Title:

Structural and Antigenic Characterization of Novel and Diverse *Henipavirus* Glycoproteins

Authors:

Aaron J. May[1,2,*], Muralikrishna Lella[1,#], Jared Lindenberger[1,#], Alex Berkman[1,#], Ujjwal Kumar[1], Moumita Dutta[1], Maggie Barr[1], Rob Parks[1], Amanda Newman[1], Xiao Huang[1], Kijun Song[1,2], Victor Ilevbare[1], Salam Sammour[1], Chan Soo Park[1], Radha Devkota Adhikari[1], Priyanka Devkota[1], Katarzyna Janowska[1], Yanshun Liu[1], Garrett Scapellato[1], Taylor N. Spence[1], Katayoun Mansouri[1], Robert J Edwards[1,3], Kevin O. Saunders[1,4,5], Barton F. Haynes[1,3,4], Priyamvada Acharya[1,2,5,*,$]

Affiliations:

[1] Duke University, Duke Human Vaccine Institute, Durham NC 27710, USA
[2] Duke University, Department of Biochemistry, Durham NC 27710, USA
[3] Duke University, Department of Medicine, Durham, NC 27710, USA
[4] Duke University, Department of Integrative Immunology, Durham NC 27710, USA
[5] Duke University, Department of Surgery, Durham NC 27710, USA
[$]Lead contact

#These authors contributed equally

[*]Correspondence: aaron.may@duke.edu (A.M.), priyamvada.acharya@duke.edu (P.A.)

Summary:
Henipaviruses, a genus within the *Paramyxoviridae* family, include the highly virulent Nipah and Hendra viruses that cause reoccurring outbreaks of deadly disease [1]. Recent discoveries of several new *Paramyxoviridae* species, including the zoonotic Langya virus [2], have revealed much higher antigenic diversity than currently characterized and prompted the reorganization of these viruses into the *Henipavirus* and *Parahenipavirus* genera [3]. Here, to explore the limits of structural and antigenic variation in both genera, collectively referred to here as HNVs, we constructed an expanded, antigenically diverse panel of HNV fusion and attachment glycoproteins from 56 unique HNV strains that better reflects global HNV diversity. We expressed and purified the fusion protein ectodomains and the attachment protein head domains and characterized their biochemical, biophysical and structural properties. We performed immunization experiments in mice leading to the elicitation of antibodies reactive to multiple HNV fusion proteins. Cryo-electron microscopy structures of diverse fusion proteins elucidated molecular determinants of differential pre-fusion state metastability and higher order contacts. A crystal structure of the Gamak virus attachment head domain revealed an additional domain added to the conserved 6-bladed, β-propeller fold. Taken together, these studies expand the known structural and antigenic limits of the HNVs, reveal new cross-reactive epitopes within both genera and provide foundational data for the development of broadly reactive countermeasures.

Introduction

Henipaviruses (HNVs) are a genus of single-stranded RNA (ssRNA) viruses that includes the highly virulent Nipah virus (NiV) responsible for reoccurring outbreaks of deadly disease. The *Henipavirus* genus is a part of the *Paramyxoviridae* family, and therefore related to other notable human pathogens such as the parainfluenza viruses, measles, and mumps [1]. The combined possibility within the family for both rapid transmission, as with measles, and high lethality, as with NiV, further highlights pandemic risk and necessitates urgent research to establish preparedness that will enable a rapid and effective response to a future emergent threat. For HNVs specifically, their identification in diverse animal reservoirs in different geographical locations, their high risk of zoonotic transmission, and current lack of approved vaccines or therapies to treat HNV infection in humans highlights the high risk and pandemic potential of this genus.

HNVs have been studied since the discovery of Hendra (HeV) and Nipah viruses in 1994 and 1998, respectively [4-6]. Additional members of the genus have been indentified since then, but it was not until 2022 with the discovery of Langya virus (LayV) that another species with pathogenicity in humans was confirmed [2]. Like HeV and NiV, LayV was of zoonotic origin. However, unlike HeV, NiV, and most of the other known HNVs at the time, LayV was found to have likely originated from a shrew reservoir, rather than from fruit bats. Several concurrent studies identified more shrew-borne HNVs over a wide geographic range, indicating the existence of a largely uncharacterized clade within the Henipavirus genus [7-9]. Among these studies were some untargeted surveillance efforts, revealing new HNV sequences among dozens of other sequences across multiple virus families [10]. While the newly discovered shrew-borne species had initially been classified as members of the *Henipavirus* genus, they have recently been reclassified into the *Parahenipavirus* genus [3]. In this study, we use the abbreviation HNV to refer collectively to species of both genera.

From the perspective of vaccine or therapeutic countermeasure development for HNV infection, the most important HNV proteins to study are the attachment (G) and fusion (F) surface glycoproteins. These are the two sole surface-exposed HNV proteins, and therefore the only targets for neutralizing antibodies [11]. Together, G and F facilitate virus entry into host cells, with G responsible for receptor binding and F mediating membrane fusion. During this process, both the G and F proteins undergo important conformational changes, including the destabilization of the prefusion conformation of F by G. The specifics of these conformational steps, including how G destabilizes the pre-fusion F conformation or “triggers” its pre-fusion to post-fusion transition remain elusive. Given dramatic structural differences observed among the *Paramyxoviridae* attachment proteins, there may be no universal model that can describe receptor-mediated F protein triggering. Although a class I fusion protein [12], the HNV-F protein, along with all *Paramyxoviridae* F proteins, differs from the canonical class I fusion mechanism in that there is the separation of the receptor attachment mechanism into a separate protein [13].

Whereas other viral fusion proteins proceed through the fusion conversion process by repositioning or removal of their attachment subunits, HNV-F proteins retain inherent metastability through F protein architecture alone.

In addition to mechanistic questions about the HNV surface glycoproteins and their role in virus entry, the antigenicity of these proteins and the possibility of cross-reactivity across the genus remains largely unexplored. While there are several known monoclonal antibodies (mAbs) that are able to neutralize both NiV and HeV [14-18], so far none of these have been found to have any reactivity with LayV [19,20]. Of the few mAbs known to bind to LayV, some have been noted to be cross-reactive with Mojiang virus (MojV), but not to NiV or HeV [20]. Whether there are any broadly reactive anti-HNV antibodies or where the limits of cross-reactivity are within the genus considering newly-discovered species remains unknown. To address these genus-wide mechanistic and antigenic questions, we assembled a curated set of HNV G and F sequences that samples the diversity of the genus as broadly as possible in order to purify and characterize these proteins, biochemically, biophysically, and structurally.

Identification and Classification of Henipavirus Species

To begin categorizing this diverse set of *Henipavirus* strains, we identified as many unique *Henipavirus* G and F proteins as could be found, assessed the phylogeny of these sequences, and developed a universal naming scheme. The NCBI BLAST tool [21] was employed to search for all available unique sequences related to the input sequence, using either the G or F amino acid sequence of either Hendra Virus (NC_001906.3, [22]) or Nipah virus, (NP_112027.1, [23]). In addition to utilizing BLAST, we carried out an extensive literature search to include other sequences that may not have been discoverable through BLAST, such as Gamak virus [7]. After identifying related sequences, we narrowed the list to include only sequences from mostly complete genomes, rather than partial DNA sequences of only G or F proteins. Almost all the identified strains had genome lengths close to the canonical 18.2 kbp for Henipaviruses [24], with some species being slightly longer. No sequence in our panel had a deposited genome shorter than 11.2 kbp. The several sequences that were shorter than the rest were either Nipah-Malaysia or Hendra strains that were not recently discovered and have been well-characterized in the literature.

Next, the Clustal Omega tool [25] was used to align the G and F amino acid sequences. As our goal was to purify soluble ectodomain constructs, we truncated the sequences, using trends in the sequence alignments to decide on exact truncation points between species. HNV-F proteins were truncated at the C-terminal end of the ectodomain, at the site equivalent to Nipah virus residue 488 to be consistent with previous purifications of F ectodomains [19]. For the G proteins, the boundaries of domains do not align as consistently across species as they do with the overall less variable F protein. For ectodomains, the N-terminal portion of the sequence up to the end of the transmembrane domain was removed, with the position set at the analogous site to Nipah

virus residue 71. For G head domains, the site analogous to Nipah virus residue 177 was selected as a suitable start point based on sequence analysis and available structural data for Nipah and Langya G proteins [14,26]. When comparing the truncated sequences, we found that the amino acid sequences of several strains differed from other panel members only in their signal sequence, transmembrane domains, cytosolic domains, or other regions not included in the final constructs. Therefore, these redundant sequences were eliminated from the panel, and the remaining sequences were assembled into a phylogenetic tree based on complete genomes (**Figure 1**).

To begin categorizing these sequences based on their phylogeny, we used the previous handling of Nipah virus naming in the literature. Previously, strains have been named based on the location the sequence was reported, specifically, Bangladesh (NiV-B) and Malaysia (NiV-M). Our tree supports the existence of potentially two further groupings, that we have also named based on location, India (NiV-I) and Cambodia (NiV-C) (**Figure 1**). With the Bangladesh strains, the phylogeny revealed the grouping of strains into two clades, prompting us to denote the two groupings as NiV-B1 and NiV-B2, respectively. From here, unique entries were designated an additional number to further classify the variant (i.e. NiV-B2.1, NiV-B2.2, NiV-B3.1, etc.). An identical scheme was applied for the India strains which previously have been referred to as members of the Bangladesh clade [27]. Malaysia strains were classified by number as well. Entry MK801755 was identified during surveillance in Cambodia in 2003, where it was isolated from *Pteropus lylei* fruit bats [28]. Appearing distinct from most other Malaysian strains in our tree, it was assigned as a new Cambodia clade (NiV-C).

Previous studies have established two clades for Hendra virus [29], in which the more recently discovered clade is referred to as HeV-g2. However, due to the use of G as an abbreviation for the attachment protein, we sought to adopt a different naming scheme to avoid confusion. Unlike Nipah, where genomic differences align with country borders, all Hendra virus discoveries have been in Australia. Therefore, the two distinct clades have been named Alpha (α) and Beta (β), with each member of the clade given a number. The other Henipaviruses that have emerged from a fruit bat reservoir have all previously been named in the literature. These include Cedar Virus (CedV, [30]), Ghana Virus (GhV, [31]), and Angavokely Virus (AngV, [32]). Our panel includes two unique Cedar virus attachment proteins; therefore, these strains were numbered 1 and 2. There are only singular strains of Ghana and Angavokely viruses identified, however some evidence indicated that the originally deposited sequence for Ghana virus F protein may have included an inaccurate N-terminal end of the protein and a corrected sequence was suggested [33]. While our phylogenetic tree only includes the one strain reported, our panel of purified proteins included both F protein sequences.

Among the species found in shrews, now officially classified as members of the *Parahenipavirus* genus, several already have established names in the literature, including Mojiang Virus (MojV), Langya Virus (LayV), Gamak Virus (GakV), Denwin Virus (DewV),

Melian Virus (MelV), and Ninorex Virus (NinV) [2,7,8,34]. For those with more than one strain reported, specifically Langya Virus (LayV) and Gamak Virus (GakV), each strain was designated with a number. However, many of the sequences were given repetitive names based on their shared collection site. To simplify these names, we began using the name "Shrew Henipavirus (SHNV)" with a number, thus establishing SHNV1, SHNV2, SHNV3, SHNV4, and SHNV5. For SHNV5 specifically, there were 3 distinct strains that the phylogenetic tree indicated were closely related enough to be considered the same species (**Figure 1**). As a result, we gave an additional number to each strain. Interestingly, these sequences were not all from the same collection event or species, indicating that SHNV5 is found in several Chinese shrew populations. Similarly, while Lee et al,. 2020 [7] previously reported the discovery of Daeryong Virus in Korea, our phylogenetic analysis revealed that Jingmen Crocidura shantungensis virus 2 (OM030315), detected in China, should be considered the same species as Daeryong Virus. Following the naming procedure used for Nipah virus, the strains were given the names Daeryong virus-Korea (DarV-K) and Daeryong virus-China (DarV-C). Again, this demonstrates that several of these newly discovered Parahenipaviruses are geographically well-distributed. In the case of Crocidura tanakae Henipavirus (OQ970176), we renamed the virus Shiyan Virus (ShyV) given the identification of the virus in Shiyan, China. A full listing of the panel, including original deposition names, accession codes, genus, and sequence ranges used is available as **Supplemental Table 1.**

Preparation of Diverse HNV-F Ectodomains

To prepare a panel of HNV F proteins we expressed constructs that encoded the ectodomains spanning residues 1 to the equivalent of 488 in NiV. Following the end of the ectodomain, we added a foldon trimerization domain [35] and a C-terminal Twin Strep tag (IBA Lifesciences). The proteins were expressed through transient transfection in HEK 293 cells and purified using their C-terminal Strep affinity tags followed by size exclusion chromatography (SEC). The yield of F proteins was variable, with some constructs expressing over 2 mg/L and others less than 0.1 mg/L (**Figure 2A**). For eight F proteins, the yields were too low to perform biochemical characterization. We found that yields could vary dramatically between strains that only differed by a few amino acids. For example, NiV-B1.1-F and NiV-B1.2-F differ by only three residues in the ectodomain, but while NiV-B1.2-F produced ~1 mg/L, NiV-B1.1-F had no yield in both of two purification attempts. The sequence identity relative to NiV-M1-F is high for both NiV and HeV strains, with all NiV sequences over 98% and HeV strains over 97% (**Figure 2A**). However, the sequence identity drops dramatically outside this grouping, with most other species at or below 40% sequence identity to NiV-M1-F. Sequence similarity also follows this drop outside the NiV and HeV clade, albeit with the percentages much higher, with no sequence dropping below 55% similar. When considering the phylogeny of HNV-F proteins on the basis of amino acid sequence, the tree remains very similar to the overall genetic-based tree with several small differences. Specifically, the distinction between NiV-C-F and NiV-M strains is more apparent, the F proteins of ShyV, SHNV3 appear to be very closely related, and AngV-F

appears to be distinct from all other HNV strains, rather than part of the bat clade as seen through full-genome phylogeny (**Supplemental Figure 1, Top**).

HNV-F ectodomains typically elute from SEC with three peaks, one that elutes immediately after the column void volume, a very broad "middle peak," and the main peak that corresponds to the molecular weight of an F trimer eluting last (**Figure 2B, Supplemental Figure 2**), with the ratio of the peaks differing from strain to strain (**Figure 2B**). Furthermore, the middle peak of some strains can sometimes be seen to have a shoulder, indicating there could be multiple sized components eluting within this peak. The SDS-PAGE profile of the middle peak components is similar to that of the main trimer peak, with a major band observed at a position identical to where the band corresponding to the main trimer peak appears and some minor higher molecular weight bands observed sometimes, indicating that the presence of the middle peak is not caused by any kind of covalent linkage, nor any differential glycosylation significant enough to change the apparent molecular weight by several hundred kDa (**Supplemental Figure 3**). This conclusion is supported by mass photometry analysis of the middle and main peaks. Both peaks individually or mixed together generated an essentially identical molecular weight profile, with each being dominated by the expected molecular weight of the trimer, and both showing a very small presence of species twice to three times the mass of the trimer (**Figure 2C**). Negative stain electron microscopy (NSEM) analysis of each middle or main peak reveals comparable sets of two-dimensional class averages. The pre- and postfusion shapes typically seen previously [19] can be seen throughout the panel for both middle and main peak components (**Supplemental Figure 4**). Thus, the middle peak appears to be indicative of transient non-covalent association of the F ectodomain trimers.

Antigenicity of HNV-F Ectodomains

We next sought to characterize the antigenicity of the F ectodomain panel. Given the novelty of recently identified HNV species and limited availability of antibodies that react with LayV and other newly categorized members of our panel, we sought to elicit F-directed antibodies from vaccination of mice with several HNV-F antigens (**Figure 2D)**. One group was given four injections of the LayV-1-F ectodomain, and for two others, the fourth injection of LayV-1-F was replaced with NiV-F protein constructs, either the wild-type NiV-M1-F, or NiVop08, a pre-fusion stabilized NiV-F construct [36]. By boosting with NiV-F constructs, we sought to elicit broadly reactive antibodies. Using hybridoma fusion technology [37], we identified two murine IgG1 antibodies, named 1C8 and 22F5, both from group 2, with reactivity to LayV-1-F.

