JOURNAL OF VIROLOGY, July 1991, p. 3813–3820
0022-538X/91/073813-08$02.00/0

Vol. 65, No. 7

Mutational Analysis of Hepatitis B Surface Antigen Particle Assembly and Secretion

VOLKER BRUSS[1]† AND DON GANEM[1,2]*

Departments of Microbiology and Immunology[1] and Medicine,[2] University of California Medical Center, San Francisco, California 94143-0502

Received 7 February 1991/Accepted 15 April 1991

Cells infected with hepatitis B virus produce both virions and 20-nm subviral (surface antigen or HBsAg) particles; the latter are composed of viral envelope proteins and host-derived lipid. Although hepatitis B virus encodes three envelope proteins (L, M, and S), all of the information required to produce an HBsAg particle resides within the S protein. This polypeptide spans the bilayer at least twice and contains three hydrophobic regions, two of which are known to harbor topogenic signal sequences that direct this transmembrane orientation. We have examined the effects of mutations in these and other regions of the S protein on particle assembly and export. Lesions in the N terminal signal sequence (signal I) can still insert into the endoplasmic reticulum bilayer but do not participate in any of the subsequent steps in assembly. Deletion of the major internal signal (signal II) completely destabilizes the chain. Deletion of the C-terminal hydrophobic domain results in a stable, glycosylated, but nonsecreted chain. However, when coexpressed with wild-type S protein this mutant polypeptide can be incorporated into particles and secreted, indicating that the chain is still competent for some of the distal steps in particle assembly. The correct transmembrane disposition of the N terminus of the molecule is important for particle formation: addition of a heterologous (globin) domain to this region impairs secretion, but the defect can be corrected by provision of an N-terminal signal sequence that restores the proper topology of this region. The resulting chimeric chain is assembled into subviral particles that are secreted with normal efficiency.

Hepatitis B virus (HBV) is an enveloped DNA virus that produces acute and chronic hepatitis and is strongly associated with the development of hepatocellular carcinoma (for a review, see reference 8). One of the most remarkable features of HBV infection is that infected cells produce multiple types of virus-related particles, including infectious virions and noninfectious subviral particles composed primarily of the envelope proteins (HBV surface antigens [HBsAg]) and host-derived lipids (9, 12, 24, 25). Subviral particles are produced in large quantities and are efficiently exported from the cell, predominantly as 20-nm spherical or disklike particles. Interest in these particles stems from the fact that they are strongly immunogenic and can induce a protective immune response; as a result, they have been used to develop an effective vaccine against hepatitis B (27).

The strong immunogenicity of these particles has raised the possibility that they could serve as a basis for novel vaccines against other pathogens, by construction of hybrid surface antigens containing additional (foreign) epitopes (28). While several examples of such chimeric particles have already been reported (3–5, 19), many other attempts to generate secreted chimeric particles have failed, in part because of our lack of understanding of the basic requirements for HBsAg particle assembly. However, recent biochemical experiments have begun to define the nature of the HBsAg assembly pathway (6, 7, 20, 21, 26), making possible a more systematic examination of structure-function relationships within HBsAg.

HBV encodes three envelope glycoproteins (L, M, and S) by alternate translational initiations at each of three in-frame AUG codons within a single open reading frame (12). Initiation at the innermost AUG generates the S protein; the M and L proteins result from initiation at AUGs located 55 and 174 codons 5′ to the S initiator, respectively. Although lesser quantities of M and L proteins can be found on circulating HBsAg particles, the S protein is their dominant component (12). It contains all of the information required to form 20-nm HBsAg particles: expression of the S-coding region in cultured cells in the absence of any other viral gene products results in efficient particle assembly and secretion (17, 18, 23). Pulse-chase experiments in such cells indicate that particle assembly occurs through a series of definable intermediates (26). The first step is insertion of the S protein into the membrane of the endoplasmic reticulum (ER). Following this, S chains are thought to aggregate in the plane of the bilayer and then extrude or bud into the ER lumen (10, 20), taking with them a complement of membrane lipids. These lumenal particles are then secreted from the cell via the constitutive pathway of vesicular transport (13).

The nucleotide sequence of the S gene predicts the existence of three hydrophobic domains, located at residues 4 to 28, 80 to 100, and 164 to 221, respectively. Separating these domains are two hydrophilic stretches (see Fig. 1, top row). Current views hold that in the transmembrane form of the S protein both the N and C termini of the chain are in the ER lumen (7, 24). Thus, the chain must cross the bilayer at least twice, and in vitro evidence indicates that transmembrane passes are mediated by the first and second hydrophobic domains (7). The N-terminal hydrophobic domain encodes an uncleaved signal sequence (termed signal I) that initiates chain translocation across the membrane, thereby delivering the N terminus to the ER lumen. The second hydrophobic region (termed signal II) is a complex internal signal sequence that inserts into the bilayer in a backwards orientation (N-terminal end facing the cytosol, C-terminal end in the

* Corresponding author.
† Present address: Department of Medical Microbiology, University of Göttingen, D-3400 Göttingen, Germany.

ER), translocating C-terminal regions into the lumen (7). As a result, most of the region between signals I and II is believed to reside in the cytoplasm. The disposition of the C-terminal hydrophobic domain has not been experimentally addressed: it might be predominently lumenal (7), or it might make two additional transmembrane loops (7, 24).

To assess the contribution of each of these regions of the S polypeptide to particle formation, we have constructed a series of deletion and insertion mutations in the S gene and examined their impact on particle assembly and secretion. Our results document in vivo the essential roles of signals I and II and provide additional information on other regions of the molecule likely to be important in this process.

