JOURNAL OF VIROLOGY, Dec. 1997, p. 9350–9357 Vol. 71, No. 12
0022-538X/97/$04.00+0
Copyright © 1997, American Society for Microbiology

A Short Linear Sequence in the Pre-S Domain of the Large Hepatitis B Virus Envelope Protein Required for Virion Formation

VOLKER BRUSS*

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

Received 17 July 1997/Accepted 10 September 1997

Envelopment of the hepatitis B virus (HBV) nucleocapsid depends on the large envelope protein L, which is expressed as a transmembrane polypeptide at the endoplasmic reticulum membrane. Previous studies demonstrated that the cytosolic exposure of the N-terminal pre-S domain (174 amino acids) of L was required for virion formation. N-terminal truncations of L up to Arg 103 were tolerated. To map sites in the remaining C-terminal part of pre-S important for virion morphogenesis, a series of 11 L mutants with linker substitutions between Asn 98 and Pro 171 was generated. The mutants formed stable proteins and were secreted in transfected cell cultures, probably as components of subviral hepatitis B surface antigen particles. All four constructs with mutations between Asn 98 and Thr 125 were unable to complement in *trans* the block in virion formation of an L-negative HBV genome in cotransfected HuH7 cells. These mutants had a transdominant negative effect on virus yield in cotransfections with the wild-type HBV genome. In contrast, all seven mutants with substitutions downstream of Ser 124 were able to envelop the nucleocapsid and to secrete HBV. The sequence between Arg 103 and Ser 124 is highly conserved among different HBV isolates and also between HBV and the woodchuck hepatitis virus. Point mutations in this region introducing alanine residues at conserved positions blocked virion formation, in contrast to mutations at nonconserved residues. These results demonstrate that the pre-S sequence between Arg 103 and Ser 124 has an important function in HBV morphogenesis.

The virion of the human hepatitis B virus (HBV) is a spherical particle of 42-nm diameter which consists of an icosahedral nucleocapsid of approximately 30-nm diameter and an envelope containing three viral surface proteins and probably lipid. Both principal virion substructures, the nucleocapsid and the envelope, are synthesized and mature separately in different cellular compartments, and they interact with each other in a subsequent step (12). Nucleocapsid assembly is initiated by binding of the viral DNA polymerase/reverse transcriptase to the pregenomic RNA, and this complex is packaged by 90 or 120 copies of core protein dimers. The pregenome is then reverse transcribed within the lumen of the capsid to minus-stranded DNA, which is replicated in a second step to yield the circular, partially double-stranded DNA genome 3.2 kb in length (13, 24, 32).

The three HBV envelope proteins are expressed from one open reading frame using three in-frame start sites (Fig. 1A and B) (18). The large surface protein (L) is the translation product of the whole open reading frame (400 amino acid residues [aa] for HBV genotype A). The middle surface protein (M) lacks the N-terminal 119 aa of L (the pre-S1 sequence), and the small surface protein (S) lacks the N-terminal 55 aa of M (the pre-S2 sequence). The proteins are synthesized at the endoplasmic reticulum (ER) and have a complex transmembrane topology. The S protein traverses the membrane at least twice with transmembrane region 1 (TM1) and TM2 and with the N terminus oriented toward the ER lumen (10). The M protein has a similar topology, and the additional N-terminal pre-S2 domain is located in the ER lumen (9). The L protein exposes the pre-S domains on the cytosolic side and traverses the ER membrane, probably with TM2 in its S sequence. Noteworthy, the transmembrane topology of approximately half of the L protein molecules changes posttranslationally, resulting in the exposure of pre-S domains on the surface of the virus particle (4, 25, 29). The mechanism of this refolding is unknown. All three HBV envelope proteins form disulfide-linked dimers with each other without any detectable preference for a certain pairing (35).

One peculiarity of HBV and related viruses is that the surface proteins not only are incorporated into virion envelopes but also form subviral spherical particles without nucleocapsids. These particles assemble at a pre-Golgi membrane (19) together with lipid, have a diameter of 20 nm, and are spherical or filamentous in shape. They appear in the luminal compartment and are secreted from infected hepatocytes or transfected cells in great excess over virions as so-called 20-nm hepatitis B surface antigen (HBsAg) particles. The molecular mechanism of the transition from the transmembrane to the particulate state of the subviral HBsAg is still enigmatic.

Thus far, no permissive permanent cell culture system is available for HBV. However, certain human liver cell lines, like HepG2 or HuH7, produce low amounts of infectious virus after transfection with suitable genomic DNA constructs (1) and have been used to study HBV particle maturation. Characterization of HBV mutants revealed that the L and S proteins, but not the M protein, were necessary for virion morphogenesis (3). This finding suggested that the pre-S domains of L have an important role in capsid envelopment. N-terminal fusion of a signal sequence to L caused the cotranslational translocation of the pre-S domains into the ER lumen, resulting in a transmembrane topology similar to that of the M protein. This fusion protein did not support virion formation (6), suggesting that the cytoplasmic exposure of pre-S sequences of L is essential for envelopment of the nucleocapsid.

* Mailing address: Department of Medical Microbiology, University of Göttingen, Kreuzbergring 57, D-37075 Göttingen, Germany. Phone: 49 551 39 5759. Fax: 49 551 39 5860. E-mail: VBRUSS@GWDG.DE.

