Molecular Cloning and Characterization of Hepatitis A Virus cDNA

John R. Ticehurst, Vincent R. Racaniello, Bahige M. Baroudy, David Baltimore, Robert H. Purcell, and Stephen M. Feinstone

This information is current as of December 2006.

Notes:

Molecular cloning and characterization of hepatitis A virus cDNA

(hepatitis A virus RNA/picornavirus)

John R. Ticehurst*, Vincent R. Racaniello†‡, Bahige M. Baroudy*, David Baltimore†,
Robert H. Purcell*, and Stephen M. Feinstone*

*Laboratory of Infectious Diseases, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Bethesda, MD 20205; and †Whitehead
Institute for Biomedical Research, Center for Cancer Research and Department of Biology, Massachusetts Institute of Technology, Cambridge, MA 02139

ABSTRACT   Double-stranded cDNA was synthesized from
hepatitis A virus (HAV) RNA and inserted into the *Pst* I site of
pBR322. Restriction endonuclease digestion and cross-hybridiza-
tion of fragments yielded a map of overlapping cloned cDNAs that
included at least 99% of the viral genome. Molecular clones con-
taining HAV cDNA were identified by hybridizing cloned cDNA
to electrophoretically resolved RNA from uninfected and HAV-in-
fected tissue culture cells. Cloned cDNA probes specifically hy-
bridized to RNA from infected cells, and the predominant species
identified had the characteristic genomic length of picornaviral
RNA ($\approx$7,500 nucleotides). A partial sequence from the 3' end of
the genome revealed 414 bases in an open reading frame followed
by two closely spaced stop codons, a 60-base noncoding region,
and a tract of poly(A).

Hepatitis A virus (HAV) is an important cause of human hep-
atitis. In the United States it has been estimated that >100,000
clinical cases occur annually (D. P. Francis, personal commu-
nication). HAV continues to be endemic in underdeveloped areas
of the world where infections usually occur in children, and
nearly all of the young adult population have antibody to HAV
(1). Clinical hepatitis A and prevalence of antibody to HAV are
decreasing in industrialized nations, resulting in increasing
numbers of adults susceptible to infection. Groups at risk in-
clude travelers to endemic areas, consumers of raw shellfish,
male homosexuals, and contacts of very young children in day
care centers.

The host range of HAV is limited to man, apes (especially the
chimpanzee), and several species of New World monkeys (2).
Viral replication is limited to the liver, and excretion into the
stool appears to occur via the biliary system. The 27-nm virion
of HAV was first observed in 1973 (3) and first isolated in tissue
culture in 1979 (4). Its cultivation remains difficult, impeding
analysis of molecular structure and events in replication.

HAV has been classified as a picornavirus (5). It has a sedi-
mentation coefficient of approximately 160 S and a primary
buoyant density of 1.34 g/ml in CsCl. Virion capsid polypep-
tides of M_r 32,000, 26,000, 22,000, and 10,000 have been de-
scribed. The single-stranded infectious RNA genome is 6,700
to 8,100 nucleotides long and contains poly(A), presumably at
the 3' terminus. By analogy with other picornaviruses the RNA
should contain an open reading frame of about 6,500 nucleo-
tides, which directs synthesis of a polyprotein that is post-trans-
lationally cleaved into virion proteins. These include the four
capsid proteins, a peptide linked to the 5' end of the genome
(VPg), a RNA-dependent RNA polymerase, and a protease (6–
9).

Previously, von der Helm *et al.* (10) described cloned cDNAs
of up to 1,000 base pairs that hybridized to RNA extracted from

HAV labeled in tissue culture. We report here the cloning of
cDNA representing at least 99% of the genome of HAV and
partial analysis of its sequence.

MATERIALS AND METHODS

Source and Propagation of HAV. The HM-175 strain of HAV
was isolated from a family outbreak in Australia (11). HAV that
had been passaged twice in marmosets (*Saguinus mystax* and
Saguinus labiatus) (11) was inoculated into eight marmosets. A
20% (wt/vol) suspension of HAV-infected liver in phosphate-
buffered saline was injected intravenously, and animals were
monitored by serum enzymes and immunofluorescence of liver
biopsy (12). When maximum immunofluorescence was reached
10–14 days after inoculation, the marmosets were killed and
their livers were immediately removed, minced, frozen in liq-
uid N_2, and stored at $-70°C$ until virus purification.

HAV was also passaged 6 times in marmosets and 19 times
in secondary African green monkey kidney (AGMK) monolay-
ers (11), which were used as a source of RNA for hybridization
(see below). Infected cultures were harvested at 21 days when
100% of the cells exhibited maximal immunofluorescence.

Virus Purification. Minced marmoset livers (total, 170 g) were
homogenized, extensively extracted with trichlorotrifluoroeth-
ane and chloroform, and treated with micrococcal nuclease (13).
Virus was monitored by radioimmunoassay (14) and by electron
microscopy during successive sucrose, CsCl, and sucrose gra-
dients.

