JOURNAL OF VIROLOGY, Sept. 1993, p. 5402–5410
0022-538X/93/095402-09$02.00/0

Vol. 67, No. 9

Hepatitis B Virus Genotype A Rarely Circulates as an HBe-Minus Mutant: Possible Contribution of a Single Nucleotide in the Precore Region

JI-SU LI,[1]† SHU-PING TONG,[1]† YU-MEI WEN,[2] LUDMILA VITVITSKI,[1] QING ZHANG,[3] AND CHRISTIAN TRÉPO[1]*

Unité de Recherche sur les Hepatites, le SIDA et les Retrovirus Humains, Institut National de la Santé et de la Recherche Médicale 271, 151 Cours Albert-Thomas, 69424 Lyon Cedex 03, France,[1] and Laboratory of Molecular Virology and Department of Microbiology, School of Basic Medical Sciences, Shanghai Medical University,[2] and Da-Hua Hospital,[3] Shanghai 200032, People's Republic of China

Received 17 March 1993/Accepted 14 June 1993

The emergence of HBe-minus hepatitis B virus (HBV) mutants, usually through a UAG nonsense mutation at codon 28 of the precore region, helps the virus to survive the anti-HBe immune response of the host. Host and viral factors that predispose to the emergence of such mutants are not well characterized. The fact that the precore region forms a hairpin structure essential for the packaging of viral pregenomic RNA may explain the extremely high prevalence of the UAG mutation at codon 28. It converts a wobble U-G pair in the packaging signal between nucleotide 3 of codon 15 (CCU) and nucleotide 2 of codon 28 (UGG) into a U-A pair. Since genotype A of HBV has a CCC sequence at codon 15, the UAG mutation would, instead, disrupt a C-G pair present in the wild-type virus. This alteration was shown by transfection experiments to greatly compromise the packaging of pregenomic RNA. The implication of this finding was elucidated by molecular epidemiological studies. Genotype A was found to be the most prevalent genotype in the wild-type virus populations in France but was found in only 1 of the 46 isolates of HBe-minus mutants found there. These mutants were contributed chiefly by genotype D, the second most prevalent genotype in France, which is characterized by a CCU sequence at codon 15. The role of the single nucleotide at codon 15 was confirmed by the finding of the single genotype A isolate in which both wild-type and mutant viruses were present. Interestingly, nearly all of the mutants had a codon 15 sequence of CCU instead of the CCC present in the wild-type viruses. Our results suggest that genotype A of HBV rarely circulates as HBe-minus mutants, probably because of a requirement for a simultaneous sequence change at codon 15. These data, together with the virtual absence of genotype A in the Chinese samples examined, may provide some insights into the uneven prevalence of HBe-minus mutants in the world.

Hepatitis B virus (HBV) can be classified into four major subtypes based on antigenic determinants present on viral surface protein (9) or into five genotypes on the basis of nucleotide sequence variation of the entire genome (19, 35). Genotype and subtype systems do not correspond to each other, since a subtype may belong to different genotypes while different subtypes may share the same genotype (Table 1). Genotype- or serotype-associated differences in the biological or pathobiological properties of HBV have not been established. The HBV genome specifies a secretory protein commonly referred to as HBe antigen (HBeAg). Expression of HBe probably helps induce persistent infection (7, 17), but this protein is essential for neither virus replication nor infectivity (1, 6, 7, 23, 26, 27, 33). Under the pressure of protective anti-HBe immunity, HBV may continue to replicate by turning off the expression of HBeAg and escaping the specific immunity at the molecular level. This is achieved through nonsense, frameshift, or initiation codon mutation in the precore region, a sequence that codes for the first 29 amino acid residues of the HBe precursor (for reviews of early reports, see references 2, 3, and 24; also see references 4, 5, 8, 11, 14, 16, 18, 31, 32, and 36). Such HBe-minus (HBe^-) mutants are found chiefly in Asia and the Mediterranean region and are only rarely reported from Western Europe and North America.

A nucleotide transition at precore codon 28 that converts a tryptophan codon (UGG) into a stop codon (UAG) is the most frequently encountered HBe^- precore mutation in nature. Recently evidence has been gained that the high mutational propensity of nucleotide 2 of codon 28 is related to the viral pregenome encapsidation (ε) signal overlapping the HBe coding sequence (12, 28, 30; see Fig. 1A). A part of the ε signal is base paired (12, 28; see Fig. 1B), and the common UGG-to-UAG mutation at codon 28 converts a wobble U-G base pair formed between nucleotide 3 of precore codon 15 (CCU) and nucleotide 2 of codon 28 into a U-A pair (see Fig. 1B, d and e). The significance of base pairing for the mutational pattern is corroborated by a rare HBe^- mutant which has a UGA (instead of a UAG) nonsense mutation at codon 28 and a CUC (instead of a CCU) sequence at codon 15 (1; see Fig. 1B, f). The simultaneous changes at codon 15 are required to maintain efficient packaging of pregenomic RNA in this mutant (28).

These findings prompted us to explore whether the base-pairing requirement would restrict the generation of HBe^- mutants in HBV isolates having uncommon precore sequences. Since genotype A of HBV differs from other genotypes in having a variant proline codon of CCC at

* Corresponding author.
† Present address: Molecular Hepatology Laboratory, Massachusetts General Hospital, Charlestown, MA 02129.

TABLE 1. Relationships among variants of precore codon 15, viral genotypes, and serological subtypes

Sequence of codon 15	Viral genotype			Subtype
	Name	Size of gene (nucleotides)		
		Core	Pre-S	
CCC	A	555	519	adw
CCU	B	549	519	adw
CCU	C	549	519	adw-adr
CCU	D	549	486	ayw
CCU	E	549	486	adw

precore codon 15 (20, 34; Table 1), a UAG or UGA nonsense mutation produced at codon 28 should disrupt a C-G pair present in the ε signal of the wild-type virus (see Fig. 1B, a, b, and c). Whether distortion of a single base pair of the ε signal hampers viral replication and restricts the emergence of HBe$^-$ mutants from this particular HBV genotype warranted further study. In this investigation, we demonstrated the poor replication capacity and rare circulation in nature of genotype A-related HBe$^-$ mutants. Our results also raised the possibility that the prevalence of HBe$^-$ mutants are inversely correlated with the prevalence of genotype A of HBV in different parts of the world.

MATERIALS AND METHODS

Precore mutants, transfection experiments, and nucleic acid analysis. Mutants α1/wt and α1/wt→M2 have been described previously (30). These two mutants served as the standard wild-type virus and a common HBe$^-$ mutant, respectively. Mutant CCCcd15 was constructed from α1/wt by changing codon 15 from CCU to CCC. This would convert a U-G pair in the ε signal into a C-G pair (see Fig. 1B, g). Mutant CCCcd15/UGAcd28 was generated from CCCcd15 by mutating codon 28 from UGG to UGA. Mutant CCCcd15/UAGcd28 was constructed from CCCcd15 by changing codon 28 to UAG and codon 29 from GGC to GAC. Both nonsense mutants disrupted a C-G pair in the ε signal (see Fig. 1B, h and i). Transfection of hepatoma cell line HepG2 has already been described (27). A flask (25 cm^2) was transfected with 10 μg of a pUC-HBV dimer. Cells were harvested at day 4 posttransfection.

