JOURNAL OF VIROLOGY, July 1985, p. 223–231
0022-538X/85/070223-09$02.00/0

Vol. 55, No. 1

Identification of Hepatitis B Virus Polypeptides Encoded by the Entire Pre-*s* Open Reading Frame

DAVID T. WONG,[1] NRAPENDRA NATH,[2] AND JOHN J. SNINSKY[1,3]†*

Departments of Microbiology and Immunology[1] and Molecular Biology,[3] Albert Einstein College of Medicine, Bronx, New York 10461, and Blood Services Laboratories, American Red Cross, Bethesda, Maryland 20014[2]

Received 14 January 1985/Accepted 30 March 1985

The open reading frame (ORF) that encodes the 226-amino-acid coat protein (hepatitis B virus surface antigen [HBsAg]) of hepatitis B virus has the potential to encode a 400-amino-acid polypeptide. The entire ORF would direct the synthesis of a polypeptide whose C-terminal amino acids represent HBsAg with an additional 174 amino acids at the N terminus (pre-*s*). Recently, virus particles have been shown to contain a polypeptide that corresponds to HBsAg with an additional 55 amino acids at the N terminus encoded by the DNA sequence immediately upstream of the HBsAg gene. A novel ORF expression vector containing the TAC promoter, the first eight codons of the gene for β-galactosidase, and the entire coding sequence for chloramphenicol acetyltransferase was used in bacteria to express determinants of the 174 amino acids predicted from the pre-*s* portion of the ORF. The resulting tribrid protein containing 108 amino acids encoded by pre-*s* was expressed as one of the major proteins of bacteria harboring the recombinant plasmid. Single-step purification of the tribrid fusion protein was achieved by fractionation on a chloramphenicol affinity resin. Polyclonal antiserum generated to the fusion protein was capable of detecting 42- and 46-kilodalton polypeptides from virus particles; both polypeptides were also shown to contain HBsAg determinants. The ability of the polyclonal antiserum to identify polypeptides with these characteristics from virus particles presents compelling evidence that the DNA sequence of the entire ORF is expressed as a contiguous polypeptide containing HBsAg. The presence of multiple promoters and primary translation products from this single ORF argues that the function and potential interaction of the encoded polypeptides play a crucial role in the life cycle of the virus. Furthermore, the procedure and vector described in this report can be applied to other systems to facilitate the generation of antibodies to defined determinants and should allow the characterization of the epitope specificity of existing antibodies.

One of the most interesting observations made after the determination of the total DNA sequence of a hepatitis B viral (HBV) genome was that the open reading frame (ORF) which encodes the viral coat protein (HBV surface antigen [HBsAg]) potentially could direct the synthesis of a 400-amino-acid polypeptide containing the 226 amino acids of HBsAg at the C terminus and an additional 174 amino acids at the N terminus (pre-*s*) (10, 43). The finding that all other HBV genomes sequenced to date (24, 25, 43), as well as viral genomes of members of this new class of viruses which infect woodchucks (woodchuck hepatitis virus [WHV]) (9), ground squirrels (ground squirrel hepatitis virus), (34), and Pekin ducks (duck hepatitis B virus) (21), contain this ORF supports the initial suggestion that a larger HBsAg-containing polypeptide may be synthesized (10). Furthermore, utilization of a conserved internal ATG in the pre-*s* region of the ORF would allow for initiation of a polypeptide with an additional 55 amino acids at the N terminus of HBsAg or 60 amino acids at the N termini of WHV and ground squirrel hepatitis virus surface antigens. Whereas the amino acid sequence of the coat protein of the hepadna viruses is conserved, the predicted amino acid sequence of pre-*s* displays little homology among viruses. Comparative computer-assisted analysis of the pre-*s* region from WHV and HBV suggested that although there is little homology at the amino acid sequence level, the overall predicted propensities to form secondary structure were similar (32). Schaeffer and Sninsky (32) went on to speculate that the sequences responsible for the marked species specificity displayed by members of the hepadna viruses may reside in the N-terminal 174 amino acids predicted by the entire ORF that encodes HBsAg.

Evidence for the polypeptide predicted to initiate from the conserved internal AUG in the pre-*s* region representing an extended coat protein has been presented (19, 23, 40). First, Stibbe and Gerlich (40) showed, using proteolytic digestions, that a 33- and a 36-kilodalton (kd) protein purified from virus particles corresponded to p24 and gp27 (commonly identified as pI and pII, respectively, from HBsAg preparations [27]) with an N-terminal extension. They further suggested that translation initiation could occur at the internal AUG in the pre-*s* region, assuming, however, the absence of transcript splicing. Second, Neurath et al. (23), using antibodies to a synthetic peptide corresponding to the first 26 amino acids of the predicted polypeptide initiating at the internal AUG, also identified a 33- and a 36-kd polypeptide from virus particles. These observations are consistent with transcriptional studies which have identified various mRNA species which could direct the translation of HBsAg-containing polypeptides larger than p24 and gp27. One series of transcripts initiates within the pre-*s* region at a simian virus 40 late promoter-like region and could be utilized for the expression of HBsAg and pre-s_{55}:HBsAg (4, 38). The other series initiates near the putative TATA box promoter upstream of the entire ORF and could direct the synthesis of an additional product, pre-s_{174}:HBsAg (18, 28, 29).

Possible insight into the function of the pre-*s* encoded region of this protein comes from studies by Machida et al. (19, 20) which show that the previously observed specific

* Corresponding author.
† Present address: Cetus Corporation, Emeryville, CA 94608.

binding of virus particles to polymerized human serum albumin occurs via recognition by the 31-kd protein but not p24 or gp27 that represent HBsAg. Although evidence for a protein initiating at the internal AUG of transcripts predicted by the ORF is overwhelming, the use of the first AUG to initiate translation is less clear. Coincident with our computer studies, we embarked on an experimental approach by using a tribrid fusion protein expression vector to synthesize determinants from the pre-*s* region. This approach was particularly attractive since determinants of HBsAg were not present in the fusion proteins, thereby allowing the unambiguous interpretation of the data. We report here evidence for the existence of virus particles of two glycosylated polypeptides of 42 and 46 kd containing determinants of both HBsAg and the pre-*s* region.

MATERIALS AND METHODS

DNA manipulation and analysis. Plasmid DNA was extracted and analyzed with the appropriate restriction endonucleases by using agarose and acrylamide gels as described previously (33). For subcloning, 5 to 25 μg of plasmid DNA was cleaved with restriction enzymes. The digestion conditions were those recommended by the suppliers (New England Biolabs and Boehringer Mannheim Biochemicals). The desired fragments were separated on 5% acrylamide gels and isolated by electroelution as previously described (33). To decrease the frequency of recircularization, plasmid DNA was treated with calf intestine alkaline phosphatase (Boehringer Mannheim Biochemicals) as described by Ullrich et al. (42). DNA fragments (5 to 20 μg/ml) were then ligated with T_4 DNA ligase (New England Biolabs) and used to transform bacteria containing a *lac* deletion, strain MC1000 (*lac* IPOZYx74) (3), by the $CaCl_2$ procedure essentially as described (5). Transformed bacteria were plated on Difco antibiotic medium no. 2 containing 25 μg of ampicillin per ml and either 5 or 25 μg of chloramphenicol per ml.

