthesis such as that we observe is not uncommon for viral regulatory proteins whose expression is regulated at the translational level. For example, the barley yellow dwarf virus (45) and the yeast L-A double-stranded RNA virus (46) use a −1 ribosomal frameshift with an efficiency of about 1 and 2%, respectively, to produce a polymerase required for their replication.

It is known that a Met-tRNA$_i$-dependent initiation of translation can occur at codons that differ from AUG by 1 nt (32–34), such as a GUG codon. However, the initiation of translation with Met-tRNA$_i$ at a GCG codon that differs from AUG by 2 nt is unprecedented. Contrasting with the fidelity of ribosomal decoding in the A site, information about the stringency of codon–anticodon pairing in the P site is rather limited. This stringency was analyzed thoroughly in translational bypassing, an event where ribosomes bearing a peptidyl-tRNA in the P site suspend translation at a given site by disrupting codon–anticodon interaction, scan the messenger and resume translation downstream. Most frequently, the codon where translation resumes is perfectly complementary to the anticodon of the peptidyl-tRNA, but it was observed that this complementarity is not always perfect [reviewed in (47)]. Depending on the local context of the codon where translation resumes, codon–anticodon complementarity can also involve 3 bp with a wobble base pair at any position or 2 out of 3 bp with a wobble base pair at one or both positions. It is interesting to observe that when initiation occurs at a GUG codon, pairing of this codon with the anticodon of Met-tRNA$_i$ (3′-UAC) implies 3 bp with a G-U wobble at the first position. With GCG, the pairing forms 2 out of 3 bp with also a G-U wobble at the first position.

We showed in this study that, when codon 26(+1) is GUG, it is selected as a translation start site whether the initiation is IRES- or cap-dependent. This indicates that there must be some peculiar features in the sequence proximal to this codon that favor translation initiation at this position. With pHCV-426-LUC, which encompasses the first 28 codons of the polyprotein sequence, the F protein was synthesized at a very low efficiency. Lengthening the HCV insertion so as to encompass the first 35 codons of the polyprotein (pHCV-447-LUC), stimulated synthesis of the F protein by ~3-fold. It can be suggested that the short segment encompassing nt 427–447 contains a signal that enhances the selection of codon 26(+1) for initiation of translation. We can propose that this signal could base pair to a complementary sequence within the 18S ribosomal RNA of the 40S subunit or participate in a specific secondary structure, causing in either case a distortion of the P site that would facilitate the decoding of a non-AUG codon by the Met-tRNA$_i$. This selection signal would be trapped in a long irregular helix in pHCV-510-LUC (see Figure 1, right), accounting for the decrease in the synthesis of the F protein observed with this construct. When the complete viral genome is present, a conformational switch between alternative RNA structures could expose codon 26(+1) and its selection signal, controlling the efficiency of initiation of the synthesis of the F protein. When codon 26(+1) is GCG, the additional presence of the HCV IRES is required for the initiation of translation of the F protein. Upon binding to the HCV IRES, the 40S subunit undergoes drastic conformational changes (48). We hypothesize that the simultaneous interaction with the HCV IRES and with a signal proximal to codon 26(+1) induces changes in the P site of the 40S subunit that stabilize the pairing between the anticodon of Met-tRNA$_i$ and the GCG codon, although this pairing is not perfect. Such an effect of the HCV IRES is supported by the fact that initiation of translation of the polyprotein is marginally altered when the AUG initiation codon is mutated to AUU or CUG (49).

Using the translational start site defined by our study, the molecular weight of the F protein can be calculated to be 14.4 kDa. As mentioned in the Introduction, Vassilaki and Mavromara (21), using a fusion of the beginning of the HCV-coding sequence to the luciferase gene, observed the synthesis of a novel HCV protein in cultured cells. This protein of ~8 kDa resulted from initiation at an AUG codon

overlapping codon 86 or 88 of the polyprotein, but in a +1 reading frame. The constructs used in the present study do not encompass the region containing these two AUG codons, but when we lengthened the inserted segment of the HCV-coding sequence so as to include these two AUGs, synthesis of luciferase in the +1 frame increased 4- to 5-fold (data not shown). This result is consistent with the observation of Vassilaki and Mavromara (21) that an additional protein, shorter than the F protein, could be synthesized from the +1 reading frame of the HCV polyprotein. Whether this short protein is actually synthesized in HCV-infected cells remains to be determined.