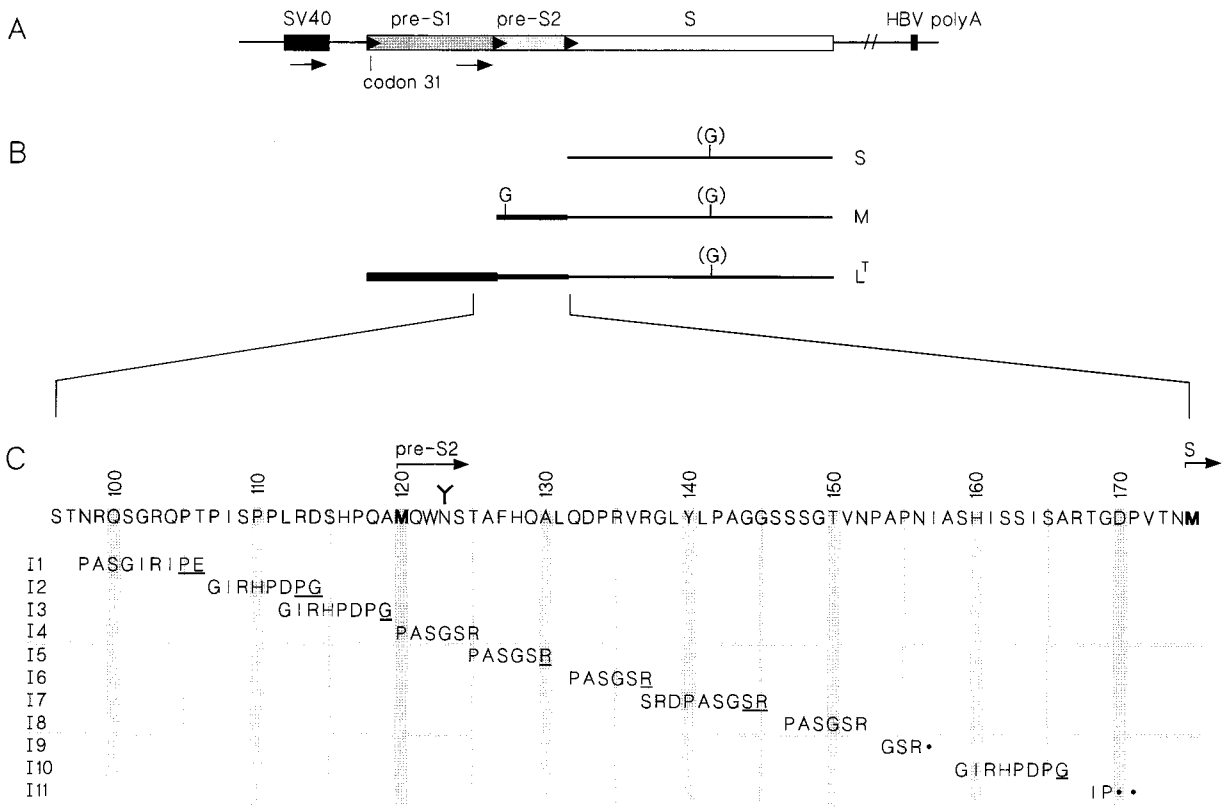

FIG. 1. Map of L mutants (I series). (A) The mutant L proteins were expressed from the SV40 early promoter (black box) construct pSV45-31 which carries the HBV *env* gene and downstream sequences including the HBV polyadenylation site (HBV polyA). The M and S proteins were expressed from their autologous promoter in the pre-S1 region (right arrow). Start codons are indicated by triangles. Pre-S1 codons 2 to 30 were deleted to destroy the inhibitory effect of the L protein on subviral HBsAg particle release. (B) The primary structures of the three HBV envelope proteins L^T (N-terminal deleted L protein), M, and S define the three domains pre-S1 (thick line), pre-S2 (line with intermediate thickness), and S (thin line). The S domain is partially N-glycosylated (G in parentheses). The pre-S2 domain of M carries a glycan (G). (C) The upper line shows the amino acid sequence of the wild-type L protein from Ser 96 to Met 175 (N terminus of the S domain) in single-letter code. Numbers indicate amino acid positions in the pre-S domain. Arrows point to the N termini of the pre-S2 and S domains. The Y stands for the glycan residue linked to Asp 4 of pre-S2 in the M protein. The lower part shows the positions and sequences of the amino acid substitutions in mutants I1 to I11. Underlined amino acids are not substitutions but insertions. Dots indicate deleted amino acids.

Furthermore, truncations of the L protein revealed that the N-terminal 102 aa of pre-S1 were dispensable for virion morphogenesis and secretion (5).

In this work, the C-terminal part of pre-S1 downstream of Gly 102 and the whole pre-S2 domain of the L protein were probed for their relevance in virion formation by a genetic approach. This region was mutagenized by linker substitutions and by double-point mutations. All mutant L proteins were demonstrated to form stable proteins and to be cosecreted with S protein probably as subviral HBsAg particles. Complementation of an HBV genome blocked in L protein expression with the mutant L constructs allowed mapping of a linear sequence of 27 aa residues in the pre-S domain which was required for virion formation. This domain is highly conserved between different HBV isolates and also present in rodent but not in avian hepadnaviruses.

MATERIALS AND METHODS

Plasmids. Plasmid pRVHBV1.5 contains 1.5 copies of the wild-type HBV genome with HBsAg subtype adw2 (34) (EMBL accession no. X02763) in a head-to-tail configuration and has been described elsewhere (3). To generate plasmid pHBV1.5LE$^-$, two nucleotide changes were introduced by site-directed in vitro mutagenesis into pRVHBV1.5, at nucleotide (nt) 196 (numbering of the plus strand of the HBV genome starts with the deoxycytidine of the unique *Eco*RI site) from T to G and at nt 3006 from G to A, resulting in two stop codons in codon 15 of the S gene and codon 51 of pre-S1, respectively (14). Both mutations are silent in the overlapping reading frame for the viral DNA polymerase. Plasmid pSVBX24H was used for expression of the HBV S protein without concomitant synthesis of M and L (15).

Plasmids pSV45-96L and pSV45-31L were used to generate the linker substitution mutants (I series) by exonuclease *Bal* 31 digestions. These plasmids carry the HBV *env* region downstream of the simian virus 40 (SV40) early promoter with deletions of pre-S1 codons 2 to 95 and 2 to 30, respectively, and were constructed from pSV45-96 and pSV45-31 (5), respectively, in two steps. First, the single *Xma*I and *Ngo*MI sites in the vector sequences of pSV45-96 and pSV45-31 were destroyed by double digestions with these enzymes, removal of the resulting 299-bp fragment, and religation. Second, the oligonucleotide 5′C CCGGGATCCGGATGCCGGC3′ was inserted as double-stranded DNA into the single *Eco*RV site (nt 1040) downstream of the S gene in the pSV45-31 derivative (in the orientation shown) and into the single *Pvu*II site of the pSV45-96 derivative 350 nt upstream of the *env* gene sequence (in the opposite orientation). The linker contains three unique sites for the restriction enzymes *Xma*I, *Bsp*EI, and *Ngo*MI. The sticky ends generated by these enzymes are compatible with each other, facilitating the construction of in-frame fusions in the following procedure. Ten micrograms of plasmid pSV45-31L was linearized with *Xba*I (nt 245) at codon 31 of the S gene and incubated with 2 U of exonuclease *Bal* 31 for 0.5 to 4 min. After removal of the exonuclease by phenol extraction, the DNA was digested with *Xma*I at the 5′ end of the linker sequence, the sticky end was filled in, and the large fragment of the plasmid was isolated and circularized by T4 DNA ligase. This generated a series of plasmids carrying the linker fused downstream to pre-S sequences between approximately nt 3140 (codon 97 of pre-S1) and 160 (start of the S gene). The plasmids of ca. 100 clones were sequenced to determine the exact fusion point. Plasmid pSV45-96L was linearized between the SV40 promoter sequence and the HBV pre-S1 sequence with *Hin*dIII and treated with *Bal* 31 in the same way as pSV45-31L. Then the plasmids were cut with *Xma*I at the 3′ end of the linker and religated after filling in of the sticky ends and removal of the small fragment. This procedure generated plasmids carrying the linker fused upstream to the pre-S sequence in a region between approximately nt 3140 and 160. Again, the fusion points of ca.

100 plasmids were determined by sequencing. In the final step, 11 pairs of pSV45-31L/pSV45-96L derivatives were selected and recombined via one of their *Xma*I/*Bsp*EI/*Ngo*MI sites in the linker sequences and their unique *Sac*II sites at nt 1449 in the HBV sequence to yield the final constructs. The resulting plasmids I1 to I11 corresponded to pSV45-31 and carried in-frame substitutions of two to seven codons of the pre-S sequence by linker sequences flanked by insertions of zero to two codons or deletions of zero to two codons (Fig. 1C). The fusion points of the 11 plasmids were sequenced to verify their genotypes.

