Structure of the Lassa virus glycan shield provides a model for immunological resistance

Yasunori Watanabe[a,b,c], Jayna Raghwani[d], Joel D. Allen[c], Gemma E. Seabright[a,c], Sai Li[b], Felipe Moser[b], Juha T. Huiskonen[b,e], Thomas Strecker[f], Thomas A. Bowden[b,1], and Max Crispin[c,1]

[a]Oxford Glycobiology Institute, Department of Biochemistry, University of Oxford, OX1 3QU Oxford, United Kingdom; [b]Division of Structural Biology, University of Oxford, OX3 7BN Oxford, United Kingdom; [c]Centre for Biological Sciences and Institute of Life Sciences, University of Southampton, SO17 1BJ Southampton, United Kingdom; [d]Big Data Institute, Li Ka Shing Centre for Health Information and Discovery, Nuffield Department of Medicine, University of Oxford, OX3 7LF Oxford, United Kingdom; [e]Helsinki Institute of Life Science and Molecular and Integrative Biosciences Research Program, University of Helsinki, 00014 Helsinki, Finland; and [f]Institute of Virology, Philipps Universität Marburg, 35043 Marburg, Germany

Lassa virus is an Old World arenavirus endemic to West Africa that causes severe hemorrhagic fever. Vaccine development has focused on the envelope glycoprotein complex (GPC) that extends from the virion envelope. The often inadequate antibody immune response elicited by both vaccine and natural infection has been, in part, attributed to the abundance of N-linked glycosylation on the GPC. Here, using a virus-like−particle system that presents Lassa virus GPC in a native-like context, we determine the composite population of each of the N-linked glycosylation sites presented on the trimeric GPC spike. Our analysis reveals the presence of underprocessed oligomannose-type glycans, which form punctuated clusters that obscure the proteinous surface of both the GP1 attachment and GP2 fusion glycoprotein subunits of the Lassa virus GPC. These oligomannose clusters are seemingly derived as a result of sterically reduced accessibility to glycan processing enzymes, and limited amino acid diversification around these sites supports their role protecting against the humoral immune response. Combined, our data provide a structure-based blueprint for understanding how glycans render the glycoprotein spikes of Lassa virus and other Old World arenaviruses immunologically resistant targets.

Lassa virus | arenavirus | glycoprotein | structure | glycosylation

Lassa virus (LASV), an arenavirus endemic to rodent populations in West Africa, causes severe hemorrhagic fever upon transmission to humans, with an estimated 300,000 infections and several thousand deaths each year (1). To date, no licensed vaccine exists against LASV and there are no approved therapeutics for treatment, although the compassionate use of the antiviral drug ribavirin alone, or in combination with favipiravir, has shown efficacy if administered early in infection (2). The inconsistent induction of potent LASV-specific antibody immunity, both during natural infection and by vaccination attempts (3–7), is thought to be partly due to the extensive glycosylation presented on the surface of the LASV glycoprotein complex (GPC) (8). While the architecture of the protein components of LASV GPC is well understood (9–12), structural knowledge of the immunologically important glycans displayed on the GPC surface has remained elusive.

The arenaviral GPC is encoded by the S segment of the single-stranded, bisegmented RNA genome. Following expression of a single polyprotein precursor, the GPC is proteolytically processed by the host cell subtilase SKI-1/S1P (13) and displayed on the lipid bilayer surface of the mature virion as a trimer of heterotrimers (11, 14). Each heterotrimer consists of a receptor-binding GP1 domain (10, 15–19), a GP2 class I membrane fusion protein (10, 19–21), and a retained myristoylated stable signal peptide (SSP), which is required for GPC processing and function (22–24). Host cell entry of Old World arenaviruses and Clade C New World arenaviruses is initiated by GPC-mediated attachment to O-mannose glycans containing xylose-glucuronic acid sugar repeats presented on the cell surface receptor, α-dystroglycan

(α-DG) (25–28). Factors including the lowered pH associated with entry into the endocytic pathway, lysosome-associated membrane protein 1 (LAMP1) recognition, and dissociation of GP1 from the GPC are responsible for triggering fusion of the viral envelope and endosomal host cell membrane (11, 29, 30).

The 11 potential N-linked glycosylation sites (PNGs) found on each SSP−GP1−GP2 protomer accounts for nearly 30% of the total mass of the LASV GPC (10, 31). It has been shown that LASV glycans promote evasion from the neutralizing antibody response and are required for viral fitness (8). Additionally, underprocessed oligomannose-type glycans on LASV GPC augment infection through interaction with the C-type lectin, dendritic cell-specific ICAM-3−grabbing nonintegrin (DC-SIGN) (32). However, the relative abundance and distribution of these underprocessed glycans and fine structure of the LASV glycome have yet to be determined.

Here, using an established virus-like−particle (VLP) system (11, 14), we provide a global and site-specific analysis of the N-linked glycosylation presented on LASV GPC. Our analysis reveals extensive global heterogeneity, ranging from highly processed complex-type glycosylation to underprocessed oligomannose-type glycans. Mapping of these glycans onto the structure of the LASV GPC ectodomain reveals that some of these carbohydrates form a

