UCSF
UC San Francisco Electronic Theses and Dissertations

Title
Studies on hepadnaviral core protein and pregnome RNA

Permalink
https://escholarship.org/uc/item/62z17578

Author
Chang, Christopher,

Publication Date
1994

Peer reviewed|Thesis/dissertation

eScholarship.org Powered by the California Digital Library
University of California

Studies on Hepadnaviral Core Protein and Pregenome RNA

by

Christopher Chang

DISSERTATION

Submitted in partial satisfaction of the requirements for the degree of

DOCTOR OF PHILOSOPHY

in

Biochemistry

in the

GRADUATE DIVISION

of the

UNIVERSITY OF CALIFORNIA

San Francisco

[redacted]

ACKNOWLEDGEMENTS

I was warned that this is the most read section of a thesis, perhaps even more closely examined that the abstract. Unfortunately, I am devoting the least time, proportional to its importance, to this than to any other section. First and foremost, I must express my gratitude to my thesis advisor, Don Ganem. From the time I first began as a rotation student in med school, he has most influenced my scientific development and set the standard for how one can balance the roles of a research scientist with that of a hands-on clinician.. I am indebted to Don for his creative insights, guidance, and undiluted support throughout my graduate school years. I also want to thank the other members of my thesis committee, Judy White and Dan Littman, for their advice and support during the past two years. David Standring and Siliang Zhou have been superb collaborators this past year and really helped make the assembly story happen.

One of the features of the HSE 4th floor that attracted me to the Ganem lab was the outstanding environment in which to do science. I am indebted to countless members, past and present, of the Ganem and Varmus labs who have helped me along the way. I especially wanted to thank several former post-docs, Lung-ji Chang, Dan Loeb, and Joel Lavine for their many helpful discussions. The current crew working on our floor have been especially generous with their ideas and insights. I particularly want to mention Jon Pollack, Sophie Roy, Frank Eng, Steve Chuck, and Ian Taylor. My baymates in 419 deserve special recognition for their tolerance of my sarcastic asides, over-the-top comments, and SI calendar. Buckwheat (JP), Dean, Shiow-her, Steve, Don, and Sophie--you guys made the day seem a lot less longer. I also want to thank all my friends outside the lab; without them, I probably would have gone insane or at least picked up a lot of gray hairs. I can't even begin to make a list, but you know who you are.

Finally, I want to thank my parents for their continued support and confidence in me to finish up this phase of my education. It's been a long haul and it isn't over yet, but I will always be grateful for their unwavering encouragement.

ABSTRACT

Hepadnaviruses are a family of small, hepatotropic DNA viruses that can cause persistent infections and liver disease, and are associated with the development of hepatocellular carcinoma. The core protein plays a major role in viral morphogenesis. Core polypeptides, the polymerase, and the pregenome RNA assemble to form the nucleocapsid during viral replication in hepatocytes.

The core gene of all hepadnaviruses is preceded by a short, in-phase open reading frame termed precore. To explore the functional importance of precore expression *in vivo*, I introduced a frameshift mutation into this region of the duck hepatitis B virus (DHBV) genome and demonstrated that the mutant DNA led to the formation of infectious progeny virus. This finding indicated that expression of the precore region and its secretory product, termed e antigen, is not essential for genome replication, viral morphogenesis, or intrahepatic spread.

I next characterized core protein interactions involved in capsid assembly. I employed a two-hybrid strategy to demonstrate hetero-oligomerization among the core proteins from the mammalian hepatitis B virus (HBV) core proteins. Consistent with this, when expressed in *Xenopus* oocytes the mammalian HBV core proteins, but not DHBV core proteins, efficiently co-assembled to form mixed capsids. Surprisingly, however, when two different core mRNAs were translated in the oocyte, the core monomers showed a marked preference for forming homodimers rather than heterodimers. This suggests that core monomers are not free to diffuse and associate with other monomers. Thus, mixed capsids result from aggregation of different species of homodimers.

Finally, I identified sequences which affect the steady state level of the multi-functional pregenomic RNA of DHBV. Through deletion analysis I mapped a small region (termed α), located near the pregenome cap site, whose presence is critical for accumulation of pregenomic RNA in transfected cells. The drastically reduced RNA levels resulting from α-deletions are suppressed by deletion of a second region located 1 kb. downstream. I further demonstrated that the activity of the α-element displays strong orientation-dependence and limited position-dependence. The implications of these findings for current models of post transcriptional regulation of pregenomic RNA are discussed.

Don Ganem

Don Ganem, Committee Chair

TABLE OF CONTENTS

Chapter 1

Introduction 1

Chapter 2

Expression of the precore region of an avian hepatitis B virus is not required for viral replication 29

Chapter 3

Phenotypic mixing of hepadnaviral nucleocapsid proteins reveals particles to be composed of homodimer precursors.......................................37

Chapter 4

Identification of sequences which affect pregenomic RNA accumulation of the duck hepatitis B virus..75

LIST OF TABLES

Chapter 3

Table 1. Full length HBV core protein oligomerizes............................56
Table 2. C-terminal and N-terminal deletions in HBV core protein 58
Table 3. Interactions among mammalian HBV core proteins..................60

LIST OF FIGURES

Chapter 1

Fig. 1. HBV genome organization 19

Fig. 2. The hepadnaviral life cycle 21

Fig. 3. Hepadnaviral DNA synthesis 23

Chapter 2

Fig. 1. Construction of precore frameshift mutation in DHBV DNA.........31

Fig. 2. Viral DNA forms in the livers of ducks transfected with mutant genomes..32

Fig. 3. Nucleotide sequence of the precore region of cloned DHBV DNA recovered from the liver of a mutant-infected duck............32

Chapter 3

Fig. 1. Analysis of cross reactivity with the anti-L mAb and the anti-HBV-C polyclonal Ab..62

Fig. 2. Assembly of a tagged and truncated HBV core protein in *Xenopus* oocytes..64

Fig. 3. Composition of capsid particles assembled in *Xenopus* oocytes co-injected with mRNA encoding core proteins from different hepadnaviruses (A)..66

Fig. 3. Analysis of mixed extracts (B)..68

Fig. 4. Composition of dimer intermediates assembled in *Xenopus* oocytes co-injected with mRNA encoding core proteins from different hepadnaviruses..70

Fig. 5. Composition of dimer intermediates formed in *Xenopus* oocytes co-expressing tagged and untagged HBV-C....................72

Chapter 4

Fig. 1. Genetic and transcriptional map of DHBV 88

Fig. 2. Northern blot analysis of mutants bearing deletions in the 5' region of pregenomic RNA................................90

Fig. 3. Poly $(A)^+$ levels of transcripts that initiate upstream of the PreS mRNA cap site....................................92

Fig. 4. Northern blot analysis of mutants bearing a deletion in the α region and an additional downstream internal deletion.....94

Fig. 5. Northern blot analysis examining position and orientation-dependence of α sequences..........................96

CHAPTER 1

THE HEPADNAVIRUS FAMILY

CLINICAL FEATURES

The human hepatitis B virus (HBV) is the prototype member of the hepadnavirus (for *hepa*totropic *DNA virus*) family, a group of small, enveloped DNA animal viruses which share common structural and biological properties (reviewed in 22). HBV infection is characterized by two cardinal features: species specificity and relative hepatotropism. Productive infections are established only in human beings and higher primates. In these hosts, viral antigens, DNA, and mature virions are found primarily in the liver. Infection is spread either horizontally (by contact with body fluids, chiefly through sexual contact) or vertically (from carrier mothers to newborns). Primary infection is usually self-limited, with the transient appearance of viral antigens and particles and their subsequent clearance from liver and blood, and results in the development of lasting immunity against future infection. However, 5-10% of primary infections fail to resolve and develop into a persistent, usually lifelong, chronic infection (21). In these chronic carriers, virions and viral antigens generally persist in the blood, and viral DNA replicative intermediates persist in the liver (17,56). Chronic infection may be asymptomatic or characterized by severe liver damage. At present, the mechanisms of hepatocellular injury and the factors modulating its severity are not well understood. Viral replication itself is not directly cytopathic; rather, the host cellular immune response has been implicated as a major cause of liver disease (8).

The clinical implications of chronic infection are significant. The over 300 million chronic HBV carriers worldwide represent the reservoir from which HBV infection is spread (65). Persistent infection and liver cirrhosis can lead to liver failure and death, a rare outcome from acute infection. Also of great significance is the association between chronic HBV infection and the development of primary hepatocellular carcinoma (HCC). Long-term carriers have a greater than 100-fold risk of developing HCC compared to agematched noncarriers (6). A molecular level mechanism to explain this association remains elusive thus far.

At present, a vaccine of recombinant HBV envelope glycoprotein is available in some countries to prevent new HBV infections. Unfortunately, effective immunization programs in Africa and Asia, where infection levels are highest, are not yet realized. Moreover, no cure exists for the millions presently infected with HBV. Future antiviral therapies will depend on a detailed understanding of the molecular mechanisms underlying the hepadnaviral life cycle.

In addition to HBV, the other identified and characterized members of the hepadnaviral family include the woodchuck hepatitis virus (WHV; 63), the ground squirrel hepatitis virus (GSHV; 34), the duck hepatitis B virus (DHBV; 35), and the heron hepatitis B virus (HHBV; 58). These viruses share a similar viral morphology, genome organization, replication strategy, and cellular life cycle. Consequently, they have been studied extensively as models for hepadnaviral biology and disease pathology.

VIRAL STRUCTURE

The HBV virion (also called the Dane particle, after its discoverer) is a 42 nanometer double-shelled particle, consisting of a 28 nm nucleocapsid surrounded by a 7 nm outer envelope of lipoprotein. In addition to the Dane particles, infected human sera also contain 22 nm spheres and filaments in great excess over virions; these structures are composed of excess viral envelope material (16). The viral envelope consists of three related viral glycoproteins termed L (large), M (middle), and S (small) in a ratio of 1:1:4 in the virion (28). The viral nucleocapsid is a T=3 icosahedral structure composed of 180 subunits of a single polypeptide, the core protein (23). Within the nucleocapsid is the partially duplex, relaxed circular viral genome, and the viral polymerase which mediates its replication. The HBV genome, at 3.2 kilobases (kb), is the smallest of any known animal DNA virus (Fig. 1). It consists of a full-length minus stand and a shorter complementary strand with a fixed 5' end and a 3' end of variable length, yielding a single-stranded gap of fixed polarity. Circularity is maintained by overlapping 5' ends of 224 bp. (62). The polymerase protein is covalently linked at the 5' end of minus strand (3,25), whereas an oligoribonucleotide is attached to the 5' terminus of plus strand (32). These asymmetries in genome strucure result from the novel reverse transcription pathway of hepadnaviruses.

GENOME ORGANIZATION

Cloning and sequencing of the hepadnaviral genomes reveals a compact and overlapping open reading frame (ORF) organization. All genomes contain three ORFs encoded by the minus strand (Fig. 1). ORF S and its upstream in-phase reading frames PreS1 and PreS2 direct the synthesis of envelope glycoproteins L, M and S. ORF C and its short upstream ORF PreC encode core, the capsid protein, and precore, a secreted variant of core which is processed into the e-antigen detected in sera. Overlapping these two ORFs is the large ORF P, which encodes the viral polymerase (P). ORF X, found only in the mammalian hepadnaviral genomes, codes for a viral transcriptional transactivator (15,57).

REPLICATION

In an elegant series of experiments published over ten years ago, Summers and Mason demonstrated that replication of hepadnaviruses, unlike all other animal DNA viruses, involves the reverse transcription of an RNA intermediate by a virus-encoded polymerase (64). This strategy is similar in several aspects with the replication schemes of RNA-containing retroviruses and cauliflower mosaic virus, a plant DNA virus (22).

Replication of the genome involves several steps (Figs. 2 and 3). Following viral entry into the hepatocyte, the viral genome is delivered to the nucleus, where it is converted to a covalently closed circular DNA (cccDNA), perhaps by host enzymes. In the nucleus, the cccDNA template is transcribed by host RNA polymerase II, yielding both genomic and subgenomic mRNAs. The shortest terminally redundant genomic RNA serves as the intermediate which templates reverse transcription (11,18). This pregenomic RNA is specifically encapsidated into cytoplasmic subviral core particles where reverse transcription and DNA synthesis occur.

Primed by P protein (3,67), minus strand synthesis begins with P protein binding to the 5' packaging signal ε and synthesis of a short primer, followed by translocation to the 3' DR1 (66,68). Direct repeats DR1 and DR2 serve as critical cis-signals for the initiation of minus and plus strand DNA synthesis, respectively (32,55,70). Reverse transcription from the 3' DR1 of pregenome generates a terminally redundant minus strand, while at the same time the RNA template is degraded by the RNase H domain of P protein. A short capped oligoribonucleotide derived from the 5' end of pregenome by RNase H is transposed to DR2 of the minus strand and primes synthesis of the plus strand (32,33). Utilizing the terminal redundancy of the minus strand, an intramolecular template switch occurs during elongation of the plus strand, resulting in the circular conformation of the hepadnaviral genome.

The mature core particle containing the DNA genome associates with viral envelope glycoproteins within the endoplasmic reticulum membrane by budding into the lumen (40). The mature virion is then secreted from the cell through the ER and Golgi apparatus pathway. Virus release thus occurs without cell lysis.

TRANSCRIPTION

Two highly abundant classes of transcripts are produced by the host transcription machinery in the acutely infected liver (11,18). These are the terminally redundant genomic transcripts (3.5 kb) and the subgenomic transcripts (2.4 and 2.1 kb). Both code for the major structural components of the virion and are unspliced. Genomic RNAs comprise the message for proteins required to form the nucleocapsid (core and P protein) and the RNA pregenome as well. Subgenomic RNAs serve as mRNAs for the expression of the envelope glycoproteins. In addition, a minor 0.7 kb message has been identified in HBV and correlates with synthesis of the X ORF gene product.

The heterogeneity of 5' start sites displayed by both classes of RNAs enables the expression of the in-phase upstream PreC, PreS1 and PreS2 ORFs. Thus, the longer genomic RNAs code for precore protein while the shorter genomic message encodes the core protein. Another feature is the common site of polyadenylation used by all hepadnaviral transcripts. The terminal redundancy of genomic transcripts arises from the initial bypassing of the polyadenylation signal encountered immediately after transcription initiation (51).

RNA PACKAGING

The HBV packaging signal, termed ε, encompasses a 94 nucleotide segment at the 5' end of pregenome (31,45). When fused to the 5' end of a heterologous transcript, ε directs its packaging into HBV capsids (30,31). Unlike HBV ε, the DHBV packaging signal appears to encompass over 1 kb of the 5' end of pregenome (30); further analysis has defined the packaging signal to the 5' end and a second region approximately 1 kb downstream (12). The ε RNA sequence forms a conserved stem-loop secondary structure both *in vitro* and *in vivo*; mutational analysis has shown that this stem-loop structure is critical for RNA encapsidation (45).

The P protein requirement for RNA packaging was demonstrated by the inability of a series of P gene nonsense mutants to encapsidate pregenomic RNA efficiently; thus, core polypeptides alone are insufficient for this step. However, packaging does not require the enzymatic activity of either the reverse transcriptase or RNase H domains of P protein (4,29). Another feature is the marked cis-preference of P protein to direct encapsidation of the pregenome message from which it has been translated (4,29). Two mechanisms have been put forth: that the encapsidation reaction mediated by P protein occurs co-translationally, and/or that P protein is expressed in limited amounts and preferentially interacts with the nearest ε sequence, i.e. the RNA from which it was translated. Recently, Pollack and Ganem have demonstrated that DHBV P protein binds to its ε stem-loop with high affinity, and that binding is necessary (but not sufficient) for packaging and DNA priming (46).

VIRAL CAPSID PROTEINS AND CAPSID ASSEMBLY

This section will briefly survey the capsid proteins and capsid assembly pathways of several viruses. These examples serve to illustrate the diversity of pathways which have evolved to construct a nucleocapsid, as well as to point out important common features.

PICORNAVIRUS

Picornaviruses are a family of small, icosahedral, non-enveloped positive-strand RNA viruses with a well-studied structure and assembly pathway. A common theme has emerged from the determination of high-resolution structures for these and a number of other positive-strand RNA viruses. The capsids are composed of 180 major "building blocks" packed in an icosahedral array. The blocks are protein domains based on an eight-strand, anti-parallel β-barrel, called an RVC domain (for *R*NA *v*irus *c*apsid). The topology of the polypeptide is common to many capsid proteins, although the particular size, shape, and disposition of the loops between the β-strands are specific for a particular virus. The 60-fold character of icosahedral symmetry implies that 180 RVC domains must be grouped into 60 identical sets. In many icosahedral virsues, capsids are composed of only one monomer subunit; packing of the subunits allows for quasi-equivalent interactions at corresponding interfaces (reviewed in 27).

The case for picornaviruses is slightly more complex. Each set is composed of three similar capsid proteins (VP0, VP1, and VP3) cleaved from a polyprotein precursor. In the assembly pathway, the polyprotein undergoes several cleavages and rearrangements to form the protomer. Latent assembly domains encoded in the protomer drive assembly and are probably controlled by configurational changes following each step (50). Five protomers associate around a five-fold axis to form a pentamer, with contacts made by the N-terminal arms of each protomer's three capsid proteins with the arms of its neighboring protomer. Twelve pentamers are then assembled into a provirion with the RNA genome. In the final step when mature virus is formed and acquires infectivity, the VP0 chains are cleaved into VP2 and VP4. Thus, picornavirus assembly proceeds in an ordered pathway involving stable intermediates.

In vitro assembly studies have shown that poliovirus-infected cell extracts possess the striking ability to promote the formation of natural empty capsids, RNA-free viral shells, from pentamers (43). Purified pentamers will self-assemble *in vitro*, but only at very high concentration. On the other hand, pentamer assembly occurs much more efficiently in cell-free extracts prepared from infected cells, even at low pentamer concentrations. Efforts to identify the assembly catalyst suggest that it is a form of the capsid protein which is bound to the rough endoplasmic reticulum (44). Despite progress made in outlining the assembly pathway, the precise mechanism underlying many steps, especially RNA packaging, remain tobe elucidated.

TOMATO BUSHY STUNT VIRUS AND COLIPHAGE R17

Although also a positive-strand RNA virus, tomato bushy stunt virus (TBSV) uses fewer components and a simpler pathway to form its capsid than picornaviruses (reviewed in 27). Its capsid monomer is a single 40 kd modular protein consisting of a central RVC domain, a C-terminal domain, a positively charged domain for neutralizing RNA, and a connecting arm. Thus unlike picornaviruses, all 180 capsid monomers are chemically identical. In addition, the virus assembles from dimer subunits rather than from larger intermediates. The complete T=3 particle can be described as an assembly of 60 A/B dimers and 30 C/C dimers, in which A, B, and C represent the location of the three polypeptides that compose a "protomer-like" identical set. Coordination of assembly involves sequential formation of an internal framework by the connecting arms of the 60 subunits in the C location; the framework in turn determines whether a dimer adopts an A/B conformation or a C/C conformation.

Like TBSV, the spherical (T=3) RNA coliphage R17 also assembles from building blocks composed of dimers of a single coat protein. In addition to its structural role, the coat protein also acts as a translational repressor of the phage replicase gene. Interestingly, the coat protein bound to a translational operator fragment is a much more active species in capsid assembly than free dimer (7). Thus, the RNA fragment does not trigger protein assembly; rather, coat protein bound to the operator fragment has greater affinity for coat protein dimers than does free coat protein. This mechanism couples the capsid formation process to the association of coat protein with RNA. Intermediates larger than dimers have not been detected for R17 assembly.

POLYOMAVIRUS

Polyomavirus is a DNA virus with a capsid design that does not fit squarely with icosahedral design. Its morphology seemed to demand an arrangment of 12 pentameric and 60 hexameric capsomeres arranged in a T=7 icosahedral surface (13). Yet, the 360 copies of its major capsid protein VP1 are assembled into 72 indistinguishable pentamers. Minor internal proteins VP2 and VP3 are also present. Given this arrangement, the subunits of the 60 hexavalent pentamers cannot form similar bonds with the subunits of its six neighboring pentamers. Therefore, nonequivalent bonding must occur among molecules of VP1 in the assembled polyomavirus capsid.

How is the bonding specificity of VP1 switched to fit into the symmetrically distinct environments in the all-pentamer capsid? Both post-translational charge modifications and interactions with VP2 and VP3 (30-60 copies each per capsid) have been proposed to regulate the VP1 bond specificity required for assembly. Interestingly, neither mechanism appears to play an essential role. Salunke and co-workers have purified VP1 expressed in *E. coli* and found that the unmodified polypeptide forms pentamers and capsid-like structures (52). Although these structures are more heterogenous in size than purified native capsids, suggesting that additional factors contribute to the accuracy of virus assembly, this self-assembly result clearly demonstrates that the potential to switch among nonequivalent bonding states is intrinsic to VP1. Since assembly proceeds spontaneously, this switching presumably is controlled by the interactions of the pentamers being added with those already assembled.

While not required for capsid formation, VP2 and VP3 are believed to help link VP1 pentamers and the viral minichromosome. In addition, VP1 in virions is acetylated and phosphorylated. Phosphorylation is essential for assembly of stable virions but not for pentamer formation (24), suggesting it could help lock in the correctly assembled structure. No stable intermediates other than VP1 pentamers have been identified, and a detailed assembly pathway has not been elucidated yet.

RETROVIRUSES

Production of retroviral core capsids occurs randomly in the cytoplasm for some morphologic types of retroviruses (type B and D). The most common type C particles mature by co-assembly of core and envelope at the plasma membrane during viral budding. In all cases, the internal structure of the particle rearranges after budding, presumably reflecting cleavage of the gag protein precursor. Surprisingly few molecular details of capsid structure or assembly are understood for any of the retroviruses; most studies have involved mutational analysis of components of the budding process.

Only the gag precursor is needed to form retroviral capsids. Assembly probably involves interaction of the three principal domains of gag: at its C-terminal nucleocapsid (NC) end with the genome; in the middle capsid (CA) with the CA region of other precursor molecules; and at the N-terminal matrix (MA) end with the cell membrane (ref

33). Deletions in CA, which form the capsid shell, prevent any particle formation (54). MA is the only other domain required for enveloped particle production. Indeed, particles form in the absence of genomes, pol, env, pro, and the NC domain. Although not essential for capsid assembly, MA mediates the membrane association step in the budding process, with its hydrophobic myristate chain playing a critical role (48). In the absence of env proteins, cores nonetheless bud efficiently; these particles are noninfectious, but can initiate infection if fused with cell membranes (69). Moreover, deletion of the short cytoplasmic domain of the transmembrane (TM) protein of env does not affect the production of wild type levels of infectious virions containing env protein (41). As this section illustrates, many questions concerning capsid assembly and virus production remain unanswered for retroviruses.