Either through ELISA or Bio-layer interferometry (BLI), we assessed the binding of these newly identified antibodies, as well as previously identified anti-NiV or anti-LayV antibodies, to our panel of HNV-F ectodomains [20,38]. 1H1 and 4B8, previously identified as anti-NiV-F antibodies, displayed broad reactivity to all NiV and HeV strains, but extremely limited

reactivity to all other species. 4G5, previously identified as an anti-LayV-F and MojV-F antibody, strongly bound not only to LayV-F and MojV-F, but also to SHNV5-1-F and SHNV5-2-F, thus indicating that antibodies targeting the 4G5 epitope have greater breadth than was previously known (**Figure 2E, Supplemental Figure 5-6**). A panel of mouse anti-NiV-F Fabs broadly reacted to nearly all NiV and HeV antigens, as expected. The two mouse antibodies identified in our vaccination study demonstrated strong reactivity not only to LayV-F proteins, but also to MojV-F and both SHNV5 strains as well (**Figure 2E, Supplemental Figure 6**). Furthermore, 22F5, which generated stronger binding signals than 1C8, bound to SHNV2-F as well.

Using cryo-EM, we determined a structure of 22F5 bound to the post-fusion form of the LayV-F protein **(Figure 2F, Supplemental Figure 7)**. 22F5 bound to the DII region [11] of LayV-F. While there have been several previously observed antibodies targeting DII epitopes in NiV-F and HeV-F proteins [38], the only other available structure for an anti-LayV-F antibody, 4G5, details an epitope that includes contributions from HRA and HRB[20]. Using Surface Plasmon Resonance (SPR), we performed a competitive binding assay with 22F5 and 4G5 that confirmed that they do not compete **(Supplemental Figure 8A).** Assessing 22F5 binding to a variety of LayV-F ectodomains using SPR **(Figure 2G, Supplemental Figure 8B),** we observed binding to mixed or uniformly post-fusion populations, as well as to two separate disulfide stabilized pre-fusion LayV-F constructs described previously, that we refer to here as DS [39] and DSv2 [20]. Thus, the 22F5 epitope is accessible in both the pre- and post-fusion conformation of LayV-F, with binding observed at low nM affinities in both cases (**Supplemental Figure 8B)**. Together with the data assessing 22F5 binding to the panel of F proteins, we can establish 22F5 as a valuable tool for studying *Parahenipavirus* F proteins, given its conformational bispecificity and greater breadth than previously identified anti-*Parahenipavirus* antibodies.

Differential Scanning Fluorimetry of HNV F ectodomains

To assess the stability of the purified F ectodomains, we performed Differential Scanning Fluorimetry (DSF), a label-free method that measures changes in intrinsic fluorescence as the protein unfolds. We first performed DSF on two wild-type proteins, NiV-M1-F and LayV-1-F, and one prefusion-stabilized NiV-F construct, NiVop08 [36]. The DSF profiles were consistent with our previous reports [19]. With each peak in a DSF profile indicating a conformational event, for the unstabilized F ectodomains we observed one transition between 45-65ºC and one or more transitions at higher temperature, up to over 90ºC (**Figure 3A**). LayV-1-F showed two transitions with a negative peak at 50ºC and a positive peak at 86ºC. NiV-M1-F showed three transitions with a negative peak 56ºC, and positive peaks at 76ºC, and 92ºC. NiVop08 showed two transitions, a negative peak at 66ºC and a positive peak at 76ºC. We next measured the the DSF profiles of the same set of proteins after pre-incubation at 60ºC for 90 minutes, based on previous observations that heat treatment could force conformational conversion in *Paramyxoviridae* F proteins [40]. Overlaying all DSF profiles revealed that the lower-temperature negative peak

present in all three proteins initially was eliminated in only the two wild-type proteins after heat incubation (**Figure 3A**). Negative Stain Electron Microscopy (NSEM) confirmed that the wild type proteins have been fully converted to the postfusion conformation, as revealed by the characteristic elongated six-helix bundle present in the 2D class averages (**Figure 3B**). The NiVop08 prefusion stabilized construct retained the more compact shape characteristic of the prefusion state (**Figure 3B**). Together, these results indicate that the negative DSF peak present around 45-60ºC is indicative of a pre- to post-fusion conformational change taking place.

We next assessed all purified F proteins for which there was sufficient material purified, typically at least 5µg, with DSF. We observed a wide range of conversion peak intensities and temperatures (**Figure 3C-F**). While all proteins generated a large positive peak at high temperature indicative of a denaturation event, not all proteins produced a detectable conversion peak. 14 proteins had essentially no detectable signal prior to the high temperature peak (**Figure 3D**). Two-dimensional class averages from NSEM analysis revealed consistent and universal postfusion conformation for these proteins, confirming DSF results. Though not exclusively, most of the proteins with no conversion peak are Parahenipaviruses, and inversely, all of the Parahenipaviruses have no conversion peak, except for Langya virus strains. Many of the Parahenipavirus F proteins could not be assessed due to insufficient purification yield, and these results may suggest it could be due to inherent instability. Among the F proteins that generated a conversion peak, there were shifts in temperature and intensity (**Figure 3E**). These changes did not appear related to species differences, with NiV-M1-F and NiV-M3-F, separated only by one amino acid mutation, being right- and left-shifted compared to the average, respectively. Both Langya virus strains were the most left-shifted of the set. Two-dimensional NSEM classes show a mix of pre- and postfusion conformations in this set, indicating that the presence of this negative peak indicates only some amount of prefusion conformation present, rather than presence of a uniform set of prefusion proteins. Between the conversion peak and the final high temperature peak is typically a broad, positive peak, present with differing intensities and shapes. Further investigation would be required to determine if this peak is either part of any natural conformational event or part of the denaturation process.

Two proteins of the panel, AngV-F and the sequence-corrected version of GhV-F that restores the signal peptide, GhV(+A)-F, display DSF profiles that are outliers to the rest of the panel (**Figure 3F**). Interestingly, both proteins display a profile that seems closer to NiVop08, with their high temperature positive peak significantly left-shifted. However, these are wild-type proteins that can be assumed to be capable of undergoing normal conformational conversion, barring unexpected and dramatic mechanistic differences. Two-dimensional NSEM classes for these two proteins show the presence of largely, but not exclusively, prefusion conformations.

Since DSF is reliant on primarily Tryptophan residues to generate signal, it is possible that differences in DSF profile could arise from differing number and position of Tryptophans,

however, the F ectodomain has only one Tryptophan that is almost universally conserved, with only CedV-1-F mutated at this spot (**Supplemental Figure 9**). Furthermore, AngV-F contains one additional Tryptophan that is unique to AngV-F. All F ectodomain constructs additionally contain one Tryptophan in the foldon tag used to assist with trimerization and two in the twin-strep tag used for purification. In general, sequence variability is likely not a significant factor in the differences observed in the DSF profiles of HNV- F proteins.

Cryo-EM Structures of Diverse HNV-F Proteins

To understand the structure of an HNV-F protein that has divergent properties from the currently characterized F proteins, we determined the structure of the AngV-F protein by cryo-EM single particle analysis (**Figure 4, Supplemental Figure 10-12, Supplemental Table 2**). The cryo-EM dataset revealed three distinct particle populations (**Figure 4A**). The most populated class was of the isolated F ectodomain trimer, followed by a dimer-of-trimers population, and finally a population where the F proteins formed a hexameric array. The isolated AngV-F trimer was refined to a resolution of 4.0 Å and used for detailed examination of the F ectodomain structure and for comparing to the previously known structures of other F proteins. In cryo-EM structures of the HNV-F proteins the stalk region is typically poorly defined [15,19,38,41,42]. Our AngV-F cryo-EM structure revealed well-defined density for the stalk, allowing atomic-level modeling.

The AngV-F ectodomain exhibited a larger central cavity compared to NiV-F and LayV-F proteins. This expanded cavity is the result of both the stalk region not inserting as far into the cavity as in other HNV-F proteins and the domains surrounding the cavity being spaced farther apart (**Figure 4B**). In the previously determined HNV-F structures, the helical stalk region begins with a Lysine at its N-terminal end. This trio of Lysines forms the bottom of the internal cavity. In AngV-F, we observe a Proline at the beginning of the stalk helix, also known as the HRB helix (**Figure 4C**). The surface electrostatics of this region are strikingly different at this position for AngV-F compared to NiV-F and LayV-F structures. The proximity of the Lysines in LayV-F and NiV-F and their electrostatic repulsion could be a source of induced instability in the pre-fusion form of the F protein, as charge repulsion may force apart the HRB helix during conformational conversion. The absence of this Lysine trio and replacement with a Proline cap at the N terminus of the HRB helix may lend greater stability to the pre-fusion form of the AngV-F protein. Notably, all species in our panel have either a Lysine or Arginine at this site, with only AngV-F substituting with a Proline (**Supplemental Data 1**).

Several glycans could be resolved (N61, N93, N353, N434 and N458), including ones analogous to those previously defined in HNV-F structures [41,42]. The glycans shield known sites of vulnerability in *Paramyxoviridae* F proteins, including at the apex (N61), near the fusion peptide (N93 in Domain III), and on the stalk domain (N458). Interestingly, the glycan proximal

to the fusion peptide utilizes a non-canonical NNV sequence (**Figure 4D**). The MojV-F and SHNV2-F proteins also have NNV sequences at this site that may be similarly glycosylated. Glycans in the apex region near N61 and the HRB/stalk region near N458 are observed in many HNV-F protein structures (**Supplemental Figure 13**). However, there is no AngV-F analogue to the N414 glycan of HeV-F and NiV-F. Instead, AngV-F contains two novel glycans in DI (N434) and DII (N353). The non-canonical NNV glycan at N93 is also found at the same site in HeV-F and NiV-F, but interestingly, not LayV-F. The differential glycosylation patterns observed across diverse HNV-F species may play a role in determining the potential for cross-reactivity, but critically, the presence of the NNV glycan in AngV-F indicates that prediction of glycosylation through sequence alone may be difficult.

Interactions between the AngV-F trimers were mediated through intermolecular contacts along the DI and DIII domains. These include hydrophobic interactions in DIII between L157 and V159 on neighboring ectodomains, both positioned on a loop that is part of an important antigenic site [11,38] (**Figure 4E**). Furthermore, there are charge interactions involving H45 from DI and T99 and F105 of DIII (**Figure 4E)**. Similar oligomeric states between pre-fusion F protein molecules (trimer, dimer-of-trimers, trimer-of-trimers) were also found in the cryo-EM dataset of a Hendra virus F protein (HeV-α2-F) (**Supplemental Figure 12**). AngV-F dimers resemble "mirror-like" protein dimers, with each exterior interaction surface of a protomer available for dimerization. These dimers, in turn, interact with protomers from three other trimers, forming hexameric structures. Each trimer is consistently aligned within the plane of the hexameric assembly, leading to a horizontally level lattice, distinct from an NiV-F crystal structure in which the lattice was not within one plane [41].

Taken together, our cryo-EM structures revealed that, while maintaining the overall conserved HNV-F architecture as observed in previously characterized HNV-F proteins, key mutations imparted notably distinct pre-fusion characteristics on the AngV-F protein. They also revealed multimerized pre-fusion F proteins adding to the growing body of literature that suggest a role for the multimeric assemblies in the context of the virion and suggest a distinct mode of F protein multimerization leading to the formation of hexameric lattices for the AngV-F protein.

Purification and Characterization of Diverse HNV-G Head Domains

To analyze our panel of HNV-G proteins, we began with a focus on the isolated head domain, which is responsible for receptor binding and is an important antigenic site. To allow for the previously described head domain sequences to be secreted in a mammalian expression system, we added an artificial secretion signal [43], and appended a C-terminal 8x His tag. The G head domains were transiently transfected and affinity purified with cobalt resin affinity chromatography, followed by SEC. The typical SEC profile for head domain proteins reflects the simple nature of their single-domain fold. All proteins eluted with a consistent single main peak (**Figure 5A**), typically with a small, higher molecular shoulder or peak that could be aggregates.

Yields for the head domains ranged from 2.5 to 35.3 mg/L (**Figure 5B**), with the majority above 10 mg/L.

Compared to F proteins, HNV-G proteins demonstrate much greater sequence variability. Sequence identity values are only ~78% between NiV and HeV, and drop to below 20% for *Parahenipavirus* G proteins, compared to NiV (**Figure 5B**). The sequence similarity scores are also rather low, with most Parahenipavirus strains at around 35% similarity to NiV (**Figure 5B**). When considering amino acid-based phylogeny, the relationships between strains remains similar. One notable difference is that ShyV-G and SHNV3-G, as with their F counterparts, appear more closely related than their full genome would suggest (**Supplemental Figure 1, Bottom**). This limited sequence similarity suggests that there could be greater antigenic, structrual, and even mechanistic differences between different HNV-G proteins than among F proteins.

We used DSF to assess the thermostability of the G head domains. Most proteins in the G head domain panel generated a profile with a single peak between 60-70ºC. Among bat-clade strains demonstrating single-peak profiles, there were some small shifts in the inflection temperature that was largely consistent within species groupings (**Figure 5C**). The NiV-M, B, and I clades generally had higher inflection temperatures than HeV strains (**Figure 5C)**. Among HeV strains, α had notably lower inflection temperatures than almost all other strains. *Parahenipavirus* strains demonstrating single-peak profiles also had a diversity of inflection temperatures that tracked with species groupings (**Figure 5D)**. The closely-related LayV and MojV showed the highest inflection temperatures for their G head domains, with the DewV and MelV G head domains having much lower inflection temperatures. However, several species generated profiles distinct from the most common result. Specifically, several species generated a rather broad and left-shifted peak, and another, GhV-G, inverting the main peak, instead generating a negative peak (**Figure 5E**). Two species, CedV-1-G, and SHNV2-G generated profiles with multiple peaks (**Figure 5F**). Unlike the relatively high conservation of Tryptophan positions in F proteins, G proteins display a much higher level of sequence diversity (**Supplemental Figure 9**). Indeed, there are significant differences in Tryptophan positions in species that generated outlier profiles, CedV, GhV, AngV, and SHNV2 especially. Despite the variability of Tryptophans in the G protein across species groupings that precludes genus-wide comparisons of DSF profiles, these results provide interesting comparative insights within species groupings that may have implications for differences in receptor binding and virus fitness between species. We have found, for example, that for the SARS-CoV-2 spike protein, thermostability of the receptor binding domain (RBD) is a property that is optimized during virus evolution [44-46]

We used ELISA to assess the antigenicity of our G head domain panel with two previously characterized anti-NiV-G antibodies, nAH1.3 and 1E5 [47,48]. We observed expected

genus-wide trends with these two antibodies demonstrating strong reactivity to all NiV and HeV constructs in the panel, but very limited reactivity outside this set. 1E5 demonstrated marginal reactivity to several species, specifically GhV, both DarV species, and MelV. Interestingly, nAH1.3 also demonstrated limited reactivity to DarV-K-G only (**Figure 5G, Supplemental Figure 14**).

Structure of Gamak Virus Attachment Protein Head Domain

HNV-G proteins are tetramers (dimer of dimers), with each monomer in the ectodomain composed of a helical stalk domain, a neck domain that is involved in oligomerization and control of quaternary structure, and a 6-bladed, β-propeller head domain where the receptor-binding site is found. A crystal structure of the GakV-2-G protein head domain determined at 1.4 Å resolution (**Figure 6A, Supplemental Table 3)** showed the expected 6-bladed β-propeller architecture that is well-conserved among HNV-G proteins (**Figure 6A**). A well-resolved glycan was observed in the region that is analagous to the Ephrin (EFN) receptor binding surface in the NiV-G head domain (**Figure 6B).** While NiV, HeV, GhV and CedV bind the Ephrin family of receptors at a common receptor binding surface, the receptors for many HNVs, including LayV, MojV, and GakV remain unknown. Though EFNB2/3 engages the same face of the head domain in NiV and HeV that sialic acid does in other Paramyxoviruses [49,50], the measles virus receptor, CD150, binds at the side surface of the head domain, making major contacts along blades 5 and 6 [51]. The glycan present in the GakV-2-G head domain structure would likely hinder binding of a protein receptor in the EFNB2/3 site, suggesting that GakV either uses a smaller receptor, binds along a side face, as in measles virus, or adopts a different, yet to be discovered mode altogether. The Ephrin binding site is an immunogenic site in NiV/HeV and the target of neutralizing antibodies. An addition of a glycan in the GakV-2-G head domain would pose a steric barrier to the elicitation of antibodies targeting this site, suggesting that the GakV-2-G protein will have an antigenic profile distinct from the Ephrin-binding HNV-G proteins.

