

Analysis of the glycosylation sites of hepatitis C virus (HCV) glycoprotein E1 and the influence of E1 glycans on the formation of the HCV glycoprotein complex

Jean-Christophe Meunier,¹ Anne Fournillier,² Amélie Choukhi,¹ Annie Cahour,³ Laurence Cocquerel,¹ Jean Dubuisson¹ and Czeslaw Wychowski¹

¹ CNRS-UMR 319, IBL/Institut Pasteur de Lille, 59021 Lille CEDEX, France

² Unité INSERM U271, 151 Cours Albert Thomas, 69003 Lyon, France

³ CERVI (Virologie), Pitié-Salpêtrière, 83 Boulevard de l'Hôpital, 75651 Paris CEDEX 13, France

The hepatitis C virus (HCV) genome encodes two membrane-associated envelope glycoproteins (E1 and E2), which are released from the viral polyprotein precursor by host signal peptidase cleavages. These glycoproteins interact to form a noncovalent heterodimeric complex, which is retained in the endoplasmic reticulum. HCV glycoproteins, E1 and E2, are heavily modified by *N*-linked glycosylation. A recent study has revealed that upon partial deglycosylation with endoglycosidase H only four of the five potential glycosylation sites of HCV glycoprotein E1 are utilized. In this work, the unused glycosylation site on the E1 glycoprotein was identified and the influence of *N*-linked glycosylation on the formation of the HCV glycoprotein complex was studied by expressing a panel of E1 glycosylation mutants in HepG2 cells. Each of the five potential *N*-linked glycosylation sites, located at amino acid positions 196, 209, 234, 305 and 325, respectively, on the HCV polyprotein, was mutated separately as well as in combination with the other sites. Expression of the mutated E1 proteins in HepG2 cells indicated that the fifth glycosylation site is not used for the addition of *N*-linked oligosaccharides and the Pro immediately following the sequon (Asn-Trp-Ser) precludes core glycosylation. The effect of each mutation on the formation of noncovalent E1E2 complexes was also analysed. As determined with the use of a conformation-sensitive monoclonal antibody, mutations at positions N2 and N3 had no, or only minor, effects on the assembly of the E1E2 complex, whereas a mutation at position N1 and predominantly at position N4 dramatically reduced the efficiency of the formation of noncovalent E1E2 complexes.

Introduction

Hepatitis C virus (HCV) is the major aetiological agent of human posttransfusion and community-acquired non-A, non-B hepatitis (Alter *et al.*, 1990; Choo *et al.*, 1989; Kuo *et al.*, 1989), infecting probably 1% of the population worldwide (Houghton, 1996; Purcell, 1994). Infected individuals have a high risk of developing chronic hepatitis, liver cirrhosis and hepatocellular carcinoma, making HCV a major cause of morbidity and mortality worldwide (Alter *et al.*, 1992; Saito *et al.*, 1990; Zhang *et al.*, 1990).

Comparative analyses of the genomes of several HCV strains that have been molecularly cloned and sequenced indicate that HCV is a member of the *Flaviviridae* family, which includes the flaviviruses and the pestiviruses (Miller & Purcell, 1990). HCV is a small, enveloped virus containing a positive-stranded linear RNA genome of approximately 9.5 kb. The RNA genome of HCV virus encodes a single polyprotein of 3010 to 3033 amino acids (Matsuura & Miyamura, 1993), which is processed co- and posttranslationally by the host cell as well as by virus-encoded proteases to create mature structural and nonstructural proteins. The HCV gene order has been determined as 5' C-E1-E2-p7-NS2-NS3-NS4A-NS4B-NS5A-NS5B 3' (reviewed in Rice, 1996). The nonstructural proteins are processed by NS2 and NS3, the two viral proteases

Author for correspondence: Czeslaw Wychowski.

Fax +33 3 20 87 11 11. e-mail wychow@infobiogen.fr

(Bartenschlager *et al.*, 1993; Grakoui *et al.*, 1993 *a, b*; Hijikata *et al.*, 1993; Manabe *et al.*, 1994; Tomei *et al.*, 1993), whereas processing of the capsid protein and the two membrane-associated glycoproteins, E1 and E2, is mediated by host signal peptidase(s) (Grakoui *et al.*, 1993 *c*; Hijikata *et al.*, 1991).

Expression of an HCV C-E1-E2 cDNA construct in a cell-free translation system (Hijikata *et al.*, 1991) or a cell culture system (Deleersnyder *et al.*, 1997; Dubuisson *et al.*, 1994; Fournillier-Jacob *et al.*, 1996; Grakoui *et al.*, 1993 *c*; Kohara *et al.*, 1992; Matsuura *et al.*, 1992; Ralston *et al.*, 1993; Spaete *et al.*, 1992) showed that E1 and E2 glycoproteins are targeted to the endoplasmic reticulum (ER) by two independently functioning amino-terminal signal peptides, whereas the capsid protein remains in the cytoplasm (Santolini *et al.*, 1994). Following translocation, Asn-linked (*N*-linked) glycosylation of E1 and E2 takes place: HCV glycoproteins are highly glycosylated. Removal of *N*-linked carbohydrates from HCV glycoproteins with endoglycosidase H or prevention of glycosylation by expression in the presence of tunicamycin results in the loss of approximately half of the molecular mass of the HCV glycoproteins, as determined by SDS-PAGE (Fournillier-Jacob *et al.*, 1996; Grakoui *et al.*, 1993 *c*). E1 and E2 contain five (or six) and eleven potential *N*-linked glycosylation sites, respectively (reviewed in Rice, 1996). E2 glycoprotein seems to display a complex processing consistent with the generation of multiple E2 species (Lin *et al.*, 1994; Mizushima *et al.*, 1994) and forms a stable complex with E1, which is co-immunoprecipitable (Dubuisson *et al.*, 1994; Grakoui *et al.*, 1993 *c*; Lanford *et al.*, 1993; Ralston *et al.*, 1993). Characterization of HCV glycoprotein complex formation using different expression systems indicates that a minority of these proteins are properly folded (Deleersnyder *et al.*, 1997; Dubuisson & Rice, 1996). Properly folded E1 and E2 interact to form a heterodimer stabilized by noncovalent interactions (Deleersnyder *et al.*, 1997). Furthermore, the formation of the noncovalent E1E2 heterodimer is very slow, probably due to slow folding of these proteins (Deleersnyder *et al.*, 1997). It has also been shown that the formation of intramolecular disulfide bonds is slow for E1, whereas it is rapid for E2 (Dubuisson & Rice, 1996; Michalak *et al.*, 1997). In addition, E1 expressed in the absence of E2 does not fold properly, and it has been hypothesized that E2 may play a chaperone-like role in the folding of E1 (Michalak *et al.*, 1997).

N-linked glycosylation, which is one of the most common posttranslational modifications of proteins translocated in the ER, plays various roles in the folding, stability or biological activity of proteins (Kornfeld & Kornfeld, 1985; Opendakker *et al.*, 1993; Rademacher *et al.*, 1988). Animal viruses utilize the host cell glycosylation machinery to synthesize and process oligosaccharides attached to their glycoproteins. The core oligosaccharide chains are preassembled on a carrier lipid precursor in the ER and covalently attached to polypeptides at asparagine in the sequence Asn-X-Ser/Thr, required for *N*-linked glycosylation (Marshall, 1974), where X is any amino

acid except Pro (Gavel & von Heijne, 1990; Kornfeld & Kornfeld, 1985; Marshall, 1974). However, not every tripeptide sequence (or sequon) in a protein sequence is used for carbohydrate addition and indeed many sequons are inefficiently glycosylated (Shakin-Eshleman *et al.*, 1992) or remain unglycosylated (Gavel & von Heijne, 1990).