HBV CORE PROTEIN AND CAPSID ASSEMBLY

ORF C encodes the core polypeptides. Its 5' end contains two conserved in-frame AUG codons separated by 28 codons in HBV. The major capsid component, core (C) protein, originates from the second of these AUGs. Translational initiation from both AUGs is accomplished by heterogenous transcription initiation of the overlength genomic RNAs. The most abundant, and shortest of the three messages initiates between the two AUGs (11,18). Expression from this second initiator produces the cytoplasmic 21 kD C protein. The C protein participates in many steps of the viral life cycle and interacts with multiple components in the production of an infectious virus particle.

PROPOSED STRUCTURE

At present, models of C protein and capsid structure based on crystallographic data have not been published. Electron micrographs of HBV nucleocapsids depict a uniform spherical structure of 28 nm diameter (38). Analysis of the micrographs suggest that 180 C polypeptides are regularly arranged on an icosahedral surface with T=3 symmetry. The C proteins do not appear to cluster into distinctive pentameric or hexameric capsomere structures.

Analysis of the 183 residue (subtype *ady*) C protein reveals two distinct regions. The first 149 residues display an unusually high content of hydrophobic amino acids. The remaining C-terminal domain is composed of an extremely basic tail region. Argos and Fuller have produced a topographic model of C protein based on the known three-dimensional structure of a picornavirus capsid protein (2). By aligning short segments of high amino acid similarity, they predict a common, conserved eight-stranded, anti-parallel β-barrel structure resembling the RVC domain found in many positive-strand RNA viruses. Although the accuracy of this theoretical model is uncertain and its resolution low, the similarity of C protein structure to other positive-strand RNA virus capsid proteins is attractive since C protein must also form a nucleocapsid around the positive-strand pregenome RNA during the packaging process. The putative β-barrel domain appears to be very tightly folded, given its extreme resistance to proteolysis, while the basic tail is very sensitive to proteases (60).

The extremely basic C-tail is composed of arginine residues (17 of 36 residues in *adw*) arranged in four clusters and bears strong sequence resemblance to protamine-like molecules. Its composition suggests that it would display nonspecific nucleic acid binding activity. This property and its role in core phosphorylation and assembly will be discussed in subsequent sections.

CORE-NUCLEIC ACID INTERACTIONS

Nucleic acid binding activity has been demonstrated using either natural or recombinant C protein and labelled RNA and DNA (20,42). The assays do not suggest any nucleic acid sequence specificity in the binding reaction. Removal of the C-terminal tail abolished this binding activity on southwestern and northwestern blots (20). Capsid particles assembled in *E. coli* from full-length C protein contain RNA; deletion of the arg-rich tail significantly diminishes the RNA content within the capsids (9).

The effects of core protein truncation mutants on viral replication has been studied in HBV- and DHBV-transfected hepatoma cell lines that are permissive for viral growth. In DHBV, removal of the C-terminal 12 residues from the full-length 262 amino acid C protein affected neither genome replication nor virus formation, although the resulting particles were noninfectious (53). Deletion of 10 C-terminal residues from HBV C protein had no observable effect on any measurable viral properties (37), but removal of another 2 residues led to incomplete plus strand synthesis (5). On the other hand, deletion of 19 residues severely reduced plus strand synthesis without affecting RNA packaging (37). Removal of the C-terminal 36 amino acids from DHBV C protein blocked DNA replication but also did not abolish RNA encapsidation (53).

CORE PROTEIN PHOSPHORYLATION

Phosphorylation is the only known modification of C protein. When partially purified core particles are incubated with gamma-[^{32}P]ATP, C protein is phosphorylated on serine residues, although only a few of the 180 subunits appear to be targets for the *in vitro* kinase activity (1,19,26). Both HBV and DHBV C protein purified from infected livers or tissue culture are also phosphorylated (47,49,53). In all cases, these phosphorylations are on serines and threonines within the tail region (53). What is the origin of the core kinase activity? Current evidence favors a host cellular enzyme rather than an intrinsic kinase activity of C protein. First, the predicted amino acid sequence of C protein does not contain homologies to any known serine/threonine kinases. Moreover, HBV and DHBV C protein expressed in *E. coli* is unphosphorylated, while it is phosphorylated when expressed in higher eukaryotic cells (49,53).

The biological function of core phosphorylation remains uncertain despite its conservation among all of the hepadnaviruses. The C protein mutants with deletions in the C-terminal tail, described in the previous section, have some or all of their phosphorylation target residues removed. However, their altered packaging or replication phenotypes may be caused by the absence of specific C-terminal residues rather than the lack of phosphorylation. Pugh, et.al. noted that cores from serum-derived virions were underphosphorylated relative to cytoplasmic cores isolated from infected cells (47). This observation raises the possibility that core dephosphorylation may be the signal for envelopment or export of mature cores.

CORE CAPSID ASSEMBLY

Several laboratories have demonstrated that core protein expressed in the absence of all other viral products is able to spontaneously assemble into capsid particles. Recombinant vectors expressing C protein in hosts as divergent as bacteria, yeast, insect cells, or *Xenopus* oocytes produce capsids which resemble cores isolated from infected livers (5,9,36,71). These observations do not support any role for other viral proteins in assembly, and rule out the requirement for pregenome RNA to nucleate proper capsid formation, as is needed for R17 phage (7). The pathway that assembles 180 core polypeptide monomers into a capsid particle remains unresolved. At present, core polypeptide dimers are the only stable intermediate identified in the xenopus oocyte assembly system (72). Standring and co-workers suggest that the highly cooperative manner of capsid formation may preclude detection of larger assembly intermediates. In this respect, HBV capsid formation more closely resembles features of TBSV or R17 phage assembly, rather than pathways involving stable oligomer intermediates that are used by picornaviruses or polyomavirus.

Efforts to identify regions of C protein that are critical for capsid formation have focused on deletions in the arginine-rich C-terminal tail. In the various heterologous systems expressing C protein, deletions that contain the N-terminal 144 residues produce stable capsids, arguing that the C-terminal tail is not essential for assembly (5,37). Deletions that extended further upstream aborted particle formation. Moreover, assembly can still proceed when heterologous sequences are placed at the N-terminus, C-terminus, and within an immunodominant loop of C protein near residue number 80 (10,14,59). Taken together, these findings suggest that the putative RVC domain may be critical for particle assembly. Although dispensible for particle formation, the arg-rich tail may in fact play a structural role in the wild-type nucleocapsid and contribute to its stability. Capsids composed of tail-less monomers are more easily dissociated by SDS, urea or high pH than wild type capsids (9,20). Standring envisages a model in which C-protein dimers linked by a disulfide bond(s) in the β-barrel domain associate by noncovalent interactions to form a capsid; these dimers are then linked to each other via disulfide bridges involving the C-terminal tail, creating a more stable, covalently closed particle (61).

PRECORE PROTEIN AND E-ANTIGEN

As stated in an earlier section, the C protein AUG is preceeded by an in-frame AUG that is 29 codons upstream; this in-phase ORF is termed PreC. Using the longer genomic transcripts that initiate upstream of pregenomic RNA, translation of the PreC ORF yields the precore protein. The addition of the precore region confers a very different fate to this polypeptide. The 29 amino acid precore region acts as a signal sequence to direct precore protein translocation into the endoplasmic reticulum. During translocation across the ER, the N-terminal 19 residue signal peptide is cleaved; the polypeptide then travels through the secretory pathway, where the C-terminal tail is cleaved at Val-149. The resultant 17 kD protein, now termed e antigen, is secreted into the medium (39,60). In HBV-infected patients, e antigen can be detected in the circulation, where it is correlated with elevated titres of Dane particles (21). Expression of the precore/core sequence in different expression systems clearly demonstrates that the immunologically related e antigen is not a breakdown product of intact HBV core particles or a denatured form of C protein. Instead, e antigen biogenesis is actually dependent on a distinct mechanism involving targeting of the precore protein to the ER by the precore signal sequence.

Several early observations were taken to suggest an essential function for precore proteins in the viral life cycle and/or viral infectivity. First, serum e antigen levels correlate with viral infectivity. Second, the precore ORF is conserved throughout hepadnavirus evolution, implying strong selection pressure to maintain its function.

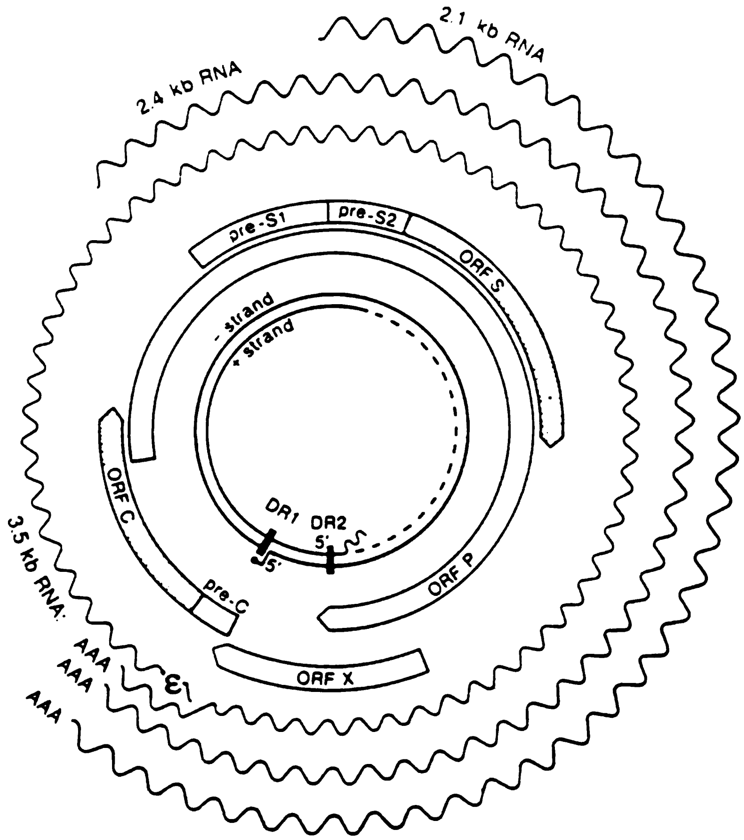
SUMMARY OF TOPICS ADDRESSED IN THE THESIS

When I began in the Ganem lab as a rotation student, I set out to determine the functional significance of precore expression *in vivo*. I engineered a frameshift mutation in the precore region of DHBV and examined its phenotype by intrahepatic inoculation into newborn ducks. I then assayed for DHBV infection to determine whether expression of precore was critical for viral replication or spread. The results of these experiments were reported in the Journal of Virology and included in Chapter 2.

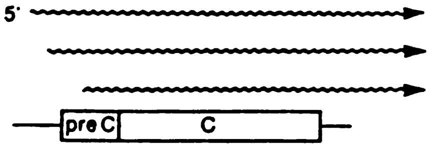
Next, I attempted to further characterize the interactions involved in capsid assembly. A two-hybrid approach was used in mammalian tissue culture to try to define regions in the core protein essential for core-core interactions. I also employed the two-hybrid assay to detect interactions among core proteins of the different hepadnaviral family members. Collaborating with the laboratory of David Standring, I assayed for the ability of cores from different species to form heterocapsids in the *Xenopus* oocyte capsid assembly system. The outcome of these experiments is reported in Chapter 3.

I also made extensive efforts to develop DHBV as a genetic vector for the delivery of foreign sequences to hepatocytes. These efforts were unsuccessful, largely because of the inability of DHBV packaging constructs to donate P *in trans* to replicate the vectors bearing the foreign gene. But in the course of this work, I noted that deletion mutants that removed sequence near the 5' end of pregenome RNA greatly reduced its steady state level in transiently transfected hepatoma cells. I more precisely defined the sequences critical to maintain pregenome levels, as well as the effects of deletions at a secondary site within the genome. I also examined the influence of relocating this 5' sequence to other sites in the genome on the level of pregenome RNA. A summary of these results is presented in Chapter 4.

FIG. 1 HBV genome organization (Top) The innermost circle represents viral DNA; shown are the direct repeats (DR1 and DR2), and the genome-linked P protein and capped oligoribonucleotide on minus and plus strands, respectively. The open reading frames encoded by the genome are indicated by the open boxes. The genomic (3.5 kb.) and subgenomic (2.1 and 2.4 kb.) viral transcripts are depicted as wavy lines; the encapsidation signal (ε) is shown. (Bottom) Structure of the 5' ends of viral transcripts. The 3.5 kb. genomic RNAs initiate either upstream or downstream of the translation start site for Pre C, thereby encoding either precore or core proteins. Precore protein contains all the amino acids of core protein plus an N-terminal extension. The analogous situation is diagrammed for the 2.1 kb. subgenomic RNAs. Reproduced, with permission (22).


2.1 kb RNA
2.4 kb RNA
pre-S1
pre-S2
ORF S
- strand
+ strand
DR1
DR2
5'
5'
ORF C
pre-C
ORF P
ORF X
3.5 kb RNA
AAA
AAA
AAA
ε

3.5 kb RNA:



2.1 kb RNA:

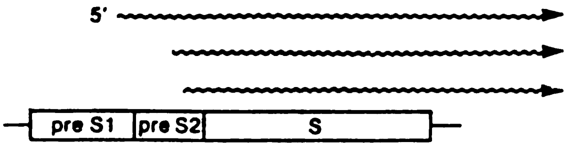
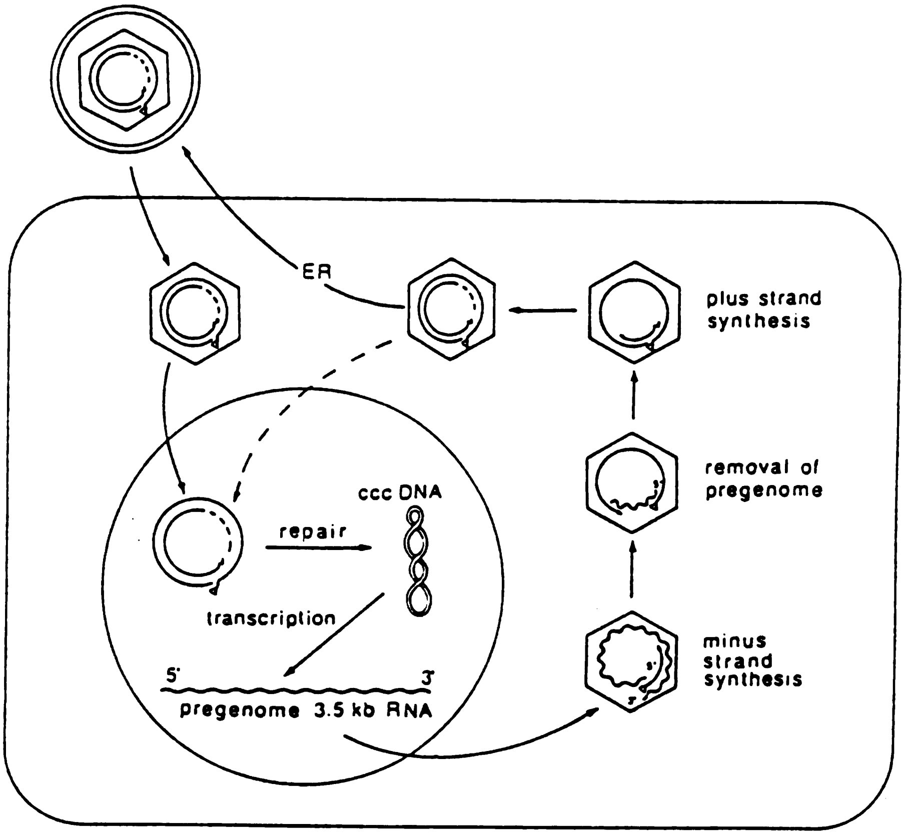
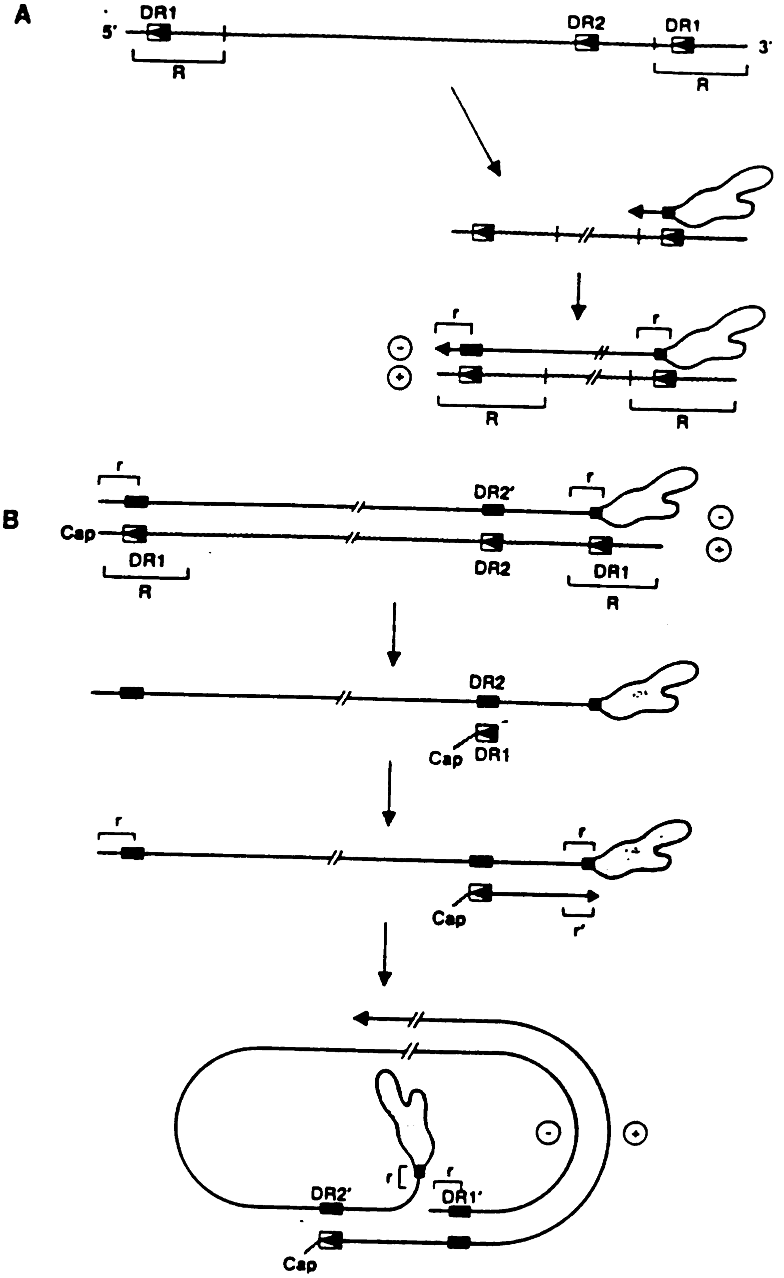

5'
pre S1
pre S2
S

FIG. 2 The hepadnaviral life cycle. Schematic diagram of the sequential events of the hepadnaviral life cycle: Hepatocyte entry, transport to the nucleus, genome repair, pregenome transcription, RNA packaging, minus and plus strand DNA synthesis, envelope acquisition and viral export. See text for further details.


ER
plus strand synthesis
removal of pregenome
ccc DNA
repair
transcription
minus strand synthesis
5'
3'
pregenome 3.5 kb RNA

UCSF LIBRARY

FIG. 3 Hepadnaviral DNA synthesis. A) Minus strand DNA synthesis. P protein serves as a protein primer for the progression of minus strand DNA synthesis from the 3' copy of DR1. Wavy lines denote RNA; straight lines denote DNA. R is the long terminal repeat in pregenome RNA; r is the short terminal repeat in minus strand DNA. B) Plus strand DNA synthesis. A short RNA containing DR1 sequences is cleaved from the 5' end of pregenome RNA and translocated to DR2 to prime plus strand DNA synthesis. Circularization occurs during plus strand synthesis by template transfer between r sequences as indicated. Reproduced, with permission (22).


A
DR1
5'
DR2
DR1
3'
R
R
r
r
−
+
R
R
B
r
DR2'
r
Cap
DR1
DR2
DR1
R
R
−
+
DR2
Cap
DR1
r
r
Cap
r'
DR2'
r
r
DR1'
Cap
−
+

REFERENCES

1. Albin,C., and W.S. Robinson. 1980. Protein kinase activity in hepatitis B virions. J.Virology 34: 297-302.

2. Argos,P., and S.D.Fuller. 1988. A model for the hepatitis B virus core protein: prediction of antigenic sites and relationship to RNA virus capsid proteins. EMBO J. 7: 819-824.

3. Bartenschlager,R., and H.Schaller. 1988. The amino-terminal domain of the hepadnaviral P-gene encodes the terminal protein (genome-linked protein) believed to prime reverse transcription. EMBO J. 7: 4185-4192.

4. Bartenschlager,R., M.Junker-Niepmann, and H.Schaller. 1990. The P gene product of the hepatitis B virus is required as a structural component for genomic RNA encapsidation. J.Virology 64: 5324-5332.

5. Beames, B., and R.Langford. 1993. Carboxy-terminal truncations of the HBV core protein affect capsid formation and the apparent size of encapsidated RNA. Virology 194: 597-607.

6. Beasley,R.P., C.C.Lin, L.Y.Hwang, and C.S.Chien. 1981. Hepatocellular carcinoma and hepatitis B virus. Lancet 2: 1129-1133.

7. Beckett, D., and O.C.Uhlenbeck. 1988. Ribonucleoprotein complexes of R17 coat protein and a translational operator analog. J.Molecular Biology 204: 927-938.

8. Bianchi,L., and F.Gudat. 1979. Immunopathology of hepatitis B. Prog. Liver Dis. 6: 371-392.

9. Birnbaum,F., and M.Nassal. 1990. Hepatitis B virus nucleocapsid assembly: primary structure requirements in the core protein. J.Virology 64: 3319-3330.

10. Borisova,G., B.Arya, A.Dislers, O.Borschukova, V.Tsibinogin, D.Skrastina, M.Eldarov, P.Pumpens, K.Skryabin, and E.Grens. 1993. Hybrid hepatitis B virus nucleocapsid bearing an immunodominant region from hepatitis B virus surface antigen. J.Virology 67: 3696-3701.

11. Buscher, M.W., W.Reiser, H.Will, and H.Schaller. 1985. Transcripts and the putative RNA pregenome of duck hepatitis B virus: implications for reverse transcription. Cell 40: 717-724.

12. Calvert, J., and J.Summers. 1994. Two regions of an avian hepadnavirus RNA pregenome are required in cis for encapsidation. J.Virology, in press.

13. Caspar,D.L.D., and A.Klug. 1962. Physical principles in the construction of regular viruses. Cold Spring Harbor Symp. Quant. Biol. 27: 1-22.

