1. Introduction

There are five different types of viral hepatitis (A, B, C, D, and E) in the world, each of which caused by a different virus. Although they all cause the disease in human, the five viruses are unrelated and are from different virus families with totally different genome structures and distinct replication mechanisms (Fauquet et al., 2005; Hochman & Balistreri, 1999; Kumar et al., 2010). Hepatitis E is an acute viral hepatitis caused by hepatitis E virus (HEV) which is transmitted primarily through a fecal-oral route (Knipe et al., 2007). Pregnancy is one of the risk factors for severe HEV infection with high mortality rate. Numerous epidemic and sporadic cases have occurred in developing countries of Asia, the Middle East, and North Africa, where sanitary conditions are not well-maintained (Okamoto, 2007; Panda et al., 2007; Vasickova et al., 2007). Recent epidemiological studies show that significant prevalence of HEV and anti-HEV antibody is found in humans and several kinds of wild and domestic animals worldwide, including industrialized countries (Meng, 2010).

Hepatitis E virus, discovered in 1983 by immune electron microscopy (Balayan et al., 1983) and first cloned in 1990 (Reyes et al., 1990), is the sole member of the genus Hepevirus within the family Hepeviridae. Based on genome sequences, five major genotypes have been identified. The circulation of genotypes 1 and 2 viruses is maintained among only humans, while those of genotypes 3 and 4 are found in human as well as animals. The viruses of genotype 5 are of avian origin, thought to be noninfectious to humans. Although four of the five genotypes infect human, only one HEV serotype has been found in human (Okamoto, 2007).

HEV is composed of a protein capsid made of a single protein and a positive-sense RNA genome of 7.2 to 7.8 kb in size. Like the hepatitis A virus (HAV), HEV does not have a viral envelope, different from the hepatitis B, C, D virus, all of which contain membrane envelopes outside their capsid. The genomic RNA of HEV is capped, polyadenylated, containing three open reading frames (ORFs). ORF1, mapped at the 5’ terminus of the genome, has about 5124 bases and encodes several viral nonstructural proteins (e.g. methyltransferase, protease, helicase and RNA-dependent RNA polymerase). ORF2 contains 1980 bases at the 3’ end of the genome and encodes the viral capsid protein (CP).
ORF3, which partially overlaps with the other 2 ORFs, encodes a 13.5kDa regulatory phosphoprotein with multifunctions (Monga, 2011; Vasickova et al., 2007). Due to the lack of a robust cell culture system, the viral structural studies have mainly relied on recombinant capsid protein (CP) from in vitro cell expression (Li et al., 2005; Li et al., 1997). Recent findings from HEV expression and purification, X-ray crystallography, and electron cryomicroscopy (cryoEM) have begun to shed new light on the structural and functional properties of this important human pathogen.

2. Hepatitis E virion morphology and capsid protein expression

Native hepatitis E viruses are a non-enveloped, spherically shaped particle. The viron is thought to be made of 180 copies of CP with an approximate diameter of 27-34nm (Guu et al., 2009; Xing et al., 1999). Both immune and negative stain electron microscopy of human stool specimens have showed that the diameter of HEV is about 32nm (Balayan et al., 1983; Bradley et al., 1988). The surface of the virion has obvious spikes that are slightly less pronounced than those of Norovirus, but is clearly distinct from the smooth, featureless surface of the hepatitis A virus (Guu et al., 2009; Xing et al., 2010). However, based on morphology alone, HEV could not be reliably distinguished from other small spherical human enteroviruses usually found in feces. The buoyant density of HEV particles was 1.35 to 1.40g/cm$^3$ in CsCl (Bradley et al., 1988).

Escherichia coli, and eukaryotic cell systems including insect, yeast and mammalian cells have been used to express the viral capsid protein(CP). Among all of these in vitro cell systems, the baculovirus-insect cell system was proved to be an excellent choice to produce virus-like particles (VLPs) for its proper post-translational modification, correct conformation and assembly (Li et al., 2005; Li et al., 1997; Xing et al., 2011). In addition to the full-length 72-kDa protein predicted from the sequence of the ORF2 gene, abundant ORF2-related polypeptides with molecular weights of 53, 56, and 64 kDa were detected in the insect Sf9 and Tn5 cells (Fig1. A). The amino terminus of the 56.5- and 63-kDa proteins was amino acid 112, while the carboxy termini of these two proteins were found to be residue 607 and 635, respectively. There were also some conflicts on the exact site where the carboxy termini of these peptides end among different research groups. The molecular masses of HEV 53- and 56-kDa proteins were larger than expected from amino acid sequence possibly due to post-translational modifications.