A recent study has revealed that one potential glycosylation site of HCV glycoprotein E1 (genotype 1a) is not utilized during the glycosylation process (Fournillier-Jacob *et al.*, 1996). In the current study, we used oligonucleotide-directed mutagenesis of the E1 gene to construct a panel of E1 glycosylation mutants, firstly to identify which of the five potential *N*-linked glycosylation sites is not utilized during the glycosylation process and secondly to analyse the influence of E1 glycosylation on the formation of the HCV glycoprotein complex. Our data indicate that site 5 is not used for *N*-linked carbohydrate addition in the E1 glycoprotein and a defect in glycosylation at position 4 has a severe effect on the formation of noncovalent E1E2 complexes. Moreover, using a cell-free translation system, the core glycosylation efficiency at each functional sequon has been determined. While the sequons at positions 1, 2 and 3 are efficiently glycosylated, the sequon at position 4 is partially glycosylated.

Methods

■ **Cell cultures and viruses.** Cell lines (CV1, HepG2 and human 143 B thymidine kinase-deficient cells) were maintained in Dulbecco's modified Eagle's medium (Gibco BRL) supplemented with 5% foetal calf serum and gentamycin (100 µg/ml). For expression studies, the CV-1 or HepG2 cells were cultured in 60 mm diameter culture dishes. The wild-type vaccinia virus Copenhagen strain and its thermosensitive *ts7* derivative (Drillien *et al.*, 1982) were amplified on HeLa cells, purified on sucrose gradients and titrated on CV1 cells. vTF7-3, a vaccinia virus recombinant expressing the T7 DNA-dependent RNA polymerase was obtained from B. Moss (National Institutes of Health, Bethesda, USA).

■ **Site-directed mutagenesis.** The cDNA encoding HCV glycoprotein E1 was cloned into the replicative-form DNA of M13mp18 by using standard techniques in such a manner that the noncoding strand of the E1 cDNA of HCV was packaged into progeny M13 phages. Five synthetic oligonucleotide primers of 24 nucleotides in length were designed so that at each consensus sequence for *N*-linked glycosylation, Asn-X-Thr/Ser, a Gln-encoding codon was substituted for the Asn-encoding codon. Mutagenesis was carried out by the method originally described by Taylor *et al.* (1985) by using a commercially available *in vitro* mutagenesis kit provided by Amersham. Mutants are named with an N (for the Asn amino acid modification) and a number (related to the position of the glycosylation site on the map) (Fig. 1). Single-site glycosylation mutants were used to generate double- and triple-site mutants and so forth. In some conditions, E1 mutants were generated by ligating different enzyme restriction fragments containing the appropriate mutant sites. Cycle sequencing to verify the mutants was performed using the ABI Prism dye terminator ready reaction kit following the manufacturer's instructions (Perkin Elmer Vetus). The mutated fragments were excised from the bacteriophage replicative-form DNA with *EcoRI* and *BamHI* and ligated to pTM1/C^ΔE1₁₃₂₋₃₈₃ (Fournillier-Jacob *et al.*, 1996) or pTM1/C^ΔE1E2p7₁₃₂₋₈₀₉. Recombinant DNA manipulations

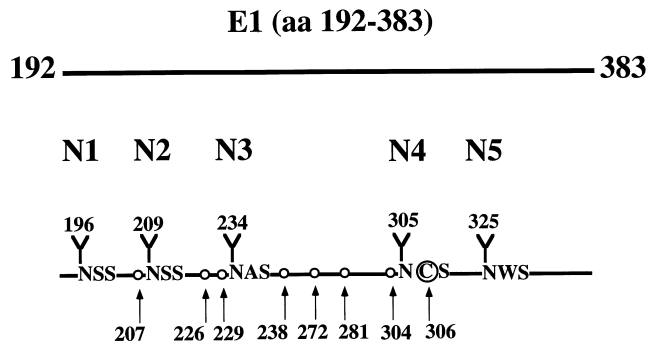


Fig. 1. Schematic representation of *N*-linked glycosylation sites in HCV glycoprotein E1. A linear diagram of the primary sequence of the E1 glycoprotein is shown. E1 is located between amino acids 192 and 383 on the HCV polyprotein. The E1 protein contains five potential *N*-linked glycosylation sites at residues 196, 209, 234, 305 and 325 as indicated by branched structures (Y). The mutants are named with an N (for the Asn amino acid modification) and a number, which is related to the position of the glycosylation site on the map (N1 to N5). The Cys residues are designated by small circles and the amino acid number of each Cys is shown below an arrow.

were made according to standard methods (Sambrook *et al.*, 1989). Recombinant plasmids were controlled by sequencing.

■ **In vitro transcription and translation.** DNA from purified recombinant pTM1 plasmids (2 µg) was used as template for *in vitro* transcription with T7 RNA polymerase (Promega). After a 60 min incubation at 37 °C followed by a 15 min RNase-free DNase RQ1 treatment (1 unit/µg), RNAs were extracted with phenol/chloroform (v/v) and ethanol-precipitated. Purified RNA transcripts were used at a concentration of 100 µg/ml for *in vitro* translation reactions using a rabbit reticulocyte lysate (15 µl) (Promega) supplemented with 15 µCi [³⁵S]methionine in the presence or absence of canine pancreatic microsomal membranes. After 60 min at 30 °C, samples were heated at 100 °C for 5 min in the presence of Laemmli buffer (Laemmli, 1970). Cell-free translation protein products were analysed by SDS-PAGE.

■ **Generation of recombinant vaccinia viruses.** Transfection and isolation of recombinant viruses were performed essentially as previously described (Kieny *et al.*, 1984).

■ **Antibodies.** Anti-HCV E1 (A4) and E2 (A11 and H2) MAbs have been described previously (Deleersnyder *et al.*, 1997; Dubuisson *et al.*, 1994) and were produced *in vitro* by using a MiniPerm apparatus (Heraeus) as recommended by the manufacturer. Anti-calnexin antibodies (SPA-860) were supplied by Stress Gen.

■ **Metabolic labelling and immunoprecipitation.** HepG2 or CV1 cells were infected with vTF7-3 alone or in combination with another recombinant vaccinia virus, each at an m.o.i. of 5 p.f.u. per cell, and metabolically labelled with ³⁵S-protein labelling mix (NEN) as previously described (Dubuisson & Rice, 1996). Labelled infected cells were then lysed with 0.5% Triton X-100 in 10 mM Tris-HCl (pH 7.5), 150 mM NaCl and 2 mM EDTA (TBS). For analyses under nonreducing conditions, the lysis buffer was supplemented with 20 mM iodoacetamide. Cell lysates were clarified by centrifugation in an Eppendorf centrifuge for 15 min.

Immunoprecipitations were carried out as described previously (Dubuisson & Rice, 1996). A 6 µl aliquot of rabbit anti-mouse IgG (Dako) was incubated with protein A-Sepharose (Pharmacia LKB) for 1 h at 4 °C in TBS containing 0.2% Triton X-100 (TBS-T). This step was omitted when polyclonal antibodies were used. The beads were then incubated

with 10 µl of MAb or polyclonal antibody, followed by the antigen (each step was performed for 1 h at 4 °C). Between each step, the beads were washed once with TBS-T. After the last step, they were washed three times with TBS-T and once with TBS. The immune complexes were boiled for 5 min in Laemmli buffer (Laemmli, 1970) (under nonreducing conditions, β-mercaptoethanol was omitted) before analysis by SDS-PAGE. After electrophoresis, the gels were treated with sodium salicylate (Chamberlain, 1979), dried and exposed at -70 °C to preflashed Hyperfilm-MP (Amersham). ¹⁴C-methylated protein molecular mass markers were purchased from Amersham.