14. Clarke,B.E., S.E.Newton, A.R.Carroll, M.J.Francis, G.Appleyard, A.D.Syred, P.E.Highfield, D.J.Rowlands, and F.Brown. 1987. Improved immunogenicity of a peptide epitope after fusion to hepatitis B core protein. Nature 330: 381-384.

UCSF LIBRARY

15. Colgrove,R., G.Simon, and D.Ganem. 1989. Transcriptional activation of homologous and heterologous genes by the hepatitis B virus X gene product in cells permissive for viral replication. J.Virology 63: 4019-4026.

16. Dane,D.S., C.H.Cameron, and M.Briggs. 1970. Virus-like particles in serum of patients with Australian antigen associated hepatitis. Lancet 1: 695-698.

17. DeJean,A., L.Bouguesleret, K.Grzeschik, and P.Tiollais. 1986. Hepatitis B virus DNA integration in a sequence homologous to v-erb and steriod receptor genes in a hepatocellular carcinoma. Nature 322: 70-72.

18. Enders,G.H., D.Ganem, and H.E.Varmus. 1985. Mapping the major transcripts of ground squirrel hepatitis virus: the presumptive template for reverse transcriptase is terminally redundant. Cell 42: 297-308.

19. Feitelson,M., P.Marion, and W.Robinson. 1982. core particles of hepatitis B virus and ground squirrel hepatitis virus II: characterization of the protein kinase reaction associated with ground squirrel hepatitis virus and hepatitis B virus. J.Virology 43: 741-748.

20. Gallina,A., F.Bonelli, L.Zentilin, G.Rindi, M.Muttini, and G.Milanesi. 1989. A recombinant hepatitis B core antigen polypeptide with the protamine-like domain deleted self-assembles into capsid particles but fails to bind nucleic acids. J.Virology 63: 4645-4652.

21. Ganem, D. 1982. Persistent infection of humans with hepatitis B virus: mechanisms and consequences. Rev. Infect. Dis. 4: 1026-1047.

22. Ganem, D., and H.E.Varmus. 1987. The molecular biology of the hepatitis B viruses. Ann. Rev. Biochem. 56: 651-694.

23. Ganem, D. 1991. Assembly of hepadnaviral virions and subviral particles. Current Topics in Microbiology and Immunology 168: 61-82.

24. Garcea, R.L., and T.L.Benjamin. 1983. Host-range transforming gene of polyoma virus plays a role in virus assembly. PNAS 80: 3613-3617.

25. Gerlich,W.H., and W.Robinson. 1980. Hepatitis B virus contains protein attached to the 5' terminus of its complete DNA strand. Cell 21: 801-809.

26. Gerlich,W.H., U.Goldmann, R.Muller, W.Stibbe, and W.Wolff. 1982. Specificity and localization of the hepatitis B virus-associated protein kinase. J.Virology 42: 761-766.

27. Harrison, S.C. 1990. Principles of virus structure., in Virology, ed. by B.Fields et al. (Raven Press, Ltd., New York) pp.37-61.

28. Heerman,K.H., I.Goldman, W.Schwartz, T.Seyffarth, H.Baumgarten, and W.H.Gerlich. 1984. Large surface proteins of hepatitis B virus containing the preS sequence. J.Virology 52: 396-402.

29. Hirsch, R., J.Lavine, L.Chang, H.Varmus, and D.Ganem. 1990. Polymerase gene products of hepatitis B viruses are required for genomic RNA packaging as well as for reverse transcription. Nature 344: 552-555.

30. Hirsch, R., D.D.Loeb, J.R.Pollack, and D.Ganem. 1991. Cis-acting sequences required for encapsidation of duck hepatitis B virus pregenome RNA. J.Virology 65: 3309-3316.

31. Junker-Niepmann, M., R.Bartenschlager, and H.Schaller. 1990. A short cis-acting sequence is required for hepatitis B virus pregenome encapsidation and sufficient for packaging of foreign RNA. EMBO J. 9: 3389-3396.

32. Lien,J.M., C.E.Aldrich, and W.S.Mason. 1986. Evidence than a capped oligonucleotide is the primer for duck hepatitis B virus plus-strand DNA synthesis. J.Virology 57: 229-236.

33. Loeb, D.D., R.Hirsch, and D.Ganem. 1991. Sequence-independent RNA cleavages generate the primers for plus strand DNA synthesis in hepatitis B viruses: implications for other reverse transcribing elements. EMBO J. 10: 3533-3540.

34. 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. PNAS 77: 2941-2945.

35. Mason,W.S., G.Seal, and J.Summers. 1980. Virus of Pekin ducks with structural and biological relatedness to human hepatitis B virus. J.Virology 36: 829-836.

36. Miyanohara, A., T.Imamura, M.Araki, K.Sugawara, N.Ohtomo, and K.Matsubara. 1986. Expression of hepatitis B virus core antigen gene in Saccharomyces cerevisiae: synthesis of two polypeptides translated from different initiation codons. J.Virology 59: 176-180.

37. Nassal, M. 1992. The arginine-rich domain of the hepatitis B virus core protein is required for pregenome encapsidation and productive viral positive-strand DNA synthesis but not for virus assembly. J.Virology 66: 4107-4116.

38. Onodera, S., H.Ohori, M.Yamaki, and N.Ishida. 1982. Electron microscopy of human hepatitis B virus cores by negative staining-carbon film technique. J. Med. Virol. 10: 147-152.

39. Ou, J.-H., O.Laub, and W.J.Rutter. 1986. Hepatitis B virus gene function: the precore region targets the core antigen to cellular membranes and causes the secretion of the e antigen. PNAS 83: 1578-1582.

40. Patzer, E.J., G.R.Nakamura, C.C.Simonsen, A.D.Levinson, and R.Brands. 1986. Intracellular assembly and packaging of hepatitis B virus surface antigen particles occur in the endoplasmic reticulum. J.Virology 58: 884-892.

41. Perez,L.G., G.L.Davis, and E.Hunter. 1987. Mutants of the rous sarcoma virus envelope glycoproteinthat lack the transmembrane anchor and cytoplasmic domains: analysis of intracellular transport and assembly into virions. J.Virology 61: 2981-2988.

42. Petit, M.A., and J.Pillot. 1985. HBc and HBe antigenicity and DNA-bindng activity of major core protein P22 in hepatitis B virus core particles isolated from the cytoplasm of human liver cells. J.Virology 53: 543-551.

43. Phillips, B.A., D.F.Summers, and J.V.Maizel. 1968. In vitro assembly of poliiovirus-related particles. Virology 88: 92-104.

44. Phillips, B.A., R.E.Lundquist, and J.V.Maizel. 1980. Absence of subviral particles and assembly activity in HeLa cells infected with defective-interfering (DI) particles of poliovirus. Virology 100: 116-124.

45. Pollack, J.R., and D.Ganem. 1993. An RNA stem-loop structure directs hepatitis B virus genomic RNA encapsidation. J.Virology 67: 3254-3263.

46. Pollack, J.R., and D.Ganem. 1994. Duck hepatitis B virus polymerase binds an RNA stem-loop structure to initiate both viral RNA encapsidation and DNA synthesis. Submitted.

47. Pugh, J.C., A.Zweidler, and J.Summers. 1989. Characterization of the major duck hepatitis B virus core particle protein. J.Virology 63: 1371-1376.

48. Rhee, S.S., and E.Hunter. 1987. Myristylation is required for intracellular transport but not for assembly of D-type retrovirus capsids. J.Virology 61: 1045-1053.

49. Roossinck, M.J., and A.Siddiqui. 1987. In vivo phosphorylation and protein analysis of hepatitis B virus core antigen. J.Virology 61: 955-961.

50. Rueckert, R.R., A.Dunker, and C.Stoltzfus. 1969. The structure of Maus-Elberfeld virus: a model. PNAS 62: 912-919.

51. Russnak, R., and D.Ganem. 1990. Sequences 5' to the polyadenylation signal mediate differential poly(A) site use in hepatitis B viruses. Genes and Development 4: 764-776.

52. Salunke, D.M., D.L.D.Caspar, and R.L.Garcea. 1986. Self-assembly of purified polyomavirus capsid protein VP_1. Cell 46: 895-904.

53. Schlict, H.J., R.Bartenschlager, and H.Schaller. 1989. The duck hepatitis B virus core protein contains a highly phosphorylated C-terminus that is essential for replication but not for RNA packaging. J.Virology 63: 2995-3000.

54. Schwartzberg, P., J.Colicelli, M.L.Gordon, and S.P.Goff. 1984. Mutations in the gag gene of Moloney murine leukemia virus: effects on production of virions and reverse transcriptase. J.Virology 49: 918-924.

55. Seeger, C.D., D.Ganem, and H.E.Varmus. 1986. Biochemical and genetic evidence for the hepatitis B virus replication strategy. Science 232: 477-484.

56. Shafritz, D.A., D.Shouval, H.Sherman, S.Hadziyannis, and M.Kew. 1981. Integration of hepatitis B virus DNA into the genome of liver cells in chronic liver disease and hepatocellular carcinoma. NEJM 305: 1067-1073.

57. Spandau, D.F., and C.H.Lee. 1988. Transactivation of viral enhancers by the hepatitis B virus X protein. J.Virology 62: 427-434.

58. Sprengel, R., E.F.Kaleta, and H.Will. 1988. Isolation and characterization of a hepatitis B virus endemic in herons. J.Virology 62: 3832-3839.

59. Stahl, S.J., and K.Murray. 1989. Immunogenicity of peptide fusions to hepatitis B virus core antigen. PNAS 86: 6283-6287.

60. Standring, D.N., J.Ou, F.Masiarz, and W.J.Rutter. 1988. A signal peptide encoded within the precore region of hepatitis B virus directs the secretion of a heterogeneous population of e antigens in Xenopus oocytes. PNAS 85: 8405-8409.

61. Standring, D.N. 1991. The molecular biology of the hepatitis B virus core protein, in Molecular Biology of the Hepatitis B Virus, ed. McLachlan, A. (CRC, Boca Raton, FL) pp.145-169.

62. Summers, J., A.O'Connel, and I.Millman. 1975. Genome of hepatitis B virus: restriction enzyme cleavage and structure of DNA extracted from Dane particles. PNAS 72: 4597-4601.

63. Summers, J., J.M.Smolec, and R.Snyder. 1978. A virus similar to human hepatitis B virus associated with hepatitis and hepatoma in woodchucks. PNAS 75: 4533-4537.

64. 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.

65. Szmuness, W. 1978. Hepatocellular carcinoma and the hepatitis B virus: evidence for a causal association. Prog. Med. Virol. 24: 40-69.

66. Tavis, J.E., S.Perri, and D.Ganem. 1994. Hepadnaviral reverse transcription initiates within the stem-loop of the RNA packaging signal and employs a novel strand transfer. Submitted.

67. Wang, G.-H., and C.Seeger. 1992. The reverse transcriptase of hepatitis B virus acts as a protein primer for viral DNA synthesis. Cell 71: 663-670.

68. Wang, G.-H., and C.Seeger. 1993. Novel mechanism for reverse transcription in hepatitis B viruses. J.Virology 67: 6507-6512.

69. Weiss, R., 1982. Experimental biology and assay of RNA tumor viruses, in RNA Tumor Viruses, vol.1, ed. R.Weiss, N.Teich, H.Varmus, and J.Coffin. (CSH Laboratory, Cold Spring Harbor, N.Y.) pp.209-260.

70. Will, H., W.Reiser, T.Weimer, E.Pfaff, M.Buscher, R.Sprengel, R.Cattaneo, and H.Schaller. 1987. Replication strategy of human hepatitis B virus. J.Virology 61: 904-911.

71. Zhou, S., S.Q.Yang, and D.N. Standring. 1992. Characterization of hepatitis B virus capsid particle assembly in xenopus oocytes. J.Virology 66: 3086-3092.

72. Zhou, S., and D.N.Standring. 1992. Hepatitis B virus capsid particles are assembled from core-protein dimer precursors. PNAS 89: 10046-10050.

CHAPTER 2

UCSF LIBRARY

JOURNAL OF VIROLOGY, Oct. 1987, p. 3322-3325
0022-538X/87/103322-04$02.00/0
Copyright © 1987, American Society for Microbiology
Vol. 61, No. 10

Expression of the Precore Region of an Avian Hepatitis B Virus Is Not Required for Viral Replication

CHRISTOPHER CHANG,[1] GREG ENDERS,[1] ROLF SPRENGEL,[2]† NICK PETERS,[3] HAROLD E. VARMUS,[1,2] AND DON GANEM[2,4]*

Departments of Microbiology and Immunology,[2] Biochemistry and Biophysics,[1] and Medicine,[4] University of California Medical Center, and Department of Biology, University of San Francisco, San Francisco, California 94143[3]

Received 21 May 1987/Accepted 26 June 1987

The core-antigen-coding region of all hepadnaviruses is preceded by a short, in-phase open reading frame termed precore whose expression can give rise to core-antigen-related polypeptides. To explore the functional significance of precore expression in vivo, we introduced a frameshift mutation into this region of the duck hepatitis B virus (DHBV) genome and examined the phenotype of this mutant DNA by intrahepatic inoculation into newborn ducklings. Animals receiving mutant DNA developed DHBV infection, as judged by the presence in hepatocytes of characteristic viral replicative intermediates; molecular cloning and DNA sequencing confirmed that the original mutation was present in the progeny genomes. Infection could be efficiently transmitted to susceptible ducklings by percutaneous inoculation with serum from mutant-infected animals, indicating that infectious progeny virus was generated. These findings indicate that expression of the precore region of DHBV is not essential for genomic replication, core particle morphogenesis, or intrahepatic viral spread.

The hepadnaviruses are a group of small, enveloped DNA viruses that produce acute and chronic infections of hepatocytes and that replicate their DNA by reverse transcription of RNA intermediates (5, 21). To date, members of this group have been recovered from hosts of several species, including humans, woodchucks, ground squirrels, and ducks (19). All of these viruses share a similar genomic organization, which includes at least three open reading frames (5, 19, 21). One of these encodes the multiple surface proteins of the envelope, another encodes the reverse transcriptase required for replication, and a third encodes the major structural protein of the viral nucleocapsid or core (core antigen). The mammalian viruses harbor an additional coding region (termed X) of unknown function.

In all hepadnaviral genomes, the coding region for core antigen is preceded by an in-phase, contiguous open reading frame of 29 to 43 codons known as precore (Fig. 1A). The expression of this region is predicted to generate core-related polypeptides with additional amino acids at their N termini. There is good evidence that precore expression does indeed occur in vivo. All hepadnavirus-infected liver cells produce at least two sets of transcripts containing core sequences. One RNA initiates within the precore region and therefore can encode only the core antigen. The other set of transcripts initiates upstream of the precore ATG codon and hence could encode the larger precore polypeptide (1, 3, 11, 24; R. Sprengel and H. Will, unpublished data).

Extensive work with human hepatitis B virus (HBV) indicates that precore polypeptides can be generated from such RNAs and that their properties differ from those of core antigen. Several studies (9, 12, 13) have used heterologous promoters to drive the expression of either precore or core gene products in cultured cells transfected with subgenomic fragments of cloned HBV DNA. When protein synthesis is initiated at the core ATG, the resulting protein is cytoplasmic in location, while initiation of translation at the precore ATG results in the targeting of the gene product to a membranous compartment (12); after further proteolytic processing (9; D. Standring, J. Ou, and W. Rutter, personal communication), the products are secreted into the medium. Similar products (collectively termed hepatitis B e antigen) are observed in the serum of HBV-infected patients, arguing strongly that this pathway is functional in vivo (5, 21). These results suggest that the HBV precore region encodes a signal sequence that can function to direct core antigen determinants to novel subcellular locales. Similarly, we have recently demonstrated that the serum of ducks infected with duck hepatitis B virus (DHBV) also contains secreted polypeptides derived from the expression of the DHBV precore region (R. Sprengel and H. Will, manuscript in preparation). However, the functional role of precore gene products in authentic viral replication in productively infected cells has not yet been examined.

Progress in dissecting the roles of the individual viral gene products in vivo has been hampered by the absence of convenient cell culture systems for viral growth. However, the observation that cloned viral DNA can initiate productive infection of susceptible hosts following intrahepatic inoculation (16, 17, 22) has allowed the development of a simple strategy for the mutational analysis of viral functions. Briefly, lesions in the region of interest are generated by site-directed mutagenesis in vitro, and the resulting mutant genomes are tested for viability in vivo by intrahepatic transfection. We have previously used this strategy to examine the role of short direct repeats (DRs) of viral DNA in genomic replication (15). In this study, we present a similar analysis of the hepadnaviral precore region. Although the limited host range of HBV (which includes only humans and higher primates) makes this approach difficult for the human virus, the animal hepadnavirus models are ideally suited to

---

* Corresponding author.

† Present address: Max Planck Institute for Biochemistry, Munich, Federal Republic of Germany.

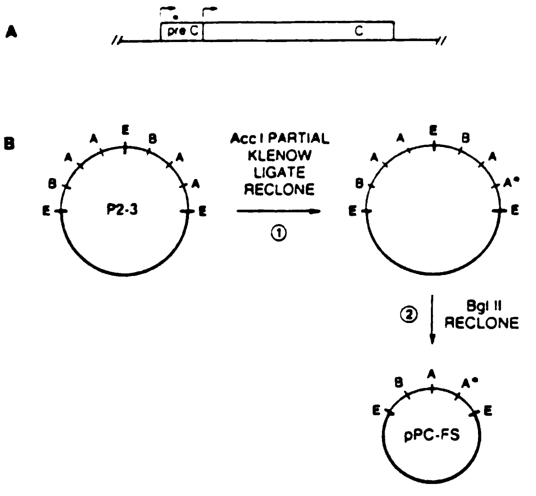


FIG. 1. Construction of precore frameshift mutation in DHBV DNA. (A) The organization of the precore and core-antigen-coding regions of DHBV. ⊏, Coding regions; →, ATG codons; PreC, precore region; C, core-antigen-coding frame; *, site of precore frameshift mutation. (B) Construction of precore frameshift lesion. Plasmid p2-3 is a tandem dimer of DHBV DNA, with restriction sites as indicated. E, *Eco*RI; A, *Acc*I; B, *Bgl*II. The manipulations at each step of the construction are described in detail in the text and are indicated next to the arrow depicting each step. A*, Precore *Acc*I site inactivated by the 2-base-pair insertion.

this form of analysis. The avian system is particularly attractive, since susceptible ducklings are readily available and the incubation period of DHBV infection is only 2 to 3 weeks (versus 2 to 3 months for the mammalian viruses). Accordingly, we used DHBV as the experimental system for these studies.

To address the role of precore sequences in vivo, we constructed a frameshift mutation in the DHBV precore region and tested the mutant DNA for infectivity. The scheme for construction of the mutant is detailed in Fig. 1B. Briefly, plasmid p2-3, bearing a tandem dimer of the DHBV genome (generous gift of W. Mason), was linearized by partial digestion with *Acc*I; one of the two *Acc*I cleavage sites in DHBV DNA lies within the precore region (8). The termini of gel-purified linear molecules were then repaired with DNA polymerase I (Klenow fragment), religated, and cloned. Clones which had lost the appropriate *Acc*I site in the precore region were identified by restriction mapping. The wild-type copy of the DHBV genome was then excised with *Bgl*II to generate clones bearing only the mutant genome. In one of these (pPC-fs), the presence of the expected two-base insertion was verified by DNA sequencing (data not shown).

For infectivity testing, mutant DHBV genomes were liberated from the plasmid vector by *Eco*RI digestion, and the released molecules were self-annealed and ligated at low DNA concentrations (1 to 10 μg/ml) to promote recircularization. A 5-μg sample of this DNA (in 0.2 ml of 500 μg of DEAE-dextran per ml) was percutaneously injected into the livers of each of six virus-free ducklings (SPAFAS, Inc., Norwalk, Conn.) 1 day posthatching. Four control ducklings were similarly exposed to wild-type DNA. Three weeks later, all ducks were sacrificed and their liver DNA was examined for progeny DHBV sequences by Southern blotting with ^{32}P-labeled DHBV DNA as a probe. Five of the six mutant-infected samples (and one of the four controls) revealed the characteristic array of replicative intermediates typical of productively infected hepatocytes (20), i.e., asymmetric, protein-linked, heterogeneous forms and fully duplex monomeric circles lacking covalently attached protein (two representative examples are shown in Fig. 2C). To verify that the replicative forms seen in mutant-infected livers were of mutant origin, two experiments were carried out. First we examined the restriction pattern of the intrahepatic circular duplex viral DNA (Fig. 2A). Liver DNA from all five infected animals was prepared by phenol extraction in the absence of proteinase K; this procedure eliminates the abundant protein-linked heterogeneous viral DNA forms (6), which would otherwise obscure the cleavage pattern. *Acc*I cleavage of the remaining viral DNA in all cases revealed the presence of only one *Acc*I site; mapping of this site relative to the unique *Eco*RI and *Bgl*II sites of DHBV confirmed that the missing *Acc*I site was that in the precore region (representative results for one of the five mutants are shown in Fig. 2B).

This result rules out contamination of the input DNA with wild-type sequences or reversion of the mutant to the wild type during replication but does not exclude the occurrence of a second-site (compensatory) frameshift mutation in precore, which could restore the reading frame to the wild type. To examine this possibility, we digested the liver DNA of one mutant-infected animal with *Eco*RI. We recovered fragments of 3.0 to 3.5 kilobases (kb) from agarose gels and cloned these into lambda gtWES. Progeny plaques annealing to ^{32}P-labeled DHBV DNA sequences were picked, and a 1.3-kb *Xba*I fragment spanning the precore region was subcloned from one of these into pSP65 (10). In this plasmid vector, the sequence of the entire precore region was determined by the chain termination method. The original two-base frameshift mutation was still present in the recovered mutant DNA (Fig. 3); further sequencing (data not shown) revealed no additional changes between the precore ATG and the first out-of-phase terminator (at position 2631) following the frameshift lesion.

The presence of replicative forms in intrahepatic DNA strongly implies that infectious virus generated in the initially transfected cells successfully spread to uninfected hepatocytes (16, 17). To confirm this inference, we prepared a cell-free homogenate from one mutant-infected liver and used this material to inoculate four uninfected 1-day-old ducklings by subcutaneous inoculation in the thigh. Examination of liver DNA from these animals 21 days later revealed the presence of typical DHBV replicative forms in all four samples; restriction analysis of two of these samples performed as outlined in Fig. 2 again confirmed the presence of the mutation (data not shown). We then tested for the presence of infectious virus in the serum of these recipients. A pool of serum from three of these mutant-infected animals was inoculated subcutaneously into three ducklings (0.1 ml per recipient). All three developed viremia, as judged by dot hybridization of recipient serum.

