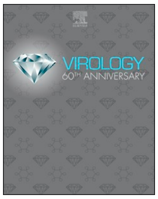
Virology 526 (2019) 203–213

Contents lists available at ScienceDirect

Virology

journal homepage: www.elsevier.com/locate/virology



Core gene insertion in hepatitis B virus genotype G functions at both the encoded amino acid sequence and RNA structure levels to stimulate core protein expression

Jing Zhang[a,1], Li Zong[a,1], Yongxiang Wang[a], Cheng Li[a], Chaoyang Chen[a], Yumei Wen[a], Jisu Li[b], Shuping Tong[a,b,*]

[a] *Key Laboratory of Medical Molecular Virology, Shanghai Medical College, Fudan University, Shanghai, China*
[b] *Liver Research Center, Rhode Island Hospital, The Alpert Warren School of Medicine, Brown University, Providence, RI, USA*

ARTICLE INFO

Keywords:
Core protein
Genotype
Hepatitis B virus
Open reading frame
RNA secondary structure
Translational control

ABSTRACT

Hepatitis B virus genotype G possesses a 36-nucleotide (nt) insertion at the 5′ end of core gene, adding 12 residues to core protein. The insertion markedly increased core protein level irrespective of viral genotype, with the effect reproducible using CMV-core gene construct. Here we used such expression constructs and transient transfection experiments in Huh7 cells to identify the structural bases. The insertion is predicted to create a stem-loop structure 14nt downstream of core gene AUG. A + 1 or + 2 frameshift into the 36nt mitigated enhancement of core protein level. Point mutations to disrupt or restore the stem-loop had opposite effects on core protein expression. Shifting the translation initiation site downstream or further upstream of the stem-loop rendered it inhibitory or no longer stimulatory of core protein expression. Therefore, both the reading frame and a properly positioned stem-loop structure contribute to marked increase in core protein expression by the 36-nt insertion.

1. Introduction

An estimated 257 million people worldwide are living with chronic hepatitis B virus (HBV) infection, who are at great risk of developing hepatic decompensation, liver cirrhosis, and hepatocellular carcinoma (HCC) (Liaw et al., 2009; Trepo et al., 2014). The high morbidity and mortality associated with chronic HBV infection have become a severe health problem (Stanaway et al., 2016). HBV has four open reading frames arranged on the small circular DNA genome of 3.2 kb: capsid (precore/core), polymerase (P), envelope (preS1/preS2/S), and X. Core protein is the translation product of the core gene alone, whereas translation initiated from the upstream precore ATG generates precore/core protein, which following proteolytic cleavage is secreted as hepatitis B e antigen (HBeAg). Whether the core protein or HBeAg is expressed depends on the transcript. Four size forms of HBV mRNAs (3.5-, 2.4-, 2.1-, and 0.7-kb) are generated from the covalently closed circular (ccc) HBV DNA in the nucleus. The 3.5-kb HBV RNA (terminally redundant) is under the transcriptional control of the core promoter and has imprecise transcription initiation sites upstream or downstream of precore ATG codon. Consequently the longer form (precore RNA or pcRNA) is responsible for HBeAg expression, while the shorter form (pregenomic RNA or pgRNA) is the mRNA for core protein expression, and also P protein translation through ribosomal leaky scanning or ribosomal shunting (Cao and Tavis, 2011). Moreover, pgRNA (but not any other HBV RNA) can be co-packaged with P protein into nucleocapsid assembled from core protein, where it is converted by P protein into partially double stranded DNA, thus fulfilling HBV genome replication.

Besides being the building block of nucleocapsid essential for HBV genome replication, the core protein can bind to cccDNA to regulate HBV RNA transcription (Bock et al., 2001; Chong et al., 2017; Guo et al., 2011). It is highly immunogenic and a major trigger of immune-mediated liver injury. In contrast, HBeAg is not essential for HBV replication but rather plays an immunomodulatory function to promote persistent infection during HBV transmission to a new host (Milich, 2003). It can modulate both innate and adaptive immunity (Tian et al., 2016; Yu et al., 2017). Thus, regulation of core vs. HBeAg expression has important biological and clinical implications. Naturally occurring core promoter mutations, with A1762T/G1764A being the most common, up regulate core protein expression and genome replication

---

[*] Correspondence to: Liver Research Center, 55 Claverick Street, Providence, RI 02903, USA.
E-mail address: shuping_tong_md@brown.edu (S. Tong).

[1] These authors contributed equally to this work.

https://doi.org/10.1016/j.virol.2018.11.002
Received 24 August 2018; Received in revised form 31 October 2018; Accepted 1 November 2018

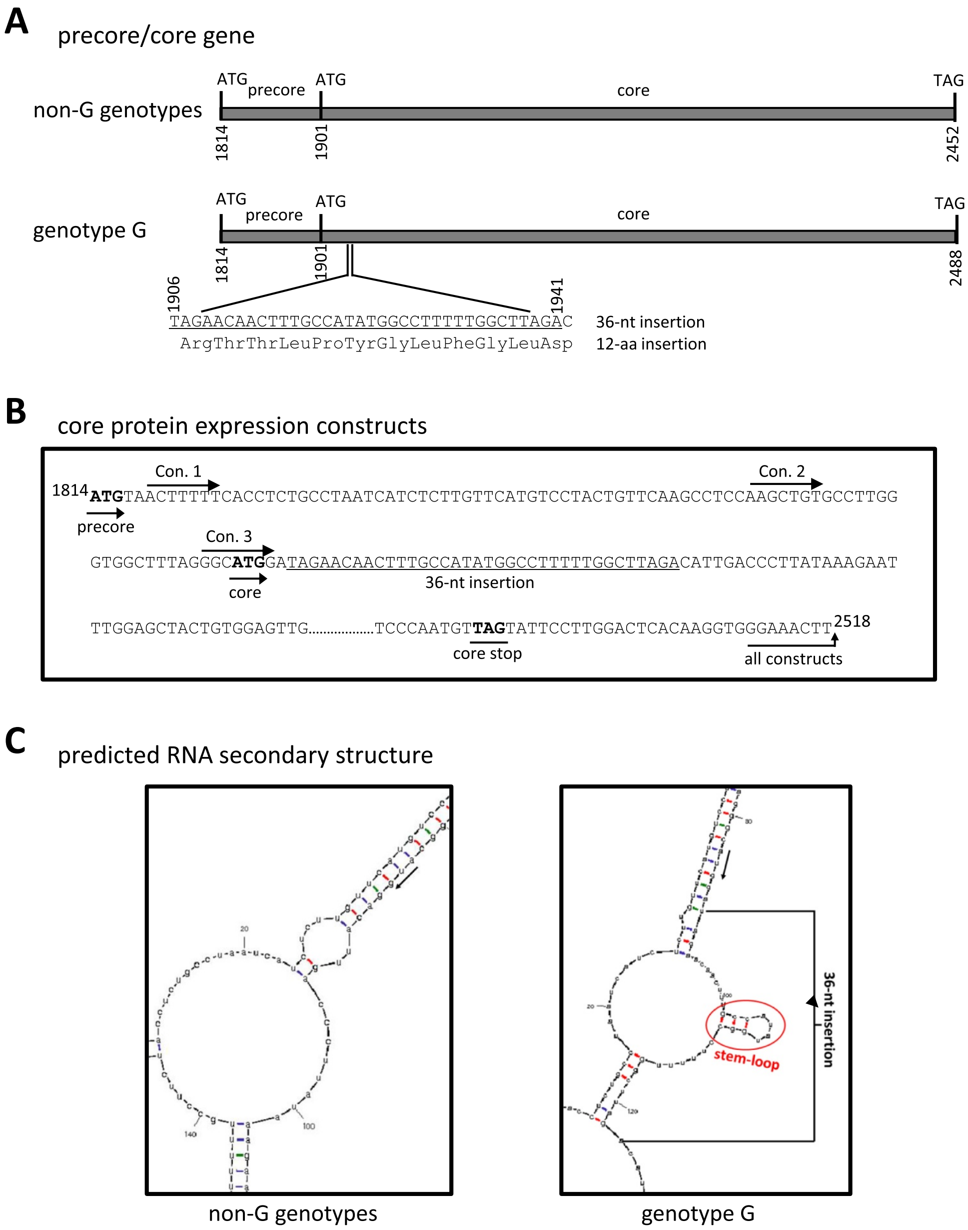


Fig. 1. The 36-nt insertion in the core gene, its impact on RNA secondary structure, and constructs used to determine its impact on core protein expression. **(A)** Schematic representation of the precore/core gene of non-G genotypes vs. genotype G. Two alternative translation initiation sites (ATGs) in the precore region (1814) and core gene (1901) drive the expression of HBeAg and the core protein, respectively. Genotype G has a 36-nt insertion (underlined) between the 5th and 6th positions of core gene, thus adding 12 extra residues to the core protein. **(B)** The three types of core protein expression constructs used. Shown are DNA sequences of the entire precore region as well as the 5′ and 3′ ends of the core gene of genotype G. The 36 extra nucleotides are underlined, while the two translation initiation sites (ATGs) as well as the termination site (TAG) are indicated. Core protein expression constructs Con. 1, Con. 2, and Con. 3 have their 5′ end located at position 1819, 1874, and 1898 in the precore region, respectively, and a shared 3′ end at position 2518 downstream of core gene. The HBV DNA fragment was inserted to pcDNA3.1zeo (-) vector for core protein expression under the CMV promoter. **(C)** Predicted RNA secondary structures of the core gene of no-G genotypes and genotype G. The Mfold program predicts that the 36-nt insertion generates a small stem-loop structure (red circle) 14nt downstream of the core gene AUG (arrowhead).