Purification of β-gal:CAT hybrid and β-gal:pre-*s*:CAT tribrid fusion proteins. Purification of β-gal (β-galactosidase): CAT (chloramphenicol acetyltransferase) hybrid and β-gal: pre-*s*:CAT tribrid fusion proteins was done as follows. Bacterial strain MC1000 harboring either pZL811 or pDW111 (see below) was grown in Luria broth (22) with 25 μg of ampicillin per ml at 37°C to an A_{650} of 1.3 and centrifuged in 500-ml bottles at 7,000 rpm for 10 min. Four liters of harvested culture resulted in 14 g of bacterial paste. All subsequent steps were performed at 4°C. Bacterial pellets were resuspended to 0.3 g/ml in cold TEP buffer (100 mM Tris-hydrochloride [pH 7.4], 10 mM EDTA, 1 mM phenylmethylsulfonyl fluoride [100 mM stock in ethanol]) and sonicated with a Heat Systems Cell Disrupter in a rosette flask on ice at setting 8 and 50% duty cycle for four 5-min intervals. The insoluble fraction was removed by centrifugation at 12,000 × *g*, and the remaining supernatant was made 4% streptomycin sulfate to precipitate nucleic acids. Bacterial proteins, including the fusion proteins, were then precipitated by the addition of neutralized saturated ammonium sulfate to 75% saturation. The pellets were washed twice, resuspended, and dialyzed against TEP before loading onto the affinity resin.

An established procedure (45) and a newly developed procedure were used to generate an affinity column with *p*-amino chloramphenicol as the ligand. Resins synthesized by either of the two methods gave similar results. Fourteen grams of activated CH-Sepharose 4B (Pharmacia Fine Chemicals) was hydrated to a gel volume of ca. 40 ml, washed extensively with cold 1 mM HCl, and coupled with 1.0 g of chloramphenicol base (Sigma Chemical Co.) dissolved in 100 ml of 0.1 M $NaHCO_3$ (pH 8.5) for 2 h at 27°C with rocking. Unbound sites were blocked with the addition of 15 ml of 1.0 M Tris-hydrochloride (pH 8.0) for an additional 2 h. Unbound chloramphenicol was removed by extensive alternating washings with 0.1 M sodium acetate (pH 4.5)–150 mM NaCl and 0.1 M Tris-hydrochloride (pH 8.0)–150 mM NaCl. Chloramphenicol was also attached by the established method of Zaidenzaig and Shaw (45); a 40-ml bed volume column substituted with chloramphenicol was achieved by mixing 2 mol of ligand per mol of carboxyl spacer arm group. To prevent irreversible binding of CAT fusions, both resins were treated with 0.1 M NaOH as described by Goldfarb et al. (12) and before use were washed with TEP containing 1 mg of chloramphenicol per ml and 0.5 M NaCl, followed by reequilibration with TEP.

The salt fractionated lysate from a 4.0-liter culture was diluted in TEP to 150 ml, loaded on the chloramphenicol affinity column, and washed until the A_{280} was 0. One hundred milligrams of bound hybrid or tribrid fusion proteins were eluted with TEP containing 1 mg of chloramphenicol per ml. The elution profiles of the hybrid and tribrid fusion proteins were determined by a colormetric CAT assay (35), by 12.75% sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE), and by Bradford protein assays (Bio-Rad Laboratories). For rapid estimates of CAT activity during protein purifications, CAT assay reactions were scaled down 10-fold (total volume, 100 μl), and the appearance of yellow color in microtiter wells was followed. The peak fractions of chloramphenicol-eluted protein were pooled and used for all experiments.

Generation of antisera. Three milligrams of tribrid fusion polypeptide purified from the chloramphenicol affinity column was separated by 3.0-mm 12.75% preparative SDS-PAGE. The band corresponding to the fusion protein was visualized in a cold 1.0 M KCl solution, excised, and grated through a medium mesh wire screen. The polypeptide was eluted in 0.3 M Tris-hydrochloride (pH 8.8) and 0.1% SDS for 1 h at 27°C and then for 10 h at 4°C. The grated acrylamide was pelleted by centrifugation, and the supernatant was mixed with an equal volume of complete Freund adjuvant (Difco Laboratories) until an emulsion formed. Rabbits were immunized with 250 μg of gel-purified tribrid protein by three intradermal injections. Four weeks later, the rabbits were given booster injections. The secondary injection was prepared as above with the substitution of incomplete Freund adjuvant. Beginning 7 days later, the rabbits were bled weekly. Sera were prepared, and the hyperimmune titers were determined by enzyme-linked immunosorbent assays. The hyperimmune titers for each rabbit serum were similar and remained high for 3 months before exsanguination. One representative immune serum was chosen and used in all experiments described in this report.

Enzyme-linked immunosorbent assay. Wells of Falcon Microtest III Flexible Assay Plates were coated with 100 μl of antigen diluted to 1 μg/ml in phosphate-buffered saline (PBS; 10 mM phosphate [pH 7.2], 150 mM NaCl) at 37°C for 2 h. The plates were blocked with 1% bovine serum albumin in PBS for 1 h at 37°C. The wells were incubated with 100 μl of antisera diluted from 10^{-1} to 10^{-6} in blocking buffer for 1 h at 37°C and washed extensively with 0.05% Tween 20 in PBS. Alkaline phosphatase-conjugated goat anti-rabbit immunoglobulin G (IgG) (Tago) was diluted 1:3,000 in blocking buffer, and 100 μl per well was incubated for 1 h at 37°C and washed as above. The plates were washed with 0.15 M NaCl

and incubated with 100 μl of 1 mg of alkaline phosphatase substrate (Sigma) per ml in 1.0 M Tris-hydrochloride (pH 9.8)–0.5 mM $MgCl_2$ until color in the positive control wells developed. The reaction was stopped with 100 μl of 5.0 N NaOH, and the plates were read at 405 nm in an ARTEK microtiter reader.

Virus preparation. The procedure used to concentrate and partially purify HBV particles was similar to that described by Gerlich and Robinson (11). Plasma samples positive for HBsAg, HBV e antigen, and DNA polymerase activity were selected for purification. One sample, ARC 6389 of the *ad* subtype (Ausria S/N 85.4; Abbott Laboratories), was found to contain the pre-*s* polypeptides and was used for all experiments described here; other sera have been shown to contain various levels of the pre-*s* polypeptides. Briefly, the plasma was clarified at 12,000 × *g* for 20 min, and 23-ml samples were layered over 15 ml of 10, 20, and 30% sucrose step gradients in TNE (0.01 M Tris-hydrochloride [pH 7.4], 0.1 M NaCl, 1.0 mM EDTA). The gradients were centrifuged in a Beckman SW27 rotor at 24,000 rpm for 24 h at 4°C, and the pellets were resuspended in 100 μl of TNE. These resuspended pellets were combined and layered over a second 10, 20, and 30% sucrose step gradient in a Beckman SW41 rotor and centrifuged at 22,000 rpm for 24 h at 4°C. The pellet was resuspended in 200 μl of TNE, resulting in a 650-fold concentration of HBV particles.

Protein blotting procedures. Viral polypeptides were identified by the method of Towbin et al. (41) with slight modifications. Specifically, the viral polypeptides were separated by 10% SDS-PAGE (17); 2.5 or 5.0 μl of a 650-fold-concentrated virus preparation was solubilized in sample buffer (2.5% SDS, 5% β-2-mercaptoethanol, 10% glycerol, 62 mM Tris-hydrochloride [pH 6.8]) and loaded per lane. ^{14}C-labeled proteins (New England Nuclear Corp.) were used as molecular weight markers. The separated polypeptides were then transferred to nitrocellulose (Schleicher & Schuell, Inc.) by using a Transblot apparatus (Bio-Rad) at a constant current (250 mA) for 6 h in a 192 mM glycine–25 mM Tris-hydrochloride (pH 8.3)–20% methanol buffer (41). Unoccupied sites on the nitrocellulose were blocked with 2% bovine serum albumin in PBS at 37°C for 1 h.