These experiments demonstrate that the expression of the precore region of DHBV is not essential for the uptake, replication, assembly, or intrahepatic spread of the virus in experimental infection in vivo. However, it is possible that precore expression might affect the efficiency of titer of virus infection or contribute to more specialized biological functions, such as tropism for extrahepatic viscera (7) or competence for vertical transmission; these possibilities are

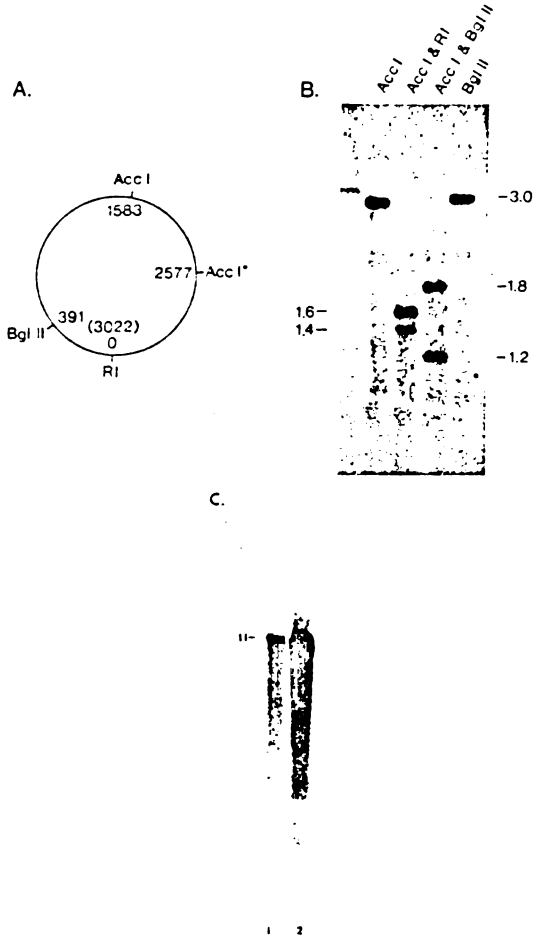


FIG. 2. Viral DNA forms in the livers of ducks transfected with mutant genomes. (A) Restriction map of relevant sites within DHBV DNA. Nucleotide positions are numbered with respect to the unique *EcoRI* site (8). *, *AccI* site eliminated by the precore frameshift lesion. (B) Cleavage analysis of duplex DHBV DNA in the liver of animal 5 at 3 weeks posttransfection with pPc-fs DNA. Liver homogenates in 1% sodium dodecyl sulfate were extracted with phenol chloroform (1:1) without prior proteinase K treatment, as previously described (16). A 10-μg portion of the resulting relaxed circular DNA was digested with the indicated enzymes; the sample in the left lane was undigested. Product DNA was electrophoresed through 1% agarose, transferred to nitrocellulose, hybridized to ^{32}P-labeled DHBV DNA, and autoradiographed as previously described (16). The sizes of the fragments (in kilobases) are indicated. (C) Liver DNA prepared from animals 5 (lane 1) and 6 (lane 2) 3 weeks posttransfection with pPC-fs DNA. Liver homogenates were treated with 1% sodium dodecyl sulfate and 500 μg of proteinase K per ml for 3 h prior to phenol extraction; 10 μg of the resulting DNA was digested with *PvuII*, which does not cleave within DHBV. Product DNA was electrophoresed through 1% agarose, transferred to nitrocellulose, and hybridized to ^{32}P-labeled DHBV DNA as in panel B. ||, Position of fully duplex open circles of DHBV DNA.

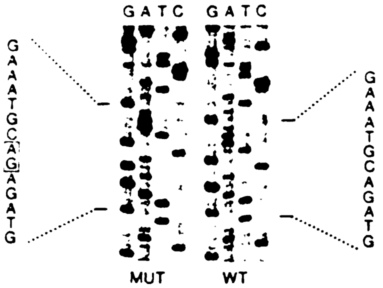


FIG. 3. Nucleotide sequence of the precore region of cloned DHBV DNA recovered from liver of a mutant-infected duck. A 1.3-kb *XbaI* fragment of DHBV DNA was excised from a lambda gt WES clone of the mutant viral DNA and subcloned into the *XbaI* site of pSP65 (10). A clone in which the precore region was adjacent to the SP6 promoter region was identified. The sequence of the entire precore region was determined by the dideoxy method by using a primer complementary to the SP6 promoter sequence. Left panel, a portion of the sequence showing the region spanning the lesion in recovered mutant (MUT) DNA; right panel, corresponding sequence of wild-type (WT) DHBV DNA. The sequences depicted are of minus-strand polarity.

currently being explored. The availability of viable precore mutant viruses will also make possible an assessment of the impact of precore proteins on viral persistence and on the pathogenic potential of DHBV.

We do not know at present whether these findings can be extrapolated to the mammalian viruses, including HBV. As noted previously, DHBV was chosen for these studies because of experimental tractability. However, DHBV is the most diverged of the hepadnaviruses at the sequence level (8, 18). Its 35-kilodalton core antigen is distinctly larger than those of the mammalian viruses (20 to 22 kilodaltons) and within the DHBV precore/core open reading frame, there is little amino acid homology to the cognate region of HBV (18). On the other hand, the DHBV precore region does contain a consensus signal peptidase recognition sequence (R. Colgrove, personal communication), and the virus does generate secretory precore products analogous to the e antigen of HBV. Additional experiments will be required to directly test whether the mammalian viruses require functional precore polypeptides for replication; we are currently constructing analogous mutations in the ground squirrel hepatitis virus precore region to examine this issue further.

Our findings with the DHBV precore gene recall an interesting parallel in murine retrovirology. The *gag* gene of Moloney murine leukemia virus encodes the nucleocapsid or core proteins of the virus, and its ATG initiator is preceded by an upstream initiator codon that is also expressed in vivo. Translation from the upstream initiator gives rise to *gag*-related proteins which enter the secretory pathway, are glycosylated, and are expressed on the cell surface and in the medium (2). Interestingly, as for DHBV, mutational ablation of these upstream sequences does not impair viral replication or infectivity (4, 14).

The experiments reported here represent the first demonstration of a nonessential region within the genome of a hepadnavirus. The existence of such a region is surprising given the small size (3.0 kb) and extremely compact coding organization of the genome: in DHBV (as in HBV), every nucleotide in the genome is in at least one coding region and 50% of the sequence is read in more than one frame. The fact

UCSF LIBRARY

that the precore region is nonessential also suggests that insertions of exogenous DNA sequences into this region may be tolerated and raises the intriguing possibility that hepadnaviruses could be developed as genetic vectors for the delivery of foreign DNA to the liver.

These experiments were supported by Public Health Service grants (AI18782 and AM26743) from the National Institutes of Health to the authors and to the University of California at San Francisco Liver Center. H.E.V. is American Cancer Society Professor of Molecular Virology.

We thank Effie Meredith for outstanding assistance with the preparation of the manuscript.

LITERATURE CITED

1. **Buscher, M., W. Reiser, H. Will, and H. Schaffer.** 1985. Transcripts and the putative RNA pregenome of duck hepatitis B virus: implications for reverse transcription. Cell 40:717–724.
2. **Edwards, S., and H. Fan.** 1980. Sequence relationship of glycosylated and unglycosylated *gag* polyproteins of Moloney murine leukemia virus. J. Virol. 35:41–51.
3. **Enders, G. H., D. Ganem, and H. E. Varmus.** 1987. 5'-Terminal sequences influence the segregation of ground squirrel hepatitis virus RNAs into polyribosomes and viral core particles. J. Virol. 61:35–41.
4. **Fan, H., H. Chute, E. Chao, and M. Feuerman.** 1983. Construction and characterization of Moloney murine leukemia virus mutants unable to synthesize glycosylated gag polyprotein. Proc. Natl. Acad. Sci. USA 80:5965–5969.
5. **Ganem, D., and H. Varmus.** 1987. The molecular biology of the hepatitis B viruses. Annu. Rev. Biochem. 56:651–695.
6. **Gerlich, W. H., and W. S. Robinson.** 1980. Hepatitis B virus contains protein attached to the 5' terminus of its complete DNA strand. Cell 21:801–810.
7. **Halpern, M. S., J. M. England, D. T. Deery, D. J. Petcu, W. S. Mason, and K. L. Molnar-Kimber.** 1983. Viral nucleic acid synthesis and antigen accumulation in pancreas and kidney of Pekin ducks infected with duck hepatitis B virus. Proc. Natl. Acad. Sci. USA 80:4865–4869.
8. **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.
9. **McLachlan, A., D. Milich, A. K. Raney, M. G. Riggs, J. Hughes, J. Sorge, and F. V. Chisari.** 1987. Expression of hepatitis B virus surface and core antigens: influences of pre-S and precore sequences. J. Virol. 61:683–692.
10. **Melton, D., P. Kreig, M. Rebogliati, T. Maniatis, K. Zinn, and M. Green.** 1984. Efficient *in vitro* synthesis of biologically active RNA and RNA hybridization probes from plasmids containing a bacteriophage SP6 promoter. Nucleic Acids Res. 12:7035–7056.
11. **Moroy, T., J. Etiemble, C. Trepo, P. Tiollais, and M. Buendia.** 1985. Transcription of woodchuck hepatitis in the chronically infected liver. EMBO J. 4:1507–1514.
12. **Ou, J.-H., O. Laub, and W. J. Rutter.** 1986. Hepatitis B gene function: the precore region targets the core antigen to cellular membranes and causes the secretion of the e antigen. Proc. Natl. Acad. Sci. USA 83:1578–1582.
13. **Roossinck, M. J., S. Jameel, S. H. Loukin, and A. Siddiqui.** 1986. Expression of hepatitis B viral core region in mammalian cells. Mol. Cell. Biol. 6:1393–1400.
14. **Schwartzberg, P., J. Colicelli, and S. P. Goff.** 1983. Deletion mutants of Moloney murine leukemia virus which lack glycosylated *gag* protein are replication competent. J. Virol. 46:538–546.
15. **Seeger, C., D. Ganem, and H. E. Varmus.** 1986. Biochemical and genetic evidence for the hepatitis B virus replication strategy. Science 232:477–484.
16. **Seeger, C., D. Ganem, and H. E. Varmus.** 1986. The cloned genome of ground squirrel hepatitis virus is infectious in the animal. Proc. Natl. Acad. Sci. USA 81:5849–5852.
17. **Sprengel, R., C. Kuhn, C. Manso, and H. Will.** 1984. Cloned duck hepatitis B virus DNA is infectious in Pekin ducks. J. Virol. 52:932–937.
18. **Sprengel, R., C. Kuhn, H. Will, and H. Schaller.** 1985. Comparative sequence analysis of duck and human hepatitis B virus genomes. J. Med. Virol. 15:323–333.
19. **Summers, J.** 1981. Three recently described animal virus models for human hepatitis B virus. Hepatology 1:179–183.
20. **Summers, J., and W. Mason.** 1982. Replication of the genome of a hepatitis B-like virus by reverse transcription of an RNA intermediate. Cell 29:403–415.
21. **Tiollais, P., C. Pourcel, and A. Dejean.** 1985. The hepatitis B virus. Nature (London) 317:489–495.
22. **Will, H., R. Cattaneo, H. G. Koch, G. Darai, and H. Schaller.** 1982. Cloned HBV DNA causes hepatitis in chimpanzees. Nature (London) 299:740–742.
23. **Will, H., R. Cattaneo, E. Pfaff, C. Kuhn, M. Roggendorf, and H. Schaller.** 1984. Expression of hepatitis B virus antigens with a simian virus 40 vector. J. Virol. 50:335–342.
24. **Will, H., W. Reiser, T. Wiemer, E. Pfaff, M. Büscher, R. Sprengel, R. Catteneo, and H. Schaller.** 1987. Replication strategy of human hepatitis B virus. J. Virol. 61:904–911.

ADDENDUM TO CHAPTER 2

The results reported in chapter 2 clearly demonstrate that expression of the precore region, and therefore secretion of e antigen, is not essential for genome replication, nucleocapsid assembly, or intrahepatic spread. Nonetheless, conservation of its ORF among all members of the hepadnaviral family suggests that precore/e antigen performs some function. One proposal is that e antigen may modulate the host cellular immune response during the course of viral infection. Since the host response has been implicated as the primary cause of liver disease during HBV infection, e antigen may play a pivotal role in viral pathogenesis.

Since the publication of chapter 2, many reports have described the identification of HBe antigen-deficient mutants in patients infected with HBV. These reports confirm our finding that e antigen expression is not required for replication. Interestingly, patients bearing these mutant viruses frequently had more severe liver injury (1,2). Additional studies have suggested a strong association between patients with fulminant hepatitis and the presence of precore nonsense mutations in the PCR-amplified HBV DNA (3,4,7). An attractive model for these results is that infected hepatocytes synthesizing e antigen are favored targets for the host T-cell response and therefore selectively eliminated. Viruses that ablate e antigen production should have a selective advantage to spread more efficiently to uninfected hepatocytes, and for this reason might be more pathogenic than wild-type virus. However, not all of the data conveniently conform to this model. Okamoto and colleagues showed that 18 of 18 HBV carriers with anti-HBe had circulating HBV mutants defective for e antigen synthesis and secretion, regardless of whether they were asymptomatic or had liver cirrhosis (6). This study strongly suggests that mutations that terminate expression of precore are not sufficient to cause severe liver injury.

In addition, variations in individual immune response must also contribute to liver injury. Another feature of the e antigen-mediated pathogenicity model deserves additional consideration. If the expression of e antigen is important in triggering the host immune response against infected hepatocytes, then it functions to restrict or curtail virus spread and should be selected against. If so, why has expression of precore been retained in viral evolution? One potential function of the secreted e antigen may be to induce immunologic tolerance *in utero*, predisposing the newborn to chronic HBV infection (5). Consequently, expression of e antigen may represent a viral strategy to persist in the host after perinatal infection, which would therefore confer a selective advantage on the conservation of the precore region. This hypothesis explains the observation that infants born to carrier mothers expressing e antigen invariably become persistently infected. E antigen-expressing transgenic mice, which are tolerant to both core and e antigens at the T-cell level, displayed an immunologic status that parallels that of neonates born to carrier mothers (5). However, because HBV cannot infect the transgenic mice with this T-cell tolerance, the role of e antigen-induced tolerance on HBV pathogenesis cannot be directly studied. Whether or not this model proves correct, a further understanding of the mechanisms involved in HBV-mediated liver pathogenesis is needed.

REFERENCES

Brunetto,M., M.Stemler, F.Bonino, F.Schodel, F.Oliveri, M.Rizetto, G.Verme, and H.Will. 1990. A new hepatitis B virus strain in patients with severe anti-HBe positive chronic hepatitis B. J.Hepatology 10: 258-261.

Carman,W.F., M.Jacyna, S.Hadziyannis, P.Karayiannis, M.McGarvey, A.Markis, and H.Thomas. 1989. Mutations preventing formation of hepatitis B e antigen in patients with chronic hepatitis B infection. Lancet ii: 588-591.

Hasegawa,K., J.Huang,J.Wands, H.Obata, and T.J.Liang. 1991. Association of hepatitis B viral precore mutations with fulminant hepatitis B in Japan. Virology 185: 460-463.

Liang,T.J., K.Hasegawa, N.Rimon, J.Wands, and E.Ben-Porath. 1991. A hepatitis B virus mutant associated with an epidemic of fulminant hepatitis. NEJM 324: 1705-1709.

Milich,D.R., J.Jones, J.Hughes, J.Price, A.Raney, and A.McLachlan. 1990. Is a function of the secreted hepatitis B e antigen to induce immunologic tolerance in utero? PNAS 87: 6599-6603.

Okamoto,H., S.Yotsumoto, Y.Akahane, T.Yamanaka, Y.Miyazaki, Y.Sugai, F.Tsuda, T.Tanaka, Y.Miyakawa, and M.Mayumi. 1990. Hepatitis B viruses with precore region defects prevail in persistently infected hosts along with seroconversion to the antibody against e antigen. J.Virology 64: 1298-1303.

Omata,M., T.Ehata, O.Yokosuka, K.Hosoda, and M.Ohto. 1991. Mutations in the precore region of hepatitis B virus DNA in patients with fulminant and severe hepatitis. NEJM 324: 1699-1704.

CHAPTER 3

PHENOTYPIC MIXING OF HEPADNAVIRAL NUCLEOCAPSID PROTEINS REVEALS PARTICLES TO BE COMPOSED OF HOMODIMER PRECURSORS

Christopher Chang[1], Siliang Zhou[2], Don Ganem[3], and David N. Standring[1,2]

Department of Biochemistry and Biophysics[1], Hormone Research Institute[2], and
Departments of Microbiology and Immunology, and Medicine[3], University of California
Medical Center, San Francisco, California 94143

ABSTRACT

Hepadnaviruses encode a single core protein (C) which assembles into a nucleocapsid containing the polymerase protein and pregenomic RNA during viral replication in hepatocytes. *In vivo* capsid assembly studies in several heterologous hosts have demonstrated that the arginine-rich C-terminal tail is dispensible for particle formation. We examined the ability of heterologous hepadnaviral C proteins to cross-oligomerize. Using a two-hybrid strategy in replication-permissive HepG2 cells, we observed cross-oligomerization among the core proteins from hepatitis B virus (HBV), woodchuck hepatitis virus (WHV), and ground squirrel hepatitis virus (GSHV). When expressed in the *Xenopus* oocyte, in which hepadnaviral C proteins form capsids, the core polypeptides from WHV and GSHV, but not duck hepatitis B virus (DHBV), can efficiently co-assemble with an epitope-tagged HBV core polypeptide to form mixed capsids. However, when two different core mRNAs are translated in the oocyte, the core monomers show a strong preference for forming homodimers rather than heterodimers. This holds true even for co-expression of two HBV C proteins differing only by an epitope tag, suggesting that core monomers are not free to diffuse and associate with other monomers. Thus, mixed capsids result from aggregation of different species of homodimers.

INTRODUCTION

The core (C) protein of hepatitis B virus (HBV) plays a major role in viral morphogenesis. Core polypeptides, the polymerase, and the pregenomic RNA assemble to form the nucleocapsid during viral replication in hepatocytes. The nucleocapsid, in turn, interacts with cellular membranes that contain HBV envelope glycoproteins. Several labortatories have demonstrated that core protein expressed in the absence of all other viral products in hosts as divergent as bacteria (3,6), yeast (13), insect cells (2), or *Xenopus* oocytes (20) can spontaneously assemble into capsid particles. Nevertheless, the pathway for capsid assembly remains uncertain. In the *Xenopus* oocyte system, Zhou and Standring have shown that unassembled core protein exists in a dimer pool that serves as a precursor to capsids. Electron microscopy and velocity sedimentation studies have failed to detect *in vivo* higher-order multimer intermediates often seen in the assembly of other animal virus capsids (21).

Efforts to identify regions of the 185 a.a. core protein (HBV subtype adw) that are critical for capsid formation have focused on the arginine-rich C-terminal tail (2,3). In various heterologous systems expressing core protein, deletions containing the N-terminal 144 a.a. produce stable capsids, demonstrating the Arg-rich tail is not necessary for capsid assembly (2,3). Moreover, assembly is tolerant of heterologous sequences placed at the N-terminus, C-terminus, and within an immunodominant loop of core protein (4,5,16). However, further deletions beyond residue 144 or into the amino terminus of core protein abrogate core particle assembly.

To further characterize C-C interactions involved in capsid assembly, we examined the ability of mammalian and avian viral C proteins to interact with each other. We employed both expression in *Xenopus* oocytes and a mammalian cell-based two-hybrid system. The two-hybrid strategy originally developed by Fields and Song to detect protein-protein interactions relies on the modular nature of the yeast Gal4 DNA-binding and transactivation domains (9). Fearon and colleagues modified this system to study interactions in mammalian cells by using the herpes simplex virus VP16 transactivation domain (8). Luban et.al. have applied this strategy to detect multimerization of gag polyprotein from both human immunodeficiency virus and Moloney murine leukemia virus (11).

In this communication, we report that HBV core proteins fused to Gal4 and VP16 domains interact in the two-hybrid assay, and this interaction behaves as predicted from *in vivo* studies of capsid formation. Using this two-hybrid assay, we demonstrate cross-oligomerization among the core proteins from HBV, woodchuck hepatitis virus (WHV) and ground squirrel hepatitis virus (GSHV). Consistent with this, when expressed in *Xenopus* oocytes the core proteins from WHV and GSHV, but not duck hepatitis B virus (DHBV), can efficiently co-assemble with HBV core protein to form mixed capsids. Surprisingly, however, when two different core mRNAs are translated in the oocyte, the core monomers show a marked preference for forming homodimers rather than heterodimers; thus, mixed capsids are aggregates of different species of homodimers.

MATERIALS AND METHODS

Materials. Restriction enzymes were purchased from New England BioLabs and used as specified by the manufacturer. Radionucleotides were purchased from Amersham Corp. MA 18/7, the monoclonal antibody directed against the L epitope, was kindly provided by Dr. Volker Bruss. Rabbit polyclonal anti-HBV core antigen antibody was purchased from Dako Corp. Anti-DHBV-C was kindly provided by Dr. Jesse Summers.

Plasmids and RNAs. All HBV nucleotide positions are numbered from the unique EcoRI site of HBV subtype adw2. Core fusion proteins with Gal4 or VP16 were made by subcloning appropriate restriction or PCR amplified fragments into the expression vectors pGAL0 and pNLVP16 which, along with the reporter plasmid pG5E1bCAT, are described earlier (8) and kindly provided by Dr. Eric Fearon. In brief, pGAL0 encodes the DNA-binding domain of GAL4, extending from a.a. 1-147, followed by polylinker sequence. pNLVP16 contains a.a. 411-455 of the HSV VP16 transactivation domain followed by polylinker sequence. Full-length HBV, WHV and GSHV core fusions were made by subcloning a 581 bp. StyI fragment into the polylinker of pGAL0 and pNLVP16 to create an in-frame fusion with the C-terminus of the GAL4 or VP16 domain. C-terminal truncations HBc1-144 and HBc1-117 were engineered by deleting sequence from BspEI and SspI, respectively, to the polylinker. HBc29-168 was constructed by inserting a 445 bp BglII fragment into the pGAL4 polylinker; HBc38-183 was constructed by inserting a 449 bp TaqI-StyI fragment into the pNLVP16 polylinker.

To create SP6 expression vectors, the fragments spanning the core ORF of WHV, GSHV, and DHBV were generated by PCR amplification so as to include an NcoI site at the 5' end and a BstEII site at the 3' end. These fragments were cloned into the NcoI and BstEII sites of a derivative of pSP64T. This strategy puts the core sequences between the translation-enhancing 5' and 3' UTR sequences of the β-globin gene. Plasmids pSP64T-C and Δ157 are described in detail elsewhere (19). Briefly, pSP64T-C contains the 581 bp. StyI fragment from HBV adw2 (nt. 1880-2461) cloned into the BglII site of pSP64T. Δ157 encodes a core deletion containing the first 157 amino acids fused in frame to the last two residues of core protein. The L epitope tag has been previously described (15). We modified the original L epitope sequence with a silent base change that would introduce a BamHI site: 5' GATCATCAGTTGGATCCTGCATTCGGAGCCAACTCA 3'. The L epitope in Δ157L and HBV-CL is bordered by core residues 78 and 80, deleting residue 79, in the immunodominant loop of core. It was introduced into this site by ligation at the BamHI site of the two PCR-amplified halves of the core ORF flanking the L epitope. SP6 expression constructs were verified by sequence analysis.

