Journal of General Virology (2003), **84**, 1751–1759 DOI 10.1099/vir.0.19065-0

Characterization of the expression of the hepatitis C virus F protein

Juliette Roussel,[1,2] André Pillez,[1] Claire Montpellier,[1] Gilles Duverlie,[2] Annie Cahour,[3] Jean Dubuisson[1] and Czeslaw Wychowski[1]

[1]CNRS-UPR 2511, IBL/Institut Pasteur de Lille, 59021 Lille Cedex, France

[2]Laboratoire de Virologie, Centre Hospitalier Universitaire-Hôpital Sud, 80054 Amiens Cedex, France

[3]CERVI (Virologie), UPRES EA 2387, Hôpital Pitié-Salpêtrière, 75651 Paris Cedex 13, France

Correspondence
Czeslaw Wychowski
czeslaw.wychowski@ibl.fr

Received 24 December 2002
Accepted 27 March 2003

Hepatitis C virus (HCV) is an important human pathogen that affects 170 million people worldwide. The HCV genome is approximately 9·6 kb in length and encodes a polyprotein that is proteolytically cleaved to generate at least 10 mature viral protein products. Recently, a new protein, named F, has been described to be expressed through a ribosomal frameshift within the capsid-encoding sequence, a mechanism unique among members of the family *Flaviviridae*. Here, expression of the F protein was investigated in an *in vitro* transcription/translation assay. Its expression in mammalian cells was confirmed using specific recombinant vaccinia viruses; under these conditions, protein expression is dependent on the HCV IRES. The F protein was tagged with firefly luciferase or the Myc epitope to facilitate its identification. Ribosomal frameshifting was dependent on the presence of mutations in the capsid-encoding sequence. No frameshifting was detected in the absence of any mutation. Furthermore, analysis of the F protein in time-course experiments revealed that the protein is very unstable and that its production can be stabilized by the proteasome inhibitor MG132. Finally, indirect immunofluorescence studies have localized the F protein in the cytoplasm, with notable perinuclear detection.

INTRODUCTION

Hepatitis C virus (HCV) infection is highly prevalent worldwide and the disease often progresses to chronic hepatitis, leading to cirrhosis and, possibly, hepatocellular carcinoma. Interferon-α (IFN-α) administration in combination with ribavirin remains the most successful treatment for HCV infection but is only effective in about 50 % of patients (McHutchison *et al.*, 1998; McHutchison & Poynard, 1999; Schalm *et al.*, 1999). The efficacy of IFN treatment can be limited in duration, leading to the selection of virus variants during the onset of infection that are resistant to IFN.

HCV is related to the flaviviruses and the pestiviruses (Lindenbach & Rice, 2001; Miller & Purcell, 1990; Takeuchi *et al.*, 1990). HCV, a positive-stranded RNA virus with a genomic size of approximately 9·6 kb (Choo *et al.*, 1989, 1991), is now classified within the genus *Hepacivirus*, family *Flaviviridae*. The viral genome contains a large open reading frame encoding a polyprotein of approximately 3000 aa that is cleaved by a combination of host signal peptidases and two virus-encoded proteases to produce the mature structural and non-structural proteins: C, E1, E2, p7, NS2, NS3, NS4A, NS4B, NS5A and NS5B (Lindenbach & Rice, 2001). The C protein, also called the core protein, is the putative capsid protein; E1 and E2 are thought to be the membrane-associated envelope glycoproteins; p7, a polypeptide of unknown function, is cleaved inefficiently from E2; NS2 through NS5B are the non-structural proteins, which, except for NS2, are involved in the replication of the viral genome (Lohmann *et al.*, 1999).

The amino acid sequence of the core protein is well conserved among different HCV isolates. Its N-terminal region is highly basic, while its C-terminal region is hydrophobic (reviewed by McLauchlan, 2000). During its maturation, the core protein undergoes two consecutive membrane-dependent cleavage events: (i) the first generates the p23 protein, the immature core protein of 191 aa, and is mediated by a signal peptidase; and (ii) the second yields the p21 protein and is mediated by a signal peptide peptidase (McLauchlan *et al.*, 2002). The C terminus of p21 has not been mapped correctly yet and different locations have been reported (Hussy *et al.*, 1996; Liu *et al.*, 1997; Lo *et al.*, 1995; Santolini *et al.*, 1994; Yasui *et al.*, 1998). Both p23 and p21 core proteins have been termed, in some early publications, as p21 and p19, respectively. In addition, a 16 kDa protein, called p16, can be expressed also from the HCV capsid protein-encoding sequence (Lo *et al.*, 1994, 1995; Yeh *et al.*, 2000). The identity of this product remained unclear until the recent observation of a novel translation mechanism

within the capsid-encoding sequence corresponding to a ribosomal frameshift, a mechanism unique among members of the family *Flaviviridae* (Varaklioti *et al.*, 2002; Walewski *et al.*, 2001; Xu *et al.*, 2001). This protein has been called F (frameshifted) (Xu *et al.*, 2001) protein or ARFP (alternative ribosomal frameshift protein) (Walewski *et al.*, 2001).

In this study, we investigated the expression of the F protein in *in vitro* and *in vivo* expression systems in the presence or absence of mutations that have been reported previously to be associated with the expression of the p16 protein (Yeh *et al.*, 2000). In the absence of any mutation, no ribosomal frameshifting was observed. In contrast, a framshifted protein was clearly identified when mutations were introduced at nt 367 and 373 (codons 9 and 11) in the capsid-encoding sequence. Furthermore, the data obtained in time-course experiments revealed that the F protein is a very short-lived protein and that its stability can be maintained by the use of the proteasome inhibitor MG132.