For the identification of specific polypeptides with pre-*s* or HBsAg determinants, these protein blots were probed with a rabbit anti-β-gal:pre-*s*:CAT or rabbit anti-HBsAg synthetic peptide (amino acids 135 to 155; kindly provided by A. R. Neurath, New York Blood Center) diluted 1:100 in the blocking buffer for 10 h at 4°C. In competition studies, the initial antibody solution was preincubated at 27°C for 10 min with either affinity-purified β-gal:CAT or β-gal:pre-*s*:CAT protein at concentrations of 5, 50, or 500 μg/ml. Nonspecifically bound antibodies were removed by extensive washing with 0.05% Tween 20 (Bio-Rad) in PBS, and the blots were incubated for 1 h at 4°C with either 10^6 cpm of ^{125}I-labeled protein A (70 to 100 μCi/μg; New England Nuclear Corp.) or a 1:1,000 dilution of horseradish peroxidase-conjugated (HRP) goat anti-rabbit immunoglobulin antibody (Bio-Rad) diluted in blocking solution. The blots were again washed extensively with 0.05% Tween 20 in PBS and then with PBS. The ^{125}I-protein A-labeled blots were dried and autoradiographed on Kodak XAR5 film at −70°C (10 h) by using a Cronex Lightning Plus intensifying screen. The HRP-labeled blots were developed 1 to 2 min in a solution containing 20 ml of 3 mg of 4-chloro-1-napthol per ml in cold methanol and 100 ml of 0.018% H_2O_2 in PBS.

To demonstrate glycosylation of certain viral polypeptides, a nitrocellulose strip containing viral polypeptides was probed with HRP-conjugated succinylated concanavalin A (ConA-HRP) (Vector Laboratories) at a 1:100 dilution in blocking solution for 30 min at 27°C, washed, and developed as described above.

RESULTS

Construction and characterization of pre-*s* expression plasmids. Figure 1 illustrates the method and construction of an expression vector that directs the synthesis of a tribrid fusion protein containing 108 amino acids of the 174 amino acids potentially encoded by the pre-*s* region from hepatitis B virus (Fig. 2). The parent ORF expression vector, pZL811, is a pBR322 derivative which contains the TAC hybrid promoter and the first eight codons of the gene for β-gal (*lacZ*) to provide the necessary signals for initiating optimal transcription and translation. The TAC promoter containing the −35 region of the *trp* promoter, the −10 region of the *lac* promoter, and the *lac* operator has been shown to be a highly efficient regulated promoter (2, 6, 7, 30). Adjoining the DNA that encodes the N terminus of β-gal is a polylinker sequence containing a unique cleavage site for *Bam*HI but devoid of translational stop codons. Immediately downstream lies the structural gene for CAT. A similar plasmid with the *lacUV5* promoter has been shown to direct the synthesis of an enzymatically active β-gal:CAT fusion protein, rendering bacteria harboring this plasmid resistant to both ampicillin and chloramphenicol. Although the start codon for CAT is retained in transcripts of these plasmids, the efficient initiation of translation at the *lacZ* start codon upstream substantially suppresses translation initiation at the CAT AUG, as shown previously (33). The hybrid protein encoded by the parent ORF vector represents one of the major proteins synthesized by the bacteria (Fig. 3, lane B) and can be purified to near homogeneity by using a chloramphenicol affinity column (Fig. 3, lane C).

We noted that a significant portion of pre-*s* was contained in a 318-base-pair *Sau*3A-*Bam*HI fragment of HBV (adw_2) (43) and could be conveniently inserted at the *Bam*HI endonuclease site of the previously described ORF vector, pZL811, since the termini generated by these two endonucleases are identical. The HBV (adw_2) genome was initially made fully double-stranded and linearized with *Eco*RI before insertion at the *Eco*RI endonuclease site of pACYC184 (36). Since the *Eco*RI site interrupts the desired *Sau*3A-*Bam*HI fragment, it was necessary to purify the DNA fragment from an *Eco*RI head-to-tail dimer of HBV DNA inserted in pBR322 (A. Siddiqui and J. J. Sninsky, unpublished data). To simplify purification, a *Bgl*II-*Bam*HI DNA fragment, which included the sequence of interest (see Fig. 1), was obtained by digestion with these endonucleases, followed by preparative gel electrophoresis and electroelution. Subsequently, this purified fragment was digested with *Sau*3A, and the 318-base-pair fragment was purified similarly. The purified DNA fragment was ligated to *Bam*HI-linearized pZL811, which was alkaline phosphatase-treated to minimize recircularization of the vector (42). After transformation and selection for ampicillin resistance, candidate colonies resistant and sensitive to chloramphenicol (25 μg/ml) were identified. Plasmid DNA was isolated from both chloramphenicol-resistant and -sensitive colonies and analyzed (data not shown). *Bam*HI endonuclease digestions were used to identify recombinant molecules that had undergone an insertion and verify the regeneration of a unique *Bam*HI endonuclease site at the downstream junction. The location of an asymmetric *Pst*I endonuclease site in the

