ABSTRACT A strain of hepatitis E virus (SAR-55) implicated in an epidemic of enterically transmitted non-A, non-B hepatitis, now called hepatitis E, was characterized extensively. Six cynomolgus monkeys (Macaca fascicularis) were infected with a strain of hepatitis E virus from Pakistan. Reverse transcription–polymerase chain reaction was used to determine the pattern of virus shedding in feces, bile, and serum relative to hepatitis and induction of specific antibodies. Virtually the entire genome of SAR-55 (7195 nucleotides) was sequenced. Comparison of the sequence of SAR-55 with that of a Burmese strain revealed a high level of homology except for one region encoding 100 amino acids of a putative nonstructural polypeptide. Identification of this region as hypervariable was obtained by partial sequencing of a third isolate of hepatitis E virus from Kirgizia.

Epidemics of enterically transmitted non-A, non-B hepatitis (hepatitis E) have been reported in Asia, Africa, and North America (for reviews, see refs. 1–4). Similar cases of sporadic hepatitis, presumed to be hepatitis E, account for up to 90% of reported hepatitis in countries where hepatitis E is endemic. Furthermore, hepatitis E has been implicated in fulminant hepatitis of pregnancy: up to 20% of pregnant women with hepatitis E die of their disease. That a viral agent was responsible for hepatitis E epidemics was first shown by Balayan et al. (5). It was provisionally called hepatitis E virus (HEV) (2). On the basis of electron microscopy and other characterization it was proposed that HEV belongs to the calicivirus family (6, 7). A partial sequence (8) and later the entire sequence (9) of an HEV genome have been reported. Comparisons of HEV sequences with sequences available in computerized data bases, including those of caliciviruses (10), showed little similarity between HEV and any other known viruses. Several species of primates have been experimentally infected with HEV (5, 11–19). Although HEV infection is less severe in monkeys than in humans, it is sufficiently similar that monkeys can serve as a reasonable and reproducible animal model to provide a more complete characterization of this infection.

Determination of the pattern of HEV shedding in feces, bile, and sera during infection is important for characterization of the disease and elucidation of the epidemiology of HEV. In some cases, the presence of HEV in a sample has been demonstrated by transmission to animals (5, 11–19) but such transmission studies cannot be used for routine investigation. Until recently, the primary method available for characterization of virus shedding was immune electron microscopy (IEM) (for reviews, see refs. 1–4). However, because the level of HEV in feces and bile is very low (5, 11–19), the sensitivity of IEM is inadequate for complete characterization of HEV infection. A more sensitive technique, detection of the virus genome by reverse transcription–polymerase chain reaction (RT–PCR), was used in this study to correlate the presence of HEV in serum, bile, and feces of an experimentally infected cynomolgus monkey with biochemical evidence of hepatitis and development of antibodies to HEV (anti-HEV). In addition, we obtained structural information about the virus genome through sequencing of the PCR products themselves or their cloned derivatives.**

MATERIALS AND METHODS

Virus Samples and Inoculation of Primates. Feces containing HEV SAR-55 were collected from a patient during a hepatitis E outbreak in Sargodha, Pakistan (20, 21). Approximately 0.5 ml per monkey of a 10% (wt/vol) stool suspension of feces in fetal calf serum was used for intravenous inoculation of six cynomolgus monkeys (Macaca fascicularis). Blood samples from these monkeys were taken approximately twice weekly before and after inoculation and tested by Metpath (Rockville, MD) for biochemical evidence of hepatitis by measuring levels of serum alanine aminotransferase, isocitric dehydrogenase, and γ-glutamyltransferase. Fecal and bile samples were also collected from one monkey. For daily collection of bile, implantation surgery was performed on the 7th day after inoculation to establish an indwelling T-tube into the bile duct.

Bile from a cynomolgus monkey infected with another strain of HEV (OSH-1852) was kindly provided by Michael Balayan (Institute of Poliomyelitis and Viral Encephalitis, Moscow). This cynomolgus monkey had been inoculated with feces collected during a hepatitis E outbreak in 1988 in Osh, Kirgizia, U.S.S.R. Bile was collected on the 14th day after infection.

