Structure of the Infective Getah Virus at 2.8 Å-resolution Determined by Cryo-electron Microscopy

Aojie Wang[1], Feng Zhou[2], Congcong Liu[1], Dongsheng Gao[2], Ruxi Qi[1], Sheng Liu[1], Yinhe Xia[2], Chuanqin Wang[2,*], and Zheng Liu[1,*]

[1] Cryo-electron Microscopy Center, Southern University of Science and Technology, Shenzhen, Guangdong, China

[2] College of Animal Veterinary Medicine, Henan Agricultural University, Zhengzhou, Henan, China

Aojie Wang, Feng Zhou, and Congcong Liu contributed equally.

** Corresponding authors: C.W. (wchuanq@163.com), Z.L. (liuz3@sustech.edu.cn)*

Summary

Getah virus (GETV) is a mosquito-borne pathogen that can cause a mild illness and reproductive losses in animals. Although antibodies to GETV have been found in humans, there are no reports of clinical symptom associated with GETV. However, antivirals or vaccine against GETV is still unavailable due to lack of knowledge of the structure of GETV virion. Here, we present the structure of mature GETV at a resolution of 2.8 Å with capsid protein, envelope glycoproteins E1 and E2. Glycosylation and S-acylation sites in E1 and E2 are identified. The surface-exposed glycans demonstrated their impact on the viral immune evasion and host cell invasion. The S-acylation sites strongly stabilize the virion. In addition, a cholesterol and phospholipid molecule are observed in transmembrane hydrophobic pocket, together with two more cholesterols surround the pocket. These structural information are helpful for structure-based antivirals and vaccine design.

Introduction

Getah virus (GETV) is a mosquito-borne arbovirus and belongs Semliki Forest group of *Alphavirus* genus within the *Togaviridae* family (Lvov et al., 2015). Alongside GETV, members in Semliki Forest Group include Chikungunya virus (CHIKV), Semliki Forest virus (SFV), Mayaro virus (MAYV), Una virus (UNAV), Bebaru virus (BEBV), and O'nyong-nyong virus (ONNV). Among them, CHIKV, ONNV, MAYV have been reported to cause severe and mortally dangerous infection diseases in human (Hua and Combe, 2017; Rezza et al., 2017; Tesh et al., 1999).

GETV was first isolated in Malaysia in 1955 from *Culex gelidus* mosquitoes (Berge, 1975), and was found to have wide distribution in worldwide. The primary hosts of GETV include pigs, cattle, foxes and horses. The first outbreak of GETV among racehorses occurred in Japan in 1978 and re-emerged several times in Japan (Bannai et al., 2015). The infected horses exhibited pyrexia, urticarial rash on various portions of the body, and edema of the hind legs (Kamada et al., 1980). The emergence of GETV in China was found in 2017 in swine farms, resulting in the death of the newborn piglets 5-10 days after birth and pregnant sows having stillbirths or foetal mummies (Yang et al., 2018). GETV was also reported to infect beef cattle (Liu et al., 2019), blue fox (Shi et al., 2019), and wild boars (Sugiyama et al., 2009). Neutralizing antibodies of GETV have been detected in a variety of vertebrate species, suggesting more domestic animals could act as reservoir hosts (Li et al., 2019). Although the pathogenicity of GETV in human has not yet been identified, seroepidemiologic investigations have shown that some individuals with a febrile illness has significantly higher GETV antibody titres than in healthy people (Li et al., 1992), suggesting the correlation of GETV with human diseases. However, no treatment or vaccine against GETV is currently available.

As a typical alphavirus, GETV is a lipid-enveloped, positive-sense single-stranded RNA virus (Lvov et al., 2015). Mature virions of alphaviruses are spherical particles with a diameter of ~70 nm. The 11-kb genome of GETV encodes two polyproteins. Among them,

one polyprotein consists of four non-structural proteins (nsP1-nsP4), and another one polyprotein composes five structural proteins, in order from the N-terminal to the C-terminal, capsid, E3, E2, 6K and E1 (Rayner et al., 2002). The structures of alphaviruses have been well studied by using cryo-electron microscopy (cryo-EM). These include the structures of WEEV, Barmah forest virus (BFV), VEEV, CHIKV, EEEV, SINV, MAYV ranging from xxx-xxx Å, and structure of CHIKV with its receptor MXRA8 (Basore et al., 2019; Chen et al., 2020; Chen et al., 2018; Hasan et al., 2018; Kostyuchenko et al., 2011; Ribeiro-Filho et al., 2021; Sherman and Weaver, 2010; Yap et al., 2017; Zhang et al., 2011). These structures reveal a typical structure for alphavirus organization: virus RNA occupies the center of the particle and extends to approximate 140-Å radially, and RNA is surrounded by the capsid protein shell ranging between 140 to 200-Å radially; the lipid membrane shell (200 to 255-Å radially) separating the capsid and the outer glycoprotein shell, the outer shell and spikes protruding outward (255 to 350-Å radially) are formed by E1 and E2 glycoproteins. In addition, high-resolution maps have assisted to build the atomic models for the capsid, glycoproteins E1 and E2 (Chen et al., 2018; Hasan et al., 2018; Ribeiro-Filho et al., 2021; Zhang et al., 2011). Moreover, Mxra8, the receptor of CHIKV, was found to bind binds into a cleft created by two adjacent CHIKV E1-E2 heterodimers (Basore et al., 2019).

In previous studies, we have isolated a new GETV strain, V1 strain, from pregnant sows had abortions, and have demonstrated the V1 strain of GETV had a strong cytopathic effect and higher proliferation titer of vaccine (Zhou et al., 2020). In this study, we utilized mice model to investigate the infectiousness and pathogenicity of mature GETV. We show that GETV was mainly a mosquito-borne arbovirus, but could also spread via animal biting. We also show that GETV induced reproductive losses in pregnant mice, including abortion, foetal mummies and stillbirths. To gain structural insight into GETV, we determined the atomic structure of the mature GETV V1 virion at 2.8 Å-resolution by cryo-EM, the highest resolution for an alphavirus. Our results offer structural knowledge for rational design and screening of drugs, antibodies and vaccine against GETV.

Results

Infectiousness of mature GETV V1 strain in mice

GETV appears to be maintained in a natural cycle between mosquitoes and various vertebrate hosts. GETV also suspected to be transmitted directly between horses, probably via aerosols or by direct contact. There is no clear evidence how GETV transmits and infects by mosquitoes. Therefore, we first investigated the transmission and infectiousness of GETV using a mouse model. As shown in **Supplemental Fig. 1A**, mice (2-day old) were inoculated oronasally with 10μl (10^6 $TCID_{50}$/mL, 50% tissue culture infectious dose) GETV V1 strain (**Supplemental Fig. 1I** and **1J**), and housed in a cage with their GETV-free littermates. Total RNA was extracted from tissue, including spleen, lung, cerebral cortex, and various lymph nodes using TRIzol reagent after 5 days post inoculation (DPI) (Kumanomido et al., 1988), and subjected to reverse transcription-polymerase chain reaction (RT-PCR) for GETV detection using specific primers (GETV-F: 5′-ACCGAAGAAGCCGAAGAA-3′, and GETV-R: 5′-GCACTCRAGGTCATACTTG-3′) (Zhou et al., 2020) (**Supplemental Fig. 1G** and **1H**). The uninfected mice remained negative for GETV (**Supplemental Fig. 1A**). When male adult infected mice (2-month old) housed with their uninfected male littermates (**Supplemental Fig. 1B**), all mice were test GETV-positive at 5 DPI. Multiple wounds were found in various part of the body, presumably leading the transmission of GETV via scratching and biting.

We designed a special device which connecting two cages by a tunnel, and placed fences in the two ends of the tunnel to separate the mice in the two cages. When mosquitoes were cooped in the device, they could freely fly cross the fences (**Supplemental Fig. 1C**). In a control test as shown in **Supplemental Fig. 1D**, GETV-free mice in the two cages live with mosquitoes, they remained test GETV-negative after five days, indicated that mosquitoes used in our study were also GETV-free. Mice inoculated with GETV, then mixed with their GETV-free littermates in the left cage, additional GETV-free littermates were raised in the right cage, and GETV-free mosquitoes were enclosed in the device (**Supplemental Fig. 1E**). At 5 DPI, all mice in the left cage were GETV-positive, and 33%

of mice in the right cage were test positive as well. In another control test, no mosquitoes were cooped into the device, all GETV-free mice were GETV-negative at 5DPI (**Supplemental Fig. 1F**). This set of experiments confirmed that GETV is a mosquito-borne arbovirus, not an airborne virus. GETV mainly spread by mosquito biting, travel in the blood between mouse and mosquito. Occasionally, GETV can transmit from animal to animal, most likely via blood.

Pathogenicity of GETV V1 strain

Previous study has shown that GETV caused reproductive disorders in pregnant sows (Yang et al., 2018). We first investigated the pathogenicity of mature GETV in pregnant mice. Mice were inoculated oronasally with 100 μl (10^6 $TCID_{50}$/mL) GETV-V1 strain in early-gestation (embryonic day 6, E6), middle-gestation (E10), and late-gestation (E14). The E6 group mice have severe abortion (**Supplemental Fig. 2A**) and delivered foetal mummies (**Supplemental Fig. 2B** and **2F**), the E10 group delivered mainly mummies or stillbirths, but also delivered small portion of health pups, however, abortion still existed (**Supplemental Fig. 2A**, **2B**, **2C** and **2F**). The E14 mice delivered mixture of mummies, stillbirths and live infants, although some newborn mice were weak (**Supplemental Fig. 2B**, **2C**, **2D** and **2F**). In the contrast, no birth defect was detected in the mice inoculation with DMEM at E6 (**Supplemental Fig. 2E** and **2F**). Interestingly, although GETV-infected pregnant mice displayed serious reproductive losses, no other disease symptom was discovered. These results suggest that when pregnant mice infected with GETV, placental and fetal infection took place in the fetuses, and eventually caused fetal damage, such as abortion, mummies or stillbirths.

Next, we investigated the pathogenicity of mature GETV in newborn mice. 3-day old mice were inoculated oronasally with 10μl (10^6 $TCID_{50}$/mL) GETV-V1 strain, and clinical signs were observed daily. At 4 DPI, mice appeared mobility impairments in pelvic limbs (**Supplemental Fig. 3A** and **Movie 1**), by 8 DPI, all mice were accompanied with paralysis, mainly in pelvic limbs. Meanwhile, from 6 DPI, there were mice died, 2 days after paralysis manifestation, and by 12 DPI, all mice were died (**Supplemental Fig. 3B**).

On the opposite, no mice had paralysis nor death in the control group (**Supplemental Fig. 3A** and **3B**). Interestingly, when experiments carried out in 7-day old newborn or 2-month old adult mice, no such clinical sign was observed. Immune-histochemistry staining revealed that abundant GETV antigens were distributed in the hippocampal dentate gyrus of the brain and in the spinal cord under lumbar vertebrae from the infected newborn mice (**Supplemental Fig. 3C**). It suggested that the mobility impairments in pelvic limbs possible caused by viral infection in the nervous system.

Overall structure of GETV

To gain structural insight into GETV, we determined the structure of the mature GETV by cryo-EM. Virion particles were purified and characterized by polyacrylamide gel electrophoresis (**Supplemental Fig. 4**), by mass spectrometry (**Supplemental Fig. 5** and **6**), by negative stain electron microscopy (**Supplemental Fig. 7**) and cryo-EM (**Supplemental Fig. 8**). The whole structure of GETV was determined to about 4.1 Å by Relion3.1 and JSPR software (**Supplemental Fig. 9 and Supplemental Table 1**). To overcome the heterogeneity of the ~70nm virus, block-based reconstruction method was adopted (**Supplemental Fig. 9** and **10**) (Zhu et al., 2018). Three blocks (5/3/2-fold symmetry axis) of GETV cryo-EM maps at 2.8-3.0 Å were acquired (**Supplemental Fig. 10** and **11**). The density maps of capsid in the pentamers are sufficient to observe the bulk residues. However, the densities for the hexamers are still intermittent. The quality of this map could be improved by averaging 15 equivalent densities from the 3- and 2-fold symmetry axis blocks. After averaging, the correlation coefficient between the densities of pentamer and hexamer is 0.92 (**Supplemental Fig. 12**). By combining these three high resolution densities, the over-all structure of GETV was obtained at 2.81 Å final resolution (**Fig. 1A and 1B**). As the same with other alphaviruses, GETV has icosahedral symmetry (T=4), comprising 60 quasi-three-fold symmetry trimers (Q-trimer) and 20 icosahedral three-fold symmetry trimers (I-trimer). There are 240 capsid proteins connecting with the corresponding E2 proteins. Each asymmetric unit (ASU, **Fig. 1C and**

1D) consists four E1-E2-capsid heterotrimers (**Fig. 1E**), with three trimers forming one unabridged Q-trimer and one trimer forming one third of I-trimer.