Northern (RNA) and Southern blot analyses were done as previously described (30), except for minor modifications. For extraction of total RNA, the DNase I digestion step was performed with 2 U of RQ1 RNase-free DNase I (Promega) at 37°C for 15 min. Virus particles were directly concentrated from 6 ml of culture medium by centrifugation through a 10 to 20% sucrose gradient. After digestion with DNase I, samples were subjected to proteinase K digestion without a heat-inactivation step.

Serum samples. HBeAg-positive French serum samples were randomly chosen from among those sent to our laboratory for routine analysis of HBV markers. They were positive for HBV DNA by dot blot hybridization. HBeAg-negative, anti-HBe-positive serum samples were collected from French patients with liver diseases. HBeAg-positive Chinese serum samples were collected from patients in Da-Hua Hospital, Shanghai, People's Republic of China. Unless otherwise stated, a single sample from each patient was analyzed.

Oligonucleotides. The oligonucleotides used in this study were synthesized in a polymerase chain reaction (PCR) MATE/391 DNA synthesizer (Applied Biosystems). Their sequences and positions on the HBV genome (according to reference 29) are listed in Table 2.

Detection of sequence variations at precore codons 15 and 28. DNA was extracted from 100 μl of serum, and 1/50 of the volume was used for amplification of a 0.4-kb sequence containing the entire precore region plus part of the core gene by PCR (13). The primer pairs used were pX32 and pC30 for French samples and pX32-Ch and pC30-Ch for Chinese samples (Table 2). HBeAg-positive and anti-HBe-positive samples were amplified for 35 and 40 cycles, respectively. For anti-HBe-positive samples that failed to be amplified by the single PCR, a nested PCR protocol was adopted. Samples were amplified by pX32 and pC30 for 35 cycles, and 2-μl volumes of the products were reamplified by a further 35 cycles by using internal primers pC13 and C126 (Table 2). Extreme care was taken to guard against contamination.

TABLE 2. Oligonucleotides used in this study

Name	Polarity	Positions	Sequence (5′ to 3′)[a]
pX32	+	1783–1806	caggaattCTGTAGGCATAAATTGGTCTGCGC
pC30	−	2171–2149	cctaagcTTAACATAACTGACTACTAGGTC
pX32-Ch	+	1783–1806	caggaattCTGTAGGCATAAATTGGTCTGTTC
pC30-Ch	−	2171–2149	cctaagcTTGACATAGCTGACTACTAATTC
pC13	+	1828–1847	CACCTCTGCCTAATCATCTC
C126	−	2028–2008	TAAGGCTTCTCGATACAGAGC
CCC	+	1851–1868	TTCATGTCCCACTGTTCA
CCU	+	1851–1868	TTCATGTCCTACTGTTCA
C2	+	1973–1992	TTCCTTCCGTCAGAGATCTC
M0	+	1887–1908	TGGGTGGCTTTGGGGCATGGAC
M1	+	1888–1908	GGGTGGCTTTAGGGCATGGAC
M2	+	1887–1908	TGGGTGGCTTTAGGACATGGAC
pS4	+	2833–2852	GGAACAAGAGCTACAGCATG
pS5	−	2955–2935	GGTTGAAGTCCCAATCTGGAT
C3	+	2314–2333	CCTATCTTATCAACACTTCC
C4	−	2383–2364	GAGTTCTTCTTCTAGGGGAC
C5	+	2301–2320	ACCACCAAATGCCCCTATCT
C6	−	2399–2380	CGTCTGCGAGGCGAGGGAGT

[a] Non-HBV sequences are shown in lowercase letters, and restriction sites are underlined.

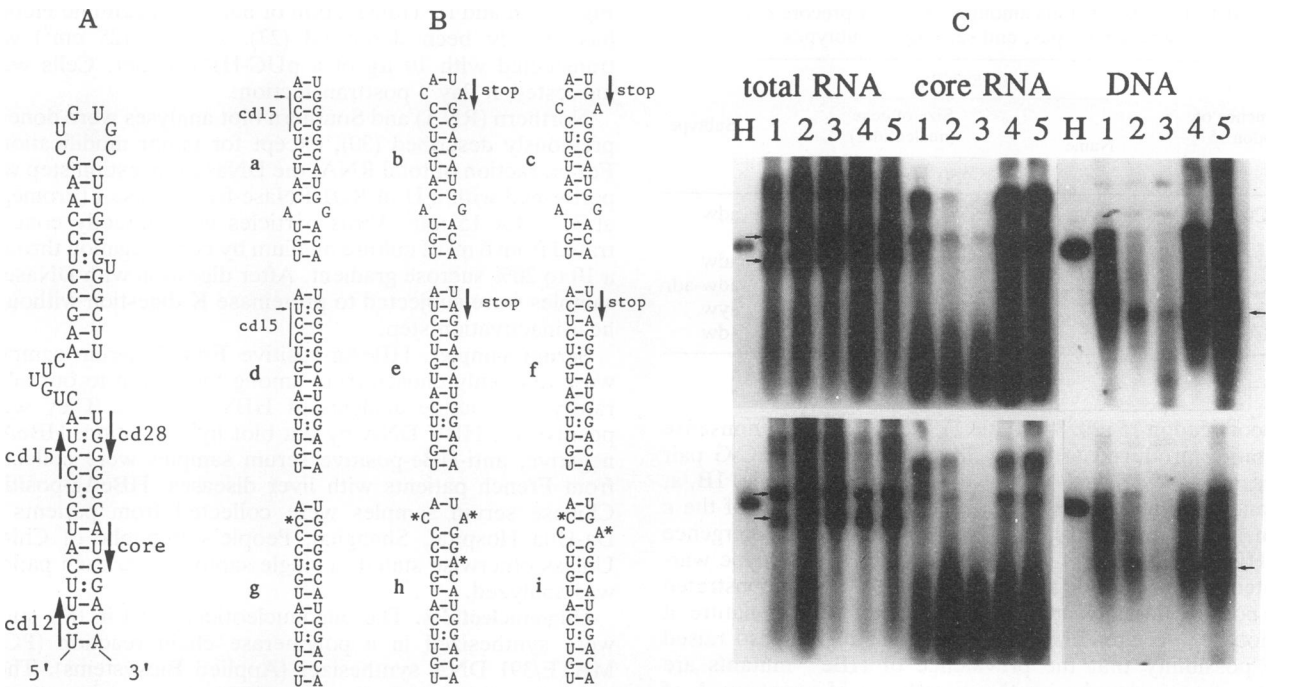

FIG. 1. Base pairing in the HBV pregenome encapsidation signal and replication capacities of artificial nonsense mutants of the precore region. (A) Proposed secondary structure of the packaging signal (12), starting from amino acid codon 12 of the precore region and ending at the first seven nucleotides of the core gene. U-G pairs are connected by colons, while C-G and A-U pairs are connected by hyphens. Note that nucleotides 2 and 3 of codon 15 pair with nucleotides 3 and 2, respectively, of codon 28. (B) Structures of a base-paired region of the packaging signal in wild-type HBV and HBe^- mutants due to a nonsense mutation at precore codon 28. a, b, and c, genotype A in the form of wild-type virus (a) and UAG (b) and UGA (c) nonsense mutants; d and e, other HBV genotypes in the form of wild-type virus (d) and a UAG mutant (e); f, rare UGA mutant of genotype A; g, h, and i, genotype D-derived artificial mutants as wild-type virus (g) and UAG (h) and UGA (i) mutants. Changed nucleotides are shown as stars. (C) Transfection results obtained with the mutants in two independent experiments. Lanes 1 to 5 correspond to mutants CCCcd15, CCCcd15/UGAcd28, CCCcd15/UAGcd28, α1/wt, and α1/wt→M2. Their secondary structures in a base-paired region of the ε signal are shown in panel B as g, i, h, d, and e. The positions of pregenomic and subgenomic RNAs and single-stranded replicative forms of viral DNA are indicated by arrows. Lane H, linearized 3.2-kb HBV DNA.

