

Structure and Glycosylation Patterns of Surface Proteins from Woodchuck Hepatitis Virus

TANJA K. TOLLE,¹ DIETER GLEBE,¹ MONICA LINDER,² DIETMAR LINDER,²
SIGRID SCHMITT,² RUDOLF GEYER,² AND WOLFRAM H. GERLICH^{2*}

*Institute of Medical Virology¹ and Institute of Biochemistry,² Clinicum and
Medical School of Justus-Liebig-University, Giessen, Germany*

Received 5 June 1998/Accepted 20 August 1998

Woodchucks chronically infected with woodchuck hepatitis virus (WHV) are a valuable model for human hepatitis B virus (HBV) in studies of pathogenesis, immunity, and antiviral therapy. For this reason, substantial efforts to characterize both the similarities and the differences between HBV and WHV are being made. The structure of the WHV surface proteins (WHs proteins) has not yet been adequately elucidated. The bands that would be expected for glycosylated and nonglycosylated small (S) WHs protein are found by sodium dodecyl sulfate gel electrophoresis of purified WHs protein, but the bands corresponding to the middle (M) and large (L) WHs proteins of HBV are not seen at the expected sizes, even though the sequences of the WHV and HBV surface protein genes are 60% homologous. By amino-terminal sequencing we have identified two bands at 41 and 45 kDa as the MWHs proteins, 8 kDa larger than expected. We have also confirmed that two bands at 24 and 27 kDa are SWHs proteins. A protein of 49 kDa was blocked at the N terminus, which using immunoblotting with an antiserum against WHV pre-S1 (positions 126 to 146) was identified, together with a part of the 45-kDa protein, as glycosylated and nonglycosylated LWHs protein of the expected size. Sialidase and O-glycosidase digestion showed that the larger size of MWHs protein results from the presence of O glycoside groups which are probably in the pre-S2 domain of MWHs protein. Since the pre-S2 domains of HBV and WHV have similar numbers of potential O glycosylation sites, it appears to be likely that the glycosyltransferases act differently on the viral proteins in woodchucks and humans.

Woodchuck hepatitis virus (WHV) (39) and human hepatitis B virus (HBV) are members of the virus family *Hepadnaviridae* (27) which cause persistent infections of hepatocytes, persistent viremia, and antigenemia (18) (for review, see reference 13). WHV infection in woodchucks serves as an important model for pathogenesis and immune reactions caused by HBV in humans (32). Furthermore, preclinical testing of antiviral treatment against HBV is done with WHV-infected woodchucks (2, 20). The basis of a novel treatment strategy for hepadnavirus infection could be the inhibition of virus secretion by glucosidase I inhibitors (3), which prevent trimming and modification of N-glycans. Middle human hepatitis B surface protein (MHBs protein) is the most extensively glycosylated protein of HBV, because the Asn4 in its pre-S2 domain is always glycosylated in secreted HBs antigen (HBsAg) (36). Trimming of the pre-S2 glycan is indeed necessary for proper intracellular transport and secretion. Thus, pre-S2 glycosylation is a potential target for treatment of HBV infections (24, 26). This approach has recently been tested successfully with WHV-infected woodchucks (2). WHV infection could, furthermore, be a model for immunotherapy of HBV infection. For these reasons it is highly relevant to know all similarities and differences between WHV surface (WHs) and HBs proteins.

Besides the complete virus, the hepadnavirus-infected hepatocytes secrete particles consisting of surplus surface proteins, named HBsAg for HBV and WHsAg for WHV. Secreted HBsAg and WHsAg form spherical and filamentous particles with a diameter of ca. 20 nm. HBsAg and the HBV envelope consist of three co-carboxy-terminal HBs proteins of small (S), middle

(M), and large (L) size. The sequence of the LHBs protein gene is, thus, divided by the three start codons into sequences encoding an N-terminal pre-S1, a central pre-S2, and a C-terminal S domain (Fig. 1). A facultative N glycosylation in the S domain generates typical doublet patterns of the three HBs proteins in sodium dodecyl sulfate (SDS) gel electrophoresis (17, 37) as shown in Fig. 2. The organization of the genes for the WHs proteins is very similar to that of the HBs, with conserved start codons for S-, M-, and LWHs proteins and conserved N glycosylation sites (12) (Fig. 1). Nevertheless, the staining pattern of purified spherical or filamentous WHs protein particles in polyacrylamide gel electrophoresis (PAGE) shows only the doublet corresponding to SWHs proteins. Bands at the expected apparent molecular masses (Fig. 1 and 2) of MWHs proteins were absent, and the expected doublet for LWHs proteins could not be clearly identified within a triplet of bands at 41, 45, and 49 kDa (Fig. 2).

The protein composition of WHsAg has been studied previously (10, 30, 33, 35), but the identification of the larger protein bands was incomplete. Since the unexplained protein pattern of WHsAg in SDS gel electrophoresis led to the question of whether WHV has a completely different expression strategy for its surface proteins, we determined their primary structure and searched for posttranslational modifications different from those in HBs proteins. We identified significant differences in the modification of MWHs protein.

MATERIALS AND METHODS

Isolation of WHs protein particles. Sera from chronically WHV-infected woodchucks were a kind gift from Maria Seifer, Bristol-Myers Squibb (Hartford, Conn.).

Ten milliliters of serum was ultracentrifuged through a discontinuous sucrose density gradient (15, 25, 35, 45, and 60% [wt/wt]) in TNE buffer (20 mM Tris-HCl [pH 7.4], 140 mM NaCl, 1 mM EDTA) for 15 h at 25,000 rpm with an SW 28.38 rotor (Beckman, Munich, Germany). The WHs protein fractions con-

* Corresponding author. Mailing address: Institute of Medical Virology, Frankfurter Strasse 107, D 35392 Giessen, Germany. Phone: 49-641-99 41200. Fax: 49-641-99 41209. E-mail: wolfram.h.gerlich@viro.med.uni-giessen.de.

