JOURNAL OF VIROLOGY, Jan. 1986, p. 173–182
0022-538X/86/010173-10$02.00/0

Vol. 57, No. 1

Identification and Localization of Pre-*s*-Encoded Polypeptides from Woodchuck and Ground Squirrel Hepatitis Viruses

ERIC SCHAEFFER,[1] ROBERT L. SNYDER,[2,3] AND JOHN J. SNINSKY[1]*†

Department of Microbiology and Immunology and Department of Molecular Biology, Albert Einstein College of Medicine, Bronx, New York 10461,[1] and Penrose Research Laboratory, Zoological Society of Philadelphia[2] and Department of Pathology, School of Medicine, and Department of Pathobiology, School of Veterinary Medicine, University of Pennsylvania,[3] Philadelphia, Pennsylvania 19104

Received 10 July 1985/Accepted 3 October 1985

A segment from the pre-*s* region of the woodchuck hepatitis virus (WHV) was inserted into an open reading frame vector allowing for the expression in *Escherichia coli* of viral determinants as part of a fusion protein. The bacterially synthesized fusion molecule contained eight amino acids from β-galactosidase (β-gal) at the N terminus, followed by 89 pre-*s*-encoded amino acids and 219 amino acids of chloramphenicol acetyltransferase (CAT) at the C terminus (β-gal:pre-*s*:CAT). This tribrid protein was used to generate antiserum which had a significant titer to the viral portion of the fusion polypeptide. Anti-β-gal:pre-*s*:CAT was used in Western blot analysis to identify viral proteins containing pre-*s*-encoded determinants. Antiserum to the tribrid molecule recognized four WHV polypeptides with molecular masses of 33, 36, 45, and 47 kilodaltons, each of which was also recognized by a monoclonal antibody to WHV surface antigen. Using the same anti-tribrid serum, we also identified analogous polypeptides from ground squirrel hepatitis virus. The antiserum was also used to immunoprecipitate virus particles containing endogenous DNA polymerase activity, indicating that pre-*s* determinants are found on the surface of mature virions. Based on previous computer studies and the location of pre-*s*-encoded molecules on the surface of virus particles, a role in hepadnavirus host cell entry is suggested for these polypeptides.

Hepatitis B virus (HBV) infection represents a serious world health problem, being associated with acute and chronic liver disease as well as hepatocellular carcinoma. In recent years, several viruses have been isolated which share morphological and genetic characteristics with HBV. These related viruses, which infect woodchucks (woodchuck hepatitis virus [WHV]) (48), beechey ground squirrels (ground squirrel hepatitis virus [GSHV]) (25), and Pekin ducks (duck hepatitis virus) (26), represent a new class termed the hepadnaviruses. Animals in which chronic liver disease and hepatocellular carcinoma have been found will likely prove to be hosts for additional viruses in the hepadna class. All of these viruses encapsidate a circular, partially single- and double-stranded DNA genome, are strongly hepatotrophic, and contain endogenous DNA polymerase activity (46). Evidence has indicated that hepadnaviruses replicate through an RNA intermediate (29, 47; M. Roggendorf and J. Summers, Abstr. Cold Spring Harbor Meet. Mol. Biol. Hepatitis B Viruses. 1985, p. 5), a unique characteristic for DNA viruses.

Classical approaches to the study of the hepadnaviruses have been hampered by the inability to propagate them in a tissue culture system or convenient laboratory animal. Progress toward the characterization of proteins encoded by the virus and their function in the viral life cycle has been slow. The application of molecular techniques to the study of this novel class of viruses has allowed the accumulation of a significant body of information. The DNA sequences of several cloned HBV genomes have been determined and examined for the presence of open reading frames (ORFs) (13, 33, 35, 52). The short or incomplete strand of the virus was consistently shown to contain four ORFs, each with the capacity to encode at least 150 amino acids, whereas the long strand had only short ORFs which were not conserved between different HBV subtypes. The genetic organization shown initially for HBV has also been found in WHV and GSHV (12, 43) and with a single modification in duck hepatitis virus (23). An interesting feature determined by DNA sequencing is that the ORF which encodes the 225-amino-acid coat protein of HBV has the capacity to encode an additional 175 amino acids (pre-*s*) upstream of, and in frame with, the surface antigen gene. A similar extended ORF (pre-*s*:S) has been found for the other hepadnaviruses. The pre-*s* region has been speculated to play a role in host cell entry (22, 41), and may be involved in the species specificity characteristic of this class of viruses.

WHV appears to be most similar to HBV in terms of the nature of the liver disease caused by the virus and the association of hepatocellular carcinoma with chronic infection (37, 44). Because of the pathogenic similarities between HBV and WHV, the latter was thought to be a useful model for studying a region which has been speculated to encode a protein important for host cell entry. If the expression of pre-*s*-could be established in WHV, the woodchuck would be valuable for studying the biological importance of this region.

The approach being taken in our laboratory involves the use of bacterial expression vectors to synthesize virally encoded determinants as fusion proteins. This report describes the cloning and expression of an internal segment from the pre-*s* region of WHV as a fusion protein in *Escherichia coli*. The bacterially synthesized pre-*s* determinants were used to generate polyclonal antiserum in rabbits, which proved useful in the identification of viral polypeptides containing pre-*s*-encoded amino acids.

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

MATERIALS AND METHODS

Bacterial strains and culture conditions. All plasmids were propagated in either *E. coli* MC1000 (Δ*lac* IPOZYx74) (3), MM294 (*thi-1 endA1 hsdR17 supE44* λ−) (27), or DG101 (MM294 derivative with *lacI*^{q} *lacZ* ΔM15). Bacteria were stored as stationary-phase frozen cultures, and for growth were diluted 1:100 in Luria broth (28) and incubated at 37°C. Bacteria harboring plasmids conferring ampicillin resistance were grown in Luria broth with 50 μg of ampicillin per ml. Density of bacterial cultures was measured by A_{650}. Bacteria were induced with 10 mM isopropyl β-D-thiogalactopyranoside (2) at an A_{650} of 0.6 to 0.8, and growth was continued for 1 to 2 h. Bacteria were plated on antibiotic medium no. 2 (Difco Laboratories) with ampicillin selection at 50 μg/ml and chloramphenicol selection at 25 or 50 μg/ml.

DNA manipulations and bacterial transformations. Plasmid DNA was purified by CsCl-EtBr density centrifugation (49) or by a minilysate procedure (16) and stored in 10 mM Tris hydrochloride (pH 7.4)–1 mM EDTA. Restriction endonuclease digestions were carried out with DNA concentrations of 0.5 to 1 mg/ml under conditions recommended by the vendor. Restriction enzymes and T4 DNA ligase were purchased from New England BioLabs, Inc. Ligations were performed in ligase buffer (50 mM Tris hydrochloride [pH 7.4], 10 mM $MgCl_2$, 10 mM dithiothreitol, 1 mM spermidine, 1 mM ATP, 0.1 mg of bovine serum albumin per ml) with 5 to 10 μg of linearized vector treated with calf intestine alkaline phosphatase (Boehringer Mannheim Biochemicals) with a 5- to 10-fold molar excess of the fragment to be inserted in a total volume of 30 μl. DNA fragments were purified from polyacrylamide gels by electroelution (42). Ligation reactions were used to transform bacteria (6), and plasmid preparations from candidate bacterial colonies were analyzed with 0.7% horizontal mini-agarose gels in Tris acetate buffer and 5% polyacrylamide gels in Tris borate buffer (24).

