JOURNAL OF VIROLOGY, Mar. 1994, p. 1643–1650
0022-538X/94/$04.00+0

Vol. 68, No. 3

Mapping a Region of the Large Envelope Protein Required for Hepatitis B Virion Maturation

VOLKER BRUSS* AND REINER THOMSSEN

Department of Medical Microbiology, University of Göttingen, D-37075 Göttingen, Federal Republic of Germany

Received 1 June 1993/Accepted 11 August 1993

The hepatitis B virion is a spherical double-shelled particle carrying three surface proteins (large [L], middle [M], and small [S]) in its envelope. All three proteins are translated from a single open reading frame by means of three different in-frame start codons from unspliced mRNAs. This organization defines three protein domains (pre-S1, pre-S2, and S). All three domains together form the L protein, whereas the M protein consists of domains pre-S2 plus S. The L and S proteins are both necessary for virion production, whereas the M protein is dispensable, suggesting an important function of the pre-S1 domain in virion morphogenesis. To investigate this point, we created a series of N-terminal-truncated L mutants and tested their ability to substitute for the wild-type L protein in virion formation. We found that the constructs fell into two classes. (i) N-terminal deletion mutants lacking up to 102 of the 119 amino acids of the pre-S1 domain still allowed virion maturation, showing that the N-terminal 5/6 of the pre-S1 sequence is dispensable for this process. (ii) Mutants lacking 110 or more N-terminal amino acids were unable to substitute for the L protein in virion assembly, although they were stably expressed and secreted as components of subviral 20-nm hepatitis B surface antigen particles. This suggests that a short C-terminal region of pre-S1 is important for virion formation. Like the wild-type L protein, the mutants of the first class were not glycosylated in their pre-S2 domains; however, this site was used for glycosylation in mutants of the second class, similar to that in the M protein. These findings can be related to a model for the function of the L protein in virion maturation.

The hepatitis B virus (HBV) is the prototype of the *Hepadnaviridae*, a family of small hepatotropic DNA viruses with a unique replication strategy (for a review, see reference 14). One of the peculiar features of HBV is that the infected cells secrete three types of viral particles: (i) the virion, which is a double-shelled sphere of 42-nm diameter with an inner icosahedral 27-nm nucleocapsid and an outer shell presumably containing lipid and three viral envelope proteins referred to as large (L), middle (M), and small (S) surface proteins; (ii) filamentous particles of variable lengths and 20-nm diameter; and (iii) spheres of 20-nm diameter (20-nm hepatitis B surface antigen [HBsAg]). The latter two particles reach much higher serum concentrations than virions for unknown reasons and consist of lipid and viral envelope proteins but contain no nucleocapsid components (17).

The HBV envelope proteins are related to each other and show a uniquely organized domain structure. All three proteins are encoded by the *env* region of the viral genome, a single continuous 400-codon-long open reading frame (Fig. 1B; the numbering of the HBV adw subtype genome used in this work starts with the deoxycytidine of the unique *Eco*RI site). Transcripts from the SPI promoter 5′ of the *env* region (45) encode the L protein, which represents the translational product of the full-length *env* region. Transcripts initiated at the SPII promoter internal in the *env* region (3) code for both the M and S proteins by two mechanisms. First, the 5′ ends of the SPII transcripts are scattered around the start codon for the M protein (45), and second, this start codon is in a weak translation initiation context (46). Therefore, protein translation from SPII transcripts mainly starts at the next downstream AUG codon and results in S protein synthesis. As a consequence of this mode of expression, the M and S proteins are like C-terminal fragments of the L protein and define three domains (Fig. 1A): the 226-amino-acid S domain, the 55-amino-acid pre-S2 domain present only in M and L, and the 119-amino-acid pre-S1 domain unique to L.

The morphogenesis of viral particles has best been studied for 20-nm HBsAg spheres consisting of S protein (1a, 6, 7, 23, 25, 28). The first step in the biosynthesis of the S protein is cotranslational insertion into the endoplasmic reticulum (ER) membrane (8). The topology of the protein is determined by at least two signal sequences, resulting in a short luminal exposed N-terminal sequence, two transmembrane regions separated by a 55-amino-acid cytosolic loop, and a luminal 70-amino-acid domain containing the major epitope of the protein and a glycosylation site (see Fig. 5A) (10). The topology of the following 50-amino-acid hydrophobic C-terminal sequence has not yet been determined experimentally. For reasons yet unknown, the glycosylation site is used only in approximately half of the S chains. The biosynthesis of the M protein seems to proceed very similarly. The pre-S2 domain contains no signal sequence but is translocated into the ER lumen by the first signal in its S domain (9), so the transmembrane topology of M and S seems to be identical (see Fig. 5A). The pre-S2 domain of M carries an additional glycan at asparagine 4. The topology of the L protein in the ER membrane is less clear. The partial usage of the glycosylation site in its S domain indicates that this site is translocated into the ER lumen by the second signal in the S domain and suggests that the topology at least of the C terminus of L is identical to that of the homologous part of M or S. The N terminus of L is myristylated (30).

The current model for the formation of subviral particles and of virions assumes that these processes take place at membranes of a pre-Golgi compartment, resulting in luminal

* Corresponding author. Mailing address: Department of Medical Microbiology, University of Göttingen, Kreuzbergring 57, D-37075 Göttingen, Federal Republic of Germany. Phone: 49 551 39 5759. Fax: 49 551 39 5860.

deposition and secretion of the particles via the constitutive secretory pathway (18, 27, 35). Which kind of particle is formed seems to be determined in part by the ratio of S and L proteins coassembling during the morphogenesis. The 20-nm HBsAg spheres can be produced by expressing the S protein alone in the absence of any other viral factor in virtually any kind of mamalian cell (6, 7, 23, 25, 28). Accordingly, the 20-nm HBsAg spheres from serum contain only traces of L protein (17). A higher proportion of L protein coassembling with S protein results in the formation of the filamentous form of HBsAg (17). An even higher content of L protein causes the inhibition of particle secretion (4, 5, 29, 36). The virion envelope also contains a relatively high proportion of L protein. In accordance with this, it was recently shown that both the L and the S proteins are necessary for virion production (2, 43). The M protein seems not to influence decisively the particle morphology, because this protein is present in all three particles forms with roughly equal ratios relative to S (17). Also, this protein is dispensable for virion formation in vitro (2) or in vivo (12) or even for infectivity in a chronically infected host (11).

All in all, the L and S proteins seem to have separate essential functions during the morphogenesis of virions and the pre-S domain has an important contribution. In this work, we shortened the pre-S1 domain of the L protein by N-terminal truncations and investigated the ability of these L derivatives to support virion morphogenesis. The results allowed us to map a short region of the pre-S1 domain necessary for virion formation and to develop a model for the L protein function in this process.

