Structures of human-infecting *Thogotovirus* fusogens support a common ancestor with insect baculovirus

Ruchao Peng[a,b], Shuijun Zhang[a,1], Yingzi Cui[a,b], Yi Shi[a,b,c,d], George F. Gao[a,b,c,d,e,2], and Jianxun Qi[a,b,2]

[a]Chinese Academy of Sciences Key Laboratory of Pathogenic Microbiology and Immunology, Institute of Microbiology, Chinese Academy of Sciences, Beijing 100101, China; [b]School of Life Sciences, University of Chinese Academy of Sciences, Beijing 101408, China; [c]Shenzhen Key Laboratory of Pathogen and Immunity, Shenzhen Third People's Hospital, Shenzhen 518112, China; [d]Center for Influenza Research and Early-Warning, Chinese Academy of Sciences, Beijing 100101, China; and [e]National Institute for Viral Disease Control and Prevention, Chinese Center for Disease Control and Prevention, Beijing 102206, China

Edited by Michael G. Rossmann, Purdue University, West Lafayette, IN, and approved September 12, 2017 (received for review April 12, 2017)

Thogotoviruses are emerging tick-borne zoonotic orthomyxoviruses infecting both humans and domestic animals with severe clinical consequences. These viruses utilize a single-envelope glycoprotein (Gp) to facilitate their entry into host cells. Here, we present the Gp structures of Thogoto and Dhori viruses, both of which are members of the *Thogotovirus* genus in the family *Orthomyxoviridae*. These structures, determined in the postfusion conformation, identified them as class III viral fusion proteins. It is intriguing that the Gp structures are similar to the envelope protein of baculovirus, although sharing a low sequence identity of ∼28%. Detailed structural and phylogenic analyses demonstrated that these Gps originated from a common ancestor. Among the structures, domain I is the most conserved region, particularly the fusion loops. Domain II showed the highest variability among different viruses, which might be related to their distinct host tropism. These findings increase our understanding of the divergent evolution processes of various orthomyxoviruses and indicate potential targets for developing antiviral therapeutics by intercepting virus entry.

Thogotovirus | *Othomyxoviridae* | crystal structure | fusion protein | evolution

The *Orthomyxoviridae* family includes seven genera—four types of influenza viruses (*Influenza virus A*, *B*, *C*, and *D*), *Thogotovirus*, *Quaranjavirus*, and *Isavirus*—all of which are segmented negative-strand RNA viruses (1, 2). Many members in this family are important animal or human pathogens, as exemplified by the epidemic or pandemic influenza A viruses that have caused hundreds of thousands of human deaths in the past century and are still discretely circulating, posing a great threat to human health and the economy.

In contrast to influenza viruses, thogotoviruses (viruses in the *Thogotovirus* genus) are transmitted mainly through tick vectors and thus are also called "tick-borne viruses" (3, 4). These viruses mainly circulate in domestic animals, such as sheep, cattle, and camels, causing neural diseases and abortion (5). Among these viruses, two main species, Thogoto virus [*Thogotovirus thogoto* (THOV)] and Dhori virus [*Thogotovirus dhori* (DHOV)], have been reported to infect humans and even cause human deaths (6, 7). Antibodies against THOV and DHOV have been identified in humans living in parts of Europe, Asia, and Africa (5, 8). Although all these viruses are transmitted mainly by tick vectors, it has been reported that DHOV caused human infections in a vector-free manner, possibly by an aerosol route (7). This further complicates its transmission route among different hosts, indicating the potential to cause epidemics among humans. In 2014, a farmer in the United States was killed by a new virus called "Bourbon virus" (BOUV), which was then identified as an unknown type of thogotovirus (9). Therefore, characterization of thogotoviruses is of great importance for public health.

There is only one glycoprotein (Gp) embedded in the envelope of thogotoviruses, which is responsible for the entire process of virus entry, including attachment, entry, and fusion (10). By bioinformatics analysis, thogotovirus Gp does not show obvious similarity with the glycoproteins of influenza viruses or isavirus, indicating a distinct mechanism for entering host cells. The closest orthomyxovirus relative of thogotovirus is quaranjavirus, whose envelope Gp shares a sequence identity of ∼26% with thogotovirus Gps (11). Thus far the receptor of thogotoviruses has not been identified, and their entry pathway is unclear. The only evidence is that the BOUV viral particles could be observed in the endosomal compartment of infected cells (9), indicating that thogotoviruses might enter host cells by endocytosis and that the membrane fusion process might occur in the endosome.

To elucidate the mechanism of thogotovirus entry and pathogenicity, we biochemically and structurally characterized the envelope Gps of THOV and DHOV. We found that both THOV and DHOV Gps undergo pH-sensitive conformational changes, which supports the hypothesis that thogotoviruses enter cells by endocytosis pathway. In addition, we determined the crystal structures of the extracellular domains of THOV and DHOV Gps in the postfusion conformation, which display features of defined class III viral fusion proteins (12). These two Gps have high structural similarities with the Gp of *Autographa californica* multiple nucleopolyhedrosis virus (*Ac*MNPV Gp64), which is a baculovirus with a double-stranded DNA genome that infects insect cells (13). It has been established that all class III viral fusion proteins are structural

Significance

Thogotoviruses belong to *Orthomyxoviridae* family and infect a variety of mammalian hosts, including humans. The emergence of these viruses poses great threats to public health and the economy. In this work, we performed structural and phylogenic analyses on the fusogenic glycoproteins of Thogoto and Dhori viruses, two representatives of the *Thogotovirus* genus that cause severe human infections. Previous studies have shown that thogotovirus glycoproteins share ∼28% sequence identity with baculovirus Gp64s. Our structural analysis confirmed their homology in evolution and identified them as class III viral fusogens, in contrast to class I members of influenza viruses. Our studies provide structural evidence to help us to understand the evolution of these viruses and indicate a potential target for antiviral drug design.

Author contributions: R.P., Y.S., G.F.G., and J.Q. designed research; R.P., S.Z., and Y.C. performed research; R.P., S.Z., Y.C., Y.S., G.F.G., and J.Q. analyzed data; and R.P., Y.S., G.F.G., and J.Q. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

Data deposition: The atomic coordinates of Thogoto virus and Dhori virus Gp extracellular domains have been deposited in the Protein Data Bank (PDB) database (PDB ID codes 5XEB and 5XEB, respectively).

[1]Present address: Program in Emerging Infectious Diseases, Duke–National University of Singapore Medical School, Singapore 169857, Singapore.

