STRUCTURAL AND FUNCTIONAL STUDIES ON THE ORF2 AND ORF3 PROTEINS OF HEPATITIS E-VIRUS

Author

Mohammad Zafrullah

Supervisor

Prof. Faizan Ahamad

Department of Biosciences
Jamia Millia Islamia

Co-Supervisor

Dr. Shahid Jameel
ICGEB, Aruna Asif Ali Marg
New Delhi

The hepatitis E virus (HEV) is the causative agent of hepatitis E, an acute form of viral hepatitis. It is transmitted primarily by the fecal-oral route, fecally contaminated drinking water being the most commonly documented source of transmission. While the infection is self-limited with no associated chronicity, a fraction of the patients progress to fulminant hepatitis, the most severe form of acute hepatitis. High mortality rates of 20-30% are reported for HEV infection during pregnancy.

The HEV genome is a ~7.2-kilobase polyadenylated, positive-sense RNA that contains three open reading frames (ORFs) designated ORF1, ORF2 and ORF3. Though only limited experimental data are available, based on homology to other positive-sense RNA viruses, ORF1 is postulated to encode the HEV nonstructural polyprotein. The open reading frame 2 (ORF2) encodes the major viral capsid protein (pORF2), which has 660 amino acids. In animal cells, it is expressed in a ~74 kDa form (pORF2) and a ~88 kDa glycosylated form. The ORF3 of HEV encodes a protein of ~13.5 kDa, called pORF3, for which no function has been assigned.

The study on HEV biology and pathogenesis have been severely restricted by the lack of reliable cell culture system and small animal models of viral infection. We have used subgenomic expression strategies to study the properties and functions of individual HEV gene products towards understanding viral replication and pathogenicity.

When expressed in COS-1 cells, pORF2 showed two forms, one corresponding to its predicted size of ~74 kDa and other considerably larger size of ~ 88kDa. That the larger form was glycosylated pORF2 (gpORF2) was confirmed through tunicamycin inhibition and endoglycosydase H digestion experiments. Tunicamycin is an inhibitor of N-linked glycosylation and endoglycosydase H specifically cleaves high-mannose residues from glycoproteins, both modifications known to occur within the endoplasmic reticulum. This was further analyzed with the drugs brefeldin A (BfA) a fungal metabolite that disrupts intracellular membrane traffic at the ER-Golgi junction and monensin, a metabolite of *Streptomyces cinnamonensis* which impairs Golgi function. Brefeldin A inhibited the synthesis of gpORF2 but monensin had no effects on gpORF2 expression

or processing.

Site-directed and deletion mutants were made to map the site of glycan addition on pORF2. Asparagine residues within the three conserved N-X-S\T sequences at amino acid positions 137, 310 and 562 were changed to alanine(s). The single, double and triple mutants were then analyzed for their glycosylation. The single site mutants ORF2 [137], ORF2 [310] and ORF2 [562] all showed the presence of gpORF2.

The dual site mutants ORF2 [137, 310], and [310,562], i.e. those carrying mutations at Asn-310 along with either one of the other two sites were found to express pORF2, but the protein was not glycosylated. In contrast, the dual site mutant ORF2 [137,562] in which the Asn-310 was unchanged, expressed the pORF2 as well as gpORF2 forms of the protein. As expected, the triple mutant ORF2 [137, 310, 562] expressed only the nonglycosylated form of the protein.

An N-terminal hydrophobic stretch resembling export signals found in proteins that are translocated across cellular membranes was also found in pORF2. A mutant of pORF2 lacking N-terminal amino acids 2 to 34 expressed well, but failed to undergo any glycosylation. Since all potential glycosylation sites were unchanged on this mutants, this can only be explained by its inability to translocate across the ER membrane. Microsomal membranes were prepared from COS-1 cells expressing the ORF2 protein and subjected to trypsin digestion in absence and presence of a detergent (NP40). While gpORF2 were protected from trypsin digestion in the absence of detergent, destruction of the microsomal vesicles with NP40 led to complete digestion of protein. The mutant of ORF2 protein lacking signal sequence was sensitive to trypsin digestion in presence of membrane. The triple mutant, pORF2[137, 310, 562] was resistant to trypsin when synthesized in the presence of membranes.

Analyses of pORF2 by chemical cross-linking and sucrose gradient sedimentation showed that the protein has the ability to homo-oligomerize. The analysis of pORF2 mutants clearly showed while an N-terminal region as large as 111 amino acids was not required for its homo-oligomerization, a C-terminal hydrophobic stretch encompassing amino acids 585-610 was critical for it.

