7 Ijaz S, Arnold E, Banks M, *et al.*: Non-travel-associated hepatitis E in England and Wales: demographic, clinical, and molecular epidemiological characteristics. *J Infect Dis* 2005; **192**:1166–1172
8 Wichmann O, Schimanski S, Koch J, *et al.*: Phylogenetic and case-control study on hepatitis E virus infection in Germany. *J Infect Dis* 2008; **198**:1732–1741
9 Matsuda H, Okada K, Takahashi K, *et al.*: Severe hepatitis E virus infection after ingestion of uncooked liver from a wild boar. *J Infect Dis* 2003; **188**:944
10 Tamada Y, Yano K, Yatsuhashi H, *et al.*: Consumption of wild boar linked to cases of hepatitis E. *J Hepatol* 2004; **40**:869–870
11 Li TC, Chijiwa K, Sera N, *et al.*: Hepatitis E virus transmission from wild boar meat. *Emerg Infect Dis* 2005; **11**:1958–1960
12 Tei S, Kitajima N, Takahashi K, *et al.*: Zoonotic transmission of hepatitis E virus from deer to human beings. *Lancet* 2003; **362**:371–373
13 Yazaki Y, Mizuo H, Takahashi M, *et al.*: Sporadic acute or fulminant hepatitis E in Hokkaido, Japan, may be food-borne, as suggested by the presence of hepatitis E virus in pig liver as food. *J Gen Virol* 2003; **84**:2351–2357
14 Matsubayashi K, Nagaoka Y, Sakata H, *et al.*: Transfusion-transmitted hepatitis E caused by apparently indigenous hepatitis E virus strain in Hokkaido, Japan. *Transfusion* 2004; **44**:934–940
15 Boxall E, Herborn A, Kochethu G, *et al.*: Transfusion-transmitted hepatitis E in a 'nonhyperendemic' country. *Transfus Med* 2006; **16**:79–83
16 Matsubayashi K, Kang JH, Sakata H, *et al.*: A case of transfusion-transmitted hepatitis E caused by blood from a donor infected with hepatitis E virus via zoonotic food-borne route. *Transfusion* 2008; **48**:1368–1375
17 Fukuda S, Sunaga J, Saito N, *et al.*: Prevalence of antibodies to hepatitis E virus among Japanese blood donors: identification of three blood donors infected with a genotype 3 hepatitis E virus. *J Med Virol* 2004; **73**:554–561
18 Gotanda Y, Iwata A, Ohnuma H, *et al.*: Ongoing subclinical infection of hepatitis E virus among blood donors with an elevated alanine aminotransferase level in Japan. *J Med Virol* 2007; **79**:734–742
19 Sakata H, Matsubayashi K, Takeda H, *et al.*: A nationwide survey for hepatitis E virus prevalence in Japanese blood donors with elevated alanine aminotransferase. *Transfusion* 2008; **48**:2568–2576
20 Li TC, Zhang J, Shinzawa H, *et al.*: Empty virus-like particle-based enzyme-linked immunosorbent assay for antibodies to hepatitis E virus. *J Med Virol* 2000; **62**:327–333
21 Li TC, Yamakawa Y, Suzuki K, *et al.*: Expression and self-assembly of empty virus-like particles of hepatitis E virus. *J Virol* 1997; **71**:7207–7213
22 Okamoto H, Takahashi M, Nishizawa T, *et al.*: Analysis of the complete genome of indigenous swine hepatitis E virus isolated in Japan. *Biochem Biophys Res Commun* 2001; **289**:929–936
23 Takahashi M, Nishizawa T, Miyajima H, *et al.*: Swine hepatitis E virus strains in Japan form four phylogenetic clusters comparable with those of Japanese isolates of human hepatitis E virus. *J Gen Virol* 2003; **84**:851–862
24 Feagins AR, Opriessnig T, Guenette DK, *et al.*: Detection and characterization of infectious Hepatitis E virus from commercial pig livers sold in local grocery stores in the USA. *J Gen Virol* 2007; **88**:912–917
25 Surya IG, Kornia K, Suwardewa TG, *et al.*: Serological markers of hepatitis B, C, and E viruses and human immunodeficiency virus type-1 infections in pregnant women in Bali, Indonesia. *J Med Virol* 2005; **75**:499–503

THE JOURNAL OF BIOLOGICAL CHEMISTRY VOL. 285, NO. 43, pp. 33175–33183, October 22, 2010
Printed in the U.S.A.

Structure of Hepatitis E Virion-sized Particle Reveals an RNA-dependent Viral Assembly Pathway*[S]

Received for publication, February 5, 2010, and in revised form, August 2, 2010 Published, JBC Papers in Press, August 18, 2010, DOI 10.1074/jbc.M110.106336

Li Xing[‡§1], Tian-Cheng Li[¶1], Naoyuki Mayazaki[‡§], Martha N. Simon[‖], Joseph S. Wall[‖], Mary Moore[‡], Che-Yen Wang[‡], Naokazu Takeda[¶], Takaji Wakita[¶], Tatsuo Miyamura[¶], and R. Holland Cheng[‡2]

From the [‡]Department of Molecular and Cellular Biology, University of California, Davis, California 95616, the [§]Structural Virology Section, Karolinska Institute, Huddinge University Hospital, SE-14186 Stockholm, Sweden, the [¶]Department of Virology II, National Institute of Infectious Diseases, Tokyo 208-0011, Japan, and the [‖]Biology Department, Brookhaven National Laboratory, Upton, New York 11973-5000

Hepatitis E virus (HEV) induces acute hepatitis in humans with a high fatality rate in pregnant women. There is a need for anti-HEV research to understand the assembly process of HEV native capsid. Here, we produced a large virion-sized and a small *T=1* capsid by expressing the HEV capsid protein in insect cells with and without the N-terminal 111 residues, respectively, for comparative structural analysis. The virion-sized capsid demonstrates a *T=3* icosahedral lattice and contains RNA fragment in contrast to the RNA-free *T=1* capsid. However, both capsids shared common decameric organization. The *in vitro* assembly further demonstrated that HEV capsid protein had the intrinsic ability to form decameric intermediate. Our data suggest that RNA binding is the extrinsic factor essential for the assembly of HEV native capsids.

Hepatitis E virus (HEV),[3] the causative agent of acute hepatitis in humans, is primarily transmitted through contaminated water and generally results in epidemic outbreaks in many developing countries. Sporadic cases have also been reported between outbreaks in HEV-endemic regions as well as in non-endemic areas, and these cases are transmitted through zoonotic route. The overall mortality rates of HEV during outbreaks range from 1 to 15% in general, and the highest mortality occurs in pregnant women, with fatality rates of up to 30% (1).

HEV consists of a non-enveloped icosahedral capsid and a single-stranded, positive-strand RNA genome of ~7.2 kb that encodes three open reading frames (ORFs) (2). The capsid protein, encoded by the ORF2, is composed of 660 amino acids and responsible for most capsid-related functions, such as virion assembly, host interaction, and immunogenicity. Like other hepatitis viruses, HEV is unable to propagate in currently available cell culture systems, and the research of HEV relies largely on the recombinant HEV capsid proteins (3–6). Virus-like particle (VLP) was obtained when the truncated HEV capsid protein was expressed in insect Tn5 cells with deletion of 52 residues from the C terminus and 111 residues from the N terminus (PORF2) (7). Our previous structural analysis of this HEV-VLP by cryo-electron microscopy (cryo-EM) provided a basic understanding of the quaternary arrangement of PORF2, where the reconstructed VLP displayed a *T=1* icosahedral particle composed of 60 copies of PORF2 (8). The essential element of PORF2 protein for *T=1* VLP assembly includes amino acids 125–600 (9). Recently, the structural information was further refined by the crystal structures of genotype-3 *T=1* VLP (10) and genotype-4 *T=1* VLP (11), which revealed the tertiary structure of PORF2 to the level of amino acids. However, the *T=1* VLPs used in these experiments were much smaller than that of the native virion, which has a diameter of 320–340 Å, as determined by immuno-EM (12). There is still a need to investigate the assembly pathway of HEV capsid.

HEV virion is hypothesized to be made of 180 copies of the capsid protein (8, 11). To test this hypothesis, we screened for HEV genotype expression and successfully produced a virion-sized VLP from the HEV genotype-3 ORF2 protein after deleting 52 residues from the C terminus. This VLP allowed us to investigate the molecular interactions that govern HEV virion assembly.

EXPERIMENTAL PROCEDURES

Production of HEV-VLPs and in Vitro Disassembly and Reassembly—HEV-VLPs were produced and purified according to the protocol described previously (7, 13). Briefly, the recombinant baculovirus Ac(G3n13ORF2), which harbored the genome of the N-terminal 13 amino-acid-deleted genotype-3 HEV ORF2, was infected into Tn5 cells with recombinant baculovirus at a multiplicity of infection of 10. The recombinant baculovirus-infected Tn5 cells were harvested at 7 days after infection. The medium and cells were separated by centrifuga-

* This work was supported, in whole or in part, by a grant from the National Institutes of Health Roadmap Project on Nanomedicine (to R. H. C.). This work was also supported by grants from the U. S. Department of Agriculture Hatch Fund, STINT Foundation, and Strategic Research Foundation (to R. H. C.). This study was also partly funded by a grant from the Swedish Research Council (to L. X.) and grants for Research on Emerging and Re-emerging Infectious Diseases, Research on Hepatitis, and Research on Food Safety from the Ministry of Health, Labor, and Welfare, Japan (to T.-C. L.).

The atomic coordinates and structure factors (codes 2ZZQ and 3IYO) have been deposited in the Protein Data Bank, Research Collaboratory for Structural Bioinformatics, Rutgers University, New Brunswick, NJ (http://www.rcsb.org/).

The EM data reported in this paper have been submitted to the Electron Microscopy Data Bank (Electron Microscopy Data Bank) with accession number EMD-5173.

[S] The on-line version of this article (available at http://www.jbc.org) contains supplemental Figs. 1–4 and supplemental Table 1.

[1] Both authors contributed equally to this work.

[2] To whom correspondence should be addressed. Dept. of Molecular and Cellular Biology, University of California, Davis CA 95616. Tel.: 530-752-2693; Fax: 530-752-5659; E-mail: rhch@ucdavis.edu.

[3] The abbreviations used are: HEV, hepatitis E virus; VLP, virus-like particle; TEM, transmission electron microscopy; STEM, scanning TEM; TMV, tobacco mosaic virus; DIG, digoxigenin.

tion at 3,000 rpm for 15 min at 4 °C. The cells were treated with a denaturation buffer containing 50 mM sodium borate, 150 mM NaCl, 1% Nonidet P-40, 0.5% sodium deoxycholate, and 5% 2-mercaptoethanol and gently rocked at room temperature for 2 h. The lysate was diluted with EX-CELL 405, and centrifuged at 32,000 rpm for 3 h in a Beckman SW32Ti rotor, and the pellet was resuspended in EX-CELL 405. The VLP was purified by multiple ultracentrifugations followed by separation on a CsCl density gradient. The final pellet was resuspended in 10 mM potassium MES buffer, pH 6.2. A homemade dialysis device was used in the disassembly and reassembly experiments because it allowed dialysis with a small amount of sample (20–40 μl). Purified VLP was disrupted by dialysis against buffer containing EDTA (10 mM) and DTT (20 mM) at different pHs. After VLP dissociation, 150 mM NaCl in Tris-HCl buffer (pH 7.5) was added, and sample was examined under the electron microscope after a 1-h incubation in the presence of the divalent ion Ca^{2+} (20 mM).

Scanning Transmission Electron Microscopy Analysis of HEV-VLPs—Scanning TEM (STEM) was performed at the Brookhaven National Laboratory STEM facility, with tobacco mosaic virus (TMV) as an internal control. The mixture of VLP and TMV was quickly frozen in liquid nitrogen and then maintained at −150 °C during data collection to eliminate contamination and reduce mass loss. The specimen was scanned by a 40-keV electron beam of 0.25 nm in size, and images were collected with a preamp gain of 10 for both large and small angle detectors (14). The image was recorded with a pixel size of 10 Å and was analyzed with the PCMass29 program. After normalizing the background, the mass of the VLPs was selected with the murine sarcoma virus (MSV) shell model provided by the program. Mass measurements for TMV and HEV-VLPs were always performed from the same image. The HEV-VLP mass was measured in MDa (mass per particle), and the TMV mass was measured in KDa/Å (mass per unit length) (15).