The double-alanine mutations (A series) were generated by PCR with antisense primers introducing the point mutations and a sense primer hybridizing in the SV40 promoter sequence, using pSV45-31 as a template. The antisense primer overlapped the unique *Eco*RI site (nt 1) in the case of mutants A1 to A7 or the unique *Bam*HI site (nt 26) in the case of mutant A8. The PCR products were cut with *Hin*dIII at the 5′ ends (at the junction between the SV40 promoter and the HBV pre-S sequence) and with *Eco*RI and *Bam*HI, respectively, at the 3′ ends and cloned into *Hin*dIII/*Eco*RI- and *Hin*dIII/*Bam*HI-digested pSV45-31, respectively. The sequences of the PCR-generated regions were determined and were as expected.

Cell culture, transfections, and immunoprecipitations. Transient transfection of COS7 cells, metabolic pulse-chase labeling with [^{35}S]methionine, cell lysis, immunoprecipitation of the HBV envelope proteins with anti-HBV surface protein (HBs) bound to protein A-Sepharose, polyacrylamide gel electrophoresis, and autoradiography were done as described elsewhere (5). Transient transfections of HuH7 cells with cloned HBV DNA or simultaneously with HBV DNA and SV40 expression vectors for envelope proteins were done as described previously (3).

Detection of cytoplasmic nucleocapsids and secreted virions. The culture medium (11 ml) from transiently transfected HuH7 cells was collected, and cells were washed with 12 ml of phosphate-buffered saline, placed on ice, and lysed by incubation with 1.5 ml of 150 mM NaCl–50 mM Tris-Cl (pH 7.5)–5 mM $MgCl_2$–0.2% (vol/vol) Nonidet P-40 for 30 min. The medium and cell lysate were cleared by low-speed centrifugation (10 min at 4,000 × *g*); 0.75 ml of the cleared lysate was stored at −20°C. Ten milliliters of the cleared medium was used to dissolve 3.5 g of CsCl and spun for 42 h at 45,000 rpm and 20°C in a Beckman Ti70.1 rotor. The upper part of the isopycnic density gradient which contained the virions (density of 1.24 g/ml) but no naked nucleocapsids (density of 1.35 g/ml) was harvested by removing and combining 1-ml portions from the top until the meniscus of the remaining gradient reached a density of between 1.29 and 1.30 g/ml. This density was measured by refractive index, using 10 μl of the remaining gradient removed from the top. Combined fractions were dialyzed in collodium bags for 2 h at room temperature against phosphate-buffered saline. Virions were immunoprecipitated from the dialyzed samples by overnight incubation with agitation at 4°C with 3 μl of goat anti-HBs (DAKO Diagnostika, Hamburg, Germany) prebound to 30 μl of swollen protein A-Sepharose CL4B beads (Sigma). The frozen cell lysates were thawed, and one-fourth (corresponding to one-eighth of the total lysate sample) was incubated in parallel with 1 μl of rabbit anti-HBV core (anti-HBc; DAKO) prebound to 10 μl of swollen protein A-beads to immunoprecipitate HBV nucleocapsids. Washing of the protein A-beads, labeling of the viral genome in nucleocapsids or virions by a radioactive endogenous polymerase reaction, preparation of the genome, agarose gel electrophoresis, and autoradiography were done as described previously (3) except that after the endogenous polymerase reaction, 2.5 μg of DNase was added to the cell lysate samples, which were then incubated for 15 min at 37°C prior to proteinase K digestion.

Sequence comparison. Thirty-two HBV sequences were retrieved from the GenEMBL data bank (accession no. D0030, D00220, D00329, D00331, D00630, D12980, D16666, D50489, L08805, M12906, M32138, M38454, M54923, M57663, V00867, X02496, X04615, X14193, X51970, X59795, X65257, X65258, X65259, X68292, X69798, X70185, X72702, X75657, X75658, X75664, X85254, and Z35716) and were compared to the sequence of the isolate used in this work (accession no. X02763) (34). The woodchuck hepatitis virus (WHV) sequence is from reference 11.

RESULTS

Generation of linker substitutions (I series) between Asn 98 and Pro 171 in the L protein. To map sites in the pre-S domains of the L protein which are essential for virion formation, a series of 11 L mutants (I1 to I11) expressed from an SV40 early promoter construct were generated (Fig. 1A and C) (for details, see Materials and Methods). In each of these mutants, a stretch of 4 to 7 consecutive aa was removed from the pre-S sequence and substituted with 2 to 9 unrelated aa by the insertion of an oligodeoxynucleotide. The 5′ boundary of the pre-S region affected by these mutations is Asn 98. This limit was chosen because N-terminal truncations of the L protein demonstrated that the region from Gly 2 to Gly 102 was dispensable for virion formation (5). The series of mutants covers the entire pre-S1 and pre-S2 domain downstream of Asn 98 with the exception of only a few amino acid residues (e.g., Leu 131 and Ser 146).

The mutations were introduced into an L protein derivative which lacks the pre-S1 region from Gly 2 to Leu 30 (L^T). This deletion has been shown to have no effect on virion morphogenesis in transfected cell cultures (5). However, it abolishes the inhibition of subviral particle release (26, 36) which is mediated by the N-terminal region of the wild-type L protein (21, 29). Using the L^T background has the advantage that the mutant L proteins could be tested not only for stability by metabolic pulse-chase labeling but also for release into the culture medium. This was used as an additional criterion for an overall correct folding of the peptide chain because incorrectly folded proteins are usually retained in the ER.

The L mutants were expressed in transiently transfected COS7 cells and labeled with [^{35}S]methionine for 1 h. After 24 h of chase, the medium was harvested, the cells were lysed, and the HBV envelope proteins were immunoprecipitated from both samples with an immune serum against HBs and depicted by polyacrylamide gel electrophoresis and autoradiography (Fig. 2). The parent plasmid produced seven visible protein bands in the cell lysate (lane 24). The lowest band represents the S protein with a molecular mass of 24 kDa. The third band from the bottom shows the M protein (30 kDa), and the sixth band depicts the L^T protein (approximately 37 kDa). The S domain contains a facultative N-glycosylation site at Asn 146 (Fig. 1B) (18). Consequently, all three proteins appear in addition as N-glycosylated derivatives (second, fourth, and seventh bands from the bottom, with molecular masses of 27, 33, and approximately 40 kDa, respectively). The pre-S2 domain contains another N-glycosylation site at Asn 4 which is used only in the case of the M protein (31). The double N-glycosylated form of M (36 kDa) as found in natural HBsAg is only barely visible (sixth band from the bottom) in this system because it is partially superimposed by the truncated L protein. The M and S proteins were expressed from their autologous promoter internal in the pre-S1 region (7). All proteins appear in the culture medium (lane 25), with the known exception of the unglycosylated M protein. The reason for the retention of this protein is unknown. The glycosylated proteins in the medium shift to slightly higher molecular weight due to modifications of their glycans in the Golgi complex during secretion.