MATERIALS AND METHODS

Plasmid constructions. Plasmid pSV24H has been previously described (16); it carries the 1.9-kb *Pst*I-*Bgl*II fragment of the HBV genome subtype adw (29) bearing the S gene and poly(A) signal inserted downstream of a simian virus 40 early promoter and replication origin. Plasmid pSVΔsI was constructed from pSV24H by deletion of a 69-nucleotide *Avr*II-*Xba*I fragment (nucleotides 176 to 245), resulting in an in-frame deletion of amino acids 9 to 31 of the S protein. For construction of pSVΔC, pSV24H was cut with *Spe*I (nucleotide 677) and *Eco*RV (nucleotide 1040), blunted by mung bean nuclease digestion, and religated, resulting in a frameshift at codon 175 of the S gene; 3 amino acids (Ile-Leu-Pro) are added before encountering a nonsense codon in this frame. Plasmids of the pSVL::ClaI series were made by partial *Bal*I digestion of pSV24H and insertion of one, two, or five copies of the 12-bp *Cla*I linker (CCCATCGATGGG), resulting in the addition of one, two, or five copies of the peptide sequence Pro-Ser-Met-Gly at position 50 of the S protein.

The S-globin fusion pSVB8* (7) bears a chimpanzee alpha globin cDNA fused to a pre-S2 codon 11 triplets 5′ to the S ATG; the entire coding region is cloned 3′ to an SP6 promoter in pSP65. To generate pSVGS, the globin-S-coding region was excised from this plasmid by partial *Hin*dIII-*Hpa*I digestion and recloned into a copy of pSV24H from which the S sequences were deleted by *Hin*dIII-*Hpa*I digestion. pSVsGS was made by digesting pSVGS with *Hin*dIII, excising a fragment extending from the 5′ noncoding region to the middle of the globin-coding region; into this was cloned the corresponding *Hin*dIII fragment from pSP125E (22), containing a coding region for the β-lactamase signal sequence fused to globin. pSVsGΔsI was made by excising from this plasmid a *Ppu*MI-*Eco*RV fragment (spanning most of the S gene) and replacing it with the corresponding fragment from pSVΔsI (1).

Cell culture and transfection. COS 7 cells were cultured in DME H16 supplemented with 10% fetal calf serum as described previously (15, 16). For transfection, cells were seeded into six-well plates (Costar) at 2×10^5 cells per well and incubated for 14 h. Cells were washed with phosphate-buffered saline (PBS), and pairs of wells were incubated for transfection in 2.0 ml of 1× DME with 0.1% $NaHCO_3$, 50 mM Tris-Cl, 500 μg of DEAE-dextran (molecular weight, 2×10^6) per ml, and 4 μg of the appropriate plasmid DNA for 10 h at 37°C. The cells were then washed with PBS and incubated for another 38 h in 2.0 ml of medium.

Metabolic labeling. Cells were washed with PBS and incubated for 40 min with 0.6 ml of methionine-free DME medium with 2% dialyzed fetal calf serum. Then 15 μl (150 μCi) of ^{35}S-methionine (10 mCi/ml; Amersham) was added. After 1 h the medium was removed. From half of the cultures (pulse samples), lysates were prepared immediately; the remaining cultures (chase samples) were incubated for 24 h in 0.6 ml of medium containing excess unlabeled methionine (300 μg/ml). For the chase samples, the medium was removed and combined with the medium harvested after the pulse.

Lysates were prepared by washing the cells with cold PBS and incubating the cells with 0.4 ml of 50 mM Tris-Cl (pH 7.5)–150 mM NaCl–5 mM $MgCl_2$–0.2% Triton X-100 on ice for 15 min. The collected lysate and medium was clarified by spinning for 5 min in an Eppendorf Microfuge. The supernatants were transferred to new tubes for immune precipitation.

Immune precipitation. Per immune precipitation, 1 μl of a polyclonal goat anti-HBs (Dako) was added to 10 μl of a slurry suspension of protein A-Sepharose CL4B (Sigma) in 40 μl of PBS and incubated overnight at 4°C with rocking (1). Fifty microliters of this suspension was added to each cleared lysate or medium sample and incubated overnight at 4°C with rocking. The beads were washed three times with 750 μl of 50 mM Tris-Cl (pH 7.5)–150 mM NaCl–20 mM EDTA–0.2% Triton X-100–0.05% sodium dodecyl sulfate (SDS)–1% sodium deoxycholate and then washed once more with PBS. The beads were then boiled for 5 min in Laemmli sample buffer, and the entire suspension was loaded onto a 12 to 17% polyacrylamide gradient gel. After electrophoresis the gel was fixed, soaked for 30 min in Amplify (Amersham), and dried for autoradiography.

RESULTS

Experimental system. To study structure-function relationships within the S protein, we constructed the mutants depicted in Fig. 1, as described in detail in Materials and Methods. In all cases, the indicated coding regions were cloned between a simian virus 40 early promoter (and replication origin) and the normal HBV polyadenylation signals. The constructs were transfected in duplicate into T-antigen-producing COS 7 cells, and 48 h later, after allowing copy number amplification, cells were labeled with ^{35}S-methionine for 1 h. Cells and media from one set of dishes (pulse) were then harvested; the remaining plates (chase) were incubated for 24 h with fresh medium containing excess unlabeled methionine before harvesting of cells and media. Labeled S protein was then precipitated from cytoplasmic lysates and media with a polyclonal anti-HBs antiserum and analyzed by SDS-polyacrylamide gel electrophoresis (PAGE). Figure 2 (lanes i to l) shows the behavior of wild-type S chains in this assay. Following the pulse, all of the chains have been inserted into the ER bilayer (26) and ca. 30 to 40% of the chains are glycosylated; over 90% of the product is still intracellular (lanes i and j). After the chase (lanes k and l), ca. 50% of the labeled chains have been secreted, with Golgi processing of their N-linked carbohydrate chains causing a slight further retardation of the mobility of the glycosylated species in the medium (lane l). No degradation of the labeled chains was evident.