When expressed in insect cells, ~72 kDa pORF2 is processed into a ~54 kDa protein and we call it naturally processed form of pORF2. We have generated a mutant of pORF2 (D 111/D585-610) and expressed these two proteins in insect cells infected with recombinant baculoviruses and have purified these proteins for structural studies.

We have carried out circular dichorism (CD) on mutant and naturally processed form of pORF2 to analyze its overall conformation and conformational stability as a function of pH, temperature and urea concentration. We have measured the far-UV CD spectra of mutant and naturally processed form of pORF2 at different pH values of 7.5, 2.5, 2.0, 1.0 and have analyzed each CD spectra for elements of secondary structure in the protein at each pH and 20^{o}C. This analysis suggested that the mutant protein at pH 7.5 contains 38% a-helix and 12%b-sheets. Lowering the pH from 7.5 to 2.5 increased the a- helical content from 38% to 58% without any significant change in the content

of b-sheet. However, a decrease in pH from 2.5 which is the pH maintained in the stomach of a healthy human slightly decreases a-helical content (a- helix = 50%). Analysis of naturally processed form of ORF2 protein showed at pH 7.5 this protein contains 20% a-helix and 46% b-sheets. Lowering the pH from 7.5 to 2.5 increased the a- helical content from 20% to 68% with decrease in b-sheet from 46% to 29%.

We studied the heat-induced denaturation of mutant and naturally processed form of pORF2 at different pH values and monitored it by $[q]_{222}$, the mean residue ellipticity at 222 nm. It was observed that the thermal denaturation of the mutant ORF2 protein was completely reversible at pH £ 3. Analysis of the thermal denaturation curve gave values of ΔH_m, the enthalpy change at T_m, ΔC_p, the heat capacity change at constant pressure, $\Delta G^0_{D,}$ the Gibbs energy change at 20°C. These measurements led us to conclude that lowering the pH stabilized pORF2 in terms of T_m and ΔG^0_D. While naturally processed form of ORF2 protein did not show any transition even at 85 ^{0}C.

We also studied the urea-induced denaturation of these proteins at pH 2.0, monitored by $[q]_{222.}$ It was observed that the protein was stable upto 2 M urea and denaturation was complete at 9 M. It was observed that the mid point of urea denaturation of was 5.4 M and and 5.37 M for mutant and naturally processed form pORF2 respectively. A comparison of the far UV-CD spectrum of the 9 M urea denatured pORF2 at 20 °C with that of the heat denatured pORF2 at 85 °C led us to conclude that the heat denatured protein is less unfolded than the urea denatured protein.

A number of potential phosphorylation sites were predicted in pORF3. To check for this modification, it was transiently expressed in COS-1 cells and labeled with ^{32}P-orthophosphate. The labeling of pORF3 and removal of this label with lambda phosphatase confirmed that this protein was phosphorylated. Similar results were obtained following expression in Huh-7 human hepatoma cells.

It was found that phosphorylation took place at one or more serine residue(s) in pORF3. To determine which of the 8 conserved serine residue(s) in pORF3 was phosphorylated, N-and C-terminal deletion mutants of pORF3 were constructed, transiently expressed and analyzed for phosphorylation. This analysis suggested Ser-80 to be the phosphorylated residue, a result that was confirmed by expression but not phosphorylation of the Ser®Ala mutant. *In vitro* experiments with hexahistidine-tagged pORF3 expressed either in *Escherichia coli* or in COS-1 cells showed efficient phosphorylation with exogenously added mitogen-activated protein kinase (MAPK). We also showed that pORF3 associates with cytoskeleton and this association was lost on deletion of hydrophobic domain I (amino acid residues 1 to 32) of the protein.

We have carried out circular dichroism (CD) on pORF3 to analyze its overall conformation and conformational stability as a function of pH, temperature and urea concentration. We have measured the far-UV CD spectra of pORF3 at different pH values of 7.5, 2.5, 2.0, 1.0 and have analyzed each CD spectra for elements of secondary structure in the protein at each pH and 20°C. This analysis suggested that the ORF3

protein at pH 7.5 contains 17.5% a-helix and 23.6% b-sheets. Lowering the pH from 7.5 to 2.5 increased the a- helical content from 17.5% to 25.1% with decrease in b-sheet from 23.5% to 8%.

We also studied heat-induced denaturation of pORF3 at different pH values and monitored it by $[q]_{222}$, the mean residue ellipticity at 222 nm. Heat-induced denaturation showed that at pH7.5 denaturation was irreversible and at pH £ 3 pORF3 did not show any transition even at 85 ^{0}C. Urea- induced denaturation of pORF3 suggested that, pORF3 is more stable than pORF2.