RNA Extraction and Characterization. Suspensions of pu-
rified HAV were incubated at 37°C for 15 min with 500 μg of
proteinase K per ml, after which NaDodSO$_4$ was added to a
concentration of 0.5%, and incubation was continued for an ad-
ditional 30 min. After extraction with phenol and then with
chloroform/isoamyl alcohol (24:1, vol/vol), RNA was precipi-
tated in ethanol, redissolved, and, after removal of a portion for
analysis, reprecipitated. RNA was characterized by UV spec-
troscopy and agarose gel electrophoresis after denaturation with
1 M glyoxal/50% dimethyl sulfoxide (vol/vol) (15). Other por-
tions were used to analyze template quality and optimal con-
ditions for cDNA synthesis as described in *Results* and Fig. 1.

RNA for hybridization studies was prepared from uninfected
and HAV-infected (see above) AGMK cells by isolation of total
cytoplasmic RNA (16). Polyadenylylated RNA was isolated by
oligo(dT)-cellulose chromatography (17). RNA was transferred
to nitrocellulose paper after electrophoretic separation through
agarose gels containing glyoxal/dimethyl sulfoxide-denatured
RNA (18).

Preparation of cDNA Clones. HAV RNA (0.8 μg) derived from marmoset liver served as a template for cDNA synthesis using reverse transcriptase (120 units/ml) for 30 min at 42.5°C in 160 μl that contained 50 mM Tris·HCl (pH 8.3)/10 mM MgCl$_2$/100 mM KCl/500 μM dATP/[α-^{32}P]dCTP (0.025 Ci/mmol; 1 Ci = 3.7 × 10^{10} Bq)/dGTP/TTP/1 mM dithioerythritol/4 mM sodium pyrophosphate/(dT)$_{12-18}$ at 30 μg/ml/RNasin (2,000 units/ml). After addition of EDTA to 20 mM, RNA·cDNA hybrids were isolated by phenol extraction, column chromatography, and ethanol precipitation. The RNA template was hydrolyzed in 0.3 M NaOH/0.7 M NaCl/5 mM EDTA for 2 hr at 37°C (8).

By using a modification of conditions optimized for poliovirus cDNA template (8), the second strand of cDNA was synthesized for 30 min at 37°C using the large (Klenow) fragment of *Escherichia coli* DNA polymerase I (28 units/ml) in 10 mM Tris·HCl (pH 7.5)/5 mM MgCl$_2$/5 mM dithioerythritol/50 μM dATP/[α-^{32}P]dCTP (0.45 Ci/mmol)/dGTP/TTP, and cDNA at 1 μg/ml.

After phenol extraction, column chromatography, and ethanol precipitation, double-stranded cDNA (ds cDNA) was digested for 1 hr at 37°C using nuclease S1 at 10 units/ml (0.1 unit per ng of ds cDNA) in 30 mM NaOAc (pH 4.5)/0.3 M NaCl/3 mM ZnCl$_2$/5% glycerol (vol/vol), followed by addition of EDTA to 35 mM, phenol and ether extraction, and dialysis against 10 mM Tris·HCl (pH 7.5) and 1 mM EDTA (8).

Homopolymer tails of dCMP were added to ds cDNA using terminal deoxynucleotidyltransferase (250 units/ml) for 20 min at room temperature in 100 μl that contained 0.14 M potassium cacodylate (pH 7.2)/1 mM CoCl$_2$/0.2 mM dithioerythritol/nuclease-free bovine serum albumin (500 μg/ml)/200 μM dCTP (8). After phenol extraction, 50% of the ds cDNA was ether-extracted and precipitated with ethanol.

The remaining tailed ds cDNA was applied to a 3-ml column of Sepharose 4B in 20 mM Tris·HCl (pH 8.0)/0.6 M NaCl/2 mM EDTA (16). The first five 65-μl fractions containing ds cDNA were pooled and precipitated in ethanol after the addition of 2 μg of yeast tRNA.

Plasmid vector pBR322, cleaved at the *Pst* I site and tailed with dGMP, was hybridized to equimolar amounts of both tailed ds cDNA preparations and used to transform *E. coli* HB101 by standard procedures (16).

Analysis of cDNA Clones. Clones containing putative HAV sequences were screened by cleaving recombinant plasmid preparations with *Pst* I and sizing by gel electrophoresis. Cloned cDNA inserts isolated from low melting point agarose (19) were labeled by nick-translation and used as probes in hybridization (*i*) to electrophoretically separated RNA bound to nitrocellulose paper (described above) for establishing the identity of cloned cDNA species, (*ii*) to DNA bound to nitrocellulose paper after lysis of bacterial colonies *in situ* for further screening, and (*iii*) to restriction fragments of DNA resolved by electrophoresis and bound to nitrocellulose paper for confirmation of tentative restriction maps constructed on the basis of single and double enzyme digests. Extent of RNA sequence not represented in cDNA clones was determined by primer extension of HAV-infected AGMK cytoplasmic RNA (8). Sequence determination of cloned cDNA was by the method of Maxam and Gilbert (20).

