Analysis of the processing and transmembrane topology of the E2p7 protein of hepatitis C virus

Beverley J. Isherwood and Arvind H. Patel

MRC Virology Unit, Church Street, Glasgow G11 5JR, UK

Hepatitis C virus C, E1, E2 and p7 proteins are cleaved from a viral polyprotein by host signal peptidases. Cleavage at the E2/p7 site is incomplete in genotype 1a strain H (resulting in E2, p7 and E2p7 species), although it has been reported to be more efficient in genotype 1b strain BK. Here, the proteolytic processing and transmembrane topology of genotype 1a strain H77c p7 was investigated when expressed in the context of E2p7. Partial processing was seen at the E2/p7 site in mammalian cells, the efficiency of which improved in the presence of nucleotide sequences downstream of p7. In insect cells, no processing at the E2/p7 site occurred and the uncleaved E2p7 species was incorporated into virus-like particles when expressed in the context of CE1E2p7c-myc. E2p7c-myc formed a heterodimer with E1, indicating that, like the well-characterized E1–E2 complex, the E1–E2p7 heterodimer may also play a functional role in virus replication. Comparison of the p7 signal peptide sequences of strains BK and H77c revealed 3 aa differences (positions 720, 733 and 742). Mutational analysis showed that the V720L change in the H77c sequence substantially increased processivity at the E2/p7 site. The p7 protein adopts a double membrane-spanning topology with both its N and C termini orientated luminally in the endoplasmic reticulum. The transmembrane topology of E2p7 species was examined by two independent means. In both cases, the C terminus of p7 in E2p7 was found to be cytoplasmically orientated, indicating that p7 adopts a dual transmembrane topology.

INTRODUCTION

Hepatitis C virus (HCV), a member of the family Flaviviridae, is an enveloped virus containing a positive-strand genomic RNA encoding a single polyprotein of approximately 3010 aa that is processed co- and post-translationally by host and viral proteases into at least 10 different proteins (C, E1, E2, p7, NS2, NS3, NS4A, NS4B, NS5A and NS5B) (Lindenbach & Rice, 2001). The HCV structural proteins, core (C) and the two envelope glycoproteins E1 and E2, are located within the N terminus of the polyprotein, whilst the non-structural proteins (NS2 to NS5B) reside within the C-terminal part. The status of the p7 protein is currently unknown, although it has been shown to have an ion-channel activity (Griffin et al., 2003, 2004; Pavlovic et al., 2003; Premkumar et al., 2004) and is essential for infectivity of HCV (Sakai et al., 2003).

The C terminus of each structural protein is composed of a hydrophobic amino acid sequence, which acts as a signal peptide to target the proteins located downstream to the endoplasmic reticulum (ER). Cleavage at the C/E1, E1/E2, E2/p7 and p7/NS2 sites is mediated by ER-resident host signal peptidase(s) (Lindenbach & Rice, 2001; Rosenberg, 2001). Although cleavage at the C/E1 and E1/E2 sites proceeds to completion rapidly after translation, cleavage at E2/p7 and p7/NS2 is delayed, resulting in an E2p7NS2 species (Dubuisson et al., 1994). Furthermore, cleavage at the E2/p7 site has been shown to be incomplete, resulting in two E2-specific species, E2 and E2p7 (Lin et al., 1994; Mizushima et al., 1994). The significance of these two forms of E2 has not been established, although it is conceivable that both play functional roles in the virus life cycle.

The inability to culture HCV remains a major obstacle to the study of virus assembly. HCV virus-like particle (VLP) assembly can occur in the absence of p7 upon expression of C, E1 and E2 in insect (but not mammalian) cells (Baumert et al., 1998, 1999; Clayton et al., 2002; Owsianka et al., 2001). A homologue of p7 (but not the E2p7 species) in the related pestivirus, bovine viral diarrhea virus (BVDV), is required for the production of infectious virus progeny (Harada et al., 2000). By analogy with BVDV p7 and its presumed role as an ion channel, it is possible that HCV p7 is important for virus particle morphogenesis.

HCV p7 is an integral membrane protein, which is translocated into the ER by the signal peptide located in the C terminus of E2 (Carrere-Kremer et al., 2002; Cocquerel et al., 2002). When expressed on its own in mammalian cells, p7 has been shown to have two membrane-spanning domains with its N and C termini luminaly disposed and a short hydrophilic loop facing the cytosol. The second transmembrane (TM) domain of p7 can act as a signal...
peptide to target NS2 to the ER, although the membrane association of NS2 can occur independently of any p7 sequences (Yamaga & Ou, 2002).