The atomic model of E1-E2-capsid heterotrimers includes the full length of E1, E2 and the residues spanning 104-268 of capsid protein. The density for residues 1-103 of pentamer capsid cannot be observed in cryo-EM map, suggesting they are in flexible architecture.

The majority of E1 ectodomain was divided into 3 subdomains (**Fig. 1E**): domain I (residues 1-36, 130-164 and 274-281), domain II (residues 37-129 and 165-273), of which residues 86-96 forming the hydrophobic fusion loop, inserting into the canyon formed by E2 A and B subdomains, and domain III (residues 282-379) (Lescar et al., 2001). Each of these three domains includes a secondary structure of antiparallel β strands. The rest ectodomain amino acid (residues 380-401) comprise the stem loop, connecting domain III and the TM helix. The stem loop become accessible after pH-induced E1/E2 dissociation, packing onto the core of E1 homotrimer to form the post-fusion conformation (Liao and Kielian, 2006). The E2 ectodomain was divided into 3 subdomains (**Fig. 1E**): domain A (residues 1-143), domain B (173-242) and domain C (270-344) (Voss et al., 2010). In addition, residues 144-172 and 243-269 form the β-ribbon, which plays a critical role in conformational changes induced by low pH conditions (Li et al., 2010).

Molecular Interaction in the E1-E2-capsid heterotrimer

The E1, E2 and capsid proteins form the basic heterotrimer (**Fig. 1E**). There are five interacting regions between E1 and E2 proteins (**Fig. 2A**). In the first region, residues of V55, S57, P58, and S66 in E1 interact with E2 residues S239, R244, Q247 and S249, as shown in **Fig. 2B**. This first region located near the tip of domain II and the β-ribbon, which is similar to the other alphaviruses, such as E1 S57 with E2 H164 in SINV (Chen et al., 2018), and E1 S57 with E2 H170 in CHIKV (Voss et al., 2010). As reported, β-ribbon can make a conformational change along with the change of pH in the environment (Chen et al., 2018). Other residues in this region contains A92 and Y93 in E1 protein and H226 in E2 protein. These three residues located at the interface between the fusion loop and domain B, as reported in the SFV (Roussel et al., 2006) and CHIKV (Voss et al., 2010).

During alphavirus fusogenic transition, the domain B cap covering the fusion loop would be removed primarily. This region acts as the acid sensitive region to be responsible for triggering conformational change during virus entry cell host.

In the second region, residues of D253, A255, F257, G258 and C259 in E1 interact with E2 residues R297, G300, R329, and R336 (**Fig. 2C**). This area of interaction locates near the tail of domain II and domain C.

In the third region, residues of S310, D385, H386, and V388 in E1 interact with E2 W338, Q340 and T342 (**Fig. 2D**). This area locates at stem loop, loop of domain III and the tail loop of domain C. Together with the second interaction region, these two areas may be important for E1/E2 dimerization. E1 and E2 proteins pulled together through these interaction areas formed near by the stem loop of E1. Residues of E1 A361 and V398 interact with E2 H348 and Y358.

The forth interaction area locates at stem loop, loop of domain II and subdomain D, nearing the hydrophobic pocket (**Fig. 2E**). This region is thought to be important for the stability of the hydrophobic pocket, thus may affect the release of the lipid in the hydrophobic pocket and result the long-range allosteric of the whole E1-E2-capside heterotrimer.

The fifth E1, E2 interaction regions is composed by residues T421 of E1 and S380/S384 of E2 (**Fig. 2F**). This area locates between E1 and E2 TM helixes to stabilize these two transmembrane helixes.

The interaction region between E2 and capsid proteins is responsible for fixing E1-E2 heterodimer on the capsid. Those residues are E2 A400, T402 and capsid K133 (**Fig. 2G**). Almost all of these residues described above are conserved in the alphaviruses (**Supplemental Fig. 13-15**).

Inter- and Intra-ASU interactions

The E1-E2-capsid heterotrimer forms the Q-trimer and I-trimer. The ASU is formed by one Q-trimer and one third of I-trimer. Within the Q-trimer, there are three regions for inter-molecular interaction (**Fig. 3A**). The E2-E2 interaction locates at the middle of domain A in the first E2 protein and the joint of domain A and domain C, where is near by the tail of β-ribbon domain, in the neighboring E2 protein (**Fig. 3B**). The E1-E2 interaction locates at the helix on domain II and the domain C to stabilize the Q-trimer (**Fig. 3C**). Residues K151 and D148 in the first capsid interact with D120 and K122 in the neighboring capsid (**Fig. 3F**).

In addition, E1 of I-trimer interact with the adjacent E1 of Q-trimer at two regions. Residues Y192 and K123 at the loop of domain II on E1 of Q-trimer and T153 and E151 at domain I on E1 of I-trimer interact with each other separately (**Fig. 3D**). And residues T41 and N43 locate at the β sheets of domain II symmetrically (**Fig. 3E**). Different from the previous research (Chen et al., 2018; Roussel et al., 2006), the conserved H125 is not involved in the interaction between Q-trimer and I-trimer. The isoelectric point of T41 is similar to the S41 in the SINV. But there is a large difference between N43 and H125. Conformational change can't be triggered by this pair of interaction residues when the pH changes during the virus entry. However, the electrostatic attraction between K123 and E151, whose pI difference is much bigger than S41 and H125 in SINV, at neutral pH conditions will change into repulsion force, which may initiate the conformational change to form the fusogenic E1 trimer to help virus entry (Chen et al., 2018).

There are six regions form the contacts between two neighboring Q-trimers (**Fig. 4A**). The same E1 of Q-trimer1 interact with adjacent heterotrimer of Q-trimer2 in three regions. The first interaction region locates between domain I and domain III (**Fig. 4C**). And within the same two E1 proteins, residues D323 and K351 distribute in the corresponding domain III (**Fig. 4D**). Residues R437 and A417 locates between the tail of E1 and C-terminal Helix of E2 (**Fig. 4G**). Beyond those, the same E1 of Q-trimer1 interact with E1 of another adjacent heterotrimer between domain I and domain II (**Fig. 4B**). Except those interaction between the envelope glycoproteins, two sets of capsid-capsid interaction between two Q-trimers are also observed (**Fig. 4E** and **4F**). Interestingly, most of those

related interacting residues pairs are between acidic and basic residues. For instance, E151 and K123, R21 and D385, R289 and D311, D323 and K351, E184 and R223, K172 and E256. These electronic interactions may also become weak when pH changes during cell entry, which will also lead to form the fusogenic E1 trimer.

Glycosylation sites identified in the E1-E2 ectodomain

Viral envelope proteins have evolved to be extensively glycosylated with versatile functions ranging from immune evasion by glycan shielding to enhancement of cell infection (Watanabe et al., 2019). Glycosylation sites have been reported in CHIKV (Voss et al., 2010), EEEV (Hasan et al., 2018), SINV (Chen et al., 2018), and MAYV (Ribeiro-Filho et al., 2021).

We first predicted the N-glycosylation and O-glycosylation sites on GETV E1/E2 using NetNGlyc (Gupta and Brunak, 2002), NetOGlyc (Steentoft et al., 2013), GlycoMine (Li et al., 2015), and YinOYang (Gupta and Brunak, 2002). List of predicted glycosylation sites were summarized in **Supplemental Table 2**. We next utilized mass spectrometry method to identify glycosylation sites in the GETV E1 and E2 glycoproteins. **Supplemental Fig. 16** and **17** displayed possible N-glycosylation sites in the E1 and E2 sequences, and list of the sites are stated in **Supplemental Table 2**. In the final 2.8 Å-resolution GETV cryo-EM density map, four densities owing to glycan were identified associated with N141 and N270 of E1, and N200 and N262 of E1 (**Fig. 5**).

The N-glycosylation of E1 N141 and E2 N262 have been also reported in MAYV structure, model of glycans were built (Ribeiro-Filho et al., 2021). In our GETV structure, densities associated with E1 N141 and E2 N262 were both clearly visible, allowing the atomic models of multiple hexosamine and hexose monomers to be fitted. The proximal part of the density associated to E2 N262 were accommodated three carbohydrates, which were modeled as classical N-glycans NAG-NG-BMA (NAG: N-Acetyl-glucosamine and BMA: β-D-mannose, **Fig. 5G**). At the position of the third carbohydrate BMA, one branch density can be observed, in which 1-6MAN (α-(1-6)-D-mannose) monomer was modeled and fitted (**Fig. 5G**). Similar model of NAG-NAG was fitted in the density proximal to E1 N141

(**Fig. 5D**). The E2 N200 has density that also fitted the typical N-glycans, NAG-NAG-BMA (**Fig. 5F**). The E1 N270 has a less defined density compared to E1 N141, E2 N200 and N262, only one NAG monomer can be modeled (**Fig. 5E**).

All four glycosylation are surface-exposed (**Fig. 5A** and **5B**), they are accessible on the GETV surface close to the trimer/trimer interface (**Fig. 5B**). The surface exposure of glycosylation sites suggests their roles in evading detection by host immune system, and in attaching to the host cells to enhance virial uptake.

S-acylation sites identified in the E1-E2 TM helixes

S-acylation (prominently palmitoylation) is a post-translational modification that attaches fatty acid, usually palmitic acid, to cysteine residue of both peripheral and integral membrane proteins. Fatty acids increases protein hydrophobicity and facilitates protein interactions with lipid bilayers, it plays important roles in modulating proteins' stability (Linder and Deschenes, 2007). Protein acylation was originally discovered with vesicular stomatitis virus (Schmidt and Schlesinger, 1979). Over the past 40 years, numerous viruses have been reported undergoing palmitoylation, such as human immunodeficiency virus (HIV), severe acute respiratory syndrome coronavirus (SARS-CoV), influenza A virus, and hepatitis C virus (HCV) (Veit, 2012). For alphaviruses, both SFV and SINV have been reported contained S-acylation sites (Ivanova and Schlesinger, 1993; Ryan et al., 1998; Veit et al., 1996; Zhao et al., 1994). In the cryo-EM structure of BFV at 6 Å-resolution, density has been found next to residue Cys-394 in the E2 TM helix, however, atomic model of the fatty acid was not built (Kostyuchenko et al., 2011).

In the present study, we first predicted the possible S-acylation sites using predictors CSS-Palm and GPS-Palm (Ning et al., 2021; Ren et al., 2008). List of predicted S-acylation sites were summarized in **Supplemental Table 3**. Next, we utilized mass spectrometry method to identify S-acylation sites in the GETV E1 and E2 proteins. **Supplemental Fig. 16** and **17** summarized possible S-acylation sites in the E1 and E2 sequences, and list of the sites are stated in **Supplemental Table 3**. Finally, in the 2.8 Å-resolution GETV map, we examined the densities next to the residues. Total three

densities have been identified, one in E1 protein, C433, and two in E2 protein, C415 and C416 (**Fig. 6**). We accommodated an atomic model of stearic acid in the straight proximal portion of the density next to E1 C433 (**Fig. 6B**), and fitted atomic models of palmitic acid in the densities adjacent to E2 C415 and C416 (**Fig. 6E** and **6F**). Different fatty acids, stearic acid acylated to E1 and palmitic acid acylated to E2 protein were previously reported in a biochemical assay of SFV (Veit et al., 1996). In another assay, stearic acid was confirmed acylated to SFV E1, while one stearic acid and three palmitic acids were attached to SFV E2 (Kordyukova et al., 2010). Moreover, we observed two additional densities next to C385 and C395 in E2 protein. In our mass spectrum assay, the C385 was not recovered and no S-acylation was detected in C395 site. However, both densities displayed similar contour as E1 C433, E2 C415 and C416. According to previous report (Kordyukova et al., 2010), one site in the TM helix was acylated with a stearic acid and three sites in the cytoplasmic tail were acylated with palmitic acid, thus we modeled and fitted stearic acid in the density next to C385 and palmitic acid to C395 (**Fig. 6C** and **6D**). As illustrated in **Fig 6A**, the five high or absolutely conserved C-terminal S-acylation sites strongly anchored the E1 and E2 TM helixes into the inner leaflet of lipid bilayer of virial envelop membrane, consider the large portion of both E1 and E2 (**Fig 1E**, **Supplemental Fig. 16** and **17**), S-acylation in the TM helixes are propitious to stabilize the E1/E2 glycoproteins.