Though the GakV-2-G head domain structure was overall similar to that of other head domains (**Figure 6C**), some differences were noted. The N-terminus of the head domain appeared well ordered and packed against the 6-blade β-propeller core (**Figure 6D**). Residues 194-199 threaded between blades 5 and 6, contributing a β-strand to blade 6 as it continued to lead into blade 1. This stretch of residues also contacted blades 2, 4, and 5 thus forming a connection between different secondary structures within the β-propeller. Interestingly, a novel minidomain was observed at the C-terminus of the head domain, spanning residues 606 to 628, that formed a three-stranded β-sheet stacked against blade 5 of the β-propeller (**Figure 6D-E**). This additional compact and ordered domain is distinct from the previosuly known structures of *Henipavirus* G protein head domains.

Taken together, the structure of the GakV G head domain illustrates the structural diversity that can be accommodated within the conserved β-propeller fold to impact receptor tropism and antigenicity.

Discussion

Over the past few years, there has been a rapid expansion of identified species and strains in the *Henipavirus* genus, ultimately leading to the separation of these species into the *Henipavirus* and *Parahenipavirus* genera [3]. Given the history of zoonotic transmission with these viruses, charcterization of these novel species is an important first step for pandemic readiness. Here we characterize a diversified panel of HNV F and G proteins to construct a panel that better reflects the global diversity of HNVs. Until now, many of the strains in our expanded panel of Henipaviruses were totally uncharacterized and their relationships to other Henipaviruses unknown. In this study we have been able to identify two new species groupings containing strains that were identified independently. Given that the two strains of Daeryong virus were from sample collection in two different countries (Korea and China), this confirms that just as with Nipah, there is likely a broad geographic range of these newly-identified Parahenipaviruses.

Purifying and characterizing the F protein ectodomains of these strains has revealed that despite low sequence identity, HNV-F proteins share many architectural and biophysical similarities, a trend we previously noted by comparing Langya virus to Nipah virus [19], and one that is apparent with our species panel. One of the most notable similarities was the essentially universal presence of some level of oligomerization between F trimers. This was apparent from the SEC profiles, but the presence of trimer multimers did not seem to correlate with other parameters measured, such as purification yield, metastability, or phylogeny. It appears instead that the formation of these complexes is transient and dependent on the concentration and environment of the F protein, as suggested by the almost exclusive presence of trimer molecular weights detected by mass photometry where sample concentration is much lower than during size-exclusion chromatography and structural analysis. Regardless, the ubiquity of this type of interaction indicates that oligomerization of F proteins likely is a conserved trait within HNVs.

In the case of Angavokely virus, the formation of a hexameric lattice of F protein trimers was observed in the cryo-EM datasets. The existence of a hexamer-of-trimers arrangment has been noted previously with NiV [41], though the biological relevance was limited by the use of crosslinking to stabilize the arrangement. Our structure demonstrates the ability of AngV-F to multimerize in this way without any such assistance and, critically, without the involvement of the transmembrane domain, viral envelope, cytoplasmic domain, or viral matrix protein. The hexameric lattice is formed purely through ectodomain interactions, validating the characterizations of inter-trimer interactions previously reported for NiV [39], yet distinct in the implicated residues, with the AngV-F interaction involving DI and DIII, while the NiV-F interaction is dominated by DII residues.

Previous studies have defined the lattice of *Paramyxoviridae* and *Pneumoviridae* F proteins, with various arrangements being determined. In one case, hexameric lattices have been observed, as with a whole-virion tomography structure of parainfluenza virus 3 [40]. In others, such as with measles virus and respiratory syncytial virus, alternate arrangements of fusion and attachment proteins have been observed [52,53]. Our results with AngV-F suggest that HNV-F proteins may adopt the hexameric arrangement, though structural studies of pseudoviruses and whole virions would help elucidate the physiological relevance of this finding. Furthermore, it cannot be ruled out that the formation of a hexameric F lattice may comprise one stage of HNV-F functionality, as the arrangment could shift during some type of maturation event or during G-mediated triggering.

Our biophysical characterizations of HNV-F proteins gives us a tool with which we can assess their metastability. Although HNV F protein pre- and postfusion conformations are visually distinct and can be readily identified in a structural dataset, obtaining structural data for a large panel can be expensive and time-consuming. Our DSF-based assay reports on F protein conformational state without relying on structure determination, and using this assay we were able to identify distinct signatures for pre-to-post fusion conversion and F protein denaturation revealing how differences in sequence or structure could impact metastability. In the case of AngV-F, the striking differences between its structure and those of other HNV-F proteins line up with the differences observed in its DSF profile. Clearly, AngV-F has greater pre-fusion state stability than other HNV-F proteins. A wider internal cavity may play a role in this, but more likely, the altered electrostatics and introduction of a proline at the beginning of the stalk domain, also referred to as the HRB region, are responsible for this change. During conformational conversion in most HNV-F proteins, it is likely that the electropositive cluster works to disperse the helix bundle, lowering the energetic barrier to conformational conversion. Instead, in AngV-F, the absence of such electrostatic repulsion at this site along with the presence of a proline at the start of the HRB helix leads to a more stable prefusion HRB arrangement, which is likely why the AngV-F stalk domain is so well resolved in our cryo-EM structure.

DSF analysis of G head domains also reveals a diversity of themostability patterns. Although G head domains are comprised of a single domain 6-stranded β-propeller fold and are not known to have any major internal conformational rearrangements as part of their receptor-binding mechanism, some studies have indicated that receptor binding by HNV-G proteins could effect subtle changes in protein flexibility that do not manifest as large-scale conformational events [54]. The differences in thermostability could potentially correlate with mechanistic differences between strains; however, more structural data detailing the structural changes that take place after receptor binding in G proteins will be needed to confirm this. From our crystal structure of the Gamak virus head domain, we can observe other likely sources of mechanistic

differences between species. Specifically, the presence of a glycan in the face of the β-propeller analagous to the binding site for the Ephrin receptor in NiV-G indicates that it is unlikely a proteinaceous receptor similar to EFNB2/B3 that binds at this site in GakV. Furthermore, the altered arrangement of the N- and C-terminal end of the head domain and the presence of a novel minidomain could suggest a different pattern for quaternary interactions in GakV-G ectodomain tetramers, different types of conformational changes within GakV-G after receptor binding, or it may even serve as a possible receptor binding site for an as of yet undiscovered receptor.

While several anti-NiV and anti-HeV neutralizing antibodies have been well-characterized, the antigenicity of novel HNV strains had not yet been explored. By assessing the binding of a panel of both previously identified anti-HNV antibodies and newly discovered antibodies elicited through vaccination, we can now better describe the antigenic landscape of HNV proteins. As expected, we saw little ability of anti-NiV/HeV antibodies to bind to non-NiV/HeV proteins, with the exception of 1E5, which demonstrated some level of reactivity to select strains outside this group. Most interestingly, we found clear antigenic overlap between the F proteins of LayV, MojV, SHNV5, and to a lesser extent, GakV. Both a previously reported anti-MojV-F antibody, 4G5, and our vaccine-elicited antibodies, 22F5 especially, bound to these strains. When considering our phylogenetic analysis, both by full genome and amino acid sequence, it can be seen this cross-reactivity spans distantly related strains. SHNV5 appears to be more distantly related from LayV than HeV is from NiV. GakV is even more distant phylogenetically, appearing to be roughly as distant from LayV as CedV is from NiV.

With a group of viruses prone to such sequence diversity, one major concern in the event of a possible widespread outbreak of disease in humans is that mutations would occur so quickly that the immunity conferred from prior infection or immunization would not be able to track with newer strains. This phenomenon can be observed with yearly outbreaks of new influenza strains or with the progression of SARS-CoV-2 evolution, where new strains emerged and quickly dominated the pool of circulating virus as they evaded prior immunity. A key focus of efforts to address this possibility have long been designing immunogens that can elicit antibodies with broad coverage of the currently exisitng and, hopefully, future diversity of strains. Being able to draw antigenic boundaries within the *Henipavirus* and *Parahenipavirus* genera is a foundation for similar efforts, and here we see that while the prospect of cross-reactivity over the entire genus remains elusive, there is already demonstrated cross-reactivity among a large portion of the Parahenipaviruses. While it will be important to assess this observed cross-reactivity in neutralization assays and animal challenge models, there is currently a lack of knowledge both of the host receptors for Parahenipaviruses and of the proper cell lines or animal models to study infection. Therefore, in the absence of appropriate experimental systems to study the virologic properties and the breadth of protection from potential HNV infections, the discoveries reported here begin to build a foundation for future exploration of *Parahenipavirus* antigenicity.

Ultimately, this broad characterization of HNV proteins provides a set of valuable biophysical and biochemical data for many species and strains that, to this point, had only been known as a sequence in a database. As these novel species are studied in more depth, we hope that our analysis can be used to provide context for the comparison of other key aspects of *Henipavirus* biochemistry, structure, and immunology.

Acknowledgements
Cryo-EM data were collected at the Duke Krios at the Duke University Shared Materials Instrumentation Facility (SMIF), a member of the North Carolina Research Triangle Nanotechnology Network (RTNN), which is supported by the National Science Foundation (award number ECCS-2025064) as part of the National Nanotechnology Coordinated Infrastructure (NNCI). We thank Dr. Nilakshee Bhattacharya for assistance with microscope alignments. This study utilized the computational resources offered by Duke Research Computing (http://rc.duke.edu; NIH 1S10OD018164-01) at Duke University. This research used resources of the National Synchrotron Light Source II, a U.S. Department of Energy (DOE) Office of Science User Facility operated for the DOE Office of Science by Brookhaven National Laboratory under Contract No. DE-SC0012704. Use of the NYX beamline 19-ID at the National Synchrotron Light Source II was supported by the New York Structural Biology Center. NYX detector instrumentation was supported by grant S10OD030394 through the Office of the Director, National Institutes of Health. This work was supported by funding from the Translating Duke Health Initiative, Duke School of Medicine, and NIH R01 AI 165147 (P.A.). We thank Hannah Violette and Nicholas Levering for technical assistance, and Nilakshee Bhattacharya for assistance with cryo-EM data collection.

Author contributions: Conceptualization – A.M. and P.A.; Methodology –A.M., A.B., B.F.H. and P.A.; Investigation – A.M., M.L., J.J.L., A.B., M.D., M.B., R.B., X.H., U.K., K.S., V.I., S.S., C.S.P., R.D.A., P.D., K.J., Y.L., R.J.E., B.F.H., P.A. Data Curation – A.M., M.L., J.J.L, A.B. and P.A. Validation – A.M., M.L., J.J.L, A.B. and P.A..; Resources – M.T. and F.W.A.; Visualization – A.M., M.L., J.J.L, A.B. and P.A.; Writing - Original Draft – A.M., A.B and P.A.; Writing - Review & Editing – all authors; Supervision – A.M., R.J.E, B.F.H. and P.A.; Project Administration –A.M. and P.A.; Funding acquisition – B.F.H. and P.A.

Declaration of interest.
A.M., M.L., J.J.L, A.B., B.F.H. and P.A. are named in provisional patents submitted on the methods for stabilizing the pre-fusion HNV F and G proteins based on the findings reported in this paper. Other authors declare no competing interests.

Data and Material Availability

Cryo-EM reconstructions and atomic models generated during this study are available at wwPDB and EMBD (www.rcsb.org; http://emsearch.rutgers.edu) under the following accession codes: AngV-F trimer - EMD ID: EMD-48423 and PDB ID: 9MNH; AngV-F dimer-of-trimer - EMD ID: EMD-48535 and PDB ID: 9MQN. The crystal structure of the GakV G protein head domain was deposited with the PDB ID: 9EHU.
All materials will be made available upon request.

Ethics Statement

All mice were cared for in a facility accredited by the Association for Assessment and Accreditation of Laboratory Animal Care International (AAALAC). All study protocol and all veterinarian procedures were approved by the Duke University Institutional Animal Care and Use Committee (IACUC).

Environment: All rooms are on a 12/12 light cycle unless otherwise requested. Heat and humidity are maintained within the parameters outlined in The Guide for the Care and Use of Laboratory Animals and animals are fed a standard rodent diet. Breeding or specialty diets will be provided upon request. Rodents and Guinea pigs are housed in individually ventilated micro-isolator caging on corn cob bedding. DLAR will provide alternative bedding based on scientific need. Mice are changed every two weeks, rats are changed twice a week, and Guinea pigs are changed three times a week or more frequently as needed. Group housing of up to 5 mice per cage is strongly encouraged. Environmental enrichment for singly housed animals will be provided unless an exception is granted by the IACUC.

Methods

Phylogenetic Analysis
Candidate sequences were identified by using the NCBI Protein BLAST tool [55]. For F proteins, each of NP_112026.1, NP_047111.2, and *UUV47205.1* were used as input sequences. For G proteins, both of NP_112027.1and NP_047112.2 were input. A distance matrix was generated for all selected sequences by using the Clustal Omega tool [25]. The final tree was built as a UPGMA based on the distance matrix using the Tree Viewer software [56].

Protein Production
For HNV proteins, human codon-optimized DNA sequences were synthesized by GeneImmune Biotechnology and cloned into a pαH plasmid vector. Fusion protein ectodomain sequences (**Supplemental Data 1**) were followed at the C-terminal end by a foldon trimerization domain, HRV3C protease site, 8x his tag, and a twin-strep tag (IBA Lifesciences) Attachment protein head domains (**Supplemental Data 2**) had an artificial secretion signal, secrecon [43], added to the N-terminal end and an 8x his tag added to the C-terminal end. HNV protein plasmids were transiently transfected into HEK 293 cell lines, either Expi293 or 293F (Thermo Fischer

Scientific) (**Supplemental Table 1**). The cells were allowed to incubate at 37ºC, 8% CO_2 and the supernatant was harvested 5 days post-transfection.

Fusion proteins were purified by first sequestering free biotin by applying BioLock blocking solution (IBA) and then using the Strep-Tactin affinity chromatography resin (IBA). Eluted protein was then further purified using a Superose 6 Increase 10/300 GL size-exclusion column (Cytiva) with PBS buffer at pH 8. The elution from SEC was spin concentrated with 100kDa filters and flash frozen with liquid nitrogen.

Attachment protein head domains were purified first by sequestering EDTA by adding 2mg of Cobalt Chloride per 1L of supernatant. TALON Cobalt Resin (Takara) was used for purification, with 20mM HEPES and 150mM NaCl at pH 8 as a buffer, adding 50mM Imidazole for elution. Proteins were further purified with a Superdex 75 Increase 10/300 GL SEC column (Cytiva), again using the HEPES-NaCl buffer. The elution from SEC was spin concentrated with 10kDa filters and flash frozen with liquid nitrogen.

Antibody human codon-optimized DNA sequences were synthesized by GeneImmune Biotechnology and cloned into pVRC8400 plasmid vectors. Heavy chains were designed to begin with the beginning of the antibody sequence and continue through the human IgG1 Fc region. Light chains included both the VL and CL domains. Depending on the sequence source, some antibody sequences either utilized mouse or human VL and VH domains. Mouse anti-NiV-F fabs were designed to end after the mouse CH1 domain and had an 8x his tag added to the C-terminal end, as described previously [38].

Plasmids encoding the heavy and light chains of the antibodies IgG 1E5, nAH1.3, 4G5, and DH851.3 were co-transfected into Expi293 cells at a 1:1 ratio, using the ExpiFectamine 293 Transfection Kit (Thermo Fisher Scientific, Cat. No. A14525), according to the manufacturer's instructions. Following a six-day incubation period post-transfection, the cells were harvested, and the culture supernatant was clarified by filtration through a 0.2 µm PES membrane filter to remove cellular debris. IgG purification was subsequently carried out using Protein A affinity chromatography. The filtered supernatant was passed through the Protein A affinity column (Pierce™ Recombinant Protein A Agarose Cat No. 20334), and the bound IgG was eluted using IgG elution buffer by adjusting pH via Tris Base. Eluted fractions were pooled and concentrated with Centricon-70 30 kDA filter (Millipore Sigma) followed by size-exclusion chromatography using a HiLoad 16/600 Superdex 200pg column (Cytiva), pre-equilibrated with 1× phosphate-buffered saline (PBS, pH 7.4) containing 0.02% sodium azide to prevent microbial contamination. The purified IgG was collected, concentrated, and analyzed for purity and integrity by SDS-PAGE.

Anti-NiV-F fabs were purified using the same Cobalt resin-based procedure as described above for head domains.