Significance

Lassa virus is a highly pathogenic arenavirus that causes severe hemorrhagic fever in humans. Currently, there are no efficacious vaccines or treatments available to combat this pathogen. An important component of any vaccine candidate against Lassa virus will likely include the highly glycosylated glycoprotein complex presented on the virion surface. Here, we determine the composition of the Lassa virus glycome, revealing that the virus presents an abundance of glycans that are not biosynthetically processed to full maturity. Such underprocessed glycans form spatially distinct clusters, which shield the proteinous surface of the Lassa virus glycoprotein spike from the humoral immune response. These data are integral for the development of humoral-based vaccines that mimic the mature Lassa virion.
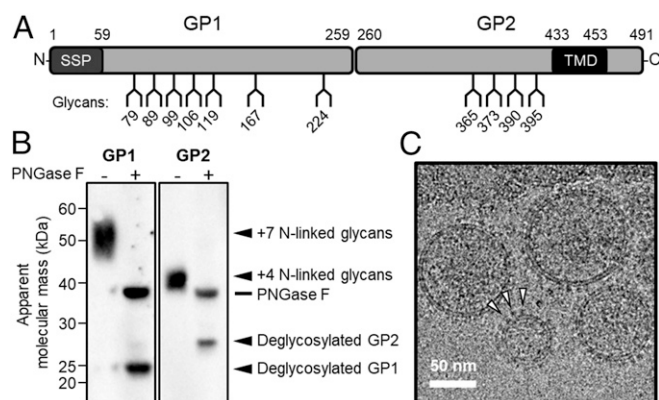
dense shield consisting of clusters of oligomannose-type glycans that are likely to be utilized during DC-SIGN−mediated entry. Sequence analysis of LASV isolates reveals a higher incidence of amino acid variation on the surface of the GPC in regions less occluded by N-linked glycosylation. Together, these data reveal how host cell-derived glycosylation renders LASV an immunologically challenging target.
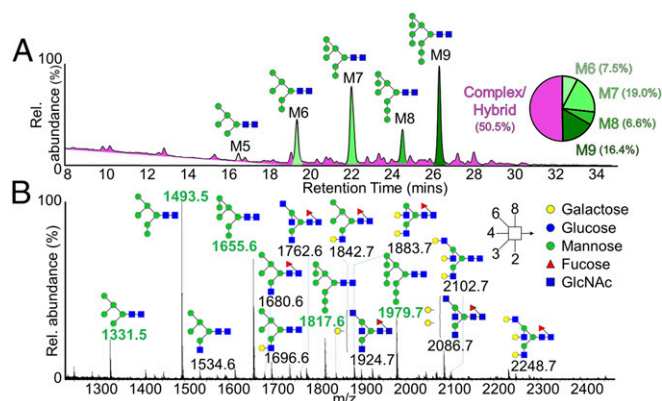
Results

Glycan Processing of Trimeric LASV GPC. LASV encodes a large number of N-linked glycosylation sequons on both the GP1 and GP2 subcomponents of the GPC (Fig. 1A). We sought to characterize the composition of the carbohydrate structures displayed on the mature LASV GPC spike using a previously established VLP production and purification strategy, which displays LASV GPC in the full native assembly (11, 14, 33). Western blot analysis of secreted VLPs using an anti-LASV GP1 (14) monoclonal antibody and anti-LASV GP2 polyclonal antibody (34) confirmed the presence and maturation of the GPC into GP1 and GP2 subunits (Fig. 1B). Electron cyromicroscopy further verified the expected display of globular GPC spikes on the surface of the purified VLPs (Fig. 1C).

To quantitatively assess the composition of the LASV GPC glycosylation, enzymatically released N-linked glycans were fluorescently labeled, and subjected to hydrophilic interaction chromatography−ultra-performance liquid chromatography (HILIC-UPLC) analysis. Cleavage of the released glycan pool with endoglycosidase H (Endo H) revealed an abundance (49%) of oligomannose-type glycans (Fig. 2A), particularly $Man_{6-9}GlcNAc_2$ glycoforms. We observed only a minor population of $Man_5GlcNAc_2$, suggesting the limited processing by Golgi-resident α-mannosidases to convert $Man_6GlcNAc_2$ to $Man_5GlcNAc_2$ at specific sites. The high proportion of oligomannose-type glycans on the LASV GPC is consistent with previous lectin-binding and glycosidase digestion studies (35, 36), which revealed underprocessed glycans on recombinant LASV GP1 constructs, albeit to varying levels.

We ascertained the precise structure of each of the glycans within the released pools by negative-ion mobility−electrospray ionization mass spectrometry (IM-ESI MS) (Fig. 2B). Consistent with the UPLC data, IM-ESI MS confirmed a dominant oligomannose-type glycan population, lacking the level of sialylation characteristic to most host cell surface glycoproteins (37, 38).

Fig. 2. Compositional analysis of LASV GPC glycosylation. (A) HILIC-UPLC profile of fluorescently labeled N-linked glycans from LASV GPC produced in VLPs. Oligomannose-type glycans (M5 to M9; $Man_5GlcNAc_2$-$Man_9GlcNAc_2$) (green) and complex-type glycans (pink) were identified by Endo H digestion, with quantification of major glycan types summarized as a pie chart. (B) Negative-ion electrospray spectrum of N-linked glycans from the GPC. Peaks are annotated with the corresponding compositions, using Consortium for Functional Glycomics symbolic nomenclature and Oxford system linkages (65), as per the key. Oligomannose-type glycan m/z values are labeled in green.

To confirm the composition of these oligomannose-type glycans, fragmentation by collision-induced dissociation was performed. This analysis reveals the expected array of underprocessed $Man_{5-9}GlcNAc_2$ species that form following processing by endoplasmic reticulum (ER)-resident α-glucosidases I and II (*SI Appendix*, Fig. S1) (39). Consistent with the biological preference of ER α-mannosidase I to remove the three-branched mannose of the six-antenna (D2) from $Man_9GlcNAc_2$ during N-linked glycan processing (40), MS/MS of $Man_8GlcNAc_2$ exhibits a single set of ions derived from a particular (D1, D3) isomer (*SI Appendix*, Fig. S1B). We observed two $Man_7GlcNAc_2$ isomers, with the seventh mannose located on either the six-antenna (D3) or three-antenna (D1) arm (*SI Appendix*, Fig. S1C). These isomers have also been observed in HIV-1 (gp120) and Semliki Forest virus E1 and E2 glycoproteins (41, 42); it is likely that the existence of high levels of oligomannose glycans among such pathobiologically diverse viruses arises from a requirement for conserved immune evasion and receptor recognition functionality, respectively.

Structural Clustering of Underprocessed Glycans on the GPC. It has been previously shown that oligomannose-type glycans displayed on the LASV GPC mediate host cell attachment to DC-SIGN presented on monocyte-derived dendritic cells (32). We sought to identify the possible location(s) of this interaction using an integrated site-specific compositional and structural approach. LASV GPC glycopeptides were generated by digestion with trypsin, chymotrypsin, and GluC and subjected to in-line liquid chromatography−mass spectrometry. This analysis revealed that each site presents a differential level of oligomannose, hybrid, and complex-type glycan populations (Fig. 3A). Consistent with our UPLC analysis (Fig. 2A), we observed that a number of these sites ($Asn79_{GP1}$, $Asn89_{GP1}$, $Asn99_{GP1}$, $Asn365_{GP2}$, and $Asn373_{GP2}$) were predominantly of the oligomannose type. The level of mannosylation varies widely across these sites, where the glycans at $Asn79_{GP1}$ and $Asn365_{GP2}$ present a dominant population of $Man_9GlcNAc_2$ and $Asn89_{GP1}$, $Asn99_{GP1}$, and $Asn373_{GP2}$ exhibit less restrictive processing and display $Man_6GlcNAc_2$ and $Man_7GlcNAc_2$.

