Proc. Natl. Acad. Sci. USA
Vol. 92, pp. 6259–6263, July 1995
Medical Sciences

The pre-S domain of the large viral envelope protein determines host range in avian hepatitis B viruses

TAKASHI ISHIKAWA AND DON GANEM*

Howard Hughes Medical Institute and Departments of Microbiology and Immunology and Medicine, University of California Medical Center, San Francisco, CA 94143-0502

Communicated by Rudi Schmid, University of California, San Francisco, CA, February 21, 1995

ABSTRACT **In addition to their well-recognized hepatotropism, all hepatitis B viruses (HBVs) display marked species specificity, growing poorly or not at all in species other than those closely related to their natural hosts. We have examined the molecular basis for this narrow host range, using duck HBV (DHBV) and heron HBV (HHBV) as a model system. HHBV virions will not infect ducks *in vivo* and infect cultured duck hepatocytes extremely inefficiently *in vitro*. Mutant HHBV genomes lacking all viral envelope proteins (HHBV *env*$^-$) can be complemented in trans with DHBV envelope proteins; the resulting pseudotyped virions can efficiently infect duck hepatocytes. Further complementation analysis reveals that of the two viral surface proteins (L and S), it is the L protein that determines host range. Pseudotyping of HHBV *env*$^-$ with DHBV/HHBV chimeric envelope proteins reveals that replacement of as few as 69 amino acids of the pre-S domain of the HHBV L protein by their DHBV counterparts is sufficient to permit infection of duck hepatocytes. These studies indicate that the species-specificity of hepadnaviral infection is determined at the level of virus entry and is governed by the pre-S domain of the viral L protein.**

Hepatitis B viruses (HBVs; hepadnaviruses) are enveloped DNA viruses that produce persistent infections of hepatocytes, often resulting in the development of chronic hepatitis, liver failure, and hepatocellular carcinoma (1). The prototype hepadnavirus is human HBV; other members of the hepadnavirus family are found in woodchucks, ground squirrels, ducks, and herons. A characteristic feature of all hepadnaviral infection is strong species-specificity. HBV, for example, grows in humans and chimpanzees but does not grow in baboons, lower primates, or other mammals (2); duck hepatitis B virus (DHBV) grows in ducks and certain strains of geese but does not grow in chicken or mammalian hosts. This narrow host range, which has been an important impediment to the experimental study of human hepatitis B, is thought to be governed at the level of virus entry into cells. For example, although chicken or human hepatocytes cannot be infected with DHBV virions, transfection of these same cells with cloned DHBV DNA initiates a productive viral life cycle (3–6). Thus, bypass of the normal entry mechanism bypasses the block to trans-species infection. Little is known, however, of the viral and host factors that account for the species-specificity of viral entry. Here we examine the viral factors that determine host range, using DHBV and heron HBV (HHBV) as experimental systems.

As for other enveloped viruses, hepadnaviral entry is governed by the viral envelope (or surface) proteins (1). In hepadnaviruses these proteins are uncommonly complex, both in number and in structure. The avian viruses encode two related surface proteins, by differential translational initiation within a single open reading frame (ORF). The larger of these two proteins (L, 36 kDa) is initiated at a 5′ AUG of the envelope ORF; initiation at an internal AUG (on a separate mRNA) generates the 17-kDa S protein (see Fig. 1). The L protein thus has two domains: the S domain and a second domain encoded by the upstream (or pre-S) sequences. Pre-S amino acid sequences are the least conserved coding sequences among different hepadnaviruses, which has prompted the speculation that they might relate to host range determination (7). Supporting an important role for pre-S sequences in infectivity is the observation that certain antibodies to them can neutralize infection (8, 9) or block binding to hepatocyte membranes (10). To date, however, no experiments have been reported that directly determine which domain(s) of the surface proteins govern viral host range. In this paper we examine the infectivity of recombinant HHBV virions bearing DHBV envelope determinants on their surface. These studies directly demonstrate that pre-S sequences are the prime determinants of the species specificity of hepadnaviral infection.