Purification of fusion proteins. Bacteria (strain DG101) harboring pWpresCAT-1, the recombinant plasmid that expresses pre-*s* determinants (see Results), were grown in 1.5 liters of Luria broth to a density of 0.6 to 0.8 A_{650}, at which time cells were induced by addition of isopropyl β-D-thiogalactopyranoside. Cultures were incubated for 1 h after induction and then centrifuged in 500-ml bottles at 7,000 rpm for 10 min in a JA-10 rotor (Beckman Instruments, Inc.). Bacterial pellets were suspended in 10 ml of cold TEP buffer (100 mM Tris hydrochloride [pH 7.4], 10 mM EDTA, 1 mM phenylmethylsulfonyl fluoride [diluted from a 100 mM solution dissolved in 100% isopropanol]) to give a density equivalent to 225 A_{650}. All subsequent steps were performed at 4°C essentially as previously described (53). Briefly, suspended bacteria were lysed by sonication and centrifuged to remove particulate material. The supernatant was brought to a final concentration of 75% saturated ammonium sulfate, and the resulting precipitate was collected by centrifugation. The pellet was suspended in and dialyzed against TEP buffer and loaded onto a column containing chloramphenicol covalently linked to a solid support (54). Nonspecifically bound material was removed by washing the column with TEP containing 0.2 M NaCl, followed by specific elution of chloramphenicol acetyltransferase (CAT)-containing fusion proteins with TEP containing 0.2 M NaCl and 2 mg of chloramphenicol per ml. Column fractions were monitored by A_{280} and analyzed by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) (18). The chloramphenicol-eluted proteins were pooled and used in subsequent experiments.

Generation of polyclonal antisera. Female New Zealand White rabbits were immunized with either the hybrid β-galactosidase (β-gal):CAT protein synthesized by pZL811 or the β-gal:pre-*s*:CAT tribrid protein. One hundred micrograms of the pooled protein eluted from the chloramphenicol column was mixed with an equal volume of Freund complete adjuvant (Difco), and an emulsion was formed. A total volume of 1 to 2 ml was injected intradermally at several sites along the back. After 2 weeks, rabbits were bled and boosted with 100 μg of protein with incomplete Freund adjuvant. Antisera from the primary bleed had significant titers of antibodies to the antigen used for immunization. Two weeks after the secondary injection, antisera to hybrid and tribrid proteins had endpoint dilutions of between 10^{-5} and 10^{-6} in enzyme-linked immunosorbent assays, which were performed as described previously (53).

Virus preparation. Sera were obtained from woodchucks housed at the Philadelphia zoo. Sera were identified as positive for WHV surface antigen (WHsAg) by a commercially available kit for the detection of human hepatitis B surface antigen (Auszyme II; Abbott Laboratories) as shown previously (30) and for DNA polymerase activity as described below. Sera were centrifuged at 9,000 rpm for 10 min to remove particulate material. Two milliliters of cleared sera was layered onto a 10-ml 15 to 30% sucrose step gradient in 10 mM Tris hydrochloride (pH 7.4)–100 mM NaCl–5 mM EDTA and centrifuged at 26,000 rpm for 16 h in an SW40 rotor (Beckman). Pellets were suspended in either a 1/10 or a 1/20 volume of 10 mM Tris hydrochloride (pH 7.4)–50 mM NaCl for use in polymerase assays or a 1/80 volume of 10 mM Tris hydrochloride (pH 7.4)–50 mM NaCl–5 mM EDTA for fractionating the virus by SDS-PAGE. Ground squirrel hepatitis virus was prepared in an identical manner with pooled infectious ground squirrel sera (gift from P. L. Marion and W. S. Robinson).

Polymerase assay. Twentyfold-concentrated and partially purified virus preparations were treated with 1.5% Nonidet P-40 (NP-40)–0.1% 2-β-mercaptoethanol at 37°C for 10 min. The detergent-treated virus preparations were incubated in 66 mM Tris hydrochloride (pH 7.4)–80 mM NH_4Cl–27 mM $MgCl_2$–0.33 mM dGTP, dCTP, and dTTP with 0.2 μM [α-^{32}P]dATP (3,000 Ci/mmol; Amersham Corp.) in a total volume of 30 μl for 2 h at 37°C. To ensure completion of the synthesized strand, dATP was added to 0.33 mM, and the reaction was continued for an additional hour. After the 3-h incubation period, acid-insoluble radioactivity was quantitated by spotting the reaction onto glass fiber filters (GF/C; Whatman, Inc.) which were soaked in 5% trichloracetic acid (TCA) containing 0.1 M sodium pyrophosphate. Filters were washed batchwise with three changes of the TCA-pyrophosphate solution, then with 5% TCA alone, and finally twice with 100% ethanol, as previously described (40). The filters were dried under a heat lamp, immersed in 5 ml of scintillant, and counted. Alternately, labeled products were fractionated by agarose gel electrophoresis and visualized by autoradiography. By this procedure, a molecule migrating as an approximately 3,300-base-pair nicked circular DNA was observed (data not shown).

Western blotting. Twenty-five microliters of partially purified, 80-fold-concentrated WHV preparation was solubilized by boiling in sample buffer (2.5% SDS, 0.5 M dithiothreitol, 5% glycerol, 62 mM Tris hydrochloride [pH 6.8], 0.0025% bromophenol blue) for 10 min and fractionated by 10% SDS-PAGE. Protein transfer from acrylamide gels to nitrocellulose was performed by a modification of the procedure used by Towbin et al. (50). Transfer was carried out

in a Trans-Blot cell (Bio-Rad Laboratories) with 0.025 M Tris base–0.192 M glycine–20% methanol buffer for 5 h at a constant current of 250 mA. A strip of nitrocellulose containing low-molecular-weight protein markers (Bio-Rad) was stained with amido black (50) to assess transfer efficiency. Unoccupied sites on the nitrocellulose were blocked with 250 ml of modified BLOTTO (17) (1 M glycine, 1% ovalbumin, 5% dry milk [Carnation], 5% fetal calf serum) for 30 min at room temperature. The blot was washed for three 10-min periods with 0.1% dry milk–0.1% bovine serum albumin–1% fetal calf serum–0.1% Tween 20 in phosphate-buffered saline (10 mM sodium phosphate [pH 7.2], 150 mM NaCl) (wash solution). All subsequent antibody incubations and washing steps were carried out with wash solution. The blots were incubated at room temperature in sealed plastic pouches with a 1:100 dilution of rabbit antisera or a 1:50 dilution of mouse monoclonal antibody for 4 h with agitation. For preadsorption experiments, neat antisera were incubated with 10 μg of hybrid or tribrid protein per ml for 30 min at room temperature, at which time sera were centrifuged at 12,000 × *g* for 4 min. Supernatants were then diluted 1:100 in wash solution, and hybrid or tribrid protein was added to a final concentration of 10 μg/ml. Diluted antisera and competing antigen were then incubated with protein blots. Nitrocellulose filters were washed as described above and incubated with either a 1:200 dilution of ^{125}I-labeled protein A (10 to 15 μg/ml, 2 to 10 μCi/μg; New England Nuclear Corp.) or a 1:500 dilution of rabbit anti-mouse immunoglobulin G (IgG)-IgA-IgM conjugated to horseradish peroxidase (Zymed). In both cases, incubations were for 30 min to 1 h at room temperature with agitation. Filters were washed as described above, and ^{125}I-labeled protein A-labeled blots were dried and exposed to XAR-5 film (Eastman Kodak Co.) with an intensifying screen at −70°C. Filters labeled with rabbit anti-mouse immunoglobulins conjugated to horseradish peroxidase were developed in a substrate solution (0.5 mg of diaminobenzidine [Sigma] and 0.03% H_2O_2 in 0.1 M Tris hydrochloride [pH 7.5]). Protein bands were visualized in approximately 10 min, and filters were washed extensively with distilled H_2O.

