Core promoter: A critical region where the hepatitis B virus makes decisions

Jorge Quarleri

Jorge Quarleri, Instituto de Investigaciones Biomédicas en Retrovirus y Sida, Universidad de Buenos Aires, Buenos Aires C1121ABG, Argentina

Author contributions: Quarleri J solely contributed to this paper.
Correspondence to: Jorge Quarleri, PhD, Instituto de Investigaciones Biomédicas en Retrovirus y Sida, Universidad de Buenos Aires, Buenos Aires C1121ABG, Argentina. quarleri@fmed.uba.ar
Telephone: +54-11-45083689 Fax: +54-11-4508 3705
Received: August 23, 2013 Revised: September 17, 2013
Accepted: October 19, 2013
Published online: January 14, 2014

Abstract
The core promoter (CP) of the viral genome plays an important role for hepatitis B virus (HBV) replication as it directs initiation of transcription for the synthesis of both the precore and pregenomic (pg) RNAs. The CP consists of the upper regulatory region and the basal core promoter (BCP). The CP overlaps with the 3'-end of the X open reading frames and the 5'-end of the precore region, and contains cis-acting elements that can independently direct transcription of the precore mRNA and pgRNA. Its transcription regulation is under strict control of viral and cellular factors. Even though this regulatory region exhibits high sequence conservation, when variations appear, they may contribute to the persistence of HBV within the host, leading to chronic infection and cirrhosis, and eventually, hepatocellular carcinoma. Among CP sequence variations, those occurring at BCP may dysregulate viral gene expression with emphasis in the hepatitis B e antigen, and contribute to disease progression. In this review these molecular aspects and pathologic topics of core promoter are deeply evaluated.

© 2014 Baishideng Publishing Group Co., Limited. All rights reserved.

Key words: Hepatitis B virus; Core promoter; Variants; Basal core promoter; Transcription regulation

Core tip: This review summarized the progress in our understanding of the core promoter of hepatitis B virus. This critical genomic region is involved in regulating hepatitis B virus (HBV) gene expression and viral replication, involving both host and virus-derived factors on its regulation. Such pivotal functions appear modified when genomic variations are detected and clinical implications are characterized. This review emphasizes several aspects of the HBV core promoter molecular biology and highlights its role on HBV life cycle. Finally, the most frequent genomic variations with their consequent clinical correlations are described.


INTRODUCTION
Hepatitis B virus (HBV) infection is a major cause of acute and chronic hepatitis, liver cirrhosis and hepatocellular carcinoma worldwide. Nearly 350 million are chronically infected carriers and are at high risk of complications from cirrhosis and primary liver cancer. Among these HBV carriers, 7%-30% are believed to be infected with HBV variants that express little or no hepatitis B e antigen (HBeAg) [1].

The HBV is an enveloped virus, containing a partially double-stranded DNA genome that is replicated via an RNA intermediate using its own encoded reverse transcriptase (RT). The HBV RT does not have proofreading or editing activity, therefore, together with this enormous number of mutations, HBV can be highly variable. This genomic variability is known as viral quasispecies and is associated with different clinical outcomes [2].

In this review-focus, we will summarize the most recent advances in our understanding of the core promoter of HBV, and its clinical implications.
daily virion production, errors inevitably occur during replication. The rate of HBV virion production (up to 10^11 virions per day) is considerably higher than for human immunodeficiency virus (HIV) and hepatitis C virus, and approximately 10-fold higher than for other DNA viruses. The half-life of HBV in plasma ranges from 1 to 3 d whilst the half-life of an infected hepatocyte is 10-100 d. The error rate of the HBV pol has been calculated as 10^-7 per nucleotide per day. Therefore, on a daily base, approximately 10^{14} nucleotides are replicated with potentially 10^3 base pairing errors.[3]

The major viral translational products include the precore and core, polymerase, large, medium and small envelope proteins and the X-protein as transcriptional regulator are encoded by four overlapping genes. The lifecycle of HBV includes viral attachment and entry, viral uncoating in the cytoplasm and entry of the partially double-stranded genome into the nucleus. In the nucleus the covalently closed circular DNA (cccDNA) is synthesized in the form of a viral minichromosome, creating a stable intermediate responsible for the persistence of the virus and rebound viraemia after withdrawal of anti-viral therapy in some patients. The lack of cccDNA in artificial host cells (e.g., hepatocytes of HBV transgenic mice) suggests that host specific factors may regulate cccDNA formation. However, in a model of hepatocyte nuclear factor 1α (HNF1α)-null HBV transgenic mice Ramey et al[10] have reported that HBV cccDNA in the nucleus as well as its potential precursor protein-free relaxed circular HBV DNA in the cytoplasm were present in the hepatocytes, suggesting that cycling of viral replication intermediates into the nucleus may occur in this in vitro model system.