but diminish HBeAg expression at the transcriptional level (Buckwold et al., 1996; Moriyama et al., 1996; Okamoto et al., 1994; Scaglioni et al., 1997). The impact of the 1762/1764 mutations can be markedly enhanced by T1753C, C1766T, and T1768A mutations (Baumert et al., 1996; Hasegawa et al., 1994; Parekh et al., 2003; Tsai et al., 2009). Mutations in the precore region abolish HBeAg expression at the translational level through nonsense or frameshift mutations, or mutated precore ATG codon (Brunetto et al., 1990; Carman et al., 1989; Okamoto et al., 1990; Tong et al., 1990; Tong and Revill, 2016). The most common HBeAg-negative precore mutation is G1896A converting penultimate precore codon from TGG to TAG. Such a nonsense mutation could enable pcRNA to produce some core protein through translational termination-reinitiation. Combination of core promoter mutations with precore mutation may cause fulminant hepatitis during acute infection (Friedt et al., 1999; Liang et al., 1991; Omata et al., 1991; Tripathy et al., 2011). In addition, core promoter mutations are implicated in liver cirrhosis and hepatocellular carcinoma during chronic infection.

According to sequence divergence of 8% or greater in the entire genome, HBV isolates worldwide can be classified into ten genotypes (Kramvis, 2014; Norder et al., 2004; Okamoto et al., 1988; Tong and Revill, 2016). Genotype G was first recognized in 2000 (Stuyver et al., 2000), although the first sequence was reported earlier (Bhat et al., 1990). It is often detected from homosexual men with HIV co-infection, and in association with another HBV genotype such as genotype A (Araujo et al., 2012; Kato et al., 2002; Osiowy et al., 2008; Sánchez et al., 2006; van der Kuyl et al., 2013). Genotype G is unable to express HBeAg due to nonsense mutations in the precore region. Another unique feature is insertion of 36 nucleotide (nt) (TAGAACAACTTTGCCA TATGGCCTTTTTGGCTTAGA) between the 5th and 6th positions of core gene, which adds 12 extra amino acid residues into the core protein (Fig. 1A). Our previous studies found that the 36-nt insertion was associated with high level of core protein expression. Removing the 36nt from a genotype G clone nearly abolished core protein expression and markedly reduced genome replication (Li et al., 2007). Conversely, inserting the 36nt into clones of genotype A and genotype D greatly enhanced core protein expression (Gutelius et al., 2011). In the present study, we aimed to dissect the relative contribution of the 36-nt insertion at the mRNA level vs. the 12-aa insertion at the protein level on core protein expression. Interestingly, the 36-nt insertion is predicted to create a small stem-loop structure 14nt downstream of the core gene AUG codon (Fig. 1C).

2. Materials and methods

2.1. DNA constructs

Core protein expression constructs used in this study are based on clone G1 of genotype G (Li et al., 2007). Con. 1, Con. 2, and Con. 3 (Fig. 1B) are CMV promoter-driven core protein expression constructs with variable precore sequence at the 5′ end. They were generated by cloning HBV genomic sequence 1819 – 2518, 1874 – 2518, and 1898–2518, respectively, into the SacI/HindIII sites (Con. 1) or XhoI/EcoRI sites (Con. 2 and Con. 3) of pcDNA3.1zeo (-) vector. The corresponding HBV DNA fragments were amplified by polymerase chain reaction (PCR) using SphI dimer of clone G1 (Li et al., 2007) as the template, with restriction sites attached to the primers (see Table S1 for primer sequences). Point mutations were introduced by overlap extension PCR followed by replacement of the SacI/HindIII or XhoI/EcoRI fragment of the wild-type construct. The envelope protein construct has a 2.3-kb HBV DNA fragment of positions 2721–3215/1–1770 inserted upstream of the SV40 polyadenylation signal and cloned into the pBluescript vector. It is capable of expressing all the three envelope proteins under endogenous HBV promoters and enhancers (Garcia et al., 2009). For the current study, the envelope gene of the expression construct was derived from geno22.5 of genotype B (Qin et al., 2011). The EcoRI dimer of clone 4B of genotype A has been described (Parekh et al., 2003), and its core-minus (core$^-$) mutant contains a C2044G nonsense mutation in the core gene to prevent core protein expression (Gutelius et al., 2011).

2.2. Transient transfection and detection of protein expression

Huh7 cells were cultured in Dulbecco's Modified Eagle's Medium (DMEM) supplemented with 10% fetal bovine serum. Transient transfection was performed on cells seeded in 6-well plates at a density of ~ 80%, using TransIL-LT1 reagent (Mirus) and 2 μg of DNA. Cells and culture supernatant were harvested 3 days later. Cells were lysed in 100 μl of lysis buffer (10 mM HEPES, pH 7.5; 100 mM NaCl; 1 mM EDTA and 1% NP40), and protein concentration in cell lysate was quantified by Pierce™ BCA Protein Assay Kit (Thermo Scientific). Core and envelope proteins in cell lysate were detected by Western blot analysis. A total of 30 μg of proteins were separated in SDS-12% polyacrylamide gel (PAGE) and transferred to a polyvinylidene fluoride (PVDF) membrane. The blot was incubated at 4 °C overnight with polyclonal rabbit anti-HBc (Dako, 1:3000), anti-HBs (Novus, 1:3000) diluted in 5% milk/TBST. The blot was washed and incubated at room temperature for 1 h with a 1:0,000 dilution of goat anti-rabbit antibody conjugated with horse radish peroxidase (HRP). Signals were revealed by enhanced chemiluminescence (ECL) and visualized by chemiluminescent imaging system (Tanon). For loading control, the blot was treated with stripping buffer and incubated successively with 1:3000 dilution of mouse antibody against glyceraldehyde-3-phosphate dehydrogenase (GAPDH) or β-actin (at 4 °C overnight) and HRP-conjugated goat anti-mouse antibody (1:20000 dilution). The grey values of signals on the blots were measured by Multi Gauge V2.2 software. Hepatitis B surface antigen (HBsAg) in culture supernatant and cell lysate was quantified by an ELISA kit (KHB) with sufficient sample dilution to avoid signal saturation.

2.3. Trans-complementation assay and Southern blot analysis

Huh7 cells seeded in 6-well plates were co-transfected with 1 μg of the core$^-$ EcoRI dimer of clone 4B of genotype A and 1 μg or 0.25 μg of core protein expression construct. Alternatively, 1 μg of the core$^-$ EcoRI dimer or wild-type EcoRI dimer was transfected alone to serve as negative and positive controls, respectively. Cells and culture supernatant were harvested at day 4 post-transfection. Details of HBV DNA analysis have been described (Garcia et al., 2009; Gutelius et al., 2011; Jia et al., 2017; Li et al., 2007; Parekh et al., 2003; Qin et al., 2011). Core particles were precipitated from 1/4th of cell lysate, and following proteinase K digestion, DNA was extracted with phenol and precipitated with ethanol. Purified DNA was dissolved in Tris buffer and separated in 1.3% agarose gel. Following overnight transfer to a nylon membrane, the blot was hybridized with α-P^{32}-labeled dCTP probe of clone 4B (purified 3.2-kb HBV DNA devoid of vector sequence). The signals on the phosphor screen were scanned by Typhoon FLA 9500 software. Virions were immunoprecipitated from culture supernatant by a combination of anti-preS1 and anti-S antibodies preconjugated to protein G-Sepharose beads, and DNA extracted from virions was subject to Southern blot analysis. Densitometric values of signals on the blot were measured by Multi Gauge V2.2 software.

2.4. Statistical analysis

All experiments were repeated for at least 3 times, and the data were expressed as mean ± SD. Statistical analysis was performed by GraphPad Prism 6 software. Differences between the groups were examined by using a Student *t*-test. $^*P < 0.05$ was considered as statistically significant.

2.5. Prediction of RNA secondary structure

RNA secondary structure was predicted using Mfold software (http://mfold.rna.albany.edu/?q=mfold), and confirmed by additional prediction softwares such as Sfold (http://sfold.wadsworth.org/cgi-bin/srna.pl) and RNAfold web server (http://rna.tbi.univie.ac.at/cgi-bin/RNAWebSuite/RNAfold.cgi). Such softwares often provide several possible secondary structures together with their free energies. The secondary structure shown in the figures for each construct is often listed at the top by the prediction software and has the lowest free energy.