MATERIALS AND METHODS

Plasmid Constructions. Plasmid pD1.5G (Fig. 1) contains an overlength (1.5-mer) copy of DHBV DNA (7) inserted into *Bam*HI and *Eco*RI sites of pBS vector. Plasmid pH1.4G (Fig. 1) contains the overlength (1.4-mer) copy of HHBV DNA (11) inserted into an *Eco*RI site of the pIBI20 vector. For the construction of DHBV L^-S^+, DHBV L^+S^-, HHBV L^-S^+, and HHBV L^+S^-, *in vitro* mutagenesis was done according to Kunkel *et al.* (12). Briefly, for L^-S^+ mutants, stop codons were introduced at nt 816–818 (TCG → TAA) in DHBV and at nt 813–815 (TCA → TAA) in HHBV. L^+S^- mutants were generated by ATG → ACG substitution at the start site of the S genes in both DHBV and HHBV. These mutations do not disrupt the amino acid sequences of the polymerase gene. To generate the HHBV *env*$^-$ genome (pHSS1, Fig. 1), a stop codon was introduced at nt 1303–1305 by TTC → TAA substitution. Pre-S/S expression vectors were constructed by ligating the *Hin*dIII-digested monomer of HHBV to the pGEM vector, or the *Eco*RI-digested monomer of DHBV to the pGEM vector (generating pD1.0G, Fig. 1).

HHBV/DHBV chimeric genes were generated by the PCR techniques described by Deminie and Emerman (13); details of the constructions are available from the authors upon request. Structures of all chimeric clones were confirmed by dideoxynucleotide chain-termination sequencing of the entire pre-S region.

Analysis of Enveloped Virions in LMH Culture Supernatant. Culture fluids of LMH cells (3) were harvested and clarified by centrifugation at 3000 rpm for 10 min. Virus was precipitated from the clarified supernatant by adding polyethylene glycol to a final concentration of 10% followed by incubation at 4°C for 1 hr. The precipitates were collected by

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviations: HBV, hepatitis B virus; DHBV, duck HBV; HHBV, heron HBV; ORF, open reading frame; PDH, primary duck hepatocytes.

*To whom reprint requests should be addressed.

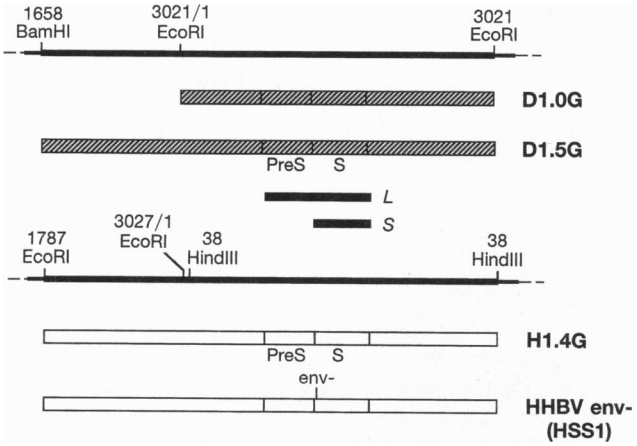

FIG. 1. Avian hepadnaviral genome organization. Hatched bar, DHBV genome in pD1.5G; open bar, HHBV genome in pH1.4G. The envelope-coding regions corresponding to the L and S proteins are shown as closed boxes. The HHBV env$^-$ mutant (generated by plasmid HSS1) was constructed by introduction of a stop codon just downstream of S AUG initiation site.

centrifugation at 10,000 × *g* for 20 min and dissolved in 90 μl of 2 mM Hepes, pH 7.5/150 mM NaCl/1 mM $CaCl_2$). Two hundred and seventy microliters of 10 mM Tris, pH 7.5/1 mM EDTA (TE buffer) (10:1) containing Pronase at 750 μg/ml was added and incubated for 1 hr at 37°C. Lenhoff and Summers (14) have shown that this digestion is sufficient to release viral DNA contained in free capsids but is insufficient to release viral DNA from enveloped particles. Viral DNA released by the Pronase digestion was then removed by the addition of DNase I at 1 mg/ml and incubation for 30 min at 37°C. Control experiments (Fig. 4*B*, lane 8, and data not shown) confirmed that these digestion conditions suffice to eliminate 90–95% of the DNA within free capsids while virion DNA was undisturbed. For the analysis of viral DNA forms present in the Pronase-resistant particles, nucleic acids were extracted with SDS and proteinase K treatment (as above) followed by phenol extraction and ethanol precipitation.

