Sequence motifs required for lipid droplet association and protein stability are unique to the hepatitis C virus core protein

R. Graham Hope and John McLauchlan

MRC Virology Unit, Division of Virology, University of Glasgow, Church Street, Glasgow G11 5JR, UK

From analysis of the primary sequence of the hepatitis C virus (HCV) core protein, we have identified three separable regions based on hydrophobicity and clustering of basic amino acids within the protein. Comparison with capsid proteins of related pesti- and flaviviruses suggested that HCV core has a unique central domain (domain 2). Previous findings have revealed that core protein can associate with lipid droplets which are intracellular storage sites for triacylglycerols and cholesterol esters. Confocal analysis of variant forms lacking regions of core indicated that most residues within the unique region are necessary for association of the protein with lipid droplets. A segment within domain 2 (from residues 125 to 144) also was required for stability of the protein and a polypeptide lacking these sequences was degraded apparently by the proteasome. In cells depleted of lipid droplets, core protein remained located in the cytoplasm. Moreover, cleavage of the protein at the maturation site and stability were not affected by inability to bind to lipid droplets.

Introduction

Hepatitis C virus (HCV) is the major causative agent of post-transfusion and sporadic non-A, non-B hepatitis (Kuo et al., 1989). In most individuals, a chronically infected state is established and this can lead to liver disease, potentially culminating in either cirrhosis or hepatocellular carcinoma (Di Bisceglie, 1998). The virus is a member of the Flaviviridae and has a positive-sense, single-stranded RNA genome of approximately 10 kb (Choo et al., 1991; Takamizawa et al., 1991; Robertson et al., 1998). The genome encodes a polyprotein of some 3000 amino acids which is post-translationally cleaved to generate ten individual gene products. At the N-terminal end of the polyprotein lie the structural components of the virus and these are generated by cleavage by host cell signalases (Hijikita et al., 1991; Grakoui et al., 1993; Selby et al., 1993). A significant difficulty with studies on HCV is the inability to produce any substantial quantities of virus particles in vitro and the low yields of virus obtained from infected sera. Thus, from comparative studies with other Flaviviridae, it has been proposed that the principal proteinaceous components of virions are the core protein, which forms the capsid, surrounded by an envelope containing glycoproteins E1 and E2 (Choo et al., 1991; Takamizawa et al., 1991).

Data from a number of studies have identified two major core species, p23 and p21 (Santolini et al., 1994; Hussy et al., 1996; Moradpour et al., 1996; Liu et al., 1997; Yasui et al., 1998). p23 is a 191 amino acid product which contains the signal sequence that directs E1 to the endoplasmic reticulum (ER). Cleavage of the polyprotein occurs between residues 191 and 192 to generate the N-terminal end of E1 (Hijikita et al., 1991). By contrast, the maturation process for producing p21 has not been precisely identified although the C terminus of p21 lies close to amino acid 174. In tissue culture cells, p21 is the predominant form of core detected and is the major species found in viral particles from infected sera, suggesting that it is the mature form of the protein (Yasui et al., 1998). A third core product, p16, has been identified in studies with the HCV-1 strain of the virus and probably results from cleavage at around residue 151 (Lo et al., 1994, 1995). The production of p16 is dependent on a lysine residue at amino acid position 9 and the absence of E1 sequences linked in cis to the 3’ end of the core coding region.

Functional analyses described in a number of reports have
shown that core protein can influence a variety of intracellular processes. In tissue culture systems, core modulates apoptosis (Chen et al., 1997; Ruggieri et al., 1997; Ray et al., 1998a; Zhu et al., 1998; Marusawa et al., 1999), signalling pathways involved in apoptotic events (Chen et al., 1997; Shrivastava et al., 1998; Zhu et al., 1998; Marusawa et al., 1999; You et al., 1999) and expression levels directed from various viral and cellular promoters (Ray et al., 1995, 1997, 1998b; Chang et al., 1998). Moreover, two lines of transgenic mice that express the protein only in the liver develop steatosis and thereafter hepatocellular carcinoma (Moriya et al., 1997, 1998). These disease states are prevalent in infected human individuals also (Scheuer et al., 1997, 1998). These studies have examined the effects of alterations to core coding sequences on its intracellular distribution (Lo et al., 1995, 1999; Suzuki et al., 1995, 1996; Liu et al., 1997, 1999), no systematic examination has been performed to identify elements that affect protein distribution. Here, we have examined the sequences that constitute core and propose that there are three distinct regions within the protein. The role of each of these regions in directing core to lipid droplets has been studied by mutational analysis. In addition, a subset of sequences within a region that was necessary for lipid droplet association and is unique to core had a role in maturation and stability of the protein. Maturation

where it is found both attached to the ER and at the surface of lipid droplets (Moradpour et al., 1996; Barba et al., 1997; Yasui et al., 1998). However, a minor proportion is present in the nucleus also (Yasui et al., 1998). Although several studies have examined the effects of alterations to core coding sequences on its intracellular distribution (Lo et al., 1995, 1999; Suzuki et al., 1995, 1996; Liu et al., 1997, 1999), no systematic examination has been performed to identify elements that affect protein distribution.