Immunoprecipitation of DNA polymerase activity. Protein A agarose (Zymed) was washed twice by pelleting and suspending in 0.5% Triton X-100–0.5 M NaCl–25 mM Tris hydrochloride (pH 8.0)–1 mg of bovine serum albumin per ml (IP-blocking buffer). Fifty microliters of a 1:1 suspension of protein A agarose in IP-blocking buffer was incubated at room temperature for 30 min with 5 to 10 μl of the appropriate antisera. Antibody to *E. coli*-synthesized HBV core antigen was purchased from Accurate Chemical Corp., and monoclonal antibody to WHsAg (101-2) was a gift from P. J. Cote and J. L. Gerin. Since the murine monoclonal antibody was not efficiently recognized by protein A, agarose protein A was first incubated for 30 min with a rabbit anti-mouse IgG antibody (Miles Scientific), followed by incubation with the monoclonal antibody. All antibody incubations were followed by three washes with 0.5 M NaCl–50 mM Tris hydrochloride (pH 8.0)–5% fetal calf serum (IP wash buffer). The protein A-antibody complex was then incubated with 20 μl of a 10-fold-concentrated and partially purified WHV preparation which had been preincubated for 30 min at 37°C with either 10 mM Tris hydrochloride (pH 7.5) or 10 mM Tris hydrochloride (pH 7.5)–1.5% NP-40–0.1% 2-β-mercaptoethanol. After a 2-h incubation at room temperature with agitation, agarose beads were pelleted and washed six times (6 ml per wash) with IP wash buffer. Pelleted agarose beads were suspended in 20 ml of 1.5% NP-40–0.1% 2-β-mercaptoethanol and incubated at 37°C for 20 min. Buffer and deoxynucleotide triphosphates were added for the endogenous polymerase reaction, and incorporated radioactivity was measured by TCA precipitation (described above).

RESULTS

Construction and characterization of plasmids expressing β-gal:pre-*s*:CAT fusion proteins. The ORF vector used in this study, pZL811, has been described previously (53). Briefly, transcription of fusion protein nucleotide sequences is under the control of a hybrid TAC promoter (8) with the ribosome binding site and the N-terminal codons contributed by the gene that encodes β-galactosidase (*lacZ*) of *E. coli*. The *lacZ* codons are followed by a polylinker devoid of translational stop codons and containing a unique *Bam*HI endonuclease site. The polylinker region is contiguous with the CAT gene from Tn*9* (1). These coding segments allow for the production of an in-frame hybrid fusion protein containing eight amino acids of β-gal at the N terminus, eight amino acids derived from linker codons, and the entire 219 amino acids of CAT at the C terminus (Fig. 1). The production of this fusion protein (β-gal:CAT) in bacteria confers resistance to at least 50 μg of chloramphenicol per ml.

An *Xho*II fragment from the pre-*s* region of WHV was chosen for insertion into the unique *Bam*HI site of pZL811 because of the identical 5′-protruding ends generated by these two endonucleases and the retention of the reading frame between β-gal and CAT. As indicated in Fig. 1, the pre-*s* region of the viral genome also encodes a portion of the largest ORF (ORF P), but in another translational reading frame. The use of an expression vector with defined reading frames at both the 5′ and 3′ junctions was necessary to unambiguously define the resulting tribrid protein as containing pre-*s* rather than ORF P-encoded determinants. The *Xho*II fragment has the potential to encode 88 pre-*s* amino acids, corresponding to amino acids 85 to 173 of a protein initiated from the first ATG of the ORF, allowing for the production of a β-gal:pre-*s*:CAT tribrid fusion protein (Fig. 1). The source of WHV DNA was the recombinant plasmid pBH20-WHV1, which is a viral genome inserted at the unique *Eco*RI site of pBH20 as described by Ogston et al. (32). Because the desired *Xho*II fragment spanned the *Eco*RI site of the WHV genome, construction of a head-to-tail dimer was necessary. Initially, pBH20-WHV1 was cleaved with *Eco*RI to release the 3.3-kilobase linearized viral genome. The viral DNA was then purified by PAGE and electroelution. Plasmid vector pACYC184 (5) was digested with *Eco*RI and then treated with calf intestine alkaline phosphatase to minimize recircularization (51). Purified WHV DNA was ligated to linearized pACYC184, and the ligation mixture was used to transform competent MC1000. The CAT gene carried by pACYC184 is interrupted by cleavage with *Eco*RI, and bacteria harboring recombinants generated at this site are chloramphenicol sensitive. Transformants were selected for resistance to ampicillin and screened on 25 μg of chloramphenicol per ml. Digestion of plasmid DNAs from ampicillin-resistant, chloramphenicol-sensitive candidates with *Pst*I, which cleaves viral DNA at a single site, would generate a monomer-length fragment from recombinants containing a head-to-tail dimer. A representative clone containing a WHV dimer was identified and designated pWHV-2.

Recombinant plasmid pWHV-2 was digested with *Xho*II, generating a 262-base-pair fragment from the pre-*s* region which was isolated by PAGE and electroelution. The ORF vector pZL811 was linearized by digestion with *Bam*HI and

