Critical Role of the 36-nucleotide Insertion in Hepatitis B Virus Genotype G on Core Protein Expression, Genome Replication, and Virion Secretion.

Ke Li, Fabien Zoulim, Christian Pichoud, Karen Kwei, Stéphanie Villet, Jack Wands, Jisu Li, Shuping Tong

Journal of Virology 81, 17 (2007) 9202-15

Critical Role of the 36-nucleotide Insertion in Hepatitis B Virus Genotype G on Core Protein Expression, Genome Replication, and Virion Secretion

Ke Li[1], Fabien Zoulim[2*], Christian Pichoud[2], Karen Kwei[1], Stéphanie Villet[2], Jack Wands[1], Jisu Li[1], and Shuping Tong[1*]

Liver Research Center, Rhode Island Hospital, Brown University, Providence, USA[1];

INSERM, U871; Université Lyon 1, IFR62 Lyon Est, Lyon, France[2]

*Corresponding authors. Shuping_Tong_MD@Brown.edu; zoulim@lyon.inserm.fr

ABSTRACT

Frequent co-infection of hepatitis B virus genotype G with genotype A suggests it might require genotype A for replication or transmission. In this regard, genotype G is unique in having a 12-amino acid extension in the core protein due to a 36-nucleotide insertion near the core gene translation initiation codon. The insertion alters base pairing in the lower stem of the pregenome encapsidation signal that harbors core gene initiator, and thus has the potential to affect both core protein translation and pregenomic RNA encapsidation. Genotype G is also unusual for possessing two nonsense mutations in the precore region, which together with the core gene encodes a secreted nonstructural protein called hepatitis B e antigen (HBeAg). We found that genotype G clones were indeed incapable of HBeAg expression, but were competent in RNA transcription, genome replication, and virion secretion. Interestingly, the 36-nucleotide insertion markedly increased core protein, which was achieved at the level of protein translation but did not involve alteration in the mRNA level. Consequently, the variant core protein was readily detectable in patient blood. The 12-amino acid insertion also enhanced genome maturity of secreted virus particles, possibly through less efficient envelopment of core particles. Co-transfection of genotypes G and A did not lead to mutual interference of genome replication or virion secretion. Considering that HBeAg is an immunotolerogen required for the establishment of persistent infection, its lack of expression rather than a replication defect could be the primary determinant for the rare occurrence of genotype G monoinfection.

INTRODUCTION

Hepatitis B virus (HBV) can be classified into 8 genotypes with distinct geographic distribution, target population, and mode of transmission (15, 19, 29). Genotype G was first recognized in French patients in 2000, although it had been described in earlier literature as viral variants (4, 38, 42). This genotype is unique in that it is frequently detected in homosexual men, who may suffer from immune suppression due to infection with HIV (4, 26, 44). At the molecular level, different isolates of genotype G display remarkable sequence conservation (>99%) (21). They harbor the A1762T/G1764A mutations in the core promoter region, which for other genotypes do not arise until at late stage of chronic infection (3, 8, 31, 33). Importantly, all the genotype G clones share a 36-nt insertion not found in any other HBV genotypes. This insertion at the 5' end of core gene adds 12 amino acid residues to the core protein immediately following the initiating methionine: DRTTLPYGLFGL. At the RNA level, the insertion is located close to a hairpin structure at the 5' end of pregenomic (pg) RNA called the encapsidation (ε) signal, which directs the pg RNA into nascent core particles for initiation of DNA replication (Fig. 1A). In fact, the first three nucleotides inserted alter base pairing at the lower stem of the ε signal (Fig. 1A), thus could potentially affect the efficiency of pg RNA encapsidation. In addition to serving as the genome precursor, the pg RNA prior to its encapsidation functions as mRNA for the translation of core protein, the building block for core particle, as well as DNA polymerase, the enzyme responsible for the conversion of pg RNA into double stranded DNA. In this regard, the core gene AUG initiator is located at the lower stem of the ε signal. Since RNA secondary structure can impede translation initiation, alteration of base pairing by the 36-nt insertion has the potential to alter the efficiency of core protein translation.

The core gene together with the preceding precore region codes for the precore/core protein, which is converted to hepatitis B e antigen (HBeAg) following cleavage of the signal peptide and arginine-rich sequence at the carboxyl terminus (32). Although other genotypes can evolve into HBeAg-negative variants at the later stage of chronic infection, genotype G always contain a defective precore region due to two nonsense mutations. Surprisingly, most genotype G patients are still HBeAg positive. This puzzling observation prompted careful molecular epidemiological studies which revealed frequent co-infection of genotype G with genotype A, the likely source of HBeAg detected in such patient sera (20). Genotype G gradually replaces genotype A as the patients seroconvert to anti-HBe (20, 21, 42). A more recent study revealed co-infection of genotype G with genotype H in Mexico (36), which seems to reinforce the requirement for a co-infecting genotype, whether in the form of genotype A or genotype H.

In the present study, we performed transfection experiments of genotype G in a human hepatoma cell line in order to characterize its genome replication, protein expression, and virion secretion properties. We found genotype G clones were rather competent in genome replication and virion secretion. They failed to express bona fide HBeAg but produced much higher level of the variant core protein, which could be detected in patient sera. The 36-nt insertion was indispensable for the replication efficiency of genotype G, and responsible for high core protein level as well as increased genome maturity of secreted virions. When co-transfected to cells, genotypes A and G did not significantly interfere with each other in terms of genome replication or virion secretion.

MATERIALS AND METHODS

HBV DNA constructs. Clone D2, containing 1.1 copies (3.5 kb) of genotype D genome under the CMV promoter, was a kind gift of Christoph Seeger, Fox Chase Cancer Institute, Philadelphia, PA. To study HBV replication driven by the strong actin promoter, HBV sequences were amplified from patient sera as two overlapping DNA fragments and cloned into the NotI - XhoI sites of pTriEx vector (Novagen) as 1.1 copies (1.1mers), as detailed elsewhere (13). Clones G1-G3 and G4-G6 of genotype G were derived from two separate patients. The complete nucleotide sequences of clones G1 and G6 were determined by customer sequencing at the Keck Laboratory, Yale University (GenBank accession numbers EF634481 and EF634480). Clone G1 in the pTriEx vector was used to generate several additional constructs. Replacement of its 2.6-kb NotI – NcoI fragment (nucleotide sequence 1806-1375) with the corresponding sequence from Kca1.6, a genotype C clone (Gewaily, unpublished) produced G1/C1 chimera. This construct expressed genotype C-specific core protein and envelope proteins. Removal of the 36 nt unique to genotype G produced G1d36. The core-minus mutants of G1 and G1d36 were generated by converting the initiation codon to ATA, while the packaging-negative mutants harbored the G1879T/T1880A double mutation in the ε signal (Fig. 1A) known to abolish genome replication (2). All the mutations (substitution, insertions, and deletions) were created by overlap-extension PCR followed by restriction fragment exchange as previously described (1, 2, 17, 22, 33). To generate an SphI dimer of G1 or G1d36, the full-length HBV genome was re-amplified from the pTriEx plasmids using primers 3 and 4 located in the precore region as previously described (33), except that a single nucleotide change was introduced into primer 4 to accommodate the TAA nonsense mutation at the second codon of the precore region. The DNA

was digested with SapI and ligated at high insert: vector ratio with a modified pUC18 vector containing compatible SapI ends to generate dimeric HBV constructs. The plasmids were further digested with SphI, and the unit-length HBV genome was religated with SphI-cut pUC18 vector to generate SphI dimers. Clones 2A, 4B, 5.4 and 6.2 of genotype A have been used previously as EcoRI dimers (33). In the present study clone 2A was converted into an SphI dimer for better comparison with SphI dimers of genotype G. Insertion of the 36 nucleotides into clone 4B generated 4Bins36, which was subsequently reconverted to an EcoRI dimer. A env$^-$, a genotype A mutant defective in virion secretion was created by a G261A nonsense mutation in the envelope gene of clone N16 described previously (22). For core protein expression independent of the ε signal, a 0.6-kb DNA fragment was amplified by sense primer 5'AGCACCTCGAGAAGCTGTGCCTTGGGTGGCTTTG-3 (XhoI site underlined) and antisense primer 5'-AAAGCGAATTCAAGTTTCCCACCTTATGAGTCC-3' (EcoRI site underlined), and cloned into the XhoI – EcoRI sites of pcDNA3.1 zeo(-) vector (Invitrogen). Such constructs direct core protein expression from the CMV promoter. For HBeAg expression under the CMV promoter, an upstream sense primer was used to include the entire precore region (5'-CCGAACTCGAGGCATAAATTGGTCTGCGCACCAGCACCATGCAACTTTTTCACCTCTGCC-3').

Transfection. Huh7 cells were maintained in MEM medium supplemented with 10% fetal bovine serum. Transient transfection was performed on cells seeded in 6-well plates or 10-cm dishes, with cells reaching 40-80% confluency at the time of transfection. A typical transfection for one well of 6-well plates consisted of 2 μg of HBV DNA, 5 ng of cDNA

encoding secreted alkaline phosphatase (SEAP) to serve as control for transfection efficiency, 6 µl of LT1 reagent (Mirus), and 200 µl of serum-free medium. Medium was changed the second day after transfection, and cells were harvested 4 days later (day 5 post-transfection). For Northern blot analysis, 1.2 µg of HBV construct was co-transfected with 0.8 µg of SEAP cDNA, and cells were harvested at day 3 rather than day 5 post-transfection.

Analysis of HBV DNA replication. HBV DNA was extracted from intracellular core particles for Southern blot analysis (1, 2, 17, 33). Briefly, cells were scrapped off each well of the 6-well plates and lyzed in 80 µl of 10 mM HEPES (pH 7.5), 100 mM NaCl, 1 mM EDTA, and 1% NP40 buffer. A 2/5th of the cell lysates was treated at 37^{o}C for 15 min with 7.5 units (u) of mung bean nuclease and 0.5 u of DNase I in a total volume of 100 µl, in the presence of 10 mM of $CaCl_2$ and 12 mM of $MgCl_2$. Next, core particles were precipitated by PEG solution. After another treatment with nucleases, HBV DNA was released by proteinase K digestion, extracted with phenol, and precipitated with ethanol. DNA was run in 1.2% agarose gels in the presence of ethidium bromide, transferred to Genescreen plus membranes (PerkinElmer), and hybridized with ^{32}P labeled full-length HBV DNA amplified by primers 3 and 4 (33). For unbiased detection of both genotype A and genotype G, genotype A and genotype G DNA was mixed at 1:1 ratio for labeling. After hybridization the blots were washed with 2x SSC / 0.1% SDS buffer at 60^{o}C for 2 hrs. For sequential hybridization of the same blots with different probes, the old probe was removed by boiling the blots for 30 min in 0.1x SSC / 1% SDS solution.

Analysis of virions and naked core particles released to culture supernatant. Viral and subviral particles secreted to culture supernatant were precipitated with a horse polyclonal anti-HBs (Ad/Ay) antibody (Abcam), which had been preconjugated to protein G-agarose beads (Roche) by rotation overnight at 4°C with a ratio of 1.5 µl antibody per 10 µl bed volume of beads. Next, 10 µl bed volume of loaded beads was incubated with 1.5 ml of percleared culture supernatant at 4°C for two overnights with rotation. The beads were brought down by low speed centrifugation, washed once with 1 ml of PBS, and spun down again. The immobilized particles were digested at 37°C for 15 min with 1 u of DNase I and 1.5 u of mung bean nuclease in 50 µl of 10 mM Tris (pH7.5), 6 mM $MgCl_2$, 8 mM $CaCl_2$ solution, followed by digestion with 1mg/ml of proteinase K. Subsequent to phenol extraction, DNA was precipitated with ethanol using glycogen (20 µg) as a carrier. Purified DNA was dissolved in TE buffer for Southern blot analysis. To detect naked core particles, 1.5 ml of percleared culture supernatant was incubated overnight at 4°C with 1.5 µl of polyclonal rabbit anti-core antibody (Dako), followed by addition of 10 µl bed volume of protein G beads and a further incubation of 3-5 hrs. The subsequent steps were identical to that for viral particles.

Differentiation between genotypes A and G during mixed transfection experiments. DNA extracted from core particles or virions was treated at 37°C for 2 hrs in high salt restriction enzyme buffer (Roche) with 1 u of Klenow fragment in the presence of 100 µM of dNTP to repair the single-stranded region. The enzyme was inactivated at 75°C for 10 min. Next, 1/3rd of the DNA extracted from intracellular core particles or extracellular virions was digested at 37°C for 4 hrs with 5 u of EcoRI, while another 1/3rd was digested

with 5 u each of BglI and XhoI. Digested and undigested DNA was heated at 85^{0}C for 10min, chilled on ice, and separated in agarose gels for Southern blot analysis.

Northern blot analysis. Cells were lyzed using Trizol (Invitrogen), and total RNA was extracted according the manufacturer's protocol. The RNA pellet was dissolved in nuclease-free water, and 8 μg was dissolved in loading buffer containing 2.1M formaldehyde, heated at 95^{o}C for 3 min, and separated in 1.2% agarose gel containing 0.8 M formaldehyde. The blot was hybridized with mixed genotype A/G probe in the same manner as for Southern blot. After stripping, the same blot was hybridized with a 1.5-kb SEAP cDNA probe, which had been PCR amplified using sense primer 5'- TGGGCCTGAGGCTACAGCTC -3' and antisense primer 5'- TATCTTATCATGTCTGCTCGAAGC -3'.