3. Results

The 36-nt insertion greatly increased core protein expression, whether in the native genotype G or in non-G genotypes (Gutelius et al., 2011; Li et al., 2007). The impact of the 36-nt insertion could be demonstrated in tandem dimer with pgRNA transcription driven by the endogenous core promoter, 1.1x HBV DNA construct where pgRNA transcription is driven by the chicken actin promoter, as well as the simple core gene construct with the precore/core sequence (lacking the precore ATG at its 5′ end) inserted to a mammalian expression vector behind the CMV promoter (Li et al., 2007). Since such CMV-core gene constructs are much easier to make than overlength constructs, in the present study we employed such constructs to further clarify the structural bases whereby the 36-nt/12-aa insertion markedly increases core protein level. Three slightly different versions of CMV-core expression construct were employed: Con. 1, Con. 2, and Con. 3. They have the same 3′ end 30nt downstream of the core gene but differ in the length of the precore sequence at the 5′ end (Fig. 1B). Most mutants (G1-G13) were made in the Con. 1 form, which has the 5′ end of the HBV sequence (position 1819) corresponding to the 5′ end of pgRNA. For Con. 1 the *Sac*I site at the 3′ end of the CMV promoter was used as the 5′ cloning site to insert the HBV sequence, thus avoiding vector sequence at the 5′ end of the transcript. Con. 2 and Con. 3 were used to shift the relative position of the stem-loop structure and had less precore sequence at the 5′ end.

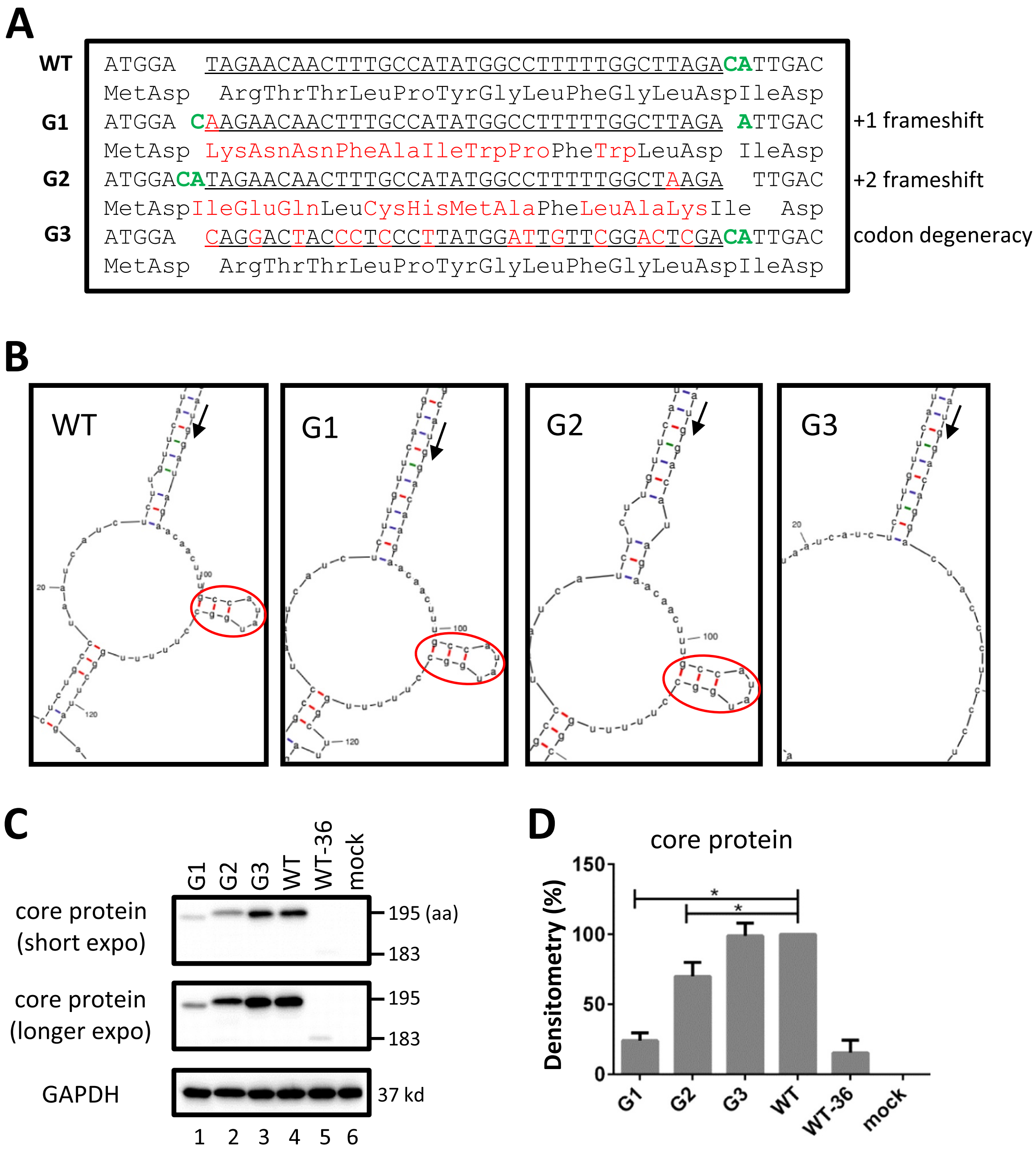


Fig. 2. Impact of frameshifting vs. synonymous mutations introduced to the 36nt on core protein expression. **(A)** Comparison of the nucleotide and amino acid sequences between the wild-type (WT) construct (Con. 1) and its three site-directed mutants. G1 and G2 had local + 1 and + 2 frameshift mutations, respectively, altering amino acid sequence of the 12 inserted residues. G3 maintained the wild-type amino acid sequence despite 14 point mutations. The red color indicates altered nucleotides and amino acids. The green color indicates two nucleotides (CA) immediately downstream of the 36-nt insertion in the WT construct. They were partly or completely moved upstream of the 36-nt insertion in G1 and G2. **(B)** The stem-loop structure created by the 36-nt insertion is maintained in G1 and G2, but lost in G3. **(C) & (D)** Core protein expression in transfected human hepatoma cells. The core protein constructs (2 μg) were transiently transfected to Huh7 cells seeded in 6-well plates, and cells were harvested three days later. **(C)** A total of 30 μg of cellular proteins was loaded for Western blot analysis with a rabbit polyclonal anti-core antibody. GAPDH served as the loading control. **(D)** Densitometric values of core protein in Western blots from three independent transfection experiments were determined and compared. The WT value is set arbitrarily at 100%.

3.1. Increased core protein expression associated with the 36-nt insertion required its translation into amino acid sequence ArgThrThrLeuProTyrGlyLeuPheGlyLeuAsp

The 36nt are inserted between the 5th and 6th positions in the core gene of HBV genotype G, thus adding amino acid sequence ArgThrThrLeuProTyrGlyLeuPheGlyLeuAsp between the 2nd and 3rd residues of core protein (Fig. 1A). An immediate question is whether the marked increase in core protein level is attributed to mutations at the nucleotide level (36-nt insertion), or depends on the sequence of the 12 extra residues. To verify this issue we generated three site-directed mutants. In mutant G1, the 36-nt insertion was shifted down by 1 nucleotide between the 6th and 7th positions of the core gene, and the + 1 frameshift changed the 12 residues into LysAsnAsnPheAlaIleTrpProPheTrpLeuArg (Fig. 2A). For this mutant an additional T-to-A substitution had to be introduced to suppress a premature stop codon (TAG). Mutant G2 had the 36nt inserted between the 7th and 8th positions, with the resultant + 2 frameshift converting the amino acid sequence into IleGluGlnLeuCysHisMetAlaPheLeuAlaLys. None of the insertions (WT, G1, or G2) altered amino acid sequence of the residue upstream (Asp) or downstream (Ile) of the insertion (Fig. 2A). In mutant G3, we took advantage of the degeneracy of the genetic code to alter 14 of the 36nt without changing the coding capacity. The expression constructs were transiently transfected to Huh7 cells, with WT-36 (having the 36nt removed) serving as a control. Western blot analysis revealed that the core protein level was reduced in G2 and especially G1, but not in G3 (Fig. 2C, lanes 1–4). Densitometric analysis of Western blots from three independent experiments found that core protein was reduced by > 80% in G1 and by 30% in G2, but unaltered in G3 relative to WT (Fig. 2D). This finding indicates that the ArgThrThrLeuProTyrGlyLeuPheGlyLeuAsp sequence associated with the 36-nt insertion was required for high core protein level.