In this study, we investigated the proteolytic processing and TM topology of p7 when expressed in the context of E2p7. We have shown that the C terminus of the E2p7 species is cytoplasmically orientated, indicating that p7 adopts a dual TM topology and that processing at the E2/p7 site is influenced by sequences within the signal peptide of p7 and also sequences in the NS2 region.

**METHODS**

**Cell culture, plasmid constructs and recombinant viruses.** *Spodoptera frugiperda* Sf21 (Sf21) insect cells and human hepatoma (Huh-7) (Nakabayashi et al., 1982) cells were cultured as described previously (Clayton et al., 2002). The cDNA sequences used here originated from an infectious cDNA clone (pCV-H77c) of strain H77c genotype 1a (Yanagi et al., 1997). The cDNA sequences from nt 336–9402 or 336–2850 encoding the entire open reading frame (ORF; aa 1–3011; Fig. 1) or aa 1–836 [representing CE1E2p7 and the N-terminal 27 aa residues of NS2 (NS2t)] of HCV, respectively, were cloned into the mammalian expression vector pCDNA3.1/Zeo (+) (Invitrogen). A recombinant vaccinia virus (rVV) expressing CE1E2p7NS2t has been described previously (Clayton et al., 2002). HCV E1E2 (aa 167–746), E1E2p7 (aa 167–809) and E1E2p7-c-myc (aa 167–809 followed by the c-myc tag) were expressed from pCDNA3.1/Zeo (+) or using rVVs generated as described previously (Clayton et al., 2002). Recombinant baculoviruses (rbacs) expressing CE1E2, CE1E2p7 and CE1E2p7c-myc were generated using the Bac-to-Bac Expression System (Invitrogen).

Site-specific mutations in the p7 signal sequence were introduced into the plasmid pCDNA/E1E2p7NS2t by fusion PCR using oligonucleotide primers carrying appropriate changes. The presence of the mutations was confirmed by nucleotide sequencing.

**Transfection of cultured cells.** Huh-7 cells were infected with vTF7-3, an rVV expressing T7 RNA polymerase (Fuerst et al., 1985), at an m.o.i. of 5 for 1 h at 37°C. The cells were then transfected with appropriate plasmids using a liposome-mediated method (Rose et al., 1991). Following incubation for 18 h at 37°C, cells were washed with PBS, lysed in lysis buffer (20 mM Tris/HCl pH 7.4, 20 mM iodoacetamide, 150 mM NaCl, 1 mM EDTA, 0.5–5% Triton X-100) and the lysate spun briefly to remove nuclei. The clarified lysates were subjected to Western immunoblotting using appropriate antibodies and bound antibodies were detected using enhanced chemiluminescence reagents (Amersham). The quantification of protein bands in the Western immunoblot was performed using Quantity One Volume Analysis software (Bio-Rad).

**Radiolabelling of proteins.** Huh-7 cells were infected with appropriate rVVs at an m.o.i. of 10 and incubated for 6 h at 37°C. Infected cells were washed with PBS, incubated in methionine-free medium containing 50 µCi (1.85 MBq) [35S]methionine ml⁻¹ for 18 h and the cell lysate prepared as described above. To immunoprecipitate proteins, radiolabelled cell lysates were incubated with anti-E2 or anti-c-myc antibodies for 2 h at 4°C and the immune complexes precipitated by incubation at 4°C for 2 h with protein A-Sepharose equilibrated with lysis buffer. Immune complexes attached to protein A-Sepharose were washed five times with lysis buffer and bound proteins were either directly released into SDS-PAGE denaturing buffer or were subjected to peptide N-glycosidase F (PNGase F) (New England Biolabs) treatment following the manufacturer’s protocol. Immune complexes were subjected to 10% SDS-PAGE. Gels were dried and exposed overnight to a phosphor screen and radiolabelled proteins were visualized with a Bio-Rad Personal FX phosphorimager.

**Antibodies.** Hybridomas secreting the mouse monoclonal antibody (mAb) 9E10 (Evan et al., 1985) to human c-myc product were obtained from the European Collection of Cell Cultures. The anti-E2 antibodies AP33 and ALP98, and the rabbit polyclonal anti-serum R646 have been described previously (Clayton et al., 2002; Owsianka et al., 2001). The anti-E2 mAb H53 was kindly supplied by J. Dubuisson (Cocquerel et al., 1998).

**VLP preparation.** VLPs from Sf21 cells infected with rbacs were prepared essentially as described by Baumert et al. (1998). The purity and quality of the VLP preparation were verified by negative-stained transmission electron microscopy (EM) as follows. VLPs (5 µl) were loaded on to Formvar-coated nickel grids, stained with Nanovan (Nanoprobe) and examined under a JEOL 100 S electron microscope. For immunogold labelling, samples loaded on to Formvar-coated nickel grids were incubated in primary antibody for 2–3 h at room temperature. The grids were washed three times in distilled water and incubated for 2 h with anti-mouse 1g conjugated to 10 nm gold particles (Nanoprobe). Following three washes as above, samples were stained with Nanovan and examined by EM.

