Virology 387 (2009) 364–372

Contents lists available at ScienceDirect

Virology

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

Chimeric constructs between two hepatitis B virus genomes confirm transcriptional impact of core promoter mutations and reveal multiple effects of core gene mutations

Adrienne Tsai [1,2], Shigenobu Kawai [1,3], Karen Kwei [4], Dina Gewaily, Alexander Hutter, David R. Tong, Jisu Li, Jack R. Wands, Shuping Tong *

Liver Research Center, Rhode Island Hospital, The Warren Alpert School of Medicine, Brown University, Providence, RI 02906, USA

ARTICLE INFO

Article history:
Received 27 November 2008
Returned to author for revision
17 December 2008
Accepted 3 March 2009
Available online 26 March 2009

Keywords:
Core gene
Core promoter
Core protein
DNA replication
Hepatitis B virus
Immature secretion
Naturally occurring mutations
Pregenomic RNA
Trans-complementation

ABSTRACT

Hepatitis B virus (HBV) clone 4B replicated much more efficiently than clone 2A of the same genotype. Introduction of its T1753C, A1762T, G1764A, and C1766T core promoter mutations into the 2A genome greatly enhanced genome replication and suppressed HBeAg expression. Here we show that these effects are mediated by transcriptional up regulation of pregenomic RNA and suppression of precore RNA. Analysis of chimeric constructs suggested that the 5′ end of the 2A core gene conferred higher level of pregenomic RNA, but less core protein and genome replication relative to the 4B sequence. Genome maturity of secreted virions was reduced by mutations present in the core protein of the 2A genome but enhanced by mutations found in the 4B core protein. The 4B core protein migrated faster than that of clone 2A. The possible links among the various phenotypes and the responsible mutations remain to be established.

Introduction

The hepatitis B virus (HBV) has a small DNA genome of 3.2 kb, which contains four open reading frames termed X, P, core, and envelope (Seeger et al., 2007). The P gene overlaps partially with the X and core genes and completely with the envelope gene. It specifies DNA polymerase (P protein), the enzyme implicated in viral genome replication. The core gene encodes core protein, which forms capsids shielding viral genomic RNA or DNA. In addition, it forms part of the coding sequence for hepatitis B e antigen (HBeAg), a secreted protein involved in modulation of host immune response. From the envelope gene three co-terminal envelope proteins: large (L), middle (M), and small (S) are generated through alternative translation initiation sites. The envelope proteins participate in virion formation. Translation of the seven HBV proteins is made possible by the generation of five transcripts with different 5′ ends but an identical 3′ end: the 3.5-kb precore RNA (pc RNA) for HBeAg, the slightly shorter 3.5-kb pregenomic RNA (pg RNA) for core and P proteins, the 2.4-kb subgenomic RNA for L protein, the 2.1-kb subgenomic RNA with heterogeneous 5′ ends for M and S proteins, and the 0.7-kb subgenomic RNA for X protein. With the exception of P gene expression from the pg RNA, the gene to be translated is located near the 5′ end of the transcript due to the low efficiency with which the downstream genes are translated. That explains why the X protein is not expressed from the 3.5-, 2.4-, and 2.1-kb transcripts despite the presence of X coding sequence in all these transcripts.

The P gene translation initiation site is located about 500 nucleotides downstream of the 5′ end of the pg RNA. It sits downstream of the core gene initiation codon but about 150 nucleotides upstream of the core gene termination codon. Therefore, P protein translation is achieved by ribosomal leaky scanning of the core gene initiator and several internal AUGs (Fouillot et al., 1993; Hwang and Su, 1998). This low rate of P protein expression is compatible with the need for just one molecule of P protein per capsid, which is assembled from 180 or 240 copies of core protein. As expected, experiments in the related duck

* Corresponding author. Liver Research Center, 55 Claverick Street, 4th Floor, Providence, RI 02906, USA. Fax: +1 401 4442939.
E-mail address: Shuping_Tong_MD@Brown.Edu (S. Tong).
[1] These authors contributed equally to this work.
[2] Current address: Department of Anesthesia, Beth Israel Deaconess Hospital, Harvard Medical School, Boston, MA, USA.
[3] Current address: Department of Medicine and Clinical Oncology, Graduate School of Medicine, Chiba University, Chiba, Japan.
[4] Current address: Boston University School of Medicine, Boston, MA, USA.