PCR products were hybridized with pairs of oligonucleotide probes (oligoprobes). Mutation at precore codon 28 was detected by mixed oligoprobes M1 and M2, and the corresponding wild-type sequence was detected by probe M0 (13; Table 2). Sequence variation at nucleotide 3 of precore codon 15 was detected by probes CCC and CCU (Table 2). Hybridization was done as previously described (13), with some modifications. (i) Southern blotting was used to facilitate comparison of signal intensities in the gel and in the autoradiograph. (ii) For controls, tandem dimers of two HBV constructs double digested with *Bgl*II and *Nco*I were used. The precore region was contained in the smallest (600-bp) fragment (see Fig. 2). Mutant HBVα1→M0 (30) was CCU^+ CCC^- $M0^+$ $M1\text{-}M2^-$, whereas mutant CCCcd15/UAGcd28 should be CCU^- CCC^+ $M0^-$ $M1\text{-}M2^+$ (see transfection section). In certain cases, hybridization of PCR products cloned into M13 mp19 or pUC18 was carried out. For M13 clones, a 200-μl volume of culture supernatant was spotted onto a nitrocellulose filter, baked, and hybridized with oligoprobes as described above. For pUC clones, the bacterial colony hybridization method was used.

Identification of HBV genotypes. Three independent methods were used to identify genotype A isolates positively. (i) The 0.4-kb PCR products of the precore-core gene were hybridized under stringent conditions with a genotype A-specific probe, C2, that covered a sequence in the core gene (Table 2). (ii) The PCR products from some samples in each clinical category were cloned and sequenced. Comparison with HBV sequences of known genotypes enabled assignment of the sequence analyzed to one of the genotypes. (iii) A simple typing method based on size polymorphism in the core gene and the pre-S region was developed. Genotype D (and E, for which only a single isolate from a chimpanzee has been reported [35]) differs from other genotypes by an in-frame deletion of 33 bp within the pre-S region (19, 29), while genotype A differs from all other genotypes in having an in-frame insertion of 6 bp near the end of the core gene (1, 20, 34; Table 1). By detecting polymorphism in both genes, known HBV genotypes can be divided into three groups: A, B-C, and D-E. Polymorphism in the pre-S region was detected by primer pair pS4 and pS5 (Table 2), which amplified fragments of 123 bp for genotypes D and E and 156 bp for genotypes A, B, and C. Size polymorphism in the core gene was detected by primers C5 and C6 (Table 2), which produced amplification products of 105 bp for genotype A and 99 bp for the other genotypes. A total of 35 or 40 cycles of amplification were carried out (95°C for 30 s, 50°C for 20 s, and 72°C for 20 s). PCR products were analyzed by electrophoresis through a 4% NuSieve agarose gel in Tris-borate buffer. For samples with extremely small amounts of viral DNA, samples were preamplified by primers C5 and C6 for 35 cycles and then subjected to the same cycles of amplification with primer pair C3 and

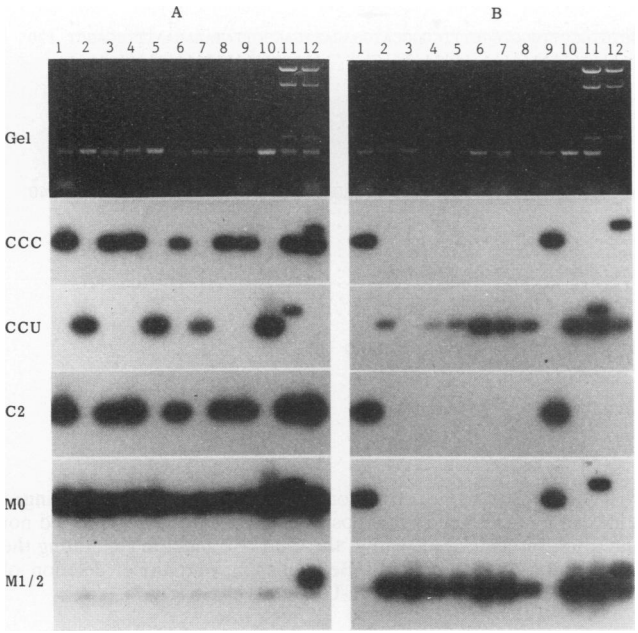

FIG. 2. Hybridization patterns of PCR products of the precore-core gene with oligoprobes for detection of variants of precore codon 15 (CCC and CCU), codon 28 mutants (M0 and M1-M2), and genotypes (C2). Panels: A, 12 HBeAg-positive French serum samples; B, 11 anti-HBe-positive French serum samples (lanes 2 to 12) and a control wild-type genotype A isolate (lane 1). The last two lanes in both panels A and B were superimposed with control DNAs: a *Bgl*II-*Nco*I-digested HBV-α1→M0 dimer in pUC18 (lane 11) and a CCCcd15/UAGcd28 dimer in pUC18 (lane 12). The precore region was contained in band 3 (just above the PCR product band). Note that the sample in panel B, lane 3, did not hybridize with either the CCC or the CCU probe. For the sequence of this isolate, see Fig. 5 (FDc).

C4 (see Table 2). This enabled classification of HBV isolates as genotype A (76 bp) or non-genotype A (70 bp).

Nucleotide sequence accession numbers. The GenBank accession numbers of the nucleotide sequences determined in this study are L12356 to L12365.

RESULTS

HBe⁻ mutants of the CCC variant exhibited markedly diminished efficiency of pregenome packaging. Natural HBV isolates exhibit sequence microheterogeneity at amino acid codon 15 of the precore region: proline codons of CCC for genotype A and CCU for the other genotypes (Table 1). The possible effect of this single-nucleotide difference on pregenome packaging of wild-type virus and codon 28 mutants was investigated by site-directed mutagenesis of genotype D mutant α1/wt coupled with transfection experiments. Quantitative results were obtained by testing the mutants in triplicate and analyzing combined cell lysates and medium, as well as by repeat experiments (Fig. 1C). Mutant CCCcd15, mimicking the codon 15 sequence of genotype A viruses, exhibited levels of total RNA, core nucleic acid, and replicating DNA similar to or slightly lower than those of parental mutant α1/wt (Fig. 1C, compare lanes 1 and 4). The packaging efficiency (relative signal intensity in the core nucleic acid fraction in comparison with the total RNA fraction) was unchanged. Introduction of a UGA or UAG stop codon at codon 28 (CCCcd15/UGAcd28 and CCCcd15/UAGcd28) resulted in significantly reduced amounts of core nucleic acid and replicating DNA, especially for mutant CCCcd15/UAGcd28 (lanes 2 and 3). This was in contrast to a UAG mutation of codon 28 coupled with a CCU sequence at codon 15, in which the packaging efficiency was not affected (a1/wt→M2, lane 5). Thus, a nonsense mutation at codon 28, in association with the CCC variant of codon 15, markedly diminished the packaging efficiency of pregenomic RNA (the two nonsense mutants were estimated to have 30 and 10%, respectively, of the packaging capacity of CCCcd15). The more drastic impairment seen in mutant CCCcd15/UAGcd28 probably reflected a positional effect. It disrupted base pair 2 rather than base pair 3 near the bulge (Fig. 1B, compare h and i) and may have more severely destabilized the stem structure.

