High-resolution cryo-EM structures of outbreak strain human norovirus shells reveal size variations

James Jung, Timothy Grant, Dennis R. Thomas, Chris W. Diehnelt, Nikolaus Grigorieff, and Leemor Joshua-Tor

Noroviruses are a leading cause of foodborne illnesses worldwide. Although GII.4 strains have been responsible for most norovirus outbreaks, the assembled virus shell structures have been available in detail for only a single strain (GI.1). We present high-resolution (2.6- to 4.1-Å) cryoelectron microscopy (cryo-EM) structures of GI.4, GI.2, GI.7, and GI.1 human norovirus outbreak strain virus-like particles (VLPs). Although norovirus VLPs have been thought to exist in a single-sized assembly, our structures reveal polymorphism between and within genogroups, with small, medium, and large particle sizes observed. Using asymmetric reconstruction, we were able to resolve a Zn²⁺ metal ion adjacent to the coreceptor binding site, which affected the structural stability of the shell. Our structures serve as valuable templates for facilitating vaccine formulations.

Noroviruses are a leading cause of foodborne illnesses, accounting for 58% of all outbreaks and over 96% of nonbacterial outbreaks, causing ~21 million cases in the United States and 685 million cases worldwide each year (1–3). Also commonly referred to as stomach flu or winter vomiting bug, the viruses cause frequent outbreaks of acute gastroenteritis in hospitals, nursing homes, day care centers, schools, and restaurants, and over 90% of diarrheal outbreaks on cruise ships. Norovirus illnesses are estimated at a cost of $2 billion for treatment and lost productivity in the United States and $60 billion worldwide annually (1–3). The viruses are highly contagious, with as few as 18 virus particles needed to cause infection (4), and spread by the fecal–oral route through direct contact with patients and aerosolized viruses from vomiting and contaminated surfaces, foods, and water supplies (5, 6).

Noroviruses are round, nonenveloped viruses with (+)ssRNA genomes that belong to the Caliciviridae family and are divided into at least 6 genogroups that are subdivided into 30 or more genotypes, of which genogroups I, II, and IV cause illnesses in humans (5, 6). The norovirus genome encodes two structural proteins, one major capsid protein (VP1) that forms the icosahedral shell enclosing the genome and a minor structural protein (VP2) that is positively charged and may interact with and stabilize its genome (7, 8). More recently, the VP2 of feline vesivirus was shown to form a portal structure upon binding of the feline receptor (9).

The major capsid protein consists of a shell (S) domain that forms the icosahedral enclosure and a protruding (P) domain that forms dimeric spikes on the virus surface and bears the antigenic features involved in host interactions (Fig. 1). The P domain consists of two subdomains: P1 emerging from the S domain and P2, which is an insertion within P1 and positioned at the outermost surface of the virus (SI Appendix, Fig. S1). The P2 subdomains have been shown to bind human histo-blood group antigens (HBGAs) as attachment factors, although some strains including GI.I2 Snow Mountain virus (SMV) show only weak or no binding to HBGAs (5, 10, 11).

Recombinant major capsid proteins assemble into virus-like particles (VLPs) that preserve the structure and antigenicity of infectious virions (7, 12). VLPs and animal noroviruses have been used as model systems, due to difficulties in propagating human virions in cell-culture systems (7, 13). There are no approved treatments available for norovirus infections, and VLPs are currently being used as candidates for vaccine trials; one candidate has reached a phase II clinical trial (12). Although crystal structures of P domains are available (10, 14), currently there is only one high-resolution VLP crystal structure of one norovirus strain (GI.1) available, which was solved at 3.4-Å resolution (7). The GI.1 VLP exhibits T = 3 icosahedral symmetry, an assembly of 180 subunits, with 90 dimeric spikes on its surface.