**Trypsin protection analysis.** Huh-7 cells were infected with appropriate rVVs and proteins radiolabelled as described above. Following incubation, cells were rinsed once with ice-cold PBS, twice with homogenization buffer (10 mM HEPES pH 7.4, 1 mM EDTA, 1 mM PMSF, 0.25 M sucrose) and lysed in this buffer using a tight-fitting Dounce homogenizer. After a brief centrifugation, the post-nuclear supernatant containing microsomal membranes was overlaid on a cushion of 0.6 M sucrose in 10 mM HEPES pH 7.4, 1 mM EDTA, 1 mM PMSF and centrifuged at 20000 r.p.m. in a Sorval TH641 rotor. The pellet containing microsomal membrane was washed and resuspended in PBS. For trypsin protection analysis, the microsomal preparation was divided into three equal aliquots. One sample was left untreated while the remaining two were treated with 50 µg freshly prepared trypsin (Sigma) ml⁻¹ in the presence or absence of 1% NP-40 for 1 h at 37°C. Proteolysis was halted by the addition of 30 µg aprotinin (Sigma) ml⁻¹ and further incubation on ice for 1 h. NP-40 was added to all samples to a final concentration of 1% and incubated on ice for 1 h. Samples were spun at

![Fig. 1. Schematic representation of constructs used to express HCV proteins. The structure of the cDNA encoding the full-length (FL) ORF is shown at the top. Numbers refer to the relevant amino acid residues. The indicated polyproteins were expressed from a mammalian expression plasmid (pCDNA3.1) or using rVVs. The rbacs used in this study additionally contained C protein-encoding sequences.](attachment:fig1.png)
13 000 r.p.m. for 5 min, the resultant supernatant subjected to immunoprecipitation using mAb ALP98 or 9E10 and the immune complexes analysed by SDS-PAGE.

**Immunofluorescence assay.** Huh-7 cells grown on coverslips were transfected as described above (but in the absence of vTF7.3 infection) with the plasmid expressing E1E2p7c-myc. Following incubation at 37 °C for 48 h, cells were washed with PBS, fixed with 2 % paraformaldehyde for 5 min at room temperature and permeabilized, either selectively or completely, as described by Crystal et al. (2003). Briefly, selective permeabilization was performed by incubating cells with 200 U streptolysin O (SLO; Sigma) ml⁻¹ for 5 min at room temperature followed by rinsing with PBS. For complete solubilization, cells were incubated with 0.1 % NP-40 at room temperature for 10 min and then washed with PBS. All cells were blocked with PBS containing 2 % FCS for 20 min at room temperature and then probed with the anti-E2 rabbit antiserum R646 and anti-c-myc mAb 9E10 for 2 h at room temperature. Following three washes with PBS, coverslips were mounted on glass slides and examined under a Zeiss laser-scanning confocal microscope. The fluorescent images were analysed using the LSM510 software.

## RESULTS

### Evaluation of proteolytic processing at the E2/p7 site in mammalian cells

Due to the lack of specific antibodies to HCV p7, antigenically tagged derivatives of this protein have mainly been used to study its properties in mammalian cells (Carrere-Kremer et al., 2002; Griffin et al., 2004; Lin et al., 1994). In this study, we inserted a c-myc tag in frame immediately downstream of the C-terminal residue of p7 to study the characteristics of p7 in mammalian and insect cell systems.

Radiolabelled proteins from Huh-7 cells infected with rVV that express various HCV sequences (Fig. 1) were immunoprecipitated with the anti-E2 mAb ALP98 or the anti-c-myc mAb 9E10 and analysed by SDS-PAGE, either directly or following treatment with PNGase F, a glycosidase that removes all N-linked sugar moieties attached to glycoproteins. As shown in Fig. 2(a, lanes 1–6), the anti-E2 mAb immunoprecipitated E2 species of different molecular masses from cells infected with various rVVs, indicating differences in glycosylation and/or processing at the E2/p7 site. The differences in the relative SDS-PAGE mobilities of the E2 species synthesized in cells infected with various rVVs were better resolved when the immunoprecipitated proteins were treated with PNGase F (Fig. 2a, lanes 9–16). Thus, rVVs expressing the full-length (FL) ORF and E1E2p7NS2t (Fig. 2a, lanes 10 and 11) produced both a fully processed E2 that co-migrated with the product synthesized by rVV-E1E2 (Fig. 2a, lane 12) and a slower-migrating form representing E2p7, indicating partial processing at the E2/p7 site. Cells infected with rVV-E1E2p7 and rVV-E1E2p7c-myc produced predominantly E2p7 and E2p7 plus E2p7c-myc, respectively, as well as a small amount of fully processed E2 (Fig. 2a, lanes 13 and 14). The relative amount of fully processed E2 was higher in

