Novel and Frequent Mutations of Hepatitis B Virus Coincide with a Major Histocompatibility Complex Class I-Restricted T-Cell Epitope of the Surface Antigen†

PEI-CHING TAI,† DANA BANIK,† GEN-IU LIN,† SHAN PAI,† KAN PAI,† MIN-HUI LIN,† GBO YUOH,† SHAOLI CHE,† SUSAN H. HSU,† TSE-CHING CHEN,† TSENG-TONG KUO,† CHUE-SHUE LEE,† CIAU-SIUNG YANG,† and CHIAHO SHIH†

Departments of Pathology and of Microbiology and Immunology, WHO Collaborating Center for Tropical Diseases, University of Texas Medical Branch, Galveston, Texas 77555-0609; Musser Blood Center, American Red Cross, Philadelphia, Pennsylvania 19123-3592; and Department of Pathology, Chang Gung College of Medicine and Technology, Kwei San, Tao Yuan, and National Taiwan University Hospital, Taipei, Taiwan

Received 24 December 1996/Accepted 10 March 1997

We examined the full-length hepatitis B virus (HBV) envelope (surface antigen or HBV small surface antigen [HBsAg]) sequences of 12 different liver samples from 10 different hepatoma-containing chronic carriers. Surprisingly, novel and frequent mutations occurred predominantly at amino acids 40 and 47 of HBsAg, in addition to within a known protective B-cell epitope (so-called group a determinant of HBsAg 124-148). Approximately 58% of chronic carriers contain mutations at the group a determinant. The mutation frequency at the hotspot codons 40 and 47 is approximately 83%, 1 order of magnitude higher than at the known polymorphic codons 122 and 160, which is approximately 4%. This new mutational domain is found to coincide with a major histocompatibility complex class I-restricted T-cell epitope. The potential biological significance of this novel mutation in the immunopathogenesis of HBV chronic carriers is discussed.

Human hepatitis B virus (HBV) small surface antigen (HBsAg), also known as Australia antigen, was discovered by B. S. Blumberg three decades ago (2). HBsAg is a 24- to 27-kDa protein, depending on the extent of glycosylation (reviewed in reference 30). Subviral 22-nm HBsAg particles have been developed as an effective vaccine for prevention of viral hepatitis B (9, 21, 29). The protective immunity of HBsAg vaccination is associated with a neutralizing antibody (anti-HBs) specific for the group a determinant (amino acids 124 to 148), which appears to be a conserved immunodominant B-cell epitope with a double-loop conformation (32). Production of neutralizing antibody to this determinant provides cross-protection against all HBV subtypes. Clearance of HBV infection is almost always associated with seroconversion from HBsAg to anti-HBs positivity. However, in some sporadic cases of vaccine failure, both HBsAg and anti-HBs can be found in the same patient simultaneously, indicating unsuccessful antibody neutralization as well as viral multiplication in the presence of humoral immunity to HBsAg. Subsequent studies revealed missense mutations occurring within the protective a epitope of HBsAg in patients from Europe, Asia, and Africa (5, 11, 34, 35). The mutated HBsAg from these patients cannot be recognized in vitro by monoclonal and polyclonal antibodies specific for the wild-type group a determinant (4, 33). Similar observations have been made with liver transplantation patients who received anti-HBs immunoglobulin treatment (14). The mutations observed in HBsAg should also be present in the medium and large surface antigens. These two envelope proteins are structurally related to HBsAg by having the same carboxyl-terminal domain and are less abundant than HBsAg (30). So far, the prevalence of similar escape mutations within the a epitope in natural populations, which have not received active or passive HBV immunization, remains unknown.