All mutants show similar patterns in the cell lysates (Fig. 2, even-numbered lanes) with the exception of I4, where the initiation codon for the M protein was changed (Fig. 1C) and consequently no M protein was expressed (Fig. 2, lane 8), and mutant I5, where the N-glycosylation signal in pre-S2 was changed (NST to NSP) (Fig. 1C) and therefore no double-glycosylated M could be produced (Fig. 2, lane 10). Note that in the case of this mutant, the unglycosylated M protein appears in the medium (lane 11). The mutant I10 produced a lower signal in the cell lysate (lane 10) for unknown reasons. The protein pattern of the mutants in the medium fractions (odd-numbered lanes) was also similar to the pattern produced by the parental plasmid (lane 25). For mutants I5 (lane 11) and I7 (lane 15), the L derivative was secreted with less efficiency. The reason for this is unknown. In summary, all mutant L proteins were produced with sufficient stability to be easily detectable and were released from the cells probably in the form of mixed subviral particles together with S and M proteins.

Complementation of an L-negative HBV genome by L protein mutants. The L and S proteins have been shown to be required for virion formation in transiently transfected human hepatoma cell lines (3). To test whether the mutant L proteins

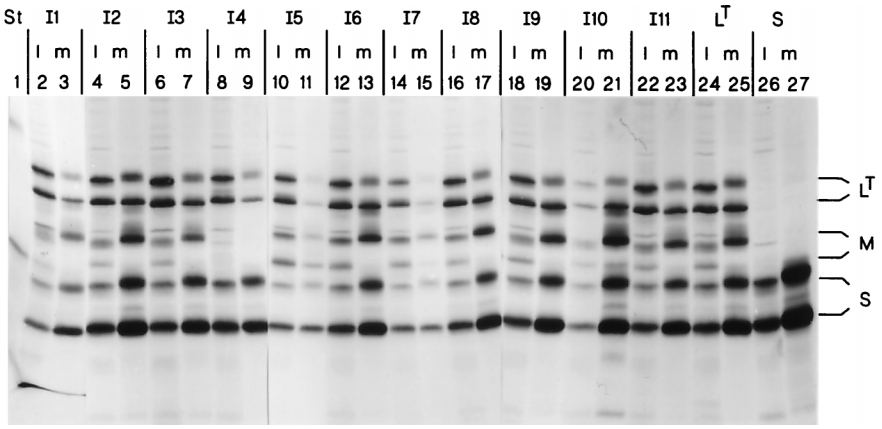

FIG. 2. Stable expression and secretion of L mutants (I series). HBV envelope proteins were immunoprecipitated from COS7 cells transiently transfected with the constructs depicted in Fig. 1C (lanes l) and from the culture medium (lanes m) after pulse-chase labeling and depicted by gel electrophoresis and autoradiography. Lanes 24 and 25 (L^T) show the wild-type S and M proteins as well as the truncated L^T protein generated from the parental construct pSV45-31 (Fig. 1A). The proteins form doublets due to the facultative N-glycosylation of the S domain (Fig. 1B). For lanes 26 and 27, the S protein was expressed without concomitant L and M synthesis. All mutants I1 to I11 formed stable proteins and were released into the culture medium (lanes 2 to 21). In mutant I4, the M protein start codon was mutated. In mutant I5, the N-glycosylation site was changed, resulting in release of the unglycosylated M protein. Lane 1, molecular weight standard (St) (from top to bottom, 46, 30, and 14.3 kDa).

described above could function in hepatitis B virion morphogenesis, HuH7 cells were transiently transfected with the corresponding SV40 vectors for L protein expression together with an HBV genome carrying a stop codon in the pre-S1 region plus a stop codon in the S region. Therefore, envelope protein synthesis was directed only by the SV40 expression vectors described above. All of these vectors support wild-type S protein expression (Fig. 1A and 2), and so differences in the phenotype could be ascribed to the L (and formally also to the M) derivative.

Nucleocapsids were immunoprecipitated from cell lysates; the viral genomes in the capsids were radioactively labeled by the viral polymerase after addition of labeled and unlabeled deoxynucleotides; the genome was isolated, separated on an agarose gel, and visualized by autoradiography (Fig. 3, upper panel). As a reference, 1 μl of a viremic human serum containing approximately 10^6 HBV genomes was directly applied to the endogenous polymerase reaction (lane 1). HBV nucleocapsids could be found in all lysate samples by this technique (lanes 2 to 14, upper panel).

Secreted virions were detected in the culture medium by the following procedure. First, naked nucleocapsids which appear in the media for unknown reasons were separated from virions by an isopycnic CsCl gradient on the basis of their different densities. The upper part of the gradient which contained the virions but no naked nucleocapsids was collected and dialyzed. The virions were then immunoprecipitated with the same anti-HBs antiserum used to bind the labeled surface proteins (Fig. 2) and were detected by the radioactive endogenous polymerase reaction (Fig. 3, lower panel) as described above. Transfection of the *env*-negative HBV genome alone did not allow the production of virions, as expected (lane 3, lower panel). Cotransfection with the parent L construct L^T resulted in the appearance of virions (lane 2, lower panel). Cotransfection with the mutants I1 to I4 failed to produce virions (lanes 4 to 7, lower panel), whereas the remaining mutants I5 to I11 were able to complement the *env*-negative HBV genome (lanes 8 to 14, lower panel). The signals from cytosolic nucleocapsids in cotransfections with mutants I5 and I6 (lanes 8 and 9, upper panel) were not consistently lower in repeated experiments relative to cotransfections with other mutants of the I series.

This analysis clearly separated the mutants into two groups, with negative and positive phenotypes. All negative mutations were clustered upstream of Thr 125 (aa 6 of the pre-S2 region), whereas all mutations with the wild-type phenotype were clustered downstream of Ser 124.

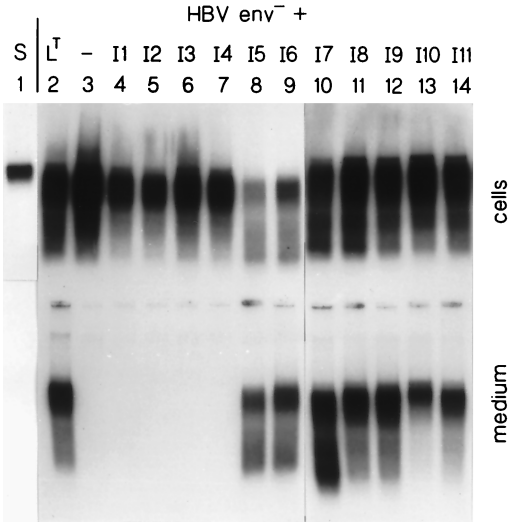

FIG. 3. Complementation of an *env*-negative HBV genome with substitution mutants of the L protein (I series). HuH7 cells were transiently cotransfected in 10-cm-diameter dishes with an *env*-negative HBV genome and SV40 expression vectors for HBV envelope proteins (Fig. 1). Virions from the medium separated from naked nucleocapsids by CsCl gradient centrifugation (lower panel) and one-eighth of the nucleocapsids from the cells (upper panel) were immunoprecipitated with anti-HBs and anti-HBc, respectively; the viral genomes were labeled by a radioactive endogenous polymerase reaction and visualized by agarose gel electrophoresis and autoradiography. One microliter of a viremic serum (S) containing approximately 10^6 virions was used for comparison (lane 1). Cytoplasmic nucleocapsids were detected in all cases (lanes 2 to 14, upper panel). The *env*-negative HBV genome was blocked in virion formation (lane 3, lower panel), and the defect was complemented by the parental construct L^T (lane 2, lower panel), as expected. Mutants I1 to I4 (lanes 4 to 7) were unable to complement, whereas mutants I5 to I11 (lanes 8 to 14) had a wild-type phenotype. The autoradiogram was exposed for 24 h.