To generate Fab fragments for 22F5 mouse IgG and 1C8 mouse IgG, 2 mg of each IgGs were passed through Zeba Spin Desalting Columns (ThermoFisher), then washed with digestion buffer containing 35 mg cysteine. IgGs were incubated with 0.250 ml of the 50% slurry of Immobilized Papain resin (ThermoFisher) followed by incubation with digestion buffer containing 35 mg cysteine at 37 °C overnight in microcentrifuge tubes. The digested Fabs were separated from the Immobilized Papain by centrifuging at 5000 x g for 1 minute, followed by washing with digestion buffer. After digestion, the mixture was subjected to protein A affinity chromatography (ThermoFisher) then monitored by SDS-PAGE to confirm Fab generation. To further purify the Fab fragments, Size-Exclusion Chromatography (SEC) was used for purification of the Fab fragments using Superose 6 Increase 10/300 column GL (Cytiva). The sample was loaded onto the column at a volume of 0.5 mL and the elution was carried out with PBS at a flow rate of 0.5 mL/min. Fractions were collected, yielding a final concentration of 3.2 mg/mL and 4.2 mg/ml for 22F5 Fab and 1C8 Fab, respectively. The purified Fab was stored at −80°C.

The ELISA control antibody Ab82 was purified as described previously [57].

DSF

Samples were analyzed using a Prometheus Panta (Nanotemper Technologies). All samples were diluted to 0.125mg/mL using PBS, pH 8. Roughly 10µL was added to glass capillaries in triplicate for measurements. The instrument was programmed to measure 350nm and 330nm fluorescence from 35.0-95.0ºC using a 6.0ºC/min temperature gradient. All curves reported represent averages from the triplicate readings.

BLI

Bio-layer interferometry (BLI) was employed to assess the binding between fusion proteins and murine monoclonal antibodies (mAbs) or fragment antigen-binding regions (Fabs). The analysis was performed using an Octet RH16 (Octet RED384) system (Sartorius) at 25°C with kinetics buffer (Sartorius) as the running buffer. HNV-F proteins with the twin-strep tag at 6 µg/mL were captured onto Streptavidin (SA) Biosensors (Sartorius) for 300s, followed by a 60s wash in kinetics buffer. The biosensors were then exposed to 20 µg/mL mAbs or 50 µg/mL Fabs for 300s to allow for association with the fusion protein. Dissociation in kinetics buffer was monitored for an additional 300s. Sensor data was analyzed using Octet Analysis Studio (Sartorius), with reference biosensor and zero-analyte control sensorgrams subtracted from the raw data to obtain the curves corresponding to specific binding events.

SPR

Binding and kinetic assays were perfomed by Surface Plasmon Resonance (SPR) on a T-200 Biacore system (Cytiva) at 25°C. All the binding and kinetic experiments were carried out in the

HBS-EP+ (10 mM HEPES, pH 7.4, 150 mM NaCl, 3 mM EDTA and 0.05% surfactant P-20) running buffer. For the binding assay, anti-Fc mouse SPR chip were used to capture the 200nM of 22F5 IgG at 10ul/min flow rate and 50nM of LayV-1-F ectodomains as analyte were run at 30ul/min flow rate and association and dissociation were monitor at 120 seconds. In the competitive assay, 200nM of 4G5 IgG was captured on anti-Fc human SPR chip at 10ul/min and analyte as 22F5 Fab with LayV-1-F ectodomain complex in 5:1 ratio was flowed at 30ul/min flow rate for 60 seconds. SPR SA chip was used to capture the 50nM of the LayV-1-F ectodomains at 5ul/min flow rate by biotin tag for kinetics analysis. 22F5 Fab 50nM with two-fold dilution (50nM, 25nM, 12.5nM, 6.25nM and 3.12nM) as analyte was flowed with 50ul/min flow rate and association and dissociation time were analyzed at 60 and 240 seconds, respectively. For the reference curves, we kept the same parameter in reference flow cells without protein capture. The sensorgrams were blank corrected and analyzed in the Biacore T-200 evaluation software.

Negative-Stain Electron Microscopy

Frozen samples from -80 °C were thawed at room temperature for 5 minutes. Samples were then diluted to 40 µg/mL with 0.02 g/dL Ruthenium Red in HBS (20 mM HEPES, 150 mM NaCl pH 7.4) buffer containing 8 mM glutaraldehyde. After 5-minute incubation, glutaraldehyde was quenched by adding sufficient 1M Tris stock, pH 7.4, to give 80 mM final Tris concentration and incubated for 5 minutes. Quenched sample was applied to a glow-discharged carbon-coated EM grid for 10-12 seconds, blotted, consecutively rinsed with 2 drops of 1/20X HBS, and stained with 2 g/dL uranyl formate for 1 minute, blotted and air-dried. Grids were examined on a Philips EM420 electron microscope operating at 120 kV and nominal magnification of 49,000x, and images were collected on a 76 Mpix CCD camera at 2.4 Å/pixel. Images were analyzed by 2D class averages using standard protocols with Relion 3.0 [58].

Mass Photometry

A TwoMP (Refeyn) mass photometer was used for analysis. The AcquireMP software was used for calibration and data acquisition. Calibration was performed with a three-point curve consisting of β-amylase, Apoferritin, and Thyroglobulin. Samples were diluted to 40 nM in PBS. To each well, 10µL of PBS was added to focus the system before adding 10µL of sample and mixing for a final concentration of 20 nM. Recorded data for 60s, with the machine counting individual adsorption events. The DiscoverMP processing software (Refeyn) was used to assign molecular weights to individual events.

Mouse Immunizations

Immunization in $V_H1\text{-}2R^{JH2}/V\kappa1\text{-}33R^{CS\Delta/hTdT}$ (SE13) mice [59] were intramuscularly immunized with LayV-F_WT (25 mcg/animal) and NiVop8 (25 mcg/animal) adjuvanted with GLA-SE (IDRI EM-082, 5 mcg/animal) at weeks 0, 4, 15 and 21. Serum titers were monitored by ELISA

as described below. Mice with high-binding antibody titers were selected for the subsequent spleen cell fusion experiments.

Hybridoma cell line generation and monoclonal antibody production

Mice were boosted with the indicated priming antigen 3 days prior to fusion. Spleen cells were harvested and fused with NS0 murine myeloma cells using PEG1500 to generate hybridomas [37]. After 2 weeks, supernatant of hybridoma clones were collected and screened by binding ELISA as described below. Hybridomas that secreted LayV-F reactive antibodies were cloned by limiting dilution until the phenotypes of all limiting dilution wells were identical. IgG mAbs were purified by protein G purification.

Indirect binding ELISA method

ELISA assays were utilized for three purposes, measuring mouse serum antibody titers, screening for antigen specific hybridoma B cell clones, and characterizing the binding of monoclonal human IgG and mouse IgG. 384 well ELISA plates (Costar #3700) were coated with 2 µg/mL streptavidin (Thermo Fisher Scientific Inc. Cat. No. S-888) or protein antigen in 0.1M sodium bicarbonate overnight at 4 ºC. Plates were washed with PBS/0.1% Tween-20 and blocked for one hour with assay diluent (PBS containing 4%(w/v) whey protein/15% Normal Goat Serum/0.5% Tween-20/ 0.05% Sodium Azide). Streptavidin coated plates were washed and followed by 10 µL Twin-Strep-tag® (IBA Lifesciences GmbH) protein at 2 µg/mL in assay diluent for one hour. All plates were then washed and samples added in 10µl volumes with dilution scheme and secondary specifics depending on sample type as follows. Mouse serum was titrated from 1 to 30 in three-fold steps and incubated for 1 hour. Plates were washed and followed with 10µL goat anti-mouse IgG-HRP secondary antibody diluted 1:10,000 (Southern Biotech #1030-05). Hybridoma supernatants were added undiluted as a single well for 1 hour, washed and followed with 10 µL goat anti-mouse IgG-HRP. Monoclonal mouse IgG and Human IgG were diluted to 100 µg/mL, titrated in three-fold steps and incubated for 1 hour. Plates were washed and followed with 10 µL goat anti-Mouse IgG Fab-HRP diluted 1:10,000 (Southern Biotech 1015-05) for mouse IgGs or 10 µL of goat anti-human IgG Fab-HRP diluted 1:15,000 (Jackson Immunoresearch, 109-035-097) for human IgG monoclonal mAbs. After 1 hour, plates were washed again and detected with 20 µL SureBlue Reserve (Seracare 5120-0081) for 15 minutes. Reactions were stopped with the addition of 20 µL HCl stop solution. Plates were read at 450 nm using a Spectramax 384 plus.

Cryo-EM grid Preparation and Imaging

To prepare grids, UltaFoil R1.2/1.3 (Cu, 300-mesh; Electron Microscopy Sciences, PA) grids were glow discharged for 15 seconds at 15 mA using a PELCO easiGlow Cleaning System (Ted Pella Inc). The final concentration of purified AngV-F and LayV-1-F ectodomains was maintained as 1.5 mg/ml in PBS. 0.005 % (w/v) of n-dodecyl β-D-maltoside (DDM) was added to the AngV-F samples and 0.005% (w/v) DDM and 2% glycerol to the LayV-1-F samples

before applying to the grid to prevent the protein form adhering to the air-water interface. 3.0 µL of the protein sample was applied to the grid and incubated for 30 seconds at >95% humidity. Excess protein was blotted away for 2.5 seconds before being plunge frozen into liquid ethane using Leica EM GP2 plunge freezer (Leica Microsystems). Cryo-grids were imaged using either an FEI Titan Krios (Thermo Scientific) or Tundra cryo TEM (Thermo Scientific).

The Krios was equipped with K3 camera (Gatan) at 81k magnification, operated at 300kV. For AngV-F, approximately 24,000 micrographs were collected at nominal defocused range between -2.8 to -1.7 µm with a dose range of 63-65 e/A^2. For LayV-1-F, 9,715 micrographs were collected at a nominal defocused range between -2.8 to -1.2 µm with a dose range of 58.6-60.6 e/A^2.

The Tundra was equipped with CETA-F camera (Thermo Scientific™) at 180k magnification, operated at 100kV. Approximately 4,000 micrographs were collected at nominal defocused range between -3.0 to -1.5 µm with dose range of 23-25 e/A2.

The structural data used to determine the AngV-F hexameric lattice structure were collected using the Tundra TEM. All other structural data was collected using the Titan Krios TEM.

Cryo-EM Data Processing: Cryo-EM image quality was monitored on-the-fly during data collection using automated processing routines. Data processing was carried out using cryoSPARC [60,61]. Micrographs were curated through contrast transfer function (CTF) where greater than >30 Å were discarded. Automated blob picker was used to assign the particle position. Different box sizes are utilized to accommodate the varying dimensions of specific molecular assemblies. For instance, monomer particles are extracted using a box size of 320 pixels, while dimer particles require a larger box size of 512 pixels. For even larger molecular assemblies, such as hexamer lattices, a box size of 1024 pixels was employed. Following particle extraction, multiple rounds of 2D classification was performed to remove junk. *ab-initio* 3D reconstruction was used to create 3D reconstructions and poor-quality particles were discarded after through heterogenous refinement. Final resulting volumes were subjected to non-uniform refinement to build a high-resolution 3D reconstruction. Phenix [62], Coot [63], Pymol, nextPYP [64] and ChimeraX [65] were used for model building and refinement.

X-ray crystallography.

Crystals were grown via the vapor diffusion method at room temperature in siting drop well format. GakV-2-G head domain protein crystals were formed by mixing 0.20 µL of protein at 15 mg/mL with 0.20 µL of well solution containing 0.2 M magnesium sulfate heptahydrate, 20% w/v PEG3350 at pH 7.1. Crystals appeared within one week.

Cryoprotection of the crystals was done by mixing glycerol to a final concentration of 50% with the well solution. Crystals were moved to the drop containing the well solution and glycerol and

soaked for one minute then flash-cooled in liquid nitrogen and data were collected at the NSLS-II, Beamline 19-ID at 100 K and a wavelength of 1 Å. All data were indexed and integrated using iMosflm [66] and scaled using AIMLESS.
Phaser [67] in the PHENIX suite was used to perform molecular replacement using a model generated from the AlphaFold3 prediction server [68] using the GakV-2-G head domain sequence. Phenix.refine [69] was used for data refinement, and manual refinement was done in Coot [63].

Figures

Figure 1: Selection and Classification of Henipavirus Species. Upper Left: A flowchart that describes the origin of sequences used and the filtering mechanisms, Genome and F/G sequence completion, used to assemble a list of non-redundant F and G proteins. **Tree:** A phylogenetic tree based on full-genome nucleotide sequences of the selected strains. Horizontal branch lengths are to scale to represent sequence divergence. Hollow boxes are drawn to represent broad groupings within the tree: Classical/Bat Henipaviruses, green, Parahenipaviruses, orange, Nipah viruses, red, Hendra viruses, dark blue. Solid highlight boxes are drawn to represent specific species or subspecies groupings: Nipah-Bangladesh, purple, Nipah-India, orange, Nipah-Malaysia, yellow, Nipah-Cambodia, salmon, Hendra-alpha, blue, Hendra-beta, green, Cedar virus, cyan, Langya virus, tan, Shrew Henipavirus 5, grey, Daeryong virus, red, Gamak virus, dark blue. The tips of each branch are labelled with the strain names and associated gen bank accession number. Each strain is colored based on which type of protein was purified: F protein only, blue, G protein only, red, F and G, yellow.

Figure 2: Purification and Antigenic Characterization of Henipavirus Fusion Proteins. (A) Table of protein yields and sequence similarities for each member of the F protein panel. Percent identities and similarities of the amino acid sequences of the full-length F proteins compared to NiV-M1-F are listed. GhV-F similarity values were excluded due to the likely inaccurate N-terminal sequence in the deposited sequence. Main peak yield is based on the final calculated protein yield after SEC. Main peak percentage of total is based on SEC area under the curve of the main peak as a fraction of all collected SEC fractions. Color scheme for HNV species continued from Figure 1. **(B)** Representative SEC profiles of HNV-F Proteins with peaks annotated. The oligomer-of-trimer peak, sometimes called the “middle peak,” and the main trimer peaks are collected separately. NiVop08 is a pre-fusion stabilized construct [36] not part of the panel but purified for a vaccination study and for comparison to wild-type proteins. **(C)** Mass photometry analysis of NiV-M1-F SEC components. Each individual measurement is assigned a calculated molecular weight by the instrument software, and individual results are binned in 10 kDa bin sizes. The graph reports the count of events for each molecular weight bin with the main, middle, and mixed peak data overlayed. **(D)** Graphical timeline of a vaccination study in mice using HNV-F antigens. Each group was vaccinated with LayV-1-F and boosted with LayV-1-F or NiV antigens as marked. Blood draws and antigen injections occurred at the indicated

weeks after the start of the study. **(E)** Heatmap of antibody binding to HNV-F proteins, on a spectrum of blue for low binding to red for high binding. The left-side graph is based on ELISA data with immobilized F proteins, antibody analytes, and anti-Human Fab-conjugated HRP as a reporter. The values for each combination are based on the log of the area under the curve for the plotted absorbance over increasing concentrations. The right-side graph is based on Bio-layer interferometry (BLI) with immobilized F proteins and direct measurement of analyte to the BLI sensors, either mouse Fabs or mouse IgG (22F5 and 1C8). The values for each combination are based on the maximum wavelength shift measured during the association phase of the BLI experiment. HNV-F antigens colored as in (A), antibody analytes colored as follows: green, known anti-NiV-F binding; salmon, known anti-LayV-F binding; gray, anti-influenza HA negative control antibody (ELISA only). **(F)** Cryo-EM map of 22F5 Fab bound to post-fusion LayV-1-F with domains colored and labelled. **(G)** SPR binding data reporting response units over time of several LayV-F ectodomain constructs binding to 22F5 IgG.

Figure 3: Thermostability of Henipavirus F Proteins. **(A)** DSF profiles of a set of HNV proteins showing the first derivative of 350/330nm emission ratio as a function of temperature. Dotted lines of the same color represent the DSF profile of the same proteins after being incubated at 60ºC for 90 minutes. **(B)** Two-dimensional NSEM class averages of a portion of the heat-treated samples from A. **(C)** DSF profiles of diverse HNV-F species. Sections of the profile are noted indicating the conformational steps occurring at those temperatures, with highlighting of peaks of interest. **(D)** A subset of DSF profiles from C, containing only those species that do not demonstrate any noticeable conformational conversion peak. Below are representative two-dimensional NSEM class averages for select members of this subset. **(E)** A subset of DSF profiles from C, containing only those species that demonstrate some presence of the standard conformational conversion signal. Below are representative two-dimensional NSEM class averages for select members of this subset. **(F)** A subset of DSF profiles from C, containing only those species that demonstrate a profile notably different from the typical HNV-F species. Below are representative two-dimensional NSEM class averages for the two members of this subset.