Cholesterol molecules and phospholipid molecule in the hydrophobic pocket

Previous study has demonstrated a hydrophobic pocket is particularly important for alphavirus assemble at the outer leaflet of envelop membrane (Chen et al., 2018; Ribeiro-Filho et al., 2021). The hydrophobic pocket composed with the TM helix from E1 protein, and the TM helix and subdomain D of E2 (see **Fig. 1E**). In the 3.5 Å-resolution SINV structure, a pocket factor was identified and predicted as one hydrophobic phospholipid tail (Chen et al., 2018). In the 4.4 Å-resolution MAYV structure, a long extra density was identified in the hydrophobic pocket, and a C18 hydrocarbon lipid molecule, octadecane, was fitted into the density (Ribeiro-Filho et al., 2021).

In our 2.8 Å-resolution GETV structure, extra densities were clearly visible in the cavity formed by E1 and E2 TM helixes, and in the region immediately adjacent to the two helixes (**Fig. 7A-7C**). Two cholesterol molecules were modeled and fitted into the densities adjacent to the helixes (**Fig. 7B**), while one cholesterol molecule was fitted in one density in the cavity (**Fig. 7C**). Cholesterol is the most abundant sterol in mammals, it is the major constituent of the plasma membrane and plasma lipoproteins. Previous study has revealed that nearly 30% of lipid class composition of baby hamster kidney cells (BHK, one of the most commonly used cell lines for expression of alphaviruses) were cholesterol, and over 40% of envelop membrane of SFV produced in BHK cells are cholesterol (Kalvodova et al., 2009). Another density in the cavity displayed a “Y” shape, and a phospholipid molecule was modeled as DOPC (**Fig. 7C**), with a hydrophilic "head" containing a phosphatidylcholine, and two hydrophobic "tails" derived from oleic acid, a mono-unsaturated omega-9 fatty acid. Both densities that fitted with two oleic acid hydrocarbon tails have a kink that fitted for the mono-unsaturated bond (**Fig. 7C**). Phosphatidylcholine is generally the most abundant lipid in animal cell membranes providing structural framework, previous study has indicated that over 23% of lipid class composition of BHK cells were phosphatidylcholine, and over 10% in the SFV envelop membrane (Kalvodova et al., 2009). Further analysis the component in the SFV revealed PC species were mainly di-unsaturated and mono-unsaturated, with 36:2 (18:1-18:1) hydrocarbon chain was the most abundant fatty acid tails (Kalvodova et al., 2009).

We further investigated the detail molecular interaction between E1/E2 proteins and cholesterol/DOPC molecules in the hydrophobic pocket. Cholesterol is a molecule derived from the sterane backbone, which contains four fused carbon rings referred to as A, B, C and D. It is the principal lipid interact with proteins through CH-π hydrogen bonds (Nishio et al., 2014). For typical cases, CH-π hydrogen bonds involving a soft acid CH and a soft base π-electron system (Nishio et al., 2014). As shown in **Fig. 7D**, in the cholesterol molecule at left, the B carbon rings interacted with aromatic ring from Tyr-359 in subdomain D, with mean distances of 4.0Å. The A ring in the right cholesterol contacted with aromatic ring from His-352 in subdomain D, with a distance of 4.4Å. This is consistent

with the establishment of the cut-off values (4.5Å) for the distance of the donor carbon atoms to the geometric center of the π -acceptor system (Nishio et al., 2014). **Fig. 7E** displayed the cholesterol molecule in the hydrophobic cavity, the carbon rings A and D interacted with Tyr-362 and Tyr-358 in subdomain D, with mean distances of 4.3 Å and 3.8 Å, respectively. CH-π hydrogen bond between a methyl group of DOPC and aromatic ring of Tyr-362 were identified, with a distance of 4.5Å. Moreover, hydrogen bonds between the phosphate group and Gln-397/Val-398/Phe-399 were affirmed. Beyond that, hydrophobic residues were found around hydrophobic tails of DOPC within 5Å. All of these three interaction regions fixed DOPC in the hydrophobic pocket (**Fig 7F**).

Discussion

Viral glycosylation has broad roles in viral pathobiology, including mediating protein folding and trafficking, facilitating viral budding and release, directing immune evasion by glycan shielding, and shaping viral infection and tropism (Watanabe et al., 2019). For alphaviruses, glycosylation is an important determinant of cellular tropism and virulence. Lack of glycosylation in RRV caused virions unable to induce alpha/beta interferon production in in myeloid dendritic cells, the professional antigen-presenting cells that rapidly respond to viruses, suggested that viral glycans could negatively regulate antiviral responses (Shabman et al., 2008). Glycosylation plays an important role for alphavirus infection and replication, lack of glycosylation in E1 protein in SINV severely impaired viral replication in vertebrate and invertebrate cells (Knight et al., 2009). Mutations of the glycosylation sites in the Salmonid alphavirus E1 and E2 proteins severely altered the virulence and production of an infectious virus (Aksnes et al., 2020).

Previous structural analyses have revealed glycosylation sites in the alphavirus. A crystal structure of the mature E3-E2-E1 glycoprotein complexes from CHIKV revealed three N-linked glycans, N12 in E3, N263 in E2, and N141 in E1 (Voss et al., 2010). The cryo-EM structure of MAYV identified two N-glycosylation sites, N141 in E1 and N262 in E2, At least five carbohydrates were modeled as NAG-NAG-BMA-MAN-MAN at E2 N262 site, and canonical NAG-NAG-BMA glycan sequence could be modeled at E1 N141 site (Ribeiro-Filho et al., 2021). Both E1 N141 and E2 N262 sites are conserved in CHIKV, MAYV, SFV, RRV and GETV, but not in other alphaviruses, such as SINDV, EEEV, WEEV, and VEEV. N-glycosylation sites were also identified in SINV by cryo-EM, N139 and N245 in E1, N283 in E2 and N14 in E3 all displayed extra density next to residue Asparagine, albeit no glycan model was built (Chen et al., 2018). In the present study, two new glycosylation sites, E1 N270 and E2 N200, have been identified, and we were able to build the atomic models of the glycans precisely into the 2.8 Å cryo-EM density map (**Fig. 5**).

Cholesterol is an essential structural component of cell membranes, and essentially an important structural component of viral envelop membrane, since the lipid membrane present in viruses acquire entirely from the host cells. For mammalian-derived SFV virions, the membrane consists high content of cholesterol, over 40% (Kalvodova et al., 2009). Interestingly, the cholesterol content in the envelope of MAYV particles obtained from mammalian cells have a 10-fold higher cholesterol content than viruses produced in mosquito cells, however, two virus samples did not significantly differ in their infectious titer (Sousa et al., 2011). Similar pattern of cholesterol content were found in SINV grown in mammalian and in insect cells (Hafer et al., 2009). In all studies, BHK cells were used as the mammalian host cells, same one we used in the current study to propagate and purify GETV. There is no doubt to have cholesterol in the envelop membrane in the GETV, but surprisingly we identified one cholesterol molecule in the hydrophobic pocket and two cholesterol molecules surround the TM helixes of E1 and E2 proteins, indicated that cholesterol bound tightly to E1 and E2 (**Fig. 7**).

Several studies have addressed the role of cholesterol present in viral envelope for viral infections, depletion of cholesterol from envelopes abolished virus infectivity, included influenza virus, human herpesvirus, herpes simplex virus, hepatitis C virus, and human immunodeficiency virus (HIV) (Aizaki et al., 2008; Campbell et al., 2002; Huang et al., 2006; Sun and Whittaker, 2003). In all studies, cholesterol was removed from virions by methyl-beta-cyclodextrin. In the current study, we have revealed that cholesterol molecules were tightly bound to GETV, it is particular interest to examine if methyl-beta-cyclodextrin could reduce GETV infectivity. For alphaviruses, primarily for CHIKV, SFV, SINV, and MAYV, cholesterol is required during viral entry to host cells, the first step of viral infection (Carvalho et al., 2017; Lu et al., 1999; Phalen and Kielian, 1991; Tsetsarkin et al., 2007). After entry of the viral particles into the host cells, the viral RNA is released and replicated. This process also depends on intracellular membranes, and cholesterol showed an impact on the viral replication. SINV replication was faster and reached higher titer in cholesterol-rich fibroblasts, virions produced in cholesterol-rich fibroblasts are more infectious than viral particles budding from normal human fibroblasts (Ng et al., 2008). Similarly, drugs that induced intracellular cholesterol accumulation and affected

cholesterol biosynthesis, displayed a strong inhibition on CHIKV replication (Wichit et al., 2017). Notably, gene expression profiling performed in cell lines that infected with M1 strain of GETV, which was isolated from *Culex* mosquitoes and known to infect horses and pigs (Zhai et al., 2008), demonstrated that more than 60% of genes associated with the cholesterol biosynthesis pathway are downregulated after GETV-M1 infection (Liang et al., 2018). Several studies have documented that cholesterol plays a critical role in SFV budding and release from the host cell (Lu and Kielian, 2000; Marquardt et al., 1993). Taken together, cholesterol is not only a molecular factor that essential for alphavirus entry, replication, budding and exit from host cells, but also a structural factor contributing to the stability of virions.

Beside cholesterol, phospholipids are other primary components of cell membrane and viral envelop membrane. Dissecting the lipid compositions of SFV by quantitative shotgun mass spectrometry revealed that viral envelop membrane comprised with phospholipids, major phospholipids include phosphatidic acid (PA), phosphatidylcholine (PC), phosphatidylethanolamine (PE), phosphatidylglycerol (PG), phosphatidylinositol (PI), and phosphatidylserine (PS). PC presented over 20% in SFV membrane produced from BHK cells (Kalvodova et al., 2009). Further analysis the component in the SFV revealed PC species were di-unsaturated and mono-unsaturated, mainly PC 36:2 (18:1-18:1), PC 34:1 (18:1-16:0), PC 32:1 (16:1-16:0), and PC 32:0 (16:0-16:0) (Kalvodova et al., 2009). This finding was later confirmed in another study of lipid components of SINV, oleic acid (18:1) was the most abundant fatty acids detected in SINV from BHK cells (Dunbar et al., 2019). Thus, it is no doubt to have DOPC (1,2-dioleoyl phosphatidylcholine) in the envelop membrane in the GETV, moreover, we identified one DOPC in the hydrophobic pocket, indicated the phospholipid molecule, together with cholesterol, bound tightly to E1 and E2 (**Fig. 7**).

Protein S-acylation is a fundamental post-translational lipid modification that attach fatty acids, usually palmitic acid, to cysteine (S-palmitoylation), and less frequently to serine and threonine (O-palmitoylation) residues of proteins. S-acylated proteins are typically membrane proteins, S-acylation enhances protein hydrophobicity and plays important

roles in the regulation of protein subcellular localization, trafficking, stability and interaction with other proteins (Veit, 2012). S-acylation of viral proteins is known to be involved in virus assembly and infection. For example, the Spike protein of severe acute respiratory syndrome coronavirus (SARS-CoV), and quite recently, of SARS-CoV-2, both have been shown to be palmitoylated that were important for cell-cell fusion that essential for viral infectivity (McBride and Machamer, 2010; Wu et al., 2021). For alphaviruses, both E1 and E2 glycoproteins of SFV have been reported contained S-acylation sites, and E1 contained mainly stearic acid, whereas E2 was acylated primarily with palmitic acid (Veit et al., 1996). Palmitoylated a cysteine residue in E2 protein of SFV was an essential determinant which were absolutely needed for virial budding (Zhao et al., 1994). In another alphavirus SINV, E1 and E2 also have palmitoylation sites, site-directed mutagenesis of cysteines in E1 and E2 abolished fatty acids binding, and displayed aberrant in assembly and particle formation, slow replication at early times post-infection and more sensitivity to detergent (Ivanova and Schlesinger, 1993; Ryan et al., 1998). These data support an earlier study indicating that mutations in SINV E2 glycoprotein led to defects in covalent palmitic acid binding, as well as defects in virus assembly and budding (Gaedigk-Nitschko and Schlesinger, 1991). In the cryo-EM structure determination of BFV at 6 Å-resolution, density has been found next to one cystine residue C394 in the E2 protein, however, atomic model of the palmitic acid was not built (Kostyuchenko et al., 2011). In another cryo-EM study of VEEV at 4.4 Å-resolution, three conserved cysteine residues in E2 protein, C396, C416, and C417 were mapped near the lipid head groups of the inner membrane, albeit, density attributed to palmitoylated fatty acids were not found (Zhang et al., 2011).