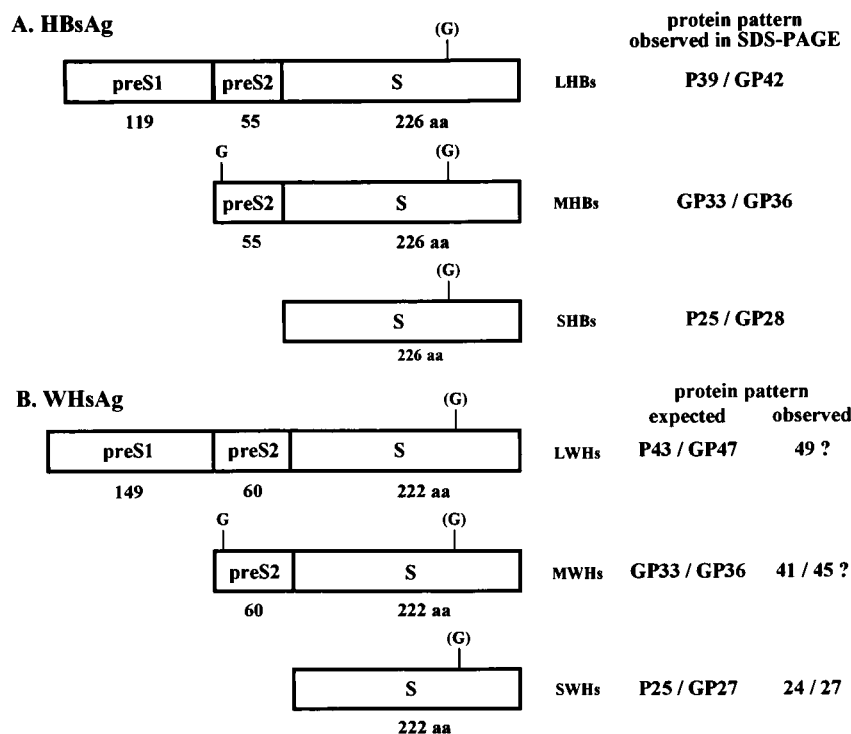


FIG. 1. Organization of the surface protein gene from HBV and WHV and nomenclature of the viral surface proteins. The observed apparent molecular masses, in kilodaltons, for the WHs proteins refer to the PAGE shown in Fig. 2A (right panel), whereas the expected molecular masses originate from the extrapolation of the size of HBs proteins to WHs proteins. aa, amino acids.

taining 35 to 45% sucrose were analyzed by SDS-PAGE and silver stained for the presence of WHsAg. Fractions were pooled, adjusted with solid CsCl to 1.30 g/ml, and layered within a CsCl density gradient ranging from 1.16 to 1.35 g/ml in TNE buffer for 36 h at 25,000 rpm in an SW 28.38 rotor (Beckman). WHsAg-containing fractions were dialyzed and purified by sedimenting in a discontinuous sucrose density gradient (5, 10, 15, 20, 25, and 60% [wt/wt]) in TNE buffer for 15 h at 25,000 rpm with an SW 28.38 rotor. The filament-rich fractions were separated from the spheres, pooled, and concentrated in a Centrifu-100 ultrafiltration device (Millipore, Eschborn, Germany). The concentration of purified WHsAg was estimated by the optical density at 280 nm, assuming a value of 5.1 for 1 mg/ml. The estimate on the specific optical density of WHsAg was based on the value of 4.3 for HBsAg (14) and the fact that SWHs protein has 17 Trp and 7 Tyr residues instead of the 14 Trp and 5 Tyr residues in SHBs protein (12). Antigens were stored at -20°C .

SDS-PAGE and immunoblotting of WHs protein particles. The WHs protein particles were analyzed in Laemmli buffer containing 4% dithiothreitol and 4% β -mercaptoethanol on 12% polyacrylamide gels. After SDS-PAGE the gels were silver stained or the proteins were blotted onto a polyvinylidene difluoride membrane (Millipore). The membranes were blocked in 3% low-fat milk powder in phosphate-buffered saline (PBS) and incubated with polyclonal rabbit anti-WHsAg serum diluted 1:1,000 or with rabbit polyclonal anti-pre-S1 peptide (amino acid residues 126 to 146) diluted 1:200 in PBS with 1% low-fat milk powder for 1 h at 37°C . Thereafter, donkey anti-rabbit antibody conjugated with horseradish peroxidase (Dianova, Hamburg, Germany) was used at a concentration of 1:2,000 and developed with diaminobenzidine- H_2O_2 substrate (Sigma, Deisenhofen, Germany) or with an enhanced chemiluminescent-light detection kit (Boehringer, Mannheim, Germany).

Synthesis of peptides. The peptides were synthesized on a model 431A peptide synthesizer (Applied Biosystems Inc., Foster City, Calif.) by using the 9-fluorenylmethoxycarbonyl protection scheme and amino acid activation with hydroxybenzotriazole-dicyclocarbodiimide (19). TentaGel PAP (Rapp-Polymer, Tübingen, Germany), a support yielding peptide conjugated to polyethylene glycol after cleavage from the resin with trifluoroacetic acid (TFA), was used for the synthesis of the four-branched multiple antigenic peptide (41).

Generation of antisera. Rabbits were injected subcutaneously with $10\text{ }\mu\text{g}$ of highly purified native WHs protein filaments in complete Freund's adjuvant and given two booster injections with incomplete Freund's adjuvant after 2 and 4 weeks. Blood was collected 10 days after the second booster injection (Eurogentec, Seraing, Belgium) and tested for reactivity to WHsAg by Western blotting.

Peptide antisera to LWHs protein were generated by immunization of rabbits with $200\text{ }\mu\text{g}$ of synthetic peptide of pre-S1 amino acid residues 126 to 146

(TNRDQGRKPTPTPLRDTHP) linked by a lysine to the four-branched carrier molecule by using the immunization protocol described above; however, four booster injections were given.

Treatment of WHsAg with different glycosidases. The N-glycans of WHsAg were removed with peptidoglycanase F (PNGase F) from *Flavobacterium meningosepticum* (New England Biolabs, Schwalbach, Germany) under native and denaturing conditions. One microgram of WHsAg was diluted in reaction buffer (50 mM sodium phosphate buffer containing 1% Nonidet P-40) and incubated with 500 U of PNGase F for 1 h at 37°C . For denaturing conditions, the WHsAg was incubated for 10 min at 100°C with 0.5% SDS and 1% β -mercaptoethanol prior to addition of reaction buffer and PNGase F at 37°C .

Three different sialidases (from *Arthrobacter ureafaciens* [Calbiochem, Bad Soden, Germany], Newcastle disease virus [Boehringer], and *Salmonella typhimurium* [New England Biolabs]) were used to examine the sialylation of N- and O-glycans. The digestion was done for 1 h at 37°C with 1 mU of enzyme per μg of WHsAg in PBS (*A. ureafaciens*) or the supplied reaction buffers specified below.

α -Fucosidase (Boehringer) and β -N-acetylhexosaminidase (Sigma) were used to remove the corresponding substituents for 1 h at 37°C .

O-Glycosidase from *Diplococcus pneumoniae* (Boehringer) was used to release Gal β (1-3)GalNAc chains from serine or threonine. Digestion was done overnight at 37°C with 0.4 mU of enzyme per μg of WHsAg in 20 mM sodium phosphate (pH 7.2)–10 mM EDTA.

Protein sequencing. Proteins (50 pmol) were electroblotted onto polyvinylidene difluoride membranes (Millipore) after separation by SDS-PAGE and were amino-terminally sequenced by automated Edman degradation on an Applied Biosystems pulsed-liquid-phase sequencer, model 477A or 471A, under standard conditions. Phenylthiohydantoin derivatives of amino acids were identified by an on-line analyzer, model 120A or 140B (Applied Biosystems), with a repetitive yield of 92 to 95%.