The cccDNA is the template for transcription of pregenomic (pg) RNA as well as subgenomic mRNAs, which are translated into the viral proteins in the cytoplasm. Encapsidation follows the binding of the polymerase and core to the pgRNA in the cytoplasm. Synthesis of the minus-strand DNA by reverse transcription and partial synthesis of the plus-strand is accomplished by the HBV polymerase within the nucleocapsid. At this point the nucleocapsids are enveloped by budding into the endoplasmic reticulum followed by secretion from the cell or return to the nucleus to amplify the cccDNA reservoir in the nucleus[6].

The typical course of hepatitis B infection involves an HBeAg-positive phase with high serum HBV DNA levels. Subsequently, patients undergo a process of seroconversion in which HBeAg is lost and antibodies to HBeAg (anti-HBe) appear. Generally this signals the decline of HBV DNA to levels that are not detectable by unamplified assays and a return of aminotransferase to normal values. Among some patients, for reasons that are not yet clear, the immune pressure associated with seroconversion selects for HBV variants that express little or no HBeAg. Although the patient may develop anti-HBc, active HBV DNA replication continues with associated liver damage.[6,7].

The core promoter region regulates transcription of the pre-core region. Therefore certain mutations in this region can affect HBeAg synthesis without adversely affecting the ability of the HBV to replicate[6,9].

The primary aim of this literature review is to describe the molecular biology, function and variants of the hepatitis B virus core promoter.

**HBV GENOMIC STRUCTURE AND ORGANIZATION**

The HBV genome is a partly double stranded DNA composed of a minus (-) and a plus (+) strand that has evolved structurally to a compact genetic organization (Figure 1). At the 5'-end of the two strands, exhibits short cohesive end regions that contain two direct repeats (DRs) termed DR1 and DR2 that allow to maintain the circular configuration and are essential for viral replication[10]. The whole genome is carrying in the full-length coding (-) strand DNA (3.2 kb), while the incomplete non-coding (+) strand DNA region extends to around two-thirds of the genome length with a variable 3'-end. The viral polymerase is covalently linked to the (-) strand DNA. This strand also contains on both, 5'- and 3'-ends, an 8-9-nucleotide-long terminal redundancy, termed “r” that is critical for the formation of relaxed circular (rc) DNA synthesis. Once the viral (-) strand DNA synthesis is completed, a residual short RNA oligomer derived from the 5'-end pgRNA remains covalently bonded to the 5'-end of the (-) strand DNA to serve as a template for (+) strand DNA synthesis. A small quantity (5%-10%) of double-stranded linear DNA can also be found packaged into the nucleocapsid instead of rcDNA. During infection, both forms of viral DNA (circular and linear) can be transformed to cccDNA in the hepatocyte nucleus.

The HBV genome contains four overlapping open reading frames (ORFs) (P, preC/C, preS/S and X) encoded by the (-) strand with six start codons, four promoters (preS1, preS2, core and X), and two enhancer elements (EN I and EN II) located upstream of the core promoter. The viral genome also contains a common 3' polyadenylation signal in the core gene and a number of cis-acting signals essential for DNA replication.

Seven viral proteins (HBc; core; large, medium and small envelopes; polymerase; and HBx) are produced by translation of the polyadenylated and capped viral RNA transcripts (3.5, 2.4, 2.1 and 0.7 kb). The open reading frame (ORF) P comprised four distinguishable domains that encode viral DNA polymerase (Pol, 90 kDa) exhibiting multiple functions and involving the terminal protein, the spacer -that exhibits no enzymatic function-, the polymerase and reverse transcriptase activity -that catalyzes viral genomic synthesis-, and the fourth domain that encodes ribonuclease H[11]. The ORF preC/C partially overlaps with ORF P and encodes the precore-core protein (HBeAg), the core protein (HBc), and the pgRNA. The pgRNA is initiated five nucleotides downstream from the precore initiation codon; it serves as a template for
the translation of core protein (p21) and the viral DNA synthesis by reverse transcription within the nascent nucleocapsids. The precore polypeptide (p25) is generated from the preC transcript, it is not packaged into viral nucleocapsids instead it is proteolytic cleaved in both N and C termini when it moves to the secretory pathway to generate the mature secreted soluble HBeAg (p17) (Figure 2). The synthesis of three integral envelope glycoproteins (large L, medium M and small S) is directed by the ORF preS/S, located within ORF P (between the spacer and the RT domains). The three glycoproteins contain HBsAg and they have distinct N-terminal domains because they use three in-frame AUG start codons[12]. The smallest ORF X overlaps with ORF P. It encodes for the small protein HBx (17 kDa) which shares no homology with any known gene. This ORF is only found in all mammalian hepadnaviruses and plays a central role in the pathogenesis of HBV-induced hepatocellular carcinoma[13,14].