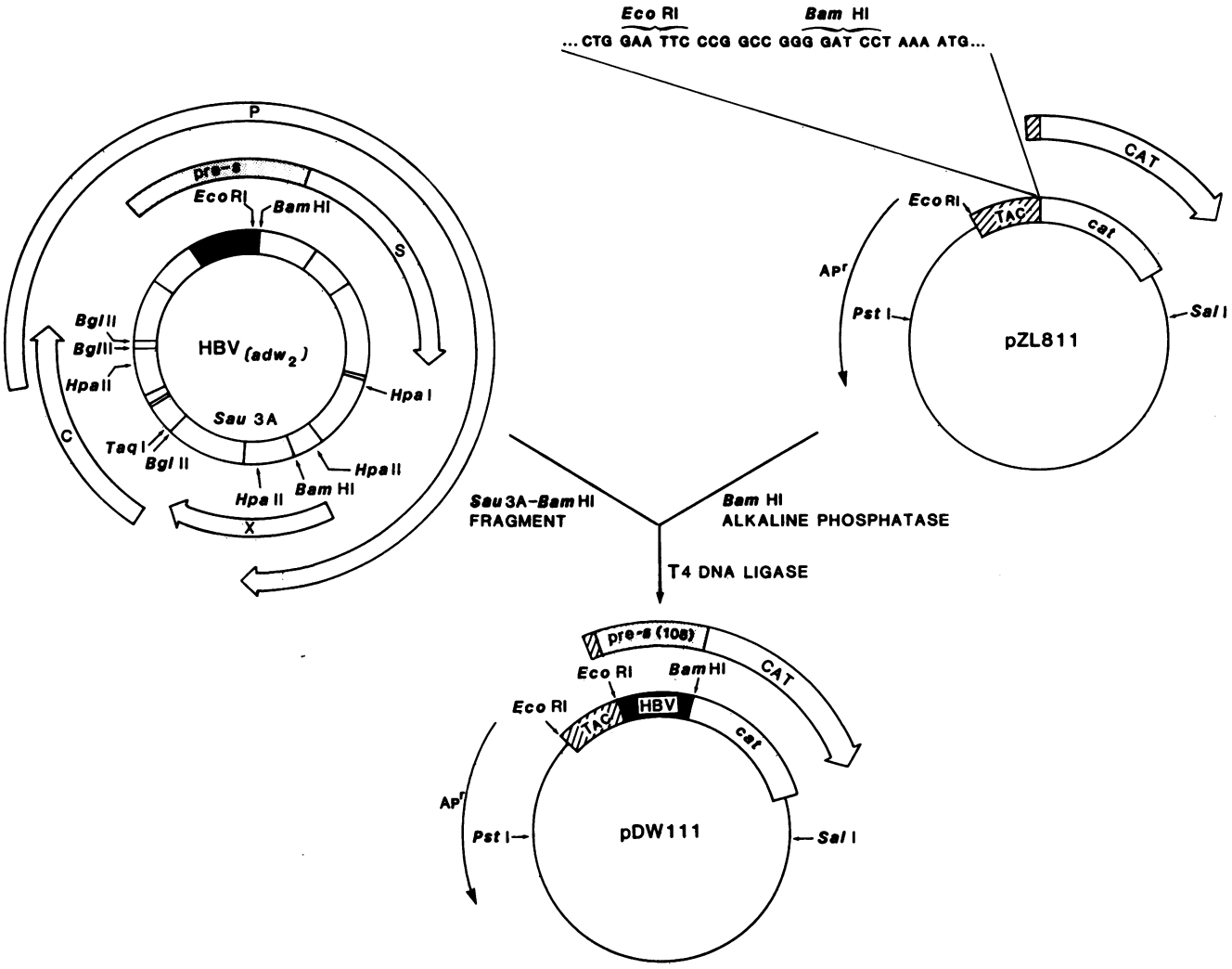

FIG. 1. Vector construction. The expression of determinants encoded by the pre-*s* region was directed by pDW111, which was constructed by the insertion of a 318-base-pair *Sau*3A-*Bam*HI DNA fragment containing pre-*s* sequences (solid block) into an ORF expression vector, pZL811. Purification of this fragment was achieved from an *Eco*RI-cloned dimer of the HBV genome (see below). However, for the sake of simplicity, only the HBV (adw_2) genome and the location of ORFs (arrows) identified from sequence data of the 3,221-base-pair genome are depicted (43). To date, viral protein products have been identified for ORF S (HBsAg) and ORF C (HBV core antigen and e antigen); the products of ORF P and ORF X have not been identified. The shaded region of ORF S, pre-s_{174}, can potentially encode 174 additional amino acids at the N terminus of the polypeptide that represents HBsAg. The parent plasmid pZL811, a pBR322 derivative, contains a TAC promoter (hatched region of pZL811)-regulated hybrid gene which includes sequences encoding the first eight amino acids of β-gal, a linker containing a unique *Bam*HI site, and the gene encoding CAT (open block in pZL811). This plasmid directs the expression of the β-gal:CAT hybrid fusion protein depicted by the hatched and open regions, respectively, of the arrow above pZL811. The sequence presented at the top of the figure represents the polylinker region between the triplet for the eighth amino acid of β-gal (CTG) and the initiation triplet (ATG) for CAT. The *Bam*HI endonuclease site used for insertion of the HBV fragment is bracketed. pDW111 directs the expression of the β-gal:pre-*s*:CAT tribrid fusion protein containing 108 amino acids of pre-s_{174}, designated by the shaded region of the arrow above pDW111. (Plasmids are not drawn to scale.)

inserted *Sau*3A-*Bam*HI fragment and the unique *Pst*I site in the parent vector allowed the identification of the orientation of the insert (data not shown). The results of such analyses indicate that 60% of the plasmid candidates contained inserts. Although bacteria containing recircularized parent plasmids were resistant to 25 μg of chloramphenicol per ml, bacteria harboring plasmids with the fragment in the incorrect orientation for the production of a tribrid protein containing pre-*s* determinants were sensitive to chloramphenicol (5 μg/ml). Bacteria carrying plasmids with the HBV fragment in the correct orientation were found to be resistant to 5 μg of chloramphenicol per ml but were sensitive to 25 μg of chloramphenicol per ml. Previous studies have shown that bacteria containing pBR322 are sensitive to 5 μg of chloramphenicol per ml (33). A representative plasmid containing the HBV fragment in the correct orientation that directed the synthesis of a fusion polypeptide of the appropriate size was designated pDW111 and was used in all subsequent studies. In addition, the candidates with pDW111 were found to have a slower growth rate than bacteria carrying either the parent plasmid or recombinant plasmids with the HBV fragment inserted in the incorrect orientation.

Expression of pre-*s* determinants in a tribrid protein. Insertion of the *Sau*3A-*Bam*HI fragment into the *Bam*HI site of

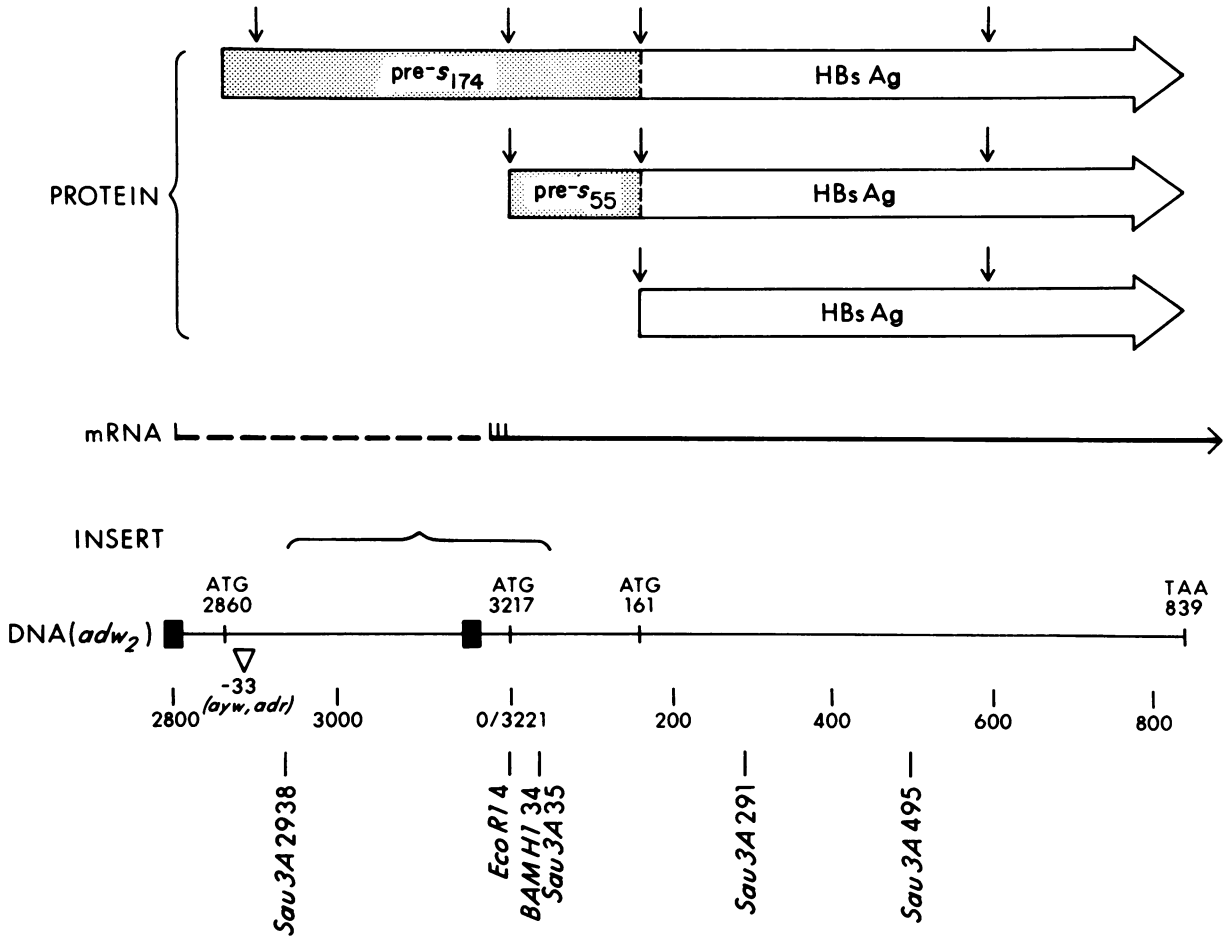

FIG. 2. Schematic diagram of ORF S. Relevant endonuclease sites and their positions in the HBV genome are presented below the line labeled DNA. The unique HBV (adw_2) *Eco*RI site is at nucleotide 1 in this numbering system. The location of translational start (ATG) and stop (TAA) signals in the ORF are indicated above the line labeled DNA. The solid blocks identify promoters potentially utilized. The pre-*s* region that is expressed by the *Sau*3A-*Bam*HI insert of pDW111 in the tribrid fusion polypeptide is defined by the bracket. The mRNA coding for pre-s_{55}:HBsAg is represented by the solid-arrow part of the line (4, 38); additional upstream sequences for pres_{174} mRNA are represented by the dashed part of the line (18, 28, 29). The translation products, all with HBsAg at the C terminus, are diagrammed, with the pre-*s*-encoded amino acids shaded. Arrows mark possible N-linked glycosylation sites as determined by the primary amino acid sequence Asn-X-Ser (Thr).