[2]To whom correspondence may be addressed. Email: gaof@im.ac.cn or jxqi@im.ac.cn.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1706125114/-/DCSupplemental.

homologs with the same origin in evolution (12–14). The structures of thogotovirus Gps and baculovirus Gp64s indicate that they diverged much later than vesicular stomatitis virus (VSV) Gp and herpes simplex virus (HSV) glycoprotein B (gB). Combined with phylogenic analysis, we found that thogotovirus Gps and baculovirus Gp64s diverged from a common origin and that thogotoviruses further evolved in two different directions, the THOV and DHOV clades. These Gps exhibit different surface electrostatic potential patterns that might result from divergent evolution. Compared with THOV Gp and *Ac*MNPV Gp64, the DHOV Gp has a more positively charged surface, indicating it is more diverged from the common ancestor. Our findings provide important clues to understand the evolution of thogotoviruses and may help in the development of antiviral therapeutics targeting the fusion proteins.

Results

Overall Structures of THOV and DHOV Gps. The ectodomains of THOV Gp (residues 18–483) and DHOV Gp (residues 21–494) were expressed using the baculovirus expression system. Initially, size-exclusion chromatography was performed to characterize the oligomerization state of THOV and DHOV Gps in solution. Judged by elution volumes, both proteins were estimated to be trimers of ∼180 kDa (Fig. S1). In addition, interchain disulfides were observed, as shown by SDS/PAGE under reducing and nonreducing conditions (Fig. S1). The crystal structures of THOV and DHOV Gps were determined at a resolution of 2.1 Å and 2.5 Å, respectively. Consistent with biochemical assays, both structures were composed of three identical chains that form an elongated bottle-shaped structure with a central helix bundle stabilizing the interchain interactions (Fig. 1). Three interchain disulfides occur at the peripheral loop regions to maintain Gps as trimers during the membrane-fusion process (Fig. 1).

The two Gp structures are very similar, with a sequence identity of 35%. Both could be divided into five domains (Fig. 1 *A*, *B*, *E*, and *F*), reminiscent of the general class III viral fusion protein structures (12, 13, 15, 16). Domain I could be further divided into subdomains Ia and Ib; subdomain Ia is located at the end near the viral envelope and contains the conserved fusion loops (Fig. 1 *B* and *F*). Domain II is a typical pleckstrin homology (PH) domain composed of antiparallel β-strands that might undergo significant conformational changes during the membrane-fusion process. Domain III is mainly composed of a helix stalk in the center of the Gp trimer, which is the main interface for interprotomer interactions (Fig. 1 *B*, *C*, *F*, and *G*). The trimerization interfaces are rich in charged residues and form extensive charged interaction networks, making the trimers sensitive to environmental pH changes (Fig. S2). There are several histidines in this interface, which might serve as pH sensors by protonization or deprotonization in response to pH changes and thus trigger conformational changes to initiate membrane fusion (17). Domain IV is located at the viral membrane-distal end and is composed of disordered loops, which are largely invisible in the density map (Fig. 1 *B*, *C*, *F*, and *G*). Domain V links the membrane-distal portion (domains III and IV) and membrane-proximal portion (domain I), further guiding the protein to stretch to the transmembrane domain. Four and two N-linked glycosylation sites can be observed in the structures of THOV and DHOV Gp, respectively. Another putative glycosylation site in both Gps is located in domain IV but was not observed in the structure due to the disorder of this region (Fig. 1 *A*, *C*, *E*, and *G*).

All these features are similar to the well-defined class III viral fusion proteins in the postfusion conformation, such as HSV gB (15), VSV Gp (16), and *Ac*MNPV Gp64 (13), indicating that thogotovirus Gps belong to class III viral fusion proteins and that the structures we obtained are in the postfusion conformation.

pH-Dependent Conformational Changes of Thogotovirus Gps. It is interesting to note that both THOV and DHOV Gp were determined in the postfusion conformation although crystallized under acidic or neutral pH conditions (see details in *SI Materials and Methods*). As members of class III viral fusion proteins, both VSV Gp and *Ac*MNPV Gp64 were reported to undergo reversible conformational changes in response to environmental pH values (13, 16, 18). This reversibility might be necessary for recovery to a high-pH conformation after passing through the acidic Golgi compartments during Gp maturation in the cell (16). As BOUV viral particles could be observed in the endosomal compartment (9), we assumed that thogotoviruses might enter cells via the endosomal pathway and that their Gps would also experience such pH-dependent structural changes during expression and infection.

By circular dichroism (CD) analysis, we observed a significant conformation shift of both THOV and DHOV Gps under different pH conditions, which was consistent with the observed pH-sensitive interfaces in Gp structures and indicated pH as the trigger that initiates member fusion (Fig. S3). Similarly, the cleaved H5 HA protein (HA1 and HA2), which is known to undergo conformational changes upon low-pH treatment, also showed different CD profiles at different pH conditions (19, 20), whereas the antibody fragment (Fab), as the negative control, showed no obvious changes in CD spectra under different pH conditions. Thus, the pH-sensitive CD profiles indicated that both THOV and DHOV Gps are expressed in the prefusion conformation but crystallized as postfusion forms. In addition, these observations also support the hypothesis that thogotoviruses might adopt the endosomal pathway to infect host cells.

Thogotovirus Gps Are Structurally Related to baculovirus Gp64s. Previous studies have reported that *Ac*MNPV Gp64 shares a low sequence identity (∼28%) with thogotovirus Gps and indicated that they are members of the same protein family (Gp64 family) (13). Our structural analysis revealed that their overall folds are very similar. The three structures could be superimposed with a certain degree of domain shifts (Fig. 2*A*). The individual domains of the three structures are highly similar, with domain Ia and domain III showing the highest similarity (Fig. 2 *B*–*D*). Notably, the putative fusion loops, which are located at the membrane-proximal end of domain Ia, are highly conserved among all proteins of the putative Gp64 family (Fig. 3 *A* and *C*). Domain IV of *Ac*MNPV Gp64 is better resolved, with more loop density observed, but basically all three structures are disordered in this region. The most striking differences among these structures are located in domain II (Fig. 2*A*). It displays the poorest alignment, as obvious domain movement was observed relative to the rest of the protein, and it is the most variable region among all proteins in Gp64 family (Figs. 2*A* and 3*A* and Table S1). When we fixed domain Ia and compared the relative orientations of the other domains, we observed a striking difference among the three structures. With THOV Gp as the reference, domain Ib and domain II of DHOV Gp and *Ac*MNPV Gp64 are rotated in opposite directions relative to domain Ia, thus leading to the poor alignment of this region (Fig. 2 *B*, *E*, and *F*). Overall, however, baculovirus Gp64 and thogotovirus Gps are closely related to each other in structure.

Distinct Electrostatic Patterns of THOV Gp, DHOV Gp, and *Ac*MNPV Gp64. It has been shown that baculovirus can enter several types of mammalian cells using Gp64 as the fusogen, and the positive charges of Gp64 favor its entry (21–23). A disordered loop connecting domain II and domain III, rich in basic amino acid residues, plays a key role in this process. Increasing the number of positively charged amino acids of this basic loop would enhance the efficiency of baculovirus entry into mammalian cells (23). Although the precise mechanism of this phenomenon is unknown, it is clear that the positive charges of Gp64 have a great impact on the ability of the virus to infect mammalian cells.