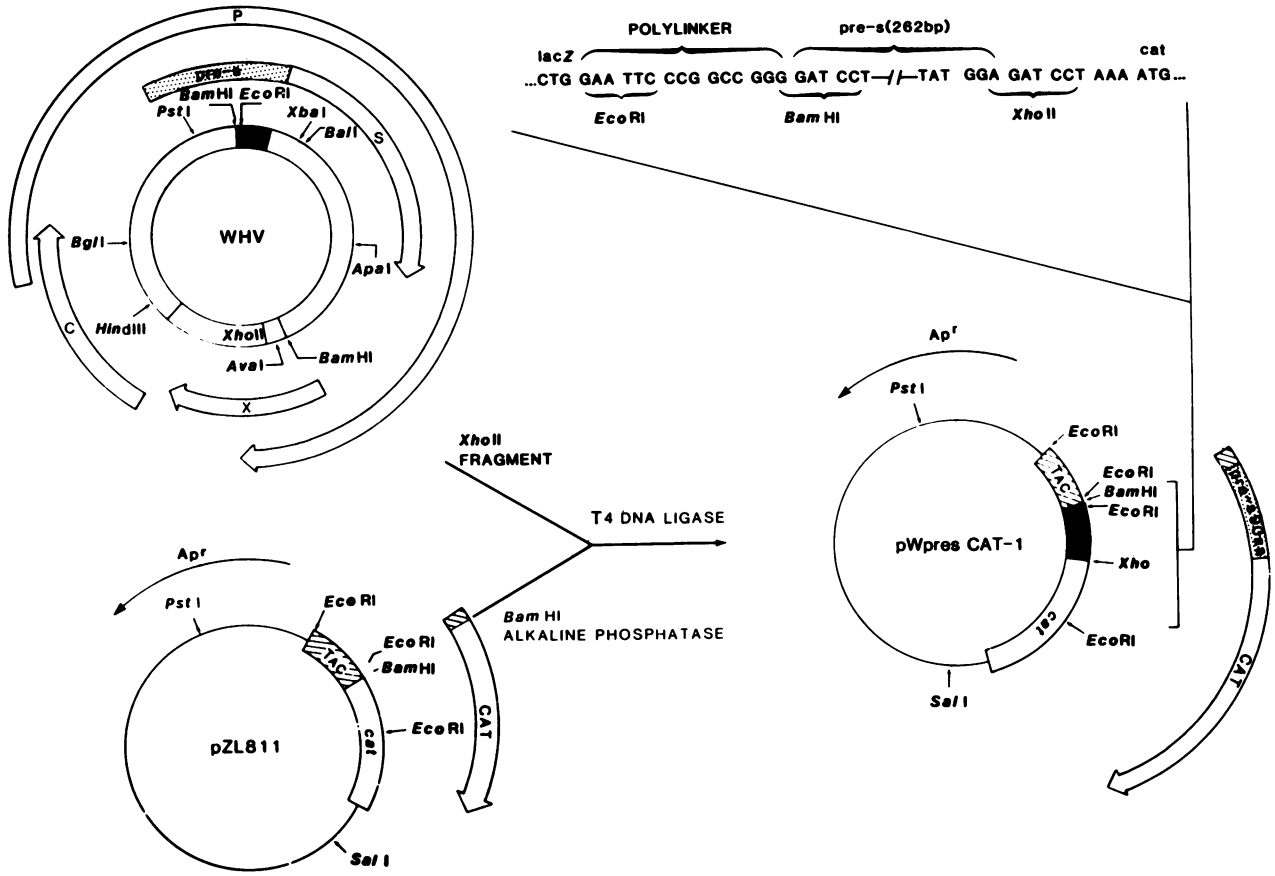

FIG. 1. Construction of tribrid protein expression plasmid. The WHV genomic map is based on the sequence of Galibert et al. (12). Arrows circling the viral map represent ORFs. Unique endonuclease sites are noted for orientation, and *Xho*II sites are indicated within the genomic circular map. Arrows labeled S and C refer to the surface and core antigen genes, respectively. The dotted segment contiguous with the S gene represents the pre-*s* portion of the pre-*s*:S ORF. Arrows labeled X and P represent the smallest and largest ORFs, respectively. The *Xho*II fragment containing pre-*s* sequences (solid block) is schematically indicated on the viral DNA for simplicity but was actually isolated from a dimer (see Results). The broad arrows associated with pZL811 and pWpresCAT-1 represent the hybrid β-gal:CAT and tribrid β-gal:pre-*s*:CAT proteins, respectively. The hatched, dotted, and open portions of the broad arrows represent β-gal, pre-*s*, and CAT amino acids, respectively. The thin arrows labeled Ap^r denoted for pZL811 and pWpresCAT-1 represent β-lactamase. The nucleotide sequence noted at the upper right corresponds to the viral DNA-vector junctions.

treated with calf intestine alkaline phosphatase. The purified WHV pre-*s* fragment was ligated to pZL811 (Fig. 1), and the ligation products were used to transform DG101. Transformants were initially selected on ampicillin-containing media. The enzymatic activity of CAT fusion proteins has been found to be variable depending on the particular fragment inserted between the *lacZ* and CAT coding sequences (11, 53). In some cases, increased sensitivity to chloramphenicol of recombinant plasmids relative to the parent plasmid is observed (53). However, all of the ampicillin-resistant candidates that were screened for chloramphenicol resistance were found to be resistant to at least 50 μg of this antibiotic per ml and thus could not be distinguished from the parental plasmid based on this criterion. It was thus necessary to prepare plasmid DNA from random ampicillin- and chloramphenicol-resistant bacterial colonies with which to carry out endonuclease analysis. Endonuclease *Pst*I cleaves within the β-lactamase gene of the vector, while *Bam*HI cleaves at the upstream WHV DNA-vector junction sequence; together they establish the presence of viral sequences and their orientation (Fig. 1).

Protein synthesized by bacterial candidates harboring the parent vector pZL811 or recombinant plasmids with the WHV insert were analyzed by SDS-PAGE (Fig. 2). Candidates which had the pre-*s* insert in the correct orientation no longer directed the expression of a β-gal:CAT fusion protein (M_r, ca. 28 kilodaltons [kDa]) characteristic of pZL811 (Fig. 2, lane 2). However, following induction of transcription under the TAC promoter with isopropyl β-D-thiogalactopyranoside, a new protein with the M_r expected for a β-gal:pre-*s*:CAT tribrid fusion molecule (ca. 35 kDa) represented one of the predominant proteins synthesized by these bacteria (Fig. 2, lane 4). The plasmid responsible for production of the tribrid protein was designated pWpresCAT-1. To verify that the new protein synthesized in bacteria harboring pWpresCAT-1 was the tribrid molecule expected and as such contained CAT determinants, we carried out Western blot analysis on total cell lysates probing with a rabbit polyclonal antibody to the β-gal:CAT hybrid fusion protein. The rabbit anti-β-gal:CAT sera recognized predominantly two proteins, a polypeptide with an M_r of ca. 35 kDa and an additional band migrating at ca. 26 kDa (see, for example, Fig. 4, lane 3), neither of which was identified by preimmune sera. The same two proteins were also specifically eluted from a chloramphenicol column by the addition of chloramphenicol to the elution buffer (Fig. 2, lane 5), as had been

observed for the hybrid β-gal:CAT polypeptide (Fig. 2, lane 3). The lower-M_r band synthesized by bacteria harboring pWpresCAT-1 is thought to be an N-terminally truncated molecule generated by internal initiation of translation or by proteolysis.

Characterization of rabbit anti-β-gal:pre-*s*:CAT sera. Use of a tribrid protein to immunize rabbits leads to the generation of antibodies which recognize multiple determinants on the molecule. The fraction of antibodies which are specific for any given segment of the tribrid protein depends on the immunogenicity of that portion of the polypeptide. As mentioned above, antisera to β-gal:CAT efficiently recognized the full-length tribrid protein as well as a smaller molecule. A similar analysis with antisera to β-gal:pre-*s*:CAT indicated that only the larger of the tribrid-related species eluted from the chloramphenicol column was recognized (see Fig. 4, lane 5). In addition, in an analogous experiment anti-β-gal:pre-*s*:CAT only weakly recognized the hybrid protein (data not shown). These results suggest that antisera generated to the tribrid molecule efficiently recognize the viral portion of the fusion protein and only weakly if at all recognize CAT determinants. Similar immunogenic characteristics have been seen for other tribrid proteins containing CAT sequences (see Discussion).

Identification of viral proteins with anti-β-gal:pre-*s*:CAT serum. Examination of the DNA sequence of WHV has indicated that there are two ATG initiation codons upstream

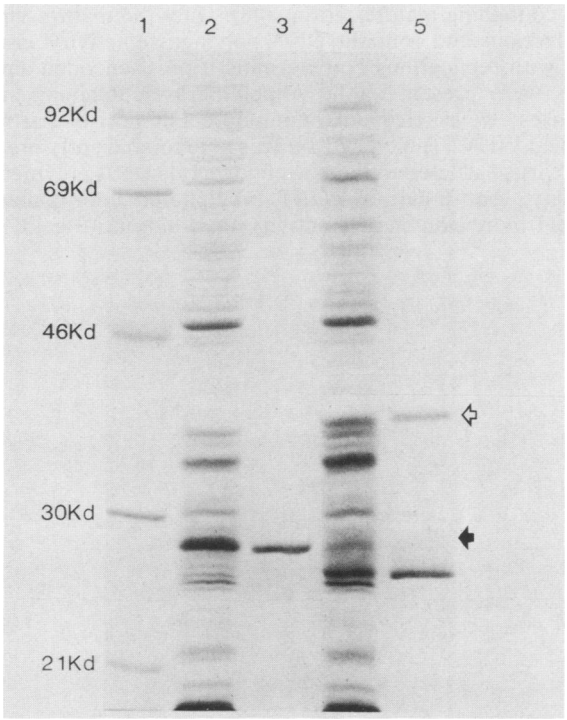

FIG. 2. Expression and purification of fusion proteins. Lane 1 contains protein standards with M_rs indicated in kilodaltons at the left. Lanes 2 and 4 contain total bacterial lysates after induction with isopropyl-β-D-thiogalactopyranoside from strain MM294 harboring pZL811 (lane 2) or strain DG101 harboring pWpresCAT-1 (lane 4). Fusion proteins purified on a chloramphenicol affinity column. Lanes 3 and 5 contain 5 μg each of β-gal:CAT (closed arrow) and β-gal:pre-*s*:CAT (open arrow), respectively. Protein bands were visualized by Coomassie blue staining. See Materials and Methods for details.