Primer extension experiments. Huh7 cells grown in 6-well plates were harvested at day 3 posttransfection with 1ml of Trizol solution (Invitrogen) and RNA was extracted. A commercial kit (AMV reverse transcription kit, Promega) was used for primer extension assay. The RNA (10 μg) was incubated at 58 ^{0}C for 20 min with 300 fmol of an anti-sense primer labeled with γ-^{32}P ATP using T4 polynucleotide kinase. The primers used were 5'-GACTCTAAGGCTTCTCGATACAGAG-3' (positions 2031-2007) for genotype A samples, and 5'- GACTCTAAGGATTCCCGGTACAAAG -3' (positions 2067-2043) for genotype G samples. The annealed oligonuclotide was extended by AMV reverse transcriptase at 42 ^{0}C for 30 min. The reaction was stopped by addition of the loading buffer and the product was heated at 90^{0}C for 10 min before separation in a 5% acrylamide gel containing 7 M urea, using 1XTBE buffer. As a molecular size marker, HaeIII-digested ϕx-174 DNA was labeled

with γ-^{32}P ATP using T4 polynucleotide kinase and run in parallel. The gel was dried and primer extension products revealed by autoradiography.

Detection of core protein and HBeAg from cell lysate or culture supernatant. HBeAg present in culture supernatant were quantified by ETI-EBK plus enzyme immunoassay (DiaSorin). Western blot analysis was performed as described previously (17). Proteins from about 30 µl of cell lysate were separated in 0.1% SDS - 12% PAGE and transferred to PVDF membranes. The blots were blocked at room temperature for 1 hr with 3% BSA in TBS/0.05% Tween 20 (TBST) buffer, and incubated overnight at 4°C with a polyclonal rabbit anti-core antibody (Dako) diluted 1: 2000 in 3% BSA/TBST. After washing at room temperature for 40 min, the blots were incubated at room temperature for 1 hr with a 1: 20,000 – 40,000 dilution of anti-rabbit antibodies conjugated with horseradish peroxidase (HRP; Amersham). The blots were washed again for 40 min and signals were revealed by enhanced chemiluminescence (ECL). The core protein expressed from clone 4B of genotype A, which was not reactive to the rabbit antibody from Dako, was detected by a monoclonal antibody (14E11, Abcam) at 1:2000 dilution followed by anti-mouse antibodies conjugated with HRP (1: 20,000 dilution). In some experiments, the blots were treated with stripping buffer (Pierce) at 37°C for 25 min with agitation, washed five times with TBST for a total of 50 min, and blocked again in 3 % BSA/TBST. The membranes were incubated with mouse GAPDH antibody (MAB374, Chemicon) at 1:5000 dilution, followed by a 1:10,000 dilution of anti-mouse antibody conjugated to HRP.

For Western blot analysis of secreted HBeAg and core protein, 1ml of culture supernatant was rotated at 4°C overnight with 1 µl of rabbit polyclonal anti-core antibody (Dako), followed by addition of 10 µl bed volume of protein G agarose beads and a further incubation of 5 hrs. The beads were brought down by low speed centrifugation, washed once with PBS, and associated proteins were separated in SDS - 15% PAGE. The subsequent Western blot analysis was identical to that for the core protein. For detection of core protein and HBeAg from patient sera, the step of immunoprecipitation was performed at room temperature (3hrs of incubation with the antibody followed by 2 hrs of incubation with protein G beads) to avoid precipitation of serum proteins.

Detection of core protein by pulse-chase experiments. Huh7 cells in 6-well plates were rinsed with Hanks solution 2 days after transfection, and starved for 2 hrs in methionine-free/cysteine-free MEM Eagle medium lacking calf serum. Next, cells were incubated for 3 hrs with 500 µl/well of the same medium supplemented with 0.12 mCi/ml of Express Protein Labeling Mix (PerkinElmer), followed by washing with complete DMEM supplemented with extra methionine (10 times higher than in the original medium). Cells were either harvested immediately by scrapping or continued to incubate in methionine-fortified complete medium before harvesting. The cell pellet was lyzed in 500 µl of lysis buffer (10 mM HEPES (pH 7.5), 100 mM NaCl, 1 mM EDTA, and 1% NP40 buffer) supplemented with protease inhibitor cocktail (Roche), and incubated with 10 µl bed volume of protein G beads at 4° for 4 hrs followed by brief spin to remove proteins that directly bound to beads. The precleared lysate was incubated at 4° overnight with 0.5 µl of rabbit polyclonal anti-core antibody (Dako), and incubated for 3 more hrs following addition of 10 µl bed volume of protein G beads. The

beads were collected by low speed centrifugation, and washed once with lysis buffer. Proteins bound to beads were separated in SDS-12% polyacrylamide gel, which was treated sequentially with 10% acetic acid/25% methanol solution and Enlightening solution (New England Nuclear). The gel was dried and radioactive signals were revealed by exposure to X-ray films. Quantitative analysis was based on scanning of lightly exposed X-ray film followed by analysis with NIH Image software.

Detection of core particles. Intracellular core particles were detected by native agarose gels (24) with minor modifications. Huh7 cell lysate (10 µl) was diluted with 1/6th volume of the 6x DNA loading buffer (0.25% bromophenol blue, 40% sucrose) and applied to 1% agarose gel. The gel was run in TAE buffer at 90 V and proteins were blotted overnight with 10x SSC solution to a PVDF membrane, which had been soaked successively in methanol, water and 2x SSC before use. Subsequent detection of core protein on the blot was identical to Western blot analysis.

RESULTS

Replication and expression constructs. Two types of constructs were employed to study the replication of genotype G: 1.1mers of the viral genome cloned in the pTriEx vector downstream of the actin promoter and dimers of the viral genome inserted into a promoter-less cloning vector (Fig. 1B). Robust transcription of pg RNA by the actin promoter makes it easy to monitor core protein expression, genome replication, and virion secretion. The biological

relevance of the 36-nt insertion was evaluated by its removal from clone G1. Furthermore, we created core protein-deficient and encapsidation defective mutants of G1 and its deletion mutant (G1d36) to separate the *cis* effect (36-nt insertion in the genome) from the trans effect (12-aa insertion in core protein). Core protein expression was prevented by mutating the core AUG into AUA, while pg RNA encapsidation was abolished by G1879T/T1880A double mutation at the loop of the ε signal (Fig. 1A, *core⁻ mutant* and *ε⁻ minus mutant*).

It should be pointed out that the degree of replication observed from 1.1mer genomes may not necessarily reflect *in vivo* situation, where viral replication is driven by the endogenous core promoter and subject to modulation by the two enhancers, HBx protein, and naturally occurring mutations in the core promoter (6, 33). Moreover, such constructs do not transcribe the precore RNA required for HBeAg expression, because their 5' ends do not include the entire precore region (Fig. 1B). Therefore, tandem dimers of clone G1 and its deletion mutant were cloned into pUC18 vector to study viral replication in a more natural setting. Core gene constructs under the CMV promoter (CMV-core, Fig. 1B) permitted us to examine the impact of the 36 nt on core protein expression without complication by the secondary structure of the ε signal (Fig. 1A). Considering that genotype G has a nonfunctional precore region, the potential impact of the 36-nt insertion on HBeAg expression was studied in genotype A clones either as tandem dimers or as CMV-precore expression constructs (Fig. 1B).

Experimental conditions for unbiased detection of the two genotypes. Thanks to a >12% sequence divergence between genotypes A and G at the nucleotide level (21), a potential complication in the comparative studies of the two HBV genotypes is preferential hybridization

of the probe to DNA or RNA of the same genotype. We prepared a blot containing 30 pg - 1 ng of HBV DNA of genotypes A, D, and G for sequential hybridization with probes of genotype A, G, and A / G at 1:1 ratio. The genotype A probe coupled with low-salt washing buffer (0.5x SSC/0.1% SDS) was quite biased against other genotypes while the mixed probe coupled with a high-salt washing buffer (2x SSC/0.1% SDS) could detect both genotypes A and G at similar efficiency (Fig. 2). These conditions were adopted for all the Southern and Northern blot analyses.

The rabbit polyclonal antibody against core protein (Dako) was capable of detecting core protein derived from different genotypes except for one genotype A clone (4B) due to a E77Q misssense mutation (Kim et al., unpublished). A mouse monoclonal antibody (Abcam) was used instead for clone 4B and its derivatives.

Genotype G clones were competent in the transcription of pregenomic and subgenomic RNAs. Huh7 cells were transfected with 1.1mers of HBV genomes cloned in the pTriEx vector, or tandem dimers (EcoRI or SphI) cloned in the pUC18 vector, together with SEAP cDNA. According to Northern blot analysis. 2 of the 5 genotype G clones in the pTriEx vector (G4, G6, lanes 6, 7, Fig. 3A) produced comparable amount of the pg RNA as genotype A and D clones (lanes 1, 2). The slightly lower HBV transcript levels associated with three clones from one patient (G1, G1d36, G2, lanes 3-5) correlated with reduced SEAP mRNA level (Fig. 3B). Converting clone G1 into tandem SphI dimer markedly reduced transcription of pg RNA (lanes 10 vs. 3). Indeed, dimers of G1 and its d36 mutant produced less precore (pc)/pg RNA and subgenomic (sg) RNA than clone 2A, a low replicating genotype A clone (33) (lanes 10, 11, 9).

Clone 4B, a high replicating genotype A clone with core promoter mutations (33), generated highest level of pc/pg RNA among the dimer constructs (lane 8). Therefore, under endogenous promoters genotype G may be less efficient in RNA transcription than genotype A.

The 36-nt insertion in genotype G was critical for efficient core protein expression. Due to the strong actin promoter the 1.1mer genomes in the pTriEx vector were generally efficient in core protein expression except a chimeric construct with core gene derived from genotype C (G1/C1) (Fig. 4A, upper panel, lane 10). The different mobility of core protein generated from various genotypes is compatible with their sizes: 183-aa for genotypes C and D, 185-aa for genotype A, and 195-aa for genotype G. The levels of core particles assembled generally correlated with core protein levels (Fig. 4A, compare upper and lower panels), suggesting that core particle assembly was not adversely affected by the 12-aa insertion unique to genotype G. In this regard, other insertions at the N-terminus of core protein were also tolerated in terms of particle formation (18, 34). Removal of the genotype G specific 36 nt from clone G1 markedly reduced core protein expression (Fig. 4A, upper panel, compare lanes 4 & 9). When core protein was expressed from tandem SphI dimers under the endogenous promoter, G1 expressed much higher level of core protein than a genotype A clone (2A) (Fig. 4B, left panel). Removal of the 36 nt from clone G1 reduced core protein to undetectable level. In this regard, the first 3 nt of the 36-nt insertion alters base pairing at the lower stem of the ε signal, which may potentially modulate efficiency of translation initiation or RNA packaging (Fig. 1A). However, encapsidation deficient mutants of G1 or G1d36 produced similar levels of core protein as their parental 1.1mer genomes (Fig. 5C, lanes 3 and 5; lanes 4 and 6). Furthermore, we generated core protein expression constructs driven by the CMV promoter, which produced core protein in the

absence of preceding ε signal (Fig. 1A & B). Again clone G1 produced much higher level of core protein than its deletion mutant (Fig. 4C, left panel).

In the complementary approach, we inserted the 36 nt into clone 4B of genotype A, either in the form of EcoRI dimer or CMV-core construct. In both cases, the insertion markedly enhanced core protein expression (Fig. 4B & 4C, right panels). The fact that a different antibody was used to detect core protein produced from clone 4B suggests that the 36-nt in genotype G increased core protein expression or stability rather than its recognition by the polyclonal antibody from Dako.

The 36-nt insertion increased core protein translation but did not enhance mRNA abundance. Pulse-chase experiments were conducted to determine whether the 36-nt insertion increased de novo core protein translation or the 12-aa insertion slowed core protein degradation. Only genotype G constructs were examined because the murine monoclonal antibody failed to immunoprecipitatate core protein from clone 4B or 4Bins36. During a short labeling period of 3 hrs about 7 times more core protein was synthesized from the 1.1mer genome of G1 than G1d36 (Fig. 6A, lanes 3, 4). Similarly, CMV-core construct of G1 produced much higher level of core protein than its deletion mutant (lanes 1, 2). Identical results were obtained following 1 hr of pulse with ^{35}S methionine (data not shown). Culturing labeled cells for 68 additional hrs resulted in reduction of labeled core protein for both G1 and G1d36 to about 50% of the original level (lanes 6, 7), suggesting similar protein half-life. Thus, up regulation of core protein by the 36-nt insertion is mediated by increased protein synthesis.

One potential mechanism whereby the 36-nt insertion increases core protein translation is to up regulate corresponding mRNA, either at the transcriptional or posttranscriptional level. However, primer extension assay excluded this possibility. For example, the 4Bins36 dimer did not produce more pg RNA than the parental clone, 4B (Fig. 6B, lanes 3, 2), nor did the corresponding CMV-core construct (lanes 5, 4). The CMV-core construct of G1 produced higher transcript level than G1d36 (lanes 11, 12), which may partly explain the much higher core protein expression. However, this was not the case for 1.1mer genome (lanes 9, 10). Consistent with the primer extension data, Northern blot analysis also failed to reveal higher pg/pc RNA in G1 than G1d36 (Fig. 3A, lanes 3, 4, 10, 11). Therefore, the 36 nucleotides unique to genotype G apparently make the corresponding mRNA more efficient for core protein translation.

The 36-nt insertion was essential for efficient replication of genotype G. The genotype G clones as 1.1mer genomes in the pTriEx vector were replication competent (Fig. 7A, upper panel, lanes 4-8). Removing the 36 nt from clone G1 not only markedly reduced core protein expression as described, but also suppressed genome replication to similar extent (lane 9 vs. lane 4). We employed trans-complementation assay in an attempt to sort out whether reduced encapsidation / replication efficiency of the G1d36 pg RNA, or the low amount of core protein produced, restricted genome replication. The core$^-$ mutants of G1 and G1d36 could replicate their respective genomes so long as core protein was provided in trans. Their ε^- mutants could not package the pgRNA but were nevertheless competent in the expression of core protein in addition to DNA polymerase and envelope proteins. The CMV-core constructs differ from the ε^- mutants in expressing the core protein alone. A representative result was shown in Fig. 5A. We found that replication of G1core$^-$ mutant (1μg) could be rescued by 1 μg of the ε^- mutant of

either G1 or G1d36 to the level achievable by 1 μg of wild-type G1 (Fig. 5A, compare lanes 7, 8 with lane 5). Rescue was less efficient by the CMV-core constructs, especially that of G1d36 (lanes 9, 10). Replication of the G1d36core$^-$ mutant was rescued more efficiently by the ε$^-$ mutant of G1 than G1d36 (lanes 11, 12), although even the latter combination reached higher replication capacity than the parental clone, G1d36 (lane 6). Therefore, the trans-complementation experiments implicate a combination of cis and trans defects in the lower replication capacity of G1d36. Certainly, the trans-complementation assay may have its limitations because the pg RNA of the core$^-$ and ε$^-$ mutants served as pregenome and mRNA respectively, rather than both pregenome and mRNA for the parental construct. With the CMV-core construct, translation of core protein was no longer coordinated with expression of DNA polymerase, another component required for the pg RNA encapsidation.