Results

Identification of the functional *N*-linked glycosylation sites in E1

One of the five potential *N*-linked oligosaccharide acceptor sites of E1 is not utilized during glycosylation (Fournillier-Jacob *et al.*, 1996). To identify the unused glycosylation site, mutant E1 proteins each lacking one of the glycosylation sites were obtained by site-directed mutagenesis.

Oligonucleotide-directed mutagenesis was employed to introduce two nucleotide changes in the codon encoding Asn, resulting in a single amino acid substitution at each potential glycosylation site. The addition of *N*-linked oligosaccharides was prevented by changing the Asn-X-Ser consensus sequence at Asn residues 196, 209, 234, 305 and 325 to Gln-X-Ser. The mutants in which the consensus sequence was altered are referred to as N1, N2, N3, N4 and N5. The positions are numbered sequentially from the N terminus of E1 (Fig. 1). The mutated cDNAs were reintroduced into the eukaryotic expression vector pTM1/C^ΔE1₁₃₂₋₃₈₃ (Fournillier-Jacob *et al.*, 1996) or pTM1/C^ΔE1E2p7₁₃₂₋₈₀₉, allowing expression of E1 or E1E2, respectively. In the first approach, the electrophoretic mobilities of the E1 mutants for a specific site of glycosylation (1, 2, 3, 4 or 5) were analysed in an *in vitro* transcription translation experiment and compared to the wild-type E1 expressed similarly *in vitro*. E1 glycoproteins with mutations at sites 1, 2, 3 or 4 (Fig. 2, E1 lanes N1 to N4 or E1E2 lanes N1

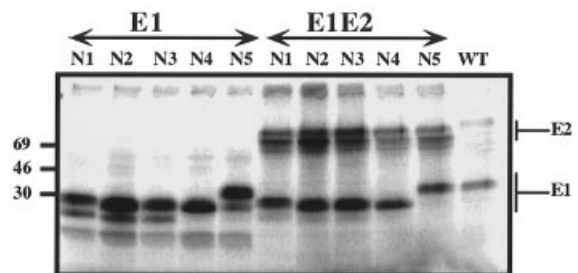


Fig. 2. SDS-PAGE analysis of single E1 glycosylation mutants translated *in vitro*. T7 RNA transcripts obtained from DNA plasmids containing sequences encoding mutants E1 N1 to N5 and E1E2 N1 to N5, and for wild-type as control, were translated in a rabbit reticulocyte lysate in the presence of microsomal membranes, with the addition of [³⁵S]methionine in the reaction mixture. The location of wild-type E1 and E2 proteins is indicated on the right. Sizes (in kDa) of protein molecular mass markers are indicated on the left.

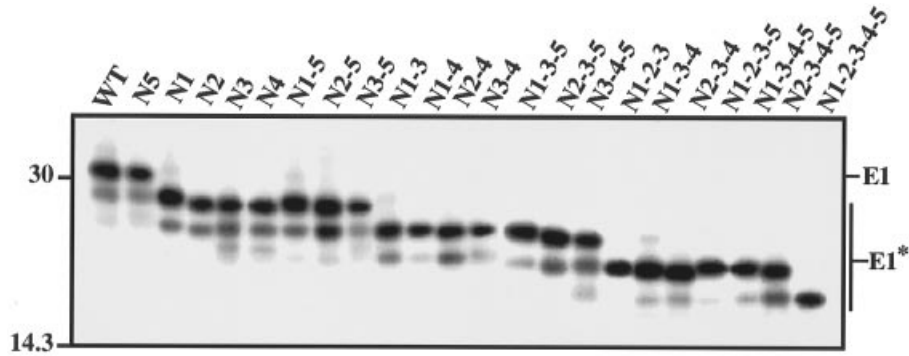


Fig. 3. SDS-PAGE analysis of wild-type and glycosylation mutants of HCV E1. HepG2 cells were coinfectd with vTF7-3 and vaccinia virus recombinants expressing either wild-type E1E2 (WT) or the different mutants of E1 (within E1E2) at a multiplicity of 5 p.f.u. per cell. Infected cells were radiolabelled with [35 S]methionine for 2 h at 18 h post-infection. Proteins present in cell extracts were immunoprecipitated with MAb A4 (anti-E1) and separated by 12% SDS-PAGE and autoradiography. E1* represents the mutant E1 proteins altered in one, two, three or all glycosylation sites. The sizes of two 14 C-labelled marker polypeptides, carbonic anhydrase (30 kDa) and lysozyme (14.3 kDa), are indicated on the left.

to N4) migrated faster than the wild-type protein, indicating that these sites are used in the wild-type protein, while the protein with a mutation at site 5 (Fig. 2, E1 or E1E2 lanes N5) comigrated with the wild-type glycoprotein, suggesting that this site is not used for carbohydrate addition.

To confirm the above-mentioned conclusions, some mutant E1 proteins missing one or several glycosylation sites were made and analysed in *in vivo* infection experiments. The E1 mutants with several mutations were referred to as N1-5, N2-5, N1-3-5 and so forth (Fig. 3). All mutated cDNAs were introduced into pTM1/C^AE1E2p7₁₃₂₋₈₀₉ as mentioned above. Recombinant vaccinia viruses, expressing the different E1 mutants, were generated by homologous recombination (Kieny *et al.*, 1984). Mutated proteins expressed by recombinant vaccinia viruses were analysed by SDS-PAGE after immunoprecipitation with an anti-E1 MAb (A4) (Fig. 3). The wild-type E1 and the different E1 mutants expressed a protein with a molecular mass which ranged between 31 kDa and 19 kDa (wild-type E1 to N1-2-3-4-5 E1 protein). As expected, the electrophoretic mobility of the mutant E1 proteins increased proportionally with the number of inactivated glycosylation sites, with the exception of the mutant proteins containing the inactivated N5 site (Fig. 3, compare N1 and N1-5, N2 and N2-5 or N3 and N3-5). The difference in molecular mass between E1 mutants and the wild-type E1 is due to the number of carbohydrate chains attached. These results indicate that the potential glycosylation sites 1 to 4 are normally used, whereas site 5 (Asn-Trp-Ser-Pro-) is not glycosylated.

Absence of glycosylation at site 5 is due to the presence of a Pro residue downstream of the sequon

A dramatic impairment of glycosylation can be obtained with some glycoproteins. Indeed, no glycosyl transfer is detectable for peptides or proteins containing a Pro residue either in position X or in position Y of the Asn sequon (Asn-

X-Ser/Thr-Y) (Bause, 1983; Gavel & von Heijne, 1990; Kornfeld & Kornfeld, 1985; Marshall, 1974). In addition, recent studies have demonstrated that the presence of large hydrophobic amino acids (i.e. Trp, Leu or Phe) present at the X position of the sequon and leading to an unfavourable local protein conformation may inhibit core glycosylation (Shakin-Eshleman *et al.*, 1996).