![Fig. 2](http://vir.sgmjournals.org)
cells expressing the FL ORF or E1E2p7NS2t compared with those lacking sequences downstream from p7 (Fig. 2a, lanes 10–14). Taken together, these results are in keeping with previously published observations that proteolytic processing at the E2/p7 site is incomplete in the genotype 1a sequence (Lin et al., 1994; Mizushima et al., 1994) and also indicated that sequences downstream of p7 may enhance cleavage efficiency at this site.

Immunoprecipitation using the anti-c-myc mAb 9E10 confirmed that the larger of the two E2 species produced by rVV-E1E2p7c-myc was E2p7c-myc (Fig. 2a, lanes 8 and 16). Interestingly, PNGase F treatment revealed a band of smaller molecular mass co-immunoprecipitated by the anti-c-myc mAb, which co-migrated with E2p7 (lanes 14 and 16), indicating partial processing at the p7/c-myc junction and possible homodimer formation between E2p7 and E2p7c-myc. The latter observation is in keeping with the E2 model proposed by Yagnik et al. (2000), which predicts that E2 may exist as a head-to-tail homodimer.

To test whether E2p7 or E2p7c-myc formed a heterodimer complex with E1, radiolabelled extracts of cells expressing various HCV sequences were immunoprecipitated with the conformation-sensitive anti-E2 mAb H53 (which specifically recognizes the native E1E2 complex but not the disulfide-linked aggregate) (Cocquerel et al., 1998; Duvet et al., 1998) or the anti-c-myc mAb 9E10. Analysis of the corresponding immune complexes confirmed that both E2p7 and E2p7c-myc formed a heterodimer with E1 (Fig. 2b, lanes 5, 6 and 8). As expected, the native E1E2 heterodimer was also detected in cells expressing E1E2, E2E2p7NS2t and the FL ORF (Fig. 2b, lanes 2–4). That E2p7c-myc interacted with E1 was further confirmed using an anti-E1 antiserum, which specifically co-immunoprecipitated this species (data not shown).

**Processing at the E2/p7 site is deficient in an insect cell system**

Previous studies have shown assembly of HCV VLPs upon expression of viral structural proteins (C, E1 and E2) in insect cells infected with rbac (Baumert et al., 1998; Owsianka et al., 2001). Here, as part of our interest in understanding the mechanisms of VLP assembly, we compared the characteristics of HCV structural proteins expressed in insect cells as CE1E2, CE1E2p7 or CE1E2p7c-myc. Radiolabelled proteins from Sf21 cells infected with different rbacs were immunoprecipitated with an anti-E2 antiserum and the immune complexes treated with PNGase F. As shown in Fig. 3(a), E2 species representing E2, E2p7 or E2p7c-myc from cells infected with rbacs encoding CE1E2, CE1E2p7 or CE1E2p7c-myc, respectively, were immunoprecipitated (Fig. 3a, lanes 2–4). Interestingly, no processed E2 was observed in cells infected with rbac-CE1E2p7 or rbac-CE1E2p7c-myc, and a substantial portion of c-myc remained a part of E2p7c-myc. As seen in Huh-7 cells (Fig. 2a, lane 16), there was processing at the p7/c-myc junction in the E2c-myc AceNPV system, and this species was also present in the Sf21 samples (Fig. 3a, lanes 2 and 3).

**Fig. 3.** Analysis of proteolytic processing at the E2/p7 site in insect cells and immunogold labelling of HCV VLPs. (a) Sf21 cells were infected with wild-type baculovirus Autographa californica nuclear polyhedrosis virus (AcNPV) or rbacs expressing HCV proteins as shown. Infected cells were radiolabelled with [35S]methionine and the proteins immunoprecipitated with the anti-E2 antiserum R646. Following treatment with PNGase F, immune complexes were analysed by 10% SDS-PAGE under reducing conditions. The positions of protein size markers are shown. (b) VLPs synthesized from Sf21 cells infected with rbacs expressing HCV CE1E2 (VLP<sub>CE1E2</sub>), CE1E2p7 (VLP<sub>CE1E2p7</sub>) or CE1E2p7c-myc (VLP<sub>CE1E2p7c-myc</sub>) were placed on EM grids and labelled with anti-E2 mAb AP33 or anti-c-myc mAb 9E10 followed by anti-mouse IgG conjugated to 10 nm gold particles as described in Methods. Grids were washed, stained and examined by EM.
E2p7c-myc is present on VLPs produced in insect cells

