JOURNAL OF VIROLOGY, Aug. 1994, p. 5063–5073
0022-538X/94/$04.00+0

Vol. 68, No. 8

Processing in the Hepatitis C Virus E2-NS2 Region: Identification of p7 and Two Distinct E2-Specific Products with Different C Termini

CHAO LIN,[1] BRETT D. LINDENBACH,[1] BÉLA M. PRÁGAI,[1] DAVID W. McCOURT,[2] AND CHARLES M. RICE[1]*

Department of Molecular Microbiology[1] and Howard Hughes Medical Institute,[2] Washington University School of Medicine, St. Louis, Missouri 63110-1093

Received 14 February 1994/Accepted 9 May 1994

The hepatitis C virus (HCV) H strain polyprotein is cleaved to produce at least nine distinct products: NH_2-C-E1-E2-NS2-NS3-NS4A-NS4B-NS5A-NS5B-COOH. In this report, a series of C-terminal truncations and fusion with a human c-*myc* epitope tag allowed identification of a tenth HCV-encoded cleavage product, p7, which is located between the E2 and NS2 proteins. As determined by N-terminal sequence analysis, p7 begins with position 747 of the HCV H strain polyprotein. p7 is preceded by a hydrophobic sequence at the C terminus of E2 which may direct its translocation into the endoplasmic reticulum, allowing cleavage at the E2/p7 site by host signal peptidase. This hypothesis is supported by the observation that cleavage at the E2/p7 and p7/NS2 sites in cell-free translation studies was dependent upon the addition of microsomal membranes. However, unlike typical cotranslational signal peptidase cleavages, pulse-chase experiments indicate that cleavage at the E2/p7 site is incomplete, leading to the production of two E2-specific species, E2 and E2-p7. Possible roles of p7 and E2-p7 in the HCV life cycle are discussed.

Hepatitis C viruses (HCV) have recently been recognized as agents of the parentally transmitted form of non-A, non-B hepatitis (10, 40). HCV causes the vast majority of transfusion-associated cases of hepatitis and a significant proportion of community-acquired hepatitis worldwide (reviewed in references 34 and 36). HCV infection results in various clinical outcomes, including acute hepatitis, chronic hepatitis, cirrhosis, and establishment of an asymptomatic carrier state which may persist for life (reviewed in reference 34). Chronic infections are frequent, perhaps universal, and have been associated with increased incidence of hepatocellular carcinoma (13, 62).

The genome structures of several HCV isolates have been elucidated (9, 11, 14, 30, 38, 39, 52, 54, 55, 66, 67), indicating the existence of several genotypes (6, 65). This group of closely related enveloped positive-strand viruses is now classified as a separate genus in the flavivirus family (22), which includes two other genera, the flaviviruses (7) and pestiviruses (12). The HCV genome RNA is approximately 9.4 kb in length and contains a highly conserved 5′ noncoding region (29) followed by a long open reading frame encoding a polyprotein of 3,010 to 3,033 amino acids.

As established by cell-free translation and cell culture expression studies (2, 16, 25–27, 31–33, 47, 64, 68), the HCV polyprotein is processed by cellular and viral proteinases to produce the putative viral structural and nonstructural (NS) proteins. The order and nomenclature of the cleavage products are as follows: NH_2-C-E1-E2-NS2-NS3-NS4A-NS4B-NS5A-NS5B-COOH. C, a basic protein, is believed to be the viral capsid protein; E1 and E2 are probable virion envelope glycoproteins; NS2 through NS5B are putative NS proteins, at least some of which may be the functional equivalents of homologous proteins encoded by flaviviruses and pestiviruses (reviewed in references 36 and 50). Previous studies indicate that the host signal peptidase of the endoplasmic reticulum (ER) catalyzes cleavages in the putative structural region (C/E1, E1/E2, and perhaps E2/NS2) (31), whereas an HCV-encoded serine proteinase located in the N-terminal one-third of the NS3 protein is responsible for four cleavages in the NS region (3/4A, 4A/4B, 4B/5A, and 5A/5B) (2, 16, 25, 32, 47, 68). Cleavage at the 2/3 site is mediated by a second HCV-encoded proteinase which encompasses the NS2 region and the NS3 serine proteinase domain (26, 32).

Although the locations of most HCV-encoded proteins and the corresponding cleavage sites have been determined, the region between structural and NS coding sequences is still poorly defined. In a previous report, we observed at least two distinct forms of N-deglycosylated E2-specific proteins and suggested that an additional cleavage product might be present between the E2 and NS2 proteins (27). In this study, we have identified this protein, called p7, by expression of a series of C-terminally truncated polyproteins and by fusion to a human c-*myc* epitope tag, which allowed isolation of the cleavage product and N-terminal sequence determination. p7 mapped between E2 and NS2. The presence of hydrophobic potential signal/anchor sequences preceding the E2/p7 and p7/NS2 cleavage sites and the results of cell-free translation analyses indicate that host signal peptidase may catalyze both of these cleavages. However, cleavage at the E2/p7 site is incomplete, leading to the production of two stable E2-specific proteins with different C termini, E2 and E2-p7.

MATERIALS AND METHODS

Plasmid constructions. Standard recombinant DNA techniques (63) were used for construction of the expression plasmids described below. Constructs for expression of polyproteins with serial C-terminal truncations in the E2-NS2 region were produced from pTM3/HCV1-966, which encodes

* Corresponding author. Mailing address: Department of Molecular Microbiology, Washington University, School of Medicine, Box 8230, 660 S. Euclid Ave., St. Louis, MO 63110-1093. Fax: (314) 362-1232. Electronic mail address: rice@borcim.wustl.edu.

TABLE 1. HCV expression constructs with C-terminal truncations in the E2-NS2 region

Construct[a]	HCV site[b]	Vector site[c]	C-terminal residues[d]
pTM3/HCV1-836	*Eco*47III (2848)	*Stu*I	1
pTM3/HCV1-786	*Bsr*BI (2695)	*Stu*I	0
pTM3/HCV1-762	***Eam*1105I** (2624)	*Stu*I	1
pTM3/HCV1-729	***Drd*I** (2523)	*Stu*I	3
pTM3/HCV1-700	*Sca*I (2440)	*Stu*I	1
pTM3/HCV1-660	***Sst*I** (2322)	*Stu*I	1

[a] Numbers after HCV refer to the portion of the HCV polyprotein encoded by each construct. Flanking residues present in the polyproteins are not included. For all of the constructs, three additional N-terminal residues (Met-Cys-Thr) are predicted to be present prior to the Met residue initiating the HCV-H polyprotein (27).

[b] Restriction sites in the HCV cDNA used for the plasmid constructs. Nucleotide numbers given in parentheses refer to the positions of these sites in the full-length HCV-H sequence (14), assuming that the 5′ noncoding regions of HCV-H and HCV-1 (11, 29) are the same length. Restriction sites in boldface indicate that protruding ends were treated with T4 DNA polymerase prior to ligation to produce blunt ends.

[c] Restriction sites in plasmid vectors used for cloning.

[d] Number of predicted non-HCV C-terminal residues prior to the first termination codon.

HCV H strain (HCV-H) amino acid residues 1 to 966, by the subcloning strategies summarized in Table 1.

Construction of pSINrep5/HCV171-379-myc and pSINrep5/HCV370-802-myc will be described in detail elsewhere (44). Briefly, a derivative of the Sindbis virus replicon expression construct pSINrep5 (3), called pSINrep5/c-myc, was assembled; it contained, in the polycloning site region, a sequence encoding an *Xba*I recognition site followed by an *Mlu*I site, the coding sequence of the c-*myc* epitope recognized by monoclonal antibody (MAb) Myc1-9E10 (EQKLISEEDL) (18, 19), and a UAA termination codon. Selected regions of the HCV coding region (as indicated by the HCV polyprotein residue numbers) were amplified by PCR to contain a 5′ *Xba*I site, an AUG start codon, and a 3′ *Mlu*I site, allowing in-frame fusion with the c-*myc* epitope tag. PCR products were digested with *Xba*I and *Mlu*I and cloned into pSINrep5/c-myc.

pTM3/HCV364-802-myc was constructed by subcloning the *Sac*I-*Xho*I fragment (847 bp) of pSINrep5/HCV370-802-myc and the *Nco*I-*Sac*I fragment (893 bp) of pTM3/HCV1-1488 (27) into pTM3 (51), which had been digested with *Nco*I and *Xho*I. pTM3/HCV364-1207 was produced by subcloning the *Nco*I-*Nhe*I fragment (1,681 bp) of pTM3/HCV1-1488 into pTM3/HCV827-1207 (26), which had been digested with *Nco*I and *Nhe*I.

Cell cultures. The BHK-21 and CV-1 cell lines were obtained from the American Type Culture Collection, and the BSC-40 cell line (5) was obtained from D. Hruby (Oregon State University). Cell monolayers were grown in Eagle's minimal essential medium (MEM) supplemented with 2 mM L-glutamine, nonessential amino acids, penicillin, streptomycin, and 10% fetal bovine serum (FBS). The A16 subclone of the human hepatoma HepG2 cell line, generously provided by Alan Schwartz (Washington University, St. Louis, Mo.), was maintained in Dulbecco's modified Eagle medium supplemented with penicillin, streptomycin, and 10% FBS.

Generation and growth of viruses. A vaccinia virus-T7 expression plasmid containing the entire HCV-H open reading frame, pBRTM/HCV1-3011, has been described previously (27). The corresponding vaccinia virus-HCV recombinant, vHCV1-3011, was generated by marker rescue on CV-1 cells (46) and identified by the *gpt* selection method (20). Recombinant viruses were plaque purified three times under selective conditions prior to growth of large-scale stocks. Stocks of vHCV1-3011, vHCV1-1488 (27), and vTF7-3, a vaccinia virus recombinant expressing the T7 DNA-dependent RNA polymerase (24), were grown in BSC-40 monolayers and partially purified (37), and titers of infectious progeny were determined by plaque assay on BSC-40 cells (37).