RESULTS AND DISCUSSION

The Host Range of HHBV. Earlier studies have shown that HHBV virions from the serum of viremic herons are incapable of infecting Pekin ducks after parenteral inoculation (11). We have confirmed this result by using cloned HHBV DNA (11). The HHBV molecular clone we used (H1.4G, Fig. 1) is replication competent: when transfected into permissive chicken LMH hepatoma cells, viral replicative intermediates appear in the cytoplasm (data not shown) and medium (Fig. 2*A*); CsCl density gradients confirm that a subset of the released particles has the buoyant density of enveloped virions (data not shown). However, as expected from earlier data, HHBV cannot initiate a spreading infection of ducks: after intrahepatic inoculation of cloned DHBV DNA none of 14 ducks developed HHBV infection, whereas two of three ducks similarly inoculated with DHBV DNA were successfully infected (data not shown).

We next examined the ability of HHBV and DHBV virions in the medium of transfected LMH cells to infect primary duck hepatocytes (PDH) in cell culture. Unconcentrated medium from such cells (containing ≈5 × 10^7–1 × 10^8 genome equivalents) was used to infect PDH cells as described (18, 19); 14 days later, cytoplasmic extracts were examined for replica-

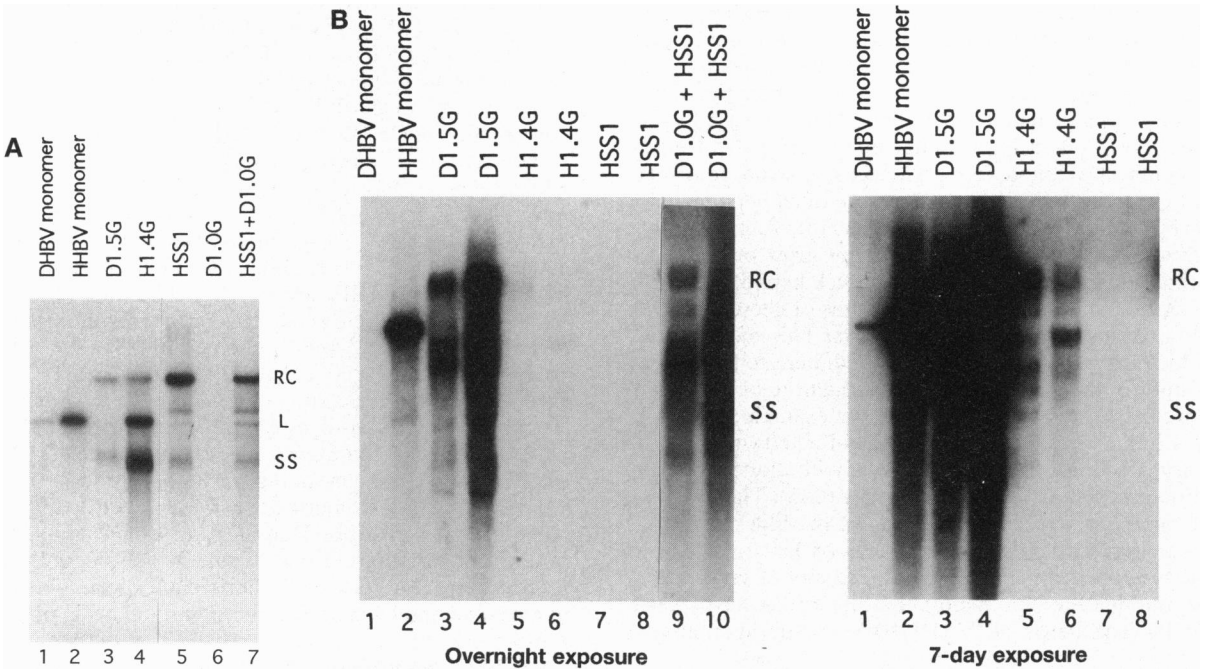

FIG. 2. (*A*) Viral DNA in medium of LMH cells transfected with wild-type and envelope mutant plasmids. Ten micrograms of the indicated plasmid DNA was used for transfection of a 100-mm-diameter dish of LMH cells. Supernatant was harvested 5 days after transfection, and viral replicative intermediates were purified as described (15) and analyzed by agarose gel electrophoresis and blot hybridization (16, 17) using ^{32}P-labeled HHBV DNA as a probe. Lanes 1 and 2 contain 3.0 kb of linear plasmid molecules of DHBV and HHBV, respectively, as molecular mass markers. The migration positions of relaxed circular (RC), linear (L), and single-stranded (SS) DNA are indicated. (*B*) Susceptibility of primary duck hepatocytes (PDH) cells to DHBV and HHBV infection. Two milliliters of the indicated LMH cell supernatant of *A* was used for infection of PDH in 60-mm dishes. Total intracellular DNA was prepared from each sample 14 days after infection, and one-fourth of this material was analyzed by Southern blot hybridization with ^{32}P-labeled HHBV DNA probe. D1.5G, wild-type DHBV infection; H1.4G, wild-type HHBV infection; HSS1, infection by HHBV stock generated by the HHBV env$^-$ genome of Fig. 1; D1.0G + HSS1, infection by virus derived from pseudotyping of HHBV env$^-$ with DHBV envelope proteins (L and S). Lanes 1–8 (*Right*) represent a long exposure of the same experiment shown in lanes 1–8 (*Left*).

