Characterization of the genome and structural proteins of hepatitis C virus resolved from infected human liver

Søren U. Nielsen,1 Margaret F. Bassendine,1 Alastair D. Burt,1,2 Debra J. Bevitt1† and Geoffrey L. Toms1

School of Clinical Medical Sciences1 and School of Clinical and Laboratory Sciences2, Medical School, Framlington Place, Newcastle upon Tyne NE2 4HH, UK

In the absence of satisfactory cell culture systems for hepatitis C virus (HCV), virtually all that is known about the proteins of the virus has been learned by the study of recombinant proteins. Characterization of virus proteins from patients with HCV has been retarded by the low virus titre in blood and limited availability of infected tissue. Here, the authors have identified a primary infection in a liver transplanted into an immunodeficient patient with chronic HCV. The patient required re-transplant and the infected liver, removed 6 weeks after the initial transplant, had a very high titre of HCV, 5 × 10^9 International Units (IU) per gram of liver. The density distribution of HCV in iodixanol gradients showed a peak at 1.04 g ml^-1 with 73 % of virus below 1.08 g ml^-1. Full-length HCV RNA was detected by Northern blotting and the ratio between positive- and negative-strand HCV RNA was determined as 60. HCV was partially purified by precipitation with heparin/Mn^2+ and a single species of each of the three structural proteins, core, E1 and E2, was detected by Western blotting. The molecular mass of core was 20 kDa, which corresponds to the mature form from recombinant sources. The molecular mass of glycoprotein E1 was 31 kDa before and 21 kDa after deglycosylation with PNGase F or endoglycosidase H. Glycoprotein E2 was 62 kDa before and 36 kDa after deglycosylation, but E2-P7 was not detected. This was in contrast to recombinant sources of E2 which contain E2-P7.

INTRODUCTION

Infection with hepatitis C virus (HCV) occurs worldwide and afflicts approximately 170 million people (Anonymous, 1997). Chronic infection, which occurs in 80 % of cases, can result in chronic active hepatitis, cirrhosis and hepatocellular carcinoma and, less commonly, extrahepatic autoimmune or immune complex diseases (reviewed by Major et al., 2001). Gene fragments from the virus were initially identified by screening a lambda phage expression library of cDNA from the nucleic acid extracted from infected chimpanzee serum against serum from a non-A non-B hepatitis patient (Choo et al., 1989). Subsequently, the whole virus genome of approximately 9500 nt was rescued, sequenced and shown to exhibit a similar genome organization to the flaviviruses and pestiviruses (Choo et al., 1991). This insight, the in vitro transcription of the genome and extensive studies of viral protein expression in eukaryotic expression systems indicate that the genome consists of a single open reading frame encoding a 3000 aa polyprotein flanked by highly conserved 5' and 3' untranslated regions. The putative structural proteins of the virus, core protein and two membrane glycoproteins, E1 and E2, lie at the amino-terminal end of the polyprotein and are released co-translationally by host cell signal peptidase. The non-structural proteins of the virus, NS2 to NS5B, are located at the C-terminal end of the polyprotein and are released by virus proteases (reviewed by Major et al., 2001; Bartenschlager & Lohmann, 2000). In between the structural and non-structural regions of the polyprotein lies a small polypeptide, P7, which is cleaved late from NS2 and subsequently from E2 by host signal peptidase. It has not been possible, from studies in recombinant expression systems, to determine whether P7 or the E2-P7, E2-P7-NS2 precursors are present in the virion (reviewed by Dubuisson, 2000). Recombinant systems are also unable to give information concerning any modifications of the virion components which may occur at virion assembly, maturation and release.

Resolution of these points requires biochemical analysis of virion structural proteins derived from mature virion particles. Virus isolates from both human and chimpanzee blood have been shown to replicate in human cells of hepatic
and lymphoid origin but replication is indolent and yields of virus transcription products and infectious virus have been too low to permit biochemical analysis (Kato & Shimotohno, 2000). Full-length infectious clones of the HCV genome, modified to enhance RNA replication, have recently been shown to express virus proteins in cell culture but no virus particles were formed (Pietschmann et al., 2002). The only proven animal model for HCV replication is the chimpanzee but, as in humans, virus yields in plasma and liver of infected chimpanzees are generally low (Bradley, 2000).