**REGULATION OF HBV DNA TRANSCRIPTION**

**HBV promoters and enhancers**

The HBV genome contains four promoters (core, preS1, preS2 and X) and two enhancers I and II (EN I and EN II) as cis-acting elements to control HBV transcription of the four 3.5, 2.4, 2.1 and 0.7 kb mRNAs, which have heterogeneous 5’-ends, with the exception of preS1, which lacks a TATA box and produces transcripts with a defined 5’-end[15].

The transcription of HBV mRNA can be modulated by both EN I and EN II[16]. EN I (approximately 200 bp long) is the main stimulating element and up-regulates significantly preC and HBx mRNAs but has little effect on surface protein mRNAs. It location partially overlaps the X promoter and exhibits binding sites for ubiquitous (NF1, AP1 and NFκB) and liver-specific [LRH-1/hB1F, HNF1, HNF3b, HNF4 and CCAAT/enhancer binding protein (C/EBP)] transcription factors[17,18]. EN I contains three distinct domains: a 5’ modulator element, a central core domain, and a 3’domain that overlaps with the HBx ORF. The central core domain contains the enhancer activity and binding sites for HNF3, RXF1, EFC, NF1, and a retinoic acid response element (RARE), which interacts with HNF4, the peroxisome proliferator-activated receptor α-retinoid X receptor α (PPAR/RXR α) heterodimers and Chicken ovalbumin upstream promoter transcription factors (COUP-TF). Interestingly, signal transducer and activator of transcription-3 (STAT-3) factor may bind to NF1 and HNF3-binding sites and in-
overlapping but genetically distinct core promoters that overlap the 3′ end of the X ORF and the 5′ end of the pre-C/C ORF and direct the transcription of both preC and pregenomic RNAs.

The core promoter (CP, nt1575-1849) of the viral genome has a pivotal role in the replication and morphogenesis of the virus [24]. It directs initiation of transcription for the synthesis of both the precore mRNA and pgRNA. The CP consists of the BCP (nt1743-1849) and the upper regulatory region (URR, nt1613-1742), the latter containing both positive and negative regulatory elements that modulate promoter activity [25]. The BCP is sufficient for accurate initiation of both pre-C mRNA and pgRNA transcription in vivo. BCP contains major nuclear binding sites which is recognized by a variety of the nuclear receptors super family, including HNF4 and PPARα-RXRα heterodimer and a series of transcription factors such as C/EBP to regulate the transcription both the pre-core RNA and the core RNA. Different biological functions of pre-core and pregenomic RNAs may have profound effects on HBV replication and pathogenesis. The BCP minimal essential sequence is a 108-bp fragment (nts 1742-1849) that contains direct repeat 1 (DR1) which is required for reverse transcription [26]. It overlaps with the 3′-end of the X ORF and the 5′-end of the precore region, and contains cis-acting elements that can increase EN I function. Despite the 5′ modulator element and the 3′ domain have no enhancer activity, the former can interact with C/EBP and HNF1 transcription factors and increase the central core domain activity. EN II is located upstream of the core promoter and increases preferentially the transcription of the preS1, preS2, and X promoters as well as the pg/pc mRNAs. EN II contains two regions, II A and II B, that act concomitantly as potent enhancer elements that bind various transcription factors such as C/EBP, RXR, PPAR, HNF4, HNF5, FTF, hepatic leukemia factor and Sp1 [19,20]. The preS1 and preS2 promoters control the envelope protein expression. The first one controls the transcription of the L mRNA (2.4 kb) while the preS2 promoter controls the transcription of S and M mRNAs (2.1 kb), and it seems to be stronger than the preS1 promoter and leads to the production of more S and M surface proteins, which is necessary for virus secretion [21].

The transcription of HBx is under the control of its own promoter X to generate an approximately 0.7 kb-long RNA. The X promoter has no TATA box, but its sequence is located upstream of the transcription initiation site and overlaps with the EN1 3′-end [13,22].

MOLECULAR BIOLOGY OF THE HBV CORE PROMOTER REGION

The 3.5 kb RNA synthesis is regulated by two partially overlapping but genetically distinct core promoters that overlap the 3′ end of the X ORF and the 5′ end of the pre-C/C ORF and direct the transcription of both preC and pregenomic RNAs [23].