Materials. The following sources were used: AGMK cells, Flow Laboratories (Rockville, MD); micrococcal nuclease and large fragment of *E. coli* DNA polymerase I, Boehringer Mannheim; proteinase K, Merck (Darmstadt, Federal Republic of Germany); HPLC grade deoxynucleoside triphosphates, ICN; [α-^{32}P]dCTP, New England Nuclear or Amersham; RNasin, JEM Research (Kensington, MD); reverse transcriptase, J. W. Beard, Life Sciences (St. Petersberg, FL), through the National Can-

cer Institute; nuclease S1, P-L Biochemicals; terminal deoxynucleotidyltransferase, nuclease-free bovine serum albumin, nick-translation kits, and certain restriction endonucleases, Bethesda Research Laboratories; other restriction enzymes, New England BioLabs; Sepharose 4B, Pharmacia; and oligo(dG)-tailed pBR322, New England Nuclear. ^{125}I-Labeled anti-HAV from the HAVAB kit (Abbott) was used for detection of HAV by radioimmunoassay (14).

RESULTS

Virus Purification and RNA Characterization. Direct electron microscopic examination of purified HAV particles revealed a homogeneous population of 27-nm virions. The yield of virion RNA from eight marmoset livers was approximately 1.0 μg and the A_{260}/A_{280} ratio was 2.0. Samples used for UV spectroscopy were also denatured and analyzed by electrophoresis through an agarose gel (15). HAV RNA contained a discrete band comigrating with poliovirus type 1 RNA (approximately 7,440 nucleotides; refs. 7–9).

Analytic cDNA Synthesis. RNAs from several sources of HAV and from poliovirus type 2 were compared for template quality in cDNA synthesis under conditions optimal for poliovirus RNA (Fig. 1). HAV RNA derived from liver or AGMK culture yielded a series of transcripts ranging in size from slightly smaller than the longest poliovirus type 2 cDNA to less than 500 nucleotides (Fig. 1, lanes a, b, and b'; lanes e and e' show poliovirus type 2 cDNA). Presumably, degradation of HAV RNA prevented more extensive synthesis of reverse transcript approaching the expected full length of 7,500 nucleotides. Other HAV cDNAs in Fig. 1 (lanes c and d) revealed evidence of more extensive RNA
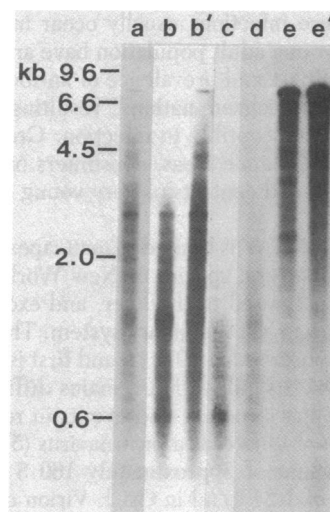

FIG. 1. Alkaline agarose gel electrophoresis of cDNA from HAV RNA and poliovirus type 2 RNA. By using conditions optimal for poliovirus cDNA synthesis (8), RNAs (5 μg/ml or less) were incubated for 60 min at 42.5°C in 10 μl containing 50 mM Tris·HCl (pH 8.3)/10 mM MgCl$_2$/50 mM KCl/500 μM dATP/[α-^{32}P]dCTP (2 Ci/mmol)/dGTP/TTP/0.4 mM dithioerythritol/4 mM sodium pyrophosphate/(dT)$_{12-18}$ (30 μg/ml)/reverse transcriptase (80 units/ml). Two reactions (cDNAs in lanes b' and e') also contained RNasin (2,000 units/ml). Reverse transcripts from RNA templates: lane a, HAV RNA derived from marmoset liver; lanes b and b', HAV RNA from AGMK cells; lane c, HAV RNA from human stool; lane d, HAV RNA from AGMK cells; lanes e and e', poliovirus type 2 RNA. HAV RNAs were isolated as described in *Materials and Methods* except that derived from human stool (lane c), which was extracted from purified virus (a gift from S. Locarnini), and an earlier preparation from AGMK cells (lane d) using modifications of standard procedures (5). Migration of λ *Hind*III fragments (in kb) is indicated on the left.

degradation, but all gave a similar banding pattern that was different from that of poliovirus cDNAs.

The effect of varying several chemical constituents and physical parameters on yield and length of HAV cDNA was analyzed by alkaline agarose gel electrophoresis and incorporation of [^{32}P]dCMP into trichloroacetic acid-precipitable product (data not shown). Incubation for 30 min, or with reverse transcriptase at 120 units/ml, or with 100 mM KCl increased size and quantity of cDNA when compared to that shown in Fig. 1 (lane b'). Denaturation of HAV RNA was not attempted because quantity was limited and, in earlier experiments, heat or methylmercury treatment of poliovirus RNA decreased yield and size of cDNA.

Preparation of cDNA Clones. A significant portion of HAV cDNA transcribed under preparative conditions was 3,000–7,500 nucleotides long (Fig. 2, lane a). However, a wide size range of ds cDNA molecules was produced (Fig. 2, lane b), and most of the preparation was less than 2,000 nucleotides long after nuclease S1 digestion (lane c). Approximately 200 ng of ds cDNA was synthesized. Ten nanograms of ds cDNA selected for large size by gel filtration (*Materials and Methods*) yielded 232 tetracycline-resistant *E. coli* transformants. Cleavage with *Pst* I demonstrated inserts of 1,000 base pairs or more in 43 of the 232 recombinant plasmids, designated pHAV$_{LB}$. An additional 2,710 clones were obtained from 9 ng of unfractionated ds cDNA. From this group, only pHAV$_L$ 1307 (described below) was extensively characterized.