The SphI dimer of clone G1 replicated to similar extent as clone 2A of genotype A, whereas G1d36 was nearly defective in genome replication (Fig. 7B, upper panel). This defect could be rescued by co-transfection of CMV-core construct of clone G1, although the rescue was incomplete (Fig. 7C, top panel, compare lane 6 with lane 1).

The 12-aa insertion in the core protein of genotype G enhanced genome maturity of secreted virus particles. Interestingly, all the genotype G clones secreted virus particles of much higher genome maturity than clones of other genotypes, as evidenced by the paucity of single stranded (SS) DNA and abundance of relaxed circular (RC) genome (Fig. 7A, lower panel, lanes 4-8). The less efficient virion secretion of clone G3 was correlated with much reduced HBsAg secretion (unpublished). Replacement of the core gene of clone G1 with

genotype C (G1/C1) reduced genome maturity of secreted particles (Fig. 7A, lower panel, lane 11 vs. lane 4). Removal of the 36 nt from G1 produced similar effect, but the efficiency of virion secretion was enhanced as evidenced by the ratio of extracellular HBV DNA / intracellular viral DNA (lane 9). Whether the genome maturity of secreted virus particles was controlled by the 36-nt insertion in the genotype G DNA or the 12-aa extension in the core protein was established by trans-complementation assay, as already described in the preceding section. We found that genome maturity was determined primarily by the core protein expression construct. Thus, the G1ε$^-$ mutant generated virions of mature genome whether co-transfected with G1core$^-$ or G1d36core$^-$ mutant (Fig. 5B, lanes 7, 11), whereas G1d36ε$^-$ ensured more efficient virion secretion with less mature genome (Fig. 5B, lanes 8, 12). Similar results were obtained with core protein expressed from CMV-core constructs (lanes 9, 13, 10, 14).

The 36-nt/12-aa insertion incrased secretion of naked core particles. By simultaneous detection of both virions and naked core particles in the culture supernatant, it became apparent that genotype G secreted more naked core particles than genotype A (Fig. 5D, lanes 1, 2, 5, 6). Moreover, naked core particles of genotype G contained less mature genome than the corresponding virions (compare lanes 6 and 2). Deletion of the 36 nucleotides not only increased virion secretion (lane 3) but also markedly suppressed core particle release (lane 7). The virus particles of G1d36 had similar genome maturity as naked core particles of G1 (compare lanes 3 and 6). Therefore, the 36-nt insertion enhanced genome maturity of secreted virions through inhibition of core particle envelopment.

Independent genome replication and virion secretion of genotypes A and G. Considering the frequent coinfection of genotype G with genotype A, we performed co-transfection experiments to test for possible mutual enhancement or interference of replication / virion secretion. The 1.1mer genomes in the pTriEx vector (A1 and G4) were employed for such studies because virion secretion from dimers was too low. Differentiation between the progeny DNA of the two genotypes was made possible by the single EcoRI site at position 1 of genotype A (but not genotype G) and single sites of BglI (position 1925 inside the 36-nt insertion) and XhoI (position 2907) on genotype G (but not genotype A) (Fig. 8A). To minimize the effect of the single-stranded region on the banding of digestion products, we repaired the single-stranded gap of DNA extracted from intracellular core particles. Furthermore, digested or undigested DNA was heated at 85^0C for 10 min prior to gel electrophoresis to convert relaxed circular DNA into a linear form. As shown in Fig. 8B & C, digestion with EcoRI converted A1 DNA into fragments of approximately 1.8 kb and 1.4 kb (lane 3), but did not affect G4 DNA (lane 15 in panel A and lane 12 in panel B). Double digestion with BglI / XhoI converted G4 DNA into fragments of 2.2 kb and 1 kb (lane 14 in panel A and lane 11 in panel B), but did not affect the migration of A1 DNA (lane 2). With ratios of input DNA at 1:3, 1:1, and 3:1, both genotypes could replicate and secrete virus particles, with their relative abundance proportional to the input (Fig. 8B & C). Assuming that most cells were transfected with both genotypes, the result suggests no competition or enhancement between the two genotypes, at least for replication driven by a strong exogenous promoter.

The 12-aa insertion into the precore protein could reduce HBeAg production. Genotype G harbors the C1817T and G1896A nonsense mutations in the precore region capable

of terminating HBeAg translation. However, it has been speculated that the 12-aa insertion at the amino terminus of core protein may serve as a signal peptide to generate a variant HBeAg from the core gene alone. First, the impact of the 12-aa insertion on HBeAg expression was studied in genotype A, which has a functional precore region. The 36-nt insertion caused about 5-fold drop in HBeAg expression from both an EcoRI dimer (clone 4B) and a CMV-precore construct (clone 6.2) (Fig. 9C, left and middle panels). Since the insertion did not down regulate the corresponding mRNA (Fig. 6B, lanes 6, 7), it is likely that the 12-aa insertion impaired HBeAg processing or secretion. IP - Western blot analysis revealed reduced rather than accelerated mobility of the variant HBeAg (Fig. 9C, right panel), suggesting that the original N-terminal signal peptide of 19 residues, rather than the 12 inserted residues, were recognized and removed by the signal peptidase.

Core protein overexpression led to its release to culture supernatant. We next investigated whether the 12-aa insertion as part of core protein could be cleaved by a signal peptide to generate variant HBeAg. The 1.1mer genomes, which could express core protein but not precore protein, were utilized. Interestingly, low level of "HBeAg" could be detected, but this was not unique to genotype G (Fig. 9A). Failure of G1core$^-$ mutant to secrete such "HBeAg" (Fig. 9A, last lane) confirmed it as a core gene product. The amount of "HBeAg" released by the 1.1mer genomes was less than 5% of HBeAg produced by the SphI dimer of a genoytpe A clone (2A). Immunoprecipitation (IP) - Western blot analysis revealed intact core protein (21-kd for most genotypes and 23-kd for genotype G) in such supernatant (Fig. 9B, lanes 1-9). If variant HBeAg is produced from genotype G core protein by cleavage around the 12-aa insertion followed by C-terminal cleavage, it should have been smaller than the classic HBeAg secreted

by the SphI dimer of clone 2A (lane 11). Since non-particulate core protein is known to possess HBe antigenicity (37), the "HBeAg" produced by the 1.1mer HBV genomes may reflect release of free core protein, or disassembly of released core / virus particles.

Natural infection with genotype G was associated with core protein release. Inability of genotype G to secrete bona fide HBeAg (Fig. 9B), coupled with recent findings of genotype G co-infection with genotype A or F (20, 36), suggests genotype A or F as the source of HBeAg detected in clinical samples. We collected 5 samples infected with genotype G, 3 of which with high HBeAg titers (Fig. 9D, lanes 1, 3, 4). Both direct sequencing and InnoLiPA test (Innogenetics, N.V.) identified sample 1 as of genotype G. Sample 3 was found to be genotype G by direct sequencing but a mixture of A/G genotypes according to InnoLiPA assay. The InnoLiPA assay also identified sample 4 as of mixed A/G infection. IP - Western blot analysis revealed strong band of HBeAg in the two HBeAg+ genotype A samples which was comparable to that from culture supernatant of transfected Huh7 cells (lanes 8, 9, 10). In contrast, only samples 1 and 3 of genotype G produced weak or barely visible band at the position of classic HBeAg. Interestingly, both samples displayed an additional band corresponding to genotype G specific core protein (lanes 1 & 3), suggesting core protein release as a consequence of increased production.

DISCUSSION

The frequent co-infection of genotype G with another HBV genotype suggests a possible defect of genotype G in replication or transmission. In this regard, the salient features of genotype G sequence include the 36-nt insertion at the 5' end of core gene, double nonsense mutation in the precore region, and double core promoter mutation. The precore and core promoter mutations do develop in other HBV genotypes, albeit at a late stage of infection in response to anti-HBe immunity (3, 7, 8, 28, 31, 40). The precore mutations that abrogate the expression of HBeAg do not necessarily impair HBV replication, and the core promoter mutations actually enhance viral replication (5, 6, 33, 39, 41, 43, 45). Thus, genotype G could be defective in genome replication due to the 36 extra nucleotides not found in other genotypes. Our experiments demonstrated that genotype G clones were replication competent whether driven by the strong actin promoter (1.1mer genomes in pTriEx vector) or by the endogenous promoter (SphI dimers in pUC18 vector), as evidenced by RNA transcription, envelope protein expression (our unpublished observation), core protein expression and particle assembly, genome replication and virion secretion. These findings confirm an earlier report of the replication capacity of an SpeI dimer of genotype G (25), and are fully compatible with a recent report of genotype G mono-infection in a blood donor and several transfusion recipients (12). These authors were unable to detect genotype A sequence in the blood samples using a variety of sensitive methods, thus demonstrating convincingly that genotype G can be transmitted and propagated in patients without the need for a helper virus.

The transfection experiments revealed several interesting biological features of genotype G, including more efficient core protein expression, secretion of virus particles with high genome maturity, and surprisingly, dependence of its replication on the 36-nt insertion.

Furthermore, the same 36-nt insertion underlies both efficient core protein expression and secretion of virions with more mature genome. The exact mechanism whereby the 36-nt insertion increases core protein levels remains to be worked out, but it has to do with more efficient protein synthesis rather than extension of protein half life (Fig. 6A). Moreover, the increased translation occurred in the absence of up regulation of the corresponding mRNA (Fig. 6B). This finding is reminiscent of our recent report revealing marked down regulation of core protein expression by several point mutations in the precore region, such as G1862T, T1863A (in the bulge of the ε signal), apparently without reducing pg RNA abundance (17). Based on experiments with the CMC-core constructs, the 36-nt insertion can impact core protein translation even in the absence of the ε signal. It will be interesting to examine whether the 36-nt insertion increases the distribution of the corresponding mRNA towards polysomes, the ribosome fraction actively involved in translation. One possible scenario is that core protein binds to its own mRNA to inhibit protein translation (for coordinated core particle assembly and pg RNA packaging), which is weakened by the 36-nt insertion on the mRNA or the 12-aa insertion in core protein. In this regard, dihydrofolate reductase protein has been found to down regulate its own translation by binding to the cognate mRNA (11, 14). Further experiments are needed to test our hypothesis.

Since the 36-nt insertion is needed for both efficient core protein expression and replication of genotype G (both as 1.1mer genomes and SphI dimers), it is natural to ask whether the two phenotypes are connected. In the present study, we performed detailed trans-complementation assay only on the 1.1mer genomes, which are certainly different from natural infection in overproduction of the pg RNA and consequently, core protein. The results were not

clear cut, but it appears that both the loss of the 36-nt in the pg RNA and loss of the 12-aa in the core protein contributed to reduced replication of G1d36. Therefore, co-transfection of G1d36core$^-$ mutant with its ε$^-$ mutant gave rise to lower replication than co-transfection of G1core$^-$ mutant with the ε$^-$ mutant of either G1 or G1d36. Another observation was the much less efficient rescue of the core$^-$ mutants by CMV-core constructs, which differ from the ε$^-$ mutants by not expressing polymerase and envelope proteins. The same G1 CMV-core construct only partially rescued replication defect of the SphI dimer of G1d36. However, this experiment differed from that of the 1.1mer genomes in that core protein with 12-aa deletion was still expressed. Considering that the SphI dimers of G1 and G1d36 produced less pg /pc RNA than a genotype A clone (2A) (Fig. 3), we favor the hypothesis that the 36-nt insertion is required to compensate for the low abundance of pg RNA to sustain genotype G replication. The fact that the SphI dimer of clone G1 replicated to similar degree as clone 2A suggests that up regulation of core protein translation alone is sufficient to augment replication. Alternatively, since both the core protein and DNA polymerase are translated from pg RNA, the 36-nt insertion may also up regulate the translation of polymerase. This possibility should be tested experimentally.

While it remains open whether the 36-nt insertion in the pg RNA or the 12-aa insertion in core protein up regulates core protein translation, we have compelling evidence to suggest that the 12-aa insertion in the core protein leads to secretion of virus particles with much maturer genome and higher RC DNA content than other genotypes. This point was demonstrated convincingly by the transcomplementation experiments. Whether for G1core$^-$ mutant or G1d36 core$^-$ mutant, cotransfection with expression constructs of G1 core protein (the ε$^-$ mutant or CMC-core construct) led to secretion of mature genome, whereas cotransfection with G1d36

core constructs ensured secretion of less mature genome, but at higher efficiency. Simultaneous analysis of naked core particles released to culture supernatant from both G1 and G1d36 mutant suggests that the 12-aa insertion reduced the efficiency of core particle envelopment, leading to 1) secretion of more naked core particles; 2) moderate increase in genome maturity of intracellular core particles; and 3) markedly increased genome maturity of particles which were eventually secreted as virions. It will be of interest to repeat the experiments in HepG2 cells, which in our hand produced more mature genomes than Huh7 cells.

Two artificial HBV mutants with 10- and 23-aa insertion at the N terminus of core protein failed to secrete virus particles (18), while a single naturally occurring P5T mutation near the N-terminus of core protein reduced virion secretion (27). A series of N-terminal insertion mutants of duck hepatitis B virus secreted virus particles of reduced rather than increased genome maturity, which was associated with preferential degradation of particles harboring mature genome (23). Therefore, N-terminal insertion in the core protein modulates virion secretion, although the exact outcome is sequence specific.