the parent vector would allow in-frame translation of the desired HBV sequence as a tribrid fusion polypeptide with β-gal at the N terminus (8 amino acids), pre-*s*-encoded internal sequences (108 amino acids), and CAT at the C terminus (219 amino acids). To produce an unambiguous tribrid protein, it is essential to take particular note of the translational reading frame of the desired *Sau*3A-*Bam*HI fragment to ensure precise in-frame insertion into the parent ORF vector because this region also encodes a portion of the largest open reading frame (ORF P) of HBV, but in another reading frame.

The 318-base-pair *Sau*3A-*Bam*HI fragment of HBV (adw_2) encodes amino acids 27 through 133 of the pre-s_{174} region. Although the nucleotide immediately downstream of the *Bam*HI recognition site is thymidine in the parent ORF vector and cytidine in the HBV sequence, due to the degeneracy of the proline codon (i.e., CCU or CCC), an additional amino acid from HBV not directed by the inserted DNA fragment is encoded in the tribrid gene. This results in an insert of 108 amino acids in the tribrid protein, representing the major hydrophilic regions of pre-s_{174} (32), and therefore, according to the method of Hopp and Woods (15), the tribrid protein should contain one or more antigenic determinants.

One of the major proteins synthesized by bacteria containing pDW111 is the predicted 40-kd polypeptide (Fig. 3, lane D). The tribrid protein is recognized by antibodies generated to CAT (data not shown) and can be purified by using a chloramphenicol affinity resin (Fig. 3, lane E). The additional lower-molecular-weight bands that copurify from the affinity resin can be removed by preparative gel electrophoresis (Fig. 3, lane F) and will be described elsewhere.

Generation of polyclonal sera to tribrid fusion protein. The gel-purified tribrid protein containing predicted pre-*s* amino acids was used as antigen for the immunization of rabbits. Polyclonal antisera were desirable for three reasons. First, we thought that polyclonal antibodies would more likely recognize the native protein. Second, due to the numerous amino acid differences in the pre-*s* region of HBV of various subtypes, a single polyclonal antiserum was thought to more likely recognize pre-*s* determinants from various HBV subtypes. Third, since our studies involve the investigation of the structure and function of pre-*s* in HBV and WHV, we decided to generate an immunological reagent that was more

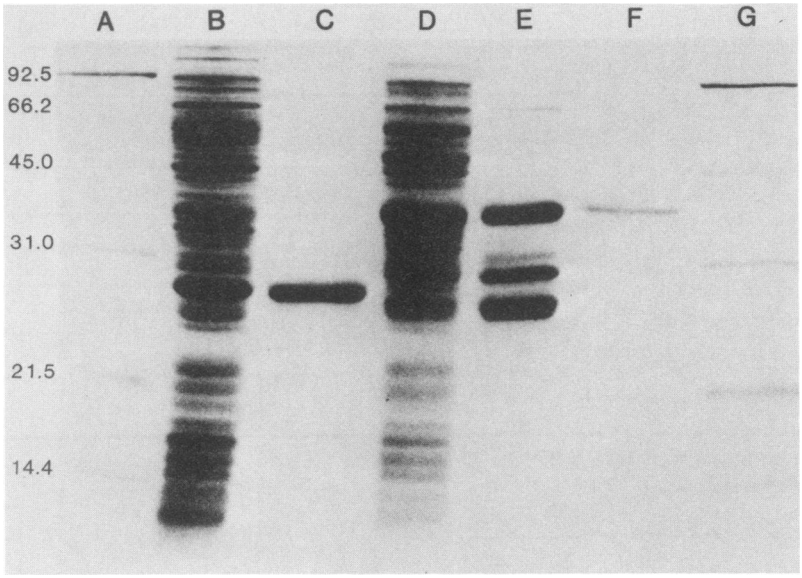

FIG. 3. Analysis of *E. coli* synthesized fusion proteins. Protein extracts and preparations were separated by 12.5% SDS-PAGE. Protein bands were visualized with Coomassie blue stain. Lanes A and G, Molecular weight standards and the sizes of proteins in kd are listed at the left. Lanes B and D, Total cell lysates of the bacterial strain MC1000 transformed with pZL811 (the parent plasmid) and pDW111 (the recombinant plasmid), respectively. Lanes C and E, Chloramphenicol affinity-purified fusion proteins encoded by pZL811 (β-gal:CAT) and pDW111 (β-gal:pre-*s*:CAT), respectively. Lane 6, *E. coli* synthesized tribrid fusion protein (β-gal:pre-*s*:CAT) purified by preparative gel electrophoresis (see the text for details).

likely to recognize determinants potentially encoded by WHV pre-*s* as well. For example, 17 of 26 amino acids show identity when comparing amino acid positions 92 to 117 of HBV (adw_2) pre-*s* with positions 116 to 141 of WHV pre-*s* (see Fig. 4 of Schaeffer and Sninsky [32]). The region of the polypeptide mentioned above is encoded by the DNA sequence that immediately precedes the internal conserved ATG noted in Fig. 2.

A procedure with complementary ORF vectors was used to verify that the antibodies generated indeed recognized the HBV-encoded portion of the tribrid protein (D. T. Wong and J. J. Sninsky, manuscript in preparation). Titration studies with the antisera generated to the β-gal:pre-*s*:CAT tribrid protein indicated an endpoint dilution of greater than 10^{-5} in enzyme-linked immunosorbent assays (see above).

Detection and characterization of polypeptides with pre-*s* determinants in infectious sera. The immune serum generated to the tribrid protein containing pre-*s* determinants was used in Western blot analyses to probe partially purified virus particles isolated from infectious sera. This antiserum was capable of detecting two polypeptides at 42 and 46 kd (Fig. 4, lane C). The high-molecular-weight band(s) seen in Fig. 4, lane C, probably represent incompletely dissociated polypeptides which have been observed previously (39). To verify the specificity of the antiserum recognition, immune sera was incubated with an excess of tribrid protein (β-gal:pre-*s*:CAT) (Fig. 4, lanes D through F) or hybrid protein (β-gal:CAT) (Fig. 4, lane G) before and during probing of nitrocellulose strips containing fractionated viral polypeptides. Although the incubation of antiserum with an excess of tribrid protein significantly decreased the detection of the 42- and 46-kd polypeptides, the hybrid protein had no effect. Competition Western blots resulted in reproducibly higher backgrounds when the β-gal:pre-*s*:CAT fusion protein was used instead of the β-gal:CAT fusion protein (Fig 4, lanes E and F). The reason for this background is unclear. In addition, preimmune serum and hyperimmune serum generated to the hybrid protein (Fig. 4, lanes A and B, respectively) did not recognize these polypeptides. Furthermore, immune serum to the tribrid protein was not capable of identifying polypeptides from normal sera processed in an manner identical to that of infectious sera (data not shown).