MATERIALS AND METHODS

Plasmid constructions. Plasmid pSV45H carries the 2.4-kb *Bst*EII (nucleotide [nt] 2819) to *Bgl*II (nt 1982) fragment containing the entire *env* region and the HBV polyadenylation site of the HBV subtype adw genome (44) cloned into the *Hin*dIII and *Bam*HI sites of the vector pSV65 (29). This vector bears a 342-bp *Pvu*II-*Hin*dIII fragment of the simian virus 40 (SV40) early promoter-origin region cloned into the *Pvu*II and *Hin*dIII sites of plasmid pSP65 (Promega, Madison, Wis.). Five micrograms of pSV45H was linearized between the SV40 promoter and the *env* region with *Hin*dIII and was digested with 2.4 U of Bal 31 in a 50-μl reaction volume. After 4 min, 25 μl was removed and extracted with phenol-chloroform. The other half was incubated for a further 2 min before being extracted with phenol-chloroform. The shortened HBV DNA was liberated by being cut with *Sac*I in the downstream part of the polylinker of pSV45 and was cloned into plasmid pSV65L. This plasmid corresponds to pSV65 but contains an ATG codon linker inserted between the *Hin*dIII and *Bam*HI sites (22). It has been digested with *Bam*HI and mung bean nuclease to create a blunt-ended ATG downstream of the SV40 promoter and with *Sac*I 3′ in the polylinker. After transformation, 72 individual clones were analyzed by dideoxy sequencing to find 11 suitable clones with pre-S1 codons fused directly to the ATG codon (plasmid pSV45-40 was kindly provided by K. Kuroki).

Plasmids pSV45-96 and pSV45-103 were constructed by PCR amplification of DNA fragments from cloned HBV DNA (plasmid pHBV1.5) (2) with 5′ primers GGGGAAGCTT ATGTCCACCAATCGGCAG3′ and GGGGAAGCTTATG AGGCAGCCTACTCCCAT3′, respectively, and the 3′ primer CCTGAACTGGAGCCACCA3′ hybridizing to nt 76 to 59. The PCR products were digested with *Hin*dIII, cutting in the 5′ primer sequences, and with *Pst*I (nt 25), cutting in the pre-S2 region, and cloned into *Hin*dIII-*Pst*I-digested pSV45H. The *Hin*dIII to *Pst*I region of the resulting plasmids was sequenced to exclude PCR-generated mutations.

The M^- derivatives were generated by digesting plasmid $pRVM^-$, containing an *Eco*RV (nt 1040) to *Bsp*MII (nt 2327) 1 1/2-mer of the HBV genome cloned into pBLUESCRIPT KS(+) with an ATG to ACG mutation of the M protein start codon (2), with *Bsu*36I at nt 3201 in front of the M^- mutation and upstream in the polylinker region with *Cla*I. To this DNA, *Taq*I (cutting upstream of the SV40 promoter region in pSP65 sequences)-*Bsu*36I (nt 3201) fragments containing the SV40 promoter region and the 5′ end of the pre-S1 region from plasmids pSV45-81, -89, -96, -103, -111, and -117 were ligated. The resulting plasmids were sequenced to verify the presence of the M^- mutation.

HBs protein expression and detection. COS7 cells were grown in 6-well plates in Dulbecco's modified essential medium–10% fetal calf serum, starting with 2×10^5 cells per well for 16 h at 37°C and 5% CO_2. Cells were washed with phosphate-buffered saline (PBS) and incubated with 2 ml of Dulbecco's modified essential medium–0.1% (wt/vol) $NaHCO_3$–50 mM Tris-Cl (pH 7.5)–0.5 mg of DEAE-dextran (molecular mass, 2×10^6 Da) per ml–2 μg of plasmid DNA for 6 to 8 h. Cells were washed twice with PBS and incubated for 40 h with 2 ml of medium. For metabolic labeling, cells were first washed with PBS and incubated for 40 min with 0.6 ml of methionine- and cysteine-free medium without fetal calf serum following the addition of 10 μl of [^{35}S]methionine/cysteine (85 μCi of methionine, 25 μCi of cysteine; 1,100 Ci/mmol; Du Pont de Nemours, Dreieich, Germany). After 30 min, cells were either harvested for subsequent immunoprecipitation and endoglycosidase H digestion or the medium was exchanged with 1 ml of fresh medium and the cells were incubated for a further 24 h following immunoprecipitation of HBV surface (HBs) proteins from the culture medium and from cell lysates. Cells were lysed after being washed with PBS by incubation with 1 ml of a buffer containing 150 mM NaCl, 50 mM Tris-Cl (pH 7.5), 5 mM $MgCl_2$, and 0.2% (vol/vol) Nonidet P-40 (NP-40) for 5 min at room temperature. Cell lysates and medium were spun for 5 min in an Eppendorf centrifuge, and supernatants were transferred to fresh tubes. One hundred microliters of a suspension of 10 μl of swollen protein A-Sepharose beads (Pharmacia, Freiburg, Germany) preincubated with 1 μl of goat anti-HBs (Dako, Hamburg, Germany) in PBS at 4°C for 6 h was added after overnight incubation at 4°C with agitation. Sepharose beads were washed three times with a mixture of 150 mM NaCl, 50 mM Tris-Cl (pH 7.5), 20 mM EDTA, 0.2% (vol/vol) NP-40, 0.05% (wt/vol) sodium dodecyl sulfate (SDS), and 1% (wt/vol) sodium deoxycholate. For electrophoresis, beads were suspended in 50 μl of sample buffer containing 3% (wt/vol) SDS and 5% (wt/vol) dithiothreitol (DTT) and incubated at 95°C for 5 min before application onto a 13% polyacrylamide gel. As molecular mass markers, ^{14}C-methylated bovine serum albumin (69 kDa), ovalbumin (46 kDa), carbonic anhydrase (30 kDa), and lysozyme (14.3 kDa) were used (Amersham, Braunschweig, Germany).

Endoglycosidase H digestion. For endoglycosidase H digestion, the Sepharose beads with bound HBs proteins were suspended after being washed in 800 μl of PBS, the suspension was divided into two halves, the PBS was removed with a thin capillary tube, and 30 μl of a mixture of 10 mM K acetate (pH 5.2), 0.02% (wt/vol) SDS, and 10 mM DTT was added. After 5 min of incubation at 95°C, 2 μl of 10% (vol/vol) NP-40–1 μl of endoglycosidase H (1 mU; Boehringer, Mannheim, Germany) was added to one-half of each sample and incubated overnight at 37°C. After denaturation with 30 μl of 2× sample buffer–

10% (wt/vol) DTT at 95°C for 5 min, the samples were used for polyacrylamide gel electrophoresis.