The core promoter (CP, nt1575-1849) of the viral genome has a pivotal role in the replication and morphogenesis of the virus [24] (Figure 3). It directs initiation of transcription for the synthesis of both the precore mRNA and pgRNA. The CP consists of the BCP (nt1743-1849) and the upper regulatory region (URR, nt1613-1742), the latter containing both positive and negative regulatory elements that modulate promoter activity [25]. The BCP is sufficient for accurate initiation of both pre-C mRNA and pgRNA transcription in vivo. BCP contains major nuclear binding sites which is recognized by a variety of the nuclear receptors super family, including HNF4 and PPARα-RXRα heterodimer and a series of transcription factors such as C/EBP to regulate the transcription both the pre-core RNA and the core RNA. Different biological functions of pre-core and pregenomic RNA, as well as differential regulation of BCP by these transcriptional factors may have profound effects on HBV replication and pathogenesis.

The BCP minimal essential sequence is a 108-bp fragment (nts 1742-1849) that contains direct repeat 1 (DR1) which is required for reverse transcription [26]. It overlaps with the 3′-end of the X ORF and the 5′-end of the precore region, and contains cis-acting elements that can
independently direct transcription of the precore mRNA and pgRNA, both of which are about 3.5 kb in length.\(^{27}\) The URR comprises the negative regulatory element (NRE, nt 1613-1636) and the core upstream regulatory sequence (CURS, nt 1636-1742).\(^{28}\) The NRE, upstream of EN\(\text{II}\), can repress both EN\(\text{II}\) and CURS in differentiated liver cells. It exhibits a minor inhibitory effect on its own but is strongly repressive in the presence of EN\(\text{II}\). At least three different subregions are distinguished: NRE-\(\alpha\), NRE-\(\beta\) and NRE-\(\gamma\) which suppress strongly the CP activity. NRE-\(\gamma\) also suppresses EN\(\text{II}\) activity. The activity of the NRE is cell-type dependent which could be ascribed to a differential regulation by cell type-dependent protein factors.\(^{23,29}\)

The CURS acts as a regulatory region neither as a promoter nor with EN\(\text{II}\) participation. It can be subdivided into two domains: CURS-A (nts 1636-1703) and CURS-B (nts 1704-1743).\(^{28}\) The former can exert a positive regulatory effect on the BCP, although to a lesser degree than does the whole CURS. It can be subdivided in several sequence motifs, namely, \(\alpha\) (nts 1646-1668) which is able to bind different transcription factors, \(\gamma\) (nts 1671-1686) and \(\delta\) (nts 1687-1703), which positively regulate BCP activity; in contrast, \(\beta\) (nts 1704-1714) regulates negatively.\(^{28}\)

The EN\(\text{II}\) element regulates the activity of the CP and partially overlaps the BCP and URR. Sequences located directly at 5’ of the BCP can positively regulate this promoter in a liver-specific manner through \(\alpha\)-acting elements in the CURS overlapping EN\(\text{II}\). Further upstream there are sequences that can negatively modulate the promoter through the NRE.

Sequences within the EN\(\text{I}\) (nts 1074-1234) can also activate the BCP which is required for augmenting the CP extremely weak promoter activity. This whole region contains nucleotide motifs that constitute transcription factor binding sites, imparting at the same time liver cell specificity for optimal function of these elements. This interaction between the \(\alpha\)-acting elements with ubiquitous and liver-specific transcription factors is absolutely necessary for liver-specific expression from the CP. Moreover, the interaction between these trans-acting factors and the \(\alpha\)-acting elements allows the virus to coordinately or differentially regulate the transcription of the two mRNAs.\(^{27}\) Moreover, the stimulation of the BCP by EN\(\text{II}\) and CURS can be repressed by the effect of the upstream NRE.\(^{30}\)

The liver tropism of HBV may involve different factors including the BCP and the enhancers as well as the interaction of viral \(\alpha\)-acting elements with transacting factors present in various cell types and during different stages of liver cell differentiation.\(^{18}\) The cooperative interaction of different liver enriched and ubiquitous factors is necessary for liver-specific expression from the CP.

---

**Figure 3 Molecular biology of the hepatitis B virus core promoter.** The core promoter (CP) overlaps the 3’-end of the X open reading frame (ORF) and the 5’-end of the pre-C/C ORF. The CP comprises the upper regulatory region (URR) and the basic core promoter (BCP). The former involves a negative regulatory element (NRE) and a core upstream regulatory sequence (CURS) which can be further subdivided into two domains: CURS-A and CURS-B. The sequence of enhancer II and binding sites hepatocyte nuclear factor 1, 3 and 4 (HNF1, HNF3, HNF4) are shown. SPI: S promoter; DR: direct repeat.
Several transcriptional factors bind to regulatory sequence elements of the CP such as C/EBP, HNF1, and HNF3. Multiple members of the nuclear receptor superfamily of transcription factors, including HNF4, RXR, PPAR, COUP-TF1 and apoA I regulatory protein (ARP1), bind this regulatory element and differentially regulate synthesis of pre-C mRNA and pgRNA\(^{[9]}\). Their effects depend on their interactions; for example, binding of HNF4 and TR2 may repress preC RNA synthesis, whereas PPAR \(\alpha\) and RXR \(\alpha\) can mediate its activation, and COUP-TF1 can inhibit both preC and pgRNA synthesis. Human Sp1 transcription factor increases the synthesis of pgRNA by binding to three binding sites, but has no effect on the synthesis of pre-C mRNA\(^{[1]}\). By binding to the AT-rich regions of the CP the ubiquitous TBP plays an important role in the regulation of transcription. Thus, the synthesis of pre-C mRNA and pgRNA occurs from separate promoters which are in turn, differentially regulated\(^{[28]}\).