Identity of Cloned cDNA. The identity of inserted DNAs in recombinant plasmids was established by hybridization to RNA bound to nitrocellulose paper after gel electrophoresis (Fig. 3). A nick-translated probe prepared from the insert of pHAV$_{LB}$ 39 specifically hybridized to RNA from HAV-infected AGMK cells (Fig. 3, lanes c' and e). Similar results (not shown) were obtained when the inserts of pHAV$_{LB}$ 93 or pHAV$_{LB}$ 228 were used as probes. The predominant band identified had the size expected for genomic HAV RNA and comigrated with poliovirus type 2 RNA. Diffuse hybridization to lanes containing RNA from infected cells was probably due to RNA degradation. Nick-translated pBR322 did not hybridize to any RNA species from either HAV-infected or uninfected AGMK cells (data not shown), thereby eliminating the possibility that a small amount of pBR322 contaminating the insert probes was responsible for specific hybridization. None of the pHAV$_{LB}$ probes tested to date hybridized to poliovirus RNA.

Restriction Map of Cloned HAV cDNAs. The restriction map shown in Fig. 4 is based on data obtained from digests and on hybridization of labeled inserts to fractionated DNA (data not shown). Hybridization of an insert fragment from pHAV$_{LB}$ 228

[corresponding to the region from 2.4 to 3.0 kilobases (kb) in Fig. 4] to DNA from bacterial colonies was used to select pHAV$_L$ 1307 (and pHAV$_{LB}$ 12 and pHAV$_{LB}$ 153, described in the legend to Fig. 4) for further analysis. Plasmids pHAV$_{LB}$ 113, pHAV$_L$ 1307, pHAV$_{LB}$ 228, pHAV$_{LB}$ 148, and pHAV$_{LB}$ 207 overlap to generate a map of about 7.4 kb, which represents at least 99% of the genome of HAV. Inserts from other clones (as indicated in Fig. 4) provided confirmation of overlap regions except for that between pHAV$_{LB}$ 228 and pHAV$_{LB}$ 148. Nick-translated pHAV$_{LB}$ 228 insert hybridized to the 1.5-kb *Pst* I fragment of pHAV$_{LB}$ 148. The overlap as drawn is based on the assumption that there is only one *Hinc*II site, present in both clones, in the area about 4 kb from the 5' terminus of the genome. No other restriction sites in this region were found by digestion with >30 other enzymes that have five- and six-base recognition sites. As noted in the legend to Fig. 4, heterogeneity in one restriction site (*Nde* I) was detected in the region about 2.5 kb from the 5' end of the viral RNA.

Primer Extension of Cloned cDNA. A labeled *Pst* I/*Bam*HI fragment (0.2–0.7 kb in Fig. 4) from pHAV$_{LB}$ 113 was used as a primer for cDNA synthesis. When the template was cytoplasmic RNA from HAV-infected AGMK cells, the primer was extended to give a discrete band (data not shown). Assuming that the 5' termini from HAV RNAs of different passage levels are conserved, we calculated that cloned cDNA extended to within 50 bases of the 5' end of the genome. There was no primer extension when RNA from uninfected AGMK cells was used as a template.

Partial Sequence of the HAV Genome. The sequence of cloned cDNA corresponding to the 3' end of HAV RNA is presented in Fig. 5. There is an open reading frame that extends from the beginning of the sequence to a pair of nonsense codons that are separated by six bases (bases 415–426). The second nonsense codon is followed by 51 bases that do not contain
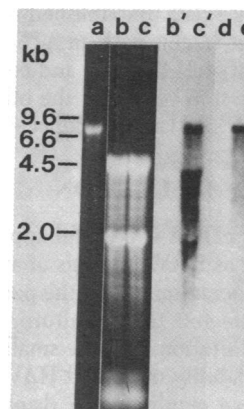
FIG. 3. Hybridization of cloned cDNA insert probe to RNA. Nucleic acids were denatured and electrophoresed (15); marker lanes (a–c) were removed and stained with ethidium bromide (2 μg/ml) in 0.05 M NaOH for 40 min, followed by ethidium bromide (0.5 μg/ml) in 40 mM Tris·HOAc (pH 7.8)/1 mM EDTA for 15 min and destaining in the same buffer for 15 min. Lanes b'–e were transferred to nitrocellulose paper (18) and hybridized for 36 hr at 42°C with [^{32}P]pHAV$_{LB}$ 39 insert (25 ng/ml; 5 × 10^8 dpm/μg) in 50% formamide (vol/vol)/0.75 M NaCl/0.075 M trisodium citrate/50 mM sodium phosphate (pH 6.5)/0.2% NaDodSO$_4$/denatured sheared salmon sperm DNA (100 μg/ml)/0.04% bovine serum albumin/polyvinyl pyrollidone/Ficoll. Lane a contained 400 ng of poliovirus type 2 RNA. Lanes b, b', and d contained cytoplasmic RNA from uninfected AGMK cells: lane b, 5 μg; lane b', 20 μg; and lane d, 1.5 μg, oligo(dT)-selected. Lanes c, c', and e contained cytoplasmic RNA from HAV-infected AGMK cells: lane c, 5 μg; lane c', 20 μg; and lane e, 1.5 μg, oligo(dT)-selected. Autoradiographic exposure for lanes b' and c' was one-eighth of that for lanes d and e. Migration of λ *Hin*dIII fragments (in kb) is indicated on the left.
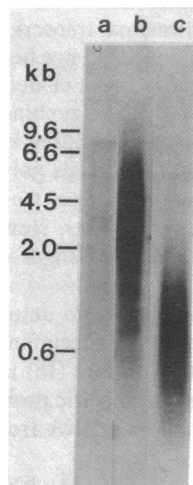