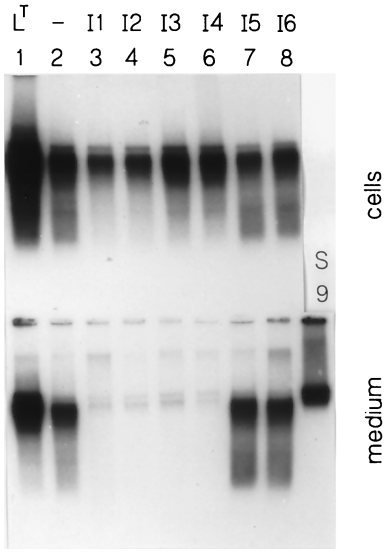

FIG. 4. Transdominant negative phenotype of L mutants I1 to I4. HuH7 cells were transfected with a wild-type HBV genome alone (lane 2) and together with the parental construct L^T (lane 1) or with mutants I1 to I6 (lanes 3 to 8). Cytoplasmic nucleocapsids (upper panel) and secreted virions (lower panel) were detected as for Fig. 3. All four mutants I1 to I4 which were not functional in virion morphogenesis (Fig. 3) had a transdominant negative phenotype and suppressed the release of virions into the medium (lanes 3 to 6, lower panel), in contrast to the two mutants I5 and I6 with wild-type phenotype (lanes 7 and 8). For lane 9, 1 μl of a viremic serum (S) with approximately 10^6 virions was used.

Transdominant negative phenotype of L mutants defective in virion morphogenesis. The L protein forms disulfide-linked homodimers and heterodimers with M and S proteins (35). Because of this dimeric state and also because of the assumption that envelopment of a nucleocapsid requires the cooperation of multiple L proteins, it was anticipated that the L mutants with the negative phenotype would act in a transdominant negative fashion. To test this hypothesis, a wild-type HBV genome was cotransfected with L mutants I1 to I6 and the secretion of virions was monitored (Fig. 4) as described above. Intracellular nucleocapsids could be found in all transfections (Fig. 4, upper panel). Cotransfection of the parental L^T construct did not alter the ratio of intracellular nucleocapsids to secreted virions (compare lanes 1 in the upper and lower panels) relative to the transfection of the genome alone (lane 2). The same result could be observed for the two mutants I5 and I6 with the wild-type phenotype used in this assay (lanes 7 and 8). However, all four mutants I1 to I4, which did not support virion formation, suppressed the release of wild-type virions into the medium (lanes 3 to 6) approximately by a factor of 10 (estimated from autoradiograms with different exposure times [data not shown]). This result demonstrates the transdominant negative character of mutants I1 to I4.

Generation of double-alanine mutants (A series) between Asn 98 and Asn 123 in the L protein. The analysis described above together with the phenotype of N-terminal truncated L proteins (5) suggested that the pre-S sequence between Arg 103 and Ser 124 is of special importance for hepatitis B virion formation. A hybrid protein consisting of the pre-S1 and pre-S2 domains of WHV (a rodent hepatitis B virus) fused to the HBV S domain could be incorporated into the envelope of an L-negative HBV. The same was not possible with a similar fusion protein consisting of the duck hepatitis B virus (DHBV) pre-S sequence and the HBV S domain (15). This demonstrated a functional homology between the pre-S sequences of HBV and WHV in virion morphogenesis which might be reflected by a corresponding sequence homology. Comparison of the pre-S sequences of HBV and WHV showed a low frequency of identical amino acid residues (approximately 20%) with the exception of one stretch in the HBV pre-S sequence between Thr 97 and Gln 129, where 21 of 33 aa (64%) are identical to the WHV sequence (Fig. 5A). This region includes and is almost identical to the domain from Arg 103 to Ser 124 which was mapped to be important for virion morphogenesis in the experiments described above. The pre-S domain of DHBV shows no detectable sequence homology to the HBV pre-S sequence (30).

To probe the importance of individual amino acids in this region for virion formation, a number of double-point mutants were generated in the L^T background by simultaneous alteration of two adjacent amino acid residues to alanine residues (A series) (Fig. 5B). In mutants A1, A2, A3, A5, and A6, both altered amino acid residues are conserved between HBV and WHV; in mutant A7, the conserved Asp 123 and the unconserved Trp 122 were changed; and in mutant A4, both affected amino acid residues were not conserved.

The seven mutant L proteins were again first analyzed for stability and secretion competence in COS7 cells (Fig. 6). All mutants were stably expressed (even-numbered lanes) and appeared in the culture medium (odd-numbered lanes) like the parental protein (lanes 2 and 3). Note that in mutant A7, the N-glycosylation signal of the M protein was changed (NST to AST), resulting in secretion of the unglycosylated M derivative (lane 17) as in the case of mutant I5 (Fig. 2, lane 11). Mutant A1 expressed less S protein (lane 4) than did the parental construct (lane 2) or mutants A2 to A7 (lanes 6, 8, 10, 12, 14, and 16) but roughly equal amounts of M protein. The reason for this is not clear.

Complementation of an L-negative HBV genome by double-alanine mutants. Only mutant A4, not mutants A1 to A3 and A5 to A7, was able to complement the *env*-negative HBV genome (Fig. 7). A4 is the only construct in which both affected amino acid residues were not conserved between HBV and WHV. Apparently, the exchange of two adjacent conserved amino acid residues in this region by two alanine residues rendered the L protein dysfunctional in virion formation without constraining its ability to be secreted.

Sequence comparison with other HBV isolates. The pre-S sequence between Val 60 and Thr 150 of the HBV strain with HBsAg subtype adw2 used in this study (34) was compared to the homologous regions of 32 HBV isolates documented in the EMBL database (see Materials and Methods) (Fig. 8). The stretch from Pro 92 to Thr 125 identified in this work as important for virion formation is relatively conserved, whereas the sequences upstream (Ile 84 to Ile 91) and downstream (Ala 126 to Gly 138) are relatively variable. Especially Asn 98, Arg 99, Gly 102, Arg 103, Pro 105, Thr 106, Arg 113, Trp 122, and Asn 123, which were shown by mutants A1 to A3 and A5 to A7 to be important for virion formation, are absolutely conserved; on the other hand, Ile 108 and Ser 109, which could be changed to alanines without blocking virion formation (mutant A4), display some variability (Fig. 8).

