Hepatitis E: Molecular Virology and Pathogenesis

Subrat K. Panda*, Satya P. K. Varma*

*Department of Pathology, All India Institute of Medical Sciences, New Delhi 110029, India

Hepatitis E virus is a single, positive-sense, capped and poly A tailed RNA virus classified under the family *Hepeviridae*. Enteric transmission, acute self-limiting hepatitis, frequent epidemic and sporadic occurrence, high mortality in affected pregnants are hallmarks of hepatitis E infection. Lack of an efficient culture system and resulting reductionist approaches for the study of replication and pathogenesis of HEV made it to be a less understood agent. Early studies on animal models, sub-genomic expression of open reading frames (ORF) and infectious cDNA clones have helped in elucidating the genome organization, important stages in HEV replication and pathogenesis. The genome contains three ORF's and three untranslated regions (UTR). The 5′ distal ORF, *ORF1* is translated by host ribosomes in a cap dependent manner to form the non-structural polyprotein including the viral replicase. HEV replicates via a negative-sense RNA intermediate which helps in the formation of the positive-sense genomic RNA and a single bi-cistronic sub-genomic RNA. The 3′ distal ORF's including the major structural protein pORF2 and the multifunctional host interacting protein pORF3 are translated from the sub-genomic RNA. Pathogenesis in HEV infections is not well articulated, and remains a concern due to the many aspects like host dependent and genotype specific variations. Animal HEV, zoonosis, chronicity in immunosuppressed patients, and rapid decompensation in affected chronic liver diseased patients warrants detailed investigation of the underlying pathogenesis. Recent advances about structure, entry, egress and functional characterization of *ORF1* domains has furthered our understanding about HEV. This article is an effort to review our present understanding about molecular biology and pathogenesis of HEV. (J Clin Exp Hepatol 2013;3:114–124)

Hepatitis E virus was first identified in 1983 as the agent responsible for the large scale waterborne epidemics of acute, self-limiting hepatitis in developing nations.[1] It was shown to be a naked 28–34 nm virus-like particle with spikes and inundations.[1,2] Successful use of non-human primate models for HEV propagation helped in its isolation, cloning and sequencing and as presumed turned out to be an RNA virus.[3,4] It had a ~7.2 kb single stranded genome with sequence homology to other positive-sense RNA viruses.[4,5] Computer based genome annotation revealed three overlapping open reading frames *ORF1*, *ORF2* and *ORF3*.[4,5] The longest ORF, *ORF1* formed the 5′ distal end of the genome and coded for the non-structural replication proteins including *methyltransferase*, *protease*, *helicase*, and RNA dependent RNA polymerase (*RdRp*). The 3′ distal ORF, *ORF2* coded for an arginine rich protein, together with immunodominant sero-reactivity and agglutination against the expressed epitopes, it was predicted to be the major structural protein. A third ORF, *ORF3* coded for a small protein thought to be a minor structural protein.[4–6]

Molecular characterization and replication model of HEV remained hypothetical till the use of strand-specific PCR and the detection of negative-sense RNA replicative intermediate.[7,8] Further understanding of the HEV replication came from the development of infectious cDNA clone of HEV which helped in elucidating the importance of 5′ end cap, RdRp activity and non-structural polyprotein processing.[9,10] Mutational and strand-specific PCR analysis using the HEV replicon helped in delineating the bi-cistronic sub-genomic RNA based expression strategy for ORF2 and ORF3 proteins and the presence of an additional untranslated region (*UTR*) in the intergenic region with probable *cis* reactive element thought to function as sub-genomic promoter.[11–14] Till date only one *cis*-acting element has been characterized, that in the 3′ *UTR*, which was shown to form stable secondary structures and help in the localization of RdRp.[15] Further the viral replicase was shown to localize to the endoplasmic reticulum membrane, which is the probable site of HEV replication.[16]

The mechanism of HEV pathogenesis especially at the molecular level is poorly understood due to lack of an

Keywords: *Hepeviridae*, open reading frames, pathogenesis
Received: 30.4.2013; *Accepted:* 1.5.2013; *Available online:* 30.5.2013
Address for correspondence. Subrat K. Panda, JC Bose Fellow, Professor & Head, Department of Pathology, All India Institute of Medical Sciences, New Delhi 110029, India. Tel.: +91 11 26594924 (off.); fax: +91 11 26588663, +91 11 26588641
E-mail: subrat@gmail.com
Abbreviations: RdRp: RNA dependent RNA polymerase; ORF: open reading frames; UTR: untranslated region; NTP: nucleoside triphosphate; ER: endoplasmic reticulum; EMSA: electrophoretic mobility shift assay; ELISPOT: enzyme-linked immunosorbent spot; DUB: deubiquitinating
http://dx.doi.org/10.1016/j.jceh.2013.05.001

in vitro cell culture system and an ideal practical animal model. However, a wide body of clinical, immunopathological and histopathological studies helped in identifying the stages of infection and the hallmark immune mediated disease outcome, similar to other viral hepatitis.[17,18] HEV pathogenesis varies, from being asymptomatic or acute, self-limiting in low risk groups to causing chronic and/or fulminant hepatic failure in high risk groups; the modus operandi remains hypothetical. Increasing reports typify route, establishment and pathogenic outcome during hepatitis E infection to host factors and genotype involved.