To further our understanding of human norovirus capsid architectures and the relationship to disease, we report high-resolution cryo-EM structures of outbreak strain norovirus virions in cell-culture systems (7, 13). There are no approved treatments available for norovirus infections. Assembled shells of the viruses without genetic materials enclosed are currently being used as candidates for vaccine trials. Although the virus shells have been thought to exist in a single-sized assembly, our structures in near-atomic detail reveal clear variations in size between different outbreak strains, and in spatial and angular arrangements of the antigenic surface spikes. The structures we present serve as valuable templates for facilitating vaccine formulations.

Significance

Despite being a leading cause of foodborne illnesses, accounting for 58% of all outbreaks and over 96% of nonbacterial outbreaks, there are no approved treatments available for norovirus infections. Assembled shells of the viruses without genetic materials enclosed are currently being used as candidates for vaccine trials. Although the virus shells have been thought to exist in a single-sized assembly, our structures in near-atomic detail reveal clear variations in size between different outbreak strains, and in spatial and angular arrangements of the antigenic surface spikes. The structures we present serve as valuable templates for facilitating vaccine formulations.

Author contributions: J.J., T.G., C.W.D., N.G., and L.J. designed research; J.J., T.G., and D.R.T. performed research; J.J., T.G., D.R.T., N.G., and L.J. analyzed data; and J.J. and L.J. wrote the paper.

Reviews: J.F.C., University of Pittsburgh; and M.W., Okinawa Institute of Science and Technology.

The authors declare no conflict of interest.

This open access article is distributed under Creative Commons Attribution-NonCommercial-NoDerivatives License 4.0 (CC BY-NC-ND).

Data deposition: The atomic coordinates and cryo-EM maps reported in this paper have been deposited in the Protein Data Bank (PDB) and Electron Microscopy Databank (EMDB) under the following ID codes: EMD-20199 for the GI.1 T = 3 symmetric reconstruction map, EMD-20205 for the GI.1 T = 3 symmetric reconstruction map, and PDB 6OTF for the GI.1 T = 3 atomic coordinates; EMD-20197 for the GI.2 T = 3 symmetric reconstruction map, EMD-20198 for the GI.2 T = 3 symmetric reconstruction map, and PDB 6OUF for the GI.2 T = 3 atomic coordinates; EMD-20195 for the GI.2 T = 3 symmetric reconstruction map and PDB 6OTF for the GI.2 T = 3 atomic coordinates; EMD-20201 for the GI.2 T = 1 symmetric reconstruction map, EMD-20202 for the GI.2 T = 1 symmetric reconstruction map, and PDB 6OUU for the GI.2 T = 1 atomic coordinates; and EMD-20206 for the GI.4 T = 4 symmetric reconstruction map and PDB 6OUU for the GI.4 T = 4 atomic coordinates.

1To whom correspondence may be addressed. Email: leemor@cshl.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1903562116/-/DCSupplemental.

Published online June 10, 2019.
cryo-EM VLP structures of four outbreak strains in the frozen hydrated state: GII.2 SMV (2.7 and 3.1 Å), GII.4 Minerva (4.1 Å), GI.7 Houston (2.9 Å), and GI.1 Norwalk (2.6 Å) strains (Fig. 2 and SI Appendix, Fig. S2 and Table S1). Notably, GII.4 strains have been responsible for 70–80% of all norovirus outbreaks, representing the most epidemiologically prevalent strain (5, 6). The VLP structures reveal polymorphism with T = 1, T = 3, and T = 4 assemblies observed. These structures serve as valuable templates for facilitating VLP vaccine formulations.

**Results and Discussion**

**Shell Domain and Size Variation.** Movies of VLPs embedded in vitreous ice on holey carbon grids were collected using a Titan Krios G3 transmission electron microscope (ThermoFisher) equipped with a K2 Summit direct electron detector (Gatan) (Materials and Methods and SI Appendix, Fig. S3). To improve the maps, we used an asymmetric focused reconstruction method that resolved features lost with symmetry imposition, using the program symmetryexpand_stack_and_par included with csTEM).