In vitro transcription. Synthetic transcripts were prepared using SP6 polymerase (Promega). Transcription reactions were treated with DNase, phenol-extracted, and ethanol precipitated and resuspended into H_2O and stored at -70° C until required and then microinjected without further purification.

Transfections. HepG2 cells were grown on 100 mm-diameter dishes in DMEH-16 medium supplemented with 10% fetal bovine serum (Gibco), 0.14% sodium bicarbonate, and 2 mM glutamine. HepG2 cells were transfected with 5µg each of the GAL4 and VP16 fusion constructs and the CAT reporter plasmid by the calcium phosphate coprecipitation procedure as described previously (10).

CAT activity assays. Cell extracts for CAT activity were prepared by freeze-thaw lysis of cells 48 hrs post-transfection. CAT activity was determined in extracts using a phase separation procedure. In brief, 5 µl of cell extract was incubated in a 100 µl volume containing 250 mM Tris-HCl pH 8.0, 0.5 mM n-butyryl CoA (Sigma) and 0.2 µCi ^{14}C-chloramphenicol (58.1 mCi/mmole) for 15-60 minutes at 37°C. Phase extraction was performed with an equal volume of 2:1 hexane/xylenes mix, and levels of n-butyrylated chloramphenicol determined by liquid scintillation counting. CAT activity is reported as fold-increase relative to CAT activity resulting from transfection of the reporter CAT plasmid alone.

Oocyte handling and microinjection. Oocytes were cultured at 17°C in modified Barth's medium containing antibiotics (MBSH) by standard procedures (20). Batches of 10-20 oocytes per experiment were injected as described previously (17) by using a programmable microinjection device (Sutter Instruments, Novato, CA) to control delivery of RNA (40 nl per oocyte). For experiments involving analysis of assembled capsids, metabolic labeling was initiated within 1 hr of injection by placing the oocytes in MBSH containing 0.5 to 1.0 mCi of ^{35}S-methionine and ^{35}S-cysteine (Expresslabel, NEN) per ml and incubating them for a further 35-40 hrs at 17°C. For analysis of the intermediates for capsid assembly, the oocytes were labelled 24 hrs after injection for approximately 20 hrs. Lysates were prepared by homogenizing the oocytes in buffer (≈400 µl) containing 50 mM Tris-HCl pH 7.5, 1% NP-40, and 10 mM EDTA. Lysates were then clarified (eppendorf centrifuge, 15,000 x g, 10 min.) before fractionation.

Sucrose gradient fractionation. Oocyte lysates were layered onto either 10 to 60% (wt/vol) or 3 to 25% sucrose gradients (20,21). For 10-60% gradients, 200 µl of lysate were layered onto step gradients (1.4 ml final volume) with six 200 µl steps of 10, 20, 30, 40, 50, and 60% sucrose in 50 mM Tris-HCl pH 7.5, 100 mM NaCl, and 1 mM EDTA (TNE), and centrifuged for 40 min at 4°C and 55,000 rpm in a TLS-55 swinging-bucket rotor on a TL-100 ultracentrifuge. Gradients of 3-25% sucrose were formed by layering 200 µl steps of 3, 5, 10, 15, 20, and 25% sucrose in TNE and immediately loaded with 200 µl of lysate and spun at 55,000 rpm for 4 hrs as described above. For both sucrose gradients, 14- 100 µl fractions were collected from the top of the gradient and analyzed by immunoprecipitation and/or SDS-PAGE.

Immunoprecipitations and SDS/PAGE. ^{35}S-labeled core proteins were immunoprecipitated from ≈30 µl of each gradient fraction under native conditions, i.e. without SDS or boiling, with either a polyclonal anti-core antibody (Dako) or the monoclonal anti-L MA 18/7 (15). Immune complexes were collected with zysobin (Zymed); after washing, the pellet was resuspended in sample buffer and core protein released by boiling for 10 min and analyzed on 13.5% polyacrylamide gels.

RESULTS

HBV core protein oligomerization. In order to use the two-hybrid system to characterize heterologous core-core interactions in replication-permissive HepG2 cells, we first demonstrated that we could detect homo-oligomerization of full-length HBV C polypeptides in this system, and that the system faithfully reproduces the known aspects of C-C oligomerization.

HepG2 cells were co-transfected with plasmids expressing HBV C protein fused to the GAL4 DNA-binding domain (GAL-core) and to the VP16 transactivation domain (VP16-core), in the presence of a CAT reporter gene bearing five GAL4 DNA-binding sites. If C-C oligomerization occurs, then the VP16 activation domain will be targeted to the DNA region 5' of CAT and will activate CAT expression. Table 1 summarizes the results of CAT assays performed following these co-transfections. When the GAL4-core fusion was co-expressed with the VP16 domain alone, no increase in CAT activity occurred. Similarly, the VP16-core fusion did not associate with the GAL4 domain to boost CAT expression. These controls indicate that HBV core does not associate with either the GAL4 or VP16 domains; furthermore, core by itself has no transactivating potential, nor can it bind to and activate the test promoter when it is fused to VP16. As a positive control, we showed that the α and β subunits of the actin-capping protein CapZ, which are known to form an α–β dimer, associated to increase CAT activity 40-fold above background levels, as shown previously (8). When GAL4-C and VP16-C clones were co-expressed, CAT gene expression increased 40-fold. This results demonstrates that HBV core protein can associate efficiently in the two-hybrid assay, thus enabling VP16 to transactivate the CAT reporter. The C-C interaction was specific, as shown by a lack of CAT activation when the core fusion was co-expressed with the CapZ fusion protein.

With this assay, we then evaluated the effect of truncations at the carboxy-terminus of C on its ability to associate with the full-length core. As summarized in Table 2, GAL4 fused to the first 144 a.a. of core and without its 41 a.a. arginine-rich tail was nonetheless able to complex with the full-length VP16-core fusion as efficiently as the full-length GAL4-core. However, when we engineered a 68 a.a. deletion from the C-terminus, this truncated core fusion failed to interact with full-length core to activate CAT expression. This behavior closely parallels the known behavior of similar deletions in C when assayed by capsid formation in bacteria or cultured hepatocyte (3,14).

Unlike C-terminal deletions in core, the effect of N-terminal deletions on core-core association has not been reported. We engineered a pair of deletions and tested their interactions with full-length core. A GAL4-core fusion lacking the first 28 residues and the C-terminal 15 residues failed to oligomerize with core, as did a VP16-core fusion missing the first 37 a.a. from core. However, we do not know whether these N-terminal residues participate in core-core interactions (either directly or indirectly) or whether their absence merely disrupts the folding or stability of the fusion proteins used in this assay.

Interactions among mammalian HBV core proteins. We next used this system to determine whether core proteins from the different mammalian hepadnaviruses could hetero-oligomerize. The mammalian C proteins display significant amino acid sequence conservation. HBV-C and WHV-C differ in 46 residues; WHV-C and GSHV-C have diverged only slightly and differ in just 11 residues (18). Failure of the different mammalian C proteins to associate would imply that amino acids critical for core-core interaction have not been conserved among the viruses, whereas hetero-oligomerization would indicate that the non-conserved residues are not essential for this process.

We constructed both GAL4 and VP16 fusion proteins with WHV or GSHV cores. The results of the two-hybrid assays are shown in Table 3. Like the HBV core fusions, the WHV constructs homo-oligomerized and activated CAT gene expression. However, the assay failed to detect homo-oligomerization between the GSHV core fusions. Interestingly, WHV-C fusions could hetero-oligomerize with their HBV counterparts. For the GSHV-C constructs, only the GAL4-GSHV-C fusion was able to hetero-oligomerize with both its HBV-C and WHV-C partners and activate CAT expression. This observation strongly argues that the reason for the failure to detect GSHV-C homo-oligomerization in this assay is that the VP16-GSHV core fusion is either unstable or improperly folded. All attempts to detect DHBV-C homo-oligomerization failed, again most likely due to improper folding or instability of the fusion proteins. The ability of the three mammalian hepadnavirus core proteins to associate with each other points to the non-conserved core residues as inessential for core oligomerization.

Assembly of an epitope-tagged HBV core protein in *Xenopus* oocytes. Our results with the two-hybrid assay indicate that hetero-oligomers can form from cores of different species, but do not establish that such mixed oligomers can proceed to capsid formation. To do this, we co-expressed different core proteins in the *Xenopus* oocyte and assayed for mixed capsids by sucrose gradient purification and immunoprecipitation. Because WHV-C and GSHV-C are also efficiently recognized by the polyclonal HBV core antibody (Fig. 1, lanes 4-6), we first constructed an epitope tagged, truncated HBV-C to serve as the HBV partner in co-expression studies with animal hepadnavirus cores; this allows an immunologic and electrophoretic distinction to be made between the two core species. Δ157 is a truncated HBV-C comprising the first 157 residues fused to the last 2 residues of HBV-C and assembles into capsids as efficiently as the full-length core. The 12 a.a. L epitope tag is from the HBV large surface antigen and was inserted into full-length HBV-C (to generate HBV-CL) and into Δ157 (generating Δ157L) between residues 78 and 80, in a predicted surface loop of the core protein that has previously been shown to be tolerant of insertions (4). To confirm that L epitope-tagged HBV C proteins could be specifically recognized by anti-L mAb, we expressed wild type C protein of GSHV and WHV and L-tagged HBV-C individually and examined their reactivity with mAb 18/7 by immunoprecipitation. As shown in Figure 1 (lanes 1-3), only HBV-CL was precipitable with this antibody.

To confirm that Δ157L can assemble into capsids that contain exposed L epitopes on their surface, we expressed this construct in *Xenopus* oocytes and fractionated the oocyte extracts on a step gradient of 10 to 60% sucrose. Fourteen fractions were collected from the gradient and analyzed by immunoprecipitation followed by SDS-PAGE . The gradient profile of Δ157L (Fig. 2A) clearly shows dimer (fractions 1-4) and capsid (fractions 7-12) peaks that closely mirror the profile produced by wild type HBV-C [Fig. 2B and (20,21)], indicating that Δ157L indeed retains assembly-competence. Immunoprecipitation under native conditions with anti-L verified that the epitope was exposed on the particle surface; core particles lacking the tag were not brought down under these conditions (data not shown). Similar gradient profiles of DHBV-C (Fig. 2C), WHV-C and GSHV-C (data not shown) confirm that they also assemble efficiently into capsids in *Xenopus* oocytes.

Mammalian hepadnaviral cores form mixed capsids. We next determined whether hybrid capsid particles composed of C polypeptides from different hepadnaviruses could be assembled in *Xenopus* oocytes. Messenger RNA encoding Δ157L was co-injected into oocytes along with equal quantities of either WHV-C, GSHV-C, or DHBV-C transcripts. ^{35}S-lysates were fractionated on 10 to 60% sucrose gradients into 14 fractions. Fractions 8-12, representing assembled capsids, were combined and immunoprecipitated with either anti-L mAb 18/7, anti-HBV-C or anti-DHBV-C antibody. The labelled proteins were analyzed by SDS-PAGE (Fig. 3A). Immunoprecipitation with anti-L clearly demonstrated that Δ157L core protein forms capsids with both WHV-C and GSHV-C (lanes 1 and 4). The ratio of WHV:HBV and GSHV:HBV chains in the anti-L precipitates (lanes 1 and 4) is similar to that in the anti-HBV-C precipitates (lanes 2 and 5) and to that in the total lysate fraction (not shown). In addition, the absolute amount of

WHV or GSHV core chains precipitated by anti-L is comparable to that precipitated by polyclonal anti-HBV-C antibodies. Together, these observations indicate that most or all of the animal hepadnaviral core polypeptides in these fractions are incorporated into mixed capsids.

Analysis of co-injections of Δ157L with DHBV-C transcripts yielded very different results. Anti-L and anti-HBV-C, neither of which recognizes DHBV-C, immunoprecipitated only Δ157L (lanes 7-8); similarly, anti-DHBV-C antibody only immunoprecipitated DHBV-C (lane 9). This shows that, unlike the WHV and GSHV core proteins, DHBV-C and HBV-C cannot co-assemble; when co-expressed, they yield separate homo-oligomeric capsid populations. This inability of mammalian and avian hepadnavirus cores to interact is not entirely unexpected given the significant divergence in their core protein sequences and the lack of immuno-cross-reactivity.

In order to rule out the formal possibility that the results supporting mixed capsid formation are due to nonspecific aggregations of different homocapsids immunoprecipitated by anti-L, we performed a mixing experiment control (Fig. 3B). Oocytes were injected with either HBV (Δ157L), WHV, or GSHV-C mRNA, so that only homo-oligomeric capsids can form. Pairs of lysates containing one tagged and one untagged C polypeptide species were then combined and subjected to sucrose gradient fractionation; capsid fractions were then immunoprecipitated as described earlier. Unlike the co-injection experiments, under these conditions L antibody was able to immunoprecipitate very little WHV-C and GSHV-C (lanes 1 and 4) when compared to anti-HBV-C immunoprecipitation (lanes 2 and 5). Not surprisingly, the mixing experiment involving DHBV capsids and HBV capsids likewise showed no interaction (lanes 7-9). This mixing experiment indicates that only a small percentage of the signal indicating mixed capsid formation in the co-injection experiment is attributable to nonspecific interactions of homo-oligomeric capsids.

Hepadnaviral core dimerization is cis-preferential. These results point to the capability of different mammalian hepadnaviral C proteins to form hetero-oligomers, but do not indicate whether these are dimers, tetramers or even higher-order multimers. In authentic capsid assembly, dimers are the initial intermediates in assembly. We next set out to determine whether heterodimers in fact form during mixed capsid assembly in oocytes. To do this, we examined gradient-purified dimer intermediates fractionated from the co-injected oocytes described in the previous section. Immunoprecipitation with anti-L revealed that Δ157L dimerized primarily with itself and formed few heterodimers with either WHV-C or GSHV-C (Fig. 4, lanes 1 and 4 vs. lanes 2 and 5). We wondered if this result was due to monomer affinity differences arising from sequence divergence that favored homodimer over heterodimer formation, or whether the dimerization process might be cis-preferential--i.e. that it might favor the association of monomers translated from the same mRNA.

To distinguish between these two possibilities, we co-injected oocytes with SP6 transcripts encoding HBV-C and HBV-CL: here, both proteins are identical except for the presence of the L epitope tag. Both proteins assemble into capsids with comparable efficiencies (Data not shown). Co-injected lysates were fractionated on 3-25% sucrose gradients into 14 fractions. Each fraction was then immunoprecipitated with anti-L or anti-HBV-C and analyzed by SDS-PAGE. Under these sedimentation conditions, dimer intermediates sediment in fractions 3-10 and capsids pellet in fraction 14 (21). Co-precipitation of both chains with anti-L in fraction 14 comfirms that mixed capsids form as expected (Fig. 5A). But precipitation of the dimer fractions with anti-L clearly demonstrated that HBV-CL forms dimers primarily with itself rather than with HBV-C (lanes 3-10), despite the presence of HBV-C in these fractions, as revealed by anti-HBV-C (Fig. 5B, lanes 3-9). Thus, the mixed capsids assembled in the co-injected oocytes (Fig. 5A and B, lanes 14) must be composed of predominantly homodimers of tagged and untagged C chains.

Using the phosphorimager to quantitate the amount of HBV-C and HBV-CL produced, we estimated that approximately 35% of dimers found in the co-injection experiment should have been heterodimers, assuming free and random association of HBV-C and HBV-CL monomers leading to a binomial distribution of dimer products. In fact, if we correct for antibody cross-reactivity and immunoprecipitation efficiencies, our results show that heterodimers account for no more than 7% of the dimer pool; moreover, in our view this figure represents an upper estimate. These findings indicate that core protein displays a marked cis-preference for dimerization in the oocyte. Perhaps the rate at which monomers associate to form dimers is so fast that it occurs immediately after monomers fold or even concurrently with translation and folding. This would explain both the preponderance of homodimers and the absence of detectable free core monomers in the oocyte system.

DISCUSSION

Sequence requirements for core-core interactions. We have used the two-hybrid system in a permissive human hepatoma cell line (HepG2) to characterize core protein interactions involved in HBV capsid assembly. We have shown that the full-length core protein fused downstream of the GAL4 or VP16 domains is able to associate with itself; this association, which normally occurs in the cytoplasm, must have stably persisted in the nucleus, where the host RNA polymerase II is located.

Core fusions with a truncation of the C-terminal 41 a.a. also oligomerized with full-length cores, as predicted from *in vivo* capsid assembly experiments. A deletion that extended 68 a.a. upstream from the C-terminus failed to interact; this result also agrees with findings from earlier assembly studies (2,3). The effect of N-terminal deletions on core-core association has not been previously reported. We engineered core fusions lacking the

first 28 and 37 residues from core; neither deletion construct oligomerized with the full-length core fusion. Unfortunately, one of the major limitations in using the two-hybrid assay is that only fusions that yield a positive CAT signal are informative. Fusions that fail to associate with core in this assay may indeed lack a domain critical for core interaction, or may have simply folded into an improper conformation or be unstable. The fact that the full-length and the arg-tail deficient core fusions were able to associate suggests that the GAL4 and VP16 domains are discrete structures that did not interfere with the folding of the C portion of the fusion. We favor the view that the assembly-defective deletions may have globally disrupted the secondary and tertiary structure of the core polypeptide. Although the core polypeptide structure has not been elucidated by crystallography, Argos and Fuller have proposed a structure based on modeling with the vp3 capsid protein of a picornavirus (1). Their sequence alignment analysis suggests that the core protein will adopt an eight-stranded antiparallel β-barrel structure, with most of the arg-rich tail located distally. Both of the N-terminal deletions and the 68 a.a. C-terminal deletion would be predicted to remove at least two of the β-strands and probably disrupted the secondary and tertiary structure important for intersubunit contact and association. The structure of core may preclude the identification of a discrete contiguous domain that mediates core association; residues critical to forming a contact surface during dimerization may be brought together only after the proper folding of core.

Assembly of mixed capsids in the *Xenopus* oocyte. Co-expression of Δ157L, a truncated and tagged HBV-C, with WHV-C or GSHV-C led to the assembly of mixed capsid particles. Not surprisingly, co-expression of Δ157L and DHBV-C, which differs considerably in sequence and size from the mammalian cores, did not lead to the formation of mixed capsids. Assembly of HBV-C into capsids in the oocyte proceeds without detectable assembly intermediates other than dimers; our co-expression experiments with the animal viral cores also failed to reveal any higher MW or blocked assembly

intermediates in hetero-oligomeric capsid formation. A number of explanations for this are possible. Most likely, such intermediates are extremely transient due to the efficient and highly cooperative assembly of capsids from dimer building blocks (21); alternatively they may be unstable in extracts or may dissociate during centrifugation.

Our results with the oocyte assembly system raise the intruiging issue of whether mixed nucleocapsids can form during hepadnaviral replication. Such a situation could occur *in vivo* for WHV and GSHV, since both can replicate in woodchuck hosts. Hybrid nucleocapsids might affect several processes in the hepadnaviral life cycle. Cores are intimately involved in RNA encapsidation, DNA synthesis, and viral budding. How a mixed nucleocapsid would affect the specificity or efficiency of any one of these process remains to be seen. There is little precedent for mixed nucleocapsid formation in other viruses with icosahedral capsids. In retroviruses, a recent report has demonstrated that mixed core particles between murine leukemia virus and human immunodeficiency virus type 1 can be made, but only when artificial chimeric Gag precursors containing HIV and MuLV determinants are co-expressed with wild-type MuLV Gag proteins (7). This suggests that homo-dimerization of the MuLV Gag proteins is the driving force for assembly, with the heterologous HIV sequences being incorporated as a result. As yet there have been no reports of mixed capsid formation between wild-type Gag proteins.

A cis-preference in dimer formation. During the mixed capsid assembly reaction in *Xenopus* oocytes, the co-expressed core proteins formed primarily homodimers. Even when the co-expressed C polypeptides are identical except for the L epitope tag, homodimers still form preferentially. These results indicate that core dimerization in the oocyte displays a strong cis preference. This suggests that core monomers are not free to diffuse and associate with other monomers, and is consistent with the notion that dimerization occurs very rapidly after monomer production. In this view, a monomer is most likely to associate with the nearest monomer available in the region of its synthesis:

i.e. those translated from the same polysome. In the extreme case, dimers may form co-translationally between adjacent, nascent core polypeptides still tethered to the elongating ribosome. This may account for the absence of detectable monomers in the oocyte assembly system (21). Once formed, dimers are free to diffuse and associate with other dimers.

We do not know if this cis-preference operates in mammalian cells, but this seems likely to be so. Our results in the mammalian two-hybrid system certainly do not exclude this possibility. Since GAL4 binds to the DNA as a dimer (12), it is likely that the GAL4-core fusion proteins active in this assay are themselves dimeric, and that all C-C interactions being detected in this system thus involve homodimers of C proteins.

ACKNOWLEDGMENTS

We thank Sophie Roy and Jon Pollack for many helpful discussions during the course of this work, and Ian Taylor for critical reading of the manuscript. This work was supported by grants from the National Institutes of Health.

TABLE 1 Full length HBV core protein oligomerizes.

For Tables 1,2, and 3) HepG2 cells were transfected with three plasmids: a fusion construct encoding the GAL4 DNA-binding domain and a downstream protein domain of interest; a second fusion construct encoding the VP16 transactivation domain and a downstream protein domain of interest; and a CAT reporter gene with five upstream GAL4 DNA-binding sites. "None" indicates the fusion construct encodes either the GAL4 or VP16 domain alone, without a downstream domain. CAT expression levels are reported as fold-increase relative to CAT activity resulting from transfection of the CAT reporter plasmid alone.

FULL LENGTH HBV CORE PROTEIN OLIGOMERIZES

GAL4 fusion	VP16 fusion	Relative CAT Activity
HBc	None	1
None	HBc	1
Actin BP beta25	Actin BP alpha 49	40
Actin BP beta25	HBc	1.5
HBc	HBc	40

TABLE 2 C-terminal and N-terminal deletions in HBV core protein.

C-TERMINAL AND N-TERMINAL DELETIONS IN HBV CORE PROTEIN

GAL4 fusion	VP16 fusion	Relative CAT Activity
HBc	HBc	22
HBc1-144	HBc	23
HBc1-117	HBc	1
HBc29-168	HBc	1
HBc	HBc38-183	1

TABLE 3 Interactions among mammalian HBV core proteins.