The recent developments in understanding the structure, entry, egress, host interactions and functional characterization of the HEV proteins helped in further providing a better insight in understanding its replication and pathogenesis. This review provides an overview of the recent advances in the understanding of the molecular biology and pathogenesis of HEV.

GENERAL CLASSIFICATION

There are four major mammalian genotypes (1–4) and one avian HEV. HEV distribution shows a socioeconomic and genotypic basis to it, with the highest seroprevalence in countries affected with HEV genotypes 1 and 2 and poor sanitation. Genotype 1 and 2 strains are mainly associated with large waterborne anthroponotic epidemics and frequent sporadic acute hepatitis in these regions. Genotype 3 and 4 HEV strains are found both in human and other mammalian reservoirs (swine, wild boar, deer, mongooses, rabbit, and cattle) and are mostly subclinical and zoonotic, however, genotype 3 is increasingly being associated with chronic infections in immunosupressed patients. Genotype 3 is found mainly in Europe, United States and Japan, and genotype 4 has been identified mainly in Asia. In these regions, the seroprevalence within the general population is 1–3% and higher in certain populations involved in animal husbandry or consuming undercooked meat. Hepatitis E-like viruses have been successfully isolated from chickens, wild rats and trout fish[19–21] and are yet to be classified.

MOLECULAR VIROLOGY

Genome Organization

HEV genome contains a short 5′untranslated region (UTR) of 27 nucleotides, followed by largest ORF (*ORF1*) of 5079 bases that code for a viral non-structural polyprotein (pORF1) (Figure 1). A second ORF (*ORF2*) begins 38 nucleotides 3′ of the termination of ORF1 consisting of 1980 nucleotides and codes for the major structural protein (pORF2). A 65-nucleotide 3′ UTR is contiguous with the termination of ORF2 and is terminated by a 3′ stretch of 150–200 or more adenosine residues. A third and smallest ORF (*ORF3*), 345 nucleotides in length, begins 23 nucleotides 3′ of the termination of ORF1 and extends into ORF2 and encodes for a phosphoprotein (pORF3).[22] HEV forms a single 3′ terminal, capped and polyadenylated ~2.2 kb bicistronic sub-genomic RNA which codes for both ORF3 and ORF2.[14] The sub-genomic RNA start site was found to be from 5124 (FJ457024) and contains a very short untranslated region of 9 nucleotides (FJ457024) followed by the reading frames of *ORF3* and *ORF2*. *ORF3* overlaps *ORF2* with their start codons just 11 nucleotides apart (FJ457024).[14]

The 5′ end of both genomic and sub-genomic RNA is capped. The m^7G cap was confirmed by reverse transcription-PCR (RT-PCR) assay based on the monoclonal antibody (MAb) to 2,2,7-trimethyl guanosine (m^3G) in case of the genomic RNA and by RNA. Ligase-mediated rapid amplification of cDNA ends (RLM–RACE) in case of both genomic and sub-genomic RNA.[11,23] The *5′ UTR* was predicted to form a hair-pin structure. Mapping of the RNA revealed that a 76-nucleotide region at the 5′ end of the HEV genome was responsible for binding the ORF2 protein,[24] and this interaction may play a role in viral encapsidation. The *3′ UTR* of HEV has been shown to form cis-active stem-loop structures which localized the viral RNA dependent RNA polymerase in *in vitro* binding studies and henceforth implicated in the initiation of virus replication.[10,15] The poly A tail at the 3′ end is approximately 150–200 nucleotides long and is crucial for RdRp binding to the 3′ UTR. An additional UTR was recently established with the elucidation of sub-genomic RNA start site downstream of the ORF1 stop site.[11,12] This junction region with similarities to the junction sequence of rubella virus and alphavirus is thought to play a role in sub-genomic RNA formation. Similarly, a highly conserved region in the *ORF2* reading frame was found to form stable stem-loop structures (ISL1 and ISL2), and were found to effect pORF2 production *in vitro* when mutated silently.[25]

Non-structural Proteins

ORF1 of HEV encodes a non-structural polyprotein with a molecular mass of ~186 kDa. Computer analysis suggested that the pORF1 contains several putative functional domains: a methyltransferase (56–240 aa); a Y domain (216–442 aa); a papain-like cysteine protease (433–592 aa); a proline-rich hinge domain (712–778 aa); a X domain (785–942 aa); an RNA helicase (960–1204 aa); and an RNA dependent RNA polymerase (1207–1693 aa).[5] (Figure 1).

The enzymatic activity of methyltransferase was experimentally demonstrated for N-terminal fragment of HEV polyprotein, produced in insect cells using recombinant baculovirus vector which catalyzed both guanine-7-methyltransferase and guanylyl-transferase activities