Finally, although the function of the F protein is unknown, it is well documented that viral proteins whose expression is regulated at the translational level, whether by frameshift, readthrough of a stop codon or alternative initiation, usually play an important role in viral replication (50–53). The F protein is not required for HCV RNA replication since HCV subgenomic RNA replicons can replicate in its absence (54–57). Interestingly, initiation of the synthesis of the F protein is reminiscent of the situation encountered with the synthesis of the L* protein in Theiler's murine encephalomyelitis virus. This picornavirus possesses a genome organization related to HCV and the L* protein is produced by an IRES-mediated initiation in the +1 reading frame of the polyprotein 13 nt downstream of the polyprotein AUG start site (58). The L* protein can be expressed from an AUG or an ACG codon and is essential for viral persistence (59). The similarity in the mechanisms accounting for the initiation of the synthesis of the L* and the F protein leads us to hypothesize that the F protein could be involved in HCV persistence.

ACKNOWLEDGEMENTS

This study was supported by a grant from the Canadian Institutes for Health Research. M.B. is a recipient of a scholarship from the Fonds de Recherche en Santé du Québec (FRSQ). We are grateful to Hugo Soudeyns for the generous gift of plasmid pCore1a-33 and to Jerry Pelletier for the gift of the edeine. We thank Hemy Chonaim for his participation to this project as a summer student. We also thank Guy Boileau, Dominic Dulude, Gerardo Ferbeyre and Nikolaus Heveker for stimulating discussions and critical reading of this manuscript. Funding to pay the Open Access publication charges for this article was provided by Canadian Institutes for Health Research.