In addition to structural proteins, the non-structural protein nsP1, which is part of the replication complex of alphaviruses, has been reported to be palmitoylated. Palmitoylated cysteine residues in CHIKV nsP1 were critical for targeting plasma membrane of host cells and had functional consequence for virial genome replication (Bakhache et al., 2020). One cysteine in SINV and three cysteines in SFV were modified by palmitoylation, when the cysteines were mutated to preventing palmitoylation, viruses remained viable and replicated to high titers, but their growth was slightly delayed (Ahola et al., 2000). Also

reported in SFV, nsP1 with a mutation in the palmitoylation site resulted in reduced synthesis of genomic and sub-genomic RNAs and their products of translation (Kiiver et al., 2008). Taken together, both non-structural protein and structural proteins of alphavirus bearing palmitoylated residues, which are function important to virial replication and assembly.

In summary, our 2.8 Å-resolution cryo-EM density map of GETV provided an unprecedented level of structural features for alphavirus, including *ab initio* atomic models of the entire E1 and E2 proteins, and 60% of the capsid protein, numerous glycosylation sites with atomic model of glycans in the E1 and E2 proteins, atomic models of cholesterol molecules and phospholipid molecule in and around the hydrophobic pocket, atomic models of fatty acid at the palmitoylated residues in the transmembrane helixes of E1 and E2 protein. The knowledge of high-resolution three-dimensional structure of biomolecular targets are essential for structure-based drug design, our map should actuate more rational design and screening for drugs and antibodies, and for vaccine development.

Methods

Virus Infection

All protocols were approved by the Institutional Animal Care and Use Committee of College of Animal Veterinary Medicine, Henan Agricultural University. Mice were apparently healthy and had no serum neutralizing antibody against GETV before the experiment. Newborn mice (2-day old or 3-day old) were inoculated oronasally with 10μl (10^6 $TCID_{50}$/mL, 50% tissue culture infectious dose) GETV-V1 strain. Pregnant mice were inoculated oronasally with 100 μl (10^6 $TCID_{50}$/mL) GETV-V1 strain in early-gestation (embryonic day 6, E6), middle-gestation (E10), and late-gestation (E14).

Virus production and purification

Mature and infective GETV virions were handled in BSL2 facilities at Cryo-Electron Microscopy Center, Southern University of Science and Technology, and at College of Animal Veterinary Medicine, Henan Agricultural University. GETV-V1 strain (GenBank Sequence Accession: KY399029.1) was incubated in monolayers Baby Hamster Kidney Fibroblast Cells (BHK-21, ATCC CCL-10) at an approximate multiplicity of infection (MOI) of 0.1 for 30 min to allow viruses bind and entry. The supernatant was removed and replaced with fresh culture medium (DMEM with 2% FBS and 1% PS). GETV were proliferated in 10 T175 culture flasks and harvested 36 h post infection. The cell debris of the supernatant was removed at 10,000 × g for 1h (Sorvall LYNX 6000 Superspeed Centrifuge). The GETV particles were concentrated by centrifuging at 80,000 × g for 1.5 h in Ultracentrifuge (Beckman Optima XPN-100) with a type 45 Ti rotor at 4 °C. After soaked in 30 ml phosphate-buffered saline (PBS, pH 7.47) for 4h, the virus suspension was initially purified by through a 20% (w/v) sucrose cushion at 80,000 × g for 1.5 h in a type SW32 Ti rotor at 4 °C. The virus pellet was gently resuspended in 2 ml PBS buffer for 4h and loaded onto a linear 20-50% (w/v) sucrose density gradient at 100,000 × g for 2 h in a SW41 Ti rotor at 4 °C. The light scattering band corresponding to the virus particles collected and resuspended in PBS. The integrity and purity of the GETV particles were examined by SDS-PAGE and mass spectrometry, and by negative-stain electron microscopy.

Cryo-EM sample preparation and data acquisition.

A 5 μl virus sample was added to the glow-discharged 400-mesh grid covered with amorphous nickel-titanium alloy film (ANTA film, R2/1) at 100% humidity and 6 °C, blotted with filter paper and frozen by plunging into liquid ethane using a Vitrobot Mark IV system (Thermo Fisher Scientific Inc.). Cryo-EM data were collected on a 300 kV Titan Krios microscope (Thermo Fisher Scientific Inc.) equipped with a BioContinuum Imaging Filter (Gatan Inc). The images were recorded on a K3 summit direct detection camera (Gatan Inc) using SerilEM software for automated image acquisition. The images were recorded at a nominal magnification of 105,000 × in super-resolution mode, yielding a calibrated pixel size of 0.83 Å. Each exposure was dose-fractionated into 32 frames leading to a total dose of 40 $e^-/Å^2$. The final defocus range of the micrographs was -0.8 to -1.5 μm.

Image processing.

A total of 16,894 movie stacks of the GETV were collected. The beam-induced drift was corrected using MotionCor2 (Zheng et al., 2017). The contrast transfer function (CTF) parameters were estimated using CTFFIND4 (Rohou and Grigorieff, 2015). In total, 171,059 particles were extracted using the EMAN2 (Tang et al., 2007). The 4× binned particle images were subjected to reference-free 2D classification in RELION-3 (Zivanov et al., 2018). After several rounds of reference-free 2D classification, a subset of 144,254 particles was isolated for 3D classification and reconstruction. Two maps at 8.94 Å-resolution and 9.54 Å-resolution were obtained with 56,573 and 43,653 good particles, respectively. The coordinates were used to extract the 2× binned particles for 3D refinement using RELION-3 with I3 symmetry imposed, which resulted in two maps at 6.64 Å-resolution and 6.69 Å-resolution were obtained with 43,775 and 62,415 good particles, respectively. The rotational and translational parameters of each particle were further refined with jalign program from JSPR, which enabled the reconstruction of a better map at 4.1 Å-resolution using the j3dr program in JSPR (Guo and Jiang, 2014). To overcome the defocus gradient and the heterogeneity of the ~70nm virus, block-based reconstruction was performed to further improve the resolution (Zhu et al., 2018). Three

types of block (five trimers near icosahedral five-fold axis, four trimers near icosahedral-three-fold axis and two trimers near icosahedral two-fold axis) were selected, 3D classified and refined separately. Maps at 3.05 Å, 3.28 Å, and 3.14 Å, were obtained using unbinned particles. At this stage, CTF refinement was carried out to better estimate the local defocus value for each block, final round of 3D refinement generated density maps in resolution of 2.81 Å, 2.92 Å, and 2.85 Å, respectively, as estimated by the gold-standard Fourier shell correlation (FSC) cut-off value of 0.143 (**Supplemental Fig. 9** and **10**). The local resolution distribution of the final reconstruction was assessed using ResMap (Kucukelbir et al., 2014) (**Supplementary Fig. 10**).

Model building and refinement.

An asymmetric subunit of VEEV model (PDB ID: 3J0C) was used as template and rigid-body fitted into the 2.8 Å GETV map using UCSF Chimera (Pettersen et al., 2004). The amino acid sequences of VEEV were then mutated to GETV and an initial 3D atomic model of E1–E2-Capsid was built using Coot (Emsley et al., 2010). The resulting GETV coordinates was then used as a starting point to a flexible refinement through the PHENIX (Adams et al., 2010). The refinement cycle was repeated, and the quality of the final 3D atomic model of E1-E2-Capsid was evaluated using MolProbity (Chen et al., 2010). The Cryo-EM data collection and processing, block-based reconstruction, model building, and refinement statistics are summarized in **Supplementary Table 1**.

Protein identification by liquid chromatography triple quadrupole mass spectrometry (LC-MS/MS)

The proteins of GETV were separated and by denaturing SDS-PAGE. After rinsing with ddH_2O for 2 times, 100μl of decolorizing solution was added and placed at room temperature to decolor overnight. 100% acetonitrile were added until the gel mass turns white. 10μl 0.01μg/μl trypsin was added in the gel to fully absorb by iced until it become transparent. The concentrated and dried trypsin-cut peptide samples were re-dissolved in Nano-RPLC mobile phase A (0.1% formic acid, 5% acetonitrile/H2O) and analyzed by on-line LC-MS/MS. The dissolved sample was fed to a C18 pre-column. The sample was

desalted and retained on the pre-column and then separated by the analytical column. An elution gradient of 5-40% solvent B (0.1% formic acid, 95% acetonitrile/H2O) was used on an analytical column. DDA (data-dependent acquisition) mass spectrum techniques were used to acquire tandem MS data on a TripleTOF 5600+ (AB SCIEX, USA) fitted with a NanoSpray®III ion source (AB SCIEX, USA). Data was acquired using an ion spray voltage of 2.3kV and an interface at 150 °C. The MS2 spectra were acquired in the ion trap in rapid mode with a maximum injection time of 80 ms. The collision cell energy setting was applicable to all precursor ion Collision-Induced Dissociation (CID). The original file of MS was imported into ProteoWizard software for data processing and protein identification in mascot (V2.3.02).

Quantitative N-glycosylation and Palmitoylation analysis by matrix-assisted laser desorption/ionization time-of-flight/time-of-flight (MALDI-TOF/TOF)

Protein bands were excised for MALDI-TOF/TOF to quantitative N-glycosylation and Palmitoylation analysis. E1 and E2 gel bands were dissolved by 200μl of 8M urea with 10mM DTT and digested by 100μl of 25mM ammonium bicarbonate containing 0.01μg/μl trypsin before collected and lyophilized in a sterile centrifuge refer to the protocol1. Dissolve the dried polypeptide in 40mM NH4HCO3 prepared with O18. 2 unit of rPNGase F was defined as the amount of enzyme that digests 100μg polypeptide. Incubated at 37 °C for 3h. The lyophilized peptide fractions were re-suspended in ddH_2O containing 0.1% formic acid. The online Chromatography separation was performed on the EASY-nLC1200 system (Thermo Fisher, USA). The trapping, desalting procedure were carried out with 20μl solvent A (0.1% formic acid). An elution gradient of 5-38% solvent B (80% acetonitrile, 0.1% formic acid) in 60 min was used on an C18 analytical column. DDA mass spectrum techniques were used to acquire tandem MS data on a Thermo Fisher™ Q Exactive™ Mass Spectrometer (Thermo Fisher, USA) fitted with a Nano Flex ion source. Data was acquired using an ion spray voltage of 1.9kV and an interface at 275 °C. For a full mass spectrometry survey scan, the target value was 3×106 and the scan ranged from 350 to 2,000m/z at a resolution of 70,000 and a maximum injection time of

100ms.For the MS2 scan, only spectra with a charge state of 2-5 were selected for fragmentation by higher-energy collision dissociation (HCD) with a normalized collision energy of 28eV. N-linked glycans can be released from the glycoprotein with the enzyme N-glycosidase F(PNGase F). Asparagine residues can cause a mass increase of 2.9882 Da in “heavy” water2 and identified the N-glycosylation sites by LC-MS/MS peptide map. Palmitoylation modification occurs on CKST amino acids which the molecular weight change after modification is 238.23 Da. The MS data were analyzed for protein modification using PEAKS Studio 8.5.

Data availability

The atomic coordinates of E1-E2-Capsid of the GETV has been deposited in the Protein Data Bank under accession code 7FD2. The cryo-EM map of the envelope glycoprotein and capsid core has been deposited in the Electron Microscopy Data Bank under accession code EMD-31533.

Declaration of Interests

The authors declare no competing financial or non-financial interests.

Acknowledgements

We thank all staff members of the Cryo-Electron Microscopy Centre, Southern University of Science and Technology, for their assistance with data collection. This work was supported by National Natural Science Foundation of China (81870246 and 82070329)

References

Adams, P.D., Afonine, P.V., Bunkóczi, G., Chen, V.B., Davis, I.W., Echols, N., Headd, J.J., Hung, L.W., Kapral, G.J., Grosse-Kunstleve, R.W.*, et al.* (2010). PHENIX: a comprehensive Python-based system for macromolecular structure solution. Acta Crystallogr D Biol Crystallogr *66*, 213-221.

Ahola, T., Kujala, P., Tuittila, M., Blom, T., Laakkonen, P., Hinkkanen, A., and Auvinen, P. (2000). Effects of palmitoylation of replicase protein nsP1 on alphavirus infection. J Virol *74*, 6725-6733.