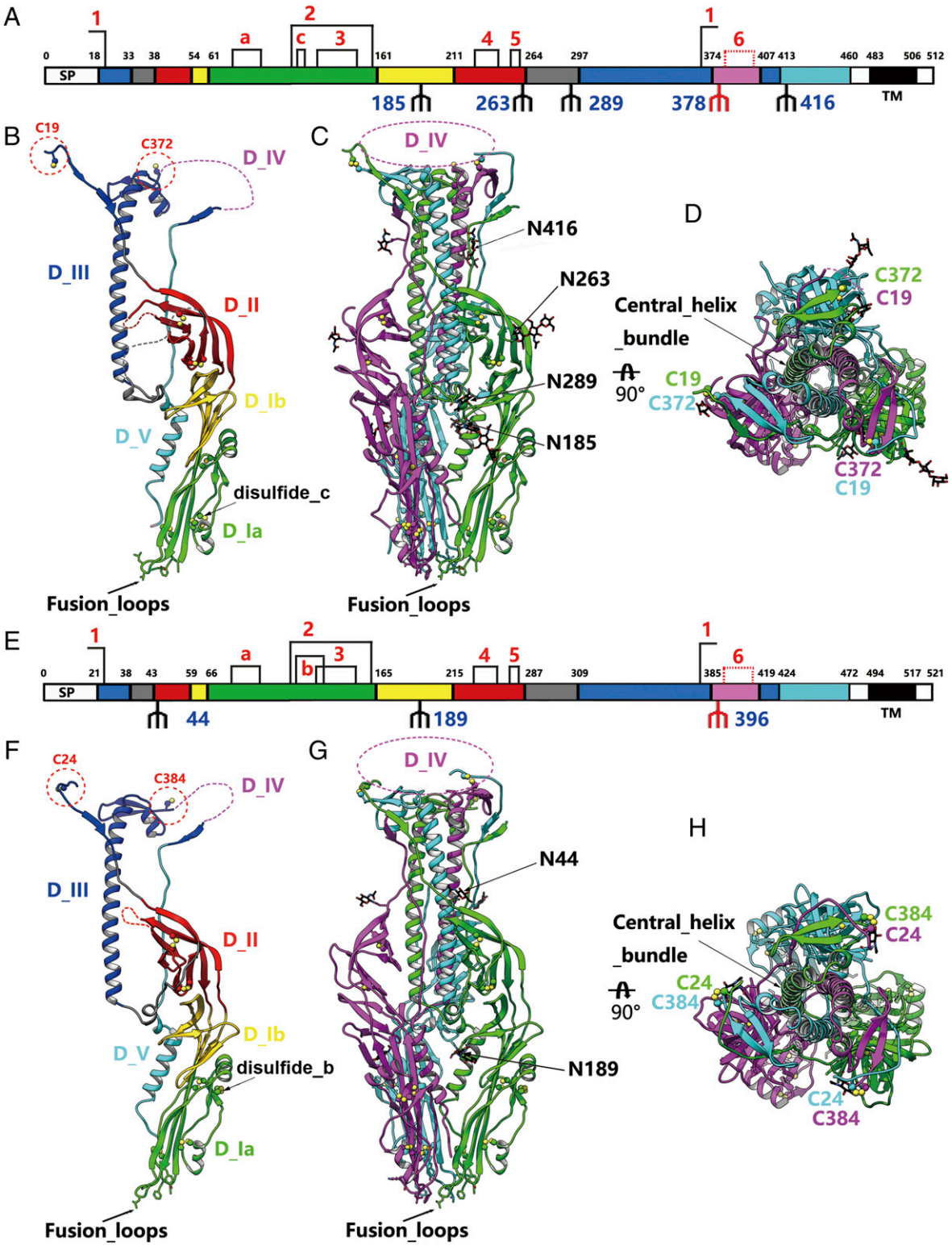

Fig. 1. Overall structure of the THOV and DHOV Gp ectodomains. (*A* and *E*) Schematic diagrams of THOV and DHOV Gp organization, colored by domains. The regions colored in gray are disordered loops not resolved in the structure. The disulfide bonds are connected by black lines. Disulfides 1–6 are conserved for all members of Gp64 protein family, but disulfide 6 is invisible in the structures. Disulfide a is shared by THOV Gp and DHOV Gp. Disulfides b and c are unique in DHOV Gp and THOV Gp, respectively. The observed glycans are shown as black branches in the diagram; red branches represent putative glycans not resolved in the structure. SP, signal peptide; TM, transmembrane region. (*B* and *F*) Ribbon representations of the THOV and DHOV Gp protomer, colored by domains as in *A* and *E*. The key residues at the putative fusion loops are shown as sticks. The unresolved disordered regions are represented by dashed lines, and disulfides are shown as spheres. Disulfides b and c are indicated by arrows. (*C*, *D*, *G*, and *H*) Two views of THOV (*C* and *D*) and DHOV (*G* and *H*) Gp trimers, colored by chains. The glycans are shown as black sticks, and the glycosylated residues are labeled. The major portion of domain IV is unresolved. The cysteines involved in interchain disulfides are labeled in the top view (*D* and *H*).