TABLE 3. Hybridization of PCR products from HBeAg-positive French and Chinese samples with oligoprobes for detection of genotypes and precore mutations

Probe(s)	No. of samples that hybridized			
	47 French samples		18 Chinese samples	
	CCC + C2	CCU	CCC + C2	CCU
M0	31	10	0	12
M0 + M1-M2	1	3	0	6
M1-M2	0	1	0	0
Neither	0	1[a]	0	0
Total	32	15	0	18

[a] This sample contained an amino acid change at precore codon 29. See FD6 in Fig. 5 for the sequence.

TABLE 4. Hybridization of single or nested PCR products from anti-HBe-positive French samples with oligoprobes for detection of genotypes and precore mutations

Probe(s)	No. of samples that hybridized					
	Single PCR products			Nested PCR products		
	CCC + C2	CCU	None	CCC + C2	CCU	None
M0	1	0	0	2	0	0
M0 + M1-M2	0	2	0	0	4	0
M1-M2	0	24	1[a]	0	10	0
Neither	0	0	0	0	0	1[b]
Total	1	26	1	2	14	1

[a] The sequence at precore codon 15 was CCU. See FDc in Fig. 5 for the sequence.

[b] The sequence at precore codon 15 was CCC, and there was no HBe-abolishing mutation.

Prevalence of codon 15 variants in Chinese and French samples. The possible implications of the above-described findings for the emergence of natural HBe⁻ mutants were explored by molecular epidemiological studies. We first investigated the relative prevalence of the CCC and CCU variants in French and Chinese patients. HBeAg-positive serum samples were used, since anti-HBe-positive samples might have a selective bias against the CCC variants. The precore-core gene from the sera of 47 French patients and 18 Chinese patients was amplified, and PCR products were hybridized with oligoprobes CCC and CCU. Some hybridization profiles are shown in Fig. 2A, and results are summarized in Table 3. Most of the French samples (32 [68%] of

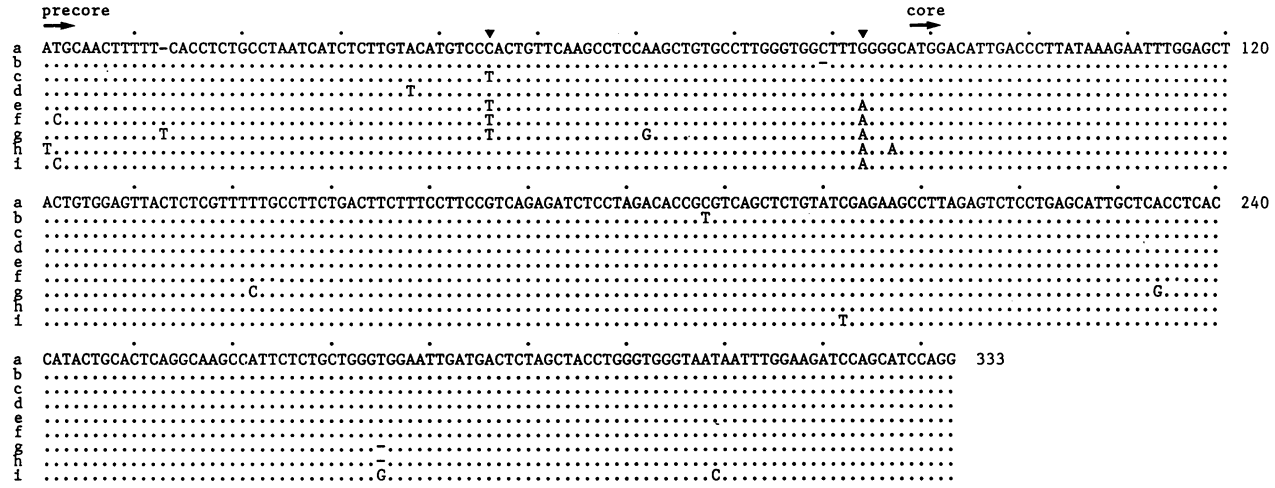

FIG. 3. Sequence of the precore-core gene from PCR clones of a patient. The complete sequence of clone a is given at the top, and changed nucleotide sequences in the other clones are written out. A dash denotes lack of a nucleotide at this position. Clones a, b, c, and d did not contain a UAG nonsense mutation at codon 28, while codons e, f, g, h, and i did. In clones e, f, and g, the G-to-A transition generating the nonsense mutation (▼) was associated with covariation of nucleotide 3 of codon 15 (▼). Other HBe-abolishing mutations: deletion of nucleotide 79 in clone b; mutated initiation codon in clones f, h, and i; and insertion of a U in the five-U stretch in clone g.

47) were infected with the CCC variants, while all of the Chinese samples were infected with the CCU variants. Hybridization with oligoprobes M0 and M1-M2 indicated that most of these HBeAg-positive samples were infected with wild-type virus, as expected, although a few samples contained both wild-type and mutant viruses. Interestingly, four of the five French samples containing the codon 28 mutation contained the CCU rather than the CCC variants, despite the higher prevalence of the CCC variants in the total samples.

Only the CCU variants were found as HBe$^-$ mutants. A total of 55 anti-HBe-positive French samples were used for amplification of the precore-core gene, and clearly visible PCR products were seen in 28 of them. Hybridization results are illustrated in Fig. 2B and summarized in Table 4. Only 1 of the 28 samples was infected with a CCC variant, which did not contain the codon 28 mutation (Fig. 2B, lane 9). DNA sequencing failed to reveal other HBe-abolishing mutations. The remaining 27 samples were infected with CCU variants, all of which contained a UAG mutation at codon 28 (Table 4). To exclude the possibility that HBe$^-$ mutants of genotype A were present in the remaining anti-HBe-positive samples in minute quantities, these samples were subjected to double PCR amplification with nested primers. Seventeen additional samples turned out positive. Fourteen samples were infected with the CCU variants, all of which contained the codon 28 mutation (some in association with wild-type virus; Table 4). Three samples were infected with the CCC variants, and all contained only wild-type virus. Thus, the codon 28 mutation was tightly associated with the CCU variants, whether in HBeAg-positive or anti-HBe-positive samples (Tables 3 and 4). On the other hand, the wild-type viruses in four anti-HBe-positive samples were all CCC variants (Table 4). Apparently contradicting this rule was a single CCC variant from an HBeAg-positive patient, in which both wild-type and mutant viruses were found (Table 3). Detailed analysis of this case as shown below revealed that in fact it obeyed this rule.