FIG. 2. Alkaline agarose gel electrophoresis of products of preparative HAV ds cDNA synthesis. Lane a, ≈3 ng of cDNA; lane b, ≈2 ng of ds cDNA prior to nuclease S1 digestion; lane c, ≈2 ng of ds cDNA after nuclease S1 digestion. Migration of λ *Hin*dIII fragments (in kb) is indicated on the left.
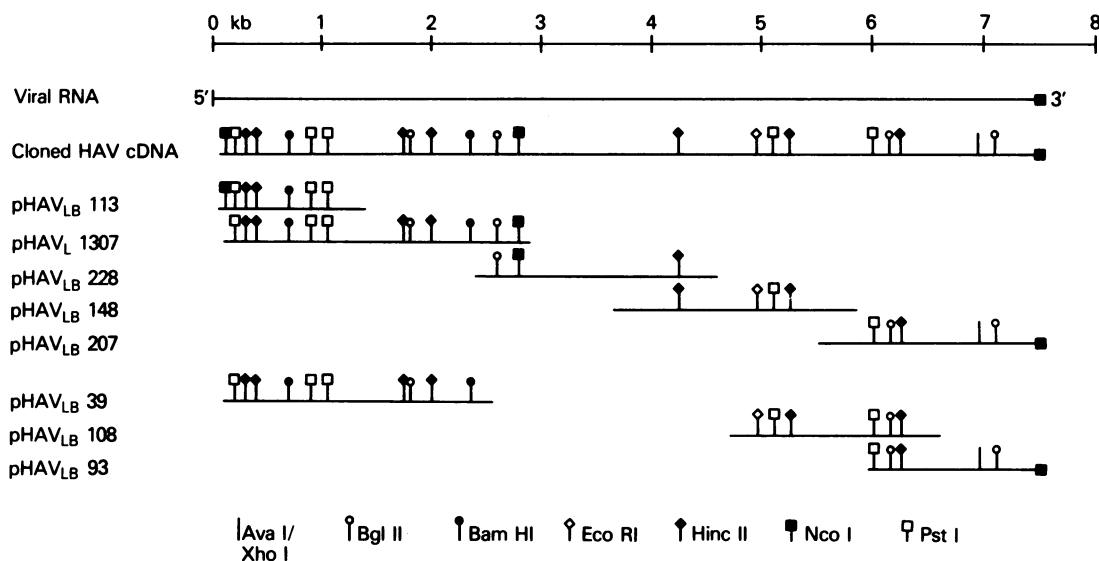
FIG. 4. Restriction map of HAV cDNA clones. Viral RNA is estimated to be 7,450 nucleotides long excluding 3' poly(A), which is shown at the right (thicker line). A composite map of cloned HAV cDNA and positions of the inserts from clones pHAV$_{LB}$ 113, pHAV$_L$ 1307, pHAV$_{LB}$ 228, pHAV$_{LB}$ 148, and pHAV$_{LB}$ 207 are shown immediately below the viral RNA. Recombinant plasmids pHAV$_{LB}$ 39, pHAV$_{LB}$ 108, and pHAV$_{LB}$ 93 have been used for DNA sequence determination, to confirm regions of overlap, or to prepare insert probes. Other plasmids that support this map include pHAV$_{LB}$ 153 (position 1.8–2.7 kb, relative to scale), pHAV$_{LB}$ 12 (2.3–3.3 kb), pHAV$_{LB}$ 87 (4.1–5.8 kb), pHAV$_{LB}$ 201 (5.2–6.7 kb), pHAV$_{LB}$ 38 (5.2–6.9 kb), pHAV$_{LB}$ 56 (5.4–7.1 kb), pHAV$_{LB}$ 185 (5.5–7.5 kb), pHAV$_{LB}$ 24 (5.6–7.5 kb), and pHAV$_{LB}$ 122 (5.7–7.5 kb). Restriction sites have also been mapped with *Acc* I (2.0 kb), *Apa* I (5.7 and 6.2 kb), *Ava* II (3.5 and 6.6 kb), *Bst*EII (1.2 kb), *Hind*III (2.2 kb), *Hpa* I (0.4 kb), *Nde* I (1.2 kb), *Pvu* II (0.8, 1.0, 3.0, 5.2, 5.5, and 7.1 kb), *Sac* I (3.0 kb), and *Xba* I (0.8 kb). There is also a site for *Nde* I (2.5 kb) in pHAV$_{LB}$ 12, pHAV$_{LB}$ 153, and pHAV$_L$ 1307, which is not present in pHAV$_{LB}$ 39. There are no sites for *Bgl* I, *Eco*RV, *Hae* II, *Kpn* I, *Mlu* I, *Nae* I, *Nar* I, *Nru* I, *Pvu* I, *Sac* II, *Sal* I, *Sma* I, *Sph* I, *Stu* I, and *Tth*111I.