Chronic infection with HBV frequently leads to the development of cirrhosis and liver cancer (25). In contrast, self-limited acute hepatitis B is rarely associated with liver cancer. The mechanism of HBV chronicity remains to be elucidated. Although the lack of adequate humoral immunity, as described above, clearly could contribute to the establishment of HBV chronic infection, the dual roles of cellular immunity in protection and pathogenesis of HBV infection are less clear. It is generally believed that viral immune escape mutations occur within major histocompatibility complex (MHC) class I-restricted cytotoxic T-cell epitopes (20). Previously, we identified frequent mutations of HBV core antigen (HBcAg) which, quite unexpectedly, accumulated within MHC class II- restricted T-cell epitopes (10). These results led to the hypothesis that some of these mutations could alter the recognition of HBcAg by CD4+ T cells and contribute to the maintenance of a chronic carrier state. It remains unknown whether the same phenomenon of frequent mutations within T-cell epitopes could also occur in other HBV-encoded proteins, such as HBsAg.

To investigate the putative immune escape mutations and the mechanism of chronicity of HBV infection, we examined systematically the entire HBsAg sequences for 10 different HBV-related hepatoma patients. Our studies uncovered a novel mutational domain of HBsAg which coincides with an MHC class I-restricted T-cell epitope. The implications of this result for the maintenance of persistent infection are discussed.

Patients and molecular typing of HLA-A locus. The clinical data for patients are summarized in Table 1. Both hepatomas (designated by the prefix T) and their cognate nontumor liver samples (designated by the prefix N) were included for screening. Samples T117 and N117, as well as samples T109 and
patient NCG35. Profiles as the replicative HBV DNA (10). Second, generally, the selective pressure by accumulating mutations (10, 25). However, in HBV is a dead-end product and not able to respond to selective pressure. In this study, we have previously screened for patterns characteristic of HBV hepatomas containing unintegrated replicative HBV DNAs, and maintain only integrated HBV DNAs (23–25, 30). To search for HBV samples of chronic hepatitis B patients (3, 12, 16). Therefore, several different research groups using HBV DNA from serum samples of hepatocellular carcinoma (HCC) patients (10) have recently been confirmed by using HBV DNA from liver samples of hepatocellular carcinoma (10) with the pGEM-T vector (Promega Biotech, Madison, Wis.).

Liver versus serum samples. The HBV DNAs used in this study were isolated from liver samples. Although serum samples from these patients are not available for further comparison of sequence data with liver samples, there are reasons to believe that similar results will be obtained. First, liver is the major, if not the only, source of HBV production. Second, several novel mutational domains of HBcAg uncovered by using HBV DNA from liver samples of hepatocellular carcinoma (HCC) patients (10) have recently been confirmed by several different research groups using HBV DNA from serum samples of chronic hepatitis B patients (3, 12, 16). Therefore, we assume that there is also no major difference in the studies of HBsAg mutations using either serum or liver samples. Integrated versus replicative forms. Most hepatomas contain only integrated HBV DNAs (23–25, 30). To search for hepatomas containing unintegrated replicative HBV DNAs, we have previously screened for patterns characteristic of HBV replicative intermediates from approximately 100 different hepatoma samples via Southern blot analysis (10). Only 10 patients contain replicative intermediates of HBV in their liver samples. We focused our current study on these 10 patients for the following reasons. First, our previous studies indicate that integrated HBV DNAs contain few mutations (10). This is in part due to the fact that most integrations occur as an early event when the preintegration substrate of the relaxed circle form is abundant (24), in part due to the fact that integrated HBV is a dead-end product and not able to respond to selective pressure by accumulating mutations (10, 25). However, in those less frequent cases in which mutations were identified in the integrated HBV DNA, they exhibited the same mutational profiles as the replicative HBV DNA (10). Second, generally, an HCC contains an average of several copies of integrated HBV DNA per cell (23, 30). Based on our estimate of the signal intensity of Southern blot data (data not shown), at least several dozen copies of HBV replicative intermediates per cell are present in liver samples of these 10 patients. Therefore, if any integrated copies of HBV DNA happened to be present in some of the liver samples of these 10 patients, the PCR product should reflect the majority of replicative forms of HBV DNA. In summary, our studies with these 10 patients represent a result from approximately 100 HCC patients.