Mass spectrometry. Purified WHs protein spheres (about $20\text{ }\mu\text{g}$) were concentrated by ultrafiltration on Microcon 10 filter units (Millipore), resuspended in 25 mM Tris-HCl (pH 8.2), washed twice, and resuspended in the same buffer with 5% β -mercaptoethanol. After a washing with water, the reduced proteins were dissolved in 30% (vol/vol) acetonitrile–0.1% (vol/vol) TFA to a concentration of about 5 pmol/ μl , based on S protein. Molecular masses were determined by matrix-assisted laser desorption/ionization–time-of-flight (MALDI-TOF) mass spectrometry on a Vision 2000 mass spectrometer (Finnegan MAT, Bremen, Germany). One microliter of protein solution was mixed with 1 μl of matrix solution (9 mg of 2,5-dihydroxybenzoic acid and 1 mg of 2-hydroxy-5-methoxybenzoic acid per ml, 0.1% [vol/vol] TFA, 30% [vol/vol] acetonitrile) and allowed

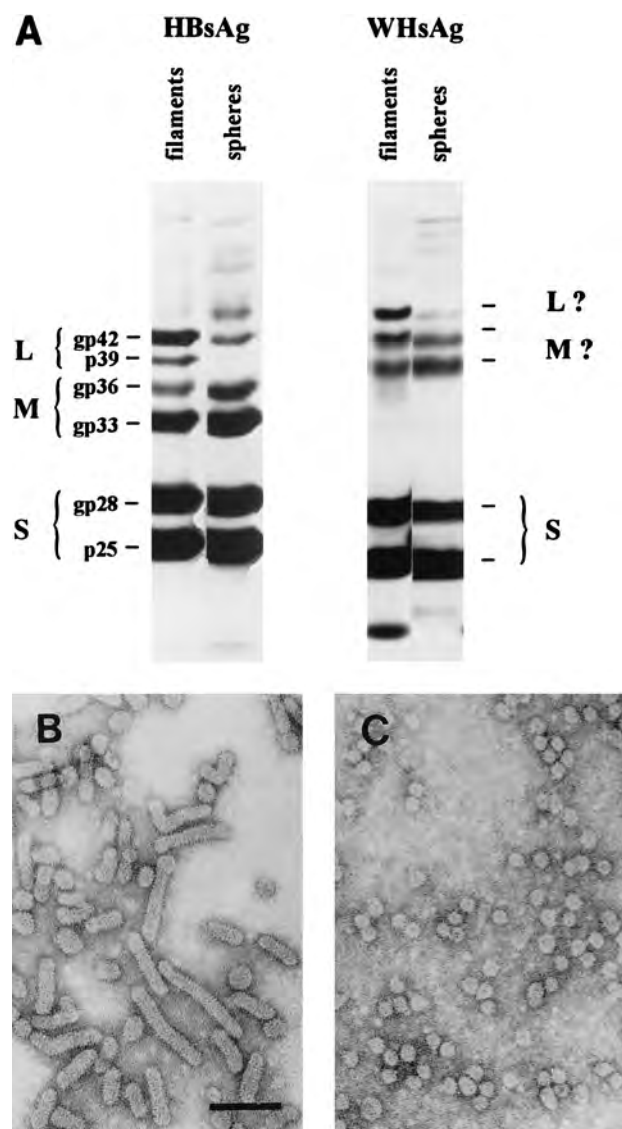


FIG. 2. Protein composition and particle morphology of WHsAg. (A) SDS gel electrophoresis and silver staining of HBsAg and WHsAg. Heavier fractions (filaments) and lighter fractions (spheres) from the first sucrose gradient were purified further as described in Materials and Methods and analyzed. (B and C) The morphology of the WHsAg preparations analyzed for panel A was analyzed by negative staining with 1% uranyl acetate and electron microscopy. The bar corresponds to 100 nm.

to air dry. Ions were generated by irradiation with a pulsed nitrogen laser (emission wavelength, 337 nm; laser power density, about 10^6 W cm $^{-2}$), and positive ions were accelerated and detected in the reflector mode. Spectra were calibrated with subtilisin (Sigma) as an external standard.

RESULTS

Protein composition of WHsAg spheres and filaments. WHs protein spheres and filaments were purified from WHV carrier serum by a combination of zonal sedimentation in sucrose, zonal flotation in a CsCl gradient, and isopycnic centrifugation in sucrose. The WHsAg was identified in the fractions of the first sucrose gradient by SDS-PAGE and silver staining. Silver is known to stain HBs proteins better than most serum proteins (38). Fractions with 35 to 45% sucrose contained, besides serum components, the five bands identified in the final prepa-

P24	SWHs	found	M S P S S L L G L L
		predicted	M S P S S L L G L L
P27	SWHs N-glycan	found	M S P S S L X G L
		predicted	M S P S S L L G L
P41	MWHs	found	X K X Q T F X L Q G
		predicted	M K N Q T F H L Q G glycan
P31	MWHs deglycosylated	found	M K X Q T F H L Q G
		predicted	M K D Q T F H L Q G
P49	LWHs	found	blocked
		predicted	G N N I K V T F N P myr

FIG. 3. N-terminal sequences of WHs proteins as obtained by Edman degradation in comparison to the predicted sequence. X, amino acids which could not be clearly identified after Edman degradation.

ration of WHs protein filaments (Fig. 2A). The heavier fractions had a more intensely stained 49-kDa band than the lighter fractions, where this band was barely visible. All WHsAg-containing fractions contained a 24-kDa band and 27-kDa band as major components and somewhat heterogeneous 41- and 45-kDa bands as minor components in constant proportions.

An analogous pattern was found with HBsAg from human serum or plasma (Fig. 2A). More prominent LHBs protein bands of 42 and 39 kDa were found together with HBs protein filaments or HBV virions, which sediment faster than the HBs protein spheres (36). SHBs protein bands of 25 and 28 kDa were major components of all HBsAg-containing fractions. MHBs protein bands of 33 and 36 kDa codistributed with SHBs proteins with no preference for filaments, virions, or spheres (17). We postulated that the 49-kDa band would correspond to LWHs protein and that it would copurify with WHs protein filaments. Therefore, the heavier and the lighter fractions of WHsAg in the first gradient were pooled separately and purified further. As shown in Fig. 2 the heavier subpopulations consisted mainly of 49-kDa-protein-rich WHs protein filaments, whereas the lighter fractions consisted of 49-kDa-protein-depleted WHs protein spheres. These data suggested that the 49-kDa protein was indeed an LWHs protein. Furthermore, the codistribution of the 45- and 41-kDa bands with the putative SWHs proteins suggested that they were MWHs proteins.