0042-6822/$ – see front matter © 2009 Elsevier Inc. All rights reserved.
doi:10.1016/j.virol.2009.03.002

hepatitis B virus (DHBV) revealed that increasing the efficiency of core protein translation reduced P protein expression, whereas ablating core protein expression increased P protein translation (Sen et al., 2004).

HBV genome replication is driven solely by the pg RNA, which serves not only as the messenger for both core and P proteins, but also as the genome precursor. The 5′ end of this 3.5-kb mRNA, which is missing in the shorter transcripts, forms a stem–loop structure that functions as the encapsidation signal. It is believed that one molecule each of pg RNA and P protein is packaged into the capsid, where the P protein synthesizes the minus strand DNA from the pg RNA template by reverse transcription, using a domain of the P protein as the protein primer. The P protein subsequently degrades pg RNA via its RNase H domain, and synthesizes plus stranded DNA through the template of minus strand DNA. The capsid is subsequently released from the cell by budding and in doing so acquires its envelope and the three envelope proteins. Virion-associated HBV genome contains full-length minus strand DNA, and a plus strand DNA of variable lengths (from 50 to 100%). In contrast, both double stranded DNA and DNA: RNA hybrid (single stranded DNA) can be detected inside intracellular capsids. This difference in genome maturity suggests the presence of a "maturation signal" for selective or preferential envelopment of capsids containing double stranded DNA genome. In this regard, virion-associated DHBV core protein is hypophosphorylated relative to the intracellular core particles (Pugh et al., 1989). However, mutating the three phosphorylation sites of HBV core protein affected the efficiencies of pg RNA encapsidation and genome replication (Gazina et al., 2000; Lan et al., 1999; Melegari et al., 2005), not necessarily the stringency of virion secretion. A more recent study in the DHBV model suggests that alanine substitutions (to mimic unphosphorylated form) blocked reverse transcription, while aspartate substitutions (to mimic phosphorylated form) blocked plus strand DNA synthesis (Basagoudanavar et al., 2007). On the other hand, a naturally occurring I97L mutation in the core protein caused an "immature secretion phenotype" of predominantly single stranded DNA (Yuan et al., 1999). We found that the extra 12-amino acid residues unique to the core protein of genotype G reduced virion secretion efficiency but increased genome maturity of secreted virions (Li et al., 2007).

Mutations that alter HBV replication capacity may have a fundamental impact on its biology and pathogenesis. In this regard, both the pg RNA, the only transcript required for genome replication, and the pc RNA (the HBeAg messenger) are under the control of the core promoter. They are products of imprecise transcription initiation, although distinct sequence elements within the core promoter can differentially regulate the transcription of pc RNA vs. pg RNA (Yu and Mertz, 1996). The core promoter can be further divided into the core upstream regulatory sequence and basal core promoter (Honigwachs et al., 1989; Yuh et al., 1992), where mutations frequently arise during the later stage of chronic HBV infection (Kidd-Ljunggren e al., 1997; McMillan et al., 1996; Okamoto et al., 1994). The hot-spot double mutation in the core promoter, A1762T/G1764A, was found by site-directed mutagenesis to moderately increase genome replication and decrease HBeAg expression in transiently transfected human hepatoma cell lines (Buckwold et al., 1996; Li et al., 1999; Scaglioni et al., 1997). In a previous study, we compared in vitro replication capacity of genotype A HBV clones isolated from patients at the HBeAg positive stage of chronic infection (Parekh et al., 2003). Remarkably, clones harboring wild-type core promoter sequence all had low replication capacity in transiently transfected Huh7 human hepatoma cell line, whereas all the high replicating clones harbored core promoter mutations. Clone 4B with quadruple core promoter mutation (T1753C, A1762T, G1764A, and C1766T) had 10–20 times higher replication capacity than clone 2A, which has a wild-type core promoter sequence. Furthermore, introducing the quadruple mutation into clone 2A was sufficient to markedly boost its replication capacity to a level comparable to that of clone 4B (Parekh et al., 2003). To systemically identify mutations that affect viral RNA transcription, protein expression, genome replication, and virion secretion we generated chimeric constructs between the 2A and 4B genomes. Functional analysis revealed many unexpected features.