METHODS

Plasmid constructs for frameshifting assays *in vitro* and *in vivo*. Plasmids were constructed using standard methods (Sambrook *et al.*, 1989). HCV sequences were amplified from clones derived from strain H (Feinstone *et al.*, 1981; Fournillier-Jacob *et al.*, 1996). The HCV 5′UTR was introduced into plasmid pTM1/CE_1E_2p7 (Fournillier-Jacob *et al.*, 1996) in place of the encephalomyocarditis virus IRES sequence by PCR with appropriate oligonucleotides and templates. Mutations $G^{367}\rightarrow A$ and $C^{373}\rightarrow A$ were introduced into pTHC/CE_1E_2p7 to generate the plasmid pTHC/CE_1E_2p7m by site-directed mutagenesis using enzymatic inverse PCR, as described by Stemmer & Morris (1992). These plasmids were used as template and also for cloning by replacing the sequence encoding the structural proteins with the new PCR products. Plasmids pTHC/C^{Δ}-Luc^{-} and pTHC/C^{Δ}m-Luc^{-} contain the wild-type and mutant capsid-encoding region (nt 342–585), respectively, from the HCV strain H isolate in which the sequence encoding the firefly luciferase is not fused in-frame with the capsid-encoding sequence or with the F-encoding sequence. The firefly luciferase sequence was amplified by PCR using appropriate primers designed to contain *Asp*718 and *Bgl*II sites at the 5′ and 3′ ends of the fragment. The restriction site *Bgl*II is an unique restriction site created at the 3′ end of the HCV structural sequence. All sequences at the junction between HCV and firefly luciferase sequences are indicated in Fig. 1. To restore a correct reading frame between the sequence encoding the first amino acids of the capsid protein and the firefly luciferase, the plasmids pTHC/C^{Δ}-Luc^{-} and pTHC/C^{Δ}m-Luc^{-} were cleaved at the *Asp*718 restriction site and the sequence was filled with Klenow and then ligated with T4 DNA ligase. The resulting plasmids pTHC/C^{Δ}-Luc^{+} and pTHC/C^{Δ}m-Luc^{+} were obtained. Plasmids pTHC/CF^{Δ}-Luc^{+} and pTHC/CF^{Δ}m-Luc^{+} contain the sequence encoding the firefly luciferase fused in-frame with the sequence encoding the F protein and not the C protein. Cloning was performed by PCR using appropriate oligonucleotides. In all constructs containing the firefly luciferase sequence, the AUG initiation codon of the luciferase gene was omitted during construction in order to avoid all internal translational of this gene in an *in vitro* transcription/translation assay.

Plasmids pTHC/CF-myc and pTHC/CFm-myc contain the entire HCV 5′UTR (nt 1–341) and part of the core-encoding sequence (nt 342–825) from HCV-H genotype 1a fused with the sequence encoding the Myc epitope (EQKLISEEDL) preceded by three Gly residues. This sequence is fused in-phase only with the sequence encoding the HCV F protein. Cloning was performed by PCR using appropriate oligonucleotides and the antisense primer was designed to contain a *Bgl*II

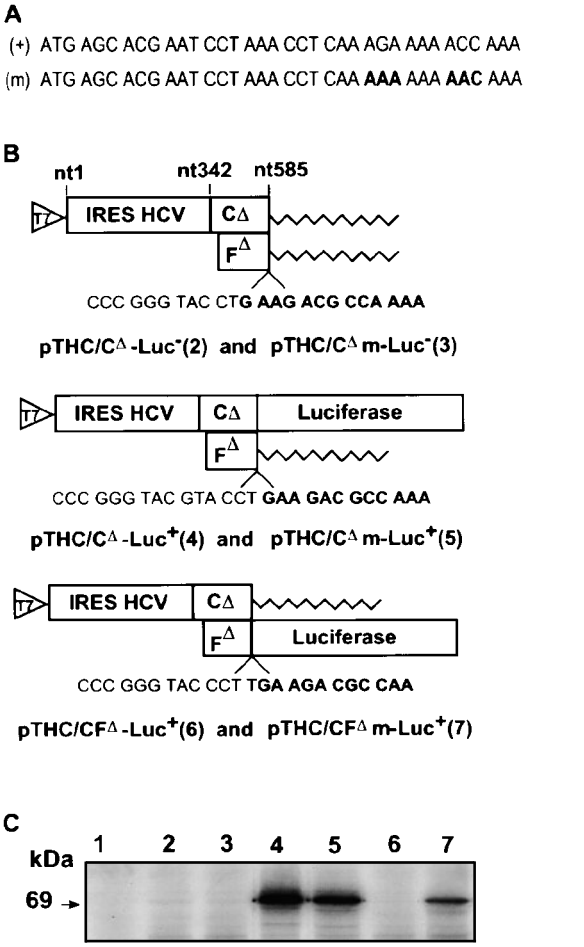

Fig. 1. Tagging experiments with the luciferase gene. (A) The first 12 codons of the capsid protein are shown. Differences between the wild-type (+) and the mutant (m) sequences are indicated. Mutated nucleotides are indicated in bold. (B) Schematic representation of the constructs used for the tagging experiments with luciferase. The entire HCV IRES and part of the HCV capsid-encoding sequences were cloned into the pTHC plasmid. Nucleotide sequences at the junction of the capsid-encoding region and the luciferase-encoding region are illustrated below. The beginning of the sequence encoding the luciferase gene is shown in bold. Broken lines indicate that the sequence encoding the luciferase protein is out of frame with the sequences encoding either the capsid protein or the F protein. (C) DNA templates were used to programme coupled transcription/translation reactions, as described previously (Meunier *et al.*, 1999). These DNAs correspond, respectively, to pTHC/C^{Δ}-Luc^{-} (lane 2), pTHC/C^{Δ}m-Luc^{-} (lane 3), pTHC/C^{Δ}-Luc^{+} (lane 4), pTHC/C^{Δ}m-Luc^{+} (lane 5), pTHC/CF^{Δ}-Luc^{+} (lane 6) and pTHC/CF^{Δ}m-Luc^{+} (lane 7). The control of the translation assay without RNA is indicated in lane 1. Translation products were analysed by SDS-PAGE (10% polyacrylamide gel). The molecular mass of the fusion protein (in kDa) is indicated on the left.

restriction site and the sequence encoding the Myc epitope. Plasmid pTM1/myc-F contains the entire sequence encoding the F protein (nt 346–825), tagged with the Myc epitope (EQKLISEEDL) at its N terminus. The nucleotide sequence AAAGAAAAA was replaced during PCR by the nucleotide sequence AAGAAGAAG in order to disrupt the A-rich region. Plasmid pTM1/myc-$C_{1–163}$ contains the sequence encoding the first 163 aa and is tagged with the Myc epitope at its N terminus. Each PCR product was cloned into the pTM1 plasmid.

Generation of recombinant vaccinia viruses. Transfection and isolation of recombinant viruses were performed essentially as described (Kieny *et al.*, 1984). The following vaccinia virus recombinant has been described previously: vTF7-3 (expressing the T7 DNA-dependent RNA polymerase) (Fuerst *et al.*, 1986).

Immunoprecipitation and Western blot assays. Cells expressing HCV proteins were metabolically labelled with [^{35}S]Protein Labelling mix ($3{\cdot}7 \times 10^6$ Bq ml^{-1}), as described previously (Meunier *et al.*, 1999). Cells were lysed with 0·5 % Igepal CA-630 in TBS (50 mM Tris/HCl, pH 7·5, and 150 mM NaCl). Immunoprecipitations were carried out as described (Dubuisson & Rice, 1996). Immune complexes were boiled for 5 min in Laemmli's buffer before analysis by SDS-PAGE. Gels were then treated as described previously (Meunier *et al.*, 1999). Proteins bound to nitrocellulose membranes (PVDF transfer membrane, NEN Life Science Products) were revealed by enhanced chemiluminescence detection (Amersham Pharmacia), as recommended by the manufacturer, with the specific monoclonal antibodies (mAbs) anti-C (diluted 1 : 5000) (Maillard *et al.*, 2001), anti-myc (diluted 1 : 200) (ATCC CRL-1725) or anti-luc (diluted 1 : 1000) (Promega).