SINrep5/HCV-myc recombinant viruses were generated by using 5′-capped SP6 RNA transcripts from pSINrep5/HCV-myc and pDH(26S)5′SIN helper templates essentially as described previously (3). The titers of these recombinants were determined by an infection assay using secondary cultures of chicken embryo fibroblasts (3, 23). Briefly, cells were seeded in 35-mm-diameter tissue culture wells to produce subconfluent monolayers and were infected with serial dilutions of the SINrep5/HCV-myc recombinants (56). The titer, expressed in infectious units, was approximated by determining the fraction of cells showing cytopathic effects by 24 h postinfection.

Transient expression using the vaccinia virus-T7 hybrid system. For expression assays utilizing vaccinia virus-HCV recombinants, monolayers of BHK-21 or HepG2-A16 cells were infected with vTF7-3 alone or in combination with vHCV1-1488 or vHCV1-3011. The multiplicity of infection for each recombinant was 10 PFU per cell (as determined on BSC-40 monolayers). After adsorption for 60 min at room temperature, the inoculum was removed and replaced with MEM containing 2% FBS. For pulse-chase experiments, monolayers were washed once with prewarmed MEM lacking methionine at 3 h postinfection and then incubated in the same medium for 20 min at 37°C. Cells were labeled by incubation for 20 min at 37°C with MEM lacking methionine and supplemented with 100 μCi of ^{35}S-protein labeling mixture (NEN) per ml. For chase experiments, the labeling mixture was replaced with MEM containing 2% FBS, 1.5 mg of methionine per ml, and 100 μg of cycloheximide per ml and incubated for the indicated periods at 37°C. For steady-state labeling, monolayers were washed once with prewarmed MEM lacking methionine at 3 h postinfection and then labeled by incubation for 4 h at 37°C with MEM containing 1/40 the normal concentration of methionine, 2% FBS, and 25 μCi of ^{35}S-protein labeling mixture (NEN) per ml. Expression assays of transfected plasmid constructs utilized subconfluent monolayers of BHK-21 cells in 35-mm-diameter dishes which had been previously infected with vTF7-3 as described above. After removal of the inoculum, cells were transfected at 37°C with a mixture consisting of 1 μg of plasmid DNA and 10 μg of Lipofectin (Bethesda Research Laboratories) in 0.5 ml of MEM. After 2 h, the transfection mixture was removed and the cells were labeled as described above.

Transient expression of HCV proteins with SINrep5/HCV-myc recombinants. BHK-21 cells were infected with SINrep5/HCV-myc recombinants at a multiplicity of infection of 5 infectious units per cell. After 60 min at 37°C, the inoculum was removed and replaced with MEM containing 2% FBS. At 3 h postinfection, monolayers were labeled as described above.

Cell lysis, immunoprecipitation, and protein analysis. After labeling, cell monolayers were washed with phosphate-buffered saline and lysed with a solution of 0.5% sodium dodecyl sulfate (SDS), 50 mM Tris-Cl (pH 7.4), 1 mM EDTA, and 20 μg of phenylmethylsulfonyl fluoride per ml (0.3 ml/10^6 cells), and cellular DNA was sheared by repeated passage through a 27.5-gauge needle. If the lysates were not used immediately, aliquots were stored frozen at −70°C. Samples were heated for 10 min at 70°C prior to dilution in the immunoprecipitation buffer containing Triton X-100 and carrier bovine serum albumin (59) and were clarified by centrifugation at 16,000 ×

g for 15 min. Portions of each lysate were incubated either with the indicated rabbit polyclonal antiserum (usually 5 μl) or with mouse MAb (usually 1 μl of ascitic fluid) followed by rabbit antiserum raised against whole mouse immunoglobulin G (usually 14 μg) (Sigma). Immune complexes were collected by using *Staphylococcus aureus* Cowan I (Calbiochem) as described previously (59).

Immunoprecipitates were solubilized and analyzed by SDS-polyacrylamide gel electrophoresis (PAGE) (41). After treatment for fluorography with Enhance (Dupont), gels were dried and exposed at −70°C with prefogged (43) X-ray film (Kodak). ^{14}C-methylated molecular mass marker proteins were purchased from Amersham.

In a previous report, endoglycosidase F (endo F) was utilized for characterization of the deglycosylated forms of HCV glycoproteins (27). However, this material was unsuitable for N-terminal sequence analysis, possibly because of proteolysis or N-terminal modification during overnight endo F digestions. In this study, we used peptide *N*-glycosidase F (PNGaseF), which cleaves between the innermost acetylglycoamine and asparaginyl residues of high-mannose, hybrid, and complex oligosaccharides from N-linked glycoproteins. PNGaseF digestions were carried out essentially according to the manufacturer's instructions (New England Biolabs). Briefly, immunoprecipitates were resuspended in 50 mM sodium phosphate (pH 7.5) containing 0.5% SDS and 1% 2-mercaptoethanol and were heated for 10 min at 99°C. The clarified supernatants were adjusted to 50 mM sodium phosphate (pH 7.5) containing 1% Nonidet P-40, 10 mM EDTA, and 100 μg of phenylmethylsulfonyl fluoride per ml and were incubated for 60 min at 37°C in the presence or absence of PNGaseF. Usually, 1,000 U (New England Biolabs) of PNGaseF was sufficient for complete digestion of samples from 1.5×10^5 cells. In the presence of 1% Nonidet P-40, up to 0.3% SDS had no significant inhibitory effect on PNGaseF digestion. Digested samples were mixed with an equal volume of 2× Laemmli sample buffer (41) and analyzed by SDS-PAGE as described above.

N-terminal sequence analysis. HepG2-A16 cells were coinfected with vHCV1-1488 and vTF7-3 as described above. At 3 h postinfection, monolayers were washed twice with prewarmed MEM lacking the amino acid used for radiolabeling and labeled for 4 h in the same medium containing 2% FBS and 200 μCi of the indicated ^{3}H-labeled amino acid per ml. The ^{3}H-amino acids used were leucine (135 Ci/mmol), valine (30 Ci/mmol), and threonine (17.7 Ci/mmol) (Amersham). After labeling, cell lysates were prepared and immunoprecipitated with either an E1-specific mouse MAb (A4; kindly provided by H. Hsu and H. Greenberg, Stanford University) or an E2-specific rabbit polyclonal antiserum (WU 105) (27) as described above. Radiolabeled E1- and E2-specific proteins were solubilized and digested with PNGaseF. The digested products were separated by SDS-PAGE, transferred to Immobilon polyvinylidene difluoride membranes (48), and localized by autoradiography. Partial amino acid sequence analyses of E1- and E2-specific proteins were performed as described previously (8).

BHK-21 cells were infected with SINrep5/HCV-myc recombinants with a multiplicity of infection of 5 infectious units per cell. At 3 h postinfection, monolayers were washed twice with prewarmed MEM lacking leucine and labeled for 4 h at 37°C in MEM lacking leucine and containing 2% FBS and 200 μCi of [^{3}H]leucine per ml. After labeling, monolayers were lysed with 0.5% SDS lysis buffer (as described above) and the proteins were immunoprecipitated with MAb Myc1-9E10 (19), specific for a c-*myc* epitope tag (18). Isolation and partial amino acid sequence analysis of [^{3}H]leucine-labeled c-*myc*-specific proteins were performed as described above.

Cell-free translation. Uncapped RNA transcripts were synthesized from linearized cDNA templates with T7 DNA-dependent RNA polymerase (Epicenter) (58). Cell-free translations, using rabbit reticulocyte lysates (Promega) and [^{35}S]methionine (Amersham), were incubated for 1 h at 30°C essentially according to the manufacturer's instructions. As indicated, canine pancreatic microsomal membranes (Promega) were added to 3.6 eq per 25-μl translation reaction. Translation reactions were terminated by the addition of RNase A to 10 μg per ml (Boehringer Mannheim) followed by continued incubation for 10 min at 30°C. Translation reactions were diluted 10-fold with the SDS lysis buffer described above and heated for 10 min at 70°C prior to immunoprecipitation. Analysis of solubilized proteins and/or PNGaseF digestion products was carried out essentially as described above.

RESULTS

Alternative forms of HCV E2 share the same N termini. In an earlier report, a vaccinia virus transient expression system was used to examine processing events in the HCV-H strain structural-NS2 region (27) (Fig. 1A). These experiments identified the 21-kDa C protein and the E1 protein, gp31, which was heavily modified by N-linked glycosylation and migrated to 21 kDa after endo F treatment (17). Glycosylated E2-specific proteins gp88 and gp70 were resolved as three species of 62, 41, and 36 kDa after digestion with endo F (reference 27 and also below). p62 (the deglycosylated form of gp88) was identified as an E2-NS2 polyprotein and proposed as a potential precursor for the E2 and NS2 proteins. Neither p41 nor p36 was recognized by E1 or NS2 region-specific antiserum, suggesting the existence of at least two alternative forms of E2 (27). To further examine the structural basis for these alternative forms of E2, we determined the N-terminal sequences of p41 and p36. Metabolically labeled lysates were prepared from HepG2-A16 cells which had been coinfected with vTF7-3 and vHCV1-1488, a vaccinia virus recombinant expressing the entire HCV structural region through the N-terminal two-thirds of NS3. Samples were immunoprecipitated with E2- or E1-specific antiserum, digested with PNGaseF, separated by SDS-PAGE, electroblotted onto Immobilon polyvinylidene difluoride membranes, and sequenced for 15 to 18 cycles of Edman degradation. For both p41 and p36, threonine residues were recovered at positions 2 and 5 (Fig. 2). This finding, along with the sizes and immunoreactivities of these species, identifies Glu-384 as the N-terminal residue of both forms of HCV-H E2. Identical N-terminal sequences were obtained for unresolved glycosylated forms of E2, produced by cell-free translation in the presence of microsomal membranes (HCV-J strain) (31) or purified from cells infected with a vaccinia virus recombinant (HCV-1 strain) (57). For the deglycosylated E1 protein, a leucine residue was recovered at position 9, valine residues were recovered at positions 3 and 12, and a threonine residue was recovered at position 13 (Fig. 2). These results establish the N terminus of HCV-H E1 as Tyr-192, which is again consistent with the previously reported E1 sequence data for the HCV-J (31) and HCV-1 (57) strains.