an open reading frame initiated by an ATG, followed by poly(A), presumed to represent the 3' end of the viral genome. The longest possible polypeptide encoded in the second and third reading frames is 29 amino acids.

Preliminary sequence data (unpublished results) indicate that an open reading frame begins with an ATG approximately 750 bases from the 5' end of HAV RNA and continues for at least 378 bases. Numerous stop codons in the other reading frames preclude a significant translation product.

DISCUSSION

Molecular cloning of cDNA is a useful approach for the study of RNA viruses such as HAV. Analysis of virion structure and HAV replication has been limited by the persistence of cellular functions in HAV-infected tissue culture, the long time required for virus proliferation, and the small quantities of HAV obtainable. The availability of cloned HAV cDNA, made possible by using efficient conditions for ds cDNA synthesis that were well characterized for another picornaviral RNA (8), should be useful for deducing RNA and amino acid sequences and elucidating the events of replication.

In an effort to produce cDNA clones that contained only HAV sequences, HAV RNA was extracted from highly purified virions. HAV RNA probably represents a very small proportion of nucleic acid from potential sources. The virus appears, from electron microscopic (21) and biophysical data (unpublished data), to be associated with intracellular membranous vesicles *in vivo* and *in vitro*. It is possible that nonviral nucleic acid species could contaminate HAV preparations by copurifying with the virus or associated membranes. In earlier experiments the template for cDNA was RNA extracted from HAV that had been incompletely purified from marmoset liver. The cDNA clones derived from this RNA were found to contain nonviral sequences that were presumably of host origin (unpublished data). In the

experiments described here, liver homogenates were hydrocarbon-extracted until an aqueous–organic interphase layer was no longer detected. Micrococcal nuclease digestion was used to eliminate unencapsidated nucleic acids. In addition, any viral probe (labeled RNA or cDNA) used to screen clones might be as contaminated with nonviral sequences as the template used for cloning. Instead of using such a probe, pHAV inserts were characterized by restriction analysis, cross-hybridization, and hybridization to HAV RNA contained in RNA from a heterologous source.

It would be difficult at present to directly determine the sequence of HAV RNA as a reference, but analysis of cloned cDNA sequences described here is consistent with a genome that has the structural organization of other picornaviruses but does not show significant homology to members of this group (7–9, 22). Although there is only one known serotype of HAV, a restriction site variation between HAV cDNA inserts was detected. It is impossible to determine whether the alteration is due to RNA sequence heterogeneity or an error in reverse transcription. However, it would not be surprising to find sequence heterogeneity, because the HAV used in this study was not cloned. The *Nde* I site variation (2.5 kb from the 5' end of the genome) occurs in an area that corresponds to the sequence coding for the NH$_2$ terminus of VP1 in poliovirus RNA (7–9). This polypeptide is thought to be important in picornaviral surface structure and in immunity (23–25). In other picornaviruses, strain and serotype differences are often manifested by changes in VP1 RNA sequence (9, 26, 27).

It may be feasible to use HAV cDNA as a probe to detect viral or virus-related sequences from a variety of specimens. Experiments (data not shown) using a dot blot assay (18) indicate that cloned HAV cDNA is a sensitive and specific probe for detecting viral RNA molecules in cytoplasmic RNA from tissue culture.