The 12-aa insertion has been proposed to drive HBeAg production in the blood of patients infected with genotype G. According to this hypothesis, a fraction of the core protein enters the secretory pathway through a novel signal peptide created by the 12-aa insertion, followed by removal of both the signal peptide and arginine-rich sequence at the carboxyl terminus. However, we failed to observe protein bands with sizes comparable to such variant HBeAg even with core protein over-expression driven by the strong actin promoter. Insertion of the same 36 nt into genotype A clones with a functional precore region suppressed HBeAg

secretion but did not alter the cleavage site. This result agrees with the lack of HBeAg or anti-HBe antibody in patients infected with genotype G alone (12). Nevertheless, high titer of HBeAg could be detected in the sera of two samples infected with genotype G, one with no indication of co-infection with another genotype (at least based on direct sequencing and Inno-LiPA assay). IP-Western blot analysis of two serum samples infected with genotype G did not reveal sufficient amount of HBeAg to account for the extremely high HBeAg titers (Fig. 9D, lanes 1 & 3). Although both samples contained core protein, our experience with 1.1mer genomes revealed extremely low HBe antigenicity of secreted core protein. An alternative possibility is that HBeAg from these two samples, derived from co-infecting genotype A isolates, harbored mutations that render it poorly recognizable by the polyclonal antibody (Dako) used in this study. We found that core protein and HBeAg from clone 4B of genotype A were not recognized by the Dako antibody (Fig. 7C) due to an E77Q mutation; this mutation is very common among genotype A clones isolated from the late HBeAg+ phase of infection.

What is responsible for the rare monoinfection by genotype G? One explanation is its relatively inefficient replication. The SphI dimer of G1 replicated to a comparable level of clone 2A, a low replicating genotype A clone, despite the presence of double core promoter mutation (A1762T/G1764A). In some transfection experiments, the 1.1mer genomes of genotype G also replicated less efficiently than the corresponding genotype A or D clones. Secondly, genotype G and genotype A (or genotype F) may complement each other for better replication or transmission. For example, high level of core protein produced by genotype G may increase the replication of another genotype. Further co-transfection experiments using tandem dimers rather than the 1.1mer genomes as shown in Fig. 8 may provide additional information. It will be also

interesting to study the fitness of genotype G alone or in a mixed infection with genotype A in the differentiaed HepRG cells (16). However, we reason that the major impediment for monoinfection by genotype G may be its inability to express HBeAg, a molecule believed to play an immunomodulatory function critical for the establishment of persistent infection (10, 30). It has been reported that children with maternal transmission of pure wild-type virus became chronically infected, while those infected with a mixture of wild-type virus and HBeAg-minus mutant resolved acute infection (35). Similarly, the precore-defective mutant of woodchuck hepatitis virus induced only transient infection in the animals (9). Consistent with the immunomodulatory role of HBeAg, genotype G frequently infects homosexual men, who due to immune suppression are less likely to clear viral infection (4, 44). In such hosts immune escape mechanisms are probably not essential for the establishment of persistent infection as opposed to immunocompetent individuals. But even in such a group, the HBeAg-producing genotype is the dominant viral species during the early stage of infection; selection of genotype G coincides with the emergence of anti-HBe (20, 21, 42).

In conclusion, genotype G is replication competent. Efficient replication of genotype G requires the 36-nt insertion in the core gene, which may modulate the efficiency of core protein expression and virion secretion, as well as genome maturity of virus particles. Our findings provide a molecular explanation for the presence of the 36-nt insertion in genotype G. Considering that this insertion has never been found in other genotypes, whether its artificial introduction into other genotypes will modulate genome replication warrants further investigation.

ACKWLEDGMENTS

This work was supported by grants CA109733, CA95490, CA35711, DK066950, and DK28614 from the National Institutes of Health, RSG 06-059-01-MBC from the American Cancer Society, and by a grant of the Agence Nationale de Recherche contre le SIDA (ANRS) to F.Z. K.L was supported in part by the Tan Yan Kee Foundation.

REFERENCES

1. **Ahn, S. H., A. Kramvis, S. Kawai, H. C. Spangenberg, J. Li, G. Kimbi, M. Kew, J. Wands, and S. Tong.** 2003. Sequence variation upstream of precore translation initiation codon reduces hepatitis B virus e antigen production. Gastroenterology **125:**1370-8.
2. **Bang, G., K. H. Kim, M. Guarnieri, F. Zoulim, S. Kawai, J. Li, J. Wands, and S. Tong.** 2005. Effect of mutating the two cysteines required for HBe antigenicity on hepatitis B virus DNA replication and virion secretion. Virology **332:**216-24.
3. **Baptista, M., A. Kramvis, and M. C. Kew.** 1999. High prevalence of 1762(T) 1764(A) mutations in the basic core promoter of hepatitis B virus isolated from black Africans with hepatocellular carcinoma compared with asymptomatic carriers. Hepatology **29:**946-53.
4. **Bhat, R. A., P. P. Ulrich, and G. N. Vyas.** 1990. Molecular characterization of a new variant of hepatitis B virus in a persistently infected homosexual man. Hepatology **11:**271-6.
5. **Blum, H. E., T. J. Liang, E. Galun, and J. R. Wands.** 1991. Persistence of hepatitis B viral DNA after serological recovery from hepatitis B virus infection. Hepatology **14:**56-63.
6. **Buckwold, V. E., Z. Xu, M. Chen, T. S. Yen, and J. H. Ou.** 1996. Effects of a naturally occurring mutation in the hepatitis B virus basal core promoter on precore gene expression and viral replication. J Virol **70:**5845-51.
7. **Carman, W. F., M. R. Jacyna, S. Hadziyannis, P. Karayiannis, M. J. McGarvey, A. Makris, and H. C. Thomas.** 1989. Mutation preventing

formation of hepatitis B e antigen in patients with chronic hepatitis B infection. Lancet **2:**588-91.

8. **Chan, H. L., M. Hussain, and A. S. Lok.** 1999. Different hepatitis B virus genotypes are associated with different mutations in the core promoter and precore regions during hepatitis B e antigen seroconversion. Hepatology **29:**976-84.
9. **Chen, H. S., M. C. Kew, W. E. Hornbuckle, B. C. Tennant, P. J. Cote, J. L. Gerin, R. H. Purcell, and R. H. Miller.** 1992. The precore gene of the woodchuck hepatitis virus genome is not essential for viral replication in the natural host. J Virol **66:**5682-4.
10. **Chen, M. T., J. N. Billaud, M. Sallberg, L. G. Guidotti, F. V. Chisari, J. Jones, J. Hughes, and D. R. Milich.** 2004. A function of the hepatitis B virus precore protein is to regulate the immune response to the core antigen. Proc Natl Acad Sci U S A **101:**14913-8.
11. **Chu, E., C. H. Takimoto, D. Voeller, J. L. Grem, and C. J. Allegra.** 1993. Specific binding of human dihydrofolate reductase protein to dihydrofolate reductase messenger RNA in vitro. Biochemistry **32:**4756-60.
12. **Chudy, M., M. Schmidt, V. Czudai, H. Scheiblauer, S. Nick, M. Mosebach, M. K. Hourfar, E. Seifried, W. K. Roth, E. Grunelt, and C. M. Nubling.** 2006. Hepatitis B virus genotype G monoinfection and its transmission by blood components. Hepatology **44:**99-107.
13. **Durantel, D., S. Carrouee-Durantel, B. Werle-Lapostolle, M. N. Brunelle, C. Pichoud, C. Trepo, and F. Zoulim.** 2004. A new strategy for studying in vitro the drug susceptibility of clinical isolates of human hepatitis B virus. Hepatology **40:**855-64.
14. **Ercikan-Abali, E. A., D. Banerjee, M. C. Waltham, N. Skacel, K. W. Scotto, and J. R. Bertino.** 1997. Dihydrofolate reductase protein inhibits its own translation by binding to dihydrofolate reductase mRNA sequences within the coding region. Biochemistry **36:**12317-22.
15. **Fung, S. K., and A. S. Lok.** 2004. Hepatitis B virus genotypes: do they play a role in the outcome of HBV infection? Hepatology **40:**790-2.
16. **Gripon, P., S. Rumin, S. Urban, J. Le Seyec, D. Glaise, I. Cannie, C. Guyomard, J. Lucas, C. Trepo, and C. Guguen-Guillouzo.** 2002. Infection of a human hepatoma cell line by hepatitis B virus. Proc Natl Acad Sci U S A **99:**15655-60.
17. **Guarnieri, M., K. H. Kim, G. Bang, J. Li, Y. Zhou, X. Tang, J. Wands, and S. Tong.** 2006. Point mutations upstream of hepatitis B virus core gene affect DNA replication at the step of core protein expression. J Virol **80:**587-95.
18. **Hui, E. K., Y. S. Yi, and S. J. Lo.** 1999. Hepatitis B viral core proteins with an N-terminal extension can assemble into core-like particles but cannot be enveloped. J Gen Virol **80 (Pt 10):**2647-59.
19. **Kao, J. H.** 2002. Hepatitis B viral genotypes: clinical relevance and molecular characteristics. J Gastroenterol Hepatol **17:**643-50.
20. **Kato, H., E. Orito, R. G. Gish, N. Bzowej, M. Newsom, F. Sugauchi, S. Suzuki, R. Ueda, Y. Miyakawa, and M. Mizokami.** 2002. Hepatitis B e antigen

in sera from individuals infected with hepatitis B virus of genotype G. Hepatology **35:**922-9.

21. **Kato, H., E. Orito, R. G. Gish, F. Sugauchi, S. Suzuki, R. Ueda, Y. Miyakawa, and M. Mizokami.** 2002. Characteristics of hepatitis B virus isolates of genotype G and their phylogenetic differences from the other six genotypes (A through F). J Virol **76:**6131-7.
22. **Khan, N., M. Guarnieri, S. H. Ahn, J. Li, Y. Zhou, G. Bang, K. H. Kim, J. R. Wands, and S. Tong.** 2004. Modulation of hepatitis B virus secretion by naturally occurring mutations in the S gene. J Virol **78:**3262-70.
23. **Kock, J., S. Wieland, H. E. Blum, and F. von Weizsacker.** 1998. Duck hepatitis B virus nucleocapsids formed by N-terminally extended or C-terminally truncated core proteins disintegrate during viral DNA maturation. J Virol **72:**9116-20.
24. **Koschel, M., D. Oed, T. Gerelsaikhan, R. Thomssen, and V. Bruss.** 2000. Hepatitis B virus core gene mutations which block nucleocapsid envelopment. J Virol **74:**1-7.
25. **Kremsdorf, D., F. Garreau, F. Capel, M. A. Petit, and C. Brechot.** 1996. In vivo selection of a hepatitis B virus mutant with abnormal viral protein expression. J Gen Virol **77 (Pt 5):**929-39.
26. **Lacombe, K., V. Massari, P. M. Girard, L. Serfaty, J. Gozlan, G. Pialoux, P. Mialhes, J. M. Molina, C. Lascoux-Combe, D. Wendum, F. Carrat, and F. Zoulim.** 2006. Major role of hepatitis B genotypes in liver fibrosis during coinfection with HIV. Aids **20:**419-27.
27. **Le Pogam, S., T. T. Yuan, G. K. Sahu, S. Chatterjee, and C. Shih.** 2000. Low-level secretion of human hepatitis B virus virions caused by two independent, naturally occurring mutations (P5T and L60V) in the capsid protein. J Virol **74:**9099-105.
28. **Li, J., S. Tong, L. Vitvitski, F. Zoulim, and C. Trepo.** 1990. Rapid detection and further characterization of infection with hepatitis B virus variants containing a stop codon in the distal pre-C region. J Gen Virol **71 (Pt 9):**1993-8.
29. **Magnius, L. O., and H. Norder.** 1995. Subtypes, genotypes and molecular epidemiology of the hepatitis B virus as reflected by sequence variability of the S-gene. Intervirology **38:**24-34.
30. **Milich, D., and T. J. Liang.** 2003. Exploring the biological basis of hepatitis B e antigen in hepatitis B virus infection. Hepatology **38:**1075-86.
31. **Okamoto, H., F. Tsuda, Y. Akahane, Y. Sugai, M. Yoshiba, K. Moriyama, T. Tanaka, Y. Miyakawa, and M. Mayumi.** 1994. Hepatitis B virus with mutations in the core promoter for an e antigen-negative phenotype in carriers with antibody to e antigen. J Virol **68:**8102-10.
32. **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. Proc Natl Acad Sci U S A **83:**1578-82.
33. **Parekh, S., F. Zoulim, S. H. Ahn, A. Tsai, J. Li, S. Kawai, N. Khan, C. Trepo, J. Wands, and S. Tong.** 2003. Genome replication, virion secretion, and e antigen expression of naturally occurring hepatitis B virus core promoter mutants. J Virol **77:**6601-12.