Amplification and cloning of HBsAg sequences via PCR. Total intracellular DNAs for Southern blot, HLA typing, and PCR analyses were prepared as described previously (24). One microgram of the extracted DNA sample was used as a template for the amplification of the surface antigen region of HBV DNA. The forward and reverse primers used for amplifying the S region are from nucleotides (nt) 2466 to 2481 (5′ CTT TGG ACT CAT AAG GT 3′) and nt 988 to 972 (5′ ACT TTC CAA TCA ATA GG 3′). The amplification reaction was performed as described previously (10) with the following modification. An initial denaturation of the DNA sample was at 94°C for 20 s followed by 25 cycles of denaturation for 1 s at 94°C, annealing for 1 s at 41°C, and extension for 40 s at 70°C in a thermocycler (Idaho Tech Co.). The PCR-amplified DNA of 1.7 kb in size was purified through a 1.0% low-melting-point agarose gel. DNA was further purified via phenol-chloroform extraction and ethanol precipitation. Standard subcloning procedures were conducted with the pGEM-T vector (Promega Co., Madison, Wis.).

### TABLE 1. Clinical data from patients with HCC

<table>
<thead>
<tr>
<th>Patient</th>
<th>Age (yr)</th>
<th>Sex</th>
<th>HBsAg</th>
<th>eAg</th>
<th>ASTa</th>
<th>ALTb</th>
<th>Tumor size (cm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>T4</td>
<td>43</td>
<td>M</td>
<td>+</td>
<td>-</td>
<td>36</td>
<td>31</td>
<td>6.5 × 5 × 5</td>
</tr>
<tr>
<td>T8</td>
<td>49</td>
<td>M</td>
<td>+</td>
<td>NA</td>
<td>30</td>
<td>28</td>
<td>18 × 17 × 14</td>
</tr>
<tr>
<td>T32</td>
<td>67</td>
<td>M</td>
<td>+</td>
<td>-</td>
<td>35</td>
<td>30</td>
<td>2.5 × 2.3</td>
</tr>
<tr>
<td>N79</td>
<td>56</td>
<td>M</td>
<td>NA</td>
<td>NA</td>
<td>30</td>
<td>32</td>
<td>13 × 10 × 3</td>
</tr>
<tr>
<td>N85</td>
<td>60</td>
<td>M</td>
<td>+</td>
<td>NA</td>
<td>25</td>
<td>24</td>
<td>8 × 6 × 5</td>
</tr>
<tr>
<td>N109</td>
<td>42</td>
<td>M</td>
<td>+</td>
<td>NA</td>
<td>40</td>
<td>35</td>
<td>5 × 4 × 3</td>
</tr>
<tr>
<td>N117</td>
<td>47</td>
<td>M</td>
<td>+</td>
<td>-</td>
<td>28</td>
<td>25</td>
<td>10 × 6 × 5</td>
</tr>
<tr>
<td>NCG13</td>
<td>62</td>
<td>M</td>
<td>+</td>
<td>NA</td>
<td>115</td>
<td>54</td>
<td>20 × 15 × 8</td>
</tr>
<tr>
<td>NCG22</td>
<td>61</td>
<td>M</td>
<td>NA</td>
<td>NA</td>
<td>31</td>
<td>22</td>
<td>5 × 5 × 4</td>
</tr>
<tr>
<td>NCG35</td>
<td>78</td>
<td>F</td>
<td>+</td>
<td>+</td>
<td>42</td>
<td>30</td>
<td>3.29</td>
</tr>
</tbody>
</table>

*NA, information not available; +, above the cutoff; −, below the cutoff.

