Discovery of a novel point mutation changing the HDAg expression of a hepatitis delta virus isolate from Central African Republic

J. R. Tang, O. Hantz, L. Vitvitski, J. P. Lamelin, R. Parana, L. Cova, J. L. Lesbordes and C. Trépo*

Institut National de la Santé et de la Recherche Médicale (INSERM) U 271, 151 Cours Albert Thomas, 69424 Lyon, France

None of the mutations so far discovered in several hepatitis delta virus (HDV) isolates appears to determine important changes in HDV specific protein (HDAg) expression, except for a putative mutation at nucleotide 1012 converting an amber stop codon (TAG) to a codon for tryptophan (TGG). Here we present the characterization of an HDV obtained from the liver of a woodchuck inoculated with sera from fulminant HDV patients in Central African Republic (CAR). By restriction enzyme analysis and sequencing of HDAg-coding region cDNA clones, we found that this HDV isolate bears a novel mutation (T to A) at nucleotide 1013 which converts the amber stop codon (TAG) to a codon for lysine (AAG). Comparison of these nucleotide sequences with those available from American, Japanese, Taiwanese, French, Italian and Nauru isolates showed a variability of 1.7 to 21.5 % and 1.9 to 28.7 % at the nucleic acid and amino acid levels, respectively. The HDAg-encoding sequence of the CAR isolate is closely related to that of the Italian HDV isolate. The in vitro expression of this HDV isolate resulted in a unique HDAg species (28K) which was identical with that characterized in vivo.

Introduction

Hepatitis delta virus (HDV) is a defective virus that requires the presence of hepatitis B virus (HBV) in primates, or woodchuck hepatitis virus (WHV) in woodchucks, to be infectious (Rizzetto et al., 1980; Ponzetto et al., 1984). The HDV genome is a single-stranded circular RNA of 1.679 kb (Wang et al., 1986). The circular RNA molecule possesses significant intramolecular complementarity such that approximately 70 % of the nucleotides (nt) can form base pairs in an unbranched rod structure (Wang et al., 1986; Makino et al., 1987). HDV RNA contains a functional open reading frame (ORF) in the antigenomic polarity that encodes the HDV-specific protein HDAg (Makino et al., 1987). Two species of HDAg are consistently found in HDV-infected humans and animals (Bonino et al., 1986). One is small, with 195 amino acids (aa) and an M_r of 27K, and the other is large, with 214 aa and an M_r of 29K. The relationship between the two species of HDAg has been elucidated: a specific modification (U to C on the genomic RNA or A to G on the antigenomic RNA) at nt 1012 converts part of the RNA species (termed the S genome) encoding the small HDAg into the RNA species (the L genome) encoding the large HDAg (Luo et al., 1990). More recently, it has been found that this base change (at the 1012 site) occurs on the genomic HDV RNA (Casey et al., 1992; Zheng et al., 1992). The consequence of this editing event is important because large HDAg both inhibits HDV replication (Glenn & White, 1991) and enables packaging of viral RNA (Chang et al., 1991).

Compared with infection with HBV alone, co-infection or superinfection with HDV significantly increases the risk of more severe liver disease, including fulminant hepatitis (Rizzetto, 1983). An outbreak of fulminant hepatitis associated with HDV has been reported in the Bangui region of Central African Republic (CAR) (Lesbordes et al., 1986). Confirmation of the aetiological role of a specific CAR HDV was reported after successful transmission of the virus to WHV-carrying woodchucks. In both woodchucks and humans, this CAR HDV showed a highly specific pathogenicity, i.e. spongiocytic hepatitis. In addition, this CAR HDV isolate was characterized by a unique protein (28K) with HDAg reactivity, found in the livers and sera of infected hosts (Faure et al., 1991).