34. **Pumpens, P., and E. Grens.** 1999. Hepatitis B core particles as a universal display model: a structure-function basis for development. FEBS Lett **442:**1-6.
35. **Raimondo, G., E. Tanzi, S. Brancatelli, S. Campo, M. A. Sardo, G. Rodino, M. Pernice, and A. R. Zanetti.** 1993. Is the course of perinatal hepatitis B virus infection influenced by genetic heterogeneity of the virus? J Med Virol **40:**87-90.
36. **Sanchez, L. V., Y. Tanaka, M. Maldonado, M. Mizokami, and A. Panduro.** 2007. Difference of hepatitis B virus genotype distribution in two groups of mexican patients with different risk factors. High prevalence of genotype h and g. Intervirology **50:**9-15.
37. **Seifer, M., and D. N. Standring.** 1995. Assembly and antigenicity of hepatitis B virus core particles. Intervirology **38:**47-62.
38. **Stuyver, L., S. De Gendt, C. Van Geyt, F. Zoulim, M. Fried, R. F. Schinazi, and R. Rossau.** 2000. A new genotype of hepatitis B virus: complete genome and phylogenetic relatedness. J Gen Virol **81:**67-74.
39. **Tong, S. P., C. Diot, P. Gripon, J. Li, L. Vitvitski, C. Trepo, and C. Guguen-Guillouzo.** 1991. In vitro replication competence of a cloned hepatitis B virus variant with a nonsense mutation in the distal pre-C region. Virology **181:**733-7.
40. **Tong, S. P., J. S. Li, L. Vitvitski, and C. Trepo.** 1990. Active hepatitis B virus replication in the presence of anti-HBe is associated with viral variants containing an inactive pre-C region. Virology **176:**596-603.
41. **Tong, S. P., J. S. Li, L. Vitvitski, and C. Trepo.** 1992. Replication capacities of natural and artificial precore stop codon mutants of hepatitis B virus: relevance of pregenome encapsidation signal. Virology **191:**237-45.
42. **Tran, A., D. Kremsdorf, F. Capel, C. Housset, C. Dauguet, M. A. Petit, and C. Brechot.** 1991. Emergence of and takeover by hepatitis B virus (HBV) with rearrangements in the pre-S/S and pre-C/C genes during chronic HBV infection. J Virol **65:**3566-74.
43. **Ulrich, P. P., R. A. Bhat, I. Kelly, M. R. Brunetto, F. Bonino, and G. N. Vyas.** 1990. A precore-defective mutant of hepatitis B virus associated with e antigen-negative chronic liver disease. J Med Virol **32:**109-18.
44. **Vieth, S., C. Manegold, C. Drosten, T. Nippraschk, and S. Gunther.** 2002. Sequence and phylogenetic analysis of hepatitis B virus genotype G isolated in Germany. Virus Genes **24:**153-6.
45. **Yuan, T. T., A. Faruqi, J. W. Shih, and C. Shih.** 1995. The mechanism of natural occurrence of two closely linked HBV precore predominant mutations. Virology **211:**144-56.

FIGURE LEGENDS

Figure 1. (A) Predicted secondary structure of the ε signal for genotypes A and G. Core gene translational initiation codon and the 5' end of the HBV sequence in the CMV-core constructs (see Fig. 1B) are indicated. Also shown for genotype G are the first three nucleotides of the 36-nt insertion, the double nucleotide change at the loop that prevents pg RNA encapsidation (ε^- *mutant*), and a G to A change of the initiation codon that abolishes core protein translation ($core^-$ *mutant*). **(B)** HBV genes, mRNAs, and constructs used for the present study. pc: precore; pg: pregenomic; sg: subgenomic; pol: polymerase; L, M, S: large, middle, and small envelope proteins. There are three in-frame translation initiation sites in the preS/S gene, leading to the expression of L protein from 2.4-kb sgRNA and M, S proteins from the 2.1-kb sgRNA. The two initiation sites in the precore/core gene lead to the expression of HBeAg from pcRNA and core protein / polymerase from the slightly shorter pgRNA. For the same reason, CMV-core constructs produce core protein, whereas CMV-precore constructs generate HBeAg. The 1.1mer constructs cannot express HBeAg because the 3.5-kb mRNA produced mimics pgRNA, not pcRNA.

Figure 2. Identification of hybridization conditions for unbiased detection of both genotypes A and G. pTriExD1 and pTriExG1 were linearized by XhoI digestion, while an EcoRI dimer of genotype A (clone 4B) in the pUC18 vector was linearized with HindIII. Increasing amounts of digested DNA (30pg, 100pg, 300pg, and 1000pg) were separated in agarose gel, and the blot was hybridized successively with probes of genotype A, genotype G, and A:G at 1:1 ratio. The blots were washed at 60^{o}C for a total of 2 hrs using the high salt (2x SSC/0.1% SDS) or low salt (0.5x SSC/0.1% SDS) buffer.

Figure 3. HBV RNA transcription in transfected Huh7 cells. The 1.1mers were cloned in the pTriEx vector with transcription of the 3.5-kb pg RNA under chicken actin promoter. The dimers were cloned into pUC18 vector, with transcription of all viral mRNAs under endogenous promoters. Huh7 cells were cotransfected with HBV constructs and SEAP cDNA, and 8µg RNA extracted at day 3 posttransfection was separated in denaturing agarose gel. (A) Hybridization of the blot with a mixed probe of genotypes A and G. (B) Hybridization of the same blot with a SEAP probe following stripping of the HBV probe. pg: pregenomic; pc: precore; sg: subgenomic.

Figure 4. Core protein expression from 1.1mer genomes (A), tandem dimers (B), and CMV-core expression constructs (C). Cells were harvested at day 5 posttransfection and core protein was detected from lysate by a rabbit polyclonal antibody (Dako) except for clone 4B and its derivatives (right B and C panels). For these samples, a mouse monoclonal antibody was used instead. The predicted core protein size is 183-aa for genotypes C & D, 185-aa for genotype A, and 195-aa for genotype G. For panel C, two independent clones of G1d36 and 4Bins36 were analyzed. The lower panel A shows intracellular core particles that have been separated in native agarose gel, and probed with the rabbit polyclonal antibody.

Figure 5. (A-C) Relative contribution of the 36-nt insertion in the genome vs. 12-aa insertion in the core protein on HBV DNA replication and virion secretion. The 1.1mer genomes were used. The core$^-$ mutants of G1 and G1d36 (1 µg) were co-transfected with same amount of core expressing constructs: ε$^-$ mutants or CMV-core (lanes 7-14). For controls, 1 µg of the ε$^-$ mutants, CMV-core constructs, or parental clones were transfected with 1 µg of pcDNA3.1 vector DNA

(lanes 1-6). About half of the cell lysate was used for detection of DNA replication (panel A), while another aliquot was used for successive detection of core protein and GAPDH (panel C). The supernatant was employed for detection of secreted virus particles (panel B). The EcoRI / RsrII digest of an EcoRI dimer served as size markers of 3.2, 1.7, and 1.5kbs. RC: relaxed circular; PDS: partially double stranded; SS: single stranded. (D) Impact of the 36-nt/12-aa insertion on release of HBV as virions or naked core particles. Pooled supernatants from transfected Huh7 cells were divided into 2 equal parts, one for immunoprecipitation with anti-S antibody and another for precipitation with anti-core antibody. The env^- mutant of genotype A (A env^-) did not express envelope proteins and therefore failed to secrete virions.

Figure 6. The 36-nt insertion in the core gene increased core protein translation without up regulating its steady state mRNA level. (A) Pulse-chase experiment. Huh7 cells were transfected with the indicated constructs, and labeled with ^{35}S methionine for 3 hrs. One set of the samples were harvested immediately, while another set was cultured for 68 additional hrs in the absence of labeled methionine. Core protein was immunoprecipitated from cell lysate and revealed by SDS-PAGE and fluorography. Positions of prestained molecular size markers are also indicated (they underestimate the size of endogenous proteins by at least 10%). The high molecular weight bands unique to pTriex constructs have the size of core protein dimers. (B) Primer extension assay. RNA (10 μg) extracted from day 3 posttransfection was annealed with an end-labeled anti-sense primer of genotype A (lanes 2-8) or genotype G (lanes 9-13), and the negative stranded cDNA was synthesized using AMV reverse transcriptase. The product was heated and separated in a 5% denaturing acrylamide gel. End labeled, HaeIII-

digested ϕx-174 DNA served as molecular size markers. The predicted sizes of the primer extension products are indicated.

Figure 7. HBV DNA replication and virion secretion from the 1.1mer constructs (A) and dimer constructs (B & C). Cells and culture supernatant were harvested at day 5 posttransfection. Core particles extracted from cell lysate were used for the detection of DNA replication, while virus particles were immunoprecipitated from culture supernatant with anti-S antibodies prior to Southern blot analysis. For lower panel B, virus particles secreted from clone 2A contained primarily single stranded genome. For panel C, SphI dimers were co-transfected with CMV-core constructs at 1:1 ratio. In addition to DNA replication and virion secretion (upper and middle panels), core protein expression was also monitored (lower panel). The rabbit polyclonal antibody used (Dako) failed to detect core protein expressed from clone 4B.

Figure 8. Lack of significant interference between genotypes A and G with regard to genome replication and virion secretion. (A) Cartoon view of the EcoRI site on genotype A, and BglI / XhoI sites on genotype G. There is no EcoRI site on genotype G, nor BglI/XhoI sites on genotype A. Shown are linear and relaxed circular (RC) forms with the single stranded region repaired. EcoRI digestion will convert linear form of genotype A into two bands of 1.8 kb and 1.4 kb, whereas double digestion with BglI and XhoI will produce 2.2-kb and 1-kb bands for the linear form of genotype G. The heating step prior to electrophoresis will melt the base pairing between the 5' ends of positive and negative strands in the RC DNA, thus generating the same migration pattern as the linear DNA. Number 1 and 2 inside box refer to DR1 and DR2 sequences. (B) HBV DNA associated with intracellular core particles. (C) HBV DNA of

extracellular virions. Huh7 cells grown in 10-cm dishes were transfected with a total of 10 µg of 1.1mer genomes using the ratios indicated. Both core particle-associated and virion-associated DNA harvested at day 5 posttransfection was treated with Klenow fragment in the presence of dNTP. A 1/3rd aliquot of the DNA was digested with EcoRI, while another aliquot with treated with BglI and XhoI. Digested and undigested DNA was heated at 85^{o}C for 10 min before Southern blot analysis with mixed A/G probe. For intracellular DNA (panel B), the smaller fragments of EcoRI and BglI / XhoI digestion (1.3 kb and 1.0 kb) were less distinct.

Figure 9. Core protein and HBeAg in the culture supernatant of transfected Huh7 cells (panels A-C) and patient blood (panel D). **(A)** "HBeAg" from cells transfected with 1.1mer viral genomes. "HBeAg" was measured from 15 µl of culture supernatant harvested at day 5 post-transfection. Values (with A1 arbitrarily set at 1) were averaged from three transfection experiments. **(B)** IP - Western blot analysis of core protein and HBeAg secreted to culture supernatant. Clones K81 and 5.4 were defective in HBeAg expression due to frameshift mutations in the precore region and the core gene, respectively. **(C)** Impact of the 36-nt insertion on HBeAg expression from two genotype A clones as EcoRI dimer (4B) or CMV-precore construct (6.2). The left and middle panels show HBeAg values of the insertion mutants relative to the parental constructs, which were arbitrarily set at 1. Results were based on three transfection experiments. The right panel shows IP – Western blot analysis. **(D)** IP - Western blot analysis of core protein and HBeAg in sera of patients infected with genotype A, G, or both. Culture supernatant of transfected Huh7 cells (sup) served as a positive control. The HBeAg values shown at the bottom were measured from 1.5 µl of serum sample or culture supernatant.