p41 and p36 differ at their C termini. We mapped the C-terminal boundaries of p41 and p36 by expressing a series of polyproteins with progressive C-terminal truncations (diagrammed in Fig. 1B) in BHK-21 cells by using the vaccinia virus-T7 system. All of the polyproteins contained the entire C and E1 sequences (presumably residues 1 to 191 and 192 to 383, respectively). Since the N terminus of NS2 has been

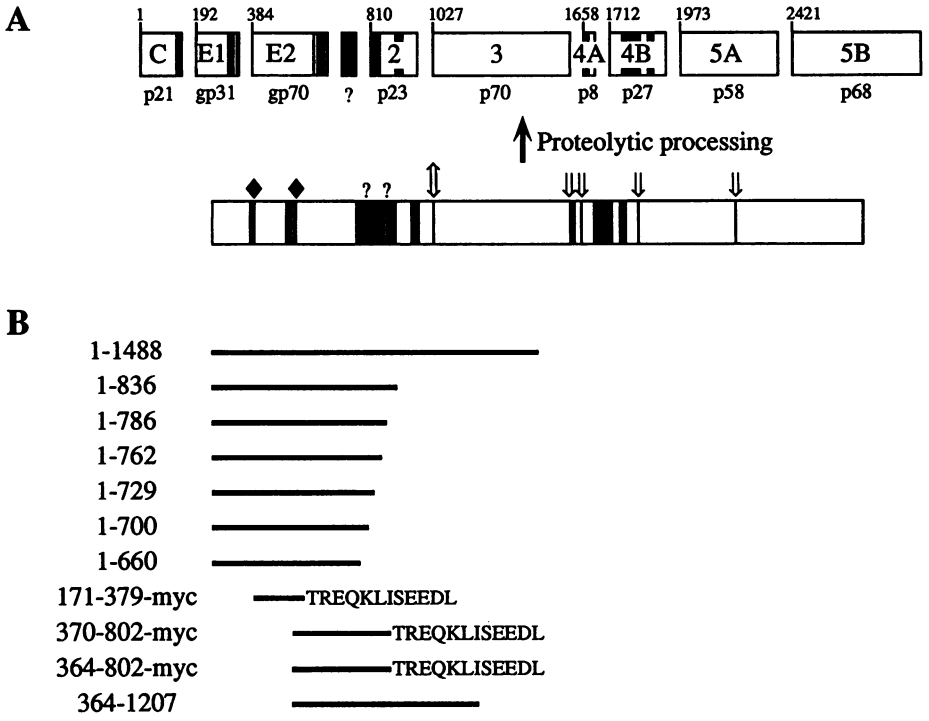


FIG. 1. HCV genome structure and expression constructs. (A) Diagram of the HCV-H polyprotein and its cleavage products shown as boxes. The identities of mature proteins, including putative structural proteins (C, E1, and E2) and NS proteins (NS2, NS3, NS4A, NS4B, NS5A, and NS5B), are indicated (27). The number at the top of each cleavage product indicates the position of its N-terminal residue in the polyprotein sequence. The apparent molecular mass of each HCV protein (p) and glycoprotein (gp) is indicated under the corresponding product (in kilodaltons). Regions containing predominantly uncharged amino acids are indicated as black bars. The region between E2 and NS2 which may represent an additional HCV-encoded protein is indicated (?). Also shown are putative cleavage sites for host signalase (◆) (31), the HCV NS2-3 proteinase (⇳) (26, 32), the NS3 serine proteinase (⇓) (2, 16, 25, 32, 47, 68), and unknown proteinases (?). (B) HCV polyprotein expression constructs used in this study. HCV polyprotein sequences present in each construct are indicated by black lines which are drawn to scale and oriented with respect to the diagram of the HCV polyprotein. Also shown is the C-terminal 10-residue c-*myc* epitope tag (EQKLISEEDL) present in some of the expressed polyproteins.

determined as residue 810 of the HCV-H polyprotein (26), polyprotein 1-836 includes the N-terminal 27 amino acids of NS2 while the remaining constructs terminate upstream of this cleavage site. Lysates of [^{35}S]methionine-labeled BHK-21 cells were immunoprecipitated with the E2-specific antiserum and digested with PNGaseF. As described previously, polyprotein 1-1488, which extends through the N-terminal two-thirds of the NS3 region, produced three E2-specific products: p62, p41, and p36 (Fig. 3). Polyproteins terminating at residues 836, 786, and 762 produced only two E2-specific proteins after PNGaseF digestion, one of which comigrated with p36 (Fig. 3). The second E2-specific product generated from polyprotein 1-836 migrated slightly more slowly than p41 and, therefore, may represent an uncleaved polyprotein containing p41 followed by the N-terminal 27 residues of NS2. The apparent molecular masses of truncated E2-specific proteins produced by polyproteins 1-786 and 1-762 were 39 and 38 kDa, respectively. These results suggest that the C terminus of p41 maps between residues 786 and 836, most likely at residue 809. Polyproteins terminating at residues 729, 700, and 660 generated only one E2-specific product of 35, 34, and 32 kDa, respectively (Fig. 3), which indicates that the C terminus of p36 is located between residues 729 and 762, probably near residue 745 (see below). On the basis of these data, we refer to the small, highly hydrophobic region between E2 and NS2 as p7 (although its existence as a stable cleavage product has not been verified). p62, p41, and p36 are hereafter called E2-NS2, E2-p7, and E2, respectively.

Determination of the E2/p7 cleavage site. Further characterization of the putative p7 cleavage product has proven difficult. To generate p7-specific antiserum, we subcloned the cDNA sequence encoding HCV residues 745 to 809 into *Escherichia coli* expression constructs and attempted to purify p7-containing proteins in sufficient quantities to generate rabbit polyclonal antiserum. Although numerous *E. coli* expression systems have been tried, this approach has not yet succeeded. We also screened antisera from HCV-positive patients by immunoprecipitation under either denaturing or nondenaturing conditions; none of these sera reacted with a distinguishable p7 protein. It is possible that the putative p7 cleavage product is unstable or poorly immunogenic because of the highly hydrophobic nature of this region.

Because of the lack of antiserum specific for the p7 region, a 10-amino-acid c-*myc* epitope tag was fused to the p7 C terminus to allow identification of this processing product (diagrammed in Fig. 1B). Construct HCV370-802-myc contains HCV residues 370 to 802, beginning in the putative signal peptide sequence at the C terminus of E1, followed by the E2 and p7 regions. As a control for the specificity of the c-*myc* MAb, we constructed HCV171-379-myc, which includes the putative signal peptide at the C terminus of the C protein followed by the E1 region and the c-*myc* peptide. For both

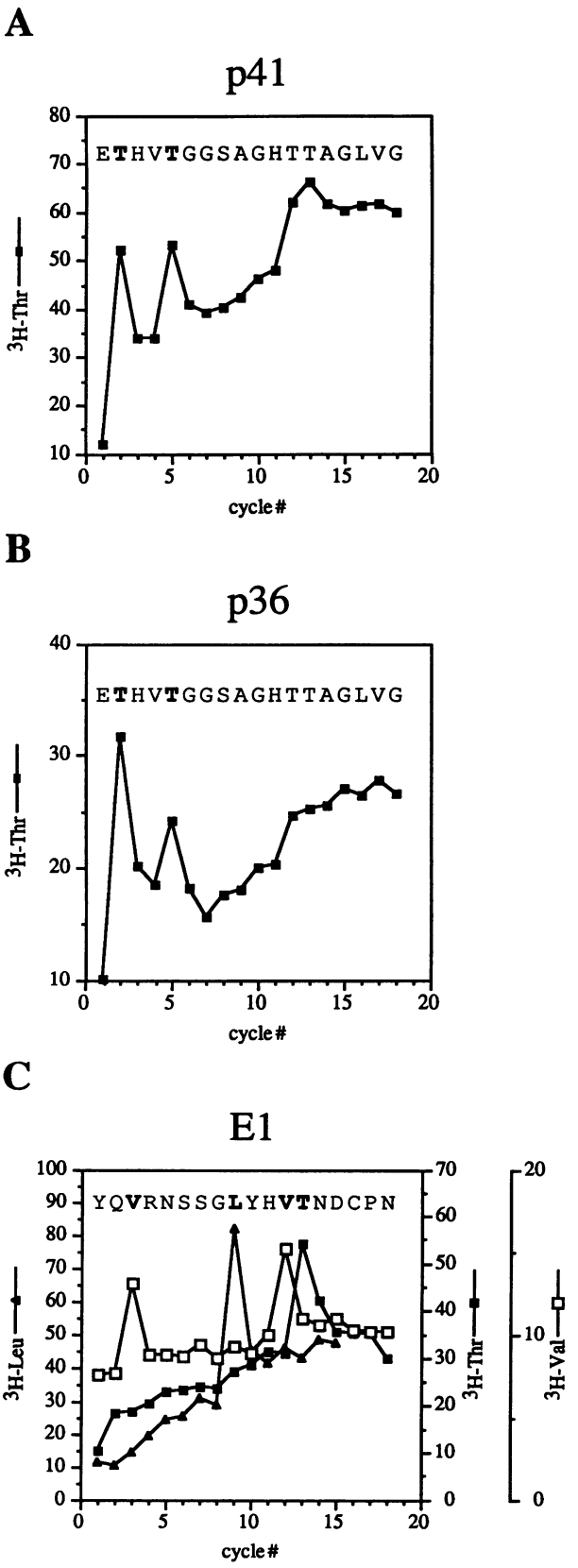


FIG. 2. N-terminal sequence analyses of the E2-specific proteins p41 (A) and p36 (B) and of E1 (C). HepG2-A16 cells were coinfected with vTF7-3 and vHCV1-1488 and labeled with the indicated ^{3}H-labeled amino acids as described in Materials and Methods. PNGaseF-digested E2-specific protein (p41 or p36) or E1 (p21) was isolated and subjected to N-terminal sequence analysis. The graphs show uncorrected counts per minute released per sequencing cycle. From these data, the sequences of p41, p36, and E1 begin at HCV-H residues 384, 384, and 192, respectively, and are shown at the top of the corresponding graph. Residues in boldface type indicate positions determined by N-terminal sequencing; the remaining residues are deduced from HCV-H nucleotide sequence data (14, 38).

constructs, the C-terminal c-*myc* peptide was fused upstream of the p7/NS2 (HCV370-802-myc) and E1/E2 (HCV171-379-myc) cleavage sites in a context which should discourage removal by signal peptidase (21, 69).