REFERENCES

1. Lefkowitch,J.H., Schiff,E.R., Davis,G.L., Perrillo,R.P., Lindsay,K., Bodenheimer,H.C., Balart,L.A., Ortego,T.J., Payne,J. and Dienstag,J.L. (1993) Pathological diagnosis of chronic hepatitis C: a multicenter comparative study with chronic hepatitis B. The Hepatitis Interventional Therapy Group. *Gastroenterology*, **104**, 595–603.
2. Saito,I., Miyamura,T., Ohbayashi,A., Harada,H., Katayama,T., Kikuchi,S., Watanabe,Y., Koi,S., Onji,M., Ohta,Y. *et al.* (1990) Hepatitis C virus infection is associated with the development of hepatocellular carcinoma. *Proc. Natl Acad. Sci. USA*, **87**, 6547–6549.
3. Scheuer,P.J., Ashrafzadeh,P., Sherlock,S., Brown,D. and Dusheiko,G.M. (1992) The pathology of hepatitis C. *Hepatology*, **15**, 567–571.
4. Choo,Q.L., Richman,K.H., Han,J.H., Berger,K., Lee,C., Dong,C., Gallegos,C., Coit,D., Medina-Selby,R., Barr,P.J. *et al.* (1991) Genetic organization and diversity of the hepatitis C virus. *Proc. Natl Acad. Sci. USA*, **88**, 2451–2455.
5. Reynolds,J.E., Kaminski,A., Carroll,A.R., Clarke,B.E., Rowlands,D.J. and Jackson,R.J. (1996) Internal initiation of translation of hepatitis C virus RNA: the ribosome entry site is at the authentic initiation codon. *RNA*, **2**, 867–878.
6. Rijnbrand,R.C. and Lemon,S.M. (2000) Internal ribosome entry site-mediated translation in hepatitis C virus replication. *Curr. Top. Microbiol. Immunol.*, **242**, 85–116.
7. 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.
8. Kato,N. (2000) Genome of human hepatitis C virus (HCV): gene organization, sequence diversity, and variation. *Microb. Comp. Genomics*, **5**, 129–151.
9. Reed,K.E. and Rice,C.M. (2000) Overview of hepatitis C virus genome structure, polyprotein processing, and protein properties. *Curr. Top. Microbiol. Immunol.*, **242**, 55–84.
10. Rosenberg,S. (2001) Recent advances in the molecular biology of hepatitis C virus. *J. Mol. Biol.*, **313**, 451–464.
11. Lo,S.Y., Selby,M., Tong,M. and Ou,J.H. (1994) Comparative studies of the core gene products of two different hepatitis C virus isolates: two alternative forms determined by a single amino acid substitution. *Virology*, **199**, 124–131.
12. Lo,S.Y., Masiarz,F., Hwang,S.B., Lai,M.M. and Ou,J.H. (1995) Differential subcellular localization of hepatitis C virus core gene products. *Virology*, **213**, 455–461.
13. Ray,R.B., Lagging,L.M., Meyer,K. and Ray,R. (1996) Hepatitis C virus core protein cooperates with ras and transforms primary rat embryo fibroblasts to tumorigenic phenotype. *J. Virol.*, **70**, 4438–4443.
14. Walewski,J.L., Keller,T.R., Stump,D.D. and Branch,A.D. (2001) Evidence for a new hepatitis C virus antigen encoded in an overlapping reading frame. *RNA*, **7**, 710–721.
15. Xu,Z., Choi,J., Yen,T.S., Lu,W., Strohecker,A., Govindarajan,S., Chien,D., Selby,M.J. and Ou,J. (2001) Synthesis of a novel hepatitis C virus protein by ribosomal frameshift. *EMBO J.*, **20**, 3840–3848.
16. Weiss,R.B., Dunn,D.M., Atkins,J.F. and Gesteland,R.F. (1987) Slippery runs, shifty stops, backward steps, and forward hops: −2, −1, +1, +2, +5, and +6 ribosomal frameshifting. *Cold Spring Harb. Symp. Quant. Biol.*, **52**, 687–693.
17. Farabaugh,P.J. (1997) Programmed +1 frameshifting. In Farabaugh,P.J. (ed.), *Programmed Alternative Reading of the Genetic Code*. R.G. Landes, Austin, TX, pp. 41–67.
18. Bain,C., Parroche,P., Lavergne,J.P., Duverger,B., Vieux,C., Dubois,V., Komurian-Pradel,F., Trepo,C., Gebuhrer,L., Paranhos-Baccala,G. *et al.* (2004) Memory T-cell-mediated immune responses specific to an alternative core protein in hepatitis C virus infection. *J. Virol.*, **78**, 10460–10469.
19. Yeh,C.T., Lo,S.Y., Dai,D.I., Tang,J.H., Chu,C.M. and Liaw,Y.F. (2000) Amino acid substitutions in codons 9–11 of hepatitis C virus core protein lead to the synthesis of a short core protein product. *J. Gastroenterol. Hepatol.*, **15**, 182–191.
20. Boulant,S., Becchi,M., Penin,F. and Lavergne,J.P. (2003) Unusual multiple recoding events leading to alternative forms of hepatitis C virus core protein from genotype 1b. *J. Biol. Chem.*, **278**, 45785–45792.
21. Vassilaki,N. and Mavromara,P. (2003) Two alternative translation mechanisms are responsible for the expression of the HCV ARFP/F/core+1 coding open reading frame. *J. Biol. Chem.*, **278**, 40503–40513.
22. Dulude,D., Baril,M. and Brakier-Gingras,L. (2002) Characterization of the frameshift stimulatory signal controlling a programmed −1 ribosomal frameshift in the human immunodeficiency virus type 1. *Nucleic Acids Res.*, **30**, 5094–5102.
23. Ho,S.N., Hunt,H.D., Horton,R.M., Pullen,J.K. and Pease,L.R. (1989) Site-directed mutagenesis by overlap extension using the polymerase chain reaction. *Gene*, **77**, 51–59.
24. Jordan,M., Schallhorn,A. and Wurm,F.M. (1996) Transfecting mammalian cells: optimization of critical parameters affecting calcium-phosphate precipitate formation. *Nucleic Acids Res.*, **24**, 596–601.
25. Eustice,D.C., Feldman,P.A., Colberg-Poley,A.M., Buckery,R.M. and Neubauer,R.H. (1991) A sensitive method for the detection of beta-galactosidase in transfected mammalian cells. *Biotechniques*, **11**, 739–740742–743.
26. Dever,T.E. (2002) Gene-specific regulation by general translation factors. *Cell*, **108**, 545–556.