Selection of the transcription initiation sites of pre-C mRNA and pgRNA is not liver-cell specific and the URR is not required for the precise initiation of either RNA type\(^{[22,30]}\). The 5'-ends of pre-C mRNAs are more varied than are those of pgRNA. The latter has been mapped to a single nt region: 1815 ± 5 whereas the 5'-ends of pre-C mRNA map to nts 1785-1786 and 1791-1797. Such discrepancy could be ascribed to the presence of three TATA-like boxes (AT-rich regions) within the BCP that are not canonical TATA box (TATAAA) and a fourth TATA-like box that controls the transcription of pgRNA. The three AT-rich sequences: TA1, 5'-AGATTA-3' (nts 1750-1755); TA2, 5'-TTAAA-3' (nts 1758-1762); and TA3, 5'-TATTA-3' (nts 1771-1775); are located 20-35 bp 5' to the pre-C mRNA start sites and can bind recombinant TATA-binding protein (TBP) as well as are required for the optimal transcription of pre-C mRNA. TA4, 5'-CATATAATT-3' (nts 1788-1795), functions as both a TBP-binding site for initiation of transcription 25-30 bp downstream at the start site of pgRNA, and as the initiator (Inr) sequence for some of the pre-C mRNAs. The Inr is a minimal promoter element that overlaps the transcription start site. The 5' boundary of the pgRNA promoter is at nt 1788, the first nt of its TATA box-like sequence. The pre-C Inr consists of 5'-CATATA-3' (nts 1788-1791), overlaps the pre-C mRNA initiation sites and resembles the optimal Inr sequence: 5'-CA(T/G)T-3'.

Using mutational studies it was shown that the 3' boundary of this Inr element is at nt 1792. On the other hand, the pgRNA Inr element is the sequence 5'-CAACT-3' (nt 1788-1802) that overlaps the transcription initiation sites of pgRNA and two of its bases match the two most important bases of the optimal Inr sequence. Therefore, the 15-nt sequence (1788-1802) is sufficient to direct the initiation of both pre-C and pgRNA because it contains the TATA-like element of the pregenome promoter and the Inr of the pre-C promoter. The transcription of pre-C mRNA and pgRNA is regulated in a coordinate manner although the transcription of the two mRNA species can be separated. Sequences in the URR can stimulate the expression of both messages to a similar extent\(^{[28]}\).

The pgRNA starts 3' to the precore AUG and is translated into the core and polymerase proteins. As well, it serves as the template for the synthesis of the negative DNA strand of the virus by reverse transcription after encapsidation within the neosynthesized core particle\(^{[10]}\). The precore mRNA, which is slightly longer than the pgRNA and is initiated upstream of the precore start codon, is the template for the translation of the precore/core protein that is post-translationally processed by proteolysis to produce HBeAg, as previously described.

**BASAL CORE PROMOTER VARIANTS**

Considering its pivotal role in viral replication, sequence variations within the CP in natural isolates is restricted. Nevertheless, several studies have focused on mutations within this region and, in particular, BCP because it may dysregulate viral gene expression and contribute to disease progression. In the CP, two highly conserved regions involved in the regulation of transcription have been identified: nts 1770-1808 and nts 1813-1849. When mutations in these regions occur, they are present in viral populations with coexistence of wild-type variants that must have a compensatory effect, overcoming the potentially lethal effect of the mutation and allowing viral replication to proceed (Table 1). The pre-C mRNA initiation sites map within the first region\(^{[26,34]}\) as does TA4 that controls pgRNA synthesis. Mutations within TA4 severely reduce the synthesis of pgRNA, leaving synthesis of the pre-C mRNAs unaffected\(^{[27,32]}\). In addition, this region has been shown to be essential for the trans-activating function of X protein\(^{[25]}\). The second conserved region contains the initiation site of pgRNA transcription\(^{[28,35]}\) as well as coding for the C-terminal of the X protein. The overlap of this region with the 5'-end of the pre-C ORF containing DR1 is a further reason for the low sequence divergence. Mutations in the CP region may have repercussions on viral gene expression and/or replication, with a concurrent impact on viral pathogenesis.