INTERACTIONS AMONG MAMMALIAN HBV CORE PROTEINS

GAL4 fusion	VP16 fusion	Relative CAT Activity
HBc	HBc	23
WHc	WHc	18
GSHc	GSHc	1
WHc	HBc	72
HBc	WHc	5
GSHc	HBc	20
HBc	GSHc	1
GSHc	WHc	5
WHc	GSHc	1

FIG. 1 Analysis of cross reactivity with the anti-L mAb and the anti-HBV-C polyclonal Ab. Lysates from [^{35}S]-labelled *Xenopus* oocytes injected with the appropriate mRNA were resolved on step gradients of 10-60% sucrose. Fractions containing assembled capsids were combined and immunoprecipitated under non-denaturing conditions with the appropriate antibody and analyzed by SDS-PAGE on a 13.5% gel. The positions of pre-stained molecular mass markers are indicated at the right of the panel.

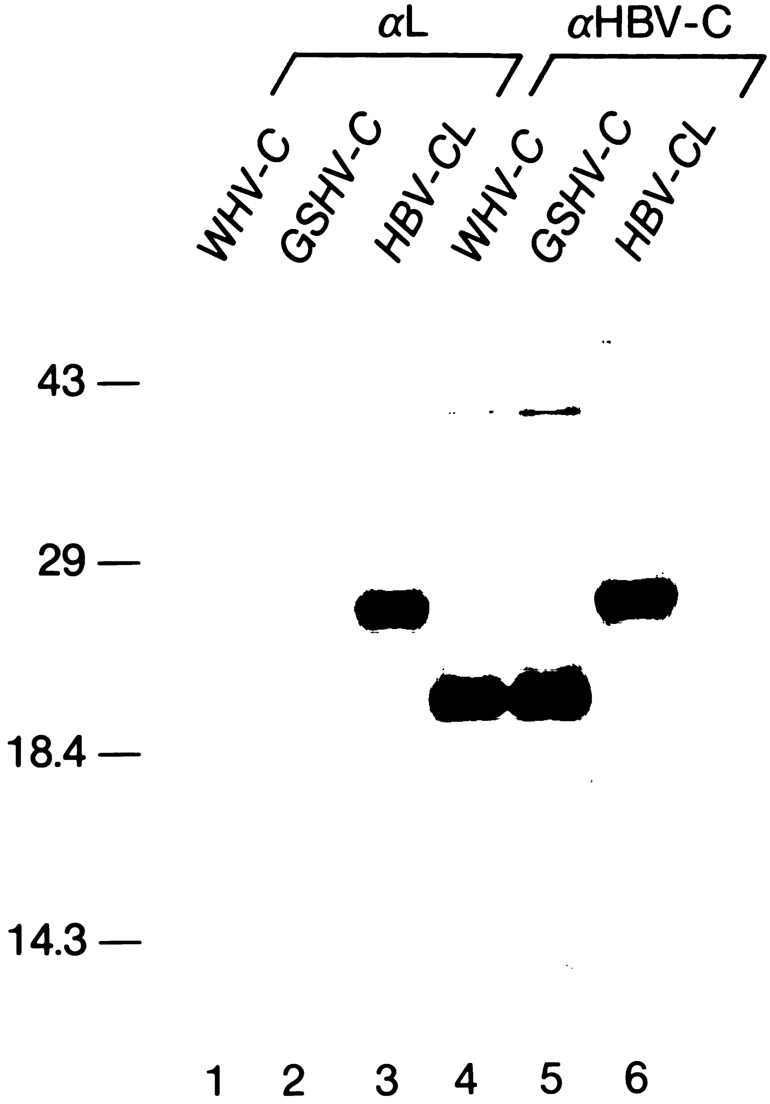

αL
αHBV-C
WHV-C
GSHV-C
HBV-CL
WHV-C
GSHV-C
HBV-CL
43
29
18.4
14.3
1
2
3
4
5
6

FIG. 2 Assembly of a tagged and truncated HBV core protein in *Xenopus* oocytes. (Top) Lysates from [^{35}S]-labelled *Xenopus* oocytes injected with mRNA encoding HBV-C, DHBV-C, or Δ157L were combined and resolved on step gradients of 10-60% sucrose. Fourteen fractions were collected from the top of each gradient and analyzed by immunoprecipitation under non-denaturing conditions with polyclonal anti-HBV-C and anti-DHBV-C antibodies, followed by SDS-PAGE on a 13.5% gel. The gradient fractions are shown in order from the top fraction at the left (fraction 1) to the bottom fraction at the right (fraction 14). The positions of the DHBV-C, HBV-C and Δ157L species are indicated on the left. (Bottom) Phosphorimager quantitation of [^{35}S] signal intensity in each fraction for the indicated core polypeptide band in the gel shown at top.

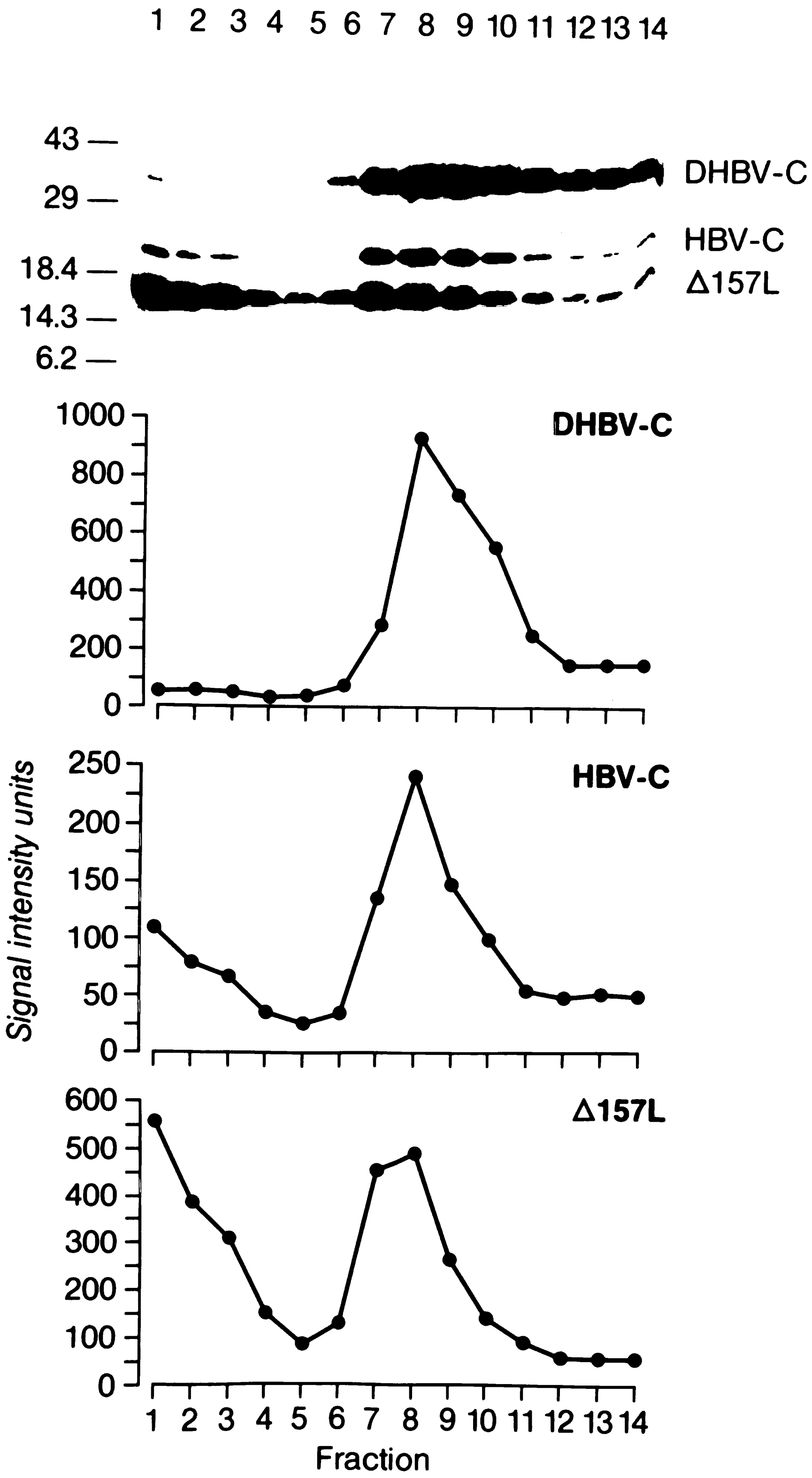

1 2 3 4 5 6 7 8 9 10 11 12 13 14
43
29
18.4
14.3
6.2
DHBV-C
HBV-C
Δ157L
Signal intensity units
1000
800
600
400
200
0
DHBV-C
250
200
150
100
50
0
HBV-C
600
500
400
300
200
100
0
Δ157L
1 2 3 4 5 6 7 8 9 10 11 12 13 14
Fraction

FIG. 3 Composition of capsid particles assembled in *Xenopus* oocytes co-injected with mRNA encoding core proteins from different hepadnaviruses. A) Oocytes were co-injected with mRNAs encoding Δ157L and either WHV-C, GSHV-C, or DHBV-C. [^{35}S]-lysates were resolved on step gradients of 10-60% sucrose. Fractions (no.8-12) containing assembled capsids were combined and immunoprecipitated under non-denaturing conditions with either anti-L mAb (αL), anti-HBV-C (αH), or anti-DHBV-C (αD) polyclonal antibodies. The labelled precipitate was analyzed by SDS-PAGE. The positions of the DHBV-C, WHV-C, GSHV-C, and Δ157L species are indicated on the right.

[illegible]

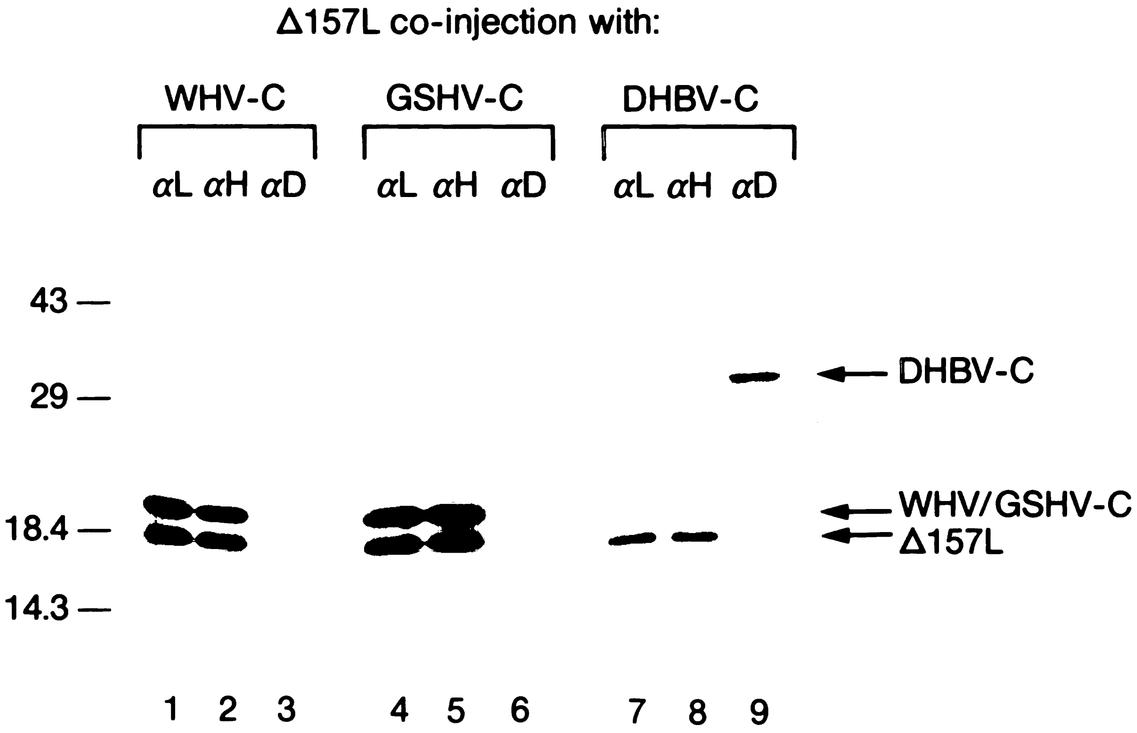

Δ157L co-injection with:
WHV-C
GSHV-C
DHBV-C
αL αH αD
αL αH αD
αL αH αD
43 —
29 —
18.4 —
14.3 —
DHBV-C
WHV/GSHV-C
Δ157L
1 2 3
4 5 6
7 8 9

FIG. 3

B) Analysis of mixed extracts. Oocytes were injected with a single transcript. [^{35}S]-lysates were combined in the indicated pairings, and subject to sucrose gradient fractionation and immunoprecipitation as described in 3A.

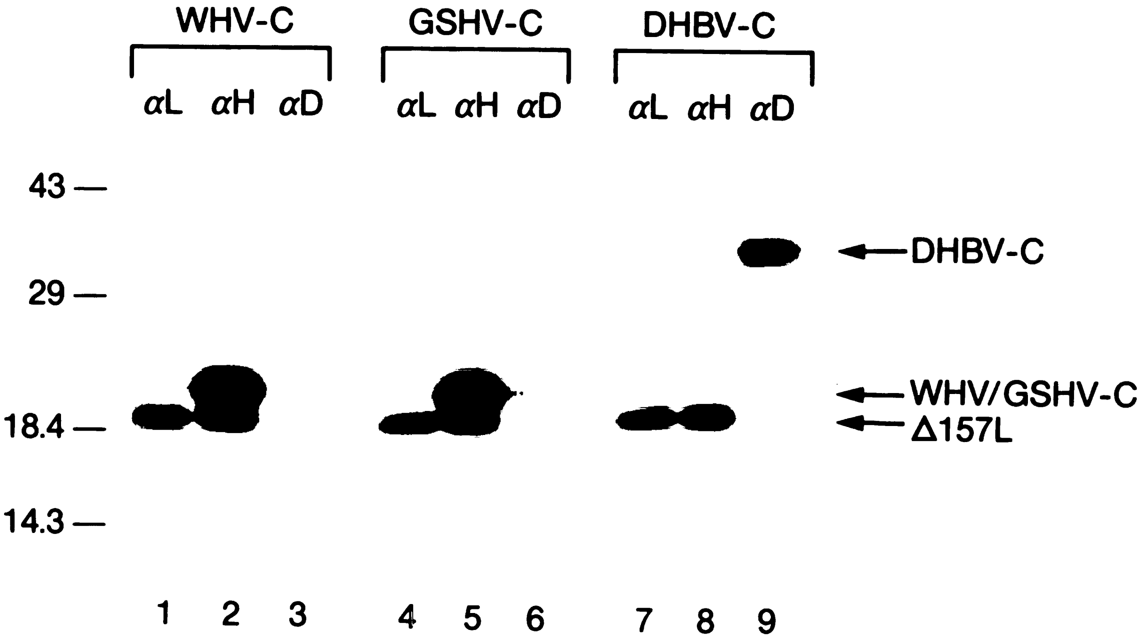
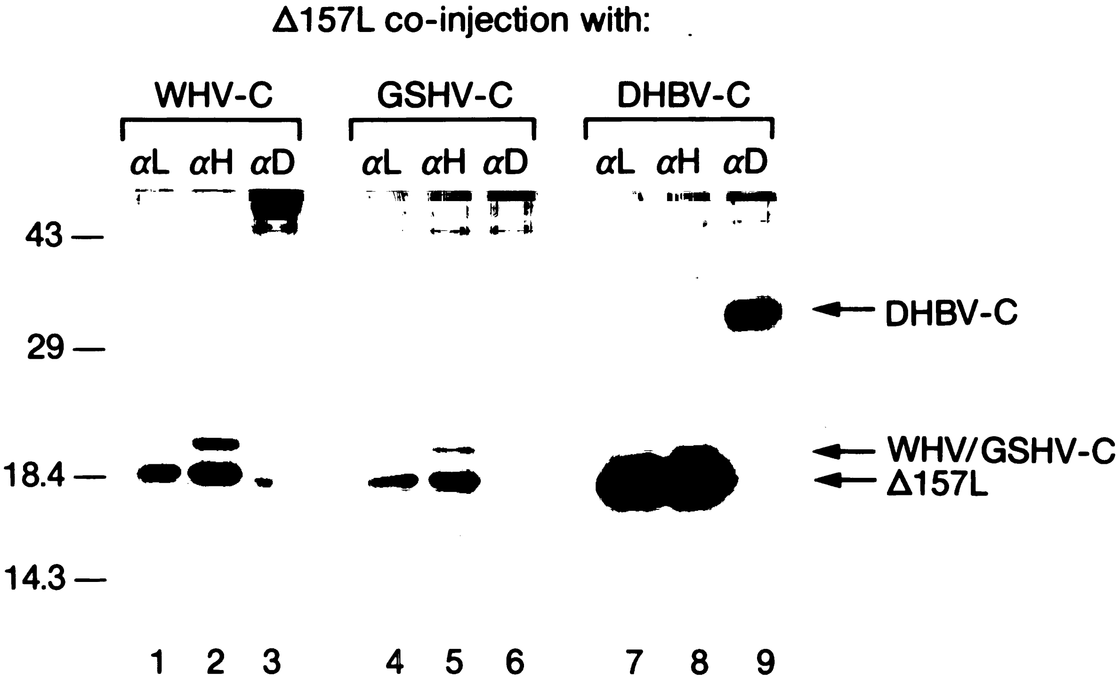

WHV-C
GSHV-C
DHBV-C
αL αH αD
αL αH αD
αL αH αD
43
29
18.4
14.3
DHBV-C
WHV/GSHV-C
Δ157L
1 2 3
4 5 6
7 8 9

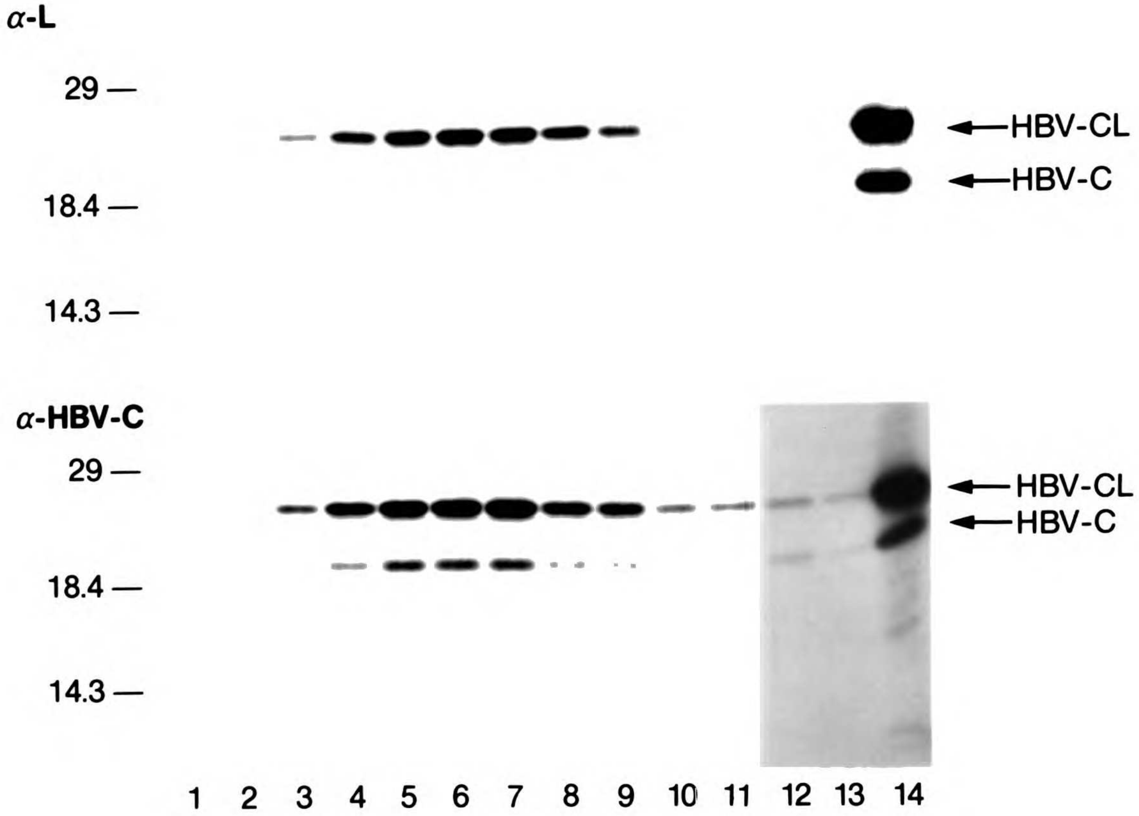
FIG. 4 Composition of dimer intermediates assembled in *Xenopus* oocytes co-injected with mRNA encoding core proteins from different hepadnaviruses. Fractions (no. 1-5, described in Fig. 3A) containing core dimer intermediates were analyzed as outlined in that section.

KGS LIBRARY


Δ157L co-injection with:
WHV-C
GSHV-C
DHBV-C
αL αH αD
αL αH αD
αL αH αD
43 —
29 —
18.4 —
14.3 —
DHBV-C
WHV/GSHV-C
Δ157L
1 2 3 4 5 6 7 8 9

UCSF LIBRARY

FIG. 5 Composition of dimer intermediates formed in *Xenopus* oocytes co-expressing tagged and untagged HBV-C. Oocytes were co-injected with mRNAs encoding HBV-C and the L-epitope tagged HBV-CL. [^{35}S]-lysates were resolved on 3-25% sucrose step gradients. Each of the fourteen fractions were immunoprecipitated under non-denaturing conditions with either A) anti-L mAb, or B) anti-HBV-C polyclonal Ab, and analyzed by SDS-PAGE. Note that dimer intermediates sediment in fractions 3-10, and capsids pellet at the bottom (fraction 14).


α-L
29
18.4
14.3
HBV-CL
HBV-C
α-HBV-C
29
18.4
14.3
HBV-CL
HBV-C
1 2 3 4 5 6 7 8 9 10 11 12 13 14

REFERENCES

1. Argos, P., and S.D. Fuller. 1988. A model for the hepatitis B virus core protein: prediction of antigenic sites and relationship to RNA virus capsid proteins. EMBO J. 7: 819-824.

2. Beames, B. and R.Langford. 1993. Carboy-terminal truncations of the HBV core protein affect capsid formation and the apparent size of encapsidated HBV RNA. Virology 194: 597-607.

3. Birnbaum, F. and M.Nassal. 1990. Hepatitis B virus nucleocapsid assembly: primary structure requirements in the core protein. J.Virology 64: 3319-3330.

4. Borisova, G., B.Arya, A.Dislers, O.Barschukova, V.Tsibinogin, D.Skrastina, M.Eldarov, P.Pumpens, K.Skryabin, and E.Grens. 1991. Hybrid hepatitis B virus nucleocapsid bearing an immunodominant region from hepatitis B virus surface antigen. J.Virology 67: 3696-3701.