Identification of the SWHs proteins. The putative SWHs proteins of 24 and 27 kDa were eluted from Coomassie blue-stained blot membranes after SDS-PAGE and analyzed by Edman degradation. Both proteins yielded unequivocally the sequence expected for SWHs proteins (Fig. 3). Purified WHs protein spheres showed in MALDI-TOF mass spectrometry a sharp peak at 25,587.8 Da, which differs only by 16.6 Da from the theoretical molecular mass of nonglycosylated SWHs pro-

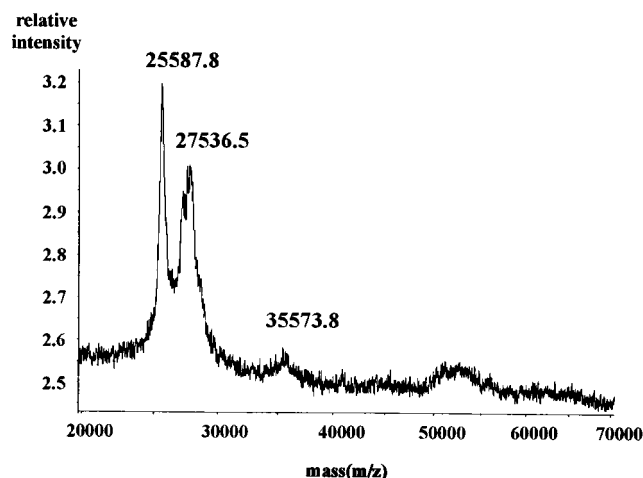


FIG. 4. MALDI-TOF mass spectrum of reduced purified WHsAg spheres. The major peaks correspond to SWHs protein with and without N-glycan, and the minor peak at m/z 35,573.8 corresponds to MWHs protein with one N-glycan. The peaks at m/z 50,000 to 55,000 are probably caused by dimers of the SWHs protein.

teins (Fig. 4) and is well within the expected standard deviation. Furthermore, a very broad peak with a maximum at 27,536.5 Da was found. This molecular mass is consistent with a modification of SWHs protein with one diantennary complex N-glycan which could contribute ca. 1,915 Da if it contains one NeuNAc residue.

Identification of MWHs proteins. Bands of the 41- and 45-kDa proteins were cut from unstained blot membranes and subjected to sequencing. The sequence of the 41-kDa band was identical with that expected for MWHs proteins (Fig. 3). The band of 45 kDa was found to contain S dimers which comigrated with the putative double N-glycosylated MWHs protein of 45 kDa. The amount of this MWHs protein was insufficient for N-terminal sequencing. As shown below, Western blotting with a pre-S2-reactive antiserum stained not only the 41-kDa but also the 45-kDa band. Asn3 of the 41-kDa band could not be identified, which suggests N glycosylation. In MALDI-TOF mass spectrometry, WHsAg particles showed, besides the peaks of the two SWHs proteins, a minor heterogeneous peak of about 35.5 kDa which probably corresponds to the 41-kDa band of MWHs protein (Fig. 4). As shown below, the 41-kDa band and a major part of the 45-kDa bands can be converted to one 31-kDa band by extensive deglycosylation. This band was also sequenced, and the sequence of MWHs protein was confirmed (Fig. 3). The difference between the true molecular mass and the excessive apparent size of glycoproteins results from overproportionally stronger retardation of electrophoretic migration by the glycan.

Identification of LWHs proteins. The 49-kDa protein band from purified WHs protein filaments was subjected to Edman degradation, but no clear signals were obtained, although the amount of protein was similar to that of the MWHs protein band from which the sequence could be determined unequivocally. In one experiment, the SWHs protein sequence was detectable at very small amounts. This is probably due to glycosylated SWHs protein dimers which may comigrate at this position. The observation of a blocked amino end is consistent with the expectation for LWHs protein, because the N-terminal methionine of LHBs protein is substituted at Gly2 of the LHBs protein by myristic acid. The myristoylation signal of the large surface protein is conserved even in avian hepadnavi-

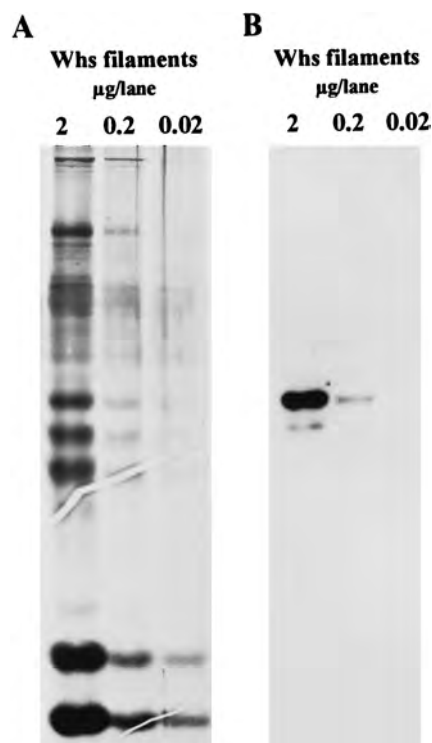


FIG. 5. Detection of LWHs protein by SDS-PAGE of purified WHsAg filaments. Three 10-fold dilutions were run in duplicate; one panel was silver stained (A), and the other one was blotted and immunostained with an anti-pre-S1 antiserum and enhanced chemiluminescence (B).

rus. The overall amino acid composition of the 49-kDa band was compatible with that of LWHs protein (data not shown). In order to identify the LWHs protein unequivocally, a rabbit antiserum against the pre-S1 (residues 126 to 146) peptide was generated. The antiserum reacted strongly with the peptide but not with the SWHs protein, the 41-kDa band of MWHs protein, or any of the HBs proteins (Fig. 5). However, the anti-pre-S1 immunoglobulin G bound very well to the 49-kDa band and less intensively to a 45-kDa band which could not be resolved from the 45-kDa bands of MWHs and SWHs protein dimers in the silver-stained SDS gels.

N Glycosylation of the WHs proteins. Purified WHs protein spheres or filaments were subjected to digestion by PNGase F. In order to distinguish between the pre-S2-linked glycan of the M proteins and N-glycan of the S domain, digestion was done with native or dithiothreitol-treated surface antigens. Previous experiments had shown that PNGase F removed the N-linked glycan from the pre-S2 domain of MHBs protein but not the N-glycan from the S domain, unless the disulfide bonds of the S domain were cleaved by reduction (24a). The digests were subjected to denaturing SDS-PAGE, and the proteins were stained with silver (Fig. 6). The 33-kDa band of undigested MHBs protein was, as would be expected, converted to a 31-kDa band, both in native and reduced HBsAg. In contrast, the 28-kDa band of glycosylated SHBs protein was converted only to the nonglycosylated 25-kDa protein in the reduced HBsAg. Likewise, the 27-kDa band of glycosylated WHs protein was converted to a nonglycosylated 24-kDa protein after reduction of WHs protein spheres or filaments. The stronger 41-kDa band and the weaker 45-kDa band of MWHs protein spheres were shifted to a stronger 37-kDa band and a weaker 41-kDa band after digestion of native particles, whereas only the 37-