3.2. Inserting the 36nt into two other reading frames rendered the core protein less efficient to support genome replication and virion secretion

Higher core protein associated with WT and G3 constructs than G1 and G2 underscores the importance of the amino acid sequence of the 12 extra residues on the core protein level, and could explain why the 36nt is not inserted between the 6th - 7th or 7th - 8th positions. Alternatively or additionally, the LysAsnAsnPheAlaIleTrp ProPheTrpLeuArg and IleGluGlnLeuCysHisMetAlaPheLeuAlaLys sequences render the core protein incapable of supporting capsid formation, or genome replication, or virion secretion. Towards this end a core promoter mutant of genotype A with high replication capacity (clone 4B) (Parekh et al., 2003) was rendered deficient in core protein expression and hence genome replication, and the core$^-$ mutant was co-transfected with core protein construct WT, G1, G2, or G3 for trans-complementation of genome replication and virion secretion. As expected, the core$^-$ mutant transfected alone failed to show evidence of genome replication or virion secretion (Fig. 3, lane 11). More efficient genome replication and virion secretion were achieved by co-transfecting 1 µg of core$^-$ 4B dimer with 1 µg rather than 0.25 µg of core protein construct (compare lanes 1–4 with lanes 6–9). Mutant G3, with unaltered amino acid sequence relative to WT construct, supported wild-type levels of genome replication and virion secretion (Fig. 3, lanes 1, 4, 6, 9). G1 and G2 supported HBV DNA replication, while G2 also supported low level of virion secretion (lanes 2 and 3). Nevertheless, the reduction in genome replication and virion secretion for G2 (Fig. 3A & B, right panels) was more striking than reduction in core protein expression (Fig. 2C), suggesting the IleGluGlnLeuCysHisMetAlaPheLeuAlaLys sequence was less efficient at supporting genome replication and virion secretion than the naturally occurring ArgThrThrLeuProTyrGlyLeuPheGlyLeuAsp sequence.

3.3. Testing the possible contribution of a downstream stem-loop structure on core protein expression through minimal point mutations

Online prediction by Mfold software suggested that the 36-nt insertion in the core gene generates a small stem-loop structure 14nt downstream of the core gene AUG (Fig. 1C). In this regard, an RNA stem-loop structure downstream of the translation initiation site could stall the scanning ribosome to promote translation initiation, with the maximum enhancement achieved by placing the stem-loop 14nt downstream of the AUG codon (because a scanning ribosome occupies 28nt) (Kozak, 1990). According to the Mfold software the stem-loop structure was maintained in G1 and G2 but lost in G3 (Fig. 2B). The sustained core protein expression by G3 but diminished core protein expression by G1 and G2 (Fig. 2C and D) seemed to question the presence of such a stem-loop structure inside the 36-nt insertion, or challenge its stimulation of protein expression. Nevertheless, G1 and especially G2 still produced higher level of core protein than WT-36 (Fig. 2C and D), which served as a control for basal level of core protein expression.

The stem-loop structure created by the 36-nt insertion consists of ^{1918}GCCATATGGC^{1927}, which forms three G:C pairs and a 4-based loop of AUAU sequence (Fig. 1C, right panel). Using wild-type Con. 1 as the template, we introduced one to four point mutations to the guanines or cytosines forming the three base pairs (Fig. 4A). Secondary structure prediction by Mfold analysis suggested that the stem-loop structure was destroyed in G4, G8 and G9, while a larger stem-loop with a longer stem (containing internal unpaired sequence) and a 11-nt loop was created in G5, G6, and G7 (Fig. 4B). To guard against differences in transfection efficiency, the core protein construct was co-transfected with an expression construct for envelope proteins. Comparable levels of intracellular small (S) envelope protein (Fig. 4C), intracellular and extracellular HBsAg (Fig. 4E), were detected suggesting no marked difference in transfection efficiency. Western blot analysis revealed reduced core protein for all the mutants, especially G8, even though that mutant still produced higher level of core protein than WT-36 (Fig. 4C and D).

3.4. Three pairs of site-directed mutants to destroy or restore the stem-loop structure confirmed its enhancement of core protein expression

Since the stem-loop structures formed by G5, G6, and G7 were different from that of the wild-type construct, and considering that mutants G4-G9 were accompanied by different amino acid substitutions in the core protein (Fig. 4A, right panel), which according to the phenotypes of mutant G1 and G2 could affect core protein level, studies described above fell short of unequivocally establishing the role of the stem-loop structure on core protein expression. We therefore chose constructs G6 and G8 for further mutagenesis. G6 was one of the three mutants creating an extended stem-loop, while G8 could no longer form the stem-loop structure and showed lowest core protein expression among the 6 mutants (Fig. 4D). A single G1934C mutation introduced to G6 destroyed the extended stem-loop (mutant G12; Fig. 5A and B). In contrast, the A1918C/T1927G double mutation introduced to G8 restored wild-type stem-loop structure (G13). In addition, through C1920G mutation we constructed G10 with an unstable (one G-C pair and one A-U pair) and shifted stem-loop (7nt downstream of AUG codon). An additional G1925C mutation restored the wild-type stem-loop (G11; Fig. 5A and B). These three pairs of constructs were transfected to Huh7 cells together with the common envelope protein construct. Western blot analysis revealed that G12 produced much less core protein than G6, the parental construct (Fig. 5C, lanes 6 and 5). In contrast, G13 showed marked increase in core protein expression relative to G8 (Fig. 5C, lanes 11 and 10). Core protein expression was moderately reduced in G10 relative to WT but restored in G11 (Fig. 5C, lanes 3, 2, 4). Simultaneous detection of intracellular S protein as well as intracellular and extracellular HBsAg (Fig. 5C and E) failed to attribute the difference in core protein expression to different transfection efficiencies. Therefore, the stem-loop structure partly accounted for enhanced core protein expression by the 36-nt insertion.

3.5. Impact of the 36nt on core protein expression depended on its position relative to the translation initiation site

We tested the positional effect of the stem-loop on core protein expression by moving the translation initiation sites upstream or further downstream of the 36nt. In construct G15 (based on Con. 2), a new optimal translation initiation site (CCATGG) was created 27nt upstream of the original ATG codon by two point mutations (A1875T/C1877G) combined with insertion of two cytosines (Fig. 6A). Consequently the stem-loop created by the 36-nt insertion became 41nt rather than 14nt downstream of the translation initiation site (Fig. 6B). Together with an A1896G back mutation to abolish the nonsense mutation in the precore region, construct G15 would produce a precore/core protein of 204 (9 + 195) residues. Transfection experiments revealed that G15 produced only about 40% of core protein of wild-type Con. 2 (Fig. 6C and D, lanes 5 & 7). Deleting the 36-nt from G15 (G15–36) reduced its size from 204 residues to 192 residues, but did not further diminish core