Figure 4. Cryo-EM Structure of the AngV-F Ectodomain. (A) Three different particle populations in the AngV-F and their representative 2D classes. Cryo-EM reconstruction of top and side view shown for AngV F trimer (left), dimer-of-trimers (middle) and hexameric lattice (right). **(B)** Comparison of the AngV-F, LayV-F and NiV-F proteins. Top. A slice through the F-protein visualizing the central cavity and the placement of the stalk region within the F-protein. Bottom. A top-down view showing measurements. **(C)** Zoom-in view of the stalk region featuring amino acid present at the start of helix bundle (top) and the electrostatics of the region (bottom), with sequence alignment at the local region between each of the three species below. The HRB start residue is marked red in the sequence alignment. **(D)** AngV-F cryo-EM structure colored by protomer with the glycans shown as spheres and colored by element. Glycans at the

different regions (Apex, non-canonical NNV, stalk) are marked with arrows. **(E)** The dimer-of-trimer interaction interface is shown in surface representation (top), along with their molecular-level interactions and zoomed-in views showing details of the interface contacts (bottom).

Figure 5: Characterization of Henipavirus G Proteins. (A) Representative SEC elution profile for G head domains with collected fractions designated. **(B)** Table of key statistics for each member of the G protein panel. Percent identities and similarities of the full amino acid sequence compared to NiV-M3-G are listed. Yields for species are marked gray and N/A for species that had a unique G ectodomain but not a unique head domain and therefore were not purified as part of this panel. Main peak yield is based on the final calculated protein yield after SEC. Color scheme for HNV species continued from Figure 1. **(C-F)** DSF analysis of G head domains, reporting the first derivative of 350/330nm ratio as a function of temperature. The panel's results are split based on their phylogeny and general DSF profile: C, *Henipavirus* proteins with a single DSF peak; D, *Parahenipavirus* proteins with a single DSF peak; E, proteins with alternative peak shapes (broad or negative); F, proteins with double peaks. In (C-D), a side plot of inflection temperatures is shown for each species in the associated graph. In (E-F), inflection temperatures are drawn directly on the plot. **(G)** Heatmap of antibody binding to G head domains, on a spectrum of blue for low binding to red for high binding. The graph is based on ELISA data with immobilized G head domains, antibody analytes, and anti-Human Fab-conjugated HRP as a reporter. The values for each combination are based on the log of the area under the curve for the plotted fluorescence over increasing concentrations. G head domain antigens colored as in (A), antibody analytes colored as follows: green, known anti-NiV-G binding; salmon, known anti-LayV-G binding; gray, anti-influenza HA negative control antibody.

Figure 6: Crystal structure of GakV G Protein Head Domain. **(A)** Three views of the GakV-G head domain colored in rainbow. **(B)** Superposition of GakV-G head domain on Ephrin-bound NiV-G head domain (PDB: 2VSM, [49]). GakV-G is colored gray, NiV-G cyan and Ephrin light blue. Ephrin is shown as a transparent surface. The GakV-G glycan 506 is shown as blue spheres. **(C)** Overlay of GakV-G head domain on the G head domains from other HNV species. PDB Codes: LayV-G, 8K80 (Unpublished); MojV-G, 5NOP [70]; GhV-G, 4UF7 [71]; NiV-G, 3D11 [72]; HeV-G, 6VY4 [16]; CedV-G, 6P72 [73]. **D.** Top. Two views of the GakV-G head domain colored in rainbow with the N terminal Asp to His218 shown as a transparent surface. The additional minidomain defined in the GakV-2-G head structure is highlighted within a dashed oval. Bottom. G head domains of LayV and NiV. **E.** Topology diagram of GakV-G protein with sequence alignment of select HNV species below.

References

1 Rima, B. *et al.* ICTV Virus Taxonomy Profile: Paramyxoviridae. *J Gen Virol* **100**, 1593-1594 (2019). https://doi.org/10.1099/jgv.0.001328

2 Zhang, X.-A. *et al.* A Zoonotic Henipavirus in Febrile Patients in China. *New England Journal of Medicine* **387**, 470-472 (2022). https://doi.org/10.1056/NEJMc2202705

3 Simmonds, P. *et al.* Changes to virus taxonomy and the ICTV Statutes ratified by the International Committee on Taxonomy of Viruses (2024). *Arch Virol* **169**, 236 (2024). https://doi.org/10.1007/s00705-024-06143-y

4 Chua, K. B. *et al.* Nipah Virus: A Recently Emergent Deadly Paramyxovirus. *Science* **288**, 1432-1435 (2000). https://doi.org/doi:10.1126/science.288.5470.1432

5 Murray, K. *et al.* A Morbillivirus that Caused Fatal Fisease in Horses and Humans. *Science* **268**, 94-97 (1995). https://doi.org/doi:10.1126/science.7701348

6 Murray, K. *et al.* in *Emerging Infections 1* 43-58 (1997).

7 Lee, S.-H. *et al.* Discovery and Genetic Characterization of Novel Paramyxoviruses Related to the Genus Henipavirus in Crocidura Species in the Republic of Korea. *Viruses* **13**, 2020 (2021).

8 Vanmechelen, B. *et al.* The characterization of multiple novel paramyxoviruses highlights the diverse nature of the subfamily Orthoparamyxovirinae. *Virus Evolution* **8** (2022). https://doi.org/10.1093/ve/veac061

9 Horemans, M., Van Bets, J., Joly Maes, T., Maes, P. & Vanmechelen, B. Discovery and genome characterization of six new orthoparamyxoviruses in small Belgian mammals. *Virus Evolution* **9** (2023). https://doi.org/10.1093/ve/vead065

10 Zhang, J. T. *et al.* Decoding the RNA viromes in shrew lungs along the eastern coast of China. *NPJ Biofilms Microbiomes* **10**, 68 (2024). https://doi.org/10.1038/s41522-024-00543-3

11 May, A. J. & Acharya, P. Structural Studies of Henipavirus Glycoproteins. *Viruses* **16**, 195 (2024).

12 White, J. M., Delos, S. E., Brecher, M. & Schornberg, K. Structures and mechanisms of viral membrane fusion proteins: multiple variations on a common theme. *Crit Rev Biochem Mol Biol* **43**, 189-219 (2008). https://doi.org/10.1080/10409230802058320

13 Iorio, R. M., Melanson, V. R. & Mahon, P. J. Glycoprotein interactions in paramyxovirus fusion. *Future Virol* **4**, 335-351 (2009). https://doi.org/10.2217/fvl.09.17

14 Wang, Z. *et al.* Architecture and antigenicity of the Nipah virus attachment glycoprotein. *Science* **375**, 1373-1378 (2022). https://doi.org/doi:10.1126/science.abm5561

15 Dang, H. V. *et al.* Broadly neutralizing antibody cocktails targeting Nipah virus and Hendra virus fusion glycoproteins. *Nature Structural & Molecular Biology* **28**, 426-434 (2021). https://doi.org/10.1038/s41594-021-00584-8

16 Dong, J. *et al.* Potent Henipavirus Neutralization by Antibodies Recognizing Diverse Sites on Hendra and Nipah Virus Receptor Binding Protein. *Cell* **183**, 1536-1550.e1517 (2020). https://doi.org/https://doi.org/10.1016/j.cell.2020.11.023

17 Avanzato, V. A. *et al.* A structural basis for antibody-mediated neutralization of Nipah virus reveals a site of vulnerability at the fusion glycoprotein apex. *Proceedings of the*

National Academy of Sciences **116**, 25057-25067 (2019). https://doi.org/10.1073/pnas.1912503116

18 Dang, H. V. *et al.* An antibody against the F glycoprotein inhibits Nipah and Hendra virus infections. *Nature Structural & Molecular Biology* **26**, 980-987 (2019). https://doi.org/10.1038/s41594-019-0308-9

19 May, A. J., Pothula, K. R., Janowska, K. & Acharya, P. Structures of Langya Virus Fusion Protein Ectodomain in Pre- and Postfusion Conformation. *J Virol* **97**, e0043323 (2023). https://doi.org/10.1128/jvi.00433-23

20 Wang, Z. *et al.* Structure and design of Langya virus glycoprotein antigens. *Proceedings of the National Academy of Sciences* **121**, e2314990121 (2024). https://doi.org/doi:10.1073/pnas.2314990121

21 Camacho, C. *et al.* BLAST+: architecture and applications. *BMC Bioinformatics* **10**, 421 (2009). https://doi.org/10.1186/1471-2105-10-421

22 Wang, L. F. *et al.* The exceptionally large genome of Hendra virus: support for creation of a new genus within the family Paramyxoviridae. *J Virol* **74**, 9972-9979 (2000). https://doi.org/10.1128/jvi.74.21.9972-9979.2000

23 Lo, M. K. *et al.* Determination of the henipavirus phosphoprotein gene mRNA editing frequencies and detection of the C, V and W proteins of Nipah virus in virus-infected cells. *J Gen Virol* **90**, 398-404 (2009). https://doi.org/10.1099/vir.0.007294-0

24 Wang, L.-F. *et al.* Molecular biology of Hendra and Nipah viruses. *Microbes and Infection* **3**, 279-287 (2001). https://doi.org/https://doi.org/10.1016/S1286-4579(01)01381-8

25 Sievers, F. *et al.* Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol Syst Biol* **7**, 539 (2011).

26 Guo, Y. *et al.* The cryo-EM structure of homotetrameric attachment glycoprotein from langya henipavirus. *Nature Communications* **15**, 812 (2024). https://doi.org/10.1038/s41467-024-45202-5

27 Whitmer, S. L. M. *et al.* Inference of Nipah virus evolution, 1999-2015. *Virus Evol* **7**, veaa062 (2021). https://doi.org/10.1093/ve/veaa062

28 Gaudino, M. *et al.* High Pathogenicity of Nipah Virus from Pteropus lylei Fruit Bats, Cambodia. *Emerg Infect Dis* **26**, 104-113 (2020). https://doi.org/10.3201/eid2601.191284

29 Wang, J. *et al.* A new Hendra virus genotype found in Australian flying foxes. *Virol J* **18**, 197 (2021). https://doi.org/10.1186/s12985-021-01652-7

30 Marsh, G. A. *et al.* Cedar virus: a novel Henipavirus isolated from Australian bats. *PLoS Pathog* **8**, e1002836 (2012). https://doi.org/10.1371/journal.ppat.1002836

31 Drexler, J. F. *et al.* Bats host major mammalian paramyxoviruses. *Nat Commun* **3**, 796 (2012). https://doi.org/10.1038/ncomms1796

32 Madera, S. *et al.* Discovery and Genomic Characterization of a Novel Henipavirus, Angavokely Virus, from Fruit Bats in Madagascar. *Journal of Virology* **96**, e00921-00922 (2022). https://doi.org/doi:10.1128/jvi.00921-22

33 Pernet, O., Beaty, S. & Lee, B. Functional rectification of the newly described African henipavirus fusion glycoprotein (Gh-M74a). *J Virol* **88**, 5171-5176 (2014). https://doi.org/10.1128/jvi.03655-13

34 Wu, Z. *et al.* Novel Henipa-like virus, Mojiang Paramyxovirus, in rats, China, 2012. *Emerg Infect Dis* **20**, 1064-1066 (2014). https://doi.org/10.3201/eid2006.131022

35 Meier, S., Güthe, S., Kiefhaber, T. & Grzesiek, S. Foldon, the natural trimerization domain of T4 fibritin, dissociates into a monomeric A-state form containing a stable beta-hairpin: atomic details of trimer dissociation and local beta-hairpin stability from residual dipolar couplings. *J Mol Biol* **344**, 1051-1069 (2004). https://doi.org/10.1016/j.jmb.2004.09.079
36 Loomis, R. J. *et al.* Structure-Based Design of Nipah Virus Vaccines: A Generalizable Approach to Paramyxovirus Immunogen Development. *Frontiers in Immunology* **11** (2020). https://doi.org/10.3389/fimmu.2020.00842
37 Haynes, B. F., Eisenbarth, G. S. & Fauci, A. S. Human lymphocyte antigens: production of a monoclonal antibody that defines functional thymus-derived lymphocyte subsets. *Proc Natl Acad Sci U S A* **76**, 5829-5833 (1979). https://doi.org/10.1073/pnas.76.11.5829
38 Byrne, P. O. *et al.* Structural basis for antibody recognition of vulnerable epitopes on Nipah virus F protein. *Nature Communications* **14**, 1494 (2023). https://doi.org/10.1038/s41467-023-36995-y
39 Byrne, P. O. *et al.* Prefusion stabilization of the Hendra and Langya virus F proteins. *J Virol* **98**, e0137223 (2024). https://doi.org/10.1128/jvi.01372-23
40 Marcink, T. C. *et al.* Subnanometer structure of an enveloped virus fusion complex on viral surface reveals new entry mechanisms. *Science Advances* **9**, eade2727 (2023). https://doi.org/doi:10.1126/sciadv.ade2727
41 Xu, K. *et al.* Crystal Structure of the Pre-fusion Nipah Virus Fusion Glycoprotein Reveals a Novel Hexamer-of-Trimers Assembly. *PLOS Pathogens* **11**, e1005322 (2015). https://doi.org/10.1371/journal.ppat.1005322
42 Isaacs, A. *et al.* Structure and antigenicity of divergent Henipavirus fusion glycoproteins. *Nature Communications* **14**, 3577 (2023). https://doi.org/10.1038/s41467-023-39278-8
43 Barash, S., Wang, W. & Shi, Y. Human secretory signal peptide description by hidden Markov model and generation of a strong artificial signal peptide for secreted protein expression. *Biochem Biophys Res Commun* **294**, 835-842 (2002). https://doi.org/10.1016/s0006-291x(02)00566-1
44 Zhang, Q. E. *et al.* SARS-CoV-2 Omicron XBB lineage spike structures, conformations, antigenicity, and receptor recognition. *Mol Cell* **84**, 2747-2764.e2747 (2024). https://doi.org/10.1016/j.molcel.2024.06.028
45 Stalls, V. *et al.* Cryo-EM structures of SARS-CoV-2 Omicron BA.2 spike. *Cell Rep* **39**, 111009 (2022). https://doi.org/10.1016/j.celrep.2022.111009
46 Gobeil, S. M. *et al.* Structural diversity of the SARS-CoV-2 Omicron spike. *Mol Cell* **82**, 2050-2068.e2056 (2022). https://doi.org/10.1016/j.molcel.2022.03.028
47 Wang, Z. *et al.* Potent monoclonal antibody-mediated neutralization of a divergent Hendra virus variant. *Proc Natl Acad Sci U S A* **119**, e2122769119 (2022). https://doi.org/10.1073/pnas.2122769119
48 Fan, P. *et al.* A potent Henipavirus cross-neutralizing antibody reveals a dynamic fusion-triggering pattern of the G-tetramer. *Nat Commun* **15**, 4330 (2024). https://doi.org/10.1038/s41467-024-48601-w
49 Bowden, T. A. *et al.* Structural basis of Nipah and Hendra virus attachment to their cell-surface receptor ephrin-B2. *Nature Structural & Molecular Biology* **15**, 567-572 (2008). https://doi.org/10.1038/nsmb.1435