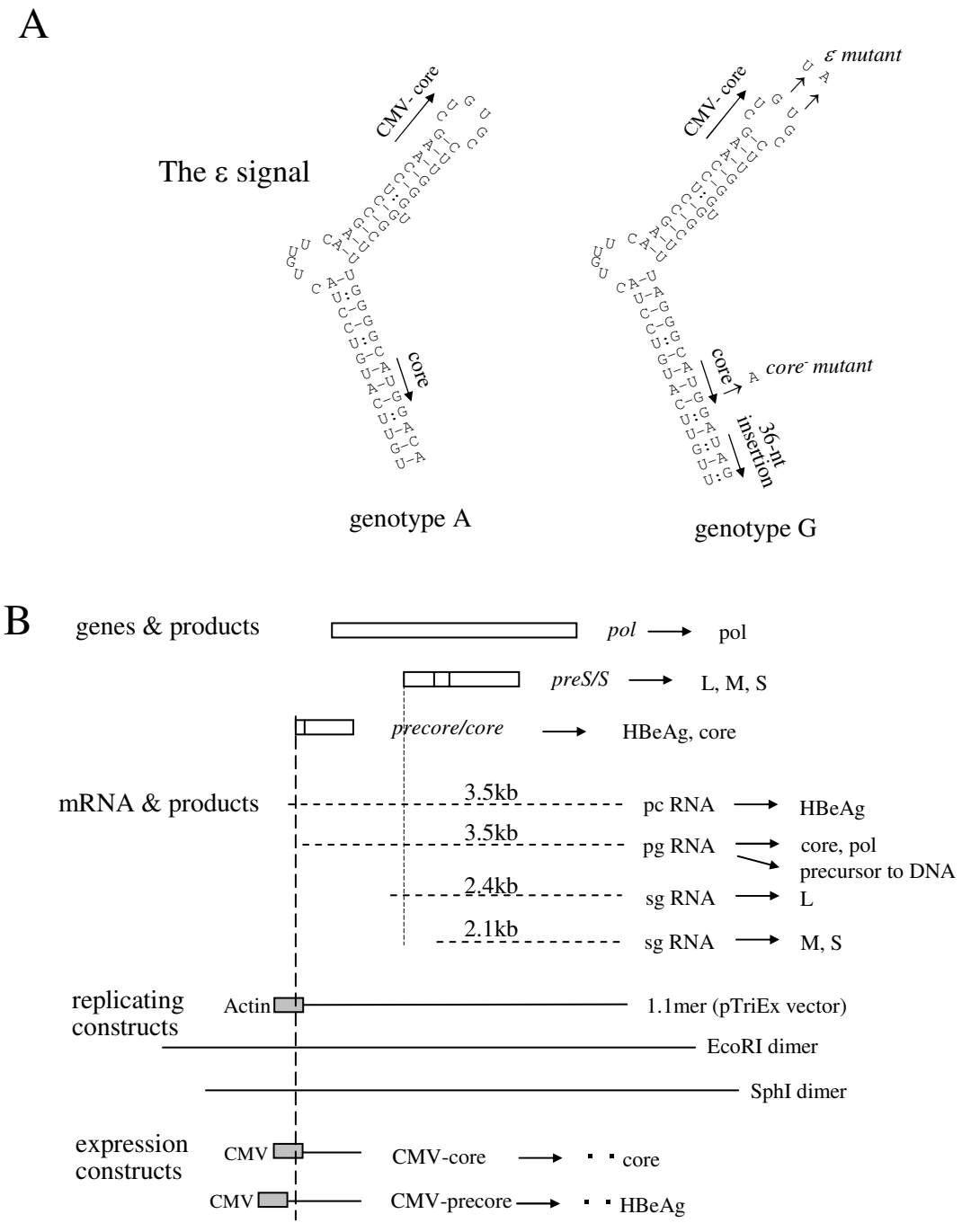


Figure 1

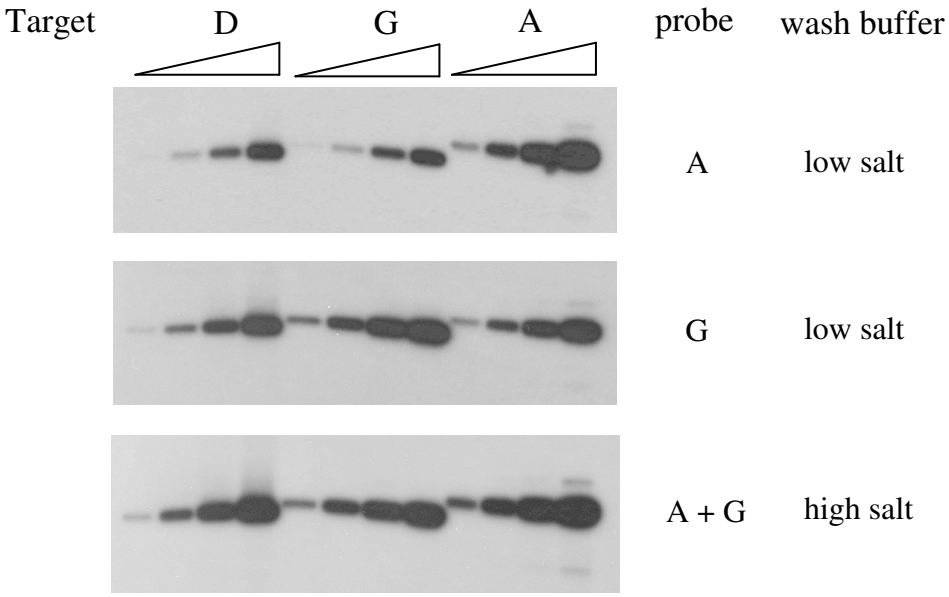


Figure 2

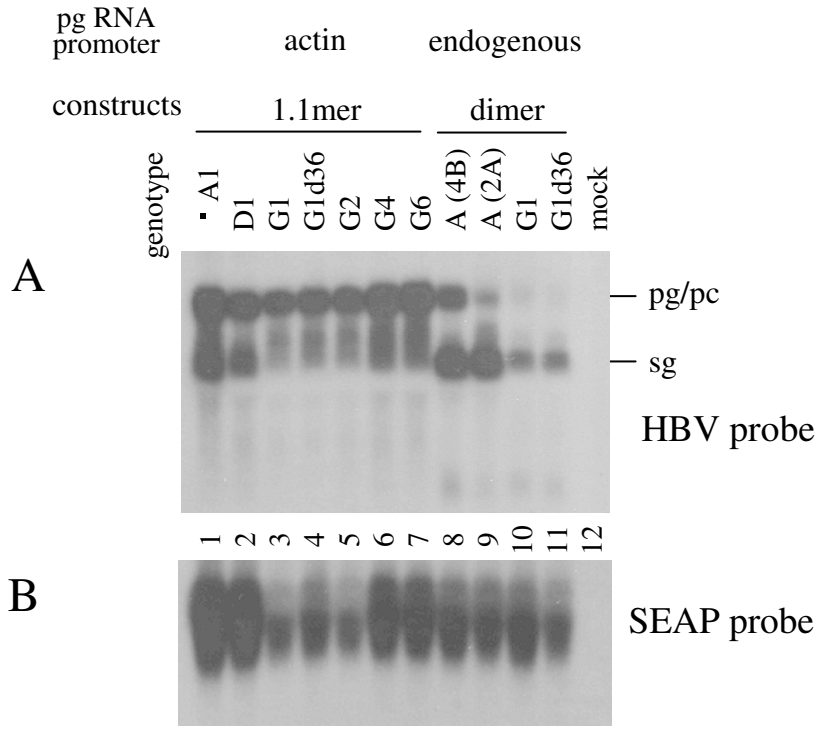


Figure 3

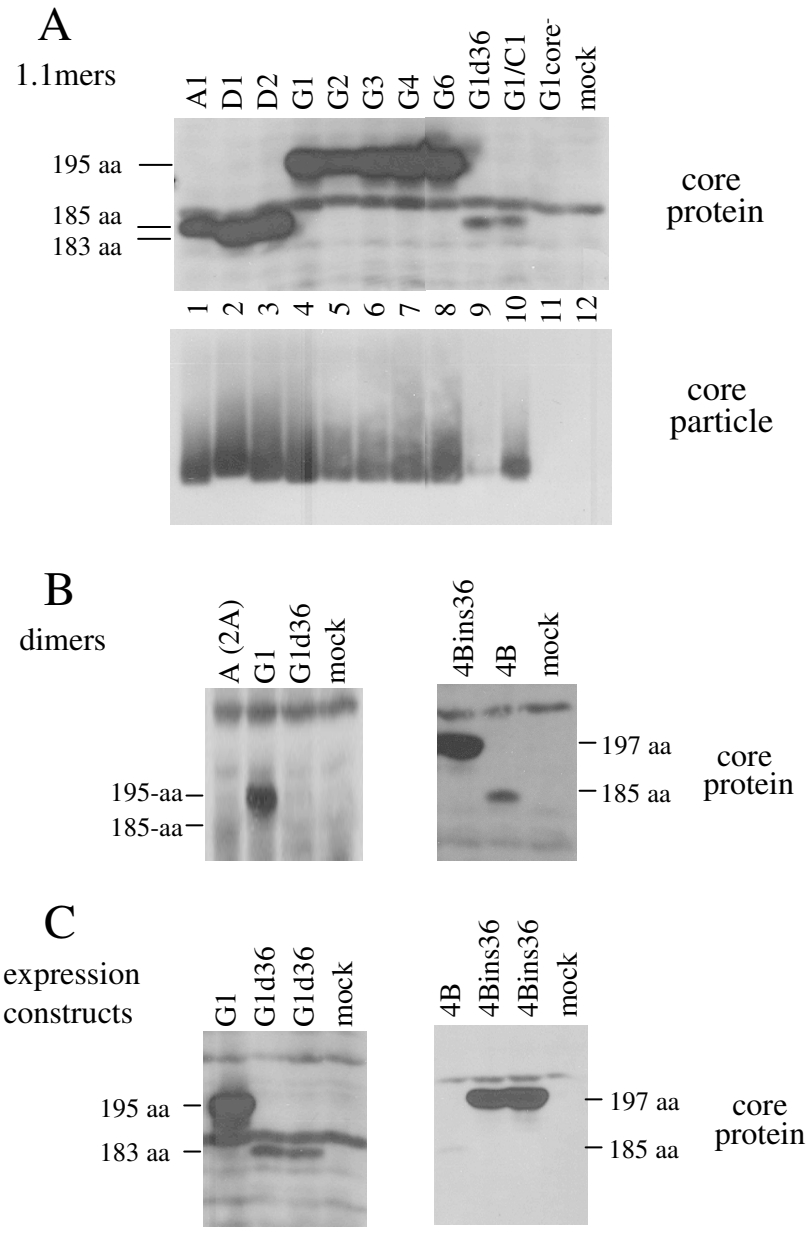


Figure 4

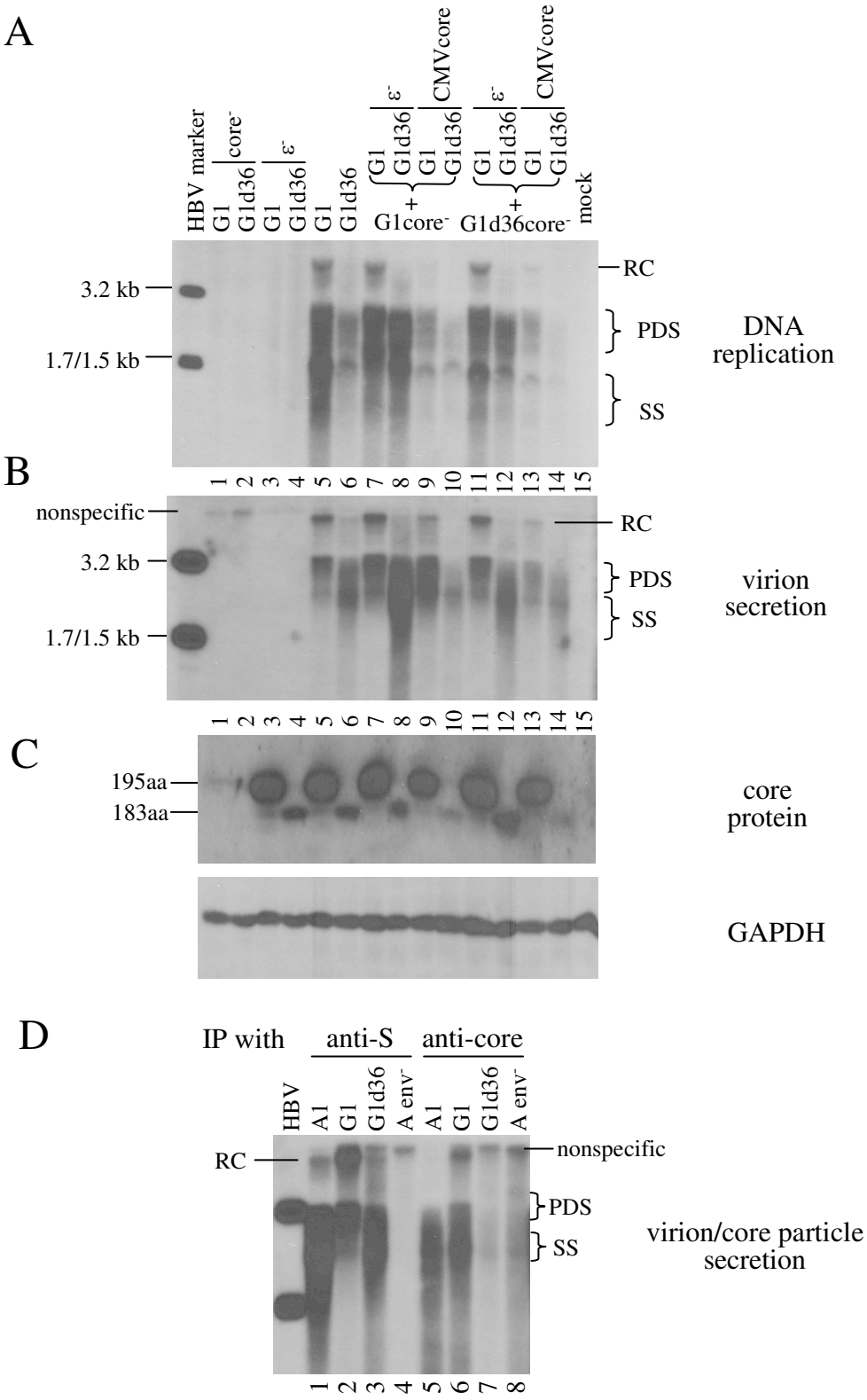


Figure 5

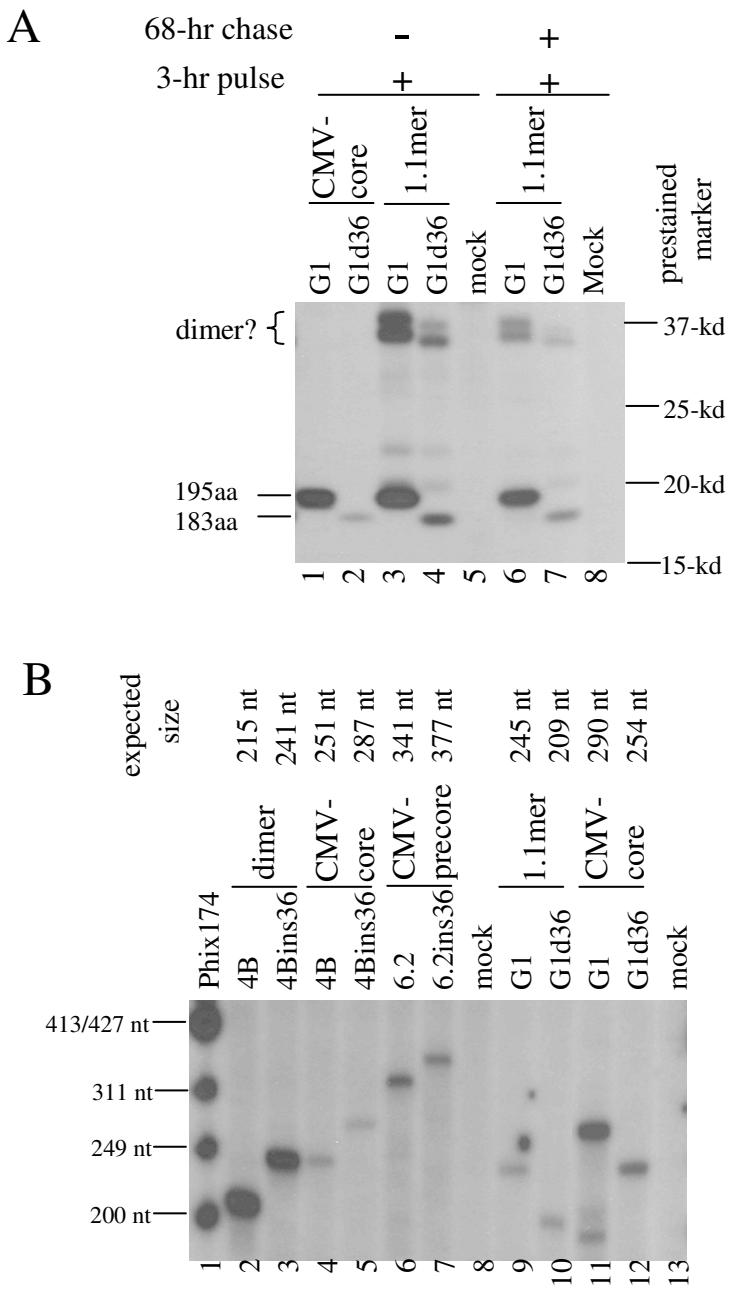


Figure 6

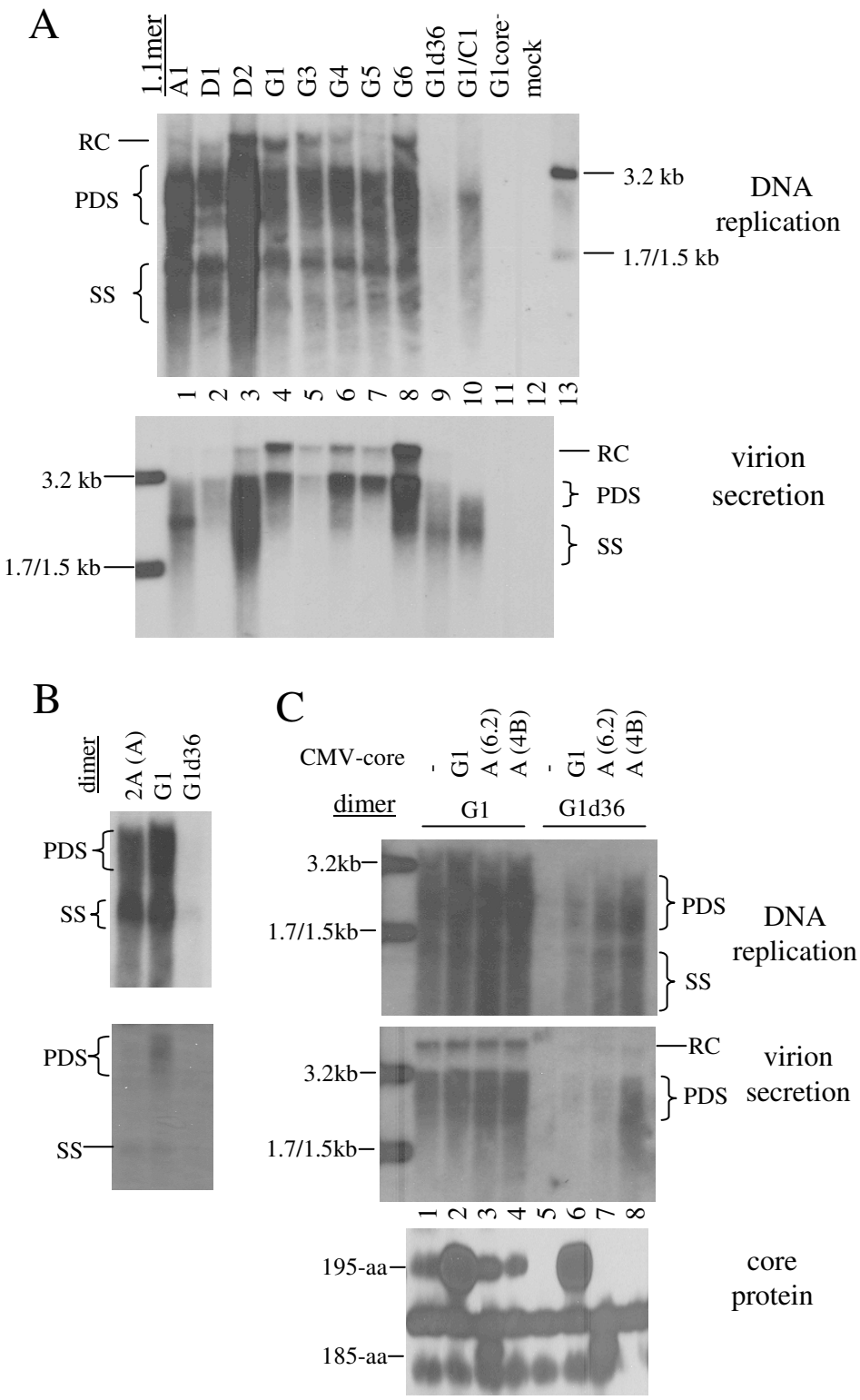
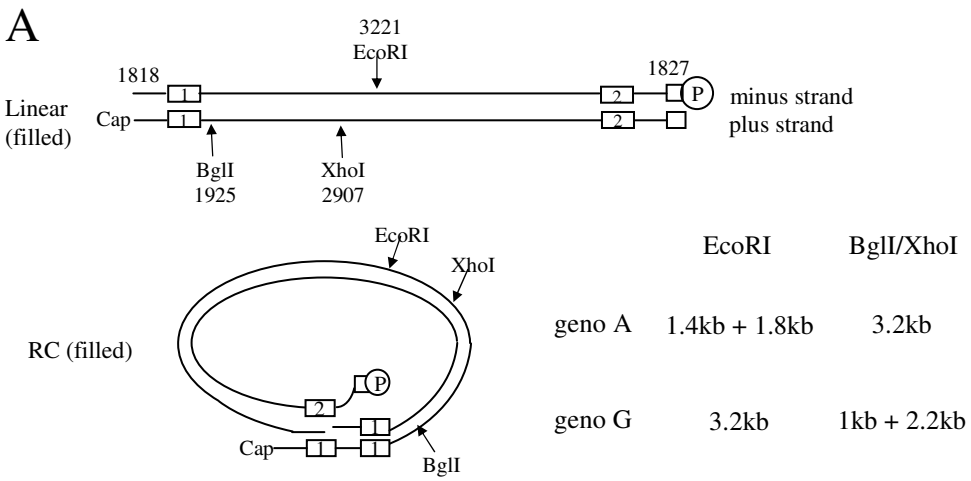
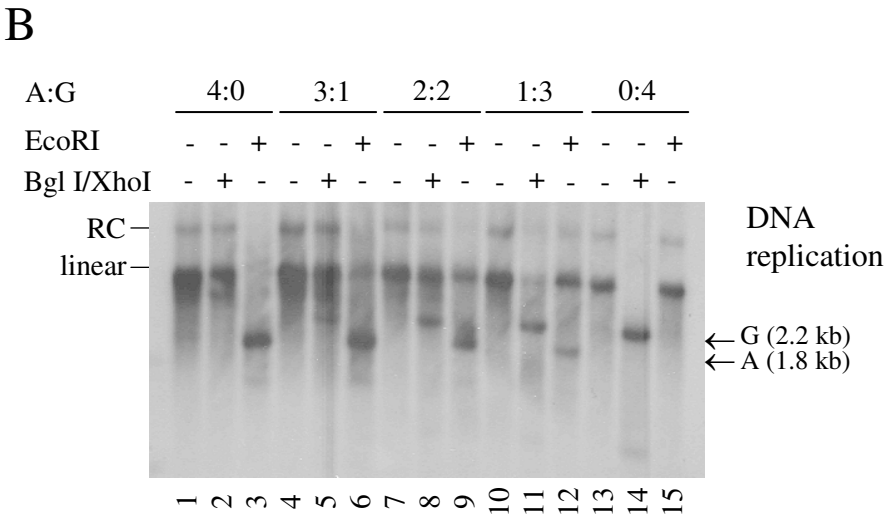


Figure 7





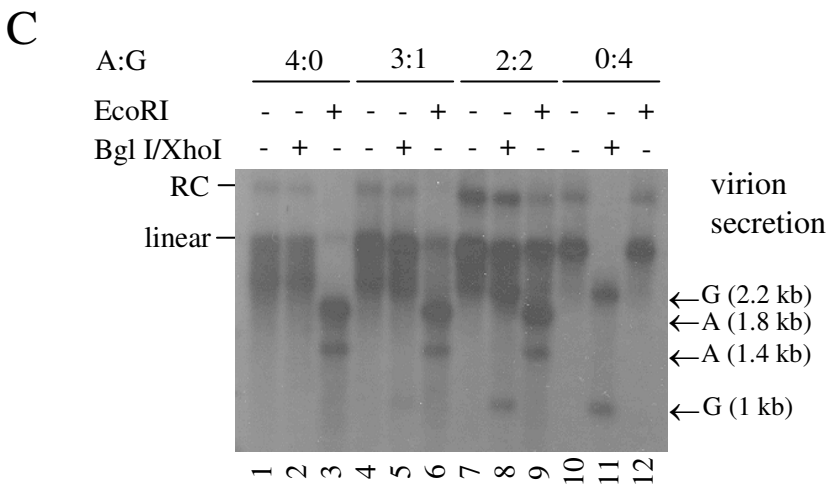


Figure 8

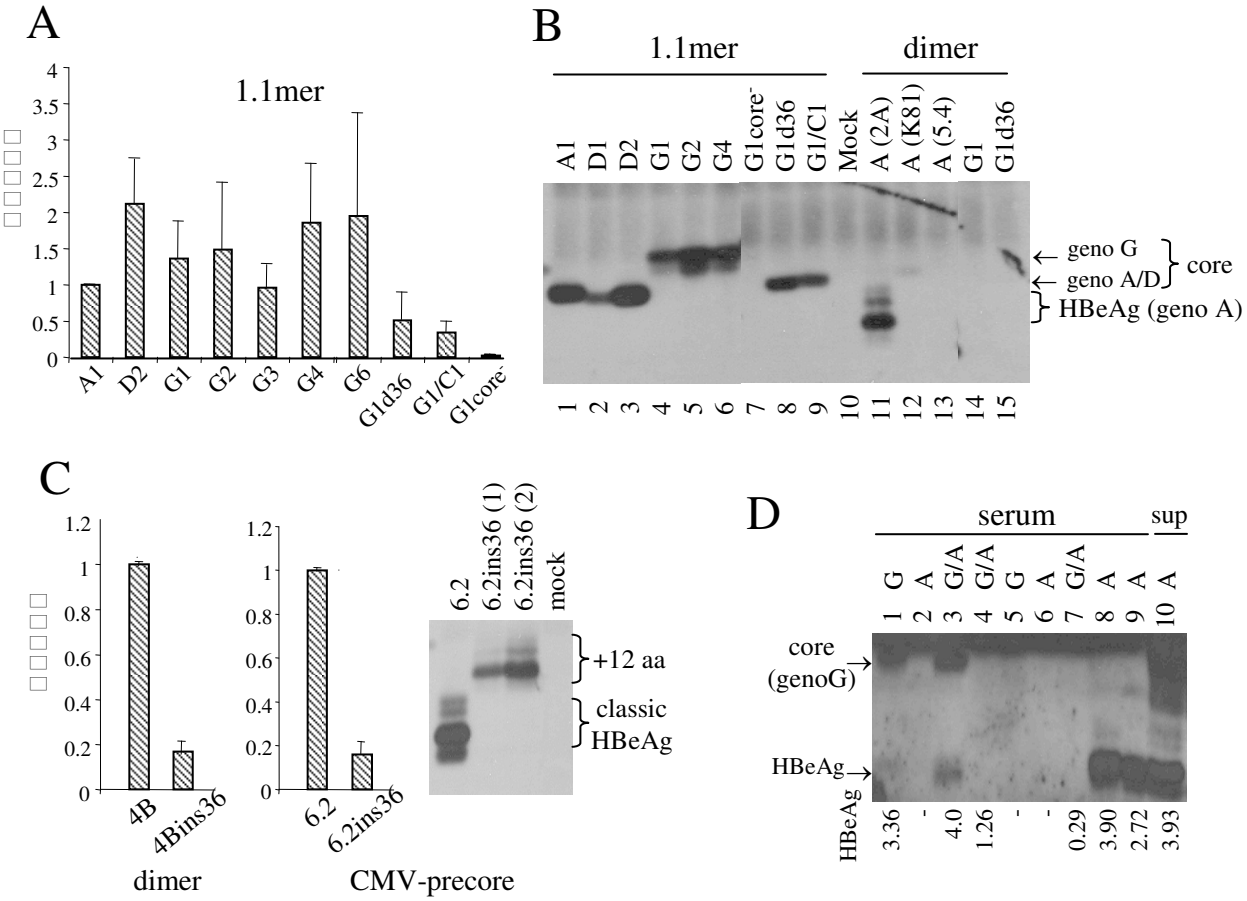


Figure 9

Dear Dr. Sandri-Goldin:

Thank you very much for giving us the opportunity to revise the manuscript entitled "Critical Role of the 36-nucleotide insertion in hepatitis B virus genotype G on core protein expression, genome replication, and virion secretion" (JVI00390-07, version 1). The two reviewers' thoughtful comments are extremely helpful and provide guidance for the improvement of the work. We have performed most of the additional experiments as suggested and modified the manuscript accordingly. The major findings are that 1), contrary to our original hypothesis, the 36-nt insertion increased core protein translation rather than its stability, which was not accompanied by elevation in mRNA level, suggesting more efficient translation; 2), the increased genome maturity of secreted virions was at least partly associated with reduced envelopment of core particles, because G1 secreted more naked core particles than its deletion mutant; 3), the reduced replication of the 1.1mer genome of G1del36 was probably related to both cis and trans effects. In addition, we have determined the complete nucleotide sequences of two genotype G clones and provided control for the Northern blot analysis. Some of the figures have been modified, deleted, added, or their numbers changed.