Although norovirus VLPs have been thought to exist only in T = 3 assemblies, our cryo-EM VLP structures revealed structural polymorphism between and within genogroups in identical buffer conditions (15, 16). Our GI.1 structure is in good agreement with the previous GI.1 crystal structure (0.8-Å rmsd). The GI.7 and GI.1 strains are also both in T = 3 assemblies (Fig. 2 and SI Appendix, Fig. S2). The external diameters of GI.7 and GI.1 particles are 420 and 410 Å, respectively, and the internal diameter of both strains is 240 Å. The GI.4 VLPs, on the other hand, are exclusively in T = 4 icosahedral symmetry consisting of 120 spikes and 240 subunits. Accordingly, the larger particles are 490 Å in external diameter and 280 Å in internal diameter. The GII.2 strain particles are observed in a T = 3 form as well as in T = 1 symmetry, the latter of which is likely to exist as an empty subviral form, with 30 spikes and 60 subunits. The external diameters of GII.2 T = 3 and T = 1 particles are 430 and 310 Å, respectively, and the internal diameters are 240 and 120 Å, respectively. We note that the T = 3 and T = 1 structures were observed together and solved from micrographs of the same sample.

The observed variances in the T numbers of norovirus VLPs are examples of quasi-equivalence, with identical capsid subunits undergoing conformational adjustments to fit multiple symmetry arrangements (17). Each T = 1 particle icosahedral asymmetric unit (ASU) consists of one subunit (A), whereas the T = 3 particle ASU consists of A, B, and C subunits in a quasi-threefold symmetry arrangement (Fig. 3). The T = 4 particle ASU consists of four subunits, A, B, C, and D, with the C–D dimer taking up a position similar to the C–C dimer of T = 3 particles. In the T = 4 particles, the icosahedral threefold axes are formed between three D subunits (Fig. 2 and SI Appendix, Fig. S2). In the T = 3 particles, B and C subunits interdigitate around the icosahedral threefold axes in quasi-sixfold symmetry. The T = 4 particle icosahedral twofold symmetry axes are formed between two sets of B, C, and D subunits in quasi-sixfold arrangements. The C–C dimer interfaces form the icosahedral twofold axes in T = 3 particles.

In the T = 3 particles of GII.2, GI.7, and GI.1, the two S domains of A–B dimers are in a bent arrangement and the C–C dimers are in plane (SI Appendix, Fig. S4). The N-terminal arm of GI.1 VP1 on the inner surface of the S domain has been suggested as a potential molecular switch that transitions between ordered and disordered states, where the ordered arm from the bent A–B dimer interacts with and stabilizes the neighboring flat C–C dimer conformation during T = 3 particle
to stabilize the flat dimer conformation in large norovirus particles. Indeed, truncation of the arm still resulted in self-assembly of GI.1 T = 3 particles, and the arm was dispensable for stabilization of the flat dimers in a previous study (18). In GI.2 small, T = 1 particles, all VP1 dimers are bent and the arms are disordered.

**Protruding Spike Domain.** Variations in the spatial and angular orientations of subunits and their domains would have consequences in host interactions, such as bivalent Ig binding. The P domain of GI.4 is lifted significantly (~4–28 Å) from the S domain through a long and flexible stalk region (Fig. 1). The P-domain densities were not as well resolved as those of the S domain. This may be due to the greater flexibility between the P domains and the shell in this assembly. Interactions between the elevated P domains form a double-layered structure with a distinct layer of spikes suspended above the icosahedral S-domain layer (Fig. 2). The dimeric P-domain spikes of A–B and C–D dimers are tilted toward one side (SI Appendix, Fig. S4). The distance between P and S domains is 16 and 24 Å on the A and B subunits, respectively, and 4 and 28 Å on the C and D subunits, respectively. On the other hand, the P domains of GI.2, GI.7, and GI.1 are placed close to the S domains, and the P1 subdomain makes contacts with the S domains of its own and of a neighboring subunit (Fig. 1). The lifted P-domain positioning and secondary-layer formation do not appear to be universally shared features among genogroup II strains. The P-domain spikes of GI.2 are twisted ~25° counterclockwise relative to the orientation of GI.1 and GI.7 (Fig. 2). The spikes of GI.4 are twisted ~40° clockwise relative to GI.1 and GI.7. Similarly lifted, twisted, and tilted P-domain positioning and orientations were previously reported in an 8-Å cryo-EM reconstruction of murine norovirus 1 native infectious particles in T = 3 form (13, 20).

