Genotyping by Multiplex Polymerase Chain Reaction for Detection of Endemic Hepatitis B Virus Transmission

REINALD REPP,† SABINE RHIEL, KLAAUS H. HEERMANN, STEPHAN SCHAEFER, CLAUDIA KELLER, PETER NDOMBE, FRITZ LAMPERT, AND WOLFRAM H. GYLICH

Department of Pediatrics and Department of Medical Virology, Justus-Liebig-University, D-6300 Giessen, and Department of Medical Microbiology, University of Göttingen, D-3400 Göttingen, Germany, and Centre Universitaire des Sciences de la Santé, University of Yaounde, Yaounde, Cameroon

Received 17 September 1992/Accepted 11 February 1993

A nested polymerase chain reaction (PCR) protocol was developed for rapid genotyping of hepatitis B virus (HBV). During the first PCR round, a universal HBV primer pair was used to amplify the entire pre-S region of the HBV genome. Within the pre-S region, many nucleotide exchanges are observed. These are partly correlated to the serological hepatitis B surface antigen subtypes. Five additional subtype-specific primers were selected from that region which, together with two universal non-group-specific primers, generated specific combinations of two to four DNA fragments of defined sizes. By this approach, 55 hepatitis B surface antigen-positive patients from a pediatric oncology unit in Germany were analyzed. Fifty-four patients who had been infected within 2 years had an identical pattern in the multiplex PCR, suggesting a common source of infection and person-to-person transmission within the unit. One child who was infected 5 years later had a different PCR pattern and, therefore, must have been infected from a different source. Furthermore, 109 serum samples taken from pregnant Cameroonian women and 25 serum samples from their babies taken 6 months after birth were analyzed. In one case, mother-to-infant transmission of the virus was demonstrated. Apart from its role in epidemiological studies on HBV, multiplex PCR may also be a useful tool for rapid genetic analysis in other fields if there is a moderate degree of sequence variation which enables the design of specific primers.

Hepatitis B virus (HBV) is a parenterally transmissible pathogen for which specific measures for protecting medical staff and hospitalized patients are required. In particular, long-term hospitalized patients undergoing hemodialysis or cancer chemotherapy are at high risk for acquiring HBV infections. Leukemic patients were among the first in whom hepatitis B surface antigen (HBsAg) was detected (4). This may be due to multiple venous punctures, transfusion of blood products, and destruction of mucous membranes in cases of polychemotherapeutic treatment for malignancies. Furthermore, these patients do not show clinical symptoms of HBV infections because of the low cytopathogenicity of the virus in the absence of a well-developed immune response (2, 8, 34, 45). Routine monitoring of patients for HBV infections, therefore, should be carried out.

If more than one patient is positive, discrimination between sporadic and endemic HBV infections is required. For such a purpose, characteristic markers of the virus need to be demonstrated as identical or different in a group of infected patients. Antigenic variations within the surface antigen (HBsAg) of HBV have been long known, and these may be used to identify different strains of the virus (40). Besides a common determinant designated "a," there are two determinant pairs, "d-y" and "w-r," which are mutually exclusive and which permit the recognition of four major subtypes of HBsAg, "adw," "adr," "ayw," and "syr." On the basis of additional minor determinants, eight subtypes of HBsAg were defined at an international workshop in Paris, France, in 1975 (5). The antisera or monoclonal antibodies for the distinction of these epitopes are not generally available and are difficult to prepare. Moreover, the definitive grouping of all known serologic subtypes by comparing their gene sequences is very complex and exceeds by far the complexity of determining the few single-base exchanges which encode the determinants d or y and w or r (26). The antigenic characteristics of HBsAg can be correlated to point mutations within the S gene (29). By using the correlations between HBsAg subtype and the S-gene sequence, the polymerase chain reaction (PCR) (36) has been previously applied to determine HBV subtypes (27, 39, 46). However, most point mutations are difficult to discriminate by PCR. Additional procedures like DNA sequencing (46) and restriction endonuclease digestion (39) are required or at least three different PCR sets are needed to type a single sample (27).