Here, we have examined the sequences that constitute core and propose that there are three distinct regions within the protein. The role of each of these regions in directing core to lipid droplets has been studied by mutational analysis. In addition, a subset of sequences within a region that was necessary for lipid droplet association and is unique to core had a role in maturation and stability of the protein. Maturation

Fig. 1. (a) Nucleotide sequence and predicted amino acid sequence for the core coding region of HCV strain Glasgow (genotype 1a; supplied by M. McElwee & R. Elliott, personal communication). Numbers at the end of each line give either the nucleotide sequence number or amino acid residue. Basic amino acids are circled. (b) Proposed domain structure and hydrophobicity plot of the predicted core amino acid sequence from HCV strain Glasgow.
was studied also in cells depleted of lipid to test whether lipid droplet association was required for cleavage and protein stability.

Methods

■ Construction of plasmids. Plasmid pgHCV/1–195 was constructed by inserting an oligonucleotide (5’ GCTGAGATCTA 3’), which had both a translational stop codon and the sequences for a BgII enzyme site, between an EspI site (numbered nucleotide residue 586 in Fig. 1a) in the HCV strain Glasgow genomic sequence and a HindIII enzyme site in the pGEM-1 vector backbone of a plasmid called pgHCV/CE1E2mix. Plasmid pgHCV/CE1E2mix encoded residues 1 to 837 of the polyprotein for HCV strain Glasgow (Patel et al., 1999). Thus, pgHCV/1–195 contained the coding sequence for the core protein (amino acids 1–191) and the N-terminal four residues of E1 protein of HCV strain Glasgow (Fig. 1a). For cloning purposes, pgHCV/1–195 also contained a BgII site introduced at the EcoRI enzyme site in the vector, which lies upstream of the HCV sequences. Other plasmids containing truncations and internal deletions of the core coding sequences were derived from pgHCV/1–195 (Table 1). Plasmid pgHCV/1–124,145–169 was made by inserting the oligonucleotides used to create pgHCV/ΔA125–144 between the ClaI/BstXI sites (Fig. 1a, residues 371 and 453 respectively) in pgHCV/1–169. Similarly, a double deletion mutant was constructed by introducing the oligonucleotides used to produce pgHCV/ΔA161–166 between the BstXI/BstEII sites (Fig. 1a, residues 453 and 502 respectively) of plasmid pgHCV/ΔA125–144 (Table 1). The resultant plasmid was called pgHCV/ΔA125–144,ΔA161–166. For expression purposes, BgII fragments from each of the constructs containing core sequences were introduced into the unique BamHI site of the Semliki Forest virus (SFV) expression vector pSFV1 (Liljestrom & Garoff, 1991). The nucleotide sequences of the regions containing HCV cDNA were determined before insertion into pSFV1. Plasmid pSFV/CA125–144,145E1E2 was made by replacing a BstIII/BstEII fragment in pSFV/CE1E2mix (Fig. 1a, nucleotide residues 141–502; Patel et al., 1999) with the corresponding DNA fragment from pgHCV/ΔA125–144.

■ Maintenance of tissue culture cells and treatment with MG132. Baby hamster kidney (BHK) C13 cells were grown and maintained in Glasgow minimal Eagle’s Medium supplemented with 10% newborn calf serum (NCS), 4% tryptose phosphate broth and 100 IU/ml penicillin–streptomycin (ETC10). Huh7 and HepG2 cells were propagated in Dulbecco’s modified Eagle’s Medium supplemented with 10% foetal calf serum, 2 mM l-glutamine, non-essential amino acids and 100 IU/ml penicillin–streptomycin. To treat BHK cells with MG132 (supplied by Boston Biochem), cells were incubated for 5 h after electroporation at 37 °C in ETC10 and the medium was replaced with fresh medium containing the protease inhibitor at a final concentration of 2.5 μg/ml. Incubation was continued at 37 °C for a further 12 h before the cells were either harvested for Western blot analysis or fixed for indirect immunofluorescence studies. For growth of BHK cells under reduced lipid conditions, NCS was replaced with 10% delipidated bovine calf serum (Sigma). Cells were cultured under reduced lipid conditions for at least 3 weeks at 37 °C and passaged minimally five times before use in experiments.

■ Immunological reagents. Monoclonal antibody JM122 was raised in mice against a fusion protein, purified from bacteria, which was composed of the N-terminal 118 amino acid residues of core protein encoded by HCV strain Glasgow linked to a histidine tag. Antisera R308