To confirm the involvement of amino acids Trp (W326) and Pro (P328) in the impairment of glycosylation site 5, these amino acids were changed by site-directed mutagenesis to Thr (T326) and Gly (G328), respectively. The amino acid Thr was selected according to Shakin-Eshleman *et al.* (1996), who reported that the presence of an amino acid with a hydroxy group such as Ser/Thr or Cys is associated with highly efficient core glycosylation. A single E1 mutant was constructed with a modification of the amino acid from Trp to Thr (W326 to T326) and a double E1 mutant with the modification of amino acids Trp to Thr (W326 to T326) and Pro to Gly (P328 to G328). The mutated cDNA sequences were introduced into pTM1/C^AE1E2₁₃₂₋₈₀₉, and recombinant vaccinia viruses were generated. The resulting mutants, named m1 and m2, correspond to the single and double mutants of E1, respectively (Fig. 4A). The mutated proteins expressed by recombinant vaccinia viruses were analysed by SDS-PAGE after immunoprecipitation with an anti-E1 MAb (A4) (Fig. 4B). The wild-type E1 and the m1E1 protein migrated as 31 kDa glycoproteins, while the m2E1 protein migrated as a 34 kDa glycoprotein, which corresponds to the addition of an N-linked oligosaccharide on amino acid Asn of the new sequon Asn-Thr-Ser-Gly. These results indicate that the replacement of amino acid W326 by T326 does not modify the profile of glycosylation of the m1E1 protein containing the sequon Asn-Thr-Ser-Pro. However, modification of P328 to G328 leads to the addition of an N-linked oligosaccharide on the Asn residue within the sequon Asn-Thr-Ser-Gly of the m2E1 protein. This analysis provides direct evidence that the presence of P328

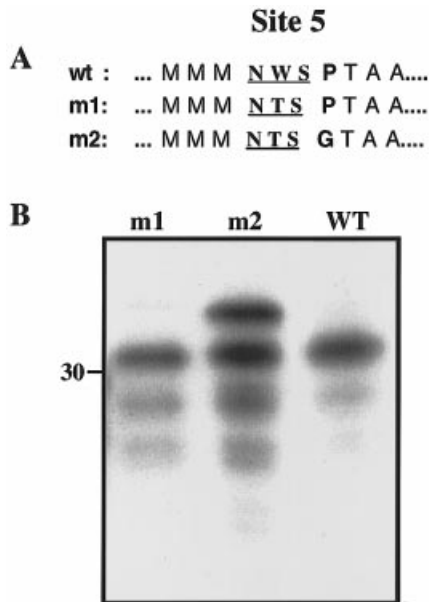


Fig. 4. Influence of the Pro residue in unglycosylated site 5. (A) The primary sequence in the region of unglycosylated site 5 and amino acid changes in the single mutant (m1) and double mutant (m2). The m1 mutant contains the modification of the amino acid Trp to Thr (W326 to T326) and the m2 mutant contains the modification of amino acids Trp to Thr (W326 to T326) and Pro to Gly (P328 to G328). (B) HepG2 cells were coinfecting with vTF7-3 and vaccinia virus recombinants expressing either wild-type E1E2 (WT) or the m1 or m2 (within E1E2) at a multiplicity of 5 p.f.u. per cell. Infected cells were radiolabelled with [³⁵S]methionine for 2 h at 18 h post-infection. Proteins present in cell extracts were immunoprecipitated with MAb A4 (anti-E1) and separated by 12% SDS-PAGE and autoradiography. The size of the ¹⁴C-labelled marker polypeptide (carbonic anhydrase, 30 kDa) is indicated on the left.

immediately following the sequon Asn-Trp-Ser precludes core glycosylation at site 5.

Analysis of the role of E1 glycosylation on the formation of HCV glycoprotein complexes

In polyprotein maturation, one of the roles of glycans is to help in the folding of glycoproteins, either directly, by increasing the solubility of a protein, or indirectly, by interacting with ER chaperones like calnexin or calreticulin (Helenius, 1994). We were therefore interested in knowing whether or not mutations in the E1 glycosylation sites would alter E1 glycoprotein folding. Since there is no conformation-sensitive MAb available against E1, its folding was assessed indirectly with the help of a conformation-sensitive E2-specific MAb (H2), which has been shown, at least in the context of wild-type proteins, to specifically precipitate properly folded E1E2 heterodimers (Deleersnyder *et al.*, 1997).

Among the single mutants (N1 to N5), only the N4 mutation affected the coprecipitation of E1, as shown by immunoprecipitation with MAb H2 under reducing conditions (Fig. 5 A, compare lane N4 with lanes N1, N2 and N3). Mutation 5 on the E1 protein has no effect on the formation of E1E2 complexes and may be considered for the following

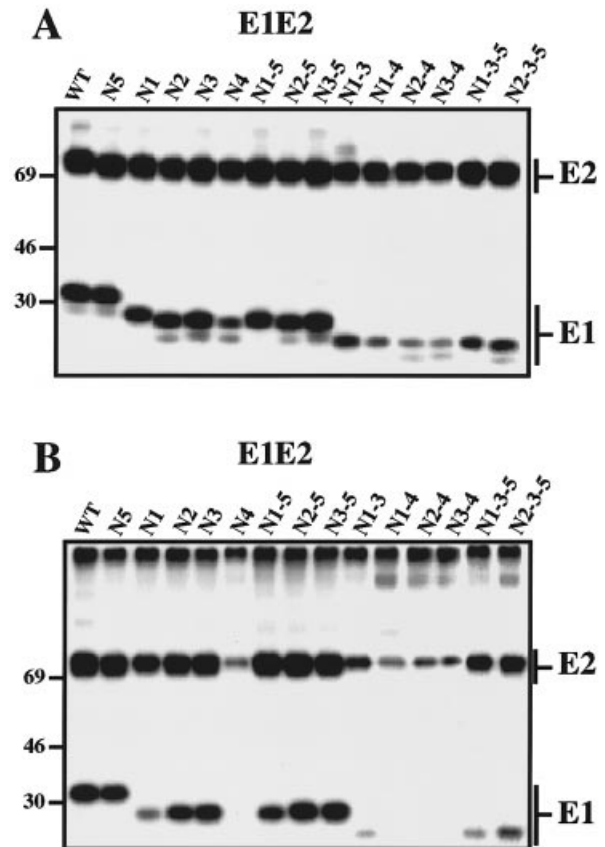


Fig. 5. Importance of E1 glycans on the formation of HCV glycoprotein complexes. HepG2 cells were coinfecting with vTF7-3 and vaccinia virus recombinants expressing either wild-type E1E2 (WT) or different mutants of E1 (within E1E2) at a multiplicity of 5 p.f.u. per cell. Infected cells were labelled with [³⁵S]methionine for 10 min and then incubated with the chase medium for 4 h. Cells were lysed in the presence of iodoacetamide (20 mM) to prevent further disulfide bond formation and the viral E1E2 complex was immunoprecipitated with MAb H2 and separated on SDS-PAGE under reducing (A) and nonreducing conditions (B). HCV-specific proteins are indicated on the right. Sizes (in kDa) of protein molecular mass markers are indicated on the left.

study as a control (Fig. 5 A, compare wild-type and N5, N2 and N2-5, and N3 and N3-5). In addition, as a general observation, the amount of E1 which coprecipitated with E2 was lower when the number of mutations in E1 increased (Fig. 5 A). In this study, one of the glycoproteins was mutated and, even if the domain bearing the H2 epitope in E2 was properly folded, we could not exclude the possibility of an alteration in the folding of another domain of E2 induced by the coexpression of the mutated form of E1. The immunoprecipitates were therefore analysed under nonreducing conditions to see whether or not the complexes formed between E2 and mutated E1 proteins were covalently linked. As shown in Fig. 5(B), it appears that E1 proteins lacking oligosaccharides at position 1 and predominantly at position 4, singly or in combination, reduced the formation of noncovalent complexes (Fig. 5 B, N1 and N4, and compare N1-4, N2-4, N3-4 with N1-3-5, N2-3-5), whereas E1 lacking an oligosaccharide at position 2 or 3 had no