tive intermediates by Southern blotting (15, 16). As shown in Fig. 2*B*, DHBV-containing medium produced an efficient infection of PDH cells (lanes 3 and 4). (Because this blot was analyzed with ^{32}P-labeled HHBV probe, which cross-hybridizes less efficiently with DHBV sequences than would homologous DHBV probe, the observed signal actually underestimates the extent of DHBV production in this sample.) Surprisingly, however, wild-type HHBV particles were not entirely devoid of PDH infectivity: viral replicative intermediates accumulated to 1–2% the level of their DHBV counterparts (lanes 5 and 6, best seen in the overexposed blot at right). Because replication of the two genomes is comparably efficient in transfected LMH cells (Fig. 2*A*, compare lanes 3 and 4), this difference is most likely due to different efficiencies of viral entry. That this represents *bona fide* virion infection is shown by the ablation of HHBV replication in PDH when a stop codon mutation (*env*$^-$, pHSS1) is introduced into the HHBV S gene. Although this mutation does not impair genomic replication in LMH (Fig. 2*A*, lane 5), the resulting medium is devoid of infectivity for PDH (Fig. 2*B*, lanes 7 and 8).

To see whether the low level of HHBV replication in PDH is truly due to inefficient entry, we asked whether infection could be enhanced by provision of DHBV envelope glycoproteins in trans. LMH cells were cotransfected with HHBV env$^-$ and an expression vector producing DHBV envelope glycoproteins (the latter was generated by cloning a monomeric DHBV genome into pGEM so as to interrupt expression of both C and P genes; because the L and S surface proteins are produced in this vector from their native viral transcripts, they should accumulate in the correct stoichiometry). As shown in Fig. 2*B* (lanes 9 and 10), infection by HHBV env$^-$ (HSS1) was efficiently rescued by provision of DHBV envelope proteins. These data indicate that the block to HHBV infection of duck cells is at the level of virus entry and is substantial (50- to 100-fold) but not absolute. Clearly, this low level of infectivity is insufficient to sustain multiple rounds of viral spread in the intact animal host. Precedents for this behavior have emerged recently from observations of others on DHBV infection of Muscovy ducks (20). Unlike Pekin ducks, Muscovy ducks cannot be infected by parental inoculation with DHBV. However, cultured Muscovy hepatocytes are susceptible to DHBV virion infection at low efficiency (20).

L Protein Governs Host Range. To explore the roles of the L and S proteins in host-range determination, we initially set out to examine the PDH infectivity of virions bearing L proteins from one virus and S proteins from the other. To do this we first constructed replication-competent (1.5-mer) HHBV and DHBV genomes harboring mutations that ablate either the L or the S protein. L mutants were generated by introducing a stop codon into the pre-S region, whereas S mutants were made by changing the S ATG initiator to ACG; neither mutation altered the product of the overlapping polymerase ORF. Thus, four separate mutants were created: DHBV L$^-$S$^+$, DHBV L$^+$S$^-$, HHBV L$^-$S$^+$, and HHBV L$^+$S$^-$. Each of the four individual mutants was shown to be replication competent in transfected LMH cells, and for each mutant the LMH cell culture medium was noninfectious for PDH (data not shown). This result confirms that both L and S proteins are required for infectivity.