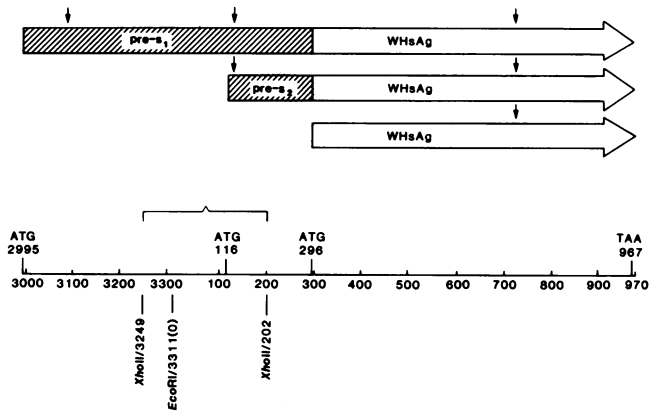

FIG. 3. Polypeptides predicted from the WHV pre-*s*:S ORF. Three horizontal arrows represent primary translation products initiating from each of the three ATGs of the pre-*s*:S ORF. The precise location of each ATG and the termination codon (TAA) is designated above the line at the bottom. Potential translation start codons distal to the putative ATG of the major coat protein are not indicated. Amino acids encoded by the pre-*s* region and surface antigen gene are denoted by hatched and open arrow segments, respectively. The longest arrow contains pre-s_1- and pre-s_2-encoded sequences, whereas the middle arrow designates a molecule with amino acids derived from only the pre-s_2 region. All three arrows contain the entire surface antigen sequence. Small vertical arrows identify potential N-linked glycosylation sites (Asn-X-Thr[Ser]). Numbers below the line correspond to the nucleotide sequence numbering system of Galibert and co-workers (12). The bracket indicates the position, with respect to the viral ORF, of the *Xho*II fragment inserted into pZL811.

of and in-frame with the ATG presumed to be used for the major coat protein WHsAg (Fig. 3) (12). The region of pre-*s* which is included in the tribrid construction has the potential to encode 89 amino acids, 88 of which are derived from amino acids 85 to 173 of a protein initiated from the first ATG of the ORF. An additional codon is contributed, since the linker region of the vector encodes the same residue as the viral sequence at position 174 (proline). The relative position of the pre-*s*-encoded segment of the fusion protein with respect to the three predicted viral polypeptides derived from the pre-*s*:S ORF is shown in Fig. 3. As indicated, the amino acids encoded by the pre-*s* segment of the tribrid are found in the putative protein translated from the first ATG as well as in the potential 33-kDa protein translated from the second ATG. Based on this information, it was predicted that the rabbit antisera generated to the bacterially synthesized fusion protein should have the ability to recognize both of the putative polypeptides from the pre-*s*:S ORF. The known mature coat protein, which has been proposed to be translated from the third ATG of the ORF (12), should not be recognized by antisera to the fusion protein. The use of anti-tribrid serum in Western blot analysis of partially purified WHV from a single woodchuck clearly identified four polypeptides (Fig. 4, lane 6) which were not recognized by preimmune sera from the same rabbit (Fig. 4, lane 2) or by antisera to the hybrid β-gal:CAT protein (Fig. 4, lane 4). Two of these polypeptides (p45 and p33) correspond approximately in M_r to molecules initiated from the first and second ATGs of the ORF. Based on the existence of only two potential initiation codons in the pre-*s* region, the other two polypeptides recognized by anti-tribrid sera (p47 and p36) are presumably posttranslationally modified. By analogy to the pre-*s*-encoded proteins in HBV (14, 45), we speculate

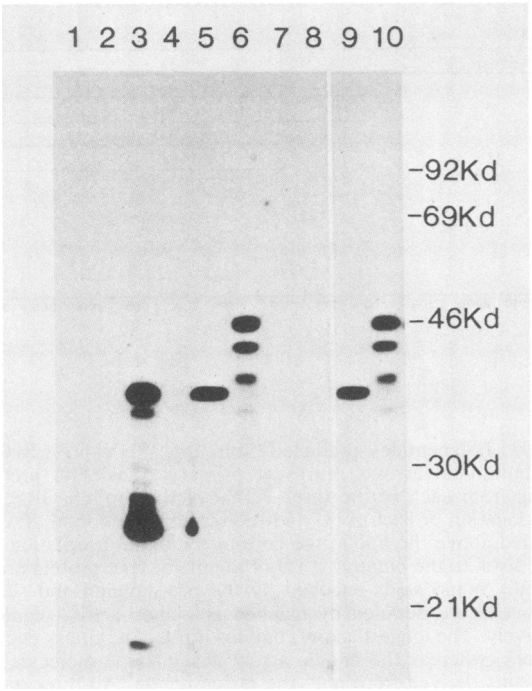

FIG. 4. Western blot identification of WHV pre-*s* polypeptides. Odd-numbered lanes contain 0.2 μg of affinity-purified β-gal:pre-*s*:CAT fusion proteins, and even-numbered lanes contain partially purified WHV from 1 ml of infectious woodchuck serum. Pairs of lanes were probed as follows. Lanes 1 and 2 were treated with preimmune rabbit serum, lanes 3 and 4 were treated with anti-β-gal:CAT serum, and lanes 5 and 6 were treated with anti-β-gal:pre-*s*:CAT serum. Lanes 7 and 8 were treated as lanes 5 and 6 were, but antiserum was preadsorbed with the tribrid protein; and lanes 9 and 10 were treated as lanes 5 and 6 were, but antiserum was preadsorbed with the hybrid protein. Following incubation with rabbit antisera, all lanes were probed with ^{125}I-labeled protein A (see Materials and Methods for details). M_rs of protein standards are indicated at the right.

that p47 and p36 are glycosylated derivatives of p45 and p33, respectively. To confirm that the recognition of four polypeptides by anti-β-gal:pre-*s*:CAT was specifically related to the pre-*s* determinants within these molecules, we preadsorbed antisera with either the hybrid or tribrid proteins before probing the Western blot. Antisera preincubated with β-gal:pre-*s*:CAT no longer recognized the viral polypeptides (Fig. 4, lane 8), whereas preincubation of the antisera with an equivalent concentration of the hybrid protein had no effect on the recognition of these polypeptides (Fig. 4, lane 10). The result demonstrated that the identification of four viral polypeptides by the anti-tribrid sera was specifically due to the presence of pre-*s*-encoded determinants on these molecules. Additionally, β-gal:CAT determinants did not appear to play a role in the recognition of viral proteins by anti-tribrid sera. Similar data have been generated with WHV preparations from the sera of other animals.

Identification of WHsAg determinants on pre-*s* polypeptides. The viral polypeptides containing determinants encoded by pre-*s* should also, in the absence of an mRNA splicing event, include the 222 amino acids of WHsAg. The M_rs of the pre-*s* molecules determined from SDS-PAGE gels are consistent with uninterrupted translation from both the first and second AUGs of pre-*s*:S ORF transcripts, giving rise to p45 and p33, respectively. The presence of surface antigen determinants was confirmed by probing Western blots of the same virus preparation with monoclonal antibody 101-2 generated to purified WHsAg (7). 101-2 recognized the major coat proteins of the virus (p22 and gp25) as well as the four higher-M_r proteins previously shown to contain pre-*s*-encoded determinants (Fig. 5, lane 1).