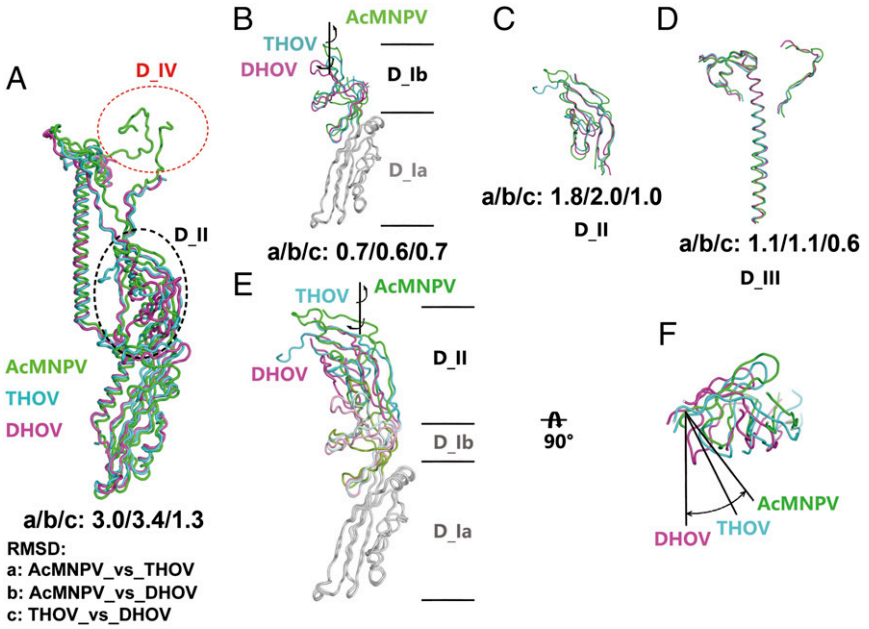

Fig. 2. Structural comparison of THOV Gp, DHOV Gp, and *Ac*MNPV Gp64. (*A*) Superimposition of THOV Gp, DHOV Gp, and *Ac*MNPV Gp64 protomers, which are colored in cyan, magenta, and green, respectively. The disordered domain IV is labeled by a red dashed ellipsoid. The poorly superimposed domain II is highlighted by a black dashed ellipsoid. (*B*) Superimposition of domain I of THOV Gp, DHOV Gp, and *Ac*MNPV Gp64 with the domain Ia fixed revealed that domain Ib shows different orientations among these structures. Domain Ia is colored in gray; domain Ib and domain II are colored as in *A*. (*C* and *D*) Superimposition of domain II and domain III of THOV Gp, DHOV Gp, and *Ac*MNPV Gp64, individually. (*E* and *F*) With domain Ia fixed, domains Ib and domain II move in different orientations relative to domain I. Domain Ia is colored in gray; domains Ib and II are colored as in *A*. The pairwise Cα rmsd values are labeled below each structure comparison figure.

The basic loop in both the *Ac*MNPV Gp64 and THOV Gp structures was largely invisible but was well resolved in the structure of DHOV Gp (Figs. 1 *A* and *E*, 3*C*, and 4 *A*–*D*). Comparing the sequences, we found that *Ac*MNPV Gp64 and THOV Gp harbor a longer basic loop than DHOV Gp (Fig. S4). As the loop of DHOV Gp is shorter, it packs much more closely to domain II and thus is less flexible. This feature also indicated that THOV Gp is much more similar to *Ac*MNPV Gp64 than to DHOV Gp in structure.

We next compared the electrostatic potential patterns of the THOV Gp, DHOV Gp, and *Ac*MNPV Gp64 structures. Interestingly, the three proteins displayed significantly different patterns of surface charge distribution (Fig. 4). On the whole, THOV and DHOV Gps have more positively charged surfaces than *Ac*MNPV Gp64, and DHOV Gp is the richest among them. Most of the positive-charge surfaces are focused on domain II, which shows the greatest variation among all proteins in Gp64 family (Figs. 3*A* and 4 *A*–*C*). Notably, a pocket formed in domain II (pocket A) is

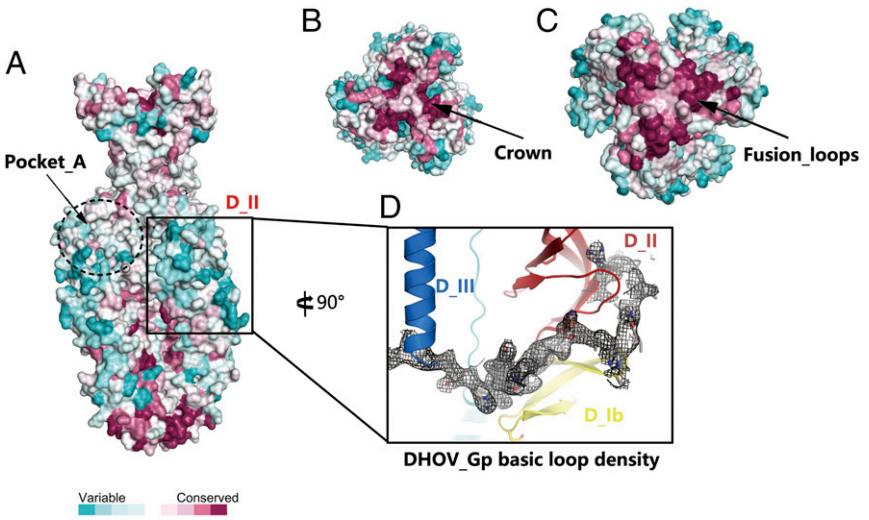

Fig. 3. Conservation analysis of thogotovirus Gps among all members of Gp64 protein family. (*A*–*C*) Side (*A*), top (*B*), and bottom (*C*) views of the conservation map. Conserved regions are colored in purple, and variable regions are cyan. (*D*) Close-up view of the well-resolved basic loop density of DHOV Gp. The $2F_o$-F_c electron density map at a contour level of 1.0 σ is shown as a mesh, and the loop residues are represented as sticks. The ribbon diagrams are colored by domains as in Fig. 2*B*.

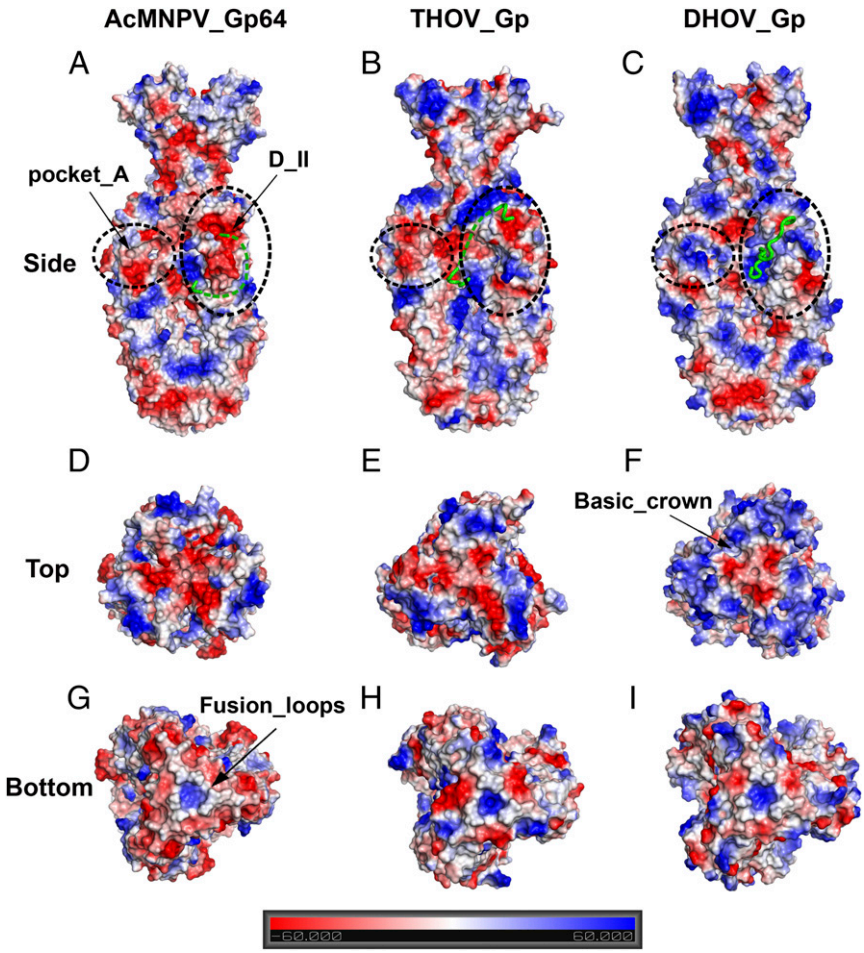

Fig. 4. Electrostatic potential maps of *Ac*MNPV Gp64 (*A*, *D*, and *G*), THOV Gp (*B*, *E*, and *H*), and DHOV Gp (*C*, *F*, *I*) in side (*A*–*C*), top (*D*–*F*), and bottom (*G*–*I*) views. The surface potential is shown from −60 KT/e in red to +60 KT/e in blue. The most variable domain II and a surface pocket (pocket_A) are highlighted by black dashed ellipsoids. The basic loops are shown as green ribbons with dashed lines representing the unresolved parts.

mainly negatively charged in *Ac*MNPV Gp64 and THOV Gp, while it is highly positively charged in DHOV Gp (Fig. 4 *A*–*C*).