Next we transfected pairwise combinations of these HHBV and DHBV mutants into LMH; the extracellular progeny viruses from each transfection were examined by Southern blotting to assure that comparable titers of particles were used for subsequent manipulations (Fig. 3*A*). These stocks were then tested for PDH infectivity. (Each viral stock resulting from the LMH cotransfection contained both HHBV and DHBV genomes bearing the same complement of envelope proteins, so each PDH infection was scored with both HHBV and DHBV ^{32}P-labeled DNA probes; as expected, in all cases

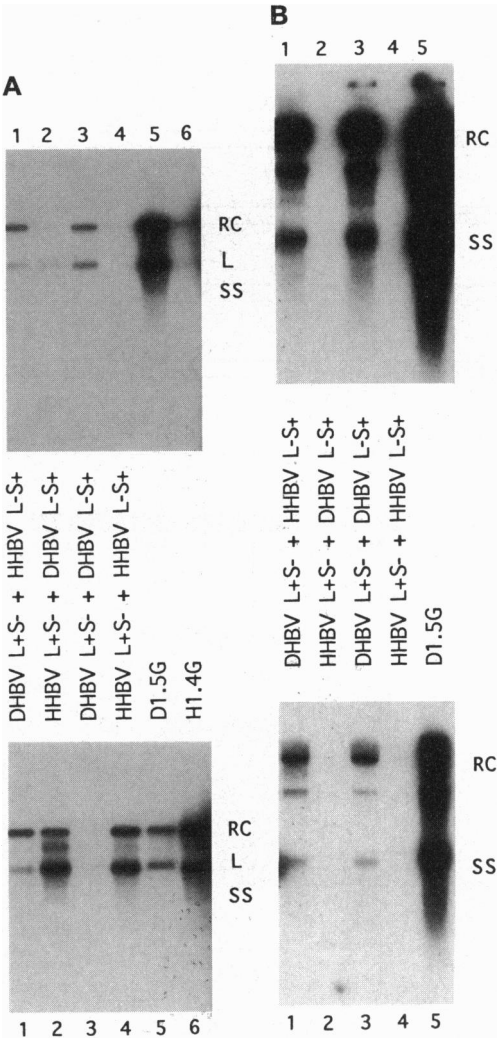

FIG. 3. L protein governs host range. (*A*) LMH cells were transfected with the indicated pairwise combination of HHBV L$^-$S$^+$, HHBV L$^+$S$^-$, DHBV L$^-$S$^+$, and DHBV L$^+$S$^-$ genomes. The extracellular progeny viruses from each transfection were then examined by Southern blotting [DHBV probe (*Upper*); HHBV probe (*Lower*)]. Each viral stock resulting from these LMH cotransfections contains both HHBV and DHBV genomes bearing the same complement of envelope proteins. (*B*) L protein is the prime determinant of species-specificity. LMH culture supernatants from the indicated cotransfections were used to infect PDH cells, and 14 days later the intracellular DNA was examined for viral replicative forms by Southern blotting with DHBV (*Upper*) or HHBV (*Lower*) probes. DHBV L$^-$S$^+$ was efficiently complemented for PDH infectivity by DHBV L$^+$S$^-$ (lane 3) but was not complemented by HHBV L$^+$S$^-$ (lane 2); conversely, HHBV L$^-$S$^+$ was efficiently complemented with DHBV L$^+$S$^-$ (lane 1) but was not complemented by HHBV L$^+$S$^-$ (lane 4).

the results with both probes were concordant.) Fig. 3*B* shows that DHBV L$^-$S$^+$ was efficiently complemented for PDH infectivity by DHBV L$^+$S$^-$ but not by HHBV L$^+$S$^-$; conversely, HHBV L$^-$S$^+$ was efficiently complemented with DHBV L$^+$S$^-$ but not by HHBV L$^+$S$^-$. Thus, only when the L protein was derived from DHBV was efficient infection of duck hepatocytes observed, indicating that L is the prime determinant of species-specificity.