Using the crystal structure of the trimeric LASV GP1−GP2 ectodomain [Protein Data Bank (PDB) ID code 5VK2 (10)], we determined the distribution of the observed glycan populations in a structural context. We generated a model of the fully glycosylated LASV GP1−GP2 ectodomain by overlaying our experimentally determined glycan structures onto crystallographically

Fig. 1. Preliminary characterization of LASV VLPs. (A) Schematic representation of LASV GPC, showing the position of N-linked glycosylation amino acid sequons (NXT/SX, where X ≠ P), SSP, GP1 attachment glycoprotein, GP2 fusion glycoprotein, and transmembrane domain (TMD). (B) Western blot analysis of purified LASV VLPs performed using anti-GP1 (14) and anti-GP2 (34) antibodies, against native and deglycosylated LASV GPC. Blots show LASV GP1 and GP2 before and after PNGase F treatment. (C) Electron cryomicroscopy image of VLPs purified from MDCK-II cells. White arrows indicate the positions of prominent LASV GPC spikes. (Scale bar, 50 nm.)
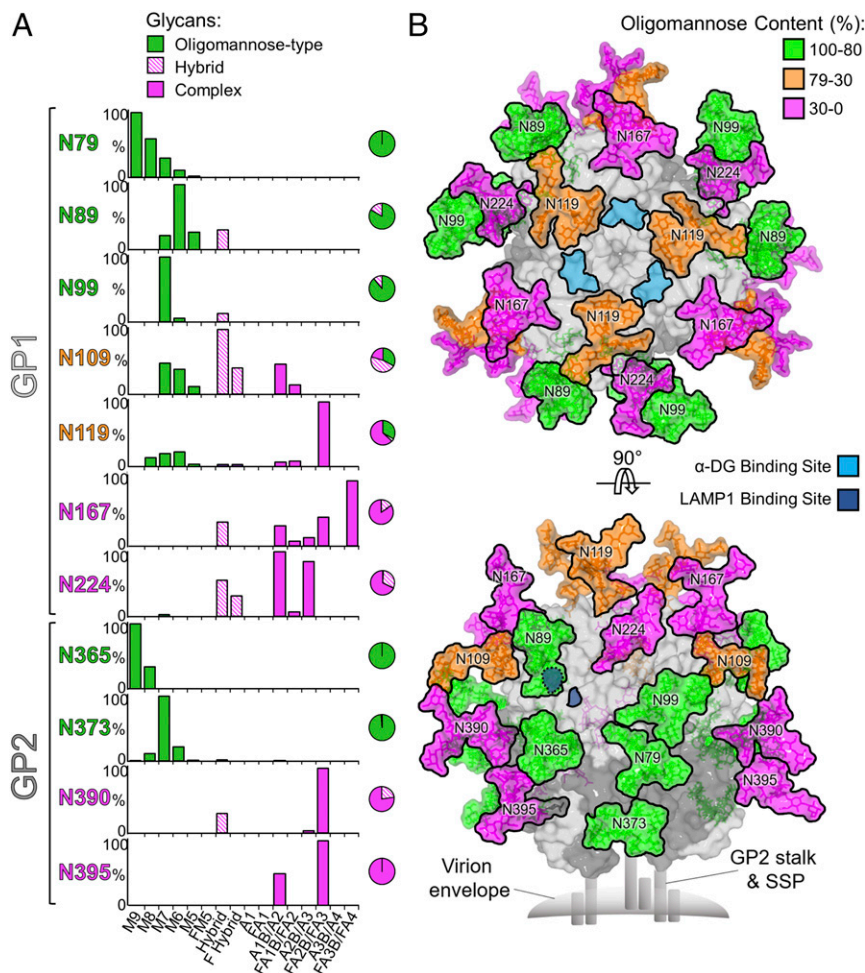
Fig. 3. Compositional analysis and structure-based mapping of LASV GPC N-linked glycans. (*A*) Quantitative site-specific N-linked glycan analysis of LASV GPC. Purified LASV GPC was digested with trypsin, chymotrypsin, and GluC, then analyzed by LC-ESI MS. Glycan compositions are based on the glycan library generated from negative-ion mass spectrometry of released N-glycans. The bar graphs represent the relative quantities of each glycan group with oligomannose-type glycan series (M9 to M5; Man_9-$GlcNAc_2$ to $Man_5GlcNAc_2$) (green), afucosylated and fucosylated hybrid glycans (Hybrid & F Hybrid) (dashed pink), and complex glycans grouped according to fucosylation and number of antennae (A1 to FA4) (pink). The pie charts summarize the quantification of these glycans. (*B*) Modeling experimentally observed glycosylation onto the crystal structure of the trimeric LASV GPC (PDB ID code 5VK2) (9). The glycans are colored according to the oligomannose content, as defined by the upper-right-hand key. The locations of α-DG and LAMP1 binding sites have been previously proposed (11, 15, 27, 66) and are annotated in cyan and dark blue, respectively. GP1 and GP2 are colored light gray and dark gray, respectively.

observed N-linked glycosylation sites (Fig. 3B). This analysis reveals how N-linked glycosylation envelopes the majority of the GPC ectodomain and provides a structure-based model for understanding glycan-mediated shielding of the proteinous LASV GPC surface. At a more detailed level, this integrated structural approach also indicates the existence of a high mannose cluster, which spans GP1 and GP2 subunits. This cluster is located on the side of the GPC trimer and includes $Asn79_{GP1}$, $Asn89_{GP1}$, $Asn99_{GP1}$, $Asn365_{GP2}$, and $Asn373_{GP2}$. We suggest that the dominant oligomannose (94%) composition observed in this region is likely due to the high spatial density of N-linked glycans, which limit the accessibility of glycan processing enzymes to these sites. The existence of such a high mannose cluster has precedence and is reminiscent of the "mannose patches" observed on the glycan shield of the HIV-1 envelope glycoprotein (43–45). Further examination reveals that the proposed LAMP1 binding site is partially occluded by this oligomannose cluster. In particular, the glycan extending from Asn89 shields the histidine triad (His92, His93, and His230) of LASV GP1 thought to be a region critical for this interaction (28, 29). The occlusion of this binding site is consistent with the proposed mechanism of LASV GPC–LAMP1 recognition (11, 15, 30), which requires pH-dependent conformational rearrangements of the GPC spike upon exposure to the late endosomes. Indeed, modeling indicates that the formation of the low-pH LASV GP1 conformation redirects the glycan extending from N89 to create a LAMP1-accessible recognition site (*SI Appendix*, Fig. S2).