27. Hershey,J.W.B., Mathews,M. and Sonenberg,N. (1996) Cold Spring Harbor monograph series, 30. *Translational Control*. Cold Spring Harbor Laboratory Press, Plainview, NY, pp. xi, 794.
28. Pestova,T.V. and Hellen,C.U. (2001) Functions of eukaryotic factors in initiation of translation. *Cold Spring Harb. Symp. Quant. Biol.*, **66**, 389–396.
29. Poulin,F. and Sonenberg,N. (2003) Mechanism of translation initiation in eukaryotes. In Lapointe,J. and Brakier-Gingras,L. (eds), *Translation Mechanisms*. Landes Bioscience, Georgetown, TX, pp. 280–297.
30. Hellen,C.U. and Pestova,T.V. (1999) Translation of hepatitis C virus RNA. *J. Viral Hepat.*, **6**, 79–87.
31. Kozak,M. (1986) Point mutations define a sequence flanking the AUG initiator codon that modulates translation by eukaryotic ribosomes. *Cell*, **44**, 283–292.
32. Kozak,M. (1999) Initiation of translation in prokaryotes and eukaryotes. *Gene*, **234**, 187–208.
33. Peabody,D.S. (1989) Translation initiation at non-AUG triplets in mammalian cells. *J. Biol. Chem.*, **264**, 5031–5035.
34. Touriol,C., Bornes,S., Bonnal,S., Audigier,S., Prats,H., Prats,A.C. and Vagner,S. (2003) Generation of protein isoform diversity by alternative initiation of translation at non-AUG codons. *Biol. Cell.*, **95**, 169–178.
35. Tate,W.P., Poole,E.S., Horsfield,J.A., Mannering,S.A., Brown,C.M., Moffat,J.G., Dalphin,M.E., McCaughan,K.K., Major,L.L. and Wilson,D.N. (1995) Translational termination efficiency in both bacteria and mammals is regulated by the base following the stop codon. *Biochem. Cell Biol.*, **73**, 1095–1103.
36. Rijnbrand,R.C., Abbink,T.E., Haasnoot,P.C., Spaan,W.J. and Bredenbeek,P.J. (1996) The influence of AUG codons in the hepatitis C virus 5′ nontranslated region on translation and mapping of the translation initiation window. *Virology*, **226**, 47–56.
37. Sasaki,J. and Nakashima,N. (2000) Methionine-independent initiation of translation in the capsid protein of an insect RNA virus. *Proc. Natl Acad. Sci. USA*, **97**, 1512–1515.
38. Wilson,J.E., Pestova,T.V., Hellen,C.U. and Sarnow,P. (2000) Initiation of protein synthesis from the A site of the ribosome. *Cell*, **102**, 511–520.
39. Schwab,S.R., Shugart,J.A., Horng,T., Malarkannan,S. and Shastri,N. (2004) Unanticipated antigens: translation initiation at CUG with leucine. *PLoS Biol.*, **2**, 1774–1784.
40. Martinez-Salas,E. (1999) Internal ribosome entry site biology and its use in expression vectors. *Curr. Opin. Biotechnol.*, **10**, 458–464.
41. Odon,O.W., Kramer,G., Henderson,A.B., Pinphanichakarn,P. and Hardesty,B. (1978) GTP hydrolysis during methionyl-tRNAf binding to 40 S ribosomal subunits and the site of edeine inhibition. *J. Biol. Chem.*, **253**, 1807–1816.
42. Kozak,M. and Shatkin,A.J. (1978) Migration of 40 S ribosomal subunits on messenger RNA in the presence of edeine. *J. Biol. Chem.*, **253**, 6568–6577.
43. Carrasco,L., Battaner,E. and Vazquez,D. (1974) The elongation steps in protein synthesis by eukaryotic ribosomes: effects of antibiotics. *Methods Enzymol.*, **30**, 282–289.
44. Szer,W. and Kurylo-Borowska,Z. (1970) Effect of edeine on aminoacyl-tRNA binding to ribosomes and its relationship to ribosomal binding sites. *Biochim. Biophys. Acta*, **224**, 477–486.
45. Barry,J.K. and Miller,W.A. (2002) A −1 ribosomal frameshift element that requires base pairing across four kilobases suggests a mechanism of regulating ribosome and replicase traffic on a viral RNA. *Proc. Natl Acad. Sci. USA*, **99**, 11133–11138.