The double mutation A1762T and G1764A in the BCP has been described in various disease states or settings of HBV infection\(^{[14-17]}\). The presence of the double mutation is clearly associated with downregulation of HBeAg production, as demonstrated by transfection studies\(^{[5-8]}\). There is evidence to suggest that the double BCP mutation results in decreased levels of the precore mRNA and therefore diminished production of HBeAg. Taking into consideration that HBV BCP contains a TATA-like box that controls transcription factor HNF4 and that of HNF1 and changes two amino acids in the overlapping X protein sequence. HNF1, which did not affect the wild type BCP, suppressed the precore RNA expres-
sion of the double mutant. Furthermore, the X protein did not affect the HNF4 activity on the core promoter and affected the HNF1 activity on the core promoter of only the double mutant[38] (Table 1).

During in vivo infection its increased presence could be temporally correlated with HBs seroconversion often preceding HBsAg clearance by many years[39] but still there are discrepant results on its relationships with the HBsAg negativity considering that some of the studies were performed at different stages of chronic liver disease. More recently its presence was also related with higher viral load levels[40] in opposite with other previous findings[41]. These discrepancies could be associated with the upregulation of pgRNA production, promoting encapsidation and core protein production[42] whereas others, however, reported decreased precore mRNA levels, but wild-type levels of replication and gene expression[43]. Such differences observed in transcript levels when BCP mutations are detected could be explained by the alteration of a nuclear receptor binding site for HNF1 that appears to be essential for this activity. By using the woodchuck animal model and the woodchuck hepatitis virus (WHV), such variations in the HNF1 binding site diminished significantly the synthesis of viral pgRNA[44].

The BCP double mutation is correlated with the L130M and V131I mutations in the overlapping X ORF gene product. The removal of the nuclear receptor binding site had no effect on the transcription of HBV mRNAs as was shown by transfection studies in Huh7 hepatoma cells, the two codon change in the X protein suppressed both transcripts, and the creation of the HNF1 site restored the pgRNA level[45]. In addition, between both, the T1762 change is decisive for the mutant phenotype as was revealed by analysis of revertants with either one or the other of the BCP mutations[3]. Studies from chronically infected patients have provided further information regarding the double A1762T/G1764A mutation in relation to genotype[45,46]. Compared with genotype A and B cases, patients with genotypes C and D have lower rates and usually delayed onset of spontaneous HBsAg seroconversion. HBV-genotype C has a higher frequency of A1762T/G1764A mutation and preS deletion, and a higher viral load than genotype B. Similarly, genotype D has a higher prevalence of BCP A1762T/G1764A mutation than genotype A[47]. The A1762T/G1764A BCP mutations in genotype C isolates correlated with increased replication capacity, while the A1752G/T mutation found in genotype B isolates correlated with low replication capacity[48]. Interestingly, genotype C isolates with wild-type BCP sequences replicated less efficiently than genotype B isolates, due to less efficient transcription of the pgRNA, but more efficient virion release was observed with the former[49]. Core promoter changes are significantly more common in patients infected with HBV genotypes exhibiting C at nucleotide 1858, while precore stop codon changes appears exclusively in those genotypes that have T at nucleotide 1858. Nevertheless, the double BCP and the precore stop codon mutations are far from being mutually exclusive[50]. Among HBV genotype C infected patients from China, those acutely infected with BCP/precore mutant viruses had higher viral load than chronic patients with the mutant virus but both showed a lower prevalence of A1762T/G1764A, G1896A, and G1899A, but higher prevalence of T1758C mutations. The T1758C and A1762T/G1764A mutations appeared as mutually exclusive[51].