The genetic differences between HBV subtypes are not limited to the S gene. They are found all over the HBV genome, with the highest variations within the pre-S region. There is no stringent correlation between the phenotypic HBsAg markers and sequence variation outside the S gene. But such a correlation between genetic and phenotypic markers is not required for epidemiological studies. Any sufficiently stable marker allows the discrimination of different strains of infectious agents. With today's gene technology, genotypic markers are often easier to determine than phenotypic markers, provided that a sufficient amount of nucleic acid is available. For detection of HBV DNA in serum, PCR has been widely used. If PCR is to be used for simultaneous amplification and subtyping, it is important that the amplified gene region exhibit a moderate degree of variation. If variation is too high, it may be impossible to guarantee a detection of any genotype with a limited number of PCR primers. On the other hand, too little genetic

† Corresponding author.
variation does not allow any characterization of endemic infections by genetic markers. The pre-S region of the HBV genome, which encodes the amino-terminal domains of the large and middle HBs proteins (14), contains moderately variable regions flanked by sequences conserved among all HBV subtypes. Therefore, it can be considered an ideal genetic target to discriminate different HBV strains by PCR. We have selected a set of PCR primers, located within the pre-S region of the HBV genome, which allows discrimination between four different HBV genotypes directly by one nested PCR run. This protocol was applied with 86 serum samples from two different epidemiologic situations.

MATERIALS AND METHODS

Sera and patients. Serum samples were obtained from 55 pediatric patients with malignancies. While under multidrug cancer chemotherapy, 74 children had been infected inadvertently in an endemic HBV outbreak at a pediatric oncology unit at the university clinics of Giessen, Germany, from 1984 to 1986 (2). Most of the patients suffered from acute lymphoblastic leukemia. Until 1989, 20 children had died because of the primary malignant disease, but an adverse prognostic influence of the HBV infection on the malignant disease was not detected (22). More than 90% of the patients who survived (49 out of 54) developed an immunotolerant HBV-carrier state without any clinical, serological, or histological signs of liver cell destruction. They are still positive for HBsAg, HBeAg, and HBV DNA serologically, even though anticancer chemotherapy has been withdrawn for more than 5 years. In all of those carriers, HBV DNA titers were higher than 10^6 genomes per ml of serum. Twenty percent have not even developed anti-hepatitis B core, although hepatitis B core antigen is very immunogenic. Liver biopsies were taken from 36 patients in 1989. These revealed normal histologies or at most minimal hepatitis with some ground-glass hepatocytes but no signs of inflammation. Furthermore, clinical signs of hepatitis or an elevation of the liver enzymes to above 200 U/liter was never detected in these patients, even though they were examined carefully in accordance with the guidelines of the cancer therapy protocols. Temporary, slight elevations of liver enzymes are a common side effect of some anticancer drugs, and these values did not differ significantly from those for noninfected patients during therapy. After therapy had been withdrawn, these levels were always normal. In contrast, nine of the patients' relatives presented with acute hepatitis B. Only those repeated cases of acute hepatitis B among the patients' relatives suggested that the children at the pediatric oncology unit were infected with HBV. When screening for HBsAg was started, the very high number of subclinical HBV infections among these patients was diagnosed. Probably, the virus had been brought to the ward by one patient's relative, an adult HBV carrier from northern Africa. Because of the high carrier rate in our patients and their high susceptibilities to HBV infection while under multidrug chemotherapy, it was very difficult to stop the endemic HBV transmission on the ward. The destruction of mucous membranes, a major side effect of anticancer chemotherapy, might have favored oral transmission of HBV with our pediatric patients. For these reasons, we started to passively immunize all of the patients attending our unit with anti-HBs hyperimmunoglobulins in addition to routinely screening for HBsAg or HBV DNA in the patients' sera.