Cryo-electron Microscopic Structure Determination of HEV T=3 VLP—The collection of cryo-EM data for image reconstruction was performed on a JEOL JEM-2100F TEM operating at 200 kV according to the procedure described in detail previously (8). Briefly, a 3-μl solution containing VLP or reassembled ORF2 complex was placed on holey carbon film-coated copper grids and then quickly plunged into liquid ethane after the removal of excess solution. The VLPs were embedded into a thin layer of vitrified ice and transferred into the EM using a Gatan 626 cryo-transferring system. The specimen was observed under 50,000× magnification, and the area of interest was recorded on a TVIPS CCD camera (TemCam-F415). The micrographs were recorded with a pixel size of 2.0 Å at a specimen space and defocus level of 0.7–3.5 Å (supplemental Fig. 1*A*). Digital images with no stigmatism or drift were selected for later image processing. Images of individual HEV *T=3* VLPs were then boxed out and processed with an established software package for icosahedral particles (16, 17). In total, 7,720 individual images were included in the process, and their defocus levels were distributed mainly within 1.0–2.5 Å.

To correct contrast transfer function effect, we applied phase flipping on each image with an in-house program. The density maps were initially reconstructed by combining 1,812 individual images to an effective resolution of 14 Å. Next, amplitude correction was applied during map reconstruction while new data were added. The final density map was reconstructed by combining images of 4,348 individual particles, and the final resolution was assessed as 10.6 Å by Fourier shell correlation with a cutoff of 0.5 (supplemental Fig. 1*B*).

Docking of the *T=1* crystal structure into the *T=3* cryo-EM density map was first done manually with the program O (18) and then refined with the Situs software package (19). The PORF2 monomer was treated as a rigid body during the initial fitting and refinement processes.

X-ray Crystallographic Structure Determination of T=1 HEV-VLP—Crystallization of the VLPs was performed according to a previously described method (20). Crystals were directly flash-frozen in liquid nitrogen, and x-ray diffraction experiments were performed. All x-ray experiments of the HEV-VLP crystals were performed at SPring-8 in Hyogo, Japan. Particle orientation in the unit cell was determined with a self-rotation function (21), and the particle position was determined by a translation search with the cryo-EM structure as the model. The asymmetric crystal unit contains one particle; as a result, 60-fold non-crystallographic symmetry averaging was enforced. The cryo-EM structure (8) was used to obtain the initial phases of Data I (supplemental Table 1) and generated the envelope (mask) used for non-crystallographic symmetry averaging. The phases were refined by real space electron density averaging with icosahedral symmetry elements and solvent flattening. The resolution was gradually extended to 8.3 Å (*R*-factor = 0.21, correlation coefficient = 0.92). This structure was used for the phasing of Data II (supplemental Table 1), and the phases were refined and extended to a 3.8 Å resolution (the overall *R*-factor and correlation coefficient were 0.18 and 0.97, respectively). The electron density map revealed a clear main chain structure. Thus, we built an atomic model into this electron density map using the program O (18) without difficulty.

Sixty icosahedrally related S-subunits were treated as identical, and strict non-crystallographic symmetry constraints were applied during refinement. The data with the resolution range of 20–3.8 Å were included in the refinement (supplemental Table 1 and supplemental Fig. 2) using the CNS program (22). The subunit position and orientation were again clarified by rigid body refinement (*R*-factor of 0.391 (R_{free} = 0.395)). After the first cycle of simulated annealing refinement, positional and *B*-factor refinement followed, and the model was enhanced to an *R*-factor of 0.261 (R_{free} = 0.264). Further positional and *B*-factor refinement, followed by manual revision of the model, resulted in an *R*-factor of 0.242 (R_{free} = 0.245) with reasonable stereochemistry (root mean square deviations in bond lengths and bond angles were 0.010 Å and 1.68°, respectively). Because of the high non-crystallographic symmetry, the *R*-factor and R_{free} factor were almost identical. After refinement, the stereochemistry of the structure was checked with Procheck (23). 98.1% of the non-glycine residues were within the most favored and the additional allowed regions of the Ramachandran plot, and none of the residues were in the additional regions. Atomic structure representations were generated using MolScript (24) and Raster3D (25).

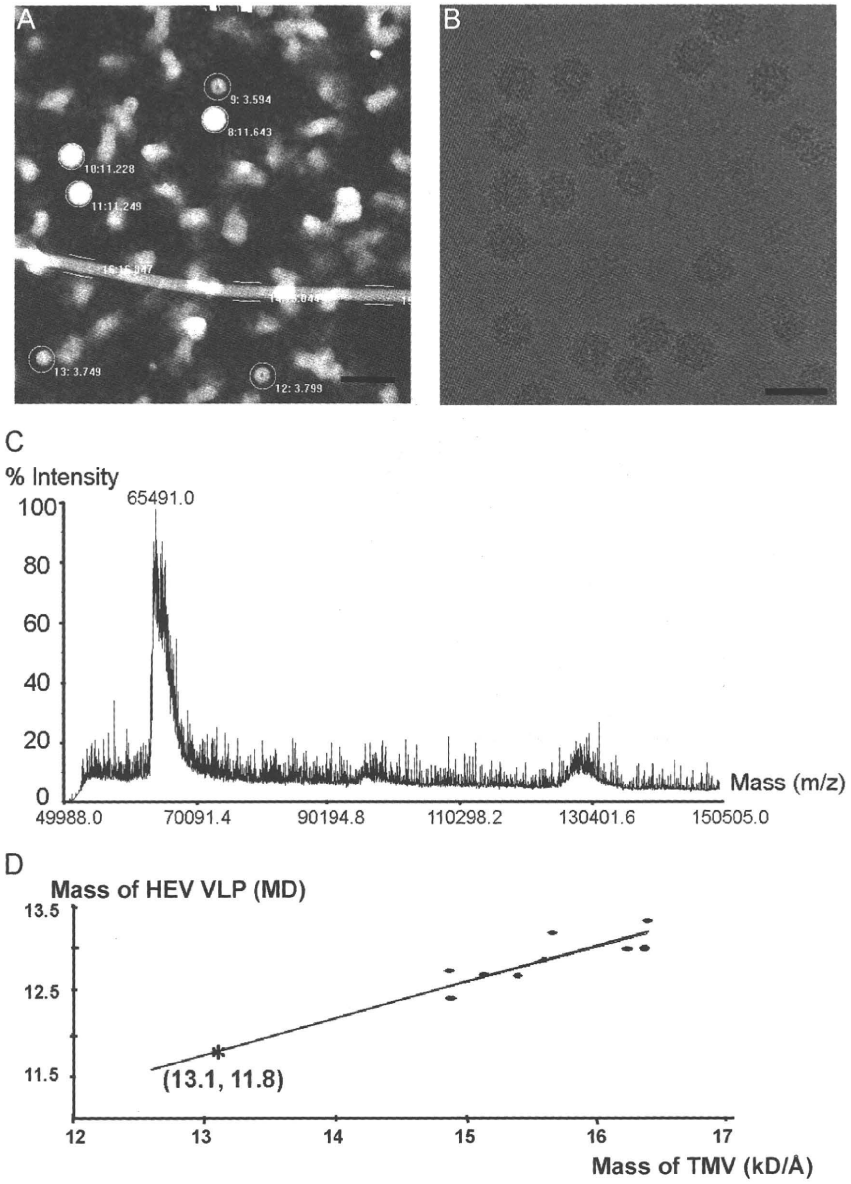


FIGURE 1. **The large HEV-VLP is composed of 180 copies of ORF2 protein.** *A*, STEM micrograph of HEV-VLPs. Both large and small *T=1* HEV-VLPs are projected as spherical images, and their corresponding particle mass was calculated. The long straight rod is TMV, which was added as an internal mass standard. (*Bar* = 1,000 Å.) *B*, the large HEV-VLP appeared as an intact particle decorated with a dot-like pattern on the surface on cryo-EM of the large HEV-VLP. (*Bar* = 500 Å.) *C*, mass spectrum of the large HEV-VLP showing that the molecular mass of the ORF2 protein is 65.5 kDa. *D*, plot showing the observed mass/length of TMV against the observed mass of the large HEV-VLP from different image conditions. For a known TMV mass/length of 13.1 kDa/Å, the mass of the large HEV-VLP was calculated as 11.8 MDa (*asterisk*). *MD*, megadalton.

Nucleic Acid Extraction and Characterization—100 μg of purified large and small VLPs were treated with DNase I (final concentration; 0.01 mg/ml) (Sigma) and RNase A (final concentration; 0.5 μg/ml) at 37 °C for 1 h and then followed by centrifugation at 50,000 rpm for 2 h at 4 °C in a Beckman TLA55Ti rotor. After removal of the supernatant, the pellets were resuspended in 200 μl of PBS minus buffer. The nucleic acids were extracted with RNAzol-LS reagent (Tel-test, Inc., Friendswood, TX) and analyzed on 1% agarose gels.

For detection of HEV RNA in VLPs, the extracted RNAs from VLPs were mixed with sample buffer (final concentration: 2% formaldehyde, 50% formamide in MOPS buffer containing 20 mM MOPS, 5 mM sodium acetate, and 1 mM EDTA (pH 7.0), Sigma-Aldrich, Tokyo, Japan). The solution was denatured at 65 °C for 10 min and cooled immediately in ice-cold water and then mixed with 6× loading buffer (80% formamide, 0.25% bromphenol blue, 0.25% xylene cyanol, 6 mM EDTA, Sigma-Aldrich). After separation on the formaldehyde-denatured agarose gel (18% formaldehyde, 1% agarose, Sigma-Aldrich) in 1× MOPS buffer, the gel was washed twice for 15 min with 20× SSC (3 M NaCl, 0.3 M trisodium citrate dehydrate (pH 7.0), Roche Applied Science, Tokyo, Japan) and transferred onto a Hybond N^+ membrane filter (GE Healthcare, Tokyo, Japan) by the capillary method. RNAs were fixed on the membrane by irradiation with UV light (Stratalinker UV crosslinker, Stratagene, Tokyo, Japan) and dried in the air. RNAs on the membrane were hybridized with DIG-labeled probe. DIG-labeled RNAs were detected by using a DIG Northern starter kit (Roche Applied Science) according to the manufacturer's protocol and visualized by the LAS-3000 imaging analyzer, Fujifilm, Tokyo, Japan). The purified template DNA for the probe RNA was prepared by digesting the plasmid carrying 584-base cDNA of HEV ORF2 (5,903–6,486 nucleotides) under the T7 promoter with BamHI. The probe RNA was prepared using the DIG Northern starter kit according to the manufacturer's manual (incorporation of DIG-UTP during RNA transcription). The norovirus-like particles that were produced by the recombinant baculovirus were used as a control, and the same procedure was used with the HEV-VLPs.

RESULTS

STEM—The virion-sized HEV-VLP was recovered when the genotype-3 ORF2 sequence was expressed in insect cells. This VLP projected as a spherical image with a diameter of ~40 nm, larger than the *T=1* VLP (27 nm in diameter) (Fig. 1*A*). In

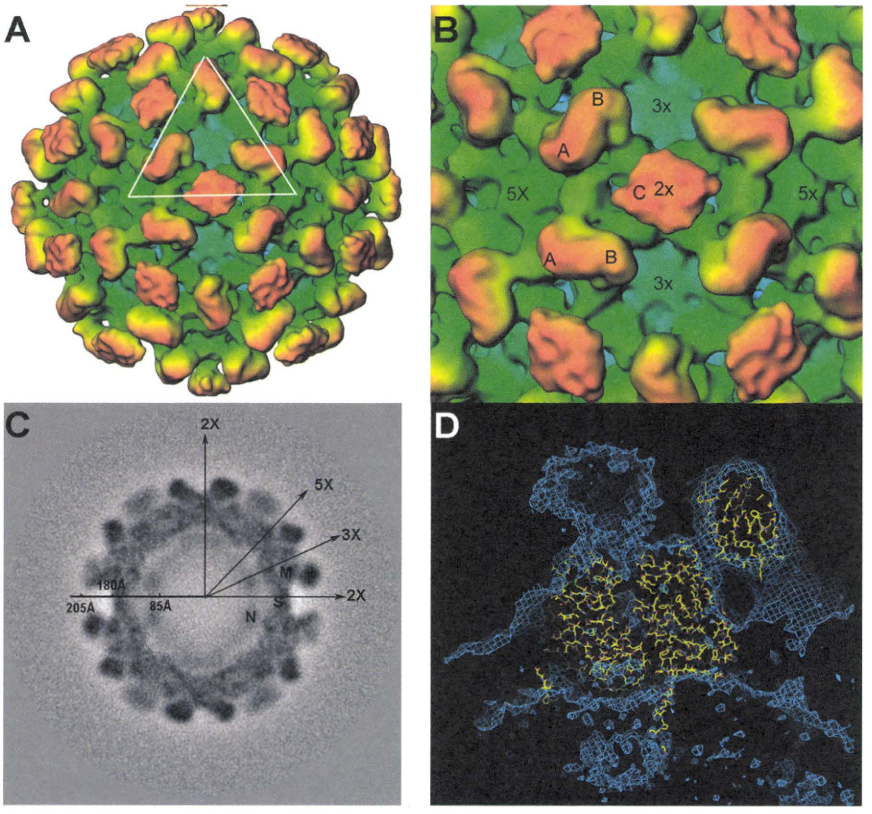


FIGURE 2. **Three-dimensional structure of HEV *T=3* VLP.** *A*, overall structure of the large HEV-VLP reveals the *T=3* icosahedral lattice of the ORF2 proteins. One icosahedral facet is defined as the *triangular area* within the three adjacent five-fold axes. *B*, there are two unique dimeric ORF2 spikes on the HEV *T=3* VLP surface. The A-B dimer is located around the five-fold axis, and the C-C dimer is located at the two-fold axis. *C*, HEV *T=3* VLP has a radius of 205 Å and contains a low density cavity with a radius of 85 Å in the particle center. The distribution of the cryo-EM density revealed four ORF2 domains, P, M, S, and N, at 50 Å from the equatorial section. *D*, the crystal structure of the HEV subunit from *T=1* VLP docks well with the cryo-EM density in the shell region of HEV *T=3* VLP, with the N-terminal loop pointing toward the center.

cryo-electron micrographs, the images of HEV-VLP are decorated with spike-like features and are homogeneous in contrast (Fig. 1*B*).