These two HCV-myc fusion cassettes were placed downstream of the 26S subgenomic promoter of a Sindbis virus-derived expression replicon, which is a useful system for rapid generation and analysis of expression constructs (3, 4). Infectious particles containing these Sindbis virus-HCV replicons were rescued (3) and used to infect BHK-21 cells, and [^{35}S]methionine-labeled cell lysates were immunoprecipitated with MAb Myc1-9E10. For HCV171-379-myc, at least five c-*myc*-specific species, ranging from 21 to 34 kDa, were immunoprecipitated (Fig. 4A). Similar patterns of E1-specific proteins have been observed when the E1 protein was expressed independently at high levels in insect cells and have been attributed to incomplete N-linked glycosylation (42, 49). For HCV370-802-myc, MAb Myc1-9E10 recognized at least two higher-molecular-mass species, which presumably represent glycosylated and unglycosylated E2-p7-myc and a protein of 6.5 kDa, whose size is consistent with the p7-myc cleavage product (Fig. 4A). [^{3}H]leucine-labeled p7-myc protein was electroblotted onto Immobilon polyvinylidene difluoride membranes, localized, and sequenced for 25 cycles of Edman degradation. As shown in Fig. 4B, leucine residues were recovered at positions 2, 5, 8, 13, 19, and 23, which tentatively establishes the N terminus of p7 as HCV-H polyprotein residue 747. Together with the definition of the p7/NS2 cleavage site (26), these data suggest that p7 is 63 amino acids in length, encompassing residues 747 to 809 of the HCV-H polyprotein, with a predicted molecular mass of 7.0 kDa. The smallest form of the E2 glycoprotein, whose deglycosylated form corresponds to p36, spans residues 384 to 746, with a polypeptide backbone of 363 residues and a predicted molecular mass of 40.0 kDa.

Cell-free processing at the E2/p7 cleavage site is dependent upon microsomal membranes. Previous cell-free translation studies have shown that processing at the HCV C/E1 and E1/E2 cleavage sites is dependent upon the addition of microsomal membranes and, therefore, is likely mediated by host signal peptidase in the ER (31). The sequences and the conserved hydrophobic character of the regions preceding the N termini of p7 and NS2 (Fig. 5) are characteristic of signal peptides (21, 69). This suggested that cleavage at the E2/p7 and p7/NS2 sites may also be mediated by host signal peptidase. This possibility was investigated by cell-free translation of two HCV expression constructs. HCV364-1207 initiates within the putative signal peptide at the C terminus of E1 and extends through the serine proteinase domain of NS3. HCV364-802-myc also begins with the signal sequence preceding E2 and terminates with the c-*myc* epitope fused in frame after p7 residue 802 (C terminus identical to that of HCV370-802-myc) (Fig. 1B).

The patterns of cell-free translation products from HCV364-802-myc are shown in Fig. 6A. In the absence of microsomal membranes, a predominant product of 41 kDa was recognized

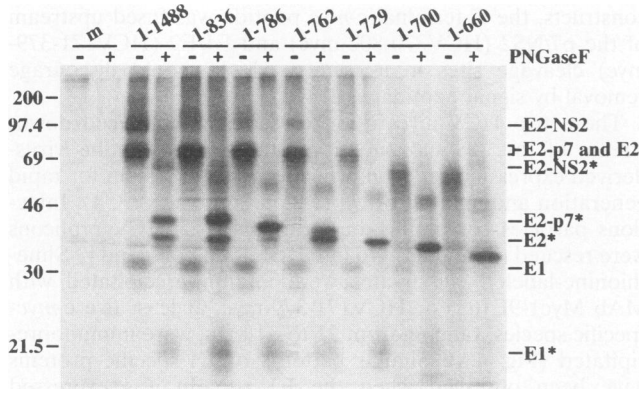


FIG. 3. C-terminal boundaries of E2-specific proteins p41 and p36. vTF7-3-infected BHK-21 cell monolayers were mock transfected (lanes m), transfected with the indicated plasmid DNAs, or coinfected with vHCV1-1488 and were labeled with ^{35}S-protein labeling mixture as described in Materials and Methods. Cell lysates were prepared and immunoprecipitated with anti-E2 antiserum WU 105 (27). Immunoprecipitated proteins were solubilized and digested in the absence (−) or presence (+) of PNGaseF. The digested products were separated by SDS-12% PAGE. HCV-specific proteins are identified at the right, and the sizes of ^{14}C-labeled protein molecular mass markers (in kilodaltons) are indicated at the left. The positions of PNGaseF-digested forms (indicated by asterisks) are also shown. As observed before, a small fraction of the E1 protein was coimmunoprecipitated with E2-specific antiserum (27).

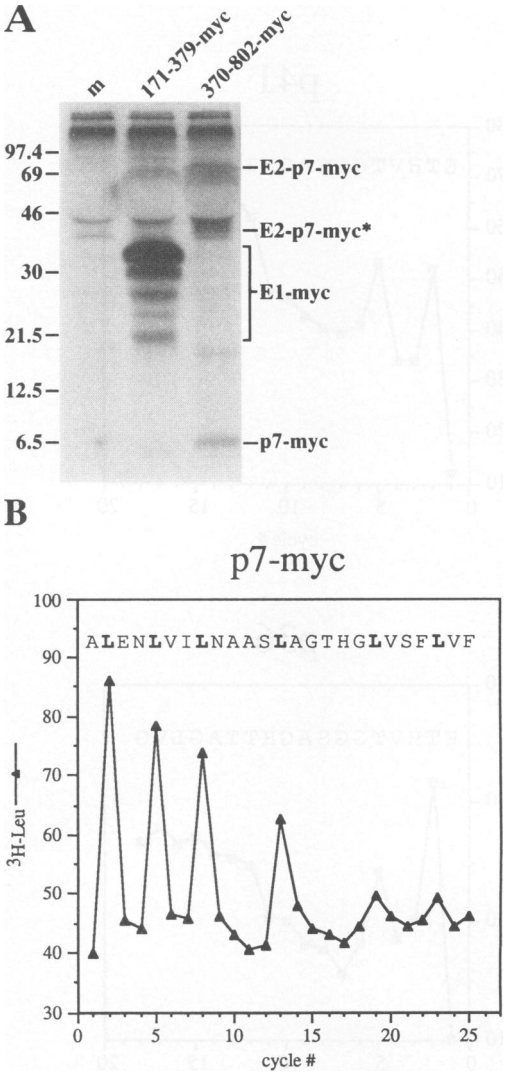


FIG. 4. Identification of the p7-myc protein. (A) BHK-21 cells were mock infected (lane m) or infected with the indicated SINrep5-HCV recombinants and then labeled with ^{35}S-protein labeling mixture as described in Materials and Methods. Cell lysates were prepared and immunoprecipitated with anti-c-*myc* MAb Myc1-9E10 (19). Immunoprecipitated proteins were solubilized and separated by SDS-14% PAGE. c-*myc*-specific proteins are identified at the right, and the sizes of ^{14}C-labeled protein molecular mass markers (in kilodaltons) are indicated at the left. The position of unglycosylated E2-p7-myc is also indicated by an asterisk. (B) N-terminal sequence analysis of [^{3}H]leucine-labeled p7-myc fusion protein. The graph shows uncorrected counts per minute released per sequencing cycle. From these data, the N-terminal sequence of p7 begins at HCV-H polyprotein residue 747 and is shown at the top of the graph. Leucine residues determined by N-terminal sequencing are in boldface type; the remaining residues are deduced from HCV-H nucleotide sequence data (14, 38).

by both anti-E2 and anti-c-*myc* antibodies and, therefore, represented the primary translation product. Upon the addition of microsomes, two glycoproteins were observed. The larger one (N-deglycosylated form, 41 kDa) was immunoprecipitated by both anti-E2 and anti-c-*myc* antibodies and identified as the E2-p7-myc glycoprotein. The smaller species (N-deglycosylated form, 36 kDa) was present in small amounts and was immunoprecipitated only with E2-specific antiserum, thus representing the processed E2 glycoprotein. Both E2-specific glycoproteins were fully protected against proteinase K digestion in the absence of detergent (data not shown), indicative of translocation into the ER lumen. In addition, a c-*myc*-specific protein of 6.5 kDa, presumably p7-myc, was observed only in the presence of microsomal membranes (data not shown).

As shown in Fig. 6B, microsomes were also required for processing of the longer polyprotein 364-1207 at the E2/p7 and p7/NS2 sites, as evidenced by the appearance of the E2 and E2-p7 glycoproteins and NS2. Both glycoproteins were fully protected against proteinase K digestion in the absence of detergent (data not shown). The E2/p7 cleavage was incomplete, in contrast to the efficient cleavage observed at the p7/NS2 site. The addition of microsomal membranes had no effect on the efficiency of cleavage at the 2/3 site in 364-1207, which is consistent with the membrane-independent autoproteolytic processing at this site by the HCV NS2-3 proteinase (26, 32).