DISCUSSION

Previous work has demonstrated that the pre-S domain of the L protein has an important function in the formation of hepatitis B virions (3, 6). N-terminal truncations of the 174-aa-long pre-S domain up to Arg 103 were tolerated (5). Further truncations changed the transmembrane topology of the L derivative, with luminal disposition of the shortened pre-S re-

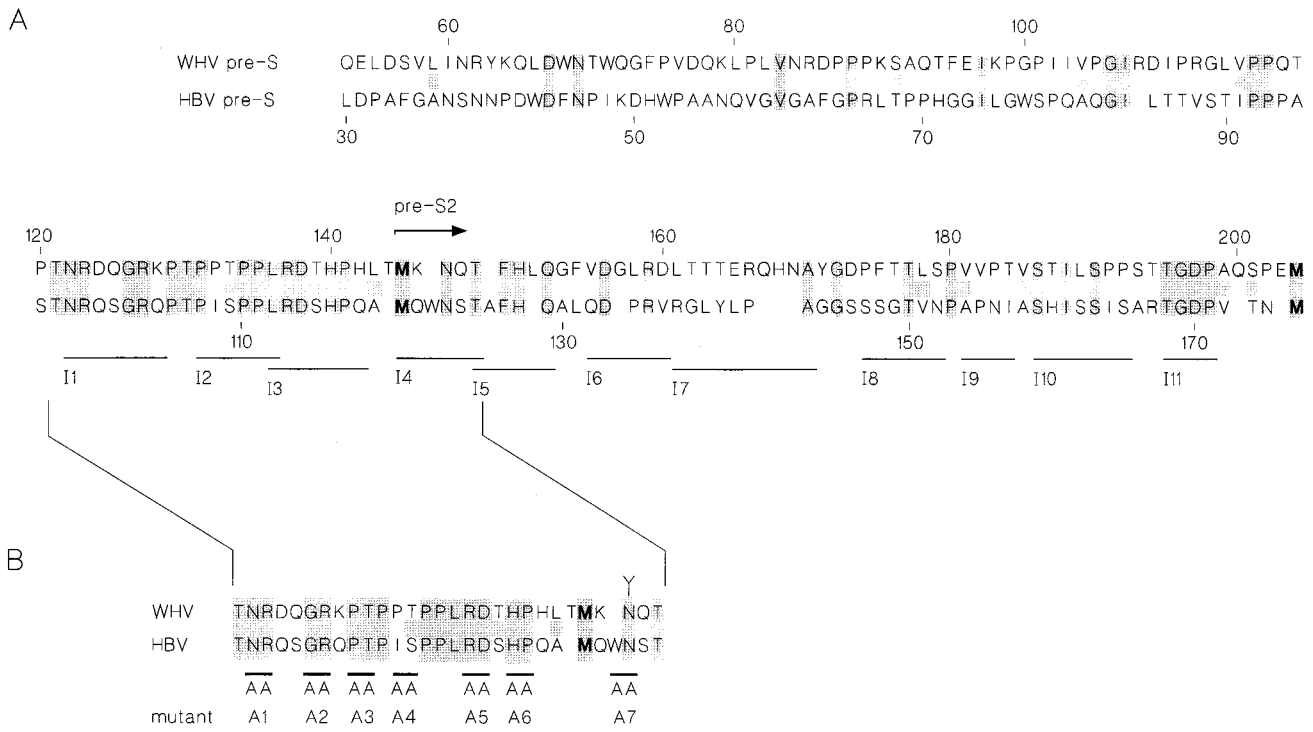

FIG. 5. Comparison of HBV and WHV pre-S sequences and map of L double-point mutants (A series). (A) The WHV pre-S amino acid sequence (single-letter code) from Gln 53 to Met 204 (upper line) is compared with the homologous HBV pre-S sequence from Leu 30 to Met 175 (lower line). Identical residues are emphasized by a gray background; similar amino acids (A = V, I, or L; E = D; S = T) are indicated by gray squares. The N termini of the M and S domains are indicated by boldfaced M. Regions affected by the substitution mutations I1 to I11 (Fig. 1C) are underlined. The HBV pre-S sequence from Thr 97 to Pro 117 is highly conserved between the two viruses. (B) In mutants A1 to A7, two consecutive amino acids of the HBV pre-S sequence were changed simultaneously to alanine residues in the L^T background (A series of mutants); e.g., in A1, asparagine 98 and arginine 99 were mutated to alanine residues. The Y indicates the glycan linked to asparagine 4 of the HBV M protein. The N terminus of the pre-S2 domain is indicated by a boldfaced M.

gion and inability to support virion morphogenesis. Here I show that mutations in the C-terminal part of cytoplasmically exposed pre-S blocked virion formation only when a small region between Asn 98 and Ser 124 was affected. The importance of this region for the HBV life cycle is emphasized by its conservation among orthohepadnaviruses (Fig. 5A and 8).

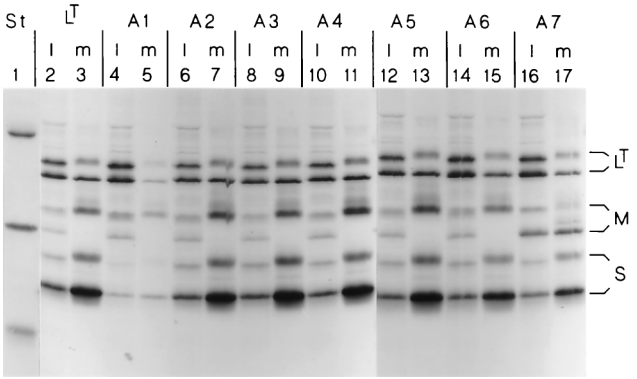

FIG. 6. Stable expression and secretion of L mutants (A series). The mutant L proteins of the A series (Fig. 5B) were expressed in COS7 cells and processed as for Fig. 2. All mutants (lanes 4 to 17) formed stable protein within the cells (lanes l) and were released into the culture medium (lanes m) like the parental construct L^T (lanes 2 and 3). Note that mutant A7 lost the N-glycosylation site in pre-S2, resulting in the secretion of unglycosylated M protein. In mutant A1, the activity of the autologous promoter for M and S protein expression seems to be impaired. Lane 1, molecular weight marker (St) as in Fig. 2.

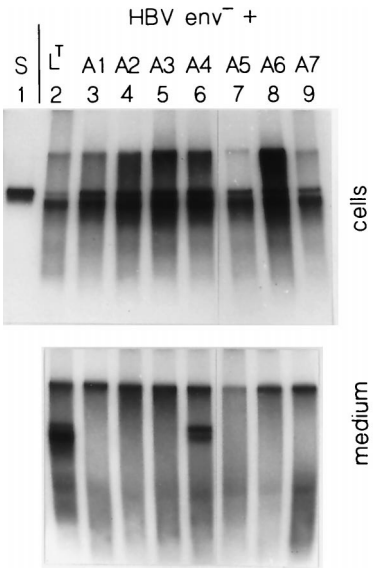

FIG. 7. Complementation of an L-negative HBV genome with double-point mutants of the L protein (A series). HuH7 cells were cotransfected with the *env*-negative HBV genome and expression vectors for the double-point mutants A1 to A7 of the L protein (Fig. 5B). Nucleocapsids from cells (upper panel) and virions from the culture medium (lower panel) were monitored as for Fig. 3. Mutant A4 allowed virion formation (lane 6) like the N-terminal truncated L protein L^T (lane 2). Mutants A1 to A3 (lanes 3 to 5) and A5 to A7 (lanes 7 to 9) showed a negative phenotype. Lane 1 shows the signal from 1 μl of viremic serum (S) containing approximately 10^6 virions.