Table 1. Construction of HCV core mutants

<table>
<thead>
<tr>
<th>Plasmid</th>
<th>Sequence of oligonucleotide inserted</th>
<th>Enzyme sites used for insertion*</th>
</tr>
</thead>
<tbody>
<tr>
<td>pgHCV/1–173</td>
<td>GTAACCTCTTGGTGTGTCCAGATCTA</td>
<td>BstII(502)/HindIII†</td>
</tr>
<tr>
<td>pgHCV/1–169</td>
<td>GTAACCTTGAGATCTA</td>
<td>BstII(502)/HindIII†</td>
</tr>
<tr>
<td>pgHCV/1–153</td>
<td>CTGGCCGATTGAGATCTA</td>
<td>BstII(453)/HindIII†</td>
</tr>
<tr>
<td>pgHCV/ΔA2–43</td>
<td>CGACCATCGTGGTGTG</td>
<td>KpnI/141/BstII</td>
</tr>
<tr>
<td>pgHCV/ΔA9–43</td>
<td>CGACCATGAGCAAGAATCCCTAACTCAGAATCTCAATTGGTGTCG</td>
<td>KpnI/141/BstII</td>
</tr>
<tr>
<td>pgHCV/ΔA49–75</td>
<td>CCAGGCTGCGCTGCTGGCA</td>
<td>BstII(141)/BglII(233)</td>
</tr>
<tr>
<td>pgHCV/ΔA80–118</td>
<td>TACGGCTTTGGAAGGTA</td>
<td>BglII(233)/ClaI(371)</td>
</tr>
<tr>
<td>pgHCV/ΔA125–144</td>
<td>CGATAGAGCAGCTGCGAGGCC</td>
<td>ClaI(371)/BstII(453)</td>
</tr>
<tr>
<td>pgHCV/ΔA125–134</td>
<td>CGATAGGAGTATACCATCCGTTGCGGCCCCCTC</td>
<td>ClaI(371)/BstII(453)</td>
</tr>
<tr>
<td>pgHCV/ΔA135–144</td>
<td>CGATAGGCTCTTACGCTGACAGGGAAGGAAGGTGC</td>
<td>ClaI(371)/BstII(453)</td>
</tr>
<tr>
<td>pgHCV/ΔA145–154</td>
<td>CATGGGGATATACATGACGGCCGCCTGCGGCCCTC</td>
<td>ClaI(371)/BstII(453)</td>
</tr>
<tr>
<td>pgHCV/ΔA155–161</td>
<td>CTGGGCCCATCCGCTGTTGACGAGAAGATCCAG</td>
<td>BstII(453)/BstII(502)</td>
</tr>
<tr>
<td>pgHCV/ΔA161–166</td>
<td>CTGGGCCCATGCGGTCCTGTTCTCTGAAGAGAGAG</td>
<td>BstII(453)/BstII(502)</td>
</tr>
</tbody>
</table>

* Numbers in parentheses refer to nucleotide positions in Fig. 1a.
† Enzyme sites located in the pGEM-1 vector.
‡ This site was incorporated into the oligonucleotide (italicized in the Table) used to construct pgHCV/ΔA135–144.
was raised in rabbits against a branched peptide, \([\{(A/P)KPQRKTKR-\text{NT}][A/N]RRPQDVKFGG\}K_A\), consisting of amino acid residues 5 to 27 of core protein (Fig. 1a). The two degenerate sites at positions 1 and 12 were introduced to obtain antisera which would be reactive against core proteins from other isolates. Monoclonal antibody ALP98, specific for HCV glycoprotein E2, has been described previously (Patel et al.,...
Fig. 2. Confocal images of the intracellular localization of core proteins and lipid droplets. BHK C13 cells were harvested 20 h after electroporation and fixed with 4% paraformaldehyde, 0.1% Triton X-100. Indirect immunofluorescence was performed with antibody JM122 and an anti-mouse secondary antibody conjugated with FITC. Since the epitope recognized by JM122 lies in the region of core removed in pSFV/Δ49–75, anti-core rabbit antisera 308 was used along with an anti-rabbit secondary antibody to analyse cells expressing this variant of core protein. Lipid droplets were stained with oil red O. Panels (a), (e), (g), (i), (k), (m), (o), (q), (s), (u), (w) and (y) show the distributions of core protein. Panels (b) and (d) show the location of lipid droplets only. Panels (c), (f), (h), (j), (l), (n), (p), (r), (t), (v), (x) and (z) are merged images of core protein and lipid droplets. Cells were electroporated with RNA from the following constructs: panels (a)–(c), pSFV/1–195; panel (d), mock-electroporated; panels (e) and (f), pSFV/1–173; panels (g) and (h), pSFV/1–169; panels (i) and (j), pSFV/1–153; panels (k) and (l), pSFV/Δ2–43; panels (m) and (n), pSFV/Δ9–43; panels (o) and (p), pSFV/Δ49–75; panels (q) and (r), pSFV/Δ80–118; panels (s) and (t), pSFV/Δ125–144; panels (u) and (v), pSFV/Δ145–154; panels (w) and (x), pSFV/Δ155–161; panels (y) and (z), pSFV/Δ161–166.

In vitro transcription and electroporation of SFV RNA into cells. RNA was transcribed in vitro from recombinant pSFV constructs linearized with SpeI and BHK cells were electroporated with transcripts as described in Patel et al. (1999). Huh7 and HepG2 cells were prepared for electroporation in the same manner as BHK cells but were electroporated at 360 V with a capacitance of 950 μF.