To determine the composition of the large HEV-VLP, we performed mass measurements by using STEM, a technique that measures the amount of electrons scattered from the objects, such as VLPs, on an EM grid. A mixture of purified large and small HEV-VLPs was freeze-dried onto EM grids for STEM mass measurement. TMV with a known mass-to-length ratio was used as an internal standard. The HEV-VLPs appeared as spherical projections with white contrast on the dark field STEM images (Fig. 1*A*). White cloud-like objects were present in the background, which might be the VLPs broken during sample preservation. The mean mass of large VLP and TMV in the images was measured to generate a plot of the mean TMV mass per unit length *versus* mean VLP mass per particle (Fig. 1*A*). A first-order fit was calculated, and the mass of the large HEV-VLP was determined to be 11.8 MDa (Fig. 1*D*). The mass of the genotype-3 ORF2 protein, which was recovered from the large VLP, was measured as 65.5 KDa by mass spectrometry (Fig. 1*C*). Therefore, the large HEV-VLP contains 180 copies of ORF2 proteins, suggesting that the large HEV-VLP is a *T=3* icosahedral particle (*T=3* VLP). We further performed element analysis with x-ray photoelectron spectroscopy. X-ray photoelectron spectroscopy, also known as electron spectroscopy for chemical analysis, determines the chemical composition of the sample with a depth of 50–70 nm. Phosphorus element, as a characteristic element of nucleic acid, was detected from the G3-VLP that was applied to the carbon-coated copper grid (supplemental Fig. 1*C*). Although the signal of phosphorus element is weak when compared with that of carbon, the phosphorus peak is not detectable in the control grid without VLP materials. This result suggests the co-existence of nucleic acid within the large G3-VLPs.

Three-dimensional Reconstruction of the HEV Virion-sized Particle—The cryo-EM structure of the large HEV-VLP revealed 90 protruding spikes on a complete icosahedral shell (Fig. 2*A*), which is consistent with the *T=3* icosahedral symmetry and the results of the STEM mass measurements. The VLP had an overall diameter of 410 Å and a central cavity of 170 Å in radius as measured from the three-dimensional density map (Fig. 2*C*). The single-layer capsid contained 180 copies of the ORF2 protein, which were 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 dimer) at each of the icosahedral two-fold axes (Fig. 2*B*). The surface lattices of ORF2 proteins in HEV *T=3* VLP were similar to the capsid arrangement of caliciviruses. When compared with the A-B dimer, the morphology of the HEV C-C dimer was less well defined, perhaps due to flexibility in the angle of the protruding domain toward the icosahedral shell.

The density map of the *T=3* VLP displayed four discrete domains, designated from the outside inward as P, M, S, and N, on a section 52 Å from the equatorial plane (Fig. 2*C*). The density profile of the P, M, and S domains displayed less variation from that observed in *T=1* HEV-VLP, and the docking of the crystal structure of the *T=1* PORF2 protein to the density map of *T=3* HEV-VLP showed a very good agreement between the two structures (Fig. 2*D*). The docking positioned N-terminal tail of the PORF2 protein at the capsid inner surface aligned well with the density linker in *T=3* VLP (Fig. 2*D*). The linker density served as a tag to connect the N domain with the icosahedral capsid, indicating the

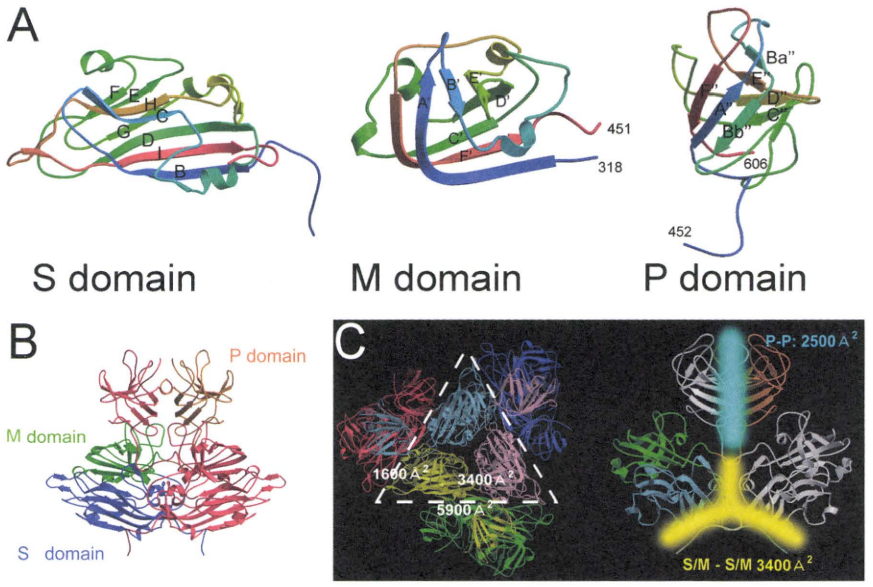


FIGURE 3. **The structure of genotype-1 PORF2 protein.** *A*, ribbon representations of S, M, and P. The structure is covered from *blue* (N terminus of the domain) to *red* (C terminus of the domain). *B*, dimer structure of the capsid protein. One subunit is colored *red*, and the other subunit is colored according to its domain structure (*blue*, S domain; *green*, M domain; *orange*, P domain). *C*, surface areas that buried at the interfaces between two adjacent subunits are overlapped with a PORF2 hexamer (*left*) and at the PORF2 dimeric interface (*right*). One icosahedral facet is defined as the *triangular area* within the three adjacent five-fold axes.

location of the N-terminal 111 amino acids of the ORF2 protein in *T=3* HEV-VLP.

Crystal Structure of the Genotype-1 T=1 HEV-VLP—The crystal structure of the truncated genotype-1 capsid protein (PORF2, containing residues 112–608) can be separated into three domains, S, M, and P, with a less resolved region covering residues 555–560. The S domain formed by residues 118–317 folds into a classical eight-stranded β-barrel with a jelly roll motif (Fig. 3*A*), as observed in many *T=3* viral capsid proteins (26, 27). Uniquely, three additional short α-helixes were observed in the S domain between strands E and F and strands G and H. The capsid shell was mainly stabilized by intersubunit interactions between the S domains. The folded M domain, consisting of residues 318–451, was a twisted antiparallel β-sheet with an α-helix between the B′ and C′ strands (Fig. 3*A*). The P domain, composed of residues 452–606, folded into a β-barrel composed of antiparallel β-sheets, F″A″Bb″ and Ba″E″D″C″ (Fig. 3*A*), and was connected with the M domain through a long proline-rich hinge (PTPSPAPSRP of residues 452–461) (Fig. 3*A*). Although both the M and the P domains existed above the S domain, the protruding spikes in the HEV cryo-EM map contain only the P domain density, which is a clear difference to those caliciviruses (supplemental Fig. 3). The PORF2 dimers have the largest buried surface area between monomers (5,900 Å^2) mainly due to the interface between P domains (Fig. 3*B*). The buried surface area is 3,400 and 1,600 Å^2 for the two adjacent PORF2 subunits around a three-fold axis and a five-fold axis, respectively (Fig. 3*B*). Moreover, the buried surface area of three molecules around a three-fold axis (9,500 Å^2) is much wider than that around a five-fold axis (4,700 Å^2).

Sequence alignment of genotype-1 PORF2 with the sequences of genotype-3 (10) and genotype-4 (11) revealed that the S domain is the most conserved region among HEV genotypes, whereas greater divergence was seen in the N-terminal region (supplemental Fig. 4*A*). Among the solved structures, genotype-3 appeared flexible at the N-terminal end and was 11 amino acids shorter than the others, whereas the structure deviation from genotype-1 PORF2 is very small (total root mean square deviation is 0.62 for the 472 equivalent amino acids) (supplemental Fig. 4*B*). Because amino acids 118–129 play an important role in bridging the N-domain to the S-domain in *T=3* VLP and serve as a docking register, we used the crystal structure of genotype-1 to decipher the *T=3* cryo-EM density map.

Consistent Interdimeric Interactions between T=3 and T=1 HEV-VLPs—To understand the mechanism of ORF2 protein transition between *T=1* and *T=3* assemblies, we docked the *T=1* decamer and hexamer into the *T=3* cryo-EM density map. The decamer of *T=1* VLP consisted of 10 adjacent PORF2 monomers corresponding to five dimers around a five-fold axis, whereas the *T=1* hexamer corresponded to three adjacent dimers around a three-fold axis. Unlike the hexamer, the coordinates of the PORF2 decamer fitted very well with the curvature of the *T=3* density map at the five-fold vertex (Fig. 4*A*) and with the domain separation (Fig. 4*B*). The curvature of *T=3* capsid at the three-fold axis did not agree with the coordinates of the PORF2 hexamer as one of the dimers appeared to be sticking out of the cryo-EM density map (data not shown). Besides, the orientation of the P domain of the C-C dimer relative to its M/S domains was 90° different from that of the A-B dimer (Fig. 5). This suggests that the molecular interactions among A-B dimers in the *T=3* icosahedron are consistent with the dimer-dimer interactions in the *T=1* icosahedral assembly, whereas the interaction between the A-B dimer and C-C dimer is unique to the *T=3* assembly.

In Vitro Reassembly of the ORF2 Protein—To understand the role of ORF2 decamer in *T=3* VLP assembly, we analyzed the self-assembly process of HEV-VLP *in vitro*. A combination of chelating (EDTA) and reducing (DTT) agents was found to disassemble *T=3* VLP in a high alkaline environment (pH 10) without denaturating the ORF2 protein (data not shown). The addition of 20 mM $CaCl_2$ into the disassembly solution led to the association of the ORF2 dimers into star-shaped complexes, and no refolded VLP was found (Fig. 4*C*). When we examined the star-shaped complexes, we found that the distance between two opposite vertices was ~18 nm, close to the diameter of TMV (Fig. 4*D*). This size was consistent with that measured from PORF2 decamers. Thus, the star-shaped complexes resembled not only the appearance but also the size of the ORF2