The Pre-S Domain of L Is the Key Host–Range Determinant. The L protein contains both Pre-S and S domains (Fig. 1*A*). In the preceding experiment, both domains of L were derived from the same virus. To more precisely localize the host–range determinant(s) we constructed envelope protein expression vectors that generate chimeric L proteins bearing

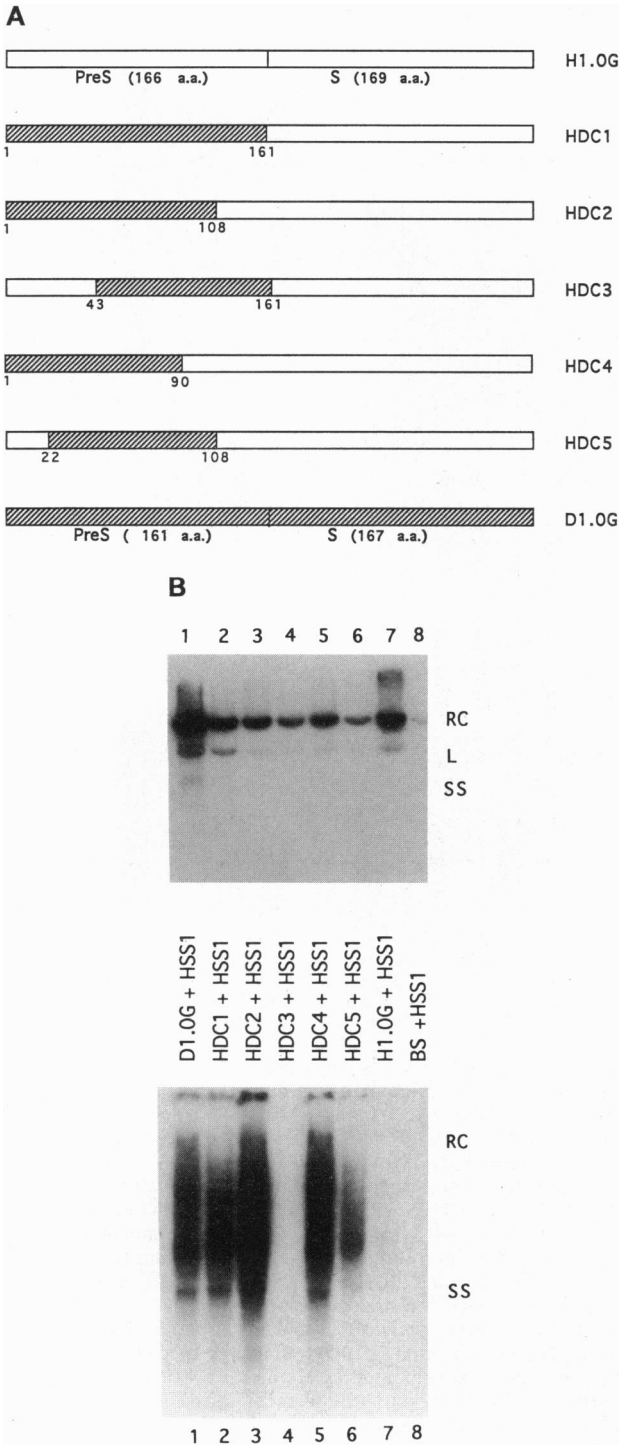

FIG. 4. (*A*) Schematic depiction of the structures of the DHBV/HHBV pre-S chimeras. In these mutants, part of the HHBV pre-S sequences (open box) were substituted with homologous DHBV sequences (hatched box). Each chimera was cloned into the pH1.0G expression vector so as to express wild-type HHBV S and the chimeric pre-S protein. These vectors [or the control Bluescript (BS) plasmid] were then used to complement the HHBV env$^-$ plasmid (HSS1). Numbers below each chimera denote the amino acid position of each end of the substituted region in DHBV pre-S. (*B*) (*Upper*) Assay of transfected cells for release of DNA-containing enveloped virus. Virus concentrated from the indicated LMH cell supernatants was digested with Pronase and then incubated with DNase I at 1 mg/ml; only DNA within enveloped virions is resistant to such treatment (ref. 14; lanes 7 and 8). Resistant DNA was then extracted as described and analyzed by Southern blotting with ^{32}P-labeled HHBV DNA probe. In general, chimeras pseudotyped cores with slightly less efficiency than wild-type envelope proteins. (*Lower*) PDH infectivity of chimeric pseudotyped virions. Equal amounts of enveloped virions from the indicated cotransfected LMH cell supernatant were used to infect PDH cells. Eight days after infection, viral replicative intermediates were analyzed by Southern blotting with HHBV DNA probes. RC, relaxed circular DNA; L, linear DNA; SS, single-stranded DNA.