50 Yuan, P. *et al.* Structural Studies of the Parainfluenza Virus 5 Hemagglutinin-Neuraminidase Tetramer in Complex with Its Receptor, Sialyllactose. *Structure* **13**, 803-815 (2005). https://doi.org/https://doi.org/10.1016/j.str.2005.02.019
51 Hashiguchi, T. *et al.* Structure of the measles virus hemagglutinin bound to its cellular receptor SLAM. *Nature Structural & Molecular Biology* **18**, 135-141 (2011). https://doi.org/10.1038/nsmb.1969
52 Ke, Z. *et al.* Promotion of virus assembly and organization by the measles virus matrix protein. *Nature Communications* **9**, 1736 (2018). https://doi.org/10.1038/s41467-018-04058-2
53 Sibert, B. S. *et al.* Assembly of respiratory syncytial virus matrix protein lattice and its coordination with fusion glycoprotein trimers. *Nature Communications* **15**, 5923 (2024). https://doi.org/10.1038/s41467-024-50162-x
54 Wong, J. J. W. *et al.* Monomeric ephrinB2 binding induces allosteric changes in Nipah virus G that precede its full activation. *Nature Communications* **8**, 781 (2017). https://doi.org/10.1038/s41467-017-00863-3
55 Camacho, C. *et al.* BLAST+: architecture and applications. *BMC Bioinformatics* **10**, 421 (2009). https://doi.org/10.1186/1471-2105-10-421
56 Bianchini, G. & Sánchez-Baracaldo, P. TreeViewer: Flexible, modular software to visualise and manipulate phylogenetic trees. *Ecology and Evolution* **14**, e10873 (2024). https://doi.org/https://doi.org/10.1002/ece3.10873
57 Whittle, J. R. *et al.* Broadly neutralizing human antibody that recognizes the receptor-binding pocket of influenza virus hemagglutinin. *Proc Natl Acad Sci U S A* **108**, 14216-14221 (2011). https://doi.org/10.1073/pnas.1111497108
58 Zivanov, J. *et al.* New tools for automated high-resolution cryo-EM structure determination in RELION-3. *Elife* **7** (2018). https://doi.org/10.7554/eLife.42166
59 Luo, S. *et al.* Humanized V(D)J-rearranging and TdT-expressing mouse vaccine models with physiological HIV-1 broadly neutralizing antibody precursors. *Proceedings of the National Academy of Sciences* **120**, e2217883120 (2023). https://doi.org/doi:10.1073/pnas.2217883120
60 Punjani, A., Zhang, H. & Fleet, D. J. Non-uniform refinement: adaptive regularization improves single-particle cryo-EM reconstruction. *Nature Methods* **17**, 1214-1221 (2020). https://doi.org/10.1038/s41592-020-00990-8
61 Punjani, A., Rubinstein, J. L., Fleet, D. J. & Brubaker, M. A. cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination. *Nature Methods* **14**, 290-296 (2017). https://doi.org/10.1038/nmeth.4169
62 Liebschner, D. *et al.* Macromolecular structure determination using X-rays, neutrons and electrons: recent developments in Phenix. *Acta Crystallographica Section D* **75**, 861-877 (2019). https://doi.org/doi:10.1107/S2059798319011471
63 Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* **60**, 2126-2132 (2004). https://doi.org/10.1107/s0907444904019158
64 Liu, H.-F. *et al.* nextPYP: a comprehensive and scalable platform for characterizing protein variability in situ using single-particle cryo-electron tomography. *Nature Methods* **20**, 1909-1919 (2023). https://doi.org/10.1038/s41592-023-02045-0
65 Meng, E. C. *et al.* UCSF ChimeraX: Tools for structure building and analysis. *Protein Sci* **32**, e4792 (2023). https://doi.org/10.1002/pro.4792

66 Battye, T. G. G., Kontogiannis, L., Johnson, O., Powell, H. R. & Leslie, A. G. W. iMOSFLM: a new graphical interface for diffraction-image processing with MOSFLM. *Acta Crystallographica Section D* **67**, 271-281 (2011). https://doi.org/doi:10.1107/S0907444910048675

67 McCoy, A. J. *et al.* Phaser crystallographic software. *Journal of Applied Crystallography* **40**, 658-674 (2007). https://doi.org/doi:10.1107/S0021889807021206

68 Abramson, J. *et al.* Accurate structure prediction of biomolecular interactions with AlphaFold 3. *Nature* **630**, 493-500 (2024). https://doi.org/10.1038/s41586-024-07487-w

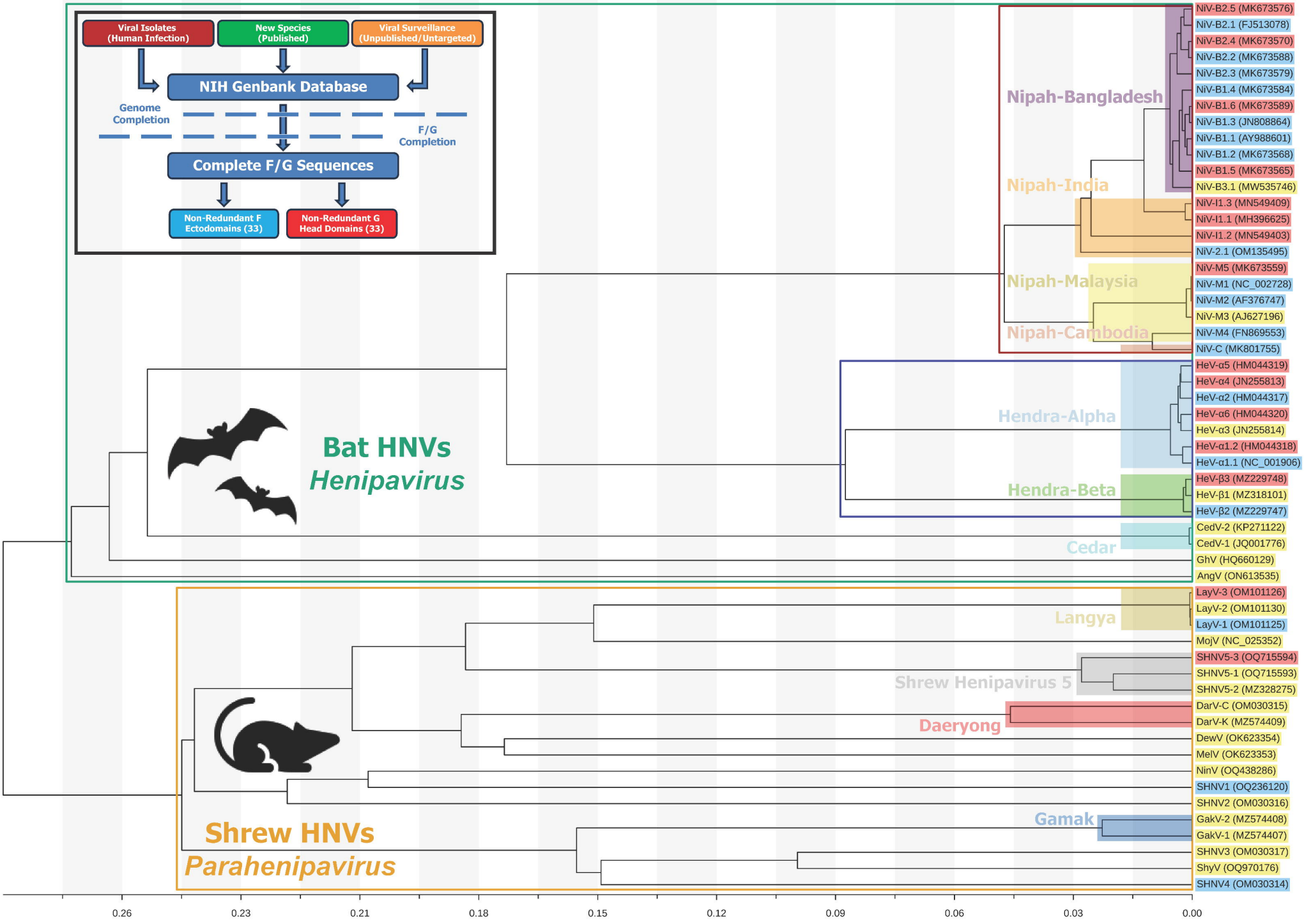
69 Adams, P. D. *et al.* PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D Biol Crystallogr* **66**, 213-221 (2010). https://doi.org/10.1107/s0907444909052925

70 Rissanen, I. *et al.* Idiosyncratic Mòjiāng virus attachment glycoprotein directs a host-cell entry pathway distinct from genetically related henipaviruses. *Nature Communications* **8**, 16060 (2017). https://doi.org/10.1038/ncomms16060

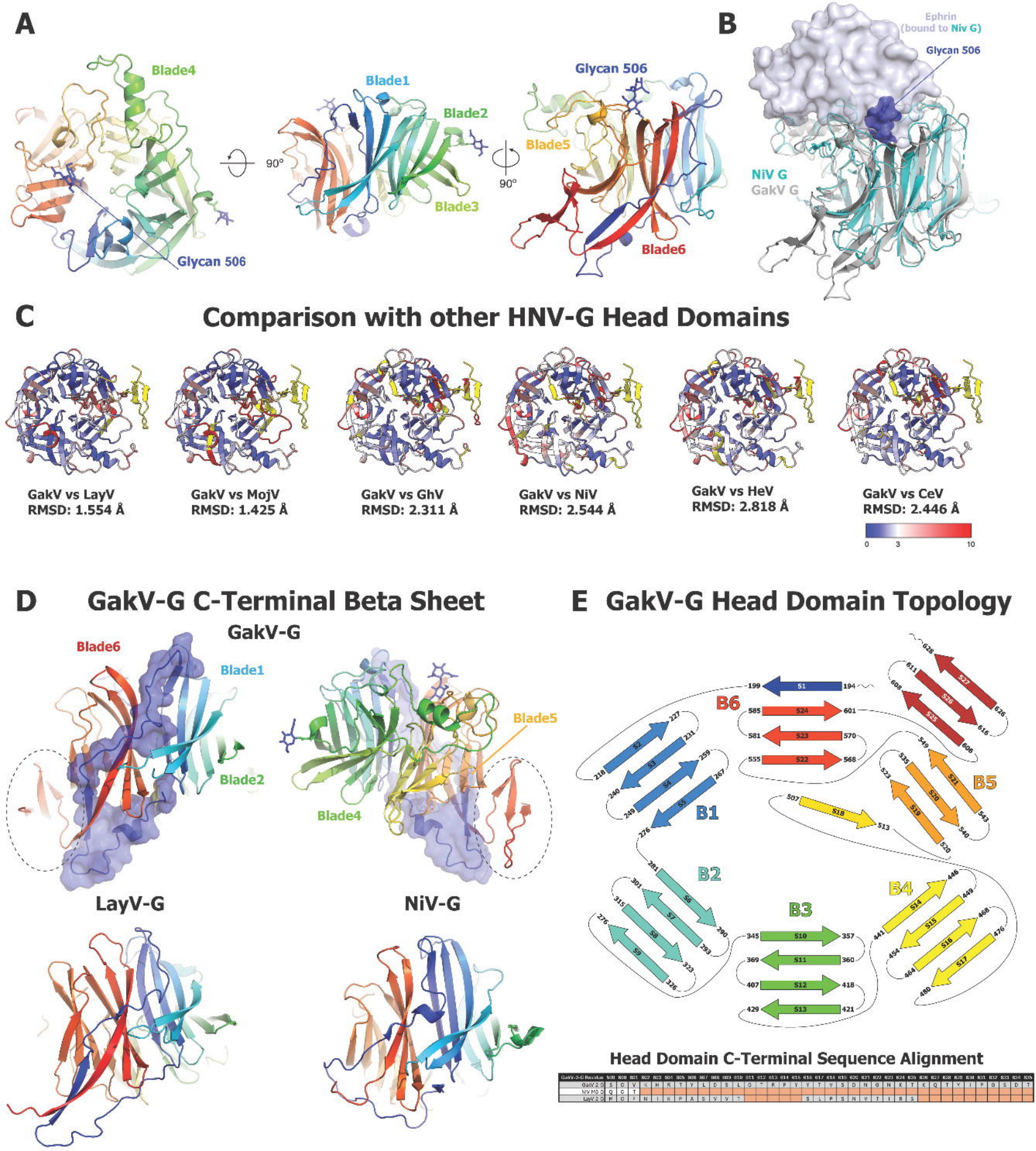
71 Lee, B. *et al.* Molecular recognition of human ephrinB2 cell surface receptor by an emergent African henipavirus. *Proceedings of the National Academy of Sciences* **112**, E2156-E2165 (2015). https://doi.org/doi:10.1073/pnas.1501690112

72 Xu, K. *et al.* Host cell recognition by the henipaviruses: Crystal structures of the Nipah G attachment glycoprotein and its complex with ephrin-B3. *Proceedings of the National Academy of Sciences* **105**, 9953-9958 (2008). https://doi.org/doi:10.1073/pnas.0804797105

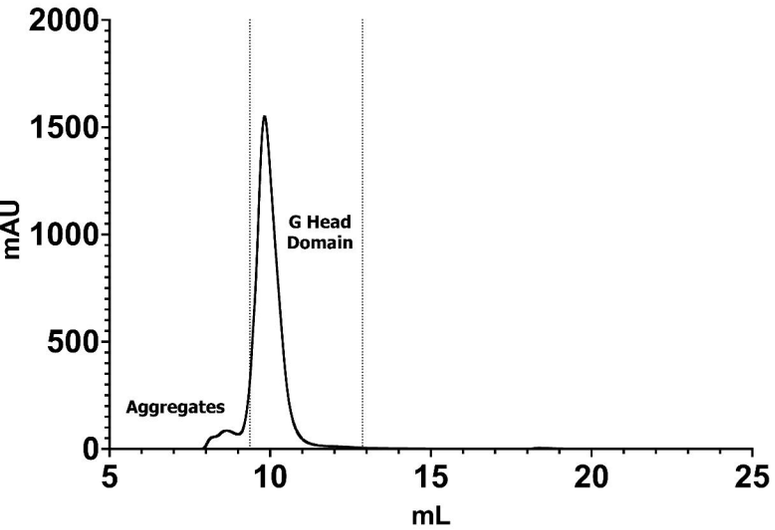
73 Laing, E. D. *et al.* Structural and functional analyses reveal promiscuous and species specific use of ephrin receptors by Cedar virus. *Proceedings of the National Academy of Sciences* **116**, 20707-20715 (2019). https://doi.org/doi:10.1073/pnas.1911773116


Viral Isolates (Human Infection)
New Species (Published)
Viral Surveillance (Unpublished/Untargeted)
NIH Genbank Database
Genome Completion
F/G Completion
Complete F/G Sequences
Non-Redundant F Ectodomains (33)
Non-Redundant G Head Domains (33)
Bat HNVs
Henipavirus
Shrew HNVs
Parahenipavirus
Nipah-Bangladesh
Nipah-India
Nipah-Malaysia
Nipah-Cambodia
Hendra-Alpha
Hendra-Beta
Cedar
Langya
Shrew Henipavirus 5
Daeryong
Gamak
NiV-B2.5 (MK673576)
NiV-B2.1 (FJ513078)
NiV-B2.4 (MK673570)
NiV-B2.2 (MK673588)
NiV-B2.3 (MK673579)
NiV-B1.4 (MK673584)
NiV-B1.6 (MK673589)
NiV-B1.3 (JN808864)
NiV-B1.1 (AY988601)
NiV-B1.2 (MK673568)
NiV-B1.5 (MK673565)
NiV-B3.1 (MW535746)
NiV-I1.3 (MN549409)
NiV-I1.1 (MH396625)
NiV-I1.2 (MN549403)
NiV-2.1 (OM135495)
NiV-M5 (MK673559)
NiV-M1 (NC_002728)
NiV-M2 (AF376747)
NiV-M3 (AJ627196)
NiV-M4 (FN869553)
NiV-C (MK801755)
HeV-α5 (HM044319)
HeV-α4 (JN255813)
HeV-α2 (HM044317)
HeV-α6 (HM044320)
HeV-α3 (JN255814)
HeV-α1.2 (HM044318)
HeV-α1.1 (NC_001906)
HeV-β3 (MZ229748)
HeV-β1 (MZ318101)
HeV-β2 (MZ229747)
CedV-2 (KP271122)
CedV-1 (JQ001776)
GhV (HQ660129)
AngV (ON613535)
LayV-3 (OM101126)
LayV-2 (OM101130)
LayV-1 (OM101125)
MojV (NC_025352)
SHNV5-3 (OQ715594)
SHNV5-1 (OQ715593)
SHNV5-2 (MZ328275)
DarV-C (OM030315)
DarV-K (MZ574409)
DewV (OK623354)
MelV (OK623353)
NinV (OQ438286)
SHNV1 (OQ236120)
SHNV2 (OM030316)
GakV-2 (MZ574408)
GakV-1 (MZ574407)
SHNV3 (OM030317)
ShyV (OQ970176)
SHNV4 (OM030314)
0.26
0.23
0.21
0.18
0.15
0.12
0.09
0.06
0.03
0.00