Cross-reactivity of pre-*s*-encoded determinants from WHV and GSHV. Of the hepadnaviruses, GSHV is known to have the highest level of nucleotide sequence homology with WHV (43). Whereas the predicted amino acid sequence of the pre-*s* region of HBV has only minimal homology with that region from WHV (12), pre-*s* from GSHV shares a high degree of homology with WHV pre-*s* (71% amino acid identity) (43). We therefore wished to determine if the antiserum generated to WHV pre-*s*-encoded amino acids would recognize polypeptides from GSHV. Western blot analysis with partially purified GSHV demonstrated a similar pattern of pre-*s*-containing polypeptides to that seen for WHV (Fig. 5, lane 3). The quantity of higher-M_r polypeptides (p45 and gp47) relative to molecules initiated from the internal ATG (p33 and gp36) is apparently greater in the GSHV preparation than in that from woodchuck sera. Upon longer exposure of this blot, polypeptides with M_rs consistent with initiation of translation from an internal ATG were evident (data not shown).

Immunoprecipitation of DNA polymerase activity. The major coat proteins from WHV, like HBV surface antigen, are associated with three distinct morphological forms, spherical particles 22 nm in diameter, long filaments, and nucleic acid-containing mature virions (48). Only the mature virion is infectious and contains DNA polymerase activity associated with replication. To determine if pre-*s*-encoded amino acids were present on DNA polymerase-containing virus particles, we carried out immunoprecipitation of partially purified WHV (Table 1). 101-2 was able to efficiently precipitate virus particles containing endogenous DNA polymerase activity. Anti-β-gal:pre-*s*:CAT precipitated approximately 14-fold more polymerase activity than did anti-β-gal:CAT.

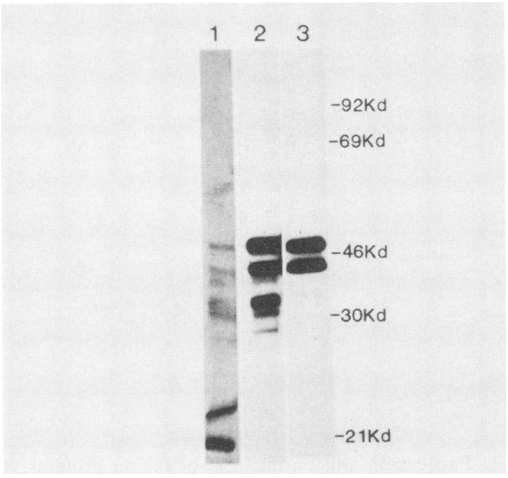

FIG. 5. Analysis of GSHV pre-*s* polypeptides and WHsAg determinants. Lanes 1 and 2 contain partially purified WHV from 0.5 ml of infectious woodchuck serum. Lane 3 contains partially purified GSHV from 0.6 ml of infectious ground squirrel sera. Lane 1 was probed with 101-2 followed by rabbit anti-mouse IgG-IgA-IgM, conjugated to horseradish peroxidase, and developed with substrate solution. Lanes 2 and 3 were probed with anti-β-gal:pre-*s*:CAT serum followed by ^{125}I-labeled protein A (see Materials and Methods for details).

To assess if the determinants being recognized by anti-β-gal:pre-s:CAT were on the virion surface, it was of interest to determine the degree to which particles were intact. Antibody to the nucleocapsid protein from HBV did not efficiently precipitate polymerase activity unless virus preparations were preincubated with NP-40-2-β-mercaptoethanol, which has been shown to solubilize the viral envelope, exposing core antigen determinants (39). These data demonstrated the existence of pre-s-related polypeptides on polymerase-containing particles and supported a surface location for these molecules.

DISCUSSION

Galibert and co-workers have shown that the surface antigen regions from WHV and HBV share considerable amino acid sequence homology, whereas the pre-s segment of the ORF has little such homology (12). In our earlier studies, computer algorithms which predict structural characteristics of a polypeptide based on its amino acid sequence were used to further compare the pre-s regions of the two viruses (41). We observed that, despite minimal amino acid sequence homology, the potential to form α-helical secondary structure, as well as the overall hydrophilicity of these coding regions, was highly conserved (41). This maintenance of structural properties suggested to us that pre-s was expressed by these viruses. In addition, the fact that pre-s was likely to exist as part of a surface antigen-containing polypeptide, along with the more hydrophilic nature of this region compared with surface antigen, suggested that it might extend from the surface of the virion.

This report describes the cloning and expression of a DNA segment from the pre-s region of WHV. We used an ORF vector to overproduce a tribrid fusion protein containing determinants encoded by the pre-s segment which was inserted between the eight N-terminal amino acids of β-gal and the entire CAT protein. The tribrid polypeptide was purified by a one-step affinity procedure in which fusion proteins were specifically bound to a chloramphenicol resin by virtue of the substrate-binding site of CAT. The purified tribrid polypeptide was used to generate antisera in rabbits, which appeared to preferentially recognize the pre-s determinants of the fusion protein. Antisera which have been generated to other β-gal:viral:CAT tribrid proteins (11, 53) have, to varying degrees, shown a greater affinity for the viral portion of the molecule than to the CAT determinants. This may be a property common to these fusion proteins and related to the N-terminal location of the viral sequences in the tribrid molecules, to the tetrameric structure known to exist for native CAT, or to both. Whatever accounts for

TABLE 1. Immunoprecipitation of polymerase activity[a]

Antibody	TCA-precipitated radioactivity	% Precipitated
101-2	57,405	37.6
Anti-HBc − NP-40	3,799	2.5
Anti-HBc + NP-40	63,431	41.6
Anti-β-gal:CAT	1,830	1.2
Anti-β-gal:pre-s:CAT	24,889	16.3
Before precipitation	152,623	

[a] Anti-HBc is a rabbit polyclonal antiserum generated to *E. coli*-synthesized HBV core antigen. Anti-β-gal:CAT and anti-β-gal:pre-s:CAT are rabbit polyclonal antisera generated to the hybrid and tribrid proteins, respectively (see Results). The designation before precipitation represents the DNA polymerase activity present in a partially purified WHV sample equivalent to that used in immunoprecipitations. TCA-precipitated radioactivity was averaged from duplicate samples from three independant experiments.

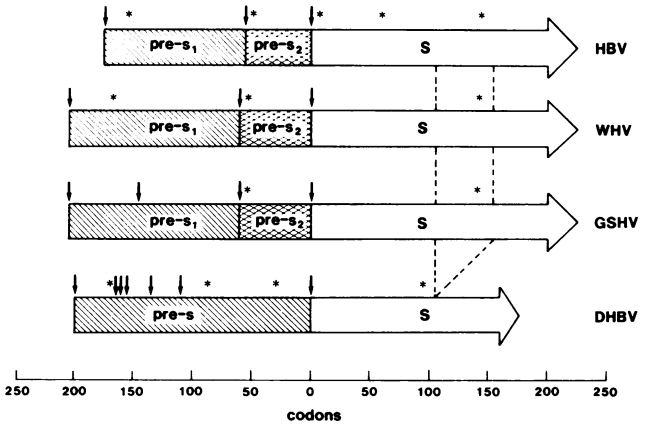

FIG. 6. Pre-s:S ORFs from hepadnaviruses. The four large arrows represent pre-s:S ORFs from the four hepadnaviruses indicated at the right. Vertical arrows denote ATG codons, and asterisks correspond to potential glycosylation sites (Asn-X-Thr[Ser]). Initiation codons distal to the known and putative major coat protein ATGs are not indicated. For the mammalian viruses, hatched areas represent the pre-s_1 region encoding N-terminal portions of proteins with M_rs ranging from 39 to 47 kDa. Cross-hatched regions represent pre-s_2, encoding amino acids found in proteins with M_rs ranging from 31 to 36 kDa. The open region of arrows labeled S are genes encoding the major coat protein of these viruses. For duck hepatitis virus (DHBV), the hatched area represents pre-s sequences, but the ATG(s) used to initiate translation has not yet been determined. The dotted lines indicate the segment of the mammalian viruses surface antigen gene which appears to be deleted in the DHBV S gene. Codons are numbered with zero (0) representing the location of the major coat protein ATG. Numbers to the right of zero are codon positions in the S gene, while to the left of zero are codon positions in the pre-s region.