Another serum sample in our multiplex PCR study was derived from one of our patients with leukemia who had been determined to be an HBsAg carrier by routine screening in 1991 during the first month of anticancer therapy. Furthermore, 25 serum samples which had been taken from mothers in Yaounde, Cameroon, at the time of their babies' births were typed by multiplex PCR. This number of sera had been positive for HBV DNA by PCR in a previous study that used a total number of 109 maternal serum samples (13). Six positive serum samples taken from the babies at the age of 6 months were also subjected to PCR genotyping. All sera used in this study had been drawn in 1990 and 1991, except for those from the patients in the pediatric oncology unit who had not become HBV carriers. With these patients, their latest HBsAg-positive sera were analyzed. Prior to being used, the sera had been stored at -20°C.

For validation of our assay, four reference serum samples of known HBsAg subtype and two negative control serum samples were used. The HBsAg determinant d was assayed by immunodiffusion with monoclonal antibody C14/02 (12b), and the subtype determinant y was determined by immunodiffusion with a polyvalent guinea pig antisera raised against purified ay HBsAg and absorbed against HBsAg of subtype ad (12a). In addition, the HBsAg subtypes of the four reference serum samples were determined by DNA sequencing of the S gene and the pre-S region of the viral genome. The subtypes of the reference sera were adw6 (serum R1), adw4 (serum R2), ayw3 (serum R3), and adr (serum R4). Every sample included in this study was analyzed twice to prove the reproducibility of the results.

Primer selection. PCR primers were synthesized with an Applied Biosystems oligonucleotide synthesizer (Applied Biosystems, Inc., Foster City, Calif.) and purified by high-performance liquid chromatography. To achieve maximal sensitivity, a nested PCR protocol was used (11, 18, 41). The PCR primer sequences shown in Table 1 were chosen on the basis of an alignment of 20 HBV DNA sequences accessible from the EMBL Data Library (Heidelberg, Germany) and the sequences of the reference sera (3, 7, 9, 10, 19-21, 28-30, 32, 35, 38-42, 44). The outer primers for the first PCR round spanned regions of the HBV genome which are highly conserved among different HBV strains flanking the pre-S segment at positions 2820 to 2845 and 191 to 168 in the case of an HBV genome size of 3,221 bp (7, 21, 32, 43). These are designated P1 and P2 in Table 1. They exhibited at most one mismatch to the different subtypes of HBV DNA sequences listed in Table 1. Therefore, all subtypes of HBV DNA should be amplified during the first round of PCR. The inner primers for the second round of the nested PCR protocol were designed to discriminate between different groups of HBV genotypes. Furthermore, the general advantages of Nested PCR resulted in increases of sensitivity and specificity (1, 11, 31). The sequences of the second-round primers shown in Table 1 were conserved within one group of HBV genomes but were very different from the other HBV sequences. Furthermore, the second-round primers generated different sizes of PCR products which were easily distinguishable by agarose gel electrophoresis. For the HBsAg subtype ayw-encoding genomes, this goal was reached with primers P3-P4, as shown by the distribution of matches and mismatches to different subtype sequences in Table 1. For genome subtypes expressing HBsAg subtype determinant r, primers P5-P6 were selected. In the case of adw, however, no primer sequences were found within the pre-S region, which could detect all adw sequences and discriminate adw from the other HBV subtypes at the same time. To overcome these difficulties, the following strategy was developed. At first, a pair of primers (P5-P6) was selected which
spanned through DNA regions complementary to some of the adw genome sequences reported so far (21, 30, 43). But there were at least four adw sequences which did not match exactly with the primer pair P5-P6, represented in our study by reference serum R2. Furthermore, there was still a risk of missing unknown HBV genotypes not reported to date, because they may not fit any of the primer pair combinations consisting of one of the three sense primers (P3, P5, and P7) together with one of the antisense primers (P4, P6, and P8). Therefore, a kind of universal HBV primer pair like P1-P2 was added to the second-round PCR. One obvious way to reach this goal would be to add the first-round primers P1-P2 to the second round also, but this could have caused a loss of the specificity which is usually achieved by the second round of a nested PCR protocol. Eventually, this problem was solved by the inadvertent nonspecificity of primer P4, which was originally designed to detect ayw sequences. However, it was also able to prime DNA synthesis with other HBV subtypes, even in the case of three nucleotide mismatches to the template sequences. This was probably due to a very high melting temperature for primer-template complexes that resulted from a high content of guanosine or cytidine residues and a low number of nucleotide mismatches at the 3' end of primer P4 in HBV subtypes other than ayw. Therefore, primer P4 could replace P2 as the universal subtype-nonspecific HBV primer, while the ability to specifically detect ayw sequences was still maintained by the highly subtype-specific primer P3 alone. To ensure a priming
by P4 in any HBV sequence in the case of more than three nucleotide mismatches, this primer was elongated by 1 nucleotide to 25 nucleotides, compared with the usual 24 nucleotide residues in the subtype-specific primers. Thus, in our final protocol a mixture of the primers P1 and P3 through P8 was used during the second-round PCR. A master mix of all seven second-round primers was prepared and was stored frozen in small aliquots to minimize possible variations in priming efficiencies due to different primer qualities. The sizes of the different PCR products to be expected and the positions of the different PCR primers P1 through P8 are shown schematically in Fig. 1. If any combination of the 4 sense primers (P1, P3, P5, and P7) and the 3 antisense primers (P4, P6, and P8) is possible, 12 different sizes of PCR products can be generated. In addition, 3 further fragments can be expected, because some of the HBV genomes have a deletion of up to 33 nucleotides close to the 3′ end of primer P1, and therefore, any PCR fragments primed by P1 differ in size depending on the presence or absence of this gap. If the limited resolution of an agarose gel is considered, however, only eight fragment sizes can be discriminated, because some of the PCR fragments to be expected are very similar in size (Fig. 1).