It is currently unclear whether p7 is a component of the HCV virion, although recombinant VLPs can form in its absence. Given that partial or no processing occurs between E2 and p7 in mammalian or insect cells, it was important to consider whether E2p7 as an unprocessed form might indeed be incorporated into virus particles. VLPs were prepared from S21 cells infected with rbacs expressing CE1E2, CE1E2p7 or CE1E2p7c-myc to identify whether E2 as an E2p7c-myc (or E2p7) form allowed particle formation. The presence of E2p7c-myc within the VLP preparation was confirmed by Western blotting using anti-E2 and anti-c-myc mAbs (data not shown). VLPs were visualized by transmission EM following immunogold labelling. Anti-E2 mAb was found to bind to all VLPs, whereas the c-myc antibody showed specific labelling of c-myc on the outer surface of the VLPs derived from CE1E2p7c-myc, but not from CE1E2p7 (Fig. 3b).

Mutational analysis of the sequence that acts as a signal peptide for p7

It has previously been reported that processing at the E2/p7 site of the genotype 1b strain BK is more efficient than that in genotype 1a strain H (Lin et al., 1994), an isolate closely related to strain H77c used in this study. The C terminus (aa 731–746) of E2 acts as a signal peptide to translocate p7 into the ER (Carrere-Kremer et al., 2002; Cocquerel et al., 2002). Presumably specific amino acid residues present in this region also play a role in cleavage between E2 and p7. Comparison of the signal peptide sequence present at the C terminus of E2 of strains BK (Takamizawa et al., 1991) and H77c (Yanagi et al., 1997) revealed 3 aa differences (Fig. 4a). To test whether the changes L720V, A733S and A742S in the H77c sequence relative to that in the BK strain were responsible for the genotype-dependent differences in the efficiency of processing at the E2/p7 site, we substituted valine, serine and alanine at residues 720, 733 and 742 for leucine (V720L), alanine (S733A) and alanine (S742A), respectively. In addition, another substitution mutation, valine to isoleucine at position 719 (V719I), was introduced inadvertently into some of the constructs. These mutations were introduced individually and in different combinations, into the construct expressing the strain H77c-encoded E1E2p7NS2t. Plasmids encoding E1E2, E1E2p7NS2t or those carrying different substitutions were transfected in rVV vTF7.3-infected Huh-7 cells. Eighteen hours after transfection, cell lysates were subjected to PNGase F treatment followed by Western immunoblotting using an anti-E2 antiserum. As seen previously in Fig. 2(a), cells expressing unmodified E1E2p7NS2t synthesized E2p7 and E2, with the latter species co-migrating with E2 expressed in cells transfected with the construct encoding E1E2 (Fig. 4b). The level of processed E2 was 55% of the total amount of the two products. A similar E2p7:E2 ratio was found with the substitution V719I, whereas slightly reduced amounts of processed E2 were seen with S733A+S742A and S742A mutations. In contrast, mutations involving aa 720, either alone (V720L) or together with the other residue substitutions (V719I+V720L, V719I+V720L+S733A, V719I+V720L+S733A+S742A), produced significantly increased amounts of E2 species relative to E2p7. This indicated that leucine at position 720 plays a crucial role in proteolytic processing at the E2/p7 site. In addition, isoleucine at position 719 and the two alanines at positions 733 and 742 acted co-operatively with L720 to enhance processivity at the E2/p7 site further.

TM topology of E2p7

The observation that the c-myc tag in our constructs was exposed externally on the surface of VLPs is in agreement with the topology predicted for a classical route of assembly via the constitutive secretory pathway. Furthermore, a small percentage of a fully processed p7 is thought to localize to the plasma membrane, where the c-myc tag can also be visualized at the cell surface (Carrere-Kremer et al., 2002) (data not shown). This indicates that the C terminus of p7

Fig. 4. Mutational analysis of the sequence upstream of p7. (a) Comparison of aa 719–746 immediately upstream from p7 of HCV genotypes 1a and 1b, strains H77c and BK, respectively. The variant amino acids at positions 720, 733 and 742 are underlined. (b) Processing at the E2/p7 site of the mutated polyproteins. vTF7.3-infected Huh-7 cells were transfected with plasmid expressing unmodified (wt) E1E2p7NS2t or its mutated derivatives as shown. Following incubation for 18 h, cell lysates were treated with PNGase F, subjected to 10% SDS-PAGE and the fractionated proteins analysed by Western immunoblotting using the anti-E2 antisemur R646. Numbers below each lane show the amount of fully processed E2 as a percentage of the total products.
is exposed to the extracellular environment, in agreement with the suggestion that signal sequences at its N and C termini should be localized towards the luminal side of the ER.