immunogenic properties of the fusion polypeptides, these proteins conveniently allow for the production of antisera with a relatively high titer to the viral sequences of interest. We used the anti-tribrid serum to identify WHV pre-s-encoded polypeptides by Western blot analysis. We also showed that the antiserum recognizes analogous polypeptides which are present in partially purified preparations of GSHV. The WHV polypeptides identified by the anti-tribrid sera were also recognized by a monoclonal antibody generated to WHs-Ag, further supporting the notion that they are products of pre-s:S ORF. Finally, the anti-tribrid serum was used to immunoprecipitate polymerase-containing virus particles. From these data we conclude the following. (i) Polypeptides are encoded by the pre-s regions of WHV and GSHV, initiating at both the first (p45 and gp47) and an internal (p33 and gp36) ATG of the pre-s:S ORF. (ii) M_rs of proteins containing pre-s determinants are consistent with a pair of differentially glycosylated molecules being derived from each primary translation product as demonstrated for HBV (14). (iii) Pre-s polypeptides are located on the surface of infectious virus particles.

From the characterization of hepadnavirus structural proteins there has emerged a family of pre-s-encoded polypeptides (Fig. 6) (9). This protein family, particularly for the mammalian members, shares the locations of initiation codons and glycosylation sites, resulting in the production of three pairs of analogous polypeptides for each virus. The purification of Machida et al. (22) of polypeptides from HBV surface antigen 22-nm particles containing the 55 amino acids encoded by the DNA upstream of the ATG for the major coat protein was the first unambiguous demonstration

that the pre-*s* region was expressed. Our laboratory (53) and others (14) have since identified molecules from HBV with M_rs ranging from ca. 39 to 45 kDa, which appear to be products of the entire pre-*s*:S ORF. In this report, we have shown that the other two hepadnavirus members which infect mammalian hosts (WHV and GSHV) also express proteins from their respective pre-*s*:S ORFs. The two ATGs which are apparently used to initiate translation of pre-*s* encoded proteins are located in analogous positions in the pre-*s*:S ORFs of these three viruses, with an additional ATG occurring in the GSHV ORF (Fig. 6). A carbohydrate attachment site has been identified at amino acid position 146 of HBV surface antigen (36). Whereas it is not known which other glycosylation sites are used by HBV or the other mammalian viruses, at least two Asn-X-Thr(Ser) sequences are located in approximately the same positions in the pre-*s*:S ORFs of each virus (Fig. 6).

From the predicted amino acid sequence of duck hepatitis virus (23), it appears that the organization of its pre-*s*:S ORF is somewhat different. There are six ATGs which precede the ATG thought to be used to initiate translation of the duck virus major coat protein (23). In addition, based on amino acid sequence comparisons with the other hepadnaviruses, there seems to be a segment which is deleted from the duck hepatitis virus surface antigen gene (S gene), and potential glycosylation site positions do not correlate as well with those from mammalian virus ORFs (Fig. 6). Using similar approaches to those described in this paper, we have recently identified a polypeptide from duck hepatitis virus which contains pre-*s* determinants (E. Schaeffer and J. Sninsky, manuscript in preparation). Further studies will be needed to determine which ATG is used to initiate translation of this molecule, as well as the presence and locations of carbohydrate moieties.

A function for pre-*s*-encoded sequences in the viral life cycle has yet to be elucidated. We have demonstrated that the pre-*s* regions of HBV (53), WHV, and GSHV are expressed and, at least for two of these viruses (for HBV, D. T. Wong and J. J. Sninsky, unpublished result), result in polypeptides which are found on the surface of virions. Computer analyses of pre-*s* sequences from WHV and HBV have indicated that structural characteristics have been conserved, whereas amino acid sequences appear to have diverged considerably (41). A polypeptide with the above properties might logically be involved in host cell recognition in the species-specific manner characteristic of this viral class. If pre-*s*-derived sequences are important for host cell recognition, one would predict that antibodies generated to this region may effectively neutralize infection. Preliminary data by D. T. Wong, A. Prince, N. Nath, and J. Sninsky (Abstr. Cold Spring Harbor Meet. Mol. Biol. Hepatitis B viruses. 1985, p. 17) have demonstrated that chimpanzees immunized with a tribrid molecule, similar to the fusion protein described in this report but containing HBV pre-*s*-encoded determinants, were protected against infection upon challenge with HBV. Analogous experiments are contemplated in which woodchucks are immunized with the WHV pre-*s* tribrid proteins before challenge with virus. The identification of pre-*s*-encoded epitopes which have the capacity to induce a neutralizing antibody response may help determine sequences which play a crucial role during infection.

Viral entry into a host cell may be mediated by specific recognition of a normal cellular protein. In the case of Epstein-Barr (10) and rabies viruses (20), tissue- or cell-specific receptors appear to determine the tropisms associated with these agents. Of course, host and tissue tropisms are also likely to be due to cytoplasmic factors which may be important in replication, transcription, and assembly processes. A hepatocyte-specific protein may be recognized by pre-*s* sequences, accounting in part for the hepatotropic nature of the hepadnaviruses. To this end, we and others (31) have initiated experiments aimed at identification of a protein from hepatocytes which specifically binds pre-*s*-encoded determinants, hoping to shed light on the receptor used by these viruses.

The interaction between a viral protein and host cell receptor may be direct but, alternatively, could require a third component. The 55 pre-*s*-encoded amino acids immediately upstream of the ATG for HBV surface antigen have been shown to have the ability to bind to polymerized human serum albumin in a species-specific manner (21). Polymerized albumin, as was first suggested by Imai et al. (15) may act as an intermediary, being recognized by both HBV and a hepatocyte membrane protein. This indirect association of the virion with hepatocytes has been postulated as the route used for viral entry. At present, this mode of virus-host cell attachment remains speculative. A specific polymerized albumin receptor on liver cells has not been characterized, although binding of polymerized albumin to the surface of rabbit hepatocytes has been observed (19). Whether the WHV pre-*s* sequences in our tribrid will bind to polymerized woodchuck serum albumin remains to be determined.

The efficient use of nucleotide sequence by this class of viruses and the existence (at least for HBV) of two dissimilar promoters associated with the pre-*s*:S ORF (4, 34, 38) support a distinct role for each pair of polypeptides translated from this ORF. Proteins initiated from the first ATG, and those translated from an internal pre-*s* ATG, may have related functions in host cell entry (i.e., a multicomponent recognition process). However, these proteins may have disparate functions, with only one pair of molecules performing a role in receptor recognition. The delineation of a function for each polypeptide (or pair of polypeptides) encoded by the pre-*s*:S ORF will undoubtedly provide insight into the life cycle of hepadnaviruses.

ACKNOWLEDGMENTS

We thank P. J. Cote and J. L. Gerin for 101-2 and P. L. Marion and W. S. Robinson for providing infectious ground squirrel sera. We also thank S. Tickton for help in typing the manuscript and Graphics Services at Cetus Corporation for graphics and photography.