Detection of Anti-HEV Antibodies. A modified ELISA protocol (22) was used in this study. Recombinant HEV antigens for use in the ELISA were derived from Mexican and Burmese strains (23, 24) and were produced in the pGEX1 vector system (25). Cynomolgus sera were diluted 1:100 in 1% gelatin/phosphate-buffered saline (PBS). Alkaline phosphatase-conjugated goat anti-human IgG was used as a second antibody.

Abbreviations: HEV, hepatitis E virus; IEM, immune electron microscopy; RT, reverse transcription; nt, nucleotide(s); ORF, open reading frame.

*To whom reprint requests should be addressed: Laboratory of Infectious Disease, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Building 7, Room 200, 9000 Rockville Pike, Bethesda, MD 20892.

**The sequences reported in this paper have been deposited in the GenBank data base (accession nos. M80581 and M81415).
Primers. Ninety-two primers, 21–40 nucleotides (nt) long, and complementary to plus or minus strands of the genome of a strain of HEV from Burma (BUR-121) (9) or the SAR-55 genome were synthesized using an Applied Biosystems model 391 DNA synthesizer. For cloning of PCR fragments, EcoRI, BamHI, or Bgl II restriction sites preceded by 3–7 nt were added to the 5' end of primers.

For detection of the HEV genome in samples of bile, sera, and feces, two sets of "nested" primers were used that represented sequences from the 3' region (ORF-2) of the SAR-55 genome. Primers for RT and the first PCR were 5'-GTATAACGATCCACATCCCTCC-3' and 5'-TACAGATCTATAAATACGTCGG-3' and for the second PCR were 5'-GCCGGAATCTCCAGACACACATTAGTA-3' and 5'-TAACTGGATCTTTATGCGGACCCTCTTAG-3'.

Preparation of Virus RNA Template for PCR. Bile (10 μl), 20% (wt/vol) SDS (to a final concentration of 1%), proteinase K (10 mg/ml; to a final concentration of 1 mg/ml), 1 μl of tRNA (10 mg/ml), and 3 μl of 0.5 M EDTA were mixed in a final volume of 250 μl and incubated for 30 min at 55°C. Total nucleic acids were extracted from bile twice with phenol/chloroform, 1:1 (vol/vol), at 65°C and once with chloroform, then precipitated by ethanol, washed with 95% ethanol, and used for RT-PCR. RT-PCR amplification of HEV RNA from feces and especially from sera was more efficient when RNA was more extensively purified. Serum (100 μl) or a 10% fecal suspension (200 μl) was treated as above with proteinase K. After a 30-min incubation, 300 μl of CHAOS buffer (4.2 M guanidine thiocyanate/0.5 N-lauroylsarcosine/0.025 M Tris-HCl, pH 8.0) was added. Nucleic acids were extracted twice with phenol/chloroform at 65°C followed by chloroform extraction at room temperature. Then 7.5 mM ammonium acetate (225 μl) was added to the upper phase and nucleic acids were precipitated with 0.68 ml of 2-propanol. The pellet was dissolved in 300 μl of CHAOS buffer and 100 μl of water was added. Chloroform extraction and 2-propanol precipitation were repeated. Nucleic acids were dissolved in water, precipitated with ethanol, washed with 95% ethanol, and used for RT-PCR.

RT-PCR. The usual 100-μl RT-PCR mixture contained template, 10 mM Tris-HCl (pH 8.4), 50 mM KCl, 2.5 mM MgCl₂, all four dNTPs (each at 0.2 mM), 50 pmol of direct primer, 50 pmol of reverse primer, 40 units of RNasin (Promega), 16 units of avian myeloblastosis virus reverse transcriptase (Promega), 4 units of AmpliTaq (Cetus), under 100 μl of light mineral oil. The mixture was incubated 1 h at 42°C and then amplified by 35 PCR cycles: 1 min at 94°C, 1 min at 45°C, and 1 min at 72°C. The PCR products were analyzed on 1% agarose gels.