Accordingly, the conserved domain I shows a similar charge pattern among all three structures, while the top region of the domain III stalk (crown) also displays an increase in positive charge in the THOV and DHOV Gps compared with *Ac*MNPV Gp64 (Fig. 4 *D*–*I*). These structural features indicate that *Ac*MNPV Gp64 and DHOV Gp likely represent two evolutionary directions in favor of different hosts. In addition, the evolving route of THOV Gp is different from that of DHOV Gp, as it showed some obvious features similar to *Ac*MNPV Gp64 rather than DHOV Gp.

Phylogenic Analysis of Thogotovirus Gps. With the hypotheses generated by structural analysis, we then compared all the gene sequences of thogotovirus Gp homologs in the National Center for Biotechnology Information (NCBI) database, which involved all available thogotovirus and quaranjavirus Gp sequences. A phylogenic tree of Gps was built with the sequences of baculovirus Gp64s as the outgroup. All sequences clustered into three main clades, which are represented by THOV (clade A), DHOV (clade B), and Quaranfil virus (clade C, representing the *Quaranjavirus* genus) (Fig. 5). Among orthomyxovirus homologs, thogotoviruses diverged early from the quaranjavirus clade and later diverged into the THOV and DHOV clades.

Three unclassified thogotoviruses, Jos virus, Upolu virus, and Aransas virus, fell into clade A, indicating much closer relation to THOV. It is interesting that Sinu virus, an unclassified orthomyxovirus newly isolated from mosquitoes (24), clustered much more closely to baculoviruses than to the other three orthomyxovirus clades. The newly identified BOUV clustered with DHOV. The Gps of DHOV and BOUV share a sequence identity of ~57%, and the basic loops are very similar to each other, i.e., shorter than that of THOV Gp and *Ac*MNPV Gp64 (Fig. S4).

Although the orthomyxovirus homologs diverged from baculovirus Gp64 before they diverged from each other, the structure of THOV Gp showed some features more similar to baculovirus Gp64 than to DHOV Gp. This indicates that the evolutionary route of the THOV clade enabled these Gps to retain certain properties of the baculovirus clade. The DHOV clade, however, evolved via a route that was substantially distinct from that of the THOV and baculovirus clades. To further test the consistency of structural features with their phylogenic relationships, we conducted structural predictions for the representative Gps of each clade. In line with this hypothesis, the predicted structure models of BOUV, Quaranfil virus, and Sinu virus Gps displayed a similar trend of evolving electrostatic potential patterns (Fig. S5). The BOUV Gp, closest to DHOV Gp, shows a dominant positively charged surface profile. In contrast, Sinu virus Gp is rich in negative charges, similar to the pattern of *Ac*MNPV Gp64. The Gp of Quaranfil virus, the type species of the *Quaranjavirus* genus, maintains some features of both BOUV and Sinu virus Gps and harbors moderate positively charged surfaces (Fig. S5).

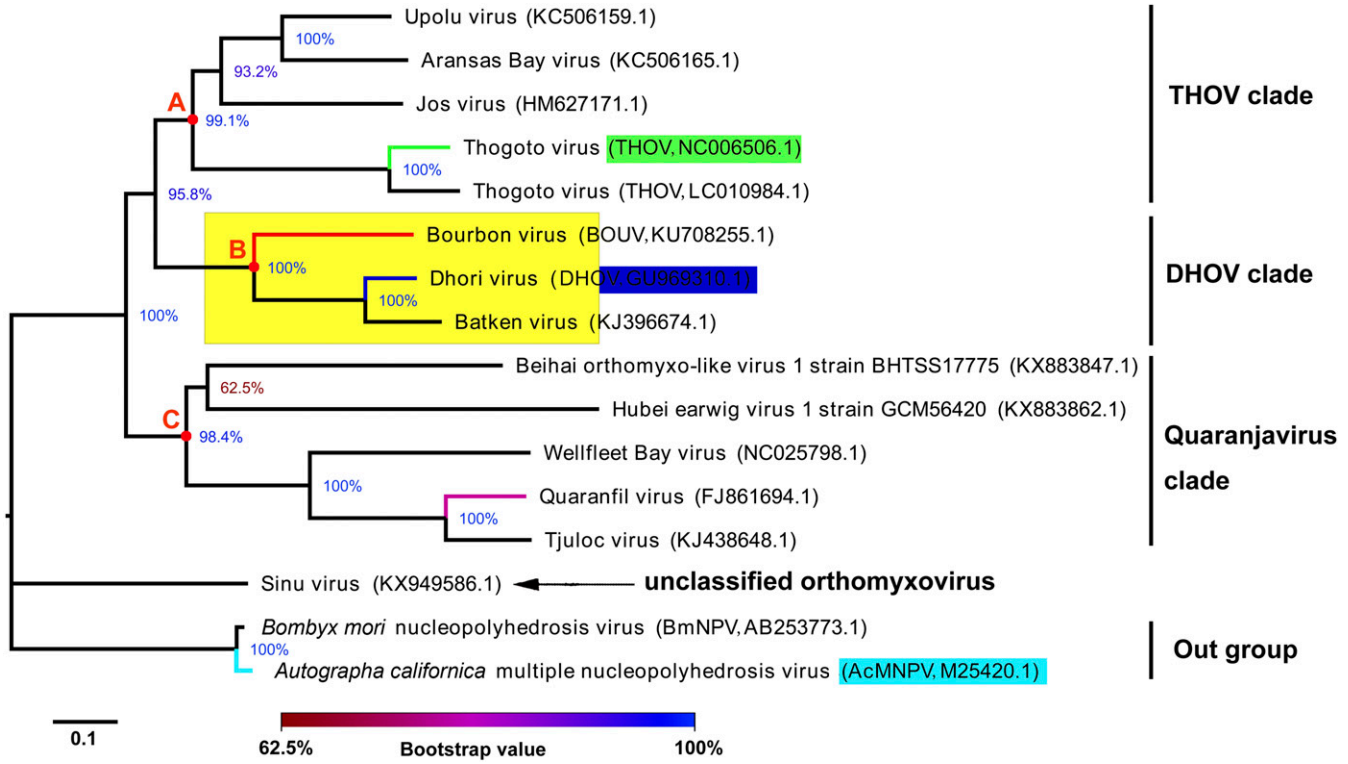

Fig. 5. Phylogenic tree of thogotovirus Gps and homologs. Two baculovirus Gp64 sequences were selected as the outgroup. Bootstrap values are shown at relevant nodes. Each taxon is labeled with the species name and the GenBank accession number. Evolutionary distances are expressed as the number of base substitutions per site. The three clade nodes A, B, and C correspond to the THOV clade, DHOV clade, and quaranjavirus clade, respectively. The DHOV clade is highlighted by a yellow rectangle. The taxons for the structures discussed in this study are highlighted in cyan (*Ac*MNPV), green (THOV), and blue (DHOV. The newly identified BOUV is colored red. Quaranfil virus, the type species of the *Quaranjavirus* genus, is magenta.