### TABLE 2. Occurrence of mutations within two different mutational domains in HBsAg and HLA-A locus alleles of patients

<table>
<thead>
<tr>
<th>Liver sample from patient with HCC</th>
<th>HLA-A locus allele</th>
<th>Value for domainc</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Amino acids</td>
<td>Amino acids</td>
</tr>
<tr>
<td></td>
<td>26-51</td>
<td>124-148</td>
</tr>
<tr>
<td>T4</td>
<td>3303/—b</td>
<td>+</td>
</tr>
<tr>
<td>T8</td>
<td>0206/1101</td>
<td>+</td>
</tr>
<tr>
<td>T32</td>
<td>3303/—b</td>
<td>−</td>
</tr>
<tr>
<td>N79</td>
<td>2402/3303</td>
<td>+</td>
</tr>
<tr>
<td>N85</td>
<td>1101/3303</td>
<td>−</td>
</tr>
<tr>
<td>N109</td>
<td>1101/3101</td>
<td>+</td>
</tr>
<tr>
<td>N117</td>
<td>0203/—b</td>
<td>+</td>
</tr>
<tr>
<td>N117</td>
<td>0203/—b</td>
<td>+</td>
</tr>
<tr>
<td>NCG13</td>
<td>0203/0207</td>
<td>−</td>
</tr>
<tr>
<td>NCG22</td>
<td>0207/1102</td>
<td>+</td>
</tr>
<tr>
<td>NCG35</td>
<td>1101/—b</td>
<td>+</td>
</tr>
</tbody>
</table>

*NA, information not available; +, more than 50% of clones contain mutations; −, 50% or less of clones contain mutations; −, no detectable mutations.

The frequency of mutation is calculated by assigning +, ≥, or 0. The frequencies of mutation in these two mutational domains relative to that of HBsAg consensus sequences are statistically significant by Student’s t-test (the P values are <0.0001 and <0.001, respectively).
Sequence analysis of recombinant HBsAg clones. The Sequense kit from United States Biochemical Corporation (Cleveland, Ohio) was used for sequencing the cloned PCR product. Sequencing reactions in both sense and antisense directions were conducted to clarify ambiguities in the data. Primers used in the DNA sequencing analysis include the following: nt 58 to 77 (5’ CCT GCT GGT GGC TCC AGT TC 3’), nt 189 to 170 (5’ GGG TGC CTA GGA ATC CTG AT 3’), nt 390 to 370 (5’ CCG CAG ACA CAT CCA GCG ATA 3’), nt 370 to 394 (5’ TAT CGC TGG ATG TGT CCG CTG ATC 3’), nt 582 to 565 (5’ AGG TTT TGT ACA GCA ACA 3’), and nt 601 to 617 (5’ TGT ATT CCC ATC CCA TC 3’). Three to four independently isolated recombinant clones per sample were sequenced. The HBsAg DNA sequences were translated by computer program into their corresponding amino acid sequences and then aligned with subtype-specific consensus sequences.

Because of the overlap between the HBV polymerase and HBsAg open reading frames, some of the HBsAg mutations could in theory simultaneously affect the polymerase sequences. However, since the predominant asparagine-to-serine mutation at HBsAg codon 40 does not affect the glutamine residue of the overlapping polymerase sequence, we focus our discussion on HBsAg.

Mutations within a known B-cell epitope (HBsAg 124-148). Consistent with previous reports (5, 11, 14, 34, 35), we found putative immune escape mutations within the group a determinant in 58% of the HBV clones from our hepatoma patients (Fig. 1 and Table 2). The mutation of threonine to histidine at codon 148 abolished the consensus sequence motif (Asn-X-Thr) required for HBsAg glycosylation at asparagine 146 (19). Most mutations at codon 129, within the first loop of the a determinant (amino acids 124 to 137), as well as the mutations at codons 147 and 148, within the second loop of the a determinant (amino acids 139 to 148), have not been reported in the literature so far. In addition, the mutation frequencies from either loop appear to be similar, at least when this limited number of liver samples was examined. No HBV variants were found to contain mutations within both loops simultaneously (Fig. 1).