In the absence of model infection systems with sufficient virus yields to permit biochemical characterization of the infectious virion, attempts have been made to recover virus and virus proteins from clinical material. Infected blood is the most readily available source of virus although titres of virus genome in human blood rarely reach $10^6$ International Units (IU) ml$^{-1}$ (Terrault et al., 1997). Hepatitis core antigen can be detected in serum of most chronically infected patients by ELISA (Aoyagi et al., 1999) and mature core protein has been detected by immunoblotting (Yasui et al., 1998), but envelope glycoproteins have not been detected in serum (André et al., 2002). The primary site of virus replication appears to be the liver and yields of $0.5 \times 10^9$–$1.5 \times 10^9$ IU have been recorded per gram of infected liver (Dailey et al., 1998; Terrault et al., 1997).

Levels of protein expression in the livers of chronically infected patients, however, are generally low and immunohistochemical detection of virus antigens in infected liver has proved difficult (Gowans, 2000). In a single study (Nakamoto et al., 1994) a 44 kDa glycoprotein was demonstrated in lysates of infected liver by immunoblotting with antibodies to the HCV E2 glycoprotein. Although this does not correspond to the molecular mass of the fully glycosylated form of E2 derived from mammalian expression systems, on deglycosylation the band was reduced to 38 kDa, which corresponds approximately to the predicted molecular mass of the unglycosylated E2 polypeptide.

Where chronically infected patients with end-stage liver disease receive a liver transplant the new liver is frequently reinfected whilst the host is heavily immunosuppressed and virus titres in blood may reach high levels (Papatheodoridis et al., 2002). The primary site of virus replication appears to be the liver and yields of $0.5 \times 10^9$–$1.5 \times 10^9$ IU have been recorded per gram of infected liver (Dailey et al., 1998; Terrault et al., 1997).

Levels of protein expression in the livers of chronically infected patients, however, are generally low and immunohistochemical detection of virus antigens in infected liver has proved difficult (Gowans, 2000). In a single study (Nakamoto et al., 1994) a 44 kDa glycoprotein was demonstrated in lysates of infected liver by immunoblotting with antibodies to the HCV E2 glycoprotein. Although this does not correspond to the molecular mass of the fully glycosylated form of E2 derived from mammalian expression systems, on deglycosylation the band was reduced to 38 kDa, which corresponds approximately to the predicted molecular mass of the unglycosylated E2 polypeptide.

Where chronically infected patients with end-stage liver disease receive a liver transplant the new liver is frequently reinfected whilst the host is heavily immunosuppressed and virus titres in blood may reach high levels (Papatheodoridis et al., 1999). We have recently studied an acutely infected transplant liver removed at re-transplant from a patient with chronic HCV infection complicated by common variable immunodeficiency (CVID). The virus titre in liver from this patient was exceptionally high and has made possible the detection of virus structural proteins by Western blotting and analysis of glycosylation.

**METHODS**

**Liver samples and maceration.** Patient S6 had CVID and was treated with intravenous immunoglobulin from December 1993 to February 1994. This patient was one of 30 in the UK who developed HCV infection from a batch of Gammanagad intravenous immunoglobulin from Baxter Healthcare. All infections were linked to batch 93F21AB11B (Healey et al., 1996) and were of genotype 1a (Christie et al., 1997). HCV RNA was first detected in serum from patient S6 in March 1994 and the titre was $1 \times 10^7$ IU ml$^{-1}$. The virus infection progressed over a period of 18 months and the native liver was removed at orthotopic liver transplant, performed because of decompensation of cirrhosis in September 1995. The native cirrhotic liver, S6a, was retained for research purposes and had a titre of $5 \times 10^6$ IU per gram of liver. However, the liver transplant failed due to a combination of rejection, a leak from the biliary anastomosis, subtotal occlusion of the hepatic artery, and re-infection by HCV. The transplant liver, S6b, was removed at re-transplant in November 1995 and had a titre of $5 \times 10^9$ IU per gram of liver. Another liver explant, TO94, was obtained from a patient with end-stage liver disease unrelated to HCV. Liver explant samples were used in research with the informed consent of patients themselves and with approval from the Research Ethics Committee of this University.

Liver tissue was stored at $-80\,\text{C}$ and blocks weighing 400–500 mg were cut while frozen. To prepare macerate, one liver block was immersed in 10 ml ice-cold homogenization buffer [50 mM Tris/HCl, pH 7.4, 150 mM NaCl, 0.2% BSA, 3 µl protease inhibitor cocktail (Sigma) ml$^{-1}$]. Homogenization was achieved by 60 strokes in a tight-fitting Dounce homogenizer and was followed by centrifugation at 8000 g (5 min, 4°C). The supernatant was passed through a 0.45 µm filter and snap-frozen in liquid nitrogen.