Aizaki, H., Morikawa, K., Fukasawa, M., Hara, H., Inoue, Y., Tani, H., Saito, K., Nishijima, M., Hanada, K., Matsuura, Y.*, et al.* (2008). Critical role of virion-associated cholesterol and sphingolipid in hepatitis C virus infection. J Virol *82*, 5715-5724.
Aksnes, I., Markussen, T., Braaen, S., and Rimstad, E. (2020). Mutation of N-glycosylation Sites in Salmonid Alphavirus (SAV) Envelope Proteins Attenuate the Virus in Cell Culture. Viruses *12*, 1071.
Bakhache, W., Neyret, A., Bernard, E., Merits, A., and Briant, L. (2020). Palmitoylated Cysteines in Chikungunya Virus nsP1 Are Critical for Targeting to Cholesterol-Rich Plasma Membrane Microdomains with Functional Consequences for Viral Genome Replication. J Virol *94*, e02183-02119.
Bannai, H., Nemoto, M., Ochi, A., Kikuchi, T., Kobayashi, M., Tsujimura, K., Yamanaka, T., and Kondo, T. (2015). Epizootiological Investigation of Getah Virus Infection among Racehorses in Japan in 2014. J Clin Microbiol *53*, 2286-2291.
Basore, K., Kim, A.S., Nelson, C.A., Zhang, R., Smith, B.K., Uranga, C., Vang, L., Cheng, M., Gross, M.L., Smith, J.*, et al.* (2019). Cryo-EM Structure of Chikungunya Virus in Complex with the Mxra8 Receptor. Cell *177*, 1725-1737.e1716.
Berge, T.O. (1975). International catalogue of arboviruses, including certain other viruses of vertebrates. [Washington]: US Dept of Health, Education, and Welfare, Public Health Service 278-279.
Campbell, S.M., Crowe, S.M., and Mak, J. (2002). Virion-associated cholesterol is critical for the maintenance of HIV-1 structure and infectivity. Aids *16*, 2253-2261.
Carvalho, C.A.M., Silva, J.L., Oliveira, A.C., and Gomes, A.M.O. (2017). On the entry of an emerging arbovirus into host cells: Mayaro virus takes the highway to the cytoplasm through fusion with early endosomes and caveolae-derived vesicles. PeerJ *5*, e3245.
Chen, C.L., Hasan, S.S., Klose, T., Sun, Y., Buda, G., Sun, C., Klimstra, W.B., and Rossmann, M.G. (2020). Cryo-EM structure of eastern equine encephalitis virus in complex with heparan sulfate analogues. Proc Natl Acad Sci U S A *117*, 8890-8899.
Chen, L., Wang, M., Zhu, D., Sun, Z., Ma, J., Wang, J., Kong, L., Wang, S., Liu, Z., Wei, L.*, et al.* (2018). Implication for alphavirus host-cell entry and assembly indicated by a 3.5Å resolution cryo-EM structure. Nat Commun *9*, 5326.
Chen, V.B., Arendall, W.B., 3rd, Headd, J.J., Keedy, D.A., Immormino, R.M., Kapral, G.J., Murray, L.W., Richardson, J.S., and Richardson, D.C. (2010). MolProbity: all-atom structure validation for macromolecular crystallography. Acta Crystallogr D Biol Crystallogr *66*, 12-21.
Dunbar, C.A., Rayaprolu, V., Wang, J.C., Brown, C.J., Leishman, E., Jones-Burrage, S., Trinidad, J.C., Bradshaw, H.B., Clemmer, D.E., Mukhopadhyay, S.*, et al.* (2019). Dissecting the Components of Sindbis Virus from Arthropod and Vertebrate Hosts: Implications for Infectivity Differences. ACS Infect Dis *5*, 892-902.
Emsley, P., Lohkamp, B., Scott, W.G., and Cowtan, K. (2010). Features and development of Coot. Acta Crystallogr D Biol Crystallogr *66*, 486-501.
Gaedigk-Nitschko, K., and Schlesinger, M.J. (1991). Site-directed mutations in Sindbis virus E2 glycoprotein's cytoplasmic domain and the 6K protein lead to similar defects in virus assembly and budding. Virology *183*, 206-214.
Guo, F., and Jiang, W. (2014). Single particle cryo-electron microscopy and 3-D reconstruction of viruses. Methods Mol Biol *1117*, 401-443.

Gupta, R., and Brunak, S. (2002). Prediction of glycosylation across the human proteome and the correlation to protein function. Pac Symp Biocomput, 310-322.
Hafer, A., Whittlesey, R., Brown, D.T., and Hernandez, R. (2009). Differential incorporation of cholesterol by Sindbis virus grown in mammalian or insect cells. J Virol *83*, 9113-9121.
Hasan, S.S., Sun, C., Kim, A.S., Watanabe, Y., Chen, C.L., Klose, T., Buda, G., Crispin, M., Diamond, M.S., Klimstra, W.B.*, et al.* (2018). Cryo-EM Structures of Eastern Equine Encephalitis Virus Reveal Mechanisms of Virus Disassembly and Antibody Neutralization. Cell Rep *25*, 3136-3147.
Hua, C., and Combe, B. (2017). Chikungunya Virus-Associated Disease. Curr Rheumatol Rep *19*, 69.
Huang, H., Li, Y., Sadaoka, T., Tang, H., Yamamoto, T., Yamanishi, K., and Mori, Y. (2006). Human herpesvirus 6 envelope cholesterol is required for virus entry. J Gen Virol *87*, 277-285.
Ivanova, L., and Schlesinger, M.J. (1993). Site-directed mutations in the Sindbis virus E2 glycoprotein identify palmitoylation sites and affect virus budding. J Virol *67*, 2546-2551.
Kalvodova, L., Sampaio, J.L., Cordo, S., Ejsing, C.S., Shevchenko, A., and Simons, K. (2009). The lipidomes of vesicular stomatitis virus, semliki forest virus, and the host plasma membrane analyzed by quantitative shotgun mass spectrometry. J Virol *83*, 7996-8003.
Kamada, M., Ando, Y., Fukunaga, Y., Kumanomido, T., Imagawa, H., Wada, R., and Akiyama, Y. (1980). Equine Getah virus infection: isolation of the virus from racehorses during an enzootic in Japan. Am J Trop Med Hyg *29*, 984-988.
Kiiver, K., Tagen, I., Žusinaite, E., Tamberg, N., Fazakerley, J.K., and Merits, A. (2008). Properties of non-structural protein 1 of Semliki Forest virus and its interference with virus replication. J Gen Virol *89*, 1457-1466.
Knight, R.L., Schultz, K.L., Kent, R.J., Venkatesan, M., and Griffin, D.E. (2009). Role of N-linked glycosylation for sindbis virus infection and replication in vertebrate and invertebrate systems. J Virol *83*, 5640-5647.
Kordyukova, L.V., Serebryakova, M.V., Baratova, L.A., and Veit, M. (2010). Site-specific attachment of palmitate or stearate to cytoplasmic versus transmembrane cysteines is a common feature of viral spike proteins. Virology *398*, 49-56.
Kostyuchenko, V.A., Jakana, J., Liu, X., Haddow, A.D., Aung, M., Weaver, S.C., Chiu, W., and Lok, S.M. (2011). The structure of barmah forest virus as revealed by cryo-electron microscopy at a 6-angstrom resolution has detailed transmembrane protein architecture and interactions. J Virol *85*, 9327-9333.
Kucukelbir, A., Sigworth, F.J., and Tagare, H.D. (2014). Quantifying the local resolution of cryo-EM density maps. Nat Methods *11*, 63-65.
Kumanomido, T., Wada, R., Kanemaru, T., Kamada, M., Hirasawa, K., and Akiyama, Y. (1988). Clinical and virological observations on swine experimentally infected with Getah virus. Vet Microbiol *16*, 295-301.
Lescar, J., Roussel, A., Wien, M.W., Navaza, J., Fuller, S.D., Wengler, G., Wengler, G., and Rey, F.A. (2001). The Fusion glycoprotein shell of Semliki Forest virus: an icosahedral assembly primed for fusogenic activation at endosomal pH. Cell *105*, 137-148.
Li, F., Li, C., Wang, M., Webb, G.I., Zhang, Y., Whisstock, J.C., and Song, J. (2015). GlycoMine: a machine learning-based approach for predicting N-, C- and O-linked glycosylation in the human proteome. Bioinformatics *31*, 1411-1419.

Li, L., Jose, J., Xiang, Y., Kuhn, R.J., and Rossmann, M.G. (2010). Structural changes of envelope proteins during alphavirus fusion. Nature *468*, 705-708.
Li, X.D., Qiu, F.X., Yang, H., Rao, Y.N., and Calisher, C.H. (1992). Isolation of Getah virus from mosquitos collected on Hainan Island, China, and results of a serosurvey. Southeast Asian J Trop Med Public Health *23*, 730-734.
Li, Y., Fu, S., Guo, X., Li, X., Li, M., Wang, L., Gao, X., Lei, W., Cao, L., Lu, Z*., et al.* (2019). Serological Survey of Getah Virus in Domestic Animals in Yunnan Province, China. Vector Borne Zoonotic Dis *19*, 59-61.
Liang, J., Guo, L., Li, K., Xiao, X., Zhu, W., Zheng, X., Hu, J., Zhang, H., Cai, J., Yu, Y*., et al.* (2018). Inhibition of the mevalonate pathway enhances cancer cell oncolysis mediated by M1 virus. Nat Commun *9*, 1524.
Liao, M., and Kielian, M. (2006). Site-directed antibodies against the stem region reveal low pH-induced conformational changes of the Semliki Forest virus fusion protein. J Virol *80*, 9599-9607.
Linder, M.E., and Deschenes, R.J. (2007). Palmitoylation: policing protein stability and traffic. Nat Rev Mol Cell Biol *8*, 74-84.
Liu, H., Zhang, X., Li, L.X., Shi, N., Sun, X.T., Liu, Q., Jin, N.Y., and Si, X.K. (2019). First isolation and characterization of Getah virus from cattle in northeastern China. BMC Vet Res *15*, 320.
Lu, Y.E., Cassese, T., and Kielian, M. (1999). The cholesterol requirement for sindbis virus entry and exit and characterization of a spike protein region involved in cholesterol dependence. J Virol *73*, 4272-4278.
Lu, Y.E., and Kielian, M. (2000). Semliki forest virus budding: assay, mechanisms, and cholesterol requirement. J Virol *74*, 7708-7719.
Lvov, D.K., Shchelkanov, M.Y., Alkhovsky, S.V., and Deryabin, P.G. (2015). Chapter 8 - Single-Stranded RNA Viruses. In Zoonotic Viruses in Northern Eurasia, D.K. Lvov, M.Y. Shchelkanov, S.V. Alkhovsky, and P.G. Deryabin, eds. (Boston: Academic Press), pp. 135-392.
Marquardt, M.T., Phalen, T., and Kielian, M. (1993). Cholesterol is required in the exit pathway of Semliki Forest virus. J Cell Biol *123*, 57-65.
McBride, C.E., and Machamer, C.E. (2010). Palmitoylation of SARS-CoV S protein is necessary for partitioning into detergent-resistant membranes and cell-cell fusion but not interaction with M protein. Virology *405*, 139-148.
Ng, C.G., Coppens, I., Govindarajan, D., Pisciotta, J., Shulaev, V., and Griffin, D.E. (2008). Effect of host cell lipid metabolism on alphavirus replication, virion morphogenesis, and infectivity. Proc Natl Acad Sci U S A *105*, 16326-16331.
Ning, W., Jiang, P., Guo, Y., Wang, C., Tan, X., Zhang, W., Peng, D., and Xue, Y. (2021). GPS-Palm: a deep learning-based graphic presentation system for the prediction of S-palmitoylation sites in proteins. Brief Bioinform *22*, 1836-1847.
Nishio, M., Umezawa, Y., Fantini, J., Weiss, M.S., and Chakrabarti, P. (2014). CH-π hydrogen bonds in biological macromolecules. Phys Chem Chem Phys *16*, 12648-12683.
Pettersen, E.F., Goddard, T.D., Huang, C.C., Couch, G.S., Greenblatt, D.M., Meng, E.C., and Ferrin, T.E. (2004). UCSF Chimera--a visualization system for exploratory research and analysis. J Comput Chem *25*, 1605-1612.
Phalen, T., and Kielian, M. (1991). Cholesterol is required for infection by Semliki Forest virus. J Cell Biol *112*, 615-623.