### Table 1  More frequent basal core promoter mutations and their clinical relevance

<table>
<thead>
<tr>
<th>Nucleotide position</th>
<th>Clinical relevance</th>
</tr>
</thead>
<tbody>
<tr>
<td>1762T + 1764A (BCP double mutant)</td>
<td>Chronic HBV Fulminant hepatitis</td>
</tr>
<tr>
<td>1762T</td>
<td>Decrease in HBsAg production and increase viral replication Diminished binding of a liver-specific transcription factor, resulting in a decrease in HBsAg, but unchanged amounts of HBV pregenomic RNA</td>
</tr>
<tr>
<td>1764A</td>
<td>Enhance viral replication through the combined effects of X gene mutations and the appearance of a HNF-1 transcription factor binding site</td>
</tr>
<tr>
<td>1762T</td>
<td>Elevated ALT (diminishing circulating HBsAg levels → augment the host immune response to HBV-infected hepatocytes → increasing hepatocyte apoptosis)</td>
</tr>
<tr>
<td>1764A</td>
<td>More often in patients with HBV genotypes that have 1858C (i.e., genotype C)</td>
</tr>
<tr>
<td>1762T</td>
<td>HBeAg seroconversion and histological inflammation</td>
</tr>
<tr>
<td>1764A</td>
<td>No suppression on HBV RNA transcription and only slightly decreases the efficiency of virus replication</td>
</tr>
<tr>
<td>1653T</td>
<td>Usually together with the 1762T + 1764A in patients with fulminant hepatitis and hepatocellular carcinoma</td>
</tr>
<tr>
<td>1753-1757</td>
<td>Together with the 1762T + 1764A mutation, have been detected in patients with fulminant hepatitis and in patients with HCC ALT levels and histological changes</td>
</tr>
<tr>
<td>1764A/T + 1766A/G</td>
<td>Found in active and inactive disease in conjunction with 1810T + 1811T double mutation</td>
</tr>
<tr>
<td>1762A1766A mutation, together with 1762T, was found in fulminant hepatitis and HCC patients</td>
<td>The 1764T + 1766G mutation was found in a patient with fulminant recurrent hepatitis after liver transplantation, but was absent in patients with fulminant hepatitis.</td>
</tr>
<tr>
<td>1766T + 1768A</td>
<td>Fulminant hepatitis Together with 1762T + 1764A, in a patient with recurrent hepatitis following liver transplantation With 1766A in a HBeAg-negative asymptomatic carrier Exacerbation of HBV infection and created two overlapping low-affinity HNF1 sites</td>
</tr>
</tbody>
</table>

BCP: Basal core promoter; HBV: Hepatitis B virus; HNF-1: Hepatocyte nuclear factor 1; ALT: Alanine transaminase; HCC: Hepatocellular carcinoma.
The double mutation has been detected with increased frequency in patients with fulminant hepatitis, including children, those with HBeAg- and anti-HBe-positive chronic hepatitis, and HCC patients, but less frequently in asymptomatic chronic carriers. In addition, it has been detected in immunosuppressed, liver transplant, and seronegative patients. Interestingly, it was recently reported that BCP T1762/A1764 mutant is an independent risk factor for progression to cirrhosis rather than hepatocellular carcinoma (HCC) in chronic HBV infection. There are differences in the prevalence of such variants between studies in various settings, and this may relate to differences between prevalent genotypes, the importance of which is not always recognized and therefore not clearly reported. Very recently Yang et al. have postulated that a quantitative analysis of the BCP double mutation (and G1896A) can predict interferon-induced HBeAg seroconversion.

The BCP mutations, as mentioned here, have been found in patients regardless of HBeAg status. However, in anti-HBe-positive patients, the double mutation was often accompanied by T1753C/G mutation. In addition, other point mutations upstream and downstream of T1762/A1764 have been described, occurring either alone or in combination with the double mutation, and in different settings, including chronic hepatitis, fulminant hepatitis B, HCC, and liver transplantation. More recent studies have identified additional mutations in the BCP region that are prevalent in HCC patients, and particularly novel individual and combination patterns of mutations in the X/precore region of HBV genotype D1 as predictors of HCC. These include T1673/G1679, G1727, C1741, C1761, A1757/T1764/G1766, T1773, T1773/G1775 and C1909 (Table 1).

The mutations G1613A (NRE) and C1653T, both of which tended to also occur simultaneously in HCC patients may suppress HBeAg production, enhanced viral DNA synthesis, and bound RFX1 with higher affinity. Since the 1762/1764 mutations overlap with the X ORF and lead to amino acid changes in the X protein (L130M and V131I, respectively), this double substitution on its own or together with G1386A (V5M) was encountered at a significant rate in HCC patients by activating NF-kB activity. Synergistic effects of A1896, C1653T and T1762/A1764 mutations in genotype C2 on development of hepatocellular carcinoma has been reported. Thus, to summarize, T1653, V1753, T1762, A1764, T1766, and A1768 have been found to occur more frequently in HCC patients.

The double BCP mutation was found to be associated with increased risk of HCC independent of genotype, while the precore A1896 mutation was associated with decreased risk. According to the Risk Evaluation of Viral Load Elevation and Associated Liver Disease/Cancer-Hepatitis B Virus study, the infection by genotype B strains exhibiting the BCP double mutation over a 10-year period increases 23% the risk of developing HCC. Likewise, the presence of family history of HCC multiplies the risk of HCC at each stage of HBV infection.

Core promoter mutations other than those at positions 1762 and 1764 can have major impact on viral DNA replication and HBeAg expression. In vitro studies reveal that CM mutants with additional mutations (T1753C and/or G1766T) replicated at high level (Table 1). When mutational complexes were assayed, the 1762/1764/1766 and 1753/1762/1764/1766 showed higher replication rates and lower HBeAg expression than 1762/1764 mutations alone, whereas the 1753/1762/1764 variant was not much different from the double BCP mutant. Single mutations including T1753C/A/G (V), 1762/1764, A1846T, G1896A, 1899 and A/G1913, as well as triple mutations 1753/1762/1764 and 1762/1764/1766 or T1768A, were more frequently detected in patients with acute-on-chronic liver failure than in patients with chronic hepatitis, and mostly associated with genotype B- than C-infected patients.