Restricted glycan processing also occurs at Asn119, albeit to a lesser extent (33% oligomannose composition), at the membrane-distal apex of the trimer (Fig. 3B). Interestingly, Asn119 is located

near, but does not occlude, residues proposed to be required for α-DG binding (His141, Asn146, Phe147, and Tyr150) (27), and a lymphocytic choriomeningitis virus GPC N119Q site-directed mutation revealed that the Asn119 glycan protects the α-DG receptor binding site from antibody-mediated neutralization (8). The use of N-linked glycosylation to obstruct immunodominant receptor-binding sites has been observed for other highly glycosylated viruses, such as the CD4 binding site on the envelope spike of HIV-1 (46) and the sialic acid binding site on the HA of influenza virus (47).

Amino Acid Sequence Diversification Occurs at Antibody-Accessible Sites on the GPC. The absence of an effective antibody response is a characteristic feature of patients who succumb to LASV infection (3–7). Indeed, extensive affinity maturation is thought to be required to generate tightly binding antibodies capable of targeting and neutralizing the highly glycosylated GPC (48). We sought to quantify the steric role that GPC-displayed glycans play in shielding the GPC surface. Using publicly available LASV GPC gene sequences, we performed a comparative nonsynonymous to synonymous nucleotide substitutions (i.e., dN/dS ratio) analysis of the two major lineages, predominantly found in Nigeria and Sierra Leone (Fig. 4A and *SI Appendix*, Fig. S3). Consistent with a recent dN/dS analysis on LASV GPC (49), the mean dN/dS ratios for the Nigeria and Sierra Leone clades were estimated at 0.060 and 0.063, respectively. Using the crystal structure of LASV GPC, we found that solvent-accessible residues were characterized by significantly higher dN/dS (mean dN/dS ratios estimated at 0.070 and 0.080 for the Nigeria and Sierra Leone lineages, respectively), compared with buried residues (mean dN/dS ratios estimated at 0.015 and 0.028 for the Nigeria
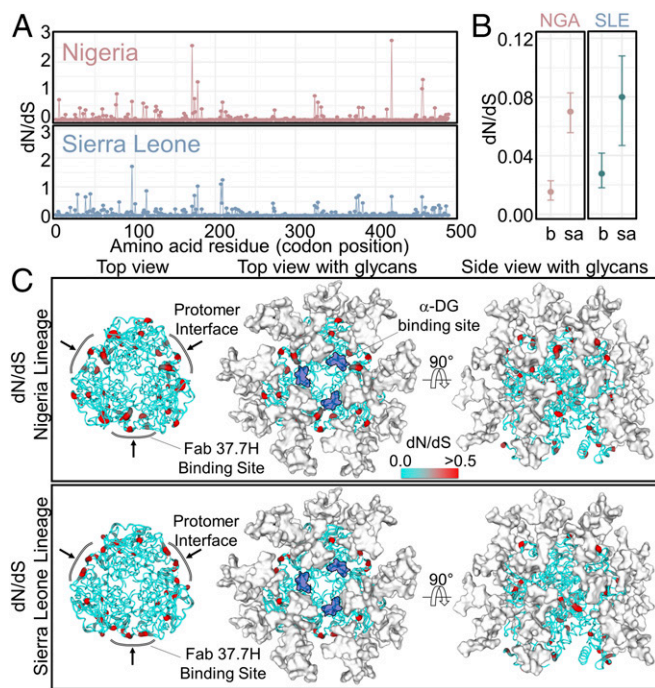
Fig. 4. Amino acid sequence diversification across the LASV GPC. (A) Site-wise dN/dS analysis of LASV GPC gene sequences from the Nigeria (red, lineage II) and Sierra Leone (blue, lineage IV) lineages. (B) Comparison of dN/dS values between buried (b) and solvent-accessible (sa) residues on the LASV GPC for Nigeria (NGA) and Sierra Leone (SLE). The error bars correspond to the 95% highest posterior density intervals, while circles indicate mean dN/dS values. (C) Mapping of per residue dN/dS values presented in A onto the crystal structure of LASV GPC (PDB ID code 5VK2) (9). LASV GPC is presented as a cartoon with residues colored according to dN/dS values. Residues with elevated dN/dS values are colored in red, and N-linked glycans are presented as white surfaces. The putative α-DG binding site is shown in dark blue (27). The interprotomer epitope of Fab 37.7H on LASV GPC is also shown.

and Sierra Leone lineages, respectively) (Fig. 4B). Interestingly, when mapped onto the structure, amino acids corresponding to nucleotides with the greatest relative number of nonsynonymous substitutions colocalized to regions of the GPC surface with lower levels of glycan density (Fig. 4C), distal from the functionally required α-DG binding site (50). These data provide a structural rationale for how protein surfaces occluded by glycan density are shielded from the humoral antibody response of the host and also rationalize previous suggestions that immune selection is a predominant driver of nonsynonymous substitutions for accessible regions of the LASV GPC (49).

On a broader level, we note the stringent conservation of LASV GPC N-linked glycosylation sequons across Old World arenaviruses, including the five glycan sites that constitute the LASV mannose cluster (Fig. 5A), suggesting that mannose patches may be inherent features of Old World arenaviral GPCs. In contrast, the glycan sites constituting the GP1 component of the mannose patch are less prevalent in New World arenaviral GPCs. This juxtaposition may be due to contrasting functional restraints (for example, the requirement to use transferrin receptor rather than α-DG). We suggest that such differential levels of N-linked glycosylation site conservation on the GP1 may at least partially account for the contrasting locations of immunodominant neutralizing epitopes on New World (the GP1 subunit) and Old World (GP1–GP2 interfaces) GPCs (10, 18, 51, 52).