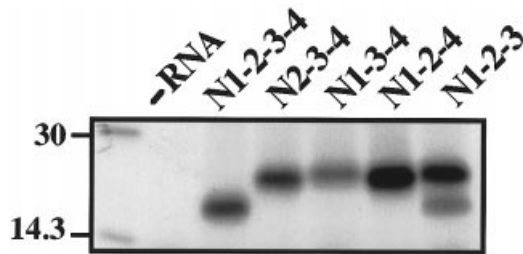


Fig. 6. Determination of *N*-linked core glycosylation efficiency for each sequon of E1 protein. RNA encoding E1 glycosylation mutants with no functional sequon (N1-2-3-4) or only one functional sequon at position 1, 2, 3 or 4 (respectively, N2-3-4, N1-3-4, N1-2-4 and N1-2-3) was prepared by *in vitro* transcription and translated in a rabbit reticulocyte lysate system containing [³⁵S]methionine in the presence of canine pancreatic microsomal membranes. A translational reaction including the canine pancreatic microsomal membranes but containing no exogenous RNA is also shown (lane -RNA). Translation products were analysed directly by SDS-PAGE and autoradiography. The positions of molecular mass markers are indicated on the left (in kDa).

impairment in noncovalent complex formation (Fig. 5 B, N2 and N3, N2-5 and N3-5). Finally, the formation of noncovalent complexes was abolished when several glycosylation sites were mutated in E1 (results not shown). Surprisingly, some aggregates can also be detected by MAb H2 (Fig. 5 B, lane 1). This may be a result of constructs that we used in this study. Indeed, neither the capsid nor the NS2 protein has been co-expressed with E1E2 as previously described (Deleersnyder *et al.*, 1997). Together, these data indicate that mutated E1 proteins interfere with the folding of E2, leading to misfolding of an E2 domain that aggregates with E1, while folding of the domain bearing the H2 epitope is not altered. These data also indicate that the formation of noncovalent E1E2 complexes can be impaired when some glycosylation sites are mutated in E1.

Determination of the efficiency of core glycosylation for each functional sequon of E1 protein

As a defect of glycosylation at position 1 and predominantly at position 4 of the E1 protein can reduce the formation of noncovalent E1E2 complexes, the core glycosylation efficiency at each of the four sequons was analysed.

A series of E1 glycosylation mutants containing only one functional sequon at each position described below was analysed using a cell-free translation system as described in Methods. Expression of E1 containing all mutated sequons produced a single major radiolabelled species with a molecular mass of 19 kDa (Fig. 6, lane N1-2-3-4), which represents the unglycosylated E1 protein. The E1 mutants containing a functional sequon at positions 1, 2 and 3 were efficiently glycosylated and produced a single major band with a molecular mass of 22 kDa (Fig. 6, lanes N2-3-4, N1-3-4 and N1-2-4), while the sequon at position 4 was glycosylated at a lower level and produced two radiolabelled products (19 kDa and 22 kDa proteins) (Fig. 6, lane N1-2-3). As determined by Phosphorimager analysis of the gels (not shown) only 66% of E1 was core glycosylated at position 4. When all these E1

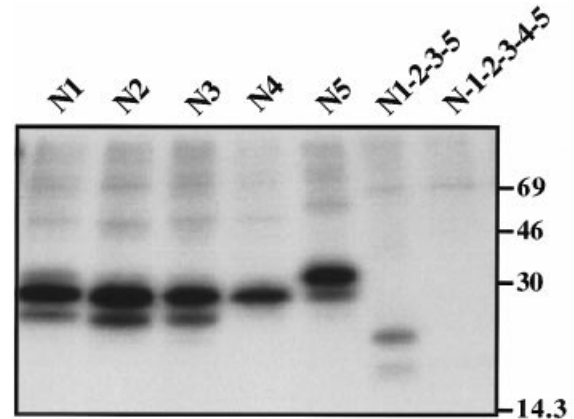


Fig. 7. Interaction of calnexin with E1 mutants lacking glycans. HepG2 cells were coinfecting with vTF7-3 and vaccinia virus recombinants expressing mutants of E1 (alone) at a multiplicity of 5 p.f.u. per cell. Mutant N5, containing all glycans, was used as a control. At 6 h post-infection, infected cells were pulse-labelled for 15 min with [³⁵S]methionine, chased for 10 min and then lysed with Triton X-100. Cell lysates were used for immunoprecipitation with anti-calnexin antibodies. The anti-calnexin immune complexes were dissociated by heating for 5 min at 37 °C in 0.5% Nonidet P-40 and reprecipitated with MAb A4 (anti-E1). Immunoprecipitates were analysed by 12% SDS-PAGE under reducing conditions. The sizes (in kDa) of molecular mass markers are indicated on the right.

mutants were expressed in a cell-free translation system in the absence of canine pancreatic microsomal membranes, only one product, with a molecular mass of 22 kDa, corresponding to the uncleaved product C^ΔN-E1₁₃₂₋₃₈₃, was expressed. This result indicates that the unglycosylated form of E1 (N1-2-3) is not due to an untranslocated product (result not shown). As revealed *in vitro*, site 4 cannot be efficiently glycosylated and this may explain why some partial glycosylation forms of E1 are expressed *in vitro* and *in vivo*.

Analysis of the association of mutant E1 proteins with calnexin

Calnexin interacts with HCV glycoproteins (Dubuisson & Rice, 1996) and this chaperone has been shown to be involved in the assembly of the noncovalent E1E2 heterodimer (Choukhi *et al.*, 1998). The low level of noncovalent E1E2 complexes formed when E2 was coexpressed with some glycosylation mutants of E1 could therefore be due to an alteration in the interaction of E1 with calnexin. To evaluate this possibility, the interaction of this chaperone with some glycosylation mutants of E1 was investigated.

The ability of calnexin to interact with mutated E1 proteins was analysed by immunoprecipitation with an anti-calnexin antibody as previously described (Dubuisson & Rice, 1996). As shown in Fig. 7, the absence of one glycan at position N1, N2, N3 or N4 did not reduce the interaction between E1 and calnexin. This suggests that the lower level of noncovalent complex formation observed for mutants N1 and N4 (Fig. 6 B, lanes N1 and N4) is not due to an alteration in their interaction with calnexin. Furthermore, the ability of calnexin to interact

with E1 mutants exhibiting an increasing number of mutated glycosylation sites was evaluated. The interaction of E1 with calnexin was erased only with the E1 protein lacking all glycans, as shown in Fig. 7 (lane N1-2-3-4-5). However, the intensity of the bands coprecipitated with calnexin decreased with the number of glycans left on E1. In conclusion, it appears that the coprecipitation of mutant E1 proteins with calnexin is dependent on the number of glycans present on the proteins and this may account for the low level of proper folding of some mutant E1 proteins.

Discussion

The addition of *N*-linked oligosaccharides to specific Asn residues has been the subject of many investigations. It appears that this modification plays an important role in regulating the activity, stability or antigenicity of the mature protein (Opdenakker *et al.*, 1993; Rademacher *et al.*, 1988). It has also been suggested that *N*-linked glycosylation is required for folding, transport, cell surface expression, secretion of glycoproteins (Helenius, 1994), protection from proteolytic degradation and enhancement of glycoprotein solubility (Doms *et al.*, 1993; Rademacher *et al.*, 1988). Although the sequon is essential for core glycosylation, numerous examples of glycoproteins have been observed containing sequons which are either unglycosylated or glycosylated at a low level (Curling *et al.*, 1990; Pohl *et al.*, 1984). In this study, the glycosylation sites used in HCV glycoprotein E1 have been defined and the role of individual oligosaccharide chains in the formation of HCV glycoprotein complexes has been investigated.