Luciferase assays. Luc activity was assayed with the Luciferase Reporter Assay system (Promega) on rabbit reticulocyte lysate expressing the appropriate constructs or on cell lysates provided by cells infected with different recombinant viruses expressing the firefly luciferase activity. HepG2 cells (10^5 cells) plated in a multiwell plate were infected with the appropriate vaccinia virus recombinant at an m.o.i. of 5 p.f.u. per cell. At 6 h post-infection (p.i.), cells were washed twice with PBS, scraped out with 120 µl 1× Reporter Lysis buffer (Promega), lysed by one freeze–thaw cycle, vortexed and spun at 25 °C to pellet cell debris. Supernatant (4 µl) in 10 µl water was placed in a luminometer (Lumat, LB9501 Berthold) and the reaction was started by injection of 20 µl Luciferase Assay reagent (Promega). Light emission was recorded for 15 s.

RESULTS

In vitro analysis of frameshifting

A recent study on the F protein has indicated that the frameshifting might not be observed in some isolates (Varaklioti *et al.*, 2002). The F protein has indeed been observed clearly in an HCV type 1 isolate but no frameshifting was detected with the HCV strain H isolate. In addition, earlier data have indicated that p16 is produced when mutations are present in the N terminus of the capsid-encoding region (Yeh *et al.*, 2000). These data suggested to us that the sequence present in the N terminus of the capsid-encoding region might influence the efficiency of frameshifting. To test this hypothesis, frameshifting has been analysed in the presence or absence of mutations.

Mutations reported previously – G^{367}→A (codon 9) and C^{373}→A (codon 11) (Yeh *et al.*, 2000) – were introduced into the N terminus of the capsid-encoding region; this construct will be referred to the mutant 'm' (Fig. 1A). To detect frameshifting and to analyse the level of expression of the framshifted protein, the F protein reading frame was fused in-frame with the Luc-encoding sequence (pTHC/CF^{Δ}-Luc$^+$ and pTHC/CF^{Δ}m-Luc$^+$ constructs). With this type of approach, the fully active luciferase protein can be detected only if a ribosomal frameshift occurs during translation. To be closer to the context of HCV expression, the sequences encoding the HCV C protein and firefly luciferase were placed under the translational control of the HCV 5′UTR. All DNA constructs used for this study are illustrated in Fig. 1(B). As expected, a band with an apparent molecular mass of 69 kDa and corresponding to the N terminus of the F protein in fusion with Luc (F^{Δ}-Luc) was detected with the *in vitro* transcription/translation assay resulting from the plasmid containing the mutations (pTHC/CF^{Δ}m-Luc$^+$, Fig. 1C). However, no band was observed when no mutation was introduced into the N terminus of the capsid-encoding region (pTHC/CF^{Δ}-Luc$^+$, Fig. 1C). To determine the efficiency of ribosomal frameshifting, the luciferase activity of the F^{Δ}-Luc fusion protein was measured and compared to that of a fusion protein in which the luciferase sequence was fused in-frame with the N-terminal sequence of the C protein (C^{Δ}-Luc). About 30 and 3 % of luciferase activity was observed for the mutated and non-mutated F^{Δ}-Luc proteins, respectively (data not shown).

To confirm that the mutations introduced into the N terminus of the capsid-encoding region have no effect on the level of expression of the capsid protein, the Luc-encoding sequence was fused in-frame with the sequence encoding the first 82 aa of the capsid protein (pTHC/C^{Δ}-Luc$^+$ and pTHC/C^{Δ}m-Luc$^+$). The fusion product, corresponding to each construct, yielded a 69 kDa protein in an *in vitro* transcription/translation assay (Fig. 1C). Levels of expression were similar and correlated with luciferase activities (data not shown). Additional constructs in which the sequence encoding the first 82 aa of the HCV capsid protein was fused out of frame with the firefly luciferase-encoding sequence (pTHC/C^{Δ}-Luc$^-$ and pTHC/C^{Δ}m-Luc$^-$) were used as negative controls. These constructs are not supposed to produce a luciferase fusion protein even if a +1 ribosomal frameshift occurs in the capsid-encoding sequence due to the presence of termination codons in the +1 frame of the luciferase-encoding sequence. This was confirmed by SDS-PAGE analysis of [^{35}S]methionine-labelled translation products obtained in an *in vitro* transcription/translation assay using rabbit reticulocyte lysates (Fig. 1C). Additionally, no firefly luciferase activity associated with these constructs was observed (data not shown).

These results indicate that a frameshifted fusion protein with an active luciferase protein is produced *in vitro* when modifications at nt 367 and 373 (codons 9 and 11) are introduced within the nucleotide sequence encoding the N terminus of the HCV capsid protein. However, in the absence of any mutation, no frameshift is detectable above background level.

In vivo analysis of the frameshifting

Due to the presence of some factors that may influence ribosome–mRNA interactions (Parkin *et al.*, 1992; Reil *et al.*, 1993), it is well known that *in vivo* frameshifting results are generally different from those observed *in vitro*. To confirm the above-mentioned conclusions by *in vivo* experiments, recombinant vaccinia viruses were generated by homologous recombination and used to infect HepG2 cells. The proteins expressed by the different recombinant vaccinia viruses were analysed by SDS-PAGE followed by Western blot analysis using anti-luc or anti-C mAbs (Fig. 2). All constructs analysed *in vitro* were examined *in vivo*. A product fused in-frame with the fully active firefly protein was detected in HepG2 cells infected with the recombinant vaccinia virus vvpTHC/CF$^{\Delta}$m-Luc^{+} (Fig. 2B, lane 7), indicating that frameshifting was observed *in vivo*. In addition, 16 % luciferase activity has been observed for this construct (Fig. 2D, lane 7). In contrast, the firefly luciferase activity was only slightly above background in the absence of any mutation in the F$^{\Delta}$-Luc fusion protein (Fig. 2D, lane 6), which correlates with the absence of detection of the corresponding product by Western blotting (Fig. 2B, lane 6). These data indicate that the mutations introduced in the capsid-encoding sequence are necessary to improve the *in vivo* expression of a frameshifted protein. To confirm that all constructs could be expressed *in vivo*, the presence of capsid-derived proteins was analysed by Western blotting with an anti-C mAb. A truncated form of the capsid protein of the expected size (10 kDa) was detected from cell extracts of HepG2 cells infected with the recombinant vaccinia viruses vvpTHC/C$^{\Delta}$-Luc^{-}, vvpTHC/C$^{\Delta}$m-Luc^{-}, vvpTHC/CF$^{\Delta}$-Luc^{+} and vvpTHC/CF$^{\Delta}$m-Luc^{+} as expected (Fig. 2C, lanes 2, 3, 6 and 7). The small size of these proteins is due to the presence of termination codons in the Luc sequence, which are not in the same frame as the capsid protein. The fused proteins resulting from the capsid protein with the luciferase were only detected when the HepG2 cells were infected with the recombinant viruses vvpTHC/C$^{\Delta}$-Luc^{+} and vvpTHC/C$^{\Delta}$m-Luc^{+} (Fig. 2A, lanes 4 and 5). The corresponding fused proteins were also detected by Western blotting with anti-luc mAb (Fig. 2B, lanes 4 and 5). These constructs gave the highest reproducible levels of luciferase activity (Fig. 2D, lanes 4 and 5).