Rayner, J.O., Dryga, S.A., and Kamrud, K.I. (2002). Alphavirus vectors and vaccination. Rev Med Virol *12*, 279-296.
Ren, J., Wen, L., Gao, X., Jin, C., Xue, Y., and Yao, X. (2008). CSS-Palm 2.0: an updated software for palmitoylation sites prediction. Protein Eng Des Sel *21*, 639-644.
Rezza, G., Chen, R., and Weaver, S.C. (2017). O'nyong-nyong fever: a neglected mosquito-borne viral disease. Pathog Glob Health *111*, 271-275.
Ribeiro-Filho, H.V., Coimbra, L.D., Cassago, A., Rocha, R.P.F., Guerra, J., de Felicio, R., Carnieli, C.M., Leme, L., Padilha, A.C., Paes Leme, A.F.*, et al.* (2021). Cryo-EM structure of the mature and infective Mayaro virus at 4.4 Å resolution reveals features of arthritogenic alphaviruses. Nat Commun *12*, 3038.
Rohou, A., and Grigorieff, N. (2015). CTFFIND4: Fast and accurate defocus estimation from electron micrographs. J Struct Biol *192*, 216-221.
Roussel, A., Lescar, J., Vaney, M.C., Wengler, G., Wengler, G., and Rey, F.A. (2006). Structure and interactions at the viral surface of the envelope protein E1 of Semliki Forest virus. Structure *14*, 75-86.
Ryan, C., Ivanova, L., and Schlesinger, M.J. (1998). Effects of site-directed mutations of transmembrane cysteines in sindbis virus E1 and E2 glycoproteins on palmitylation and virus replication. Virology *249*, 62-67.
Schmidt, M.F., and Schlesinger, M.J. (1979). Fatty acid binding to vesicular stomatitis virus glycoprotein: a new type of post-translational modification of the viral glycoprotein. Cell *17*, 813-819.
Shabman, R.S., Rogers, K.M., and Heise, M.T. (2008). Ross River virus envelope glycans contribute to type I interferon production in myeloid dendritic cells. J Virol *82*, 12374-12383.
Sherman, M.B., and Weaver, S.C. (2010). Structure of the recombinant alphavirus Western equine encephalitis virus revealed by cryoelectron microscopy. J Virol *84*, 9775-9782.
Shi, N., Li, L.X., Lu, R.G., Yan, X.J., and Liu, H. (2019). Highly Pathogenic Swine Getah Virus in Blue Foxes, Eastern China, 2017. Emerg Infect Dis *25*, 1252-1254.
Sousa, I.P., Jr., Carvalho, C.A., Ferreira, D.F., Weissmüller, G., Rocha, G.M., Silva, J.L., and Gomes, A.M. (2011). Envelope lipid-packing as a critical factor for the biological activity and stability of alphavirus particles isolated from mammalian and mosquito cells. J Biol Chem *286*, 1730-1736.
Steentoft, C., Vakhrushev, S.Y., Joshi, H.J., Kong, Y., Vester-Christensen, M.B., Schjoldager, K.T., Lavrsen, K., Dabelsteen, S., Pedersen, N.B., Marcos-Silva, L.*, et al.* (2013). Precision mapping of the human O-GalNAc glycoproteome through SimpleCell technology. EMBO J *32*, 1478-1488.
Sugiyama, I., Shimizu, E., Nogami, S., Suzuki, K., Miura, Y., and Sentsui, H. (2009). Serological survey of arthropod-borne viruses among wild boars in Japan. J Vet Med Sci *71*, 1059-1061.
Sun, X., and Whittaker, G.R. (2003). Role for influenza virus envelope cholesterol in virus entry and infection. J Virol *77*, 12543-12551.
Tang, G., Peng, L., Baldwin, P.R., Mann, D.S., Jiang, W., Rees, I., and Ludtke, S.J. (2007). EMAN2: an extensible image processing suite for electron microscopy. J Struct Biol *157*, 38-46.
Tesh, R.B., Watts, D.M., Russell, K.L., Damodaran, C., Calampa, C., Cabezas, C., Ramirez, G., Vasquez, B., Hayes, C.G., Rossi, C.A.*, et al.* (1999). Mayaro virus disease: an emerging mosquito-borne zoonosis in tropical South America. Clin Infect Dis *28*, 67-73.

Tsetsarkin, K.A., Vanlandingham, D.L., McGee, C.E., and Higgs, S. (2007). A single mutation in chikungunya virus affects vector specificity and epidemic potential. PLoS Pathog *3*, e201.
Veit, M. (2012). Palmitoylation of virus proteins. Biol Cell *104*, 493-515.
Veit, M., Reverey, H., and Schmidt, M.F. (1996). Cytoplasmic tail length influences fatty acid selection for acylation of viral glycoproteins. Biochem J *318*, 163-172.
Voss, J.E., Vaney, M.C., Duquerroy, S., Vonrhein, C., Girard-Blanc, C., Crublet, E., Thompson, A., Bricogne, G., and Rey, F.A. (2010). Glycoprotein organization of Chikungunya virus particles revealed by X-ray crystallography. Nature *468*, 709-712.
Watanabe, Y., Bowden, T.A., Wilson, I.A., and Crispin, M. (2019). Exploitation of glycosylation in enveloped virus pathobiology. Biochim Biophys Acta Gen Subj *1863*, 1480-1497.
Wichit, S., Hamel, R., Bernard, E., Talignani, L., Diop, F., Ferraris, P., Liegeois, F., Ekchariyawat, P., Luplertlop, N., Surasombatpattana, P*., et al.* (2017). Imipramine Inhibits Chikungunya Virus Replication in Human Skin Fibroblasts through Interference with Intracellular Cholesterol Trafficking. Sci Rep *7*, 3145.
Wu, Z., Zhang, Z., Wang, X., Zhang, J., Ren, C., Li, Y., Gao, L., Liang, X., Wang, P., and Ma, C. (2021). Palmitoylation of SARS-CoV-2 S protein is essential for viral infectivity. Signal Transduct Target Ther *6*, 231.
Yang, T., Li, R., Hu, Y., Yang, L., Zhao, D., Du, L., Li, J., Ge, M., and Yu, X. (2018). An outbreak of Getah virus infection among pigs in China, 2017. Transbound Emerg Dis *65*, 632-637.
Yap, M.L., Klose, T., Urakami, A., Hasan, S.S., Akahata, W., and Rossmann, M.G. (2017). Structural studies of Chikungunya virus maturation. Proc Natl Acad Sci U S A *114*, 13703-13707.
Zhai, Y.G., Wang, H.Y., Sun, X.H., Fu, S.H., Wang, H.Q., Attoui, H., Tang, Q., and Liang, G.D. (2008). Complete sequence characterization of isolates of Getah virus (genus Alphavirus, family Togaviridae) from China. J Gen Virol *89*, 1446-1456.
Zhang, R., Hryc, C.F., Cong, Y., Liu, X., Jakana, J., Gorchakov, R., Baker, M.L., Weaver, S.C., and Chiu, W. (2011). 4.4 Å cryo-EM structure of an enveloped alphavirus Venezuelan equine encephalitis virus. EMBO J *30*, 3854-3863.
Zhao, H., Lindqvist, B., Garoff, H., von Bonsdorff, C.H., and Liljeström, P. (1994). A tyrosine-based motif in the cytoplasmic domain of the alphavirus envelope protein is essential for budding. EMBO J *13*, 4204-4211.
Zheng, S.Q., Palovcak, E., Armache, J.P., Verba, K.A., Cheng, Y., and Agard, D.A. (2017). MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy. Nat Methods *14*, 331-332.
Zhou, F., Wang, A., Chen, L., Wang, X., Cui, D., Chang, H., and Wang, C. (2020). Isolation and phylogenetic analysis of Getah virus from a commercial modified live vaccine against porcine reproductive and respiratory syndrome virus. Mol Cell Probes *53*, 101650.
Zhu, D., Wang, X., Fang, Q., Van Etten, J.L., Rossmann, M.G., Rao, Z., and Zhang, X. (2018). Pushing the resolution limit by correcting the Ewald sphere effect in single-particle Cryo-EM reconstructions. Nat Commun *9*, 1552.
Zivanov, J., Nakane, T., Forsberg, B.O., Kimanius, D., Hagen, W.J., Lindahl, E., and Scheres, S.H. (2018). New tools for automated high-resolution cryo-EM structure determination in RELION-3. Elife *7*, e42166.

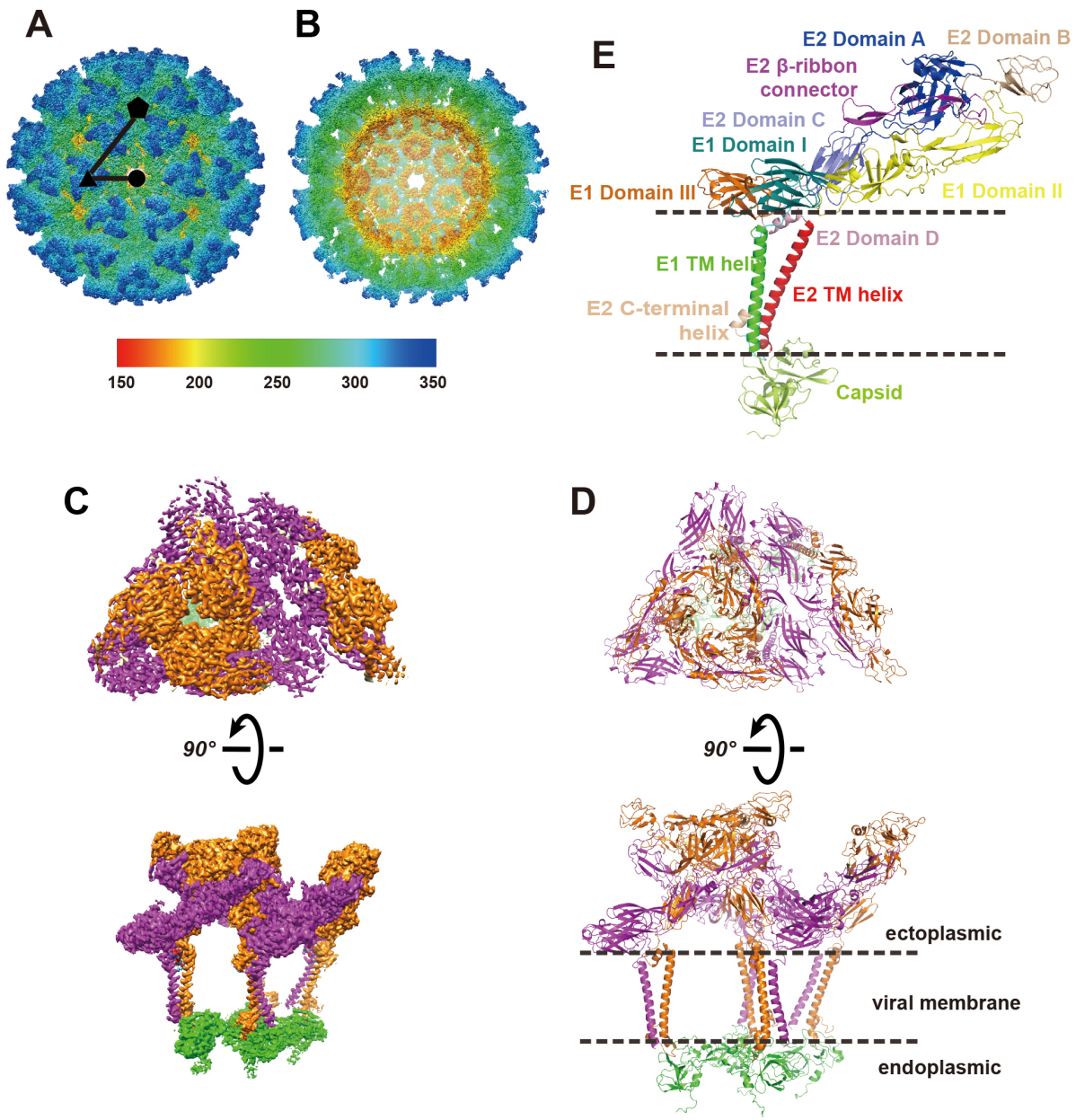

Figure 1. Overall structure of the mature and infectious GETV determined by cryo-EM at 2. 8 Å-resolution.

A, GETV virion showing the external surface with assigned symmetry axes, and **B**, central cross-section of the GETV density map. **C** and **D**, density map and atomic model of one asymmetric unit. E1, E2 and capsid proteins are colored separately by magenta, orange and green. **E**, atomic model of E1-E2-capsid heterotrimer. Different subdomains of E1 (Domain I, II, III and TM helix) and E2 (Domain A, B, C, D, TM helix and C-terminal helix) are shown, following the previous definition (Lescar et al., 2001; Voss et al., 2010).