Virion expression and detection. Cotransfection of a cloned L^- HBV genome with SV40 expression vectors for L protein derivatives and detection of viral nucleocapsids and released virions was done as described previously (2). Briefly, HepG2 cells were seeded in 6-well plates and transfected for 9 h by the calcium phosphate method with 0.5 μg of a plasmid carrying the 1 1/2-mer of the HBV genome with a stop codon in the pre-S1 region ($pRVL^-$) together with 0.5 μg of an SV40 expression plasmid. After 5 days, medium and NP-40 lysates were harvested from each well and cleared. Viral nucleocapsids from one-fifth of the cell lysate (125 μl) were immunoprecipitated with a rabbit antiserum against the HBV nucleocapsid antigen (anti-HBc; Dako, Hamburg, Germany). In addition, four samples of the medium (each 400 μl or one-fifth of the total volume) were immunoprecipitated with anti-HBc or anti-HBs bound to protein A-Sepharose beads both in the presence and in the absence of NP-40. The beads were washed once with PBS and incubated in a buffer containing [α-^{32}P] dCTP to allow the endogenous polymerase of the immunoprecipitated nucleocapsids or virions to label the viral genome radioactively. After proteinase K digestion, phenol-chloroform extraction, and two ammonium acetate-ethanol precipitations, the viral DNA was separated on an agarose gel which was dried and used for autoradiography.

RESULTS

Construction of N-terminal pre-S1 deletion mutants. To identify pre-S1 sequences of the large envelope protein L which are important for virion morphogenesis, we constructed a series of 5′ truncations of the pre-S1 region by Bal 31 exonuclease digestion and ligated the remaining 3′ part of the *env* region in frame to an ATG codon downstream of an SV40 early promoter. The deletion end points of the selected constructs were distributed almost equally throughout the pre-S1 sequence (Fig. 1C). The resulting expression plasmids were named pSV45-N, where N indicates the pre-S1 triplet fused to the start codon. To fill a gap between construct pSV45-89 and pSV45-111, two further deletion mutants (pSV45-96 and pSV45-103) were created by PCR. Transcription of the *env* region from all plasmids is initiated at the SV40 early promoter instead of at the autologous HBV SPI promoter. In addition, the wild-type M and S proteins are expressed from some of the plasmids since the SPII promoter is still present. Larger pre-S1 deletions diminished the SPII promoter activity (see below).

Complementation of the viral genome carrying a pre-S1 amber mutation. HepG2 cells secrete infectious virions after transient transfection of cloned overlength copies of the HBV genome (34, 40, 42, 45). A genome carrying an amber mutation in codon 52 of the pre-S1 frame (L^- mutant), which prevents L protein expression without influencing other functions, fails to produce virions. This defect can be complemented in *trans* by cotransfection of an expression vector for the missing wild-type L protein (2). By using the series of pre-S1 deletion constructs as the cotransfecting plasmids and looking for virions in the culture medium, we were able to investigate the ability of the L mutants to replace the wild-type L protein in virion morphogenesis.

Virions were identified and discriminated from naked cores or enveloped cores containing no HBs proteins in their envelopes by a two-step procedure (2). First, viral particles were immunoprecipitated with antibodies against envelope proteins (anti-HBs) or against the nucleocapsid (anti-HBc) in the absence and the presence of a mild detergent. Then, the viral genomes within the immunoprecipitated particles were labeled by a radioactive endogenous polymerase reaction (19) following isolation of the genome, agarose gel electrophoresis, and autoradiography. The labeled genome appeared as a smear around the position of a 3-kb linear DNA on the gel (Fig. 2, WT-L). A distinctive feature of virions is that they are immunoprecipitable with anti-HBs in the absence but not in the presence of detergent (compare third and fourth lanes) and with anti-HBc only in the presence of detergent (compare first and second lanes). As a control, nucleocapsids from an equivalent amount of cell lysate were immunoprecipitated with anti-HBc and included (fifth lane).

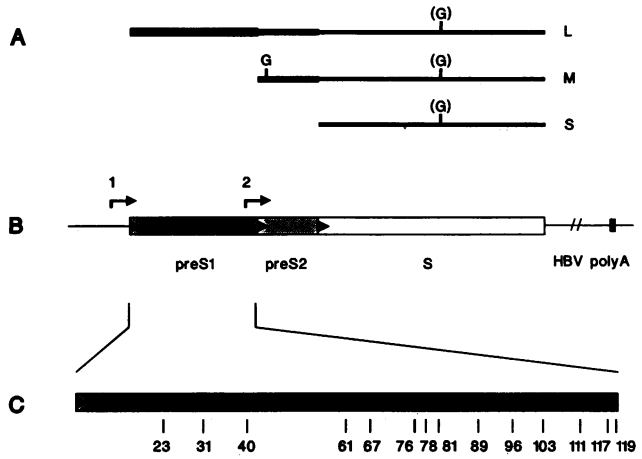

FIG. 1. (A) Domain structure of the three coterminal HBV envelope proteins S, M, and L. The 226-amino-acid S domain, the 55-amino-acid pre-S2 domain at the N terminus of M and central in L, and the 119-amino-acid pre-S1 domain at the N terminus of L are shown as lines of different thickness. The S protein and the S domains of M and L are glycosylated in approximately half of the chains, as indicated by a G in parentheses. The pre-S2 domain of M but not that of L is glycosylated in naturally secreted viral particles. (B) The *env* region on the HBV genome. The whole box represents the *env* region, a single open reading frame on the HBV genome which encodes all three surface proteins S, M, and L. The triangles indicate ATG start codons for surface protein synthesis which divide the open reading frame into the pre-S1, pre-S2, and S regions. The arrows above the box mark the SPI and SPII transcription start sites. In the expression plasmids used in this work, the SPI promoter is replaced by the SV40 early promoter. (C) Enlargement of the pre-S1 domain. Each number represents a 5′-terminal deletion mutant constructed for this work. The numbers refer to the pre-S1 codon fused to the translation initiation codon, e.g., mutant 61 has a deletion of codons 2 to 60 of the pre-S1 domain.

As previously shown (2), the L^- mutation could be complemented in *trans* by expression of the wild-type L protein (Fig. 2, WT-L). Cotransfection with the N-terminal-truncated L versions clearly showed that mutants 23 to 103 were able to complement the L^- defect in virion morphogenesis, whereas mutants 111, 117, and 119 were unable to do so although comparable amounts of nucleocapsids could be found in cell lysates (Fig. 2). The mutants were divided by this test into two distinct classes. All mutants lacking up to 5/6 of the pre-S1 N terminus still allowed for virion formation, whereas the additional deletion of only eight or more amino acids drastically changed the phenotype to deficiency for virion formation.