Recently reported co cryocrystal structures of murine norovirus P domains in complex with their receptor, CD300lf, revealed the receptors bind at the interfaces between three P-domain spikes when superimposed on the map of the assembled virus (21–23). In GI.4 T = 4 particles, the interface between dimeric P-domain spikes forms around the 20 icosahedral threefold axes between D subunits, and the 60 quasi-threefold axes between A, B, and C subunits (Fig. 2). In the T = 3 particles of GI.2, GI.7, and GI.1, the equivalent interface surrounds the 60 quasi-threefold axes. Thus, a consequence of the larger, T = 4 form is the presence of a larger number of potential receptor attachment sites at the interfaces between three spikes, compared with those in T = 3 particles. On the other extreme, the assembly (7, 17, 18). The morphologically closely related tomato bushy stunt virus capsid with T = 3 symmetry, the arm of the C subunit wedges the two S domains apart in the C–C dimer interface to control the bent-to-flat conformation switching (17, 19). In our cryo-EM structures of GI.2, GI.7, and GI.1, the arm is ordered only on the B subunit that latches onto the neighboring C subunit.

In the large, T = 4, GI.4 particles, the two S domains of A–B dimers are bent as in the A–B dimers of T = 3 particles, and the C–D dimers are flat like the C–C dimers of those particles (SI Appendix, Fig. S4). However, the arm is disordered in all four subunits. It appears, therefore, that an ordered arm is not required.
spikes are placed significantly farther apart (∼10–25 Å) in the GII.2 T = 1 particle architecture, and no contacts are made.

**Asymmetric Reconstruction Reveals a Zn^{2+} Ion That Affects Shell Stability.** Positioned at the outermost surface of the virus, the P2 subdomain bears many of the antigenic features involved in host interactions. Two hypervariable loops (loop A residues 378–381 and loop B residues 295–298) on the outermost apex of P2 in GII.2 are not observed in the existing P-domain crystal structure (11). These loops are only partially resolved in our reconstructions with symmetry imposed. However, subsequent symmetry expansion assigning and extracting 60 icosahedrally related views of the ASU from each whole-particle image, and signal subtraction outside of each ASU, followed by asymmetric focused reconstruction of the ASU enabled near-complete observation of the two missing loops (Materials and Methods and SI Appendix, Fig. S5).

In HBGA-binding genogroup II strains, the conserved Asp382 (loop A) binds the fucose moiety of human HBGAs (14). In the GII.2 P-domain crystal structure, Asp382 was the last residue observed and the side chain was turned away ∼180° from the fucose binding site, which was proposed along with the dynamic loop movement as the possible molecular basis of weak or no HBGA binding by the GII.2 SMV strain. In our cryo-EM structures in the vitrified hydrated state, however, the Asp382 side chain is correctly oriented toward the fucose binding site on all subunits observed in a self-consistent manner (Fig. 4 and SI Appendix, Fig. S6).