HEV CP could self-assemble into two populations of virus-like particles, a small one with a diameter of about 23 nm, and a large one with a diameter of about 42 nm (Fig. 1C and D). Although multiple polypeptides could be generated by expression of the full length ORF2 gene, only the 53-Kd peptide from aa 112 to aa 608 could self-assemble into the small particles. In contrast, the big particle was formed by the peptides from aa 14 to aa 608. The buoyant density of the small VLPs is only 1.285 g/cm$^3$ in CsCl gradient, and the large VLPs has a density of 1.31 g/cm$^3$ (Xing et al., 2010), both smaller than the buoyant densities of 1.35 g/cm$^3$ and 1.39 to 1.40 g/cm$^3$ reported for the native HEV particles extracted from human stool possibly due to the lack of nucleic acids. In addition to the gene truncations at the N and C termini, the yield of HEV VLPs also depends on the insect cell lines used for overexpression. When using the Tn5 insect cells, the capsid proteins with 111 or 13 amino acids truncation at the N terminus could produce small or large particles at high yield (Li et al., 2005; Li et al., 1997; Xing et al., 2011).
Fig. 1. Capsid protein and virions of HEV. (A) 10% SDS-PAGE of the full ORF2 proteins of genotype 4 from overexpression in Tn 5 cell. Target proteins were detected by western blot analysis using the serum of a patient with acute hepatitis E. M, molecular weight standard. (B) Native virions of hepatitis E virus under electron microscopy (Courtesy of Angusmclellan, CDC of United States). (C) Electron microscopy of small virus-like-particles (VLPs) (~23.7nm) purified from Tn5 cells by CsCl equilibrium centrifugation. The VLPs were stained with 2% phosphotungstic acid solution (buffered to pH 7 using sodium hydroxide). (D) Electron microscopy of large VLPs (~41.5nm) purified from Tn5 cells with CsCl equilibrium centrifugation. The VLPs were stained with 2% phosphotungstic acid solution (buffered to pH 7 using sodium hydroxide) Bar=100 Å.

3. The structure of small and large hepatitis E virus capsid

HEV VLPs made by the baculovirus-insect cell expression system are easy to purify to high concentration, thus making it possible for high resolution structural analyses. The structure of small VLPs has been analyzed first by cryoEM (Xing et al., 1999) and then by X-ray crystallography (Guu et al., 2009; Li et al., 2009; Yamashita et al., 2009). In the later case, the cryoEM map has served as a phasing model for crystal structure determination. In addition,
the structure of large HEV VLPs has been extensively solved by electron cryomicroscopy (Xing et al., 2010). Our understanding of HEV capsid greatly benefited from the combination of medium-resolution cryoEM maps with atomic resolution X-ray crystal structures of these VLPs.

The cryoEM density map of the small VLP shows that the capsid consists of 60 copies of HEV CP arranged on T=1 lattice of icosahedral symmetry. The capsid is decorated by a total of 30 dimeric protrusions each measured to be 56 Å long and 43 Å wide. These protrusions are found at twofold axes, with each protrusion surrounded by another four protrusions related by adjacent threefold axes (Fig 2.B). Small plateaus are also found on threefold axes. Based on the cryoEM map, each CP molecule can be clearly divided into two domains, a shell domain and a protrusion domain (Fig 2.B) (Li et al., 2005; Xing et al., 1999).

Using the cryoEM reconstruction as phase model, the structure of the HEV small VLP was determined to atomic resolution by X-ray crystallography (Guu et al., 2009; Yamashita et al., 2009). In agreement with the low resolution density determined by cryoEM, the crystal structure of the $T=1$ VLPs shows that each CP can be divided into 3 domains: the $S$ domain (aa 118–313), the $P1$ domain (aa 314–453), and the $P2$ domain (aa 454–608) (Fig. 3A and B), with the latter two domains connected by a long, flexible hinge linker. The $S$ domain forms a continuous capsid shell that is reinforced by 3-fold protrusions formed by $P1$ and 2-fold spikes formed by $P2$. The $S$ domain adopts the jelly-roll, $\beta$-barrel fold that is most closely related to plant $T=3$ viruses. Both $P1$ and $P2$ contain compact, 6-stranded barrels that resemble the $\beta$-barrel domain of phage sialidase and the receptor-binding domain of norovirus respectively (Guu et al., 2009). Although HEV CP contains 3 domains like the
calicivirus coat protein, the organization of the 3 domains and their structural details are different (Fig. 3B). In calicivirus, the P2 domain is a large insertion in the P1 domain, while the 3 domains S, P1, and P2 are arranged in a linear sequence in $T=1$ VLPs (Guu et al., 2009; Li et al., 2009; Yamashita et al., 2009).