Discussion

Among the seven genera in the *Othomyxoviridae* family, five (*Influenza virus A*, *B*, *C*, and *D* and *Isavirus*) possess class I fusogenic proteins to mediate membrane fusion and utilize sialic acid-related receptors for attachment to the host cell surface and subsequent entry (25–30). Meanwhile, they also harbor receptor-destroying enzymes to facilitate the release of progeny virions (31, 32). Precisely, influenza *A* and *B* viruses utilize HA proteins to bind to α2–3–, α2–6–, or α2–8–linked sialic acid receptors and mediate membrane fusion in the endosome, and they use neuraminidase proteins for receptor cleavage (25). For influenza C and *D* viruses, all three processes (i.e., attachment, fusion, and release) are mediated by a single protein, hemagglutinin-esterase-fusion (HEF) protein, and they recognize 9-O-acetylated sialic acid receptors (28–30, 33). In isaviruses, however, two envelope proteins are involved, the HE (hemagglutinin-esterase) and F (fusion) proteins, which are responsible for binding and destroying 4-O-acetylated sialic acid receptors and mediating membrane fusion, respectively (26, 27, 32). It is intriguing that thogotoviruses, as a genus in the *Othomyxoviridae* family, utilize a single-envelope Gp for entry into host cells, which is strikingly different from all the proteins mentioned above. To date, the receptors governing thogotoviruses' entry have not been identified. Previous studies have shown that Gp can induce hemagglutination but lacks neuraminidase or esterase activities (10). All these unique features demonstrate that thogotoviruses represent a type of virus significantly distinct from other orthomyxoviruses.

Interestingly, our current study indicates that the closest homolog of thogotovirus Gps is *Ac*MNPV Gp64, a class III viral fusogen (13). Although sharing a low sequence identity (~28%) (Fig. S3), the overall folds of THOV/DHOV Gps and *Ac*MNPV Gp64 are quite similar. Their individual domains can be superimposed well, but their relative orientations vary among different species (Fig. 2 *A*–*F*). Because of this variation, the electrostatic potential patterns of these proteins also show obvious diversity. The most variable domain, domain II, is obviously the result of the divergent evolution of the different viruses, which might be related to their distinctions in host tropism.

Baculoviruses can enter several mammalian cells using Gp64 as the fusion machine but cannot replicate (21, 22), which is reminiscent of the bitropism of thogotoviruses in mammalian hosts and tick vectors. Our data further provide strong support for the functional correlation between thogotovirus Gps and baculovirus Gp64s. Our structural and phylogenic analyses indicated that the *Ac*MNPV Gp64 and DHOV Gp represent two evolving directions that favor insect and mammal hosts, respectively. Although THOV is closer to DHOV than to baculovirus, its Gp evolved in a direction different from that of DHOV and retains obvious similarities to baculovirus Gp64.

Among all orthomyxoviruses using thogotovirus Gp homologs for cell entry, only members of the *Thogotovirus* genus were reported to infect humans (5–7, 9). In particular, DHOV is reported to be capable of transmitting through an aerosol route and causing severe infections (7), implying the potential to cause human epidemics. Notably, the newly identified BOUV clustered within the DHOV clade. These observations indicate that the DHOV clade likely represents a group of well-adapted and dangerous mammalian-tropic viruses that pose great threats to human health. Thus, special care should be taken for the surveillance of potential epidemics of DHOV clade viruses in humans. Our study here indicates that structural information could also provide important clues for early warnings of potential human infections of emerging pathogens.

Although thogotovirus Gps and baculovirus Gp64s are clearly of the same origin, it is difficult to tell which is more similar to the common ancestor, i.e., whether the Gp64 of baculoviruses evolved from a thogotovirus-like ancestor or the Gp of thogotoviruses originated from a baculovirus-like prototype. We noticed that the group I baculoviruses using Gp64 as the fusogen appear later in evolution than those in group II utilizing an F protein. There is only one Gp, F protein, in the envelope of group II baculoviruses that facilitates their entry (34). Group I baculoviruses, however, have two envelop Gps, Gp64 and a vestigial F homolog that is incapable of mediating virus infection (34, 35). Thus, group I baculoviruses evolved from a group II-like ancestor that obtained the prototype of Gp64, which gradually replaced the function of F protein. Because thogotoviruses have only one envelope Gp, the most likely scenario might be that an ancient group II-like baculovirus obtained a Gp64 prototype from a thogotovirus-like ancestor and evolved into group I baculoviruses. The Gp64 prototype is conceivably quite similar to THOV Gp, which possesses the features of both baculovirus Gp64 and DHOV Gp.

The structural and phylogenic analyses in this study provide important insight into the evolution of viruses from different families: *Orthomyxoviridae* (segmented negative-strand RNA viruses) and *Baculoviridae* (double-stranded DNA viruses). Our findings are of great significance for understanding the origin and evolving process of orthomyxoviruses, many of which, including some of the thogotoviruses discussed above, are multitropism pathogens that infect both humans and animals.

In addition, the two Gp structures could potentially offer important guidance for antiviral drug design. Based on our analysis, the fusion loops of Gps are highly conserved among all Gp64 family members. By analogy to VSV Gp, domain I, where the fusion loop is located, would retain the same fold before and after membrane fusion (16, 18). Therefore, it would be a suitable target for the development of broad-spectrum antiviral drugs that block virus entry by interfering with the membrane-fusion process.

Despite these advances, some important questions remain to be answered to understand the mechanism of thogotovirus entry. It is noteworthy that all three structures were determined in the postfusion conformation, although they were crystallized under different pH conditions varying from 3.5 to 7.5 (see details in *SI Materials and Methods*). Both *Ac*MNPV Gp64 and DHOV Gp were crystallized at acidic pHs, 5.0 and 3.5, respectively (13). However, THOV Gp crystals were obtained at nearly neutral conditions, pH 7.0–7.5, where Gps are thought to exist in the prefusion conformation. As *Ac*MNPV Gp64 undergoes a pH-induced conformational change (13) and THOV and DHOV Gps also displayed significant conformational shifts at different pH conditions (Fig. S2), we infer that the prefusion structures of thogotovirus Gps are metastable and unfavorable for crystallization. In neutral pH solutions, Gp can also adopt a postfusion conformation, but at a low rate, thereby offering the opportunity for crystallization by displacing the equilibrium between different conformations. A similar phenomenon was also observed for the Gp of VSV, which could be crystallized in the postfusion conformation at pH 6.0–7.0 (16, 18). Thus far, the receptors for both baculoviruses and thogotoviruses are unknown. To fully uncover the mechanism of thogotovirus entry, more efforts will be required to resolve the Gp structures at the prefusion state and to identify their potential receptors.