variable portions of DHBV pre-S replacing the corresponding HHBV pre-S sequences; in all cases the S gene is derived from HHBV (Fig. 4*A*). By cotransfecting LMH cells with these vectors and the HHBV env$^-$ plasmid described earlier, we created pseudotyped virions bearing chimeric L proteins and wild-type HHBV S proteins. Culture supernatants were assayed for enveloped virus as described; Fig. 4*B* (*Upper*) shows that comparable titers of virions were produced by each recombinant. Supernatants containing comparable quantities of HHBV virions were then tested for their ability to infect PDH cells. Fig. 4*B* (*Lower*) shows the results of this analysis. The chimera with the full DHBV pre-S region (HDC 1) is infectious for PDH. Although varying somewhat in their specific infectivity, all recombinants save HDC 3 allowed efficient infection of PDH. Inspection of the structures of the infectious chimeras (Fig. 4*A*) reveals that the critical region for host–range determination can be narrowed to aa 22–90 of pre-S; the amino acid sequence in this region of both viruses is summarized in Fig. 5.

These data indicate that the block to cross-species infection by hepadnaviruses is at the level of viral entry and is governed principally by the pre-S domain of the viral L protein. Formally, the entry block could be at the level of virus binding or at a postbinding, internalization step; our experiments do not distinguish between these alternatives. The region defined by these experiments can be instructively compared with those defined in other studies of DHBV infectivity. Lenhoff and Summers (14) have examined a series of linker-substitution mutations across DHBV pre-S for their effects on virus assembly and infectivity. All lesions between aa 5 and aa 115 were assembly-competent but noninfectious—a region that extensively overlaps that defined herein. It is not surprising that the two studies do not define identical regions: the differences define envelope sequences that are important for infectivity but that do not contribute selectively to species-specificity. That is, whatever function in PDH entry is supplied by these regions can be supplied equally well by either viral sequence. Such a result is not unexpected, given the sequence similarity between DHBV and HHBV in the pre-S region. The host-range-determining region also includes one of the two binding sites for neutralizing monoclonal antibodies to pre-S and overlaps (but is not fully coextensive with) the pre-S region involved in the binding of gp180 (aa 43–108). The latter is a host-cell glycoprotein that has been suggested as a possible component of the entry mechanism (21, 22). The fact that both host range and infectivity determinants include sequences not involved in gp180 binding (aa 20–43) strongly suggests that gp180 cannot be the sole determinant of viral entry, an inference that has recently been directly sustained by cDNA transfection experiments (K. Kuroki, T.I., and D.G., unpublished work).

We emphasize that our results by no means exclude an important role for the S protein in viral entry; such roles could be direct (e.g., to mediate membrane fusion) or indirect (e.g., to allow the proper display of the pre-S sequences). However, our findings indicate that whatever functions they supply in internalization can be supplied equally well across species barriers and thus do not contribute to the species-specificity of entry. Our findings also raise the possibility that the narrow host range of hepadnaviruses could be experimentally modulated by mutational alteration of pre-S sequences. However, in practice such alterations will be severely constrained by the need to preserve the functionality of the overlapping P ORF.

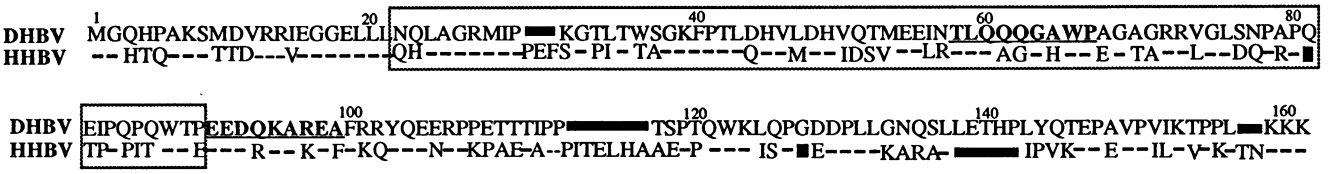

FIG. 5. DHBV/HHBV pre-S sequences. The amino acid sequence of the 69-aa region implicated in host-range determination in avian hepadnaviruses (box) is shown within the context of the whole pre-S region. The corresponding DHBV and HHBV sequences are also shown. Dashes in HHBV sequence denote residues identical to those in corresponding positions of DHBV. Black bars denote positions of deletions required for optimal sequence alignment. Boldface type indicates locations of epitopes of neutralizing monoclonal antibodies to DHBV pre-S (8).