**Iodixanol density gradients.** Iodixanol (Axis-Shield) is isotonic and isosmotic with blood. Self-forming iodixanol gradients were prepared as described by Graham et al. (1996). A Ti-50 tube (Beckman) received 0.2 ml 0.5 M Tris/HCl, pH 8.0, 0.1 ml 0.5 M EDTA, 4.2 ml iodixanol, 3.3 ml 0.25 M sucrose, 2.2 ml liver macerate and was spun at 50,000 r.p.m. for 24 h at 4°C. Gradients were harvested from the top into 18 fractions.

**Sequencing and real-time PCR.** RNA from clarified liver macerate or from fractions of iodixanol density gradients was extracted using the QiAamp viral RNA kit (Qiagen). Such RNA served as template for reverse transcription and PCR as described by Kolykhlov et al. (1998) using primers within the structural region of the HCV 1a genome (Table 1). PCR products were subcloned into pBluescript SK+ (Stratagene) for sequencing and aligned using CLUSTAL_X (EMBL). RNA from blocks of liver tissue was extracted using an Ambion Midi kit. Real-time quantitative PCR was performed on a TagMan 7000 (Applied Biosystems) with the default PCR conditions. The primer set for positive-strand HCV RNA (5F, 5R; Table 1) was complementary to a region between nt 120 and 290 in the 5'-NTR of HCV genotype 1a and was adapted from Mercier et al. (1999). The probe was labelled with FAM and TAMRA (SN1, Table 1). This assay was calibrated against WHO International Standard for HCV no. 96/790. Negative-strand HCV RNA was detected by real-time PCR using a tagged primer for reverse transcription (TagR, Table 1). cDNA was digested with RNase A and H (Lin et al., 2002) to remove RNA and primer plus reverse transcriptase removed using a QiAamp kit (Qiagen). An internal control, comprising RNA from position 120–290 of the HCV 1a genome, was prepared by PCR and in vitro transcription. Control RNA had the annealing site for probe SN1 replaced by a nonsense sequence with similar $T_m$, 5'-gtcgtgtagacaagattcgtgat-3', and was detected by another probe, SN2 (Table 1).

**Northern blotting.** RNA isolated from liver macerate was denatured by heating to 56°C for 60 min in glyoxal/dimethyl sulfoxide buffer. The sample was applied to a 0.8% agarose gel followed by Northern blotting as described by Sambrook et al. (1989). A biotinylated RNA probe spanning nucleotide position 278–1375 in the structural region of the HCV 1a genome was synthesized by *in vitro* transcription (TagR, Table 1). cDNA was digested with RNase A and H (Lin et al., 2002) to remove RNA and primer plus reverse transcriptase removed using a QiAamp kit (Qiagen). An internal control, comprising RNA from position 120–290 of the HCV 1a genome, was prepared by PCR and in vitro transcription. Control RNA had the annealing site for probe SN1 replaced by a nonsense sequence with similar $T_m$, 5'-gtcgtgtagacaagattcgtgat-3', and was detected by another probe, SN2 (Table 1).
transcription. This probe was hybridized with Northern blots at a concentration of 60 pg ml\(^{-1}\) using the Brightstar Biodetect kit (Ambion).

**SDS-PAGE and Western blotting.** PAGE was performed with 3–18 % gradient gels on a Bio-Rad Protean II system. Western blots were made onto PVDF as described by Simpson & Reid (1998). Monoclonal anti-HCV core antibody, clone P126 (Biogenesis) was used at 1 μg ml\(^{-1}\). The epitope is between aa 1 and 120 of the polyprotein. Monoclonal anti-HCV E1 antibody clone 11B7D8 was kindly provided by Dr E. Depla from Innogenetics and used at 0.5 μg ml\(^{-1}\). The epitope is between aa 212 and 224 of the polyprotein. Monoclonal anti-HCV E2 antibodies, clones AP33 and ALP98 were kindly provided by Dr A. Patel (Glasgow, UK) and used at a dilution of 1:2000. The epitopes for AP33 and ALP98 were kindly provided by Dr A. Patel (Glasgow, UK) and used at 1 μg ml\(^{-1}\). Western blots were developed using ECL-Plus (Amersham). Densitometric quantification with an Alphalmager model 2200 (Flowgen) was used to correlate the staining achieved with anti-core antibody with the amount of core protein. Calibration was done using samples containing from 0.5 pg to 33 ng purified core protein (Biogenesis).

**Precipitation methods.** Chemical precipitation of LDL/VLDL was developed from protocols by Mackness & Durrington (1992) and Burstein et al. (1970). All procedures were performed at 4 °C. To 840 μl of clarified liver macerate was added 120 μl of 1 M Tris/HC1, pH 7–4. The solution was centrifuged at 10,000 g for 5 min. The supernatant was mixed with an equal volume of 110 mM MnCl\(_2\), 154 mM NaCl, 2 mg heparin (187 UCS units mg\(^{-1}\); Sigma) ml\(^{-1}\) in water. After 1 h incubation, a small pellet was recovered by centrifugation (3000 g, 15 min) and washed twice in PBS. Resolubilization of this pellet required a denaturing solution containing either SDS or guanidinium thiocyanate.