Deletions of the hydrophobic domains. To study the functions of the hydrophobic domains in assembly, we constructed in-frame deletion mutants that precisely excised each of these regions. Deletion of the C-terminal hydrophobic region was achieved by construction of a frameshift mutation at codon 175 of the S gene (mutant ΔC); this results in a 3-amino-acid extension before termination at a UAA codon. Transfection of COS 7 cells with this mutant resulted

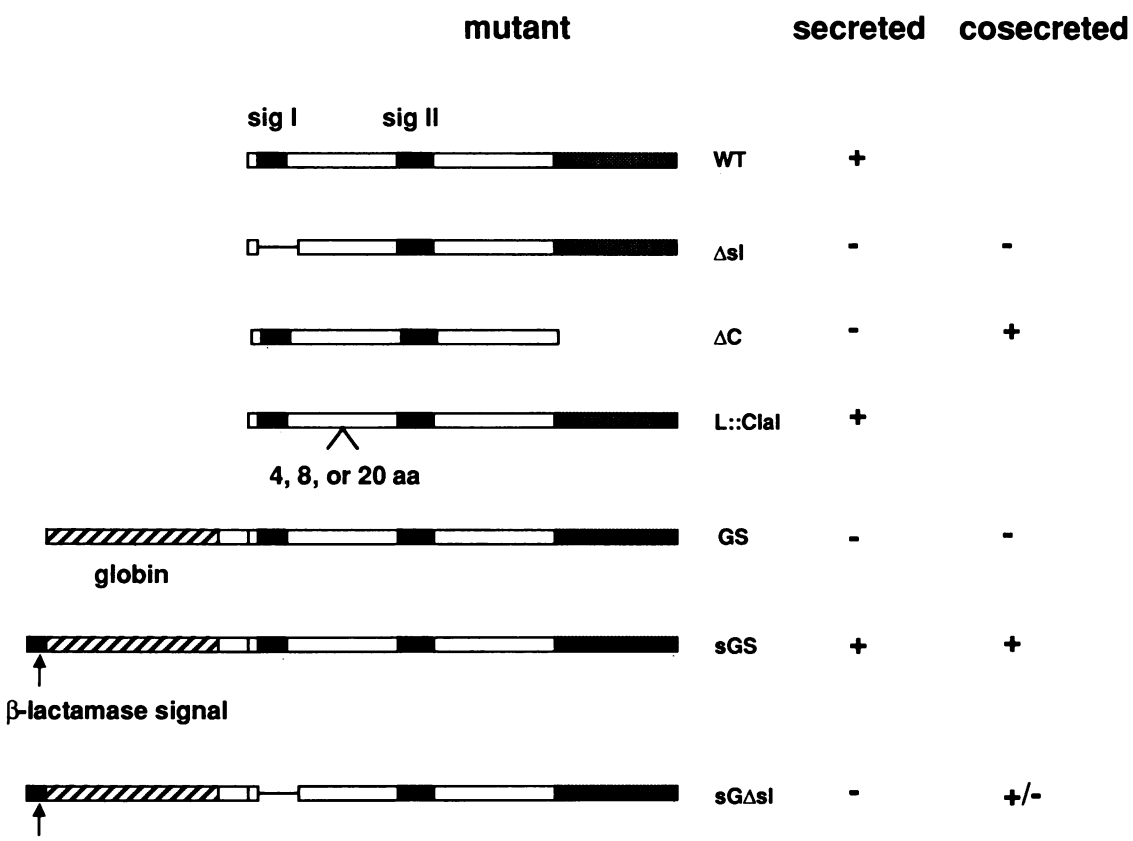

FIG. 1. Coding organization of the HBV S gene and mutant genes used in this study. At the left are schematic depictions of the S-gene derivatives referred to in the text. The top line shows the wild-type (WT) S gene product; the remaining lines show individual mutants. White boxes denote hydrophilic S domains; black boxes denote hydrophobic signal I (sig I, residues 4 to 28) and signal II (sig II, residues 80 to 100); stippled boxes denote the C-terminal hydrophobic domain of S protein (residues 164 to 221). *Cla*I linkers are denoted by /\; each linker inserts 4 amino acids. Hatched boxes indicate the first 100 amino acids of chimpanzee alpha globin. The β-lactamase signal sequence is indicated by an arrow and black box. Columns at the right summarize the secretory phenotype of the indicated mutant. "Secreted" refers to the ability of the mutant chain to be secreted by itself; "cosecreted" refers to the ability of the chain to be secreted following cotransfection of the mutant with wild-type S-protein expression vector.

in production of a labeled immunoreactive protein of ca. 18 kDa, in accord with the size predicted from the location of the mutation; a small amount of glycosylated protein is also present (Fig. 2, lane a). This protein is somewhat less stable than its wild-type parent, with ca. 60 to 70% of the chains undergoing degradation during the chase period (compare lanes a and c). But even allowing for this, no secretion of the mutant was detectable (lane d), even on prolonged exposure of the autoradiogram. Thus, the mutant appears competent for membrane insertion but must be blocked at some more distal step in the assembly pathway. The next step in the pathway is the interaction with other S chains. To determine whether the chain was still competent for such interactions, we cotransfected the ΔC mutant with a plasmid encoding the wild-type S protein and examined particle export as described above (Fig. 2, lanes e to h). Strikingly, export of mutant chains was rescued by coexpression of full-length S chains, and the efficiency of export of mutant chains was within twofold that of the wild-type. In addition, the stability of the mutant chains was also restored to nearly normal levels by the presence of wild-type S protein. This behavior is most readily explained by the formation of mixed aggregates of wild-type and mutant chains and suggests that the assembly defect of ΔC is distal to the aggregation step. (Direct evidence that such mixed aggregates form under these cotransfection conditions will be presented in Fig. 4B, below.) From this we infer that the C-terminal region may be required for either the budding reaction itself or the intracellular transport of budded particles through the vesicular transport pathway (see Discussion).