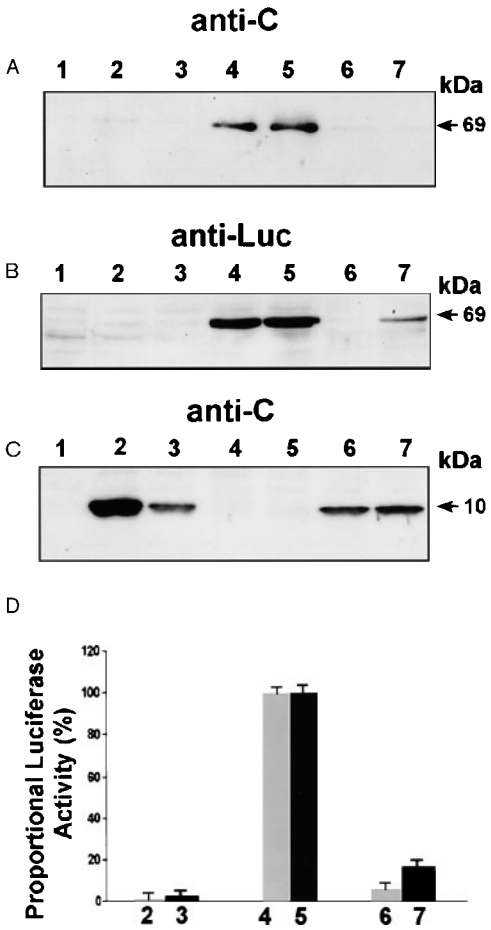

Fig. 2. Detection in HepG2 cells of different products resulting from the tagging experiments with the luciferase gene. HepG2 cells were infected with vTF7-3 alone (lane 1) or co-infected with vTF7-3 and pTHC/C$^{\Delta}$-Luc^{-} (lane 2), vvpTHC/C$^{\Delta}$m-Luc^{-} (lane 3), vvpTHC/C$^{\Delta}$-Luc^{+} (lane 4), vvpTHC/C$^{\Delta}$m-Luc^{+} (lane 5), vvpTHC/CF$^{\Delta}$-Luc^{+} (lane 6) or vvpTHC/CF$^{\Delta}$m-Luc^{+} (lane 7) at an m.o.i. of 5 p.f.u. per cell. At 6 h p.i., cells were lysed in luciferase cell culture lysis reagent. Then, the proteins of the cell lysates were separated by SDS-PAGE (10 % polyacrylamide gel) and revealed by Western blotting using anti-C mAb (A, C) and anti-luc mAb (B). Of the cell extracts, 4 μl was measured for luciferase activity, according to the manufacturer's protocol (Promega), and the resulting values in percentages are illustrated in (D). The molecular masses (kDa) of the different proteins are indicated on the right.

Detection of the F protein in HepG2 cells

Taking advantage of the data described above, and to characterize further the expression of the F protein in mammalian cells, plasmids containing the sequence encoding the first 163 aa of the capsid protein downstream of the HCV 5′UTR were constructed with or without mutations in the sequence of the HCV capsid protein. Because no antibodies were available against the F protein, the sequence of the Myc epitope was introduced in-frame with the C-terminal sequence of the predicted F protein (Xu *et al.*, 2001). Plasmids with wild-type (+) and mutant (m) sequences, corresponding to pTHC/CF-myc and pTHC/CFm-myc, respectively, were constructed (Fig. 3A). To analyse the expression of the F protein in mammalian cells, HepG2 cells were infected with recombinant vaccinia viruses vvpTHC/CF-myc and vvpTHC/CFm-myc. The products expressed by these viruses were analysed by SDS-PAGE followed by Western blotting using anti-C or anti-myc mAbs. As shown in Fig. 3(C, lane 3), a 17 kDa product corresponding to the size of the F protein was revealed with the anti-myc mAb in cells infected with vvpTHC/CFm-myc. In contrast, no product was detected with this mAb when the cells were infected with vvpTHC/CF-myc (Fig. 3C, lane

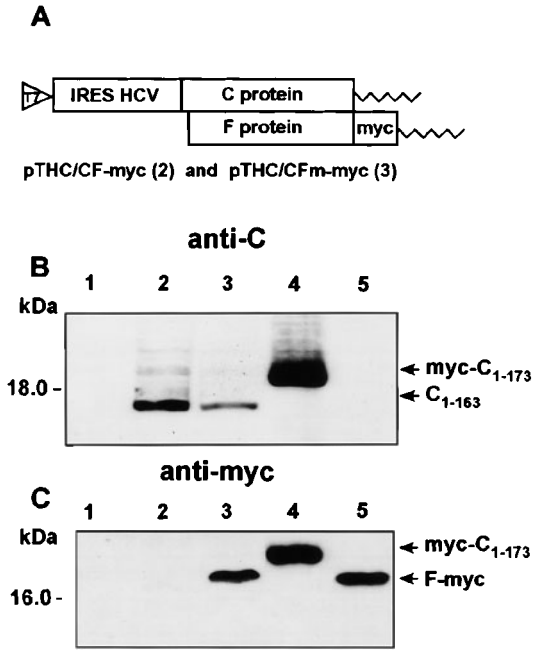

Fig. 3. Tagging experiments with the Myc epitope. (A) Schematic representation of the constructs used for the tagging experiments with *myc*. The entire HCV IRES and part of the HCV capsid-encoding sequences were cloned into pTHC. The C terminus of the F protein was tagged with the Myc epitope. (B, C) HepG2 cells were infected with vTF7-3 alone (lane 1) or co-infected vTF7-3 and vvpTHC/CF-myc (lane 2), vvpTHC/CFm-myc (lane 3), vvmyc-C_{1-173} (lane 4) or vvmyc-F_{1-161} (lane 5) at an m.o.i. of 5 p.f.u. per cell. Infected cells were lysed at 6 h p.i. Proteins were separated by SDS-PAGE and then revealed by Western blotting using anti-C mAb (B) and anti-myc mAb (C).