DNA extraction. One milliliter of the patients’ sera was diluted with 3 ml of phosphate-buffered saline, and the HBV particles were pelleted by ultracentrifuging through a cushion of 0.5 ml of 10% sucrose (Beckman SW60Ti rotor, 2 h, 4°C, 50,000 rpm). The pellets were treated with proteinase K in the presence of sodium dodecyl sulfate as previously described (33). HBV DNA was purified by ion-exchange chromatography with a Quiagen 20 column (Diagen, Düsseldorf, Germany) according to the manufacturer’s protocol (6). However, more simplified methods of HBV DNA purification would have been applicable (27) if samples had not been heparinized (17). The amount of DNA corresponding to 1 ml of serum was dissolved in 20 μl of sterile water, and 1 μl was subjected to PCR. A 10−4 to 10−7 dilution of these samples was used in experiments to determine the sensitivity of the PCR protocol and to rule out any effects of different template concentrations on the final agarose gel band pattern.

**DNA amplification.** PCR was performed with a Bio-Med Thermocycler 60 (Bio-Med, Theres, Germany) with a GeneAmp kit (Perkin Elmer, Überlingen, Germany). After an initial melting step (5 min at 94°C), the enzyme was added and 35 amplification cycles of 60 s at 94°C, 120 s at 60°C, and 120 s at 74°C were carried out in a 50-μl final volume containing 4 pmol of each primer during the first round of the nested PCR. The rather high number of 35 cycles was chosen during the first round to drive the PCR into the linear phase of amplification when the amount of primers or enzyme became limiting. By this method, different template concentrations should become compensated prior to the second-round PCR, since smaller amounts of template will be amplified exponentially during more cycles before the plateau phase of the PCR is reached. Such a compensation step should reduce the possible influence of differences in the primary template concentrations on the final results after the second-round multiplex PCR. One microliter of the first-round product was subjected to the second round of PCR. This round differed from the first round of PCR by the number of cycles, which was reduced to 25, the amount of primer (20 pmol), and the annealing temperature (68°C). PCR products were analyzed on a 2.5% agarose gel and visualized by ethidium bromide staining.

**DNA sequencing.** In addition to being serologically subtyped, HBV subtypes in samples used as reference sera were determined more accurately by DNA sequencing of the S
RESULTS

All sera from the 55 children with malignancies were subjected successfully to nested PCR with the multiplex primer set in the second round. By agarose gel analysis of the amplificate, an identical pattern with two bands was obtained from all 54 samples taken from the children who had been infected during the hospital endemia. The results for three of these samples are shown in Fig. 2. This pattern was identical to the result from the reference serum R3 expressing the HBsAg subtype ayw. The sizes of these two bands corresponded well to the expected sizes of 486 and 521 bp for reference serum R3, suggesting that the fragment amplification indeed was primed by P3-P4 and P1-P4, respectively (Fig. 1). Furthermore, the identities of the PCR products obtained from these 54 children were confirmed by DNA sequencing of the pre-S region in 10 cases.