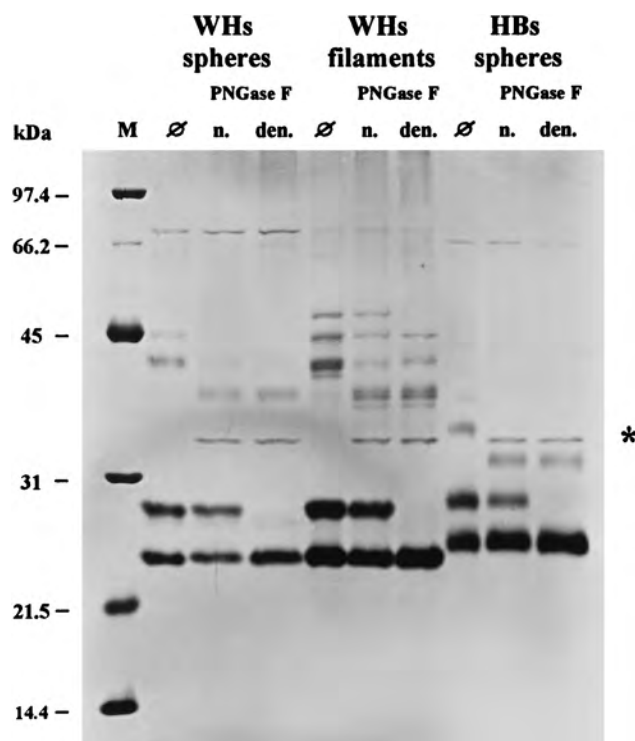


FIG. 6. Effect of PNGase F on the WHsAg spheres and filaments and HBsAg spheres. Purified antigens either left untreated (Ø), digested with PNGase F in the native state (n.), or digested with PNGase F after denaturation (den.) were subjected to SDS gel electrophoresis and silver staining as described in Materials and Methods. The sharp band at 34 kDa is PNGase F (*).

kDa band was visible in the digest of reduced WHs protein spheres. With WHs protein filaments, the same shifts of SWHs and MWHs proteins were observed by digestion with PNGase F. The 41-kDa band of the PNGase-digested filaments in Fig. 6 is probably due to dimeric nonglycosylated SWHs protein. In a Western blot this 41-kDa band was not stained with an anti-pre-S2-reactive antiserum (data not shown). The sharp 49-kDa band of glycosylated LWHs protein was completely converted to a sharp deglycosylated 45-kDa band, but only after reduction. This suggests that the N-glycan of LWHs protein is linked to the S domain, as has been described for LHBs protein (17). The appearance of a sharp 45-kDa band in the digest of unreduced WHs protein filaments suggests that, in agreement with the data shown in Fig. 5, WHs protein filaments contain glycosylated and nonglycosylated LWHs protein. In untreated WHs protein filaments, this LWHs protein band is obscured by double N-glycosylated MWHs protein and glycosylated SWHs protein dimers.

O glycosylation of MWHs protein. In order to account for the apparent oversize of the MWHs proteins, we tested the hypothesis that they contain O-linked glycans. Direct digestion of WHsAg with *O*-glycosidase did not cause any shift in the protein migration pattern. Since it is known that this enzyme can cleave only O-glycans without terminal sialic acid, WHs protein filaments were first digested with sialidase. This digestion generated a complex mixture of faster-migrating LWHs and MWHs proteins (Fig. 7A). Most prominent was the shift of the strong 41-kDa band of MWHs protein to a strong 36-kDa band. The 45-kDa band of MWHs protein was shifted to a 39-kDa band. This interpretation was confirmed by immunostaining a Western blot of the SDS gels with an antiserum

against native WHsAg. Similar to the antisera against HBsAg (17), this antiserum reacted in Western blotting preferably with the pre-S2 domain of MWHs protein (Fig. 7B). Subsequent digestion with *O*-glycosidase shifted the MWHs protein bands further down by 2 kDa to 34 and 37 kDa. Additional digestion with PNGase F generated finally one band of 31 kDa, which is the expected size of nonglycosylated MWHs protein. As mentioned above, this band had the N-terminal sequence predicted for MWHs protein. The antiserum seemed to recognize the pre-S2 sequence of MWHs protein preferentially in its N-glycosylated state, because the immunostaining of the 31-kDa band in Fig. 6B was overproportionally weakened after PNGase F digestion, compared to the silver staining in Fig. 7A. Digestion with *O*-glycosidase, even after sialidase treatment, did not have a detectable effect on the apparent size of MHBs protein in SDS-PAGE (data not shown).

Structure and number of O-glycan groups. *O*-Glycosidase is known to cleave only glycans which contain the mucin-type core 1 structure Gal β 1-3GalNAc α 1-Ser/Thr. Since prior sialidase treatment was necessary for removal of the O-glycan, it is obvious that these glycans carry NeuNAc. In order to elucidate the nature of this binding, WHsAg particles were digested in parallel with three sialidases specific for either α 2-6, α 2-8, or α 2-3 and α 2-8 bonds. As shown in Fig. 7C, only sialidase from *A. ureafaciens*, which cleaves NeuNAc α 2-6 bonds, removed sialic acid from the two MWHs protein species and allowed further cleavage by *O*-glycosidase.

α -Fucosidase and β -*N*-acetylhexosaminidase did not change the apparent size of the MWHs proteins (data not shown). Thus, the well-known structure (4) shown in Fig. 7D may be postulated from the data. This O-glycan is probably present in all MWHs protein subunits of serum-derived WHsAg.

The MALDI-TOF-mass spectrum shown in Fig. 4 indicates an average molecular mass of about 35,573 Da for the MWHs protein with one N-linked glycan. The polypeptide mass of MWHs protein is 32,125.6 Da; the N-glycan with one terminal NeuNAc residue would account for 1,915 Da, leaving 1,530 Da for O-linked carbohydrates. One O-glycan residue with the putative structure shown in Fig. 7D has a molecular mass of 656.7 Da. This suggests that the majority of the MWHs protein molecules contains two or three O-glycan groups. The pre-S2 domain of WHs proteins contains 10 threonine and 5 serine residues, 10 of which are sites for potential O glycosylation (16).

DISCUSSION

The protein composition of WHsAg has been previously studied by several groups. Feitelson et al. suggested that besides the major protein bands of SWHs protein, larger proteins with the pre-S sequences were present, but a clear identification was not possible (10). Pohl et al. reported that the protein composition of WHsAg from serum was quite similar to those of the pre-S-containing proteins gp33, gp36, p39, and gp42 of HBsAg, but actually the WHs protein band in their silver-stained SDS gel corresponding to gp33 of HBsAg was very weak, and the other proteins were not clearly resolved and possibly larger (30). Schaeffer et al. identified by immunoblotting with an antiserum against WHV pre-S peptide (residues 85 to 173) four proteins in partially purified WHsAg of 33, 36, 45, and 47 kDa. The 47-kDa protein formed the major band, and the 33-kDa protein formed the minor band. It appears that the 47-kDa protein corresponds to the 49-kDa LWHs band that we identified, whereas the identities of the other bands are not clear (33). Using a WHV anti-pre-S2 antiserum, Shamoon et al. detected 36-, 45-, and 47-kDa bands, and they suggested