Preparation of cell extracts, polyacrylamide gel electrophoresis and Western blot analysis. To prepare extracts, electroporated cells were harvested by removing the growth medium and washing the cell monolayers with PBS. Cells were scraped into PBS and pelleted by centrifugation at 100 g for 5 min at 4 °C. The cell pellet was solubilized in sample buffer (160 mM Tris, pH 6.7, 2% SDS, 700 mM β-mercaptoethanol, 10% glycerol, 0.004% bromophenol blue). Alternatively, sample buffer was added directly to cells that had been washed with PBS. Cells were solubilized at a concentration of approximately 4 × 10^6 cell equivalents per ml sample buffer. Samples were heated to 100 °C for 5 min to fully denature proteins and nucleic acids. Gel electrophoresis was performed on 12.5% polyacrylamide gels cross-linked with 2.5% (w/w) N,N'-methylene bisacrylamide.

For Western blot analysis, proteins separated on polyacrylamide gels were transferred to nitrocellulose membrane. After blocking with 3% gelatine, 4 mM Tris–HCl, pH 7.4, 100 mM NaCl, membranes were incubated with monoclonal antibodies (diluted to 1/500) in 1% gelatine, 4 mM Tris–HCl, pH 7.4, 100 mM NaCl, 0.05% Tween 20. After washing, bound antibody was detected using a horseradish peroxidase-conjugated secondary antibody followed by enhanced chemiluminescence (Amersham).

Indirect immunofluorescence and staining of lipids. Cells on 13 mm coverslips were fixed for 30 min in either methanol at −20 °C or 4% paraformaldehyde, 0.1% Triton X-100 (prepared in PBS) at 4 °C. After washing with PBS and blocking with PBS–1% NCS, cells were incubated with primary antibody (diluted in PBS–NCS at 1/200 for JM122 antibody, 1/1000 for R308 antiserum) for 2 h at room temperature. Cells were washed extensively with PBS–NCS and then incubated with conjugated secondary antibody (either anti-mouse or anti-rabbit IgG raised in goat) for 2 h at room temperature. Cells were washed extensively in solutions of PBS–NCS followed by PBS and finally H_2O before mounting on slides using Citifluor. Samples were analysed using a Zeiss LSM confocal microscope.

After incubation with both antibodies and washing, lipid droplets were stained in paraformaldehyde-fixed cells by briefly rinsing coverslips in 60% propan-2-ol followed by incubation with 0.5 ml 60% propan-2-ol containing oil red O (final concentration approximately 0.6%) for 1.5 to 2 min at room temperature. Coverslips were briefly rinsed with 60% propan-2-ol, washed with PBS and H_2O and mounted as described above. The oil red O staining solution was prepared from a saturated stock of approximately 1% oil red O (Sigma) dissolved in propan-2-ol. Before staining, the stock was diluted with H_2O and then filtered.

1999); the anti-ADRP antibody, AP125, was obtained from Cymbus Biotechnology.
Results

Predicted domain structure of core protein

Previous studies have highlighted the overall basic nature of the HCV core protein. Closer examination of the predicted amino acid sequences for strain Glasgow and other HCV strains reveals that the basic residues are clustered at the N-terminal end of the protein (Bukh et al., 1994). These clusters correspond to hydrophilic regions which are found between residues 1 to 118 as predicted from hydropathy profiles; the overall basic amino acid content in this region is approximately 24% (Fig. 1a, b). Beyond residue 118, there are very few basic residues and the hydrophobicity is generally higher, in particular between amino acids 119 and 145 (Fig. 1a, b). Another area of yet higher hydrophobicity is found from residue 170 to 190. This latter region corresponds to the signal sequence for E1 which is removed during processing of p23 to p21. Based on this analysis, we propose that core protein can be regarded as consisting of three domains (Fig. 1b): the N-terminal region (domain 1; residues 1–118), a central domain (domain 2; residues 119–174) and the signal sequence for E1 (domain 3; residues 175–191). These limits for the domains are somewhat arbitrary: for example the C-terminal end of the central domain would be defined by the events which generate p21 and, as yet, the precise processes that give rise to p21 have not been unambiguously characterized.

Characterization of sequences within core protein necessary for lipid droplet association

Expression of core protein from HCV strain Glasgow in BHK-21 C13 cells from an SFV construct, pSFV/1–195, revealed that a high proportion of the protein was associated with globular vesicles. By combining indirect immunofluorescence for detection of core protein with a lipid-specific stain, oil red O, we identified the globular structures as lipid droplets that are storage compartments containing triacylglycerols and cholesterol esters (Londos et al., 1999; Murphy & Vance, 1999). Representative images of core associated with lipid droplets are shown in Fig. 2(a–c) and 5(a, panel i) at 18 h after electroporation. In cells propagated in normal growth media, core was invariably detected at the surface of lipid droplets and this distribution was apparent from early times (9 h) after electroporation (data not shown). This localization confirms previous findings from electron microscopy (EM) studies that core protein is present at the surface of such intracellular structures (Moradpour et al., 1996; Barba et al., 1997). The sizes of lipid droplets in BHK cells are highly variable (approx. 0.4–3.0 µm in diameter; Fig. 2d). In BHK cells expressing core protein, lipid droplets were rarely greater than 1.4 µm in diameter (Fig. 2a–c), suggesting that association of core with these structures reduces the upper limits of their sizes and hence their size heterogeneity. From Western blot analysis, 95% of the core protein expressed in cells from pSFV/1–195 corresponded to p21 (see Fig. 3a, b, lane 1) and presumably is therefore the predominant species associated with droplets.