In order to characterize the CAR HDV genome encoding the unique HDAg, we made a comparison between the CAR HDV and the reference HDV by cloning and sequencing the HDAg-coding region of HDV isolated from liver of woodchuck inoculated with CAR HDV or with reference HDV.
Methods

Animals. WHV-carrying woodchucks (Marmota monax) were used in all inoculation experiments. A pool of sera and homogenized liver from three CAR patients with fulminant HDV hepatitis was inoculated into two woodchucks (first passage). Sera and homogenized liver from these two woodchucks were subsequently pooled and inoculated into two other recipients (second passage). A similar protocol was followed with an HDV reference isolate (Abbott Laboratories) instead of the CAR isolate. Serological markers of HDV (HDAg, total anti-HDV) and HDAg in liver were detectable in all woodchucks (Faure et al., 1991). Liver samples derived from woodchuck W310, one of the second passage woodchucks inoculated by CAR HDV, and W318, one of the second passage animals infected by the HDV reference isolate, were used for the study.

Extraction of RNA. Extraction of RNA from liver tissue frozen at 
−80°C was performed using the guanidinium/hot phenol method (Maniatis et al., 1982). To extract HDV RNA from serum, 300 µl was incubated at 60°C for 3 h in an extraction buffer [50 mM-Tris-HCl pH 7.5, 10 mM-EDTA, 200 mM-NaCl, 1% SDS, 1 mg proteinase K per ml and 10 µg of carrier poly(A) per ml]. HDV RNA was then extracted with phenol–chloroform, followed by isopropanol precipitation. Nucleic acids were then redissolved in diethylpyrocarbonate-treated water and stored in batches at −80°C.

Northern blotting. Purified RNA (20 µg) from liver specimens was electrophoresed in agarose-formaldehyde denaturing gel and transferred to nitrocellulose membrane as described previously (Maniatis et al., 1982). HDV RNA was detected using a purified HDV cDNA probe 1:3 kb long (kindly provided by Dr J. Taylor, Fox Chase Cancer Centre, Philadelphia, Pa., U.S.A.) labelled by nick translation. Prehybridization was performed at 42°C for 4 h in a solution containing 50% formamide, 0.1% SDS, 5 × SSC, 5 × Denhardt's solution, 250 µg/ml sheared salmon sperm DNA. Hybridization was carried out at 42°C for 12 h in a similar solution which contained a 32P-labelled probe. The filter was washed twice at 65°C in 2 × SSC, 0.1% SDS, then twice in 1 × SSC, 0.1% SDS and finally twice in 0.5 × SSC, 0.1% SDS, dried and autoradiographed at −80°C for 18 h.

cDNA synthesis and PCR amplification. RNA extracted from liver or serum was used as the template for reverse transcription (RT). RT was performed at 42°C for 45 min in the presence of Moloney murine leukaemia virus reverse transcriptase (BRL) and its corresponding buffer, containing 150 µM-dNTPs, 150 µM antisense primer and sense primer. HDV-derived cDNA was then amplified by PCR using 1:25 units of Taq DNA polymerase (Perkin-Elmer Cetus). PCR was cycled 40 times, each cycle comprising 94°C for 1 min, 42°C for 1 min, 72°C for 1.5 min, followed by a 10 min extension at 72°C in a DNA thermal cycler (Perkin-Elmer Cetus). To obtain sufficient cDNA for the subsequent restriction enzyme analysis and cloning in bacteriophage M13, the PCR product (10 to 15 µl) was re-amplified either with the same primer pairs or with internal primer pairs (nested PCR). To avoid false positive results, the risk of contamination was reduced by following the usual preventive measures (Kwok & Hitch, 1989). Synthetic oligonucleotides used as RT/PCR primer and as hybridization probe were deduced from the known sequences of HDV genome previously reported (Kuo et al., 1988; Saldanha et al., 1990; Wang et al., 1986).