Sequence covariation at codons 15 and 28 in a genotype A isolate. Sample FB-5 displayed a weak hybridization signal with probe M1-M2 (10 to 20% of the signal obtained with probe M0; data not shown). The specificity of hybridization with probe M1-M2 was confirmed by cloning of the PCR product into M13mp19 and the hybridization of M13 clones with probes M0 and M1-M2. Eight of the 45 clones examined turned out positive with probe M1-M2. DNA sequencing of several M0$^+$ clones confirmed this isolate as genotype A (Fig. 3, a, b, and d). Sequencing of three M1-M2$^+$ clones revealed a UAG stop codon at precore codon 28 (Fig. 3, e, f, and g). Interestingly, all had a codon 15 sequence changed from CCC to CCU. (The facts that these sequences differ from that of the CCU probe by a single nucleotide at codon 13, ACA versus UCA and that stringent washing conditions were adopted may explain why the CCU variant was not detected by the probe.) Otherwise, the entire 333-bp sequence was nearly identical between the wild-type and mutated virus populations. Some of the M0$^+$ (b) and M1-M2$^+$ (f and g) clones had a mutated precore initiation codon or frameshift mutation, which should also prevent expression of HBeAg.

To investigate the association of sequence changes at codons 15 and 28 further, hybridization experiments involving large numbers of clones were conducted. PCR products from this sample and four previous serial samples collected from this patient over a 1-year period were cloned into pUC18, and pUC clones were hybridized with probes CCC, CCU, M0, and M1-M2 (less stringent hybridization conditions were adopted for probe CCU). As a result, 591 clones belonged to the CCC variant, of which 585 were in the form of wild-type virus and the remaining 6 were codon 28 mutants. Eighty-four clones belonged to the CCU variant, of which 71 were codon 28 mutants and the remainder were wild-type virus. This result confirmed the sequence covariation at codons 15 and 28 in most of the mutant clones. One CCU$^+$ M0$^+$ clone and two CCC$^+$ M1-M2$^+$ clones were sequenced (Fig. 3, c, h, and i). The latter two clones contained a UAG mutation at precore codon 28, with or without a concomitant GGC-to-GAC change at codon 29.

HBe$^-$ mutants and viral genotypes. The rare prevalence of the CCC variants in the HBe$^-$ mutant populations did not

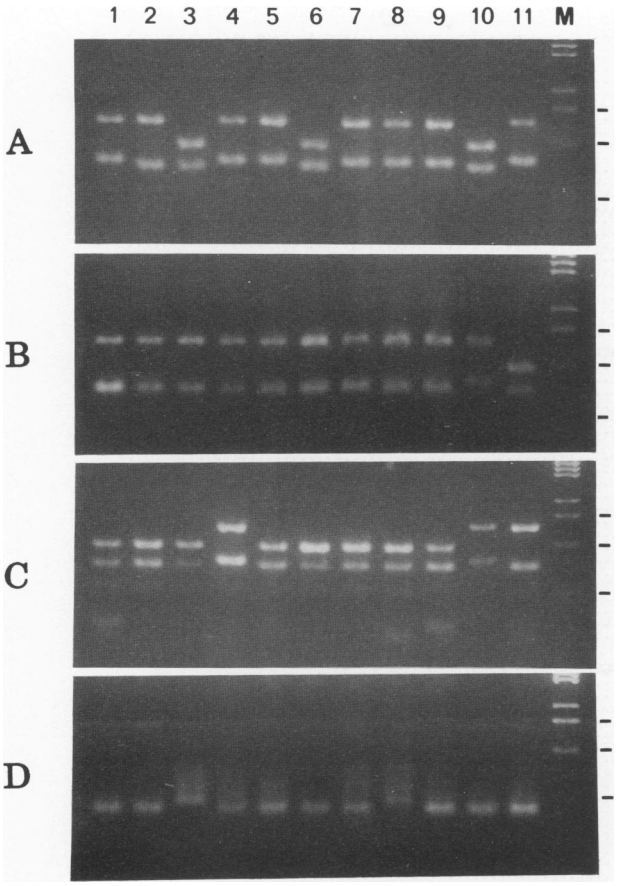

FIG. 4. Genotyping of PCR products of the pre-S region and the core gene by size polymorphism. In panels A to C, the pre-S1 region and the core gene were amplified by primer pairs pS4-pS5 and C5-C6, respectively. The expected amplification products were 156 and 105 bp for genotype A, 156 and 99 bp for genotypes B and C, and 123 and 99 bp for genotypes D and E. Panels: A, 11 HBeAg-positive French samples; B, 9 HBeAg-positive Chinese samples (lanes 1 to 9) and controls of genotypes A (lane 10) and D (lane 11); C, 9 anti-HBe-positive French samples (lanes 1 to 9) and controls of genotypes A (lane 10) and C (lane 11); D, 11 anti-HBe-positive samples amplified by nested PCR with internal primer pair C3-C4. The amplification products were expected to be 76 bp for genotype A and 70 bp for the other genotypes. Lanes 3 and 8 had the genotype A pattern, and the others had a non-A genotype. Lane M, pGEM DNA markers (Promega). The three bands indicated correspond to 179, 126, and 75 bp, respectively.

necessarily indicate a low occurrence of genotype A mutants, since some genotype A isolates may undergo a CCC-to-CCU change at codon 15, as exemplified above. The genotypes of the HBV isolates studied were therefore examined by using several independent methods: hybridization of the PCR products with genotype A-specific oligoprobe C2, sequence analysis of the PCR products, and typing by analysis of DNA size polymorphism in the core gene and the pre-S region (see Materials and Methods). Patterns of hybridization of some clinical samples with probe C2 are shown in Fig. 2, and results are summarized in Table 3 for HBe-positive samples and in Table 4 for anti-HBe-positive samples. All of the C2-positive samples were positive for a CCC sequence at codon 15, and vice versa. Size patterns of the PCR products of the core gene and the pre-S region from some samples are shown in Fig. 4, and results are summarized in Table 5. These data suggested that all of the CCC variants belonged to genotype A, while the CCU variants belonged to either genotype D-E (French samples) or B-C (Chinese samples). Sequencing of the 0.4-kb precore-core gene in some samples (Fig. 5) not only confirmed the association between CCC variants and genotype A (FA1-6) but also determined the exact genotypes present in the CCU variants. The French CCU variants sequenced were all contributed by genotype D (FD1-6 from HBeAg-positive samples, FDa-h from anti-HBe-positive samples). The Chinese CCU variants were chiefly infected with genotype C (CC1-5) and occasionally infected with genotype B (CB1). Thus, emergence of HBe^- mutants of genotype A through a CCC-to-CCU change at precore codon 15 was a rare event.