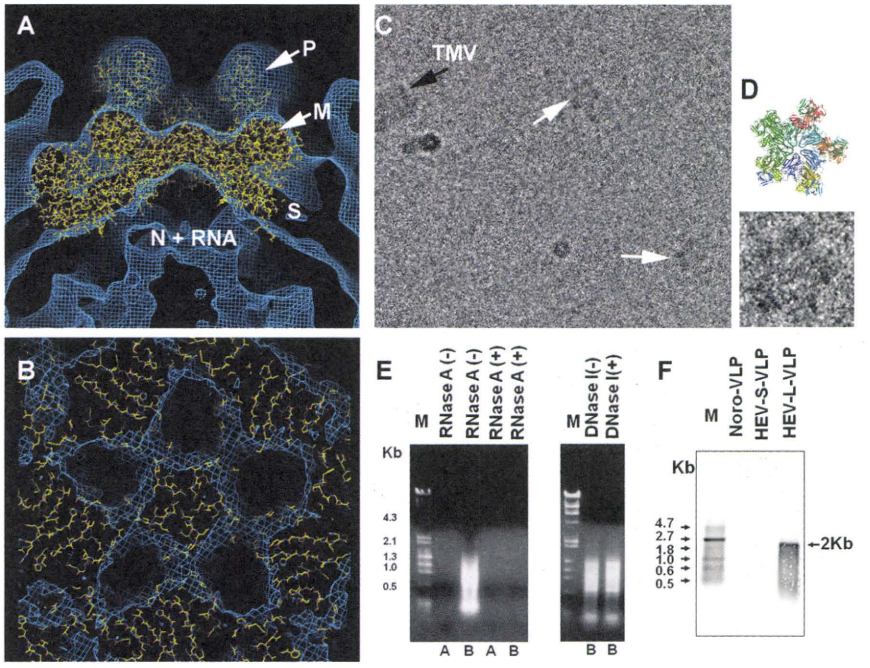


FIGURE 4. **Structure of ORF2 decamer is consistent in both *T=1* and *T=3* VLPs.** *A*, the cryo-EM density of HEV *T=3* VLP agrees well with the coordinates of *T=1* VLP at the region around the five-fold axis. The *P*, *M*, *S*, and *N + RNA* mark the corresponding density layer. *B*, consistent features between the decamer in HEV *T=3* and *T=1* VLP were revealed by the crystal structure docking of PORF2 dimers into the cryo-EM density map of HEV *T=3* VLP. *C*, reassembly of the ORF2 protein *in vitro* led to ORF2 star-shaped complex formation (*white arrows*). The size of this complex fits well into that of the ORF2 decamer after calibration with TMV (*black arrow*) as an internal standard. *D*, one ORF2 complex was zoomed in twice and is displayed in comparison with the crystal structure of PORF2 decamer (*ribbon drawing*). *E*, nucleic acids were extracted from HEV *T=1*VLP (*lane A*) and *T=3* VLP (*lane B*) and degraded in the presence of RNase A but remained as intact when incubated with DNase I. No nucleic acid was detected from *T=1* VLP (*lane A*). DNA maker (*lane M*): λ-HindIII + ϕ-174 HincII. *F*, Northern blotting to detect HEV RNA. RNA marker (*lane M*): 0.5–10-kb RNA ladder. RNA extract from large HEV *T=3* VLP, small HEV *T=1* VLP, and noravirus VLP was loaded in lane marked with *HEV-L-VLP*, *HEV-S-VLP*, and *nora-VLP*, respectively.

decamer (a pentamer of dimers). Although the overall buried surface area around the three-fold axis is larger than that around the five-fold axis, we did not find any complexes that could fit with the PORF2 hexamer.

The *in vitro* disassembly and reassembly suggested that other factors than ORF2 protein contribute to *T=3* VLP assembly. Considering the electropositivity of the ORF2 N-terminal 111 amino acids, we performed nucleic acid extraction from both the *T=3* and the *T=1* VLPs. Electrophoresis results demonstrated the presence of nucleic acids in the *T=3* extract, whereas the *T=1* VLP extract was negative for nucleic acids (Fig. 4*E*). The extracted nucleic acids were sensitive to RNase treatment and resistant to DNase treatment. To further characterize the nucleic acid, we performed Northern blotting according to the protocol described under "Experimental Procedures." An RNA band about 2 kb in size was detected from large particles, whereas no band was detected from the small particles and the control norovirus-like particles (Fig. 4*F*). This result is consistent with the VLP profiles observed from the cryo-electron micrographs and further indicated that the *T=3* particles encapsulated the HEV gene encoding the ORF2 protein.

DISCUSSION

Hepatitis E virus is a human pathogen that causes acute liver failure. Like other hepatitis viruses, HEV cannot be propagated with currently available cell culture techniques. The capsid protein of genotype-3 HEV can be expressed in insect cells as PORF2 protein, including amino acids 112–608 that self-assemble into *T=1* VLP, and as ORF2 protein, including amino acids 14–608 that form *T=3* VLP.

The crystal structures of PORF2 revealed three functional domains, S, M, and P, and the function of each domain constrained its sequence flexibility. The S domain formed an icosahedral shell that served as the base for arranging M and P domains; hence, the subunit surface should be highly conserved among genotypes. Sequence alignment agreed very well with this function, identifying the S domain as the most conserved region among HEV genotypes (28). The P domain serves as the putative binding site for both neutralizing antibody and cellular receptor (29) and contains 19 divergent amino acids across four genotypes (supplemental Fig. 4). Only nine of these amino acids were exposed at the surface of the P domain. Inspection of the binding footprint of antibodies on the cryo-EM density map indicated that only one amino acid was buried within the antibody-binding interface[4]. This explains why the HEV serotype is non-divergent despite sequence variation among HEV genotypes. The direct correlation between sequence variability and domain functionality may be necessary for the HEV capsid to carry multiple functions and to ensure error-free assembly. It also explains why the transition of HEV-VLP from the *T=3* to the *T=1* lattice does not disturb its antigenicity and why *T=1* VLP can be disassembled and reassembled *in vitro* to carry foreign antigenic epitopes (30) or DNA plasmids (31).

The *T=3* HEV-VLP has a similar morphology to that of calicivirus; however, the crystal structures of PORF2 revealed a distinctive M domain arrangement, although the folding of the HEV M domain is similar to the folding of the P1 domain in caliciviruses (supplemental Fig. 3). In HEV, the P domain is located at the C-terminal end of the M domain, whereas the P2 domain of caliciviruses is inserted into the P1 domain at the region between the A′ and B′ strands (32, 33). Furthermore, the M domain of HEV interacts strongly with the S domain and connects to the P domain via a long proline-rich hinge, whereas the P1 domain in caliciviruses is a subdomain of the protrusion

[4] L. Xing, C.-Y. Wang, T.-C. Li, Y. Yasutomi, J. Lara, Y. Khudyakov, D. Schofield, S. Emerson, R. Purcell, N. Takeda, T. Miyamura, and R. H. Cheng, manuscript in preparation.

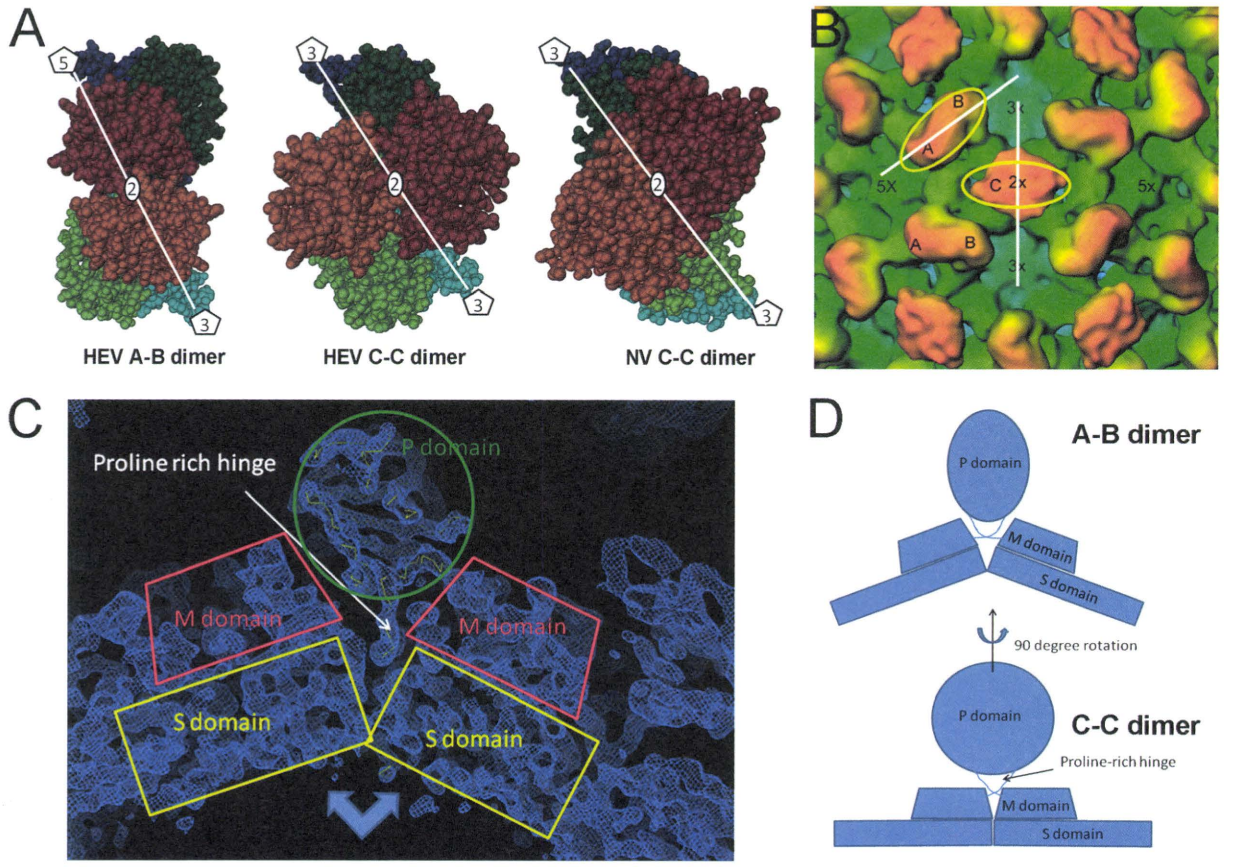


FIGURE 5. **The orientation of P domain relative to the M and S domain in the C-C dimer appears different from that in A-B dimer.** *A*, the dimeric interactions in the HEV A-B dimer, the HEV C/C dimers, and the Norwalk virus C-C dimer are shown in space-filling models as observed from the outside of the particles. S, M (P1) and P (P2) domains in a subunit are highlighted in *dark/light blue*, *dark/light green*, and *brown/chocolate*, respectively. *B*, in the *T=3* cryo-EM density map, the orientation of the P domain is shown as the angle between the platform of the spike (*yellow circles*) relative to the M/S domains in the C-C dimer (*white line* passing through the two adjacent three-fold axes) or the M/S domains in the A-B dimer (*white line* passing through a five-fold axis and the neighboring three-fold axis). *C*, the crystal density map showing the position of the proline-rich hinge within the cleft of the M domains. *D*, the cleft in between the M/S domains provides sufficient room to accommodate the proline-rich hinge in the A-B dimer, where the domains take a bent conformation. The cleft is narrowed down in the C-C dimer due to the flat conformation between the M/S domains, thus pushing the hinge up and out of the cleft.

spike (supplemental Fig. 3). This seems to have an impact on VLP stability; the spike of the HEV C-subunits appeared weakly defined when compared with that in the A-B dimer, whereas the spike of the Norwalk virus C-subunit appeared rigid and similar to that in the A-B dimer in the cryo-EM structure (34). Additionally, deletion of the N-terminal positively charged amino acids from the Norwalk virus capsid protein does not induce *T=1* VLP because the Norwalk virus capsid protein only contains a short N-terminal tail of 20 amino acids (35).The HEV C-C dimer is profoundly different from the HEV A-B dimer in the orientation of the P domain relative to the M/S domain (Fig. 5). Conformational difference between the A-B dimer and C-C dimer has been reported earlier *T=3* viruses. In tomato bushy stunt virus, binding of RNA plays an important role to differentiate the C-C dimer from the A-B dimer. The N-terminal arm of the C-C dimer is well ordered and interacts with the RNA genome, whereas the A-B dimer is disordered and free from RNA interactions (36). In the flock house virus, the C-C dimeric contact acquires a flat conformation to accommodate the RNA duplex, whereas the A-B dimer is in a bent conformation and involves no RNA (37). The different orientation observed between the HEV C-C dimer and A-B dimer may result from the difference in RNA occupancy. The A-B dimers do not interact with RNA and have a bent conformation. As a result, the angled contact of the M/S domains accommodated the proline-rich hinge within the V-shaped cleft, similar to that in the *T=1* VLP, thus solidifying the orientation of the P domain (Fig. 5*C*). In contrast, the contact with the RNA led the C-C dimer to a flat conformation that pushes the hinge out of the cleft. Thus, the P domain in the C-C dimer is flexible and could take a 90° rotation from the orientation in the A-B dimer (Fig. 5*D*).