Kinetics of processing in the structural-NS2 region. To identify possible precursors involved in the structural-NS2 region processing, pulse-chase experiments were carried out in HepG2-A16 cells with a vaccinia virus-HCV recombinant, vHCV1-3011. As shown in Fig. 7A, the E1 protein was readily apparent after a 20-min pulse and was not associated with any higher-molecular-weight polyprotein precursors, which is consistent with the proposed cotranslational cleavage at both C/E1 and E1/E2 sites by ER signal peptidase (31). No shift in deglycosylated E1 migration was observed during the chase period (Fig. 7A), suggesting that E1 does not undergo further proteolytic processing under these conditions. However, N-glycosylated forms of E1 did show a gradual increase in migration over the chase period (Fig. 7A), which probably results from trimming of the N-linked, high-mannose glycans. It should be noted that the level of E1 increased slightly during

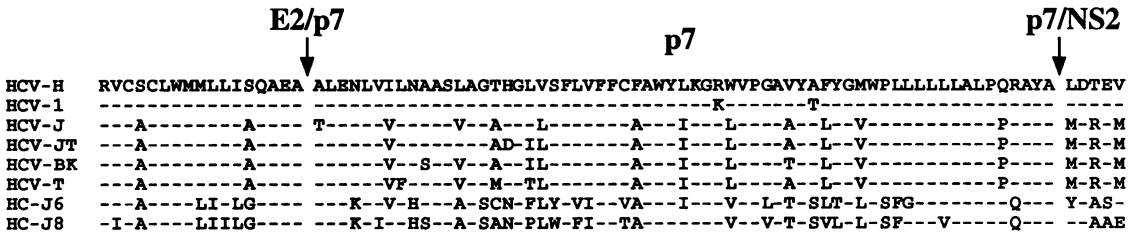
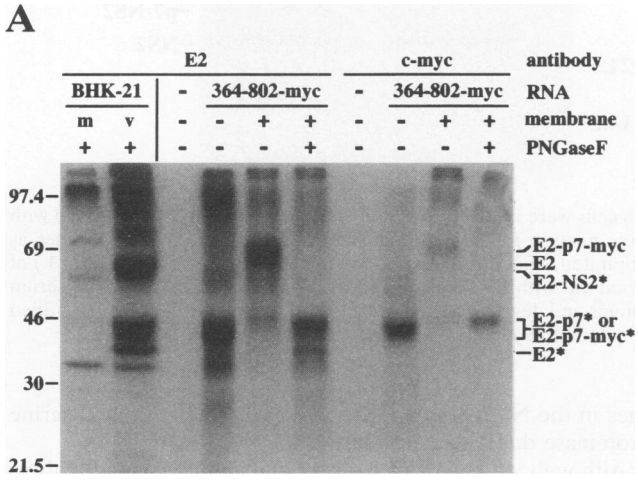


FIG. 5. Alignment of the HCV E2/p7 and p7/NS2 cleavage sites. The amino acid sequences around the E2/p7 and p7/NS2 cleavage sites are aligned for the following HCV isolates: HCV-H (14, 38), HCV-1 (11), HCV-J (39), HCV-JT (67), HCV-BK (66), HCV-T (9), HCV-J6 (55), and HCV-J8 (54). The single-letter code for amino acids is used. A hyphen indicates a residue identical to that of the HCV-H sequence.

the chase period (Fig. 7A), in apparent conflict with the absence of any E1-containing polyprotein precursors. This has been observed before and may be due to a delay in the formation of the epitope recognized by the anti-E1 MAb used for immunoprecipitation (15).

Processing in the E2-NS2 region was more complex. After a 20-min pulse, three E2-specific proteins, including E2, E2-p7, and E2-NS2, were apparent after PNGaseF digestion (Fig. 7B) while the anti-NS2 antiserum recognized two products, E2-NS2 and NS2 (Fig. 7C). These results indicate that cleavages at the E1/E2 and 2/3 sites occur rapidly, perhaps cotranslationally, and are followed by delayed processing within the E2-NS2 region (see Discussion). As seen in Fig. 7B and C, the level of E2-NS2 decreased concomitantly with an increase in the levels of E2 and NS2 during the chase. This provides evidence that E2-NS2 is a discrete precursor of E2 and NS2 (27), which has also been demonstrated recently for the HCV-BK strain (15). The precursor-product relationship between E2-p7 and E2, if any, is less clear, and the results so far do not distinguish between a kinetic mechanism and an alternative pathway for the inefficient cleavage at the E2/p7 site. Cleavage at the p7/NS2 site relative to that at the E2/p7 site was efficient as evidenced by the disappearance of E2-NS2 and the low levels of p7-NS2 detected during the chase. This observation, to-



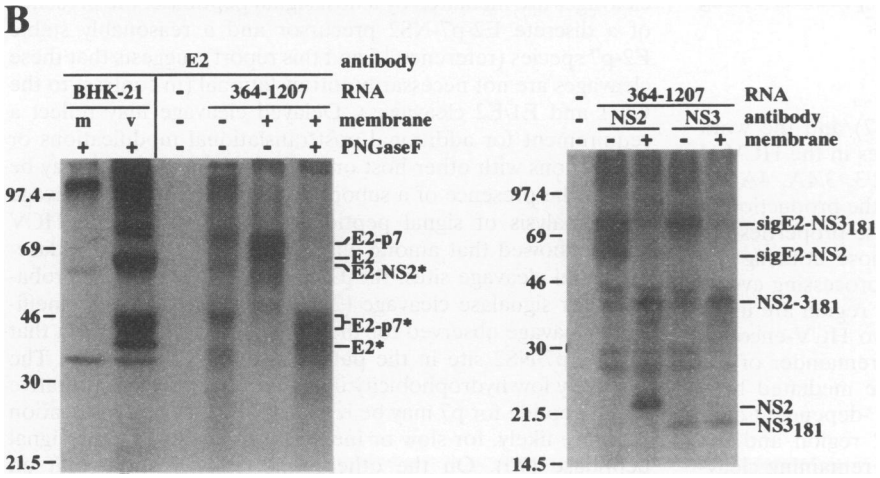


FIG. 6. In vitro processing of HCV polyproteins. Uncapped RNA transcripts were translated in reticulocyte lysates in the presence (+) or absence (−) of microsomal membranes as described in Materials and Methods. These transcripts encoded polyprotein 364-802-myc (A) and polyprotein 364-1207 (B). [^{35}S]methionine-labeled translation products were denatured and immunoprecipitated with the following antisera: anti-c-*myc* MAb (Myc1-9E10) (19), anti-E2 (WU 105), anti-NS2 (WU 107), or anti-NS3 (WU 110) (27). Also shown are control reactions without added RNA (−) and the HCV E2-specific proteins produced in BHK-21 cells by infection with vTF7-3 alone (lane m) or in combination with vHCV1-1488 (lane v). Some immunoprecipitated proteins were solubilized and digested in the presence (+) or absence (−) of PNGaseF. Immunoprecipitated and PNGaseF-digested products were separated by SDS-14% PAGE. HCV- or c-*myc*-specific proteins are identified at the right, and the sizes of ^{14}C-labeled protein molecular mass markers (in kilodaltons) are indicated at the left. The positions of the PNGaseF-digested forms (indicated by asterisks) are also shown.