TABLE 5. DNA fragment length polymorphism in the pre-S region and the core gene: correlation with sequence polymorphism at precore codon 15[a]

Samples	No. of samples showing length patterns of genotype:			
	A	B-C	D-E	Non-A
CCC at codon 15				
French HBeAg positive	17	0	0	NA
French anti-e positive, 1 PCR	1	0	0	NA
French anti-e positive, 2 PCR	3	0	0	NA
CCU at codon 15				
French HBeAg positive	0	1	13	1
Chinese HBeAg positive	0	18	0	NA
French anti-e positive, 1 PCR	0	0	27	NA
French anti-e positive, 2 PCR	0	NA	NA	14

[a] Abbreviations: NA, nonapplicable; 1 PCR, single PCR; 2 PCR, nested PCR.

DISCUSSION

The present work was stimulated by our recent study demonstrating the influence of base pairing in a stem region of the viral pregenome encapsidation signal on the mutational patterns of two natural HBe^- mutants. Both mutants had an additional nucleotide change(s), other than that needed to abolish HBe expression, which would restore base pairing and was found to be required for efficient packaging of viral pregenomic RNA (28). Additional mutagenesis work confirmed the presence of the proposed stem structure. These findings led us to speculate about whether or not genotype A of HBV, with a precore codon 15 sequence of CCC rather than CCU, could develop a replication-competent UAG or UGA mutation at codon 28. We were encouraged by the fact that in none of the published sequences of HBe^- mutants was the codon 28 UAG stop codon associated with a codon 15 sequence of CCC.

Molecular epidemiological studies were carried out to compare the prevalence of genotype A in two parts of the world and between HBeAg-positive and anti-HBe-positive patients. We used genotyping methods that included detection of DNA length polymorphism and hybridization with oligoprobes. Oligoprobes CCC and C2 were highly efficient in the identification of genotype A isolates, as evidenced by the concordant results obtained with the two probes, as well as confirmation by typing through length polymorphism and by DNA sequencing. This is certainly related to the extremely high degree of sequence conservation among iso-

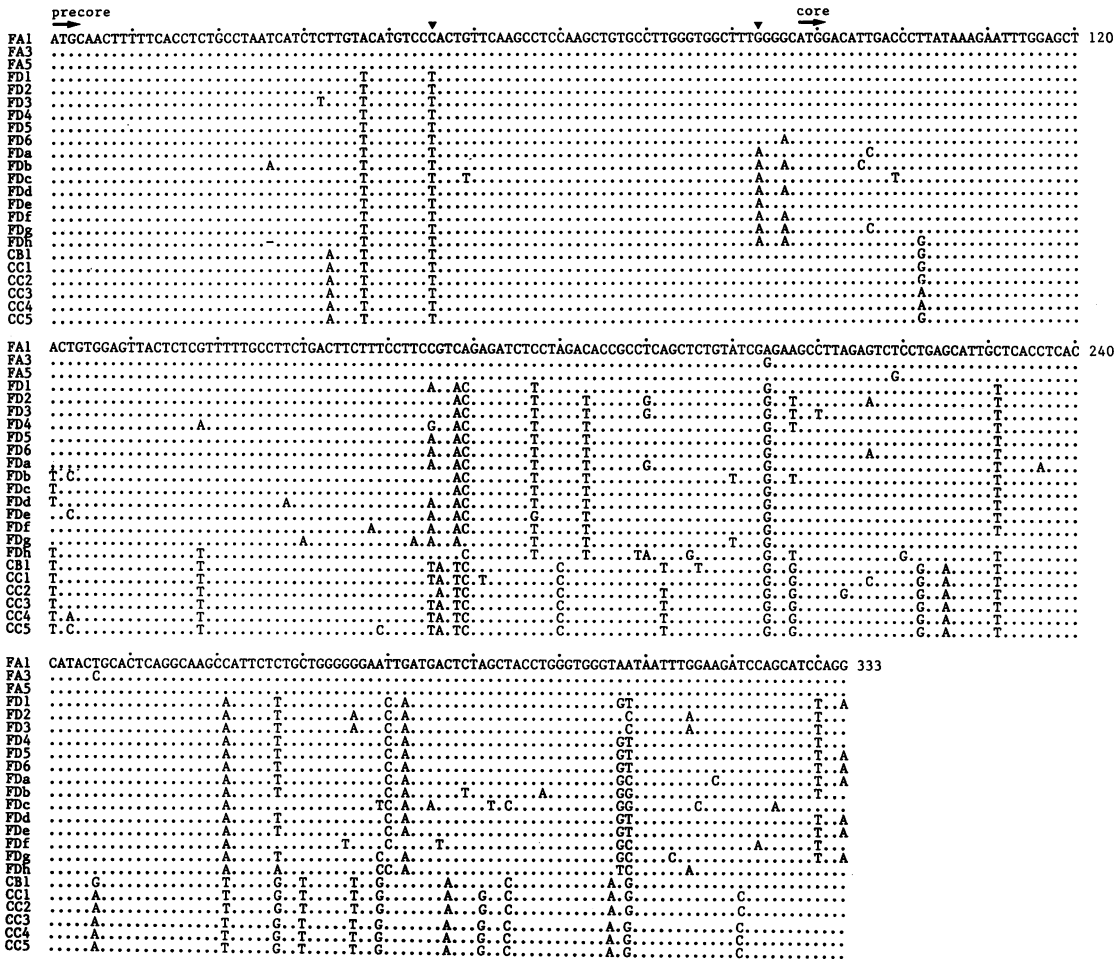

FIG. 5. A 333-bp sequence of the precore-core gene from 26 clinical samples. Two or more clones were sequenced for each sample. The sequences were from HBeAg-positive French patients as genotype A (FA1-6) or D (FD1-6), from anti-HBe-positive French patients as genotype D (FDa-h), and from HBeAg-positive Chinese patients as genotype B (CB1) or C (CC1-5). The complete sequence of FA1 is shown at the top. The initiation codons of the precore region and the core gene are marked by arrows. The two triangles indicate nucleotide 3 of precore codon 15 and nucleotide 2 of codon 28. The sequences of isolates FA2, FA4, and FA6 are identical to that of FA1 and are not shown. For other isolates, only changed nucleotides are shown. A deleted nucleotide is denoted as a dash. Note that FDh contains a deletion at nucleotide 26 and the common nonsense mutation at codon 28 (although it was no longer in frame following the deletion).

lates of this genotype (Fig. 5, FA1-6; also, compare three completely sequenced genomes of genotype A in references 1, 20, and 34). The biological implication of this phenomenon remains unknown. From sequence profiles shown in Fig. 5, it appears that type-specific oligoprobes for other genotypes, especially B-C, are also possible. As a complement to the oligoprobe method, we developed a typing method by analyzing DNA size polymorphisms in the pre-S region and the core gene. We believe that this typing method is superior to the oligoprobe method in both simplicity and reliability. While a single nucleotide change (which changes the hybridization pattern when located in regions covered by the oligoprobes) during prolonged infection is possible, insertion or deletion within the polymorphic site should be extremely rare. Analysis of 95 samples in this study indicated that size polymorphism segregates nicely with genotypes. Thus, the small deletion in the pre-S1 region of one genotype A isolate (20) is quite uncommon.