Results

The 2A and 4B genomes differ at 46 positions of the 3221-nt genome (1.4%). They had been cloned to the EcoRI site of pUC18 vector as tandem dimers (Parekh et al., 2003). Initially, seven chimeric constructs were generated by fragment exchange using the unique AvrII (position 177), EcoRV (1039), RsrII (1570), and ApaI (2599) restriction sites. As shown schematically in Fig. 1A, three of the seven constructs are 4B based with a shorter restriction fragment derived from 2A (chi1–chi3), while the other four are 2A based (chi4–chi7). Availability of both types of constructs permitted us to achieve reliable results. Virtually all the phenotypes were mapped to the 1-kb RsrII–ApaI restriction fragment, which covers the core promoter region, the entire core gene, and the 5′ end of the P gene (Fig. 1B). The determinants for these phenotypes were further narrowed down using 9 additional constructs for which the basal core promoter region (positions 1570–1813), 5′ 2/3rd of the core gene (1814–2330), or the 3′ end of core gene/5′ end of P gene (2331–2599) was selectively exchanged by overlap extension PCR (Fig. 1A, chi8–chi16).

Sequence variations in the core gene in addition to the core promoter mutations affect the steady state levels of pc RNA and pg RNA

Fig. 1C shows the result of a representative primer extension assay of total RNA extracted from Huh7 cells at day 4 posttransfection. Clone 2A displayed a pc RNA/pg RNA ratio of about 1:1.5, while clone 4B produced 2–4 times higher levels of the pg RNA but a negligible amount of pc RNA. The transcriptional difference between the two genomes could be mapped to the 1-kb RsrII–ApaI restriction fragment covering nucleotides 1570–2599 (Fig. 1C, compare chi1, chi2 with chi3 in the left panel; compare chi4–chi6 with chi7 in the right panel). Further study established that the 1570–1813 segment of the 4B genome, which harbors the T1753C, A1762T, G1764A, and C1766T core promoter mutations, up regulated pg RNA but nearly abolished pc RNA (chi9 vs. 4B; chi11 vs. 2A). Analysis of site-directed mutants of clone 2A revealed comparable impact of the A1762T/G1764A/C1766T triple mutation (mu 2a), slightly lower effect of the T1753C/A1762T/G1764A triple mutation (mu 4), and least effect of the A1762T/G1764A double mutation (mu 1) (Fig. 2B). Western blot analysis of core protein expression revealed a very similar pattern (Fig. 2C). These results correlate well with the HBeAg expression and genome replication phenotypes of the mutants (Parekh et al., 2003), suggesting that these core promoter mutations suppress HBeAg expression and augment genome replication at the transcriptional level.

Surprisingly, three chimeric constructs with core promoter mutations produced more pg RNA than clone 4B (Fig. 1C, chi8, chi11, chi15), and they share a 2A-derived 5′ core gene (1814–2330 fragment) (Fig. 1A). The impact of the 1814–2330 fragment can also be observed among constructs lacking core promoter mutations, as evidenced by higher pc RNA and pg RNA levels produced by chi3 than chi9 and chi10 among 4B-based constructs (Fig. 1C, left panel), as well as lower RNA levels produced by chi12 and chi13 relative to 2A itself (right panel).

Both core promoter and core gene mutations regulate core protein expression

We used two antibodies to detect core protein expression. A polyclonal rabbit antibody (Dako) efficiently detects the wild-type core protein, but it is less reactive with the core protein of clone 4B due to an E77Q mutation (Kim K et al., unpublished). Binding of this antibody to core protein of 4B origin is subject to the washing