Further characterization of the two polypeptides containing pre-*s* determinants employed antisera to HBsAg and ConA-HRP. Figure 5, lane A, is identical to Fig. 4, lane C, except that HRP-conjugated goat anti-rabbit IgG was used instead of ^{125}I-labeled protein A to detect rabbit IgG. As expected, the two bands at 42 and 46 kd were detected. The band at ca. 53 kd in lanes A and B represents contaminating human immunoglobulin heavy chain that is recognized by cross-species reactivity with goat anti-rabbit IgG; the same band in lane D represents contaminating human immunoglobulin heavy chain recognized by ConA-HRP. Antiserum to a synthetic peptide that recognizes the denatured form of HBsAg (lane B) identified 25- and 27-kd polypeptides (identified previously as pI and pII) (27), as well as polypeptides at 33, 42, and 46 kd. Antisera to native HBsAg does not efficiently recognize denatured HBsAg in Western blots (44). These data suggest that the 42- and 46-kd polypeptides contain determinants of both pre-*s* and HBsAg. A 33-kd polypeptide corresponding to HBsAg with an additional 55 amino acids at the N terminus has been previously characterized (23, 40) (see Fig. 2). The tribrid protein used to generate antisera contains only 14 amino acids encoded by the pre-s_{55} region, and therefore the 33-kd polypeptide was not recognized efficiently. Virus preparations from other patient sera contain significantly more 33-kd protein and can be identified with our immune sera (data not shown). Figure 5, lane C, represents partially purified 25- and 27-kd HBsAg probed with an antiserum that recognizes denatured HBsAg, as shown in lane C. Figure 5, lane D, shows the same virus preparation used in lanes A and B but probed with ConA-HRP to identify glycoproteins containing α-D-mannopyranosyl and α-D-glucopyranosyl residues (13).

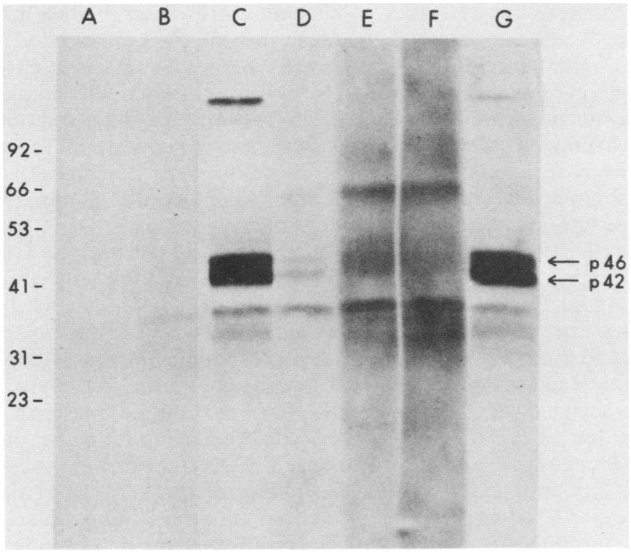

FIG. 4. Viral polypeptide immunotransfer. In each lane, 5 μl of 650-fold-concentrated virus particles was separated by 10% SDS-PAGE and transferred to nitrocellulose. Each strip was then probed with rabbit antisera and labeled with ^{125}I-labeled protein A. Lane A, Preimmune serum; lane B, anti-hybrid (β-gal:CAT) fusion protein; lanes C through G, anti-tribrid (β-gal:pre-*s*:CAT) fusion protein. The antiserum was preincubated with affinity-purified tribrid and hybrid fusion polypeptides at the designated concentrations. Lane D, 5 μg of tribrid protein per ml; lane E, 50 μg of tribrid protein per ml; lane F, 500 μg of tribrid protein per ml; lane G, 500 μg of hybrid protein per ml. The positions of p42 and p46 containing pre-*s* determinants are noted with arrows to the right of the figure. The locations of ^{14}C-labeled protein size markers (New England Nuclear Corp.) and their respective molecular masses in kd are designated at the left of the figure.

Polypeptides at 27, 33, 42, and 46 kd were detected, suggesting that both polypeptides identified by our immune serum are glycoproteins. However, due to the partial purity of the virus particle preparation, we cannot eliminate the possibility of comigrating host glycoproteins. The glycosylation site at amino acid 146 of HBsAg has been shown to contain covalently attached carbohydrate (26), and two potential glycosylation sites (Asn-X-Ser [Thr]) occur in the predicted amino acid sequence of pre-*s* (see Fig. 2). We suggest that the 42- and 46-kd polypeptides represent differential glycosylation of a unique polypeptide, but until amino acid sequence information is available for the N and C termini of these molecules, this remains speculation.

DISCUSSION

A prerequisite to understanding how the HBV genome directs the synthesis of polypeptides is the delineation of those regions of the genome that encode the various structural components of the virion (37). Proteolytic processing, RNA splicing, and overlapping ORFs are strategies of viruses that prevent the simple correlation of a protein and the region(s) of the genome that encodes it. The small quantities of some polypeptides in virions and the frequent contamination by cellular or serum proteins necessitates the use of antibodies to defined determinants to characterize virion polypeptides. To this end, we used a novel ORF expression vector to facilitate the synthesis and purification from *Escherichia coli* of large quantities of a tribrid protein containing a defined region of the pre-s_{174}:HBsAg ORF. Polyclonal antiserum generated to this fusion protein was capable of detecting two polypeptides of ca. 42 and 46 kd in partially purified virus particle preparations. Furthermore, these polypeptides contain HBsAg determinants. These data strongly support and extend the observations of others that HBV encodes multiple polypeptides of higher molecular weight than HBsAg which contain HBsAg sequences. Specifically, our studies unambiguously identified that sequences upstream and in-frame with the HBsAg gene are expressed in polypeptides with a molecular weight consistent with that predicted from the total ORF.

Sanchez et al. (31) were able to show that the minor proteins p27, p31, p35, and p40 of virus particles, subsequent to purification by preparative SDS-PAGE, generated antibodies capable of recognizing the two major proteins pI and pII. Feitelson et al. (8) presented evidence that the minor viral proteins p43, p35, and p32, upon tryptic proteolysis, generated both unique peptides and shared peptides with pI and pII. The first studies capable of unambiguously identifying the region of HBV and the translational reading frame used to synthesize such higher-molecular-weight forms of HBsAg were carried out by Neurath et al. (23). These investigators generated antibodies to a synthetic peptide predicted from pre-s_{55} that recognized gp33 and gp36. Machida et al. (19, 20) not only identified p31 and p35 polypeptides of virus particles but also correlated these polypeptides with the previously observed receptor for polymerized human serum albumin. Studies concurrent with ours have demonstrated the presence of gp33 and gp36 (40), as well as p39 and gp42 (14), polypeptides that share proteolytic fragments with similarly treated p24 and gp27. The polypeptides identified in this report as gp42 and gp46 appear to correspond to the polypeptides of 39 and 42 kd described by Heerman et al. (14). Although Heermann et al. (14) presented evidence for glycosylation of only the higher-molecular-weight polypeptide, our studies suggest that both

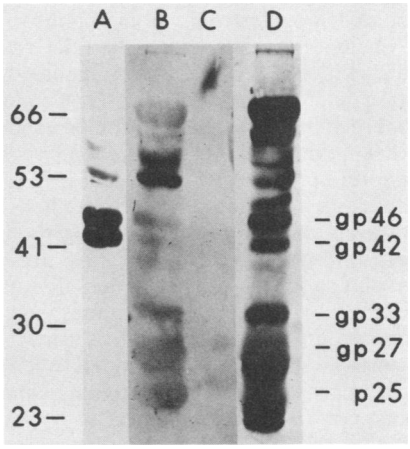

FIG. 5. Viral polypeptide transfer. Concentrated virus preparations (2.5 μl; lanes A, B, and D) and HBsAg (0.1 μg, kindly given by A. R. Neurath; lane C) were separated by 10% SDS-PAGE and transferred to nitrocellulose. Nitrocellulose strips corresponding to these lanes were probed with antisera and labeled with HRP-conjugated goat anti-rabbit IgG. Lane A, Rabbit anti-tribrid protein; lanes B and C, rabbit anti-HBsAg synthetic peptides 138 to 149. Lane D, Glycoproteins were identified with ConA-HRP. These HRP-labeled blots were visualized with 4-chloro-1-napthol and H_2O_2. The pre-s_{174}:HBsAg (gp46 and gp42), pre-s_{55}:HBsAg (gp33), and HBsAg (gp27 and p25) products of ORF S are designated. The positions of ^{14}C-labeled protein size markers and their respective molecular masses in kd are designated at the left of the figure.