We acknowledge financial support from Cetus Corporation for a portion of this study. E.S. acknowledges predoctoral support from National Institutes of Health training grant 5T32 CA09060; J.J.S. acknowledges Public Health Service grant A108295 from the National Institutes of Health and an award (JFRA-42) by the American Cancer Society.

LITERATURE CITED

1. **Alton, N. K., and D. Vapnek.** 1979. Nucleotide sequence analysis of the chloramphenicol resistance transposon Tn9. Nature (London) **282**:864–869.
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. **Chang, A. C. Y., and S. N. Cohen.** 1978. Construction and characterization of amplifiable multicopy DNA cloning vehicles

derived from the P15A cryptic miniplasmid. J. Bacteriol. **134**:1141–1156.
6. **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. Sci. USA **69**:2110–2114.
7. **Cote, P. J., Jr., G. M. Dapolito, J. W.-K. Shih, and J. L. Gerin.** 1982. Surface antigenic determinants of mammalian "hepadnavirus" defined by group- and class-specific monoclonal antibodies. J. Virol. **42**:135–142.
8. **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 Publishers, New York.
9. **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.
10. **Fingeroth, J. D., J. J. Weis, T. F. Tedder, J. L. Strominger, P. A. Biro, and D. T. Fearon.** 1984. Epstein-Barr virus receptor of human B lymphocytes is the C3d receptor CR2. Proc. Natl. Acad. Sci. USA **81**:4510–4514.
11. **Friefeld, B. R., R. Korn, P. J. deJong, J. J. Sninsky, and M. S. Horwitz.** 1985. The 140-kDa adenovirus DNA polymerase is recognized by antibodies to *Escherichia coli*-synthesized determinants predicted from an open reading frame on the adenovirus genome. Proc. Natl. Acad. Sci. USA **82**:2652–2656.
12. **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.
13. **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.
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. **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.
16. **Ish-Horowicz, D., and J. F. Burke.** 1981. Rapid and efficient cosmid cloning. Nucleic Acids Res. **9**:2989–2998.
17. **Johnson, D. A., J. W. Gautsch, J. R. Sportsman, and J. H. Elder.** 1984. Improved technique utilizing nonfat dry milk for analysis of proteins and nucleic acids transferred to nitrocellulose. Gene Anal. Technol. **1**:3–8.
18. **Laemmli, U. K.** 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature (London) **227**:680–685.
19. **Lenkei, R., D. Onica, and V. Ghetie.** 1977. Receptors for polymerized albumin on liver cells. Experientia **33**:1046–1047.
20. **Lentz, T. L., T. G. Burrage, A. L. Smith, J. Crick, and G. H. Tignor.** 1982. Is the acetylcholine receptor a rabies virus receptor? Science **215**:182–184.
21. **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.
22. **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 of polymerized human as well as chimpanzee albumins. Gastroenterology **85**:268–274.
23. **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.
24. **Maniatis, T., E. F. Fritsch, and J. Sambrook.** 1982. Molecular cloning: a laboratory manual, p. 156–174. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
25. **Marion, P. L., L. S. Oshiro, D. C. Regnery, G. H. Scullard, and W. S. Robinson.** 1980. A virus in beechey ground squirrels that is related to hepatitis B virus of humans. Proc. Natl. Acad. Sci. USA **77**:2941–2945.
26. **Mason, W. S., G. Seal, and J. Summers.** 1980. Virus of Pekin ducks with structural and biological relatedness to human hepatitis B virus. J. Virol. **36**:829–836.
27. **Meselson, M., and R. Yuan.** 1968. DNA restriction enzyme from *E. coli*. Nature (London) **217**:1110–1114.
28. **Miller, J. H.** 1972. Experiments in molecular genetics, p. 433. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
29. **Miller, R. H., P. L. Marion, and W. S. Robinson.** 1984. Hepatitis B viral DNA-RNA hybrid molecules in particles from infected livers are converted to viral DNA molecules during an endogenous DNA polymerase reaction. Virology **139**:64–72.
30. **Millman, I., T. Halbherr, and H. Simmons.** 1982. Immunological cross-reactivities of woodchuck and hepatitis B viral antigens. Infect. Immun. **35**:752–757.
31. **Neurath, A. R., S. B. H. Kent, N. Strick, P. Taylor, and C. E. Stevens.** 1985. Hepatitis B virus contains pre-S gene-encoded domains. Nature (London) **315**:154–156.
32. **Ogston, C. W., G. J. Jonak, C. E. Rogler, S. M. Astrin, and J. Summers.** 1982. Cloning and structural analysis of integrated woodchuck hepatitis virus sequences from hepatocellular carcinomas of woodchucks. Cell **29**:385–394.
33. **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.
34. **Ou, J.-H., and W. J. Rutter.** 1985. Hybrid hepatitis B virus-host transcripts in a human hepatoma cell. Proc. Natl. Acad. Sci. USA **82**:83–87.
35. **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.
36. **Peterson, D. L., N. Nath, and F. Gavilanes.** 1982. Structure of hepatitis B surface antigen. J. Biol. Chem. **257**:10414–10420.
37. **Popper, H., J. W.-K. Shih, J. L. Gerin, D. C. Wong, B. H. Hoyer, W. T. London, D. L. Sly, and R. H. Purcell.** 1981. Woodchuck hepatitis and hepatocellular carcinoma: correlation of histologic with virologic observations. Hepatology **1**:91–98.
38. **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.
39. **Robinson, W. S., and R. L. Greenman.** 1974. DNA polymerase in the core of human hepatitis B virus candidate. J. Virol. **13**:1231–1236.
40. **Robinson, W. S., and H. L. Robinson.** 1971. DNA polymerase in defective Rous sarcoma virus. Virology **44**:457–462.
41. **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. USA **81**:2902–2906.
42. **Schottel, J. L., M. J. Bibb, and S. N. Cohen.** 1981. Cloning and expression in *Streptomyces lividans* of antibiotic resistance genes derived from *Escherichia coli*. J. Bacteriol. **146**:360–368.
43. **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.
44. **Snyder, R. L., and J. Summers.** 1980. Woodchuck hepatitis virus and hepatocellular carcinoma, p. 447–457. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
45. **Stibbe, W., and W. H. Gerlich.** 1983. Structural relationships between minor and major proteins of hepatitis B surface antigen. J. Virol. **46**:626–628.
46. **Summers, J.** 1981. Three recently described animal virus models for human hepatitis B virus. Hepatology **1**:179–183.
47. **Summers, J., and W. S. Mason.** 1982. Replication of the genome

of a hepatitis B-like virus by reverse transcription of an RNA intermediate. Cell **29**:403–415.

48. **Summers, J., J. M. Smolec, and R. L. Snyder.** 1978. A virus similar to human hepatitis B virus associated with hepatitis and hepatoma in woodchucks. Proc. Natl. Acad. Sci. USA **75**:4533–4537.
49. **Timmis, K. F., F. Cabello, and S. N. Cohen.** 1978. Cloning and characterization of *Eco*RI and *Hin*dIII restriction endonuclease-generated fragments of antibiotic resistance plasmids R6-5 and R6. Mol. Gen. Genet. **162**:121–137.
50. **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. USA **76**:4350–4354.
51. **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.
52. **Valenzuela, 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.
53. **Wong, D. T., N. Nath, and J. J. Sninsky.** 1985. Identification of hepatitis B virus polypeptides encoded by the entire pre-*s* open reading frame. J. Virol. **55**:223–231.
54. **Zaidenzaig, Y., and W. V. Shaw.** 1976. Affinity and hydrophobic chromatography of three varients of chloramphenicol acetyltransferases specified by R factors in *Escherichia coli*. FEBS Lett. **62**:266–271.