To investigate the TM topology of inefficiently processed p7, trypsin protection assays of radiolabelled microsomes prepared from Huh-7 cells infected with rVV-E1E2p7 or rVV-E1E2p7c-myc were performed. Microsomes were either untreated or treated with trypsin in the presence or absence of 1% NP-40 prior to immunoprecipitation with anti-E2 or anti-c-myc mAbs and analysed. As shown in Fig. 5(a, lanes 1, 2, 4 and 5), mAb ALP98 recognized E2 (and co-precipitated E1) in untreated or trypsin-treated microsomes of cells infected with both rVVes. Interestingly, the E2 band precipitated by mAb ALP98 from the microsomes of rVV-E1E2p7c-myc-infected cells treated with trypsin in the absence of detergent migrated at a slightly faster rate than that from the corresponding untreated sample (Fig. 5a, lanes 5 and 4, respectively) (the difference in the molecular

**Fig. 5.** Analysis of the TM topology of E2p7. (a) [35S]methionine-labelled crude microsomes of Huh-7 cells infected with rVV expressing E1E2p7 or E1E2p7c-myc were either left untreated or treated with trypsin in the presence or absence of NP-40. Following inactivation of trypsin with aprotinin, the solubilized microsomal proteins were immunoprecipitated with anti-E2 mAb ALP98 (lanes 1–6) or anti-c-myc mAb 9E10 (lanes 7–9) and the immune complexes analysed by 10% SDS-PAGE. The arrow shows the ~18 kDa protected fragment in lane 8. Positions of the molecular mass markers are indicated. Microsomes of Huh-7 (b) and Sf21 (c) cells infected with rVV and rbac, respectively, expressing E1E2p7 or E1E2p7c-myc were subjected to trypsin treatment or not as above followed by Western immunoblotting using anti-E2 mAb ALP98 (b, lanes 1–6) or anti-c-myc mAb 9E10 (b, lanes 7–9; c, lanes 1–6). The arrow indicates the ~18 kDa protected fragment in lanes 8 and 5 in (b) and (c), respectively. The E1E2, E1E2p7 and E1E2p7c-myc precursor bands in (a) and (b) are indicated by asterisks. (d) Huh-7 cells transfected with a plasmid encoding E1E2p7c-myc were fixed and either permeabilized completely with NP-40 or semi-permeabilized with SLO. Cells were then probed with anti-c-myc mAb 9E10 and anti-E2 rabbit polyclonal serum R646 and the bound antibodies detected using Alexa Fluor 488 goat anti-mouse IgG and Alexa Fluor 633 goat anti-rabbit IgG. Labelled cells were analysed by confocal microscopy.
mass was more apparent at low exposure). This indicated possible trimming of E2p7c-myc, most probably at the C-terminal end. As expected, neither E1 nor E2 was observed in the trypsin-treated detergent-solubilized microsomes (Fig. 5a, lanes 3 and 6). The epitope recognized by mAb ALP98 is located within the E2 ectodomain (Clayton et al., 2002), which is luminaly disposed, and therefore is expected to be available for trypsin degradation only in detergent-solubilized microsomes. Our mAb ALP98 results shown in Fig. 5(a) are in accordance with this hypothesis. In contrast to the mAb ALP98 data, immunoprecipitation with the anti-c-myc mAb showed that the E2p7 species produced by rVV-E1E2p7c-myc was degraded upon treatment of microsomes in the absence as well as in the presence of detergent (Fig. 5a, lanes 8 and 9), indicating the c-myc epitope tag to be cytoplasmically exposed. Interestingly, a trypsin-resistant fragment of approximately 18 kDa was seen following treatment of microsomes from cells expressing E1E2p7c-myc with trypsin in the absence of detergent (Fig. 5a, lane 8).

Similarly, Western immunoblotting of microsomes of Huh-7 and Sf21 cells expressing CE1E2p7 or CE1E2p7c-myc showed that the C-terminal c-myc tag of E2p7 was susceptible to trypsin digestion in the absence as well as in the presence of NP-40 (Fig. 5b and c, lanes 8 and 9, and 5 and 6, respectively), again indicating that this portion of E2p7 is cytoplasmically disposed. In contrast, the anti-E2 mAb recognized the bulk of full-length E2 in trypsin-treated microsomes of Huh-7 (Fig. 5b, lanes 2 and 5) and Sf21 cells (data not shown). As with the immunoprecipitation experiment above, the anti-c-myc mAb-reactive trypsin-resistant fragment of approximately 18 kDa was present in microsomes treated with trypsin alone (Fig. 5b and c, lane 8 and 5, respectively). Together, this indicated a protease cleavage site within the C terminus of E2. Furthermore, several trypsin-resistant fragments reactive with anti-E2 mAb were also consistently seen (Fig. 5b, lanes 2, 3, 5 and 6), surprisingly even in the absence of membrane solubilization with NP-40 (Fig. 5b, lanes 2 and 5). The reason for the latter observation is at present unclear; it may reflect a possible mixed TM topology for E2, although the immunofluorescence data of selectively permeabilized cells do not support this hypothesis (see Fig. 5d and unpublished observations). Clearly, further work is warranted to determine the nature of the protected species, which may shed more light on the TM topology of E2 and E2p7.