The ability to simultaneously detect core protein and lipid droplets in cells offered the possibility of examining the sequences which direct the protein to these structures. Hence, we constructed a series of variants from pSFV/1–195 that either truncated the coding region or lacked internal segments in domains 1 and 2 (Fig. 1; Table 1). Prior reports had indicated that truncation of core protein to either residues 173 (Liu et al., 1997) or 152 (Suzuki et al., 1995) altered intracellular distribution. In our experiments, truncation of the strain Glasgow core coding sequence to residue 173 (Fig. 2e, f) did not prevent lipid droplet association and the distribution of the protein was indistinguishable from that obtained with pSFV/1–195. Removal of a further four amino acids to residue 169 (Fig. 2g, h) also gave identical data to pSFV/1–195. Identical results were obtained for core proteins truncated at residues 173 and 169 in Huh7 cells (data not shown), a liver-derived cell line of human origin. However, a variant truncated at residue 153 did not associate with droplets but had a diffuse distribution throughout the cytoplasm (Fig. 2i, j). In BHK cells fixed with methanol, which permits permeabilization of the nuclear membrane, this core mutant localized also either in the nucleus or at the nuclear membrane (data not shown). Similarly, in Huh7 cells, the core mutant truncated at residue 153 did not locate to droplets although the protein was found either primarily in the nucleus or at the nuclear membrane (data not shown). Hence data from two independent cell lines indicate that residues between 153 and 169 are required for lipid droplet association but the region corresponding to domain 3 that contains the E1 signal peptide is dispensable. Since the truncated form made by pSFV/1–169 is unlikely to contain the processing site that gives rise to p21, we also conclude that, for strain Glasgow, maturation of core protein is not essential for localization to lipid droplets.

Further analysis with variants containing internal deletions showed that lipid droplet association was not abolished by removing various segments spanning residues 2 and 118 (Fig. 2k–r). The most notable features observed with these variants were firstly that there was an increase in the relative sizes of some lipid droplets in a proportion of BHK cells as compared to the sizes found with wild-type core protein. An example is shown for mutant Δ2–43 (panels k and l) and there was apparent fusion of droplets also with mutant Δ9–43 (panels m and n). These phenomena were reproducible in BHK cells but to varying degrees over several experiments. Secondly, removal of residues 80 to 118 did reduce but did not abolish association of core with lipid droplets (Fig. 2q, r). Thirdly, in Huh7 cells, protein levels of the mutant lacking residues 49 to 75 were barely detectable, thus precluding conclusions on the behaviour of this mutant in those cells (data not shown). Overall, localization of core protein at the surface of lipid droplets was not eliminated with the domain 1 variants in either BHK or Huh7 cells. By contrast, lipid droplet association
was not detected for proteins which lacked various regions between residues 125 and 166 (Fig. 2a–2c). Indeed, the only sequences dispensable in this portion of core protein lay between amino acids 155 and 161 (panels w and x). Identical behaviour for each of these mutants was observed in Huh7 cells (data not shown). From these data we conclude that domain 2 contains elements that are essential for lipid droplet association.

Sequences within core required for lipid droplet association also facilitate maturation of precursor protein

During our investigations into the residues required for lipid droplet association, we analysed the intracellular levels of protein made by the core constructs that had been processed at the p21 maturation site. With mutants in domain 2, we found that deletion of residues between 125 and 144 dramatically reduced the level of mature product to only about 25% of the abundance of unprocessed precursor protein (Fig. 3a, lane 4). To further define the sequences in this region which affected maturation, two other constructs were made and these showed that deletion of residues from 125 to 134 reduced maturation efficiency to a greater extent than removal of residues from 135 to 144 (~ 30% as compared to 50%, Fig. 3a, lanes 6 and 7). The removal of other sequences between residues 145 and 166 also gave reduced amounts of mature protein but the reductions were less pronounced (Fig. 3a, lanes 8–10). Furthermore, the levels of mature protein from one mutant which lacked residues 125 to 144 and 161 to 166 were extremely low (approximately 10%; lane 11), suggesting that more than one element may be important in domain 2 for protein maturation. Analysis of a number of these mutants in Huh7 (Fig. 3b) and HepG2 cells (data not shown) also indicated that residues between amino acids 125 and 144 had the most dramatic effect on maturation of core protein.