Discussion

While the LASV GPC constitutes a key therapeutic target, vaccine design efforts have been hindered by the host-derived

glycosylation that encompasses the protein surface (8). Here, using an established VLP expression strategy, we performed a comprehensive compositional investigation of the LASV glycome. Our LASV GPC was produced in MDCK-II cells and forms a native-like conformation on the surface of the VLP (Fig. 1 B and C) (11, 14, 33). Although MDCK-II cells do not represent a natural host cell type, due to the stringent conservation of the N-linked glycan biosynthesis processing pathway (39) across mammalian cell types (53, 54), we expect only limited variations in the composition and abundance of the N-linked glycosylation of LASV GPC. Additionally, we note that the HIV-1 glycome is determined by structural constraints and is compositionally well conserved when produced in varied mammalian cell types (55). We suggest that future studies on the glycosylation state of circulating viruses in LASV-infected patients should corroborate our findings.

We present a structure-based model for rationalizing glycan-mediated shielding of LASV GPC from the humoral host response, revealing a cluster of underprocessed glycans that spans both the GP1 and GP2 subunits of the trimeric spike complex (Fig. 3). While the presence of mannose-rich glycosylation is unusual for cellular proteins, it has often been observed on enveloped viruses, such as orthomyxoviruses, phleboviruses, alphaviruses, and flaviviruses (56). The principles rationalizing the extent of glycan processing of viral proteins have yet to be fully elucidated and should be considered in the context of the route of egress from the host cell. In the case of LASV, however, we suggest that the formation of underprocessed moieties is likely
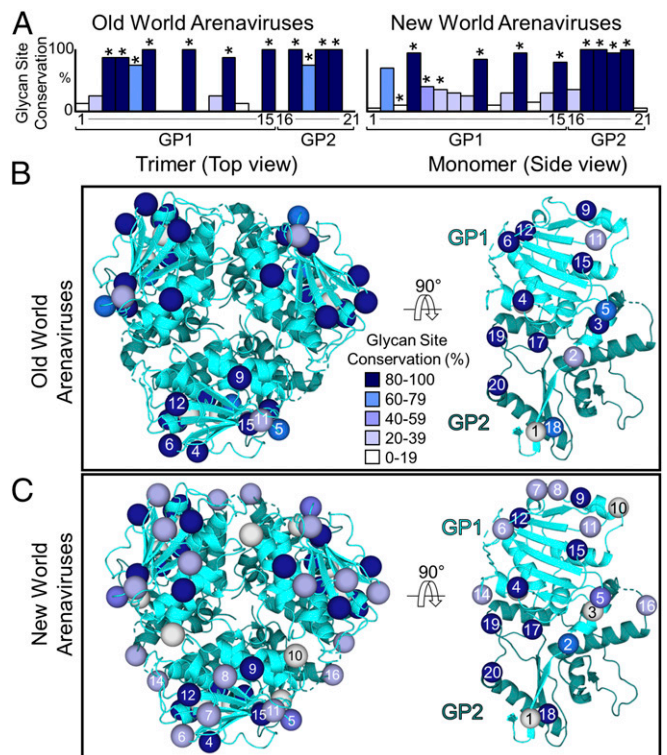
Fig. 5. N-linked glycosylation site conservation across Old and New World arenaviruses. (A) Glycan site conservation across arenaviruses following sequence alignment with Jalview (67). GenBank accession codes for the arenavirus glycoprotein sequences used in this analysis are listed in *SI Appendix*, Fig. S4. Glycan sites are numbered according to the nomenclature established by Sommerstein et al. (8), with those present on LASV GPC labeled with asterisks. (B and C) The positions of (B) Old World and (C) New World arenavirus N-linked glycosylation sequons mapped onto the structure of LASV GPC (PDB ID code 5VK2) (9). Spheres are colored according to the level of conservation of each glycosylation site.

enabled by the local structural environment and that the proximity of glycans extending from $Asn79_{GP1}$, $Asn89_{GP1}$, $Asn99_{GP1}$, $Asn365_{GP2}$, and $Asn373_{GP2}$ sterically impede accessibility to glycan-processing enzymes (*SI Appendix*, Fig. S5). Oligomannose-type glycans were also observed at sites outside of this cluster, albeit at lower levels. The presence of oligomannose at multiple locations across the GPC is suggestive of a promiscuous model of DC-SIGN usage, where numerous *N*-linked glycans may contribute as attachment factors for LASV entry.

The existence of a large population of oligomannose-type structures on LASV GPC has important consequences in the development and production of recombinant vaccine candidates. On recombinant glycoproteins, oligomannose-type glycans are considered to impart elevated clearance rates and may influence immunogenicity (57, 58). In addition, from immunogen design efforts for HIV, we have learned that the levels of oligomannose-type glycans on recombinant mimics of the envelope spike are highly sensitive to the structural constraints imposed by native-like trimeric oligomeric architecture (44, 55). Given the importance of both glycan processing states and native quaternary architecture, it is likely that the current recombinant expression technologies for LASV GP1 (15, 35, 36) and GPC (10) monomeric ectodomains may not reproduce native-like glycosylation. We envisage that the next generation of full-length LASV GPC ectodomain immunogens will both exhibit trimeric architecture and suppress the exposure of natively inaccessible epitopes through the presentation of more-native-like glycosylation.

This study also reveals that amino acid diversification occurs primarily at glycan-devoid regions on the LASV GPC (Fig. 4), providing important clues to what sites may be most vulnerable to the antibody immune response. By analogy to template-based vaccinology approaches against "evasion-strong" pathogens (59), such as HIV-1, it is possible that these regions are viable targets for the development of immunogens capable of effectively neutralizing LASV.

Continued emergence and high pathogenicity render LASV an important human pathogen and underscore the necessity for reagents capable of preventing and treating infection. Given the prominence of *N*-linked glycosylation on the LASV surface, we suggest that the compositional and structural data presented here will be integral for the creation of a humoral immunity-based vaccine capable of mimicking the mature LASV virion.