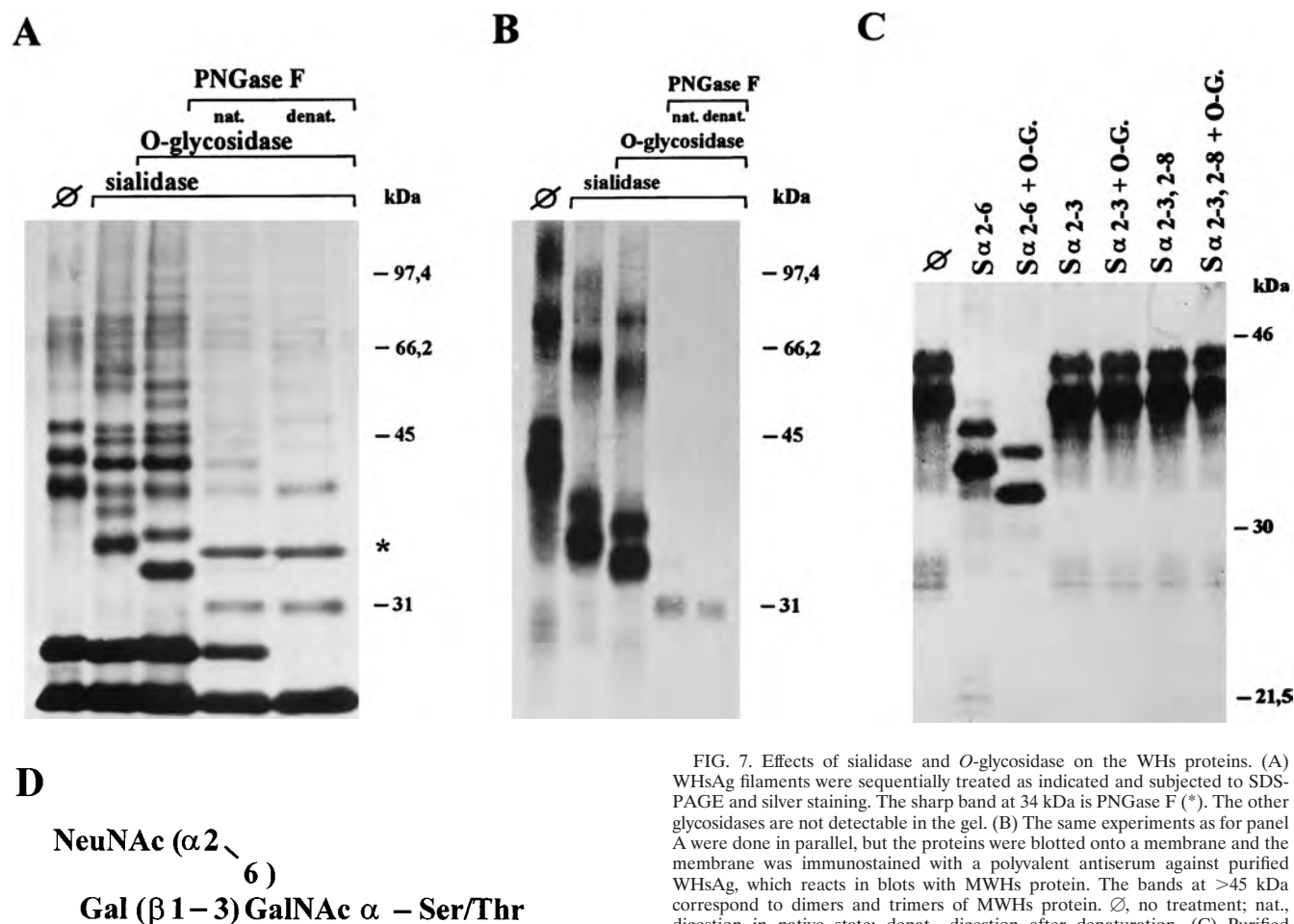


FIG. 7. Effects of sialidase and O-glycosidase on the WHs proteins. (A) WHsAg filaments were sequentially treated as indicated and subjected to SDS-PAGE and silver staining. The sharp band at 34 kDa is PNGase F (*). The other glycosidases are not detectable in the gel. (B) The same experiments as for panel A were done in parallel, but the proteins were blotted onto a membrane and the membrane was immunostained with a polyvalent antiserum against purified WHsAg, which reacts in blots with MWHs protein. The bands at >45 kDa correspond to dimers and trimers of MWHs protein. $\emptyset$, no treatment; nat., digestion in native state; denat., digestion after denaturation. (C) Purified WHsAg filaments were treated with the indicated glycosidases and subjected to SDS-PAGE, blotting, and immunostaining with the anti-WHs protein serum. (D) Proposed structure of the MWHs protein-linked O-glycan as derived from the glycosidase cleavage pattern.

that the 45- and 47-kDa bands were LWHs proteins, whereas gp36 was believed to be a completely (i.e., two times) N-glycosylated MWHs protein (35). In view of our data, it appears that the predominant anti-pre-S2-reactive band at 45 kDa seen by Shamoon et al. contained the mono-N-glycosylated MWHs protein corresponding to our 41-kDa band. In contrast to these previous studies, we used highly purified preparations of WHsAg, separated WHs protein filaments and spheres, and applied N-terminal sequencing for unequivocal identification of the WHs proteins. We found that a direct comparison of WHsAg and HBsAg in PAGE is not appropriate in the case of MWHs protein and, furthermore, that the unexpected large size of MWHs protein obscures the pattern of LWHs protein. We analyzed for the first time the glycosylation pattern of the WHs proteins and found, surprisingly, O glycosylation of MWHs protein.

WHV and HBV share genome organization, replication strategy, organ tropism, and major aspects of transmission and pathogenesis. Nevertheless, some important differences exist. (i) WHV is more oncogenic due to integration of its DNA into the N-myc2 locus of woodchucks (11, 43). A similarly frequent integration in a relevant locus does not exist in human genomes (8). (ii) The human enhancer I of HBV is obviously not present in the WHV genome (9). (iii) MWHs protein does not bind woodchuck serum albumin (30, 34), whereas MHBs protein binds in a species-specific manner modified natural human serum albumin (21). (iv) The LWHs protein is 42 (or 31)

amino acids longer than the LHBs protein, due to a larger pre-S1 domain. This point is consistent with the identification of 45- and 49-kDa proteins as nonglycosylated or mono-N-glycosylated LWHs proteins instead of 39- and 42-kDa proteins as LHBs proteins.

With the exception of the difference in size, all other properties of LWHs protein determined in this study are analogous to those of LHBs protein. LWHs protein is prevalent in the filamentous form of WHsAg protein particles, as is LHBs protein in HBsAg filaments. One may postulate that LWHs protein is also—in analogy to HBV (17)—prevalent in WHV virions and essential for virion formation (5, 7). The dual topology of LHBs protein with half of the pre-S domain inside the envelope and the other half outside the envelope (6) is probably also present in WHV, because even in the very distantly related duck HBV is this topology preserved (15, 40), although formal proof is not provided in this report. The observed N-terminal blockade of LWHs protein is consistent with the modification of the conserved Gly2 by myristic acid, which is known to be present in HBV (29) and duck HBV (25). The cotranslational retention of the pre-S domain of LHBs protein in the cytosol (6, 7, 23, 28, 31) probably also occurs with LWHs protein, because no N glycosylation of the pre-S domain of LWHs protein was found. LWHs, MWHs, and SWHs proteins

coassemble probably soon after translation to WHs protein filaments. This is suggested by the observation of Abe et al. (1) that WHV-infected liver contains WHs protein filaments within the cisternae of the endoplasmic reticulum. The presence of O glycosylation in the pre-S2 domain of MWHs protein suggests that the liver of woodchucks would also be able to glycosylate the pre-S domains of LWHs protein, if these domains were accessible to the respective glycosyltransferases in the Golgi apparatus. Theoretically, the pre-S1 sequences of LWHs protein may sterically hinder the accessibility of the O glycosylation site in pre-S2. An alternative explanation would be that in the Golgi apparatus the pre-S domains of LWHs protein are still in an internal position and translocation to the surface of the particles does not occur before the *trans*-Golgi network or even extracellularly. The slightly acidic environment within the *trans*-Golgi network is a likely place for the partial posttranslational translocation of pre-S domains to the viral surface, because acidification to pH 5 was shown to increase the number of surface-accessible pre-S domains of HBV in vitro (6).