Identification of sequences that modulate the stability of mature core protein

From analysis of the effect of sequences on levels of p21 produced, we observed that the amount of core protein detected in extracts from cells electroporated with pSFV/1–124,145–169 RNA was lower in a number of experiments than that made in cells electroporated with pSFV/1–169 RNA (Fig. 3a, compare lanes 2 and 5). This could suggest that either translation is inhibited as a result of the removal of these sequences or that the protein produced was unstable and proteolytically degraded. This prompted us to examine the effect of removing amino acids 125 to 144 from the coding sequences of the HCV structural proteins to determine whether protein levels of the glycoproteins as well as core were affected by these residues. Thus, a plasmid called pSFV/CΔ125–141E1E2 (see Methods) was constructed that contains sequences encoding residues 1 to 124 and 145 to 191 of core protein linked to those for E1 and E2. Given that removal of residues 125 to 144 apparently reduced the amount of p21 derived from p23 (Fig. 3a, compare lanes 1 and 2; Fig. 3b, compare lanes 1 and 2), this construct would address also whether generation of p21 was affected by these sequences on linked expression with E1 and E2. For comparative purposes, levels of HCV structural proteins were determined from pSFV/CΔ125–141E1E2.
CE1E2<sub>Gla</sub> RNA, which expresses unmodified core, E1 and E2 from strain Glasgow.

Western blot analysis showed that p21 and E2 proteins were detected in extracts from BHK, Huh7 and HepG2 cells electroporated with pSFV/CE1E2<sub>Gla</sub> RNA (Fig. 4a, panels i and ii, lanes 1, 3 and 5). The amount of E2 made by pSFV/C<sub>Gla</sub><sub>Δ125-144</sub>E1E2 in each cell type was similar to that for pSFV/CE1E2<sub>Gla</sub> (Fig. 4a, panel ii, lanes 2, 4 and 6); however, no core protein was detected (panel i, lanes 2, 4 and 6). Similar relative patterns of expression for core and envelope glycoproteins were observed visually in BHK cells by indirect immunofluorescence (Fig. 4c, panels i, ii, iv and v); in pSFV/C<sub>Gla</sub><sub>Δ125-144</sub>E1E2-electroporated cells, glycoprotein was readily identified but there was almost no fluorescent signal for core protein. This indicates that translation of the structural proteins is not inhibited by deletion of residues 125 to 144. To examine whether protein stability could account for the different levels of core protein detected, electroporated BHK cells were seeded in duplicate and one set of cultures was treated with the protease inhibitor MG132. Compared with untreated cells, there was no significant effect on the amount of E2 detected in MG132-treated samples (Fig. 4b, panel ii, lanes 1–4). Comparable levels of p21 were found also in treated and untreated samples from pSFV/CE1E2<sub>Gla</sub>-electroporated cells (Fig. 4b, panel ii, lanes 3 and 4). In pSFV/C<sub>Gla</sub><sub>Δ125-144</sub>E1E2-electroporated cells, core protein could be detected in MG132-treated but not untreated cells (Fig. 4b, panel i, lanes 1 and 2). Indirect immunofluorescence analysis of pSFV/CE1E2<sub>Gla</sub>...
Localization and stability of HCV core protein

Fig. 5. Properties of core protein in cells grown in delipidated sera. BHK cells were electroporated with RNA from pSFV/1–195 and SFV/1–153 and processed for indirect immunofluorescence (a) or Western blot analysis (b) after 20 h. (a) Electroporated cells had been grown in ETC10 (panels i and iii) and ETC10 prepared with delipidated serum (panel ii). Panels show cells electroporated with pSFV/1–195 (panels i and ii) and pSFV/1–153 RNAs (panel iii). Indirect immunofluorescence was performed with antibody JM122 and lipid droplets were stained with oil red O. (b) Samples were from extracts electroporated with the following constructs: Lanes 1 and 4, pSFV/1–195; lanes 2 and 5, pSFV/1–153; lanes 3 and 6, no RNA. Electroporated cells were grown in ETC10 (lanes 1–3) and ETC10 prepared with delipidated serum (lanes 4–6). Western blot analysis was performed with antibody JM122 and the position of core is indicated. (c) Western blot analysis of endogenous ADRP in cells grown in ETC10 (lane 1) and ETC10 prepared with delipidated serum (lane 2). Samples in lanes 1 and 2 correspond to the mock-electroporated cell extracts in lanes 3 and 6 respectively in (b). Anti-ADRP antibody AP125 was used to analyse samples and the position of ADRP is arrowed.

electroporated cells also revealed that considerably more core protein could be found in MG132-treated than untreated cells (Fig. 4c, panels ii and iii). Thus, residues between 125 and 144 are critical for the stability of p21. Finally, the size of core protein detected in the pSFV/CΔ125–144/E1E2-electroporated cells treated with MG132 corresponded to that of a product which had been processed to give a mature protein with a C terminus at approximately residue 172 (Fig. 4b, panel i, lane 2; data not shown). This is in contrast to the data obtained with construct pSFV/A125–144 in which the HCV coding region is terminated at residue 195, and lower relative levels of processed to unprocessed product were detected (Fig. 3a, lane 4; Fig. 3b, lane 2). This would suggest that the reduced amounts of mature product made by pSFV/A125–144 are a consequence of the removal of residues 125 to 144 as well as the absence of any substantial E1 and E2 coding sequences.