Materials and Methods

Protein Expression, Purification, and Crystallization. The ectodomains of THOV Gp (GenBank accession code: NC006506, residues 18–483) and DHOV Gp (GenBank accession code: GU969310, residues 21–494) fused with a C-terminal T4 fibritin trimerization domain and a 6× histidine affinity tag were expressed using the Bac-to-Bac baculovirus expression system (*SI Materials and Methods*). The soluble Gp ectodomains were purified by tandem immobilized metal affinity chromatography (IMAC) and size-exclusion chromatography (*SI Materials and Methods*). The eluted products in the buffer containing 20 mM Tris·HCl (pH 8.0) and 150 mM NaCl were concentrated to ∼10 mg/mL for crystallization. Both THOV and DHOV Gps were crystallized by the sitting-drop vapor-diffusion method at 18 °C (*SI Materials and Methods*).

Diffraction Data Collection and Structure Determination. Diffraction data were collected with cryoprotected [in a reservoir solution containing 20% (vol/vol) glycerol] crystals at the Shanghai Synchrotron Radiation Facility (SSRF) beamline BL17U (36). All datasets were processed with HKL2000 software (37). The THOV and DHOV Gps datasets were refined to 2.1 and 2.5 Å, respectively. Both structures were solved by the molecular-replacement method (38). The initial phase of THOV Gp was obtained using the N-terminal 160-amino acid partial model of baculovirus Gp64 [Protein Data Bank (PDB) ID: 3DUZ] as the search input. Density modification (39) was performed with noncrystallographic symmetry restraints to improve the density of other parts. Iterative autobuilding and manual adjustments were conducted to complete the model (*SI Materials and Methods*). The DHOV Gp structure was solved with the THOV Gp structure as the search input. The final structures of the THOV and DHOV Gps were refined to an R_{free} of 20.9% and 25.7%, respectively (*SI Materials and Methods*). Data collection, processing, and refinement statistics are summarized in Table S2. All structural figures were generated with either Chimera (40) or PyMOL (41).

Phylogenic Analysis. The Gp coding sequences of thogotoviruses and quaranjaviruses (all annotated and unclassified species in the two genera available in the NCBI database) were analyzed. Two sequences of baculovirus Gp64, *Ac*MNPV and *Bombyx mori* nucleopolyhedrosis virus (*Bm*NPV), were selected as the outgroup. The sequences were aligned using MUSCLE (42), and the phylogenetic tree was inferred using the Minimum Evolution (ME) method (*SI Materials and Methods*) (43). The resulting phylogenetic trees were displayed and annotated with FigTree.

Data Availability. The atomic coordinates of the THOV and DHOV Gp extracellular domains have been deposited in the PDB under the ID codes 5XEA and 5XEB, respectively.

ACKNOWLEDGMENTS. We thank Haixia Xiao for excellent suggestions regarding this study and comments on the manuscript and the staff at the SSRF beamline 17U for assistance in diffraction data collection. This work was supported by Strategic Priority Research Program of the Chinese Academy of Sciences Grant XDPB03, China Ministry of Science and Technology National 973 Project Grants 2013CB531502 and 2014CB542503, National Key Plan for Scientific Research and Development of China Grant 2016YFD0500305, and China National Science and Technology Special Project Grant 2017ZX10303403 of the National Natural Science Foundation of China (NSFC). Y.S. is supported by NSFC Excellent Young Scientist Program Grant 81622031 and Youth Innovation Promotion Association of the Chinese Academy of Sciences Grant 2015078. G.F.G. is partly supported as a leading principal investigator of the NSFC Innovative Research Group supported by NSFC Grant 81621091.

1. Kawaoka Y, Palese P (2005) Family *Orthomyxoviridae*. *Virus Taxonomy: Eighth Report of the International Committee on Taxonomy of Viruses*, eds Fauquet CM, Mayo MA, Maniloff J, Desselberger U, Ball L (Elsevier Academic, San Diego), pp 681–693.
2. Perez D, et al. (2011) Orthomyxoviridae. *Virus Taxonomy*, eds King AMQ, Adams MJ, Carstens EB, Lefkowitz EJ (Elsevier, Oxford), pp 749–762.
3. Haig DA, Woodall JP, Danskin D (1965) Thogoto virus: A hitherto underscribed agent isolated from ticks in Kenya. *J Gen Microbiol* 38:389–394.
4. Anderson CR, Casals J (1973) Dhori virus, a new agent isolated from *Hyalomma dromedarii* in India. *Indian J Med Res* 61:1416–1420.
5. Hubálek Z, Rudolf I (2012) Tick-borne viruses in Europe. *Parasitol Res* 111:9–36.
6. Moore DL, et al. (1975) Arthropod-borne viral infections of man in Nigeria, 1964-1970. *Ann Trop Med Parasitol* 69:49–64.
7. Butenko AM, Leshchinskaia EV, Semashko IV, Donets MA, Mart'ianova LI (1987) [Dhori virus–A causative agent of human disease. 5 cases of laboratory infection]. *Vopr Virusol* 32:724–729. Russian.
8. Filipe AR, Calisher CH, Lazuick J (1985) Antibodies to Congo-Crimean haemorrhagic fever, Dhori, Thogoto and Bhanja viruses in southern Portugal. *Acta Virol* 29: 324–328.
9. Kosoy OI, et al. (2015) Novel thogotovirus associated with febrile illness and death, United States, 2014. *Emerg Infect Dis* 21:760–764.
10. Portela A, Jones LD, Nuttall P (1992) Identification of viral structural polypeptides of Thogoto virus (a tick-borne orthomyxo-like virus) and functions associated with the glycoprotein. *J Gen Virol* 73:2823–2830.
11. Presti RM, et al. (2009) Quaranfil, Johnston Atoll, and Lake Chad viruses are novel members of the family *Orthomyxoviridae*. *J Virol* 83:11599–11606.