Next we deleted each of the known translocation signals residing in the remaining hydrophobic regions. As a first step we engineered a precise deletion of signal II using oligonucleotide mutagenesis (14) to delete codons 80 to 98. This mutant polypeptide was grossly unstable in COS 7 cells, being barely detectable even during short pulse-labeling with ^{35}S-methionine and disappearing entirely during the chase (data not shown). It could not be stabilized by coexpression of wild-type S chains (data not shown). This precluded definition of the precise role of signal II in particle assembly in vivo by this approach. The instability of the mutant is not surprising, since signal II is believed to be the dominant element that determines the transmembrane orientation of the bulk of the chain (7). Its deletion would be expected to globally disrupt the transmembrane structure of the protein, and degradation is often the fate of misfolded polypeptides

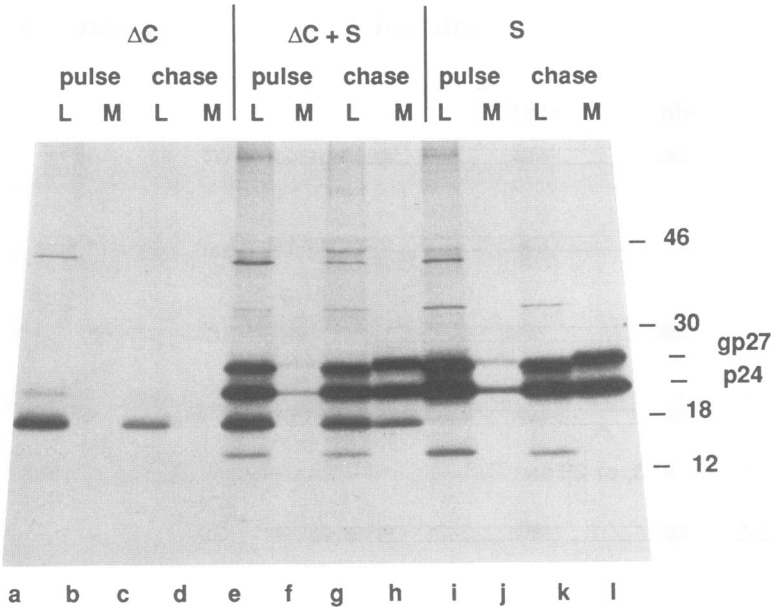

FIG. 2. Mutant ΔC is competent for export only in the presence of wild-type S chains. Duplicate dishes of COS 7 cells were transfected with pΔC alone (lanes a to d), pSV24H alone (lanes i to l), or both plasmids together (lanes e to h). Two days later, dishes were labeled with ^{35}S-methionine for 1 h, then from one dish from each transfection, cell lysates (L) and media (M) were harvested (pulse lanes). To the remaining dishes, excess unlabeled methionine was added, and incubation was continued for 24 h; then the lysates and media were harvested as described in Materials and Methods (chase lanes). Equal fractions of each sample were immunoprecipitated with anti-HBs, and the precipitates were analyzed by SDS-PAGE as described in Materials and Methods. Positions of molecular weight standards are indicated at right. p24 and gp27 denote the positions of the unglycosylated and glycosylated S proteins, respectively.

(11). In another attempt to explore the functions of signal II, we asked whether assembly could occur in a mutant in which signal II had been replaced by the transmembrane domain of the transferrin receptor. This region of the transferrin receptor contains an internal signal sequence that mediates the same type of topology as signal II (N terminus in cytoplasm, C terminus in ER lumen [30]). Our goal was to see whether correct orientation of the chain was all that was required for the subsequent assembly steps or whether signal II made some additional, more specific contribution to this process. But this chimera, too, was unstable in vivo (2).

Next we constructed an in-frame deletion of signal I (mutant ΔsI), created by excision of the *Avr*II-*Xba*I region of the S gene. We have previously shown that this lesion destroys the function of signal I in a coupled in vitro translation-translocation assay (7). However, in the in vitro system the ΔsI mutant chain can still enter the ER bilayer by virtue of the presence of signal II, which can still function to insert itself and direct the correct translocation of the C terminal domains (2, 7). Consistent with this, when ΔsI is expressed in vivo (Fig. 3, lanes a to d) it is glycosylated with normal efficiency (the glycosylation site [24] is in the second hydrophilic domain, just downstream of signal II). The protein is as stable as the wild type (compare lanes a and c) but is also completely defective for secretion. Unlike mutant ΔC, however, coexpression of ΔsI with wild-type S chains does not restore secretion; conversely, ΔsI chains do not significantly interfere with the export of wild-type S chains (lanes e to h). This suggests that although signal I is not required for the insertion of S chains into the bilayer (a function that can be subsumed by signal II), in its absence the topology of the N terminus of the chain is incorrect and the resulting polypeptides are incapable of associating with each other (or with full-length S chains). As a result, they cannot participate in any of the later steps in particle formation.

Chimeric polypeptides with N-terminal extensions. The behavior of ΔsI suggested to us that the correct orientation of the N terminus of the S protein may be critical for assembly. Because one strategy for producing chimeric HBsAg particles involves addition of foreign epitopes to the N terminus of S (28), understanding the impact of such additions on S protein topogenesis and assembly could be of practical significance. Accordingly, we constructed a model chimera in which 100 amino acids of alpha globin sequences were fused to the pre-S2 region, 11 codons 5′ to the S AUG codon. This chimera (GS) therefore contains a fully intact S region. We had earlier shown (6, 7) that the addition of the bulky globin domain interferes with the translocation of the N terminus into the ER lumen in vitro (in this respect it resembles the phenotype of ΔsI). Consistent with this, the mutant behaves similarly to ΔsI in vivo (Fig. 4A, lanes a to d): the chains enter the secretory pathway and are glycosylated, but secretion of the chain is barely detectable. When wild-type S chains are coexpressed, there is some stabilization of the polypeptide but only minimal impact upon its secretion (lanes e to h). Thus, this chain, too, appears defective for aggregation with other S proteins.