A small central portion (aa 779–781) of p7 has been shown to be cytoplasmically orientated (Carrere-Kremer et al., 2002). Therefore, it is possible that trypsin in the biochemical assays above may have cleaved at this cytoplasmic loop, thus generating a fragment too small to be observed on the gel, hence allowing the possibility of the c-myc tag being disposed on the luminal side of the membrane. To rule out this possibility, immunofluorescence microscopy was performed under different permeabilization conditions. Huh-7 cells infected with rVV-E1E2p7c-myc were fixed and treated with 1% NP-40 to permeabilize all cell membranes (to detect both cytoplasmically and luminally disposed proteins) or selectively permeabilized with SLO leaving the ER and Golgi membranes intact (thus allowing detection of only cytoplasmically disposed proteins). Fixed cells were then co-probed with the rabbit anti-E2 antisemur R646 and the anti-c-myc mAb 9E10. The antiserum R646 (raised against the luminally disposed E2 ectodomain) recognized E2 in cells permeabilized with NP-40 but not with SLO (Fig. 5d). In contrast, the anti-c-myc mAb recognized its epitope in E2p7c-myc in cells permeabilized with both NP-40 and SLO (Fig. 5d), consistent with the biochemical data above (Fig. 5a–c). Together, these results confirmed that the C-terminal portion of E2p7c-myc is cytoplasmically orientated. Given that this portion of an epitope-tagged p7, when expressed alone, is luminaly disposed (Carrere-Kremer et al., 2002), it seems that p7 is capable of adopting different topologies at various stages in virion morphogenesis.

**DISCUSSION**

We examined the proteolytic processing and TM topology of the HCV p7 protein encoded by genotype 1a strain H77c, when expressed in the context of E2p7. In keeping with previous reports (Lin et al., 1994; Mizushima et al., 1994), incomplete processing at the E2/p7 site was seen in mammalian cells. Our results further showed that the efficiency of processing at this site was enhanced in the presence of a portion of NS2 sequences downstream of p7, although they were not sufficient for complete cleavage to occur. This increase in cleavage efficiency may be due to additional sequence requirements for accessibility by host enzymes to the E2/p7 site during or shortly after translation. Further extension of the NS2 sequences in the relevant construct did not alter the efficiency of cleavage (data not shown). This was in keeping with our observation that processing at the E2/p7 site in strain H77c was also inefficient when the entire viral polyprotein was expressed (Fig. 2a). Together, these results indicated that amino acid residues in the p7 signal peptide (which is located at the C terminus of E2) (Carrere-Kremer et al., 2002; Cocquerel et al., 2002) and/or in p7 itself may also influence the processing events at the E2/p7 site. Indeed, comparison of the p7 signal peptide sequence of HCV genotype 1b strain BK (in which E2p7 processing is more efficient) (Lin et al., 1994; Mizushima et al., 1994) and genotype 1a strain H77c revealed three amino acid differences at positions 720, 733 and 742. Mutational analysis of this region of H77c p7 revealed that the proteolytic cleavage still remained incomplete suggested that amino acid residues in the p7 sequence itself might also be necessary. Carrere-Kremer et al. (2004) recently reported that the E2/p7 cleavage site, together with the N terminus...
of p7, might form an α-helix, which could influence processing at this site. Indeed, destabilization of this structure by insertion of threonine between 1 and 2 (P1’ and P2’) of p7 (or insertion of an epitope linker sequence between P3’ and P4’ of the E2/p7 cleavage site) led to an improvement in the efficiency of cleavage at the E2/p7 site, indicating that structural elements located at the P’ side of the cleavage site also impose constraints on processing at this site. The significance of the different ratios of E2p7:p7 produced between genotypes is unclear, but it is interesting that in a chimpanzee infection model the N and/or C termini of p7 contained sequences with genotype-specific functions (Sakai et al., 2003).

The conservation of suboptimal cleavage at the E2/p7 site in related pestiviruses has been suggested to indicate a common function (Elbers et al., 1996). Although the E2p7 species in these viruses is not essential for replication (Elbers et al., 1996; Harada et al., 2000), it may still play an important role in virus morphogenesis. The results presented here have identified for the first time the formation of a heterodimer between E2p7 and E1, indicating that, like E1–E2, the E1–E2p7 heterodimer may also play a functional role in virus replication.