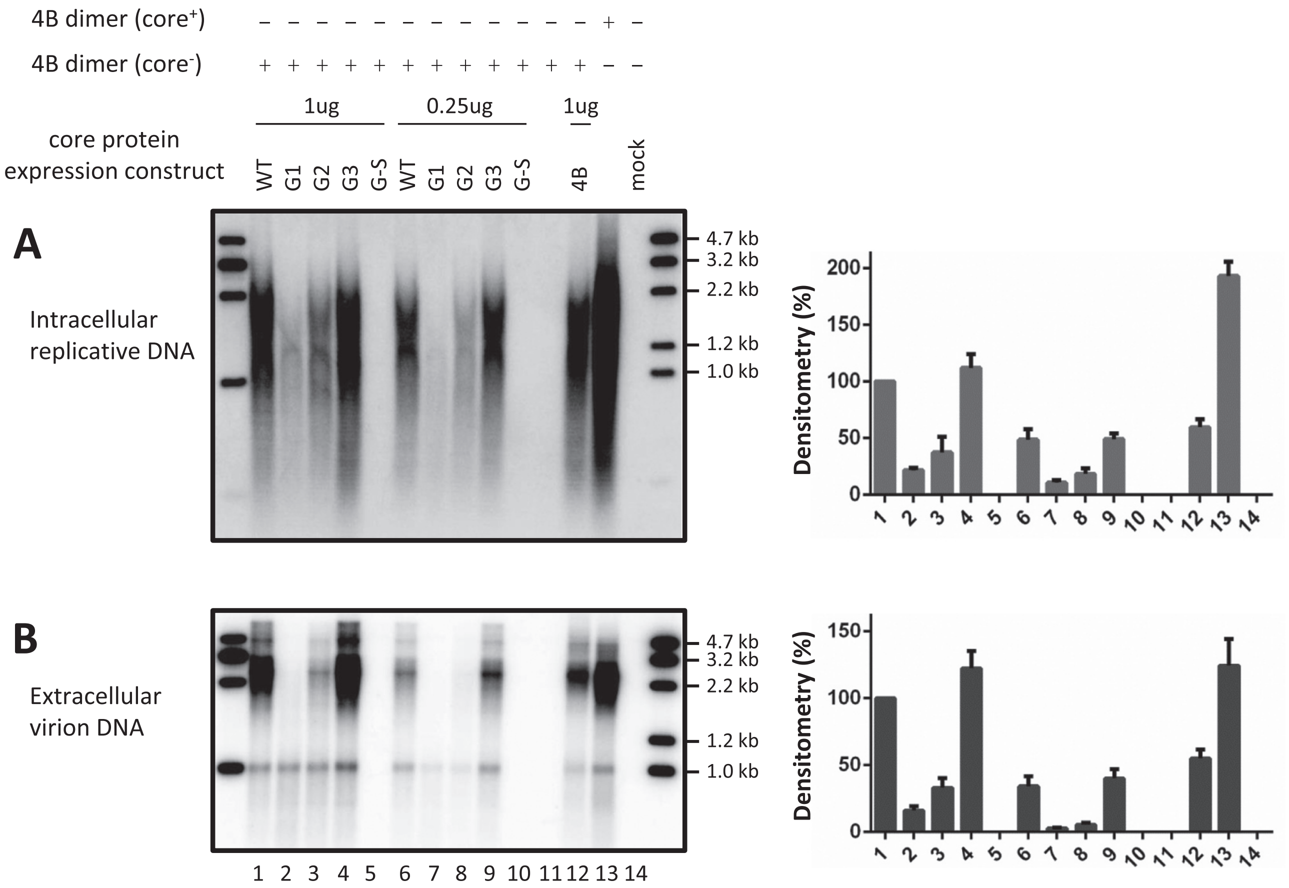


Fig. 3. Ability of various core protein constructs to support genome replication and virion secretion through trans-complementation. Huh7 cells were co-transfected with 1 μg of the core^{-} 4B dimer and 1 μg or 0.25 μg of core protein expression construct. Core^{-} 4B dimer transfected alone served as a negative control, while the wild-type 4B dimer (1 μg) transfected alone served as the positive control. Intracellular replication DNA **(A)** and extracellular virion DNA **(B)** were subjected to Southern blot analysis. Cloned HBV DNA cut with different restriction enzymes was pooled to serve as size markers. Construct G-S has the 36nt from genotype G replaced with a similar sequence from Streptomyces (CCGCACCTCACTGCCGTACGGCCTGTTCGGCCTCTC) encoding amino acid sequence ArgThrSerLeuProTyrGlyLeuPheGlyLeuSer. These 36nt are predicted to form a stem-loop structure similar to the 36nt found in genotype G but expressed negligible amount of core protein (data not shown). Southern blots from one transfection experiment are shown in the left. The right panels are densitometric analysis of signals from Southern blots derived from three independent transfection experiments, with value for 1υg of the WT construct arbitrarily set at 100%.

protein level (Fig. 6C and D, lanes 5 and 6). Thus, a stem-loop structure placed 41nt downstream of AUG codon apparently no longer affected core protein expression.

Through an A1901T mutation to destroy the original ATG codon and T1945G mutation to create a new ATG codon 42nt downstream, mutant G14 (based on Con. 3) would produce a shortened core protein of 181 residues (Fig. 6A). The stem-loop structure created by the 36-nt insertion became 5nt upstream of the translation initiation site (Fig. 6B). Western blot analysis showed that G14 produced merely 20% of core protein of wild-type Con. 3 (Fig. 6C and D, lanes 1 and 3). Deleting the 36nt from G14 (G14–36) did not alter protein size but increased core protein expression (lanes 1 and 2). Together, these results are consistent with the presence of a stem-loop structure, which when placed upstream of translation initiation site moderately reduced core protein expression.

4. Discussion

HBV genotype G has unique clinical features including its target population (homosexual men), co-infection with HIV, co-infection with and gradual takeover of another HBV genotype. It harbors A1762T/G1764A core promoter mutations as well as G1896A precore mutation, which for other HBV genotypes do not arise until the immune clearance phase of infection. Its unique genomic features include an extra precore nonsense mutation (C1817T converting codon 2 to TAA) and a 36-nt insertion at the 5′ end of the core gene. It is very likely that the additional C1817T nonsense mutation (which increases the distance from translational termination site to the reinitiation site from 1 to 27 codons) combined with core promoter mutations (which diminish pcRNA transcription) minimize core protein expression from pcRNA. On the other hand, our previous studies found that the 36-nt insertion greatly increased core protein expression, whether in genotype G or when artificially introduced to other HBV genotypes (Gutelius et al., 2011; Li et al., 2007). Therefore, besides the core promoter mutations and some precore mutations (such as mutated precore AUG and G1896A nonsense mutation) the 36-nt insertion in the core gene represents a novel type of mutations capable of up regulating core protein expression. Since the effect of the 36-nt insertion was reproducible in dimeric HBV construct, 1.1x construct with pgRNA driven by an exogenous promoter, and CMV- core gene construct with most of the HBV genomic sequence deleted (Li et al., 2007), ability of the 36nt to increase core protein level is apparently independent of the core promoter as well as other three genes and their products. Indeed, primer extension assay revealed that the insertion did not increase mRNA abundance, for any of the three types of constructs mentioned above (Li et al., 2007). Thus, an effect of the 36-nt insertion should be most likely mediated at the translational or post-translational level. In this regard the 36-nt insertion is predicted to create a small stem-loop structure composed of three G:C pairs and a loop of AUAU sequence, 14nt downstream of translational initiation site of the core gene (Fig. 1C).

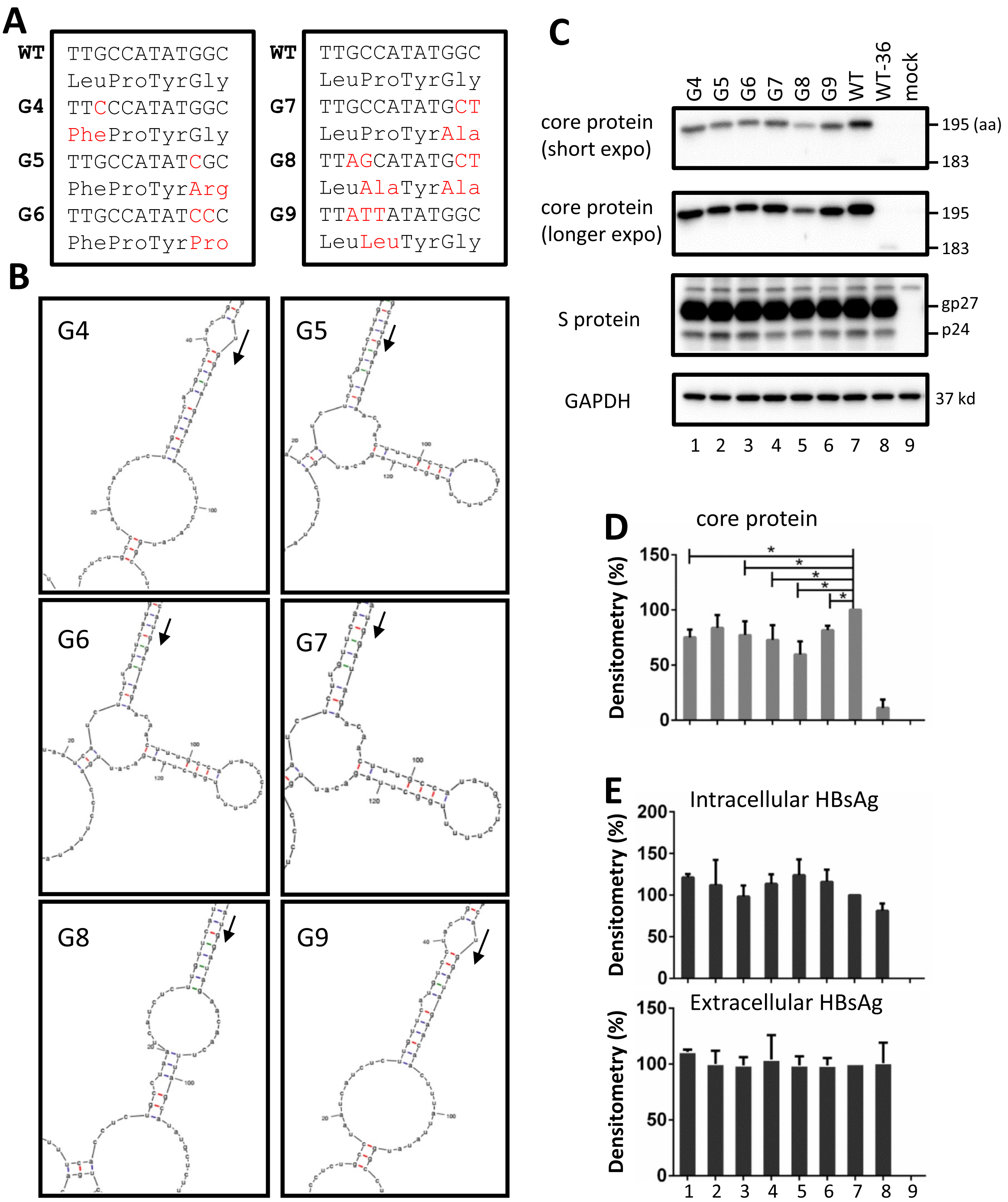


Fig. 4. Impact of destroying the original RNA stem-loop structure associated with the 36-nt insertion on core protein expression. Point mutations were introduced to position 13–15 and 20–22 of the 36-nt insertion to disrupt the original stem-loop structure, using WT Con. 1 construct as the template. **(A)** Mutations of constructs G4-G9 at the nucleotide level and amino acid level. Only positions 11–22 of the 36nt and 4–7 of the 12aa are shown. The red font indicates the changes. **(B)** Secondary structures of these mutants according to Mfold, with the translation initiation site indicated by an arrowhead. Panels C-E show results of functional characterization following transient transfection experiments. Huh7 cells in 6-well plates were transfected with 1.6 µg of individual core protein construct together with 0.4 µg of envelope protein expression construct. Cells and culture supernatant were harvested at day 3 post-transfection. **(C)** Western blot analysis of intracellular core and S proteins, with GAPDH serving as a loading control. **(D)** Core protein quantification following densitometric analysis of Western blots from three independent transfection experiments. The value for wild-type construct was set arbitrarily at 100%. $^*P < 0.05$. **(E)** HBsAg averaged from the three independent transfection experiments (measured after 1:4000 dilution of cell lysate and 1:500 dilution of culture supernatant), with values for wild-type construct set at 100%.