Restriction enzyme analysis and Southern blot hybridization. For the amplified cDNA fragment A (Fig. 1), PCR products were extracted by phenol–chloroform and then digested with NcoI (500 units/ml) (Boehringer Mannheim) at 37°C for 6 h. PCR products were analysed on 2% agarose gels, and the samples were then transferred onto a nylon membrane, hybridized with a 32P-labelled specific oligonucleotide probe and autoradiographed at −80°C for 1 day.

cDNA cloning and sequencing. Two techniques were used to clone HDV cDNA fragments into M13 mp18, or mp19 phage vector. One method was the direct cloning of PCR products in ddT-tailed M13 vector (Holton & Graham, 1991). A second method consisted of polishing the termini of purified PCR products with large fragment DNA polymerase (BRL) and then ligation the blunt-end fragment with T4 DNA ligase (BRL) into SmaI-linearized M13. The ligation mixture was used to transform competent Escherichia coli cells which were either prepared using a protocol described elsewhere (Nishimura et al., 1990) or purchased from BRL. Plaques were screened with the 32P-end-labelled specific oligonucleotide probes. Three or four independent clones were sequenced for each cDNA fragment of CAR HDV by the dideoxynucleotide chain termination method using the United States Biochemical sequencing kit. The sequencing reaction products were analysed in a 6% polyacrylamide gel containing 7 M-urea.

PCR-based overlap extension for HDAg-coding ORF reconstruction. To synthesize the T7 expression promoter and then ligate with HDAg-coding ORF, the PCR-based overlap extension (OE-PCR) (Ho et al., 1989; Horton et al., 1989) method was used to reconstruct a full-length HDAg-coding ORF under the control of the T7 promoter.

The universal T7 promoter (UP) was synthesized by primer-dimer formation using 100 pmol each of the primers Up-1 and Up-2 described by Kain et al. (1991). Twenty-five cycles of PCR were carried out, each consisting of 1 min at 94°C, 1 min at 30°C and 1 min at 72°C. The complete HDAg-coding ORF was synthesized from three overlapping DNA fragments as shown in Fig. 7(a).

To eliminate the interfering effect of primers on the annealing of the strands with the matching sequence at their 3' ends during the OE-PCR reaction, PCR products larger than 250 bp were precipitated with PEG (Kusukawa et al., 1990), whereas DNA fragments smaller than 250 bp were purified by electrophoresis through low melting point agarose gel. Approximately 200 ng each of two overlapping fragments were joined by OE-PCR reaction in the presence of 200 µM of each dNTPS and 15 units of Taq polymerase. The OE-PCR comprised a joining step without primers (first denaturation performed at 94°C for 5 min, then 1:5 min in the subsequent five cycles, each consisting of 2 min at 45°C and 15 min at 72°C), and an amplifying step with sense and antisense primers (30 cycles, each consisting of 1 min at 94°C, 2 min at 45 or 55°C, and 3 min at 72°C). For UP addition, the joining step was performed for 15 cycles with an annealing temperature of 25°C followed by 25 cycles of the amplifying step. The ORF coding for HDAg cDNA with UP was extracted with phenol–chloroform, precipitated with ethanol and resuspended in 10 µl RNase-free water.

In vitro transcription and translation of HDAg-coding ORF. The UP- joined cDNA of the HDAg-encoding ORF was transcribed with T7 RNA polymerase using the SP6/T7 Transcription Kit (Boehringer Mannheim). The transcription was performed at 37°C for 60 min in
the presence of 1000 units RNase inhibitor per ml reaction mixture. The DNA template was removed with RNase-free DNase I. Finally, the RNA transcripts were subjected to phenol extraction, ethanol precipitation and resuspension in 10 µl RNase-free water. The synthetic HDV mRNA was translated in the presence of $^35$S-methionine in a rabbit reticulocyte lysate system (Boehringer Mannheim) according to the manufacturer’s specifications. Tobacco mosaic virus (TMV) RNA was used in a control translation.