**Expression of recombinant HCV structural proteins.** A cDNA containing the entire HCV 1a structural region was kindly provided by Dr M. Houghton (Chiron Diagnostics). This clone was constructed from the prototype HCV cDNA (Choo et al., 1991) and spans aa 1–906. The cDNA was subcloned into a recombinant vaccinia virus (vvHCV3240) and used to express HCV structural proteins (Mackett, 1995). Confluent HeLa cells were co-infected in T-225 flasks with recombinant vaccinia viruses vVT7-3 and vvHCV3240 at an m.o.i. of 3 for each. Cells were harvested after 22 h infection and the extent of infection was checked by indirect immunofluorescence using anti-HCV core antibody.

**Deglycosylation with PNGase F and endoglycosidase H.** Recombinant glycoproteins from vaccinia virus expression were purified on *Galanthus nivalis* lectin (GNL)/agarose (Sigma) as described by Ralston et al. (1993). Glycoproteins were eluted from beads by boiling in denaturation buffer (0.5 % SDS, 1 % β-mercaptoethanol) (NEB). Native HCV protein from a heparin/Mn\(^2\)+ precipitate of liver macerate was resuspended in Tris/EDTA buffer. NEB was added and the solution was boiled for 10 min. Prior to deglycosylation, samples were dialysed in Slide-A-Lyser mini units (Pierce) at 4 °C. The dialysis buffer contained 50 mM sodium phosphate, pH 7–4, 0.2 % SDS, 0.2 % β-mercaptoethanol, 1 mM EDTA for PNGase F digestion and 50 mM sodium citrate, pH 5–5, 0.2 % SDS, 0.2 % β-mercaptoethanol, 1 mM EDTA for endoglycosidase H digestion. Incubation with PNGase F at 50 units (μg substrate protein)\(^{-1}\) or endoglycosidase H at 30 units μg\(^{-1}\) was for 16 h at 37 °C.

**RESULTS**

**Characterization of the HCV RNA genome in liver macerate**

A sample of RNA from liver S6b containing 5 × 10\(^7\) IU of HCV RNA was analysed by Northern blotting. The RNA

<table>
<thead>
<tr>
<th>Name</th>
<th>Purpose</th>
<th>Sequence*</th>
<th>Segment of HCV amplified†</th>
</tr>
</thead>
<tbody>
<tr>
<td>1R</td>
<td>Reverse</td>
<td>CGCGGATCCCCCCCTGTGAGAATCTAGTCTCCAC</td>
<td>42–331</td>
</tr>
<tr>
<td>2R</td>
<td>Reverse</td>
<td>CGCGAATTCCCTTGGTGATCTGGCCTGATAGGTG</td>
<td>278–1375</td>
</tr>
<tr>
<td>3R</td>
<td>Reverse</td>
<td>CGCGAATTCAATGATGAACTGGTCCCCTACGGCAG</td>
<td>1305–2611</td>
</tr>
<tr>
<td>4R</td>
<td>Reverse</td>
<td>CGCGAATTCCGTCCAGCCCCGAAACCCACGTCA</td>
<td>148–2786</td>
</tr>
<tr>
<td>5R</td>
<td>Reverse</td>
<td>CCCCCCTCCGGGGAGAGCCAT</td>
<td>120–290</td>
</tr>
<tr>
<td>5RN</td>
<td>Reverse</td>
<td>CGTCGGCAATGTCGTAATTC</td>
<td>120–290‡</td>
</tr>
<tr>
<td>1F</td>
<td>Forward</td>
<td>GCCGAGGTTCCTACGAGACCTCCCGGGGCAC</td>
<td>42–331</td>
</tr>
<tr>
<td>2F</td>
<td>Forward</td>
<td>GCCGAGGTTCATGATGGCTTGGGAGATCCGGAG</td>
<td>278–1375</td>
</tr>
<tr>
<td>3F</td>
<td>Forward</td>
<td>GCCGAGGTTCATGATGGCTTGGGAGATCCGGAG</td>
<td>1305–2611</td>
</tr>
<tr>
<td>4F</td>
<td>Forward</td>
<td>GCCGAGGTTCCTACGAGACCTCCCGGGGCAC</td>
<td>1482–2786</td>
</tr>
<tr>
<td>5F</td>
<td>Forward</td>
<td>ACCACAGGGCTCTGCGGACCCAC</td>
<td>120–290</td>
</tr>
<tr>
<td>TagF</td>
<td>RT</td>
<td>GCGTCGGCAATGATGTCGTCCTACGAGCCCCTCCCGGGGAGAGCCAT</td>
<td>120–290‡</td>
</tr>
<tr>
<td>SN1</td>
<td>Probe</td>
<td>FAM–ATTCCGGTGTACTCACGCTGCTTGGCATGC</td>
<td>174–204</td>
</tr>
<tr>
<td>SN2</td>
<td>Probe</td>
<td>VIC–ATACACAGGGATTTGCTGCAGCACAGC–TAMRA</td>
<td>174–204</td>
</tr>
</tbody>
</table>