Cloning of PCR Fragments. PCR fragments containing restriction sites at the ends were digested with EcoRI and BamHI or EcoRI and Bgl II restriction enzymes and cloned in EcoRI/BamHI-digested pBR322 or pGEM-3Z (Promega). Alternatively, PCR fragments were cloned into pCR1000 using the TA cloning kit (Invitrogen, San Diego).

Sequencing of PCR Fragments and Plasmids. PCR fragments were excised from 1% agarose gels and purified by GeneClean (Bio 101, La Jolla, CA). Double-stranded PCR fragments were sequenced by using Sequenase (United States Biochemical) as described (26). Double-stranded plasmids purified through CsCl gradients were sequenced with a Sequenase kit (United States Biochemical).

Computer Analysis of Sequences. Nucleotide sequences of HEV strains were compared using the Genetics Computer Group (Madison, WI) software package (27), version 7.5, on a VAX 8650 computer (at the National Cancer Institute, Frederick, MD). This software package was also used to generate the hydropathy plots.

RESULTS

Biological, Serological, and Virological Characterization of Infection. Six cynomolgus monkeys were intravenously inoculated with a 10% suspension of human feces containing the SAR-55 strain of HEV. Biochemical and serological assays confirmed that all six inoculated animals were infected. Although data are shown for only one monkey (Fig. 1), all six monkeys displayed an increase in alanine aminotransferase, isocitric dehydrogenase, and γ-glutamyltransferase activity, indicative of acute hepatitis. All six animals also developed antibody to recombinant-derived HEV antigens. These data demonstrated that the SAR-55 strain of

![Fig. 1. Evidence for SAR-55 infection of Cyno-376. Serum alanine aminotransferase (ALT) levels are plotted as units/liter and the presence (+) or absence (−) of viral genomes or antibody to HEV is indicated. Viral genomes in the feces and serum were detected by RT-PCR. The presence of total serum antibodies to HEV was monitored by ELISA using recombinant antigens from both the Mexican and Burmese strains of HEV (23, 24).](image-url)
HEV was able to consistently infect and induce hepatitis E in cynomolgus monkeys.

To correlate the distribution of virus with an increase in virus-specific antibody and alanine aminotransferase levels, a sensitive nested RT–PCR protocol was performed on fecal, serum, and bile samples from cynomolgus monkey 376 (Cyno-376). Control experiments in which we assayed dilutions of feces of known infectivity demonstrated that on average the HEV genome could be detected by RT–PCR in feces, with a sensitivity approximately equal to that of an assay based on transmission of HEV to cynomolgus monkeys (unpublished results). All bile, serum, and fecal samples were tested by RT–PCR multiple times to ensure that positive samples were identified.

Although Cyno-376 was inoculated with at least $1 \times 10^3$ infectious doses intravenously, we were not able to detect virus in serum on the day of infection or for several days thereafter (Fig. 1). However, we were able to detect viral genomes in feces and bile as early as day 6 and day 7, respectively, confirming that excretion of virus was an early indicator of infection. All samples of bile collected (days 7–41) were positive for HEV RNA by RT–PCR. Unfortunately, due to the indwelling catheter method of bile collection, residual contamination from earlier samples could not be eliminated so the end point for virus presence in the bile could not be determined by PCR. However, when the less sensitive IEM technique was used in the analysis of the bile, virus was detected sporadically from day 26 to day 37 with a peak on day 32 (data not shown). The HEV particles resembled those described previously (Fig. 2).

Viral excretion in feces could be documented more precisely by PCR, and excretion began on day 6 and ended by day 33 (Fig. 1). Three samples, taken between days 10 and 15, were repeatedly negative for viral RNA. The inability to detect viral RNA in these three samples suggested that the level of excreted particles was low. Virus was first detected in the serum on day 9, appeared to be present at higher titer (based on the quantity of PCR product) from days 14 to 23, then disappeared by day 28 (Fig. 1). A similar pattern of viremia in a cynomolgus monkey infected with HEV was observed by Uchida et al. (28).