Deletions within the CP region -mostly at BCP- and varying in length from 1 to 21 bp have been reported once again in different settings. These include fulminant hepatitis, chronic hepatitis, asymptomatic infection, serologically silent infection, HCC, renal dialysis with atypical HBsAg-negative infections characterized by either the presence of viral antibodies alone or no serological markers, and liver and renal transplantation, HIV coinfection; they are also found in patients who have survived hematological malignancies or solid tumors. Such deletion variants are often characterized by low viremia levels, and may need help from the wild-type virus for survival. Likewise, reduced levels of transcription and progeny virus production were obtained by transfection with clones having an 8 bp deletion in the BCP (nt 1768-1775). It appears, therefore, that the strain genetic background in which the BCP and G1896A mutations arise, in relation to additional ones, determine replication rate, expression of HBeAg, and pathogenicity. A limited number of reports have shown the presence of deletions in the URR or in the region of the CP overlapping the pre-C ORF.

Insertions within the BCP have been described in patients with CH and in those undergoing orthotopic liver transplantation. One patient who had a fulminant exacerbation of CH after heart transplantation was reported to be infected with a high-replicating HBV variant carrying an 11-bp insertion in the CP. This insertion occurred shortly before, or during, FH and created a binding site for HNF1. Thus, the emergence or presence of a novel HNF1, or putative HNF3, site may be related to fulminant exacerbating hepatitis in immunosuppressed patients. This concurs with the presence of such strains in renal transplant patients with severe liver disease but not in those with mild disease. Simultaneous detection of deletion and insertion at BCP was recently reported in a single patient with HCC from South Africa with concomitant presence of other variations that could...
accelerate the HCC development[6].

The G1862T substitution affecting codon 17 of the precore protein of the virus was only observed in patients infected with genotype B. This variant was five times more common in patients with fulminant hepatitis than in chronic carriers. The G1862T mutation leads to substitution of valine for phenylalanine, affecting the -3 position in the signal peptide recognition motif, impairing the processing of the precore/core protein into HBsAg[7]. A similar behavior was described for viral variants carrying the C1856T mutation. It leads to P15S substitution at precore. This amino acid change affects the signal peptide also, and may therefore affect production of HBsAg from the precursor precore/core, but there are no experimental data supporting this proposition at the moment[7]. This variant was initially found by chance while sequencing the complete HBV genome of an HIV positive patient with a strange serological profile[7], and then in Hong Kong Chinese patients with chronic hepatitis[7].

HBV possesses the smallest genome of any virus known to infect man. Therefore, it is not surprising that HBV utilizes its genetic material economically. The core promoter is an emblematic model of this strategy whose DNA harbors several pivotal functions. Its sequence involves three critical elements (EN I, EN II, BCP, and the overlapping X ORF) which are able to execute critical viral functions (stimulate the S, X, and CP; regulate the transcription of both pre-C mRNA and pgRNA; encode the functions (stimulate the S, X, and CP; regulate the transcription by its two enhancers). Its integrity appears essential to maintain viral replication, and variation may contribute to the persistence of HBV within the host, leading to chronic infection and, eventually, cellular transformation.

REFERENCES
5 Locarnini S, Zoulim F. Molecular genetics of HBV infection. Antivir Ther 2010; 15 Suppl 3: 3-14 [PMID: 21041899 DOI: 10.3851/IMP1619].
16 Su H, Yee JK. Regulation of hepatitis B virus gene expression by its two enhancers. Proc Natl Acad Sci USA 1992; 89: 2708-2712 [PMID: 1313564 DOI: 10.1073/pnas.89.7.2708].
Quarleri J. HBV core promoter


40 Li DH, Newbold JE, Cullen JM. Natural populations of woodchuck hepatitis virus contain variant precore and core sequences including a premature stop codon in the epsilon motif. Virology 1996; 220: 256-262 [PMID: 9659124 DOI: 10.1006/viro.1996.0311]


47 Chen HL, Hussain M, Lok AS. Different hepatitis B virus genotypes are associated with different mutations in the core promoter and precore regions during hepatitis B e antigen seroconversion. Hepatology 1999; 29: 976-984 [PMID: 10051506 DOI: 10.1002/hep.10290352]


January 14, 2014 | Volume 20 | Issue 2 |

Quarleri J. HBV core promoter


P- Reviewers: Sener A, Shlomai A, Taylor J; S- Editor: Wen LL; L- Editor: Logan S; E- Editor: Wu HL