5. Clarke, B.E., S.E.Newton, A.R.Carrol, M.J.Francis, G.Appleyard,A.D.Syred, P.E.Highfield, D.J.Rowlands, and F.Brown. 1987. Improved immunogenicity of a peptide epitope after fusion to hepatitis B core protein. Nature 330: 381-384.

6. Cohen, B., and J.Richmond. 1982. Electron microscopy of hepatitis B core antigen synthesized in E. coli. Nature 296: 677-678.

7. Deminie,C.A., and M.Emerman. 1993. Incorporation of human immunodeficiency virus type 1 gag protein into murine leukemia virus virions. J.Virology 67: 6499-6506.

8. Fearon, E.R., T.Finkel, M.L.Gillison, S.P.Kennedy, J.F.Casella, G.F.Tomaselli, J.S.Morrow, and C.V.Dang. 1992. Karyoplasmic interaction selection strategy: a general strategy to detect protein-protein interactions in mammalian cells. PNAS 89: 7958-7962.

9. Fields, S., and O.Song. 1989. A novel genetic system to detect protein-protein interactions. Nature 340: 245-246.

10. Hirsch, R., R.Colgrove, and D.Ganem. 1988. Replication of duck hepatitis B virus in two differentiated human hepatoma cell lines after transfection with cloned viral DNA. Virology 167: 136-142.

11. Luban, J., K.B.Alin, K.L.Bossolt, T.Humaran, and S.P.Goff. 1992. Genetic assay for multimerization of retroviral gag polyproteins. J.Virology 66: 5157-5160.

12. Marmorstein, R., M.Carey, M.Ptashne, and S.C.Harrison. 1992. DNA recognition by GAL4: structure of a protein-DNA complex. Nature 356: 408-414,

13. Miyonahara, A., T.Imamura, M.Araki, K.Sugawara, N.Ohtomo, and K.Matsubara. 1986. Expression of hepatitis B virus core antigen gene in Saccharomyces cerevisiae: synthesis of two polypeptides translated from different initiation codons. J.Virology 59: 176-180.

USGS LIBRARY

14. Nassal, M. 1992. The arginine-rich domain of the hepatitis B virus core protein is required for pregenome encapsidation and productive viral positive-strand DNA synthesis but not for virus assembly. J.Virology 66: 4107-4116.

15. Seifer, M., and D.N.Standring. 1993. Recombinant human hepatitis B virus reverse transcriptase is active in the absence of the nucleocapsid or the viral replication origin, DR1. J.Virology 67: 4513-4520.

16. Stahl, S.J., and K.Murray. 1989. Immunogenicity of peptide fusions to hepatitis B virus core antigen. PNAS 86: 6283-6287.

17. Standring, D.N., J.-H.Ou, and W.J.Rutter. 1986. Assembly of viral particles in Xenopus oocytes: pre-surface antigens regulate secretion of the hepatitis B viral surface envelope particle. PNAS 83: 9338-9342.

18. Standring, D.N. 1991. The molecular biology of the hepatitis B virus core protein, in Molecular Biology of the Hepatitis B Virus, ed.McLachlan,A. (CRC, Boca Raton, FL) pp.145-169.

19. Zhou, S., and D.N.Standring. 1991. Production of hepatitis B virus nucleocapsidlike core particles in Xenopus oocytes: assembly occurs mainly in the cytoplasm and does not require the nucleus. J.Virology 65: 5457-5464.

20. Zhou, S., S.Q.Yang, and D.N.Standring. 1992. Characterization of hepatitis B virus capsid particle assembly in Xenopus oocytes. J.Virology 66: 3086-3092.

21. Zhou, S., and D.N.Standring. 1992. Hepatitis B virus capsid particles are assembled from core-protein dimer precursors. PNAS 89: 10046-10050.

CHAPTER 4

IDENTIFICATION OF SEQUENCES WHICH AFFECT PREGENOMIC RNA ACCUMULATION OF THE DUCK HEPATITIS B VIRUS

Christopher Chang[1] and Don Ganem[2,3]

Department of Biochemistry and Biophysics[1], Howard Hughes Medical Institute[2], and Departments of Microbiology and Immunology, and Medicine[3], University of California Medical Center, San Francisco, California 94143-0502.

ABSTRACT

The pregenomic RNA of hepadnaviruses serves as both the mRNA for the core and polymerase proteins and the RNA template for reverse transcription and genome replication. We have identified a region through deletion analysis in the duck hepatitis B virus (DHBV) pregenomic RNA that is critical for its accumulation in a transfected chicken hepatoma cell line, as well as in cell lines of non-hepatic origin. This 80 nt. region, termed α, is located near the pregenomic cap site. α-deletion mutants yielded drastically reduced steady state levels of pregenomic RNA. However, this phenotype is suppressed by deletion of a second, larger region located 1 kb. downstream. The activity of the α element displays strong orientation dependence and is influenced by its position within the transcript.

INTRODUCTION

Hepadnaviruses are a family of small, hepatotropic DNA viruses that replicate primarily in hepatocytes and can cause persistent infection in a variety of animal hosts (reviewed in 6). Members of this family include human hepatitis B virus, woodchuck hepatitis virus (19), ground squirrel hepatitis virus (11), and duck hepatitis B virus (DHBV) (12). During an infection, their DNA genome is replicated via a reverse transcription pathway involving an RNA intermediate (20). The viral genome (Fig. 1) is very small (3.0-3.2 kb.) and covered with extensively overlapping open reading frames (ORFs). The DHBV genome contains 3 ORFs, encoding the viral surface antigen (S), the core protein (C), and the viral polymerase (P). Both the S and C ORFs are preceded by short, in-frame ORFs (termed PreS and PreC) which specify larger related polypeptides. Following infection of the hepatocyte, the partially double stranded DHBV genome is delivered to the nucleus, where it is converted into a covalently closed circular DNA structure. This 3.0 kb molecule serves as the template for the transcription of two major classes of viral RNAs, a 3.3 kb pregenomic RNA (pgRNA) and 2.1 and 1.8 kb subgenomic transcripts (3)(Fig. 1). The pregenome RNA encodes the C and P proteins and is selectively encapsidated along with the P protein into virus cores (1,2), where reverse transcription to produce the DNA genome takes place (20). The subgenomic transcripts serve as mRNA for the expression of the PreS and S envelope glycoproteins. The 3.3 kb. genomic transcripts comprise the entire genome length and include a short terminal redundancy which duplicates sequence elements that are critical for replication and DNA circularization (10,16,23). The sequence of the subgenomic transcripts are completely contained within the pregenomic RNA, and both are polyadenylated at the same site.

In this report, we identify a region through deletion analysis that is critical for pgRNA accumulation in transfected LMH cells. This 80 nt. region is located near the pregenome cap site; mutants lacking this element yield drastically reduced steady state

INST LIBRARY

levels of pgRNA. However, this phenotype is suppressed when a second region of about 170 nt. located 1 kb downstream is deleted.

MATERIALS AND METHODS

Materials. Enzymes were purchased from New England BioLabs, unless otherwise noted, and used according to the manufacturer's protocols. Radionucleotides were purchased from Amersham Corp.

Plasmids and plasmid construction. DHBV nucleotide positions are numbered from the unique EcoRI site of DHBV3 (18). In this numbering scheme, nt. 2529 is the transcription initiation site of pregenome RNA (3). Standard molecular biology techniques were used in constructing the various plasmids. Ligation of incompatible ends was carried out after removal of protruding 3' ends or filling in of recessed 3' ends with the Klenow fragment of DNA polymerase I. Overlength DHBV genomes are generally constructed in two steps, with a less than full length genome construct as an intermediate.

pD1.5G is an overlength DHBV genome between nt.1658-3021. CMVDHBV is also an overlength construct driven by the CMV immediate early promoter. It has been described earlier and was kindly provided by Dr.W. Mason (4). pPGpol was constructed by digesting pD0.5G, a partial DHBV genome between nt. 1658-3021, with AflII and EcoRI, and then ligated with a 3 kb. HindIII-EcoRI fragment. pCMVpol has been described elsewhere (8) and is essentially an EcoRI DHBV monomer located downstream of the CMV promoter.

CMVΔ1 was made by inserting a 1.7 kb. XbaI fragment from DHBV (nt. 2662-1358) into the XbaI site in the multiple cloning site (MCS) of pcDNA-1 (Invitrogen Corp.). This intermediate, termed CMVXbaI, was cut with BglII and ApaI(in MCS) and ligated

[illegible] LIBRARY

with the 2.6 kb. BglII-EcoRI fragment (nt. 391-3021). CMVΔ2 was generated by subcloning the AflII-EcoRI fragment (nt. 2526-3021) into the vector fragment of a BamHI/EcoRI digest of pcDNA-1. This intermediate, termed CMVAfl-RI, then had an NsiI DHBV monomer subcloned into its XbaI site (nt. 2662). CMVΔ3 was constructed by subcloning an EcoRI DHBV monomer into its NsiI site (nt. 2845) of CMVAfl-RI. CMVΔ4 was made by cutting CMVAfl-RI with AflII and AccI to generate a deletion between these sites, and then subcloning an EcoRI monomer into the EcoRI site at nt. 3021. CMVΔ5 was generated by a partial XbaI and complete AccI digest of CMVAfl-RI to generated the deletion between nt. 2581-2662; an EcoRI monomer was then subcloned into the EcoRI site at nt. 3021.

CMVΔ5/Δ6 and CMVΔ5/Δ7 were generated by partial digestion of CMVΔ5 with SspI and complete digestion with BglII, ligation, and selection for clones bearing the appropriate deletions between nt. 391-561 and nt. 391-675, respectively. CMVΔ1/NsiI was made by subcloning the 140 bp. AflII-XbaI fragment (nt. 2526-2666) into the NsiI site of CMVXbaI. This intermediate was then cut with BglII and ApaI and ligated with the BglII-EcoRI fragment described above. CMVΔ1/BglII and CMVΔ1/BglIIinv were constructed by subcloning the same 140 bp. fragment into the BglII site (nt.391) of CMVΔ1. CMVΔ1/KpnI was generated by subcloning the same 140 bp. fragment into the KpnI site (nt. 1290) of CMVXbaI. A HindIII-EcoRI DHBV fragment (nt.1334-3021) was then cloned into the ApaI site of the MCS of this intermediate. This generated a tandem 25 bp. repeat at the site of insertion (nt.1334-1358).

Cell culture and transfections. LMH chicken hepatoma cells (4) were grown in Dulbecco's minimum essential medium-Ham's nutrtient mixture F12 supplemented with 10% fetal calf serum and were passaged every 3-4 days. DNA transfections were performed by the calcium phosphate coprecipitation method as described previously (7). Transfections were carried out with 13 μg of each construct DNA and 2 μg of a human

UCSF LIBRARY

growth hormone or β-galactosidase reporter plasmid per 100mm diameter tissue culture dish. The concentration of growth hormone secreted into the media was measured using a commercial kit (Nicholls). β-galactosidase activity was measured as previously described (15).

RNA preparation and hybridization analysis. Poly (A)$^+$ total cellular RNA was extracted from LMH cells 48 hours post-transfection as described previously (7). In brief, cells were lysed in 1% SDS and proteinase K (250mg/ml), and their DNA sheared by multiple passages through a 26- guage needle. Poly (A)$^+$ RNA was then selected by annealing to oligo(dT)-cellulose. Appropriate fractions of RNA were electrophoresed through a 1% agarose/2.2 M formaldehyde gel, transferred to a nylon membrane (Hybond; Amersham) and then annealed to radiolabled DNA probes prepared by random primed DNA synthesis of the DHBV EcoRI monomer fragment.

RESULTS

Deletion analysis of 5' pregenome sequence. We first appreciated the significance of sequences at the 5' end of pgRNA when we constructed deletion mutants designed to supply the polymerase gene product (P) *in trans* to P-deficient DHBV mutants. We reasoned that DHBV mRNA which lacked both the upstream AUGs in the core ORF and the packaging signal ε might translate P more efficiently and supply its product *in trans* more readily than the wild type pregenome. Initially, we engineered a construct bearing a deletion of the first 505 nt. of pregenome, driven by the endogenous DHBV pregenomic promoter (PGpol) (Fig. 2A). LMH cells were cotransfected with this mutant and reporter plasmids expressing β-gal or human growth hormone; expression of the reporter genes was used to normalize for differences in transfection efficiency. Northern blot analysis of

[illegible] LIBRARY

normalized quantities of total cellular poly (A)+ RNA in LMH cells transfected with this mutant revealed a dramatic reduction in the level of deleted pgRNA transcripts when compared to wild-type pgRNA (Fig. 2B, lanes 1 and 2). To determine if the selective loss of these transcripts was due to deletion of an essential enhancer for the pregenomic promoter, we constructed a similar deletion in a pgRNA transcription unit driven by the CMV IE promoter, which is not dependent upon any hepadnaviral enhancers. Again, CMV-driven pgRNA accumulated to high levels, while the deleted transcript once again was selectively depleted (Fig. 2B, lanes 3 and 4). These findings suggest that the loss of pgRNA transcripts in this mutant was due to a defect subsequent to transcriptional initiation.

Further deletion analysis of the 5' region of pregenome localized the critical sequences to the extreme 5' end of the transcript. Removal of the first 133 nt. of pregenome (CMVΔ1, lane 5) significantly diminished pgRNA levels, while non-overlapping deletions further downstream had no effect (CMVΔ2 and CMVΔ3, lanes 6 and 7). Removal of the first 48 nt. had no effect (CMVΔ4, lane 8), but deletion of the next 85 nt., encoded by nt. 2581-2662 in mutant CMVΔ5 (lane 9) again reduced steady state pgRNA levels dramatically. Pregenome transcription of these deletion mutants was driven by the CMV promoter; identical results were obtained with similar constructs driven by the DHBV pregenomic promoter (data not shown). We note that deletions that significantly diminish pregenome message have no major effect on subgenomic mRNA levels. These results indicate that sequences between nt. 2581-2662 in the DHBV genome (referred to hereafter as α) play a critical role in allowing accumulation of pgRNA in transfected LMH cells. This behavior is not cell-type-specific: the CMV-driven deletion constructs yielded similar findings when expressed in HeLa cells or mouse L cells (data not shown).

Deletion of downstream sequences restore pregenome RNA levels. Deletion mutants CMVpol and PGpol both initiate transcription approximately 500 nt. downstream

[illegible] LIBRARY

from the pregenome cap site and direct production of transcripts that fail to accumulate. However, the PreS message of these constructs, which initiates some 1175 nt. downstream of the pgRNA cap site, and is otherwise co-extensive with pgRNA (and PGpol RNA), was present in normal quantities. This raises the possibility that sequences in this 675 nt. interval might be responsible for the phenotype of the α deletion, and that their absence in PreS mRNA accounted for its accumulation to wild type levels.

To test this idea, we first examined accumulation of CMV-driven transcripts that initiated at several sites in this interval (Fig. 3A). While RNA that initiated 660 nt. from the 5' end of pgRNA (at nt. 165) was selectively lost (Fig. 3B, lane 3), messages that initiated 220 nt. further downstream (at nt. 391) accumulated to intermediate levels (lane 4). In contrast, transcripts beginning 1050 nt. and 1170 nt. from the pregenome cap site (at nt. 561 and 675) were detected at levels comparable to subgenomic mRNA (lanes 5 and 6). We next constructed several internal deletions in CMVΔ5, which has reduced levels of pgRNA (Fig. 4A). Transfection of these deletion mutants in LMH cells and northern blot analysis of poly $(A)^+$ RNA showed that removal of either nt. 391-561 or nt. 391-675 from CMVΔ5 resulted in levels of pgRNA comparable to wild type (Fig. 4B)(Note that mutant CMVΔ5/Δ7 deleted the PreS promoter, thus accounting for the diminished PreS mRNA compared to CMVΔ5/Δ6). These results suggest that sequences approximately 1 kb. downstream from the 5' end of pgRNA contain elements (termed β) which, when deleted, restore accumulation of pregenome transcripts in α-deletion mutants.

Position and orientation-dependence of α sequences. Consideration of the structure of the above constructs suggested that the activity of the α element might display position dependence. Recall that in the terminally redundant pregenome RNA (and in our overlength plasmid DNAs), a second copy of the α sequence is located at the 3' end of this transcription unit. Despite the presence of this second copy, deletion of the 5' copy of α caused a reduction in steady state RNA levels. We decided to further examine the position-

dependence of α by inserting the DNA fragment spanning nt. 2526-2666 into several sites in CMVΔ1, which is missing the first 143 nt. of pregenome-encoding sequence. An α-containing 140 bp. fragment was inserted at three restriction sites in the pregenome: NsiI (nt. 2845), located 315 nt. downstream from the 5' end; BglII (nt. 391), located 883 nt. downstream; and KpnI (nt.1290), located 1782 nt. downstream (Fig. 5A). Northern blot analysis of poly $(A)^+$ RNA from transfected LMH cells demonstrated that this fragment could restore pgRNA levels when inserted at downstream sites (Fig. 5B). CMVΔ1/NsiI and CMVΔ1/BglII both consistently yielded wild type levels of pregenomic RNA (lanes 2 and 4). Results with CMVΔ1/KpnI were equivocal; insertion at KpnI at least partially restored pregenome RNA levels (lane 5), but the extent of restoration was highly variable, ranging from 10% - 70% of wild type pgRNA levels. CMVΔ1/BglIIinv has the 140 bp. fragment inserted in the inverse orientation. The greatly diminished level of its pgRNA (lane 3) demonstrates that α function is orientation-dependent.

DISCUSSION

We have identified through deletion analysis a region critical for pregenome RNA accumulation in transfected LMH cells. This region (α) is encoded by nt. 2581-2662 and is located between nt. +52 and +133 from the pregenome cap site. Constructs which lacked this element yield very low levels of pregenome RNA if sequences 1 kb. downstream, encoded by nt. 391-561, were present. Deletion of this second element (β) restored α-deleted pregenomes to wild type levels. Huang and Summers have also recently reported closely similar observations (9).

Our deletion mutants suggest that the α element overlaps but is unrelated to the RNA stem-loop structure ε present at the 5' end of pgRNA. The ε structure is conserved in all hepadnaviruses and plays critical roles in pregenome RNA encapsidation (13), P protein binding (14), and initiation of viral minus strand DNA synthesis (21,22). We observed that removal of the first 48 nt. of pgRNA (CMVΔ4), which deletes almost the entire 5'

UCSF LIBRARY

stem of the proposed ε structure, did not affect pgRNA levels; thus, if the α element requires RNA secondary structure, it must fold into an alternative structure different from that of ε. Unlike ε, the α and β elements also mediate their effects in the absence of all viral proteins. Moreover, the phenotype of the α deletion mutants was reproduced in HeLa and L-cells, two cell lines of non-hepatic origin. In sum, these results suggest that the mechanisms regulating pregenome accumulation involve cellular proteins and general regulatory processes.

What mechanisms regulate the accumulation of pgRNA levels? The dramatic reduction in pgRNA following deletion of the α element is observed whether transcription is driven by the endogenous pregenomic promoter or by the CMV IE promoter. This promoter nonspecificity argues against transcription initiation as a likely mode of action. Huang and Summers have proposed a model (9) in which pgRNA elongation is blocked in a region located just upstream of the subgenomic promoters (corresponding at least in part to β). They suggest that a positive effector of transcription (pet) located near the 5' end of pregenome (corresponding to our α) overcomes this block to transcriptional elongation through this (β) region, perhaps by forming a structure in the 5' region of the nascent transcript, much like the TAR element located at the 5' end of HIV-1 mRNA (reviewed in 5). In support of this mechanism of premature transcription termination suppression, they point to increased levels of subgenomic message in pet (-) deletions, which is suggested to be a consequence of relieving the occlusion of subgenomic promoters by the pregenomic promoter. Furthermore, their *in vitro* transcription studies showed an inhibition of nascent RNA elongation in pet (-) templates, though this finding was not consistently observed with all HeLa cell extracts.

Our results are not sufficient to establish a mechanistic model, but do bear on the proposal of Huang and Summers. The α and pet elements defined by the two groups map virtually identically. The downstream (β) element we delineated by deletion mutants to lie between nt. 391-561 is within the region of transcription termination they broadly define

between nt. 1-720. However, our northern blot analysis did not reveal consistent increases in subgenomic messages in the presence of α deletions, in contrast to their findings. Moreover, our finding that downstream reinsertions of α restored pgRNA levels in 5' α deletions has implications for their model of pet action. Our observation that α/pet can function with normal efficiency when located 315 nt. or 883 nt. downstream from its original site argues against a mechanism for pet that is strictly analogous to that of HIV-1 TAR, which displays strong position dependence relative to the cap site (17). In addition, the variable but sometimes extensive restoration of pregenome levels when α/pet was inserted 1782 nt. downstream is difficult to reconcile with their model; in this construct, the α/pet sequence is positioned downstream of the proposed region of transcription termination. While our results by no means refute the model proposed by Huang and Summers, they suggest that other models should also be considered.