Knowing that the disposition of the N-terminus was abnormal in GS and having inferred from ΔsI that correct topogenesis of this region is important to allow proper subunit interactions, we attempted to correct the topology of GS genetically. If the problem was the failure to translocate the N terminus into the ER lumen, we reasoned that this

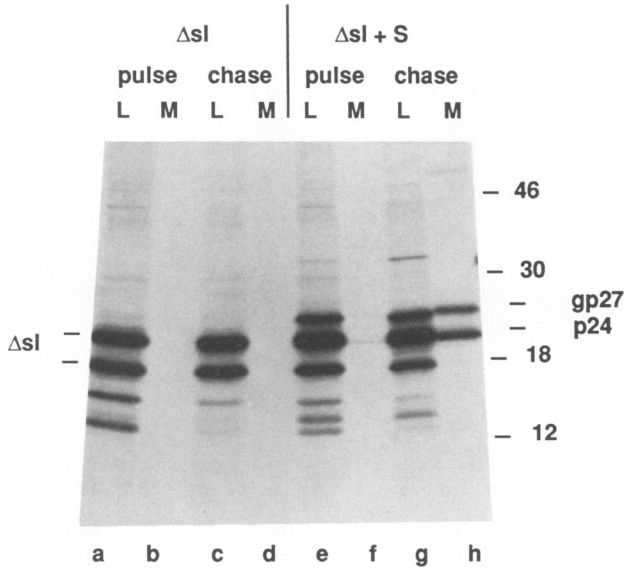

FIG. 3. Secretory phenotype of mutant ΔsI. Duplicate COS 7 cell plates were transfected with either pΔsI alone (lanes a to d) or with pΔsI plus pSV24H (lanes e to h). Two days later, dishes were labeled with ^{35}S-methionine for 1 h. From one set of dishes, lysates (L) and media (M) were harvested immediately (pulse); from the other set, similar harvests followed a 24-h incubation in unlabeled medium (chase). Then equal amounts of all samples were immunoprecipitated with anti-HBs, and the products were displayed by SDS-PAGE. The positions of the molecular weight standards are indicated at the right, as are the positions of p24 and gp27. The positions of the glycosylated and unglycosylated ΔsI mutant chains are indicated at left. (The faster-migrating bands in each cell lysate represent the products of internal translation initiation at each of the internal AUG codons of the S gene.)

might be rectified by provision of a proper N-terminal signal sequence. Accordingly, we constructed mutant sGS, in which the signal sequence of *Escherichia coli* β-lactamase was fused in frame to the N terminus of GS. Transfection of COS 7 cells with sGS revealed a striking phenotype: the chimeric chains were now secreted with wild-type efficiency, even in the absence of cotransfected S chains (Fig. 4A, lanes i to l). CsCl gradient analysis (data not shown) demonstrated that the exported chimera had the buoyant density (1.2 g/ml) characteristic of HBsAg lipoprotein particles. As expected, cotransfection with wild-type S chains did not further enhance sGS export (Fig. 4A, lanes m to p).

The availability of sGS chimeric chains allowed us to prove directly that mixed particles form in cells cotransfected with both S and sGS constructs. As shown in Fig. 4B, when particles harvested from the medium of sh cotransfected cells were precipitated with either anti-HBs or anti-globin antibodies, both S and sGS chains were coprecipitated (lanes a and b). As a control for nonspecific aggregation, particles produced by cultures singly transfected with either S or sGS clones were admixed in vitro and then precipitated with either anti-HBs (lane c) or anti-globin (lane d). As expected, in this case both chains were precipitated by anti-HBs, since both contain an intact S region; however, only sGS chains were precipitated by anti-globin. Thus, under our cotransfection conditions the majority of transfected cells are producing mixed particles with subunits derived from both transfected genes. Identical results were obtained when sGS was cotransfected with mutant ΔC (2).

The ability of the β-lactamase signal to restore secretion to GS led us to ask whether it could now render the topogenesis of the protein independent of signal I. This might be expected if signal I's only role is to ensure disposition of the N terminus into the ER lumen; however, if signal I plays an