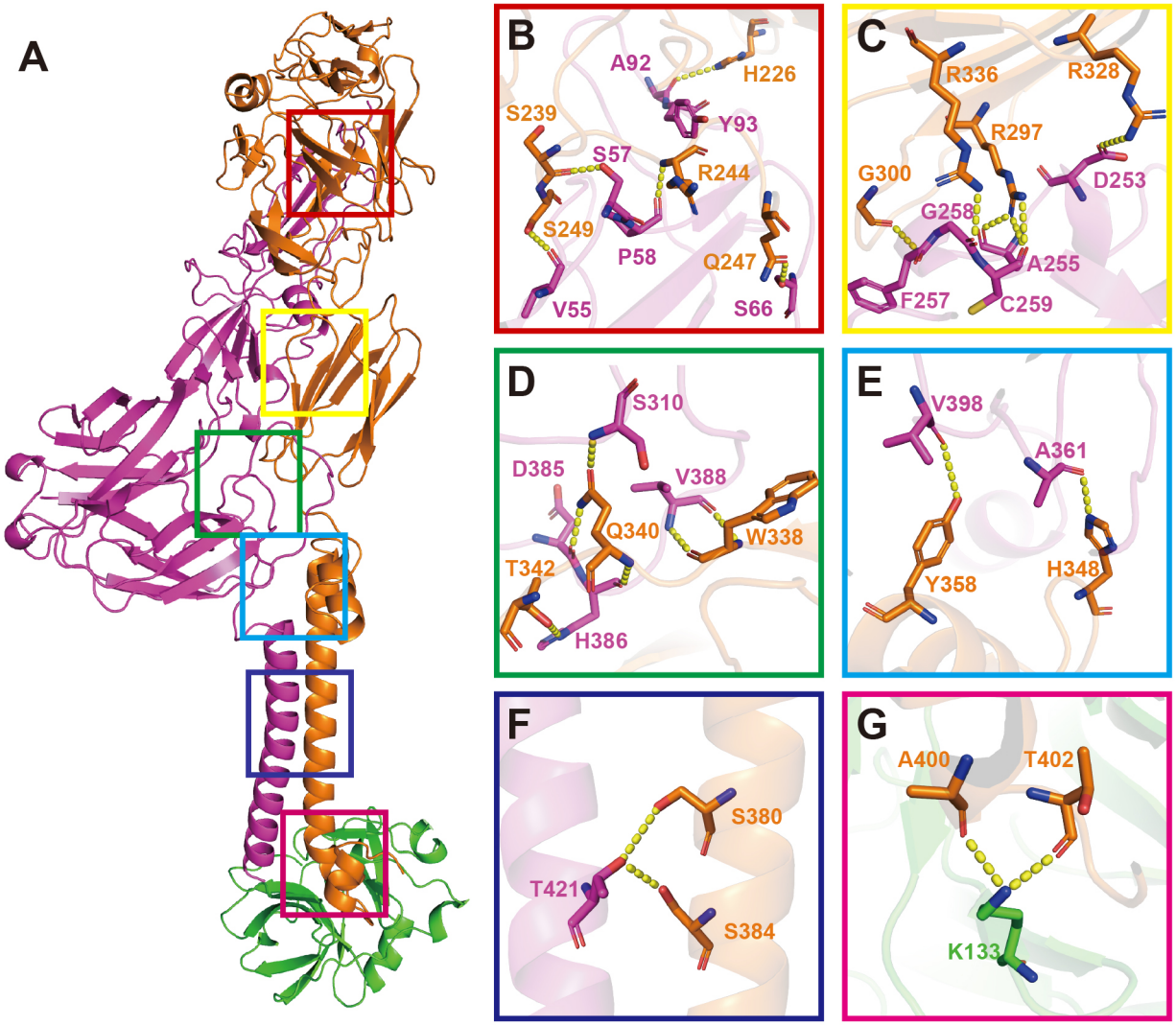

Figure 2. Interaction of GETV protein subdomains in the E1-E2-Capsid heterotrimer. **A**, Distribution of interaction regions between E1-E2 protein and E2-capsid protein. Proteins are colored with the same color in Figure 2. **B-F**, zoom views of different interaction regions in **A** between E1 and E2 protein. **G**, zoom view of interaction regions in **A** between E2 and capsid protein. The yellow dash lines indicate the distance between the atoms involved in the interaction (the cut off distance is 3.5Å). Residues names are labeled by the side of corresponding residues with the same color of proteins.

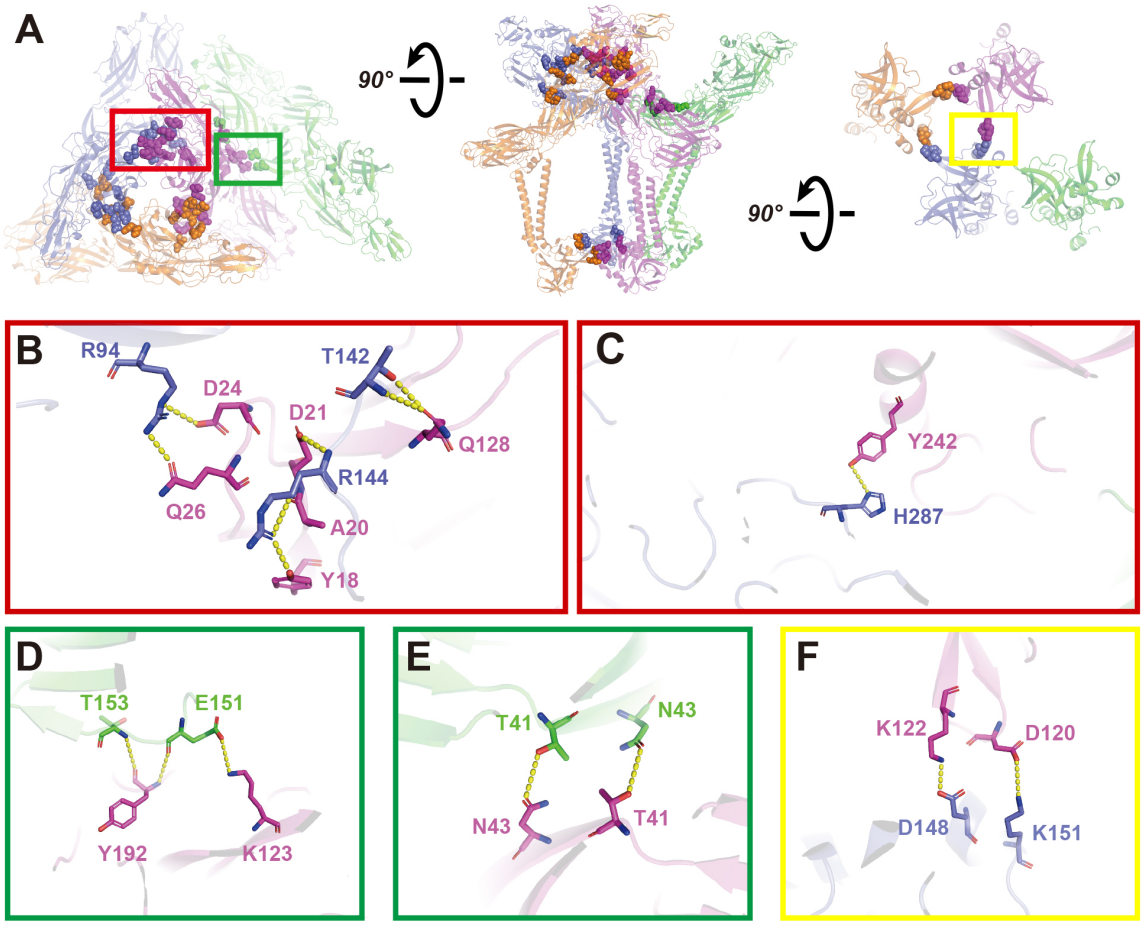

Figure 3. Interaction of GETV protein domains in the asymmetric unit.

A, The top, side and bottom view of the asymmetric unit (ASU). Four E1-E2-capsid heterotrimers in the AUS are colored slate, magenta, orange and green respectively. Residues that involved in the interaction are displayed as spheres with the same color of the corresponding heterotrimers. The slate, magenta and orange heterotrimers form the Q-trimer, and the green heterotrimer is one third of the I-trimer. **B**, zoom view of the interaction between E2 (magenta) and the neighboring E2 (slate) proteins. **C**, the interaction between E1 and the neighboring E2 proteins and **F**, the interaction between capsid (magenta) and the neighboring capsid (slate) proteins in the homotrimer formed by the E1-E2-capsid heterotrimers. **D** and **E**, zoom views of the interaction between E1 (magenta) and the neighboring E1 (green) proteins, which also represents the interaction of Q-trimer and I-trimer. The yellow dash lines indicate the distance between the atoms involved in the interaction (the cut off distance is 3.5Å). Residues names are labeled by the side of corresponding residues with the same color of proteins.

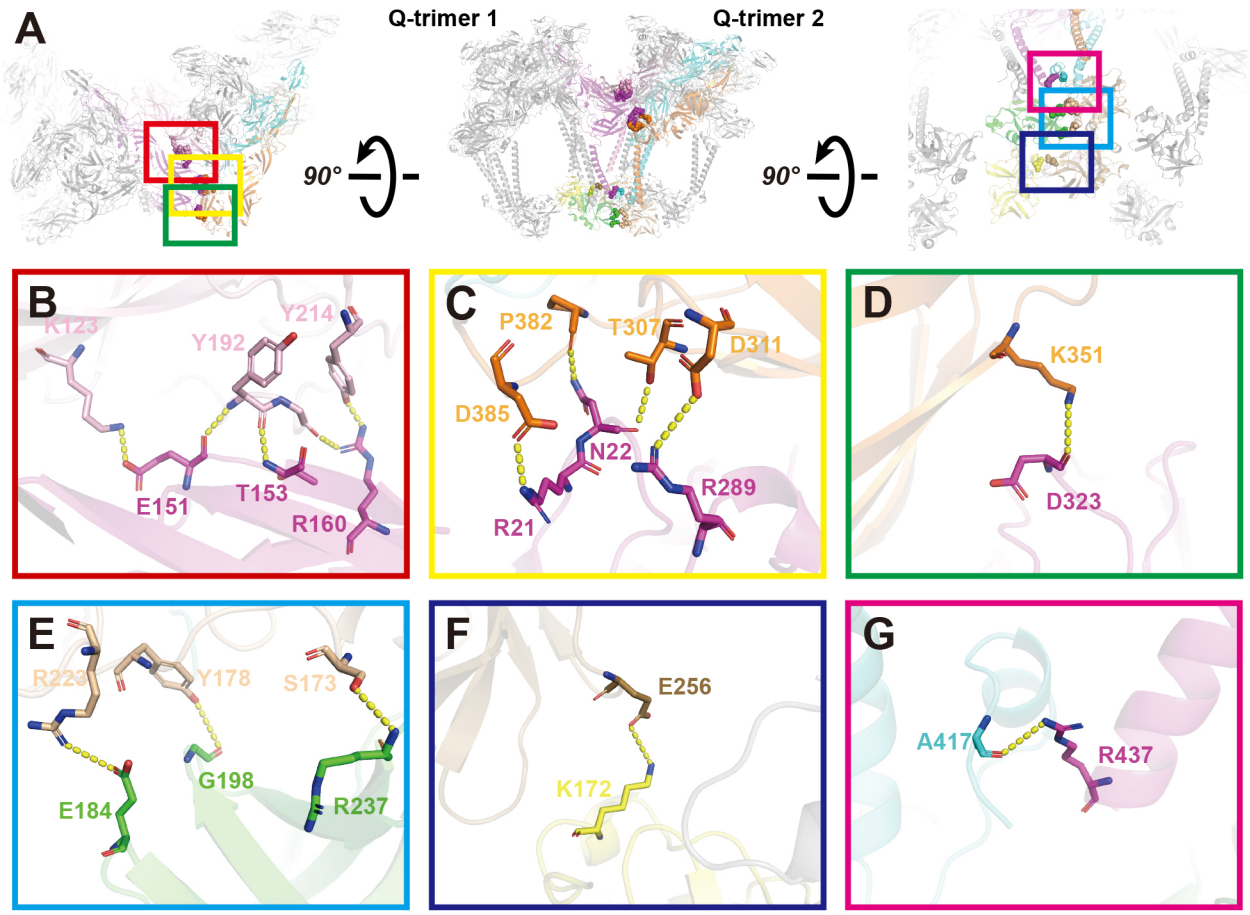

Figure 4. Interaction of GETV protein domains between two Q-trimers.