The two flexible loops appear to be held by a metal-ion density on each subunit. Although partial disorder around the dynamic loop regions limits precise positioning, the metal appears to be bound between the His293 and His299 side chains and a water molecule. The metal-binding site is immediately adjacent to the fucose-binding Asp382 residue. To confirm the identity and significance of this observed metal binding, we added EDTA, EGTA, or 1,10-phenanthroline to GII.2 VLP samples at the same concentration. The addition of either EDTA or 1,10-phenanthroline that preferentially chelates Zn^{2+} caused an over threefold decrease in the number of intact VLPs observed, and the metal density disappeared in the asymmetric and symmetric reconstructions from the remaining intact particles, whereas EGTA, which preferentially chelates Ca^{2+}, did not cause such changes. The infectivities of murine norovirus and feline caliciviruses that do not bind HBGAs have been reported to be metal dependent (21, 22, 24). Moreover, GII.2 VLPs were reported to bind HBGAs when mixed with GII.2-positive patient stool, underscoring the involvement of coligands (25). This opens a possibility that metal ions may be required for the infectivity and stability of GII.2 SMV and possibly other strains by coordination between the flexible loops as an allosteric coligand, enhancing the human HBGA recognition and virus attachment.

The high-resolution structures of the human outbreak norovirus strain VLPs we have presented here serve as valuable templates for vaccine formulations and development of antivirals that inhibit host attachment and encapsidation. The structural polymorphism between and within genogroups should be valuable in guiding vaccine formulations that cover the broad spectrum of structural differences observed. Careful isolation of specific symmetry forms or a mixed sampling of all forms for vaccine manufacture may help to ensure maximum immunogenicity. The metal ions we identified present the surface antigen loops in the HBGA-binding conformations and allow raising neutralizing immunoglobulins. The addition of metal ions in the vaccine formulations would also be important in ensuring long-term structural integrity of the VLPs during storage of vaccines.

**Materials and Methods**

**VLP Preparation.** The G1.1 Norwalk, G1.4 Minerva, and GII.2 Snow Mountain virus strain VLPs were expressed in tobacco leaves. The G1.1 and G1.4 VLPs were prepared commercially by Kentucky Bioprocessing, and the GII.2 VLPs were produced by Hugh S. Mason and Andrew G. Diamos, Arizona State University, Tempe, AZ (26). The G1.7 VLPs were a kind gift from Robert L. Atmar, Baylor College of Medicine, Houston, TX. The VLPs were purified further by size-exclusion chromatography using a Superose 6 Increase 10/300 column (GE) in 20 mM MES OH (pH 5.75) and 50 mM NaCl, and the VLPs were collected from the void-volume fractions. The samples were concentrated to 3–5 mg/mL using centrifugal concentrators (Amicon).

**Cryo-EM Sample Preparation.** Lacey carbon grids (300 mesh; Electron Microscopy Sciences) were glow discharged using an easiGlow glow discharge (Pelco). Aliquots (4 μL) of VLP samples were applied to the grids and blotted for 1.7 s using an EM GP2 blotting system (Leica) at 22 °C and 95% humidity level with 10-s preblot incubation time before plunging in liquid ethane for vitrification. For metal-free experiments, the blotting papers were pretreated with 5 mM EDTA and then thoroughly washed in ultrapure water and dried before blotting.

**Data Collection.** Movies of VLPs embedded in vitreous solution were collected at liquid-nitrogen temperature using a Titan Krios G3 transmission electron microscope (ThermoFisher) equipped with a K2 Summit direct electron detector (Gatan) and a GIF Quantum LS imaging filter (Gatan). The movies were recorded in supersolution mode using EPU acquisition software (ThermoFisher) at 130,000x magnification with a pixel size of 0.535 Å per pixel, later resampled twofold to 1.07 Å per pixel, and nominal defocus range of 1.0–2.8 μm. The total electron exposure of each movie was ∼70–85 e/Å² with a nominal exposure rate of 2.0 electrons Å⁻² s⁻¹ per frame, fractionated into 35 movie frames with 200 ms per frame exposure time.