Crystal Structure of GakV-G Head Domain



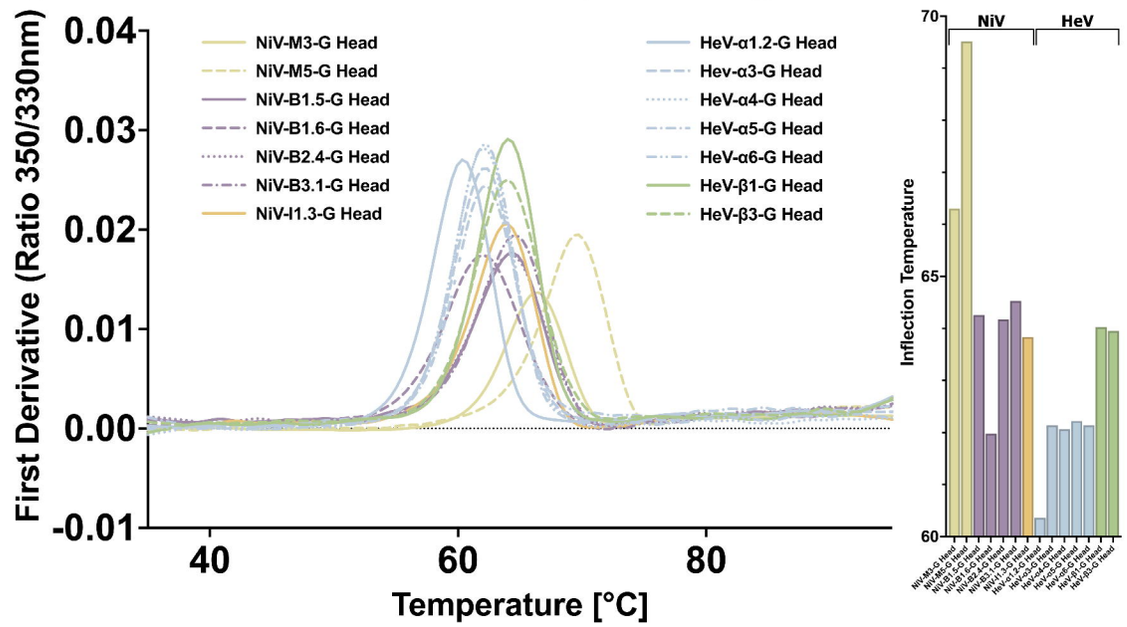
A G Head Domain SEC



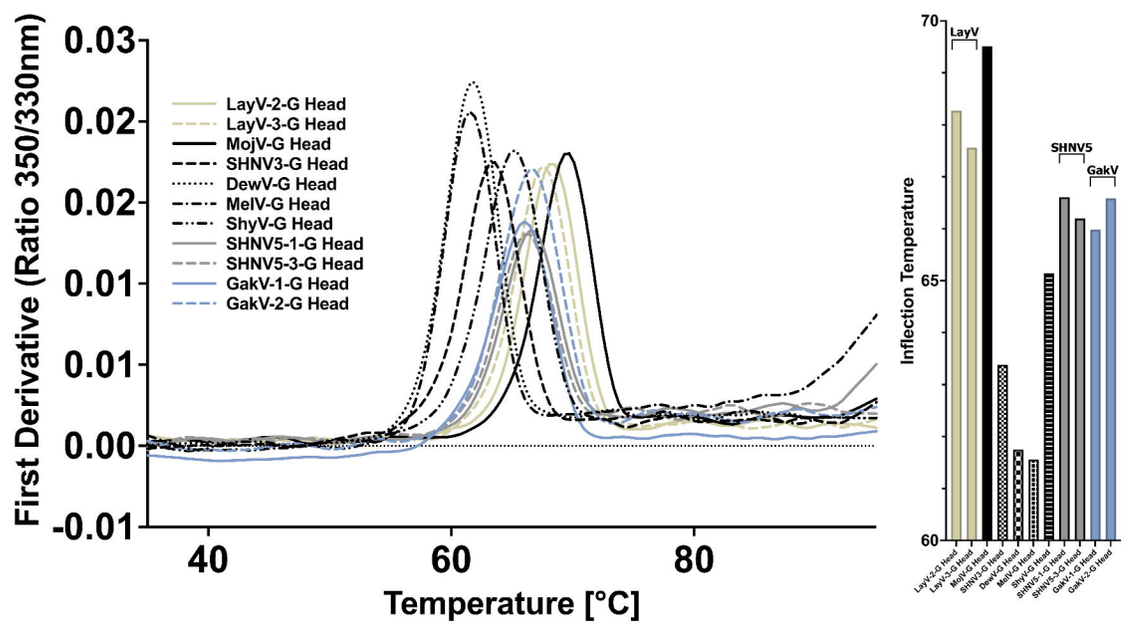
B G Protein Similarity and Yield

Name	Identity to NiV-M3-G	Similarity to NiV-M3-G	Head Domain Yield/Liter (mg/L)
NiV-M3-G	N/A	N/A	26.52
NiV-M5-G	99.5	99.5	18.75
NiV-B2.5-G	95.85	97.01	N/A
NiV-B1.5-G	95.68	96.84	26.7
NiV-B3.1-G	95.51	96.84	26.7
NiV-B2.4-G	95.35	96.84	31.2
NiV-B1.6-G	95.35	96.51	25.2
NiV-I1.3-G	95.18	96.68	13.45
HeV-β3-G	78.97	87.58	15.4
HeV-β1-G	78.64	87.25	12.75
HeV-α6-G	78.31	87.58	17
HeV-α5-G	78.31	87.58	16.15
HeV-α3-G	78.31	87.58	12.75
HeV-α1.2-G	78.15	87.58	12.75
HeV-α4-G	77.98	87.43	6.95
CedV-2-G	29.02	44.16	N/A
CedV-1-G	28.86	44.01	16.07
GhV-G	24.92	40.97	6.75
MojV-G	19.57	36.02	7.1
LayV-2-G	19.28	35.46	17.5
LayV-3-G	19.13	35.46	14
DarV-K-G	18.97	38.72	5.2
DarV-C-G	18.97	38.72	2.45
SHNV5-1-G	18.88	34.14	35.3
SHNV5-3-G	18.73	33.84	19.95
MelV-G	18.73	36.69	6.4
DewV-G	18.04	38.26	7.75
SHNV3-G	17.23	35.52	14.6
GakV-1-G	17.07	34.6	14.8
GakV-2-G	16.92	34.45	19.3
ShyV-G	16.92	35.06	6.55
NinV-G	16.89	34.39	3.085
AngV-G	16.86	31.14	8.15
SHNV2-G	16.28	36.74	3.25

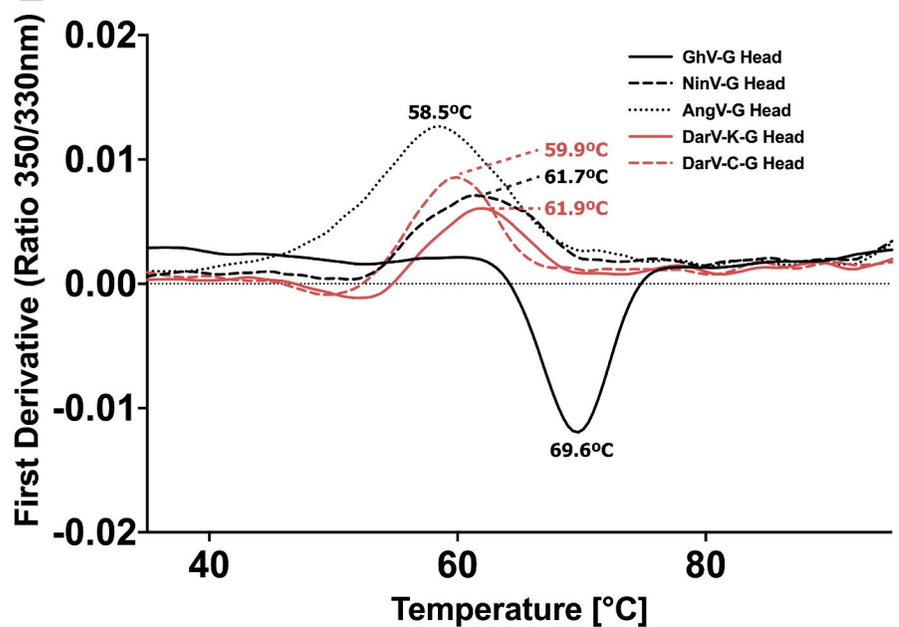
C G Head Domain Henipavirus Single Peak



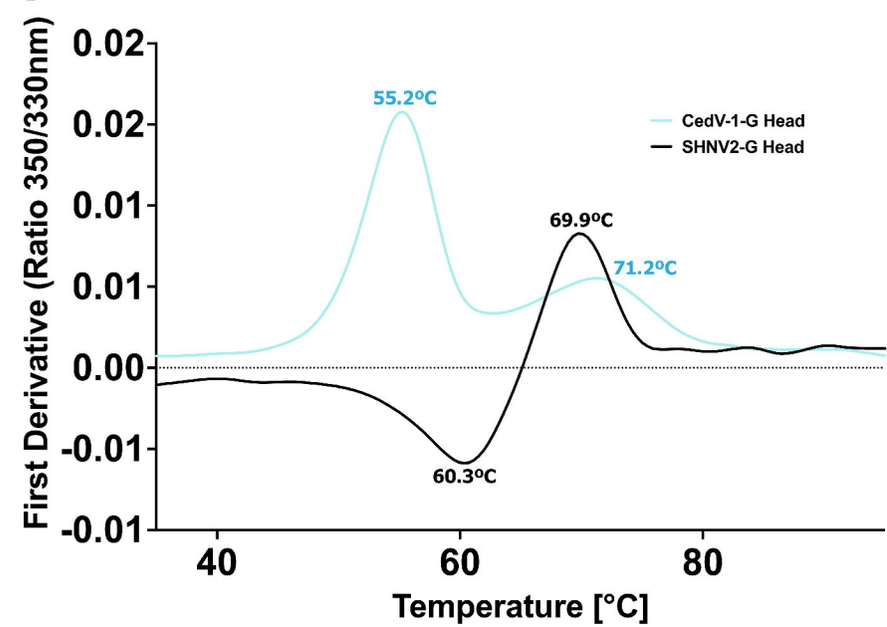
D G Head Domain Parahenipavirus Single Peak



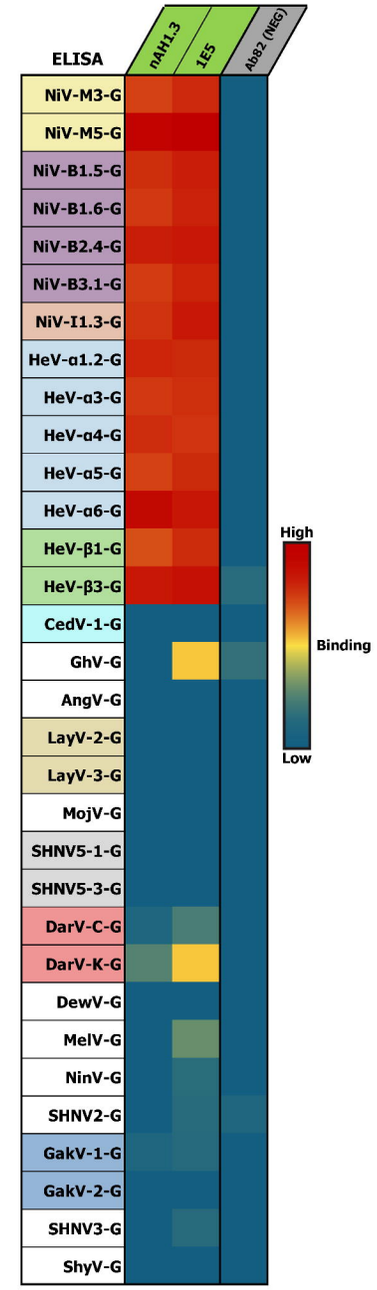
E G Head Domain Alternate Peak

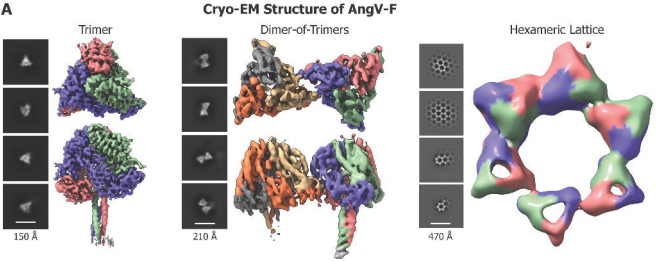


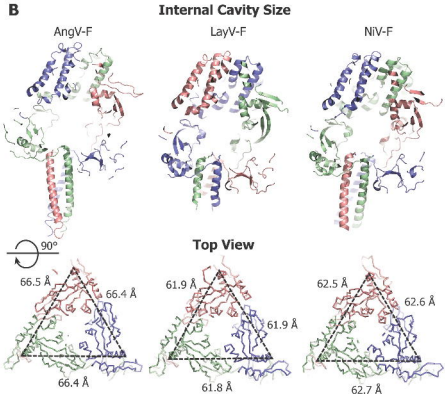
F G Head Domain Double Peak

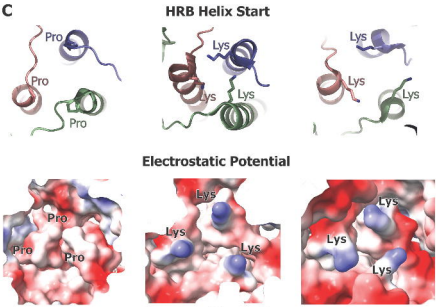


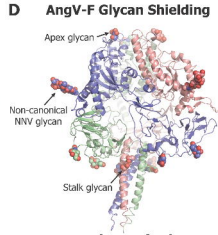
G G Head Domain Antibody Binding

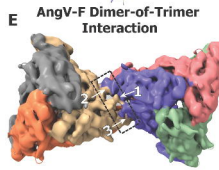


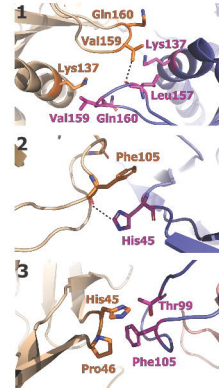












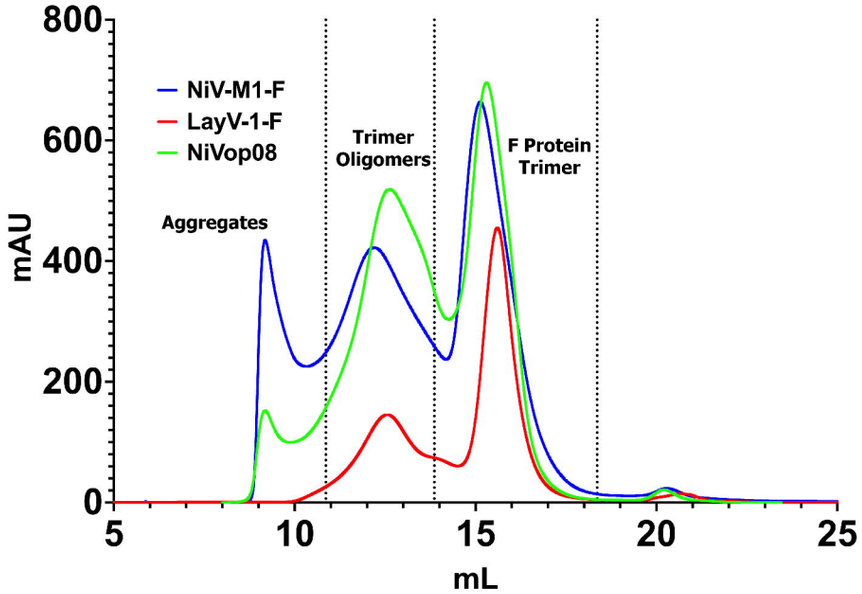
Sequence Comparison

AngV-F Residue	440	441	442	443	444	445	446	447	448	449	450	451	452	453	454	455	456	457	458	459	460
AngV-F	G	P	P	V	Y	T	E	P	V	D	I	T	N	Q	L	G	N	I	N	E	T
NiV-M1-F	G	P	P	V	F	T	D	K	V	D	I	S	S	Q	I	S	S	M	N	Q	S
LayV-1-F	G	P	P	V	V	I	D	K	I	D	I	G	N	Q	L	A	G	I	N	Q	T