2). In addition, a band of 18 kDa corresponding to a truncated form of the capsid protein (C_{1-163}) was revealed with the anti-C mAb (Fig. 3B, lanes 2 and 3). Recombinant vaccinia viruses expressing HCV C_{1-173} or F_{1-161} proteins tagged at their N terminus with the Myc epitope were used as controls. The sequence encoding the F_{1-161} protein was based on its reading frame without the first amino acids of the HCV capsid protein. Modifications in the stretch of A residues were introduced in the sequence in order to avoid a possible ribosomal frameshift (codons 9–14). As shown in Fig. 3(B, C, lanes 5), the F_{1-161} protein (17 kDa) was detected only by anti-myc mAb, while the tagged capsid protein was detected with the anti-C and anti-myc mAbs (Fig. 3B, C, lane 4).

These data are in agreement with the *in vitro* experiments and confirm that a frameshift occurs *in vivo* at least when some mutations are introduced in the capsid-encoding sequence. In this particular case, modifications of G→A and C→A at nt 367 and 373, respectively, lead to the formation of a perfect heptanucleotide sequence, which could be considered as the first element necessary for the ribosomal frameshift and could be assimilated to the slippery sequence (see Discussion).

Analysis of F protein stability

We also wanted to determine whether the F protein is stable during its translation in mammalian cells. The expression of the F protein was analysed in pulse–chase experiments by infecting HepG2 cells with the recombinant vaccinia virus vvpTHC/CFm-myc. As shown in Fig. 4(A), the level of expression of the F protein decreased very quickly. After 15 min of chase, the intensity of the band was already very low. The estimated half-life of the F protein was approximately 10 min (Fig. 4A). A decrease in the amount of the F protein was also observed when the HepG2 cells were infected with the recombinant virus vvmyc-F_{1-161} (Fig. 4B). However, after 2 h of chase, 20 % of the F protein was still detected, suggesting that overexpression of the F protein might reduce its degradation. The capsid protein tagged with the Myc epitope at its N terminus (myc-C_{1-173}) was used as a control. As shown in Fig. 4(C), this protein was very stable during the same pulse–chase conditions.

Degradation of most proteins in mammalian cells occurs via the proteasome (Lee & Goldberg, 1998), which degrades both short-lived proteins ($t_{1/2} < 3$ h) and more stable proteins ($t_{1/2}$ of h or days). Investigation of protein turnover has been facilitated by specific proteasome inhibitors like lactacystin or MG132. In this study, we wanted to address

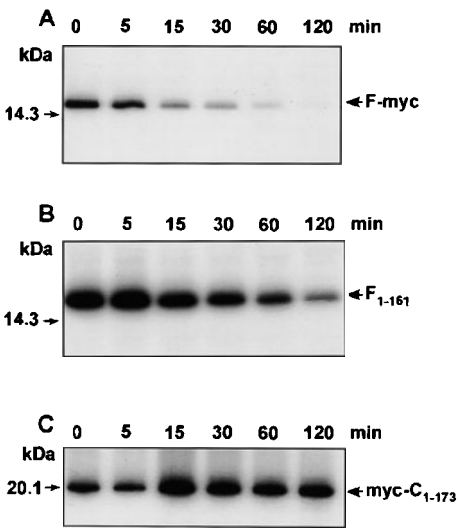

Fig. 4. Expression of F-myc, myc-C_{1-173} and myc-F_{1-161} analysed in pulse–chase experiments. HepG2 cells were co-infected with vTF7-3 and (A) vvpTHC/CFm-myc, (B) vvmyc-F_{1-161} or (C) vvmyc-C_{1-173} at an m.o.i. of 5 p.f.u. per cell. Infected cells were pulse-labelled for 10 min and chased for the times indicated. Cell lysates were immunoprecipitated with anti-myc mAb. Samples were separated by SDS-PAGE (15 % polyacrylamide gel).

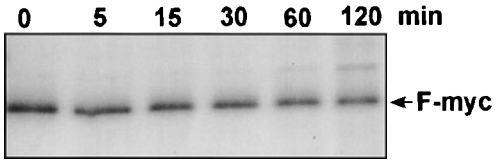

Fig. 5. Pulse–chase analysis of the F-myc protein with the proteasome inhibitor MG132. HepG2 cells were co-infected with vTF7-3 and vvpTHC/CFm-myc at an m.o.i. of 5 p.f.u. per cell. At 1 h before pulse labelling, MG132 (2 μM) was added to the infected cells. Infected cells were then pulse-labelled for 10 min and chased, as indicated before, in the presence of inhibitor. Cell lysates were immunoprecipitated with anti-myc mAb. Samples were separated by SDS-PAGE (15 % polyacrylamide gel). A control experiment in the absence of MG132 was also carried out (data not shown).

the question of whether the proteasome pathway could be involved in the degradation of this protein. The proteasome inhibitor MG132 was used during infection of HepG2 cells with both recombinant vaccinia viruses vTF7-3 and vvpTHC/CFm-myc and cells were pulse-labelled as described above. As shown in Fig. 5, the degradation of the F protein was markedly inhibited by adding the proteasome inhibitor. Indeed, the amount of F protein was stable during the pulse–chase experiment. Thus, these results seem to indicate a role of the proteasome pathway in the turnover of this protein. Based on these results, we reinvestigated the expression of a frameshifted protein in the absence of any mutation in the presence of the proteasome inhibitor MG132. Indeed, it is possible that the absence of detection of the F protein with the wild-type construct could be due to a rapid degradation of this protein during its expression. The pulse–chase experiment was done under the same conditions as described above. However, even in the presence of the proteasome inhibitor, the F protein was not detected (data not shown). These data indicate that the lack of detection of the F protein in the absence of any mutation is not due to a rapid degradation of this protein.

Intracellular localization of F protein

As a next step in the characterization of the F protein, we analysed its subcellular localization. Fig. 6 illustrates the immunostaining of vvpTHC/CF-myc (Fig. 6C, D) and vvpTHC/CFm-myc (Fig. 6E, F) infected HepG2 cells observed by immunofluorescence using the anti-C mAb (Fig. 6A, C and E) or anti-myc mAb (Fig. 6B, D and F). As expected, immunofluorescence staining with anti-myc mAb was detected with the vvpTHC/CFm-myc recombinant virus (Fig. 6F). Reactivity was localized mainly in the cytoplasm and in some cells the reactivity had a perinuclear distribution. The immunostaining experiment with the anti-C mAb was used as a control of the expression of the capsid protein. Reactivity against the capsid protein was localized in the cytoplasm. Reactivity of the anti-myc and anti-C mAbs was also performed with control vTF7-3-infected cells in order to evaluate the level of background immunostaining (Fig. 6A, B).