As a side observation, we noted that the level of endogenous polymerase-reactive nucleocapsids in the cell lysates and viri-

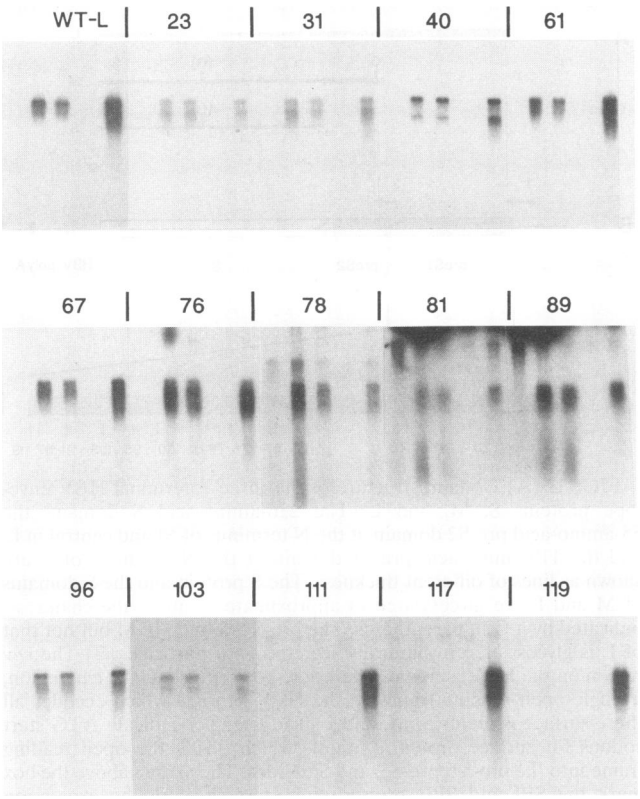

FIG. 2. Complementation of the L$^-$ HBV genome. Wild-type L protein (panel WT-L) as well as the L mutants were coexpressed in HepG2 cells with an L$^-$ mutant HBV genome. Secretion of enveloped virions into the culture medium was detected by immunoprecipitation of viral particles from the medium with (from left to right for each panel) anti-HBc minus detergent, anti-HBc plus detergent, anti-HBs minus detergent, anti-HBs plus detergent, and NP-40 cell lysates with anti-HBc. After a radioactive endogenous polymerase reaction, the labeled HBV genomes were isolated from the immunoprecipitated particles and separated on an agarose gel which was used for autoradiography. The pattern indicative for completely enveloped virions can be seen in panel WT-L (for details, see the text). Mutants 23 to 103 showed the wild-type phenotype. Mutants 111, 117, and 119 were unable to substitute for the L protein in virion morphogenesis. Autoradiograms for mutants 78, 81, and 89 were exposed for 5 to 7 days, the other autoradiograms were exposed between 16 and 36 h.

ons in the medium was approximately five times lower for the cotransfection experiments with mutants 78, 81, and 89 (Fig. 2). HBsAg secretion was not comparably impaired (data not shown). This effect of one mutant (mutant 78) tested further was dominant since cotransfection of mutant 78 with wild-type HBV DNA also showed the negative influence. This was dependent on the expression of mutant 78 protein, since cotransfection of mutant 78 carrying an amber mutation in its pre-S1 sequence did not express the inhibitory phenotype (data not shown). At this time, we do not know the reason for this surprising behavior.

Expression and glycosylation of the mutant L proteins. One trivial reason for the inability of mutants 103, 111, and 117 to complement the L$^-$ HBV genome could be that these proteins were expressed inefficiently. To investigate this possibility, we transiently transfected the SV40 expression plasmids into COS7 cells and tried to detect the proteins after labeling them with [^{35}S]methionine by immunoprecipitation with a polyclonal anti-HBs serum from NP-40 lysates, gel electrophoresis, and autoradiography of the labeled proteins. The anti-HBs serum mainly reacts with the major HBsAg epitope in the S domain which is also present in the M protein and L derivatives so that all envelope proteins could be detected by this antiserum. The resulting picture is fairly complex due to the expression of up to seven proteins from the *env* region, because all three envelope proteins are expressed with variable glycosylation. To facilitate interpretation of the band patterns, half of each sample was digested with endoglycosidase H to remove the glycoside residues and was run adjacent to the undigested samples on the gel (Fig. 3).

All plasmids drove the synthesis of S protein, generating the two lowest bands which represent the unglycosylated p24 and the glycosylated gp27 forms of the S protein. The three constructs with the largest pre-S1 deletions, pSV45-111, -117, and -119, expressed less S protein probably because essential parts of the internal SPII promoter driving M and S expression have been deleted. The residual S expression might be due to internal initiation on the SV40-driven transcript. The M protein was synthesized as non (p30)-, once (gp33)-, and twice (gp36)-glycosylated peptides (e.g., wild-type M [WT-M] or mutant 23, 31, or 40). The L proteins appeared as un- and monoglycosylated forms for mutants 23 to 103 and as un-, mono-, and diglycosylated forms for mutants 111 to 119. Interpretation of the band pattern is possible considering that one glycoside residue enlarged the apparent molecular mass equally by 3 kDa (37). For example, the seven bands appearing in mutant 103 were (from bottom to top) unglycosylated S, glycosylated S, unglycosylated M, unglycosylated mutant 103, once-glycosylated M, once-glycosylated mutant 103, and double-glycosylated M.

This analysis allows two statements. First, mutants 111 to 119 were expressed in comparable amounts relative to the longer L derivatives at least in COS7 cells so that low expression as a trivial explanation for their failure in the complementation assay is not likely to be true. And, second and more interestingly, all mutants allowing virion morphogenesis were, like the wild-type L protein, glycosylated only once, probably in the S domain and not in the pre-S2 domain, whereas all mutants incompetent for virion formation were, like the M protein, partially double glycosylated, showing that the pre-S2 glycosylation site was used.

Secretion of L mutants. Mutants 111 to 119, which were unable to substitute for the wild-type L protein in virion formation, were stably expressed. As another reason for their negative phenotype, an incorrect folding of these protein chains could be considered. To test this theory, we investigated another feature of the HBV envelope proteins which is likely to depend on correct folding: the secretion of the proteins as components of 20-nm HBsAg particles. Former studies on S mutants showed that even small mutations can render the protein unstable or incompetent for secretion (1a). Also, small insertions in the pre-S1 domain destroyed the ability of L to be secreted (1). If L mutants 111 to 119 were secreted, this would argue for a correct folding at least of protein domains important for 20-nm particle formation.