Genotypes A and D were found to be the predominant HBV genotypes in the HBeAg-positive French serum samples examined, while genotypes B and C were the prevalent HBV genotypes in the Chinese samples. Genotype A was not found in any of the 18 Chinese samples tested. Previously, the three completely sequenced genotype A genomes originated from the United States (20, 34) and Germany (1), respectively, and HBV isolates with sequence profiles of genotype A were found in Germany and France (11, 31). In contrast, genotype A was not found in any of a dozen Asian HBV genomes isolated from Japan (see reference 19; 18, 25), Indonesia (19), South Korea (22), and the People's Republic of China (10, 15, 21) which belong to genotype B or C. Therefore, genotypes A and D seem to represent the major HBV genotypes in Western Europe and North America, while genotypes B and C are mainly Asian genotypes.

Remarkably, genotype A was not found in any of the 41 isolates of HBe^- mutants obtained from serum samples of chronic hepatitis patients positive for anti-HBe; these isolates were chiefly or entirely of genotype D. Although we did not have the chance to check for the possible persistence of genotype A-related HBe^- mutants in the liver, our data

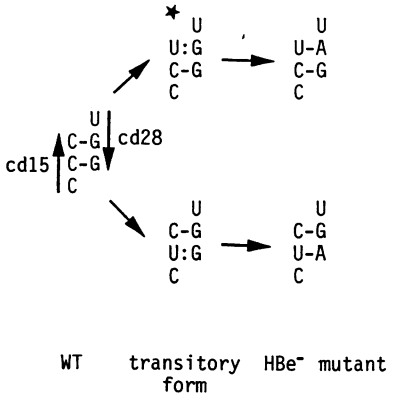

FIG. 6. Possible pathways in the emergence of HBe$^-$ precore mutants of genotype A. Shown are precore codons 15 and 28 forming part of the stem region of the pregenome encapsidation signal. The first step could be a C-to-U transition at nucleotide 3 or 2 of codon 15, converting a C-G pair into a wobble U-G pair. The second step is a G-to-A transition at nucleotide 2 or 3 of codon 28 that generates a UAG or UGA stop codon and converts the wobble pair into another Watson-Crick pair (U-A). Note that one of the proposed transitory forms (star) is in fact present in the wild-type (WT) viruses of all other genotypes.

suggest that genotype A rarely circulates as HBe$^-$ precore mutants. This is the first example of genotype-associated difference in biological properties of HBV. Since genotype A is prevalent in regions where HBe$^-$ mutants are seldom detected (Western Europe and North America) but scarcely found in regions where such mutants prevail (the Far East), our findings may provide an important clue to the uneven geographic distribution of HBe$^-$ mutants. Examination of the distribution of genotype A in additional countries with high (Italy and Greece) and low (United States) prevalences of HBe$^-$ mutants is crucial to test this hypothesis.

The base-pairing requirement of the ε signal may help to explain the rare detection of genotype A HBe$^-$ mutants in serum, as suggested by two lines of evidence. By sequentially mutating an HBV genome of a non-A genotype, we found that a CCC sequence at codon 15 markedly reduced the packaging and replication capacities of HBVs with a nonsense mutation at codon 28, especially with the UAG mutation. Detailed analysis of an HBeAg-positive case harboring both codon 28 mutants and wild-type virus clearly pinpointed the significance of a simultaneous CCC-to-CCU change at codon 15 for the codon 28 mutations. Detection of a few clones without an altered codon 15 sequence may be explained by the low packaging efficiency of such mutant genomes (as suggested by transfection experiments) or by artifactual "shuffle clones" produced by PCR. The significance of sequence covariation at codon 15 was also implied by the work of Blum et al. (1), Tran et al. (31), Gunter et al. (11), and (possibly) Carman et al. (5). Thus, genotype A can develop either a UAG or a UGA nonsense mutation at precore codon 28 which is tightly associated with a simultaneous sequence change at codon 15 to CCU or CUC, respectively. It is conceivable that in both circumstances the sequence change at codon 15 occurs prior to the change at codon 28, so that a transitory state with a wobble pair in the stem of the ε signal is produced (Fig. 6). On the basis of the base-pairing requirement of the ε signal, we also predict that genotypes B-E would rarely develop a UGA nonsense mutation at codon 28.

If the requirement for a simultaneous sequence change at codon 15 explains the rare detection of HBe$^-$ mutants of genotype A in serum, then we should be able to detect such mutants in the livers of some patients as unpackaged mRNAs (supposing that a G-to-A transition at nucleotide 2 of codon 28 arises with similar frequency in genotype A). Since a CCC-to-CCU change at precore codon 15 neither changes the encoded amino acid nor disrupts base pairing of the ε signal, it remains to determine the proportion of the CCU variant during the HBeAg-positive phase of replication and the efficiency with which the UAG mutation at codon 28 is generated from such a variant. A decreased chance of a nonsense mutation at codon 28 might promote other types of HBe-abolishing mutations or cryptic nonsense mutations at other amino acid codons (30). All of these issues should be tested before we can accept the significance of a single base pair in the ε signal as the key determinant of HBe$^-$ precore mutation.

ACKNOWLEDGMENTS

We thank C. Pichoud for synthesis of the oligonucleotides used in this study.

This work was supported by grant CII*920058 from the Commission of The European Community.

REFERENCES

1. **Blum, H. E., J. Liang, E. Galun, and J. R. Wands.** 1991. Persistence of hepatitis B viral DNA after serological recovery from hepatitis B virus infection. Hepatology **14:**56–63.
2. **Bonino, F., M. Brunetto, M. Rizzetto, and H. Will.** 1991. Hepatitis B virus unable to secrete e antigen. Gastroenterology **100:**1138–1141.
3. **Brown, J. L., W. F. Carman, and H. C. Thomas.** 1992. The clinical significance of molecular variation within the hepatitis B virus genome. Hepatology **15:**144–148.
4. **Brunetto, M. R., M. M. Giarin, F. Oliveri, E. Chiaberge, M. Baldi, A. Alfarano, A. Serra, G. Saracco, G. Verme, H. Will, and F. Bonino.** 1991. Wild-type and e antigen-minus hepatitis B viruses and course of chronic hepatitis. Proc. Natl. Acad. Sci. USA **88:**4186–4190.
5. **Carman, W. F., M. Ferrao, A. S. F. Lok, O. C. K. Ma, C. L. Lai, and C. Thomas.** 1992. Precore sequence variation in Chinese isolates of hepatitis B virus. J. Infect. Dis. **165:**127–133.
6. **Chang, C., G. Enders, R. Sprengel, N. Peters, H. E. Varmus, and D. Ganem.** 1987. Expression of the precore region of an avian hepatitis B virus is not required for viral replication. J. Virol. **61:**3322–3325.
7. **Chen, H.-S., M. C. Kew, W. E. Hornbuckle, B. C. Tennant, P. J. Cote, J. L. Gerin, R. H. Purcell, and R. H. Miller.** 1992. The precore gene of the woodchuck hepatitis virus genome is not essential for viral replication in the natural host. J. Virol. **66:**5682–5684.
8. **Chen, Z.-P., Y.-M. Wen, J.-R. Guo, D.-F. Wan, and G.-B. Yao.** 1990. Detection of HBV gene variation in hepatocellular carcinoma. Chin. J. Virol. **6:**192–196.
9. **Courouce-Pauty, A. M., A. Plancon, and J. P. Soulier.** 1983. Distribution of HBsAg subtypes in the world. Vox Sang. **44:**197–211.
10. **Gan, R. B., M. J. Chu, L. P. Shen, S. W. Qian, and Z. P. Li.** 1987. The complete nucleotide sequence of the adr subtype hepatitis B virus DNA in pADR-1. Sci. Sin. Ser. B **30:**507–521.
11. **Gunter, S., H. Meisel, A. Reip, S. Miska, D. H. Kruger, and H. Will.** 1992. Frequent and rapid emergence of mutated pre-C sequences in HBV from e-antigen positive carriers who seroconvert to anti-HBe during interferon treatment. Virology **187:** 271–279.
12. **Junker-Niepmann, M., R. Bartenschlager, and H. Schaller.** 1990. A short cis-acting sequence is required for hepatitis B virus pregenome encapsidation and sufficient for packaging of foreign RNA. EMBO J. **9:**3389–3396.