The native HEV capsid was predicted to possess *T=3* icosahedral symmetry (8, 11), and Guu *et al.* (11) suggested that such *T=3* capsids require the dimer to have both flat and inwardly bent conformation at quasi-two-fold and icosahedral two-fold positions, respectively. Here, with detailed features, our cryo-EM structure of *T=3* VLP provided direct observation of

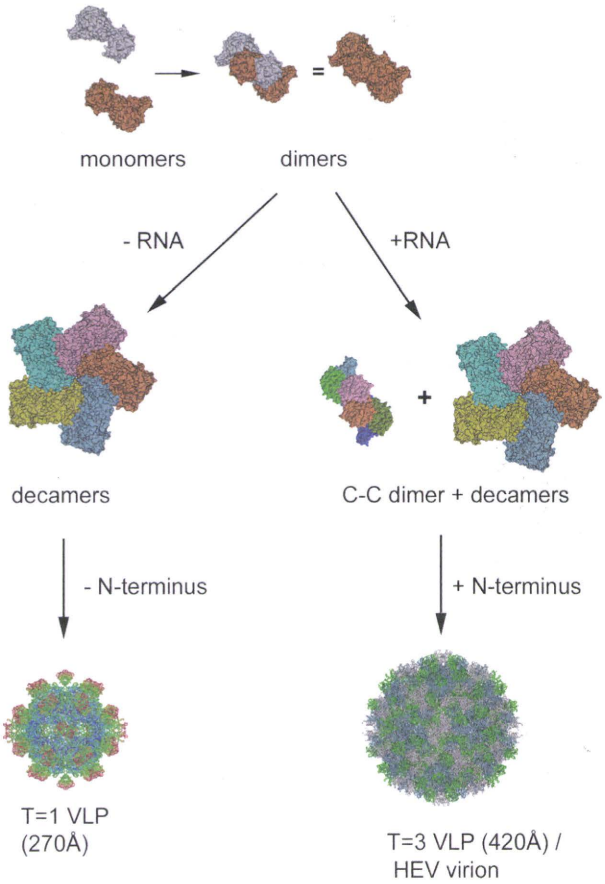


FIGURE 6. **Diagram showing the putative assembly process of HEV *T=1* and *T=3* VLP.** The ORF2 subunit encodes information that governs the assembly of decamers. Interaction with the RNA fragment induces flat dimeric contact and the formation of C-C dimers, which guides the assembly of the complete icosahedral capsid.

the two unique conformations between the A/B and C/C dimer at the foothold of S-domain contact but also in the arrangement of the protruding P-domain.

It is suggested by molecular simulation that *T=3* icosahedral capsid assembly utilizes a mechanism in which preformed aggregates of intermediates combine in contrast to the formation of the *T=1* icosahedral capsid that includes the addition of predominately monomers (38). The ORF2 decamer is therefore the assembly intermediate of *T=3* HEV capsid and is located at each of the five-fold vertexes. The appearance of a hexameric ring at icosahedral three-fold positions is the critical step in *T=3* capsid assembly and depends on the C-C dimer. The *in vitro* disassembly and reassembly also indicates the involvement of an extrinsic factor other than the ORF2 protein in the assembly of *T=3* VLP and the C-C dimer is in a flat conformation that is concomitant with RNA binding. The induction of C-C conformation has been reported with bacteriophage MS2, where the complete assembly of capsid requires the presence of synthetic RNA fragment (39). Therefore, the pentamer of dimer (Fig. 4*C*) may be the assembly intermediate that is common for the *T=3* virus. The interaction of RNA with the N-terminal end of ORF2 is the driving force leading the C-C dimer to the flat formation and ultimately full capsid formation through the integration of 30 copies of C-C dimers with 12 copies of A-B decamers (Fig. 6).

The existence of the N-terminal amino acids 14–111 prevents ORF2 proteins from forming *T=1* VLP. The capsid of *T=1* VLP encloses a central cavity with a volume allowing a maximum of 55 additional residues on each copy of PORF2 protein, if the average protein density is considered to be 1.30 g/ml. The maximum diameter of the central cavity is about 340 Å for the *T=3* VLP and thus is sufficient to accommodate both the HEV genome and the ORF2 N-terminal domains. By characterizing the size and the sequence of the encapsidated RNA, we found that the *T=3* HEV VLP selectively encapsidated the RNA fragment that encodes the ORF2 protein. Thus, it is very possible that the native HEV capsid is the *T=3* icosahedron. There, the encapsulated genomic RNA may play a direct role in the assembly of the HEV infectious virion. Our data demonstrated here that HEV was different from caliciviruses in its assembly pathway, protein domain arrangement, and genome organization, although both viruses are *T=3* icosahedral particles with dimeric spikes. Hepatitis E virus showed a high similarity to some plant viruses in its assembly pathway: the utilization of a long electropositive N-terminal domain to interact with genomic RNA. Although the evolutionary origin for such similarities requires further investigation, our data place the HEV structure in a unique position, deviating from that of human caliciviruses and approaching that of *T=3* small plant viruses.

Acknowledgments—We are grateful to Masaaki Kawano and Carlos G. Moscoso for fruitful discussion on the technical aspects of the manuscript and to Akifumi Higashiura and Atsushi Nakagawa for technical support. We thank Kai Sun from the University of Michigan for technical assistance on x-ray photoelectron spectroscopy analysis. The Brookhaven National Laboratory STEM facility is supported by the Department of Energy, Office of Health and Environmental Research.

REFERENCES

1. Naik, S. R., Aggarwal, R., Salunke, P. N., and Mehrotra, N. N. (1992) *Bull. World Health Organ* **70,** 597–604
2. Tam, A. W., Smith, M. M., Guerra, M. E., Huang, C. C., Bradley, D. W., Fry, K. E., and Reyes, G. R. (1991) *Virology* **185,** 120–131
3. Schofield, D. J., Purcell, R. H., Nguyen, H. T., and Emerson, S. U. (2003) *Vaccine* **22,** 257–267
4. Li, T. C., Suzaki, Y., Ami, Y., Dhole, T. N., Miyamura, T., and Takeda, N. (2004) *Vaccine* **22,** 370–377
5. Purdy, M. A., McCaustland, K. A., Krawczynski, K., Spelbring, J., Reyes, G. R., and Bradley, D. W. (1993) *J. Med. Virol.* **41,** 90–94
6. Riddell, M. A., Li, F., and Anderson, D. A. (2000) *J. Virol.* **74,** 8011–8017
7. Li, T. C., Yamakawa, Y., Suzuki, K., Tatsumi, M., Razak, M. A., Uchida, T., Takeda, N., and Miyamura, T. (1997) *J. Virol.* **71,** 7207–7213
8. Xing, L., Kato, K., Li, T., Takeda, N., Miyamura, T., Hammar, L., and Cheng, R. H. (1999) *Virology* **265,** 35–45
9. Li, T. C., Takeda, N., Miyamura, T., Matsuura, Y., Wang, J. C., Engvall, H., Hammar, L., Xing, L., and Cheng, R. H. (2005) *J. Virol.* **79,** 12999–13006
10. Yamashita, T., Mori, Y., Miyazaki, N., Cheng, R. H., Yoshimura, M., Unno, H., Shima, R., Moriishi, K., Tsukihara, T., Li, T. C., Takeda, N., Miyamura, T., and Matsuura, Y. (2009) *Proc. Natl. Acad. Sci. U.S.A.* **106,** 12986–12991
11. Guu, T. S., Liu, Z., Ye, Q., Mata, D. A., Li, K., Yin, C., Zhang, J., and Tao,

Y. J. (2009) *Proc. Natl. Acad. Sci. U.S.A.* **106,** 12992–12997
12. Balayan, M. S., Andjaparidze, A. G., Savinskaya, S. S., Ketiladze, E. S., Braginsky, D. M., Savinov, A. P., and Poleschuk, V. F. (1983) *Intervirology* **20,** 23–31
13. Li, T. C., Scotti, P. D., Miyamura, T., and Takeda, N. (2007) *J. Virol.* **81,** 10890–10896
14. Wall, J. S., Hainfeld, J. F., and Simon, M. N. (1998) *Methods Cell Biol.* **53,** 139–164
15. Wall, J. S., and Simon, M. N. (2001) *Methods Mol. Biol.* **148,** 589–601
16. Baker, T. S., and Cheng, R. H. (1996) *J. Struct. Biol.* **116,** 120–130
17. Ji, Y., Marinescu, D. C., Zhang, W., Zhang, X., Yan, X., and Baker, T. S. (2006) *J. Struct. Biol.* **154,** 1–19
18. Jones, T. A., Zou, J. Y., Cowan, S. W., and Kjeldgaard, M. (1991) *Acta Crystallogr. Sect. A* **47,** 110–119
19. Chacón, P., and Wriggers, W. (2002) *J. Mol. Biol.* **317,** 375–384
20. Wang, C. Y., Miyazaki, N., Yamashita, T., Higashiura, A., Nakagawa, A., Li, T. C., Takeda, N., Xing, L., Hjalmarsson, E., Friberg, C., Liou, D. M., Sung, Y. J., Tsukihara, T., Matsuura, Y., Miyamura, T., and Cheng, R. H. (2008) *Acta Crystallogr. Sect. F Struct. Biol. Cryst. Commun.* **64,** 318–322
21. Blow, D. M., Rossmann, M. G., and Jeffery, B. A. (1964) *J. Mol. Biol.* **8,** 65–78
22. Brünger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J. S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T., and Warren, G. L. (1998) *Acta Crystallogr. D Biol. Crystallogr.* **54,** 905–921
23. Laskowski, R. A., Rullmannn, J. A., MacArthur, M. W., Kaptein, R., and Thornton, J. M. (1996) *J. Biomol. NMR.* **8,** 477–486
24. Kraulis, P. (1991) *J. Appl. Cryst.* **24,** 946–950
25. Merritt, E. A., and Murphy, M. E. (1994) *Acta Crystallogr. D Biol. Crystallogr.* **50,** 869–873
26. Rossmann, M. G., and Johnson, J. E. (1989) *Ann. Rev. Biochem.* **58,** 533–573
27. Harrison, S. C. (2001) *Curr. Opin. Struct. Biol.* **11,** 195–199
28. Zhai, L., Dai, X., and Meng, J. (2006) *Virus Res.* **120,** 57–69
29. He, S., Miao, J., Zheng, Z., Wu, T., Xie, M., Tang, M., Zhang, J., Ng, M. H., and Xia, N. S. (2008) *J. Gen. Virol.* **89,** 245–249
30. Niikura, M., Takamura, S., Kim, G., Kawai, S., Saijo, M., Morikawa, S., Kurane, I., Li, T. C., Takeda, N., and Yasutomi, Y. (2002) *Virology* **293,** 273–280
31. Takamura, S., Niikura, M., Li, T. C., Takeda, N., Kusagawa, S., Takebe, Y., Miyamura, T., and Yasutomi, Y. (2004) *Gene Ther.* **11,** 628–635
32. Chen, R., Neill, J. D., Estes, M. K., and Prasad, B. V. (2006) *Proc. Natl. Acad. Sci. U.S.A.* **103,** 8048–8053
33. Prasad, B. V., Hardy, M. E., Dokland, T., Bella, J., Rossmann, M. G., and Estes, M. K. (1999) *Science* **286,** 287–290
34. Prasad, B. V., Matson, D. O., and Smith, A. W. (1994) *J. Mol. Biol.* **240,** 256–264
35. Bertolotti-Ciarlet, A., White, L. J., Chen, R., Prasad, B. V., and Estes, M. K. (2002) *J. Virol.* **76,** 4044–4055
36. Timmins, P. A., Wild, D., and Witz, J. (1994) *Structure* **2,** 1191–1201
37. Fisher, A. J., and Johnson, J. E. (1993) *Nature* **361,** 176–179
38. Nguyen, H. D., Reddy, V. S., and Brooks, C. L., 3rd (2009) *J. Am. Chem. Soc.* **131,** 2606–2614
39. Stockley, P. G., Rolfsson, O., Thompson, G. S., Basnak, G., Francese, S., Stonehouse, N. J., Homans, S. W., and Ashcroft, A. E. (2007) *J. Mol. Biol.* **369,** 541–552

SUPPLEMENTARY MATERIAL

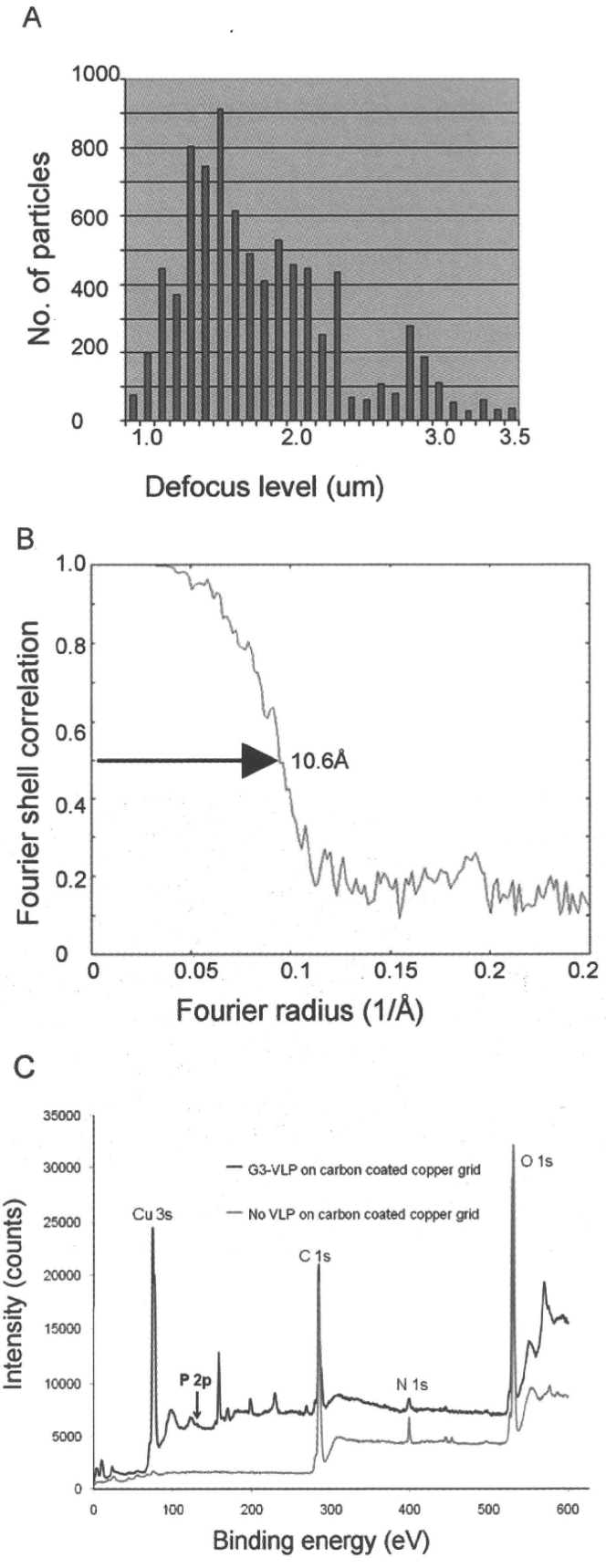
Structural basis of the RNA-dependent assembly pathway of hepatitis E virion-sized *T=3* icosahedral particles