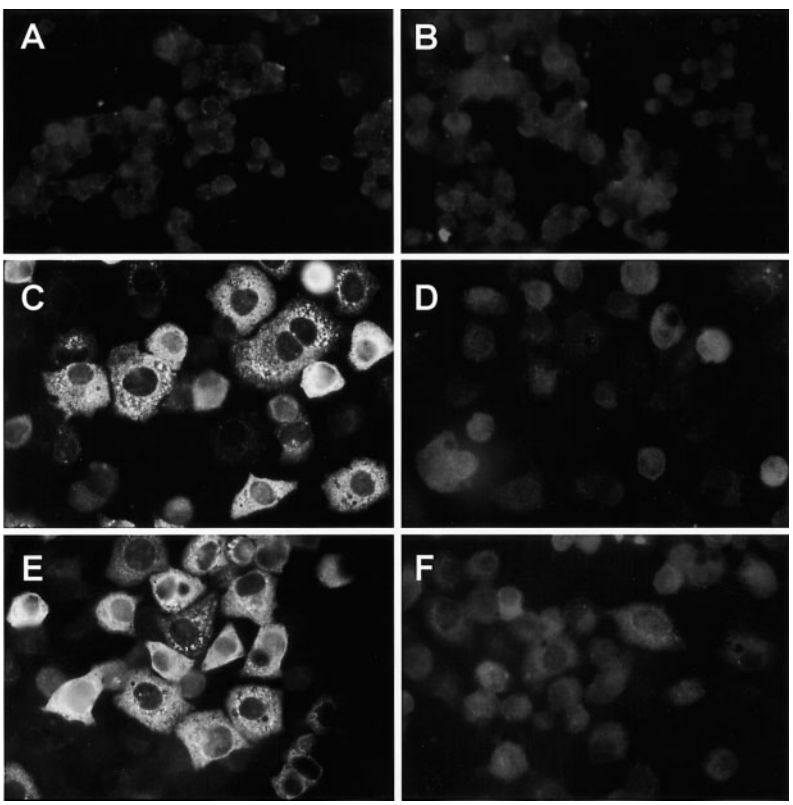

Fig. 6. Detection of the F protein by indirect immunofluorescence. Subconfluent HepG2 cells grown on coverslips were infected with vTF7-3 alone (A, B), vTF7-3 and vvpTHC/CF-myc (C, D) or vTF7-3 and vvpTHC/CFm-myc (E, F) at a m.o.i. of 5 p.f.u. per cell. At 8 h p.i., cells were fixed for 10 min with paraformaldehyde (4 % in PBS), permeabilized for 30 min with Triton X-100, and labelled with anti-C mAb (diluted 1 : 300) [secondary antibody: Rhodamine Red-X-conjugated $F(ab')_2$ fragment donkey anti-mouse IgG (A, C and E) or anti-myc mAb (diluted 1 : 100) (B, D and F)].

DISCUSSION

Our data reveal that the F protein, produced by a ribosomal frameshift mechanism, is expressed in mammalian cells. Translation of this protein is favoured by mutations located in the capsid-encoding sequence. When expressed alone, this protein is short-lived compared to the other HCV proteins (Pietschmann *et al.*, 2001) and its degradation involves the proteasome pathway.

Reporter enzymes, such as luciferase, CAT and β-galactosidase, are used commonly *in vivo* and *in vitro* (Naylor, 1999). Such enzymes are not expressed naturally in mammalian cells, making them useful to study gene expression in cells. Luciferase-tagging experiments used in this study have revealed that the ribosomal frameshift is more efficient when mutations at nt 367 (codon 9) and 373 (codon 11) are introduced in the capsid-encoding sequence compared to the results obtained with the wild-type sequence. The luciferase activity determined *in vivo* was about 16 % relative to the control construct and was less than that observed *in vitro* (30 %). This lower efficiency observed in mammalian cells is not unusual and can depend on the genes and the expression systems used (Ivanov *et al.*, 2000). In our study, tagging the C terminus of the F protein has made its detection easier in mammalian cells by immunoblot or immunoprecipitation analyses. Our results have shown that, besides its expression in an *in vitro* system, the F protein can be expressed also in mammalian cells. These data are reinforced by the fact that antibodies directed against the F protein are present in some patient sera, indicating that the F protein is produced during a natural HCV infection in patients (Varaklioti *et al.*, 2002; Xu *et al.*, 2001).

The molecular mechanism of frameshifting leading to the translation of the HCV F protein remains to be determined. However, given the data presented in this study, it could be concluded that the modifications at nt 367 and 373 do not simply increase efficiency but are necessary for the production of the F protein. Interestingly, a similar observation was made using the sequence of a HCV genotype 1b (data not shown). Modifications of codons Arg→Lys (codon 9) and Thr→Asn (codon 11) lead to the modification of a G→A (nt 367) and a C→A (nt 373), generating an A-rich region (10 A residues) between nt 363 and 374 of the HCV sequence. In this context, a slippery sequence can emerge from this region. Indeed, defined initially by *in vitro* translation assays, two structural motifs in mRNA have been characterized as important for an efficient −1 ribosomal frameshift (Brierley, 1995; Dinman, 1995; Gesteland & Atkins, 1996). One is the slippery sequence, the heptanucleotide XXXYYYZ (X is any base, Y is A or U and Z is not G) (Jacks *et al.*, 1988), and the other component is a downstream RNA structural element, either a simple hairpin structure or, more frequently, a pseudoknot (Chamorro *et al.*, 1992; Jacks *et al.*, 1988; ten Dam *et al.*, 1990). In contrast, the +1 ribosomal frameshift in the yeast retrotransposon Ty requires only a short slippery sequence with a rare codon in the original reading frame (Belcourt & Farabaugh, 1990). Frameshifting in human immunodeficiency virus requires only the short slippery sequence but not the 3′ sequence with its predicted stem–loop structure (Wilson *et al.*, 1988). This unusual process occurs also among other viruses, including coronaviruses (Brierley, 1995; Herold & Siddell, 1993) and human astrovirus serotype 1 (Marczinke *et al.*, 1994). In HCV, a single mutation of the codon AGA (Arg) to AAA (Lys) is also sufficient to generate *in vitro* the expression of the F protein (Lo *et al.*, 1994). This mutation creates a heptanucleotide sequence between nt 364 and 372 or 368 and 374. As mutations introduced in this region may recreate a functional slippery sequence, as described for other viruses, it can be suggested that this sequence constitutes the first control element for the translation of the F protein. Comparative studies of the sequence of the codons 8–14 reveal that these codons are also very conserved among the HCV sequences, reinforcing the importance of this region (Rijnbrand & Lemon, 2000). However, if the slippery sequence can be suggested, no available data allow us to determine whether an RNA structural element is also necessary for the ribosomal frameshift.