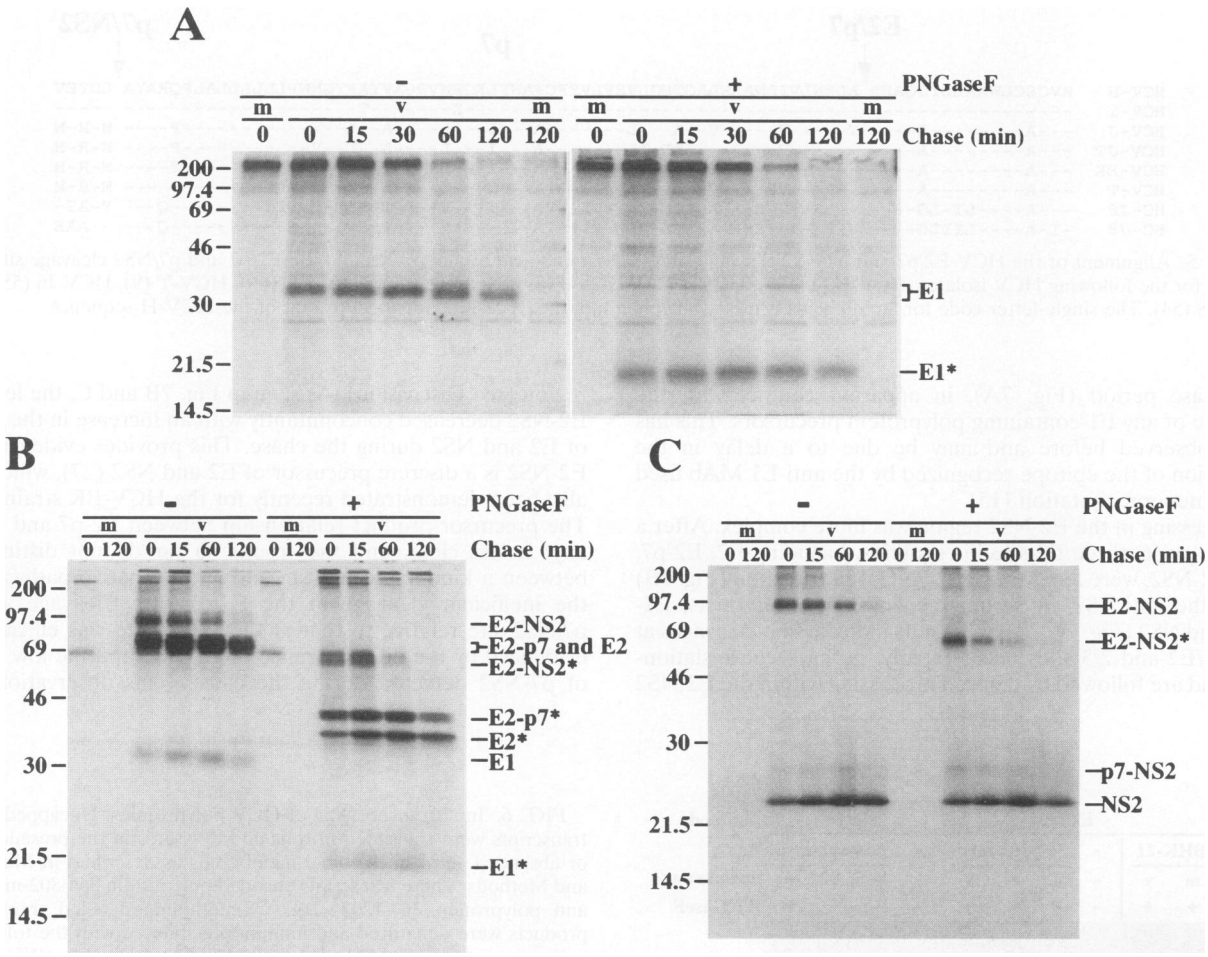


FIG. 7. Processing kinetics in the HCV structural-NS2 region. HepG2-A16 cells were infected with vTF7-3 alone (lanes m) or coinfected with vTF7-3 and vHCV1-3011 (lanes v) and pulse-labeled with ^{35}S-protein labeling mixture for 20 min and then chased for the indicated times as described in Materials and Methods. Cell lysates were prepared, immunoprecipitated, solubilized, digested in the absence (−) or presence (+) of PNGaseF, and separated by SDS-12% PAGE. Immunoprecipitation was carried out with E1-specific MAb A4 (A), E2-specific rabbit antiserum WU 105 (B), or NS2-specific rabbit antiserum WU 107 (C). HCV-specific proteins and deglycosylated forms (indicated by asterisks) are identified at the right, and the sizes of ^{14}C-labeled protein molecular mass markers (in kilodaltons) from Amersham are indicated at the left.

gether with the decrease in the ratio of E2-p7 to E2 during the chase, suggests that E2-p7 may be cleaved slowly to produce E2. However, further experiments are needed to substantiate this conclusion, given the instability of these products during long chases in the presence of cycloheximide.

DISCUSSION

Earlier transient expression studies (25–27) and the work reported here have defined nine cleavage sites in the HCV-H polyprotein: C/E1, E1/E2, E2/p7, p7/NS2, 2/3, 3/4A, 4A/4B, 4B/5A, and 5A/5B. These cleavages result in the production of at least 10 nonoverlapping polypeptides whose properties are summarized in Table 2. Our studies and those utilizing the HCV-J strain (31) suggest that the primary processing events in the putative structural region and the NS2 region are likely to be catalyzed by host signal peptidase. Two HCV-encoded proteinases are required for processing the remainder of the polyprotein. The 2/3 cleavage appears to be mediated by a novel autoproteolytic activity, perhaps a Zn^{2+}-dependent metalloproteinase, which encompasses the NS2 region and the NS3 serine proteinase domain (26, 32). The remaining cleavages in the NS region are dependent on an active NS3 serine proteinase domain (2, 16, 25, 32, 47, 68).

Although we believe it is likely that the E2/p7 and p7/NS2 cleavages are mediated by a host signal peptidase, the presence of a discrete E2-p7-NS2 precursor and a reasonably stable E2-p7 species (reference 15 and this report) suggests that these cleavages are not necessarily cotranslational (in contrast to the C/E1 and E1/E2 cleavages). Delayed cleavage may reflect a requirement for additional posttranslational modifications or interactions with other host or viral components, or it may be due to the presence of a suboptimal signal peptidase cleavage site. Analysis of signal peptide sequences of several HCV strains showed that among all four putative signal peptidase-mediated cleavage sites, the E2/p7 site has the lowest probability for signalase cleavage (21, 69). This fits with the inefficient cleavage observed for the E2/p7 site compared with that for the p7/NS2 site in the pulse-chase analyses (Fig. 7). The relatively low hydrophobicity in the core region of the putative signal peptide for p7 may be responsible for slow translocation or, more likely, for slow or inefficient processing by the signal peptidase (70). On the other hand, it is possible that an

TABLE 2. Properties of HCV proteins

Protein	Amino acid sequence of the HCV polyprotein[a]	Predicted polypeptide mol mass (kDa)[a]	Observed product mol mass (kDa)[b]	Similarity[c]	Function(s)[d]
C	1–191	20.8	21	1.454	Nucleocapsid protein (?)
E1	192–383	20.9	21 (gp31)	1.300	Virion envelope glycoprotein (?)
E2	384–746	40.0	36 (gp68-72)	1.300	Virion envelope glycoprotein (?)
E2-p7	384–809	47.0	41 (gp68-72)		(?)
p7	747–809	7.0	ND[e]	1.309	(?)
NS2	810–1026	24.0	23	1.273	NS2-3 proteinase component
NS3	1027–1657	67.3	70	1.409	NS2-3 proteinase component, serine proteinase, helicase
NS4A	1658–1711	5.8	8	1.391	Membrane-associated replicase component (?)
NS4B	1712–1972	27.2	27	1.373	Membrane-associated replicase component (?)
NS5A	1973–2420	49.1	58	1.236	Replicase component (?)
NS5B	2421–3011	65.4	68	1.373	RNA-dependent RNA polymerase (?)

[a] Based on the HCV-H strain (14) and the assumption that there is no trimming at the C termini of these proteins.
[b] For glycoproteins (E1, E2, and E2-p7), the molecular weights of N-deglycosylated forms are given, with the molecular weights of glycosylated forms (gp) in brackets.
[c] Similarity is the arithmetic mean of comparison scores from several full-length HCV polypeptide sequences, calculated by using the modified Dayhoff table (28) and averaged over the length of each protein. An average score of 1.5 represents complete conservation while a score of 0 indicates lack of similarity. Pairwise alignments of 19 full-length HCV polypeptide sequences were analyzed with the Genetics Computer Group suite of programs (version 7). These 19 sequences include HCV-1 (11), HCV-H (14, 38), HC-J1 (D10749), HC-G9 (53), HC-J4/83 and J4/91 (52), HCV-J (39), HCV-JT and JT' (67), HCV-BK (66), HCV-T (9), HCV-JK1 (S18030), HCV-N (30), HCV-Unkcds (M96362), HCV-L2 (U01214), HCV-HeBei (L02836), HC-J6 (55), and HC-J8 (54). For unpublished sequences, accession numbers are given in parentheses.
[d] (?), unknown or predicted function.
[e] ND, not determined.

alternative microsomal or viral proteinase is responsible for cleavage at this site.

This processing scheme for the HCV structural-NS2 region has features similar to as well as distinct from those exhibited by flaviviruses and pestiviruses. For members of the flavivirus genus, the gene order is NH_2-anchC-prM-E-NS1-NS2A-NS2B-NS3-NS4A-NS4B-NS5-COOH and the mature virion proteins include C, M, and E (see reference 60 for a review). Cleavages generating the N termini of prM, E, and NS1, a nonstructural glycoprotein, are mediated by host signal peptidase. prM is later processed to pr (the N-terminal part of prM) and M (the C-terminal fragment), presumably by a Golgi-associated proteinase, shortly before virus release. Although earlier models proposed that cleavage at a C-terminal dibasic site of anchC (by the virus-encoded NS2B-3 serine proteinase) was responsible for production of mature virion C protein, this view has recently been challenged by evidence which indicates that cleavage at the dibasic site may be a prerequisite for signal peptidase cleavage (1, 45, 72). For pestiviruses, the first protein of the open reading frame, p20, is an autoproteinase (71) responsible for p20/C cleavage to produce the N terminus of the nucleocapsid protein, C, which is followed by three virion envelope glycoproteins, E0, E1, and E2 (61). Cotranslational cleavage at the C/E0 site and the C terminus of E2 generate an E0-E1-E2 precursor. Processing at the E1/E2 site, although slightly delayed, is also thought to be mediated by host signal peptidase. The E0/E1 cleavage site, similar to the NS1/2A site of flaviviruses (7), has a sequence that fulfills the (−1, −3) rule for signal peptidase cleavage sites, but it lacks an upstream hydrophobic region. Proteinases of novel specificity have been invoked for both cleavages (35, 60, 61), although cleavage by signal peptidase has not been excluded. Thus, while signalase-mediated cleavages appear to play an important role for processing in the structural region for all three genera in the flavivirus family, delayed cleavages catalyzed by proteinases other than host signal peptidase, such as the prM cleavage (flaviviruses) and perhaps the E0/E1 cleavage (pestiviruses), have not been identified for HCV.

Assuming that the C terminus of HCV E2 (polyprotein residue 746) does not undergo additional trimming, the C-terminal region of the E2 protein is characterized by stretches of hydrophobic amino acid residues punctuated by occasional charged residues which are highly conserved. Although short, the best candidates for potential membrane-spanning segments appear to be residues 699 to 714 and 731 to 744. A similar arrangement of sequences is found near the predicted C terminus of the E1 protein (residue 383), where uncharged residues 347 to 369 are followed by a conserved lysine residue and then a second hydrophobic region consisting of residues 371 to 381. These elements are reminiscent of those found near the C termini of flavivirus virion proteins M and E which appear to function as signal/anchor sequences and may also play a role in mediating specific interactions important for virion assembly. The 63-residue p7 protein is composed mainly of uncharged polar or hydrophobic residues punctuated by a few conserved charged residues (Fig. 5). One model for the topology of p7, constrained by lumenal signal peptidase cleavages at the N and C termini, predicts that residues 764 to 778 and 782 to 803 constitute transmembrane segments which are separated by a few charged residues localized on the cytoplasmic side of the membrane. Whether p7 corresponds to a structural or NS protein remains to be determined.