**Molecular Characterization of the SAR-55 Genome.** RT–PCR amplification of bile from Cyno-376 was employed to generate six cDNA fragments encompassing the entire genome of SAR-55. The PCR-generated fragments were either sequenced directly or were cloned into individual plasmids and sequenced subsequent to amplification in *Escherichia coli*. Direct sequencing of the PCR product provided 45% of the genome sequence as a consensus sequence. Both strands of cloned cDNAs representing 87% of the genome were sequenced to provide the remainder of the sequence and to confirm the consensus sequence. The sequence of the entire 7.195-kilobase genome, with the exception of 30 and 27 nt of the 5' and 3' termini, respectively, was obtained.

Since one other isolate of HEV, BUR-121 (9), had been totally sequenced, computer analyses were performed to determine the relatedness of SAR-55 to BUR-121. Unique nucleotide insertions or deletions were not detected and the same three ORFs identified in BUR-121 were found in SAR-55. Overall, the genomes were quite similar and differed by only 6.7% in nucleotide sequence and by 1% in deduced amino acid sequence (Table 1). As might be expected, the most conserved region was located in the overlap of ORF-2 and ORF-3 (bases 5147–5477). Against this overall pattern of sequence relatedness, the region between bases 2011 and 2325 in ORF-1 appeared unique (Fig. 3). Although the sum of the nucleotide differences in this region was only 2% higher than the average for the entire genome, the sum of amino acid differences was 14% or 14 times higher than the average. Because the sequence of this region in the SAR-55 genome differed so greatly from that of BUR-121, it was also amplified by RT–PCR directly from the original human feces. The sequence obtained from virus in the human fecal sample was identical to that obtained from virus in the cynomolgus bile.

**Identification of a Hypervariable Region.** To determine if this extreme divergence of sequence between nt 2011 and 2325 was unique to these two isolates or had a more universal significance, sequences from two regions of a third strain of HEV were derived. Fragments of cDNA from the RT–PCR amplification of the OSH-1852 strain of HEV corresponding to positions 2002–2424 and 4424–4800 of the genome were sequenced. Both the nucleotide and amino acid identities in the region (positions 4424–4800) encoding the putative RNA-directed RNA polymerase were high and did not differ substantially among the three strains (Table 2 and Fig. 3). In contrast, between nt 2002 and 2424 there was significant divergence in nucleotide sequences and even greater divergence in amino acid sequences among the three strains.

**Table 1. Comparison of BUR-121 and SAR-55 sequences**

<table>
<thead>
<tr>
<th>Region</th>
<th>Nucleotides</th>
<th>Amino acids</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Length, no.</td>
<td>Identity, %</td>
</tr>
<tr>
<td>ORF-1</td>
<td>5079</td>
<td>93.1</td>
</tr>
<tr>
<td>ORF-2</td>
<td>1980</td>
<td>93.8</td>
</tr>
<tr>
<td>ORF-3</td>
<td>369</td>
<td>98.9</td>
</tr>
<tr>
<td>Total</td>
<td>7138*</td>
<td>93.3</td>
</tr>
</tbody>
</table>

*Length of the primers is not included.
Non structural region (nucleotide binding protein?)

RNA-dependent RNA polymerase region

Fig. 3. Partial cDNA nucleotide sequence and the deduced amino acid sequence of SAR-55. Nucleotides and amino acids that differ in the BUR-121 (boldface type) (9) or OSH-1852 (boldface underlined type) strains are indicated.

Therefore, the region between nt 211 and 2325 probably represents a relatively hypervariable region of the HEV genome.

DISCUSSION

HEV Infection of Cynomolgus Monkeys. Previous studies have shown that cynomolgus monkeys are useful for experimental HEV infection (5, 11-13, 16, 18, 19). Extension of our original study demonstrated that all six cynomolgus monkeys inoculated with human feces containing HEV SAR-55 developed hepatitis E, as defined by liver enzyme elevation, and developed anti-HEV antibodies. Since this virus was able to consistently infect cynomolgus monkeys and cause hepatitis, it appeared to be suitable for selection as a prototype virus for biological and molecular characterization. Therefore, infection of a cynomolgus monkey with the SAR-55 strain of HEV was intensively monitored. Detection of the viral genome in bile and feces suggested that the virus replicated in the liver during the first week after infection. This finding is in good agreement with previous studies in which virus was detected by IEM in feces and bile as early as 9 days after infection (5, 11, 19). Although the monkey was inoculated intravenously with the virus, we were not able to detect virus in serum on