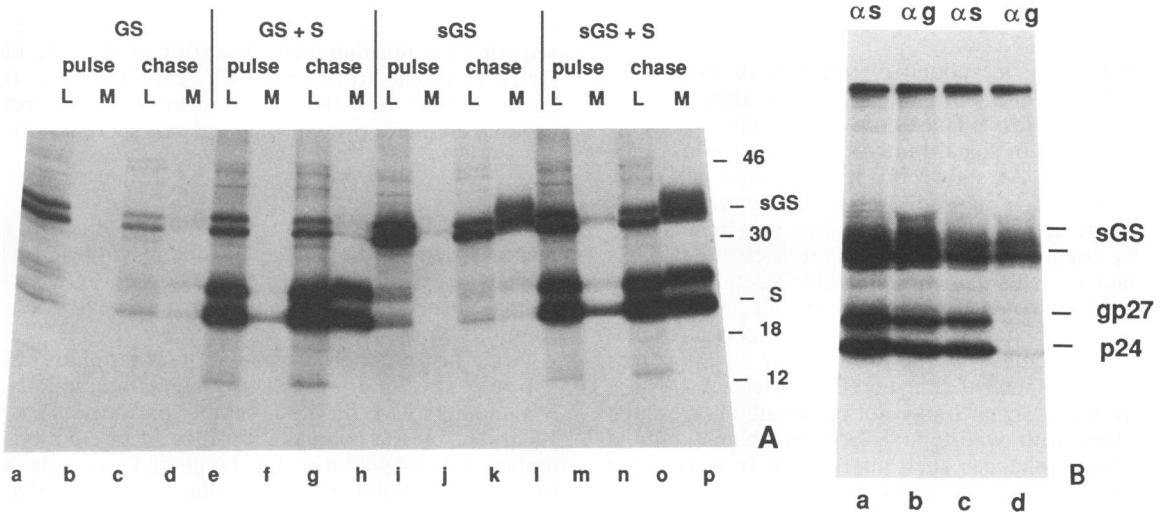

FIG. 4. (A) Secretion of a globin-S chimera is rescued by provision of an N-terminal signal sequence. Duplicate dishes of COS 7 cells were transfected with pGS (lanes a to d), pGS plus pSV24H (lanes e to h), psGS (lanes i to l), or psGS plus pSV24H (lanes m to p). Two days later, dishes were labeled with ^{35}S-methionine for 1 h. From one set of dishes, lysates (L) and media (M) were harvested immediately (pulse); from the other set, similar harvests followed a further 24-h incubation in unlabeled medium (chase). Then equal amounts of all samples were immunoprecipitated with anti-HBs, and the products were displayed by SDS-PAGE. The positions of the molecular weight standards are indicated at the right, as are the positions of unglycosylated S and sGS chains. The bands immediately above the S and sGS species represent their glycosylated derivatives. (B) Coexpression of S and sGS results in mixed particle formation. Cells were cotransfected with S and sGS plasmids (lanes a and b), radiolabeled as described above and medium precipitated with anti-HBs (lane a) or anti-globin (lane b). In parallel, labeled media from cells separately transfected with each plasmid singly were pooled and then precipitated with anti-HBs (lane c) or anti-globin (lane d). Precipitates were then examined by SDS-PAGE.

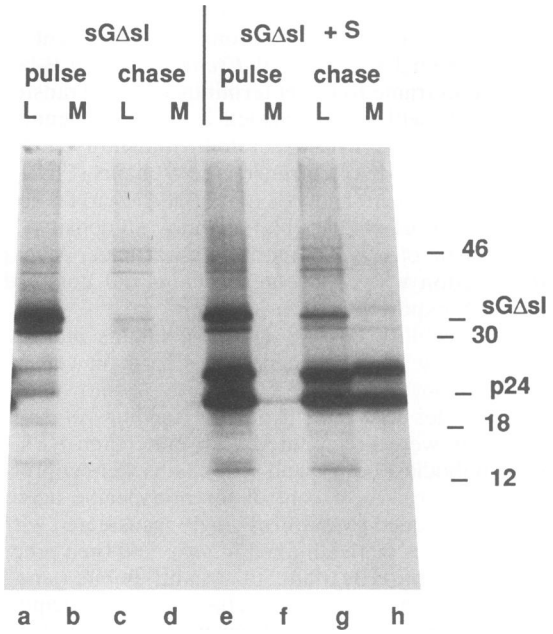

FIG. 5. Effect of signal I deletion on sGS secretion. Duplicate dishes of COS 7 cells were transfected with either psGΔsI (lanes a to d) or psGΔsI plus pSV24H (lanes e to h). Two days later, dishes were labeled with ^{35}S-methionine for 1 h. From one set of dishes, lysates (L) and media (M) were harvested immediately (pulse); from the other set, similar harvests followed a 1-day incubation in unlabeled medium (chase). Then equal aliquots of all samples were immunoprecipitated with anti-HBs, and the products were displayed by SDS-PAGE. The positions of the molecular weight standards are indicated at the right, as are the positions of p24 and sGΔsI polypeptides. The bands immediately above the S and sGΔsI species represent their glycosylated derivatives.

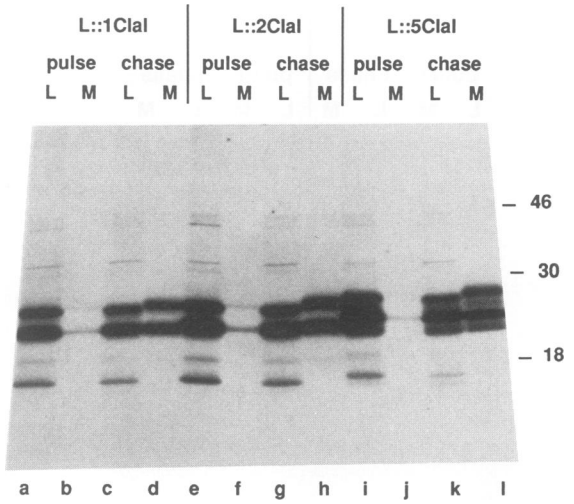

FIG. 6. Linker insertions within the first hydrophilic loop do not interfere with particle assembly or export. Duplicate dishes of COS 7 cells were transfected with derivatives of pSV24H carrying one (L::1ClaI), two (L::2ClaI), or five (L::5ClaI) *Cla*I linkers inserted in the *Bal*I site at codon 50 of the S gene. Two days later, dishes were labeled with ^{35}S-methionine for 1 h. From one set of dishes, lysates (L) and media (M) were harvested immediately (pulse); from the other set, similar harvests followed a 24-h incubation in unlabeled medium (chase). Then equal amounts of all samples were immunoprecipitated with anti-HBs, and the products were displayed by SDS-PAGE. Positions of molecular weight standards are indicated at the right.

additional role (e.g., in subunit interactions or as a membrane-spanning region), then the N-terminal signal in sGS might not substitute for it. To explore this, we introduced the *Avr*II-*Xba*I deletion of signal I into sGS (mutant sGΔsI) and assayed for export. As shown in Fig. 5 (lanes a to d), the resulting polypeptide was quite unstable (compare lanes a and c) and not detectably secreted. But its stability was increased by coexpression with wild-type S chains (lanes e to h), in which case a small but detectable fraction (ca. 5 to 10%) of the pulse-labeled chains that were still stable following the chase were exported into the medium (compare lanes g and h). (The instability of the intracellular chains makes this a minimal estimate of the secretion efficiency.) These results suggest that signal I may not be absolutely essential for aggregation with wild-type S protein chains, but its presence clearly enhances such interactions in a way that cannot be fully recapitulated by simple provision of an exogenous signal sequence.