Interestingly, the F protein is very unstable when expressed in mammalian cells. The degradation of most proteins in mammalian cells occurs via the ubiquitin-proteasome pathway (Lee & Goldberg, 1998). Under these circumstances, ubiquitin is linked covalently to the proteins and is then targeted to a large multiproteinase complex (700 kDa) that constitutes the catalytic core of the 26S proteasome used as an intracellular protein-degrading machine of eukaryotic organisms. The proteasome is essential for the normal turnover of regulatory proteins that controls cell growth and metabolism or is necessary for the removal of damaged or mutated proteins (Molinari *et al.*, 1999). Since the discovery that lactacystin is a potent inhibitor of the 26S proteasome (Fenteany *et al.*, 1995), different novel proteasome inhibitors, like MG132 or proteasome inhibitor I, were developed. In our study, pulse–chase assays indicate that the decreased protein level of the HCV F protein in mammalian cells is due to proteasome degradation. A specific inhibitor of 26S proteasome activity blocks degradation of the F protein during its translation and stabilizes the level of its expression. It is possible that the F protein expressed under our experimental conditions is not folded properly and is, therefore, targeted directly for degradation, as observed for some damaged or mutated proteins (Molinari *et al.*, 1999). In this way, it would be very interesting also to determine whether HCV proteins could be involved in the stabilization of the F protein. Alternatively, it is not the F protein itself but some of its degradation product(s) that might play a functional role in the HCV life cycle.

No function has been attributed to the F protein yet. Its localization in the cytosol suggests that this protein plays a functional role in this compartment. However, we cannot exclude that it moves to another compartment, e.g. the nucleus, after interacting with a cellular component. Its very

low level of expression in the absence of any mutation suggests that the F protein might also be expressed at very low levels in HCV-infected cells, which is not in favour for a role of this protein in virus assembly. Interestingly, mutations leading to higher levels of expression of the F protein *ex vivo* can be observed in HCV isolated from chronically infected patients with HCV-related hepatocellular carcinoma, leading to the hypothesis that higher levels of F protein expression might be linked to virus pathogenesis. Further studies on the F protein should lead to a better understanding of the role of this protein in the HCV life cycle and in pathogenesis.

ACKNOWLEDGEMENTS

We would like to thank Dr Delagneau for the generous gift of anti-C monoclonal antibody. This work was supported by the CNRS (Centre National de la Recherche Scientifique) and the Institut Pasteur of Lille. Juliette Roussel is recipient of funding from the ANRS (Agence Nationale de la Recherche sur le SIDA).

REFERENCES

Belcourt, M. F. & Farabaugh, P. J. (1990). Ribosomal frameshifting in the yeast retrotransposon Ty: tRNAs induce slippage on a 7 nucleotide minimal site. *Cell* **62**, 339–352.

Brierley, I. (1995). Ribosomal frameshifting viral RNAs. *J Gen Virol* **76**, 1885–1892.

Chamorro, M., Parkin, N. & Varmus, H. E. (1992). An RNA pseudoknot and an optimal heptameric shift site are required for highly efficient ribosomal frameshifting on a retroviral messenger RNA. *Proc Natl Acad Sci U S A* **89**, 713–717.

Choo, Q.-L., Kuo, G., Weiner, A. J., Overby, L. R., Bradley, D. W. & 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., Richman, K. H., Han, J. H. & 11 other authors (1991). Genetic organization and diversity of the hepatitis C virus. *Proc Natl Acad Sci U S A* **88**, 2451–2455.

Dinman, J. D. (1995). Ribosomal frameshifting in yeast viruses. *Yeast* **11**, 1115–1127.

Dubuisson, J. & Rice, C. M. (1996). Hepatitis C virus glycoprotein folding: disulfide bond formation and association with calnexin. *J Virol* **70**, 778–786.

Feinstone, S. M., Alter, H. J., Dienes, H. P., Shimizu, Y., Popper, H., Blackmore, D., Sly, D., London, W. T. & Purcell, R. H. (1981). Non-A, non-B hepatitis in chimpanzees and marmosets. *J Infect Dis* **144**, 588–598.

Fenteany, G., Standaert, R. F., Lane, W. S., Choi, S., Corey, E. J. & Schreiber, S. L. (1995). Inhibition of proteasome activities and subunit-specific amino-terminal threonine modification by lactacystin. *Science* **268**, 726–731.

Fournillier-Jacob, A., Cahour, A., Escriou, N., Girard, M. & Wychowski, C. (1996). Processing of the E1 glycoprotein of hepatitis C virus expressed in mammalian cells. *J Gen Virol* **77**, 1055–1064.

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

Gesteland, R. F. & Atkins, J. F. (1996). Recoding: dynamic reprogramming of translation. *Annu Rev Biochem* **65**, 741–768.

Herold, J. & Siddell, S. G. (1993). An 'elaborated' pseudoknot is required for high frequency frameshifting during translation of HCV 229E polymerase mRNA. *Nucleic Acids Res* **21**, 5838–5842.

Hussy, P., Langen, H., Mous, J. & Jacobsen, H. (1996). Hepatitis C virus core protein: carboxy-terminal boundaries of two processed species suggest cleavage by a signal peptide peptidase. *Virology* **224**, 93–104.

Ivanov, I. P., Matsufuji, S., Murakami, Y., Gesteland, R. F. & Atkins, J. F. (2000). Conservation of polyamine regulation by translational frameshifting from yeast to mammals. *EMBO J* **19**, 1907–1917.

Jacks, T., Madhani, H. D., Masiarz, F. R. & Varmus, H. E. (1988). Signals for ribosomal frameshifting in the Rous sarcoma virus *gag–pol* region. *Cell* **55**, 447–458.

Kieny, M.-P., Lathe, R., Drillien, R., Spehner, D., Skory, S., Schmitt, D., Wiktor, T., Koprowski, H. & Lecocq, J.-P. (1984). Expression of rabies virus glycoprotein from a recombinant vaccinia virus. *Nature* **312**, 163–166.

Lee, D. H. & Goldberg, A. L. (1998). Proteasome inhibitors: valuable new tools for cell biologists. *Trends Cell Biol* **8**, 397–403.

Lindenbach, B. D. & Rice, C. M. (2001). *Flaviviridae*: the viruses and their replication. In *Fields Virology*, 4th edn, pp. 991–1042. Edited by D. M. Knipe & P. M. Howley. Philadelphia: Lippincott Williams & Wilkins.