Fig. 3: the experiment repeated, and the lower panel (EB staining) replaced with hybridization with a SEAP probe;

Fig. 4A: clone G5 deleted as suggested;

old Fig. 5 = new Fig. 7, with clone G2 deleted as suggested;

old Fig. 6 = new Fig. 5, with the experiment redone to include G1del36core$^-$ mutant; a new panel (D) is included to show DNA pattern associated with extracellular naked core particles.

old Fig. 7 = new Fig. 8, with addition of a cartoon showing restriction sites;

old Fig. 8 = new Fig. 9, with clone G5 deleted;

old Fig. 9: deleted as suggested. A supplementary figure is attached for reviewer 1.

New Fig. 6 describes results of metabolic labeling (panel A) and primer extension assay (panel B).

G1del36 has been shortened to G1d36.

The following are point-by-point responses to the questions from the two reviewers, with the reviewers' comments in italics.

REVIEWER 1.

Major points:

1. In Fig. 3, rRNA was used as a control. Since the SEAP cDNA was used for co-transfection to monitor the transfection efficiency, Northern-blot of SEAP RNA will be a more appropriate control. We routinely use only 5ng of the SEAP cDNA for co-transfection. To use SEAP mRNA as a better measurement of transfection efficiency, we have repeated the experiment using 1.2 μg of HBV constructs and 0.8 μg of the SEAP vector. The lower pg /pc RNA and envelope RNA levels of tandem dimers of

genotype G (G1, G1d36) compared to genotype A (2A) was not a consequence of lower transfection efficiency.

2. In Fig. 4A, it is unclear why G1del36 produced a significantly lower level of core particles. Is this due to its enhanced secretion? This needs to be clarified.

The much lower level of core particles inside cells transfected with the G1del36 mutant than G1 is consistent with similar low level of core protein, which appears a direct consequence of reduced protein translation according to the new metabolic labeling experiment. In theory increased protein export could diminish intracellular pool of the protein, yet according to our experience less than 1/10th of viral genome is secreted as virions or naked core particles. Therefore, increased particle secretion should not markedly reduce intracellular pool of core protein/particle. Conversely, we found that ablation of envelope protein expression by a nonsense mutation failed to increase the intracellular pool of HBV DNA. Although G1del36 displayed increased virion secretion than G1, this was achieved at the expense of naked core particles (new Fig. 5D). That also explains why virions of G1del36 contain much higher ratio of single-stranded DNA.

3. In Fig. 4B and 4C, right panels: it is important to show whether 4B and 4Bins36 produced similar levels of pg/pc RNA. This is important for ruling out the possibility that the 36-nt insertion affects viral RNA transcription. We have performed primer extension assays on both dimers and CMV-core constructs, and found no evidence for increased pg RNA level in 4Bins36 relative to 4B (pc RNA was too low to be detected under the sensitivity because 4B is a core promoter mutant with HBeAg expression at a quarter of the wild-type level), or 4Bins36CMV-core than 4BCMV-core. Therefore, the higher core protein level associated with the 36-nt insertion is primarily mediated at the level of translation. This important finding is shown in the new Fig. 6.

4. In Fig. 5B, G1 and A(2A) dimers had similar DNA replication efficiency. However, G1 produced significantly less pg/pc RNA than A(2A) (Fig. 3) and much more core protein than A(2A) (Fig. 4B). These results do not make sense as DNA replication efficiency and the core protein level should both be positively correlated with the pgRNA level. The authors need to provide explanations for these seemingly contradictive results. (Also see comment #7.) We fully agree with the general rule that within the same genotype the pg RNA level correlates with core protein expression and with genome replication. However, a major finding of the present study is the marked up regulation of core protein expression by the 36-nt insertion. In another word comparison between G1del36, but not G1, and other genotypes is valid. Therefore, G1del36 produced less pg RNA than 2A, and consequently its replication was under the detection limit (less efficient than 2A). That also explains why core protein was detected from serum samples of genotype G infection. If, among the three components for replication (core protein, polymerase, the pregenome), core protein is

the limiting factor, then increased core protein expression can augment genome replication despite lower level of pg RNA. At present we don't know whether the 36-nt insertion also modulate polymerase translation.

5. *The quality of Fig. 8D is poor. It is difficult to tell whether the indicated genoG core protein band was indeed the core protein band. A genotype G core protein expressed in cell cultures should be used as a marker in this figure.* We did observe co-migration of core protein expressed from cell culture with one of the genotype G-infected serum sample in the pilot experiment, which did not include all the serum samples shown in Fig. 8D. For Fig. 8D, position of the band relative to the prestained protein size markers is consistent with it being the genotype G core protein. Unfortunately we could not repeat this experiment due to the depletion of precious serum samples.

6. *It is curious that A2 and G1 dimers produced similar levels of p24 and gp27 (Fig. 9), in spite of the observation that G1 produced a much lower level of the S RNA (Fig. 3A). This cannot be explained by the reduced secretion of the G1 S proteins, as the monomers of genotype G and other genotypes had similar S RNA (Fig. 3A) and protein levels (Fig. 9A). This requires some explanations.* The S protein levels were not similar between G1 and 2A in some other experiments. It appears that insufficient dilution of the secondary antibody ablates quantitative differences among samples. With lower concentration of the secondary antibody, 2A clearly displayed stronger p24/gp27 bands than G1. The results shown in Fig. 9A were based on 2μg HBV DNA/well. We have now used serial dilution of HBV DNA for transfection and found that both the intracellular and secreted HBsAg plateaued at 1μg/well for 2A but not for G1. The new result is included as a supplementary figure for review purpose only. In lieu of reviewer 2's comment #7 (see below), we have deleted this figure to concentrate on the impacts of the 36-nt insertion in the core gene.

7. *As indicated in the first paragraph of p.19 by the authors, it is unclear whether the 36-nt insertion affects the core protein translation or its stability. This issue can be easily resolved by a simple pulse-chase labeling experiments.* We have performed the pulse-chase experiments using both 1.1mer genomes and CMV-core constructs of genotype G (the monoclonal antibody used for the detection of 4B core protein by Western blot failed to work in immunoprecipitation). Contrary to our hypothesis, the 36 nucleotides affected core protein translation, rather than its stability. The marked difference between G1 and G1del36 was obvious following not only 3hrs of metabolic labeling (new Fig. 6), but also as short as 1 hr of labeling (data not shown). This, when coupled with the lack of the effect of the insertion on transcript levels, suggest that the insertion anhences translation efficiency. This is reminiscent of our recent findings that point mutations in the precore region can influence core protein translation, apparently without altering the mRNA level

(Guarnieri et al., J. Virol., 2006, 80: 587-595). Certainly, how sequence alterations in the core mRNA or pg RNA modulate core protein translation warrants further studies.

8. The term "stringency of virion secretion" throughout the text is rather confusing. It is unclear whether the more stringent secretion means more secretion or less secretion. The authors also seemed to use "stringency" to imply "DNA maturity" (e.g., p14, lines 4-5 from the bottom). In the last sentence of p.14, the authors indicated "the 12-aa insertion in core protein…was responsible for stringent virion secretion unique to genotype G." However, in Fig. 5A, G2 DNA was secreted efficiently whereas G3 DNA was secreted poorly. This is very confusing. We apologize for the confusion. By "stringency of virion secretion" we really meant "DNA maturity", although based on the phenotype of G1del36 the insertion also appears to reduce the efficiency of secretion. Interestingly, while G1del36 displayed increased virion secretion and reduced genome maturity, it secreted much less naked core particles in comparison to G1. The reviewer is right in pointing out the anomaly of G2 and G3 in Fig. 5. Of the two clones, the poor secretion of G3 has been consistently observed, which is correlated with its poor HBsAg secretion (Fig. 9B). The low DNA replication of G2 is not typical. We have deleted clone G2 from this figure. Irrespective of quantitative anomaly of clone G3, all the genotype G clones secreted virus particles of more mature genome than other genotypes.

Minor comments:

9. In the last paragraph of p.21, the authors indicated that "subviral particles, which are secreted less efficiently, may be important for the establishment of persistent infection". It is unclear what the basis is for this statement is. It has been demonstrated by Bruns et al. (J. of Virol. 72:1462-8, 1998) that subviral particles could enhance HBV infection. The enhancement observed by Bruns et al occurred in DHBV, where viral and subviral particles contain similar proportion of large envelope protein. Moreover, the enhancement only occurred at low M.O.I. of 0.01 and 0.1. Our hypothesis is that higher level of S protein will delay the rise of anti-HBs neutralizing antibodies. Since we have deleted results on HBsAg expression, this sentence has been removed from the Discussion.

10. In p.2, first line of the abstract, "101" appears to be a mistake. We have deleted this from the abstract.

11. References or the manufacturer of the InnoLiPA test should be provided. Manufacture has been provided.

REVIEWER 2:

1. Multiple clones of genotype G were reported, with some differences in results (e.g., G5 vs. the other clones in figure 4A; G2 vs. the other clones in figure 5A). Are all these cones sequenced and what are the sequence differences among the different clones? If this information is unavailable, it may be better to focus on one or few clones with complete sequence information. Otherwise, it is difficult to interpret the results from the different clones. Clones G1-G3 were derived from one patient, while clones G4-G6 were obtained from another patient. We have now determined complete nucleotide sequences of clone G1 and G6. We agree with this reviewer that some of the clones may represent viral variants/mutants, but have not sequenced all the clones to understand the structural basis of the abnormality. As mentioned in response to Reviewer 1, we usually observe much higher replication of G2 than shown in this blot. The high mobility of G5 core protein was consistently observed but the structural basis is unknown. We have deleted G2 from the Southern blot analysis and G5 from the Western blot as suggested.

2. The results from figure 5 and 6 are consistent with the notion that the 36 nt insertion increased genotype G DNA replication in cis. However, as the authors pointed out rightly in the discussion, a critical construct to be tested is G1del36 core-, complemented by G1core and G1del36core, in order to more vigorously differentiation the effect in cis vs. in trans. Also, the authors suggested the cis effect may be at the level of pgRNA packaging but no data were provided. Differentiating the cis effect on pgRNA packaging vs. DNA synthesis would improve the manuscript significantly. We have made G1del36core^{-} mutant, but the results were more complex than we predicted because co-transfection of this mutant with G1ε- mutant still led to efficient genome replication. Co-transfection of G1del36 core- mutant with its ε-minus mutant produced less efficient genome replication, but still more efficient than the G1del36 parental clone. Our interpretation is that a threshold of core protein expression level is required for efficient genome replication. The trans-complementation assay failed to fully reproduce the low replication phenotype of the G1del36 mutant possibly because under such conditions the pg RNA was devoted to a single function of protein translation or packaging / replication. Nevertheless, results of the additional trans-complementation assay validate the trans effect of the insertion (core protein rather than pg RNA) on genome maturity of secreted virus particles.

3. A comparison of viral DNA in intracellular cores and extracellular virions (figure 5 and 6) indicates that the effect of the genotype G insertion, at the core protein level, on the pattern of viral DNA (levels of RC DNA vs. less mature DNA intermediates), whether intracellular or extracellular, was similar. This suggests that the effect of the insertion was exerted at the stage of intracellular DNA maturation, rather than the stage of virion secretion as the authors concluded. A careful quantitative analysis from multiple experiments will be required to differentiate these possibilities. We fully agree with the reviewer that the 36 nt not only impact

genome maturity of extracellular virions, but also intracellular core particles. Such a concerted effect is most consistent with the role of the 12-aa insertion in reducing the efficiency of core particle envelopment. In this regard, we have now compared G1 with G1del36 in terms of extracelluar naked core particles, with genotype A as control. G1 secreted much higher level of naked core particles than G1del36, and also higher proportion than genotype A. The naked core particles of G1 have less mature genome than corresponding virions, while naked core particles of G1del36 had similar DNA pattern as corresponding virions (new Fig. 5D).

4. *It was stated that digestion with Bgl II and Xho I converted G4 DNA into the same two fragments, whether RC or linear DNA (figure 7). This is puzzling. A restriction map of genotype G should be provided with these restriction sites marked. Digestion of the linear DNA with the two single cutters should have generated three, not two, fragments. Indeed, given the unusual pattern of the DNA replicative intermediates (very little to no mature "RC" DNA, much more abundant "PDS" DNA, and the inconsistent migration of the "SS" DNA; figure 5, 6, 7), it may be necessary to verify the identities of these presumptive DNA species before a firm conclusion can be drawn regarding the effect of the core protein insertion on DNA synthesis.* We apologize for not providing information on the restriction enzyme sites in the HBV genome. BglI (not BglII) site is located just downstream of DR1 (position 1925), and XhoI site is located at position 2907. The EcoRI site of genotype A is at position 1. Thus, all the sites are within the double-stranded region of the linear or relaxed circular DNA and cleavable by a restriction enzyme, although that is irrelevant here since the single-stranded region was filled in by Klenow fragment. The fill-in reaction also enabled the 3' half of the genome (in terms of the positive strand) to migrate as a discrete band rather than smear following restriction enzyme digestion. We apologize for not mentioning that for data shown in this figure (but not in other figures), all the DNA samples were heated at 85^{o}C for 10 min in order to break the cohesive ends responsible for the different migration between RC DNA and linear DNA. That explains why in this figure, there was very little genotype G DNA remaining at the RC position. With the double digestion and heating step, both RC and linear DNA should generate two big fragments of 1 kb and 2.2 kb. Only the third, much smaller band would have different migration pattern between RC and linear DNA forms (0.34-kb for the former but 0.1-kb for the latter form). The digestion patterns of both genotypes A and G in the figure are consistent with our interpretation.

5. *Since it is still unknown whether the unusual migration pattern of HBsAg on SDS/PAGE and the reduced HBsAg secretion (figure 9) are related to the 36 nt insertion or some other unknown genetic changes, it may be more appropriate to delete these results so as to focus better on the effects of the insertion.* We have deleted figure 9 as suggested, and removed corresponding sections in the Discussion section.

In conclusion, we have addressed all the major questions raised from both reviewers, and reached a better understanding of how the 36-nt insertion affects core protein expression and the 12-aa insertion regulates virion secretion. We hope you find the extensively revised manuscript much improved and acceptable to Journal of Virology.