At least for our clone of HCV-H, inefficient processing at the E2/p7 site leads to the production of two polypeptides containing E2 sequences, E2 and E2-p7. Although it is generally believed that the E2 glycoprotein is a structural protein, the existence of E2-p7 raises the possibility that this protein could also be a virion component or possibly a distinct cell-associated form of the E2 protein. However, although E2-p7 was found to be relatively stable in our experiments (reference 15 and this report), preliminary studies using constructs derived from a distinct HCV subtype, HCV-BK (66), suggest that processing at the E2/p7 site can be more efficient (data not shown). Further studies, with additional independent cDNA constructs, are required to determine if strain-specific differences in processing at this site exist. In any case, roles for

these proteins in HCV virion assembly or RNA replication will remain speculative until it becomes possible to examine virus-encoded proteins in authentic HCV-infected cells and virions and to assess their importance in the life cycle via genetic and biochemical studies.

ACKNOWLEDGMENTS

We thank Arash Grakoui for his participation in the early stages of this study; Ilya Frolov for help with the SIN replicon expression system; Henry Hsu and Harry Greenberg for HCV E1-specific MAb A4; JoAnn Suzich for c-*myc*-specific MAb Myc1-9E10; and Sean M. Amberg, Jean Dubuisson, and Alexander Kolykhalov for critical reading of the manuscript.

This work was supported by a grant from the Public Health Service (CA57973). C.L. is a predoctoral candidate and was supported in part by the Division of Biology and Biomedical Sciences at Washington University. B.D.L. is a graduate student in the Lucille P. Markey Special Emphasis Pathway in Human Pathobiology. B.M.P. was supported in part by the Ministry of Welfare (T076-ETT [90-93]), Hungary.

REFERENCES

1. **Amberg, S. M., A. Nestorowicz, D. W. McCourt, and C. M. Rice.** 1994. NS2B-3 proteinase-mediated processing in the yellow fever virus structural region: in vitro and in vivo studies. J. Virol. **68:**3794–3802.
2. **Bartenschlager, R., L. Ahlborn-Laake, J. Mous, and H. Jacobsen.** 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. J. Virol. **67:**3835–3844.
3. **Bredenbeek, P. J., I. Frolov, C. M. Rice, and S. Schlesinger.** 1993. Sindbis virus expression vectors: packaging of RNA replicons by using defective helper RNAs. J. Virol. **67:**6439–6446.
4. **Bredenbeek, P. J., and C. M. Rice.** 1992. Animal RNA virus expression systems. Semin. Virol. **3:**297–310.
5. **Brockman, W. W., and D. Nathans.** 1974. The isolation of simian virus 40 variants with specifically altered genomes. Proc. Natl. Acad. Sci. USA **71:**942–946.
6. **Bukh, J., R. H. Purcell, and R. H. Miller.** 1993. At least 12 genotypes of hepatitis C virus predicted by sequence analysis of the putative E1 gene of isolates collected worldwide. Proc. Natl. Acad. Sci. USA **90:**8234–8238.
7. **Chambers, T. J., C. S. Hahn, R. Galler, and C. M. Rice.** 1990. Flavivirus genome organization, expression, and replication. Annu. Rev. Microbiol. **44:**649–688.
8. **Chambers, T. J., D. W. McCourt, and C. M. Rice.** 1989. Yellow fever virus proteins NS2A, NS2B, and NS4B: identification and partial N-terminal amino acid sequence analysis. Virology **169:** 100–109.
9. **Chen, P.-J., M.-H. Lin, K.-F. Tai, P.-C. Liu, C.-J. Lin, and D.-S. Chen.** 1992. The Taiwanese hepatitis C virus genome: sequence determination and mapping the 5′ termini of viral genomic and antigenomic RNA. Virology **188:**102–113.
10. **Choo, Q.-L., G. Kuo, A. J. Weiner, L. R. Overby, D. W. Bradley, and M. Houghton.** 1989. Isolation of a cDNA clone derived from a blood-borne non-A, non-B viral hepatitis genome. Science **244:**359–362.
11. **Choo, Q.-L., K. H. Richman, J. H. Han, K. Berger, C. Lee, C. Dong, C. Gallegos, D. Coit, A. Medina-Selby, P. J. Barr, A. J. Weiner, D. W. Bradley, G. Kuo, and M. Houghton.** 1991. Genetic organization and diversity of the hepatitis C virus. Proc. Natl. Acad. Sci. USA **88:**2451–2455.
12. **Collett, M. S.** 1992. Molecular genetics of pestiviruses. Comp. Immunol. Microbiol. Infect. Dis. **15:**145–154.
13. **Colombo, M., G. Kuo, Q.-L. Choo, M. F. Donato, E. D. Ninno, M. A. Tommasini, N. Dioguardi, and M. Houghton.** 1989. Prevalence of antibodies to hepatitis C virus in Italian patients with hepatocellular carcinoma. Lancet **ii:**1006–1008.
14. **Daemer, R., C. Wychowski, A. Grakoui, C. M. Rice, and S. M. Feinstone.** Unpublished data.
15. **Dubuisson, J., H. H. Hsu, R. C. Cheung, H. Greenberg, D. G. Russell, and C. M. Rice.** Formation and intracellular localization of hepatitis C virus envelope glycoprotein complexes expressed by recombinant vaccinia and Sindbis viruses. Submitted for publication.
16. **Eckart, M. R., M. Selby, F. Masiarz, C. Lee, K. Berger, K. Crawford, C. Kuo, G. Kuo, M. Houghton, and Q.-L. Choo.** 1993. The hepatitis C virus encodes a serine protease involved in processing of the putative nonstructural proteins from the viral polyprotein precursor. Biochem. Biophys. Res. Commun. **192:** 399–406.
17. **Elder, J. H., and S. Alexander.** 1982. *endo*-β-*N*-Acetylglucosaminidase F: endoglycosidase from *Flavobacterium meningosepticum* that cleaves both high-mannose and complex glycoproteins. Proc. Natl. Acad. Sci. USA **79:**4540–4544.
18. **Ellison, M. J., and M. Hochstrasser.** 1991. Epitope-tagged ubiquitin. J. Biol. Chem. **266:**21150–21157.
19. **Evan, G. I., G. K. Lewis, G. Ramsay, and J. M. Bishop.** 1985. Isolation of monoclonal antibodies specific for human c-*myc* proto-oncogene product. Mol. Cell. Biol. **5:**3610–3616.
20. **Falkner, F. G., and B. Moss.** 1988. *Escherichia coli gpt* gene provides dominant selection for vaccinia virus open reading frame expression vectors. J. Virol. **62:**1849–1854.
21. **Folz, R. J., and J. I. Gordon.** 1987. Computer-assisted predictions of signal peptidase processing sites. Biochem. Biophys. Res. Commun. **146:**870–877.
22. **Francki, R. I. B., C. M. Fauquet, D. L. Knudson, and F. Brown (ed.).** 1991. Classification and nomenclature of viruses: fifth report of the International Committee on Taxonomy of Viruses. Arch. Virol. **1991**(Suppl. 2)**:**223.
23. **Frolov, I.** Unpublished data.
24. **Fuerst, T. R., E. G. Niles, F. W. Studier, and B. Moss.** 1986. Eukaryotic transient-expression system based on recombinant vaccinia virus that synthesizes bacteriophage T7 RNA polymerase. Proc. Natl. Acad. Sci. USA **83:**8122–8126.
25. **Grakoui, A., D. W. McCourt, C. Wychowski, S. M. Feinstone, and C. M. Rice.** 1993. Characterization of the hepatitis C virus-encoded serine proteinase: determination of proteinase-dependent polyprotein cleavage sites. J. Virol. **67:**2832–2843.
26. **Grakoui, A., D. W. McCourt, C. Wychowski, S. M. Feinstone, and C. M. Rice.** 1993. A second hepatitis C virus-encoded proteinase. Proc. Natl. Acad. Sci. USA **90:**10583–10587.
27. **Grakoui, A., C. Wychowski, C. Lin, S. M. Feinstone, and C. M. Rice.** 1993. Expression and identification of hepatitis C virus polyprotein cleavage products. J. Virol. **67:**1385–1395.
28. **Gribskov, M., and R. R. Burgess.** 1986. Sigma factors from E. coli, B. subtilis, phage SP01, and phage T4 are homologous proteins. Nucleic Acids Res. **14:**6745–6763.
29. **Han, J. H., V. Shyamala, K. H. Richman, M. J. Brauer, B. Irvine, M. S. Urdea, P. Tekamp-Olson, G. Kuo, Q.-L. Choo, and M. Houghton.** 1991. Characterization of the terminal regions of hepatitis C viral RNA: identification of conserved sequences in the 5′ untranslated region and poly(A) tails at the 3′ end. Proc. Natl. Acad. Sci. USA **88:**1711–1715.
30. **Hayashi, N., H. Higashi, K. Kaminaka, H. Sugimoto, M. Esumi, K. Komatsu, K. Hayashi, M. Sugitani, K. Suzuki, O. Tadao, C. Nozaki, K. Mizuno, and T. Shikata.** 1993. Molecular cloning and heterogeneity of the human hepatitis C virus (HCV) genome. J. Hepatol. **17**(Suppl. 3)**:**S94–S107.
31. **Hijikata, M., N. Kato, Y. Ootsuyama, M. Nakagawa, and K. Shimotohno.** 1991. Gene mapping of the putative structural region of the hepatitis C virus genome by *in vitro* processing analysis. Proc. Natl. Acad. Sci. USA **88:**5547–5551.
32. **Hijikata, M., H. Mizushima, T. Akagi, S. Mori, N. Kakiuchi, N. Kato, T. Tanaka, K. Kimura, and K. Shimotohno.** 1993. Two distinct proteinase activities required for the processing of a putative nonstructural precursor protein of hepatitis C virus. J. Virol. **67:**4665–4675.
33. **Hijikata, M., H. Mizushima, Y. Tanji, Y. Komoda, Y. Hirowatari, T. Akagi, N. Kato, K. Kimura, and K. Shimotohno.** 1993. Proteolytic processing and membrane association of putative nonstructural proteins of hepatitis C virus. Proc. Natl. Acad. Sci. USA **90:**10773–10777.
34. **Hollinger, F. B.** 1990. Non-A, non-B hepatitis viruses, p. 2239–2273. *In* B. N. Fields (ed.), Virology. Raven Press, New York.