Liu, Q., Tackney, C., Bhat, R. A., Prince, A. M. & Zhang, P. (1997). Regulated processing of hepatitis C virus core protein is linked to subcellular localization. *J Virol* **71**, 657–662.

Lo, S. Y., Selby, M., Tong, M. & 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.

Lo, S. Y., Masiarz, F., Hwang, S. B., Lai, M. M. & Ou, J. H. (1995). Differential subcellular localization of hepatitis C virus core gene products. *Virology* **213**, 455–461.

Lohmann, V., Korner, F., Koch, J., Herian, U., Theilmann, L. & Bartenschlager, R. (1999). Replication of subgenomic hepatitis C virus RNAs in a hepatoma cell line. *Science* **285**, 110–113.

Maillard, P., Krawczynski, K., Nitkiewicz, J. & 7 other authors (2001). Nonenveloped nucleocapsids of hepatitis C virus in the serum of infected patients. *J Virol* **75**, 8240–8250.

Marczinke, B., Bloys, A. J., Brown, T. D., Willcocks, M. M., Carter, M. J. & Brierley, I. (1994). The human astrovirus RNA-dependent RNA polymerase coding region is expressed by ribosomal frameshifting. *J Virol* **68**, 5588–5595.

McHutchison, J. G. & Poynard, T. (1999). Combination therapy with interferon plus ribavirin for the initial treatment of chronic hepatitis C. *Semin Liver Dis* **19**, 57–65.

McHutchison, J. G., Gordon, S. C., Schiff, E. R. & 7 other authors (1998). Interferon α2b alone or in combination with ribavirin as initial treatment for chronic hepatitis C. Hepatitis Interventional Therapy Group. *N Engl J Med* **339**, 1485–1492.

McLauchlan, J. (2000). Properties of the hepatitis C virus core protein: a structural protein that modulates cellular processes. *J Viral Hep* **7**, 2–14.

McLauchlan, J., Lemberg, M. K., Hope, G. & Martoglio, B. (2002). Intramembrane proteolysis promotes trafficking of hepatitis C virus core protein to lipid droplets. *EMBO J* **21**, 3980–3988.

Meunier, J. C., Fournillier, A., Choukhi, A., Cahour, A., Cocquerel, L., Dubuisson, J. & Wychowski, C. (1999). Analysis of the glycosylation sites of hepatitis C virus (HCV) glycoprotein E1 and the influence of

E1 glycans on the formation of the HCV glycoprotein complex. *J Gen Virol* **80**, 887–896.

Miller, R. H. & Purcell, R. H. (1990). Hepatitis C virus shares amino acid sequence similarity with pestiviruses and flaviviruses as well as members of two plant virus supergroups. *Proc Natl Acad Sci U S A* **87**, 2057–2061.

Molinari, E., Gilman, M. & Natesan, S. (1999). Proteasome-mediated degradation of transcriptional activators correlates with activation domain potency *in vivo*. *EMBO J* **18**, 6439–6447.

Naylor, L. H. (1999). Reporter gene technology: the future looks bright. *Biochem Pharmacol* **58**, 749–757.

Parkin, N. T., Chamorro, M. & Varmus, H. E. (1992). Human immunodeficiency virus type 1 *gag–pol* frameshifting is dependent on downstream mRNA secondary structure: demonstration by expression *in vivo*. *J Virol* **66**, 5147–5151.

Pietschmann, T., Lohmann, V., Rutter, G., Kurpanek, K. & Bartenschlager, R. (2001). Characterization of cell lines carrying self-replicating hepatitis C virus RNAs. *J Virol* **75**, 1252–1264.

Reil, H., Kollmus, H., Weidle, U. H. & Hauser, H. (1993). A heptanucleotide sequence mediates ribosomal frameshifting in mammalian cells. *J Virol* **67**, 5579–5584.

Rijnbrand, R. C. & Lemon, S. M. (2000). Internal ribosome entry site-mediated translation in hepatitis C virus replication. *Curr Top Microbiol Immunol* **242**, 85–116.

Sambrook, J., Fritsch, E. F. & Maniatis, T. (1989). *Molecular Cloning: a Laboratory Manual*, 2nd edn. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory.

Santolini, E., Migliaccio, G. & La Monica, N. (1994). Biosynthesis and biochemical properties of the hepatitis C virus core protein. *J Virol* **68**, 3631–3641.

Schalm, S. W., Weiland, O., Hansen, B. E. & 9 other authors (1999). Interferon-ribavirin for chronic hepatitis C with and without cirrhosis: analysis of individual patient data of six controlled trials. Eurohep Study Group for Viral Hepatitis. *Gastroenterology* **117**, 408–413.

Stemmer, W. P. C. & Morris, S. K. (1992). Enzymatic inverse PCR: a restriction site independent, single-fragment method for high-efficiency, site directed mutagenesis. *Biotechniques* **13**, 214–220.

Takeuchi, K., Kubo, Y., Boonmar, S. & 7 other authors (1990). The putative nucleocapsid and envelope protein genes of hepatitis C virus determined by comparison of the nucleotide sequences of two isolates derived from an experimentally infected chimpanzee and healthy human carriers. *J Gen Virol* **71**, 3027–3033.

ten Dam, E. B., Pleij, C. W. & Bosch, L. (1990). RNA pseudoknots: translational frameshifting and readthrough on viral RNAs. *Virus Genes* **4**, 121–136.

Varaklioti, A., Vassilaki, N., Georgopoulou, U. & Mavromara, P. (2002). Alternate translation occurs within the core coding region of the hepatitis C viral genome. *J Biol Chem* **277**, 17713–17721.

Walewski, J. L., Keller, T. R., Stump, D. D. & Branch, A. D. (2001). Evidence for a new hepatitis C virus antigen encoded in an overlapping reading frame. *RNA* **7**, 710–721.

Wilson, W., Braddock, M., Adams, S. E., Rathjen, P. D., Kingsman, S. M. & Kingsman, A. J. (1988). HIV expression strategies: ribosomal frameshifting is directed by a short sequence in both mammalian and yeast systems. *Cell* **55**, 1159–1169.

Xu, Z., Choi, J., Yen, T. S., Lu, W., Strohecker, A., Govindarajan, S., Chien, D., Selby, M. J. & Ou, J. (2001). Synthesis of a novel hepatitis C virus protein by ribosomal frameshift. *EMBO J* **20**, 3840–3848.

Yasui, K., Wakita, T., Tsukiyama-Kohara, K., Funahashi, S. I., Ichikawa, M., Kajita, T., Moradpour, D., Wands, J. R. & Kohara, M. (1998). The native form and maturation process of hepatitis C virus core protein. *J Virol* **72**, 6048–6055.

Yeh, C. T., Lo, S. Y., Dai, D. I., Tang, J. H., Chu, C. M. & 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.