Because many of the same limitations for obtaining the RNA

```
                           27                                              54
TCT CGA GAT GTT CAG ATT GAT AAT CTT GAT TTG ATT GGA CAA AAA ATT GTA GAT
Ser Arg Asp Val Gln Ile Asp Asn Leu Asp Leu Ile Gly Gln Lys Ile Val Asp
 Leu Glu MET Phe Arg Leu Ile Ile Leu Ile  .  Leu Asp Lys Lys Leu  .  MET
   Ser Arg Cys Ser Asp  .   .  Ser  .  Phe Asp Trp Thr Lys Asn Cys Arg  .

                          81                                             108
GAG TTT AAG AAA CTT GGC ATG ACA GCT ACT TCT GCT GAC AAG AAT GTA CCT CAG
Glu Phe Lys Lys Leu Gly MET Thr Ala Thr Ser Ala Asp Lys Asn Val Pro Gln
 Ser Leu Arg Asn Leu Ala  .  Gln Leu Leu Leu Leu Thr Arg MET Tyr Leu Ser
   Val  .  Glu Thr Trp His Asp Ser Tyr Phe Cys  .  Gln Glu Cys Thr Ser Ala

                          135                                            162
CTG AAA CCA GTT TCG GAA TTG ACT TTT CTC AAA AGA TCT TTC AAT TTG GTA GAG
Leu Lys Pro Val Ser Glu Leu Thr Phe Leu Lys Arg Ser Phe Asn Leu Val Glu
 .  Asn Gln Phe Arg Asn  .  Leu Phe Ser Lys Asp Leu Ser Ile Trp  .  Arg
   Glu Thr Ser Phe Gly Ile Asp Phe Ser Gln Lys Ile Phe Gln Phe Gly Arg Gly

                          189                                            216
GAT AGA ATT AGA CCT GCA ATT TCG GAA AAA ACA ATT TGG TCT TTA ATA GCA TGG
Asp Arg Ile Arg Pro Ala Ile Ser Glu Lys Thr Ile Trp Ser Leu Ile Ala Trp
 Ile Glu Leu Asp Leu Gln Phe Arg Lys Lys Gln Phe Gly Leu  .   .  His Gly
   .  Asn  .  Thr Cys Asn Phe Gly Lys Asn Asn Leu Val Phe Asn Ser MET Ala

                          243                                            270
CAG AGA AGT AAC GCT GAG TTT GAG CAG AAT TTA GAA AAT GCT CAG TGG TTT GCT
Gln Arg Ser Asn Ala Glu Phe Glu Gln Asn Leu Glu Asn Ala Gln Trp Phe Ala
 Arg Glu Val Thr Leu Ser Leu Ser Arg Ile  .  Lys MET Leu Ser Gly Leu Leu
   Glu Lys  .  Arg  .  Val  .  Ala Glu Phe Arg Lys Cys Ser Val Val Cys Phe

                          297                                            324
TTT ATG CAT GGC TAT GAG TTT TAT CAG AAA TTT TAT TAT TTT GTT CAG TCC TGT
Phe MET His Gly Tyr Glu Phe Tyr Gln Lys Phe Tyr Tyr Phe Val Gln Ser Cys
 Leu Cys MET Ala MET Ser Phe Ile Arg Asn Phe Ile Ile Leu Phe Ser Pro Val
   Tyr Ala Trp Leu  .  Val Leu Ser Glu Ile Leu Leu Phe Cys Ser Val Leu Phe

                          351                                            378
TTG GAG AAA GAG ATG ATA GAA TAC AGA CTT AAA TCT TAT GAT TGG TGG AGA ATG
Leu Glu Lys Glu MET Ile Glu Tyr Arg Leu Lys Ser Tyr Asp Trp Trp Arg MET
 Trp Arg Lys Arg  .   .  Asn Thr Asp Leu Asn Leu MET Ile Gly Gly Glu  .
   Gly Glu Arg Asp Asp Arg Ile Gln Thr  .  Ile Leu  .  Leu Val Glu Asn Glu

                          405                                            432
AGA TTT TAT GAC CAG TGT TTC ATT TGT GAC CTT TCA TGA TTT GTT TAA ACA AAT
Arg Phe Tyr Asp Gln Cys Phe Ile Cys Asp Leu Ser  .  Phe Val  .  Thr Asn
 Asp Phe MET Thr Ser Val Ser Phe Val Thr Phe His Asp Leu Phe Lys Gln Ile
   Ile Leu  .  Pro Val Phe His Leu  .  Pro Phe MET Ile Cys Leu Asn Lys Phe

                          459                                            486
TTT CTT AAA ATT TCT GAG GTT TGT TTA TTT CTT TTA TCA GTA AAT AAA AAA AAA
Phe Leu Lys Ile Ser Glu Val Cys Leu Phe Leu Leu Ser Val Asn Lys Lys Lys
 Phe Leu Lys Phe Leu Arg Phe Val Tyr Phe Phe Tyr Gln  .  Ile Lys Lys Lys
   Ser  .  Asn Phe  .  Gly Leu Phe Ile Ser Phe Ile Ser Lys  .  Lys Lys Lys

                          513
AAA AAA
Lys Lys
 Lys
   Lys
```

FIG. 5. Nucleotide sequence of cloned cDNA corresponding to the HAV genome 3' terminus. Sequence was determined for the insert from pHAV_LB 93 (20), starting near the *Ava* I/*Xho* I site (Fig. 4). Possible translation products are listed in all three reading frames below the nucleotide sequence, with the frame probably used for viral protein uppermost. Following poly(A) (bases 478–492), poly(G) links cloned HAV cDNA to pBR322 (data not shown).

of HAV apply to its proteins, there is no information about amino acid sequences and little is known about the immunologic determinants that elicit the polyclonal antibody response to infection. Antisera directed against specific HAV polypeptides have not yet been identified. It may be possible to prepare such antisera from synthetic peptides or from cDNAs expressed in prokaryotes (28) or eukaryotes (29) as has been done for other picornaviruses (24, 25). Immunogenic peptides or polypeptides would have obvious potential for vaccine development.