35. **Hori, H., and C.-J. Lai.** 1990. Cleavage of dengue virus NS1-NS2A requires an octapeptide sequence at the C terminus of NS1. J. Virol. **64**:4573–4577.
36. **Houghton, M., A. Weiner, J. Han, G. Kuo, and Q.-L. Choo.** 1991. Molecular biology of the hepatitis C viruses: implications for diagnosis, development and control of viral disease. Hepatology **14**:381–388.
37. **Hruby, D. E., L. A. Guarino, and J. R. Kates.** 1979. Vaccinia virus replication. I. Requirement for the host-cell nucleus. J. Virol. **29**:705–715.
38. **Inchauspe, G., S. Zebedee, D.-H. Lee, M. Sugitani, M. Nasoff, and A. M. Prince.** 1991. Genomic structure of the human prototype strain H of hepatitis C virus: comparison with American and Japanese isolates. Proc. Natl. Acad. Sci. USA **88**:10292–10296.
39. **Kato, N., M. Hijikata, Y. Ootsuyama, M. Nakagawa, S. Ohkoshi, T. Sugimura, and K. Shimotohno.** 1990. Molecular cloning of the human hepatitis C virus genome from Japanese patients with non-A, non-B hepatitis. Proc. Natl. Acad. Sci. USA **87**:9524–9528.
40. **Kuo, G., Q.-L. Choo, H. J. Alter, G. L. Gitnick, A. G. Redeker, R. H. Purcell, T. Miyamura, J. L. Dienstag, M. J. Alter, C. E. Stevens, G. E. Tegtmeier, F. Bonino, M. Colombo, W.-S. Lee, C. Kuo, K. Berger, J. R. Shuster, L. R. Overby, D. W. Bradley, and M. Houghton.** 1989. An assay for circulating antibodies to a major etiologic virus of human non-A, non-B hepatitis. Science **244**:362–364.
41. **Laemmli, U. K.** 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature (London) **227**:680–685.
42. **Lanford, R. E., L. Notvall, D. Chavez, R. White, G. Frenzel, C. Simonsen, and J. Kim.** 1993. Analysis of hepatitis C virus capsid, E1, and E2/NS1 proteins expressed in insect cells. Virology **197**:225–235.
43. **Laskey, R. A., and A. D. Mills.** 1975. Quantitative film detection of ^{3}H and ^{14}C in polyacrylamide gels by fluorography. Eur. J. Biochem. **56**:335–341.
44. **Lindenbach, B., S. M. Feinstone, and C. M. Rice.** Unpublished data.
45. **Lobigs, M.** 1993. Flavivirus premembrane protein cleavage and spike heterodimer secretion requires the function of the viral proteinase NS3. Proc. Natl. Acad. Sci. USA **90**:6218–6222.
46. **Mackett, M., and G. L. Smith.** 1986. Vaccinia virus expression vectors. J. Gen. Virol. **67**:2067–2082.
47. **Manabe, S., I. Fuke, O. Tanishita, C. Kaji, Y. Gomi, S. Yoshida, C. Mori, A. Takamizawa, I. Yoshida, and H. Okayama.** 1994. Production of nonstructural proteins of hepatitis C virus requires a putative viral protease encoded by NS3. Virology **198**:636–644.
48. **Matsudaira, P.** 1987. Sequence from picomole quantities of proteins electroblotted onto polyvinylidene difluoride membranes. J. Biol. Chem. **262**:10035–10038.
49. **Matsuura, Y., S. Harada, R. Suzuki, Y. Watanabe, Y. Inoue, I. Saito, and T. Miyamura.** 1992. Expression of processed envelope protein of hepatitis C virus in mammalian and insect cells. J. Virol. **66**:1425–1431.
50. **Matsuura, Y., and T. Miyamura.** 1993. The molecular biology of hepatitis C virus. Semin. Virol. **4**:297–304.
51. **Moss, B., O. Elroy-Stein, T. Mizukami, W. A. Alexander, and T. R. Fuerst.** 1990. New mammalian expression vectors. Nature (London) **348**:91–92.
52. **Okamoto, H., M. Kojima, S.-I. Okada, H. Yoshizawa, H. Iizuka, T. Tanaka, E. E. Muchmore, D. A. Peterson, Y. Ito, and S. Mishiro.** 1992. Genetic drift of hepatitis C virus during an 8.2 year infection in a chimpanzee: variability and stability. Virology **190**:894–899.
53. **Okamoto, H., M. Kojima, M. Sakamoto, H. Iizuka, S. Hadiwandowo, S. Suwignyo, Y. Miyakawa, and M. Mayumi.** 1994. The entire nucleotide sequence and classification of a hepatitis C virus isolate of a novel genotype from an Indonesian patient with chronic liver disease. J. Gen. Virol. **75**:629–635.
54. **Okamoto, H., K. Kurai, S.-I. Okada, K. Yamamoto, H. Iizuka, T. Tanaka, S. Fukuda, F. Tsuda, and S. Mishiro.** 1992. Full-length sequence of a hepatitis C virus genome having poor homology to reported isolates: comparative study of four distinct genotypes. Virology **188**:331–341.
55. **Okamoto, H., S. Okada, Y. Sugiyama, K. Kurai, H. Iizuka, A. Machida, Y. Miyakawa, and M. Mayumi.** 1991. Nucleotide sequence of the genomic RNA of hepatitis C virus isolated from a human carrier: comparison with reported isolates for conserved and divergent regions. J. Gen. Virol. **72**:2697–2704.
56. **Pierce, J. S., E. G. Strauss, and J. H. Strauss.** 1974. Effect of ionic strength on the binding of Sindbis virus to chick cells. J. Virol. **13**:1030–1036.
57. **Ralston, R., K. Thudium, K. Berger, C. Kuo, B. Gervase, J. Hall, M. Selby, G. Kuo, M. Houghton, and Q.-L. Choo.** 1993. Characterization of hepatitis C virus envelope glycoprotein complexes expressed by recombinant vaccinia viruses. J. Virol. **67**:6753–6761.
58. **Rice, C. M., R. Levis, J. H. Strauss, and H. V. Huang.** 1987. Production of infectious RNA transcripts from Sindbis virus cDNA clones: mapping of lethal mutations, rescue of a temperature-sensitive marker, and in vitro mutagenesis to generate defined mutants. J. Virol. **61**:3809–3819.
59. **Rice, C. M., and J. H. Strauss.** 1982. Association of Sindbis virion glycoproteins and their precursors. J. Mol. Biol. **154**:325–348.
60. **Rice, C. M., and J. H. Strauss.** 1990. Production of flavivirus polypeptides by proteolytic processing. Semin. Virol. **1**:357–367.
61. **Rümenapf, T., G. Unger, J. H. Strauss, and H.-J. Thiel.** 1993. Processing of the envelope glycoproteins of pestiviruses. J. Virol. **67**:3288–3294.
62. **Saito, I., T. Miyamura, A. Ohbayashi, H. Harada, T. Katayama, S. Kikuchi, Y. Watanabe, S. Koi, M. Onji, Y. Ohta, Q.-L. Choo, M. Houghton, and G. Kuo.** 1990. Hepatitis C virus infection is associated with the development of hepatocellular carcinoma. Proc. Natl. Acad. Sci. USA **87**:6547–6549.
63. **Sambrook, J., E. F. Fritsch, and T. Maniatis.** 1989. Molecular cloning: a laboratory manual, 2nd ed. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
64. **Selby, M. J., Q.-L. Choo, K. Berger, G. Kuo, E. Glazer, M. Eckart, C. Lee, D. Chien, C. Kuo, and M. Houghton.** 1993. Expression, identification and subcellular localization of the proteins encoded by the hepatitis C viral genome. J. Gen. Virol. **74**:1103–1113.
65. **Simmonds, P., E. C. Holmes, T.-A. Cha, S.-W. Chan, F. McOmish, B. Irvine, E. Beall, P. L. Yap, J. Kolberg, and M. S. Urdea.** 1993. Classification of hepatitis C virus into six major genotypes and a series of subtypes by phylogenetic analysis of the NS5 region. J. Gen. Virol. **74**:2391–2399.
66. **Takamizawa, A., C. Mori, I. Fuke, S. Manabe, S. Murakami, J. Fujita, E. Onishi, T. Andoh, I. Yoshida, and H. Okayama.** 1991. Structure and organization of the hepatitis C virus genome isolated from human carriers. J. Virol. **65**:1105–1113.
67. **Tanaka, T., N. Kato, M. Nakagawa, Y. Ootsuyama, M.-J. Cho, T. Nakazawa, M. Hijikata, Y. Ishimura, and K. Shimotohno.** 1992. Molecular cloning of hepatitis C virus genome from a single Japanese carrier: sequence variation within the same individual and among infected individuals. Virus Res. **23**:39–53.
68. **Tomei, L., C. Failla, E. Santolini, R. De Francesco, and N. La Monica.** 1993. NS3 is a serine protease required for processing of hepatitis C virus polyprotein. J. Virol. **67**:4017–4026.
69. **von Heijne, G.** 1986. A new method for predicting signal sequence cleavage sites. Nucleic Acids Res. **14**:4683–4690.
70. **von Heijne, G.** 1990. The signal peptide. J. Membr. Biol. **115**:195–201.
71. **Wiskerchen, M., S. K. Belzer, and M. S. Collett.** 1991. Pestivirus gene expression: the first protein product of the bovine viral diarrhea virus large open reading frame, p20, possesses proteolytic activity. J. Virol. **65**:4508–4514.
72. **Yamshchikov, V. F., and R. W. Compans.** 1993. Regulation of the late events in flavivirus protein processing and maturation. Virology **192**:38–51.