Li Xing, [1, 2, #] Tian-cheng Li, [3, #] Naoyuki Mayazaki, [1, 2] Martha N Simon, [4] Joseph S Wall, [4] Mary Moore [1], Che-Yen Wang, [1] Naokazu Takeda [3], Takaji Wakita[3], Tatsuo Miyamura[3], R Holland, Cheng[1,*]

1. Department of Molecular and Cellular Biology, University of California, Davis CA95616
2. Karolinska Institute Structural Virology, F68 University Hospital, SE-14186 Stockholm, Sweden.
3. Biology Department, Brookhaven National Lab, Upton, NY 11973-5000
4. Department of Virology II, National Institute of Infectious Diseases, Tokyo 208-0011, Japan.

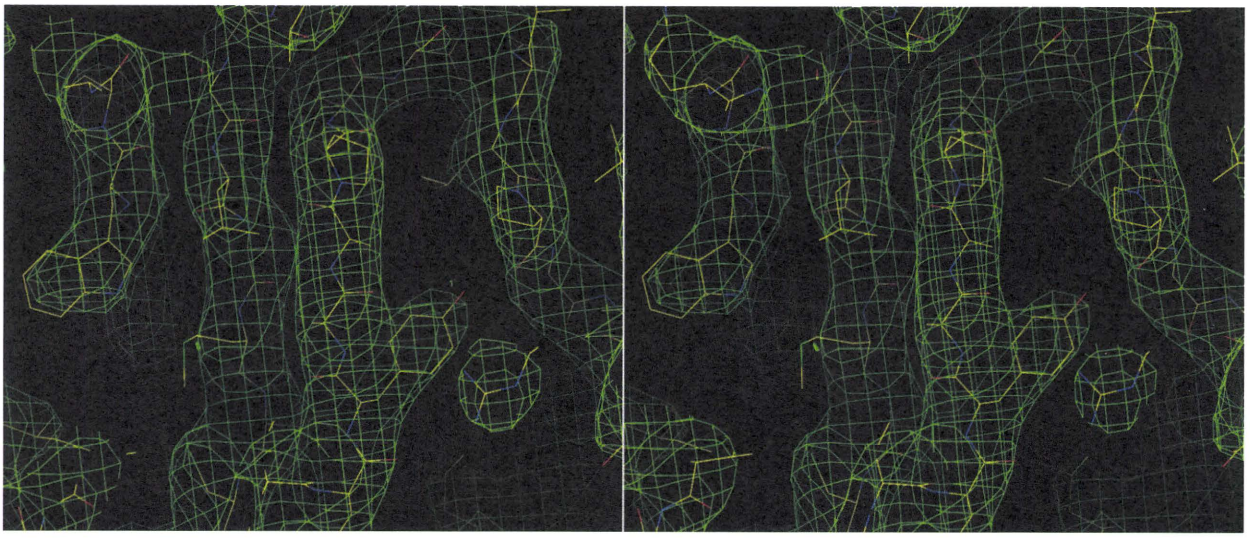
*Author to whom all correspondence should be addressed: R. Holland Cheng, Department of Molecular and Cellular Biology, University of California, Davis, CA 95616

#These authors contribute equally

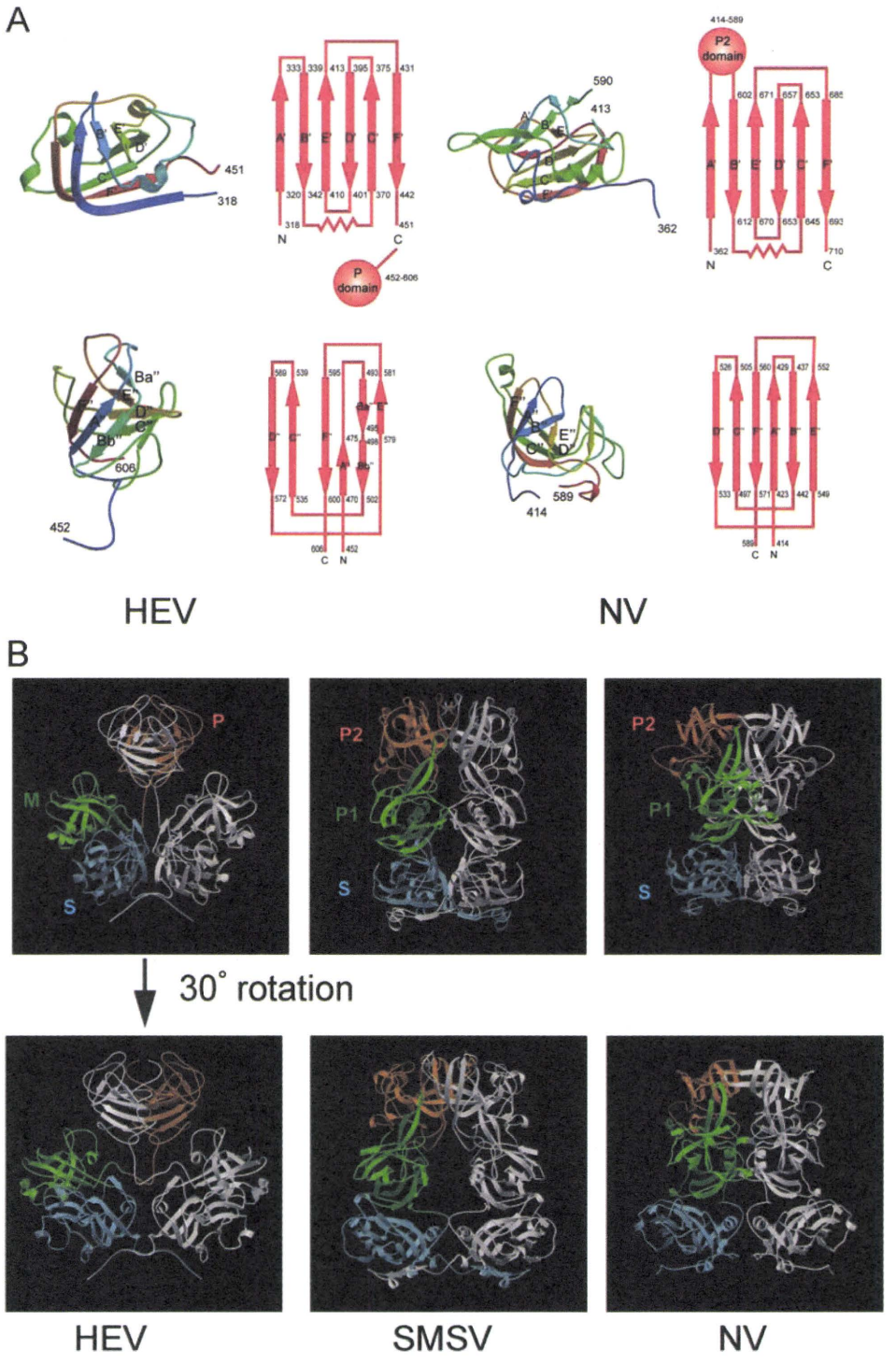

A
1000
800
600
400
200
0
No. of particles
1.0
2.0
3.0
3.5
Defocus level (um)
B
1.0
0.8
0.6
0.4
0.2
0
Fourier shell correlation
10.6Å
0
0.05
0.1
0.15
0.2
0.2
Fourier radius (1/Å)
C
35000
30000
25000
20000
15000
10000
5000
0
Intensity (counts)
G3-VLP on carbon coated copper grid
No VLP on carbon coated copper grid
Cu 3s
C 1s
O 1s
P 2p
N 1s
0
100
200
300
400
500
600
Binding energy (eV)

2

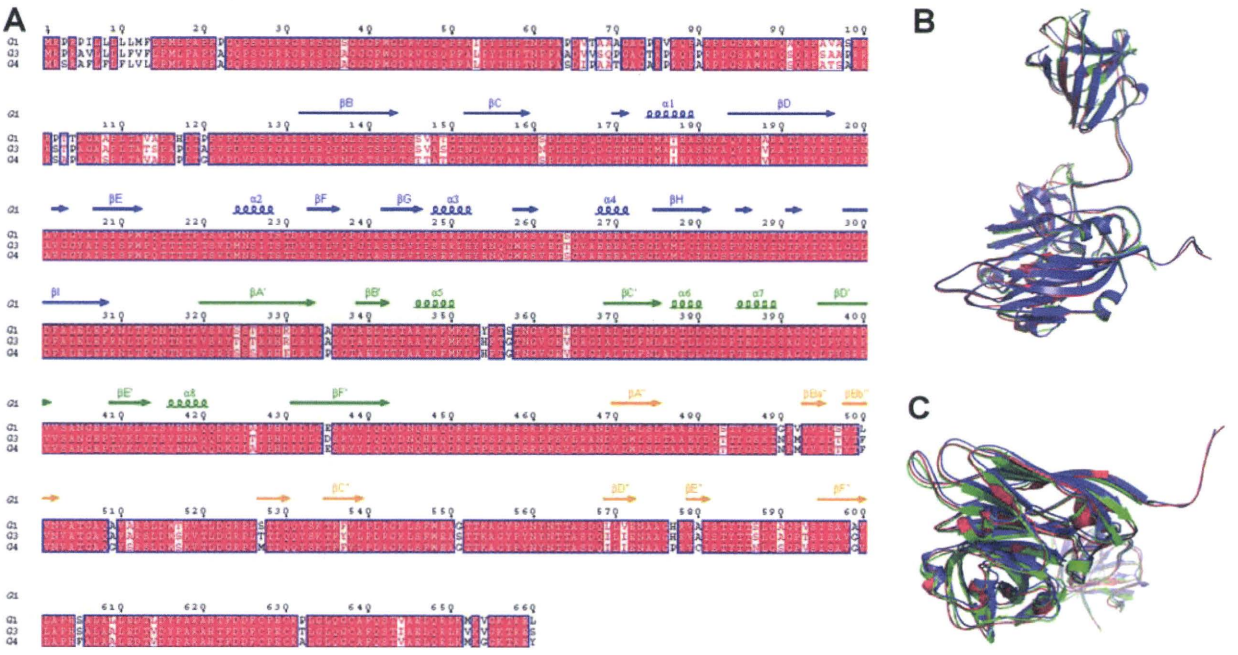
Supplementary Figure 1: Electron microscopic data for large HEV-VLP. (a) Distribution of cryo-EM with a defocus level of 0.7-3.5 μm. (b) Fourier shell correlation indicating that the resolution of the final density map is 10.6Å (cutoff of 0.5). (c): X-ray photoelectron spectra of HEV-G3 large VLPs. 3uL of large *T=3* VLPs was applied on glow discharged carbon coated copper grid and was blotted with a filter paper to remove excess liquid. No chemical staining was applied. The grid was placed in an ultra high vacuum chamber and irradiated with a monoenergetic soft X-rays. The emitted electrons were collected and was plotted relative to its electron binding energy.



Supplementary Figure 2. Stereoimage of the *T=1* VLP electron density map with modeled amino acid residues in the S domain.