Materials and Methods

Expression and Purification of LASV VLPs. MDCK-II cells stably expressing LASV GPC (Josiah strain, GenBank ID code AY628203) were cultured in DMEM supplemented with 10% FCS for 96 h. The medium was replaced with DMEM with 2% FCS for VLP expression, as previously described (14). After 96 h, the cell supernatant was cleared by centrifugation (4,000 × *g* for 30 min). VLPs were pelleted through 10 mL of 20% sucrose cushion by ultracentrifugation (170,000 × *g* for 3 h). The pellets were resuspended in SPG buffer at pH 8 for 12 h at 4 °C.

Western Blot Analysis. Proteins from LASV VLPs were analyzed by Western blotting, as previously described (11). Monoclonal mouse antibody AC1 against GPC and GP1 and polyclonal rabbit antibody GP4 against GPC and GP2 were used as primary antibodies.

Electron Cyromicroscopy of VLPs. A 3-µL aliquot of purified VLP samples was applied on a plasma-cleaned grid coated with holey carbon (C-flat; Protochips). The grids were floated on a drop of 200-µL PBS buffer to remove sucrose and were blotted for 3 s followed by plunge-freezing into a mixture of liquid ethane (37%) and propane (63%) using a plunger device (CP3; Gatan). Micrographs were collected at liquid nitrogen temperature using a 300-kV transmission electron microscope (TF30 Polara; FEI) equipped with a CCD camera (Ultrascan 4000SP; Gatan) at a calibrated magnification of 75,000×, resulting in a final pixel size of 2 Å/pixel.

Release and Labeling of *N*-Linked Glycans. Excised LASV GPC gel bands were washed alternately with acetonitrile and water before drying in a vacuum centrifuge. The bands were rehydrated with 100 µL of water and incubated with PNGase F at 37 °C overnight. Released *N*-linked glycans were

fluorescently labeled with procainamide. A 100-µL aliquot of labeling mixture (110 mg/mL of procainamide and 60 mg/mL of sodium cyanoborohydrate in 70% DMSO and 30% glacial acetic acid) was added to the sample and incubated at 65 °C for 4 h. Labeled glycans were purified using Spe-ed Amide 2 columns (Applied Separations), as previously described (43).

Glycan Analysis by HILIC-UPLC. Labeled glycans were analyzed using a 2.1 mm × 10 mm Acquity BEH Glycan column (Waters) on a Waters Acquity H-Class UPLC instrument as performed in ref. 43, with wavelengths $\lambda_{ex} = 310$ nm and $\lambda_{em} = 370$ nm. Endo H digestions of labeled glycans were used to measure abundance of oligomannose-type glycans, as previously described (43).

Mass Spectrometry of Glycans. IM-ESI MS and tandem MS of released *N*-linked glycans were performed on a Synapt G2Si instrument (Waters) as previously described (44). Glycans were purified on a Nafion membrane before injection. Data acquisition and processing were carried out using MassLynx v4.11 and Driftscope version 2.8 software (Waters).

Mass Spectrometry of Glycopeptides. Aliquots of 150 µg to 200 µg of the VLPs were proteolytically digested using trypsin, chymotrypsin, and GluC (Promega). ProteoExtract Glycopeptide Enrichment Kit (Merck Millipore) was used to enrich glycopeptides for LC-ESI MS. For site-specific analysis of LASV GPC glycosylation, enriched glycopeptides were analyzed on an Orbitrap Fusion Tribrid mass spectrometer (Thermo Fisher Scientific) coupled to an EASY-Spray nano-LC system (Thermo Fisher Scientific). MS data were acquired with XCalibur 4.0 (Thermo Fisher Scientific). Glycopeptides were fragmented using both higher-energy collisional dissociation and energy transfer dissociation. Data analysis and glycopeptide identification were performed using Byonic and Byologic software (Protein Metrics).

Phylogenetic and Molecular Evolution Analysis. Publicly available sequences encoding the GPC spike were downloaded from GenBank and manually aligned. The sequence alignment comprised 217 sequences, spanning 1,473 bp of GPC region. A maximum likelihood phylogeny was inferred using randomized axelerated maximum likelihood (RAxML), where the model parameters included general time reversible model with rate categories (GTRCAT) substitutional model and rapid bootstrap analysis with 100 bootstrap samples (60). The resulting phylogeny indicated two distinct lineages that were supported by high bootstrap score (i.e., >0.95). Specifically, one lineage was dominated by sequences from Nigeria, while the other was predominantly observed in Sierra Leone. Since these two lineages likely represent independent epidemics, we estimated site-wise dN/dS ratios and mean amino acid diversity for each lineage. For the dN/dS analysis, Bayesian molecular clock phylogenies were estimated using Bayesian evolutionary analysis and sampling trees (BEAST) v1.8.4 (61). We assumed a strict molecular clock, Bayesian Skyline coalescent prior, and a codon-structured substitution model (62). Multiple independent Markov chain Monte Carlo (MCMC) runs of 10 million steps were executed to ensure stationarity and convergence had been achieved. An empirical distribution of 4,500 molecular clock phylogenies was obtained by combining (after the removal of burn-in) the posterior tree distributions from the separate runs, used to estimate dN/dS ratios using renaissance counting (63) implemented in BEAST v 1.8.4. Site-wise mean amino acid diversity was calculated using the Python script shown in *SI Appendix, Supplementary Method*.

Model Construction. Structural models of *N*-linked glycan presentation on LASV GPC were created using the trimeric crystal structure [PDB ID code 5VK2 (10)], complex-, hybrid-, and high mannose-type *N*-linked glycans (PDB ID codes 4BYH, 4B7I, and 2WAH, respectively). The most-dominant glycoform presented at each site, as observed on LASV GPC when produced in MDCK cells, was modeled onto the *N*-linked carbohydrate attachment sites in Coot (64).