In contrast, the serum from the child who had been diagnosed as an HBV carrier 5 years later revealed a different pattern with three bands: two major bands and a third band of a lower intensity. In Fig. 2, it is shown as reference serum R1, which had been typed as adw serologically. As there are two serum’s babies at 6 months. The exact sizes of the bands (in base pairs) are listed at the left.

gene and the pre-S region of the HBV genome. Furthermore, the pre-S region was sequenced in 10 out of 54 serum samples from endemic infections among children with malignancies to prove the identity of the HBV genome in these patients and also in a serum sample from a Cameroonian mother with a unique gel pattern after PCR typing. For this purpose, PCR was performed according to the first-round protocol, but the primer concentration was increased 10-fold and recognition sites for the restriction enzymes SalI and BclI were added to the PCR primers to facilitate cloning. PCR products were cloned into the plasmid vector pUC19 with the SalI and BamHI restriction sites within the polylinker region (24). Plasmid preparations from single colonies were sequenced in both directions with a Sequenase kit (Pharmacia, Freiburg, Germany) (37). In order to avoid PCR artifacts, two different clones were sequenced from each patient’s serum.

FIG. 1. Schematic diagram of the different fragment sizes to be generated by multiplex PCR. Primer positions are indicated by P1 through P8, and the sizes of their products are shown at the right. Because of a gap of up to 33 nucleotides near the 3' end of P1, fragments initiated by this primer may vary in size. +GAP indicates the presence of 33 nucleotides, whereas −GAP indicates their absence. Gaps of fewer than 33 nucleotides can also be found in some HBV genomes, which lead to further size variation in PCR products primed by P1. The gel patterns obtained from the reference sera and from the Cameroonian serum M61 are shown schematically as a boxed set of lanes (Fig. 2). LHBs, large HBs protein.

FIG. 2. Ethidium bromide staining of PCR products separated over an agarose gel. Lanes: MW, molecular size markers; NE, negative control serum; R1 to R4, reference sera; P1 to P3, sera from endemic infections among children with malignancies; M61, M70, and M94, sera taken from pregnant women in Cameroon; B61 and B62 were taken from the Cameroonian mother with the sera subjected successfully. The fragment amplification indeed was primed by P3-P4 and P1-P4, respectively (Fig. 1). Furthermore, the identities of the PCR products obtained from these 54 children were confirmed by DNA sequencing of the pre-S region in 10 cases.
three serum samples. One child and her mother revealed a unique pattern different from those of all other sera analyzed in this study. This suggests mother-to-infant transmission of the virus. The identity of these two isolates was confirmed by DNA sequencing of the pre-S region amplified from this HBV genome (Table 1, serum M61).

It appeared possible that different template concentrations would alter the agarose gel pattern seen after PCR. Therefore, two endpoint dilution series were performed with each of the four reference serum samples, R1 to R4, after determination of their respective HBV DNA titers (47). No differences were seen at template concentrations corresponding to $10^8$ to $10^9$ HBV genomes per ml of serum. For R3, the sensitivity seemed to be even slightly higher, because there was still a typical gel pattern at a concentration of $10^8$ genomes per ml of serum in one of the two dilution series, but these differences cannot be considered significant. When only the two subtype-specific primers were used during the second-round PCR (e.g., P3 and P4 with serum R3), the sensitivity was higher than $10^5$ HBV genomes per ml in each of the reference sera. The sensitivity of a single first-round PCR as described for DNA sequencing was $10^3$ genomes per ml of serum in all four reference serum samples. However, the very low concentrations of P1 and P2 used during the first round of the nested PCR did not reveal enough product to be clearly visible on an agarose gel (data not shown).