The reasons for the complete lack of cleavage at the E2/p7 site in insect cells is at present not clear. The host-derived signal peptidase(s) that cleave HCV structural proteins from the polyprotein remain unidentified. Given that efficient proteolytic processing at the core/E1 and E1/E2 sites occurs in insect cells, it is tempting to speculate that a novel proteinase, not present in Sf cells, may be responsible for cleavage at the E2/p7 site. Identification of the host signal peptidase(s) responsible for cleavage of CE1E2p7 may reveal a mechanism for the peculiar processing events observed.

HCV p7 is an integral membrane protein possessing a double membrane-spanning topology with both of its termini luminally disposed (Carrere-Kremer et al., 2002) (Fig. 6b). Our investigation into the TM topology of E2p7 by two independent means intriguingly identified a cytoplasmically disposed C-terminal portion of p7 present both in mammalian and insect cells. Taken together, this indicates a dual TM topology for p7 (Fig. 6a and b). It is not unprecedented for viral membrane proteins to adopt mixed TM topology. For example, the N terminus of HCV NS4B has been reported to assume dual TM topology, with possible distinct functions, on each side of the ER membrane (Lundin et al., 2003). The hepatitis B virus surface L glycoprotein can act in virus assembly as a matrix-like protein and in virus entry as a receptor-binding protein by adopting different TM topologies (Lambert & Prange, 2001). Moreover, the Newcastle disease virus fusion protein exists in two topological forms with respect to membranes, one of which has been proposed to be involved in cell-to-cell fusion (McGinnes et al., 2003).

It is premature to assign E2p7 as a virus structural protein. The fact that HCV VLP CE1E2p7c-myc produced in insect cells was indistinguishable from VLP CE1E2 may suggest that E2p7 is a component of the HCV virion. However, given that available data with homologues in pestiviruses indicate otherwise, this issue is likely to remain unresolved until a fully productive HCV cell-culture system and specific antibodies to p7 become available. It is currently unclear how HCV VLP assembly occurs in insect cells, where accumulation of the particles has been observed in cytoplasmic vesicles or vacuoles (Baumert et al., 1998), although the particles are not secreted. Flavivirus virions accumulate in intracellular vesicles derived from the ER before leaving by the exocytic pathway (Mackenzie & Westaway, 2001; Pettersson, 1991). Similarly, an accumulation of HCV particles has been observed intracellularly by EM in a human T-cell line (Shimizu et al., 1996) and complex N-linked glycans have been found on the surface of virus particles isolated from patients, suggesting virus egress through the Golgi (Sato et al., 1993). If it is assumed that VLP assembly in insect cells occurs by budding of the viral structural proteins into the ER, then given its cytoplasmic orientation the c-myc portion of E2p7c-myc would be expected to be located inside the particles. However, our immuno-EM analysis of VLP CE1E2p7c-myc identified c-myc as exposed on the surface of the particles. This may be due to an artefact of the system in as much as the VLPs failed to egress from the cells, reflecting an incomplete assembly process.

Carrere-Kremer et al. (2004) postulated that the existence of an E2p7 precursor may be a means of keeping p7 inactive in terms of its ion-channel function during glycoprotein maturation and particle assembly or in regulation of cleavage at E2/p7 and/or p7/NS2 cleavage sites. Another possibility is that inefficient cleavage of the E2/p7 site may relate to a mechanism to prevent premature fusion of E2 during virus egress. Similar processing events have been reported in other viruses, including that of the flavivirus prM/E site, whose cleavage is delayed until the stage of

---

**Fig. 6.** Model for the TM topology of E2p7 (a) and p7 (b). The filled circle denotes the presence of the antibody recognition tag position at the C terminus of E2p7 in this study (a) or p7 in a previous study (b) (Carrere-Kremer et al., 2002).
virion release, hence preventing E from undergoing acid-catalysed conformational changes during transport of immature virus through the acidic environment of the intermediate compartment (Chambers et al., 1990; Guirakhoo et al., 1991, 1992; Heinz & Allison, 2000). Inhibition of this cleavage does not prevent infectious virus from being produced, but is thought to be necessary for the production of highly infectious virus particles (Chambers et al., 1990).

ACKNOWLEDGEMENTS

We thank Professor Duncan McGeoch for critical reading of this manuscript, and Drs Jens Bukh and Jean Dubuisson for provision of the HCV strain H77c cDNA construct and mAb H53, respectively. This work was supported by a BBSRC-Pfizer CASE studentship and the Medical Research Council, UK.

REFERENCES