Because all the hepatoma patients used in this study are HBsAg positive, by inference, they lack anti-HBs. Therefore, we expected that few immune escape HBsAg mutants can predominate in the absence of immune selective pressure from the neutralizing anti-HBs. As shown in Fig. 1, we were surprised to find frequent mutations within the a determinant. Perhaps the lack of anti-HBs in these patients does not indicate that they are truly immunologically unresponsive. Using a new assay via immune complex dissociation, Maruyama et al. reported that virtually all the chronic carriers with liver disease and approximately 50% without liver disease demonstrated anti-HBs in their sera (13). Our results argued that, despite the apparent presence of HBsAg in the sera of chronic hepatitis patients, anti-HBs was probably present in some of these patients at some point in the past and could exert the immunosuppressive pressure to influence the evolution of HBV variants.

A novel mutational domain (HBsAg 28-51) coincides with a T-cell epitope. In addition to the previously documented escape mutations in B-cell epitopes (the a determinant), to our surprise we found a higher incidence of mutations (83%) occurring within HBsAg amino acids 28 to 51 (Fig. 2 and Table 2). Mutations within this region appear to occur predominantly at positions 40 (asparagine to serine) and 47 (valine/threonine to alanine/glutamic acid/lysine). It should be noted here that position 40 is highly evolutionarily conserved. It remains an invariable asparagine in 28 of 28 available full-length HBV sequences in existing databases as well as in 88 of 88 published HBsAg sequences (18), irrespective of the subtypes and geographic origins (approximately one-third are from Asia, including China and Japan). Furthermore, the occurrence of these mutations is not random. For example, position 40 almost always changed into a serine, not into any other kind of amino acid (Fig. 2).

The novel mutations within domain 28-51 should not be taken as directly oncogenic or even hepatoma specific since they are present at almost equal frequencies in both HCCs (four of five) and their adjacent nontumor livers (six of seven) (Fig. 2 and Table 2). They could alter either an unknown protective B- or T-cell epitope. Indeed, this region coincides with a mapped class I HLA-A2-restricted T-cell epitope (HBsAg 41-49) (Fig. 3). As shown in Table 2, only 40% of our patients are HLA-A2 positive. However, some patients who are not
HLA-A2 positive still have mutations within HBsAg 28-51. We speculate that, in addition to HLA-A2, other HLA-A2-related haplotypes, including HLA-A11 and -Aw33, could also present peptides from HBsAg 28-51 (Table 2). Both A11 and Aw33 are common alleles in Asians (40 and 17%, respectively). Since the phenotype frequency of the HLA-A2 allele is rather high in Caucasians (49%), blacks (33%), and Asians (55%), our findings have general significance in different ethnic groups.

Few mutations at subtype-specific determinants. In contrast to the frequent mutations clustering at HBsAg amino acids 28 to 51 and 124 to 148 (Fig. 1 and 2), few mutations were found outside these two domains (data not shown). For example, very few mutations were found at codons 122 and 160, which are known to dictate the d/y and w/r subtypic determinants, respectively (around 4.1% at codon 122 and 0% at codon 160). These subtype-specific positions are supposed to be rather tolerant to random mutations.

Association of the novel mutations with HBV subtypes? It should be pointed out that, due to the limited sampling size and the preliminary nature of our data, it is too early to conclude whether the mutation at codon 40 preferentially occurs in patients infected with the adw subtype and the mutation at codon 47 preferentially occurs in patients infected with the adr subtype. However, if such an association is true, this result probably reflects different selective pressures against this same T-cell epitope in the different genetic contexts of HBV subtypes. Alternatively, it could reflect a different immunogenetic makeup among different ethnic groups. It is known that the adr subtype is often associated with chronic carriers among the northern Chinese while the adw subtype is often associated with carriers among the southern Chinese (27, 28).