A, The top, side and bottom view of two adjacent Q-trimers. The E1, E2 and capsid proteins participating in the interaction are colored, and the others are gray. Residues that involved in the interaction are displayed as spheres with the same color of the corresponding proteins. **B**, **C** and **D**, zoom views of the interaction between E1 (magenta) of Q-trimer1 and the E1(pink) of the neighboring E1-E2-capsid heterotrimer along with another E1 (orange) of heterotrimer in Q-trimer2. **E** and **F**, Two sets of capsid-capsid interaction between the contiguous capsids of Q-trimer1 and Q-trimer2. Green and yellow capsids are in the Q-trimer1, and the other two capsids are in the Q-trimer2. **G**, interaction region between E1(magenta, the same E1 protein with **B**, **C** and **D**) and E2 (cyan, the same heterotrimer with E1 in **C** and **D**). The yellow dash lines indicate the distance between the atoms involved in the interaction. Residues names are labeled by the side of corresponding residues with the same color of proteins.

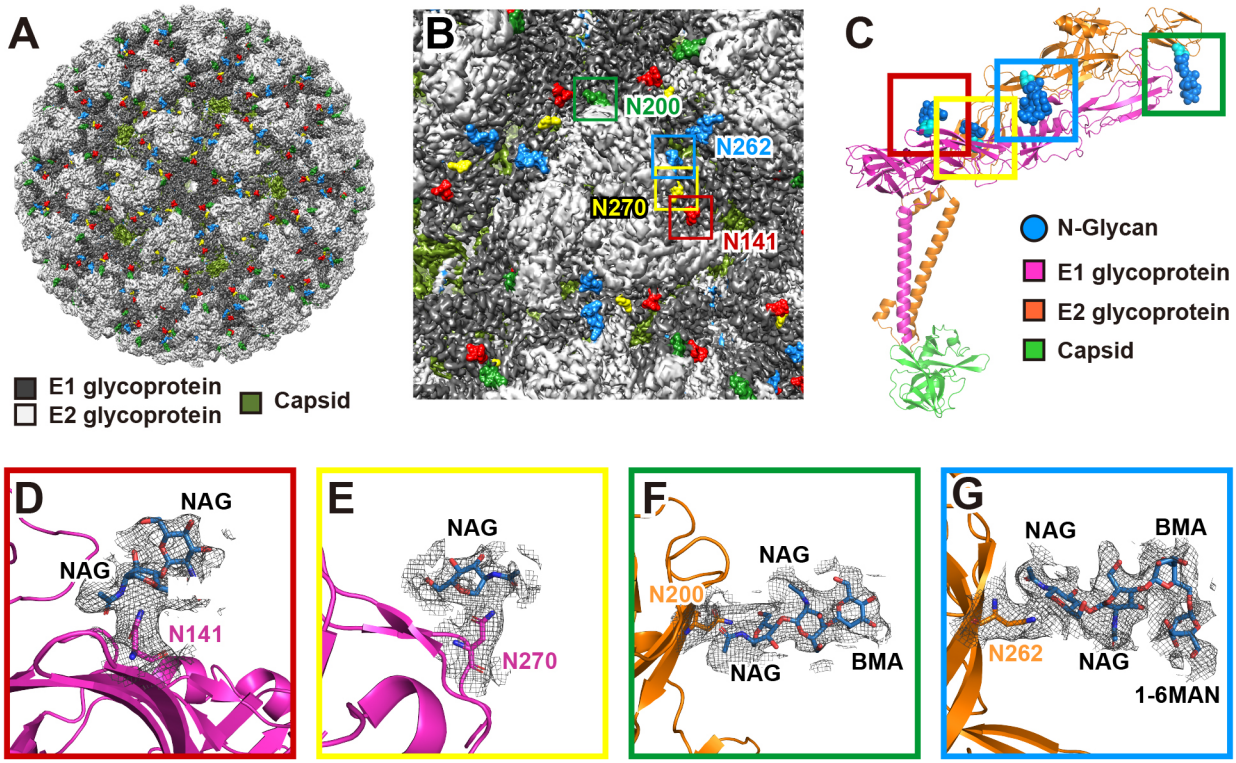

Figure 5. Glycosylation sites identified in the E1-E2 ectodomain.

A, GETV virion showing the external surface with glycosylation sites displayed in color (E1 N141 in red, E1 N270 in yellow, E2 N200 in green, and E2 N262 in blue, respectively). **B**, zoom view of external surface close to the icosahedral 3-fold vertices, glycosylation sites are framed and labeled in same color as shown in A. **C**, glycosylation sites in the atomic model of E1-E2-Capsid heterotrimer. **D** - **G**, zoom view of N-glycosylation sites with residues shown as sticks (N141 and N270 in E1 are in purple, and N200 and N262 in E2 in orange), cryo-EM densities attributed to glycans are shown in mesh, and atomic model of glycans are shown as blue sticks. Name of carbohydrate monomers are labeled in black. N-Acetyl-glucosamine (NAG) and mannose (BMA: β-D-mannose, 1-6MAN: α-(1-6)-D-mannose) were assigned as the hexosamine and hexose monomers based on cryo-EM density map interpretation and models previously built in MAYV (Ribeiro-Filho et al., 2021).

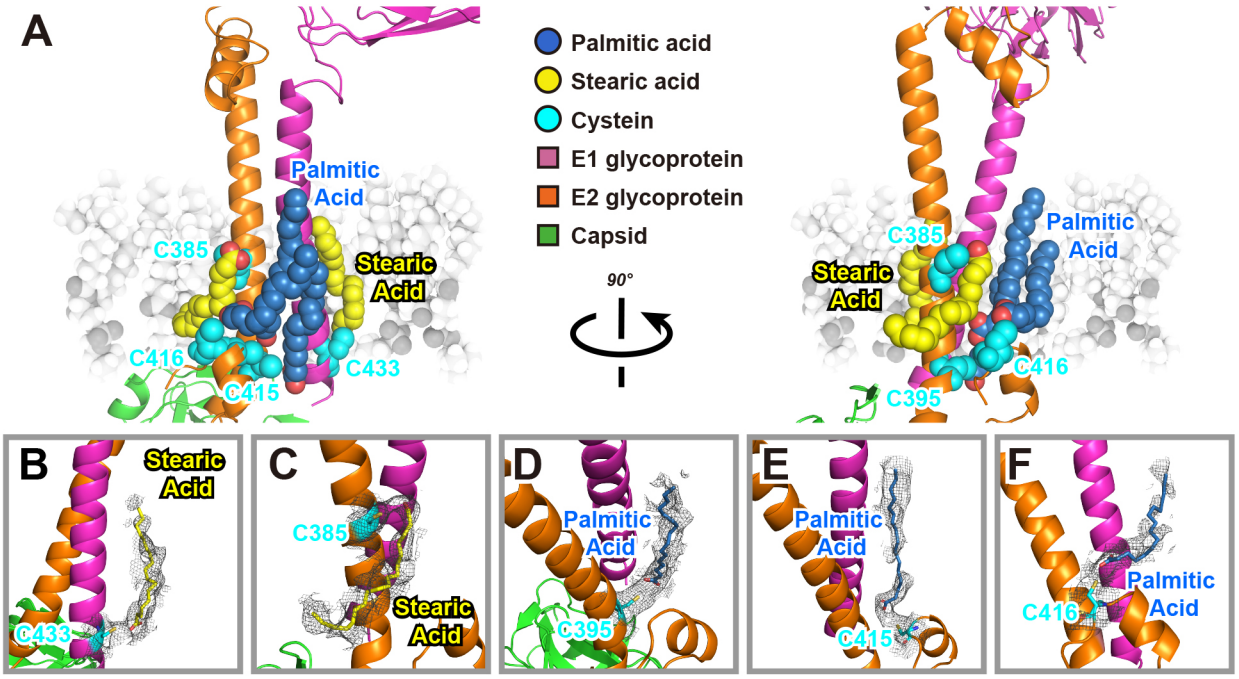

Figure 6. S-acylation sites identified in the E1 and E2 transmembrane helixes and in the E2 cytoplasmic tail.

A, Overview of S-acylation sites of E1 and E2 proteins. E1, E2 are presented as cartoon in Pymol, and colored with the same color in Figure 2. The involved cysteines, palmitic acids and stearic acids are presented as sphere, and colored with cyan, steal blue and yellow separately. The inner leaflet of virial envelop membrane are shown as sphere in the background. **B-F**, zoom views of S-acylation sites. Palmitic acids and stearic acids are shown as sticks. Cryo-EM densities attributed to the cysteines, palmitic acids and stearic acids are shown as mesh.

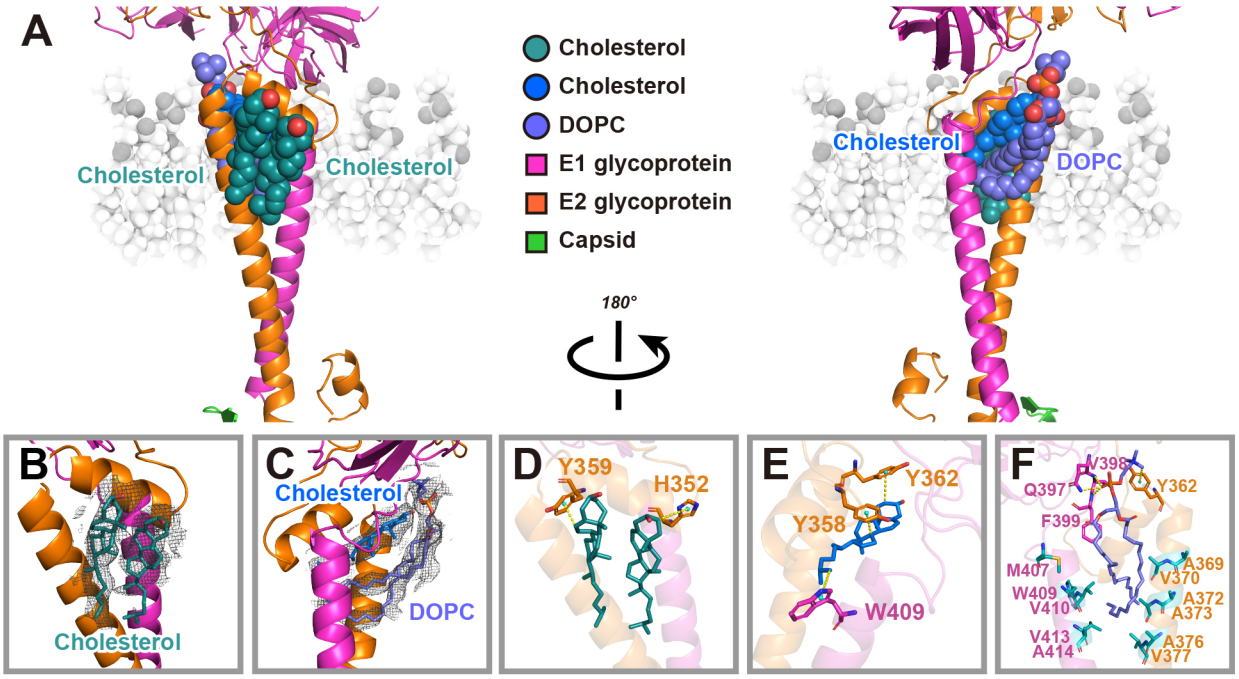

Figure 7. Cholesterol molecules and phospholipid molecule in the hydrophobic pocket.

A, Overview of the distribution of cholesterols and phospholipid molecule. A cholesterol molecule (blue) and a phospholipid molecule (slate) are in the hydrophobic pocket formed by E1 TM helix, E2 TM helix and E2 domain D. The other two cholesterol molecules (deep teal) are on the other side of domain D. E1 and E2 are presented as cartoon in Pymol. The involved cholesterols and phospholipid are presented as sphere, and outer leaflet of virial envelop membrane are shown as sphere in the background. **B** and **C**, zoom views of cholesterols and phospholipid molecules, which are shown as sticks. Cryo-EM densities attributed to the cholesterols and phospholipid are shown as mesh. **D-F**, zoom views of the interaction between cholesterol or phospholipid and E1 or E2 proteins. The aromatic residues nearing cholesterol or phospholipid are shown as sticks with the same color of E1 or E2. The hydrophobic residues nearing the tail of DOPC are shown as sticks and colored cyan. The center of mass of aromatic ring are shown as nb_spheres, colored with cyan. Yellow dashes represent the hydrogen bonds. Part of the hydrogen bond are CH-π hydrogen bonds between the soft acids CH and a soft base π-electron system. The other part are common hydrogen bonds between phosphate group of DOPC and Q397, V398 or F399.