*The restriction enzyme sites for EcoRl, BamHI and HindIII are underlined.
†The start position of each primer in the HCV RNA 1a genome is underlined.
‡Primer 5RN was utilized for negative-strand amplification and is identical to the double-underlined tag sequence in primer TagF.
probe, which was complementary to a region of the HCV genome encoding core and part of E1, detected a single band of HCV RNA (Fig. 1a). This band co-migrated with a single-stranded RNA marker at 9.5 kb, which corresponds to the expected length of the HCV RNA genome (Choo et al., 1989). The nucleotide sequence of 2744 bp of cDNA from HCV RNA derived from liver S6b was determined. The region sequenced began at position 42 in the 5′ non-translated region and spans the complete coding region for the structural proteins. This sequence was compared with HCV isolated from a Swedish patient who contracted the virus from another batch of Gammagard intravenous immunoglobulin, 93D19AB11A (Widell et al., 1997), and with the prototype HCV 1a sequence (H77; Choo et al., 1991). Differences were observed throughout the sequence but particularly in the hypervariable region 1 (HVR-1) with 10 aa differences between S6b and H77 and 5 aa differences between S6b and 93D19AB11A (Fig. 1b). The S6b sequence contained the same number of potential glycosylation sites in E1 as strain H77, but one site less (10 compared with 11) in E2. Comparison between HCV sequence from liver S6a and liver S6b revealed two conservative (Ile34→Val and Lys115→Arg) and two non-conservative (Ser190→Trp and Thr510→Ala) changes. The HVR-1 region was found to be identical between HCV isolated from these two livers (Fig. 1b); thus no mutations had occurred in HVR-1 between the first and the second liver transplant. This is consistent with observations by Gaud et al. (2003) in immunodeficient patients.

The titre of positive- and negative-strand HCV RNA was determined for liver S6a and S6b (Table 2). The presence of negative-strand HCV RNA is a sign of active virus replication (Lin et al., 2002) and a ratio between positive- and negative-strand of 60 is consistent with values determined by others at between 10 and 100 (Sakamoto et al., 1994).

**Density distribution of HCV in liver macerate on iodixanol gradients**

Prince et al. (1996), Thomssen et al. (1992) and Thomssen et al. (1993) have described an association between HCV and LDL/VLDL in the serum of patients infected with HCV. Although variations exist in the degree of association with LDL/VLDL between patients, in some sera all HCV RNA was associated with LDL/VLDL and the virus had a low density (<1.06 g ml⁻¹). HCV in macerate from transplant liver S6b was analysed by buoyant density-gradient centrifugation in isotonic iodixanol gradients (Fig. 2). The gradient was evaluated by quantitative real-time PCR for...
HCV RNA and with an assay for protein. The peak fraction of HCV RNA was detected with a density of $1.04 \text{ g ml}^{-1}$ and the top four fractions contained 73% of HCV RNA. An internal control was detected with a $C_T$ value of $30 \pm 1$ in all fractions, showing the absence of PCR inhibitors. Protein content was lowest in the four top fractions, which contained only 2% of overall protein. ApoB-100, the major protein component of LDL/VLDL, was detected in fractions 14–18 (data not shown).

Precipitation of LDL/VLDL and HCV from liver macerate

Two LDL/VLDL precipitation methods described by Thomssen et al. (1992) for use with serum were applied to liver macerates. Immunoprecipitation with a polyclonal anti-human apoB-100 precipitated 50% of HCV RNA from S6b liver macerate. In comparison, immunoprecipitation with anti-apoB-100 from serum of the same patient precipitated 98% of HCV RNA (S. U. Nielsen & G. L. Toms, unpublished results). Precipitation of HCV RNA from liver S6b was more effective with heparin/Mn$^{2+}$ and 99.7% of the HCV RNA was precipitated. Both pellets and supernatants from immunoprecipitation with anti-apoB-100 antibody and from precipitation with heparin/Mn$^{2+}$ were analysed by SDS-PAGE (Fig. 3). Precipitation of apoB-100 was complete with both methods but heparin/Mn$^{2+}$ precipitated other proteins as well. The protein concentration in macerate from liver S6b was determined to be $5800 \text{ mg l}^{-1}$ and 500 mg (8%) was precipitated by heparin/Mn$^{2+}$.