**Data Processing.** Many of the software packages used for data processing were available through SBGrid (27). Beam-induced motion of particles was corrected using UNBLUR (cistEM) on whole frames (28, 29). Contrast transfer function (CTF) parameters were estimated from sums of three movie frames using CTFFIND4 (cistEM) (29, 30). Particles were automatically picked ab initio using soft-edged disk templates internally generated in cistEM (29, 31). The picked particle images were boxed, extracted, and 2D-classified ab initio using RELION (32).
The GII.2 S-domain atomic models were built manually into the EM densities using Coot. The partial crystal structure of GII.2 (PDB ID code 4RPB) was fitted using MOLREP, modified to fit the EM densities, and manually built in previously missing regions using Coot. The manually built GII.2 S-domain model was fitted into the GII.4 EM densities using MOLREP, and manually rebuilt with the correct amino acid sequence for GII.4 using Coot. A crystal structure of the GII.4 P domain (PDB ID code 5IYN) was fitted and modified into the P-domain EM densities using MOLREP and Coot.

All protein models were real space-refined using PHENIX (38), and evaluated using Coot and the MolProbity server (39). Water molecules were added based on identifying EM density peaks with appropriate shape and hydrogen-bond interactions. The symmetric and asymmetric reconstruction cryo-EM maps were deposited in the Electron Microscopy Databank, and the coordinates of the atomic models were deposited in the Protein Data Bank (40, 41). The figures were generated using UCSF Chimera and ChimeraX (33).

Further details are provided in SI Appendix.

ACKNOWLEDGMENTS: We thank Kentucky BioProcessing (GI.1 Norwalk and GI.7 Houston strain) for producing the norovirus vaccine candidate in healthy adults: A randomized, controlled, double-blind, phase 2b study. J. Virol. 76, 66–73 (2002).

14. B. K. Singh, M. M. Leuthold, G. S. Hansman, Human noroviruses
15. T. Kilic, A. Koromyslova, V. Malak, G. S. Hansman, Atomic structure of the murine norovirus protruding domain and soluble CD300f receptor complex. J. Virol. 92, e00413-18 (2018).

16. C. A. Nelson et al., Structural basis for murine norovirus engagement of bile acids and the CD300f receptor. Proc. Natl. Acad. Sci. U.S.A. 113, E9201–E9210 (2018).

17. S. C. Harrison, The familiar and the unexpected in structures of icosahedral viruses. Nature 433 (2005), 151–152.

18. G. Leroux-Roels, J. M. Novembre, Outbreak management and disease prevention guidelines. MMWR Recomm. Rep. 59, 1 (2010).

19. T. Kilic, A. Koromyslova, V. Malak, G. S. Hansman, The Protein Data Bank.

20. B. V. Prasad et al., Structural basis for broad detection of genogroup II noroviruses in like particles in plants using modified geminiviral vectors. Proc. Natl. Acad. Sci. U.S.A. 105, 9175–9180 (2008).

21. S. Vongpunsawad, B. V. Venkataram Prasad, M. K. Estes, Norwalk virus minor capsid binding to histo-blood group antigens. J. Virol. 82, 421–433 (2016).

22. E. F. Donaldson, L. C. Lindsmith, A. D. Lobue, R. S. Baric, Virus shape-shifting: Norovirus evasion of the human immune system. Nat. Rev. Microbiol. 8, 231–241 (2010).

23. B. V. V. Prasad et al., X-ray crystallographic structure of the Norwalk virus capsid. Science 286, 287–290 (1999).

24. S. Vongpunsawad, B. V. Venkataram Prasad, M. K. Estes, Norwalk virus minor capsid protein VP2 associates within the VP1 shell domain. J. Virol. 87, 4818–4825 (2013).

25. M. I. Conley et al., Calicivirus VP2 forms a portal-like assembly following receptor mediated entry. J. Virol. 84, 2817–2821 (2010).

26. J. M. Choi, A. M. Hutson, M. K. Estes, B. V. Prasad, Atomic resolution structural characterization of recognition of histo-blood group antigens by Norovirus. Proc. Natl. Acad. Sci. U.S.A. 105, 9175–9180 (2008).

27. B. K. Singh, M. M. Leuthold, G. S. Hansman, Structural constraints on human norovirus binding to histo-blood group antigens. mSphere, 1 e00049-16 (2016).