A large number of site-directed mutants were generated to destroy or restore the stem-loop structure, and the impact on core protein expression was determined in transiently transfected Huh7 cells. In additional mutants, the translation initiation site was moved upstream or downstream to check for the positional impact of this RNA secondary structure. Three mutants with the stem-loop eliminated by 1, 3, or 4 point mutations (G4, G9, G8) displayed diminished core protein expression than WT, with lowest core protein level found in G8 (Fig. 4C and D). In this regard G8 is structurally most similar to WT except for the loss of the stem-loop structure (compare Fig. 1C with Fig. 4B). For the wild-type virus the translation initiation site is located in a stem structure, with the stem downstream of the AUG codon maintained in

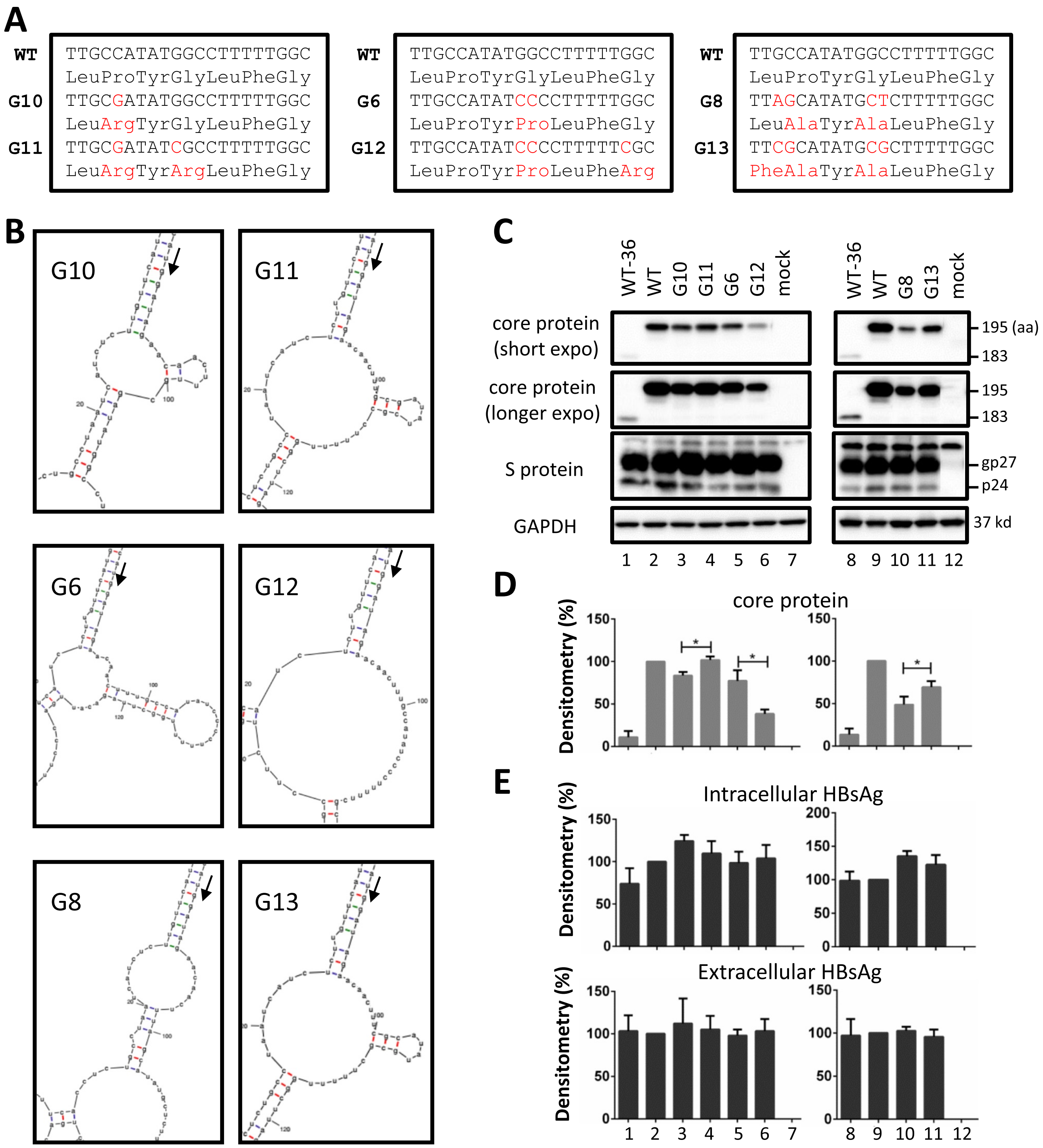


Fig. 5. Comparison of paired constructs with or without a strong stem-loop structure downstream of translation initiation site on core protein expression. Three pairs of constructs were compared: G6 vs. G12; G8 vs. G13; and G10 vs. G11. All were based on wild-type Con. 1 construct. **(A)** Mutations at the nucleotide and amino acid levels. The red color indicates altered nucleotides and amino acids. **(B)** Predicted RNA secondary structures. Huh7 cells were co-transfected with the WT or mutant construct and envelope protein expression construct at 1.6 µg/0.4 µg ratio. Cells and culture supernatant were harvested at day 3 post-transfection. **(C)** Intracellular core and S proteins were detected by Western blot. **(D)** Densitometric values of core protein from three transfection experiments were determined, with the value for wild-type construct set at 100%. $^{*}P < 0.05$. **(E)** ELISA measurement of HBsAg in the cell lysate (after 1:4000 dilution) and culture supernatant (after 1:500 dilution). Values were averaged from three independent experiments.

G8 but extended in G4 and G9. The downstream stem probably impedes the scanning ribosome to increase translation initiation at the AUG, which might explain the higher core protein translation by G4 and G9 than G8. As shown in Fig. 5, changing two nucleotides in G8 to restore the original stem-loop structure markedly enhanced core protein expression (G15; but not to the level of WT), whereas destroying a new type of stem-loop structure created in G6 by a single point mutation greatly diminished core protein level (G12; but still higher than WT-36). Finally, comparison of the phenotypes of G14 with WT and G14–36, as well as G15 with WT and G15–36, validated presence of a stem-loop structure which no longer promoted core protein expression when placed 41nt downstream of translation initiation site, and rather suppressed core protein expression when located 5nt upstream of the AUG codon (Fig. 6). The inhibitory effect of an upstream stem-loop