Mutagenesis studies of E1 glycosylation sites indicate that site 5 in position 325 of the HCV polyprotein is non-glycosylated. While sequon Asn-X-Ser/Thr is absolutely required for the attachment of *N*-linked oligosaccharides to a glycoprotein (Marshall, 1974), its presence does not always result in glycosylation and some such sequons in glycoproteins can remain unglycosylated (Nakai & Kanehisa, 1988). Large hydrophobic amino acids (e.g. Trp, Leu, Phe and Tyr) (Shakin-Eshleman *et al.*, 1996) or Pro (Bause, 1983; Gavel & von Heijne, 1990; Kornfeld & Kornfeld, 1985; Shakin-Eshleman *et al.*, 1996) present at position X of the sequon have been described as affecting core glycosylation by producing unfavourable local protein conformation, or blocking the accessibility of the oligosaccharyl transferase or the dolichol oligosaccharide donor to the sequon. Inhibition of core glycosylation can also be observed when a Pro is present at position Y of the sequon (Asn-X-Ser/Thr-Y) (Bause, 1983; Mellquist *et al.*, 1998). The unused glycosylation site 5 with the sequence Asn-Trp-Ser-Pro presents two unfavourable amino acids. Our results indicate that the modification of amino acid Trp to an amino acid with a hydroxy group, such as Thr, which is associated with highly efficient core glycosylation (Shakin-Eshleman *et al.*, 1996), does not restore core glycosylation. However, when both Trp at position X and Pro at position Y were mutated,

attachment of a carbohydrate to glycosylation site 5 was observed. These data demonstrate that the inhibitory effect of Trp and Pro in the sequence Asn-Trp-Ser-Pro can be overcome by replacing them with Thr and Gly, respectively. However, the addition of a carbohydrate moiety to the glycosylation site is not blocked when Pro is located on the N-terminal side of the sequon, as shown by the glycosylation of site 2 of HCV glycoprotein E1 (Pro-Asn-Ser-Ser). These results confirm that glycosylation can be severely affected, depending on the position of Pro in the sequon (Bause, 1983).

A comparison of HCV nucleotide and amino acid sequences obtained from different parts of the world shows considerable diversity, and the existence of multiple HCV genotypes has been determined (Simmonds *et al.*, 1993). Analysis of the Japanese HCV strain BK has revealed six potential glycosylation sites (compared to the five of the prototype HCV-1) (Choo *et al.*, 1991), five of these sites being very conserved. Recent studies on the BK strain have shown that among the six potential sites, only five are glycosylated (De Martynoff *et al.*, 1997). The unused glycosylation site has not been localized for the BK strain but site 6, located at the same position as site 5 of HCV-1, has the same unfavourable sequence, Asn-Trp-Ser-Pro. This observation strongly suggests that site 6 of the BK strain is not recognized during glycosylation. The conservation among the glycosylation sites suggests that a selective pressure ensures the availability of these sites for carbohydrate attachment.

Inactivation of some functional glycosylation sites by site-directed mutagenesis reduces the formation of noncovalent HCV glycoprotein complexes and the severity of the defect depends on the location of the mutated glycosylated site. Mutation of glycosylation sites 2 and 3 of HCV E1 has no effect on the formation of noncovalent HCV glycoprotein complexes, in contrast to mutations at sites 1 and 4. The effect on noncovalent complex formation was more dramatic when site 4 was modified. It is likely that the presence of glycans at sites 1 and 4 is important for stabilizing the E1 structure or favouring its proper folding. The presence of a Cys residue at glycosylation site 4 in position X (Asn-Cys-Ser) suggests that the Cys cannot be involved in intramolecular disulfide bond formation if the site is glycosylated. However, in the absence of glycosylation, the Cys residue located at amino acid position 306 on the HCV polyprotein (Fig. 1) probably exhibits a free SH group that can interact with other proteins. In the context of coexpression with E2, an E1 protein mutated at position 4 may interact preferentially with E2. Indeed, folding of HCV glycoproteins involves the formation of intermediate E1E2 complexes and folding of E1 is thought to occur in these complexes (Choukhi *et al.*, 1998). The disulfide bridge potentially formed between the Cys present in site 4 and another Cys residue in E2 is probably very stable. Immunoprecipitation of aggregates by MAb H2 suggests that the N4 mutant is misfolded, causing the exposure of normally concealed Cys residues which then become available for disulphide bridging.

This may cause correctly folded E2 protein to become involved in aggregates. Sequon 1 does not contain any Cys residues, but mutating site 1 probably has some effect on the E1 folding which could result in disulfide bridge formation with a domain of E2. However, the effect on noncovalent E1E2 complex formation was less dramatic with the N1 mutant than with N4. HCV glycoprotein E1 lacking an oligosaccharide at position 2 or 3 had no impairment in noncovalent complex formation, indicating that the glycans present in these positions do not help in the folding and/or stability of the structure of E1. However, these glycans could play other roles, such as impeding recognition by neutralizing antibodies or protecting the glycoprotein from degradation in the extracellular environment. Due to the problems of replication in tissue culture, these potential functions cannot be studied for HCV.

Glycosylation mutants of E1 with a single modification did not show any impairment in their interaction with calnexin. Calnexin is an ER chaperone which binds selectively and transiently to newly synthesized glycoproteins (reviewed in Bergeron *et al.*, 1994) according to a mechanism based on a lectin-like affinity for monoglucosylated *N*-linked oligosaccharides (Hammond *et al.*, 1994; Hebert *et al.*, 1995; Ware *et al.*, 1995). It has been reported that calnexin interacts with newly synthesized HCV glycoproteins (Dubuisson & Rice, 1996) and more recent data show that this chaperone is involved in their folding (Choukhi *et al.*, 1998). However, since the glycosylation mutants N1 to N4 did not show any modification in their interaction with calnexin, the alteration of folding observed for some of them (N1 and N4) does not seem to involve calnexin. As discussed above, it may instead be due to exposure of free Cys residues. The amount of mutated E1 protein coprecipitated with calnexin depended on the number of glycans added to E1. Misfolding of the E1 protein with several mutations might therefore lead to a decrease in the E1 chaperoning activity of calnexin.

HCV glycoproteins are heavily glycosylated proteins. In this work, we have shown that some glycans present on E1 play a role in the folding and/or stability of the structure of E1. A similar approach will be necessary for a functional analysis of the glycans present on E2.

We thank André Pillez and Sophana Ung for excellent technical assistance. This work was supported by the CNRS, the Institut Pasteur de Lille and grant 1039 from the ARC. M.J.-C. was supported by a grant from the ARC and by Dominique Stehelin's Jeantet Prize.