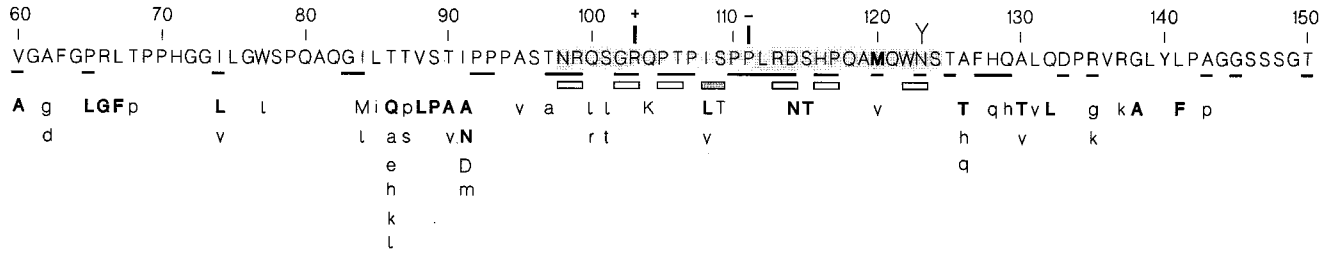

FIG. 8. Sequence variability in the pre-S region among different HBV isolates. The pre-S amino acid sequence from Val 60 to Thr 150 of the isolate used in this work (HBsAg subtype adw2) (34) is shown in one-letter code. Residues conserved in the WHV sequence are underlined (see also Fig. 5A). Residues affected in the A series of L mutants are marked by boxes. Open boxes indicate mutants with negative phenotype; the dark box indicates the A4 mutation with wild-type phenotype. The Y stands for the glycan linked to Asp 4 of the M protein. The N-terminal Met residue of M is in boldface. The region from Asn 98 to Ser 124 mapped by the substitution mutants to be important for virion formation (Fig. 3) is emphasized by a gray background. The two vertical bars point to the N termini of two truncated L proteins used in another study (5). The mutant labeled with a plus sign supported virion formation, in contrast to the one labeled with a minus sign. The pre-S region was compared to the homologous sequences of 32 published HBV isolates (see Materials and Methods). Lowercase (uppercase; boldfaced) letters indicate that not more than two (three, four, or five; more than five) sequences carry the corresponding residue at the indicated position.

Correspondingly, the WHV pre-S sequence was able to replace the HBV pre-S sequence in virion formation. The pre-S domain of avihepadnaviruses does not contain a homologous sequence on a primary structural level and could not substitute for the HBV pre-S sequence in virion morphogenesis (15). However, mutations in the region from Pro 116 to Leu 132 of the 161-aa-long DHBV pre-S domain by linker substitutions had a deleterious effect on virion formation (22). The similarity in position of the two regions in HBV and DHBV pre-S which were sensitive to mutations with respect to virion formation suggests that they have similar functions in this process.

The experiments presented here do not allow the function which is lost in the L mutants with negative phenotype to be defined. Several possibilities, however, can be envisioned. For example, the transmembrane topology of the mutant L proteins could be altered. Prange and Streeck (29) showed that a deletion of aa 81 to 118 of pre-S1 (numbers adjusted to HBsAg subtype ad used in this work) caused the cotranslational translocation of pre-S into the ER lumen. In this mutant, the potential N-glycosylation site at Asn 4 in pre-S2 was used, resulting in additional protein bands of higher molecular weight in polyacrylamide gels. Former experiments demonstrated that this N-glycosylation site can be used as an indirect indicator for the topology of the protein (5, 6). Inspection of the protein bands in Fig. 2 and 6 shows that the mutant L proteins were expressed in unglycosylated and monoglycosylated form and that a higher glycosylated species is missing (e.g., in Fig. 2, compare lane 24 with even-numbered lanes 2 to 22). This finding suggests that the topology of the mutant L chains is wild type and makes it unlikely that an altered transmembrane topology is the cause of the negative phenotype. Also, the overall folding of the L mutants is probably normal because they were stable and released from the cells, whereas misfolded proteins are usually retained in the ER and degraded (2, 17). It is also not likely that the secretion of completed virions from the ER lumen to the outside of the cells was inhibited because the L derivatives were efficiently secreted from COS7 cells (e.g., Fig. 2).

This leaves the envelopment of the nucleocapsid as a potentially blocked step in virion morphogenesis. A plausible interpretation of the results presented here is that the pre-S region between Arg 103 and Ser 124 directly contacts the nucleocapsid during budding, similar to the model developed for alphavirus formation (37). Both virus families do not have a matrix protein and require the expression of envelope proteins for nucleocapsid envelopment (33), in contrast to, e.g., retroviruses (16). This view is supported by the recent finding that the binding of synthetic pre-S peptides to core particles isolated from human liver was strongest for a peptide corresponding to the pre-S domain from Pro 107 to Phe 127 (numbering adjusted from HBsAg subtype ay to subtype ad) (27).

The screening of a phage peptide display library for hexapeptides binding to bacterially expressed core particles did not select sequences from this region (8). Possibly, hexapeptides are too short to confer sufficient binding. It is also possible that core particles expressed in bacteria cannot interact with envelope protein domains in a relevant manner. This is not unlikely since immature nucleocapsids were shown to be incompetent for envelopment and the formation of virions was dependent on viral DNA synthesis in the nucleocapsid (14). Recently, a mutant virus genome with an in-frame deletion of pre-S1 aa 47 to 107 (corresponding to aa 57 to 118 in HBsAg subtype ad) was isolated from a chronically infected patient (23). Cotransfection of HepG2 cells with this genome and an expression vector for S protein resulted in the secretion of viral particles into the medium which could be immunoprecipitated with a monoclonal anti-pre-S1 antibody. The reason for the discrepancy of these results and the analysis presented here is not clear.

Two of the mutations which blocked virion formation (I4 and A7) affected not only the L but also the M protein, and so formally their negative phenotype cannot be attributed to L alone. However, in mutant I4, the expression of M was blocked due to a missense mutation of the start codon (Fig. 1C), and previous work has shown that M is not required for virion formation. In mutant A7, Trp 2 and Asp 3 of the M protein have been altered to alanine residues (Fig. 5B). This part of the molecule is on the luminal side of the ER membrane and exposed on the surface of virions. It should therefore not be directly involved in envelopment of the nucleocapsid or budding.

Based on the work presented here, the pre-S1 domain can be separated into two subdomains. The N-terminal region of pre-S1 probably serves as a ligand for a viral receptor (20). This function is realized by L proteins exposing the pre-S domains to the outside of the particle. The C-terminal part between Arg 103 and Ser 124 has a pivotal function in virion morphogenesis which is realized by L proteins with internal pre-S domains.