1. Ogbu O, Ajuluchukwu E, Uneke CJ (2007) Lassa fever in West African sub-region: An overview. *J Vector Borne Dis* 44:1–11.
2. Raabe VN, et al.; Emory Serious Communicable Diseases Unit (2017) Favipiravir and ribavirin treatment of epidemiologically linked cases of lassa fever. *Clin Infect Dis* 65: 855–859.
3. Baize S, et al. (2009) Early and strong immune responses are associated with control of viral replication and recovery in Lassa virus-infected cynomolgus monkeys. *J Virol* 83: 5890–5903.
4. Fisher-Hoch SP, Hutwagner L, Brown B, McCormick JB (2000) Effective vaccine for Lassa fever. *J Virol* 74:6777–6783.
5. Lukashevich IS, et al. (2008) Safety, immunogenicity, and efficacy of the ML29 re-assortant vaccine for Lassa fever in small non-human primates. *Vaccine* 26:5246–5254.
6. Johnson KM, et al. (1987) Clinical virology of Lassa fever in hospitalized patients. *J Infect Dis* 155:456–464.
7. Jahrling PB, et al. (1980) Lassa virus infection of rhesus monkeys: Pathogenesis and treatment with ribavirin. *J Infect Dis* 141:580–589.
8. Sommerstein R, et al. (2015) Arenavirus glycan shield promotes neutralizing antibody evasion and protracted infection. *PLoS Pathog* 11:e1005276.
9. Crispin M, Zeltina A, Zitzmann N, Bowden TA (2016) Native functionality and thera-peutic targeting of arenaviral glycoproteins. *Curr Opin Virol* 18:70–75.
10. Hastie KM, et al. (2017) Structural basis for antibody-mediated neutralization of Lassa virus. *Science* 356:923–928.
11. Li S, et al. (2016) Acidic pH-induced conformations and LAMP1 binding of the Lassa virus glycoprotein spike. *PLoS Pathog* 12:e1005418.
12. Zeltina A, Bowden TA (2017) Human antibody pieces together the puzzle of the trimeric Lassa virus surface antigen. *Nat Struct Mol Biol* 24:559–560.
13. Lenz O, ter Meulen J, Klenk HD, Seidah NG, Garten W (2001) The Lassa virus glyco-protein precursor GP-C is proteolytically processed by subtilase SKI-1/S1P. *Proc Natl Acad Sci USA* 98:12701–12705.
14. Schlie K, et al. (2010) Characterization of Lassa virus glycoprotein oligomerization and influence of cholesterol on virus replication. *J Virol* 84:983–992.
15. Cohen-Dvashi H, Cohen N, Israeli H, Diskin R (2015) Molecular mechanism for LAMP1 recognition by Lassa virus. *J Virol* 89:7584–7592.
16. Bowden TA, et al. (2009) Unusual molecular architecture of the machupo virus at-tachment glycoprotein. *J Virol* 83:8259–8265.
17. Shimon A, Shani O, Diskin R (2017) Structural basis for receptor selectivity by the Whitewater Arroyo mammarenavirus. *J Mol Biol* 429:2825–2839.
18. Zeltina A, et al. (2017) Convergent immunological solutions to argentine hemorrhagic fever virus neutralization. *Proc Natl Acad Sci USA* 114:7031–7036.
19. Hastie KM, et al. (2016) Crystal structure of the prefusion surface glycoprotein of the prototypic arenavirus LCMV. *Nat Struct Mol Biol* 23:513–521.
20. Parsy ML, Harlos K, Huiskonen JT, Bowden TA (2013) Crystal structure of Venezuelan hemorrhagic fever virus fusion glycoprotein reveals a class 1 postfusion architecture with extensive glycosylation. *J Virol* 87:13070–13075.
21. Igonet S, et al. (2011) X-ray structure of the arenavirus glycoprotein GP2 in its post-fusion hairpin conformation. *Proc Natl Acad Sci USA* 108:19967–19972.
22. York J, Nunberg JH (2016) Myristoylation of the arenavirus envelope glycoprotein stable signal peptide is critical for membrane fusion but dispensable for virion mor-phogenesis. *J Virol* 90:8341–8350.
23. York J, Nunberg JH (2006) Role of the stable signal peptide of Junín arenavirus en-velope glycoprotein in pH-dependent membrane fusion. *J Virol* 80:7775–7780.
24. Eichler R, et al. (2003) Identification of Lassa virus glycoprotein signal peptide as a trans-acting maturation factor. *EMBO Rep* 4:1084–1088.
25. Oppliger J, Torriani G, Herrador A, Kunz S (2016) Lassa virus cell entry via dystroglycan involves an unusual pathway of macropinocytosis. *J Virol* 90:6412–6429.
26. Yoshida-Moriguchi T, Campbell KP (2015) Matriglycan: A novel polysaccharide that links dystroglycan to the basement membrane. *Glycobiology* 25:702–713.
27. Acciani M, et al. (2017) Mutational analysis of Lassa virus glycoprotein highlights regions required for alpha-dystroglycan utilization. *J Virol* 91:e00574-17.
28. Cao W, et al. (1998) Identification of alpha-dystroglycan as a receptor for lymphocytic choriomeningitis virus and Lassa fever virus. *Science* 282:2079–2081.
29. Jae LT, et al. (2014) Virus entry. Lassa virus entry requires a trigger-induced receptor switch. *Science* 344:1506–1510.
30. Cohen-Dvashi H, Israeli H, Shani O, Katz A, Diskin R (2016) Role of LAMP1 binding and pH sensing by the spike complex of Lassa virus. *J Virol* 90:10329–10338.
31. Eichler R, Lenz O, Garten W, Strecker T (2006) The role of single N-glycans in proteolytic processing and cell surface transport of the Lassa virus glycoprotein GP-C. *Virol J* 3:41.
32. Goncalves AR, et al. (2013) Role of DC-SIGN in Lassa virus entry into human dendritic cells. *J Virol* 87:11504–11515.
33. Schlie K, et al. (2010) Viral protein determinants of Lassa virus entry and release from polarized epithelial cells. *J Virol* 84:3178–3188.
34. Lenz O, ter Meulen J, Feldmann H, Klenk HD, Garten W (2000) Identification of a novel consensus sequence at the cleavage site of the Lassa virus glycoprotein. *J Virol* 74:11418–11421.
35. Illick MM, et al. (2008) Uncoupling GP1 and GP2 expression in the Lassa virus glycoprotein complex: Implications for GP1 ectodomain shedding. *Virol J* 5:161.