Stability and maturation of core protein is not dependent on lipid droplet association

The above data indicate that sequences which are required for lipid droplet association could also play a role in maturation and stability of core protein. Therefore, it is possible that these properties of core protein are inter-dependent. To investigate this possibility, core protein was analysed in BHK cells which had been grown in delipidated serum. After several passages, cells grown under such conditions had considerably reduced numbers of lipid droplets as observed by staining with oil red O and, in a high proportion of cells, lipid droplets were absent (Fig. 5a, compare panels i and ii). In addition, Western blot analysis (Fig. 5c, lanes 1 and 2) and indirect immunofluorescence (data not shown) showed that the level of adipocyte differentiation related protein (ADRP; Brasaemle et al., 1997), a cellular protein found at the surface of lipid droplets and whose abundance reflects the quantity of lipid droplets in a cell, was reduced by several-fold.

Electroporation of cells with RNA from pSFV/1–195 and pSFV/1–153 showed that the core proteins made from both constructs in cells grown in normal medium had distributions consistent with those shown in Fig. 2 (Fig. 5a, panels i and iii). However, in cells grown in media containing delipidated media, core protein made by pSFV/1–195 remained in the cytoplasm and had a similar localization to the truncated protein produced from pSFV/1–153 (panels ii and iii). From Western blot analysis, the efficiency of cleavage of precursor molecules from pSFV/1–195 to generate p21 was similar in cells grown either under normal conditions or in lipid-deficient serum (Fig. 5b, lanes 1 and 4). Moreover, growth conditions had no effect on the abundance of the protein. Hence, we conclude that association with lipid droplets is not required for
maturation of core protein and reduced amounts of lipid droplets in cells do not significantly affect protein stability.

Discussion

From analysis of hydrophobicity profiles and the nature of its amino acid composition, we have suggested that core protein can be separated into three domains. Domain 1 contains clusters of lysine and arginine residues while domains 2 and 3 have very few and no basic residues respectively. Domain 3 is highly hydrophobic and is the segment, removed during processing of core, that acts as the signal sequence for the E1 glycoprotein. Thus, the mature p21 protein would consist of domains 1 and 2. Compared with the capsid proteins of pesti- and flaviviruses, p21 is considerably larger. For example, the mature capsid proteins of classical swine fever virus (CSFV) and Kunjin virus are 100 (Rümenapf et al., 1993; Stark et al., 1993) and 105 amino acids (Speight & Westaway, 1989) respectively. There is no significant amino acid identity between HCV core and its pesti- and flavivirus equivalents. Nonetheless, domain 1 of core is highly basic (24% of residues), and almost identical proportions of basic residues are found also in the pesti- and flavivirus capsid proteins. Therefore, we propose that domain 2 constitutes a region that is unique to the HCV core protein.

In our experiments, processing of core was not required for lipid droplet association since variants truncated at residues 173 and 169 had identical distributions to the processed p21 form of the protein; these results were reproducible in both liver- (Huh7 cells) and non-liver-derived (BHK cells) cell lines. Our results differ from those of Liu et al. (1997) who presented evidence that processing of core is linked to intracellular localization and that protein truncated at residue 173 accumulated in the nucleus. The difference may be a consequence of the different cell types and HCV strains used in the two studies. We also note that the truncated construct used in their study had an additional isoleucine residue immediately following residue 173 which is not present in the natural HCV sequence. It is possible that this additional residue could influence the intracellular localization of their truncated protein. In agreement with our data, Sabile et al. (1999) also have shown that truncation of the core coding region to residue 173 does not induce nuclear localization of the protein. Further truncation of core protein to residue 153 did abolish lipid droplet association, a result which was consistent in both Huh7 and BHK cells. However, in BHK cells, the truncated protein had both nuclear and cytoplasmic distributions whereas only nuclear localization was observed in Huh7 cells (unpublished data). Other reports have shown that proteins truncated at residues 151 to 153 have a nuclear localization (Lo et al., 1995; Suzuki et al., 1995; Marusawa et al., 1999; Sabile et al., 1999). The cytoplasmic distribution of the protein truncated at residue 153 in BHK cells may reflect differences in nuclear/cytoplasmic trafficking events or other intracellular processes between these cells and other cell types. Nevertheless, there was a clear correlation between the behaviour of the core mutants in BHK and Huh7 cells, indicating that the processes involved in directing core to lipid droplets are similar in cell types of different lineages.