After pulse-labeling of the L derivatives in transfected COS7 cells, the envelope proteins were immunoprecipitated after a 24-h chase period with anti-HBs from an NP-40 cell lysate (Fig. 4, lanes L) and additionally from the culture medium (Fig. 4, lanes M) and analyzed by gel electrophoresis and autoradiography. The cell lysates showed basically the same patterns of protein bands as those in the experiment described above (Fig. 3). The S proteins were secreted into the medium in all cases, as expected. During export, the glycoside of gp27 is modified

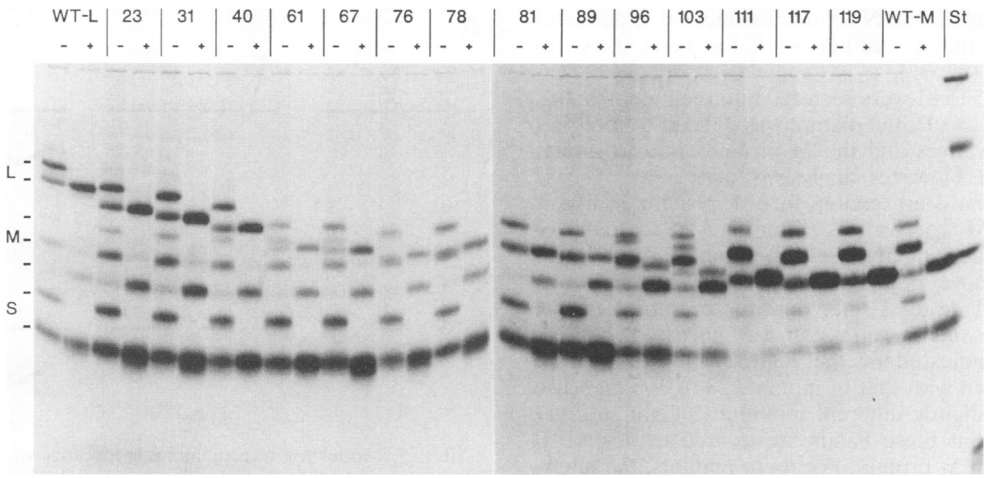

FIG. 3. Expression and glycosylation of L mutants. Expression plasmids driving the synthesis of wild-type L (WT-L), wild-type M (WT-M), and N-terminal-truncated L proteins by the SV40 promoter and M and S proteins by their autologous SPII promoters were transiently transfected into COS7 cells. HBV envelope proteins were immunoprecipitated from cell lysates after radioactive pulse-labeling. The samples were divided, half remained untreated (−), and the other half was digested with endoglycosidase H (+). The resulting proteins were depicted by polyacrylamide gel electrophoresis and autoradiography. The S protein was expressed as unglycosylated and once-glycosylated peptides. The M protein showed an additional double-glycosylated form due to an additional glycosylation site in the pre-S2 domain (Fig. 1A). The wild-type L protein is not glycosylated in its pre-S2 domain. Mutants 23 to 103 were glycosylated like the L protein, whereas mutants 111 to 119 were glycosylated like the M protein. The addition of one glycoside resulted in a molecular mass shift of 3 kDa (37), e.g., for mutant 103 the seven bands in the untreated sample represent (from bottom to top) unglycosylated S, glycosylated S, unglycosylated M, unglycosylated mutant 103, once-glycosylated M, once-glycosylated mutant 103, twice-glycosylated M. Lane St, molecular mass markers (69, 46, 30, and 14.3 kDa).

from an endoglycosidase H sensitive to a resistant form, resulting in a slightly higher molecular mass of secreted versus intracellular gp27, which is the reason for the slower mobility of this protein from the medium relative to the lysate fraction (28). This characteristic offers a convenient control to exclude release of cytoplasmic proteins into the medium by cell lysis. The three forms of the M protein were secreted differently (best seen in Fig. 4, WT-M lanes). While the unglycosylated form was mainly retained in the cell, the once- and twice-glycosylated peptides were secreted as efficiently as the S protein. The reason for this behavior is unknown, but natural HBsAg from serum also contains only the glycosylated versions of the M protein. The wild-type L protein inhibited the secretion of envelope proteins (Fig. 4, lanes WT-L) (4, 29, 36).

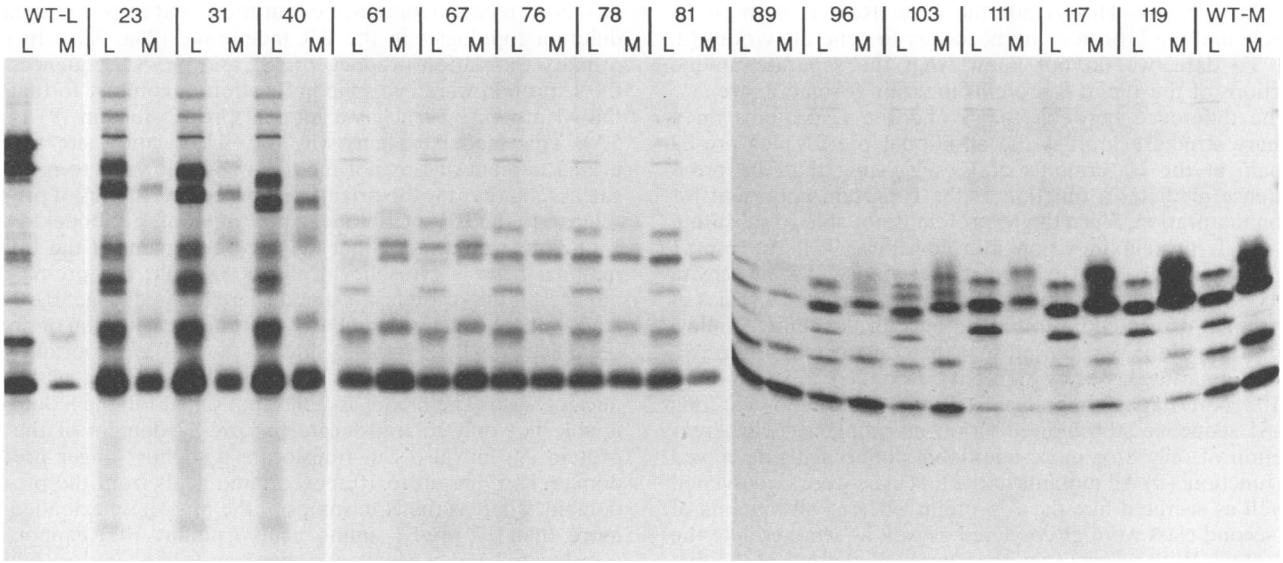

FIG. 4. Secretion of the L deletion mutants. Wild-type L (WT-L), wild-type M (WT-M), and L mutants were expressed in COS7 cells. The proteins were radioactively pulse-chase labeled, immunoprecipitated with polyclonal anti-HBs from NP-40 cell lysates (L) and medium (M), and detected by polyacrylamide gel electrophoresis and autoradiography. L mutants 89 to 103 can hardly be distinguished from the M protein. All envelope proteins were secreted well except the unglycosylated versions of the M protein and of the L mutants 111, 117, and 119. Again, L mutants 111 to 119 were similar to the M protein.

This feature was lost in the N-terminal-truncated L derivates as expected, since the N terminus of L has been shown to mediate this effect (22, 32).