**Analysis of translated products.** The $^35$S-methionine-labelled proteins produced by the translation system were immunoprecipitated and analysed by SDS–12.5 % PAGE (Sambrook et al., 1989). Briefly, the mixture of rabbit reticulocyte lysate was incubated with an anti-HDV serum from a fulminant CAR HDV hepatitis patient at 4 °C for 2 h under gentle rotation. After addition of staphylococcal Protein A (Sigma), incubation at 4 °C for 1 h and centrifugation, the pellet of Protein A and antibody-bound protein was washed three times and then boiled for 5 min in the presence of 1% SDS to release the target protein. The viral protein was analysed for HDAg by ELISA (Deltassay Ag Tests, Noctech) and by SDS–12.5 % PAGE. After electrophoresis, the gel was dried under vacuum and exposed to Hyperfilm β-max (Amersham) at room temperature for 6 days.

**Results**

**Characteristic features of CAR HDV**

A highly pathogenic strain of HDV responsible for an outbreak of fulminant hepatitis in CAR (Lesbordes et al., 1986) was successfully transmitted to woodchucks as demonstrated by the appearance of serological HDV markers in infected woodchucks (Faure et al., 1991). As compared with the reference isolate of HDV, the CAR HDV infection in woodchuck was associated with the same particular pathological event, i.e. spongioscyt hepatitis. The replication of CAR HDV seemed to be very low as evidenced by the negative result of the Northern blot in the detection of HDV RNA in liver of woodchuck infected with the CAR HDV using an HDV cDNA probe covering 77% of the HDV genome. As shown in Fig. 2, within just 18 h of exposure, a high intensity band of the expected size, approximately 1.679 kb, was detected in the liver of reference HDV-infected woodchuck W318 (lane 2), whereas within the same exposure time this band was not detectable in the liver from W310 infected with the CAR HDV (lane 1). This result, in combination with the positive result in RT/PCR (described below), strongly suggests that the CAR HDV replicates more weakly than the reference HDV.

Another characteristic of the CAR HDV is the detection in Western blot analysis of a single 28K protein. In the same experiment, both the 27K and 29K proteins classically associated with HDV were detected in sera of woodchucks infected with the reference HDV strain. In addition, the CAR HDV was better recognized by anti-HDV antibody from the CAR HDV-infected patients’ and woodchucks’ sera.

**Restriction enzyme analysis and Southern blot hybridization**

The CAR HDV appears to encode a unique HDAg (28K) in infected hosts, in contrast to the prototype HDV which encodes the 27K and 29K HDAgs. To understand the molecular basis of this phenomenon, the analysis of the CAR HDV genome was performed initially by testing for the S and L HDV genomes using NcoI enzyme digestion of a PCR-amplified cDNA fragment A (Fig. 1) containing the hexanucleotide sequence CCATGG (nt 1011 to 1016, as shown in Fig. 3a). The CAR HDV isolate showed only one band in the position of the S genome (Fig. 3b, lanes 1 and 2) whereas the prototype French (F) HDV had two bands at the positions of the S and L genomes (Fig. 3b, lane 3). This result suggests that a variation exists in the CAR HDV genome.

**A novel point mutation at position 1013**

Following the above discovery, the precise nature of the mutation was determined by sequence analysis with cDNA clones of the PCR-amplified fragment A of the
HDV genome (Fig. 1) around positions 1011 to 1016 for both F HDV and CAR HDV isolates. As shown in Fig. 4, this hexanucleotide sequence (antigenomic sense) was CCAAAG in CAR HDV, but CCATAG (containing an amber codon) and CCATGG (the target sequence for NcoI) in the F HDV S and L genomes, respectively. Thus an unusual characteristic of the CAR HDV was the point mutation (T to A) at position 1013 when compared with the other isolates. Twenty-four cDNA clones derived from two independent manipulations (from RNA extraction to RT/PCR) showed identical results. Neither of the genomes of prototype HDV was detectable. The T_{1013} to A_{1013} transversion means that codon number 196 in CAR HDV encodes lysine, in contrast to all other HDV isolates so far examined in which it codes for either an amber stop codon or tryptophan.