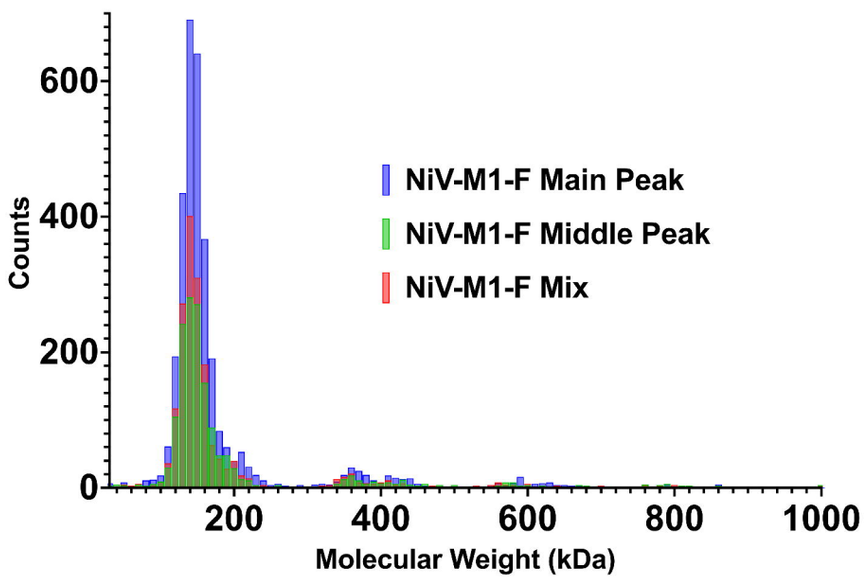
A HNV F Protein Yields and Similarities

Name	Identity to NiV-M1-F	Similarity to NiV-M1-F	Main Peak Yield/Liter (mg/L)	Main Peak Yield (Percentage of Total)
NiV-M1-F	N/A	N/A	0.525	50.7%
NiV-M2-F	99.82	99.82	0	N/A
NiV-M3-F	99.82	99.82	1.8	35.2%
NiV-M4-F	99.45	99.45	0	N/A
NiV-B1.2-F	99.08	99.63	1	26.8%
NiV-C-F	98.9	99.63	0.5	70.3%
NiV-B1.3-F	98.9	99.45	0.425	46.6%
NiV-B1.4-F	98.72	99.45	0.03	50.8%
NiV-B2.3-F	98.53	99.27	0.14	65.0%
NiV-B1.1-F	98.53	99.08	0	N/A
NiV-B3.1-F	98.17	98.72	0.38	25.1%
HeV-α2-F	88.3	94.33	0.8	35.1%
HeV-α1.1-F	88.12	94.15	0.305	57.1%
HeV-β1-F	87.91	93.22	0.65	55.9%
HeV-β2-F	87.73	93.41	0.1	65.7%
GhV(+A)-F	51.5	69.31	0.147	58.5%
GhV-F	N/A	N/A	0	N/A
CedV-1-F	42.93	61.72	0.05	63.7%
MojV-F	41.24	61.86	0.335	38.0%
SHNV5-1-F	40.8	61.75	5.7	23.8%
SHNV5-2-F	40.8	61.57	2.1	25.0%
LayV-1-F	40.77	60.33	1.17	65.6%
LayV-2-F	40.77	60.33	0.85	53.1%
SHNV2-F	40.61	61.55	2.1	36.9%
DarV-K-F	40.59	61.61	0	N/A
SHNV4-F	40.56	57.32	0.165	41.6%
DarV-C-F	40.4	61.06	0	N/A
GakV-1-F	39.86	58.02	0.44	58.6%
GakV-2-F	39.86	57.85	0.445	61.0%
DewV-F	39.85	62.52	0	N/A
MelV-F	39.71	61.93	0	N/A
SHNV3-F	39.51	57.14	0.31	41.1%
AngV-F	39.27	57.64	1.35	72.2%
ShyV-F	38.7	55.48	0.2	40.7%
NinV-F	38.02	55.11	0.9	59.1%
SHNV1-F	37.39	56.26	2.15	31.1%

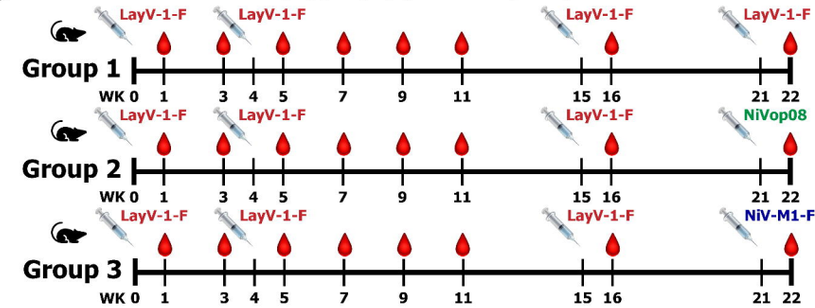
B Representative HNV F SEC Profiles



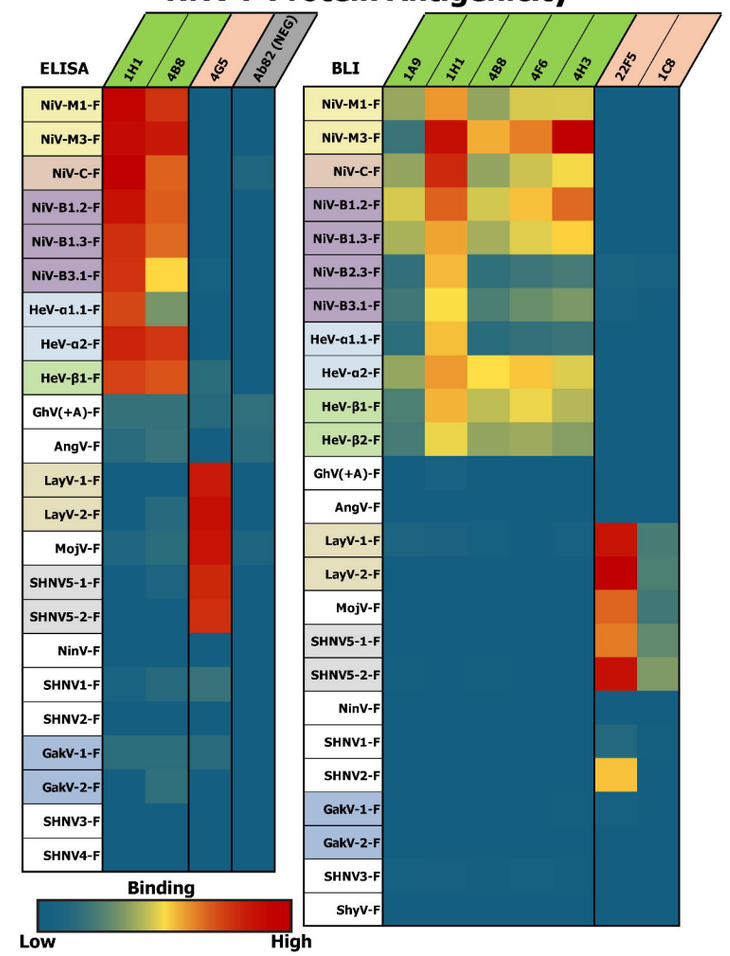
C HNV F Mass Photometry



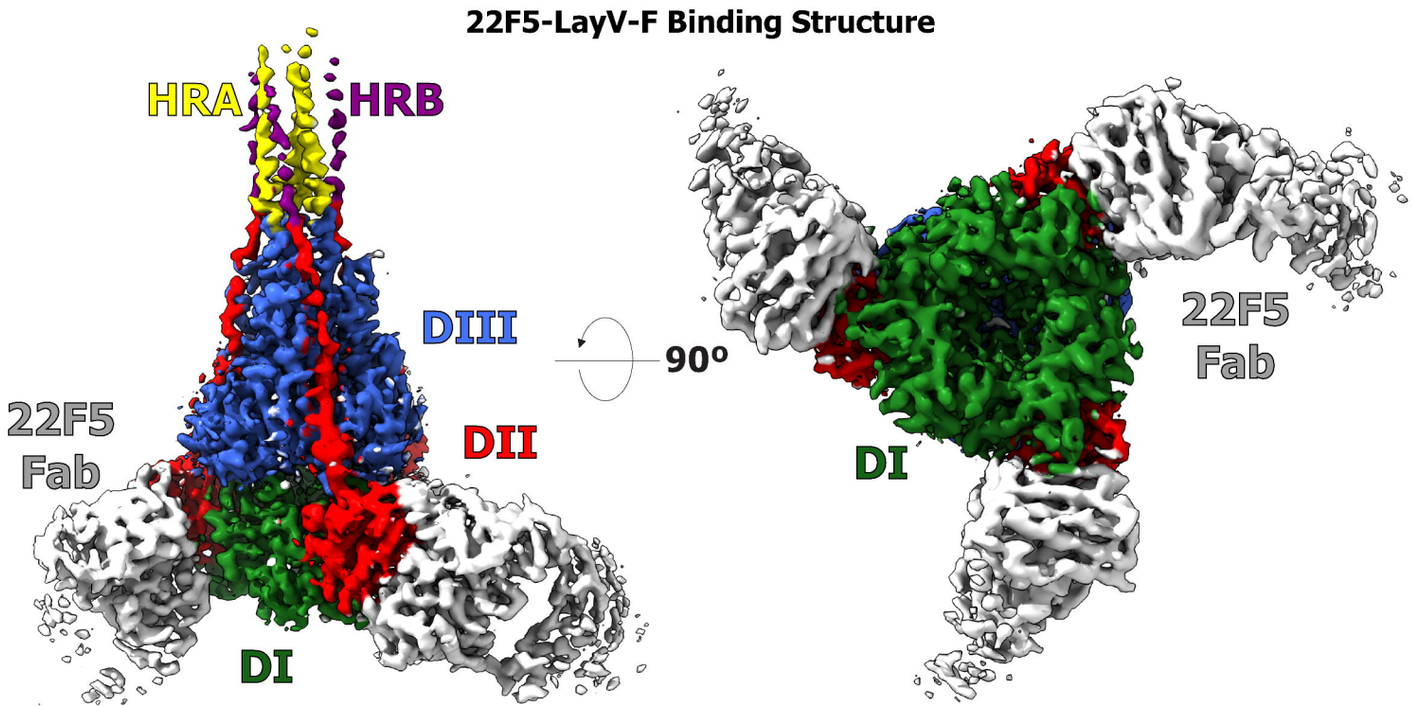
D Vaccination Timeline



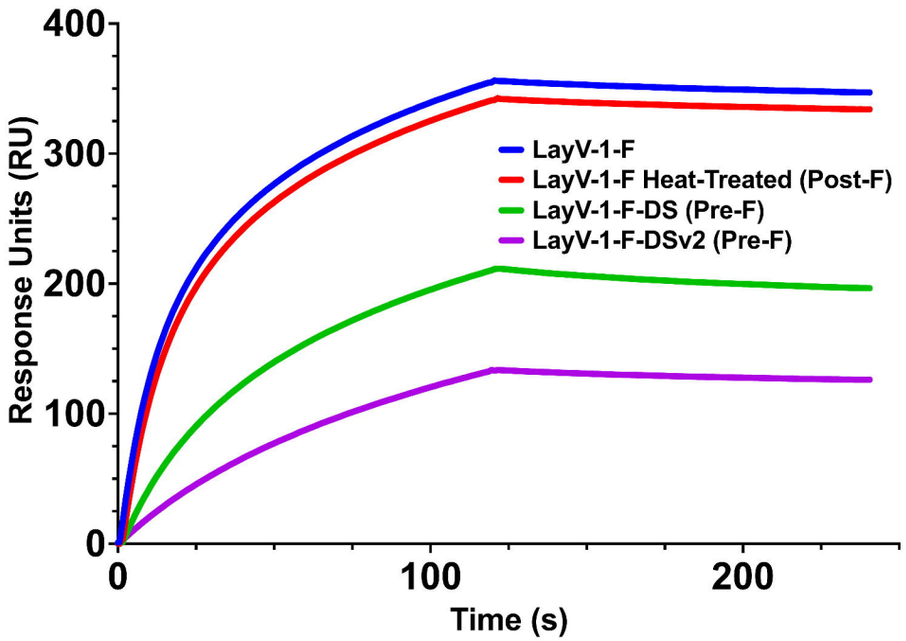
E HNV F Protein Antigenicity

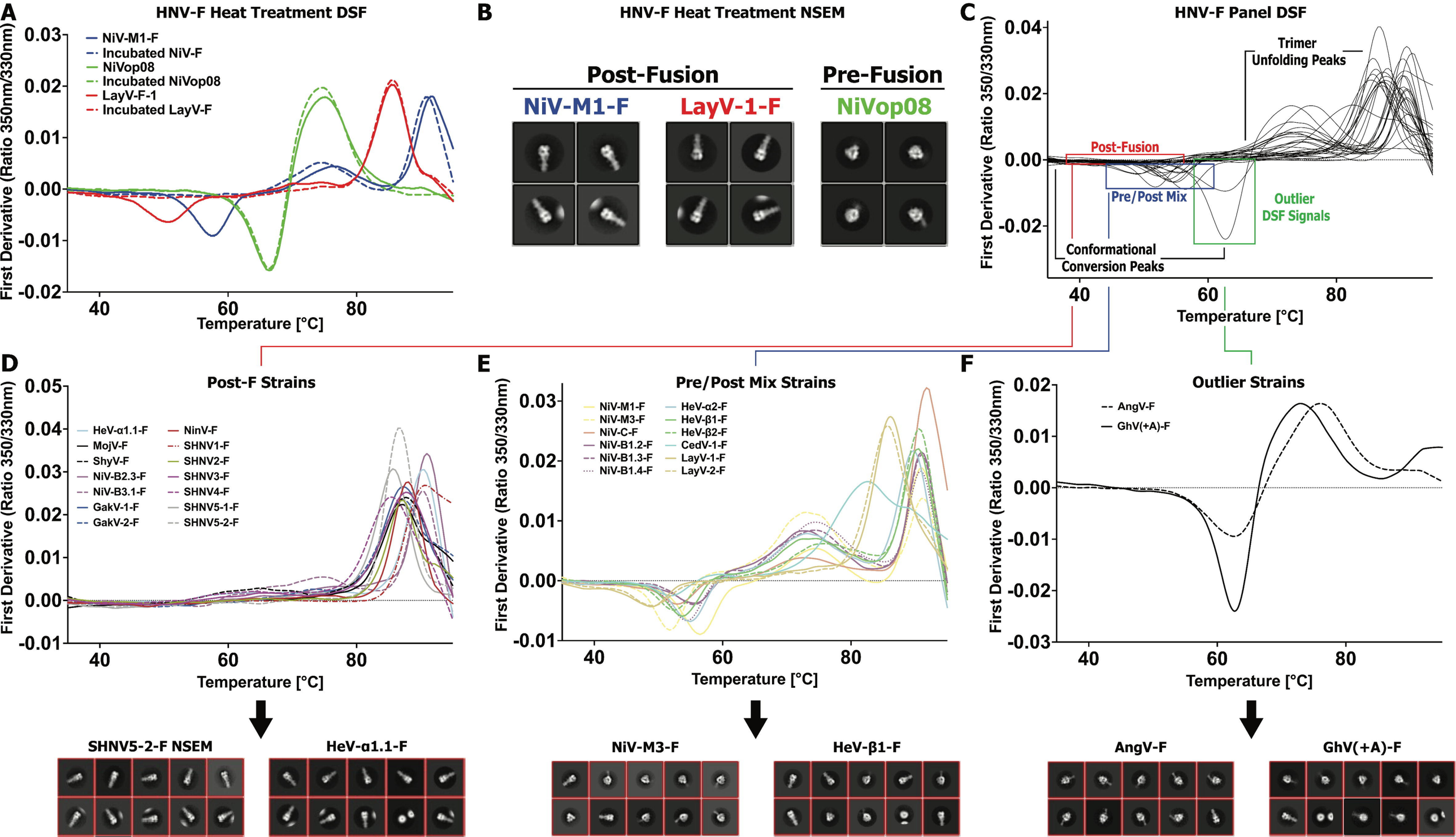


F 22F5-LayV-F Binding Structure



G 22F5-LayV-F Binding SPR




A
HNV-F Heat Treatment DSF
NiV-M1-F
Incubated NiV-F
NiVop08
Incubated NiVop08
LayV-F-1
Incubated LayV-F
First Derivative (Ratio 350nm/330nm)
Temperature [°C]
B
HNV-F Heat Treatment NSEM
Post-Fusion
Pre-Fusion
NiV-M1-F
LayV-1-F
NiVop08
C
HNV-F Panel DSF
Trimer Unfolding Peaks
Post-Fusion
Pre/Post Mix
Outlier DSF Signals
Conformational Conversion Peaks
First Derivative (Ratio 350/330nm)
Temperature [°C]
D
Post-F Strains
HeV-α1.1-F
MojV-F
ShyV-F
NiV-B2.3-F
NiV-B3.1-F
GakV-1-F
GakV-2-F
NinV-F
SHNV1-F
SHNV2-F
SHNV3-F
SHNV4-F
SHNV5-1-F
SHNV5-2-F
First Derivative (Ratio 350/330nm)
Temperature [°C]
E
Pre/Post Mix Strains
NiV-M1-F
NiV-M3-F
NiV-C-F
NiV-B1.2-F
NiV-B1.3-F
NiV-B1.4-F
HeV-α2-F
HeV-β1-F
HeV-β2-F
CedV-1-F
LayV-1-F
LayV-2-F
First Derivative (Ratio 350/330nm)
Temperature [°C]
F
Outlier Strains
AngV-F
GhV(+A)-F
First Derivative (Ratio 350/330nm)
Temperature [°C]
SHNV5-2-F NSEM
HeV-α1.1-F
NiV-M3-F
HeV-β1-F
AngV-F
GhV(+A)-F