References

- Alter, M. J., Hadler, S. C., Judson, F. N., Mares, A., Alexander, W. J., Hu, P. Y., Miller, J. K., Moyer, L. A., Fields, H. A. & Bradley, D. W. (1990). Risk factors for acute non-A non-B hepatitis in the United States and association with hepatitis C virus infection. *Journal of the American Medical Association* **264**, 2231–2235.
- Alter, M. J., Margolis, H. S., Krawczynski, K., Judson, F. N., Mares, A., Alexander, W. J., Hu, P. Y., Miller, J. K., Gerber, M. A., Sampliner, R. E., Meeks, E. L. & Beach, M. J. (1992). The natural history of community-acquired hepatitis C in the United States. The sentinel counties chronic non-A, non-B hepatitis study team. *New England Journal of Medicine* **327**, 1899–1905.
- Bartenschlager, R., Ahlbornlaake, L., Mous, J. & Jacobsen, H. (1993). Nonstructural protein 3 of the hepatitis C virus encodes a serine-type proteinase required for cleavage at the NS3/4 and NS4/5 junctions. *Journal of Virology* **67**, 3835–3844.
- Bause, E. (1983). Structural requirements of *N*-glycosylation of proteins. *Biochemical Journal* **209**, 331–336.
- Bergeron, J. J., Brenner, M. B., Thomas, D. Y. & Williams, D. B. (1994). Calnexin: a membrane-bound chaperone of the endoplasmic reticulum. *Trends in Biochemical Sciences* **19**, 124–128.
- Chamberlain, J. P. (1979). Fluorographic detection of radioactivity in polyacrylamide gels with the water-soluble fluor, sodium salicylate. *Analytical Biochemistry* **98**, 132–135.
- Choo, Q.-L., Kuo, G., Weiner, A. J., Overby, L. R., Bradley, D. W. & Houghton, M. (1989). Isolation of a cDNA clone derived from a blood-borne non-A, non-B viral hepatitis genome. *Science* **244**, 359–362.
- Choo, Q.-L., Richman, K. H., Han, J. H., Berger, K., Lee, C., Dong, C., Gallegos, C., Coit, D., Medina-Selby, A., Barr, P. J., Weiner, A. J., Bradley, D. W., Kuo, G. & Houghton, M. (1991). Genetic organisation and diversity of the hepatitis C virus. *Proceedings of the National Academy of Sciences, USA* **88**, 2451–2455.
- Choukhi, A., Ung, S., Wychowski, C. & Dubuisson, J. (1998). Involvement of endoplasmic reticulum chaperones in the folding of hepatitis C virus glycoproteins. *Journal of Virology* **72**, 3851–3858.
- Curling, E. M., Hayter, P. M., Baines, A. J., Bull, A. T., Gull, K., Strange, P. G. & Jenkins, N. (1990). Recombinant human interferon-gamma. Differences in glycosylation and proteolytic processing lead to heterogeneity in batch culture. *Biochemical Journal* **272**, 333–337.
- Deleersnyder, V., Pillez, A., Wychowski, C., Blight, K., Xu, J., Hahn, Y. S., Rice, C. M. & Dubuisson, J. (1997). Formation of native hepatitis C virus glycoprotein complexes. *Journal of Virology* **71**, 697–704.
- De Martynoff, G., Venneman, A. & Maertens, G. (1997). Analysis of post-translational modifications of HCV structural proteins by using the vaccinia virus expression system. In *Viral Hepatitis and Liver Disease*. Edited by M. Rizzetto, R. H. Purcell & H. Gerin. Turin: Minerva Medica.
- Doms, R. W., Lamb, R. A., Rose, J. R. & Helenius, A. (1993). Folding and assembly of viral membrane proteins. *Virology* **193**, 545–562.
- Drillien, R., Spehner, D. & Kirn, A. (1982). Complementation and genetic linkage between vaccinia virus temperature-sensitive mutants. *Virology* **68**, 372–381.
- Dubuisson, J. & Rice, C. M. (1996). Hepatitis C virus glycoprotein folding: disulfide bond formation and association with calnexin. *Journal of Virology* **70**, 778–786.
- Dubuisson, J., Hsu, H. H., Cheung, R. C., Greenberg, H. B., Russell, D. G. & Rice, C. M. (1994). Formation and intracellular localization of hepatitis C virus envelope glycoprotein complexes expressed by recombinant vaccinia and sindbis viruses. *Journal of Virology* **68**, 6147–6160.
- Fournillier-Jacob, A., Cahour, A., Escriou, N., Girard, M. & Wychowski, C. (1996). Processing of the E1 glycoprotein of hepatitis C virus expressed in mammalian cells. *Journal of General Virology* **77**, 1055–1064.
- Gavel, Y. & von Heijne, G. (1990). Sequence differences between glycosylated and non-glycosylated Asn-X-Thr/Ser acceptor sites: implications for protein engineering. *Protein Engineering* **3**, 433–442.
- Grakoui, A., Mccourt, D. W., Wychowski, C., Feinstone, S. M. & Rice, C. M. (1993a). Characterization of the hepatitis C virus-encoded serine