12. Backovic M, Jardetzky TS (2009) Class III viral membrane fusion proteins. *Curr Opin Struct Biol* 19:189–196.
13. Kadlec J, Loureiro S, Abrescia NG, Stuart DI, Jones IM (2008) The postfusion structure of baculovirus gp64 supports a unified view of viral fusion machines. *Nat Struct Mol Biol* 15:1024–1030.
14. Harrison SC (2008) Viral membrane fusion. *Nat Struct Mol Biol* 15:690–698.
15. Heldwein EE, et al. (2006) Crystal structure of glycoprotein B from herpes simplex virus 1. *Science* 313:217–220.
16. Roche S, Bressanelli S, Rey FA, Gaudin Y (2006) Crystal structure of the low-pH form of the vesicular stomatitis virus glycoprotein G. *Science* 313:187–191.
17. Kampmann T, Mueller DS, Mark AE, Young PR, Kobe B (2006) The role of histidine residues in low-pH-mediated viral membrane fusion. *Structure* 14:1481–1487.
18. Roche S, Rey FA, Gaudin Y, Bressanelli S (2007) Structure of the prefusion form of the vesicular stomatitis virus glycoprotein G. *Science* 315:843–848.
19. Lu X, et al. (2013) Structure and receptor-binding properties of an airborne transmissible avian influenza A virus hemagglutinin H5 (VN1203mut). *Protein Cell* 4: 502–511.
20. Zhang W, et al. (2013) An airborne transmissible avian influenza H5 hemagglutinin seen at the atomic level. *Science* 340:1463–1467.
21. Ho YC, et al. (2006) Baculovirus transduction of human mesenchymal stem cell-derived progenitor cells: Variation of transgene expression with cellular differentiation states. *Gene Ther* 13:1471–1479.
22. Zeng J, Du J, Zhao Y, Palanisamy N, Wang S (2007) Baculoviral vector-mediated transient and stable transgene expression in human embryonic stem cells. *Stem Cells* 25:1055–1061.
23. O'Flynn NM, Patel A, Kadlec J, Jones IM (2012) Improving promiscuous mammalian cell entry by the baculovirus *Autographa californica* multiple nuclear polyhedrosis virus. *Biosci Rep* 33:23–36.
24. Contreras-Gutiérrez MA, et al. (2017) Sinu virus, a novel and divergent orthomyxovirus related to members of the genus *Thogotovirus* isolated from mosquitoes in Colombia. *Virology* 501:166–175.
25. Bouvier NM, Palese P (2008) The biology of influenza viruses. *Vaccine* 26:D49–D53.
26. Aspehaug V, Mikalsen AB, Snow M, Biering E, Villoing S (2005) Characterization of the infectious salmon anemia virus fusion protein. *J Virol* 79:12544–12553.
27. Rimstad E, Mjaaland S, Snow M, Mikalsen AB, Cunningham CO (2001) Characterization of the infectious salmon anemia virus genomic segment that encodes the putative hemagglutinin. *J Virol* 75:5352–5356.
28. Muraki Y, Hongo S (2010) The molecular virology and reverse genetics of influenza C virus. *Jpn J Infect Dis* 63:157–165.
29. Song H, et al. (2016) An open receptor-binding cavity of hemagglutinin-esterase-fusion glycoprotein from newly-identified influenza D virus: Basis for its broad cell tropism. *PLoS Pathog* 12:e1005411.
30. Hause BM, et al. (2014) Characterization of a novel influenza virus in cattle and swine: Proposal for a new genus in the *Orthomyxoviridae* family. *MBio* 5:e00031-14.
31. Cohen M, et al. (2013) Influenza A penetrates host mucus by cleaving sialic acids with neuraminidase. *Virol J* 10:321.
32. Kristiansen M, Frøystad MK, Rishovd AL, Gjøen T (2002) Characterization of the receptor-destroying enzyme activity from infectious salmon anaemia virus. *J Gen Virol* 83:2693–2697.
33. Muchmore EA, Varki A (1987) Selective inactivation of influenza C esterase: A probe for detecting 9-O-acetylated sialic acids. *Science* 236:1293–1295.
34. Lung O, Westenberg M, Vlak JM, Zuidema D, Blissard GW (2002) Pseudotyping *Autographa californica* multicapsid nucleopolyhedrovirus (*Ac*MNPV): F proteins from group II NPVs are functionally analogous to *Ac*MNPV GP64. *J Virol* 76:5729–5736.
35. Lung OY, Cruz-Alvarez M, Blissard GW (2003) Ac23, an envelope fusion protein homolog in the baculovirus *Autographa californica* multicapsid nucleopolyhedrovirus, is a viral pathogenicity factor. *J Virol* 77:328–339.
36. Wang Q, et al. (2015) The macromolecular crystallography beamline of SSRF. *Nucl Sci Tech* 26:10102.
37. Otwinowski Z, Minor W (1997) Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol* 276:307–326.
38. Read RJ (2001) Pushing the boundaries of molecular replacement with maximum likelihood. *Acta Crystallogr D Biol Crystallogr* 57:1373–1382.
39. Cowtan K (1994) Joint CCP4 and ESF-EACBM newsletter on protein. *Crystallography* 31:34–38.
40. Pettersen EF, et al. (2004) UCSF Chimera–A visualization system for exploratory research and analysis. *J Comput Chem* 25:1605–1612.
41. DeLano WL (2002) PyMOL molecular graphics system. Available at www.pymol.org. Accessed January 15, 2017.
42. Edgar RC (2004) MUSCLE: Multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res* 32:1792–1797.
43. Rzhetsky A, Nei M (1992) A simple method for estimating and testing minimum-evolution trees. *Mol Biol Evol* 9:945–967.
44. Collaborative Computational Project, Number 4 (1994) The CCP4 suite: Programs for protein crystallography. *Acta Crystallogr D Biol Crystallogr* 50:760–763.
45. Adams PD, et al. (2010) PHENIX: A comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D Biol Crystallogr* 66:213–221.
46. Emsley P, Cowtan K (2004) Coot: Model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* 60:2126–2132.
47. Chen VB, et al. (2010) MolProbity: All-atom structure validation for macromolecular crystallography. *Acta Crystallogr D Biol Crystallogr* 66:12–21.
48. Sievers F, et al. (2011) Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol Syst Biol* 7:539.
49. Landau M, et al. (2005) ConSurf 2005: The projection of evolutionary conservation scores of residues on protein structures. *Nucleic Acids Res* 33:W299–W302.
50. Konagurthu AS, Whisstock JC, Stuckey PJ, Lesk AM (2006) MUSTANG: A multiple structural alignment algorithm. *Proteins* 64:559–574.
51. Johnson M, et al. (2008) NCBI BLAST: A better web interface. *Nucleic Acids Res* 36: W5–W9.
52. Gouet P, Robert X, Courcelle E (2003) ESPript/ENDscript: Extracting and rendering sequence and 3D information from atomic structures of proteins. *Nucleic Acids Res* 31:3320–3323.
53. Kumar S, Stecher G, Tamura K (2016) MEGA7: Molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Mol Biol Evol* 33:1870–1874.
54. Tamura K, Nei M, Kumar S (2004) Prospects for inferring very large phylogenies by using the neighbor-joining method. *Proc Natl Acad Sci USA* 101:11030–11035.
55. Schwede T, Kopp J, Guex N, Peitsch MC (2003) SWISS-MODEL: An automated protein homology-modeling server. *Nucleic Acids Res* 31:3381–3385.