Insertions in the first hydrophilic loop. To examine the impact of insertions in the first hydrophilic region we inserted in-frame *Cla*I linkers into the *Bal*I site at codon 50 of the S gene, in the central region of the loop. Each *Cla*I linker encodes a 4-amino-acid insertion (Gly-Pro-Ser-Met); clones with one, two, or five tandem copies of the linker (Fig. 1) were isolated and tested. As shown in Fig. 6, the S gene products of all three clones were efficiently secreted into the medium. (Clone L::5Cla, bearing five copies of the linker, frequently underwent homologous recombination within the inserted linkers to generate heterogenous products; lanes i to l.) This indicates that this region of the loop will tolerate small insertions, as has also been found by others (3, 4). However, the capacity of this region to accept foreign sequences is not unlimited: insertion at this site of a 69-bp DNA fragment from the HBV pre-S1 region (encoding residues 26 to 48 of the L protein) resulted in expression of a stable chimeric protein that was not secreted (2).

DISCUSSION

In these studies we have examined the effects of deletions and insertions in the S gene on the biosynthesis and release of HBsAg particles. In general, two new inferences about the possible roles of various domains of the S protein emerge from this work.

First, the proper orientation of the N terminus of the chain appears to be important to allow lateral interactions between S monomers early in the assembly pathway. This was first suggested by the complete absence of interaction between mutant ΔsI and wild-type S subunits; it is strongly supported by the observation that the similar assembly defect caused by the addition of a bulky foreign domain known to impair membrane translocation of the N terminus (6, 7) can be reversed by provision of an N-terminal signal sequence. The simplest model consistent with these results is that the N-terminal region (amino acids 1 to 80) is directly involved in interchain recognition. But it is equally possible that this region is involved in the recognition process indirectly, e.g., by maintaining the proper tertiary structure of the protein so that aggregation domain(s) located elsewhere in the chain can be functional.

Second, the phenotype of mutant ΔC suggests possible functions for the C-terminal hydrophobic region, a domain whose function(s) have hitherto been obscure. Chains lacking this region are clearly still capable of lateral interactions with other subunits, since they can co-oligomerize with wild-type S proteins. Thus, they must retain at least one aggregation-recognition function, though the presence of a second such function within the C-terminal region is not excluded. However, their failure to be secreted without accompanying S chains implies that they lack additional information required for either budding or intracellular transport (or both). From the fact that expression of ΔC does not impair secretion of S we can infer that the mutant does not create a structure that dominantly inhibits these processes (e.g., by competitively binding cellular structures required for their completion). The simplest interpretation is that ΔC lacks the ability to recognize structures (e.g., cellular lipids or membrane proteins) required for these steps. Again, this defect could be due either to direct effects of the deletion (i.e., loss of the binding domain) or to indirect effects arising from an altered tertiary structure of the chain.

Our results also have implications for the design of chimeric HBsAg particles bearing foreign epitopes. First, the behavior of mutant ΔC indicates that some mutants that are assembly or export defective can be rescued by provision of wild-type S protein in *trans*. We therefore asked whether mutants in which the C-terminal domain is replaced by foreign epitopes could be similarly rescued by wild-type S chains, in an effort to develop a general strategy for chimeric particle formation. We constructed a series of S-globin fusion proteins bearing replacements of various portions of the C-terminal hydrophobic region with alpha globin; these were expressed in COS 7 cells either alone or in the presence of wild-type S chains. Unfortunately, the majority of these fusion proteins were unstable in vivo, and the few that were not were inefficiently complemented by the wild type (2).

Similarly, our success with genetically correcting the topology of the GS chimera led us to explore whether provision of N-terminal signal sequences might allow fusion of any protein domain to the N terminus of intact S chains. Accordingly, we constructed additional chimeras in which the N-terminal portions of either the HBV precore (pre-C) protein or the T-cell surface protein CD4 were fused to HBsAg determinants; these foreign proteins provided their own signal sequences at the N terminus of the hybrid. Fusions of the first 150 or 243 residues of CD4 to codon 11 of pre-S2 both resulted in unstable chains (2). Likewise, a chimera bearing 143 residues of HBV pre-C protein fused to codon 11 of pre-S2 was highly unstable in COS 7 cells and could not be assessed for secretion; when the same pre-C segment was fused to codon 11 of the S gene, the resulting chimera was more stable but was neither secreted nor cosecreted with the wild type (2). Thus, although provision of an N-terminal signal sequence may be necessary for the export of large heterologous N-terminal fusions, it is clearly not sufficient.

These results emphasize that the addition of bulky foreign sequences can have unpredictable deleterious effects on global protein structure. Indeed, in our experience the twin problems of misfolding and instability have been the main impediments to the development of generally applicable strategies for chimeric particle formation. These problems can perhaps be minimized by the use of very small foreign epitopes so as to less radically disrupt protein structure.

ACKNOWLEDGMENTS

We thank Vishwanath Lingappa for help with in vitro translocation experiments and for many stimulating discussions and Wolfram Gerlich for helpful comments on the manuscript.

This work was supported by grants from the National Institutes of Health. V.B. was a scholar of the Deutsche Forschungsgemeinschaft.