polypeptides are glycosylated. The reason for this difference is unclear. Gerlich and coworkers further concluded that the additional amino acid sequences present on these molecules represent N-terminal extensions of HBsAg. A monoclonal antibody to virus particles that specifically binds the p39 and gp42 polypeptides was also described (14). The utility of our tribrid protein is further exemplified by information resulting from an exchange of antibody reagents with these investigators (data not shown). The p39- and gp42-specific monoclonal antibody, MA 18/7, and the polyclonal antiserum described in this report both bound the same two polypeptides in the two virus particle preparations. In addition, this monoclonal antibody was capable of detecting the tribrid fusion protein. This unambiguously identifies the epitope recognized by the previously described monoclonal antibody as being between amino acids 27 and 119 of pre-*s*, as predicted from the ORF of HBV (adw_2). Furthermore, we recently generated monoclonal antibodies to the tribrid protein and demonstrated the specific recognition of the 42- and 46-kd polypeptides in Western blots (unpublished data).

Studies of RNA transcripts of the viral genome are in agreement with information accumulating on the polypeptides of the virion. Transcripts that contain the entire ORF have been identified in monkey cells infected with simian virus 40-HBV recombinants (18), transcription experiments in vitro (29), and mouse cells transfected with head-to-tail HBV dimers (28). The initiation sites have been located near the putative TATA box promoter just upstream of the pre-*s* region. However, more recent transfection studies of mouse cells with the long terminal repeats of Rous sarcoma virus and subgenomic fragments of HBV DNA were unable to detect stable RNA species initiating upstream of the entire ORF (38). Since hepatitis viral DNA is only expressed in human hepatocytes, those transcription studies employing transfected cell lines will require confirmation in infected liver cells in vivo. Several laboratories have now shown both in vivo and in tissue culture that an efficient non-TATA box promoter partially homologous to the simian virus 40 late promoter occurs within the ORF (ca. 200 base pairs upstream of the ATG of the HBsAg gene). An interesting observation made by Laub et al. (18) concerning the transcriptional regulation of the two promoters identified for this ORF is that simian virus 40-HBV recombinants containing both promoters produce 40- to 50-fold less HBsAg than those recombinants that contain subgenomic HBV DNA fragments carrying only the non-TATA box promoter.

Several arguments suggest that the pre-*s*-encoded polypeptides play a crucial role in the life cycle of the hepadna viruses. First, all hepadna genomes isolated to date (HBV, WHV, ground squirrel hepatitis virus, and duck hepatitis B virus) contain this extended ORF, with the surface antigen representing the C-terminal sequence of the predicted polypeptide. Second, whereas the amino acid sequence of the coat protein of the various viruses is strongly conserved, only the overall hydrophilicity and propensity to form specific secondary structure is conserved for pre-*s*, even though the amino acid sequence has diverged (32). Third, Machida et al. (19, 20) not only identified p31 and p35 polypeptides of virus particles but also correlated these polypeptides with the previously observed species-specific binding site for polymerized serum albumin. Alberti et al. (1) identified antibodies in sera from patients which recognize antigenic sites on virus particles containing polymerase activity and inhibit the ability for these virus particles to bind to polymerized human serum albumin. Whether the antibody described by Alberti et al. (1) with anti-polymerized albumin receptor characteristics is the same as those antibodies that recognize the pre-s_{55} region described by Neurath et al. (23) is not known. It is also not known whether the gp42 and gp46 polypeptides reported here bind to polymerized human serum albumin. Even though Machida et al. (19) showed that a cyanogen bromide-generated polypeptide corresponding to the N terminus of p31 binds to polymerized human serum albumin, one cannot exclude the possibility that regions of the gp42 and gp46 polypeptides not represented in p31 or p35 participate in a modulation of this binding. Imai et al. (16) proposed that the binding of HBV particles to polymerized human serum albumin may mediate virus access to hepatocytes, since liver cells also appear to have the capacity to bind the polymerized form of albumin. Further studies are required to elucidate the exact function of the higher-molecular-weight polypeptides in the maturation and propagation of HBV. Towards this goal, we are continuing the biochemical and functional characterization of these polypeptides, determining the nature of the immunological response of infected hosts, and conducting experiments to determine whether antibodies raised to pre-*s* determinants can neutralize HBV.

E. Schaeffer and J. J. Sninsky (manuscript in preparation), using the same technique, identified similar pre-*s*-encoded polypeptides in woodchuck hepatitis virions, providing further support for the existence of these molecules in the hepadna viruses.

ACKNOWLEDGMENTS

We kindly thank Z. Lowey for the ORF expression vector, pZL811; A. R. Neurath for immunological reagents; E. Schaeffer for critical discussions and review of manuscript; and Kathy Levenson (Cetus Corp.) for excellent manuscript preparation.

D.W. and the original research described in this study were supported by Public Health Service grant AI 08295 from the National Institutes of Health. J.J.S. acknowledges an award by the American Cancer Society (JFRA-42).