Partial sequence data of CAR HDV and a comparison with other HDV isolates

In preliminary experiments, we failed to amplify a complete HDAG-encoding ORF even with specific primers and a high viral titre serum sample. Thus the HDAG-encoding ORF was amplified by three overlapping fragments A, B and C, which had been amplified and sequenced (Fig. 1). A 638 bp sequence including the putative initial codon (nt 1597 to 1595) and the distant ochre stop codon (nt 952 to 950) was determined for the CAR HDV. The nucleotide sequence and deduced amino acid sequence are shown and compared with those of six other known isolates (Fig. 5a and b). When compared with the Italian, Nauru and F HDV isolates, the HDV CAR isolate showed only point mutations but no deletion or insertion of genetic material. The sequence of CAR HDV was 98.3%, 92.2%, 90.7%, 90.7%, 90% and 78.3% similar to the Italian, F, Southern Californian, Taiwanese, Nauru and one of Japanese isolates, respectively. The predicted aa sequence of the CAR HDV HDAG-encoding ORF revealed 98.1%, 90.8%, 87%, 86.4%, 89% and 71.8% identities, respectively, with the six known HDV isolates (Fig. 5b). As schematically shown in Fig. 6, the similarity between the CAR HDV and the other HDV isolates is not strictly related to their geographical distribution. Nevertheless,
the CAR HDV is closely related to the Italian HDV isolate and distantly to one of the Japanese isolates.

In vitro translation of HDV RNA and analysis of translated products

As shown in Fig. 7(a), a three-step OE-PCR technique was used to construct a DNA fragment corresponding to the HDAg-coding ORF of both CAR and F HDV isolates under the control of the T7 promoter.

mRNA templates produced by in vitro transcription of the DNA fragments were then used to programme a cell-free translation reaction in a rabbit reticulocyte lysate. The immunoprecipitated protein synthesized from the F HDV RNA and the CAR HDV RNA showed HDAg reactivity in an ELISA (Noctech), but that from the control TMV RNA did not. Moreover, the immunoprecipitated protein synthesized from the F HDV RNA showed two protein bands with HDAg reactivity which migrated at positions corresponding to apparent Mₐs of...
Fig. 6. Nucleic acid and amino acid similarity between HDV isolates from different geographic areas. HDAg-encoding ORFs of previously known HDV isolates were compared with that of CAR HDV at the nucleotide (hatched boxes) and amino acid (open boxes) levels. The designations I, F, A, T, N and J are as in Fig. 5.

27K and 29K (Fig. 8a, lane 3), whereas that of the CAR HDV showed a unique protein band with HDAG reactivity at a position corresponding to an Mr of 28K (Fig. 8a, lane 2). The translation product of TMV RNA was used as a negative control (Fig. 8a, lane 1). The pattern of PAGE analysis for in vitro produced HDAG was similar to that of HDAG derived in vivo (Fig. 8b).

Discussion

Compared to all of the HDV isolates derived from different geographical areas, the CAR HDV showed biological and virological peculiarities. CAR HDV was highly pathogenic, causing the appearance of spongioscysic hepatitis in infected hosts (Faure et al., 1991; Lesbordes et al., 1986), but showed extremely low genome replication, as demonstrated by previous and present results from Northern blot and RT/PCR assays. Although restriction enzyme analysis of PCR-amplified products detected both the S and L HDV genomes in F HDV-infected woodchuck liver, this approach failed to demonstrate the co-existence of two RNA species in CAR HDV-infected woodchuck liver. This was confirmed by sequence analysis which demonstrated the existence of a single detectable RNA species of CAR HDV in two independent experiments using RT/PCR products. Comparison of the HDAG-encoding ORF nt sequence between CAR HDV and other isolates (Chao et al., 1990, 1991; Imazeki et al., 1990; Kuo et al.,