- proteinase – determination of proteinase-dependent polyprotein cleavage sites. *Journal of Virology* **67**, 2832–2843.
- Grakoui, A., Mccourt, D. W., Wychowski, C., Feinstone, S. M. & Rice, C. M. (1993b).** A second hepatitis-C virus-encoded proteinase. *Proceedings of the National Academy of Sciences, USA* **90**, 10583–10587.
- Grakoui, A., Wychowski, C., Lin, C., Feinstone, S. M. & Rice, C. M. (1993c).** Expression and identification of hepatitis C virus polyprotein cleavage products. *Journal of Virology* **67**, 1385–1395.
- Hammond, C., Braakman, I. & Helenius, A. (1994).** Role of N-linked oligosaccharides, glucose trimming and calnexin during glycoprotein folding in the endoplasmic reticulum. *Proceedings of the National Academy of Sciences, USA* **91**, 913–917.
- Hebert, D. N., Foellmer, B. & Helenius, A. (1995).** Glucose trimming and reglucosylation determine glycoprotein association with calnexin in the endoplasmic reticulum. *Cell* **81**, 425–433.
- Helenius, A. (1994).** How N-linked oligosaccharides affect glycoprotein folding in the endoplasmic reticulum. *Molecular Biology of the Cell* **5**, 253–265.
- Hijikata, M., Kato, N., Ootsuyama, Y., Nakagawa, M. & Shimotohno, K. (1991).** Gene mapping of the putative structural region of the hepatitis C virus genome by in vitro processing analysis. *Proceedings of the National Academy of Sciences, USA* **88**, 5547–5551.
- Hijikata, M., Mizushima, H., Akagi, T., Mori, S., Kakiuchi, N., Kato, N., Tanaka, T., Kimura, K. & Shimotohno, K. (1993).** Two distinct proteinase activities required for the processing of a putative non-structural precursor protein of hepatitis C virus. *Journal of Virology* **67**, 4665–4675.
- Houghton, M. (1996).** Hepatitis C viruses. In *Fields Virology*, 3rd edn, pp. 1035–1058. Edited by B. N. Fields, D. M. Knipe & P. M. Howley. Philadelphia: Lippincott–Raven.
- Kieny, M.-P., Lathe, R., Drillien, R., Spohner, D., Skory, S., Schmitt, D., Wiktor, T., Koprowski, H. & Lecocq, J.-P. (1984).** Expression of rabies virus glycoprotein from a recombinant vaccinia virus. *Nature* **312**, 163–166.
- Kohara, M., Tsukiyama-Kohara, K., Maki, N., Asano, K., Yamaguchi, K., Miki, K., Tanaka, S., Hattori, N., Matsuura, Y., Saito, I., Miyamura, T. & Nomoto, A. (1992).** Expression and characterization of glycoprotein gp35 of hepatitis C virus using recombinant vaccinia virus. *Journal of General Virology* **73**, 2313–2318.
- Kornfeld, R. & Kornfeld, S. (1985).** Assembly of asparagine-linked oligosaccharides. *Annual Review of Biochemistry* **54**, 631–664.
- Kuo, G., Choo, Q.-L., Alter, H. J., Gitnick, G. L., Redecker, A. G., Purcell, R. H., Miyamura, T., Dienstag, J. L., Alter, M. J., Stewens, C. E., Tegtmeier, G. E., Bonino, F., Colombo, M., Lee, W. S., Kuo, C., Berger, K., Shuster, J. R., Overby, L. R., Bradley, D. W. & Houghton, M. (1989).** An assay for circulating antibodies to a major etiologic virus of human non-A, non-B hepatitis. *Science* **244**, 362–364.
- Laemmli, U. K. (1970).** Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227**, 680–685.
- Lanford, R. E., Notvall, L., Chavez, D., White, R., Frenzel, G., Simonsen, C. & Kim, J. (1993).** Analysis of hepatitis C virus capsid, E1, and E2/NS1 proteins expressed in insect cells. *Virology* **197**, 225–235.
- Lin, C., Lindenbach, B. D., Pragai, B. M., Mccourt, D. W. & Rice, C. M. (1994).** Processing in the hepatitis C virus E2–NS2 region: identification of p7 and two distinct E2-specific products with different C termini. *Journal of Virology* **68**, 5063–5073.
- Manabe, S., Fuke, I., Tanishita, O., Kaji, C., Gomi, Y., Yoshida, S., Mori, C., Takamizawa, A., Yosida, I. & Okayama, H. (1994).** Production of nonstructural proteins of hepatitis C virus requires a putative viral protease encoded by NS3. *Virology* **198**, 636–644.
- Marshall, R. D. (1974).** The nature and metabolism of the carbohydrate–peptide linkages of glycoproteins. *Biochemical Society Symposia* **40**, 17–26.
- Matsuura, Y. & Miyamura, T. (1993).** The molecular biology of hepatitis C virus. *Seminars in Virology* **4**, 297–304.
- Matsuura, Y., Harada, S., Suzuki, R., Watanabe, Y., Inoue, Y., Saito, I. & Miyamura, T. (1992).** Expression of processed envelope protein of hepatitis C virus in mammalian and insect cells. *Journal of Virology* **66**, 1425–1431.
- Mellquist, J. L., Kasturi, L., Spitalnik, S. L. & Shakin-Eshleman, S. H. (1998).** The amino acid following an Asn-X-Ser/Thr sequon is an important determinant of N-linked core glycosylation efficiency. *Biochemistry* **37**, 6833–6837.
- Michalak, J.-P., Wychowski, C., Choukhi, A., Meunier, J.-C., Ung, S., Rice, C. M. & Dubuisson, J. (1997).** Characterization of truncated forms of hepatitis C virus glycoproteins. *Journal of General Virology* **78**, 2299–2306.
- Miller, R. H. & Purcell, R. H. (1990).** Hepatitis C virus shares amino acid sequence similarity with pestiviruses and flaviviruses as well as members of two plant virus supergroups. *Proceedings of the National Academy of Sciences, USA* **87**, 2057–2061.
- Mizushima, H., Hijikata, M., Asabe, S. I., Hirota, M., Kimura, K. & Shimotohno, K. (1994).** Two hepatitis C virus glycoprotein E2 products with different C termini. *Journal of Virology* **68**, 6215–6222.
- Nakai, K. & Kanehisa, M. (1988).** Prediction of in vivo modification sites of proteins from their primary structures. *Journal of Biochemistry* **104**, 693–699.
- Opendakker, G., Rudd, P. M., Ponting, C. P. & Dwek, R. A. (1993).** Concepts and principles of glycobiology. *FASEB Journal* **7**, 1330–1337.
- Pohl, G., Kallstrom, M., Bergsdorf, N., Wallen, P. & Jornvall, H. (1984).** Tissue plasminogen activator: peptide analyses confirm an indirectly derived amino acid sequence, identify the active site serine residue, establish glycosylation sites, and localize variant differences. *Biochemistry* **23**, 3701–3707.
- Purcell, R. H. (1994).** Hepatitis C virus: historical perspective and current concepts. *FEMS Microbiology Review* **14**, 181–191.
- Rademacher, T. G., Parekh, R. B. & Dwek, R. A. (1988).** Glycobiology. *Annual Review of Biochemistry* **57**, 786–830.
- Ralston, R., Thudium, K., Berger, K., Kuo, C., Gervase, B., Hall, J., Selby, M., Kuo, G., Houghton, M. & Choo, Q. L. (1993).** Characterization of hepatitis C virus envelope glycoprotein complexes expressed by recombinant vaccinia viruses. *Journal of Virology* **67**, 6753–6761.
- Rice, C. M. (1996).** Flaviviridae: the viruses and their replication. In *Fields Virology*, 3rd edn, pp. 931–959. Edited by B. N. Fields, D. M. Knipe & P. M. Howley. Philadelphia: Lippincott–Raven.
- Saito, I., Miyamura, T., Ohbayashi, A., Harada, H., Katayama, T., Kikuchi, S., Watanabe, Y., Koi, S., Onji, M., Ohta, Y., Choo, Q. L., Houghton, M. & Kuo, G. (1990).** Hepatitis C virus infection is associated with the development of hepatocellular carcinoma. *Proceedings of the National Academy of Sciences, USA* **87**, 6547–6549.
- Sambrook, J., Fritsch, E. F. & Maniatis, T. (1989).** *Molecular Cloning: A Laboratory Manual*, 2nd edn. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory.
- Santolini, E., Migliaccio, G. & Lamonica, N. (1994).** Biosynthesis and biochemical properties of the hepatitis C virus core protein. *Journal of Virology* **68**, 3631–3641.
- Shakin-Eshleman, S. H., Remaley, A. T., Eshleman, J. R., Wunner, W. H. & Spitalnik, S. L. (1992).** N-linked glycosylation of rabies virus

glycoprotein. Individual sequons differ in their glycosylation efficiencies and influence on cell surface expression. *Journal of Biological Chemistry* **267**, 10690–10698.

Shakin-Eshleman, S. H., Spitalnik, S. L. & Kasturi, L. (1996). The amino acid at the X position of an Asn-X-Ser sequon is an important determinant of N-linked core glycosylation efficiency. *Journal of Biological Chemistry* **271**, 6363–6366.

Simmonds, P., Holmes, E. C., Cha, T.-A., Chan, S.-W., McOmish, F., Irvine, B., Beall, E., Yap, P. L., Kolberg, J. & Urdea, M. S. (1993). Classification of hepatitis C virus into six major genotypes and a series of subtypes by phylogenetic analysis of the NS-5 region. *Journal of General Virology* **74**, 2391–2399.

Spaete, R. R., Alexander, D., Rugroden, M. E., Choo, Q. L., Berger, K., Crawford, K., Kuo, C., Leng, S., Lee, C., Ralston, R., Thudium, K., Tung, J. W., Kuo, G. & Houghton, M. (1992). Characterization of the hepatitis E2/NS1 gene product expressed in mammalian cells. *Virology* **188**, 819–830.

Taylor, J. W., Ott, J. & Eckstein, F. (1985). The rapid generation of oligonucleotide-directed mutations at high frequency using phosphorothioate-modified DNA. *Nucleic Acids Research* **13**, 8764–8785.

Tomei, L., Failla, C., Santolini, E., Defrancesco, R. & Lamonica, N. (1993). NS3 is a serine protease required for processing of hepatitis C virus polyprotein. *Journal of Virology* **67**, 4017–4026.

Ware, F. E., Vassilakos, A., Peterson, P. A., Jackson, M. R., Lehrman, M. A. & Williams, D. B. (1995). The molecular chaperone calnexin binds Glc1Man9GlcNAc2 oligosaccharide as an initial step in recognizing unfolded glycoproteins. *Journal of Biological Chemistry* **270**, 4697–4704.

Zhang, W. H., Liu, C. B., Sun, Y. D., Alter, H. J. & Shih, J. W. (1990). Hepatitis C virus causing non-A, non-B hepatitis in plasmapheresis center. *Lancet* **335**, 353.

Received 22 October 1998; Accepted 24 December 1998