36. Galan-Navarro C, et al. (2017) Oxidation-sensitive polymersomes as vaccine nano-carriers enhance humoral responses against Lassa virus envelope glycoprotein. *Virology* 512:161–171.
37. Rudd PM, Dwek RA (1997) Glycosylation: Heterogeneity and the 3D structure of proteins. *Crit Rev Biochem Mol Biol* 32:1–100.
38. Qian R, Chen C, Colley KJ (2001) Location and mechanism of alpha 2,6-sialyltransferase dimer formation. Role of cysteine residues in enzyme dimerization, localization, ac-tivity, and processing. *J Biol Chem* 276:28641–28649.
39. Kornfeld R, Kornfeld S (1985) Assembly of asparagine-linked oligosaccharides. *Annu Rev Biochem* 54:631–664.
40. Tremblay LO, Herscovics A (1999) Cloning and expression of a specific human alpha 1,2-mannosidase that trims Man9GlcNAc2 to Man8GlcNAc2 isomer B during N-glycan biosynthesis. *Glycobiology* 9:1073–1078.
41. Crispin M, et al. (2014) Structural plasticity of the Semliki Forest virus glycome upon interspecies transmission. *J Proteome Res* 13:1702–1712.
42. Pritchard LK, Harvey DJ, Bonomelli C, Crispin M, Doores KJ (2015) Cell- and protein-directed glycosylation of native cleaved HIV-1 envelope. *J Virol* 89:8932–8944.
43. Pritchard LK, et al. (2015) Glycan clustering stabilizes the mannose patch of HIV-1 and preserves vulnerability to broadly neutralizing antibodies. *Nat Commun* 6:7479.
44. Behrens AJ, et al. (2016) Composition and antigenic effects of individual glycan sites of a trimeric HIV-1 envelope glycoprotein. *Cell Rep* 14:2695–2706.
45. Coss KP, et al. (2016) HIV-1 glycan density drives the persistence of the mannose patch within an infected individual. *J Virol* 90:11132–11144.
46. Jardine J, et al. (2013) Rational HIV immunogen design to target specific germline B cell receptors. *Science* 340:711–716.
47. Wei CJ, et al. (2010) Cross-neutralization of 1918 and 2009 influenza viruses: Role of glycans in viral evolution and vaccine design. *Sci Transl Med* 2:24ra21.
48. Robinson JE, et al. (2016) Most neutralizing human monoclonal antibodies target novel epitopes requiring both Lassa virus glycoprotein subunits. *Nat Commun* 7: 11544.
49. Andersen KG, et al.; Viral Hemorrhagic Fever Consortium (2015) Clinical sequencing uncovers origins and evolution of Lassa virus. *Cell* 162:738–750.
50. Jae LT, et al. (2013) Deciphering the glycosylome of dystroglycanopathies using haploid screens for Lassa virus entry. *Science* 340:479–483.
51. Brouillette RB, Phillips EK, Ayithan N, Maury W (2017) Differences in glycoprotein complex (GPC) receptor binding site accessibility prompt poor cross-reactivity of neutralizing antibodies between closely related arenaviruses. *J Virol* 91:e01454-16.
52. Mahmutovic S, et al. (2015) Molecular basis for antibody-mediated neutralization of new world hemorrhagic fever mammarenaviruses. *Cell Host Microbe* 18:705–713.
53. Gagneux P, Aebi M, Varki A (2015) Evolution of glycan diversity. *Essentials of Glycobiology*, eds Varki A, et al. (Cold Spring Harb Lab, Cold Spring Harbor, NY), 3rd Ed.
54. Huai G, Qi P, Yang H, Wang Y (2016) Characteristics of α-Gal epitope, anti-Gal anti-body, α1,3 galactosyltransferase and its clinical exploitation (Review). *Int J Mol Med* 37:11–20.
55. Pritchard LK, et al. (2015) Structural constraints determine the glycosylation of HIV-1 envelope trimers. *Cell Rep* 11:1604–1613.
56. Van Breedam W, Pöhlmann S, Favoreel HW, de Groot RJ, Nauwynck HJ (2014) Bitter-sweet symphony: Glycan-lectin interactions in virus biology. *FEMS Microbiol Rev* 38: 598–632.
57. Dey AK, et al. (2018) cGMP production and analysis of BG505 SOSIP.664, an extensively glycosylated, trimeric HIV-1 envelope glycoprotein vaccine candidate. *Biotechnol Bioeng* 115:885–899.
58. Reusch D, Tejada ML (2015) Fc glycans of therapeutic antibodies as critical quality attributes. *Glycobiology* 25:1325–1334.
59. Burton DR (2017) What are the most powerful immunogen design vaccine strategies? Reverse vaccinology 2.0 shows great promise. *Cold Spring Harb Perspect Biol* 9: a030262.
60. Stamatakis A (2014) RAxML version 8: A tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* 30:1312–1313.
61. Drummond AJ, Suchard MA, Xie D, Rambaut A (2012) Bayesian phylogenetics with BEAUti and the BEAST 1.7. *Mol Biol Evol* 29:1969–1973.
62. Shapiro B, Rambaut A, Drummond AJ (2006) Choosing appropriate substitution models for the phylogenetic analysis of protein-coding sequences. *Mol Biol Evol* 23: 7–9.
63. Lemey P, Minin VN, Bielejec F, Kosakovsky Pond SL, Suchard MA (2012) A counting renaissance: Combining stochastic mapping and empirical Bayes to quickly detect amino acid sites under positive selection. *Bioinformatics* 28:3248–3256.
64. Emsley P, Cowtan K (2004) Coot: Model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* 60:2126–2132.
65. Harvey DJ, et al. (2009) Proposal for a standard system for drawing structural dia-grams of N- and O-linked carbohydrates and related compounds. *Proteomics* 9: 3796–3801.
66. Israeli H, Cohen-Dvashi H, Shulman A, Shimon A, Diskin R (2017) Mapping of the Lassa virus LAMP1 binding site reveals unique determinants not shared by other old world arenaviruses. *PLoS Pathog* 13:e1006337.
67. Waterhouse AM, Procter JB, Martin DM, Clamp M, Barton GJ (2009) Jalview Version 2–A multiple sequence alignment editor and analysis workbench. *Bioinformatics* 25: 1189–1191.

BIOCHEMISTRY