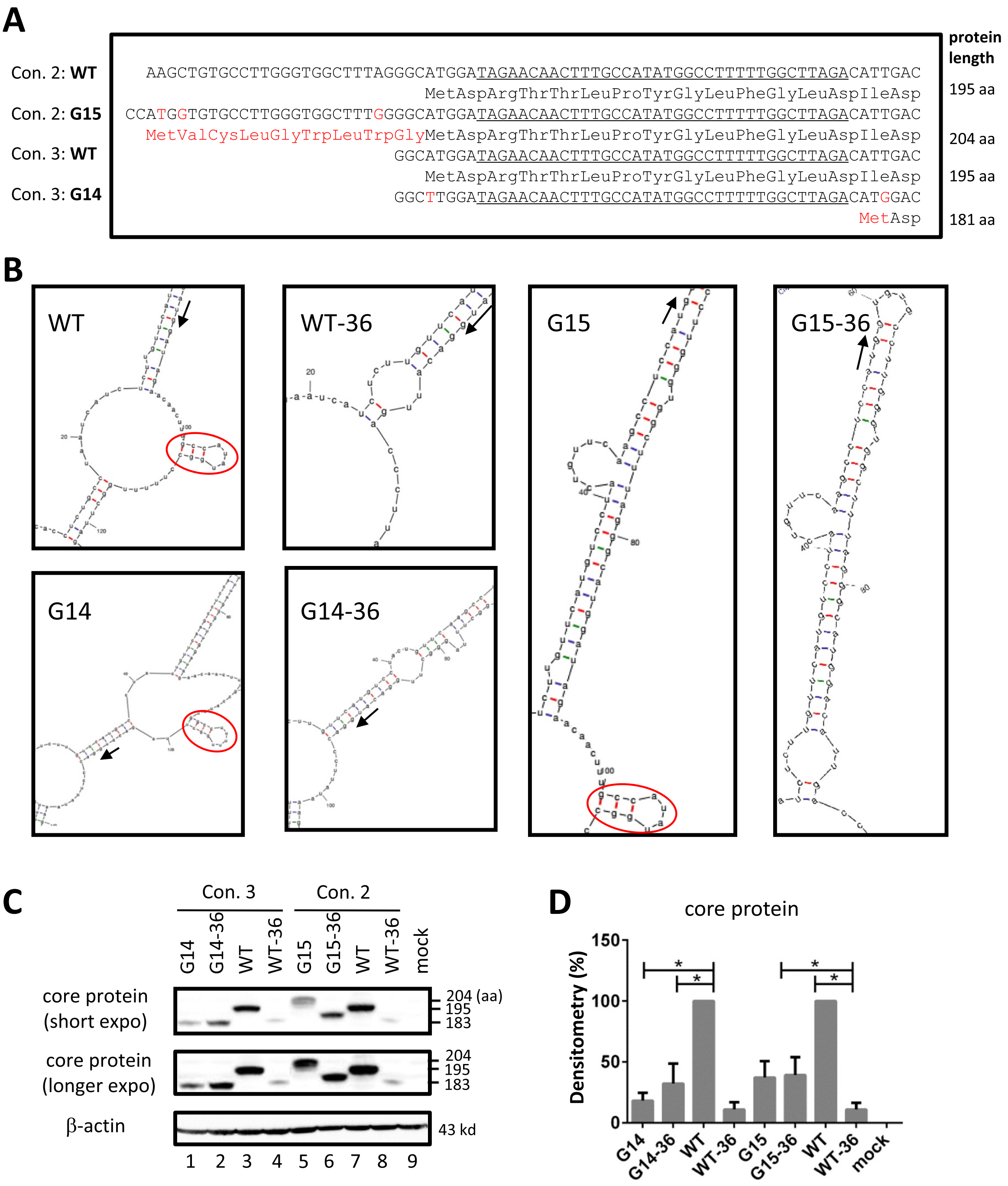


Fig. 6. Impact of shifting the translation initiation site relative to the RNA stem-loop structure on core protein expression. **(A)** Schematic representation of the mutant constructs. The red font indicates altered nucleotides and amino acids. G14 was derived from WT construct Con. 3 by mutating the core gene ATG into TTG and the ATT codon downstream of the 36-nt insertion into ATG, thereby moving down the translation initiation site by 14 codons and shortening the core protein to 181 residues. Consequently the stem-loop structure associated with the 36-nt insertion became upstream of the translation initiation site. G15 was derived from WT Con. 2 construct by converting an upstream AAG codon into ATG, thus extending the core protein at the amino terminus by 9 residues (to 204aa) and rendering the stem-loop structure 41nt downstream of the translation initiation site. G14–36 and G15–36 had the 36nt removed from G14 and G15, respectively, and would respectively express a core protein of 181aa and 192aa. **(B)** Predicted RNA secondary structures of these constructs. The arrowheads indicate the translation initiation site, and the stem-loop structure is circled. **(C)** Western blot analysis of core protein from transiently transfected Huh7 cells. β-actin served as a loading control. **(D)** Densitometric values obtained from Western blots from three independent transfection experiments, with values of WT Con. 3 (relative to G14) and WT Con. 2 (relative to G15) set at 100%.

structure on translation efficiency is well documented (Kozak, 1986). Taken together, these results are mostly compatible with the presence of an RNA secondary structure which enhanced core protein expression when placed 14nt downstream of translation initiation site. Nevertheless, the stem-loop structure alone cannot explain the huge difference in core protein level between WT and WT-36.

Still, we cannot exclude the possibility that some of the mutations introduced affected RNA production or stability to alter the level of core protein expression. In addition, for the 10 mutants aimed at destroying or restoring the stem-loop structure (G4-G13), 1–3amino acid changes were introduced as well (Figs. 4A & 5A). Such sequence changes may alter protein stability to affect the steady-state level of the core protein in transfected cells, thus complicating data interpretation. One way to study the impact of RNA secondary structure on the efficiency of core protein translation without complication of protein stability is coupled in vitro transcription/translation system using rabbit reticulocyte lysate. However these mutants were all made with Con. 1 construct, in which the vector sequence downstream of the SacI site, including the promoter for T7 RNA polymerase, has been removed to avoid non-HBV sequence at the 5′ end of mRNA transcribed by the CMV promoter.

To examine whether amino acid sequence of the 12 inserted residues contribute to the high core protein expression associated with the 36-nt insertion, we generated G1, G2, and G3 mutants. G1 and G2 differed from wild-type construct in that the site of insertion was shifted one or two positions down, thus causing a + 1 or + 2 frameshift mutation (Fig. 2A). Consequently, the 12aa inserted were completely different yet the stem-loop structure was maintained. In contrast, G3 had 14 of the 36nt mutated to alter RNA secondary structure without changing amino acid sequence. The fact that the high core protein level was mitigated in G2 and especially in G1 suggested that the naturally occurring ArgThrThrLeuProTyrGlyLeuPheGlyLeuAsp sequence is essential for the high core protein level. We suspect that the IleGluGlnLeuCysHisMetAlaPheLeuAlaLys sequence found in G2 and especially LysAsnAsnPheAlaIleTrpProPheTrpLeuArg sequence found in G1 accelerate core protein degradation. This can be verified in future study using metabolic labeling (pulse-chase) experiments to measure the protein half lives of the mutant core proteins. The sustained core protein level associated with mutant G3 seemed to rule out the presence of an RNA secondary structure or question its stimulation of core protein translation. However, 11 of the 14 point mutations introduced to G3 lie outside the predicted stem-loop structure, which may complicate data interpretation. There is also the possibility that the synonymous mutations increased translation efficiency by switching to more abundant tRNAs for amino acid attachment to the elongating chain of polypeptide. In summary, the marked up regulation of core protein level by the 36-nt insertion of TAGAACAACTTTGCCATATGGCCTTTTTGGCTTAGA requires its translation into ArgThrThrLeuProTyrGlyLeuPheGlyLeuAsp sequence, and is at least partly mediated by a stem-loop structure which most likely enhances translation initiation.

pgRNA serves as the mRNA for P protein as well, through ribosomal leaky scanning past the core gene AUG codon or through ribosomal shunting (Cao and Tavis, 2011). Increased translation initiation from the core gene AUG will diminish ribosomal scanning further down towards the P gene AUG, thus reducing P protein expression. We previously found that the 36-nt insertion was essential for efficient genome replication of genotype G, whereas its introduction into HBV genotypes A and D rather impaired genome replication despite marked increase in core protein expression (Gutelius et al., 2011; Li et al., 2007). A core particle is assembled from 240 copies of core protein and packages just one molecule of P protein. Apparently in G genotype but not other HBV genotypes, the core/P protein ratio from pgRNA is too low for genome replication without the 36-nt insertion. Although from the CMV-core gene construct the 36-nt insertion markedly increased core protein expression, we previously found that tandem dimer of the genotype G clone transcribed much less 3.5-kb RNA than genotype A clones when transiently transfected to Huh7 cells (Li et al., 2007). Thus, the insertion can be considered as a mechanism to rescue core protein expression and genome replication for genotype G despite low pgRNA transcript level. Moreover, genotype G patients are often immune suppressed due to HIV co-infection, which tolerates the high expression of core protein, otherwise a strong stimulant of immune clearance mechanisms. Together, these factors could explain why the 36-nt insertion is found in genotype G but not in other HBV genotypes.

Acknowledgements

This work was supported by NIH grants AI116639, AI107618, and also by grants 31370195, 81672017, 81672064, and ZX10202203 from National Science Foundation of China.

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.virol.2018.11.002

References

Araujo, N.M., Araujo, O.C., Silva, E.M., Villela-Nogueira, C.A., Nabuco, L.C., Parana, R., Bessone, F., Gomes, S.A., Trepo, C., Kay, A., 2012. Identification of novel recombinants of hepatitis B virus genotypes F and G in human immunodeficiency virus-positive patients from Argentina and Brazil. J. Gen. Virol. 94, 150–158.

Baumert, T.F., Rogers, S.A., Hasegawa, K., Liang, T.J., 1996. Two core promotor mutations identified in a hepatitis B virus strain associated with fulminant hepatitis result in enhanced viral replication. J. Clin. Investig. 98, 2268–2276.

Bhat, R.A., Ulrich, P.P., Vyas, G.N., 1990. Molecular characterization of a new variant of hepatitis B virus in a persistently infected homosexual man. Hepatology 11, 271–276.

Bock, C.T., Schwinn, S., Locarnini, S., Fyfe, J., Manns, M.P., Trautwein, C., Zentgraf, H., 2001. Structural organization of the hepatitis B virus minichromosome. J. Mol. Biol. 307, 183–196.

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

Buckwold, V.E., Xu, Z., Chen, M., Yen, T.S., Ou, J.H., 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, 5845–5851.

Cao, F., Tavis, J.E., 2011. RNA elements directing translation of hepatitis B virus polymerase via ribosomal shunting. J. Virol. 85, 6343–6352.

Carman, W.F., Jacyna, M.R., Hadziyannis, S., Karayiannis, P., McGarvey, M.J., Makris, A., Thomas, H.C., 1989. Mutation preventing formation of hepatitis B e antigen in patients with chronic hepatitis B infection. Lancet 2, 588–591.