**DISCUSSION**

Proof of an endemic infection often requires detailed subtyping of the infectious agent, especially if there is a high number of sporadic infections in parallel. With HBV, several methods based on genetic or antigenic differences among HBV subtypes have already been applied for this purpose (15, 25, 27, 39, 40, 46). Methods using the highly sensitive PCR possess some advantages over other procedures, as they are still able to detect and to type HBV DNA in low-viremic subjects (41). The nested PCR protocol with a multiplex primer set described in this paper further simplifies screening for endemic HBV transmission. It combines highly sensitive detection of HBV infectivity and HBV genotyping without the need for further time-consuming procedures such as digestion by restriction endonucleases (46) or DNA sequencing (39). Furthermore, these data show that a multiprimer PCR protocol generates more differentiated results in HBV genotyping than other PCR-based methods because of the competitive effects between different primers which not only result in the presence or absence of bands on an agarose gel but also lead to different band intensities (Fig. 1 and 2). Certainly, this is an advantage in epidemiologic studies, but, on the other hand, it means that the gel patterns generated by our protocol do not exactly reflect HBsAg subtypes or the four HBV genotypes, A to D, previously defined (26, 29). On the basis of published sequence data (Table 1), the highest correlations can be expected for the reference sera R1 and R3. An R1 pattern indicates the HBsAg subtype adw corresponding to genotype A, and the R3 pattern is obtained from an ayw serotype corresponding to genotype D. Other adw genotypes corresponding to genotype B or C may result in variable gel patterns with similarities to that for the reference serum R2 or R4. Finally, an r determinant serum will show an R4 pattern (29).

In a multiplex PCR protocol, there are several points that have to be considered which may affect primer-template affinities and competition between different primers of the same sense. First, the melting temperature of primer-template complexes has a major influence on the efficiency of amplification. It is dependent on the length of the primer, the relative amount of G or C nucleotides in comparison to A or T nucleotides, and the number of mismatches to the target sequence. As there are two primers required in each PCR amplification, the primer-template affinities of both primers have to be multiplied with regard to the efficiency of amplification. This means that a reduction in primer-template affinity to 50% in each primer should theoretically result in only 25% product amplification efficiency. Second, the positions of primer-template mismatches and the kinds of mismatches are important. Mismatches at the 3' end are more relevant than mismatches at the 5' terminus of a primer. A mismatch caused by a thymidine residue within the 3'-terminal primer sequence has fewer consequences than mismatches caused by other nucleotides (12). Third, the relative position of the primers to one another is of relevance. If, for instance, there are two primers of the same sense and the same affinity to the template competing with one antisense primer, the one sense primer located more closely to the antisense primer has an advantage if the reaction is not limited by too-low primer concentrations (Fig. 2, reference serum R4). Third, if the primers bind at one template molecule at the same time, the strand elongation initiated by the outer primer will stop at the position of the inner primer, as this primer already blocks the template molecule. Fourth, the annealing temperature used in a multiplex PCR has to be calculated well, so that different primer-template affinities clearly result in different amounts of final amplification products. An annealing temperature that is too low will mask the discriminatory effects of different primers. An annealing temperature that is too high will prevent any amplification at all. In summary, an optimal multiplex PCR primer set cannot only be calculated theoretically; it needs further empirical evaluation.

The application of such a multiplex PCR protocol will not be limited to subtyping of HBV in populations at high risk for HBV infections. Such a protocol might also be adapted to genomes of other infectious agents showing a suitable degree of sequence variability.

**ACKNOWLEDGMENTS**

Subtype adw HBV serum was kindly provided by Angela Uy, University of Göttingen. Subtype adw was kindly provided by Clara Yoshida, Fundação Oswaldo Cruz, Rio de Janeiro, Brazil. This work was supported by the Forschungshilfe Station Peiper and by the Deutsche Forschungsgemeinschaft, with a Habilitandenstipendium to R. Repp (re 841/1-1).

**REFERENCES**


38. Vaudin, M., J. A. Wollenholler, K. N. Tsiquaye, A. J. Zacker,man, and T. J. Harrison. 1988. The complete nucleotide sequence of the genome of a hepatitis B virus isolated from a...