The close similarity between the WHs and HBs proteins is also documented by the typical pattern of N glycosylation in the S domains. Approximately one-half of the SWHs and SHBs protein molecules contain N-glycan substitutions. The proportion is smaller in MWHs and MHBs proteins but in the LWHs and LHBs proteins the majority of the S domains is glycosylated. Given this quantitative similarity in N glycosylation, it is even more striking that the O glycosylation seems to be so different in MWHs and MHBs proteins. Three factors may contribute to this difference. (i) The woodchuck may have different glycosyltransferases, since species-specific variations in glycoconjugate structures are well documented (42). (ii) WHV may encode more efficient sites for O glycosylation. This explanation is possibly less relevant, because O glycosylation has in fact been found in MHBs protein-associated pre-S2 domains which were expressed in yeast cells (22) or mammalian cells. The extent of O glycosylation in COS cells is such that it is probably similar to that of MWHs protein (44). While one could argue that O glycosylation of MHBs protein in artificial expression systems is an artifact, the WHsAg analyzed in this study is derived from naturally infected woodchuck liver and corresponds to the secreted circulating WHsAg particles. (iii) The O glycosylation sites in human pre-S2 are partially inaccessible. This explanation may be relevant, because the pre-S2 domain of MHBs protein binds natural human serum albumin (21). This binding may possibly occur intracellularly during transport from the endoplasmic reticulum to the Golgi apparatus, thus preventing full accessibility of O glycosylation sites, whereas the cotranslational N glycosylation occurs in the endoplasmic reticulum and is not impaired. In fact, a covalently bound, disulfide-linked fraction of serum albumin was found in HBsAg particles (21). Such a linkage can only be formed by molecules with free SH groups, and these occur in the case of HBsAg only shortly after biosynthesis within the cytosol in the liver cells.

The role of the facultative N-glycan in the S proteins or S domains of mammalian hepadnaviruses is not understood, because this glycosylation can be deleted without effect on assembly or secretion (44). The large proportion of N glycosylation in LHBs and LWHs proteins suggests that glycosylation of LHBs and LWHs proteins in the S domain may be more essential than that of SHBs and SWHs proteins. The N-glycan of pre-S2 in MHBs protein is involved in calnexin-associated quality control of folding and in secretion (44).

The significance of the newly discovered O glycosylation of natural WHs protein particles is not yet clear. Many serum

proteins are O glycosylated with terminal sialic acid. Furthermore, the O-glycoside may contribute to binding of WHV to target cells. It remains to be determined whether O-glycoside-deficient WHs protein particles have a shorter half-life in the circulation or are less infectious. The conservation of M surface proteins in mammalian hepadnaviruses suggests that MWHs protein or MHBs protein confers a selective advantage for survival of these viruses in their host population, but the manner by which this advantage is achieved may differ in the two host species studied in this investigation.

ACKNOWLEDGMENTS

We thank B. Boschek for electron microscopy, S. Broehl for technical assistance, U. Friedrich for peptide synthesis, M. Seifer and D. Standing for woodchuck sera, and G. Caspari for HBV plasma.

The work was supported by SFB 535 projects A2 and Z1.

REFERENCES

1. Abe, K., T. Kurata, and T. Shikata. 1988. Localization of woodchuck hepatitis virus in the liver. *Hepatology* 8:88–92.
2. Block, T. M., X. Lu, A. S. Mehta, B. S. Blumberg, B. Tennant, M. Ebling, B. Korba, D. M. Lansky, G. S. Jacob, and R. A. Dwek. 1998. Treatment of chronic hepadnavirus infection in a woodchuck animal model with an inhibitor of protein folding and trafficking. *Nat. Med.* 4:610–614.
3. Block, T. M., X. Lu, F. M. Platt, G. R. Foster, W. H. Gerlich, B. S. Blumberg, and R. A. Dwek. 1994. Secretion of human hepatitis B virus is inhibited by the imino sugar N-butyldeoxynojirimycin. *Proc. Natl. Acad. Sci. USA* 91:2235–2239.
4. Brockhausen, I. 1995. Biosynthesis of O-glycans of the N-acetylgalactosamine- α -Ser/Thr linkage type, p. 201–259. In J. Montreuil, J. F. G. Vliegthart, and H. Schachter (ed.), *Glycoproteins*. Elsevier, Amsterdam, The Netherlands.
5. Bruss, V., and D. Ganem. 1991. The role of envelope proteins in hepatitis B virus assembly. *Proc. Natl. Acad. Sci. USA* 88:1059–1063.
6. 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.
7. 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.
8. Buendia, M. A., and P. Paterlini. 1993. Liver cancer, p. 137–164. In A. J. Zuckermann and H. C. Thomas (ed.), *Viral hepatitis—scientific basis and management*. Churchill Livingstone, Edinburgh, Scotland.
9. Di, Q., J. Summers, J. B. Burch, and W. S. Mason. 1997. Major differences between WHV and HBV in the regulation of transcription. *Virology* 229: 25–35.
10. Feitelson, M. A., P. L. Marion, and W. S. Robinson. 1983. The nature of polypeptides larger in size than the major surface antigen components of hepatitis B and like viruses in ground squirrels, woodchucks, and ducks. *Virology* 130:76–90.
11. Fouriel, G., C. Trepo, L. Bougueleret, B. Henglein, A. Ponzetto, P. Tiollais, and M. A. Buendia. 1990. Frequent activation of N-myc genes by hepadnavirus insertion in woodchuck liver tumours. *Nature* 347:294–298.
12. Galibert, F., T. N. Chen, and E. Mandart. 1981. Localization and nucleotide sequence of the gene coding for the woodchuck hepatitis virus surface antigen: comparison with the gene coding for the human hepatitis B virus surface antigen. *Proc. Natl. Acad. Sci. USA* 78:5315–5319.
13. Ganem, D. 1996. Hepadnaviridae and their replication, p. 2703–2807. In B. N. Fields, D. M. Knipe, and P. M. Howley (ed.), *Fields virology*, 3rd ed., vol. 2. Lippincott-Raven Publishers, Philadelphia, Pa.
14. Gerlich, W., and R. Thomssen. 1975. Standardized detection of hepatitis B surface antigen: determination of its serum concentration in weight units per volume. *Dev. Biol. Stand.* 30:78–87.
15. Guo, J.-T., and J. C. Pugh. 1997. Topology of the large envelope protein of duck hepatitis B virus suggests a mechanism for membrane translocation during particle morphogenesis. *J. Virol.* 71:1107–1114.
16. Hansen, J. E., O. Lund, J. Engelbrecht, H. Bohr, J. O. Nielsen, and J. E. Hansen. 1995. Prediction of O-glycosylation of mammalian proteins: specificity patterns of UDP-GalNAc:polypeptide N-acetylgalactosaminyltransferase. *Biochem. J.* 308:801–813.
17. Heermann, K. H., 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.
18. Kann, M., and W. H. Gerlich. 1998. Hepatitis B, p. 745–773. In L. Collier, A. Balows, and M. Sussmann (ed.), *Microbiological and microbial infections*, 9th ed., vol. 1. Arnold, Hodder Headline Group, London, England.
19. Koenig, W., and R. Geiger. 1970. A new method for synthesis of peptides: activation of the carboxyl group with dicyclohexylcarbodiimide using 1-hy-