Detection of three structural proteins and one non-structural protein of HCV in liver macerate

Western blots of heparin/Mn$^{2+}$ precipitate from HCV-infected livers S6a, S6b and uninfected control liver TO94 were probed with monoclonal antibodies to viral structural proteins, core, E1 and E2, and non-structural protein, NS3 (Fig. 4). Immunostaining with these antibodies was only observed in liver S6b with the highest titre of HCV; no specific staining was observed in the sample from liver S6a with the lower titre of HCV, or the HCV-negative liver TO94. The anti-core antibody, clone W126, detected a single protein with a molecular mass of 20 kDa in liver S6b. This band co-migrated in polyacrylamide gels with the recombinant core protein from vaccinia virus expression using plasmid vvHCV$^{3240}$ (Fig. 4). The amount of HCV core protein was evaluated using semi-quantitative densitometry.
of Western blots. The content of core was estimated as 350 ng per gram of liver S6b. Immunostaining with anti-glycoprotein E1 antibody, clone 11B7D8, detected a single band with a molecular mass of 31 kDa in liver S6b. In the vaccinia virus expression system a ladder of glycosylation variants between 21 and 31 kDa was observed and E1 from liver S6b co-migrated with the fully glycosylated recombinant E1 protein. Immunostaining with anti-glycoprotein E2 antibody, clone AP33, showed a single band with a molecular mass of 62 kDa in liver S6b. E2 with molecular mass 44 kDa (Nakamoto et al., 1994) was not observed. The molecular mass was approximately 6 kDa below recombinant E2 from vaccinia virus expression using vvHCV3240. The difference could be due to differential processing of the polyprotein. Grakoui et al. (1993) have shown that E2 from vaccinia virus expression was a mixture of incompletely processed E2-P7, which was cleaved at position 809 of the polyprotein and completely processed E2 which was cleaved at position 746 of the polyprotein. Deglycosylation was performed to clarify this issue (see below). The antibody against non-structural protein NS3 detected a protein with a molecular mass of 68 kDa, which corresponds to the expected size of full-length NS3. A second protein was detected, with a molecular mass of 44 kDa, which is reminiscent of a protease cleavage product described previously in recombinant NS3 (Shoji et al., 1999).

**Deglycosylation of HCV glycoproteins E1 and E2**

Native HCV proteins from liver S6b and recombinant HCV proteins from vaccinia virus expression were treated with PNGase F or endoglycosidase H (Fig. 5). PNGase F cleaves both complex and hybrid type glycans but endoglycosidase H cleaves only hybrid type glycans. After treatment of proteins from liver with PNGase F a single band of E1, at 21 kDa, was detected. A comparison was made with recombinant HCV glycoproteins purified by GNL/agarose. Such samples were enriched with HCV glycoproteins but also seemed to contain aggregated or unprocessed protein variants with higher molecular mass. PNGase F treatment curtailed recombinant glycoprotein E1 to a single band at 21 kDa. This analysis showed the protein backbone of native E1 was of similar length to recombinant E1 protein.

The staining intensity achieved with E2 antibody, clone AP33, was reduced after deglycosylation. The epitope for AP33 contains a glycosylation site at position 418 and it is conceivable that deglycosylation alters the binding affinity of this antibody. Therefore, deglycosylated E2 was visualized with ALP98, which has an epitope that does not include a glycosylation site. Deglycosylated recombinant E2 protein showed a major band at 40 kDa and a less intense band at 36 kDa (Fig. 5). Lin et al. (1994) have reported variants of E2 with identical molecular mass from vaccinia virus expression in BHK-21 and HepG2 cells. These authors identified the 40 kDa species as incompletely processed E2-P7 and the 36 kDa species as completely processed E2. The deglycosylated E2 species detected in liver was 36 kDa, corresponding to the completely processed variant. Similar deglycosylation patterns were observed using endoglycosidase H (data not shown).