LITERATURE CITED

1. **Alberti, A., P. Pontisso, E. Schiavon, and G. Realdi.** 1984. An antibody which precipitates dane particles in acute hepatitis type B: relation of receptor sites which bind polymerized human serum albumin on virus particles. Hepatology **4**:220–226.
2. **Amann, E., J. Brosius, and M. Ptashne.** 1983. Vectors bearing a hybrid *trp-lac* promoter useful for regulated expression of cloned genes in *Escherichia coli*. Gene **25**:167–178.
3. **Casadaban, M. J., and S. N. Cohen.** 1980. Analysis of gene control signals by DNA fusion and cloning in *Escherichia coli*. J. Mol. Biol. **138**:179–207.
4. **Cattaneo, R., H. Will, N. Hernandez, and H. Schaller.** 1983. Signals regulating hepatitis B surface antigen transcription. Nature (London) **305**:336–338.
5. **Cohen, S. N., A. C. Y. Chang, and L. Hsu.** 1972. Nonchromosomal antibiotic resistance in bacteria: genetic transformation of *E. coli* by R factor DNA. Proc. Natl. Acad. Sci. U.S.A. **69**:2110–2114.
6. **DeBoer, H. A., L. J. Comstock, and M. Vasser.** 1983. The *tac* promoter: a functional hybrid derived from the *trp* and *lac* promoters. Proc. Natl. Acad. Sci. U.S.A. **80**:21–25.
7. **DeBoer, H. A., L. J. Comstock, D. G. Yansura, and H. L. Heynecker.** 1982. Construction of a tandem *trp-lac* promoter and a hybrid *trp-lac* promoter for efficient and controlled expression of the human growth hormone gene in *Escherichia coli*, p. 462–481. *In* R. L. Rodriguez and M. J. Chamberlin (ed.), Promoters, structure and function. Praeger, New York.
8. **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.
9. **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.
10. **Galibert, F., E. Mandart, F. Fitoussi, P. Tiollais, and P. Charnay.** 1979. Nucleotide sequence of the hepatitis B virus genome (subtype ayw) cloned in *E. coli*. Nature (London) **281**:646–650.
11. **Gerlich, W. H., and W. S. Robinson.** 1980. Hepatitis B virus contains protein attached to the 5′ terminus of its complete DNA strand. Cell **21**:801–809.
12. **Goldfarb, D. S., R. L. Rodriguez, and R. H. Doi.** 1982. Translational block to expression of the *Escherichia coli TN*9-derived chloramphenicol-resistance gene in *Bacillus subtilis*. Proc. Natl. Acad. Sci. U.S.A. **79**:5886–5890.
13. **Hawkes, R.** 1982. Identification of concanavalin A-binding proteins after sodium dodecyl sulfate-gel electrophoresis and protein blotting. Anal. Biochem. **123**:143–146.
14. **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.
15. **Hopp, T. P., and K. R. Woods.** 1981. Prediction of protein antigenic determinants from amino acid sequences. Proc. Natl. Acad. Sci. U.S.A. **78**;3824–3828.
16. **Imai, M., Y. Yanase, T. Nojiri, Y. Miyakawa, and M. Mayumi.** 1979. A receptor for polymerized human and chimpanzee albumins on hepatitis B virus particles co-occurring with HBeAg. Gastroenterology **76**:242–247.
17. **Laemmli, U. K.** 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature (London) **227**:680–685.
18. **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.
19. **Machida, A., S. Kishimoto, H. Ohnuma, K. Baba, Y. Ito, H. Miyamoto, G. Funatsu, K. Oda, S. Usuda, S. Togami, T. Nakamura, Y. Miyakawa, and M. Mayumi.** 1984. A polypeptide containing 55 amino acid residues coded by the pre-*s* region of hepatitis B virus deoxyribonucleic acid bears the receptor for polymerized human as well as chimpanzee albumins. Gastroenterology **86**:910–918.
20. **Machida, A., S. Kishimoto, H. Ohnuma, H. Miyamoto, K. Baba, K. Oda, T. Nakamura, Y. Miyakawa, and M. Mayumi.** 1983. A hepatitis B surface antigen polypeptide (p31) with the receptor for polymerized human as well as chimpanzee albumins. Gastroenterology **85**:268–274.
21. **Mandart, E., A. Kay, and F. Galibert.** 1984. Nucleotide sequence of a cloned duck hepatitis B virus genome: comparison with woodchuck and human hepatitis B virus sequences. J. Virol. **49**:782–792.
22. **Miller, J. H.** 1972. Experiments in molecular genetics. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
23. **Neurath, A. R., S. B. Kent, and N. Strick.** 1984. Location and chemical synthesis of a pre-*s* gene coded immunodominant epitope of hepatitis B virus. Science **224**:392–395.
24. **Ono, Y., H. Onda, R. Sasada, K. Igarashi, Y. Sugino, and K. Nishioka.** 1983. The complete nucleotide sequence of the cloned hepatitis B virus DNA; subtypes adr and adw. Nucleic Acids Res. **11**:1747–1757.
25. **Pasek, M., T. Goto, W. Gilbert, B. Zink, H. Schaller, P. Mackay, G. Leadbetter, and G. Murray.** 1979. Hepatitis B virus genes and their expression in *E. coli*. Nature (London) **282**:575–579.
26. **Peterson, D. L.** 1981. Isolation and characterization of the major protein and glycoprotein of hepatitis B surface antigen. J. Biol. Chem. **256**:6975–6983.
27. **Peterson, D. L., I. M. Roberts, and G. N. Vyas.** 1977. Partial amino acid sequence of two major component polypeptides of hepatitis B surface antigen. Proc. Natl. Acad. Sci. U.S.A. **74**:1530–1534.
28. **Pourcel, C., A. Louise, M. Gervais, N. Chenciner, M.-F. DuBois, and P. Tiollais.** 1982. Transcription of the hepatitis B surface antigen gene in mouse cells transformed with cloned viral DNA. J. Virol. **42**:100–105.
29. **Rall, L. B., D. N. Standring, O. Laub, and W. J. Rutter.** 1983. Transcription of hepatitis B virus by RNA polymerase II. Mol. Cell. Biol. **3**:1766–1773.
30. **Russell, D. R., and G. N. Bennett.** 1982. Construction and analysis of in vivo activity of *E. coli* promoter hybrids and promoter mutants that alter the −35 to −10 spacing. Gene **20**:231–243.
31. **Sanchez, Y., I. Ionescu-Matiu, G. R. Dreesman, F. B. Hollinger, and J. L. Melnick.** 1981. Evidence for the presence of repeating antigenic determinants in the major and minor polypeptides derived from hepatitis B surface antigen. Virology **114**:71–80.
32. **Schaeffer, E., and J. J. Sninsky.** 1984. Predicted secondary structure similarity in the absence of primary amino acid sequence homology: hepatitis B virus open reading frames. Proc. Natl. Acad. Sci. U.S.A. **81**:2902–2906.
33. **Schottel, J. L., J. J. Sninsky, and S. N. Cohen.** 1984. Effects of alterations in the translation control region on bacterial gene expression: use of *cat* gene constructs transcribed from the *lac* promoter as a model system. Gene **28**:177–193.
34. **Seeger, C., D. Ganem, and H. E. Varmus.** 1984. Nucleotide sequence of an infectious molecularly cloned genome of ground squirrel hepatitis virus. J. Virol. **51**:367–375.
35. **Shaw, W. V., and R. F. Brodsky.** 1968. Characterization of chloramphenicol acetyltransferase from chloramphenicol-resistant *Staphylococcus aureus*. J. Bacteriol. **95**:28–36.
36. **Sninsky, J. J., A. Siddiqui, W. S. Robinson, and S. N. Cohen.** 1979. Cloning and endonuclease mapping of the hepatitis B viral genome. Nature (London) **279**:346–348.
37. **Sninsky, J. J., C. Zraly, M. L. Meyers, Z. G. Loewy, D. T. Wong, and E. Schaeffer.** 1984. Approaches in the study of hepatitis B virus open-reading frames, p. 21–29. *In* F. V. Chisari (ed.), Advances in hepatitis research. Masson Publishing USA, Inc., New York.
38. **Standring, D. N., W. J. Rutter, H. E. Varmus, and D. Ganem.** 1984. Transcription of the hepatitis B surface antigen gene in cultured murine cells initiates within the presurface region. J. Virol. **50**:563–571.
39. **Stibbe, W., and W. H. Gerlich.** 1982. Variable protein composition of hepatitis B surface antigen from different donors. Virology **123**:436–442.
40. **Stibbe, W., and W. H. Gerlich.** 1983. Structural relationships between minor and major proteins of hepatitis B surface antigen. J. Virol. **46**:626–628.
41. **Towbin, H., T. Staehelin, and J. Gordon.** 1979. Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: procedure and some applications. Proc. Natl. Acad. Sci. U.S.A. **76**:4350–4354.
42. **Ullrich, A., J. Shine, J. Chirgwin, R. Pictet, E. Tischer, W. J. Rutter, and H. M. Goodman.** 1977. Rat insulin genes: construction of plasmids containing the coding sequences. Science **196**:1313–1319.
43. **Valenzula, P., M. Quiroga, J. Zaldivar, P. Gray, and W. J. Rutter.** 1980. The nucleotide sequence of the hepatitis B virus genome and the identification of the major viral genes, p. 57–70. *In* B. Fields, R. Jaenisch, and C. F. Fox (ed.), Animal virus genetics. Academic Press, Inc., New York.
44. **Vyas, G. N., K. R. Rao, and A.B. Ibrahim.** 1972. Australia antigen (hepatitis B antigen): a conformational antigen dependent on disulfide bonds. Science **178**:1300–1301.
45. **Zaidenzaig, Y., and W. V. Shaw.** 1976. Affinity and hydrophobic chromatography of three variants of chloramphenicol acetyltransferases specified by R factors in *Escherichia coli*. FEBS Lett. **62**:266–271.