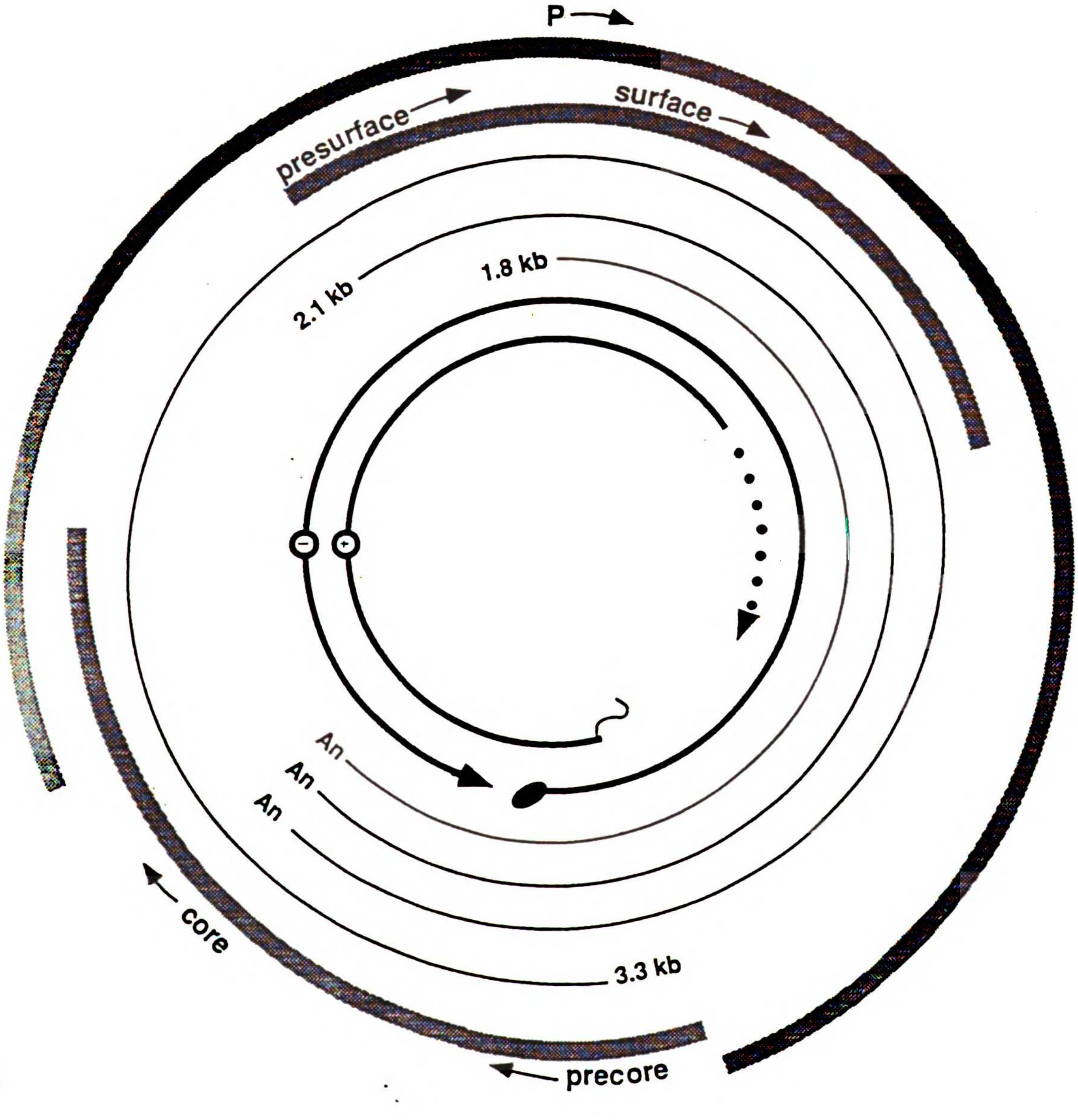
For example, reduced pregenome levels from α/pet deletion could be caused by a defect in RNA transport from the nucleus, by inefficient RNA processing, or by mRNA instability. Reversal of this defect by downstream (β) deletion could occur if the positive signal contributed by α sequences required an interaction with sequences within β to generate a structure necessary for effector function. Clearly, further studies are required to classify the molecular basis of these effects. Irrespective of their mechanism, these phenomena further underscore the complexity of hepadnaviral gene regulation.

ACKNOWLEDGMENTS

This work was supported by grants from the National Institutes of Health. We thank Sophie Roy for critical reading of the manuscript.

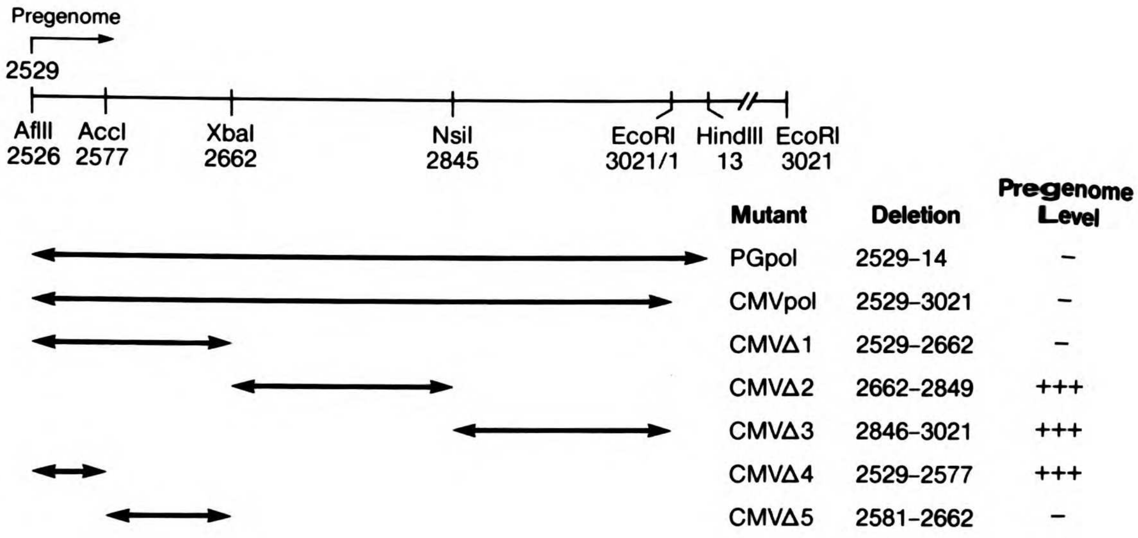
IIST LIBRARY

FIG. 1 Genetic and transcriptional map of DHBV. The open reading frames are shown in thick outer bars. The three thin lines represent the major transcripts in the infected cell: The 2.1 kb. and 1.8 kb. subgenomic transcripts, and the 3.3 kb. pregenomic RNA. An indicates the mRNA poly (A) tail. The inner circle represents the partially double stranded DNA genome. The 5' end of the minus strand is covalently linked to a terminal protein derived from the N-terminus of P protein. The 5' end of the variable length plus strand is fixed relative to minus strand and is linked to an oligoribonucleotide.

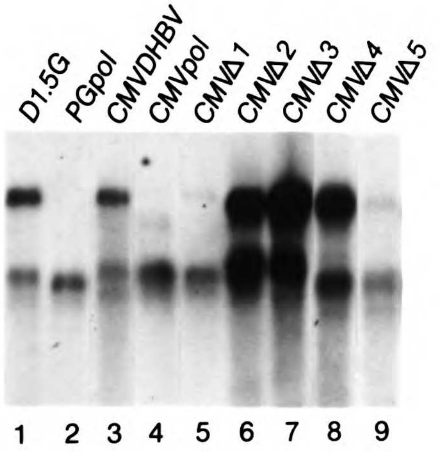

P
presurface
surface
2.1 kb
1.8 kb
An
An
An
core
3.3 kb
precore

[illegible] LIBRARY

FIG. 2 Northern blot analysis of mutants bearing deletions in the 5' region of pregenomic RNA. A) Schematic representation of 5' end deletion mutants and their effect on pregenome RNA levels. The extent of each deletion is denoted by the double-headed arrows and defined by nucleotide position. DHBV nucleotide positions are numbered from the unique EcoRI site of DHBV-3. Pregenome RNA levels detected for each mutant is defined as follows: -, no detectable RNA; +, approximately 20% or less of WT; ++, approximately 20-50% of WT; +++, approximately equivalent to WT. B) Northern blot of DHBV-specific poly(A)$^+$ RNA from transfected LMH cells. RNA was prepared 48 hours post-transfection as described in the materials and methods, and electrophoresed through a 1% agarose-2.2M formaldehyde gel. RNA was then transferred onto a nylon membrane and hybridized to [^{32}P] DHBV DNA. To normalize for differences in transfection efficiency for each mutant, expression of a β-galactosidase reporter gene was measured to calculate the approximate quantity of RNA to load in each lane. Note that RNA samples in lanes 8 and 9 were prepared from a separate transfection experiment than samples in lanes 1-7. D1.5G is an overlength DHBV genome plasmid that initiates pregenome transcription from the endogenous pregenome promoter. CMVDHBV is also an overlength DHBV genome plasmid that initiates pregenome transcription from the CMV immediate-early promoter. Transfection of either constructs into LMH cells leads to production of infectious virus.


Pregenome
2529
AflII
2526
AccI
2577
XbaI
2662
NsiI
2845
EcoRI
3021/1
HindIII
13
EcoRI
3021

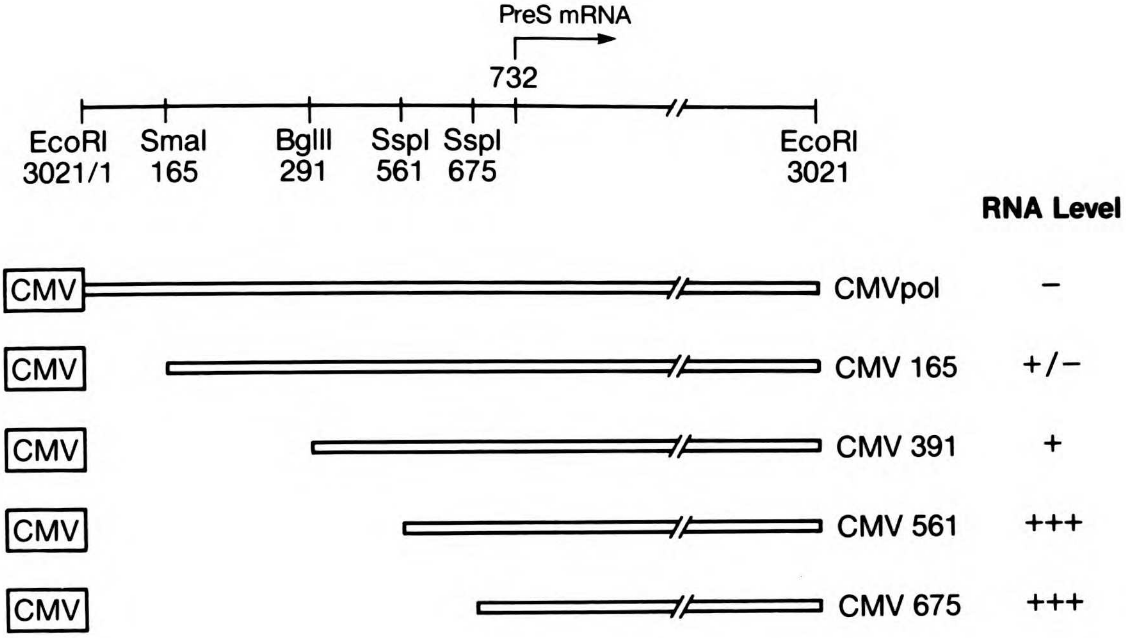
Mutant	**Deletion**	**Pregenome Level**
PGpol	2529–14	–
CMVpol	2529–3021	–
CMVΔ1	2529–2662	–
CMVΔ2	2662–2849	+++
CMVΔ3	2846–3021	+++
CMVΔ4	2529–2577	+++
CMVΔ5	2581–2662	–

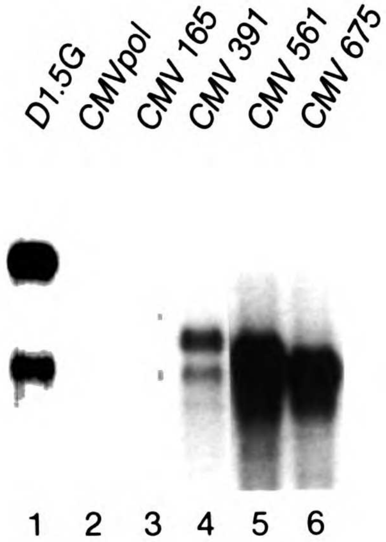

D1.5G
PGpol
CMVDHBV
CMVpol
CMVΔ1
CMVΔ2
CMVΔ3
CMVΔ4
CMVΔ5
1
2
3
4
5
6
7
8
9

IISc LIBRARY

FIG. 3 Poly (A)$^+$ RNA levels of transcripts that initiate upstream of the PreS mRNA cap site. A) Schematic diagram of CMV-driven constructs that initiate at sites between nt. 1-675 (between -731 to -56 from the PreS mRNA cap site). The CMV I.E. promoter is located immediately upstream of the DHBV fragment indicated by the open bar. The same subjective scale is used to denote RNA levels as in fig. 2A. B) Northern blot of DHBV-specific poly (A)$^+$ RNA from transfected LMH cells. Methods are identical to those described in fig. 2B.

USGS LIBRARY

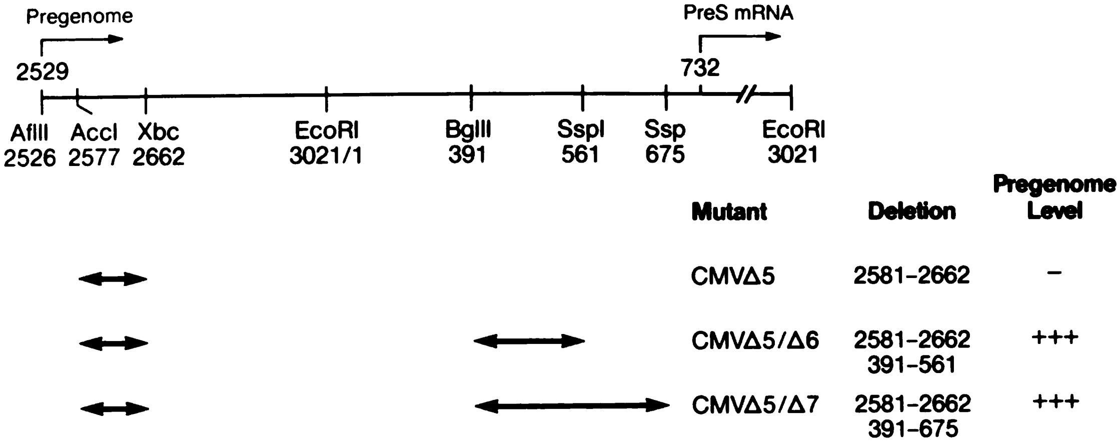

PreS mRNA
732
EcoRI 3021/1
SmaI 165
BglII 291
SspI 561
SspI 675
EcoRI 3021
RNA Level
CMV
CMVpol
–
CMV 165
+/–
CMV 391
+
CMV 561
+++
CMV 675
+++


D1.5G
CMVpol
CMV 165
CMV 391
CMV 561
CMV 675
1
2
3
4
5
6

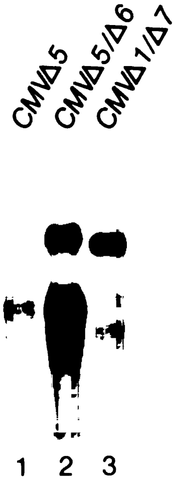
UCSF LIBRARY

FIG. 4 Northern blot analysis of mutants bearing a deletion in the α region and an additional downstream internal deletion. A) Schematic diagram of deletion mutants CMVΔ5, and of CMV5/Δ6 and CMV5/Δ7, which bear additional deletions further downstream. The deletions are denoted by the double-headed arrows and defined by nucleotide position. The same subjective scale is used to denote pregenome levels as in fig. 2A. B) Same as fig. 2B.

UCSF LIBRARY

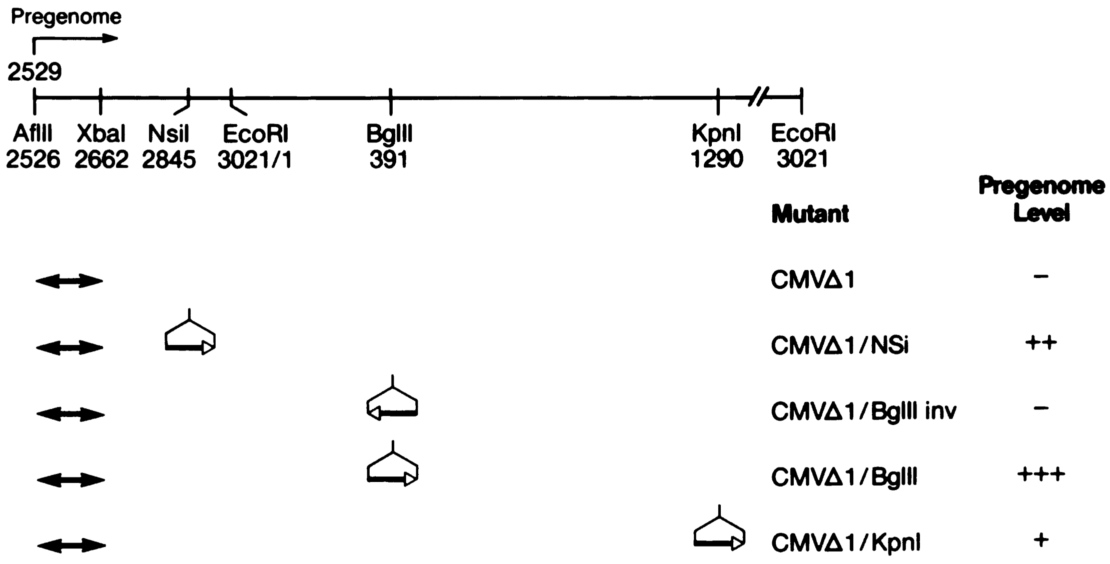

Pregenome
2529
PreS mRNA
732
AflII
2526
AccI
2577
Xbc
2662
EcoRI
3021/1
BglII
391
SspI
561
Ssp
675
EcoRI
3021

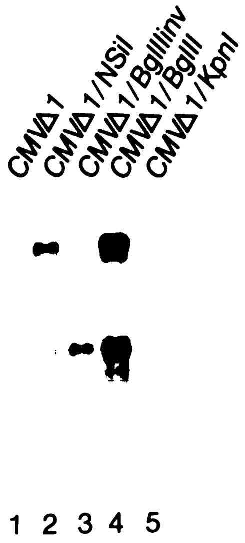
Mutant	**Deletion**	**Pregenome Level**
CMVΔ5	2581-2662	–
CMVΔ5/Δ6	2581-2662 391-561	+++
CMVΔ5/Δ7	2581-2662 391-675	+++


CMVΔ5
CMVΔ5/Δ6
CMVΔ1/Δ7
1
2
3

USGS LIBRARY

FIG. 5 Northern blot analysis examining position and orientation-dependence of α sequences. A) Schematic diagram of deletion mutant CMVΔ1, and four additional constructs in which a 140 nt. region containing α is inserted at three restriction sites in the pregenome: NsiI, BglII, and KpnI. Deleted regions are denoted by double-headed arrows. The location and orientation of the inserted α region is represented by the open arrow. Note that in CMVΔ1/BglIIinv, the region bearing α is inserted in the opposite orientation relative to its WT position. The same subjective scale is used to denote pregenome levels as in fig. 2A. B) Same as fig. 2B.


Pregenome
2529
AflII 2526
XbaI 2662
NsiI 2845
EcoRI 3021/1
BglII 391
KpnI 1290
EcoRI 3021
Mutant
Pregenome Level
CMVΔ1 –
CMVΔ1/NSi ++
CMVΔ1/BglII inv –
CMVΔ1/BglII +++
CMVΔ1/KpnI +


CMVΔ1
CMVΔ1/NSiI
CMVΔ1/BglIIinv
CMVΔ1/BglII
CMVΔ1/KpnI
1 2 3 4 5

MCZ LIBRARY

REFERENCES

1. Bartenschlager, R., and H.Schaller. 1988. The amino-terminal domain of the hepadnaviral P-gene encodes the terminal protein (genome-linked protein) believed to prime reverse transcription. EMBO J. 7: 4185-4192.

2. Bartenschlager, R., and H.Schaller. 1992. Hepadnaviral assembly is initiated by polymerase binding to the encapsidation signal in the viral RNA genome. EMBO J. 11: 3413-3420.

3. Buscher,M.W., W.Reiser, H.Will, and H.Schaller. 1985. Transcripts and the putative RNA pregenome of duck hepatitis B virus: implications for reverse transcription. Cell 40: 717-724.

4. Condreay, L.D., C.E.Aldrich, L.Coates, W.S.Mason, and T.-T.Wu. 1990. Efficient duck hepatitis B virus production by an avian liver tumor cell line. J.Virology 64: 3249-3258.

5. Cullen, B.R. 1992. Mechanisms of action of regulatory proteins encoded by complex retroviruses. Microbiological Reviews 56: 375-394.

6. Ganem, D., and H.E.Varmus. 1987. The molecular biology of the hepatitis B viruses. Ann. Rev. Biochem. 56: 651-694.

7. Hirsch,R., R.Colgrove, and D.Ganem. 1988. Replication of duck hepatitis B virus in two differentiated human hepatoma cell lines after transfection with cloned viral DNA. Virology 167: 136-142.

8. Hirsch,R., J.Lavine, L.Chang, H.E.Varmus, and D.Ganem. 1990. Polymerase gene products of hepatitis B viruses are required for genome RNA packaging as well as for reverse transcription. Nature 344: 552-555.

9. Huang,M., and J.Summers. 1994. Pet, a small sequence distal to the pregenome cap site is required for expression of the duck hepatitis B virus pregenome. Submitted.

10. Lien, J.M., C.E.Aldrich, and W.S.Mason. 1986. Evidence that a capped oligonucleotide is the primer for duck hepatitis B virus plus-strand DNA synthesis. J.Virology 57: 229-236.

11. 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. PNAS 77: 2941-2945.

12. Mason, W.S., G.Seal, and J.Summers. 1980. Virus of Pekin ducks with structural and biological relatedness to human hepatitis B virus. J.Virology 36: 829-836.

13. Pollack, J.R., and D.Ganem. 1993. An RNA stem-loop structure directs hepatitis B virus genomic RNA encapsidation. J.Virology 67: 3254-3263.

14. Pollack, J.R., and D.Ganem, 1994. Duck hepatitis B virus polymerase binds an RNA stem-loop structure to indicate both viral RNA encapsidation and DNA synthesis. Submitted.

15. Sambrook,J., E.F.Fritsch, and T.Maniatis. 1989. Molecular Cloning: A Laboratory Manual. 2nd ed. Cold Spring Harbor Press, Cold Spring Harbor, N.Y.

16. Seeger, C., D.Ganem, and H.E.Varmus. 1986. Biochemical and genetic evidence for the hepatitis B virus replication strategy. Science 232: 477-484.

17. Selby, M., E.Bain, P.Luciw, and B.M.Peterlin. 1989. Structure, sequence, and position of the stem-loop in tar determine transcriptional elongation by tat through the HIV-1 long terminal repeat. Genes and Development 3: 547-558.

18. Sprengel, R., C.Kuhn, H.Will, and H.Schaller. 1985. Comparative sequence analysis of duck and human hepatitis B virus genomes. J. Med. Virol. 15: 323-333.

19. Summers, J., J.M.Smolec, and R.Snyder. 1978. A virus similar to human hepatitis B virus associated with hepatitis and hepatoma in woodchucks. PNAS 75: 4533-4537.

20. 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.

21. Tavis, J.E., S.Perri, and D.Ganem. Hepadnaviral reverse transcription initiates within the stem-loop of the RNA packaging signal and employs a novel strand transfer. Submitted.

22. Wang, G.-H., and C.Seeger. 1993. Novel mechanism for reverse transcription in hepatitis B viruses. J.Virology 67: 6507-6512.

23. Will, H., W.Reiser, T.Weimer, E.Pfaff, M.Buscher, R.Sprengel, R.Cattaneo, and H.Schaller. 1987. Replication strategy of human hepatitis B virus. J.Virology 61: 904-911.

UCSF LIBRARY

118GR

IIEST LIBRARY

For reference

Not to be taken
from the room.

629744

3 1378 00629 7447