Hierarchy of immune protection. Eight distinct T-cell epitopes of HBsAg have been identified by a cytotoxic T-lymphocyte assay (7) (Fig. 3). Our results should not be taken as an indication that the other seven different T-cell epitopes are less important in vivo. It is possible, for example, that mutations within these seven epitopes cannot be tolerated because of their lethal effect on HBV replication. Similarly, in part due to the small sampling size, it is too early to conclude whether the T-cell epitope 28-51 is higher in the hierarchy of protective immunity than the B-cell epitope 124-148 (Table 2).

De novo versus preexisting mutation. Our present study is cross-sectional rather than longitudinal. Therefore, it is still unclear whether the observed mutations of HBsAg occurred in the very same individual or occurred and were selected previously in another individual and transmitted subsequently to the patients being examined here. This issue is still under active investigation by long-term follow-up of chronic carriers. One way or the other, it should be pointed out here that mutation as a term used in this paper is not defined by comparing HBV sequences at different time points for the same individual. Rather, it is defined by its occurrence at highly conserved positions.

Considering that chronic carriers are the major reservoir of HBV natural infection, it is unclear why these putative immune escape mutations have not evolved to become the predominant population in natural acute hepatitis. It is likely that some kind of disadvantage is also associated with these mutations, such as attenuated infectivity and replication. To our knowledge, this is the first report of a frequent HBsAg mutational domain occurring outside the well-documented B-cell epitope 124-148. The fact that this hotspot mutational domain of HBsAg 28-51 coincides with a T-cell epitope suggests that this domain could contribute to the protective cellular immunity in natural infection with HBV. We hypothesize that mutations within this

FIG. 1. Amino acid sequence comparison of the mutational domain (amino acids 28 to 51) in HBsAg between HBV consensus and naturally occurring mutants in hepatomas. Data are determined by the T-cell epitope. The asterisk indicates a mutation position with a frequency of 80% or higher. The dynamic range of HBV consensus (that is, the range of frequencies of mutations at various positions is shown in parentheses) is indicated by the brackets.

FIG. 2. Comparison of the potential immunodominant domains of HBsAg mapped by genetic or immunological approaches. For the sake of clarity, this figure is not drawn to scale. Amino acids 28 to 51 coincide with either an HLA-A2-restricted or a mouse H-2 Ld-restricted (asterisk) T-cell epitope. Amino acids 124 to 148 represent a group a determinant. Data for class I-restricted T-cell epitopes are taken from references 7, 17, 22, and 26. Data for class II-restricted T-cell epitopes are from references 1, 6, and 15.

FIG. 3. Comparison of the potential immunodominant domains of HBsAg mapped by genetic or immunological approaches. For the sake of clarity, this figure is not drawn to scale. Amino acids 28 to 51 coincide with either an HLA-A2-restricted or a mouse H-2 Ld-restricted (asterisk) T-cell epitope. Amino acids 124 to 148 represent a group a determinant. Data for class I-restricted T-cell epitopes are taken from references 7, 17, 22, and 26. Data for class II-restricted T-cell epitopes are from references 1, 6, and 15.

VOL. 71, 1997 NOTES 4855
domain, either generated by a de novo event or acquired by infection with a survival-of-the-fittest variant from other individuals, probably could predispose the hosts to chronic infection. Consistent with this hypothesis, this novel mutational domain has not been found in acute hepatitis patients. Further functional studies of these mutations should lead to a better understanding of the complex virus-host interactions.

We thank T. Yuan and W. Whitehead as well as D. Milich, D. Walker, S. Baron, N. Roberts, and V. Reyes for discussion and careful reading of the manuscript. We also thank A. Nouraiedeen and A. Ahmed for statistical analysis.

P.-C. Tai is supported by a James W. McLaughlin Predoctoral Fellowship. C. Shih is a recipient of an NIH Research Career Development Award (K04 CA 60163).

REFERENCES