Chong, C.K., Cheng, C.Y.S., Tsoi, S.Y.J., Huang, F., Liu, F., Seto, W., Lai, C., Yuen, M., Wong, D.K., 2017. Role of hepatitis B core protein in HBV transcription and recruitment of histone acetyltransferases to cccDNA minichromosome. Antivir. Res. 144, 1–7.

Friedt, M., Gerner, P., Lausch, E., Trubel, H., Zabel, B., Wirth, S., 1999. Mutations in the basic core promotor and the precore region of hepatitis B virus and their selection in children with fulminant and chronic hepatitis B. Hepatology 29, 1252–1258.

Garcia, T., Li, J., Sureau, C., Ito, K., Qin, Y., Wands, J., Tong, S., 2009. Drastic reduction in the production of subviral particles eoes not impair hepatitis B virus virion secretion. J. Virol. 83, 11152–11165.

Guo, Y.H., Li, Y.N., Zhao, J.R., Zhang, J., Yan, Z., 2011. HBc binds to the CpG islands of HBV cccDNA and promotes an epigenetic permissive state. Epigenetics 6, 720–726.

Gutelius, D., Li, J., Wands, J., Tong, S., 2011. Characterization of the pleiotropic effects of the genotype G-specific 36-nucleotide insertion in the context of other hepatitis B virus genotypes. J. Virol. 85, 13278–13289.

Hasegawa, K., Huang, J., Rogers, S.A., Blum, H.E., Liang, T.J., 1994. Enhanced replication of a hepatitis B virus mutant associated with an epidemic of fulminant hepatitis. J. Virol. 68, 1651–1659.

Jia, H., Qin, Y., Chen, C., Zhang, F., Li, C., Zong, L., Wang, Y., Zhang, J., Li, J., Wen, Y., Tong, S., 2017. The envelope gene of hepatitis B virus is implicated in both differential virion secretion and genome replication capacities between genotype B and genotype C isolates. Viruses 9, 62.

Kato, H., Orito, E., Gish, R.G., Sugauchi, F., Suzuki, S., Ueda, R., Miyakawa, Y., Mizokami, M., 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, 6131–6137.

Kozak, M., 1986. Influences of mRNA secondary structure on initiation by eukaryotic ribosomes. Proc. Natl. Acad. Sci. USA 83, 2850–2854.

Kozak, M., 1990. Downstream secondary structure facilitates recognition of initiator codons by eukaryotic ribosomes. Proc. Natl. Acad. Sci. USA 87, 8301–8305.

Kramvis, A., 2014. Genotypes and genetic variability of hepatitis B virus. Intervirology 57, 141–150.

Li, K., Zoulim, F., Pichoud, C., Kwei, K., Villet, S., Wands, J., Li, J., Tong, S., 2007. Critical role of the 36-nucleotide insertion in hepatitis B virus genotype G in core protein expression, genome replication, and virion secretion. J. Virol. 81, 9202–9215.

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

Liaw, Y.F., Chu, C.M., 2009. Hepatitis B virus infection. Lancet 373, 582–592.

Milich, D., 2003. Exploring the biological basis of hepatitis B e antigen in hepatitis B virus infection. Hepatology 38, 1075–1086.

Moriyama, K., Okamoto, H., Tsuda, F., Mayumi, M., 1996. Reduced precore transcription and enhanced core-pregenome transcription of hepatitis B virus DNA after replacement of the precore-core promoter with sequences associated with e antigen-seronegative persistent infections. Virology 226, 269–280.

Norder, H., Couroucé, A., Coursaget, P., Echevarria, J.M., Lee, S., Mushahwar, I.K., Robertson, B.H., Locarnini, S., Magnius, L.O., 2004. Genetic diversity of hepatitis B virus strains derived worldwide: genotypes, subgenotypes, and HB_sAg subtypes. Intervirology 47, 289–309.

Okamoto, H., Tsuda, F., Akahane, Y., Sugai, Y., Yoshiba, M., Moriyama, K., Tanaka, T., Miyakawa, Y., Mayumi, M., 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, 8102–8110.

Okamoto, H., Tsuda, F., Sakugawa, H., Sastrosoewignjo, R.I., Imai, M., Miyakawa, Y., Mayumi, M., 1988. Typing hepatitis B virus by homology in nucleotide sequence: comparison of surface antigen subtypes. J. Gen. Virol. 69, 2575–2583.

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

Omata, M., Ehata, T., Yokosuka, O., Hosoda, K., Ohto, M., 1991. Mutations in the precore region of hepatitis B virus DNA in patients with fulminant and severe hepatitis. N. Engl. J. Med. 324, 1699–1704.

Osiowy, C., Gordon, D., Borlang, J., Giles, E., Villeneuve, J.P., 2008. Hepatitis B virus genotype G epidemiology and co-infection with genotype A in Canada. J. Gen. Virol. 89, 3009–3015.

Parekh, S., Zoulim, F., Ahn, S.H., Tsai, A., Li, J., Kawai, S., Khan, N., Trepo, C., Wands, J., Tong, S.P., 2003. Genome replication, virion secretion, and e antigen expression of naturally occurring hepatitis B virus core promoter mutants. J. Virol. 77, 6601–6612.

Qin, Y., Tang, X., Garcia, T., Hussain, M., Zhang, J., Lok, A., Wands, J., Li, J., Tong, S., 2011. Hepatitis B virus genotype C isolates with wild-type core promoter sequence replicate less efficiently than genotype B isolates but possess higher virion secretion capacity. J. Virol. 85, 10167–10177.

Sánchez, L.V., Tanaka, Y., Maldonado, M., Mizokami, M., Panduro, A., 2006. Difference of hepatitis B virus genotype distribution in two groups of Mexican patients with different risk factors. Intervirology 50, 9–15.

Scaglioni, P.P., Melegari, M., Wands, J.R., 1997. Biologic properties of hepatitis B viral genomes with mutations in the precore promoter and precore open reading frame. Virology 233, 374–381.

Stanaway, J.D., Flaxman, A.D., Naghavi, M., Fitzmaurice, C., Vos, T., Abubakar, I., Abu-Raddad, L.J., Assadi, R., Bhala, N., Cowie, B., Forouzanfour, M.H., Groeger, J., Hanafiah, K.M., Jacobsen, K.H., James, S.L., MacLachlan, J., Malekzadeh, R., Martin, N.K., Mokdad, A.A., Mokdad, A.H., Murray, C.J.L., Plass, D., Rana, S., Rein, D.B., Richardus, J.H., Sanabria, J., Saylan, M., Shahraz, S., So, S., Vlassov, V.V., Weiderpass, E., Wiersma, S.T., Younis, M., Yu, C., El Sayed Zaki, M., Cooke, G.S., 2016. The global burden of viral hepatitis from 1990 to 2013: findings from the Global Burden of disease study 2013. Lancet 388, 1081–1088.

Stuyver, L., De Gendt, S., Van Geyt, C., Zoulim, F., Fried, M., Schinazi, R.F., Rossau, R., 2000. A new genotype of hepatitis B virus: complete genome and phylogenetic relatedness. J. Gen. Virol. 81, 67–74.

Tian, Y., Kuo, C.F., Akbari, O., Ou, J.H., 2016. Maternal-derived hepatitis B virus e antigen alters macrophage function in offspring to drive viral persistence after vertical transmission. Immunity 44, 1204–1214.

Tong, S.P., Li, J.S., Vitvitski, L., Trepo, C., 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.

Tong, S., Revill, P., 2016. Overview of hepatitis B viral replication and genetic variability. J. Hepatol. 64, S4–S16.

Trepo, C., Chan, H.L., Lok, A., 2014. Hepatitis B virus infection. Lancet 384, 2053–2063.

Tripathy, A.S., Das, R., Chadha, M.S., Arankalle, V.A., 2011. Epidemic of Hepatitis B with high mortality in India: association of fulminant disease with lack of CCL4 and natural killer T cells. J. Viral Hepat. 18, e415–e422.

Tsai, A., Kawai, S., Kwei, K., Gewaily, D., Hutter, A., Tong, D.R., Li, J., Wands, J.R., Tong, S., 2009. Chimeric constructs between two hepatitis B virus genomes confirm transcriptional impact of core promoter mutations and reveal multiple effects of core gene mutations. Virology 387, 364–372.

van der Kuyl, A.C., Zorgdrager, F., Hogema, B., Bakker, M., Jurriaans, S., Back, N.K., Berkhout, B., Zaaijer, H.L., Cornelissen, M., 2013. High prevalence of hepatitis B virus dual infection with genotypes A and G in HIV-1 infected men in Amsterdam, the Netherlands, during 2000–2011. BMC Infect. Dis. 13, 540.

Yu, Y., Wan, P., Cao, Y., Zhang, W., Chen, J., Tan, L., Wang, Y., Sun, Z., Zhang, Q., Wan, Y., Zhu, Y., Liu, F., Wu, K., Liu, Y., Wu, J., 2017. Hepatitis B virus E antigen activates the suppressor of cytokine signaling 2 to repress interferon action. Sci. Rep. 7, 1729.