The lipid droplets to which p21 is directed are storage compartments for triacylglycerols and cholesterol esters (Londos et al., 1999; Murphy & Vance, 1999). In mammalian cells, very few proteins that bind to lipid droplets have been identified and the sequences required to direct proteins to these storage compartments are not known. However, a family of plant proteins, called oleosins, found at the surface of oil bodies, have been studied in greater detail (reviewed in Huang, 1992; Murphy & Vance, 1999). Oleosins are considered to play an important structural role in maintaining stability and preventing coalescence of these storage structures (LePrince et al., 1997). They are small proteins that consist of three domains, a central hydrophobic segment of about 70 amino acids which is bounded by N- and C-terminal amphipathic segments (Huang, 1992). Targeting studies have established that the N-terminal and central hydrophobic domains are necessary for association with oil bodies but the C-terminal region may be dispensable (van Rooijen & Moloney, 1995). The N-terminal region has positively charged residues that are presumed to aid neutralization of the negative charges on phospholipids at the oil body surface while the central hydrophobic domain appears to interact with the lipid matrix. A similar series of interactions may occur between core protein and lipid droplets. Thus, domain 2 of core protein may form hydrophobic interactions with lipid droplets while the basic residues in domain 1 could contact with phospholipid at the droplet surface. Such surface contacts need not be sequence-specific. This would be consistent with our findings since there are no specific elements in domain 1 that apparently are essential for lipid droplet association. However, this region is not entirely dispensable (data not shown) and does contribute to efficient localization to droplets. Moreover, removal of portions of domain 1 does alter the morphology and size of droplets detected.

A protein that is truncated at residue 195 and lacks residues 125 to 144 is inefficiently processed to yield a mature core species. Previous studies have indicated that maturation of core is carried out by host cell proteases which are located at the ER (Santolini et al., 1994; Hussy et al., 1996). Hence, the reduced levels of cleaved protein made by the above construct could result from inefficient trafficking of the ribosome/nascent polypeptide complex to the rough ER. It may be significant that the data were obtained with constructs in which the HCV coding region was truncated at residue 195 and therefore core was not expressed as part of a longer polypeptide. Therefore, another possibility is that the combined absence of residues 125 to 144 as well as any extended sequence downstream from the core/E1 cleavage site may not permit stable interaction of the nascent protein–ribosome complex with components at
the ER that are required for processing. In the context of expression as part of a longer polyprotein that contains E1 and E2, removal of residues 125 to 144 does not appear to affect maturation of core but does have a dramatic effect on stability of the mature protein. Based on our studies with MG132, an inhibitor of protein degradation (Lee & Goldberg, 1996), mature core protein lacking this stretch of amino acids is degraded by post-translational processes and it is likely that proteolysis is performed by the proteasome. Since this region of core is required also for lipid droplet association, it may be that, following maturation, inability of the protein to associate with the appropriate cellular components targets the protein for proteolysis. Previously described control mechanisms with other cellular proteins show similar characteristics. For example, apolipoprotein B (apoB) is assembled with cholesterol and triacylglycerol into lipoprotein particles which are secreted by hepatocytes. Intracellular levels of apoB are regulated by the availability of lipid and excess protein is degraded by the ubiquitin/proteasome pathway (Chen et al., 1998, and references therein). In addition, the levels of a lipid droplet binding protein, ADRP, are regulated by post-transcriptional events in which the abundance of the protein is controlled by the apparent availability of a lipophilic surface (Brasaemle et al., 1997). Experiments are in progress to elucidate further the role of residues 125 to 144 in the interactions of core protein with cellular processes.

Our results with cells grown in lipid-deficient serum, and therefore do not accumulate lipid droplets, indicate that the distribution of core is different from that in cells maintained in normal serum. Under these conditions, cleavage of core is efficient and the abundance of the protein is not substantially diminished. Thus, core does not require association with lipid droplets for stability. A number of reports have shown both punctate and reticular localizations for core (Lo et al., 1995; Suzuki et al., 1995; Moradpour et al., 1996; Barba et al., 1997; Yatsu et al., 1999), and from EM studies the protein is present at the surface of lipid droplets and on ER membranes (Moradpour et al., 1996; Barba et al., 1997). In the absence of lipid droplets, core assumes a reticular cytoplasmic distribution in BHK cells and we suggest that this represents association with ER membranes. Therefore, the different extents to which core is found attached either to lipid droplets or membranes may be dependent on the presence of lipid droplets in the various cell types used for analysis. The processes involved in lipid droplet formation are not well understood but it is thought that they bud from ER membranes (Murphy & Vance, 1999). Core may bind to a common lipid or protein component present both on lipid droplets and ER membranes. One example could be the apoAII protein that associates with core, an interaction that requires residues 160 to 173 contained within domain 2 (Sabile et al., 1999).

In conclusion, the HCV core protein contains a domain which is not present in related proteins from pesti- and flaviviruses. This region of the protein plays a vital role in the localization of core to lipid droplets and has a motif which is important for stability of the mature protein. It also may impart other properties to core protein which are relevant to virion morphogenesis and viral pathogenesis and thus provide unique features to the life-cycle of HCV.

We wish to thank Professor R. Elliott and Drs M. McElwee and A. Patel for supplying materials. We are extremely grateful to Dr S. Graham for assistance with the production of antibodies. We are also grateful to D. McGeoch for helpful comments on the manuscript.

References


Kuo, G., Choo, Q.-L., Alter, H. J., Gitnick, G. L., Redeker, A. G., Purcell, R. H., Miyamura, T., Dienstag, J. L., Alter, M. J., Stevens, C. E.,


Localization and stability of HCV core protein


Received 1 December 1999; Accepted 20 April 2000