A
451
318
333 339 413 395 375 431
320 342 410 401 370 442
318
451
N
C
P domain
452-606
590
413
362
414-589
P2 domain
602 671 657 653 685
612 670 653 645 693
362
710
N
C
Ba''
Bb''
606
452
589 539 595 483 581
572 535 600 470 502
606
452
C N
589
414
526 505 560 429 437 552
533 497 571 423 442 549
589
414
C N
HEV
NV
B
P
M
S
P2
P1
S
P2
P1
S
30° rotation
HEV
SMSV
NV

Supplementary Figure 3: Structure of HEV M and P domain and their difference from calciviruses. (A) Ribbon representations (*Top left*) of the M domain in HEV and the P1 domain in Nowalk virus (NV) along with their respective topology diagrams (*Top right*). The polypeptide chain colored from blue (the amino terminus) to red (the carboxyl terminus), while the β-strands are labeled from A' to F'. Ribbon representations (*Bottom left*) of the P and P2 domains in HEV, SMSV and NV along with their respective topology diagrams (*Bottom right*). The polypeptide chain colored from blue (the amino terminus) to red (the carboxyl terminus) and the β-strands are labeled from A" to F", following the nomenclature of NV. (B) Dimer interactions and interdomain orientations. The dimeric interactions in the dimer of HEV structure and the C/C dimers of San Miguel sea lion virus (SMSV) and NV structures are shown in ribbon representations. S, M and P domains in a subunit are highlighted in skyblue, green and orange, respectively. The dimer as viewed along the line joining the icosahedral three- and five-fold axes (*Upper*) and after 30° rotation (*Lower*).



Supplementary Figure 4. Comparison between HEV capsid proteins of genotype-1, -3, and -4. (A) Sequence alignment of HEV genotypes-1, -3, and -4. Amino acid residues are boxed according to the alignment generated by CLUSTAL-X. Red background indicates identical amino acids and a red character with a white background indicates residues that are similar according to a Risler matrix. Secondary structural elements are labeled above the sequence in blue for the S domain (residues 118-317), green for M domain (residues 318-451), and orange for P domain (residues 452-606). (B) Ribbon representations of the monomer structures of the HEV-VLP PORF2 protein of genotype-1 (red), genotype-3 (green) and genotype-4 (blue).Note the position of the genotype-3 N-termini (green arrow). (C) Structures of HEV PORF2 proteins after a 90° rotation to show the locations of the N-termini.

Supplementary Table 1: Data collection and refinement statistics (Molecular Replacement)

	Crystal 1	Crystal 2
Data collection		
No. Crystals	1	2
Space group	$P2_12_12_1$	$P2_12_12_1$
Cell dimensions		
a, *b*, *c* (Å)	337.00,343.00,346.00	337.00,347.00,354.00
α, β, γ (°)	α=β=γ=90°	α=β=γ=90°
Resolution (Å)	70.0-8.30 (8.60-8.30)	70.0-3.80 (3.94-3.80)
R_{sym} or R_{merge}	0.136 (0.500)	0.136 (0.626)
$I / \sigma I$	8.3 (2.2)	6.3 (1.0)
Completeness (%)	87.6 (89.4)	68.4 (24.8)
Redundancy	3.0 (3.0)	3.4 (1.4)
Refinement		
Resolution (Å)		20.00-3.80
No. reflections		275,008
R_{work} / R_{free}		0.242 / 0.245
No. atoms		
Protein		3,659
Ligand/ion		0
Water		0
B-factors		
Protein		140.7
Ligand/ion		
Water		
R.m.s deviations		
Bond lengths (Å)		0.010
Bond angles (°)		1.7

JOURNAL OF VIROLOGY, Jan. 2011, p. 1117–1124 Vol. 85, No. 2
0022-538X/11/$12.00 doi:10.1128/JVI.00657-10

Spatial Configuration of Hepatitis E Virus Antigenic Domain▽

Li Xing,[1,2]† Joseph C. Wang,[1]† Tian-Cheng Li,[3] Yasuhiro Yasutomi,[4] James Lara,[5] Yury Khudyakov,[5] Darren Schofield,[6] Suzanne U. Emerson,[6] Robert H. Purcell,[6] Naokazu Takeda,[3] Tatsuo Miyamura,[3] and R. Holland Cheng[1]*

Molecular and Cellular Biology, University of California, Davis, California 95616[1]; Karolinska Institute Structural Virology, F68 University Hospital, SE-14186 Stockholm, Sweden[2]; Department of Virology II, National Institute of Infectious Disease, Tokyo 162, Japan[3]; Tsukuba Primate Research Center, National Institute of Biomedical Innovation, Ibaraki 305-0843, Japan[4]; Division of Viral Hepatitis, Centers for Disease Control and Prevention (CDC), Atlanta, Georgia 30333[5]; and Hepatitis Virus Section, National Institute of Allergy and Infectious Diseases, Bethesda, Maryland 208892[6]

Received 26 March 2010/Accepted 28 October 2010

Hepatitis E virus (HEV) is a human pathogen that causes acute hepatitis. When an HEV capsid protein containing a 52-amino-acid deletion at the C terminus and a 111-amino-acid deletion at the N terminus is expressed in insect cells, the recombinant HEV capsid protein can self-assemble into a T=1 virus-like particle (VLP) that retains the antigenicity of the native HEV virion. In this study, we used cryoelectron microscopy and image reconstruction to show that anti-HEV monoclonal antibodies bind to the protruding domain of the capsid protein at the lateral side of the spikes. Molecular docking of the HEV VLP crystal structure revealed that Fab224 covered three surface loops of the recombinant truncated second open reading frame (ORF2) protein (PORF2) at the top part of the spike. We also determined the structure of a chimeric HEV VLP and located the inserted B-cell tag, an epitope of 11 amino acids coupled to the C-terminal end of the recombinant ORF2 protein. The binding site of Fab224 appeared to be distinct from the location of the inserted B-cell tag, suggesting that the chimeric VLP could elicit immunity against both HEV and an inserted foreign epitope. Therefore, the T=1 HEV VLP is a novel delivery system for displaying foreign epitopes at the VLP surface in order to induce antibodies against both HEV and the inserted epitope.

Hepatitis E virus (HEV) is a causative agent of acute hepatitis in humans and is primarily transmitted via the fecal-oral route. HEV is thus resistant to the low pH and digestive enzymes associated with the stomach and gastrointestinal tract. HEV regularly causes epidemics in many tropical and subtropical countries. In India, 101 outbreaks were confirmed by serological analysis in the state of Maharashtra in the last 5 years (6), and the lifetime risk of HEV infection exceeds 60% (28). Sporadic cases have also been reported in regions where HEV is endemic, as well as in areas where it is not endemic. Although some of these cases were associated with travel, many cases involved patients without a history of travel to regions where HEV is endemic. Accumulating evidence suggests that sporadic infection occurs through a zoonotic route and is not limited to developing countries. Seroprevalence suggests hepatitis E infection may also be prevalent in high-income countries (21), such as the United States (17), the United Kingdom (3), and Japan (18). The overall mortality rate of HEV infection during an outbreak generally ranges from 1 to 15%, and the highest mortality occurs in pregnant women, with fatality rates of up to 30% (19).

The HEV virion is composed of a 7.2-kb single-stranded RNA molecule and a 32- to 34-nm icosahedral capsid. The HEV genome contains three open reading frames (ORFs). The capsid protein, encoded by the second open reading frame (ORF2), located at the 3′ terminus of the genome, comprises 660 amino acids and is responsible for most capsid-related functions, such as assembly, host interaction, and immunogenicity. Recombinant ORF2 proteins can induce antibodies that block HEV infection in nonhuman primates (12, 27). Four major antigenic domains were predicted to be located within the C-terminal 268 amino acids of the ORF2 protein; one domain was experimentally identified as a neutralization epitope in the Sar-55 ORF2 capsid protein (25, 26). However, the minimal peptide needed to induce anti-HEV neutralizing antibodies contains residues 459 to 607 of the ORF2 protein (33), which is much longer than a linear antigenic epitope, suggesting that the neutralization epitope is conformational. Therefore, the detailed structure of the HEV capsid protein is required in order to understand the organization of HEV epitopes.

Currently, there are 1,600 HEV genomic sequences available through the International Nucleotide Sequence Database Collaboration. They are classified into four genotypes which vary by geographic distribution and host range (10). In contrast, only a single serotype has been identified, suggesting that the immunodominant domain of HEV is highly conserved among genotypes. Antibodies from any one of the four genotypes cross-react with the capsid protein of genotype 1 (7).

Like other hepatitis viruses, HEV does not propagate well in currently available cell culture systems. Hepatitis E preventive strategies so far rely on the use of ORF2-derived recombinant protein (16). When expressed in insect cells, recombinant truncated ORF2 protein (PORF2), with 52 residues deleted from

* Corresponding author. Mailing address: Molecular and Cellular Biology, University of California, 1 Shields Ave., Davis, CA 95616. Phone: (530) 752-5659. Fax: (530) 752-3085. E-mail: rhch@ucdavis.edu.

† These authors contributed equally.

▽ Published ahead of print on 10 November 2010.

the C terminus and 111 residues deleted from the N terminus, self-assembles into virus-like particles (VLPs) (15). Our previous structural analysis of recombinant HEV VLP by cryoelectron microscopy (cryo-EM) provided the first understanding of the quaternary arrangement of PORF2.

The essential assembly element of the PORF2 protein contained amino acids 125 to 600 (13), and the reconstructed VLP displayed a T=1 icosahedral particle composed of 60 copies of truncated PORF2 (30). Recently, crystal structures were reported for genotype 1 T=1 VLPs (31), genotype 3 T=1 VLPs (32), and genotype 4 T=1 VLPs (8), revealing that PORF2 is composed of three domains, the S domain, M domain, and P domain. The T=1 icosahedral shell is composed of 60 copies of S domains, while the M domain binds tightly to the S domain and interacts with two 3-fold-related M domains to form a surface plateau at each of the 3-fold axes. Two P domains are tightly associated as a dimeric spike that protrudes from each of the icosahedral 2-fold axes. As a result, on a low-resolution cryo-EM density map, the HEV T=1 VLP appears as an icosahedral particle with 30 spikes (30).

Although these VLPs are smaller (270 Å in diameter) than the native HEV virion (320 to 340 Å), oral administration of HEV VLPs to experimental animals can induce anti-HEV antibodies that bind to native HEV (14). When a B-cell tag of 11 amino acids on glycoprotein D of herpes simplex virus was covalently coupled to the C-terminal end of PORF2 (after residue 608), the fusion protein retained the ability of PORF2 to assemble and form chimeric T=1 icosahedral VLPs that were capable of eliciting systemic and mucosal antibodies against both HEV capsid protein and the attached B-cell tag (20). Therefore, the HEV T=1 VLP is a potential carrier for delivering not only HEV antigen but also foreign antigens or antiviral drugs to the host immune system. However, rational design of HEV-based delivery vectors requires detailed information on HEV VLP structure, as well as on HEV immunodominant domains.

Here, we identified antigenic structures using cryo-EM and three-dimensional reconstruction. Our results indicate that the binding footprint of a neutralizing antibody covers the lateral side of the P domain, while a B-cell tag at the C terminus does not alter the assembly of T=1 HEV VLP.

MATERIALS AND METHODS

Production and purification of anti-HEV monoclonal antibody (MAb) MAb224. Eight-week-old female BALB/c mice were immunized at 0 and 4 weeks by intraperitoneal inoculation with HEV VLPs (100 μg/ml). Four weeks later, a final boost containing an equal volume of antigen was administered. Three days after the final boost, mouse spleen cells were fused with P3U1 mouse myeloma cells using polyethylene glycol 1500 (50% [wt/vol]) (Boehringer, Mannheim, Germany) essentially as described by Adler and Faine (1). Supernatants from microplate wells positive for hybridoma growth were screened by enzyme-linked immunosorbent assay (ELISA) using recombinant HEV VLPs as the antigen. Hybridomas that secrete antibodies specific for HEV were subcloned three times by limiting dilution, after which they were considered to be monoclonal. Antibodies in the supernatants were isotyped using a mouse monoclonal antibody isotyping kit (Amersham, Little Chalfont, Buckinghamshire, United Kingdom) in accordance with the manufacturer's protocol. Hybridomas were grown in bulk in stationary flasks (Nunc, Roskilde, Denmark) using RPMI 1640 with 15% fetal calf serum. Antibodies were purified from cell supernatants using HiTrap protein G affinity columns (Pharmacia Biotech AB, Uppsala, Sweden) and stored at −80°C. Among all of the antibodies that were generated, MAb224, an immunoglobulin G1 (IgG1) isotype, was chosen for structural analysis.