The L mutants 23 to 81 were secreted into the medium (Fig. 4). For mutants 89 to 103, the picture was difficult to interpret since these L derivatives and the M proteins ran on similar positions in the gel. However, analysis of derivatives of these mutants lacking the start codon for M protein synthesis revealed that the M^- derivatives of mutants 89 to 103 were indeed secreted (data not shown). For mutants 111, 117, and 119, the bands above gp27 probably represent L protein mutants and not M proteins since the SPII promoter activity responsible for M and S expression was impaired in these constructs, as is indicated by the relatively low level of S expression compared with that of mutants 23 to 103 (see also Fig. 3). Also, the slightly different mobilities of mutants 111 and 117 confirm that these bands are related to deleted L proteins rather than M protein. For these mutants, the unglycosylated L derivative was only inefficiently secreted whereas the once- and double-glycosylated forms were secreted well. This is the same characteristic the M protein shows.

This analysis showed that all L mutants were secreted into the medium, suggesting that no mutant was drastically denatured by the truncation. Interestingly, the unglycosylated versions of those mutants deficient in virion assembly (mutants 111, 117, and 119) were only poorly secreted in contrast to the remaining L mutants. This difference between the two classes of mutants reflects the difference between the L and M proteins: the unglycosylated L protein is secreted in contrast to the unglycosylated M protein.

DISCUSSION

The three HBV envelope proteins, L, M, and S, have a uniquely organized domain structure. The M and S protein sequence is completely preserved at the C terminus of their larger relative L. Naively speaking, one would expect that the L protein can serve all functions of the envelope proteins during the generation of virions since it contains all envelope protein domains. This is not the case. Rather, both the S protein and the L protein are necessary to generate virions (2, 43). To date, we do not know what the separate unique functions of the L and S proteins in virion formation are.

The difference between the S and the L proteins on a primary structure level is the additional pre-S1 plus pre-S2 domain at the N terminus of L, suggesting that the pre-S sequence mediates a function of the L protein important for virion maturation. Since the M protein is not able to substitute for the L protein in virion morphogenesis (2), the pre-S1 domain seems to be of particular importance. To investigate the contribution of the pre-S1 domain, we analyzed a series of N-terminal-truncated L protein mutants for the ability to allow virion formation. We showed that the constructs fall into two clearly separate classes by different criteria. (i) Deletion of up to 102 N-terminal amino acids of the 119-amino-acid-long pre-S1 sequence still allowed virion morphogenesis, whereas deletion of only 8 or more additional amino acids destroyed this function. (ii) All mutants of the first class were glycosylated as well as secreted like the L protein, whereas all mutants of the second class were glycosylated as well as secreted like the M protein. With respect to both criteria, all mutants of the first class were L-like and all mutants of the second class were M-like.

Deduced from the mapping of pre-S epitopes exposed on the surface of viral particles (21) and from the mapping of protease-sensitive sites of L on the virion, most models predict

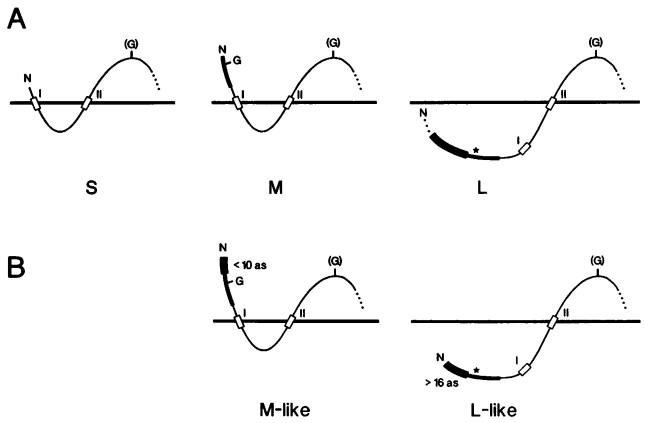

FIG. 5. Model for transmembrane topology of wild-type and N-terminal-truncated L proteins. The horizontal lines represent the ER membrane, with the luminal side above the line. (A) The S protein traverses the membrane at least twice with the N terminus (N) in the ER lumen and contains two signals (I and II, shown as open boxes) (10). The luminal loop is partially glycosylated (G). The location of the C terminus has not yet been determined. The topology of the M protein is very similar, with a luminal, glycosylated pre-S2 domain shown as a thicker line (9). The topology of the L protein is proposed to have a cytosolic pre-S2 domain which is consequently not glycosylated (indicated by an asterisk) (26a, 32a). (B) Proposed model for transmembrane topology of N-terminal-truncated L mutants. Mutants carrying fewer than 10 pre-S1 amino acids show an M-like topology, as indicated by the glycosylation of their pre-S2 domains, whereas mutants carrying more than 17 pre-S1 amino acids have an L-like topology. Only the L-like mutants were able to substitute for the wild-type L protein in virion maturation. This argues for the importance of the cytosolic presentation of pre-S sequences for virion formation.

that the transmembrane L protein topology is similar to the topology of M and S proteins, with the pre-S1 domain in the ER lumen and consequently external on viral particles (13). However, recent data indicate that the L and M proteins have different topologies in the ER membrane (26a, 32a). In the primary translation product, pre-S2 and pre-S1 sequences of the L protein were cytosolic in location in contrast to that of the M protein which has a luminal pre-S2 domain (9) (Fig. 5A). This model explains why the glycosylation site in the pre-S2 domain of L is not modified and allows us to use this site as a marker for the orientation of the N terminus of pre-S2 relative to the ER membrane. Application of this model to the L mutants would mean that the pre-S2 domains of the L-like mutants have a cytosolic localization, while the pre-S2 sequences of the second class mutants are luminal (Fig. 5B). This leads to the attractive hypothesis that the presentation of pre-S sequences on the cytosolic side of the ER is important for virion formation, perhaps for establishing contact with the nucleocapsid. The model assumes that signal I in the S domain is able not only to translocate the pre-S2 domain of the M protein (9) but also can translocate a slightly longer pre-S2 domain carrying up to 10 more amino acids from the pre-S1 domain, whereas the N terminus of the M protein extended by more than 17 pre-S1 amino acids (mutant 103) cannot be translocated and consequently shows an L-like topology which would be sufficient to allow virion formation.

Alternative models are possible, and some observations do not fit well with the proposal. For example, it has been shown that the correct luminal orientation of the N terminus of the S protein or foreign domains fused to the N terminus of S was

important for secretion as a component of 20-nm particles (1a). In addition, the pre-S2 domain at least of a subset of L proteins is exposed on the surface of virions, since mutant virions expressed in HepG2 cells which lack the M protein can be immunoprecipitated with monoclonal antibodies binding to the C-terminal portion of the pre-S2 domain of L (1). This would argue for a luminal localization of the pre-S2 domain of L which is topologically equivalent to the surface of the virion. To date, no conclusive model combining these apparently contradictory findings has been found.