**Absence of disulphide bond linked dimers or heterodimers in HCV structural proteins**

PAGE was performed under non-reducing conditions to assess the tertiary structure of HCV structural proteins. All three structural proteins have multiple cysteine residues and intramolecular disulphide bonds are thought to be present in E1 and E2 (Dubuisson, 2000). The core protein from liver macerate S6b migrated as a monomer at 20 kDa (Fig. 6). The E1 protein migrated at ~28 kDa under non-reducing conditions, but a protein at 40 kDa was also observed. However, this band was also detected in HCV-negative liver TO94 and therefore deduced to be unrelated to E1. A difference of ~3 kDa was observed between the apparent molecular mass of reduced and non-reduced E1. This is likely due to intramolecular disulphide bonds which render
the protein more compact and faster migrating in acrylamide gels (Dubuisson, 2000). The E2 protein migrated as a broad band at 60–62 kDa under non-reducing conditions, similar to or slightly below the reduced E2 species. It was deduced that neither core, E1 nor E2 engage in disulphide bond linked homodimers or heterodimers in infected liver.

DISCUSSION

Patient S6, suffering CVID, was most probably infected with HCV from the 93F21AB11B batch of Gammagard intravenous immunoglobulin, now known to have been contaminated with a 1a strain of HCV. In this case, as in some others where HCV infection complicated CVID (Bjøro et al., 1999), the disease progressed unusually rapidly and within 2 years resulted in decompensated liver failure requiring liver transplant. The transplant was unsuccessful and was replaced after 6 weeks, a sample from the explant being stored at −80 °C. During that time the transplant liver, S6b, had become extensively infected with hepatitis C with a titre of $5 \times 10^9$ IU per gram. The human liver contains approximately $250 \times 10^9$ hepatocytes (Miyai, 1991). This number allows us to calculate the mean HCV RNA titre per cell to 25 IU for the positive-strand and 0.4 IU for the negative-strand. Schlaak et al. (1997) have determined HCV viral load in a highly infected human liver to 10 copies of HCV RNA per cell. One IU of the 96/790 WHO standard contains between 3 and 5 copies of HCV RNA (Puig et al., 2002; Sarrazin et al., 2001). Thus, liver S6b is extensively infected with HCV and this observation correlates with immunohistochemistry experiments, in which most cells in sections were stained by the NS3 antibody (data not shown). HCV reinfection in liver transplant patients with CVID has been described by Gow & Mutimer (2002) (genotype 1a), Bjøro et al. (1999) (genotype 3a) and Smith et al. (1995). In all cases the reinfection was aggressive and rapid, with liver failure after 5–23 months. Hepatic HCV RNA titres were not reported but a serum titre of $1 \times 10^7$ copies ml$^{-1}$ (Bjøro et al., 1999) may reflect a much higher titre in liver (Di Bisceglie et al., 1994; present work).

To characterize HCV structural proteins we took advantage of progress in the identification and physico-chemical characterization of HCV virions derived from serum of patient S6. All virus from this serum had a density below 1.13 g ml$^{-1}$ in iodixanol gradients and anti-apoB-100 polyclonal antibody precipitated virus with density below 1.10 g ml$^{-1}$. This showed that HCV of low density was associated with LDL/VLDL. Such complexes are assumed to represent mature virions (Bradley, 2000; Hijikata et al., 1993) and have been shown to be infectious for chimpanzees. HCV from liver S6b had a density profile similar to that seen in serum and 99.7 % was precipitated with heparin/Mn$^{2+}$. This prompted us to use heparin/Mn$^{2+}$ in the purification of HCV structural proteins. The method precipitated 99.7 % of HCV RNA from liver and we have shown that HCV proteins expressed from recombinant vaccinia virus were precipitated, although this system does not contain virus particles. Structural proteins of HCV are

![Fig. 4. Comparison of molecular masses of recombinant HCV proteins core, E1 and E2 with the structural proteins from HCV-infected human liver. The protein sample from vaccinia virus expression was whole-cell lysate from HeLa cells infected with strain vHCV3240. The protein samples from liver macerate were heparin/Mn$^{2+}$ precipitates from liver S6a or liver S6b. Lanes 1, 4 and 7 received 8 μg protein from vaccinia virus expression. Lanes 2, 5, 8 and 10 received 30 μg protein from liver macerate S6a with $5 \times 10^9$ IU of HCV RNA. Lanes 3, 6, 9 and 11 received 30 μg protein from liver macerate S6b with $1 \times 10^7$ IU of HCV RNA. Sections of the Western blot were immunostained with monoclonal antibodies against core protein, glycoprotein E1, glycoprotein E2 (clone AP33) and non-structural protein NS3. M, molecular mass markers.](http://vir.sgmjournals.org)
associated with microsomal membranes in the vaccinia virus system (Dubuisson et al., 1994); thus heparin/Mn\(^{2+}\) precipitate from liver may contain HCV proteins originating from mature virus particles as well as membrane-associated non-assembled HCV proteins.