Preparation of Fab224 fragments. Isolated Fab224 fragments were prepared from purified mouse monoclonal antibodies by papain cleavage. A reducing L-cysteine buffer was used to activate the papain, and MAb224 was mixed with papain at a molar ratio of 100:1. The mixture was incubated overnight at 30°C. The reaction was stopped by the addition of iodoacetamide, and the product was analyzed by SDS-PAGE. The Fab224 fragments were purified using a 5-ml prepacked protein A chromatography column (Pierce Protein Research) according to the manufacturer's instructions. The Fc fragments and uncleaved MAb224 antibodies were trapped in the column due to their affinity for protein A, while the Fab224 fragments were collected in the flowthrough fraction.

Production and purification of anti-HEV Fab4. Fab4 was prepared by phage display and purified according to the protocol described previously (25). Briefly, chimpanzee 1441 was infected with HEV strain SAR-55. Bone marrow was aspirated from the iliac crest of this animal, and the antibody κ-chain gene and γ1-chain gene were amplified and cloned into the pComb3H phage display vector and pGEM-T cloning vector (Promega), respectively, and transformed into *Escherichia coli* XL-1 Blue. The bacteria were then amplified and infected with helper phage VCS M13 at a multiplicity of infection of 50 to produce a library displayed on the surfaces of phage particles. Phage was panned on SAR-55 ORF2-coated ELISA wells; four rounds of panning were performed. After amplification of the selected library, the phagemid DNA was extracted and the vector was modified to remove the bacteriophage coat protein III-encoding region of the phage. The phagemid DNAs were religated and transformed into *E. coli* XL-1 Blue to produce soluble Fabs. The vector pComb3H was constructed to encode a six-histidine tail at the end of the Fab fragment, thus facilitating Fab purification. Fab4 purity was determined by SDS-PAGE, followed by colloidal Coomassie brilliant blue staining.

Production and purification of HEV VLPs. The production and purification of HEV VLPs were conducted as described previously (13, 15, 20, 30). Briefly, DNA fragments encoding the N-truncated ORF2 protein (for the wild-type VLP) and the chimeric ORF2 protein (for VLP-C-tag) were cloned using the baculovirus transfer vector pVL1393 to yield pVLORF2. Insect Sf9 cells (Riken Cell Bank, Tsukuba, Japan) were used to produce recombinant baculovirus. Tn5 insect cells were infected with the recombinant baculoviruses at a multiplicity of infection of 5 and incubated in Ex-Cell 405 medium (JRH Biosciences, Lenexa, KS) for 6 days at 26.5°C. The supernatant was collected after the removal of cell debris by centrifugation at 10,000 × *g* for 90 min. The HEV VLPs were pelleted at 100,000 × *g* for 2 h in a Beckman SW32 Ti rotor and resuspended in 4.5 ml Ex-Cell 405. The VLPs were further purified by centrifugation through a CsCl density gradient (1.31 g/ml) at 110,000 × *g* for 24 h at 4°C in a Beckman SW 55 Ti rotor. The white virus band was collected and diluted 4 times with Ex-Cell 405 to decrease the CsCl concentration, and then the VLPs were centrifuged for 2 h in a Beckman TLA 55 rotor at 100,000 × *g*. The VLPs were resuspended in 100 to 500 μl of 10 mM potassium-MES (morpholineethanesulfonic acid) buffer (pH = 6.2) and stored at 4°C. To construct chimeric VLP-C-tag, recombinant baculoviruses were prepared by inserting the B-cell tag epitope from herpes simplex virus glycoprotein D (QPELAPEDPED) at amino acid position 608 (20).

Western blotting. A series of DNA fragments were constructed to encode truncated ORF2 residues 112 to 660, 112 to 608, 112 to 602, 112 to 601, 112 to 600, 112 to 596, and 112 to 589. These recombinant ORF2 genes were inserted into a baculovirus vector and expressed in insect cells using the protocol for VLP production, except that the recombinant proteins were recovered from the cytoplasm after lysis of the cell. Recombinant proteins were heated in 4× Laemmli sample buffer and electrophoresed under reducing conditions in a 10% SDS–polyacrylamide gel. After transfer of proteins to a polyvinylidene difluoride (PVDF) membrane, the membrane was blocked with TBS buffer (20 mM Tris, pH 7.6, NaCl) containing 0.5% Tween 20 (vol/vol) prior to overnight incubation with Fab224 fragments at a 1:10 dilution. After extensive washing with TBS buffer containing 0.05% Tween 20 (vol/vol), alkaline phosphatase-conjugated anti-mouse IgG (Fab specific) was incubated with the membrane for 1 h at room temperature. The blot was then washed and developed with the *p*-nitroblue tetrazolium-5-bromo-4-chloro-3-indolylphosphate (NBT-BCIP) reaction.

Preparation of VLP-Fab complexes for cryoelectron microscopy. The VLP-Fab complexes were prepared by incubating Fabs with VLPs at a molar ratio exceeding 1:300 (VLP versus Fabs) at 4°C overnight. To reduce the background density in the subsequent structural determination, highly pure VLP-Fab complexes were obtained using a short column containing Sephacryl 300, which resulted in the removal of the unbound Fab from the sample. The fractions containing VLP-Fab complexes were collected based on their optical density readings at a wavelength of 280 nm. The Fab binding occupancy was roughly estimated by performing SDS-PAGE (8-to-25% gradient) on the purified VLP-Fab complexes at a constant voltage using the Phast system (Pharmacia). The

particle morphology of VLP-Fab complexes was examined by negative-stain electron microscopy using 2% uranyl acetate.

Cryoelectron microscopy. Sample preparation and cryo-EM were performed following previously described, well-established procedures (13, 30). Briefly, a drop containing 3.5 μl of the sample was applied to a glow-discharged holey carbon-coated copper grid, blotted with a piece of filter paper for 3 s to remove the extra liquid, and quickly plunged into liquid ethane cooled by liquid nitrogen. Samples were frozen in a thin layer of vitrified ice. The grid was then transferred into a Gatan 626DH cryo holder and kept at a low temperature (−178°C) during the subsequent data collection. Micrographs were collected under low-dose conditions (<10 $e^-/Å^2$) using Kodak SO163 film at a magnification of ×45,000 on an FEI CM-120 electron microscope operated at 120 kV, and particles were photographed at a defocus range of 1,000 to 3,000 nm. Micrographs were visually inspected and selected based on a suitable particle concentration, optimal ice thickness, and minimal specimen drift. Only micrographs fulfilling these criteria were analyzed.

Image processing. Selected micrographs were digitized using a Heidelberg Primescan D8200 (Heidelberg, Germany) at a 14-μm scanning step size, corresponding to 3.11 Å per pixel of specimen space. Particles were manually picked and centered by cross-correlating each one against the circular average image. The astigmatism and defocus value were evaluated by the superimposed power spectra from all particles within a single micrograph. The contrast transfer function's first zero was approximately within the range of 17 to 20 $Å^{-1}$ for the data used for the structural determination. The self-common-lines algorithm (4) was used to yield the initial models for VLP-C-tag, VLP-Fab4, and VLP-Fab224. The origin and orientation search for each particle was carried out iteratively using the polar Fourier transformation (PFT) algorithm running on an AMD MP1800 MHz dual-processor Linux workstation (2). Three-dimensional reconstructions were computed by combining a set of particles with orientations that spread evenly in an icosahedral asymmetric unit using the Fourier-Bessel algorithm and by superimposing 5-3-2 icosahedral symmetry. To examine the reliability of the three-dimensional reconstruction, the data set was evenly divided into two parts at the final refinement step and two three-dimensional reconstructions were computed. The resolution was estimated using Fourier shell correlation (FSC) by assessing the agreement between these two reconstructions in Fourier space. Using a coefficient value of 0.5 as the criteria, the estimated resolutions of the three-dimensional reconstructions of VLP-C-tag, VLP-Fab224, and VLP-Fab4 were computed as 17.5 Å, 18.5 Å, and 24 Å, respectively.

The three-dimensional reconstructions were rendered and visualized using the Chimera program (22). The contour level was chosen at a value corresponding to 100% of the mass of the PORF2 protein. The electron density map was displayed in the isosurface mode, which builds a barrier to contour the density about a certain threshold.

Fitting the crystal structure into cryo-EM density maps. The density of the bound Fab molecule was determined from a difference density map, which was calculated by subtracting the cryo-EM map of unbound HEV T=1 VLP from the density map of the Fab-VLP complex. The cryo-EM map of unbound HEV VLP was published previously (30). Because the cryo-EM data for unbound VLP and the Fab-VLP complex were collected with the same FEI CM-120 electron microscope under similar imaging conditions, the difference density map was calculated by direct subtraction of the density of unbound VLP from the reconstruction of the Fab-VLP complex after normalizing the contrast between the two maps. The calculated difference map was used as a constraint in model fitting. Manual fitting was carried out by translational and rotational movement of the three-dimensional crystal structure of the PORF2 protein (PDB ID 2ZZQ) (31) into the cryo-EM density maps using program O (9). To obtain the best fit, the atomic model of the PORF2 subunit was treated as a rigid body. The fitting was first manually refined by minimizing the crashes between symmetry-related PORF2 molecules and then evaluated based on the cross correlation coefficient (CC value) between the cryo-EM density and the density computed from the fitted PORF2 coordinates. Fitting was halted when the CC value reached 80%. The figures were prepared using the program PyMOL (5), and the surface stereographic projection of the HEV VLP was prepared using the program RIVEM (29).

RESULTS

Binding of antibody MAb224 to PORF2. The binding of the monoclonal antibody Fab224 to PORF2 was examined via immunoblot analysis. A series of recombinant ORF2 proteins with C-terminal truncations were separated by SDS-PAGE on a 10% gel under reducing conditions and blotted with Fab224 (Fig. 1A). Fab224 recognized both reduced and denatured recombinant ORF2 proteins that contained amino acids 112 to 660, 112 to 608, 112 to 602, and 112 to 601. In contrast, recombinant ORF2 proteins composed of residues 112 to 600, 112 to 596, and 112 to 589 did not bind to Fab224. These data indicate that residues 597 to 601 are critical for Fab224 binding to PORF2. Because the recombinant ORF2 proteins were recovered from cell cytoplasm where multiple forms of PORF2 were reported (15), the positive bands observed at a low molecular weight may be the proteolytic products or degraded forms of ORF2 that contain the Fab224 binding sequence.

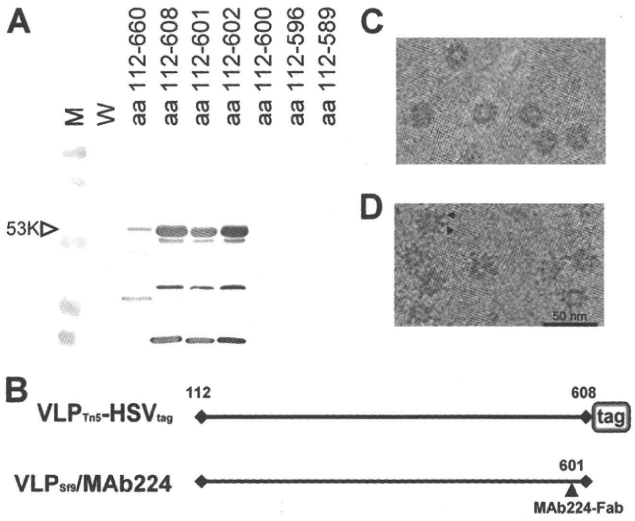


FIG. 1. Characterization of VLP-C-tag and VLP-Fab224. (A) Western blot assay of the C-terminally truncated ORF2 proteins with Fab224. M, molecular weight markers; W, peptides recovered from baculovirus-infected cells. (B) Diagram of the C-terminal markers. (C) Electron micrograph of frozen-hydrated VLP-C-tag. (D) Electron micrograph of frozen-hydrated VLP-Fab224. Black arrowheads indicate the Fab molecules attached to the VLP. Both particles showed an absence of density in the center. Note that the surface spikes in VLP-Fab224 appeared as longer thorn-like densities compared to those of VLP-C-tag.

Two-dimensional electron cryomicrographs. The chimeric VLPs (Fig. 1C) and the Fab224-conjugated VLP complex (Fig. 1D) showed circular profiles with spike-like densities that extended from the surface. As we observed previously (15, 30), they appeared to have a white, contrasting center, indicating that they are empty particles lacking RNA (data not shown). The sizes of both VLPs were approximately 27 nm without taking into account the extra densities that extended from the VLP-Fab224 surface (Fig. 1D).

Binding site of antibodies. The cryo-EM structure of HEV-Fab224 was reconstructed from 615 images of individual particles and displayed T=1 icosahedral symmetry with 60 protein subunits that were arranged into 30 dimeric protruding spikes located at each icosahedral 2-fold axis (Fig. 2A). Sixty Fab molecules were observed around each VLP particle, bound to the shoulder of the P domain. The Fab density extended ~57 Å radially away from the spike surface. The density corresponding to the Fab was approximately equal in magnitude to that of the HEV VLP, indicating that most or all of the 60