Fig. 3. The X-ray structure of small VLPs. (A) The ribbon diagram of a capsid monomer (PDB accession code: 2ZTN). The P2 (blue), P1 (green) and S (magenta) domains are at the top, middle, and bottom, respectively. (B) Secondary structure assignment. $\alpha$-helices are shown by rectangles, $\beta$-strands by arrows, loops by thick lines, and disordered regions by dotted lines. Regions from the S, P1, and P2 domains are colored the same as in (A). The conventional naming scheme for the 8 $\beta$-strands (BIDG and CHEF) from the jelly-roll $\beta$-barrel is shown in parentheses. (C) The whole crystal structure of HEV $T=1$ VLP. The three domains, S, P1, and P2 are similarly colored as in (A). (D) The crystal structure of one CP dimer fits well into the cryoEM density of $T=3$ VLP.

Meanwhile, the large VLPs have been proved that it is made up of 180 copied of coat protein. According to the cryoEM reconstruction of $T=3$ VLPs (Xing et al., 2010), the density map reveals four discrete domains, P2, P1, S and N from outside to inward (Fig. 4C). The density profile of the P2, P1 and S domains displayed less variation from that observed in $T=1$ VLPs, thus the CPs was able to be grouped into three unique monomers according to their geometric environments. Although monomers A and B formed dimeric spikes (A-B dimers) around each of the five-fold axes, two two-fold related C monomers formed a spike (C-C dimers) at each of the icosahedral two-fold axes (Fig. 4B). The surface lattices of CPs in $T=3$ HEV-LP were similar to the capsid arrangement of caliciviruses. Comparing with the A-B dimer, the morphology of the HEV C-C dimer was less well defined, perhaps due to the flexibility in the angle of the protruding domain toward the icosahedral shell (Guu et al., 2009; Xing et al., 2010).

Although there is no crystal structure for the $T=3$ particles, a pseudo-atomic structure of the $T=3$ particles was obtained by docking the atomic structures of the CP from $T=1$ particle into the $T=3$ cryoEM density map. The docking of the crystal structure of the $T=1$ CP to the density map of $T=3$ VLP showed very good agreement between the two structures (Fig. 3D). The docking positioned N-terminal tail of the HEV CP at the capsid inner surface and
aligned well with an internal density linker in $T=3$ VLPs. The linker density served as a tag to connect the N domain with the icosahedral capsid shell, highlighting the location of the N-terminal 111 amino acids of the HEV CP in the $T=3$ VLP.

Fig. 4. CryoEM structure of the large VLPs. (A) CryoEM micrographs of large VLPs embedded in vitreous ice (Bar=1000 Å). (B) Shade surface of the large VLP reconstruction at about 11 Å resolution as viewed along the twofold axis (Electron Microscopy Data Bank accession number: EMD-5173). Different domains are colored differently, the S domain colored magenta, P1 domain green, P2 domain blue. (C) A central slice view along the twofold axis. (D) Shaded surface representation of the 3D reconstruction after its upper half was removed to reveal the internal view; the N domain is colored yellow.

### 4. Assembly mechanism of HEV capsids

After the structure of the $T=1$ and $T=3$ capsids of HEV were established, it became possible to propose a model for the assembly of the capsids. Because HEV CP containing amino acids 112–608 self-assemble into $T=1$ VLP and CP containing amino acids 14–608 form $T=3$ VLP, it is likely that the N-arm acts as a molecular switch in regulating capsid assembly. Based on previous studies on the $T=3$ viruses such as small plant RNA viruses, insect nodaviruses, and caliciviruses (Ban et al., 1995; Chen et al., 2006; Prasad et al., 1999), two different forms of dimers (A-B dimer and C-C dimer) are necessary for the formation of the $T=3$ particles. Indeed, the $T=3$ density map of HEV VLP confirms that it is composed of two types of dimers: the angled A-B dimer and the flat C-C dimer. *In vitro*, dimer and decamer intermediates were found during the purification, disassembly and reassembly process. The finding of these CP assemblies helped us to understand the HEV capsid assembly process in some detail (Guu et al., 2009; Xing et al., 2010).