Table 2. Comparison of SAR-55, BUR-121, and OSH-1852 sequences in two regions of ORF-1

<table>
<thead>
<tr>
<th>Region</th>
<th>Nucleotides</th>
<th>Amino acids</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>BUR-121</td>
<td>OSH-1852</td>
</tr>
<tr>
<td>nt 2002-2424</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SAR-55</td>
<td>92.0</td>
<td>92.9</td>
</tr>
<tr>
<td>BUR-121</td>
<td>100</td>
<td>91.0</td>
</tr>
<tr>
<td>nt 4423-4878</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SAR-55</td>
<td>93.2</td>
<td>96.3</td>
</tr>
<tr>
<td>BUR-121</td>
<td>100</td>
<td>92.6</td>
</tr>
</tbody>
</table>

Regions compared are nt 2002-2424 (predicted nonstructural proteins) and nt 4423-4878 (predicted RNA-directed RNA polymerase).
RNA polymerase region could be used for designing PCR primers to detect a spectrum of HEV strains with different origins.

We are grateful to Mrs. Marianne Lewis, Mr. Terry Popkin, and Ms. Katherine Gabor for technical assistance and to Drs. Michael Balayan, Joe Bryan, and Stephen Denny for providing material containing HEV. Excellent animal care was provided by the staff of Bioqual, Gaithersburg, MD. We acknowledge the National Cancer Institute for allocation of computing time and staff support at the Advanced Scientific Computing Laboratory of the Frederick Cancer Research and Development Center. This study was supported in part by a grant from the World Health Organization Programme for Vaccine Development and Contract N01-AI-05069.


**Fig. 4.** Hydropathy plot of the hypervariable region shown in Fig. 3 was plotted for each strain using the method of Kyte and Doolittle (30) and profiles were superimposed for comparison. SAR-55, bold line; BUR-121, dotted line; OSH-1852, dashed line.

day of infection or thereafter until 2–3 days after virus was first detected in bile and feces. Virus excretion started by 1 week after infection and 3 weeks before evidence of disease was observed. The amount of virus excreted at this time was not great, which probably explains why we were not able to detect the HEV genome in fecal samples taken 12–14 days after infection even though samples taken prior to and after this time were positive. Viremia was first observed on the 9th day of the infection and appeared to intensify between 2 and 3 weeks after infection. Elevation of liver enzymes was first noted as viremia ceased. It is probable that viremia is the result of viral replication in the liver although nonhepatic sites of replication have not been ruled out.

**Comparison of SAR-55 Genome with That of Other Strains.** We found that the sequence of SAR-55 was quite similar to that of BUR-121 over the entire length of the genome (Table 1). This result is not unexpected since the two viruses were isolated from geographically contiguous regions at about the same time. Differences between the two sequences were distributed approximately uniformly throughout the genome except for one hypervariable domain located in the putative nonstructural ORF 9 (Fig. 3 and Table 2). This region of hypervariability preceded the nucleotide triphosphate binding domain and also the RNA-directed RNA polymerase motifs recognized in ORF-1 (8, 9, 29). That this region was hypervariable was confirmed when another strain from Asia was partly sequenced. The genome of the OSH-1852 strain from Kirgizia differed in the hypervariable region to approximately the same extent from the Pakistani and Burmese strains as those strains differed from each other (Table 2). However, all three strains did not differ significantly from each other in the region coding for RNA-directed RNA polymerase. The finding of a hypervariable region in a presumed nonstructural protein is surprising, since such regions are usually found in structural proteins of viruses, where they are exposed to selection by the immune response of the host. A hydropathy analysis of the region revealed that all three strains displayed a similar profile in spite of their extensive sequence differences (Fig. 4). However, the significance of this pattern remains to be determined. We predict that the existence of this hypervariable region will allow even closely related strains to be distinguished by the application of molecular epidemiological approaches. Presumably, the sequence of the hypervariable region could be used for more detailed molecular classification of HEV strains, and the sequence from the much more conserved...