ACKNOWLEDGMENTS

I am grateful to Stephanie Grethe for help with the sequence comparison. I thank Wolfram Gerlich for critical reading of the manuscript and Reiner Thomssen for support.

This work was supported by Deutsche Forschungsgemeinschaft grant Br 1018/2-3.

REFERENCES

1. **Acs, G., M. A. Sells, R. H. Purcell, P. Price, R. Engle, M. Shapiro, and H. Popper.** 1987. Hepatitis B virus produced by transfected Hep G2 cells causes hepatitis in chimpanzees. Proc. Natl. Acad. Sci. USA **84:**4641–4644.
2. **Bruss, V., and D. Ganem.** 1991. Mutational analysis of hepatitis B surface antigen particle assembly and secretion. J. Virol. **65:**3813–3820.
3. **Bruss, V., and D. Ganem.** 1991. The role of envelope proteins in hepatitis B virus assembly. Proc. Natl. Acad. Sci. USA **88:**1059–1063.
4. **Bruss, V., X. Lu, R. Thomssen, and W. H. Gerlich.** 1994. Post-translational alterations in transmembrane topology of the hepatitis B virus large envelope protein. EMBO J. **13:**2273–2279.
5. **Bruss, V., and R. Thomssen.** 1994. Mapping a region of the large envelope protein required for hepatitis B virion maturation. J. Virol. **68:**1643–1650.
6. **Bruss, V., and K. Vieluf.** 1995. Functions of the internal pre-S domain of the large surface protein in hepatitis B virus particle morphogenesis. J. Virol. **69:**6652–6657.
7. **Cattaneo, R., H. Will, N. Hernandez, and H. Schaller.** 1983. Signals regulating hepatitis B surface antigen transcription. Nature (London) **305:**336–338.
8. **Dyson, M. R., and K. Murray.** 1995. Selection of peptide inhibitors of interactions involved in complex protein assemblies: association of the core and surface antigens of hepatitis B virus. Proc. Natl. Acad. Sci. USA **92:**2194–2198.
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 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. **Galibert, F., T. N. Chen, and E. Mandart.** 1982. Nucleotide sequence of a cloned woodchuck hepatitis virus genome: comparison with the hepatitis B virus sequence. J. Virol. **41:**51–65.
12. **Ganem, D.** 1991. Assembly of hepadnaviral virions and subviral particles. Curr. Top. Microbiol. Immunol. **168:**61–83.
13. **Ganem, D., and H. E. Varmus.** 1987. The molecular biology of the hepatitis B viruses. Annu. Rev. Biochem. **56:**651–693.
14. **Gerelsaikhan, T., J. E. Tavis, and V. Bruss.** 1996. Hepatitis B virus nucleocapsid envelopment does not occur without genomic DNA synthesis. J. Virol. **70:**4269–4274.
15. **Gerhardt, E., and V. Bruss.** 1995. Phenotypic mixing of rodent but not avian hepadnavirus surface proteins into human hepatitis B virus particles. J. Virol. **69:**1201–1208.
16. **Gheysen, D., E. Jacobs, F. de Foresta, C. Thiriart, M. Francotte, D. Thines, and M. De Wilde.** 1989. Assembly and release of HIV-1 precursor $Pr55^{gag}$ virus-like particles from recombinant baculovirus-infected insect cells. Cell **59:**103–112.
17. **Hebert, D. N., B. Foellmer, and A. Helenius.** 1995. Glucose trimming and reglycosylation determine glycoprotein association with calnexin in the endoplasmic reticulum. Cell **81:**425–433.
18. **Heermann, K., U. Goldmann, W. Schwartz, T. Seyffarth, H. Baumgarten, and W. H. Gerlich.** 1984. Large surface proteins of hepatitis B virus containing the pre-S sequence. J. Virol. **52:**396–402.
19. **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.
20. **Ishikawa, T., K. Kuroki, R. Lenhoff, J. Summers, and D. Ganem.** 1994. Analysis of the binding of a host cell surface glycoprotein to the pre-S protein of duck hepatitis B virus. Virology **202:**1061–1064.
21. **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.
22. **Lenhoff, R. J., and J. Summers.** 1994. Coordinate regulation of replication and virus assembly by the large envelope protein of an avian hepadnavirus. J. Virol. **68:**4565–4571.
23. **Melegari, M., P. P. Scaglioni, and J. R. Wands.** 1997. The small envelope protein is required for secretion of a naturally occurring hepatitis B virus mutant with pre-S1 deleted. J. Virol. **71:**5449–5454.
24. **Nassal, M., and H. Schaller.** 1993. Hepatitis B virus replication. Trends Microbiol. **1:**221–228.
25. **Ostapchuk, P., P. Hearing, and D. Ganem.** 1994. A dramatic shift in the transmembrane topology of a viral envelope glycoprotein accompanies hepatitis B viral morphogenesis. EMBO J. **13:**1048–1057.
26. **Persing, D. H., H. E. Varmus, and D. Ganem.** 1986. Inhibition of secretion of hepatitis B surface antigen by a related presurface polypeptide. Science **234:**1388–1391.
27. **Poisson, F., A. Severac, C. Hourioux, A. Goudeau, and P. Roingeard.** 1997. Both pre-S1 and S domains of hepatitis B virus envelope proteins interact with the core particle. Virology **228:**115–120.
28. **Prange, R., A. Clemen, and R. E. Streeck.** 1991. Myristylation is involved in intracellular retention of hepatitis B virus envelope protein. J. Virol. **65:**3919–3923.
29. **Prange, R., and R. E. Streeck.** 1995. Novel transmembrane topology of the hepatitis B virus envelope proteins. EMBO J. **14:**247–256.
30. **Sprengel, R., C. Kuhn, H. Will, and H. Schaller.** 1985. Comparative sequence analysis of duck and human hepatitis B virus genomes. J. Med. Virol. **15:**323–333.
31. **Stibbe, W., and W. H. Gerlich.** 1983. Structural relationships between minor and major proteins of hepatitis B surface antigen. J. Virol. **46:**626–628.
32. **Summers, J., and W. S. Mason.** 1982. Replication of the genome of a hepatitis B-like virus by reverse transcription of an RNA intermediate. Cell **29:**403–425.
33. **Suomalainen, M., P. Liljeström, and H. Garoff.** 1992. Spike protein-nucleocapsid interactions drive the budding of alphaviruses. J. Virol. **66:**4737–4747.
34. **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.
35. **Wunderlich, G., and V. Bruss.** 1996. Characterization of early hepatitis B virus surface protein oligomers. Arch. Virol. **141:**1191–1205.
36. **Xu, Z., V. Bruss, and T. S. B. Yen.** 1997. Formation of intracellular particles by hepatitis B virus large surface protein. J. Virol. **71:**5487–5494.
37. **Zhao, H., B. Lindqvist, H. Garoff, C. von Bonsdorff, and P. Liljeström.** 1994. A tyrosine-based motif in the cytoplasmic domain of the alphavirus envelope protein is essential for budding. EMBO J. **13:**4204–4211.