Western blotting showed a single band of core protein in the heparin/Mn\(^{2+}\) precipitate. Full-length core protein is 191 aa, but a shorter species, the mature form, has been demonstrated previously when the core gene was expressed in vitro (Huüssy et al., 1996; Liu et al., 1997). Expression in the baculovirus system gave two forms of core with molecular masses of 20 and 22 kDa (Watson, 1996). In the vaccinia virus expression system both forms were detected after 3 h infection, but only the mature form was present after 24 h infection (Yasui et al., 1998). Recombinant core protein is cleaved at position 191 by signal peptidase (Huüssy et al., 1996) and at position 179 or 182 by signal peptide peptidase (McLauchlan et al., 2002) but it is not known if this occurs in vivo. Core protein from liver S6b co-migrated in SDS-polyacrylamide gels with the single band of core protein observed in HeLa cells infected with vvHCV\(_{3240}\) for 22 h.

This suggests that the core protein with a molecular mass of 20 kDa in human liver corresponds to the mature form. Yasui et al. (1998) have published a similar result with HCV from human serum, detecting only the smaller mature form of core.

Fully glycosylated E1 protein expressed in the vaccinia virus system had a molecular mass of 31 kDa and comprised aa 174–383 of the polyprotein (Fournillier-Jacob et al., 1996). Partial deglycosylation with endoglycosidase H resulted in three partially deglycosylated products and the aglycosyl form, indicating that only four of the possible five N-linked glycosylation sites in the polypeptide are utilized. It was subsequently established that the site at residue 325 is blocked by an adjacent proline molecule (Meunier et al., 1999). In Western blots of E1 expressed from vvHCV\(_{3240}\) here, four heavy bands were apparent corresponding to the fully glycosylated form and three partially deglycosylated products. On deglycosylation with PNGase F a single protein with molecular mass 21 kDa was resolved, confirming that the multiple E1 bands are glycosylation variants, as has also been reported by Dumonceaux et al. (2003). In infected liver, a single E1 band was resolved which co-migrated with the vaccinia virus system 31 kDa band before and with the 21 kDa band after deglycosylation. The splicing variant of E1 with molecular mass...
20 kDa, described by Dumonceaux et al. (2003) in a plasmid expression system, was not observed in human liver. Our results suggest that the structural form of the E1 protein comprises aa 174–383 and is glycosylated at four sites, but this conclusion requires confirmation by direct determination of the terminal amino acids.

Grakoui et al. (1993) demonstrated that, following expression of a vaccinia virus recombinant encoding the first 1488 aa of the HCV polyprotein, an anti-E2 serum precipitated E2 with molecular mass 70 kDa and E2-P7-Ns2 with molecular mass 88 kDa. Digestion with PNGase F reduced the E2 band to two bands of 41 and 36 kDa. These bands were subsequently demonstrated to be aglycosyl E2 (36 kDa) and aglycosyl E2-P7 (41 kDa), the glycosylated forms of which run together in a diffuse band at approximately 70 kDa (Lin et al., 1994). Host signalases mediate the cleavage between E2-P7, but the efficiency varies with the virus strain. Western blots of the expression products of vvHCV3240 presented here visualize a diffuse band of approximately 68 kDa. On deglycosylation with PNGase F, two proteins with molecular masses of 36 and 40 kDa appeared, corresponding to the deglycosylated forms of E2 and E2-P7 described previously (Lin et al., 1994). A single anti-E2 antibody staining band was resolved from heparin/Mn²⁺ precipitates of infected liver of apparently lower mean size than the E2 band resolved from the crude lysate of recombinant vaccinia virus infected cells. On deglycosylation, this was replaced with a single band of 36 kDa which co-migrated with the E2 recombinant vaccinia virus band. This indicates that E2 but not the E2-P7 protein is present in the particles isolated from liver. Our results are similar to those of Koike et al. (1995), who detected E2 with molecular mass 58 kDa in transgenic mice, reduced to 35 kDa after deglycosylation with PNGase F. It remains unclear whether the P7 protein is present in the virion, as suitable antibodies for blotting P7 are currently unavailable.

Two types of E1–E2 complex have been demonstrated in cell lysates. They were either disulphide-linked aggregates, which are probably misfolded complexes (Choukhi et al., 1999), or non-covalently linked heterodimers, which may be the functional subunit of the HCV virus particle (reviewed by Dubuisson, 2000). In line with this, all three structural proteins from infected human liver S6b ran as monomers in SDS-polyacrylamide gels under non-reducing conditions.

ACKNOWLEDGEMENTS

This work was supported by The Wellcome Trust (Grant no. 059376). The research was initiated by Dr W. Pumeechochok. The technical assistance of Mrs F. Fenwick is appreciated.

REFERENCES