As is apparent from Fig. 3 and 4, the border separating L- and M-like mutants is located at a position in the pre-S1 sequence which contains important elements of the SPII promoter driving the expression of M and S proteins. So, for all L-like mutants (mutants 23 to 103), the M and S proteins were still synthesized at high levels from the expression plasmids, whereas for all M-like mutants (mutants 111 to 119), the SPII activity was markedly decreased. Thus, in cotransfection experiments using the SV40 expression plasmids and the L^- HBV genome, different ratios of L derivatives to M and S protein were present. The fact that these different L to M plus S ratios might influence virion morphogenesis cannot be overlooked. However, the coexpression of mutant 31 with a wild-type HBV genome did not inhibit virion export (1).

Mutants 81, 89, and 96 showed a negative influence on nucleocapsid formation and virion production (Fig. 2). It is enigmatic to us how a transmembrane protein in the ER could influence the formation of cytoplasmic nucleocapsids in our system. The duck hepatitis B virus L protein has been shown to regulate the intracellular pool of viral covalently closed circular DNA (38). However, this study was done in an infection system in which regulation of intracellular DNA levels could be studied. In our system using transient transfections, it is unlikely that such a mechanism of regulation could be observed. In another report, the transcription activation activity of C-terminal-truncated M proteins has been described (20). Also in this case the mechanism is unknown.

The evolutionary conservation of the N-terminal part of pre-S1 argues for an important function of this region apart from virion formation in the viral life cycle. Binding studies of the L protein (31) or pre-S1 peptides (26) to hepatocytes suggested that this region might be the ligand for an HBV receptor. Also, for the duck hepatitis B virus the N-terminal pre-S region was important for infectivity and a C-terminal part was important for virion formation (22a, 24, 39).

Recently, several reports on the occurrence of HBV variants with mutations in the *env* region in chronic HBV carriers have been published (12, 15, 33, 41). Some mutations abolished expression of the L protein by stop codons (41) or represented deletions of C-terminal parts of pre-S1 (15, 33) which has been shown in this study to be important for virion morphogenesis. For two of these studies (33, 41), it is unclear whether the mutant genomes were isolated from enveloped virions. In the other report (15), virions have been purified from serum by isopycnic sucrose gradient centrifugation before DNA extraction. For these mutants, it would have to be assumed that the L^- defect was complemented by coinfection of the virus-producing cell with an HBV genome expressing a functional L protein.

ACKNOWLEDGMENTS

We thank Kazujuki Kuroki for plasmid pSV45-40, referred to as plasmid 37 in reference 22. We are grateful to Don Ganem for strong support at the onset of this study. We also thank Wolfram Gerlich and Don Ganem for critical reading of the manuscript and for many valuable suggestions.

This work was supported by the Deutsche Forschungsgemeinschaft, grant number Br 1018/2-1.

REFERENCES

1. **Bruss, V.** Unpublished data.
1a. **Bruss, V., and D. Ganem.** 1991. Mutational analysis of hepatitis B surface antigen particle assembly and secretion. J. Virol. **65:**3813–3820.
2. **Bruss, V., and D. Ganem.** 1991. The role of envelope proteins in hepatitis B virus assembly. Proc. Natl. Acad. Sci. USA **88:**1059–1063.
3. **Cattaneo, R., H. Will, N. Hernandez, and H. Schaller.** 1983. Signals regulating hepatitis B surface antigen transcription. Nature (London) **305:**336–338.
4. **Cheng, C., K. Smith, and B. Moss.** 1986. Hepatitis large surface protein is not secreted but is immunogenic when selectively expressed by recombinant vaccinia virus. J. Virol. **60:**337–344.
5. **Chisari, F. V., P. Filippi, A. MacLachlan, D. Milich, M. Riggs, S. Lee, R. Palmiter, C. Pinkert, and R. Brinster.** 1986. Expression of hepatitis B large envelope polypeptide inhibits hepatitis B surface antigen secretion in transgenic mice. J. Virol. **60:**880–887.
6. **Crowley, C. W., C. C. Liu, and A. D. Levinson.** 1983. Plasmid directed synthesis of hepatitis B surface antigen in monkey cells. Mol. Cell. Biol. **3:**44–45.
7. **Dubois, M. F., C. Pourcel, S. Pousset, C. Chany, and P. Tiollais.** 1980. Excretion of hepatitis B surface antigen particles from mouse cells transformed with cloned viral DNA. Proc. Natl. Acad. Sci. USA **77:**4549–4553.
8. **Eble, B., V. Lingappa, and D. Ganem.** 1986. Hepatitis B surface antigen: an unusual secreted protein initially synthesized as a transmembrane polypeptide. Mol. Cell. Biol. **6:**1454–1463.
9. **Eble, B. E., V. R. Lingappa, and D. Ganem.** 1990. The N-terminal (pre-S2) domain of a hepatitis B virus surface glycoprotein is translocated across membranes by downstream signal sequences. J. Virol. **64:**1414–1419.
10. **Eble, B. E., D. R. MacRae, V. R. Lingappa, and D. Ganem.** 1987. Multiple topogenic sequences determine the transmembrane orientation of hepatitis B surface antigen. Mol. Cell. Biol. **7:**3591–3601.
11. **Fernholz, D., P. R. Galle, M. Stemler, M. Brunetto, F. Bonino, and H. Will.** 1993. Infectious hepatitis B virus variant defective in pre-S2 protein expression in a chronic carrier. Virology **194:**137–148.
12. **Fernholz, D., M. Stemler, M. Brunetto, F. Bonino, and H. Will.** 1991. Replicating and virion secreting hepatitis B mutant virus unable to produce preS2 protein. J. Hepatol **13**(Suppl. 4)**:**S102–S104.
13. **Ganem, D.** 1991. Assembly of hepadnaviral virions and subviral paricles. Curr. Top. Microbiol. Immunol. **168:**61–83.
14. **Ganem, D., and H. E. Varmus.** 1987. The molecular biology of the hepatitis B viruses. Annu. Rev. Biochem. **56:**651–693.
15. **Gerken, G., D. Kremsdorf, F. Capel, M. A. Petit, C. Dauguet, M. R. Manns, K. Meyer zum Büschenfelde, and C. Brechot.** 1991. Hepatitis B defective virus with rearrangements in the preS gene during chronic HBV infection. Virology **183:**555–565.
16. **Gerlich, W. H., and V. Bruss.** 1993. Functions of hepatitis B virus proteins and molecular targets for protective immunity, p. 41–82. *In* R. E. Ellis (ed.), Hepatitis B vaccines in clinical practice. Marcel Dekker, Inc., New York.
17. **Heermann, K., U. Goldmann, W. Schwartz, T. Seyffarth, H. Baumgarten, and W. H. Gerlich.** 1984. Large surface proteins of hepatitis B virus containing the preS sequence. J. Virol. **52:**396–402.
18. **Huovila, A. J., A. M. Eder, and S. D. Fuller.** 1992. Hepatitis B surface antigen assembles in a post-ER, pre-Golgi compartment. J. Cell Biol. **118:**1305–1320.
19. **Kaplan, P. M., R. L. Greenman, J. L. Gerin, R. H. Purcell, and W. S. Robinson.** 1973. DNA polymerase associated with human hepatitis B antigen. J. Virol. **12:**995–1005.
20. **Kekule, A. S., U. Lauer, M. Meyer, W. H. Caselmann, P. H. Hofschneider, and R. Koshy.** 1990. The preS/S region of integrated hepatitis B virus DNA encodes a transcriptional transactivator. Nature (London) **343:**457–461.