46. Dinman,J.D. and Wickner,R.B. (1992) Ribosomal frameshifting efficiency and gag/gag-pol ratio are critical for yeast M1 double-stranded RNA virus propagation. *J. Virol.*, **66**, 3669–3676.
47. Herr,A.J., Atkins,J.F. and Gesteland,R.F. (2000) Coupling of open reading frames by translational bypassing. *Annu. Rev. Biochem.*, **69**, 343–372.
48. Spahn,C.M., Kieft,J.S., Grassucci,R.A., Penczek,P.A., Zhou,K., Doudna,J.A. and Frank,J. (2001) Hepatitis C virus IRES RNA-induced changes in the conformation of the 40s ribosomal subunit. *Science*, **291**, 1959–1962.
49. Reynolds,J.E., Kaminski,A., Kettinen,H.J., Grace,K., Clarke,B.E., Carroll,A.R., Rowlands,D.J. and Jackson,R.J. (1995) Unique features of internal initiation of hepatitis C virus RNA translation. *EMBO J.*, **14**, 6010–6020.
50. Brierley,I. and Pennel,S. (2001) Structure and function of the stimulatory RNAs involved in programmed eukaryotic −1 ribosomal frameshifting. In Stillman,B. (ed.), *The Ribosome*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, pp. 233–248.
51. Farabaugh,P.J. (1997) Programmed readthrough of translational termination codons. In Farabaugh,P.J. (ed.), *Programmed Alternative Reading of the Genetic Code*. R.G. Landes, Austin, TX, pp. 149–182.
52. Farabaugh,P.J. (2000) Translational frameshifting: implications for the mechanism of translational frame maintenance. *Prog. Nucleic Acid Res. Mol. Biol.*, **64**, 131–170.
53. Sarnow,P. (2003) Viral internal ribosome entry site elements: novel ribosome-RNA complexes and roles in viral pathogenesis. *J. Virol.*, **77**, 2801–2806.
54. Blight,K.J., Kolykhalov,A.A. and Rice,C.M. (2000) Efficient initiation of HCV RNA replication in cell culture. *Science*, **290**, 1972–1974.
55. Guo,J.T., Bichko,V.V. and Seeger,C. (2001) Effect of alpha interferon on the hepatitis C virus replicon. *J. Virol.*, **75**, 8516–8523.
56. Ikeda,M., Yi,M., Li,K. and Lemon,S.M. (2002) Selectable subgenomic and genome-length dicistronic RNAs derived from an infectious molecular clone of the HCV-N strain of hepatitis C virus replicate efficiently in cultured Huh7 cells. *J. Virol.*, **76**, 2997–3006.
57. Lohmann,V., Korner,F., Koch,J., Herian,U., Theilmann,L. and Bartenschlager,R. (1999) Replication of subgenomic hepatitis C virus RNAs in a hepatoma cell line. *Science*, **285**, 110–113.
58. van Eyll,O. and Michiels,T. (2002) Non-AUG-initiated internal translation of the L* protein of Theiler's virus and importance of this protein for viral persistence. *J. Virol.*, **76**, 10665–10673.
59. van Eyll,O. and Michiels,T. (2000) Influence of the Theiler's virus L* protein on macrophage infection, viral persistence, and neurovirulence. *J. Virol.*, **74**, 9071–9077.
60. Honda,M., Beard,M.R., Ping,L.H. and Lemon,S.M. (1999) A phylogenetically conserved stem-loop structure at the 5′ border of the internal ribosome entry site of hepatitis C virus is required for cap-independent viral translation. *J. Virol.*, **73**, 1165–1174.
61. Zuker,M. (2003) Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Res.*, **31**, 3406–3415.
62. Honda,M., Brown,E.A. and Lemon,S.M. (1996) Stability of a stem-loop involving the initiator AUG controls the efficiency of internal initiation of translation on hepatitis C virus RNA. *RNA*, **2**, 955–968.