REFERENCES

1. **Bruss, V., and D. Ganem.** 1991. The role of envelope proteins in hepatitis B virus assembly. Proc. Natl. Acad. Sci. USA **88:** 1059–1063.
2. **Bruss, V., and D. Ganem.** Unpublished observations.
3. **Delpeyroux, F., N. Chenciner, A. Lim, M. Lambert, Y. Malpiece, and R. E. Streeck.** 1987. Insertions in the hepatitis B surface antigen: effect on assembly and secretion of 22-nm particles from mammalian cells. J. Mol. Biol. **195:**343–350.
4. **Delpeyroux, F., N. Chenciner, A. Lim, Y. Malpiece, B. Blondel, R. Crainic, F. van der Werf, and R. E. Streeck.** 1986. A poliovirus neutralization epitope expressed on hybrid hepatitis B surface antigen particles. Science **233:**472–475.
5. **Delpeyroux, F., N. Peillon, B. Blondel, R. Crainic, and R. E. Streeck.** 1988. Presentation and immunogenicity of the hepatitis B surface antigen and poliovirus neutralization antigen on mixed empty envelope particles. J. Virol. **62:**1836–1839.
6. **Eble, B. E., V. R. Lingappa, and D. Ganem.** 1986. Hepatitis B surface antigen: an unusual secreted protein initially synthesized as a transmembrane polypeptide. Mol. Cell. Biol. **6:**1454–1463.
7. **Eble, B. E., D. R. Macrae, V. R. Lingappa, and D. Ganem.** 1987. Multiple topogenic sequences determine the transmembrane orientation hepatitis B surface antigen. Mol. Cell. Biol. **7:**3591–3601.
8. **Ganem, D., and H. E. Varmus.** 1987. The molecular biology of the hepatitis B viruses. Annu. Rev. Biochem. **56:**651–694.
9. **Gavilanes, F., A. Gonzales-Ros, and D. Peterson.** 1982. Structure of hepatitis B surface antigen: characterization of the lipid components and their association with the viral proteins. J. Biol. Chem. **257:**7770–7777.
10. **Gerber, M., S. Hadziyannis, C. Vissoulis, F. Schaffner, F. Paronetto, and H. Popper.** 1974. Electron microscopy and immunoelectron microscopy of cytoplasmic hepatitis B antigen in hepatocytes. Am. J. Pathol. **75:**489–502.
11. **Goldberg, A., and A. St. John.** 1976. Intracellular protein degradation in mammalian and bacterial cells. Annu. Rev. Biochem. **45:**747–803.
12. **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.
13. **Kelly, R.** 1985. Pathways of protein secretion in eukaryotes. Science **230:**25–31.
14. **Kunkel, T. A., J. D. Roberts, and R. A. Zabour.** 1987. Rapid and efficient site-specific mutagenesis without phenotypic selection. Methods Enzymol. **154:**376–382.
15. **Kuroki, K., M. Floreani, L. Mimms, and D. Ganem.** 1990. Epitope mapping of the pre S1 domain of the hepatitis B virus large surface antigen. Virology **176:**620–624.
16. **Kuroki, K., R. Russnack, and D. Ganem.** 1989. A 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.
17. **Laub, O., L. B. Rall, M. Truett, Y. Shaul, D. N. Standring, P. Valenzuela, and W. J. Rutter.** 1983. Synthesis of hepatitis B surface antigen in mammalian cells: expression of the entire gene and the coding region. J. Virol. **48:**271–280.
18. **Liu, C. C., D. Yansura, and A. Levinson.** 1982. Direct expression of hepatitis B surface antigen in monkey cells from an SV40 vector. DNA **1:**213–221.
19. **Michel, M. L., M. Mancini, E. Sobczak, V. Favier, D. Guetard, E. M. Bahraoui, and P. Tiollais.** 1988. Induction of anti-human immunodeficiency virus (HIV) neutralizing antibodies in rabbits immunized with recombinant HIV-hepatitis B surface antigen

particles. Proc. Natl. Acad. Sci. USA **85**:7957–7961.
20. **Patzer, E. J., G. R. Nakamura, G. C. Simonsen, A. D. Levinson, and R. Brands.** 1986. Intracellular assembly and packaging of hepatitis B surface antigen particles occur in the endoplasmic reticulum. J. Virol. **58**:884–892.
21. **Patzer, E. J., G. R. Nakamura, and A. Yaffe.** 1984. Intracellular transport and secretion of hepatitis B surface antigen in mammalian cells. J. Virol. **51**:346–353.
22. **Perara, E., R. E. Rothman, and V. R. Lingappa.** 1986. Uncoupling translocation from translocation: implications for transport of proteins across membranes. Science **232**:348–352.
23. **Persing, D. H., H. E. Varmus, and D. Ganem.** 1985. A frameshift mutation in the pre-S region of the human hepatitis B virus genome allows production of surface antigen particles but eliminates binding to polymerized albumin. Proc. Natl. Acad. Sci. USA **82**:3440–3444.
24. **Peterson, D.** 1987. The structure of hepatitis B surface antigen and its antigenic sites. Bioessays **6**:258–262.
25. **Peterson, D. L.** 1981. Isolation and characterization of the major protein and glycoprotein of hepatitis B surface antigen. J. Biol. Chem. **256**:6975–6983.
26. **Simon, K., V. Lingappa, and D. Ganem.** 1988. Secreted hepatitis B surface antigen polypeptides are derived from a transmembrane precursor. J. Cell. Biol. **107**:2163–2168.
27. **Szmuness, W., C. E. Stevens, E. J. Harley, E. A. Zang, W. R. Oleszko, D. C. William, R. Sadovsky, and J. M. Morrison.** 1980. Hepatitis B vaccine. N. Engl. J. Med. **303**:833–841.
28. **Valenzuela, P., D. Coit, A. Medina-Selby, C. Kuo, G. Van Nest, R. L. Burke, P. Bull, M. Urdea, and P. Graves.** 1985. Antigen engineering in yeast: synthesis and assembly of hybrid hepatitis B surface antigen-herpes simplex gD particles. Bio/Technology **3**:323–326.
29. **Valenzuela, P., L. Rall, M. Zaldivar, M. Quiroga, P. Gray, and W. Rutter.** 1980. The nucleotide sequence of the hepatitis B viral genome and the identification of the major viral genes. ICN-UCLA Symp. Mol. Cell. Biol. **18**:57–70.
30. **Zerial, M., P. Melancon, C. Schneider, and H. Garoff.** 1986. The transmembrane segment of the human transferrin receptor functions as a signal peptide. EMBO J. **5**:1543–1550.