21. **Kuroki, K., M. Floreani, L. Mimms, and D. Ganem.** 1990. Epitope mapping of the hepatitis B virus large surface antigen. Virology **176:**620–624.
22. **Kuroki, K., R. Russnak, and D. Ganem.** 1989. Novel N-terminal amino acid sequence required for retention of a hepatitis B virus glycoprotein in the endoplasmic reticulum. Mol. Cell. Biol. **9:**4459–4466.
22a. **Lenhoff, R., and J. Summers.** Personal communication.
23. **Liu, C. C., D. Yansura, and A. D. Levinson.** 1982. Direct expression of hepatitis B surface antigen in monkey cells from an SV40 vector. DNA **1:**213–221.
24. **MacRae, D., V. Bruss, and D. Ganem.** 1991. Myristylation of a duck hepatitis B virus envelope protein is essential for infectivity but not for virus assembly. Virology **181:**359–361.
25. **Moriarty, A. M., B. H. Hoyer, J. W. Shih, J. L. Gerin, and D. H. Hamer.** 1981. Expression of the hepatitis B virus surface antigen gene in cell culture using a simian virus 40 vector. Proc. Natl. Acad. Sci. USA **78:**2606–2610.
26. **Neurath, A. R., S. B. H. Kent, N. Strick, and K. Parker.** 1986. Identification and chemical synthesis of a host cell receptor binding site on the hepatitis B virus. Cell **46:**429–436.
26a. **Ostapchuk, P., and D. Ganem.** Personal communication.
27. **Patzer, E. J., G. R. Nakamura, C. C. Simonsen, A. D. Levinson, and R. Brands.** 1986. Intracellular assembly and packaging of hepatitis B surface antigen particles occur in the endoplasmic reticulum. J. Virol. **58:**884–892.
28. **Patzer, E. J., G. R. Nakamura, and A. Yaffee.** 1984. Intracellular transport and secretion of hepatitis B surface antigen in mammalian cells. J. Virol. **51:**346–353.
29. **Persing, D., H. Varmus, and D. Ganem.** 1986. Inhibition of secretion of hepatitis B surface antigen by a related presurface polypeptide. Science **234:**1388–1392.
30. **Persing, D. H., H. E. Varmus, and D. Ganem.** 1987. The preS1 protein of hepatitis B virus is acylated at its amino terminus with myristic acid. J. Virol. **61:**1672–1677.
31. **Pontisso, P., M. A. Petit, M. J. Bankowski, and M. E. Peeples.** 1989. Human liver plasma membranes contain receptors for the hepatitis B virus pre-S1 region and, via polymerized human serum albumin, for the pre-S2 region. J. Virol. **63:**1981–1988.
32. **Prange, R., A. Clemen, and R. E. Streeck.** 1991. Myristylation is involved in intracellular retention of hepatitis B virus envelope proteins. J. Virol. **65:**3919–3923.
32a. **Prange, R., and R. E. Streeck.** Personal communication.
33. **Santantonio, T., M. Jung, R. Schneider, D. Fernholz, M. Milella, L. Monno, G. Pastore, G. R. Pape, and H. Will.** 1992. Hepatitis B virus genomes that cannot synthesize pre-S2 proteins occur frequently and as dominant virus population in chronic carriers in Italy. Virology **188:**948–952.
34. **Seels, M. A., M. L. Chen, and G. Acs.** 1987. Production of hepatitis B virus particles in HepG2 cells transfected with cloned hepatitis B virus DNA. Proc. Natl. Acad. Sci. USA **84:**1005–1009.
35. **Simon, K., V. R. Lingappa, and D. Ganem.** 1988. Secreted hepatitis B surface antigen polypeptides are derived from a transmembrane precursor. J. Cell Biol. **107:**2163–2168.
36. **Standring, D. N., J. Ou, and W. J. Rutter.** 1986. Assembly of viral particles in Xenopus oocytes: pre-surface-antigens regulate secretion of the hepatitis B viral surface envelope particles. Proc. Natl. Acad. Sci. USA **83:**9338–9342.
37. **Stibbe, W., and W. H. Gerlich.** 1982. Variable protein composition of hepatitis B surface antigen from different donors. Virology **123:**436–442.
38. **Summers, J., P. M. Smith, and A. L. Horwich.** 1990. Hepadnavirus envelope proteins regulate covalently closed circular DNA amplification. J. Virol. **64:**2819–2824.
39. **Summers, J., P. M. Smith, M. Huang, and M. Yu.** 1991. Morphogenetic and regulatory effects of mutations in the envelope proteins of an avian hepadnavirus. J. Virol. **65:**1310–1317.
40. **Sureau, C., J. L. Romet-Lemonne, J. I. Mullins, and M. Essex.** 1986. Production of hepatitis B virus by a differentiated human hepatoma cell line after transfection with cloned circular HBV DNA. Cell **47:**37–47.
41. **Tran, A., D. Kremsdorf, F. Capel, C. Housset, C. Dauguet, M. A. Petit, and C. Brechot.** 1991. Emergence of and takeover by hepatitis B virus (HBV) with rearrangements in the Pre-S/S and Pre-C/C genes during chronic HBV infection. J. Virol. **65:**3566–3574.
42. **Tsurimoto, T., A. Fujiyama, and K. Matsubara.** 1987. Stable expression and replication of hepatitis B virus genome in an integrated state in human hepatoma cell line transfected with the cloned viral DNA. Proc. Natl. Acad. Sci. USA **84:**444–448.
43. **Ueda, K., T. Tsurimoto, and K. Matsubara.** 1991. Three envelope proteins of hepatitis B virus: large S, middle S, and major S proteins needed for the formation of Dane particles. J. Virol. **65:**3521–3529.
44. **Valenzuela, P., M. Quiroga, J. Zaldivar, R. Gray, and W. Rutter.** 1980. The nucleotide sequence of the hepatitis B viral genome and the identification of the major viral genes. UCLA Symp. Mol. Cell. Biol. **18:**57–70.
45. **Yaginuma, K., Y. Shirakata, M. Kobayashi, and K. Koike.** 1987. Hepatitis B virus (HBV) particles are produced in a cell culture system by transient expression of transfected HBV DNA. Proc. Natl. Acad. Sci. USA **84:**2678–2682.
46. **Yi Sheu, S., and S. J. Lo.** 1992. Preferential ribosomal scanning is involved in the differential synthesis of the hepatitis B viral surface antigens from subgenomic transcripts. Virology **188:**353–357.