- droxybenzotriazoles as additives. *Chem. Ber.* **103**:788–798.
20. **Korba, B. A., H. Xie, K. N. Wright, W. E. Hornbuckle, J. L. Gerin, B. C. Tennant, and K. Y. Hostetler.** 1996. Liver-targeted antiviral nucleosides: enhanced antiviral activity of phosphatidyl-dideoxyguanosine versus dideoxyguanosine in woodchuck hepatitis virus infection in vivo. *Hepatology* **23**: 958–963.
 21. **Krone, B., A. Lenz, K. H. Heermann, M. Seifer, X. Y. Lu, and W. H. Gerlich.** 1990. Interaction between hepatitis B surface proteins and monomeric human serum albumin. *Hepatology* **11**:1050–1056.
 22. **Langley, K. E., K. M. Egan, J. M. Barendt, C. G. Parker, and G. A. Bitter.** 1988. Characterization of purified hepatitis B surface antigen containing pre-S(2) epitopes expressed in *Saccharomyces cerevisiae*. *Gene* **67**:229–245.
 23. **Loeffler, M. H., M. Werr, and R. Prange.** 1997. Sequence-specific repression of cotranslational translocation of the hepatitis B virus envelope proteins coincides with binding of heat shock protein Hsc70. *Virology* **235**:144–152.
 24. **Lu, X., A. Mehta, M. Dadmarz, R. Dwek, B. S. Blumberg, and T. M. Block.** 1997. Aberrant trafficking of hepatitis B virus glycoproteins in cells in which N-glycan processing is inhibited. *Proc. Natl. Acad. Sci. USA* **94**:2380–2385.
 - 24a. **Lu, X. Y., and W. H. Gerlich.** Unpublished data.
 25. **Macrae, D. R., 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**:359–363.
 26. **Mehta, A., X. Lu, T. M. Block, B. S. Blumberg, and R. A. Dwek.** 1997. Hepatitis B virus (HBV) envelope glycoproteins vary drastically in their sensitivity to glycan processing: evidence that alteration of a single N-linked glycosylation site can regulate HBV secretion. *Proc. Natl. Acad. Sci. USA* **94**:1822–1827.
 27. **Murphy, F. A., C. M. Fauquet, D. H. L. Bishop, S. A. Ghabrial, A. W. Jarvis, G. P. Martelli, M. A. Mayo, and M. D. Summers (ed.).** 1995. Virus taxonomy. Sixth report of the International Committee on Taxonomy of Viruses, p. 179–184. Springer-Verlag, Vienna, Austria.
 28. **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.
 29. **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.
 30. **Pohl, C. J., P. J. Cote, R. H. Purcell, and J. L. Gerin.** 1986. Failure to detect polyalbumin-binding sites on the woodchuck hepatitis virus surface antigen: implications for the pathogenesis of hepatitis B virus in humans. *J. Virol.* **60**:943–949.
 31. **Prange, R., and R. E. Strebeck.** 1995. Novel transmembrane topology of the hepatitis B virus envelope proteins. *EMBO J.* **14**:247–256.
 32. **Roggendorf, M., and T. K. Tolle.** 1995. The woodchuck: an animal model for hepatitis B virus infection in man. *Intervirology* **38**:100–112.
 33. **Schaeffer, E., R. L. Snyder, and J. J. Sninsky.** 1986. Identification and localization of pre-S-encoded polypeptides from woodchuck and ground squirrel hepatitis viruses. *J. Virol.* **57**:173–182.
 34. **Shamoon, B., E. Mandart, A. Kay, and F. Galibert.** 1994. Woodchuck hepatitis virus surface antigen produced in vitro fails to bind polymerized woodchuck serum albumin. *J. Gen. Virol.* **75**:2081–2084.
 35. **Shamoon, B. M., A. Kay, S. Dupont de Dinechin, F. Galibert, and E. Mandart.** 1991. Production of polyclonal antibodies against the S and preS2 regions of woodchuck hepatitis virus: lack of detectable low glycosylated preS2 protein (GP33) in sera from infected animals. *J. Gen. Virol.* **72**:421–425.
 36. **Stibbe, W., and W. H. Gerlich.** 1983. Characterization of pre-s gene products in hepatitis B surface antigen. *Dev. Biol. Stand.* **54**:33–43.
 37. **Stibbe, W., and W. H. Gerlich.** 1983. Structural relationships between minor and major proteins of hepatitis B surface antigen. *J. Virol.* **46**:626–628.
 38. **Stibbe, W., and W. H. Gerlich.** 1982. Variable protein composition of hepatitis B surface antigen from different donors. *Virology* **123**:436–442.
 39. **Summers, J., J. M. Smolec, and R. Snyder.** 1978. A virus similar to human hepatitis B virus associated with hepatitis and hepatoma in woodchucks. *Proc. Natl. Acad. Sci. USA* **75**:4533–4537.
 40. **Swameye, I., and H. Schaller.** 1997. Dual topology of the large envelope protein of duck hepatitis B virus: determinants preventing pre-S translocation and glycosylation. *J. Virol.* **71**:9434–9441.
 41. **Tam, J. P.** 1988. Synthetic peptide vaccine design: synthesis and properties of a high-density multiple antigenic peptide system. *Proc. Natl. Acad. Sci. USA* **85**:5409–5413.
 42. **Varki, A.** 1993. Biological roles of oligosaccharides: all the theories are correct. *Glycobiology* **3**:97–130.
 43. **Wei, Y., G. Fourel, A. Ponzetto, M. Silvestro, P. Tiollais, and M.-A. Buendia.** 1992. Hepadnavirus integration: mechanisms of activation of the N-myc2 retrotransposon in woodchuck liver tumors. *J. Virol.* **66**:5265–5276.
 44. **Werr, M., and R. Prange.** 1998. Role for calnexin and N-linked glycosylation in the assembly and secretion of hepatitis B virus middle envelope protein particles. *J. Virol.* **72**:778–782.