13. **Li, J.-S., S.-P. Tong, L. Vitvitski, F. Zoulim, and C. Trépo.** 1990. Rapid detection and further characterization of infection with hepatitis B virus variants containing a stop codon in the distal pre-C region. J. Gen. Virol. **71:**1993–1998.
14. **Ljunggren, K., and A. H. Kidd.** 1991. Enzymatic amplification and sequence analysis of precore/core DNA in HBsAg-positive patients. J. Med. Virol. **34:**179–183.
15. **Loncarevic, I. F., H. Zentgraf, and C. H. Schroder.** 1990. Sequence of a replication competent hepatitis B virus genome with a preX open reading frame. Nucleic Acids Res. **18:**4940.
16. **Manzin, A., S. Menzo, P. Bagnarelli, P. V. Varaldo, I. Bearzi, G. Carloni, F. Galibert, and M. Clementi.** 1992. Sequence analysis of the hepatitis B virus pre-C region in hepatocellular carcinoma (HCC) and nontumoral liver tissues from HCC patients. Virology **188:**890–895.
17. **Milich, D. R., J. E. Jones, J. L. Hughes, J. Price, A. K. Raney, and A. McLachlan.** 1990. Is a function of the secreted hepatitis B e antigen to induce immunologic tolerance in utero? Proc. Natl. Acad. Sci. USA **87:**6599–6603.
18. **Ogata, N., R. H. Miller, K. G. Ishak, and R. H. Purcell.** 1993. The complete nucleotide sequence of a pre-core mutant of hepatitis B virus implicated in fulminant hepatitis and its biological characterization in chimpanzee. Virology **194:**263–276.
19. **Okamoto, H., F. Tsuda, H. Sakugawa, R. Sastrosoewignjo, M. Imai, Y. Miyakawa, and M. Mayumi.** 1987. Typing hepatitis B virus by homology in nucleotide sequence: comparison of surface antigen subtypes. J. Gen. Virol. **69:**2573–2583.
20. **Ono, Y., H. Onda, R. Sasada, K. Igarashi, Y. Sugino, and K. Nishioka.** 1983. The complete nucleotide sequences of the cloned hepatitis B virus DNA: subtype adr and adw. Nucleic Acids Res. **11:**1747–1757.
21. **Qi, Z. H., J. Yang, W. J. Xiong, J. G. Yuan, S. Song, C. Q. Xue, and L. W. Cai.** 1989. Determination of the complete nucleotide sequence of HBV adr NC-1 DNA. Sci. Sin. Ser. B **32:**1082–1086.
22. **Rho, H. M., K. Kim, S. W. Hyun, and Y. S. Kim.** 1989. The nucleotide sequence and reading frames of a mutant hepatitis B virus subtype adr. Nucleic Acids Res. **17:**2124.
23. **Schlicht, H. J., J. Salfeld, and H. Schaller.** 1987. The duck hepatitis B virus pre-C region encodes a signal sequence which is essential for synthesis and secretion of processed core proteins but not for virus formation. J. Virol. **61:**3701–3709.
24. **Shafritz, D. A.** 1992. Variants of hepatitis B virus associated with fulminant liver disease. N. Engl. J. Med. **324:**1737–1739.
25. **Takemura, F., T. Ishii, N. Fujii, and T. Uchida.** 1990. Complete nucleotide sequence of hepatitis B virus. Nucleic Acids Res. **18:**4587.
26. **Tong, S.-P., B. Brotman, J.-S. Li, L. Vitvitski, D. Pascal, A. M. Prince, and C. Trépo.** 1991. In vitro and in vivo replication capacity of the precore region defective hepatitis B virus variants. J. Hepatol. **13**(Suppl. 4):S68–S73.
27. **Tong, S.-P., C. Diot, P. Gripon, J.-S. Li, L. Vitvitski, C. Trépo, and C. Guguen-Guillouzo.** 1991. In vitro replication competence of a cloned hepatitis B virus variant with a nonsense mutation in the distal pre-C region. Virology **181:**733–737.
28. **Tong, S.-P., J.-S. Li, L. Vitvitski, A. Kay, and C. Trépo.** 1993. Evidence for a base-paired region of hepatitis B virus pregenome encapsidation signal which influences the patterns of precore mutations abolishing HBe protein expression. J. Virol. **67:**5651–5655.
29. **Tong, S.-P., J.-S. Li, L. Vitvitski, and C. Trépo.** 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.
30. **Tong, S.-P., J.-S. Li, L. Vitvitski, and C. Trépo.** 1992. Replication capacities of natural and artificial precore stop codon mutants of hepatitis B virus: relevance of pregenome encapsidation signal. Virology **191:**237–245.
31. **Tran, A., D. Kremsdorf, F. Capel, C. Housset, C. Dauguet, M.-A. Petit, and C. Bréchot.** 1991. Emergence of and takeover by hepatitis B virus (HBV) with rearrangements in the pre-S/S and pre-C/C genes during chronic HBV infection. J. Virol. **65:**3566–3574.
32. **Tur-Kaspa, R., A. Klein, and S. Aharonson.** 1992. Hepatitis B virus precore mutants are identical in carriers from various ethnic origins and are associated with a range of liver disease severity. Hepatology **16:**1338–1342.
33. **Ulrich, P. P., R. A. Bhat, M. R. Brunetto, F. Bonino, and G. N. Vyas.** 1990. A precore-defective mutant of hepatitis B virus associated with e antigen-negative chronic liver disease. J. Med. Virol. **32:**109–118.
34. **Valenzuela, P., M. Quiroga, J. Zaldivar, P. Gray, and W. J. Rutter.** 1980. The nucleotide sequence of the hepatitis B viral genome and the identification of the major viral genes, p. 57–70. *In* B. N. Fields, R. Jaenisch, and C. F. Fox (ed.), Animal virus genetics. Academic Press, Inc., New York.
35. **Vaudin, M., A. J. Wolstenholme, K. N. Tsiquaye, A. J. Zuckerman, and T. J. Harrison.** 1988. The complete nucleotide sequence of the genome of a hepatitis B virus isolated from a naturally infected chimpanzee. J. Gen. Virol. **69:**1383–1389.
36. **Wands, J. R., T. J. Liang, H. E. Blum, and D. A. Shafritz.** 1992. The molecular pathogenesis of liver disease during persistent hepatitis B infection. Semin. Liver Dis. **12:**250–258.