We thank Benson Fenn and Kun-Bee Chang for their skillful assistance in PDH preparation; Hans Will for cloned HHBV DNA; and Kazuyuki Kuroki, Chris Chang, and Dan Loeb for plasmids DHBV L^-S^+, DHBV L^+S^-, and H1.4G, respectively. Special thanks go to Frank Eng for helpful comments on the manuscript. This work was supported by grants from the National Institutes of Health to D.G. and to the University of California at San Francisco Liver Center. D.G. is an Investigator of the Howard Hughes Medical Institute.

1. Ganem, D. & Varmus, H. E. (1987) *Annu. Rev. Biochem.* **56,** 651–694.
2. Baker, L. F., Maynard, L. E., Purcell, R. H., Hoofnagle, J. H., Berquist, K. R. & London, W. T. (1975) *Am. J. Med. Sci.* **270,** 189–196.
3. Condreay, L. D., Cats, C. E., Mason, W. S. & Wu, T. T. (1990) *J. Virol.* **64,** 3249–3258.
4. Hirsch, R., Colgrove, R. & Ganem, D. (1988) *Virology* **167,** 136–142.
5. Galle, P., Schlicht, H.-J., Fisher, M. & Schaller, H. (1988) *J. Virol.* **62,** 1736–1740.
6. Pugh, J., Yaginuma, K., Koike, K. & Summers, J. (1988) *J. Virol.* **62,** 3513–3516.
7. Sprengel, R. C., Kuhn, H., Will, H. & Schaller, H. (1985) *J. Med. Virol.* **15,** 323–333.
8. Cheung, R. C., Trujillo, D. E., Robinson, W. S., Greenberg, P. L. & Marion, P. L. (1990) *Virology* **176,** 546–552.
9. Lambert, V., Fernhorz, D., Sprengel, L., Fourel, I., Deleage, G., Winder, G., Peyret, C., Trepo, C., Cova, L. & Will, H. (1990) *J. Virol.* **64,** 1290–1297.
10. Neurath, A. R., Kent, S. B. H., Strick, N. & Parker, K. (1986) *Cell* **46,** 429–436.
11. Sprengel, R., Kalata, E. F. & Will, H. (1988) *J. Virol.* **62,** 3832–3839.
12. Kunkel, T. A., Roberts, J. D. & Zabour, R. A. (1987) *Methods Enzymol.* **154,** 367–382.
13. Deminie, C. A. & Emerman, M. (1993) *J. Virol.* **67,** 6499–6506.
14. Lenhoff, R. J. & Summers, J. W. (1994) *J. Virol.* **68,** 4565–4571.
15. Staprans, S., Loeb, D. & Ganem, D. (1991) *J. Virol.* **65,** 1255–1262.
16. Maniatis, T., Fritsch, E. F. & Sambrook, J. (1982) *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Lab. Press, Plainview, NY).
17. Church, G. M. & Gilbert, W. (1984) *Proc. Natl. Acad. Sci. USA* **81,** 1991–1995.
18. Pugh, J. C. & Summers, J. W. (1989) *Virology* **172,** 564–572.
19. Tuttleman, J. S., Pugh, J. C. & Summers, J. W. (1985) *J. Virol.* **58,** 17–25.
20. Pugh, J. & Simmons, S. H. (1994) *J. Virol.* **68,** 2487–2494.
21. Kuroki, K., Cheung, R., Marion, P. & Ganem, D. (1994) *J. Virol.* **68,** 2091–2096.
22. Ishikawa, T., Kuroki, K., Lenhoff, R., Summers, J. & Ganem, D. (1994) *Virology* **202,** 1061–1064.