1. Gust, I. D. (1982) in *Viral Hepatitis, 1981 International Symposium*, eds. Szmuness, W., Alter, H. J. & Maynard, J. E. (Franklin Institute Press, Philadelphia), pp. 129–143.
2. Holmes, A. W., Wolfe, L., Rosenblate, H. & Deinhardt, F. (1969) *Science* 165, 816–817.
3. Feinstone, S. M., Kapikian, A. Z. & Purcell, R. H. (1973) *Science* 182, 1026–1028.
4. Provost, P. J. & Hilleman, M. R. (1979) *Proc. Soc. Exp. Biol. Med.* 160, 213–221.
5. Coulepis, A. G., Locarnini, S. A., Westaway, E. G., Tannock, G. A. & Gust, I. D. (1982) *Intervirology* 18, 107–127.
6. Putnak, J. R. & Phillips, B. A. (1981) *Microbiol. Rev.* 45, 287–315.
7. Kitamura, N., Semler, B. L., Rothberg, P. G., Larsen, G. R., Adler, C. J., Dorner, A. J., Emini, E. A., Hanecak, R., Lee, J. J., van der Werf, S., Anderson, C. W. & Wimmer, E. (1981) *Nature (London)* 291, 547–553.
8. Racaniello, V. R. & Baltimore, D. (1981) *Proc. Natl. Acad. Sci. USA* 78, 4887–4891.
9. Nomoto, A., Omata, T., Toyoda, H., Kuge, S., Horie, H., Kataoka, Y., Genba, Y., Nakano, Y. & Imura, N. (1982) *Proc. Natl. Acad. Sci. USA* 79, 5793–5797.
10. von der Helm, K., Winnacker, E. L., Deinhardt, F., Frösner, G. G., Gauss-Müller, V., Bayerl, B., Scheid, R. & Siegl, G. (1981) *J. Virol. Methods* 3, 37–43.
11. Daemer, R. J., Feinstone, S. M., Gust, I. D. & Purcell, R. H. (1981) *Infect. Immun.* 32, 388–393.
12. Mathiesen, L. R., Feinstone, S. M., Purcell, R. H. & Wagner, J. (1977) *Infect. Immun.* 18, 524–530.
13. Pelham, H. R. B. & Jackson, R. J. (1976) *Eur. J. Biochem.* 67, 247–256.
14. Purcell, R. H., Wong, D. C., Moritsugu, Y., Dienstag, J. L., Routenberg, J. A. & Boggs, J. D. (1976) *J. Immunol.* 116, 349–356.
15. McMaster, G. K. & Carmichael, G. G. (1977) *Proc. Natl. Acad. Sci. USA* 74, 4835–4838.
16. Maniatis, T., Fritsch, E. F. & Sambrook, J. (1982) *Molecular Cloning* (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY).
17. Varmus, H. E., Quintrell, N. & Ortiz, S. (1981) *Cell* 25, 23–36.
18. Thomas, P. S. (1980) *Proc. Natl. Acad. Sci. USA* 77, 5201–5205.
19. Cummings, I. W., Browne, J. K., Salser, W. A., Tyler, G. V., Snyder, R. L., Smolec, J. M. & Summers, J. (1980) *Proc. Natl. Acad. Sci. USA* 77, 1842–1846.
20. Maxam, A. M. & Gilbert, W. (1980) *Methods Enzymol.* 65, 499–560.
21. Provost, P. J., Wolanski, B. S., Miller, W. J., Ittensohn, O. L., McAleer, W. J. & Hilleman, M. R. (1975) *Proc. Soc. Exp. Biol. Med.* 148, 532–539.
22. Fellner, P. (1979) in *The Molecular Biology of Picornaviruses*, ed. Pérez-Bercoff, R. (Plenum, New York), pp. 25–47.
23. Chow, M. & Baltimore, D. (1982) *Proc. Natl. Acad. Sci. USA* 79, 7518–7521.
24. Kleid, D. G., Yansura, D., Small, B., Dowbenko, D., Moore, D. M., Grubman, M. J., McKercher, P. D., Morgan, D. O., Robertson, B. H. & Bachrach, H. L. (1981) *Science* 214, 1125–1129.
25. Bittle, J. L., Houghten, R. A., Alexander, H., Shinnick, T. M., Sutcliffe, J. G., Lerner, R. A., Rowlands, D. J. & Brown, F. (1982) *Nature (London)* 298, 30–33.
26. Boothroyd, J. C., Highfield, P. E., Cross, G. A. M., Rowlands, D. J., Lowe, P. A., Brown, F. & Harris, T. J. R. (1981) *Nature (London)* 290, 800–802.
27. Minor, P. D., Schild, G. C., Bootman, J., Evans, D. M. A., Ferguson, M., Reeve, P., Spitz, M., Stanway, G., Cann, A. J., Hauptmann, R., Clarke, L. D., Mountford, R. C. & Almond, J. W. (1983) *Nature (London)* 301, 674–679.
28. Küpper, H., Keller, W., Kurz, C., Forss, S., Schaller, H., Franze, R., Strohmaier, K., Marquardt, O., Zaslavsky, V. G. & Hofschneider, P. H. (1981) *Nature (London)* 289, 555–559.
29. Racaniello, V. R. & Baltimore, D. (1981) *Science* 214, 916–919.