As for the $T=1$ capsid, the CP monomers form stable dimers through the interactions largely mediated by P2 domain through an extended loop (550–566) and three β-strands from the central β-barrel. These CP dimers are designated A-B dimers and they are found on the icosahedral two-fold axis. The decamers, each assembled from five A-B dimers, are located at the 5-fold axes, and stabilized by interactions from four loops between the β-strands in
the S domain. After formation of the decamers, they begin to interact around the three-fold axes to assemble the whole icosahedral capsid. Interactions around the three-fold protrusions are largely maintained by the P1 domains. In contrast to the formation of the $T=1$ icosahedral capsid, $T=3$ icosahedral capsid assembly utilizes a mechanism that requires the formation of the flat C-C dimers help to assemble hexamers at icosahedral three-fold positions (Xing et al., 2010). The CP of $T=3$ capsid has residues 14-118 at N terminus with an RNA binding activity. The induction of the C-C conformation was also noted in the bacteriophage MS2, where the complete assembly of capsid requires the presence of
synthetic RNA fragment (Stockley et al., 2007). The interaction of RNA with the N-terminal end of ORF2 may act as the driving force leading the formation of flat C-C dimers and ultimately the full capsid formation through the integration of 30 copies of C-C dimmers and 12 copies of A-B decamers. One putative assembly model for $T=1$ and $T=3$ HEV capsids is depicted Figure 5.

5. Biological functions of HEV capsid and pathogenesis

Viral capsids help to provide structural support for the virion and to contain or protect the viral genome. In non-enveloped viruses, they also mediate cell receptor binding, virus internalization and viral genome release. In addition, viral capsids bind to antibodies and play important roles in inducing host immune reactions.

The three domains of VLP play different roles in organizing the icosahedral capsid. The S domain, which adopts the jelly-roll, β-barrel fold, forms a continuous capsid shell. The P1 domain interacts near three-fold axes, forming isolated trimetric protrusions. Thus, the P1 domain stabilizes only the trimetric interactions. The P2 domain forms dimeric spikes that stabilize capsid protein interactions across the two-folds (Guu et al., 2009; Li et al., 2009; Xing et al., 2010; Yamashita et al., 2009).

The two P domains play additional roles other than maintaining structural stability. The peptide of the HEV capsid protein (amino acids 368–606), which consists of a part of the P1 and entire P2 domain, was shown to be capable of binding and penetrating different cell lines susceptible to HEV, thus is capable of inhibiting HEV infection (He et al., 2008). The potential sugar binding sequence from the P1 domain, 376ADTLLGGLPTELISSA391, is strictly conserved among all HEV genotypes. In contrast, loops 550-566 and 583-593, the potential sugar binding site of the P2 domain, contain three and four hyper-variable amino acids, respectively, indicating that these regions are instead likely to mediate antibody recognition and immune escape. The putative sugar binding motif in the P1 domain (376–391) forms a hidden pocket at the interface between two capsid proteins around three-fold axes, suggesting receptor binding to P1 domain may potentially lead to the destabilization of the VLP trimer, resulting in conformational changes that eventually lead to membrane penetration and genome release into the infected cell. Further mutagenesis studies targeting these two potential sugar-binding sites will determine which domain functions in host-cell binding and virus internalization. Many studies also implicated the P1 and P2 domains in antibody binding (Guu et al., 2009; Li et al., 2009; Xing et al., 2010; Xing et al., 2011; Yamashita et al., 2009).

There are three potential N-glycosylation sites in the capsid protein that are well-conserved, Asn-137-Leu-Ser, Asn-310-Leu-Thr, and Asn-562-Thr-Thr. The N137 and N310 are hidden in the inside of the capsid. In the CP dimmer structure, these two sites are mapped on the horizontal surface of the S domain, suggesting that, if it occurs at all, N-glycosylation in these sites may inhibit assembly of VLPs. Indeed, Graff et al. reported that HEV carrying mutations in Asn-137 or Asn-310 to Glu lost infectivity to cells or rhesus macaques due to a defect in the virion assembly (Graff et al., 2008). N562 located at P2 domain is exposed to solvent at the very top of the surface spike, and could potentially be subjected to receive glycosylation in ER (Graff et al., 2008; Torresi et al., 1999).