Fig. 7. Illustration of the OE-PCR method used for the synthesis of the full-length HDAG coding region preceded by a functional T7 promoter sequence (UP). (a) Four DNA fragments were joined together, two at a time, to produce the HDAG-encoding ORF with an expression promoter. A and B, B and C, C and UP have an overlapping region in each pair. Under PCR conditions, the strands having the matching sequence at their 3' ends overlap and act as primers for each other. The B and C fragments were first joined together to yield BC in the first OE-PCR; then A and BC were joined together in the second OE-PCR to produce ABC; and finally ABC and UP were joined together to produce the target fragment ABCUP. (b, c) Analysis of PCR product by ethidium bromide-stained agarose gel electrophoreses: in gel (b) the primer-dimer formation of UP (73 bp) after 25 cycles (lane 1) or only five cycles (lane 2); in gel (c) the PCR fragments A (lane 1), B (lane 2), C (lane 3), BC (lane 4), ABC (lane 5), ABCUP (lane 6) corresponding to 255 bp, 320 bp, 390 bp, 590 bp, 650 bp and 714 bp, respectively. Ma and Mb are the DNA size markers V and VI (Boehringer Mannheim).
reported in other HDV isolates. The possibility of an artefact mutation can be excluded, since in a previous study (J. R. Tang et al., unpublished), the same PCR products amplified from a number of serum samples were susceptible to the NcoI enzyme. This suggests a high fidelity of PCR reaction under our experimental conditions. In addition, the reliability of the PCR was also assessed by sequencing the 680 nt-long fragment of the F HDV genome including the HDAg-encoding ORF. This sequence proved identical to that previously reported using conventional cloning procedures (Saldanha et al., 1990) (only the sequence data around position 1012 are shown in Fig. 4b and c). Furthermore, among the base mutations caused by Taq polymerase, most are either A to G or T to C transition, and none is a transversion (Tindall & Kunkel, 1988). The finding of a single detectable RNA species was consistent with the finding of only one species of CAR HDAg, the 28K protein, in CAR HDV-infected woodchucks (Faure et al., 1991). The unique detectable polypeptide (28K HDAg) encoded by CAR HDV was characterized by in vivo (Faure et al., 1991) and in vitro studies. However, at the level of genome analysis, the CAR HDV could encode a polypeptide of 214 aa similar to that of the prototype HDV L genome. It remains unclear why the HDAg of CAR HDV manifested a slightly faster electrophoretic mobility (28K) than predicted (29K). The initiation role of an ATG triplet at nt 1366 to 1364 can be excluded, because of its poor content for an initiation codon (Kozak, 1983) and its capacity for initiating a polypeptide of only 136 aa (with a theoretical M, of 21K). Furthermore, none of the three types of premature stop codon (amber, opal and ochre) was discovered upstream of the putative stop codon (nt 952 to 950) in the HDAg-encoding ORF of CAR HDV. Because the HDAg is a phosphorylated protein (Chang et al., 1988; Hwang et al., 1992), it is possible that the degree of phosphorylation may account for the difference in HDAg electrophoretic mobility.

Although the CAR HDV RNA species differs from the prototype HDV L genome in the T1013 to A1913 transversion, the HDAg of CAR HDV seems to share biological functions of the prototype large HDAg, such as the replication inhibition of HDV RNA (Glenn & White, 1991). It has also been shown that large HDAg could be packaged into HBsAg particles in the absence of the HDV genome or small HDAg. This suggests the possible formation of empty HDV-like particles (Chang et al., 1991), which might explain why the unique 28K HDAg was more readily detected than the CAR HDV genome in the sera of infected hosts (Faure et al., 1991).

In other systems, minor changes in the sequence of a viral genome may result in major changes in the biology
and pathogenicity of the virus (Evans et al., 1985; Liang et al., 1991; Pritchard et al., 1992; Skinner et al., 1989); whether the high pathogenicity of CAR HDV may be related to the mutation at nt 1013 or to other changes remains to be proven. The existence of the CAR HDV isolate will be of special interest, however, in unravelling the mechanisms of pathogenicity in HDV infection and the occurrence of specific editing events during viral replication.

We thank Daniel Lepot for synthesis of oligonucleotides and Pascal Faure for his contribution to this work.

References


(Received 4 January 1993; Accepted 29 April 1993)