The inner surface of the capsid shell is covered with a large number of basic amino acid side chains (R128, R133, R186, R189, R193, and R195, a total of six from each subunit), remarkably
different from dsRNA viruses in which a large number of negatively charged residues on the inner surface are used to facilitate the movement of dsRNA genome during particle-associated transcription (Pan et al., 2009). These arginine side chains from capsids presumably help to neutralize the negative charges of the genomic RNA. Around the five-fold axes is a ring of five tyrosine residues (Y288) that are hydrogen bonded to five serine residues (S200), which are also positioned around the five-fold axes, but closer to the particle interior (Guu et al., 2009).

Presumably because the structural features of small HEV VLPs resemble that of the native virion, small VLPs possess antigenity similar to that of an authentic HEV particle and function as an antigen to detect HEV-specific immunoglobulin G (IgG) and immunoglobulin (IgM) responses (He et al., 2008; Xing et al., 1999; Xing et al., 2011; Yamashita et al., 2009). When VLPs are mixed and incubated with patient serum, the resultant antigen-antibody complexes are able to be examined under EM. Moreover, the VLPs can be also used as the antigen to detect HEV-specific antibodies elicited in an experimentally infected monkey, and obvious IgM and IgG antibody responses were observed during the clinical course of acute hepatitis. Oral administration of recombinant VLPs was immunogenic and stimulated an immune response in mice in the absence of an adjuvant, and all the antibody responses including IgM, IgG and IgA were obtained (Li et al., 2004; Li et al., 2001). Because of its large size and ability to encapsulate the viral genome, it is generally accepted that the T=3 VLPs have the same capsid structure with native HEV. Therefore, the T=3 VLPs should have the same antigenic and immunological functions as the T=1 VLPs as well as the native HEV virion.

6. Conclusion

Major advances have been achieved in the last few years in studying the structure and assembly of the hepatitis E virus capsid. Our structural and functional studies of HEV capsid have largely benefitted from the ability to produce large quantities of recombinant proteins, and under certain conditions to induce the assembly of the HEV-like particles. Many questions still remain unanswered, however. For example, what is the biological function of the C-terminal domain of the capsid protein? The C-terminal domain, or the last ~50 residues of the capsid protein, is not present in any of the virus-like particles that have been analyzed so far. It is possible that this sequence has an important role in HEV assembly but may not be needed for virus infectivity. The HEV cell receptor, which is recognized by the capsid protein, has not been identified either. Other interesting questions are how the infectious virion gets internalized into host cells, and subsequently how the capsid structure changes to release the viral genome into the host cytoplasm to initiate infection. It is anticipated that molecular virology together with structural biology methods such as cryo-EM and X-ray crystallography will continue to make important contributions to help answer these question. In the end, our ultimate understanding of the HEV capsid function requires an effective cell culture system where recombinant viruses can be generated and analyzed for virus entry, replication and growth.

7. Acknowledgments

We thank Drs. Ilker Donmez and Jinmin Cui for valuable discussions and Xiaokang Zhang and Jian He for preparing figure 5. Research in Drs. Zhang and Tao’s laboratories was
supported by Grants NO. 10274106 from Nation natrual funding of China (to JZ) and C-1565 from the Welch Foundation (to YJT).

8. References


There are a lot of important issues related to viral hepatitis studies: molecular biology of viruses, laboratory diagnostics, epidemiology, treatment etc. However, there is a number of special textbooks and monographs on the subject. Considering this fact and rather fast progress in our understanding of the problem this book focuses on the important sections of the problem immune pathogenesis of parenterally transmitted viral hepatitis and some aspects of hepatitis diagnostics. Seven chapters were prepared by several groups of researchers to share information and results of studies with specialists working in the field and persons who are interested to learn about the viral hepatitis issue. The Nobel Prize Committee (the field of physiology and medicine, 2011) awarded Bruce A. Beutler and Jules A. Hoffmann for their discoveries concerning the activation of innate immunity whilst Ralph M. Steinman was awarded for his discovery of the dendritic cell and its role in adaptive immunity. We are proud to say that our book is in line with these discoveries, because 3 chapters cover the problems of innate and adaptive immune response in case of viral hepatitis.

How to reference
In order to correctly reference this scholarly work, feel free to copy and paste the following:
