Defense Threat Reduction Agency
8725 John J. Kingman Road, MS 6201
Fort Belvoir, VA 22060-6201

DTRA-TR-20-37

Effects of Differential Glycosylation on Filovirus Glycoprotein Immunogenicity

HDTRA1-15-1-0061

Dr Steven Bradfute
University of New Mexico Health Sciences Center

TECHNICAL REPORT

June 2020

DESTRUCTION NOTICE:

Destroy this report when it is no longer needed.
Do not return to sender.

PLEASE NOTIFY THE DEFENSE THREAT REDUCTION AGENCY, ATTN: DTRIAC/ RD-NTF, 8725 JOHN J. KINGMAN ROAD, MS-6201, FT BELVOIR, VA 22060-6201, IF YOUR ADDRESS IS INCORRECT, IF YOU WISH IT DELETED FROM THE DISTRIBUTION LIST, OR IF THE ADDRESSEE IS NO LONGER EMPLOYED BY YOUR ORGANIZATION.

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YYYY)*	2. REPORT TYPE	3. DATES COVERED *(From - To)*
05/15/2020	Final Report	June 30, 2015-March 31, 2020

4. TITLE AND SUBTITLE
Effects of differential filovirus glycoprotein glycosylation on immunogenicity

5a. CONTRACT NUMBER

5b. GRANT NUMBER
HDTRA1-15-1-0061

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)
Clarke, Elizabeth C, Collar, Amanda L, Ye, C, and Bradfute, Steven B.

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
University of New Mexico Health Sciences Center
Department of Internal Medicine
Center for Global Health
MSC10 5550, Albuquerque, NM 87131

8. PERFORMING ORGANIZATION REPORT NUMBER
FP5455, Fund 3Y590

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
DEFENSE THREAT REDUCTION AGENCY
8725 JOHN J. KINGMAN ROAD, MSC 6201
FORT BELVOIR VA 22060-6201

10. SPONSOR/MONITOR'S ACRONYM(S)
DTRA

11. SPONSOR/MONITOR'S REPORT NUMBER(S)
DTRA-TR-20-37

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release: distribution unlimited.

13. SUPPLEMENTARY NOTES

14. ABSTRACT
All current filovirus vaccine platforms utilize the highly glycosylated filovirus attachment protein (glycoprotein, or GP1,2) as an antigen. However, the different vaccine platforms are expressed in various cell types that endow differential glycosylation on GP1,2s, and the importance of the various types of glycosylation in promoting immunogenicity of filovirus GPs is not well known. In this project, we 1) determined N- and O-linked glycan patterns on filovirus GP1,2 expressed in mammalian, insect, and plant cells; 2) determined that insect cell-generated glycosylated GP1,2s had increased in vitro immunogenicity in dendritic cells; and 3) found that glycosylation of GP1,2 affected in vitro antibody and T cell responses.

15. SUBJECT TERMS
Filovirus, glycoprotein, immunogenicity, vaccine, glycosylation

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			Lt Col Krystal Walker
U	U	U	UU	123	19b. TELEPHONE NUMBER *(Include area code)* 571-616-6319

Standard Form 298 (Rev. 8/98)
Prescribed by ANSI Std. Z39.18

INSTRUCTIONS FOR COMPLETING SF 298

1. REPORT DATE. Full publication date, including day, month, if available. Must cite at least the year and be Year 2000 compliant, e.g. 30-06-1998; xx-06-1998; xx-xx-1998.

2. REPORT TYPE. State the type of report, such as final, technical, interim, memorandum, master's thesis, progress, quarterly, research, special, group study, etc.

3. DATE COVERED. Indicate the time during which the work was performed and the report was written, e.g., Jun 1997 - Jun 1998; 1-10 Jun 1996; May - Nov 1998; Nov 1998.

4. TITLE. Enter title and subtitle with volume number and part number, if applicable. On classified documents, enter the title classification in parentheses.

5a. CONTRACT NUMBER. Enter all contract numbers as they appear in the report, e.g. F33315-86-C-5169.

5b. GRANT NUMBER. Enter all grant numbers as they appear in the report. e.g. AFOSR-82-1234.

5c. PROGRAM ELEMENT NUMBER. Enter all program element numbers as they appear in the report, e.g. 61101A.

5e. TASK NUMBER. Enter all task numbers as they appear in the report, e.g. 05; RF0330201; T4112.

5f. WORK UNIT NUMBER. Enter all work unit numbers as they appear in the report, e.g. 001; AFAPL30480105.

6. AUTHOR(S). Enter name(s) of person(s) responsible for writing the report, performing the research, or credited with the content of the report. The form of entry is the last name, first name, middle initial, and additional qualifiers separated by commas, e.g. Smith, Richard, J, Jr.

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES). Self-explanatory.

8. PERFORMING ORGANIZATION REPORT NUMBER. Enter all unique alphanumeric report numbers assigned by the performing organization, e.g. BRL-1234; AFWL-TR-85-4017-Vol-21-PT-2.

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES). Enter the name and address of the organization(s) financially responsible for and monitoring the work.

10. SPONSOR/MONITOR'S ACRONYM(S). Enter, if available, e.g. BRL, ARDEC, NADC.

11. SPONSOR/MONITOR'S REPORT NUMBER(S). Enter report number as assigned by the sponsoring/ monitoring agency, if available, e.g. BRL-TR-829; -215.

12. DISTRIBUTION/AVAILABILITY STATEMENT. Use agency-mandated availability statements to indicate the public availability or distribution limitations of the report. If additional limitations/ restrictions or special markings are indicated, follow agency authorization procedures, e.g. RD/FRD, PROPIN, ITAR, etc. Include copyright information.

13. SUPPLEMENTARY NOTES. Enter information not included elsewhere such as: prepared in cooperation with; translation of; report supersedes; old edition number, etc.

14. ABSTRACT. A brief (approximately 200 words) factual summary of the most significant information.

15. SUBJECT TERMS. Key words or phrases identifying major concepts in the report.

16. SECURITY CLASSIFICATION. Enter security classification in accordance with security classification regulations, e.g. U, C, S, etc. If this form contains classified information, stamp classification level on the top and bottom of this page.

17. LIMITATION OF ABSTRACT. This block must be completed to assign a distribution limitation to the abstract. Enter UU (Unclassified Unlimited) or SAR (Same as Report). An entry in this block is necessary if the abstract is to be limited.

FINAL REPORT

Grant/Award #: HDTRA1-15-1-0061

Project Title: Effects of Differential Glycosylation on Filovirus Glycoprotein Immunogenicity

PI: Steven Bradfute

Institution: University of New Mexico Health Sciences Center

TABLE OF CONTENTS

Section **Page**

EXECUTIVE SUMMARY

Background

Filoviruses are highly virulent pathogens listed as Centers for Disease Control and Prevention (CDC) Category A Bioterrorism Agents. The glycoproteins (GP1,2) of filoviruses are the only virally expressed proteins on the virion surface and are required for receptor binding and cell entry; as such, they are the main candidate vaccine antigen. Although the filovirus glycoproteins are extensively glycosylated with N- and O-linked glycans, the extensive characterization of glycans present on GP1,2 from the different filoviruses is not well-studied, and the immunogenic impact of changing the character of the glycans on a vaccine is unknown. Both mammalian and insect cell lines are commonly used to produce filovirus proteins used in vaccine studies. We determined the glycosylation of the GP1,2 of all known pathogenic ebolaviruses (Ebola, Tai Forest, Bundibugyo, Sudan), as well as a Marburgvirus (Marburg) and a cuevavirus (Lloviu).

Data

We demonstrated that the N-linked glycosylation imparted by the mammalian (HEK293T) cell line was mostly complex type, terminally sialyated and displaying a large diversity of glycans. By contrast, the N-linked glycans from insect cells (Sf9) for all the GP1,2s were overwhelming fucosylated paucimannosidic glycans, with one particular glycan, (Fuc)1(Man)3(GlcNAc)2, making up the majority of all the glycans found on the insect GP1,2s. The paucimannosidic glycans were not part of the diversity of glycans found in the mammalian GP1,2s. We also investigated O-linked glycans, as the filovirus GP1,2s are known to contain mucin-like domains that are extensively O-glycosylated. We found a diversity of O-linked glycans present on mammalian GP1,2, but the insect GP1,2s carried either no, or extremely limited, O-linked glycans. We also determined the site-occupancy of each predicted N-linked glycan site of Ebola-Makona GP1,2, showing that 15 of 17 predicted sites on were occupied in the mammalian system 13 occupied in the insect system.

We then investigated the impact that these changes in glycosylation have on vaccine immunogenicity. To do this, we used an in vitro approach using virus like particles (VLPs), and an in vivo approach using purified GP1,2 subunit vaccines. To examine immunogenicity in vitro, murine bone marrow derived dendritic cells (DCs) were stimulated with VLPs generated from either insect or mammalian cell lines, and activation of the DCs was ascertained. Insect-derived VLPs from all tested filoviruses stimulated plasmacytoid DCs and CD8α-positive myeloid DCs to a greater extent than did mammalian-derived VLPs.

For in vivo investigations, we determined the impact of the filovirus GP1,2 glycans on the immunogenicity of the transmembrane-deleted GP1,2 (GPdTM) subunit vaccine in mice using insect-produced, mammalian-produced, and deglycosylated filovirus GPdTMs. Antibody titer was higher in insect-derived vaccines after one and two injections than in the mammalian-

derived vaccine. This effect was mediated by glycosylation as all differences between GPdTMs were lost after enzymatic deglycosylaton of the protein. However, the differences were no longer significant after the third and final vaccine dose. Notably, in neutralizing assays using VSV-EBOV GP, there were no differences in the neutralizing ability of the vaccinated mice. Since T-cell mediated immunity may also be important for filovirus protection, we investigated the presence of specific T cells using peptide mapping. We show that the specific T cells are induced by all vaccines, but while patterns of epitopes are different between the different groups, it is unclear if the patterns are statistically significant. To investigate this further, we are developing T-cell receptor sequencing that would allow us to show changes in diversity of T cells after vaccination. Finally, we have received approval to have mice vaccinated with the Ebola and Marburg viruses subunit vaccines (mammalian-derived, insect-derived and deglycosylated) and challenged with the respective mouse-adapted viruses at the National Institutes of Health, which will allow us to determine whether glycosylation mediated differences in antibodies and T cells correlate to changes in protection. These studies have been put on hold due to the coronavirus outbreak, but we plan to pursue these studies regardless of the ending of the grant.

Personnel supported

Twenty personnel were trained and/or supported by this grant. Of these, 6/23 trainees were from underrepresented minority groups, and 16/23 were women. This grant trained and supported faculty, medical students, graduate rotation students, graduated students, post-baccalaureate students, technicians, and undergraduate students. Included in this total were minority post-baccalaureate female students under the Premedical Enrichment Program and Post-Baccalaureate Research and Education Program, which focuses on minority and educationally disadvantaged students. Furthermore, after attending a Chemical and Biological Defense Science and Technology (CS&T) Defense Threat Reduction Agency (DTRA) meeting, we spearheaded our institution's entry into Minority-Serving Institute Science Technology Engineering and Mathematics Research & Development Consortium (MSRDC). The work reported here comprises the bulk of one student's Ph.D. project and compiled preliminary data that was used for multiple grant applications.

Dissemination of data

The work accomplished under this grant resulted in 4 primary manuscripts, 2 reviews, and 1 commentary, for a total of 7 peer-reviewed publications. Data from this work was presented in a poster format at 13 scientific meetings and in an oral format at 17 meetings.

CHAPTER 1. INTRODUCTION AND BACKGROUND

Filoviruses. Filoviruses (mononegaviral family Filoviridae) are a group of highly virulent pathogens that can cause viral hemorrhagic fevers in humans and/or nonhuman primates. Currently, ten recognized filoviruses are classified into four genera: Ebolavirus and Marburgvirus, and two genera, Cuevavirus and Dianlovirus, which contain single viruses of unknown human pathogenicity. The members of the genus Ebolavirus, i.e., ebolaviruses, are Bundibugyo (BDBV), Ebola virus (EBOV), Reston virus (RESTV), Sudan virus (SUDV), Bombali virus (BOMV) and Taï Forest virus (TAFV). All ebolaviruses except RESTV cause Ebola virus disease (EVD) in humans. The genus Marburgvirus, i.e., marburgviruses, contains Marburg virus (MARV) and Ravn virus (RAVV), both of which cause Marburg virus disease (MVD) in humans. The genus Cuevavirus has a single member, Lloviu virus (LLOV), which has been associated with lethal disease in bats but has unknown pathogenicity for primates. Finally, the genus Dianlovirus also has a single member, Mengla virus (MLAV), which was found in bat species in China, but has unknown pathogenicity [2]. There are additionally fish viruses, striavirus and thamnovirus genera, that are phylogenetically classified as filoviruses [3]), although both genera lack homologues of significant Ebolavirus and Marburgvirus proteins including the glycoprotein and VP30, and are evolutionarily distal to the human pathogens. Infection with pathogenic mammalian filoviruses can cause severe hemorrhagic fever, with high levels of proinflammatory cytokines and hematological alterations that culminate in multiple organ failure and death [4]. The filoviruses are composed of 7 proteins, including the glycoprotein (GP1,2) which is involved in cellular attachment and entry. At the beginning of this study, there were no U.S.-approved therapeutics or vaccines for filovirus disease, although in late 2019 an EBOV vaccine was approved by the FDA.

Glycosylation. Glycosylation is the addition of sugars to protein or lipid molecules in the cell. The interaction between cells of any multicellular organism relies of on the extensively glycosylated extracellular matrix (ECM) [5]. Additionally, of all proteins that sit on the cell membrane, or are secreted, the two sets of proteins most vital for cell-to-cell communication, are proteins that are overwhelmingly glycosylated in higher organisms. Unlike proteins and nucleic acids, which are made according to templates, glycans are altered by complex pathways of biosynthesis, and there is a stochastic nature in their addition to proteins. This allows glycans to enable responses within cells to environmental changes, as well as contribute to protein structure and stability.

There are two main types of glycosylation, N-linked and O-linked glycosylation. N-linked glycosylation is the addition of a glycan, or sugar moiety, to a protein structure at a nitrogen atom of an asparagine amino acid at a conserved Asn-X-Ser/Thr, where X is any amino acid except Proline. In the cell, N-linked glycosylation occurs in the endoplasmic reticulum (ER) and Golgi network, and a well-defined series of glycosylation enzymes interact with sites on the

protein to produce glycosylated proteins. However, while glycosylation is considered a post-translational modification, the sugars added in N-linked glycosylation are not a binary presence/absence event as occurs with phosphorylation. The same protein may have different glycans present at the same glycosylation site even when produced in the same cells under the same conditions as a result of the stochasticity in the system. The initial step of N-linked glycan biosynthesis involves the en bloc attachment of the N-glycan precursor by oligosaccharyl transferase (OST) enzyme in the endoplasmic reticulum, and this begins a series of trimming and adding enzymes modifying that original glycan to produce a range of glycans that are generally divided into three groups: high mannose, hybrid and complex.

In O-glycosylation, the glycan is bound to the oxygen atom of serine or threonine amino acid in the protein in the cis-Golgi, although, unlike N-linked glycosylation, without a conserved amino acid motif to guide the enzymatic addition of the O-glycans. This process occurs once the protein is folded, and is more difficult to predict than N-linked glycosylation, and less well-studied. Most types of O-glycosylation are carried out in the secretory pathway, primarily in the cis-Golgi, but cytosolic and nuclear proteins can also be glycosylated with N-acetylglucosamine (GlcNAc) by a cytosolic glycosyltransferase [6].

Glycosylation and viruses. Enveloped viruses co-opt the host glycosylation pathways, and a variety of viral proteins are glycoproteins. The functions of viral glycosylation can be divided into two broad groups: glycosylation that is essential for the correct structure and folding of the protein (intrinsic functions), and glycosylation that changes the way in which the viral glycoprotein interacts with other molecules, including binding to chaperones, receptors, lectins and interactions with the immune system (extrinsic functions) [7].
The former example, where glycosylation is essential for structure, can be determined by mutations of the glycan sites of the proteins. The loss of many N-glycan sites in viral proteins has rendered them unable to fold. In Lassa virus, the glycoprotein is synthesized as a precursor, preGP-C in the ER, which has 11 N-linked glycosylation sites. When systematically mutated, six of the sites were determined to be necessary for pre-GP-C cleavage into mature GP-1 and GP-2, suggesting the sites are necessary for the folding of the viral protein into a conformation sensitive to host cleavage enzymes SKI-1/S1P [8]. Newcastle disease virus, a bird pathogen, has a hemagglutinin-neuraminidase (HN) protein with four N-glycosylation sites. The loss of one of the four sites (G4) results in loss of virulence and decreased HN protein levels on the virion, and prevents the formation of the mature HN protein [9] [10]. The hantavirus Hantaan (HTNV) glycoprotein Gn-Gc also relies on glycosylation to fold correctly. The N glycan site at residue N134 was found to be crucial for protein folding, and a double mutant at glycosylation sites N235 and N399 also resulted in Gn misfolding. The endoplasmic reticulum chaperones calnexin and calreticulin were found to be involved in HTNV glycoprotein folding [11].

The best studied example of viral glycosylation that alters interactions with extrinsic molecules comes from the HIV glycoprotein, Env, which has extensive N-linked glycosylation making up half the proteins molecular weight [12]. The glycans in HIV are suggested to function as a “glycan shield”, sterically preventing the interaction of antibodies with the protein core of the Env, and therefore making neutralization less likely [13]. Removal of N-linked glycans on the HIV Env makes the virus more sensitive to neutralization by antibodies raised against wild-type Env [14]. The “glycan shield” theory relies on the fact that broadly neutralizing antibodies are generally not against glycans, but while this is generally true, there are many examples of

broadly neutralizing antibodies that form against glycans on viruses [12]. This may seem counterintuitive, but although glycans on viruses generally look like the glycans on host proteins, viral infections that are producing a lot of particularly extensively glycosylated proteins, can result in glycosylation that is not typically seen on host proteins. In the case of HIV, the glycan shield is made up of a higher proportion of high-mannose, the least modified and processed N-glycans, than most host proteins, and so specific high-mannose motifs can become a target of antibodies [15]. Regardless of type of glycan, most host proteins do not have the level of glycosylation seen in HIV, and so again during chronic infection people can develop broadly neutralizing antibodies directly against the glycan shield [15].

O-linked glycosylation in viruses is characteristically less well-studied. However, there are emerging studies on the role of O-linked glycans in viral lifecycles. In Herpes simplex virus type 1 (HSV1), the glycoprotein gC-1 contains numerous O-linked glycans. Two basic amino acids in a O-glycan rich region of the protein were mutated to alanines and increased the number of O-glycans in the region considerably, and resulted in a virus with impaired binding to cells expressing chondroitin sulfate (herpes auxiliary receptor) as well as reduced plaque size [16]. Herpes simplex virus type 1 glycoproteins have also been demonstrated to require elongated O-glycans for propagation of the virus [17].

Filovirus glycosylation. Filovirions enter target cells through interaction of their only particle surface protein, glycoprotein GP1,2, with cell-surface attachment factors and Niemann Pick C1 (NPC1) as the common endosomal entry receptor [18, 19]. GP1,2, is a class I fusion type 1 transmembrane protein that is expressed from the fourth of the seven filoviral genes, GP. In the case of marburgviruses, GP1,2 is the only GP expression product. In the case of cuevaviruses and ebolaviruses, the primary GP expression product is a non-structural secreted glycoprotein of unknown function (sGP) [20].

GP1 contains a mucin-like domain, a region of the protein that is modified with the addition of O-glycans, and the glycan cap, a region particularly dense in N-glycans. As with the viral proteins discussed above, there are both intrinsic (folding, trafficking and quaternary structure) and extrinsic (interactions with other molecules including receptors and the immune system) roles for the glycans on filovirus glycoproteins. The intrinsic role relates to protein folding and trafficking. Systematic mutation of the N-linked glycosylation sites in EBOV GP1 and GP2 determined which sites were essential for the folding of the glycoprotein, all 15 putative N-glycosylation sites of EBOV GP1 could be removed without compromising the expression of GP [21], but the removal of both sites in GP2 together dramatically decreased GP expression and abolished entry [22]. The loss of the mucin-like domain, GPdMuc, which will be discussed below, made the protein more sensitive to the loss of additional N-glycan sites outside the mucin-like domain, suggesting the mucin-like domain provides stability that renders the protein less sensitive to removal of glycans [21]. A similar systematic mutation of the 19 putative N-linked sites in MARV GP has not been performed, although the two N-linked glycosylation sites in GP2 are predicted to be conserved in MARV [22].

Loss of the N-linked glycans of EBOV were also shown to have extrinsic effects influencing immune interactions. The removal of all GP1 N-glycans outside the mucin-like domain enhanced antibody sensitivity [21], suggesting a role for EBOV glycans as a glycan shield. Loss of either of the two GP2 sites individually did not result in changes in antibody neutralization [22].

Certain extrinsic effects of glycosylation have been of interest to filovirus virologists for some time, as (prior to the discovery of NPC1 as the common endosomal receptor) various lectins, glycan-binding proteins, were discussed as potential filovirus entry receptors. Various C-type lectins were reported as entry sites for EBOV and MARV, mediated by the interaction of GP1,2. For example, the dendritic cell-specific ICAM3 (intercellular adhesion molecule 3)-grabbing non-integrin (DC-SIGN) and DC-SIGN-related receptor (DC-SIGNR) were shown to promote Ebola entry into dendritic cells and other myeloid cells [23-25]. Liver/lymph node specific ICAM-3-grabbing nonintegrin (L-SIGN), the DC-SIGN homologue, expressed not on dendritic cells but expressed on the surface of endothelial cells in the liver, lymph node sinuses, and placental villi, was also published as a potential Ebola cell attachment factor [26]. The human macrophage galactose- and N-acetylgalactosamine-specific C-type lectin (hMGL) was also shown to enhances the infectivity of filoviruses, with a lower affinity for Marburg than Ebola [27]. The interaction with hMGL depended on the presence of the mucin-like domain. Similarly, for Marburg virus, the asialoglycoprotein receptor was studied as a potential liver-specific receptor [28].

Perhaps surprisingly given the number of lectin interaction studies, the loss of the 15 glycosylation sites in EBOV GP1 significantly enhanced VSV pseudovirion transduction in Vero cells. However, Vero cell binding of EBOV is thought to be mediated by the phosphatidyl serine receptor TIM-1, and therefore it is possible that *in vivo* or with other cell types the binding of EBOV would be impacted by N-glycan loss [21].

The individual C-type lectins have particular affinity to different glycans, for example, DC-SIGN(R) and L-SIGN bind preferentially to high mannose glycans, whereas hMGL binds to glycans with galactose sugars, and asialoglycoprotein receptor binds galactose sugars without sialic acid termini. However, despite this information from binding studies, the actual character of the filovirus glycans themselves has not been well-studied. An early study of the Marburg virus glycoprotein determined glycan types associated with the virus produced in Vero E6 cells, by using radioactive D-[6-3H]glucosamine to label the N-acetylglucosamine, N-acetylgalactosamine and galactose sugars present on the GP [29]. Interestingly, using lectins that specifically bind to sialic acid, the same study also determined that sialic acid was not present in Marburg virus GP, but was present in Ebola virus GP [29]. However, this study was limited to using lectins to probe the proteins, rather than more precise mass spectrometry methods. Additionally, the impact that treatment with deoxynojiromycin (DNJ) an inhibitor of the formation of hybrid and complex glycans, had on Ebola binding to DC-SIGN(R), was investigated and it was determined that binding of the virus increased with the increase in high-mannose sugars [24].

Filoviruses are often used as an example of a virus family with extensive O-linked glycosylation [17], but the role of the O-linked glycans remains controversial. One theory is that the O-linked glycans act as part of the filovirus glycan shield on infected cells. The expression of filovirus glycoproteins prevents antibody access to host-proteins and can be used as a measure of “shielding”. Loss of the mucin-like domain reduced the shielding effects that were observed on the cells [30]. The steric shielding of host-cell proteins may have a biological role, as expression of filovirus glycoproteins prevents CD8 T cell access to major histocompatibility class I molecules [31].

EBOV GP1,2 is by far the most-studied of the filovirus GP1,2. EBOV GP1,2 contains multiple N-linked glycan sites; some sites are highly conserved in the other ebolavirus species, while others vary considerably between species [21]. EBOV GP1,2 also contains a mucin-like domain composed of both O- and N-linked glycans, and the glycosylation in this region is quite diverse between filovirus species. Detailed mapping of N-linked glycosylation of EBOV GP1,2 produced in a human cell line revealed complex bi-, tri-, and tetra-antennary glycans [32]. At the beginning of this study, however, there were no in-depth studies of the exact glycan structure of non-EBOV filovirus GP1,2s (such as SUDV, BDBV, TAFV, and RAVV). MARV GP1,which also contains N- and O-linked glycosylation [33], but appears to lack the sialylation found in EBOV GP [34]. Overall, there is currently a lack of understanding of glycan structure and linkage on most filovirus GP1,2s.

Filovirus vaccines. Although there were no approved treatments or vaccines against these viruses at the beginning of this study, several vaccines have shown protection in non-human primates, including those based on vesicular stomatitis virus, Venezuelan Equine Encephalitis (VEE) replicon particles, adenovirus, virus-like particles (VLPs), rabies vaccine, and human parainfluenza virus type 3 platforms [35-37]. Importantly, each of these vaccines include the GP1,2 of the filovirus species in question, with GP1,2 being the only immunogen in most of these platforms. The importance of the GP1,2 in vaccines is thought to be due to the generation of antibodies that block GP-mediated cell entry [35]; however, non-neutralizing GP1,2 antibodies have also been found to be efficacious in animal models [38]. Furthermore, CD8+ and CD4+ T cell responses are generated against GP1,2 after vaccination in rodents [39], non-human primates [40, 41], and humans [42, 43], and can be protective in animal model experiments [39, 40]. Thus, both antibody and T cell responses must be measured to appropriately gauge anti-filovirus GP1,2 immune responses. In late 2019, a vesicular stomatitis virus (VSV)-based vaccine was approved for use in humans for EBOV.

Filovirus GP1,2 glycosylation and immune responses. The different filovirus vaccine platforms are made in different cells, including mammalian [40, 41, 44-46], insect [46-48], and plant cells [49]. Wild-type insect cell-derived vaccines have altered N-linked glycosylation of GP1,2 compared to mammalian cell-derived GP1,2, in that they generate less complex N-linked glycosylation that is not trimmed or sialylated. O-linked glycosylation is similar but not identical to mammalian cells. The result is altered glycosylation of filovirus GP1,2s from the same vaccine platform generated in insect versus mammalian cells [46]. Plants (such as Nicotiana benthamiana, which has been used to generate many different VLP platforms [50]), have similar but not identical N-glycosylation as mammalian cells, and have limited O-glycosylation machinery, especially of the mucin-type O-glycosylation that is found in filovirus GP1,2s produced in mammalian cells. The mucin-like domain appears to be important for generation of immune responses, as many neutralizing antibodies target this site. n addition, it has been shown that the mucin-like domain of EBOV GP1,2 is required for potent induction of type I IFN and proinflammatory cytokine production in DCs *in vitro* [51], likely due to the differential interaction of GP1,2 with pattern recognition receptors [52]. The link between differential glycosylation and type I IFN responses in other systems has been well-documented. There is only limited, conflicting data available comparing the immunogenicity of insect cell-derived and mammalian cell-derived EBOV VLP [46, 53]. Conversely, glycosylation of filovirus GP1,2 is

also thought to have immune-evasive properties, due to shielding of GP1,2 epitopes; and EBOV GP1,2 lacking the mucin-like domain may induce antibody responses that cross-react with other ebolavirus species [54]. Therefore, understanding of filovirus GP1,2 glycosylation and subsequent immunogenicity is currently incomplete.

Together, these data suggest that differential glycosylation of filovirus GP1,2 has an important effect on immunogenicity in vaccination. Given that filovirus vaccines are being generated in mammalian, insect, and plant cells (resulting in very different glycosylation patterns), along with evidence suggesting glycosylation can affect immune responses, there is a need for understanding the effects of differential glycosylation of filovirus GP1,2s on induction of immune responses after vaccination.

All filovirus vaccines that have been evaluated at advanced stages are based on GP1,2 as the only (or primary) immunogen. Therefore, this proposal on the effect of GP1,2 glycosylation on immunity will apply to all filovirus vaccine platforms currently in advanced development. Furthermore, this work will be applicable to other vaccines based on glycoproteins from biological WMD, such as non-filoviral hemorrhagic fever viruses, as well as bacterial pathogens.

Gaps. There are few published data directly comparing immunogenicity of differentially glycosylated GP1,2 in filoviruses. The glycosylation patterns of filovirus GP1,2s are incompletely understood, since there is a profound lack of data on GP1,2 glycosylation for BDBV, SUDV, MARV, RAVV, and TAFV. This proposal will address these gaps and potentiate advanced filovirus vaccine development. Additionally, data on *in vitro* and *in vivo* immunogenicity of differentially glycosylated glycoproteins has broad applications for vaccine and therapeutic approaches directed against glycoproteins of other pathogens (including other hemorrhagic fever viruses). Therefore, this proposal will specifically address gaps for filovirus vaccine development and can serve as a reference for development of countermeasures against other biological Weapons of Mass Destruction (WMD).

Work to be performed. In Task 1, GP1,2 from the 6 known human pathogenic filoviruses (EBOV, SUDV, BDBV, TAFV, MARV, and RAVV) will be produced in mammalian, insect, and plant cells and assayed for O- and N-linked glycosylation patterns. These experiments will be performed since most of the published glycosylation mapping is for EBOV, and information is lacking for the glycosylation patterns of other filovirus GP1,2s; furthermore, there is little known about the glycosylation patterns on filovirus GP1,2s generated in insect or plant cells for vaccine purposes. In Task 2, *in vitro* immunogenicity will be performed on differentially glycosylated filovirus GPs by incorporating them into virus-like particles (VLPs) generated in mammalian, insect, or plant cells. The VLPs bearing differentially glycosylated GPs will then be incubated with dendritic cells (DCs), and activation of DCs will be assessed. In Task 3, mice will be vaccinated with these differentially glycosylated GP1,2s, and antibody titers and T cell responses will be assessed following vaccination to determine *in vivo* antigenicity of differentially glycosylated GPs. Figure 1 summarizes the experiments to be performed.

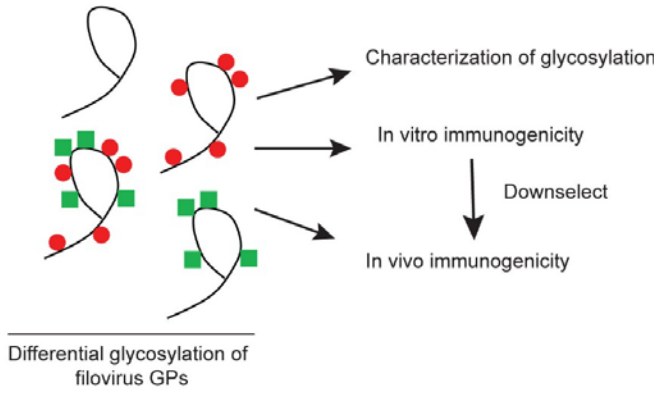


Figure 1. Overview of proposal. Filovirus GP glycosylation will be defined, and differentially glycosylated GPs will be assessed for *in vitro* and *in vivo* immunogenicity.

CHAPTER 2. ELUCIDATION OF THE GLYCAN STRUCTURE AND LINKAGE OF EBOV, BDBV, SUDV, TAFV, MARV, AND RAVV GP1,2 PRODUCED IN MAMMALIAN, INSECT, AND PLANT CELLS

Background. All current filovirus vaccine platforms utilize the highly glycosylated filovirus attachment protein (glycoprotein, or GP1,2) as an antigen. However, the different vaccine platforms are expressed in various cell types that endow differential glycosylation on GP1,2s, and the importance of the various types of glycosylation in promoting immunogenicity of filovirus GP1,2s is not well known. To date, there has been no direct comparison in GP1,2 immunogenicity between these vaccine platforms, and glycosylation pattern analysis of filovirus GP has primarily been limited to only 1 of the 6 species known to be pathogenic in humans. Therefore, the overall goal of this proposal is to delineate the importance of the various glycosylation patterns of filovirus GP in vaccine immunogenicity in *in vitro* and *in vivo.* The first step was to analyze N- and O-linked glycosylation on filovirus GP1,2s expressed in mammalian, insect, and plant cells.

Production of filovirus GP1,2 in mammalian, insect, and plant cells.

Construction of filovirus GP expression plasmids for expression in mammalian 293T cells. In order to complete the proposed tasks, we first had to establish protocols to express and purify filovirus GPs in a manner that allowed for acceptable levels and purity for mass spectrometry analysis. GPs from the pathogenic ebolaviruses (Zaire ebolavirus (EBOV), Sudan ebolavirus

(SUDV), Bundibugyo ebolavirus (BDBV), and Tai Forest ebolavirus (TAFV)) are expressed primarily as soluble GP (sGP). sGP's function is unknown, and only full-length GP is found on the virion surface and is the version of the protein used in vaccination. Full-length GP is generate in infection when the viral polymerase stutters and adds an additional adenosine in a 7-adenosine stretch of GP, causing a frame-shift that results in generation of full-length GP. This full-length GP is called preGP. PreGP is then cleaved by furin into GP1 and GP2, which dimerize. Three of these dimers form full-length, transmembrane-bound GP1,2.

First, we had to design GP1,2s that were amenable to expression at significant levels. Early experiments consisting of expression full-length EBOV-Mayinga GP1,2 were successful, but did not result in high enough yields for the types of analysis we needed. To mitigate this, we altered the GPs as follows (see Figure 2):

a) The transmembrane region of GP1,2 was deleted. This allows for secretion of GP into the media, which allows for higher expression levels, decreased cellular toxicity, and simplification of protein harvest.
b) A 6xHis tag was added to the C-terminus of GP1,2 to aid in purification.
c) The furin cleavage site (which induces cleavage of the preGP polypeptide into two proteins, GP1 and GP2, which dimerize) was mutated with a single arginine-to-lysine change. This blocks the cleavage of the preGP into GP1 and GP2, but, importantly, does not affect the folding or function of the protein. In this manner we minimized concerns about the GP1 and GP2 subunits dissociating during purification.
d) An additional adenosine was added to make 8A GP, which generates only full-length GP1,2.

GPs were constructed in this manner for EBOV-Mayinga (the original 1976 strain), EBOV-Makona (the 2014-2015 strain), SUDV, BDBV, and TAFV. MARV and RAVV were made similarly, but without the 7A/8A change (marburgviruses do not have this property). GP1,2s were cloned into the pcDNA3.1+ mammalian expression vector.

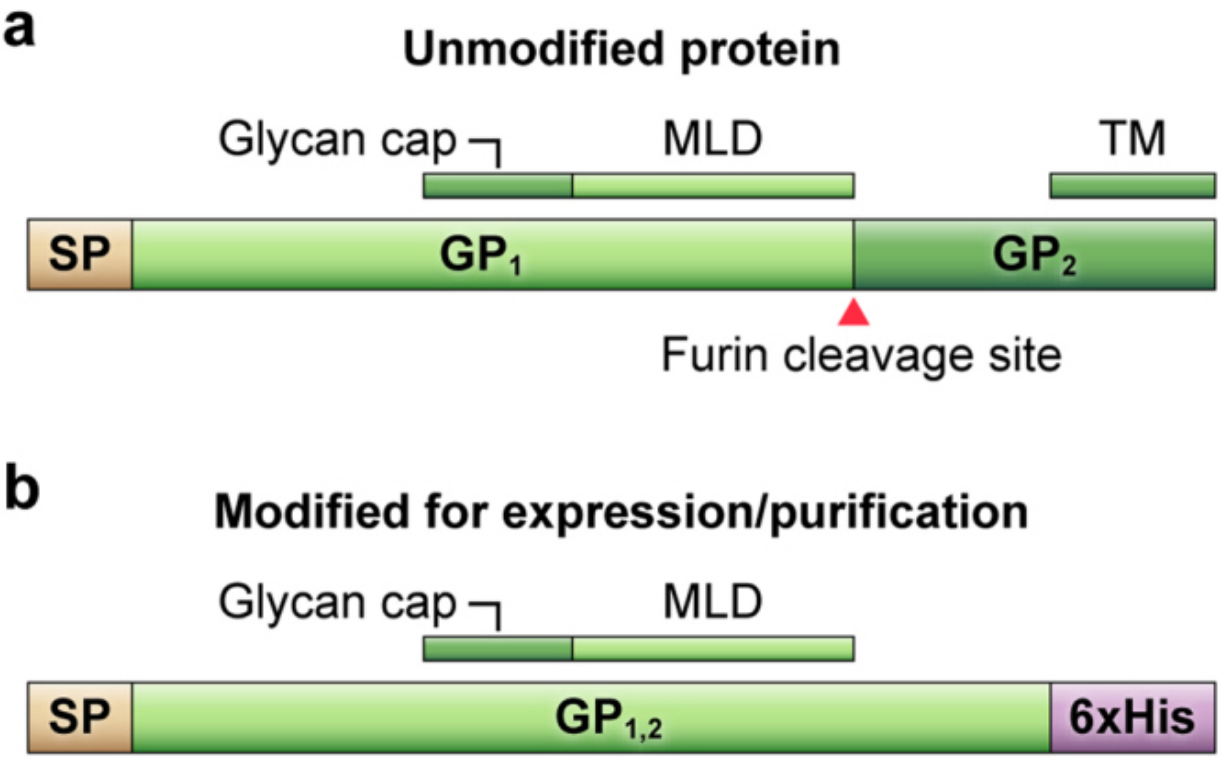


Figure 2. Modifications made to the filovirus *GP* genes for expression and purification of $GP_{1,2}$s. (**a**) *GP* gene co-transcriptional mRNA editing results in an 8A mRNA encoding a preproprotein consisting of a signal peptide (SP) and preGP (later cleaved by furin into subunits GP_1 and GP_2). The GP_1 portion of preGP contains a glycan cap and mucin-like domain (MLD), whereas the GP_2 portion of the protein anchors the $GP_{1,2}$ trimer in membranes via a transmembrane domain (TM). (**b**) Constructs used for this study are expressed from engineered genes to only express the $GP_{1,2}$ ORF in the absence of the furin cleavage site and TM replaced with 6xHis tag. Published in [1].

Establishment of transfection protocols for optimal expression of filovirus GPs. A number of different transfection reagents and protocols were tested for optimal expression of EBOV-Mayinga GP1,2 in 293T cells:

a) jetPrime and lipofectamine transfection reagents
b) different amounts of transfection reagent
c) different amounts of plasmid DNA
d) different harvesting times post-transfection

These experiments led us to choose jetPrime transfection reagent (similar efficacy, less toxicity, lower price) with 50 ug of filovirus GP plasmid with 90 uL of jetPrime reagent per 150 cm2 flask for 3 days as our standard protocol (data not shown).

Purification of GP1,2 from supernatants of transfected mammalian 293T cells. Using the above reagents and protocols, we were able to generate sufficient quantities of GP1,2 for analysis. However, the samples need to be very pure for proper mass spectrometry analysis. We began by utilizing nickel columns attached to a peristaltic pump to purify the 6xHis-tagged GP1,2. The nickel bound the 6xHis-tagged GP1,2, and after washing with excess buffer, the GP1,2 was eluted with buffer containing imidazole, which binds the nickel and causes elution of GP1,2. Early experiments showed that this resulted in significant purification of supernatants containing EBOV-Mayinga GP1,2. As shown in Figure 2, the GP1,2 was not visible by coomassie staining in supernatants; however, purification with a nickel column resulted in quite efficient recovery and purification of GP1,2. However, a contaminating protein of approximately 80 kDa co-eluted with EBOV-Mayinga GP1,2; this protein did not react to the anti-EBOV antibody, suggesting a non-EBOV GP1,2 contaminant (Figure 3b). This contaminating protein was glycosylated (as evidenced by glycoprotein-specific staining with the Periodic Acid-Schiff (PAS) method, data not shown), therefore would complicate analysis of glycans by mass spectrometry.

Multiple protocol modifications were attempted to remove this contaminant from the elution. The simplest method would be to simply excise the GP1,2 band from the gel; however, upon consultation with our mass spectrometry collaborators, this would likely reduce the ability to ascertain the presence of sialic acids and could reduce accuracy of the glycan analysis. Therefore, the following protocol modifications were attempted:

a) Dialysis of the elution with 100 kDa, 300 kDa, or 1,000 kDa pore size membranes
b) Doubling the wash buffer volume
c) Increasing imidazole concentration in wash buffer
d) Increasing imidazole concentration in sample
e) Using a his-streptavidin tag-enrichment kit with a streptavidin column
f) Replacing the nickel column with a cobalt column
g) Use of 8M Urea to dissociate a possible protein-protein interaction
h) analysis on a native gel to detect interaction between GP1,2 and the contaminant (no interaction was found)
i) Use of a size exclusion column to remove the contaminant

None of these approaches removed the contaminating protein. Mass spectrometry analysis revealed the contaminant was bovine serotransferase, found in the fetal bovine serum used for cell growth. This protein did not have a 6xHis region, but is known to have an affinity for metals, thus explaining the contamination.

We found a commercially available affinity column made of a slurry with an anti-6xHis antibody. When using this column, we found that we could obtain high-purity filovirus GP1,2 recovery from supernatants of transfected 293T cells. Although this product was expensive and had reduced yield, we opted for its use and were able to generate samples of sufficient purity for

analysis. Figure 3 outlines the protocol we developed for production and purification of filovirus GP1,2 in mammalian cells.

In our initial experiments, MARV/Musoke $GP_{1,2}$ could not be expressed using pcDNA3.1^{+} in mammalian cells. The sequence of the MARV/Musoke GP gene was truncated at residues encoding amino acid 636 (preGP numbering), i.e., upstream of the predicted transmembrane region [55]. This sequence was identical to the sequence used in the MARV/Musoke GP gene cloned into the baculovirus that successfully produced MARV $GP_{1,2}$ in the insect cell system. Likewise, expression of the MARV/Musoke GP gene encoding a $GP_{1,2}$ truncated to encode only amino acid residues 1–648 or 1–644 was unsuccessful in the pcDNA3.1^{+} background (Table 1). In contrast to commercially available MARV $GP_{1,2}$ expression plasmids, which use codon-optimized GP genes for expression in mammalian cells, all our modified sequences were based on the wild-type GP sequences. Transfer of the MARV GP gene sequence from pcDNA3.1^{+} into the pCAGGs expression vector did not lead to successful expression of MARV $GP_{1,2,}$ despite successful expression of unmodified full-length MARV GP by pCAGGs and not pcDNA3.1^{+} (Table 1). Similarly, we had difficulty expressing RAVN GP in mammalian cells. However, MARV/Angola (1–648) $GP_{1,2}$ could be successfully expressed from pCAGGs. Therefore, we purified MARV/Angola GP1,2 and analyzed glycan patterns on this protein.

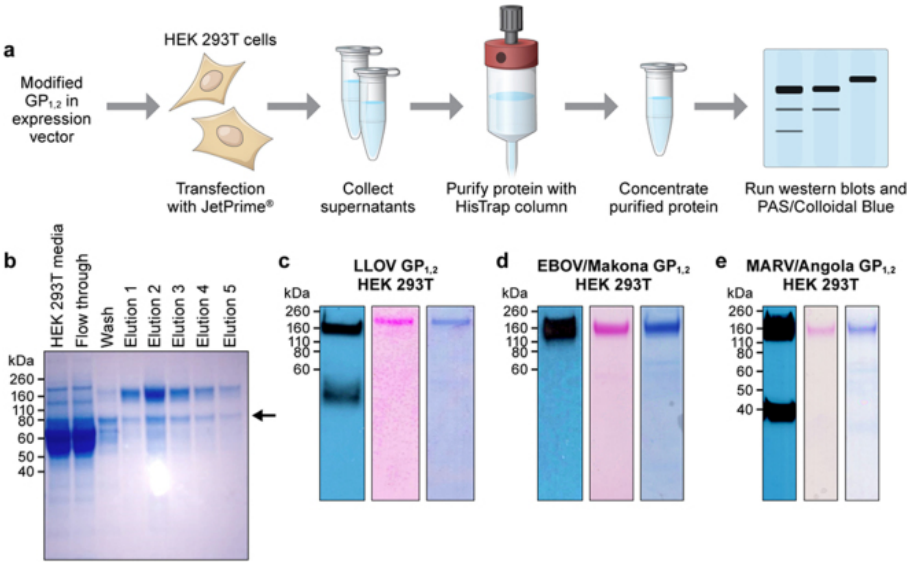


Figure 3. Flowchart of scheme for expression and purification of filovirus GP1,2 in mammalian 293T cells. Western blots verifying purity of $GP_{1,2}$s obtained from HEK 293T cells. (**a**) Schematic of the filovirus $GP_{1,2}$ production workflow from HEK 293T cells. (**b**) EBOV/Makona $GP_{1,2}$ from HEK 293T cells was loaded onto a HisTrap Excel column. A contaminant was detected by western blot and PAS and colloidal blue stains, as shown by an arrow, of the elutes from the column of media only, flow through wash, and five 1-minute dilutions. (**c**) Western blot and PAS and colloidal blue staining of LLOV $GP_{1,2}$ from HEK 293T cells. (**d**) Western blot and PAS and colloidal blue staining of purified EBOV/Makona $GP_{1,2}$ from HEK 293T cells. (**e**) Western blot and PAS and colloidal blue staining of MARV-Angola $GP_{1,2}$ from HEK 293T cells. Published in [1].

Expression vector	*Expressed constructs*	*HEK 293T expression*
pcDNA3.1$^+$	full-length MARV/Musoke $GP_{1,2}$	No
pcDNA3.1$^+$	MARV/Musoke $GP_{1,2}$ (1–648)[a]	No
pcDNA3.1$^+$	MARV/Musoke $GP_{1,2}$ (1–644)[a]	No
pcDNA3.1$^+$	MARV/Musoke $GP_{1,2}$ (1–636)[a]	No
pCAGGS	full-length MARV/Musoke $GP_{1,2}$	Yes
pCAGGS	MARV/Musoke $GP_{1,2}$ (1–636)[a]	No
pCAGGS	MARV/Angola $GP_{1,2}$ (1–648)[a]	Yes

Table 1. Differential expression of MARV GP1,2s in mammalian cells.

Production of filovirus GP1,2 in insect Sf9 cells. Insect Sf9 cells have lower transfection rates than mammalian 293T cells; usually protein expression in these cells is achieved by use of an infectious baculovirus carrying the gene of interest. We have constructed these baculoviruses by cloning the filovirus GP1,2s above into a shuttle vector that is recombined with a baculovirus bacmid (Bac-to-bac system, ThermoFisher). These baculoviruses carrying filovirus GP1,2 genes are then used to infect Sf9 cells, which generate GP1,2 protein that is secreted in the media.

Purification of GP1,2 from supernatants of baculovirus-infected Sf9 insect cells. The insect cell media we used does not require fetal bovine serum; therefore, we were able to use the more efficient and less expensive nickel columns for GP1,2 purification. These systems are now routinely used in our lab.

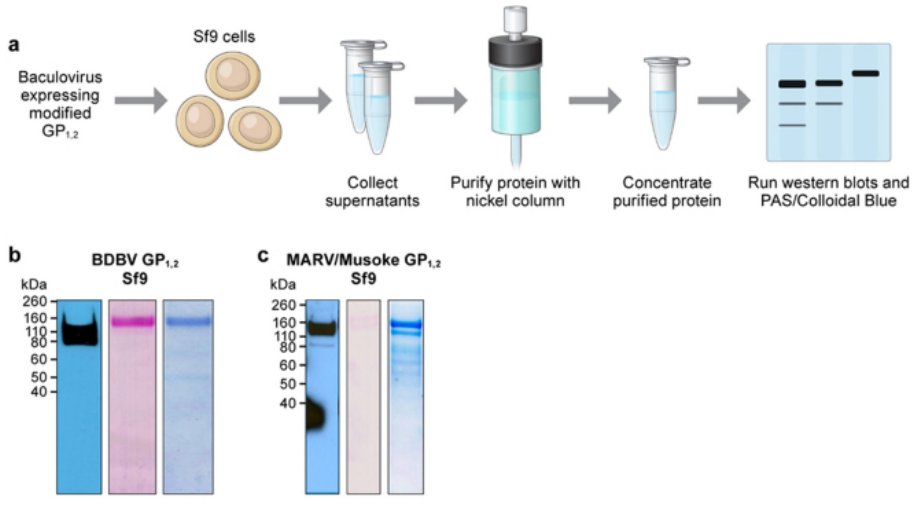


Figure 4. Flowchart of scheme for expression and purification of filovirus GP1,2 in insect Sf9 cells. Western blot gels verifying the purity of modified $GP_{1,2}$s obtained from Sf9 insect cells. (**a**) Schematic of the Sf9 filovirus $GP_{1,2}$ production workflow. (**b**) Western blot, PAS, and colloidal blue staining of purified BDBV $GP_{1,2}$ from Sf9 cells. (**c**) Western blot, PAS, and colloidal blue staining of MARV $GP_{1,2}$.

Construction of filovirus GP expression plasmids for expression in plants. In order to complete the proposed tasks, we first had to establish protocols to express and purify filovirus GPs in plants. GP1,2s were constructed as above for EBOV-Mayinga (the original 1976 strain), EBOV-Makona, SUDV, BDBV, TAFV, RAVV, and MARV, and sent to PlantForm Corporation for cloning of these genes into a suitable expression system for plants. Unique Polymerase Chain Reaction (PCR) primers were designed for each coding sequence. Each forward primer included a restriction enzyme cut site, a plant Kozak sequence (CAACA) and an *Arabidopsis* basic chitinase signal peptide. Each reverse primer included a stop codon followed by a restriction enzyme cut site. The DNA sequence of each glycoprotein was amplified using PCR then digested with the appropriate restriction enzymes and cloned into PlantForm's *vivo*XPRESS® vector system. Two genes were transiently co-expressed in wild-type (TW16) *N. benthamiana* to produce each individual filovirus GP1,2: a filovirus GP1,2 and P19, a suppressor of gene silencing (Garabagi et al., 2012a). Each gene was expressed using a single binary vector (Figure 5). In addition, an antibody was coexpressed with P19 as a positive control for expression while P19 was expressed alone as a negative control.

Agrobacterium tumefaciens strain At564 was individually electroporated with each of the GP1,2

DNA vectors. Colony PCR was used to confirm plasmid identity of transformants. Glycerol stocks were prepared for each *A. tumefaciens* strain and stored at -80°C. The glycoprotein expression vectors were verified after transformation into *A. tumefaciens* stain At564. Transformants were screened using colony PCR and primers corresponding to eachindividual glycoprotein coding sequence (Figure 6).

Transient expression was performed, as described in Garabagi et al. (2012b), using the wild-type *N. benthamiana* host line TW16. Briefly, *Agrobacteria* strains containing the glycoprotein and control expression vectors were grown to exponential phase in LB-Miller broth with rifampicin and carbenicillin (50 μg/mL each). *Agrobacteria* cultures were diluted to an OD600 of 0.2 in infiltration buffer (10 mM MES pH 5.5, 10 mM MgSO4). The foliar areas of 12 plants per treatment were submerged in the *Agrobacteria* solution and exposed to a reduced atmospheric pressure for abouttwo minutes before slowly returning the chamber to normal atmospheric pressure. Plants were returned to the greenhouse for five, seven or nine days, after which all foliar material was removed from the plants and stored at -20°C in sealed freezer bags until extraction.

Total soluble protein extraction and small-scale purification of His-tagged glycoproteins**.** Frozen plant tissue was homogenized with a Polytron homogenizer in 100mM Tris-HCl, 150mM NaCl, 10% glycerol & 0.1% TritonX-100 (1:3 w/v). The extract was clarified by centrifugation at 14,000 rpm for 30 minutes at 4°C. The biomass from the biological replicates were pooled for each treatment. Four grams of biomass from each timepoint for each treatment were homogenized and clarified as described above. Clarified total soluble protein (TSP) extracts were incubated for one hour with 20 mM imidazole-equilibrated Ni sepharose from His SpinTrap columns (GE Healthcare Life Sciences). The resin was washed three times with 20 mM imidazole followed by a single 100 mM imidazole wash. Glycoproteins were eluted in two fractions using 500 mM imidazole. The two eluates from each treatment were pooled.

A Bradford assay was used to determine the concentration of protein in the eluates. An equal amount of protein from each purified treatment was resolved by SDS-PAGE and either stained

with Coomassie brilliant blue or transferred to polyvinylidene difluoride (PVDF) membrane for

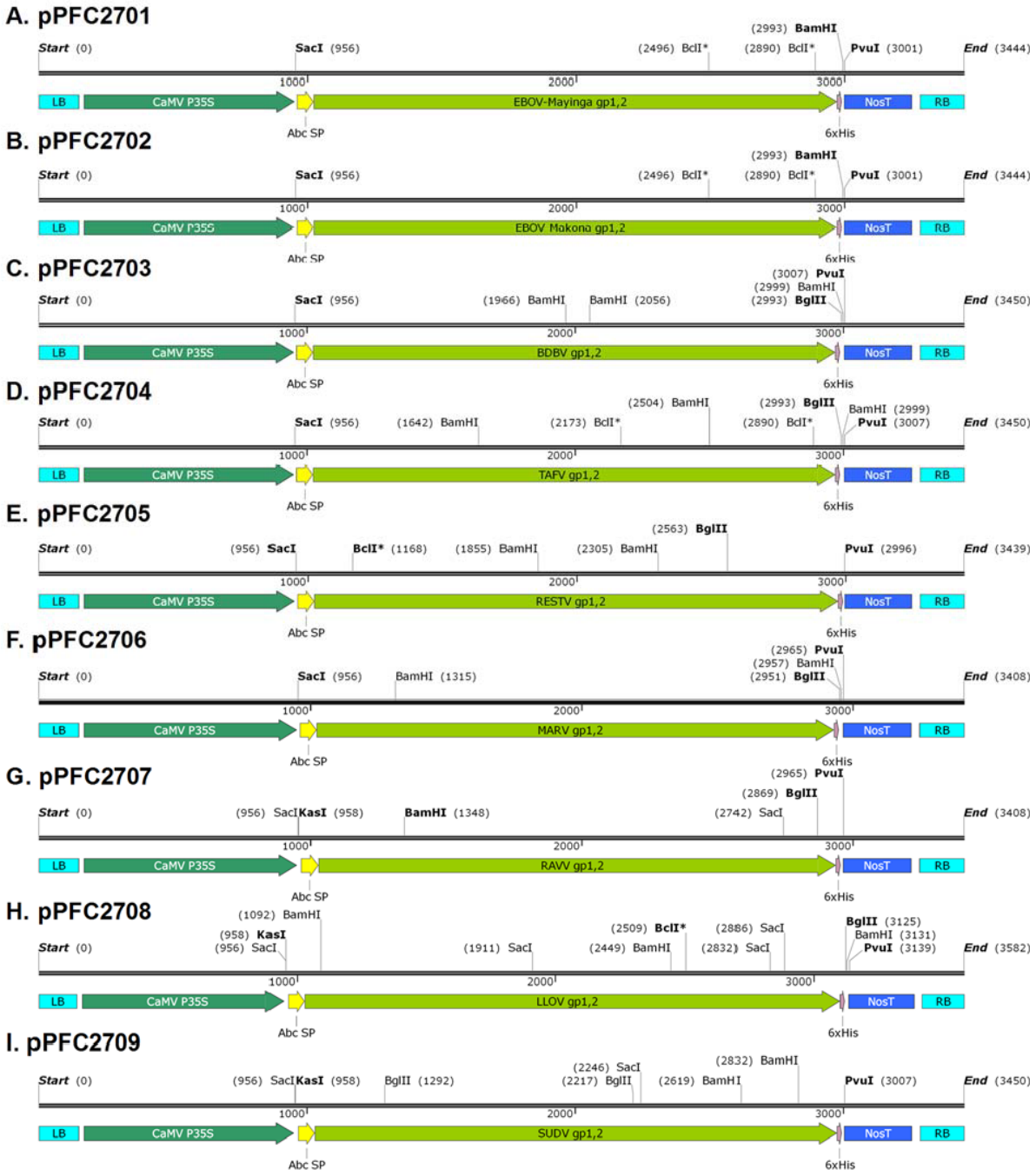


Figure 5. Construction of filovirus GP1,2 expression vectors for plant production.

Western blotting. Glycoprotein yield was estimated using densitometry. Briefly, Image J densitometry was used to estimate the percentage of the GP1,2s in the respective eluate.

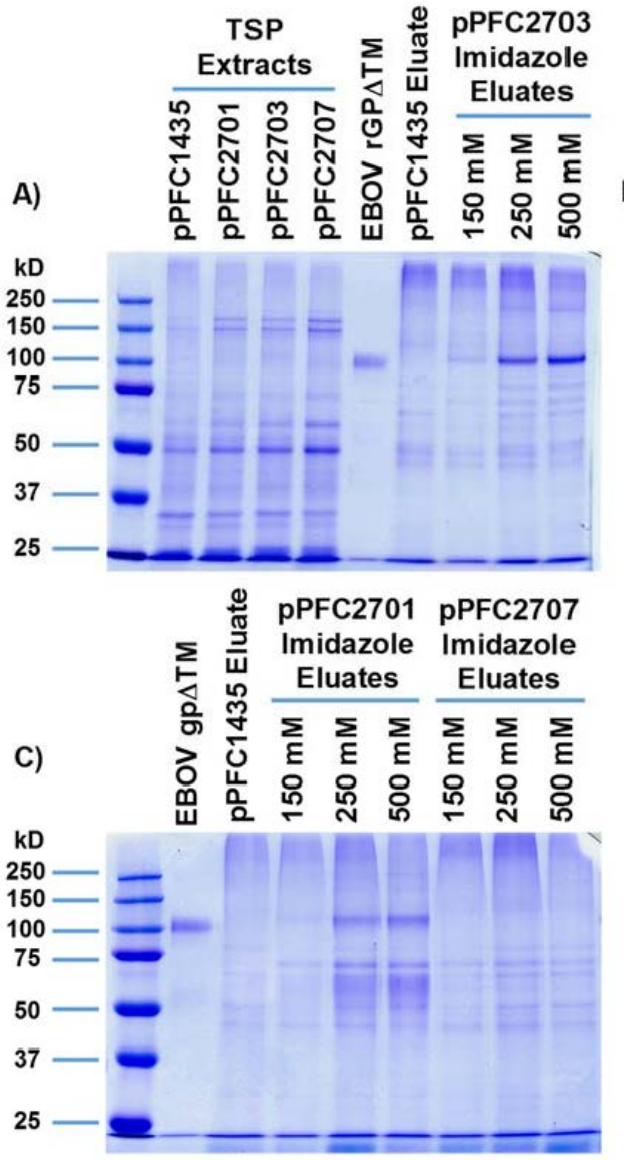


Figure 7. SDS-PAGE analysis of trial purification of three glycoproteins. EBOV-Yambuku, BDBV and RAVV (pPFC2701, pPFC2703 and pPFC2707, respectively) day five samples were extracted and subjected to IMAC using His SpinTraps with three sequential elution steps using increasing concentrations of imidazole, 150, 250 and 500 mM. Samples were electrophoresed under reducing conditions and stained with Coomassie brilliant blue (A and C). Commercially available recombinant EBOV gpΔTM was used as a positive control (2 µg for gel).

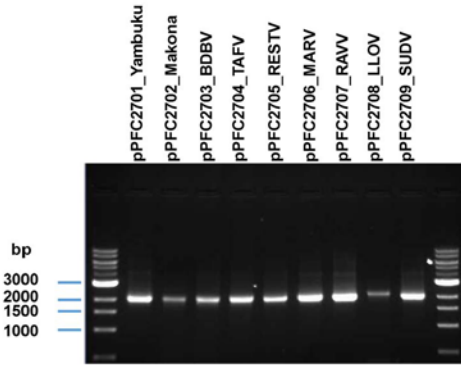


Figure 6. Confirmation of filovirus GP1,2 expression vectors.

<u>Expression analysis of control antibody and glycoproteins.</u> An antibody was co-expressed with P19 as a control for general expression. Antibody expression was quantified in the TSP extract using the BLItz biosensor assay. Antibody expression was an average of 378 mg/kg (*n=4*), which is typical for this antibody.

<u>Analysis of glycoprotein expression.</u> The filovirus GP1,2s were co-expressed with P19, a suppressor of gene silencing. Biomass was harvested 5, 7 and 9 days post-infiltration. Three of the eight glycoproteins were selected for initial evaluation and optimization: EBOV-Yambuku, BDBV and RAVV (pPFC2701, pPFC2703 and pPFC2707, respectively). Clarified TSP extracts were analyzed by SDS-PAGE (Figure 7 A & C), Western blotting, and enzyme linked immunosorbent assay (ELISA) (data not shown). Glycoprotein accumulation of appropriate molecular weight was not visible on the gel or Western blot. The positive control,

EBOV rGPΔTM, was visualized on both the gel and Western blot.

In parallel, the three extracts were subjected to a trial purification (Figure 7A and C). Three different elution conditions were evaluated: 150, 250 and 500 mM imidazole. BDBV glycoprotein was detected in each eluate on the gel but only the 250 and 500 mM eluates on the Western blot, indicating the anti-His antibody was not very sensitive. In addition, the EBOV-Yambuku glycoprotein was visible in all three eluates. However, the RAVV glycoprotein was not visible in the eluates. An ELISA was also performed using the anti-His antibody. Unfortunately, the anti-His antibody reacted non-specifically to plant TSP rendering the data unreliable. Based on these preliminary analyses the recovery levels, from highest to lowest, were BDBV > EBOV-Yambuku > RAVV. Unfortunately, the Ni-beads recovered a large amount of endogenous proteins from TSP extracts. Further optimization was required to minimize the binding of endogenous plant proteins.

PlantForm optimized the purification of the glycoproteins using BDBV (pPFC2703) as a model. Briefly, two different concentrations of imidazole were evaluated for both the equilibration and wash steps. In Figure 8A, Ni-beads were incubated with TSP in the presence of either 75 or 100

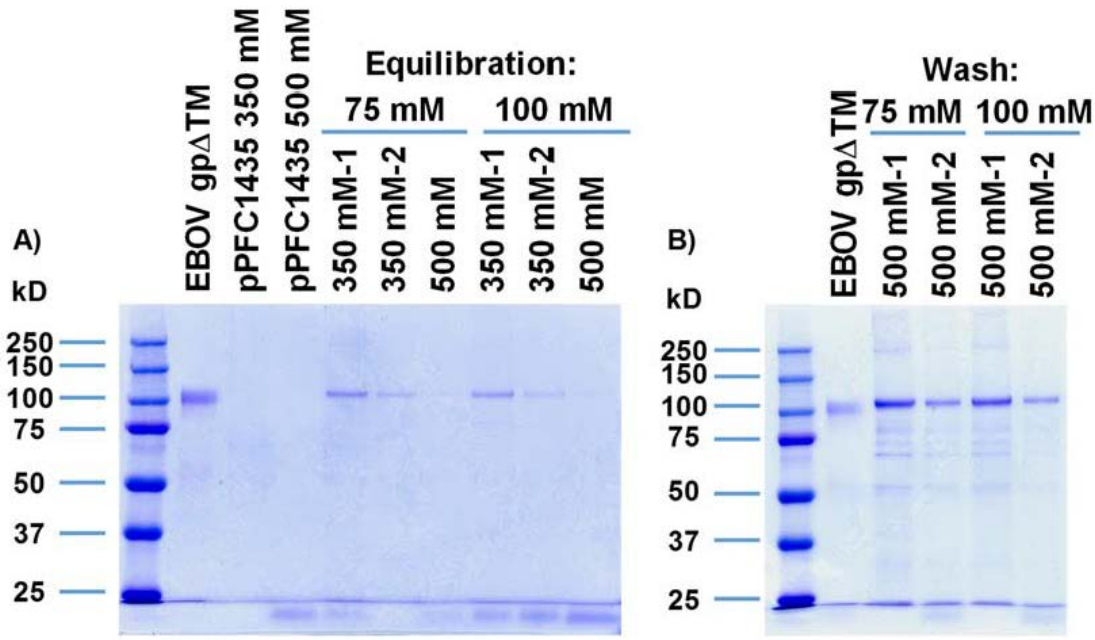


Figure 8. SDS-PAGE analysis of IMAC optimization. P19 and BDBV glycoprotein (pPFC1435 and pPFC2703, respectively) were purified using Ni-beads. (A) TSP was incubated with Ni-beads in the presence of 75 or 100 mM imidazole and washed with 100 mM imidazole. Two sequential elution steps were performed at 350 mM imidazole followed by a 500 mM elution. (B) TSP was incubated with Ni-beads in the presence of 40 mM imidazole followed by three washes with 75 or 100 mM imidazole. Two sequential elutions were performed with 500 mM imidazole. Eluates were electrophoresed under reducing conditions and stained with Coomassie Brilliant blue. Commercially available recombinant EBOV gpΔTM was used as a positive control (2 µg).

mM imidazole to minimize non-specific binding of plant proteins. Beads were washed with 100

mM imidazole and eluted twice with 350 mM imidazole followed by a single 500 mM imidazole elution. Both equilibration conditions resulted in highly pure fractions of glycoprotein. However, most of the BDBV glycoprotein was detected in the first 350 mM imidazole eluate and overall yield was lower than expected (Figure 8A). For both equilibration conditions, approximately 10 µg of glycoprotein was obtained in the first 350 mM elution, which translates to a yield of ~5 mg/kg.

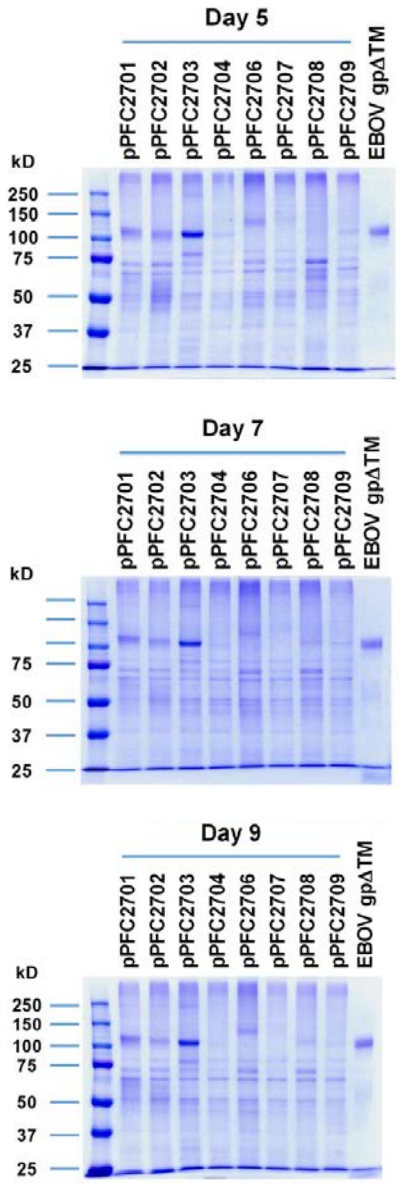


Figure 9. SDS-PAGE analysis of partially purified Filovirus glycoproteins. Nickel beads were incubated with clarified total soluble protein extracts in the presence of 20 mM imidazole for one hour at 4°C, followed by three washes with 20 mM imidazole and one 100 mM imidazole wash. Bound proteins were eluted twice with 500 mM imidazole. Eluates were combined and electrophoresed under reducing conditions and stained with Coomassie brilliant blue. Commercially available recombinant EBOV gpΔTM (2 µg) was used as a positive control.

In Figure 8B, Ni-beads were incubated with TSP in the presence of 40 mM imidazole and washed with either 75 or 100 mM imidazole in an attempt to increase recovery while minimizing nonspecific binding of plant proteins. Two elutions at 500 mM imidazole were performed and resulted in an increase in glycoprotein recovery and non-specific proteins. For the 75 mM and 100 mM imidazole washes a total of 84.8 and 71.4 µg, respectively, of glycoprotein was recovered, indicating a yield of approximately 21.2 and 17.4 mg/kg, respectively. Unfortunately, equilibration in only 40 mM imidazole was not sufficient to reduce non-specific binding even when washed with a higher concentration of imidazole (100 mM).

Based on the optimization results, PlantForm selected the higher stringency purification conditions in order to obtain a purer product for delivery to the Univeristy of New Mexico. Ni-beads were incubated with TSP in the presence of 75 mM imidazole for one hour at 4°C and washed three times with 100 mM imidazole. Two elutions using 350 mM imidazole were conducted. Unfortunately, the recovery of glycoprotein was exceptionally low for nearly all treatments. Therefore, PlantForm performed all remaining purifications using a less stringent process. Briefly, Ni-beads were incubated with TSP in 20 mM imidazole, washed three times with 20 mM and once with 100 mM imidazole, then eluted twice with 500 mM imidazole. Figure 9 shows the eluates from day 5, 7, and 9 materials analyzed by SDS-PAGE. Due to the high non-specific binding of the anti-His antibody, densitometry was used to estimate the recovery of each glycoprotein at each time point (Table 2). In all cases, yield was greatest on day 7. The best yields were achieved with BDBV (pPFC2703) and EBOV-Yambuku, 21.8 and 12.8 mg/kg, respectively. The majority of treatments resulted in yields of <5 mg/kg.

Filovirus GP1,2	Vector	Yield (mg/kg) Day 5	Day 7	Day 9
EBOV Yambuku	pPFC2701	6.3	12.8	12.3
EBOV Makona	pPFC2702	3.0	6.8	5.6
BDBV	pPFC2703	15.3	21.8	19.1
TAFV	pPFC2704	0.8	0.9	0.4
RESTV	pPFC2705	n/a*	n/a*	n/a*
MARV	pPFC2706	3.6	4.4	3.2
RAVV	pPFC2707	0.3	0.6	0.1
LLOV	pPFC2708	0.3	2.3	2.1
SUDV	pPFC2709	1.2	1.7	0.8

Table 2. GP1,2 yield after purification.

Analysis of filovirus GP1,2 glycosylation produced in mammalian, plant, and insect cells.

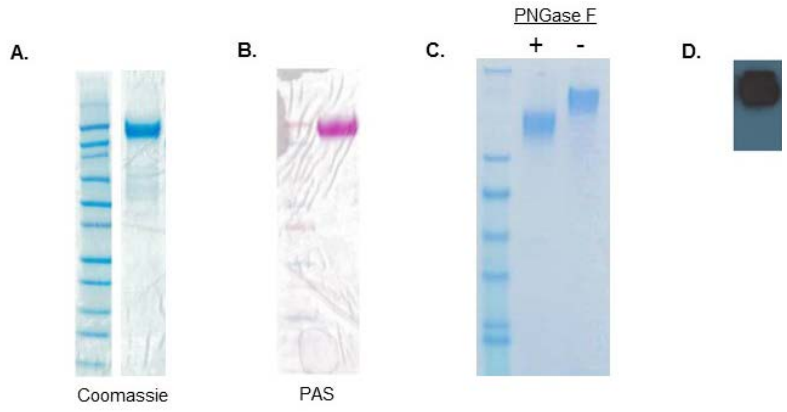


Figure 10. Purification of EBOV-Mayinga GP1,2 produced in mammalian 293T cells. A. Coomassie stain. B. PAS glycoprotein-specific stain. C. PNGase analysis showing presence of glycans in GP1,2. D. Western blot.

Glycan analysis of EBOV-Mayinga GP1,2 produced in mammalian 293T cells.

EBOV-Mayinga GP1,2 (RefSeq GeneID: 911829) was modified as in Figure 2 and expressed in mammalian 293T cells. Purified GP1,2 was confirmed for purity and specificity by coomassie staining (Figure 10A), which stains all proteins; periodic acid schiff stain (PAS), which stains only glycosylated proteins (Figure10B); and Western blotting with an anti-EBOV antibody (Figure 10D). Two individual experiments were performed, and the purified proteins were combined and analyzed by MALDI-TOF mass spectrometry (Proteodynamics). N-linked glycans on GP1,2 were released by incubation with PNGase F, which was confirmed by reduced protein mass (Figure 10D). Mass spectrometry analysis of the released glycans revealed a large number of different N-linked glycan structures with a wide range of frequency in GP1,2 (Figure 11).

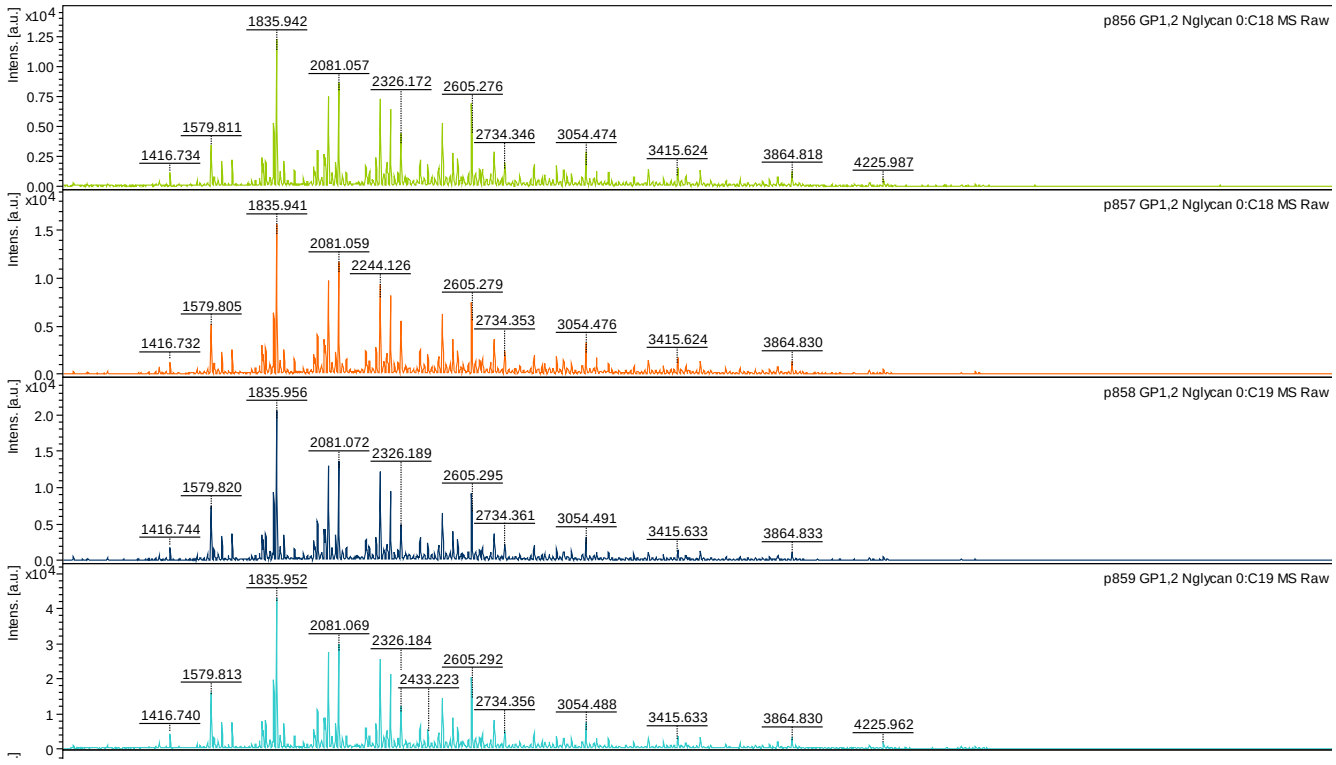


Figure 11. MALDI-TOF MS spectra of N-linked glycans from EBOV-Mayinga GP1,2. Each set of peaks is taken from a different section of the sample, showing consistancy of the assay.

Analysis of these data revealed 49 different N-linked glycan structures, including high mannose type structures from $(Man)_4(GlcNAc)_2$ to $(Man)_8(GlcNAc)_2$, hybrid glycans having only one GlcNAc residue in the antenna, and bi-, tri- and tetra-antennary complex glycans (Figure 12 shows graphical representation; Table 3 shows the data in a list format). This antennary profile is in good agreement with data published for EBOV-Mayinga GP1 (Ritchie et al., Rapid Commun Mass Spectrom. 2010 March 15; 24(5): 571-585). Dominant individual *N*-glycan structures are bi-antennary glycans: $(GlcNAc)_2(Fuc)_1$ + $(Man)_3(GlcNAc)_2$ (m/z 1835) at 12%, $(GlcNAc)_3(Fuc)_1$ + $(Man)_3(GlcNAc)_2$ (m/z 2081) at 8%, $(Gal)_1(GlcNAc)_2(Fuc)_1$ + $(Man)_3(GlcNAc)_2$ (m/z 2040) at 7.5% and $(Gal)_2(GlcNAc)_2(Fuc)_1$ + $(Man)_3(GlcNAc)_2$ (m/z 2244) at 6%.

The sialylation level of EBOV-Mayinga GP1,2 expressed in 293T cells was at 14%, which is significantly higher than the value for GP1 indicated in the above mentioned publication (2.3%). This could be due to the fact that the Ritchie et al. manuscript removed the N-glycans after SDS-PAGE of the samples; this procedure has been known to significantly reduce the ability to detect sialic acid as compared to the method we have used.

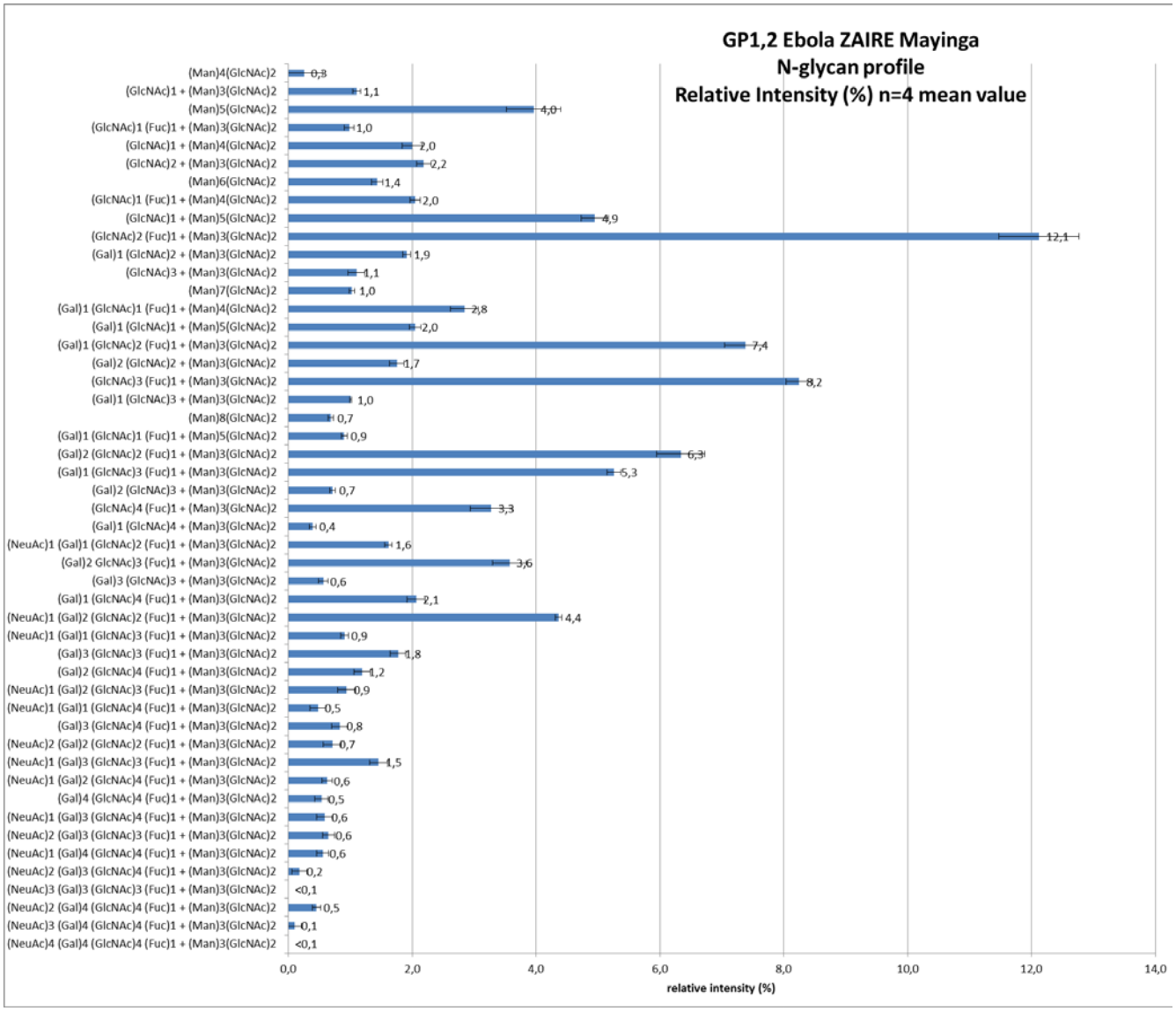


Figure 12. MALDI-TOF MS analysis of N-linked glycans from EBOV-Mayinga GP1,2. Relative intensity (normalized to 100%) of 49 individual N-glycans are shown. Error bars represent standard error of 4 separate spots of the pooled GP1,2 from two separate experiments.

		GP1,2 Ebola ZAIRE Maying		
Structure N°	$[M+Na]^+$ theoretical monoisotopic mass	Rel. Intensity (%),mean value (n=4)	Rel. Intensity (%) stand. Deviat.	N -glycan structure
1	1375,7	0,3	0,3	$(Man)_4(GlcNAc)_2$
2	1416,7	1,1	0,1	$(GlcNAc)_1 + (Man)_3(GlcNAc)_2$
3	1579,8	4,0	0,4	$(Man)_5(GlcNAc)_2$
4	1590,8	1,0	0,1	$(GlcNAc)_1\ (Fuc)_1 + (Man)_3(GlcNAc)_2$
5	1620,8	2,0	0,2	$(GlcNAc)_1 + (Man)_4(GlcNAc)_2$
6	1661,8	2,2	0,1	$(GlcNAc)_2 + (Man)_3(GlcNAc)_2$
7	1783,9	1,4	0,1	$(Man)_6(GlcNAc)_2$
8	1794,9	2,0	0,1	$(GlcNAc)_1\ (Fuc)_1 + (Man)_4(GlcNAc)_2$
9	1824,9	4,9	0,2	$(GlcNAc)_1 + (Man)_5(GlcNAc)_2$
10	1835,9	12,1	0,6	$(GlcNAc)_2\ (Fuc)_1 + (Man)_3(GlcNAc)_2$
11	1865,9	1,9	0,1	$(Gal)_1\ (GlcNAc)_2 + (Man)_3(GlcNAc)_2$
12	1907,0	1,1	0,1	$(GlcNAc)_3 + (Man)_3(GlcNAc)_2$
13	1988,0	1,0	0,0	$(Man)_7(GlcNAc)_2$
14	1999,0	2,8	0,2	$(Gal)_1\ (GlcNAc)_1\ (Fuc)_1 + (Man)_4(GlcNAc)_2$
15	2029,0	2,0	0,1	$(Gal)_1\ (GlcNAc)_1 + (Man)_5(GlcNAc)_2$
16	2040,0	7,4	0,3	$(Gal)_1\ (GlcNAc)_2\ (Fuc)_1 + (Man)_3(GlcNAc)_2$
17	2070,0	1,7	0,1	$(Gal)_2\ (GlcNAc)_2 + (Man)_3(GlcNAc)_2$
18	2081,1	8,2	0,2	$(GlcNAc)_3\ (Fuc)_1 + (Man)_3(GlcNAc)_2$
19	2111,1	1,0	0,0	$(Gal)_1\ (GlcNAc)_3 + (Man)_3(GlcNAc)_2$
20	2192,1	0,7	0,0	$(Man)_8(GlcNAc)_2$
21	2203,1	0,9	0,0	$(Gal)_1\ (GlcNAc)_1\ (Fuc)_1 + (Man)_5(GlcNAc)_2$
22	2244,1	6,3	0,4	$(Gal)_2\ (GlcNAc)_2\ (Fuc)_1 + (Man)_3(GlcNAc)_2$
23	2285,2	5,3	0,1	$(Gal)_1\ (GlcNAc)_3\ (Fuc)_1 + (Man)_3(GlcNAc)_2$
24	2315,2	0,7	0,0	$(Gal)_2\ (GlcNAc)_3 + (Man)_3(GlcNAc)_2$
25	2326,2	3,3	0,3	$(GlcNAc)_4\ (Fuc)_1 + (Man)_3(GlcNAc)_2$
26	2356,2	0,4	0,1	$(Gal)_1\ (GlcNAc)_4 + (Man)_3(GlcNAc)_2$
27	2401,2	1,6	0,1	$(NeuAc)_1\ (Gal)_1\ (GlcNAc)_2\ (Fuc)_1 + (Man)_3(GlcNAc)_2$
28	2489,2	3,6	0,3	$(Gal)_2\ GlcNAc)_3\ (Fuc)_1 + (Man)_3(GlcNAc)_2$
29	2519,3	0,6	0,1	$(Gal)_3\ (GlcNAc)_3 + (Man)_3(GlcNAc)_2$
30	2530,3	2,1	0,1	$(Gal)_1\ (GlcNAc)_4\ (Fuc)_1 + (Man)_3(GlcNAc)_2$
31	2605,3	4,4	0,1	$(NeuAc)_1\ (Gal)_2\ (GlcNAc)_2\ (Fuc)_1 + (Man)_3(GlcNAc)_2$
32	2646,3	0,9	0,1	$(NeuAc)_1\ (Gal)_1\ (GlcNAc)_3\ (Fuc)_1 + (Man)_3(GlcNAc)_2$
33	2693,4	1,8	0,1	$(Gal)_3\ (GlcNAc)_3\ (Fuc)_1 + (Man)_3(GlcNAc)_2$
34	2734,4	1,2	0,1	$(Gal)_2\ (GlcNAc)_4\ (Fuc)_1 + (Man)_3(GlcNAc)_2$
35	2850,4	0,9	0,1	$(NeuAc)_1\ (Gal)_2\ (GlcNAc)_3\ (Fuc)_1 + (Man)_3(GlcNAc)_2$
36	2891,5	0,5	0,1	$(NeuAc)_1\ (Gal)_1\ (GlcNAc)_4\ (Fuc)_1 + (Man)_3(GlcNAc)_2$
37	2938,5	0,8	0,1	$(Gal)_3\ (GlcNAc)_4\ (Fuc)_1 + (Man)_3(GlcNAc)_2$
38	2966,5	0,7	0,1	$(NeuAc)_2\ (Gal)_2\ (GlcNAc)_2\ (Fuc)_1 + (Man)_3(GlcNAc)_2$
39	3054,5	1,5	0,1	$(NeuAc)_1\ (Gal)_3\ (GlcNAc)_3\ (Fuc)_1 + (Man)_3(GlcNAc)_2$
40	3095,5	0,6	0,1	$(NeuAc)_1\ (Gal)_2\ (GlcNAc)_4\ (Fuc)_1 + (Man)_3(GlcNAc)_2$
41	3142,6	0,5	0,1	$(Gal)_4\ (GlcNAc)_4\ (Fuc)_1 + (Man)_3(GlcNAc)_2$
42	3299,7	0,6	0,1	$(NeuAc)_1\ (Gal)_3\ (GlcNAc)_4\ (Fuc)_1 + (Man)_3(GlcNAc)_2$
43	3415,7	0,6	0,1	$(NeuAc)_2\ (Gal)_3\ (GlcNAc)_3\ (Fuc)_1 + (Man)_3(GlcNAc)_2$
44	3503,8	0,6	0,1	$(NeuAc)_1\ (Gal)_4\ (GlcNAc)_4\ (Fuc)_1 + (Man)_3(GlcNAc)_2$
45	3660,8	0,2	0,1	$(NeuAc)_2\ (Gal)_3\ (GlcNAc)_4\ (Fuc)_1 + (Man)_3(GlcNAc)_2$
46	3776,9	0,0	0,0	$(NeuAc)_3\ (Gal)_3\ (GlcNAc)_3\ (Fuc)_1 + (Man)_3(GlcNAc)_2$
47	3864,9	0,5	0,1	$(NeuAc)_2\ (Gal)_4\ (GlcNAc)_4\ (Fuc)_1 + (Man)_3(GlcNAc)_2$
48	4226,1	0,1	0,1	$(NeuAc)_3\ (Gal)_4\ (GlcNAc)_4\ (Fuc)_1 + (Man)_3(GlcNAc)_2$
49	4587,3	0,0	0,0	$(NeuAc)_4\ (Gal)_4\ (GlcNAc)_4\ (Fuc)_1 + (Man)_3(GlcNAc)_2$
		100,0		

Table 3. MALDI-TOF MS spectra of N-linked glycans from EBOV-Mayinga GP1,2. Relative intensity (normalized to 100%) of 49 individual N-glycans are shown.

O-linked glycans were removed from GP1,2 beta-elimination for and analyzed by MALDI-TOF mass spectrometry. Spectra revealed 8 O-glycan structures (Figure 13). These structures are shown in Figure 14 and Table 4.

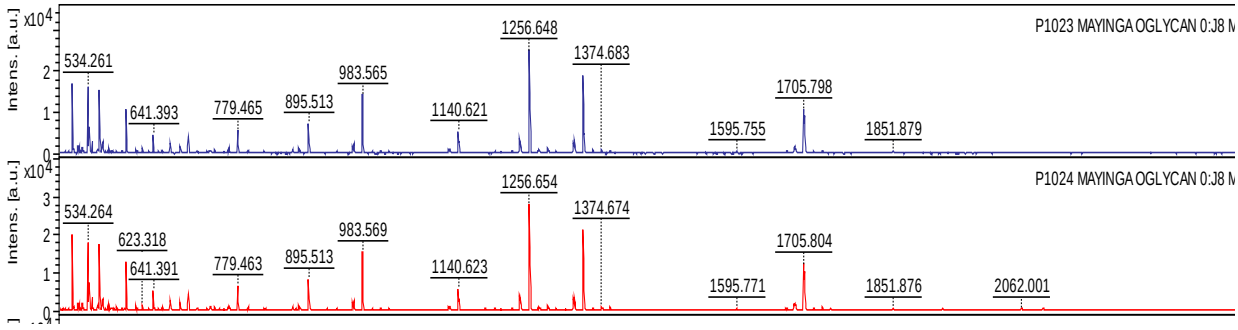


Figure 13. MALDI-TOF MS spectra of O-linked glycans from EBOV-Mayinga GP1,2. Each set of peaks is taken from a different section of the sample, showing consistancy of the assay.

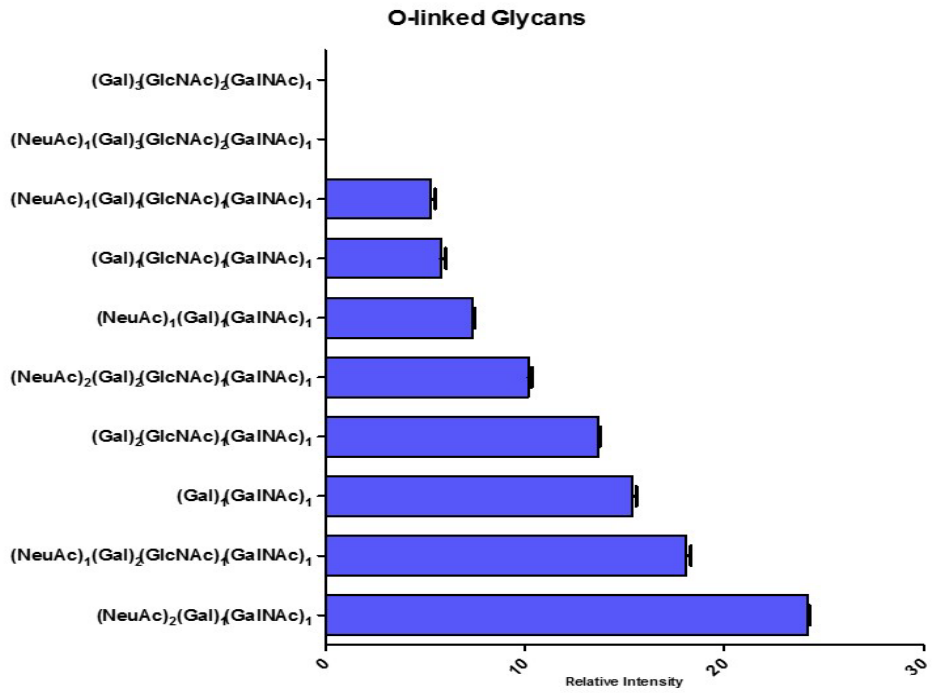


Figure 14. O-linked glycans from EBOV-Mayinga GP1,2. Relative intensity (normalized to 100%) of 10 individual O-glycans are shown. Error bars represent standard error of 2 separate spots of the pooled GP1,2 from two separate experiments.

O-glycan	Relative intenstiy	SD
$(NeuAc)_2(Gal)_1(GalNAc)_1$	24.2	0.07
$(NeuAc)_1(Gal)_2(GlcNAc)_1(GalNAc)_1$	18.1	0.23
$(Gal)_1(GalNAc)_1$	15.4	0.2
$(Gal)_2(GlcNAc)_1(GalNAc)_1$	13.7	0.08
$(NeuAc)_2(Gal)_2(GlcNAc)_1(GalNAc)_1$	10.2	0.16
$(NeuAc)_1(Gal)_1(GalNAc)_1$	7.4	0.1
$(Gal)_1(GlcNAc)_1(GalNAc)_1$	5.8	0.26
$(NeuAc)_1(Gal)_1(GlcNAc)_1(GalNAc)_1$	5.3	0.2
$(NeuAc)_1(Gal)_3(GlcNAc)_2(GalNAc)_1$	0	0
$(Gal)_3(GlcNAc)_2(GalNAc)_1$	0	0

Table 4. MALDI-TOF MS spectra of O-linked glycans from EBOV-Mayinga GP1,2. Relative intensity (normalized to 100%) of 10 individual N-glycans are shown.

Glycan analysis of EBOV-Makona GP1,2 produced in mammalian 293T cells. Although EBOV-Makona (responsible for the 2014-2015 West Africa Ebola outbreak) is the same viral species as EBOV-Mayinga (the founding 1976 strain of EBOV), we chose to analyze its GP1,2 glycan pattern for comparison, to detect any differences between these two EBOV strains. Therefore, EBOV-Makona GP1,2 (NCBI Accession #KJ660347) was modified as in Figure 2 and expressed in mammalian 293T cells. Purified GP1,2 was confirmed for purity and specificity by coomassie staining (Figure 15A), which stains all proteins; periodic schiff stain (PAS), which stains only glycosylated proteins (Figure 15B); and Western blotting with an anti-EBOV antibody (Figure 15D). Two individual experiments were performed, and the purified proteins were combined and analyzed by MALDI-TOF mass spectrometry (Proteodynamics). N-linked glycans on GP1,2 were released by incubation with PNGase F, which was confirmed by reduced protein mass (Figure 15C). Similar to EBOV-Mayinga, mass spectrometry analysis of the released glycans revealed a large number of different N-linked glycan structures with a wide range of frequency in EBOV-Makona GP1,2 (Figure 16).

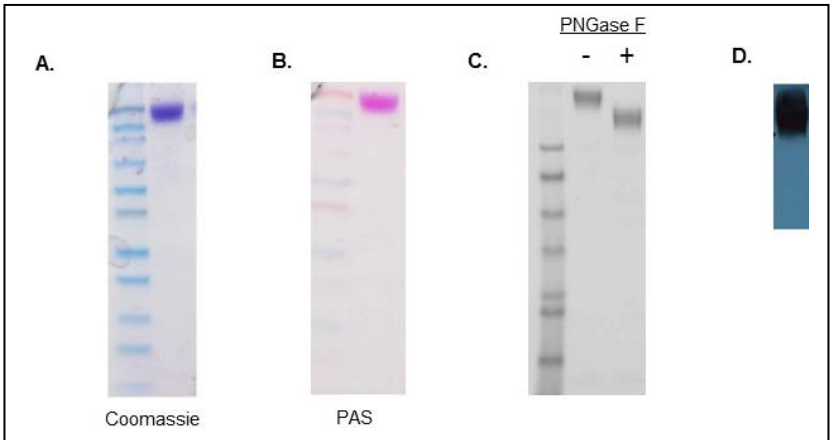


Figure 15. Purification of EBOV-Makona GP1,2 produced in mammalian 293T cells. A. Coomassie stain. B. PAS glycoprotein-specific stain. C. PNGase analysis showing presence of glycans in GP1,2. D. Western blot.

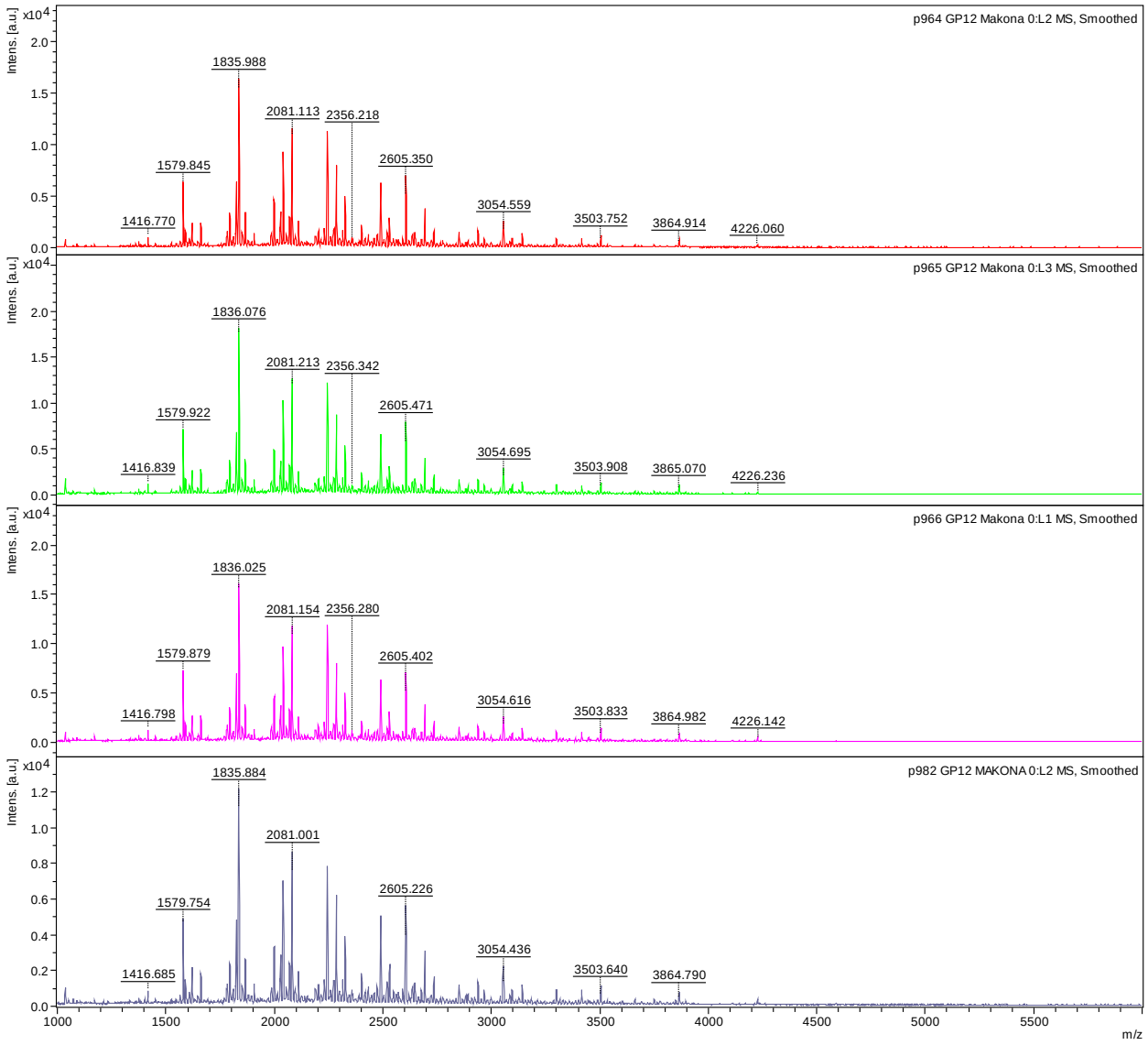


Figure 16. MALDI-TOF MS spectra of N-linked glycans from EBOV-Makona GP1,2. Each set of peaks is taken from a different section of the sample, showing consistancy of the assay.

Analysis of these data revealed 47 different N-linked glycan structures, including high mannose type structures from $(Man)_4(GlcNAc)_2$ to $(Man)_8(GlcNAc)_2$, hybrid glycans having only one GlcNAc residue in the antenna, and bi-, tri- and tetra-antennary complex glycans (Figure 17 shows graphical representation; Table 6 shows the data as in a list format).

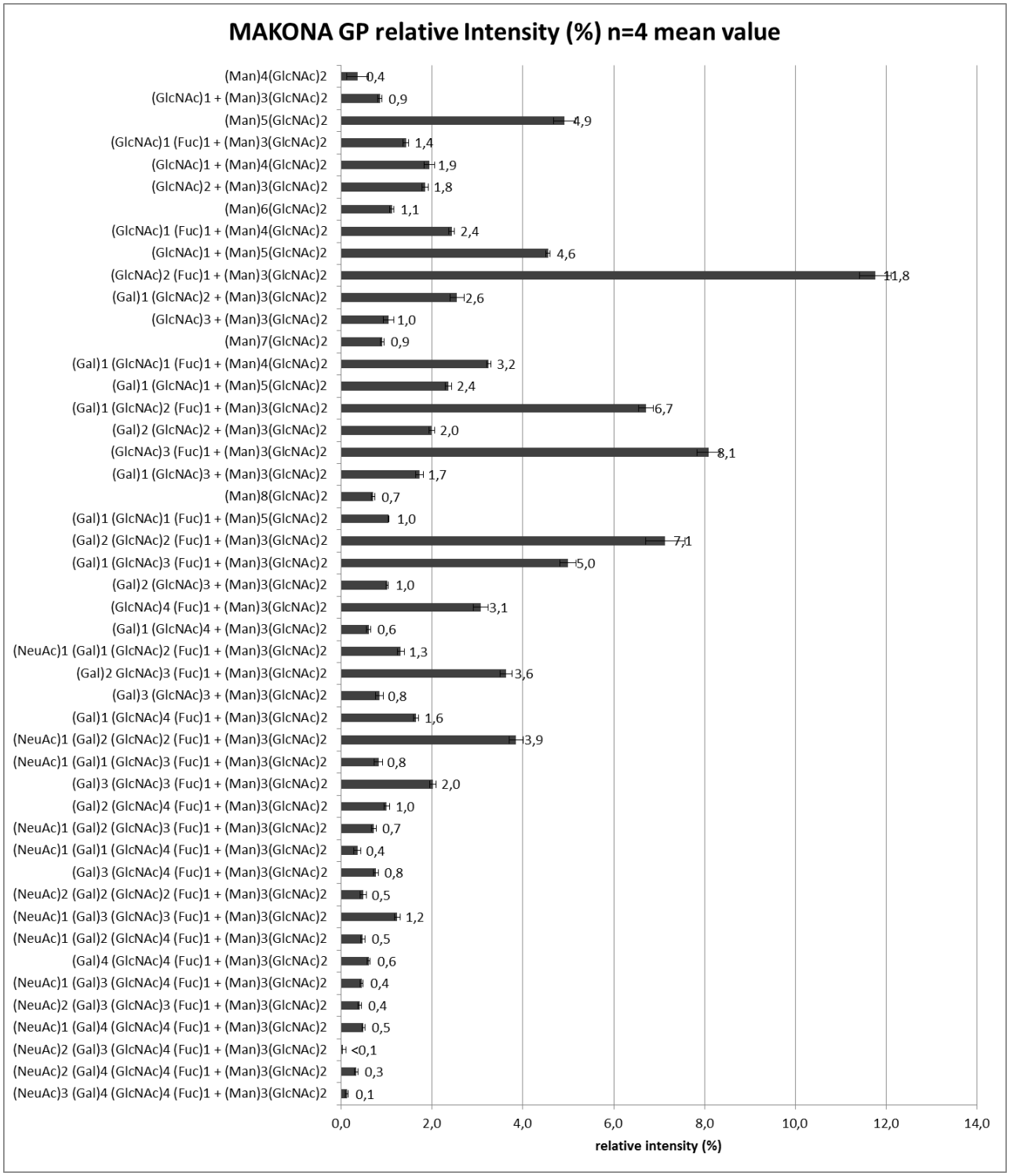


Figure 17. MALDI-TOF MS spectra of N-linked glycans from EBOV-Makona GP1,2. Relative intensity (normalized to 100%) of 47 individual N-glycans are shown. Error bars represent standard error of 4 separate spots of the pooled GP1,2 from two separate experiments.

		MAKONA GP		
Structure N°	$[M+Na]^+$ theoretical monoisotopic mass	mean relative. Intensity %	standard deviation	Interpretation
1	1375,7	0,4	0,2	$(Man)_4(GlcNAc)_2$
2	1416,7	0,9	0,0	$(GlcNAc)_1 + (Man)_3(GlcNAc)_2$
3	1579,8	4,9	0,2	$(Man)_5(GlcNAc)_2$
4	1590,8	1,4	0,1	$(GlcNAc)_1\ (Fuc)_1 + (Man)_3(GlcNAc)_2$
5	1620,8	1,9	0,1	$(GlcNAc)_1 + (Man)_4(GlcNAc)_2$
6	1661,8	1,8	0,1	$(GlcNAc)_2 + (Man)_3(GlcNAc)_2$
7	1783,9	1,1	0,0	$(Man)_6(GlcNAc)_2$
8	1794,9	2,4	0,1	$(GlcNAc)_1\ (Fuc)_1 + (Man)_4(GlcNAc)_2$
9	1824,9	4,6	0,0	$(GlcNAc)_1 + (Man)_5(GlcNAc)_2$
10	1835,9	11,8	0,3	$(GlcNAc)_2\ (Fuc)_1 + (Man)_3(GlcNAc)_2$
11	1865,9	2,6	0,2	$(Gal)_1\ (GlcNAc)_2 + (Man)_3(GlcNAc)_2$
12	1907,0	1,0	0,1	$(GlcNAc)_3 + (Man)_3(GlcNAc)_2$
13	1988,0	0,9	0,0	$(Man)_7(GlcNAc)_2$
14	1999,0	3,2	0,0	$(Gal)_1\ (GlcNAc)_1\ (Fuc)_1 + (Man)_4(GlcNAc)_2$
15	2029,0	2,4	0,1	$(Gal)_1\ (GlcNAc)_1 + (Man)_5(GlcNAc)_2$
16	2040,0	6,7	0,2	$(Gal)_1\ (GlcNAc)_2\ (Fuc)_1 + (Man)_3(GlcNAc)_2$
17	2070,0	2,0	0,1	$(Gal)_2\ (GlcNAc)_2 + (Man)_3(GlcNAc)_2$
18	2081,1	8,1	0,3	$(GlcNAc)_3\ (Fuc)_1 + (Man)_3(GlcNAc)_2$
19	2111,1	1,7	0,1	$(Gal)_1\ (GlcNAc)_3 + (Man)_3(GlcNAc)_2$
20	2192,1	0,7	0,0	$(Man)_8(GlcNAc)_2$
21	2203,1	1,0	0,0	$(Gal)_1\ (GlcNAc)_1\ (Fuc)_1 + (Man)_5(GlcNAc)_2$
22	2244,1	7,1	0,4	$(Gal)_2\ (GlcNAc)_2\ (Fuc)_1 + (Man)_3(GlcNAc)_2$
23	2285,2	5,0	0,2	$(Gal)_1\ (GlcNAc)_3\ (Fuc)_1 + (Man)_3(GlcNAc)_2$
24	2315,2	1,0	0,0	$(Gal)_2\ (GlcNAc)_3 + (Man)_3(GlcNAc)_2$
25	2326,2	3,1	0,2	$(GlcNAc)_4\ (Fuc)_1 + (Man)_3(GlcNAc)_2$
26	2356,2	0,6	0,0	$(Gal)_1\ (GlcNAc)_4 + (Man)_3(GlcNAc)_2$
27	2401,2	1,3	0,1	$(NeuAc)_1\ (Gal)_1\ (GlcNAc)_2\ (Fuc)_1 + (Man)_3(GlcNAc)_2$
28	2489,2	3,6	0,1	$(Gal)_2\ GlcNAc)_3\ (Fuc)_1 + (Man)_3(GlcNAc)_2$
29	2519,3	0,8	0,1	$(Gal)_3\ (GlcNAc)_3 + (Man)_3(GlcNAc)_2$
30	2530,3	1,6	0,1	$(Gal)_1\ (GlcNAc)_4\ (Fuc)_1 + (Man)_3(GlcNAc)_2$
31	2605,3	3,9	0,1	$(NeuAc)_1\ (Gal)_2\ (GlcNAc)_2\ (Fuc)_1 + (Man)_3(GlcNAc)_2$
32	2646,3	0,8	0,1	$(NeuAc)_1\ (Gal)_1\ (GlcNAc)_3\ (Fuc)_1 + (Man)_3(GlcNAc)_2$
33	2693,4	2,0	0,1	$(Gal)_3\ (GlcNAc)_3\ (Fuc)_1 + (Man)_3(GlcNAc)_2$
34	2734,4	1,0	0,1	$(Gal)_2\ (GlcNAc)_4\ (Fuc)_1 + (Man)_3(GlcNAc)_2$
35	2850,4	0,7	0,1	$(NeuAc)_1\ (Gal)_2\ (GlcNAc)_3\ (Fuc)_1 + (Man)_3(GlcNAc)_2$
36	2891,5	0,4	0,1	$(NeuAc)_1\ (Gal)_1\ (GlcNAc)_4\ (Fuc)_1 + (Man)_3(GlcNAc)_2$
37	2938,5	0,8	0,1	$(Gal)_3\ (GlcNAc)_4\ (Fuc)_1 + (Man)_3(GlcNAc)_2$
38	2966,5	0,5	0,1	$(NeuAc)_2\ (Gal)_2\ (GlcNAc)_2\ (Fuc)_1 + (Man)_3(GlcNAc)_2$
39	3054,5	1,2	0,1	$(NeuAc)_1\ (Gal)_3\ (GlcNAc)_3\ (Fuc)_1 + (Man)_3(GlcNAc)_2$
40	3095,5	0,5	0,0	$(NeuAc)_1\ (Gal)_2\ (GlcNAc)_4\ (Fuc)_1 + (Man)_3(GlcNAc)_2$
41	3142,6	0,6	0,0	$(Gal)_4\ (GlcNAc)_4\ (Fuc)_1 + (Man)_3(GlcNAc)_2$
42	3299,7	0,4	0,0	$(NeuAc)_1\ (Gal)_3\ (GlcNAc)_4\ (Fuc)_1 + (Man)_3(GlcNAc)_2$
43	3415,7	0,4	0,0	$(NeuAc)_2\ (Gal)_3\ (GlcNAc)_3\ (Fuc)_1 + (Man)_3(GlcNAc)_2$
44	3503,8	0,5	0,0	$(NeuAc)_1\ (Gal)_4\ (GlcNAc)_4\ (Fuc)_1 + (Man)_3(GlcNAc)_2$
45	3660,8	0,0	0,1	$(NeuAc)_2\ (Gal)_3\ (GlcNAc)_4\ (Fuc)_1 + (Man)_3(GlcNAc)_2$
46	3864,9	0,3	0,0	$(NeuAc)_2\ (Gal)_4\ (GlcNAc)_4\ (Fuc)_1 + (Man)_3(GlcNAc)_2$
47	4226,1	0,1	0,0	$(NeuAc)_3\ (Gal)_4\ (GlcNAc)_4\ (Fuc)_1 + (Man)_3(GlcNAc)_2$
		100		

Table 6. MALDI-TOF MS spectra of N-linked glycans from EBOV-Makona GP1,2. Relative intensity (normalized to 100%) of 49 individual N-glycans are shown.

Similar to EBOV Mayinga GP1,2 dominant individual *N*-glycan structures are bi-antennary glycans: $(GlcNAc)_2(Fuc)_1 + (Man)_3(GlcNAc)_2$ (m/z 1835) at 12%, $(GlcNAc)_3(Fuc)_1 + (Man)_3(GlcNAc)_2$ (m/z 2081) at 8%, $(Gal)_1(GlcNAc)_2(Fuc)_1 + (Man)_3(GlcNAc)_2$ (m/z 2040) at 7% and $(Gal)_2(GlcNAc)_2(Fuc)_1 + (Man)_3(GlcNAc)_2$ (m/z 2244) at 7%.

O-linked glycans were removed from GP1,2 beta-elimination for and analyzed by MALDI-TOF mass spectrometry. Spectra revealed 8 O-glycan structures (Figure 18). These structures are shown in Figure 19 and Table 7.

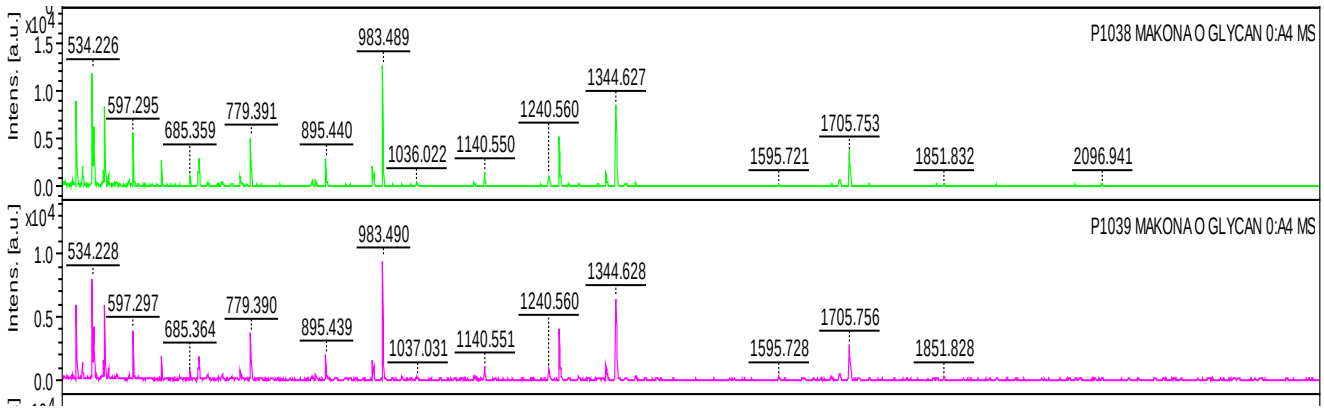
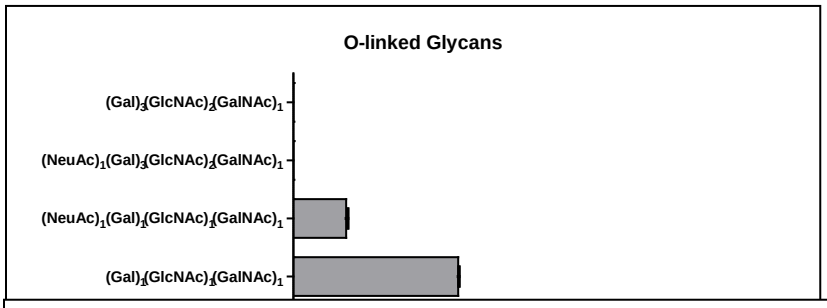


Figure 18. MALDI-TOF MS spectra of O-linked glycans from EBOV-Makona GP1,2. Each set of peaks is taken from a different section of the sample, showing consistancy of the assay.



O-glycan	Relative intenstiy	SD
$(NeuAc)_2(Gal)_1(GalNAc)_1$	11.5	0.48
$(NeuAc)_1(Gal)_2(GlcNAc)_1(GalNAc)_1$	17.5	0.87
$(Gal)_1(GalNAc)_1$	22.9	1.33
$(Gal)_2(GlcNAc)_1(GalNAc)_1$	25.4	0.82
$(NeuAc)_2(Gal)_2(GlcNAc)_1(GalNAc)_1$	6.8	0.72
$(NeuAc)_1(Gal)_1(GalNAc)_1$	2.3	0.05
$(Gal)_1(GlcNAc)_1(GalNAc)_1$	10.3	0.06
$(NeuAc)_1(Gal)_1(GlcNAc)_1(GalNAc)_1$	3.3	0.11
$(NeuAc)_1(Gal)_3(GlcNAc)_2(GalNAc)_1$	0	0
$(Gal)_3(GlcNAc)_2(GalNAc)_1$	0	0

Table 7. MALDI-TOF MS spectra of O-linked glycans from EBOV-Makona GP1,2.
Relative intensity (normalized to 100%) of 10 individual N-glycans are shown.

Glycan analysis of SUDV GP1,2 produced in mammalian 293T cells. SUDV GP1,2 (GeneID: 3160774)) was modified as in Figure 2 and expressed in mammalian 293T cells. Purified GP1,2 was confirmed for purity and specificity by coomassie staining (Figure 20A), which stains all proteins; periodic schiff stain (PAS), which stains only glycosylated proteins (Figure 20B); and Western blotting with an anti-SUDV antibody (Figure 20D). Two individual experiments were performed, and the purified proteins were combined and analyzed by MALDI-TOF mass spectrometry (Proteodynamics). N-linked glycans on GP1,2 were released by incubation with PNGase F, which was confirmed by reduced protein mass (Figure 20C). Similar to EBOV GP1,2, mass spectrometry analysis of the released glycans revealed a large number of different N-linked glycan structures with a wide range of frequency in SUDV GP1,2 (Figure 21).

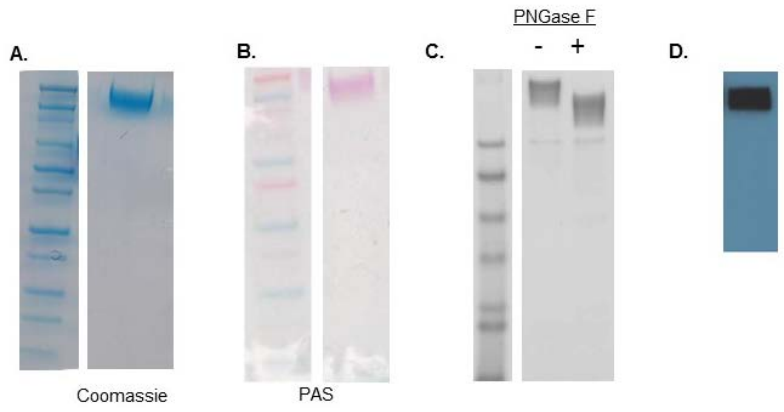


Figure 20. Purification of SUDV GP1,2 produced in mammalian 293T cells. A. Coomassie stain. B. PAS glycoprotein-specific stain. C. PNGase analysis showing presence of glycans in GP1,2. D. Western blot.

Analysis of these data revealed 56 different N-linked glycan structures, a larger number than EBOV-Mayinga (49) or EBOV-Makona (47), including high mannose type structures from $(Man)_4(GlcNAc)_2$ to $(Man)_8(GlcNAc)_2$, hybrid glycans having only one GlcNAc residue in the antenna, and bi-, tri- and tetra-antennary complex glycans (Figure 22 shows graphical representation; Table 8 shows the data in a list format). Similar to EBOV Mayinga GP1,2 dominant individual *N*-glycan structures are bi-antennary glycans: $(GlcNAc)_2(Fuc)_1$ + $(Man)_3(GlcNAc)_2$ (m/z 1835) at 7.5%, $(GlcNAc)_3(Fuc)_1$ + $(Man)_3(GlcNAc)_2$ (m/z 2081) at 11%.

However, different from EBOV, the other two dominant N-glyans were $(GlcNAc)_4(Fuc)_1$ + $(Man)_3(GlcNAc)_2$ (m/z 2326) at 11% and $(Man)_5(GlcNAc)_2$ (m/z 1580) at 6%.

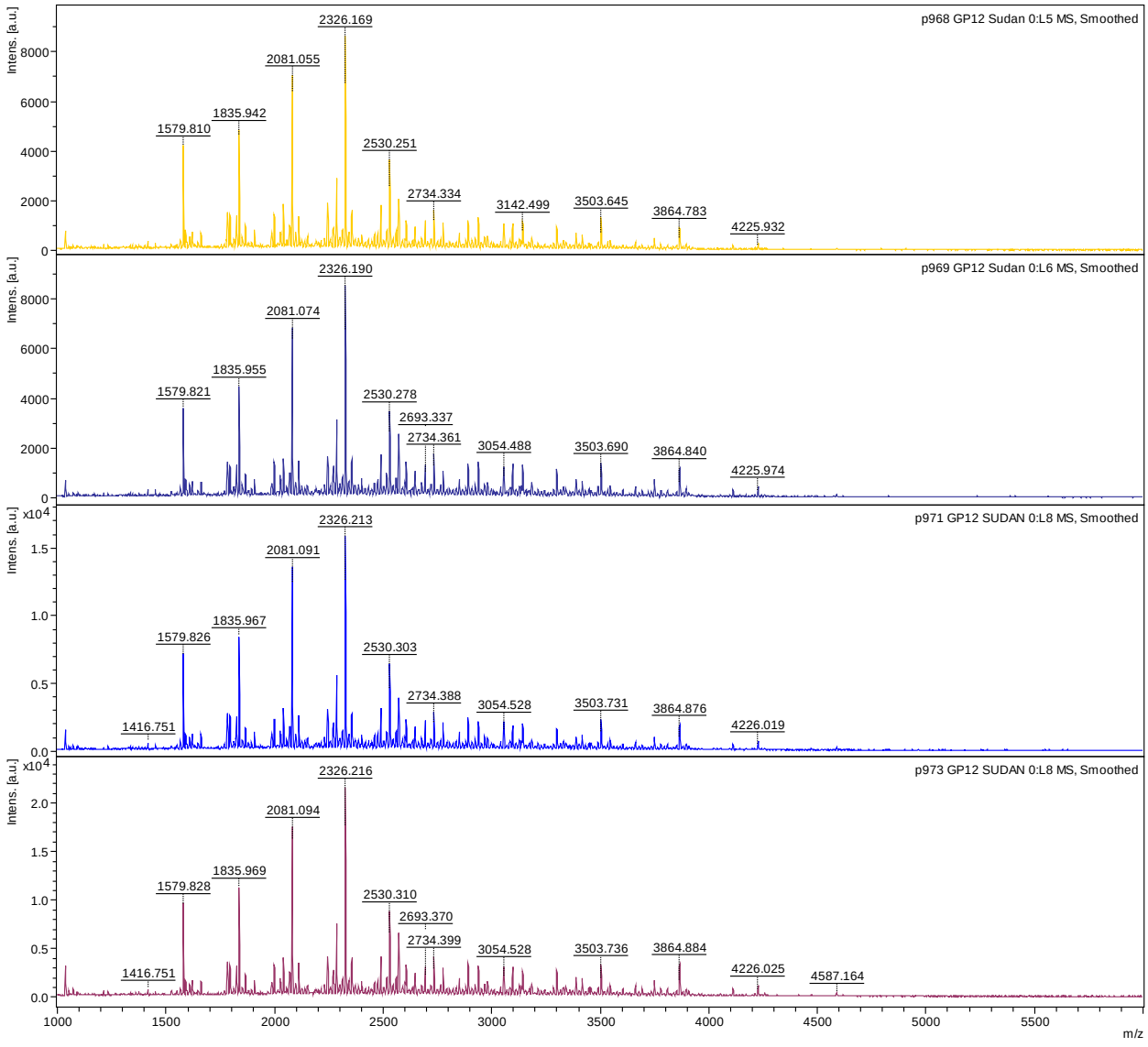


Figure 21. MALDI-TOF MS spectra of N-linked glycans from SUDV GP1,2. Each set of peaks is taken from a different section of the sample, showing consistancy of the assay.

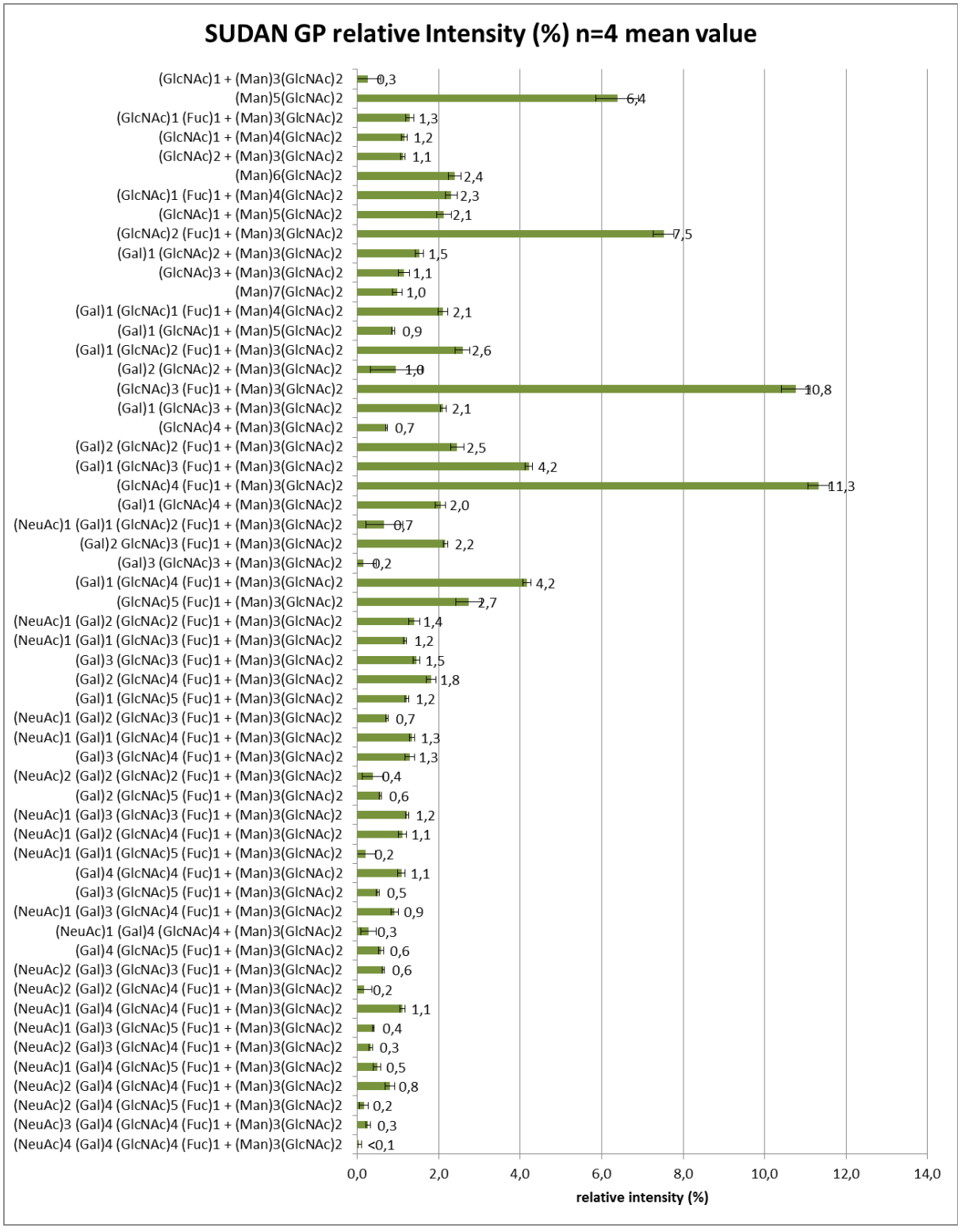


Figure 22. MALDI-TOF MS spectra of N-linked glycans from SUDV GP1,2. Relative intensity (normalized to 100%) of 56 individual N-glycans are shown. Error bars represent standard error of 4 separate spots of the pooled GP1,2 from two separate experiments.

		SUDAN GP		
Structure N°	$[M+Na]^+$ theoretical monoisotopic mass	mean relative. Intensity %	standard deviation	Interpretation
1	1416,7	0,3	0,3	$(GlcNAc)_1 + (Man)_3(GlcNAc)_2$
2	1579,8	6,4	0,5	$(Man)_5(GlcNAc)_2$
3	1590,8	1,3	0,1	$(GlcNAc)_1\ (Fuc)_1 + (Man)_3(GlcNAc)_2$
4	1620,8	1,2	0,1	$(GlcNAc)_1 + (Man)_4(GlcNAc)_2$
5	1661,8	1,1	0,1	$(GlcNAc)_2 + (Man)_3(GlcNAc)_2$
6	1783,9	2,4	0,1	$(Man)_6(GlcNAc)_2$
7	1794,9	2,3	0,2	$(GlcNAc)_1\ (Fuc)_1 + (Man)_4(GlcNAc)_2$
8	1824,9	2,1	0,2	$(GlcNAc)_1 + (Man)_5(GlcNAc)_2$
9	1835,9	7,5	0,3	$(GlcNAc)_2\ (Fuc)_1 + (Man)_3(GlcNAc)_2$
10	1865,9	1,5	0,1	$(Gal)_1\ (GlcNAc)_2 + (Man)_3(GlcNAc)_2$
11	1907,0	1,1	0,1	$(GlcNAc)_3 + (Man)_3(GlcNAc)_2$
12	1988,0	1,0	0,1	$(Man)_7(GlcNAc)_2$
13	1999,0	2,1	0,1	$(Gal)_1\ (GlcNAc)_1\ (Fuc)_1 + (Man)_4(GlcNAc)_2$
14	2029,0	0,9	0,0	$(Gal)_1\ (GlcNAc)_1 + (Man)_5(GlcNAc)_2$
15	2040,0	2,6	0,2	$(Gal)_1\ (GlcNAc)_2\ (Fuc)_1 + (Man)_3(GlcNAc)_2$
16	2070,0	1,0	0,6	$(Gal)_2\ (GlcNAc)_2 + (Man)_3(GlcNAc)_2$
17	2081,1	10,8	0,3	$(GlcNAc)_3\ (Fuc)_1 + (Man)_3(GlcNAc)_2$
18	2111,1	2,1	0,1	$(Gal)_1\ (GlcNAc)_3 + (Man)_3(GlcNAc)_2$
19	2152,1	0,7	0,0	$(GlcNAc)_4 + (Man)_3(GlcNAc)_2$
20	2244,1	2,5	0,2	$(Gal)_2\ (GlcNAc)_2\ (Fuc)_1 + (Man)_3(GlcNAc)_2$
21	2285,2	4,2	0,1	$(Gal)_1\ (GlcNAc)_3\ (Fuc)_1 + (Man)_3(GlcNAc)_2$
22	2326,2	11,3	0,3	$(GlcNAc)_4\ (Fuc)_1 + (Man)_3(GlcNAc)_2$
23	2356,2	2,0	0,1	$(Gal)_1\ (GlcNAc)_4 + (Man)_3(GlcNAc)_2$
24	2401,2	0,7	0,4	$(NeuAc)_1\ (Gal)_1\ (GlcNAc)_2\ (Fuc)_1 + (Man)_3(GlcNAc)_2$
25	2489,2	2,2	0,1	$(Gal)_2\ GlcNAc)_3\ (Fuc)_1 + (Man)_3(GlcNAc)_2$
26	2519,3	0,2	0,3	$(Gal)_3\ (GlcNAc)_3 + (Man)_3(GlcNAc)_2$
27	2530,3	4,2	0,1	$(Gal)_1\ (GlcNAc)_4\ (Fuc)_1 + (Man)_3(GlcNAc)_2$
28	2571,3	2,7	0,3	$(GlcNAc)_5\ (Fuc)_1 + (Man)_3(GlcNAc)_2$
29	2605,3	1,4	0,1	$(NeuAc)_1\ (Gal)_2\ (GlcNAc)_2\ (Fuc)_1 + (Man)_3(GlcNAc)_2$
30	2646,3	1,2	0,0	$(NeuAc)_1\ (Gal)_1\ (GlcNAc)_3\ (Fuc)_1 + (Man)_3(GlcNAc)_2$
31	2693,4	1,5	0,1	$(Gal)_3\ (GlcNAc)_3\ (Fuc)_1 + (Man)_3(GlcNAc)_2$
32	2734,4	1,8	0,1	$(Gal)_2\ (GlcNAc)_4\ (Fuc)_1 + (Man)_3(GlcNAc)_2$
33	2755,4	1,2	0,0	$(Gal)_1\ (GlcNAc)_5\ (Fuc)_1 + (Man)_3(GlcNAc)_2$
34	2850,4	0,7	0,0	$(NeuAc)_1\ (Gal)_2\ (GlcNAc)_3\ (Fuc)_1 + (Man)_3(GlcNAc)_2$
35	2891,5	1,3	0,1	$(NeuAc)_1\ (Gal)_1\ (GlcNAc)_4\ (Fuc)_1 + (Man)_3(GlcNAc)_2$
36	2938,5	1,3	0,1	$(Gal)_3\ (GlcNAc)_4\ (Fuc)_1 + (Man)_3(GlcNAc)_2$
37	2966,5	0,4	0,3	$(NeuAc)_2\ (Gal)_2\ (GlcNAc)_2\ (Fuc)_1 + (Man)_3(GlcNAc)_2$
38	2979,5	0,6	0,0	$(Gal)_2\ (GlcNAc)_5\ (Fuc)_1 + (Man)_3(GlcNAc)_2$
39	3054,5	1,2	0,0	$(NeuAc)_1\ (Gal)_3\ (GlcNAc)_3\ (Fuc)_1 + (Man)_3(GlcNAc)_2$
40	3095,5	1,1	0,1	$(NeuAc)_1\ (Gal)_2\ (GlcNAc)_4\ (Fuc)_1 + (Man)_3(GlcNAc)_2$
41	3136,6	0,2	0,2	$(NeuAc)_1\ (Gal)_1\ (GlcNAc)_5\ (Fuc)_1 + (Man)_3(GlcNAc)_2$
42	3142,6	1,1	0,1	$(Gal)_4\ (GlcNAc)_4\ (Fuc)_1 + (Man)_3(GlcNAc)_2$
43	3183,6	0,5	0,0	$(Gal)_3\ (GlcNAc)_5\ (Fuc)_1 + (Man)_3(GlcNAc)_2$
44	3299,7	0,9	0,1	$(NeuAc)_1\ (Gal)_3\ (GlcNAc)_4\ (Fuc)_1 + (Man)_3(GlcNAc)_2$
45	3329,7	0,3	0,2	$(NeuAc)_1\ (Gal)_4\ (GlcNAc)_4 + (Man)_3(GlcNAc)_2$
46	3387,7	0,6	0,1	$(Gal)_4\ (GlcNAc)_5\ (Fuc)_1 + (Man)_3(GlcNAc)_2$
47	3415,7	0,6	0,0	$(NeuAc)_2\ (Gal)_3\ (GlcNAc)_3\ (Fuc)_1 + (Man)_3(GlcNAc)_2$
48	3456,7	0,2	0,2	$(NeuAc)_2\ (Gal)_2\ (GlcNAc)_4\ (Fuc)_1 + (Man)_3(GlcNAc)_2$
49	3503,8	1,1	0,1	$(NeuAc)_1\ (Gal)_4\ (GlcNAc)_4\ (Fuc)_1 + (Man)_3(GlcNAc)_2$
50	3544,8	0,4	0,0	$(NeuAc)_1\ (Gal)_3\ (GlcNAc)_5\ (Fuc)_1 + (Man)_3(GlcNAc)_2$
51	3660,8	0,3	0,0	$(NeuAc)_2\ (Gal)_3\ (GlcNAc)_4\ (Fuc)_1 + (Man)_3(GlcNAc)_2$
52	3748,9	0,5	0,1	$(NeuAc)_1\ (Gal)_4\ (GlcNAc)_5\ (Fuc)_1 + (Man)_3(GlcNAc)_2$
53	3864,9	0,8	0,1	$(NeuAc)_2\ (Gal)_4\ (GlcNAc)_4\ (Fuc)_1 + (Man)_3(GlcNAc)_2$
54	4110,1	0,2	0,1	$(NeuAc)_2\ (Gal)_4\ (GlcNAc)_5\ (Fuc)_1 + (Man)_3(GlcNAc)_2$
55	4226,1	0,3	0,1	$(NeuAc)_3\ (Gal)_4\ (GlcNAc)_4\ (Fuc)_1 + (Man)_3(GlcNAc)_2$
56	4587,3	0,0	0,1	$(NeuAc)_4\ (Gal)_4\ (GlcNAc)_4\ (Fuc)_1 + (Man)_3(GlcNAc)_2$
		100,0		

Table 8. MALDI-TOF MS spectra of N-linked glycans from SUDV GP1,2. Relative intensity (normalized to 100%) of 56 individual N-glycans are shown.

O-linked glycans were removed from GP1,2 beta-elimination for and analyzed by MALDI-TOF mass spectrometry. Spectra revealed 10 O-glycan structures (Figure 23). These structures are shown in Figure 24 and Table 9.

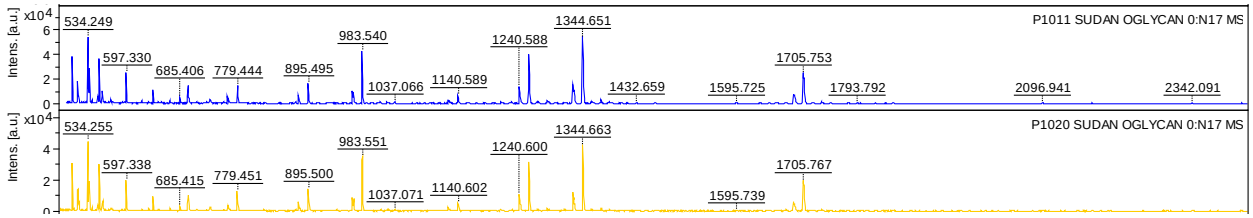


Figure 23. MALDI-TOF MS spectra of O-linked glycans from SUDV GP1,2. Each set of peaks is taken from a different section of the sample, showing consistancy of the assay.

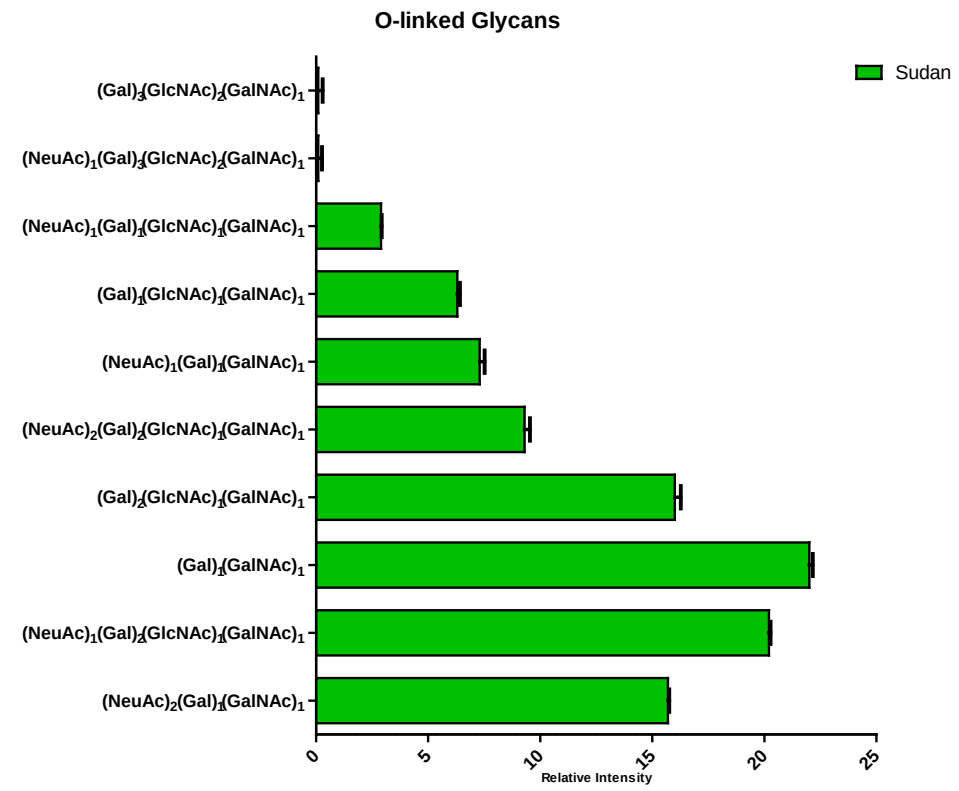


Figure 24. O-linked glycans from SUDV GP1,2. Relative intensity (normalized to 100%) of 10 individual N-glycans are shown. Error bars represent standard error of 2 separate spots of the pooled GP1,2 from two separate experiments.

O-glycan	Relative intenstiy	SD
$(NeuAc)_2(Gal)_1(GalNAc)_1$	15.7	0.06
$(NeuAc)_1(Gal)_2(GlcNAc)_1(GalNAc)_1$	20.2	0.07
$(Gal)_1(GalNAc)_1$	22.0	0.15
$(Gal)_2(GlcNAc)_1(GalNAc)_1$	16.0	0.27
$(NeuAc)_2(Gal)_2(GlcNAc)_1(GalNAc)_1$	9.3	0.24
$(NeuAc)_1(Gal)_1(GalNAc)_1$	7.3	0.22
$(Gal)_1(GlcNAc)_1(GalNAc)_1$	6.3	0.12
$(NeuAc)_1(Gal)_1(GlcNAc)_1(GalNAc)_1$	2.9	0.03
$(NeuAc)_1(Gal)_3(GlcNAc)_2(GalNAc)_1$	0.1	0.17
$(Gal)_3(GlcNAc)_2(GalNAc)_1$	0.1	0.19

Table 9. MALDI-TOF MS spectra of O-linked glycans from SUDV GP1,2. Relative intensity (normalized to 100%) of 10 individual N-glycans are shown.

Glycan analysis of BDBV GP1,2 produced in mammalian 293T cells. BDBV GP1,2 (GeneID: 9487265)) was modified as in Figure 2 and expressed in mammalian 293T cells. Purified GP1,2 was confirmed for purity and specificity by coomassie staining (Figure 25A), which stains all proteins; periodic schiff stain (PAS), which stains only glycosylated proteins (Figure 25B); and Western blotting with an anti-BDBV antibody (Figure 25D). Four individual experiments were performed, and the purified proteins were combined and analyzed by MALDI-TOF mass spectrometry (Proteodynamics). N-linked glycans on GP1,2 were released by incubation with PNGase F, which was confirmed by reduced protein mass (Figure 25C). Similar to EBOV GP1,2, mass spectrometry analysis of the released glycans revealed a large number of different N-linked glycan structures with a wide range of frequency in BDBV GP1,2 (Figure 26).

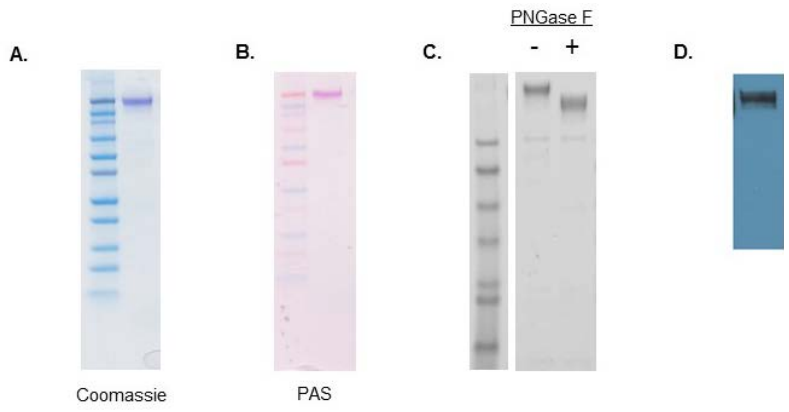


Figure 25. Purification of BDBV GP1,2 produced in mammalian 293T cells. A. Coomassie stain. B. PAS glycoprotein-specific stain. C. PNGase analysis showing presence of glycans in GP1,2. D. Western blot.

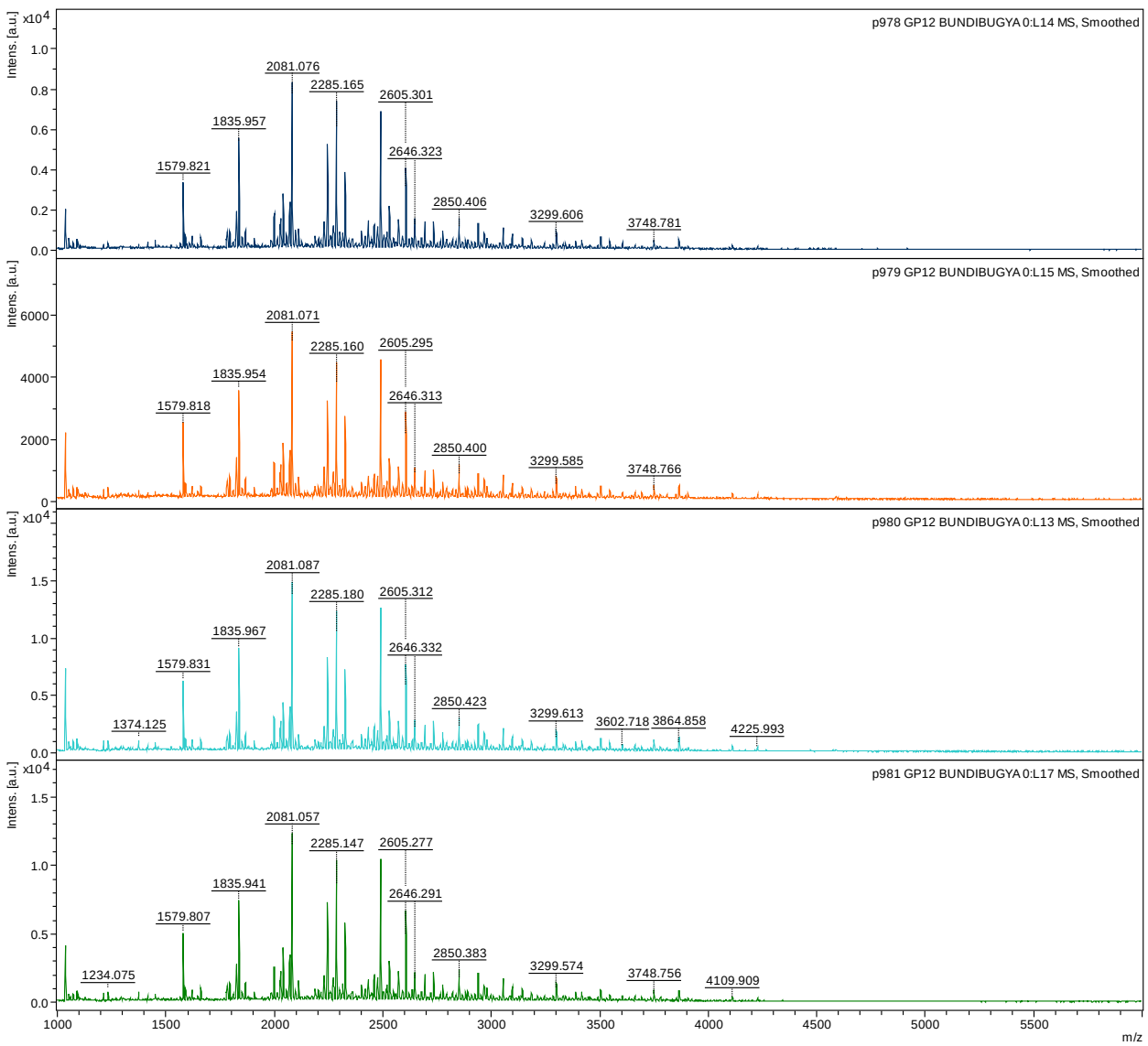


Figure 26. MALDI-TOF MS spectra of N-linked glycans from BDBV GP1,2. Each set of peaks is taken from a different section of the sample, showing consistancy of the assay.

Analysis of these data revealed 52 different N-linked glycan structures, a different number than EBOV-Mayinga (49), EBOV-Makona (47) or SUDV (56), including high mannose type structures from $(Man)_4(GlcNAc)_2$ to $(Man)_8(GlcNAc)_2$, hybrid glycans having only one GlcNAc residue in the antenna, and bi-, tri- and tetra-antennary complex glycans (Figure 27 shows graphical representation; Table 10 shows the data in a list format). Similar to EBOV and SUDV

GP1,2, dominant

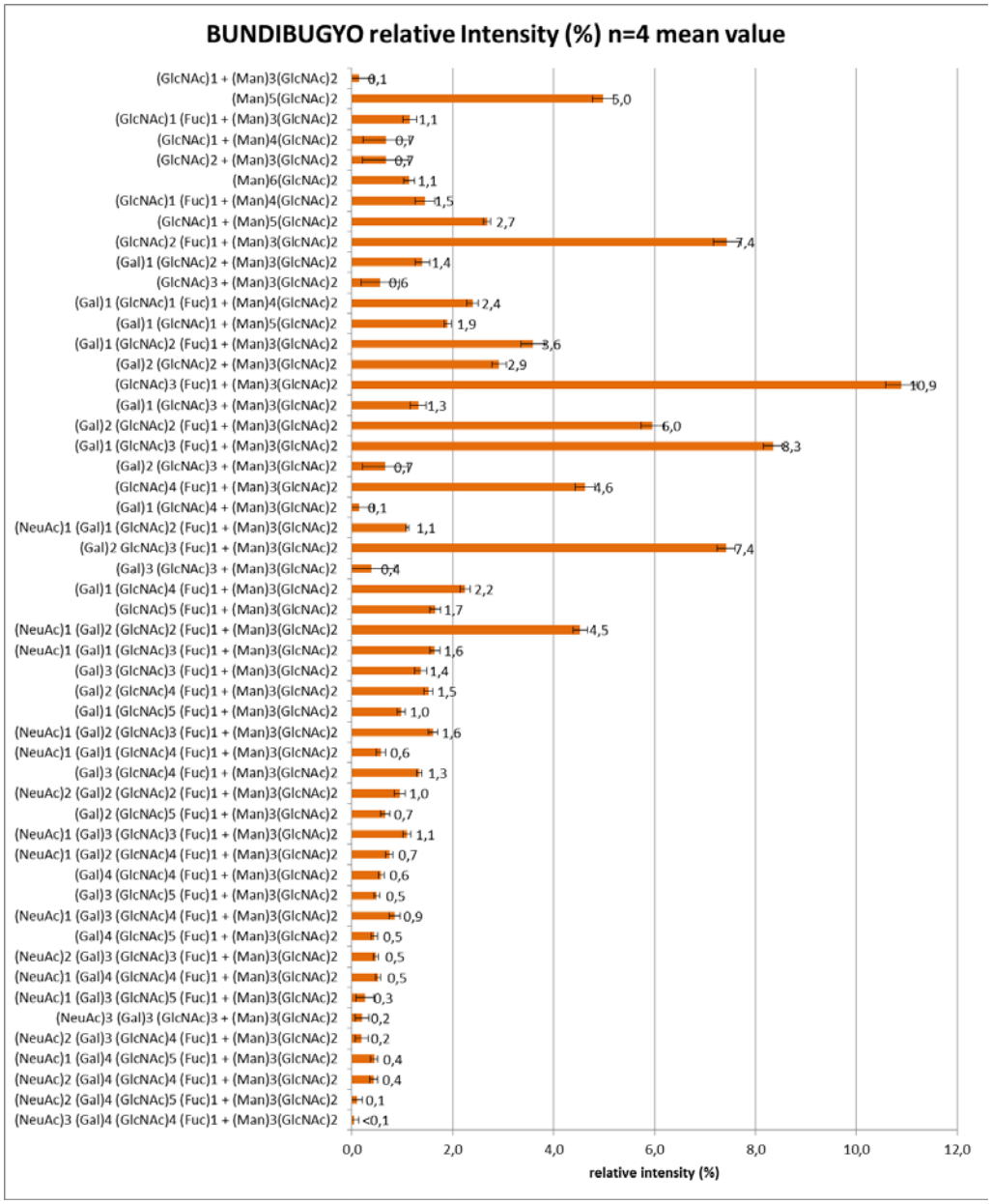


Figure 27. MALDI-TOF MS spectra of N-linked glycans from BDBV GP1,2. Relative intensity (normalized to 100%) of 52 individual N-glycans are shown. Error bars represent standard error of 4 separate spots of the pooled GP1,2 from two separate experiments.

individual *N*-glycan structures are bi-antennary glycans: $(GlcNAc)_2(Fuc)_1 + (Man)_3(GlcNAc)_2$ (m/z 1835) at 7%, $(GlcNAc)_3(Fuc)_1 + (Man)_3(GlcNAc)_2$ (m/z 2081) at 11%. However, different from EBOV and SUDV, the other two dominant N-glyans were $(Gal)_1(GlcNAc)_3(Fuc)_1 +$

$(Man)_3(GlcNAc)_2$ (m/z 2285) at 8% and $(Gal)_2(GlcNAc)_3(Fuc)_1$ + $(Man)_3(GlcNAc)_2$ (m/z 2489) at 7%.

		BUNDIBUGYO		
Structure N°	$[M+Na]^+$ theoretical monoisotopic mass	mean relative. Intensity %	standard deviation	Interpretaton
1	1416,7	0,1	0,3	$(GlcNAc)_1$ + $(Man)_3(GlcNAc)_2$
2	1579,8	5,0	0,2	$(Man)_5(GlcNAc)_2$
3	1590,8	1,1	0,1	$(GlcNAc)_1$ $(Fuc)_1$ + $(Man)_3(GlcNAc)_2$
4	1620,8	0,7	0,5	$(GlcNAc)_1$ + $(Man)_4(GlcNAc)_2$
5	1661,8	0,7	0,5	$(GlcNAc)_2$ + $(Man)_3(GlcNAc)_2$
6	1783,9	1,1	0,1	$(Man)_6(GlcNAc)_2$
7	1794,9	1,5	0,2	$(GlcNAc)_1$ $(Fuc)_1$ + $(Man)_4(GlcNAc)_2$
8	1824,9	2,7	0,1	$(GlcNAc)_1$ + $(Man)_5(GlcNAc)_2$
9	1835,9	7,4	0,3	$(GlcNAc)_2$ $(Fuc)_1$ + $(Man)_3(GlcNAc)_2$
10	1865,9	1,4	0,1	$(Gal)_1$ $(GlcNAc)_2$ + $(Man)_3(GlcNAc)_2$
11	1907,0	0,6	0,4	$(GlcNAc)_3$ + $(Man)_3(GlcNAc)_2$
12	1999,0	2,4	0,1	$(Gal)_1$ $(GlcNAc)_1$ $(Fuc)_1$ + $(Man)_4(GlcNAc)_2$
13	2029,0	1,9	0,1	$(Gal)_1$ $(GlcNAc)_1$ + $(Man)_5(GlcNAc)_2$
14	2040,0	3,6	0,2	$(Gal)_1$ $(GlcNAc)_2$ $(Fuc)_1$ + $(Man)_3(GlcNAc)_2$
15	2070,0	2,9	0,1	$(Gal)_2$ $(GlcNAc)_2$ + $(Man)_3(GlcNAc)_2$
16	2081,1	10,9	0,3	$(GlcNAc)_3$ $(Fuc)_1$ + $(Man)_3(GlcNAc)_2$
17	2111,1	1,3	0,2	$(Gal)_1$ $(GlcNAc)_3$ + $(Man)_3(GlcNAc)_2$
18	2244,1	6,0	0,2	$(Gal)_2$ $(GlcNAc)_2$ $(Fuc)_1$ + $(Man)_3(GlcNAc)_2$
19	2285,2	8,3	0,2	$(Gal)_1$ $(GlcNAc)_3$ $(Fuc)_1$ + $(Man)_3(GlcNAc)_2$
20	2315,2	0,7	0,5	$(Gal)_2$ $(GlcNAc)_3$ + $(Man)_3(GlcNAc)_2$
21	2326,2	4,6	0,2	$(GlcNAc)_4$ $(Fuc)_1$ + $(Man)_3(GlcNAc)_2$
22	2356,2	0,1	0,3	$(Gal)_1$ $(GlcNAc)_4$ + $(Man)_3(GlcNAc)_2$
23	2401,2	1,1	0,0	$(NeuAc)_1$ $(Gal)_1$ $(GlcNAc)_2$ $(Fuc)_1$ + $(Man)_3(GlcNAc)_2$
24	2489,2	7,4	0,2	$(Gal)_2$ $GlcNAc)_3$ $(Fuc)_1$ + $(Man)_3(GlcNAc)_2$
25	2519,3	0,4	0,5	$(Gal)_3$ $(GlcNAc)_3$ + $(Man)_3(GlcNAc)_2$
26	2530,3	2,2	0,1	$(Gal)_1$ $(GlcNAc)_4$ $(Fuc)_1$ + $(Man)_3(GlcNAc)_2$
27	2571,3	1,7	0,1	$(GlcNAc)_5$ $(Fuc)_1$ + $(Man)_3(GlcNAc)_2$
28	2605,3	4,5	0,1	$(NeuAc)_1$ $(Gal)_2$ $(GlcNAc)_2$ $(Fuc)_1$ + $(Man)_3(GlcNAc)_2$
29	2646,3	1,6	0,1	$(NeuAc)_1$ $(Gal)_1$ $(GlcNAc)_3$ $(Fuc)_1$ + $(Man)_3(GlcNAc)_2$
30	2693,4	1,4	0,1	$(Gal)_3$ $(GlcNAc)_3$ $(Fuc)_1$ + $(Man)_3(GlcNAc)_2$
31	2734,4	1,5	0,1	$(Gal)_2$ $(GlcNAc)_4$ $(Fuc)_1$ + $(Man)_3(GlcNAc)_2$
32	2755,4	1,0	0,1	$(Gal)_1$ $(GlcNAc)_5$ $(Fuc)_1$ + $(Man)_3(GlcNAc)_2$
33	2850,4	1,6	0,1	$(NeuAc)_1$ $(Gal)_2$ $(GlcNAc)_3$ $(Fuc)_1$ + $(Man)_3(GlcNAc)_2$
34	2891,5	0,6	0,1	$(NeuAc)_1$ $(Gal)_1$ $(GlcNAc)_4$ $(Fuc)_1$ + $(Man)_3(GlcNAc)_2$
35	2938,5	1,3	0,0	$(Gal)_3$ $(GlcNAc)_4$ $(Fuc)_1$ + $(Man)_3(GlcNAc)_2$
36	2966,5	1,0	0,1	$(NeuAc)_2$ $(Gal)_2$ $(GlcNAc)_2$ $(Fuc)_1$ + $(Man)_3(GlcNAc)_2$
37	2979,5	0,7	0,1	$(Gal)_2$ $(GlcNAc)_5$ $(Fuc)_1$ + $(Man)_3(GlcNAc)_2$
38	3054,5	1,1	0,1	$(NeuAc)_1$ $(Gal)_3$ $(GlcNAc)_3$ $(Fuc)_1$ + $(Man)_3(GlcNAc)_2$
39	3095,5	0,7	0,1	$(NeuAc)_1$ $(Gal)_2$ $(GlcNAc)_4$ $(Fuc)_1$ + $(Man)_3(GlcNAc)_2$
40	3142,6	0,6	0,1	$(Gal)_4$ $(GlcNAc)_4$ $(Fuc)_1$ + $(Man)_3(GlcNAc)_2$
41	3183,6	0,5	0,1	$(Gal)_3$ $(GlcNAc)_5$ $(Fuc)_1$ + $(Man)_3(GlcNAc)_2$
42	3299,7	0,9	0,1	$(NeuAc)_1$ $(Gal)_3$ $(GlcNAc)_4$ $(Fuc)_1$ + $(Man)_3(GlcNAc)_2$
43	3387,7	0,5	0,1	$(Gal)_4$ $(GlcNAc)_5$ $(Fuc)_1$ + $(Man)_3(GlcNAc)_2$
44	3415,7	0,5	0,0	$(NeuAc)_2$ $(Gal)_3$ $(GlcNAc)_3$ $(Fuc)_1$ + $(Man)_3(GlcNAc)_2$
45	3503,8	0,5	0,1	$(NeuAc)_1$ $(Gal)_4$ $(GlcNAc)_4$ $(Fuc)_1$ + $(Man)_3(GlcNAc)_2$
46	3544,8	0,3	0,2	$(NeuAc)_1$ $(Gal)_3$ $(GlcNAc)_5$ $(Fuc)_1$ + $(Man)_3(GlcNAc)_2$
47	3602,8	0,2	0,1	$(NeuAc)_3$ $(Gal)_3$ $(GlcNAc)_3$ + $(Man)_3(GlcNAc)_2$
48	3660,8	0,2	0,1	$(NeuAc)_2$ $(Gal)_3$ $(GlcNAc)_4$ $(Fuc)_1$ + $(Man)_3(GlcNAc)_2$
49	3748,9	0,4	0,1	$(NeuAc)_1$ $(Gal)_4$ $(GlcNAc)_5$ $(Fuc)_1$ + $(Man)_3(GlcNAc)_2$
50	3864,9	0,4	0,1	$(NeuAc)_2$ $(Gal)_4$ $(GlcNAc)_4$ $(Fuc)_1$ + $(Man)_3(GlcNAc)_2$
51	4110,1	0,1	0,1	$(NeuAc)_2$ $(Gal)_4$ $(GlcNAc)_5$ $(Fuc)_1$ + $(Man)_3(GlcNAc)_2$
52	4226,1	0,0	0,1	$(NeuAc)_3$ $(Gal)_4$ $(GlcNAc)_4$ $(Fuc)_1$ + $(Man)_3(GlcNAc)_2$

Table 10. MALDI-TOF MS spectra of N-linked glycans from BDBV GP1,2. Relative intensity (normalized to 100%) of 52 individual N-glycans are shown.

O-linked glycans were removed from GP1,2 beta-elimination for and analyzed by MALDI-TOF mass spectrometry. Spectra revealed 10 O-glycan structures (Figure 28). These structures are shown in Figure 29 and Table 11.

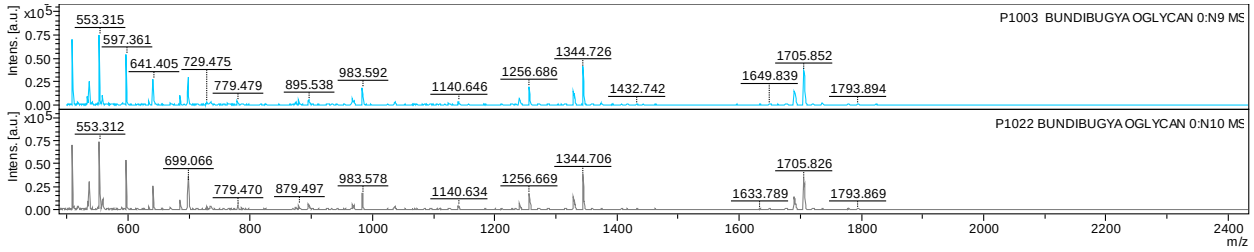


Figure 28. MALDI-TOF MS spectra of O-linked glycans from BDBV GP1,2. Each set of peaks is taken from a different section of the sample, showing consistancy of the assay.

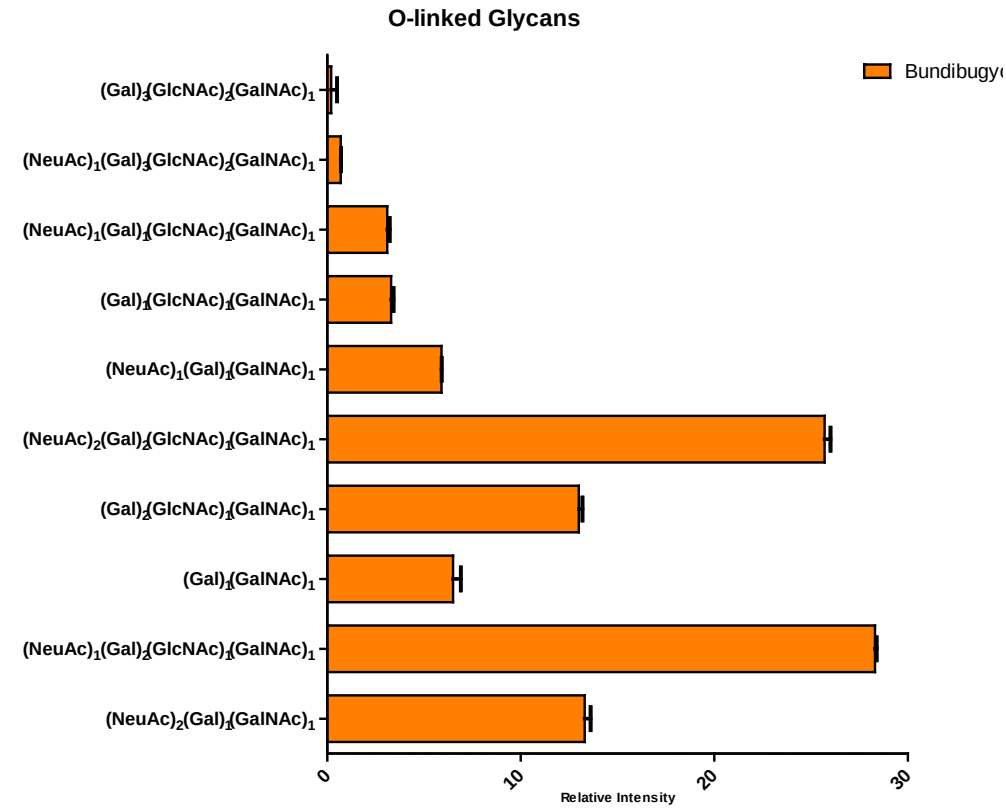


Figure 29. O-linked glycans from BDBV GP1,2. Relative intensity (normalized to 100%) of 10 individual N-glycans are shown. Error bars represent standard error of 2 separate spots of the pooled GP1,2 from four separate experiments.

O-glycan	Relative intenstiy	SD
$(NeuAc)_2(Gal)_1(GalNAc)_1$	13.3	0.31
$(NeuAc)_1(Gal)_2(GlcNAc)_1(GalNAc)_1$	28.3	0.09
$(Gal)_1(GalNAc)_1$	6.5	0.40
$(Gal)_2(GlcNAc)_1(GalNAc)_1$	13.0	0.19
$(NeuAc)_2(Gal)_2(GlcNAc)_1(GalNAc)_1$	25.7	0.31
$(NeuAc)_1(Gal)_1(GalNAc)_1$	5.9	0.01
$(Gal)_1(GlcNAc)_1(GalNAc)_1$	3.3	0.13
$(NeuAc)_1(Gal)_1(GlcNAc)_1(GalNAc)_1$	3.1	0.13
$(NeuAc)_1(Gal)_3(GlcNAc)_2(GalNAc)_1$	0.7	0.02
$(Gal)_3(GlcNAc)_2(GalNAc)_1$	0.2	0.31

Table 11. MALDI-TOF MS spectra of O-linked glycans from BDBV GP1,2. Relative intensity (normalized to 100%) of 10 individual N-glycans are shown.

Glycan analysis of TAFV GP1,2 produced in mammalian 293T cells. TAFV GP1,2 (GeneID: 9487535) was modified as in Figure 2 and expressed in mammalian 293T cells. Purified GP1,2 was confirmed for purity and specificity by coomassie staining (Figure 30A), which stains all proteins; periodic schiff stain, which stains only glycosylated proteins (Figure 30B); and Western blotting with an anti-TAFV antibody (Figure 30D). Four individual experiments were performed, and the purified proteins were combined and analyzed by MALDI-TOF mass spectrometry (Proteodynamics). N-linked glycans on GP1,2 were released by incubation with PNGase F, which was confirmed by reduced protein mass (Figure 30D). Mass spectrometry analysis of the released glycans revealed a large number of different N-linked glycan structures with a wide range of frequency in TAFV GP1,2 (Figure 31).

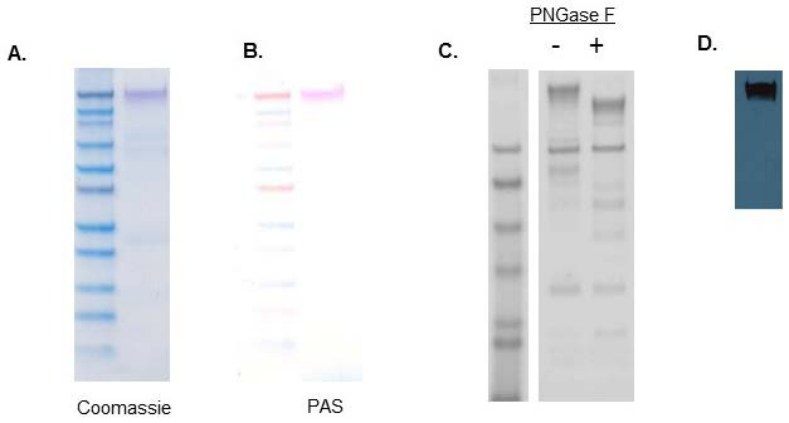


Figure 30. Purification of TAFV GP1,2 produced in mammalian 293T cells. A. Coomassie stain. B. PAS glycoprotein-specific stain. C. PNGase analysis showing presence of glycans in GP1,2. D. Western blot.

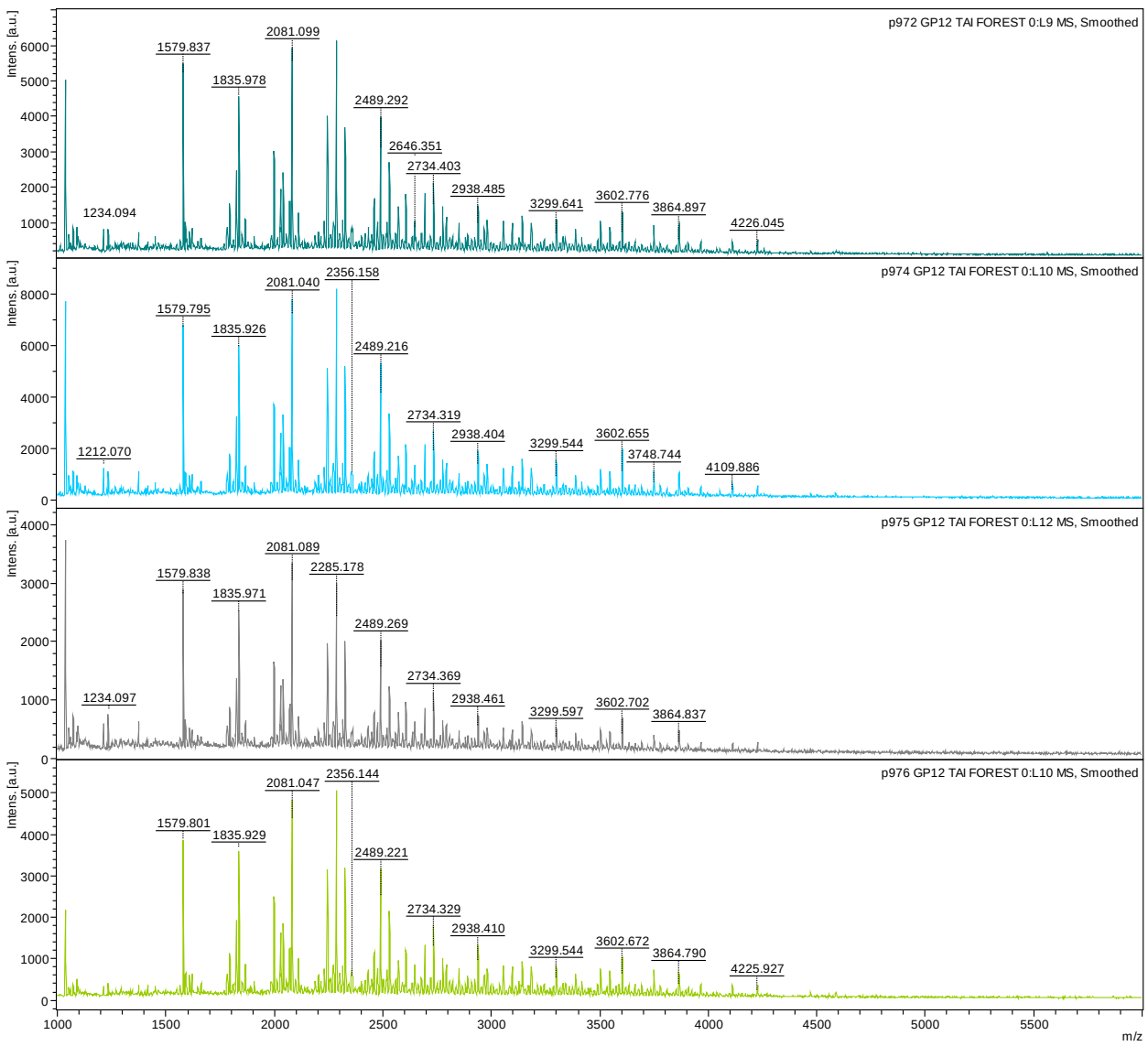


Figure 31. MALDI-TOF MS spectra of N-linked glycans from TAFV GP1,2. Each set of peaks is taken from a different section of the sample, showing consistancy of the assay.

Analysis of these data revealed 47 different N-linked glycan structures, a different number than EBOV-Mayinga (49), SUDV (56), or BDBV (52), including high mannose type structures from $(Man)_4(GlcNAc)_2$ to $(Man)_8(GlcNAc)_2$, hybrid glycans having only one GlcNAc residue in the antenna, and bi-, tri- and tetra-antennary complex glycans (Figure 32 shows graphical representation; Table 12 shows the data in a list format). Similar to EBOV, SUDV, and BDBV

GP1,2, dominant individual *N*-glycan structures are bi-antennary glycans: $(GlcNAc)_2(Fuc)_1$ + $(Man)_3(GlcNAc)_2$ (m/z 1835) at 7%, $(GlcNAc)_3(Fuc)_1$ + $(Man)_3(GlcNAc)_2$ (m/z 2081) at 8.5%. However, different from EBOV, SUDV, and BDBV, one other dominant N-glyan was

$(Gal)_1(GlcNAc)_3(Fuc)_1$ + $(Man)_3(GlcNAc)_2$ (m/z 2285); similar to SUDV, a dominant N-glycan was $(Man)_5(GlcNAc)_2$ (m/z 1580) at 8%.

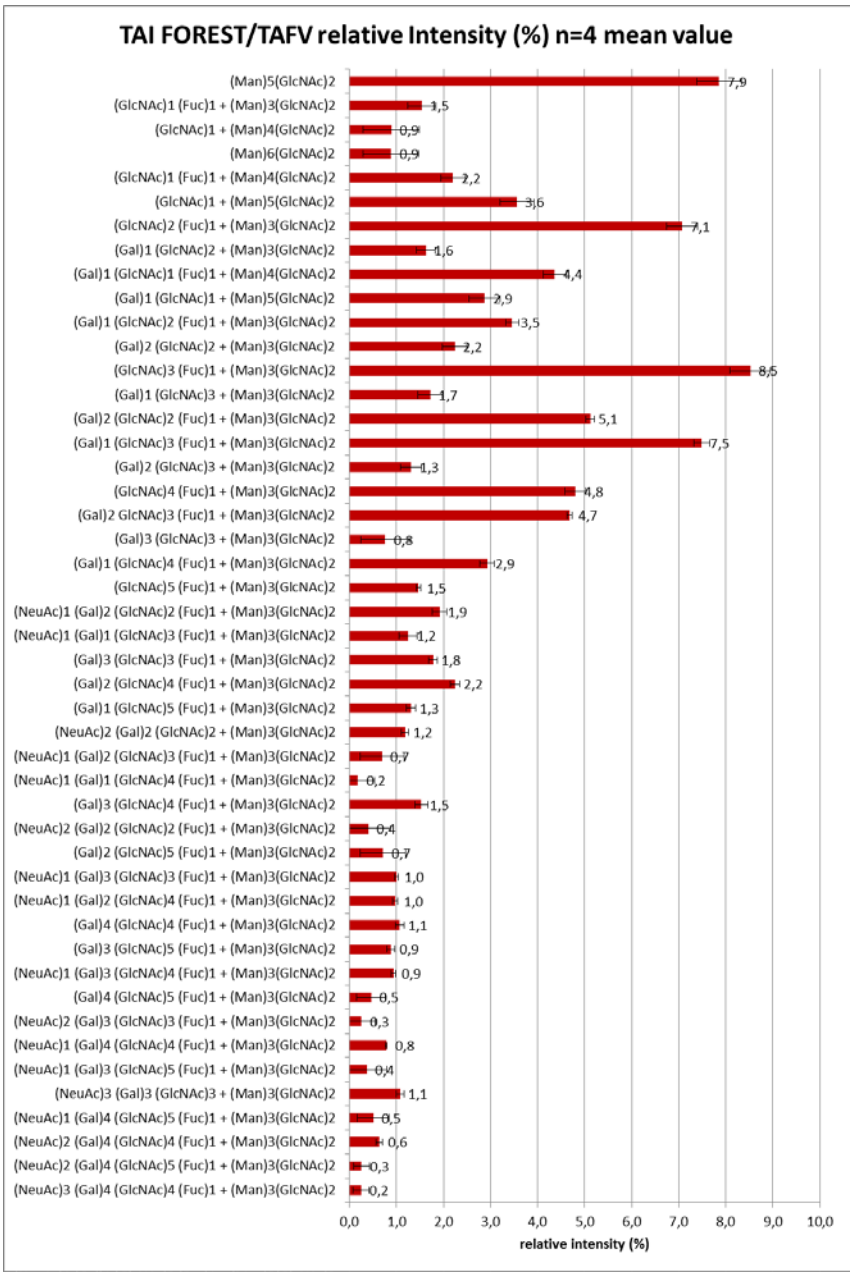


Figure 32. MALDI-TOF MS spectra of N-linked glycans from TAFV GP1,2. Relative intensity (normalized to 100%) of 47 individual N-glycans are shown. Error bars represent standard error of 4 separate spots of the pooled GP1,2 from four separate experiments.

		TAI FOREST/TAFV		
Structure N°	$[M+Na]^+$ theoretical monoisotopic mass	mean relative. Intensity %	standard deviation	Assignment
1	1579,8	7,9	0,5	$(Man)_5(GlcNAc)_2$
2	1590,8	1,5	0,3	$(GlcNAc)_1$ $(Fuc)_1$ + $(Man)_3(GlcNAc)_2$
3	1620,8	0,9	0,6	$(GlcNAc)_1$ + $(Man)_4(GlcNAc)_2$
4	1783,9	0,9	0,6	$(Man)_6(GlcNAc)_2$
5	1794,9	2,2	0,3	$(GlcNAc)_1$ $(Fuc)_1$ + $(Man)_4(GlcNAc)_2$
6	1824,9	3,6	0,4	$(GlcNAc)_1$ + $(Man)_5(GlcNAc)_2$
7	1835,9	7,1	0,3	$(GlcNAc)_2$ $(Fuc)_1$ + $(Man)_3(GlcNAc)_2$
8	1865,9	1,6	0,2	$(Gal)_1$ $(GlcNAc)_2$ + $(Man)_3(GlcNAc)_2$
9	1999,0	4,4	0,2	$(Gal)_1$ $(GlcNAc)_1$ $(Fuc)_1$ + $(Man)_4(GlcNAc)_2$
10	2029,0	2,9	0,3	$(Gal)_1$ $(GlcNAc)_1$ + $(Man)_5(GlcNAc)_2$
11	2040,0	3,5	0,1	$(Gal)_1$ $(GlcNAc)_2$ $(Fuc)_1$ + $(Man)_3(GlcNAc)_2$
12	2070,0	2,2	0,3	$(Gal)_2$ $(GlcNAc)_2$ + $(Man)_3(GlcNAc)_2$
13	2081,1	8,5	0,4	$(GlcNAc)_3$ $(Fuc)_1$ + $(Man)_3(GlcNAc)_2$
14	2111,1	1,7	0,3	$(Gal)_1$ $(GlcNAc)_3$ + $(Man)_3(GlcNAc)_2$
15	2244,1	5,1	0,1	$(Gal)_2$ $(GlcNAc)_2$ $(Fuc)_1$ + $(Man)_3(GlcNAc)_2$
16	2285,2	7,5	0,2	$(Gal)_1$ $(GlcNAc)_3$ $(Fuc)_1$ + $(Man)_3(GlcNAc)_2$
17	2315,2	1,3	0,2	$(Gal)_2$ $(GlcNAc)_3$ + $(Man)_3(GlcNAc)_2$
18	2326,2	4,8	0,2	$(GlcNAc)_4$ $(Fuc)_1$ + $(Man)_3(GlcNAc)_2$
19	2489,2	4,7	0,1	$(Gal)_2$ $GlcNAc)_3$ $(Fuc)_1$ + $(Man)_3(GlcNAc)_2$
20	2519,3	0,8	0,5	$(Gal)_3$ $(GlcNAc)_3$ + $(Man)_3(GlcNAc)_2$
21	2530,3	2,9	0,2	$(Gal)_1$ $(GlcNAc)_4$ $(Fuc)_1$ + $(Man)_3(GlcNAc)_2$
22	2571,3	1,5	0,0	$(GlcNAc)_5$ $(Fuc)_1$ + $(Man)_3(GlcNAc)_2$
23	2605,3	1,9	0,2	$(NeuAc)_1$ $(Gal)_2$ $(GlcNAc)_2$ $(Fuc)_1$ + $(Man)_3(GlcNAc)_2$
24	2646,3	1,2	0,2	$(NeuAc)_1$ $(Gal)_1$ $(GlcNAc)_3$ $(Fuc)_1$ + $(Man)_3(GlcNAc)_2$
25	2693,4	1,8	0,1	$(Gal)_3$ $(GlcNAc)_3$ $(Fuc)_1$ + $(Man)_3(GlcNAc)_2$
26	2734,4	2,2	0,1	$(Gal)_2$ $(GlcNAc)_4$ $(Fuc)_1$ + $(Man)_3(GlcNAc)_2$
27	2775,4	1,3	0,1	$(Gal)_1$ $(GlcNAc)_5$ $(Fuc)_1$ + $(Man)_3(GlcNAc)_2$
28	2792,4	1,2	0,1	$(NeuAc)_2$ $(Gal)_2$ $(GlcNAc)_2$ + $(Man)_3(GlcNAc)_2$
29	2850,4	0,7	0,5	$(NeuAc)_1$ $(Gal)_2$ $(GlcNAc)_3$ $(Fuc)_1$ + $(Man)_3(GlcNAc)_2$
30	2891,5	0,2	0,4	$(NeuAc)_1$ $(Gal)_1$ $(GlcNAc)_4$ $(Fuc)_1$ + $(Man)_3(GlcNAc)_2$
31	2938,5	1,5	0,1	$(Gal)_3$ $(GlcNAc)_4$ $(Fuc)_1$ + $(Man)_3(GlcNAc)_2$
32	2966,5	0,4	0,5	$(NeuAc)_2$ $(Gal)_2$ $(GlcNAc)_2$ $(Fuc)_1$ + $(Man)_3(GlcNAc)_2$
33	2979,5	0,7	0,5	$(Gal)_2$ $(GlcNAc)_5$ $(Fuc)_1$ + $(Man)_3(GlcNAc)_2$
34	3054,5	1,0	0,0	$(NeuAc)_1$ $(Gal)_3$ $(GlcNAc)_3$ $(Fuc)_1$ + $(Man)_3(GlcNAc)_2$
35	3095,5	1,0	0,1	$(NeuAc)_1$ $(Gal)_2$ $(GlcNAc)_4$ $(Fuc)_1$ + $(Man)_3(GlcNAc)_2$
36	3142,6	1,1	0,1	$(Gal)_4$ $(GlcNAc)_4$ $(Fuc)_1$ + $(Man)_3(GlcNAc)_2$
37	3183,6	0,9	0,1	$(Gal)_3$ $(GlcNAc)_5$ $(Fuc)_1$ + $(Man)_3(GlcNAc)_2$
38	3299,7	0,9	0,0	$(NeuAc)_1$ $(Gal)_3$ $(GlcNAc)_4$ $(Fuc)_1$ + $(Man)_3(GlcNAc)_2$
39	3387,7	0,5	0,3	$(Gal)_4$ $(GlcNAc)_5$ $(Fuc)_1$ + $(Man)_3(GlcNAc)_2$
40	3415,7	0,3	0,3	$(NeuAc)_2$ $(Gal)_3$ $(GlcNAc)_3$ $(Fuc)_1$ + $(Man)_3(GlcNAc)_2$
41	3503,8	0,8	0,0	$(NeuAc)_1$ $(Gal)_4$ $(GlcNAc)_4$ $(Fuc)_1$ + $(Man)_3(GlcNAc)_2$
42	3544,8	0,4	0,4	$(NeuAc)_1$ $(Gal)_3$ $(GlcNAc)_5$ $(Fuc)_1$ + $(Man)_3(GlcNAc)_2$
43	3602,8	1,1	0,1	$(NeuAc)_3$ $(Gal)_3$ $(GlcNAc)_3$ + $(Man)_3(GlcNAc)_2$
44	3748,9	0,5	0,3	$(NeuAc)_1$ $(Gal)_4$ $(GlcNAc)_5$ $(Fuc)_1$ + $(Man)_3(GlcNAc)_2$
45	3864,9	0,6	0,1	$(NeuAc)_2$ $(Gal)_4$ $(GlcNAc)_4$ $(Fuc)_1$ + $(Man)_3(GlcNAc)_2$
46	4110,1	0,3	0,2	$(NeuAc)_2$ $(Gal)_4$ $(GlcNAc)_5$ $(Fuc)_1$ + $(Man)_3(GlcNAc)_2$
47	4226,1	0,2	0,2	$(NeuAc)_3$ $(Gal)_4$ $(GlcNAc)_4$ $(Fuc)_1$ + $(Man)_3(GlcNAc)_2$
		100,0		

Table 12. MALDI-TOF MS spectra of N-linked glycans from TAFV GP1,2. Relative intensity (normalized to 100%) of 47 individual N-glycans are shown.

O-linked glycans were removed from GP1,2 beta-elimination for and analyzed by MALDI-TOF mass spectrometry. Spectra revealed 10 O-glycan structures (Figure 33). These structures are shown in Figure 34 and Table 13.

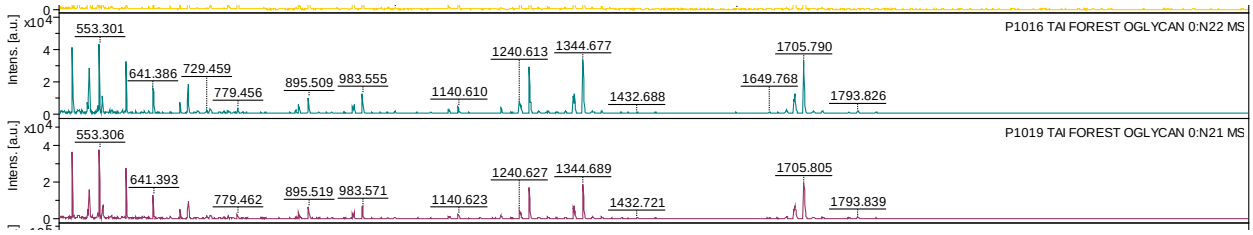


Figure 33. MALDI-TOF MS spectra of O-linked glycans from TAFV GP1,2. Each set of peaks is taken from a different section of the sample, showing consistancy of the assay.

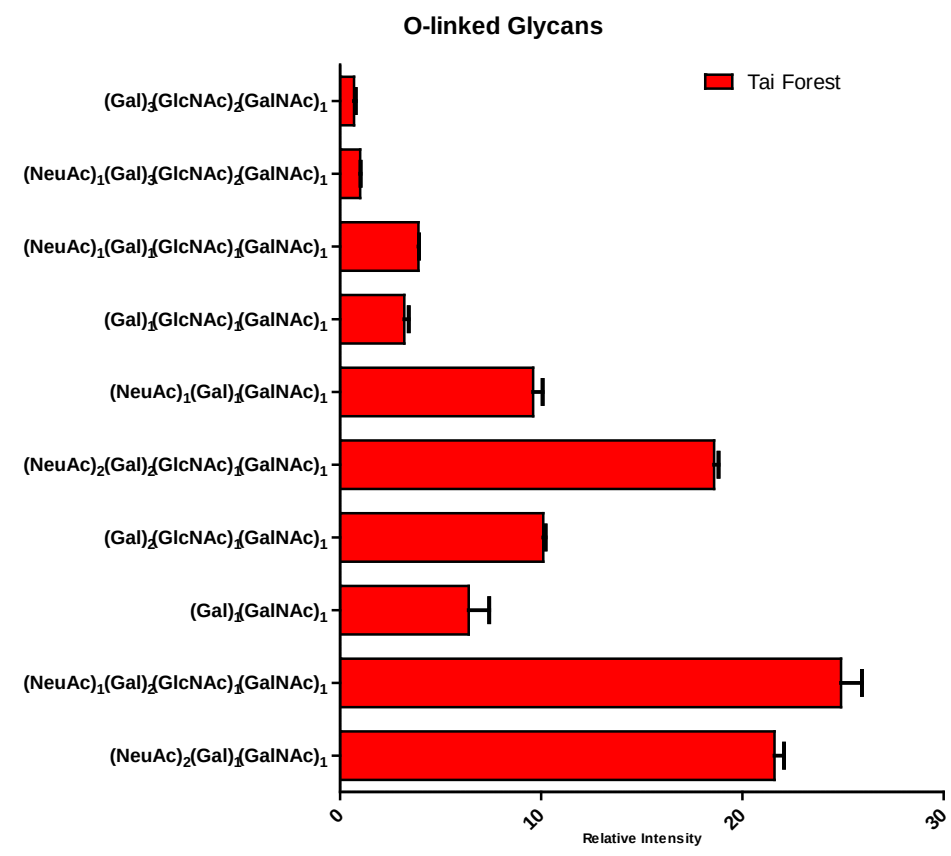


Figure 34. O-linked glycans from TAFV GP1,2. Relative intensity (normalized to 100%) of 10 individual N-glycans are shown. Error bars represent standard error of 2 separate spots of the pooled GP1,2 from four separate experiments.

O-glycan	Relative intenstiy	SD
$(NeuAc)_2(Gal)_1(GalNAc)_1$	21.6	0.47
$(NeuAc)_1(Gal)_2(GlcNAc)_1(GalNAc)_1$	24.9	1.04
$(Gal)_1(GalNAc)_1$	6.4	1.02
$(Gal)_2(GlcNAc)_1(GalNAc)_1$	10.1	0.12
$(NeuAc)_2(Gal)_2(GlcNAc)_1(GalNAc)_1$	18.6	0.22
$(NeuAc)_1(Gal)_1(GalNAc)_1$	9.6	0.47
$(Gal)_1(GlcNAc)_1(GalNAc)_1$	3.2	0.22
$(NeuAc)_1(Gal)_1(GlcNAc)_1(GalNAc)_1$	3.9	0.03
$(NeuAc)_1(Gal)_3(GlcNAc)_2(GalNAc)_1$	1.0	0.02
$(Gal)_3(GlcNAc)_2(GalNAc)_1$	0.7	0.08

Table 13. MALDI-TOF MS spectra of O-linked glycans from TAFV GP1,2. Relative intensity (normalized to 100%) of 10 individual N-glycans are shown.

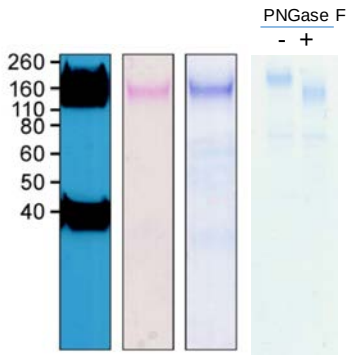


Figure 35. Purification of MARV-Angola GP1,2 expressed in mammalian cells. Purified MARV-Angola GP1,2 was analyzed by western blot (1st lane), PAS (2nd lane), colloidal blue (3rd lane), and PNGase F cleavage (4th lane).

Glycan analysis of MARV-Angola GP1,2 produced in mammalian 293T cells. MARV-Angola GP1,2 was modified as in Figure 2 and expressed in mammalian 293T cells. Purified GP1,2 was confirmed for purity and specificity by western blot (Figure 35, 1st lane), periodic acid schiff stain (PAS), which stains only glycosylated proteins (Figure 35, 2nd lane), coomassie staining (Figure 35, 3rd lane), which stains all proteins; and PNGaseF treatment, which removes N-glycans (Figure 35, 4th lane). Two individual experiments were performed, and the purified proteins were combined and analyzed by MALDI-TOF mass spectrometry (Proteodynamics). N-linked glycans on GP1,2 were released by incubation with PNGase F, which was confirmed by reduced protein mass. Mass spectrometry analysis of the released glycans revealed a large number of different N-linked glycan

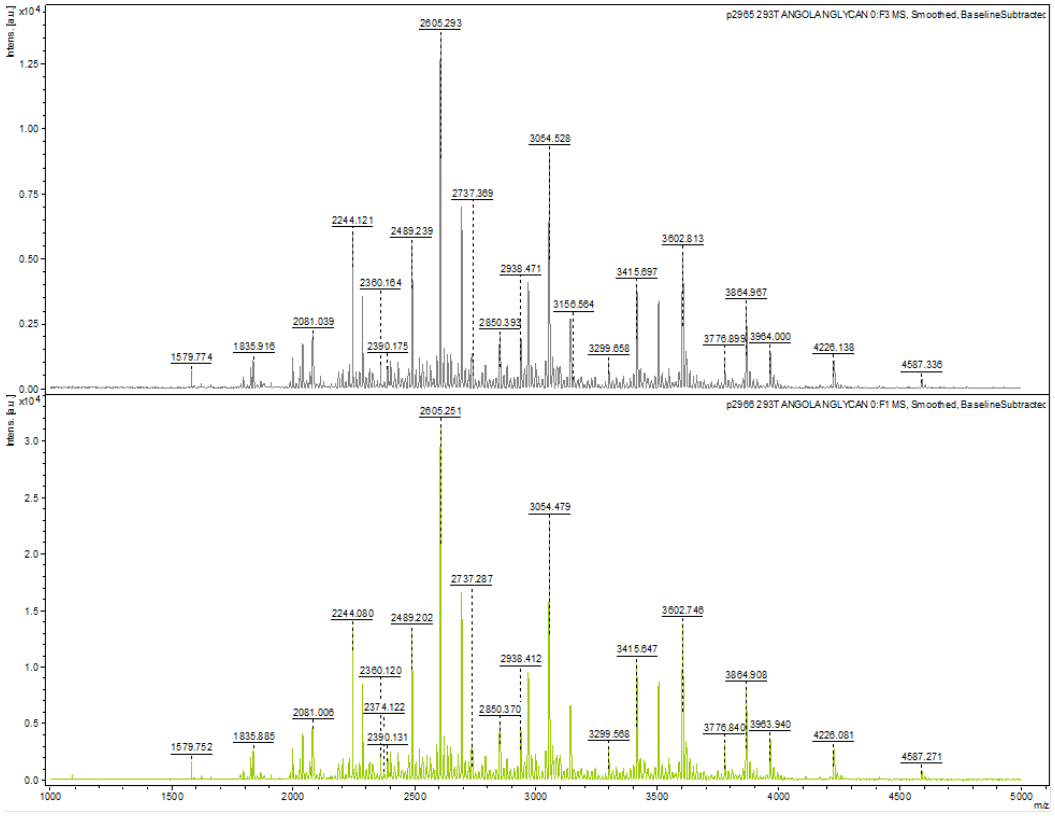


Figure 36. MALDI-TOF MS spectra of N-linked glycans from MARV-Angola GP1,2. Each set of peaks is taken from a different section of the sample, showing consistancy of the assay.

structures with a wide range of frequency in GP1,2 (Figure 36).

Analysis of these data revealed 37 different N-linked glycan structures, including high mannose type structures from $(Man)_4(GlcNAc)_2$ to $(Man)_8(GlcNAc)_2$, hybrid glycans having only one GlcNAc residue in the antenna, and bi-, tri- and tetra-antennary complex glycans (Figure 37 shows graphical representation; Table 14 shows the data in a list format). Dominant individual *N*-glycan structures included $(NeuAc)_1(Gal)_2(GlcNAc)_2(Fuc)_1$ + $(Man)_3(GlcNAc)_2$ at 15.4%;

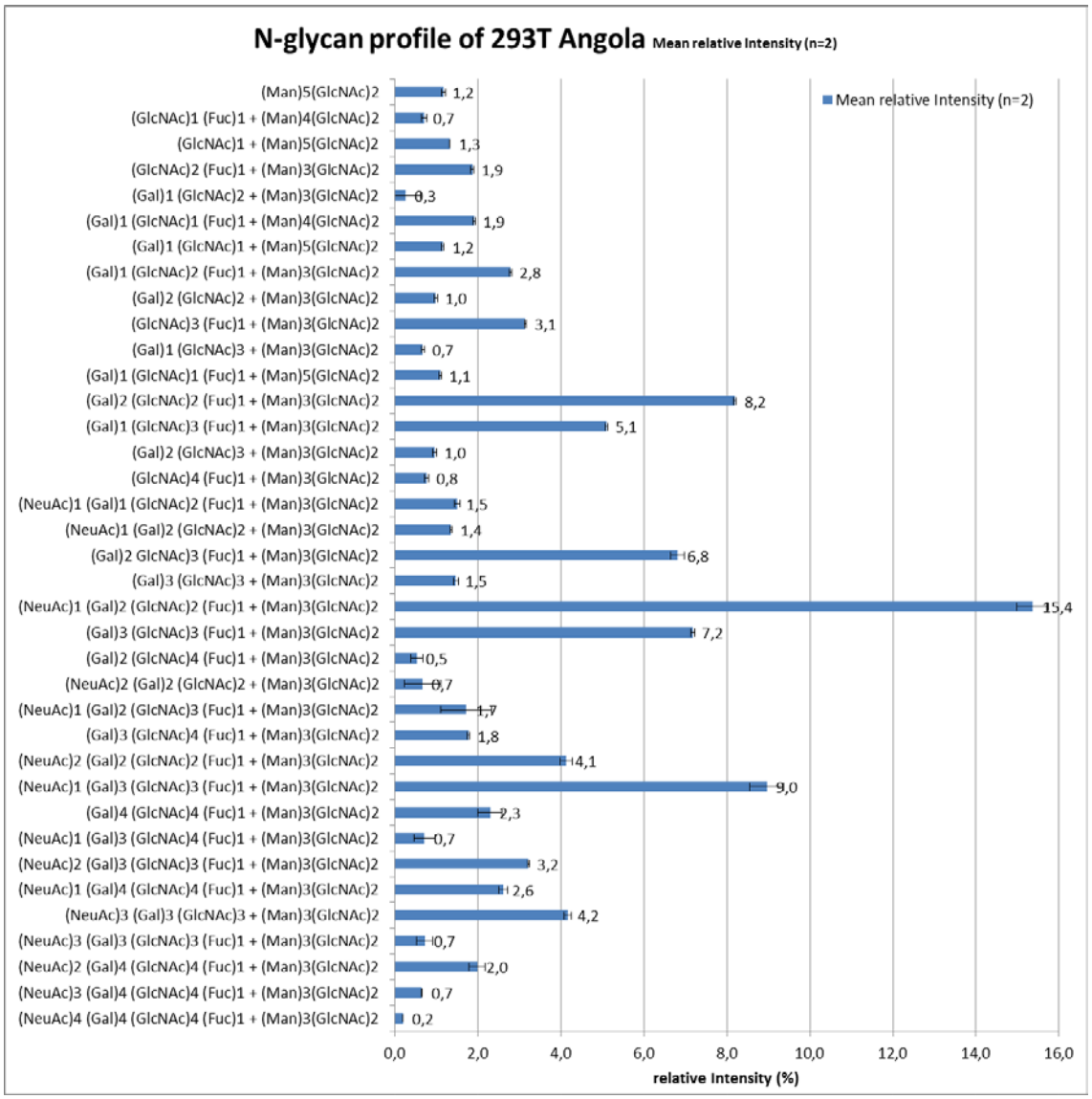


Figure 37. MALDI-TOF MS analysis of N-linked glycans from MARV-Angola GP1,2. Relative intensity (normalized to 100%) of 49 individual N-glycans are shown. Error bars represent standard error of 4 separate spots of the pooled GP1,2 from two separate experiments.

$(NeuAc)_1(Gal)_3(GlcNAc)_3(Fuc)_1$ + $(Man)_3(GlcNAc)_2$ at 9.0%; $(Gal)_2(GlcNAC)_2(Fuc)1$ + $(Man)_3(GlcNAc)_2$ at 8.2%; $(Gal)_3(GlcNAc)_3(Fuc)_1$ + $(Man)_3(GlcNAc)_2$ at 7.2%; and $(Gal)_2(GlcNAc)_3(Fuc)_1$ + $(Man)_3(GlcNAc)_2$ at 6.8%.

Structure N°	$[M+Na]^+$ theoretical monoisotopic mass	293T Angola P2965 rel. Intensity	293T Angola P2966 rel. Intensity	Mean relative Intensity (n=2)	Standard deviation	Assignment/Interpretation
1	1579,8	1,1	1,2	1,2	0,1	$(Man)_5(GlcNAc)_2$
2	1794,9	0,7	0,7	0,7	0,1	$(GlcNAc)_1$ $(Fuc)_1$ + $(Man)_4(GlcNAc)_2$
3	1824,9	1,3	1,3	1,3	0,0	$(GlcNAc)_1$ + $(Man)_5(GlcNAc)_2$
4	1835,9	1,8	1,9	1,9	0,0	$(GlcNAc)_2$ $(Fuc)_1$ + $(Man)_3(GlcNAc)_2$
5	1865,9	0,5	0,0	0,3	0,4	$(Gal)_1$ $(GlcNAc)_2$ + $(Man)_3(GlcNAc)_2$
6	1999,0	1,9	1,9	1,9	0,0	$(Gal)_1$ $(GlcNAc)_1$ $(Fuc)_1$ + $(Man)_4(GlcNAc)_2$
7	2029,0	1,2	1,1	1,2	0,0	$(Gal)_1$ $(GlcNAc)_1$ + $(Man)_5(GlcNAc)_2$
8	2040,0	2,8	2,8	2,8	0,0	$(Gal)_1$ $(GlcNAc)_2$ $(Fuc)_1$ + $(Man)_3(GlcNAc)_2$
9	2070,0	1,0	1,0	1,0	0,0	$(Gal)_2$ $(GlcNAc)_2$ + $(Man)_3(GlcNAc)_2$
10	2081,1	3,2	3,1	3,1	0,0	$(GlcNAc)_3$ $(Fuc)_1$ + $(Man)_3(GlcNAc)_2$
11	2111,1	0,7	0,6	0,7	0,0	$(Gal)_1$ $(GlcNAc)_3$ + $(Man)_3(GlcNAc)_2$
12	2203,1	1,1	1,1	1,1	0,0	$(Gal)_1$ $(GlcNAc)_1$ $(Fuc)_1$ + $(Man)_5(GlcNAc)_2$
13	2244,1	8,2	8,2	8,2	0,0	$(Gal)_2$ $(GlcNAc)_2$ $(Fuc)_1$ + $(Man)_3(GlcNAc)_2$
14	2285,2	5,1	5,1	5,1	0,0	$(Gal)_1$ $(GlcNAc)_3$ $(Fuc)_1$ + $(Man)_3(GlcNAc)_2$
15	2315,2	1,0	0,9	1,0	0,1	$(Gal)_2$ $(GlcNAc)_3$ + $(Man)_3(GlcNAc)_2$
16	2326,2	0,7	0,8	0,8	0,1	$(GlcNAc)_4$ $(Fuc)_1$ + $(Man)_3(GlcNAc)_2$
17	2401,2	1,5	1,4	1,5	0,1	$(NeuAc)_1$ $(Gal)_1$ $(GlcNAc)_2$ $(Fuc)_1$ + $(Man)_3(GlcNAc)_2$
18	2431,2	1,4	1,3	1,4	0,0	$(NeuAc)_1$ $(Gal)_2$ $(GlcNAc)_2$ + $(Man)_3(GlcNAc)_2$
19	2489,2	6,7	6,9	6,8	0,2	$(Gal)_2$ $GlcNAc)_3$ $(Fuc)_1$ + $(Man)_3(GlcNAc)_2$
20	2519,3	1,5	1,4	1,5	0,1	$(Gal)_3$ $(GlcNAc)_3$ + $(Man)_3(GlcNAc)_2$
21	2605,3	15,6	15,1	15,4	0,4	$(NeuAc)_1$ $(Gal)_2$ $(GlcNAc)_2$ $(Fuc)_1$ + $(Man)_3(GlcNAc)_2$
22	2693,4	7,1	7,2	7,2	0,1	$(Gal)_3$ $(GlcNAc)_3$ $(Fuc)_1$ + $(Man)_3(GlcNAc)_2$
23	2734,4	0,4	0,6	0,5	0,1	$(Gal)_2$ $(GlcNAc)_4$ $(Fuc)_1$ + $(Man)_3(GlcNAc)_2$
24	2792,4	0,4	1,0	0,7	0,4	$(NeuAc)_2$ $(Gal)_2$ $(GlcNAc)_2$ + $(Man)_3(GlcNAc)_2$
25	2850,4	1,3	2,2	1,7	0,6	$(NeuAc)_1$ $(Gal)_2$ $(GlcNAc)_3$ $(Fuc)_1$ + $(Man)_3(GlcNAc)_2$
26	2938,5	1,7	1,8	1,8	0,0	$(Gal)_3$ $(GlcNAc)_4$ $(Fuc)_1$ + $(Man)_3(GlcNAc)_2$
27	2966,5	4,0	4,2	4,1	0,1	$(NeuAc)_2$ $(Gal)_2$ $(GlcNAc)_2$ $(Fuc)_1$ + $(Man)_3(GlcNAc)_2$
28	3054,5	8,7	9,2	9,0	0,4	$(NeuAc)_1$ $(Gal)_3$ $(GlcNAc)_3$ $(Fuc)_1$ + $(Man)_3(GlcNAc)_2$
29	3142,6	2,5	2,1	2,3	0,3	$(Gal)_4$ $(GlcNAc)_4$ $(Fuc)_1$ + $(Man)_3(GlcNAc)_2$
30	3299,7	0,9	0,5	0,7	0,2	$(NeuAc)_1$ $(Gal)_3$ $(GlcNAc)_4$ $(Fuc)_1$ + $(Man)_3(GlcNAc)_2$
31	3415,7	3,2	3,2	3,2	0,0	$(NeuAc)_2$ $(Gal)_3$ $(GlcNAc)_3$ $(Fuc)_1$ + $(Man)_3(GlcNAc)_2$
32	3503,8	2,7	2,5	2,6	0,1	$(NeuAc)_1$ $(Gal)_4$ $(GlcNAc)_4$ $(Fuc)_1$ + $(Man)_3(GlcNAc)_2$
33	3602,8	4,1	4,2	4,2	0,1	$(NeuAc)_3$ $(Gal)_3$ $(GlcNAc)_3$ + $(Man)_3(GlcNAc)_2$
34	3776,9	0,9	0,6	0,7	0,2	$(NeuAc)_3$ $(Gal)_3$ $(GlcNAc)_3$ $(Fuc)_1$ + $(Man)_3(GlcNAc)_2$
35	3864,9	2,1	1,8	2,0	0,2	$(NeuAc)_2$ $(Gal)_4$ $(GlcNAc)_4$ $(Fuc)_1$ + $(Man)_3(GlcNAc)_2$
36	4226,1	0,6	0,7	0,7	0,0	$(NeuAc)_3$ $(Gal)_4$ $(GlcNAc)_4$ $(Fuc)_1$ + $(Man)_3(GlcNAc)_2$
37	4587,3	0,2	0,2	0,2	0,0	$(NeuAc)_4$ $(Gal)_4$ $(GlcNAc)_4$ $(Fuc)_1$ + $(Man)_3(GlcNAc)_2$
		100,0	100,0	100,0		

Table 14. MALDI-TOF MS spectra of N-linked glycans from MARV-Angola GP1,2. Relative intensity (normalized to 100%) of 49 individual N-glycans are shown.

O-linked glycans were removed from GP1,2 beta-elimination for and analyzed by MALDI-TOF mass spectrometry. Spectra revealed 8 O-glycan structures (Figure 38). These structures are shown in Figure 39 and Table 15.

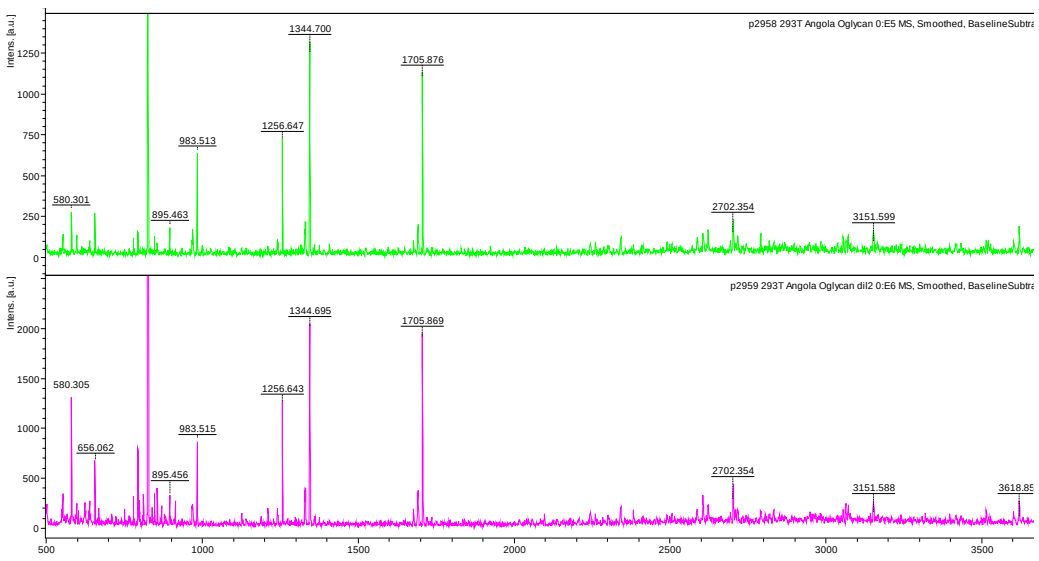


Figure 38. MALDI-TOF MS spectra of O-linked glycans from MARV-Angola GP1,2. Each set of peaks is taken from a different section of the sample, showing consistancy of the assay.

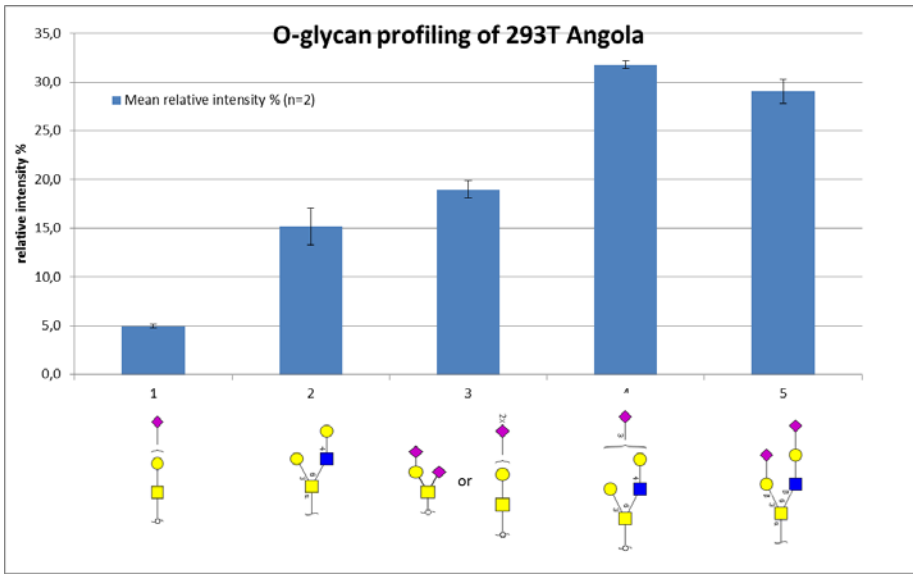


Figure 39. O-linked glycans from MARV-Angola GP1,2. Relative intensity (normalized to 100%) of 5 individual O-glycans are shown. Error bars represent standard error of 2 separate spots of the pooled GP1,2 from two separate

O-glycan	Relative intenstiy	SD
$(NeuAc)_2(Gal)_1(GalNAc)_1$	19.0	0.9
$(NeuAc)_1(Gal)_2(GlcNAc)_1(GalNAc)_1$	31.8	0.4
$(Gal)_2(GlcNAc)_1(GalNAc)_1$	15.2	1.9
$(NeuAc)_2(Gal)_2(GlcNAc)_1(GalNAc)_1$	29.0	1.2
$(NeuAc)_1(Gal)_1(GalNAc)_1$	5.0	0.2

Table 15. MALDI-TOF MS spectra of O-linked glycans from MARV-Angola. Relative intensity (normalized to 100%) of 5 individual N-glycans are shown.

Glycan analysis of EBOV-Mayinga GP1,2 produced in insect Sf9 cells. EBOV-Mayinga was modified, incorporated into baculovirus, and expressed in insect Sf9 cells by infection. Purified GP1,2 was confirmed for purity and specificity by coomassie staining (Figure 40A), which stains all proteins; periodic schiff stain, which stains only glycosylated proteins (Figure 40B); and Western blotting with an anti-EBOV antibody (Figure 40D). Two individual experiments were performed, and the purified proteins were combined and analyzed by MALDI-TOF mass spectrometry (Proteodynamics). N-linked glycans on GP1,2 were released by incubation with PNGase F, which was confirmed by reduced protein mass (Figure 40C). Mass spectrometry analysis of the released glycans revealed eight N-linked glycan structures for EBOV-Mayinga

GP1,2 (Figure 41. However, only one N-glycan, $(Fuc)_1 + (Man)_3(GlcNAc)_2$, was dominant, comprising 76% of total N-glycans (Figure 42 and Table 16).

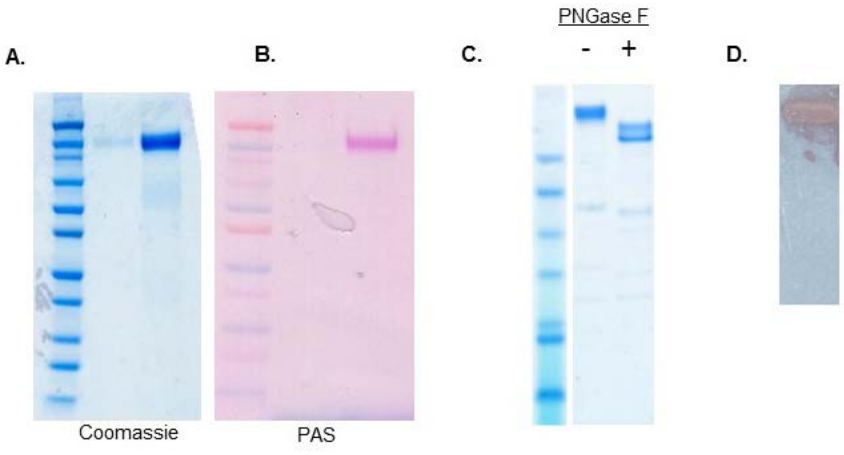


Figure 40. Purification of EBOV-Mayinga GP1,2 produced in insect Sf9 cells. A. Coomassie stain. B. PAS glycoprotein-specific stain. C. PNGase analysis showing presence of glycans in GP1,2. D. Western blot.

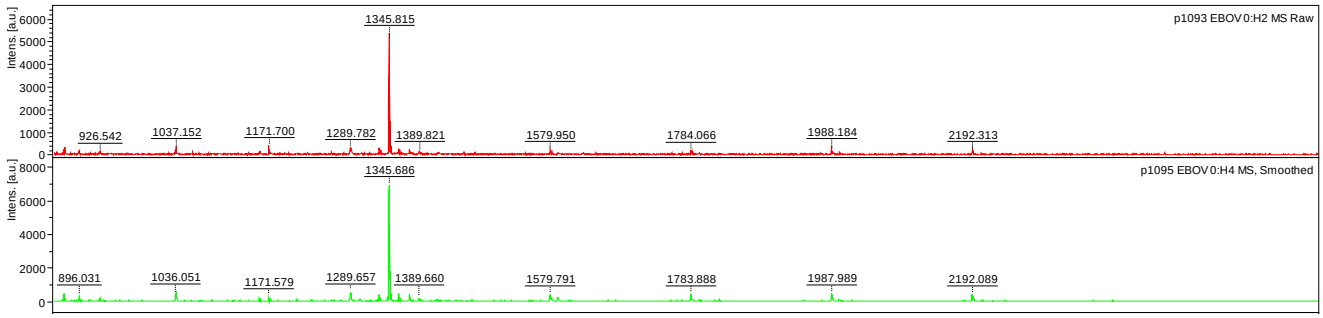


Figure 41. MALDI-TOF MS spectra of N-linked glycans from insect cell-derived EBOV-Mayinga GP1,2. Each set of peaks is taken from a different section of the sample, showing consistancy of the assay.

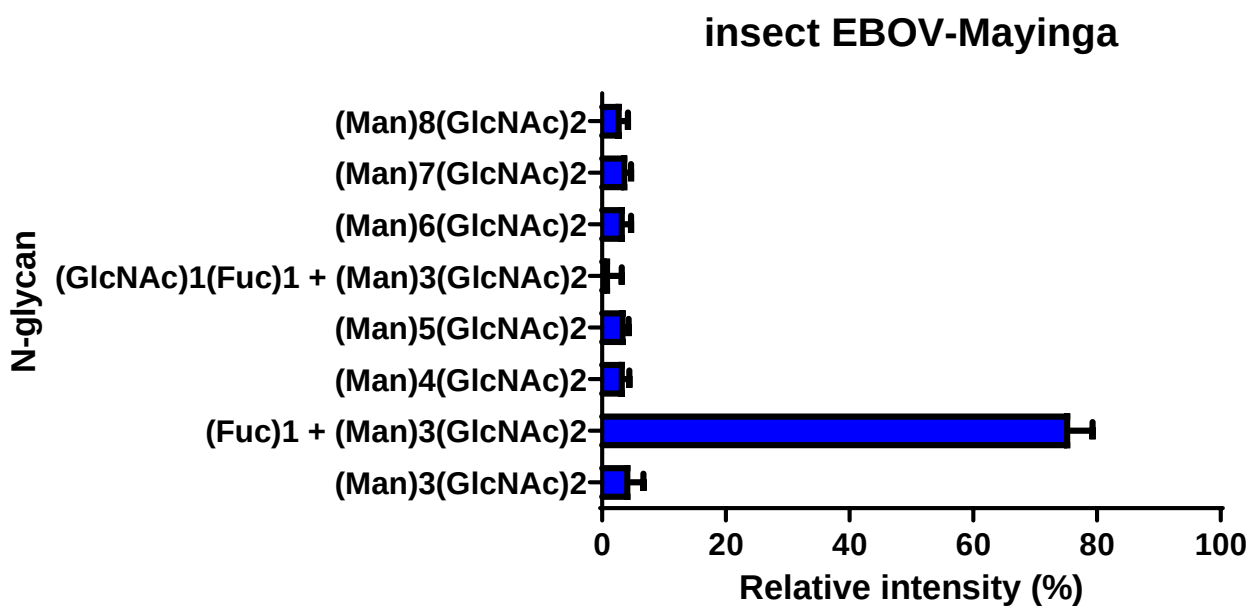


Figure 42. MALDI-TOF MS spectra of N-linked glycans from insect cell-derived EBOV-Mayinga GP1,2. Relative intensity (normalized to 100%) of 8 individual N-glycans are shown. Error bars represent standard error of 2 separate spots of the pooled GP1,2 from two separate experiments.

N-glycan	Relative intenstiy	SD
$(Man)_3\ (GlcNAc)_2$	4.6	2.1
$(Fuc)_1 + (Man)_3\ (GlcNAc)_2$	75.7	3.6
$(Man)_4\ (GlcNAc)_2$	3.7	0.7
$(Man)_5(GlcNAc)_2$	3.8	0.5
$(GlcNAc)_1\ (Fuc)_1 + (Man)_3(GlcNAc)_2$	1.3	1.9
$(Man)_6(GlcNAc)_2$	3.7	1.0
$(Man)_7(GlcNAc)_2$	4.1	0.6
$(Man)_8(GlcNAc)_2$	3.2	1.0

Table 16. MALDI-TOF MS spectra of N-linked glycans from insect cell-derived EBOV-Mayinga GP1,2. Relative intensity (normalized to 100%) of 8 individual N-glycans are shown.

O-linked glycans were removed from GP1,2 beta-elimination for and analyzed by MALDI-TOF mass spectrometry. However, no O-linked glycans were detected in EBOV-Mayinga GP1,2.

Glycan analysis of SUDV GP1,2 produced in insect Sf9 cells. SUDV was modified as in Figure 2, incorporated into baculovirus, and expressed in insect Sf9 cells by infection. Purified GP1,2 was confirmed for purity and specificity by coomassie staining (Figure 43A), which stains all proteins; periodic schiff stain, which stains only glycosylated proteins (Figure 43B); and Western blotting with an anti-SUDV antibody (Figure 43D). Two individual experiments were performed, and the purified proteins were combined and analyzed by MALDI-TOF mass spectrometry (Proteodynamics). N-linked glycans on GP1,2 were released by incubation with PNGase F, which was confirmed by reduced protein mass (Figure 43C). Mass spectrometry analysis of the released glycans revealed eight N-linked glycan structures for SUDV GP1,2 (Figure 44). However, only one N-glycan, $(Fuc)_1 + (Man)_3(GlcNAc)_2$, was dominant, comprising 77% of total N-glycans (Figure 45 and Table 17).

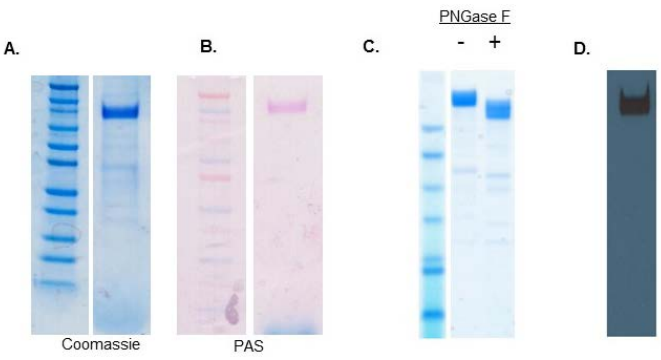


Figure 43. Purification of SUDV GP1,2 produced in insect Sf9 cells. A. Coomassie stain. B. PAS glycoprotein-specific stain. C. PNGase analysis showing presence of glycans in GP1,2. D. Western blot.

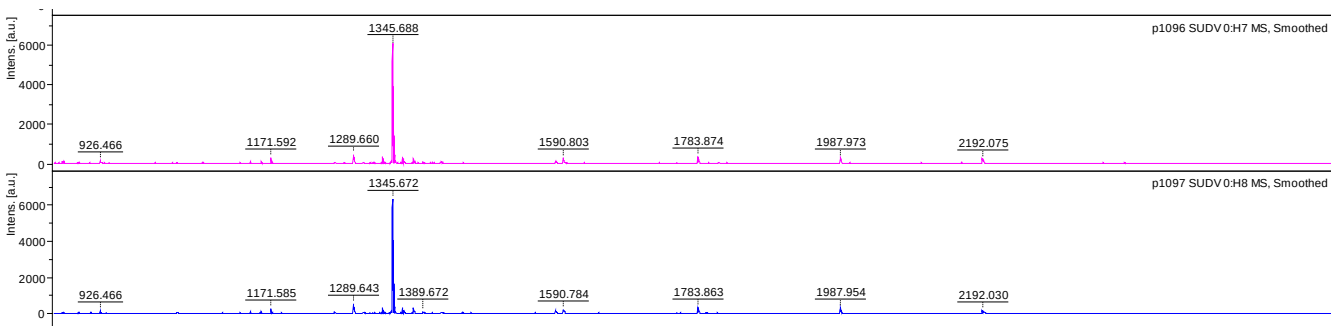


Figure 44. MALDI-TOF MS spectra of N-linked glycans from insect cell-derived SUDV GP1,2. Each set of peaks is taken from a different section of the sample, showing consistancy of the assay.

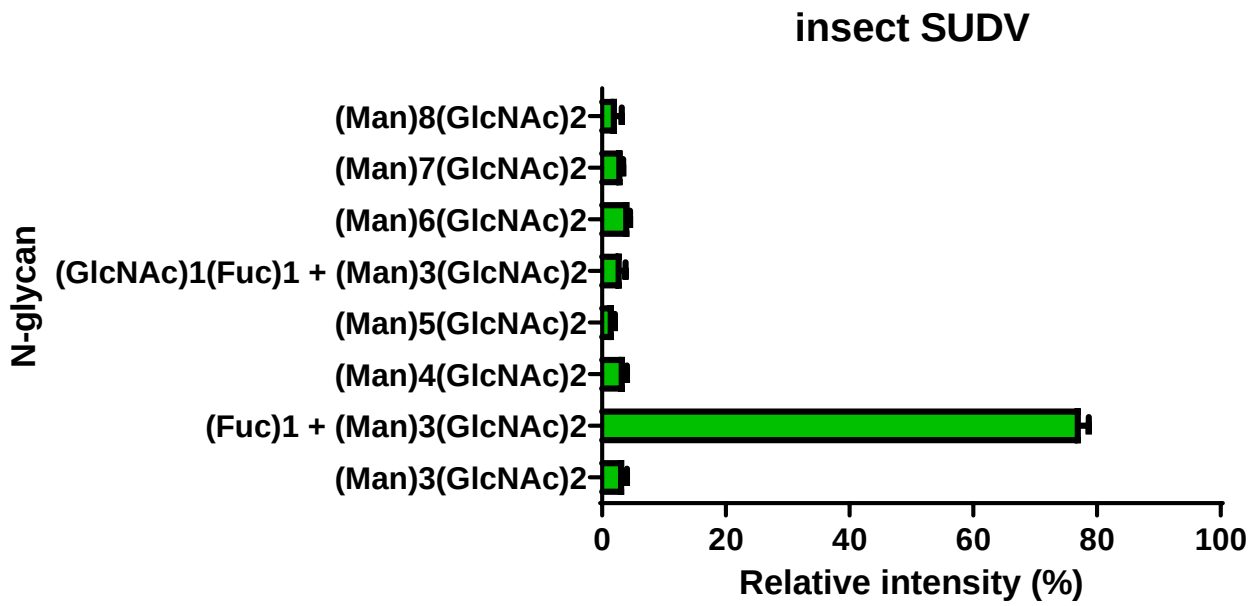


Figure 45. MALDI-TOF MS spectra of N-linked glycans from insect cell-derived SUDV GP1,2. Relative intensity (normalized to 100%) of 8 individual N-glycans are shown. Error bars represent standard error of 2 separate spots of the pooled GP1,2 from two separate experiments.

N-glycan	Relative intenstiy	SD
$(Man)_3\ (GlcNAc)_2$	3.6	0.4
$(Fuc)_1 + (Man)_3\ (GlcNAc)_2$	77.4	1.3
$(Man)_4\ (GlcNAc)_2$	3.7	0.3
$(Man)_5(GlcNAc)_2$	1.9	0.2
$(GlcNAc)_1\ (Fuc)_1 + (Man)_3(GlcNAc)_2$	3.2	0.6
$(Man)_6(GlcNAc)_2$	4.4	0.1
$(Man)_7(GlcNAc)_2$	3.3	0.1
$(Man)_8(GlcNAc)_2$	2.4	0.8

Table 17. MALDI-TOF MS spectra of N-linked glycans from insect cell-derived SUDV GP1,2. Relative intensity (normalized to 100%) of 8 individual N-glycans are shown.

O-linked glycans were removed from GP1,2 beta-elimination for and analyzed by MALDI-TOF mass spectrometry. However, no O-linked glycans were detected in SUDV GP1,2.

Glycan analysis of TAFV GP1,2 produced in insect Sf9 cells. TAFV was modified as in Figure 2, incorporated into baculovirus, and expressed in insect Sf9 cells by infection. Purified GP1,2 was confirmed for purity and specificity by coomassie staining (Figure 46A), which stains all proteins; periodic schiff stain, which stains only glycosylated proteins (Figure 46B); and Western blotting with an anti-TAFV antibody (Figure 46D). Two individual experiments were performed, and the purified proteins were combined and analyzed by MALDI-TOF mass spectrometry (Proteodynamics). N-linked glycans on GP1,2 were released by incubation with PNGase F, which was confirmed by reduced protein mass (Figure 46C). Mass spectrometry analysis of the released glycans revealed eight N-linked glycan structures for TAFV GP1,2 (Figure 47). However, only one N-glycan, $(Fuc)_1 + (Man)_3(GlcNAc)_2$, was dominant, comprising 75% of total N-glycans (Figure 48 and Table 18).

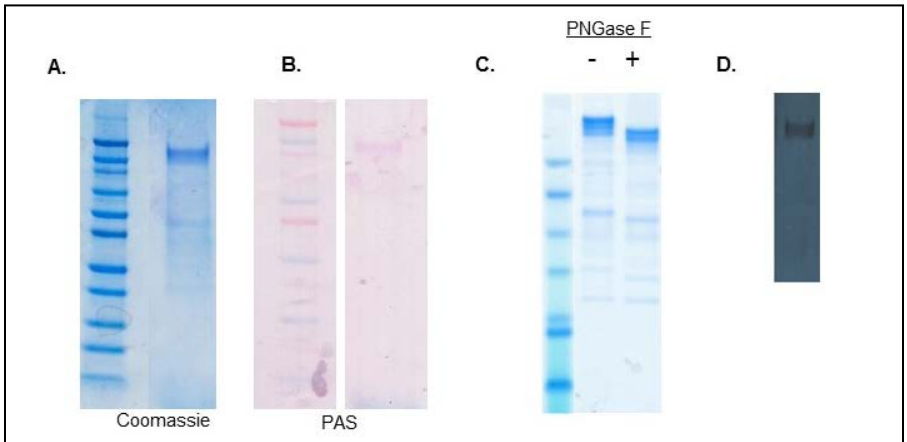


Figure 46. Purification of TAFV GP1,2 produced in insect Sf9 cells. A. Coomassie stain. B. PAS glycoprotein-specific stain. C. PNGase analysis showing presence of glycans in GP1,2. D. Western blot.

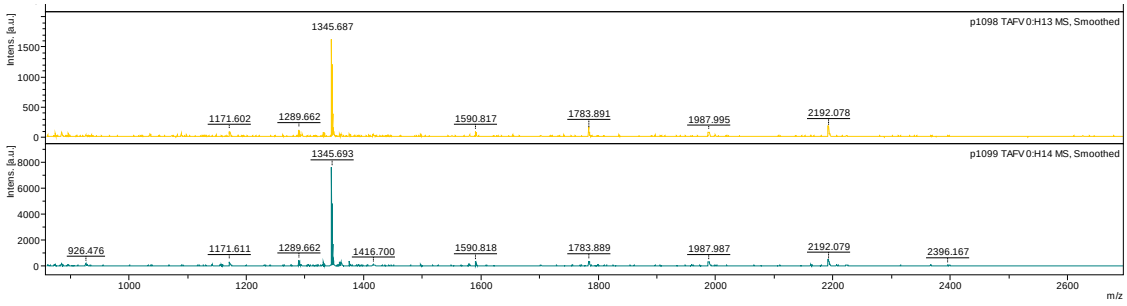


Figure 47. MALDI-TOF MS spectra of N-linked glycans from insect cell-derived TAFV GP1,2. Each set of peaks is taken from a different section of the sample, showing consistancy of the assay.

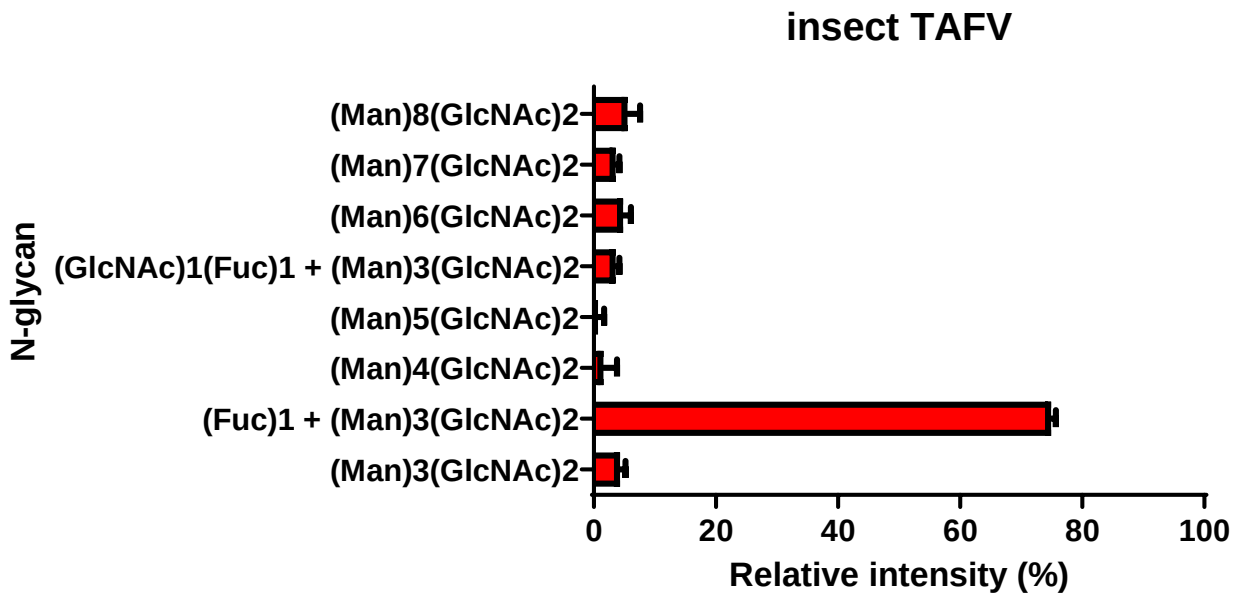


Figure 48. MALDI-TOF MS spectra of N-linked glycans from insect cell-derived TAFV GP1,2. Relative intensity (normalized to 100%) of 8 individual N-glycans are shown. Error bars represent standard error of 2 separate spots of the pooled GP1,2 from two separate experiments.

N-glycan	Relative intenstiy	SD
$(Man)_3\ (GlcNAc)_2$	4.3	0.9
$(Fuc)_1 + (Man)_3\ (GlcNAc)_2$	74.9	0.8
$(Man)_4\ (GlcNAc)_2$	1.6	2.2
$(Man)_5(GlcNAc)_2$	0.7	1.0
$(GlcNAc)_1\ (Fuc)_1 + (Man)_3(GlcNAc)_2$	3.6	0.6
$(Man)_6(GlcNAc)_2$	4.8	1.3
$(Man)_7(GlcNAc)_2$	3.6	0.6
$(Man)_8(GlcNAc)_2$	5.5	2.1

Table 18. MALDI-TOF MS spectra of N-linked glycans from insect cell-derived TAFV GP1,2. Relative intensity (normalized to 100%) of 8 individual N-glycans are shown.

O-linked glycans were removed from GP1,2 beta-elimination for and analyzed by MALDI-TOF mass spectrometry. However, no O-linked glycans were detected in TAFV GP1,2.

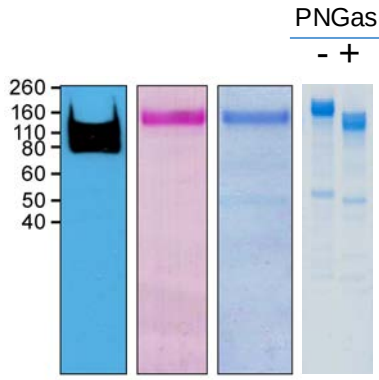


Figure 49. Purification of BDBV GP1,2 expressed in insect cells. Purified MARV-Angola GP1,2 was analyzed by western blot (1st lane), PAS (2nd lane), colloidal blue (3rd lane), and PNGase F cleavage (4th lane).

Glycan analysis of BDBV GP1,2 produced in insect Sf9 cells**.** BDBV was modified as in Figure 2, incorporated into baculovirus, and expressed in insect Sf9 cells by infection. Purified GP1,2 was confirmed for purity and specificity by coomassie staining, periodic schiff stain, and Western blotting with an anti-BDBV antibody (Figure 49). Two individual experiments were performed, and the purified proteins were combined and analyzed by MALDI-TOF mass spectrometry (Proteodynamics). N-linked glycans on GP1,2 were released by incubation with PNGase F, which was confirmed by reduced protein mass (Figure 49). Mass spectrometry analysis of the released glycans revealed eight N-linked glycan structures for BDBV GP1,2 (Figure 50). However, only one N-glycan, $(Fuc)_1 + (Man)_3(GlcNAc)_2$, was dominant, comprising 64.3% of total N-glycans (Figure 51 and Table 19).

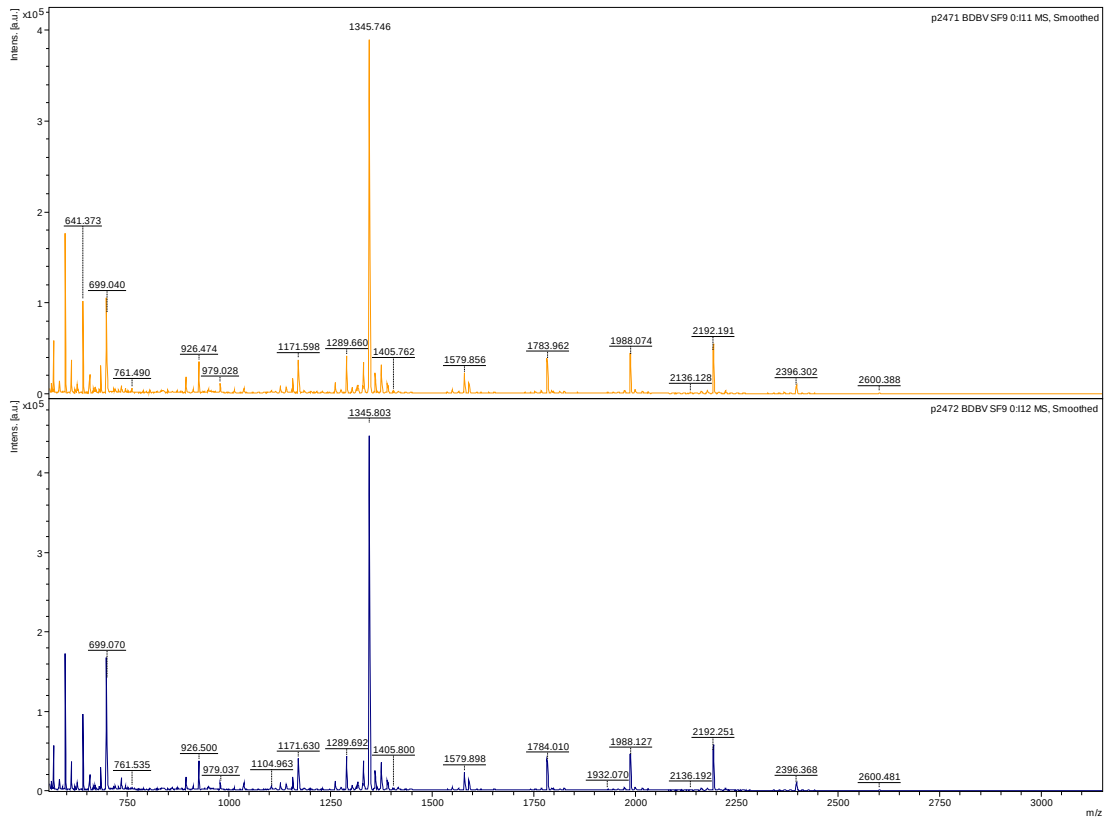


Figure 50. MALDI-TOF MS spectra of N-linked glycans from insect cell-derived BDBV GP1,2. Each set of peaks is taken from a different section of the sample, showing consistancy of the assay.

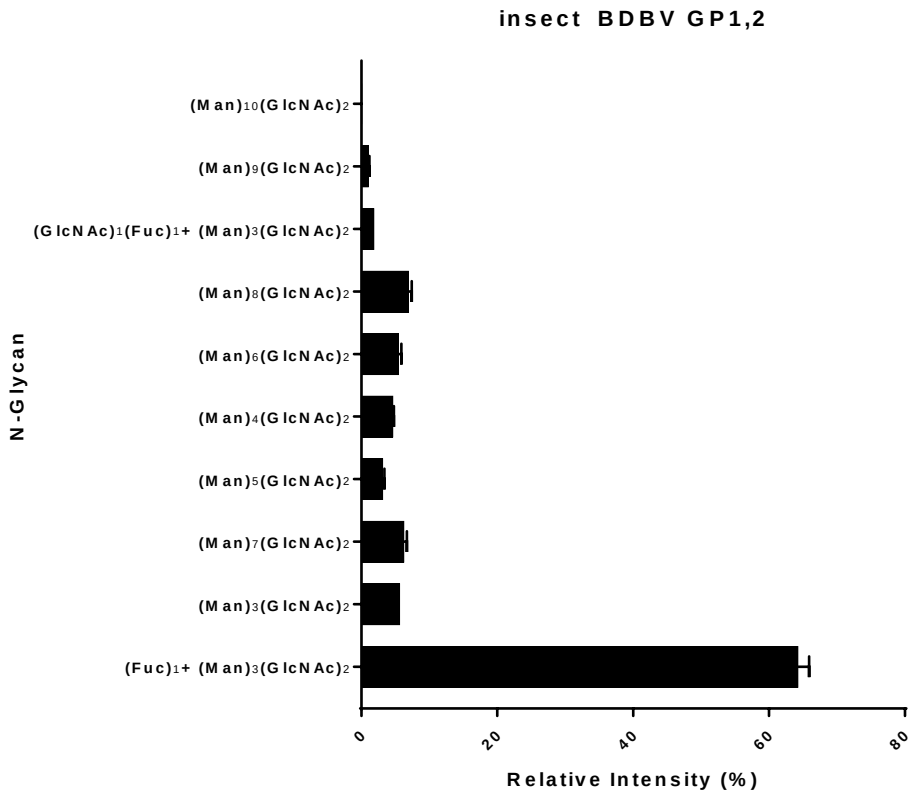


Figure 51. MALDI-TOF MS spectra of N-linked glycans from insect cell-derived BDBV GP1,2. Relative intensity (normalized to 100%) of 8 individual N-glycans are shown. Error bars represent standard error of 2 separate spots of the pooled GP1,2 from two separate experiments.

O-linked glycans were removed from GP1,2 beta-elimination for and analyzed by MALDI-TOF mass spectrometry. However, no O-linked glycans were detected in BDBV GP1,2.

N-glycan	Relative intenstiy	SD
$(Man)_3\ (GlcNAc)_2$	5.7	0.2
$(Fuc)_1 + (Man)_3\ (GlcNAc)_2$	64.3	1.6
$(Man)_4\ (GlcNAc)_2$	4.7	0.1
$(Man)_5(GlcNAc)_2$	3.2	0.2
$(GlcNAc)_1\ (Fuc)_1 + (Man)_3(GlcNAc)_2$	1.9	0.0
$(Man)_6(GlcNAc)_2$	5.5	0.4
$(Man)_7(GlcNAc)_2$	6.3	0.4
$(Man)_8(GlcNAc)_2$	7.0	0.4
$(Man)_9(GlcNAc)_2$	1.1	0.1
$(Man)_{10}(GlcNAc)_2$	0.1	0.0

Table 19. MALDI-TOF MS spectra of N-linked glycans from insect cell-derived BDBV GP1,2. Relative intensity (normalized to 100%) of 10 individual N-glycans are shown.

Glycan linkage and structure of BDBV GP1,2 produced in plant cells

To assess glycan patterns on plant-derived filovirus glycoprotein we analyzed BDBV GP1,2 due to the higher amount of protein available. BDBV GP1,2 was modified as in Figure 2, incorporated into *A. tumefaciens*, expressed in *N. benthamiana* plants, and purified using nickel columns. Two individual experiments were performed, and the purified proteins were combined and analyzed by MALDI-TOF mass spectrometry (Proteodynamics). The N-glycan profile was analyzed using two independent deglycosylation approaches, using either deglycosylation of the intact protein by both PNGase A plus F (Figure 52) or proteolytic digestion of the protein with trypsin followed by deglycosylation of the generated peptides by PNGase A (Figure 53). These figures showed MALDI-TOF MS spectra of permethylated N-glycans after deglycosylation of plant BDBV GP1,2 with PNGase A plus F or PNGase A alone. For each deglycosylation approach two N-glycan profiles were acquired on separate spots.

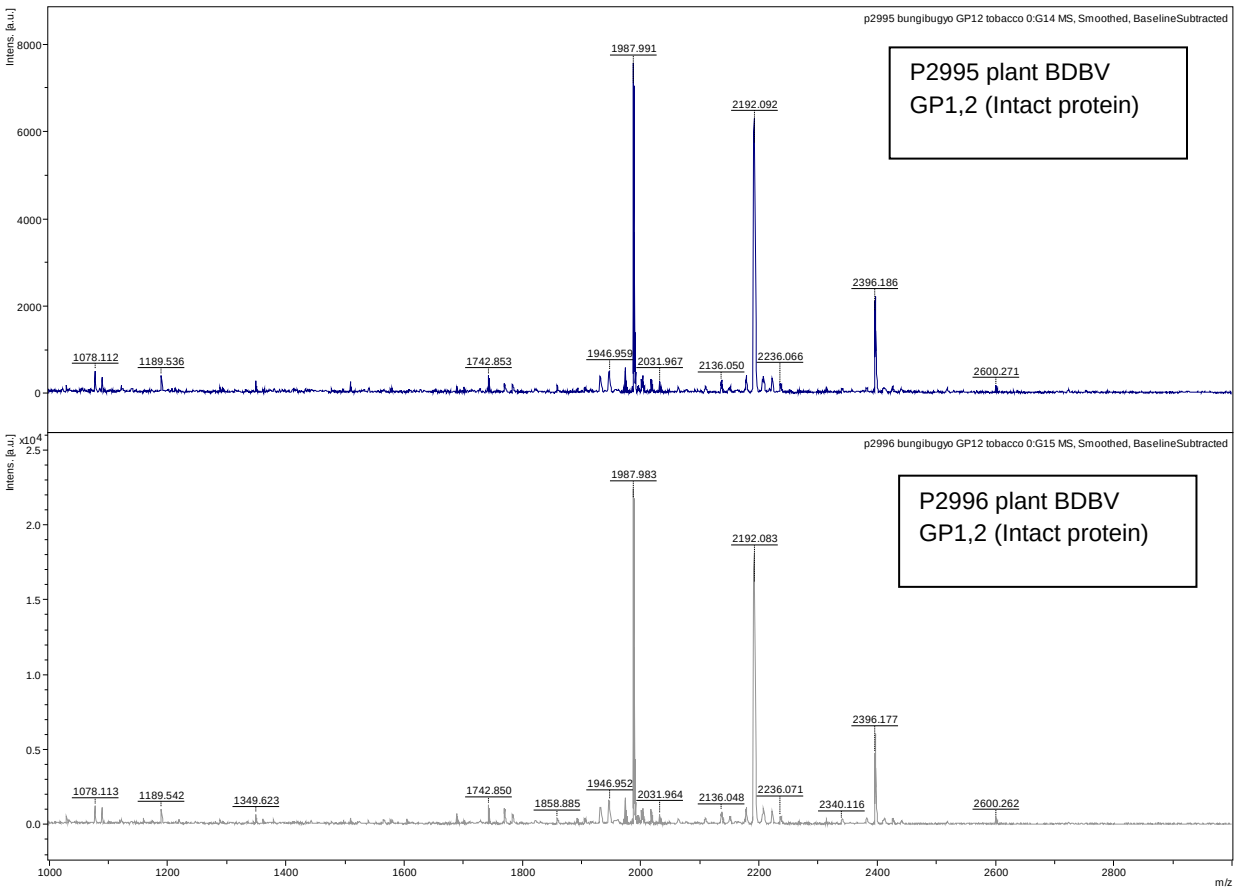


Figure 52: MALDI-TOF MS spectra of N-glycans after deglycosylation of intact protein from plant BDBV GP1,2 by PNGase A plus F.

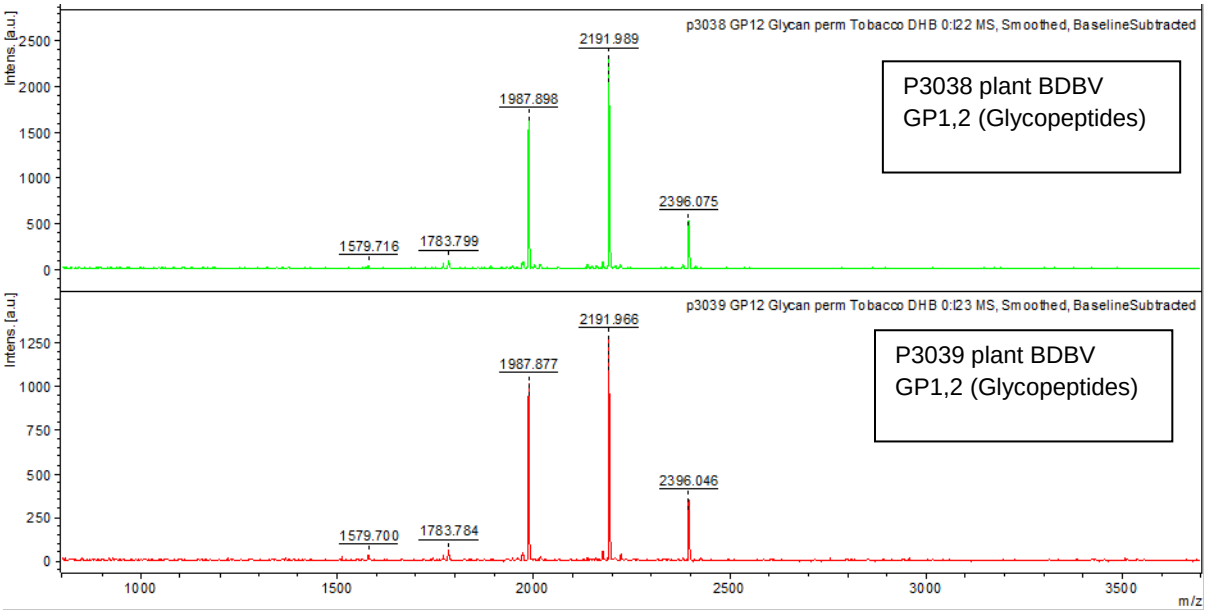


Figure 53: MALDI-TOF MS spectra of N-glycans after deglycosylation of tryptic peptides from plant BDBV GP1,2 by PNGase A.

A total of five individual N-glycan structures were assigned to ions within the different spectra. All of them correspond to high mannose type glycans. The semi-quantitative *N*-glycan profiles for the spectra as well as average values are represented in Tables 20 and 21.

Theoretical monoisotopic mass	plant BDBV GP1,2				Assignment/Interpretation
	Intact protein		Peptides		
[M+Na]+	rel. Intens. %	rel. Intens. %	rel. Intens. %	rel. Intens. %	
1579,8	0,0	0,7	1,2	1,6	$(Man)_5(GlcNAc)_2$
1783,9	1,3	1,5	2,5	2,4	$(Man)_6(GlcNAc)_2$
1988,0	48,4	49,9	37,8	39,1	$(Man)_7(GlcNAc)_2$
2192,1	37,6	37,0	48,2	45,5	$(Man)_8(GlcNAc)_2$
2396,2	12,6	10,9	10,4	11,4	$(Man)_9(GlcNAc)_2$
	100,0	100,0	100,0	100,0	

Table 20: Quantification of N-glycan structures

Theoretical monoisotopic mass	plant BDBV GP1,2				Assignment/Interpretation
	Intact protein		Peptides		
$[M+Na]^+$	Mean rel. Intens. %	standard deviation	Mean rel. Intens. %	standard deviation	
1579,8	0,3	0,5	1,4	0,3	$(Man)_5(GlcNAc)_2$
1783,9	1,4	0,1	2,5	0,1	$(Man)_6(GlcNAc)_2$
1988,0	49,2	1,1	38,5	1,0	$(Man)_7(GlcNAc)_2$
2192,1	37,3	0,5	46,8	1,9	$(Man)_8(GlcNAc)_2$
2396,2	12,6	1,2	10,9	0,7	$(Man)_9(GlcNAc)_2$

Table 21: Quantification of N-glycan structures. Average values and standard deviations (n=2) for each of the different experimental approaches

Average values calculated for the two independent experiments are presented in Figure 54.

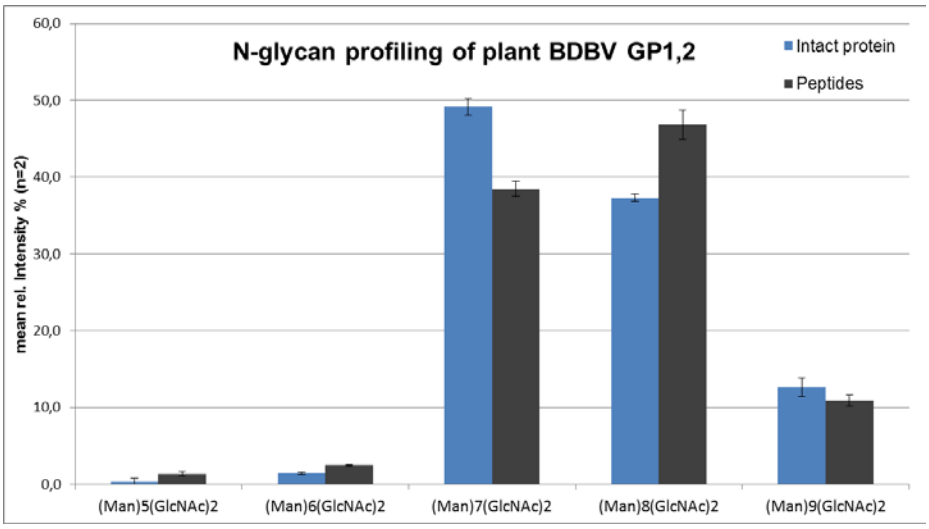


Figure 54: Relative Intensities (%) and standard deviations of *N*-glycan structures (n=2) for each of the different experimental approaches.

Conclusions on N-glycan profiling of plant-derived BDBV GP1,2. The obtained profiles comprising exclusively high mannose type glycans from $(Man)_5$ $(GlcNAc)_2$ to $(Man)_9$ $(GlcNAc)_2$. Surprisingly, there was no xylose or fucose residues found, which is unusual for plant-derived proteins. In both experimental approaches, no typical plant specific glycans containing Xylose and/or Fucose were present. No mass spectroscopy analysis was conducted on other filovirus GP1,2s other than BDBV due to low levels of protein produced in this approach.

Summary of N-linked glycan differences between filovirus GP1,2s expressed in mammalian 293T cells.

The N-linked glycan profiles for EBOV-Mayinga, EBOV-Makona, SUDV, BDBV, and SUDV were overall quite similar, with heavy glycosylation with a total of 47-56 glycan species. Figure 50 shows a direct comparison of all the N-linked glycans detected as a a grouping of the N-linked glycans according to individual residues, demonstrating the overall similarity between the GP1,2s studied. However, it is important to note a few differences, which are depicted in Figure 56. In comparison to profiles previously obtained for the different ebolavirus GP1,2 proteins we noted the following differences:

- The profile of 293T-produced MARV GP1,2 contains significantly less high mannose glycans than the previous ebolavirus samples
- The relative intensity of biantennary $(GlcNAc)_2$ structures is increased
- Sialic acid containing N-glycans are increased
- Fewer glycan species (37) were found in MARV GP1,2

It is surprising that sialic acid residues (NeuAc) were found at high levels, as previous reports found sialylation in ebolavirus GP1,2 but not MARV GP1,2 after expression in mammalian VeroE6 cells.

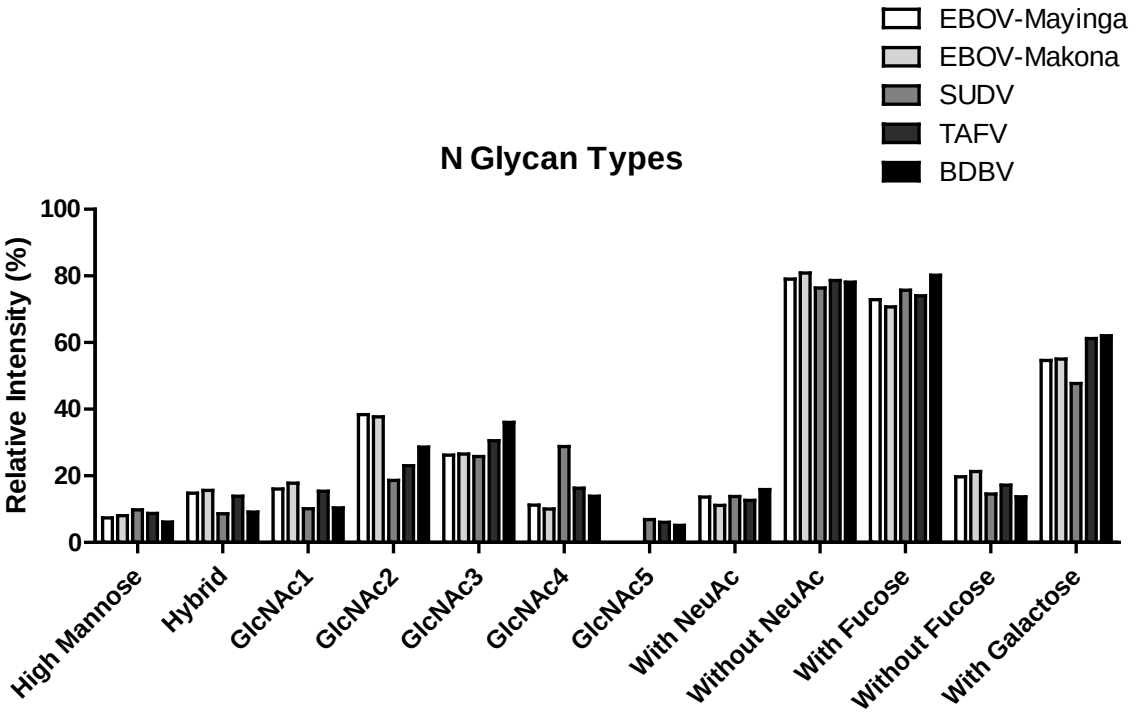


Figure 55. Comparison of N-linked glycan groups, showing overall similarity between GP1,2s.

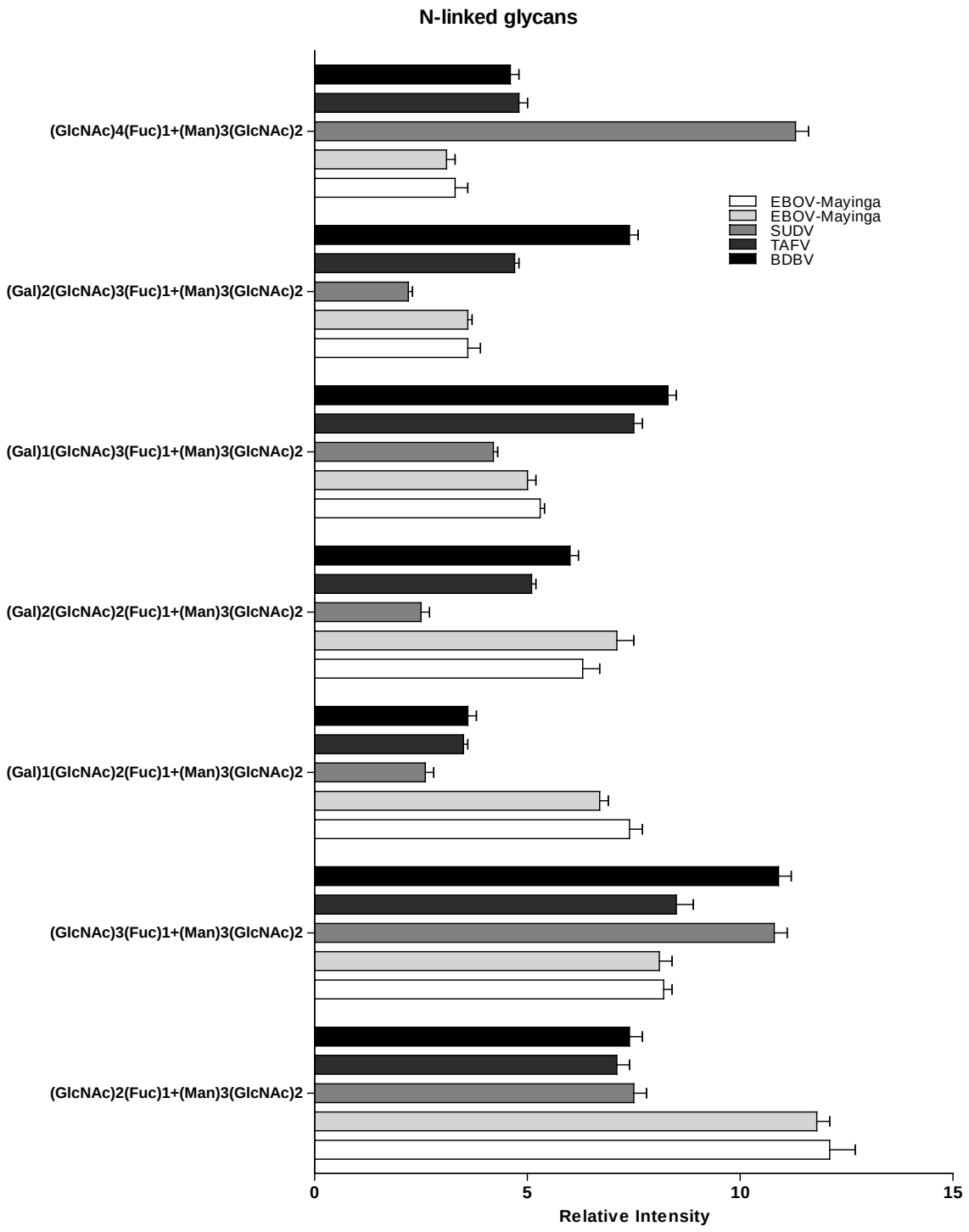


Figure 56. N-linked glycans with different frequency in GP1,2s.

Summary of O-linked glycan differences between filovirus GP1,2s expressed in mammalian 293T cells.

O-linked glycans were different between the GP1,2s of EBOV-Mayinga, EBOV-Makona, SUDV, TAFV, and BDBV. Figure 57 shows a comparison of the 8-10 O-linked glycans found in these proteins, and Figure 58 shows differences based on the presence, absence, or relative number of sialic acid residues. Another interesting finding was the difference in intensity of O-linked glycans between EBOV-Mayinga and EBOV-Makona, even though they are in the same viral species. MARV-Angola contained 5 O-glycans, with NeuAc1 and NeuAc2 species predominating.

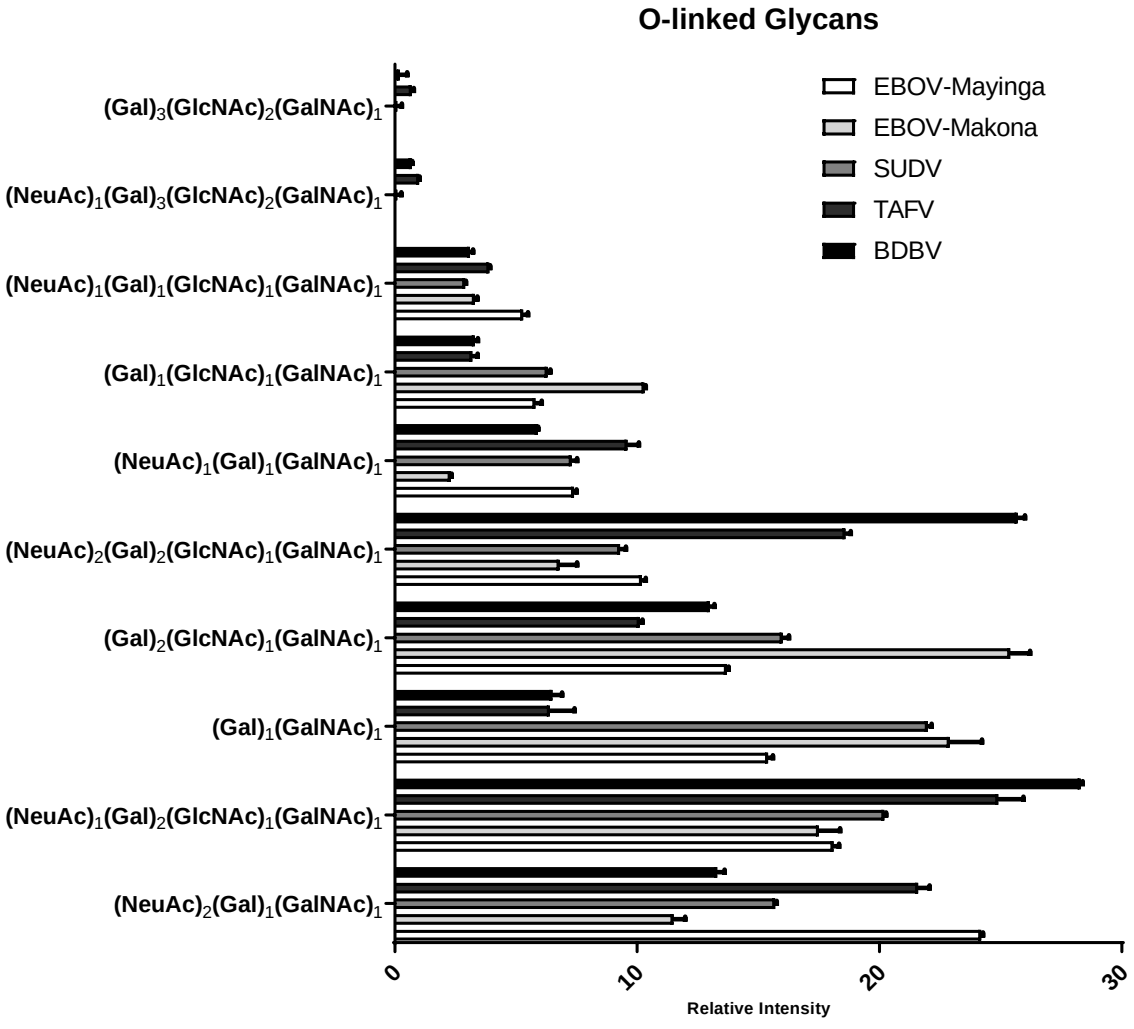


Figure 57. Comparison of O-linked glycans between GP1,2s expressed in mammalian cells.

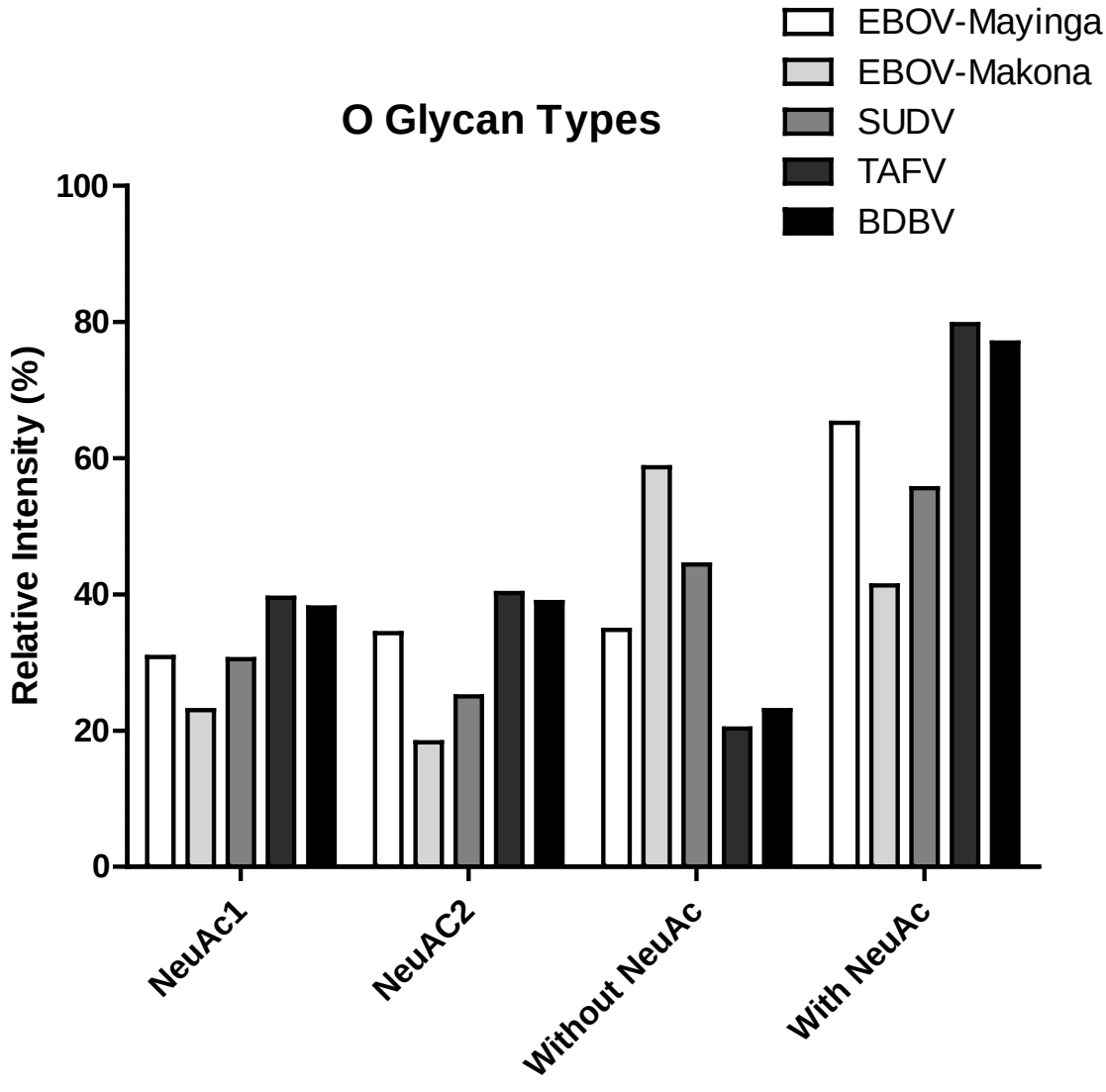


Figure 58. Comparison of O-linked glycan groups found in GP1,2s expressed in mammalian cells.

Summary of N- and O-linked glycan differences between filovirus GP1,2s expressed in insect Sf9 cells.

As opposed to GP1,2s expressed in mammalian cells, insect Sf9 cell-expressed GP1,2 from EBOV-Mayinga, SUDV, and TAFV had a single dominant N-linked glycan ($(Fuc)_1$ + $(Man)_3(GlcNAc)_2$) comprising ~75 of all N-glycans and no O-linked glycans. Mammalian cell-derived EBOV-Mayinga, SUDV, and TAFV contained ~50 N-linked glycans and 8-10 O-linked glycans. We were unable to complete the production and analysis of MARV-GP1,2 in insect cells.

Determination of the percentage of molecular weight contributed by glycans on filovirus GP1,2s made in mammalian, insect, and plant cells

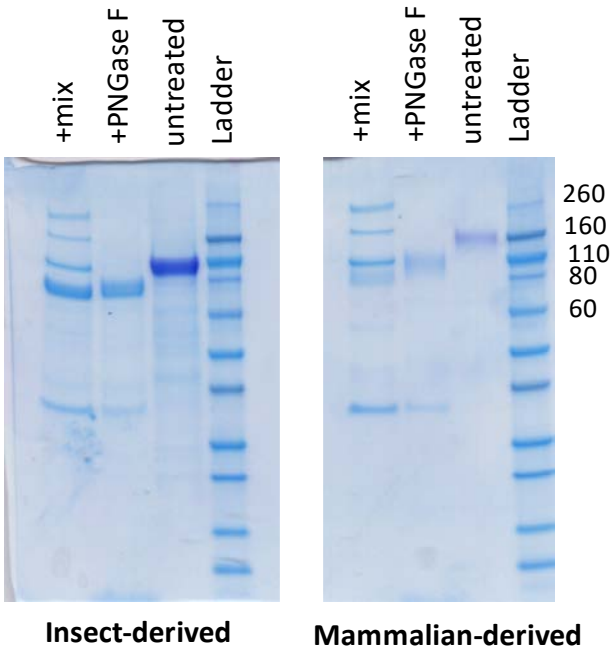


Figure 59. Quantitation of amount of N- and O-glycans on EBOV-Makona GP1,2 made in insect or mammalian cells. GP1,2 was incubated with enzymes that cleave N-glycans (PNGaseF) or both N- and O-glycans (mix) and molecular weight was compared with uncleaved protein by SDS-PAGE and colloidal blue staining. The extra bands in "mix" lanes are the glycosidases.

It is unknown how much glycosylation is present on different filovirus GP1,2s expressed in mammalian, insect, or plant cells. To address this gap in knowledge, we utilized enzymes that cleave N- or O-glycans. By measuring the molecular weight of the protein before and after glycan cleavage, we are able to estimate the contribution of N- and O-glycans to the molecular weight of the proteins.

We incubated filovirus GP1,2 made in mammalian, insect, or plant cells with PNGase F (which cleaves N-glycans), or PNGase F with O-glycosidases. These samples were separated by SDS-PAGE and analyzed by colloidal blue staining, PAS staining (for glycoproteins), or western blot, and molecular weights were compared to proteins that were not de-glycosylated.

First, EBOV-Makona GP1,2 from mammalian or insect cells was treated with de-glycosidases and analyzed by colloidal blue staining for molecular weight. Based on these experiments, we found that approximately 31% of mammalian-derived EBOV-Makona GP1,2 was comprised of N-glycans, whereas ~24% was due to O-glycans. For insect-derived protein, ~29% of the molecular weight was due to N-linked glycans, whereas no O-glycans were detected (Figure60). This is in agreement with our previous findings of only very simple and minimal GlcNAc residues on insect-derived GP1,2. A similar finding was observed for LLOV GP1,2; ~31% of the molecular weight of mammalian-derived protein was due to N-glycans and ~20% due to O-glycans; whereas ~29% of the molecular weight of insect-derived protein was comprised on N-glycans (Figure 59 and 60).

For plant-derived GP1,2 analysis, we used western blotting due to the low amounts of proteins present. As shown in Figure 61, approximately 33% of the molecular weight of the GP1,2s

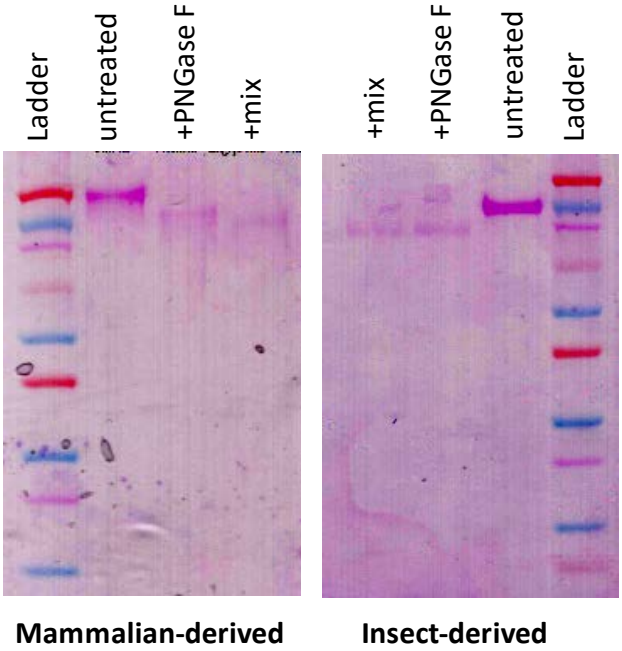


Figure 60. Quantitation of amount of N- and O-glycans on LLOV GP1,2 made in insect or mammalian cells. GP1,2 was incubated with enzymes that cleave N-glycans (PNGaseF) or both N- and O-glycans (mix) and molecular weight was compared with uncleaved protein by SDS-PAGE and PAS glycoprotein-specific staining.

analyzed was due to N-glycans. No O-glycans were detected, as expected, since O-glycosylation is limited in *N. benthamiana.*

In silico prediction of glycosylation patterns in filovirus GP1,2s

To ascertain the predicted differences in *N*- and *O*-linked glycosylation sites, $GP_{1,2}$ sequences were analyzed using two different prediction algorithms, NetOGlyc or GlycoEP. As shown in Figures 62 and 63, differences in the predictions of *N*- and *O*-linked glycosylation between the two algorithms were appreciable. However, the overall predicted *N*-glycosylation sites were similar between EBOV Yambuku and EBOV Makona, but different for *O*-glycosylation sites, which matches the glycosylation patterns found in the above-described data. The predicted *O*-glycosylation sites for SUDV $GP_{1,2}$ are considerably distinct from those of other ebolaviruses. Whether the differences in *O*-glycan profiles reported here are due to site-specific glycosylation variation or other factors is not known.

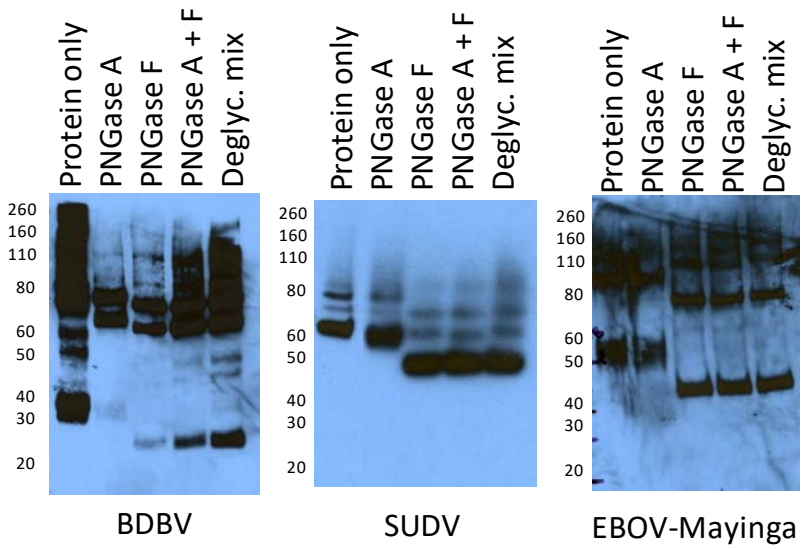


Figure 61. Quantitation of amount of N- and O-glycans on BDBV, SUDV, and EBOV-Mayinga GP1,2 made plants. Plant-derived GP1,2 was incubated with enzymes that cleave N-glycans (PNGaseF) or both N- and O-glycans (mix) and molecular weight was compared with uncleaved protein by SDS-PAGE and western blot.

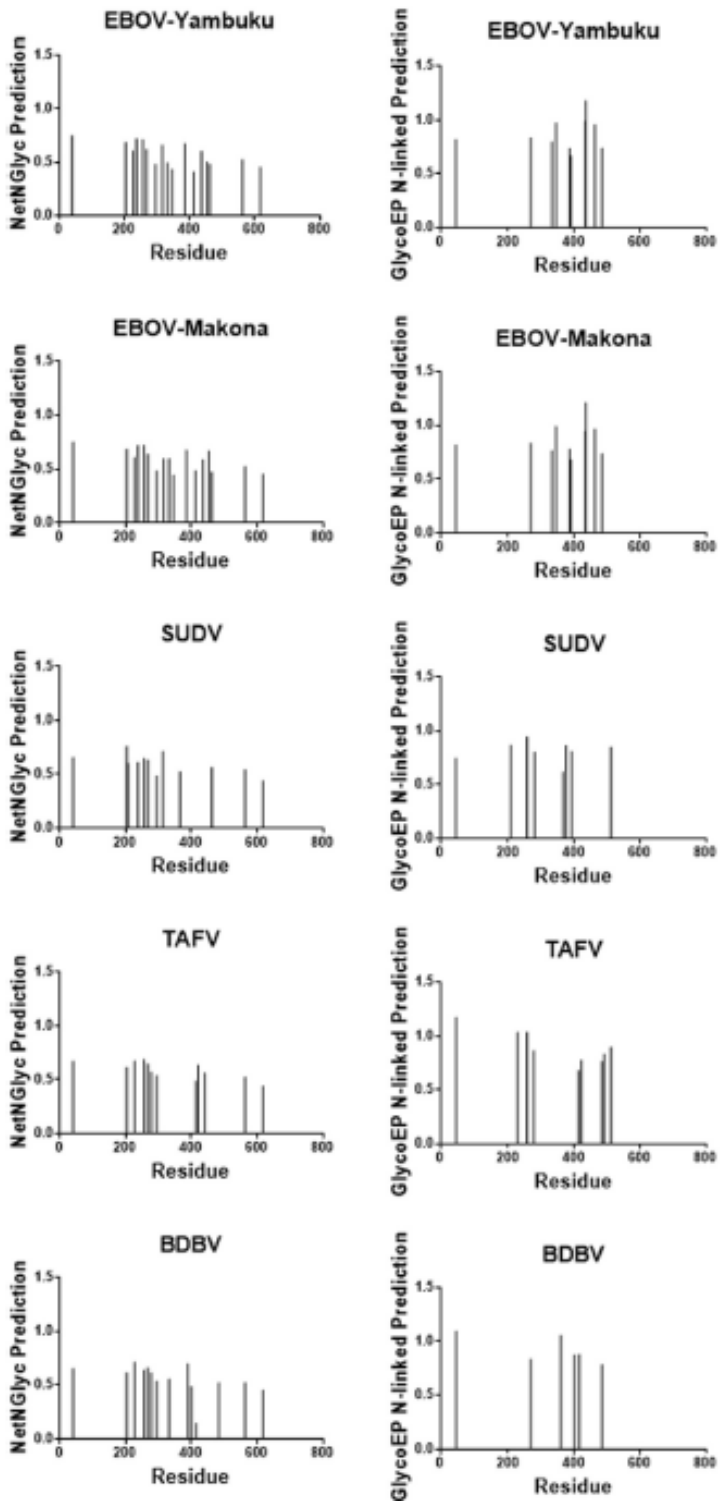


Figure 62. Comparison of predicted N-glycosylation sites found in ebolavirus GP1,2. N-glycan glycosylation site predictions were ascertained using NetNGlyc (left) or GlycoEP (right) algorithms.

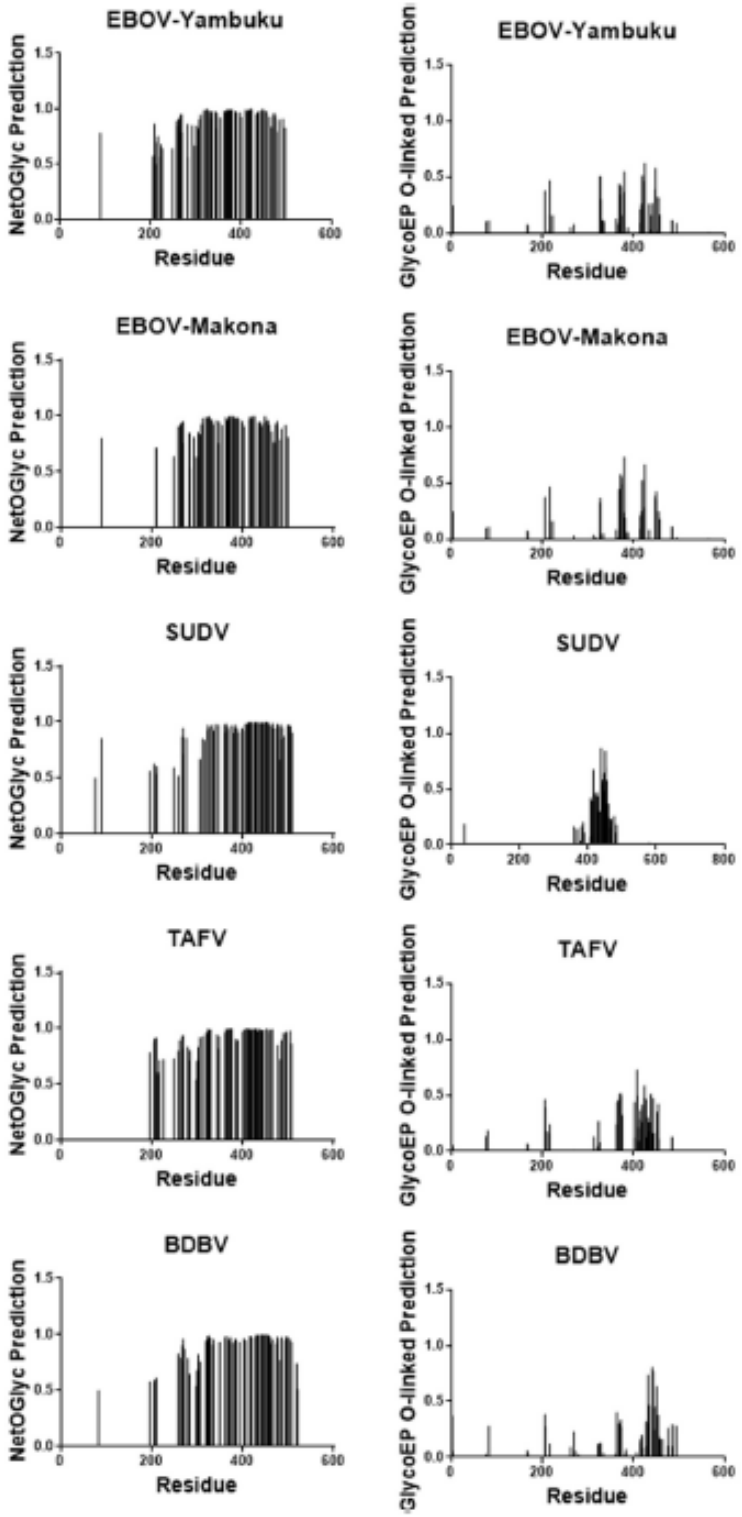


Figure 63. Comparison of predicted O-glycosylation sites found in ebolavirus GP1,2. O-glycan glycosylation site predictions were ascertained using NetOGlyc (left) or GlycoEP (right) algorithms.

e) Determination of glycan site occupation of EBOV GP1,2 produced in insect and mammalian cells

Having determined the theoretical site occupation of filovirus GP1,2, we set out to test how this compared with the actual site occupation of EBOV-Makona GP1,2 produced in mammalian or insect cells. First, the analyses of the fractions after deglycosylation by PNgase A +F were performed. Figure 64 represents MALDI-TOF MS spectra of deglycosylated glycopeptides.

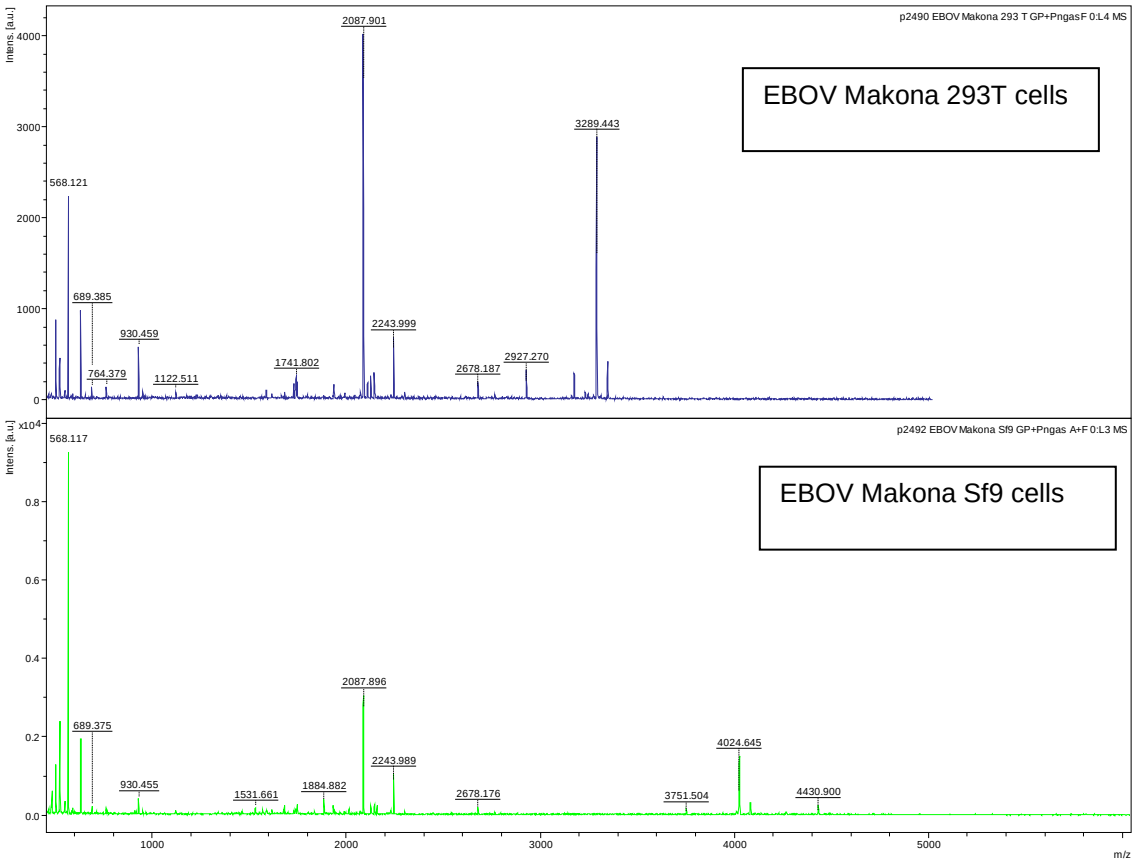


Figure 64: MALDI-TOF MS spectra of deglycosylated glycopeptides

Table 22 summarizes the results obtained.

N-glycosylation site	Position	MC	Sequence	Theo. $[M+H]^+$ Monoiso deglycosylated peptide mass	EBOV Makona 293T P2490	EBOV Makona Sf9 P2492
Asn 8	1 to 18	0	IPLGVIHNSTLQVSDVDK	1936,0	1936,0	1936,0
Asn 172	160-187	1	DFFSSHPLREPVNATEDPSS GYYSTTIR	3174,5	3174,4	nd
Asn 172	169-187	0	EPVNATEDPSSGYYSTTIR	2087,9	2087,9	2087,9
Asn 196/206	188-215	0	YQATGFGTNETEYLFEVDNL TYVQLESR	3289,5	3289,4	nd
Asn 225	216-233	0	FTPQFLLQLNETIYASGK	2071,1	nd	nd
Asn 236	234-240	1	RSNTTGK	764,4	764,4	764,4
Asn 264	263-267	1	KNLTR	632,4	632,4	632,4
Asn 264	264-267	0	NLTR	504,3	504,3	504,3
Asn 264	264-268	1	NLTRK	632,4	632,4	632,4
Asn 285	285-293	0	NISGQSPAR	930,5	930,5	930,5
Asn 301	294-308	0	TSSDPETNTTNEDHK	1676,7	1676,6	1676,6
Asn 314	309-326	0	IMASENSSAMVQVHSQGR	1932,9	1932,8	1932,9
Asn 354	350-363	0	TGPDNSTHNTPVYK	1531,7	nd	1531,7
Asn 381	378-399	1	RADNDSTASDTPPATTAAGP LK	2158,0	nd	2158,1
Asn 381	379-399	0	ADNDSTASDTPPATTAAGPL K	2001,9	nd	nd
Asn 404	400-407	0	AENTNTSK	865,4	nd	nd
Asn 422/430	408-446	0	SADSLDLATTTSPQNYSETA GNNNTHHQDTGEESASSGK	4024,7	nd	4024,7
Asn 531	528-542	0	QLANETTQALQLFLR	1746,9	1746,9	1746,9
Asn 586	586-590	0	NITDK	591,3	591,3	591,3

Table 22: Identified deglycosylated glycopeptides: MC= trypsin missed cleavage

The analysis of deglycosylated peptides reveals the occupation of several of the 17 theoretical N-glycosylation sites with N-glycans. It is important to note that deglycosylation by PNGase leads to a change of the Asn residue within the protein sequence to Asp. This induces a mass shift of +1 Da for the corresponding deglycosylated peptide as compared to the non-glycosylated sequence. The spectra contain in addition some ions corresponding to normal peptides of the protein that passed through the glycopeptide enrichment step. Deglycosylated peptides were searched assuming up to 3 trypsin miss cleavages. As summarized in the table, 11 out of the 17 potential glycosylation sites have been found occupied in this assay for EBOV Makona GP1,2 from 293T cells. Deglycosylation of EBOV Makona GP1,2 from Sf9 cells identifies 13 out of 17 potential N-glycosylation sites as occupied. Please note that this result is a first indication for glycosylation site occupation, since it may happen that a deglycosylated glycopeptide for a given site is not detected in this assay while the corresponding site can be found occupied with glycans when analyzing the spectra of the corresponding glycopeptide fraction.

Figure 65 represents spectra of the glycopeptide fractions in a linear positive mode.

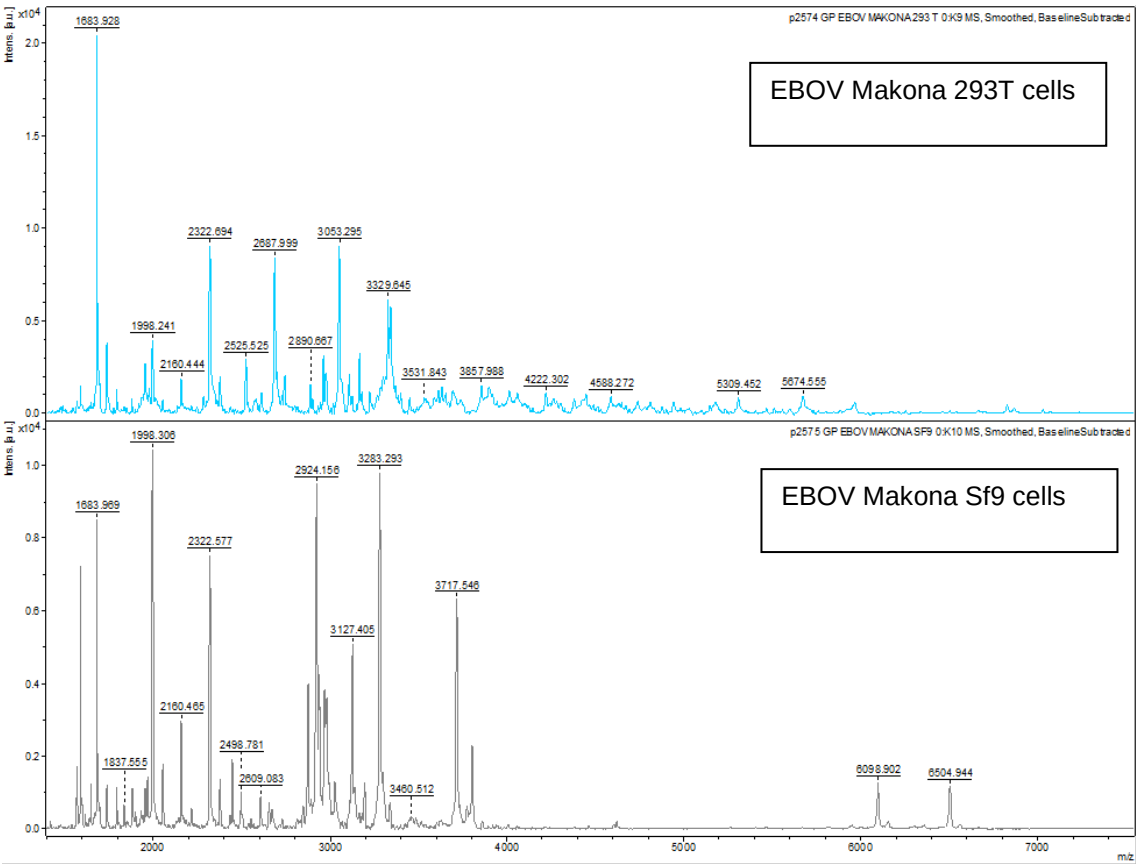


Figure 65: MALDI-TOF MS spectra of occupied glycopeptides

Spectra were analyzed to identify occupied N-glycopeptides and the glycan structures bound to the respective peptides. In addition spectra were examined to obtain a first indication on the presence of O-glycopeptides. The analysis is based on assigning measured masses to theoretical masses for combinations of N-glycans in the overall permethylated profile and glycopeptides harboring a given theoretical N-glycosylation site. Measured masses from mass lists obtained by MALDI-TOF analysis were matched to the theoretical masses of occupied glycopeptides. A mass difference of +/- 1.5 Da was used as selection criteria and mass peaks within this error range were assigned to theoretical glycopeptide masses. Manual inspection of data and spectra was then used to complete the analysis. In this manual inspection a mass tolerance of up to 3 Da was accepted for assigning glycopeptide masses especially for ions within the higher mass range. Resolution of average ion peaks in complex linear MALDI spectra may lie within this mass difference as compared to the theoretical value.

Seeing the complexity of EBOV Makona GP1.2 glycosylation there is a certain redundancy of theoretical masses for glycan/glycopeptide combinations. It is thus not always possible to clearly assign a mass to only one glycan/peptide combination. In the following tables redundant glycopeptide masses that cannot be clearly assigned to one glycopeptide/glycan combination are marked with an asterisk. The current analysis does not allow distinguishing between the different possible matches. In general all potential matches were therefore considered as positive in the result tables.

Minor glycan structures identified in the permethylated profiles may not be detected by this analysis due to the sensitivity of the assay. This is notably the case for sialic acid containing glycans from 293T cell derived GP1.2.

For several N-glycosylation sites N-glycan/glycopeptide combinations can be matched to ions in the spectra for glycopeptides that were not visible in their deglycosylated form after deglycosylation with PNGase F. In the tables these glycosylation sites are marked by #.

The result of this analysis shows that glycosylation of each site is heterogeneous which means that different N-glycans can be bound to the same glycosylation site on different individual GP1,2 proteins of one lot. We first present the determination of site specific N-glycosylation of EBOV Makona GP1.2 from Sf9 cells, since the global N-glycan profile is simpler and assignment of specific glycopeptide ions is less complex. The following Table 23 shows glycan structures bound to each of the theoretical N-glycosylation sites.

Asn 8	Glycan mass	theoretical glycopeptide mass	measured glycopeptide mass
N-linked glycan on peptide 1-28	[M] average	[M+H]+ average	[M+H]+ average
peptide ILPGVIHNSTLQVSDVDK		1936,2	
$(Fuc)_1 + (Man)_3 (GlcNAc)_2$	1057,0	2975,2	2975,2
Asn172	Glycan mass	theoretical glycopeptide mass	measured glycopeptide mass
N-linked glycan on peptide 169-187	[M] average	[M+H]+ average	[M+H]+ average
peptide EPVNATEDPSSGYYSTTIR		2088,2	
$(Man)_3 (GlcNAc)_2$	910,8	2981,0	2981,7
$(Fuc)_1 + (Man)_3 (GlcNAc)_2$	1057,0	3127,2	3127,4
$(GlcNAc)_1 (Fuc)_1 + (Man)_3 (GlcNAc)_2$	1260,2	3330,4	3329,5

Asn196/206 #	Glycan mass	theoretical glycopeptide mass	measured glycopeptide mass
N-linked glycan on peptide 188-215	[M] average	[M+H]+ average	[M+H]+ average
peptide YQATGFGTNETEYLFEVDNLTYVQLESR		3289,5	
no occupied peptides found			

Asn 225 #	Glycan mass	theoretical glycopeptide mass	measured glycopeptide mass
N-linked glycan on peptide 216-233	[M] average	[M+H]+ average	[M+H]+ average
peptide FTPQFLLQLNETIYASGK		2071,4	
$(Fuc)_1 + (Man)_3 (GlcNAc)_2$	1057,0	3110,4	3112,3
$(Man)_6 (GlcNAc)_2$	1397,3	3450,7	3450,4
$(Man)_8 (GlcNAc)_2$	1721,5	3774,9	3774,3

Asn 236 1MC	Glycan mass	theoretical glycopeptide mass	measured glycopeptide mass
N-linked glycan on peptide 234-240	[M] average	[M+H]+ average	[M+H]+ average
peptide RSNTTGK		763,8	
$(GlcNAc)_1 (Fuc)_1 + (Man)_3 (GlcNAc)_2$	1260,2	2006,0	2007,3

Asn 264 1MC	Glycan mass	theoretical glycopeptide mass	measured glycopeptide mass
N-linked glycan on peptide 263-267/264/268	[M] average	[M+H]+ average	[M+H]+ average
peptide KNLTR/NLTRK		631,8	
$(Fuc)_1 + (Man)_3 (GlcNAc)_2$	1057,0	1670,8	1670,3
$(Man)_6 (GlcNAc)_2$	1397,3	2011,1	2013,6
$(Man)_8 (GlcNAc)_2$	1721,5	2335,3	2334,5
$(Man)_9 (GlcNAc)_2$	1883,7	2497,5	2497,5

Asn 285	Glycan mass	theoretical glycopeptide mass	measured glycopeptide mass
N-linked glycan on peptide 285-293	[M] average	[M+H]+ average	[M+H]+ average
peptide NISGQSPAR		930,0	
$(Fuc)_1 + (Man)_3 (GlcNAc)_2$	1057,0	1969,0	1968,9

Asn 301	Glycan mass	theoretical glycopeptide mass	measured glycopeptide mass
N-linked glycan on peptide 294-308	[M] average	[M+H]+ average	[M+H]+ average
peptide TSSDPETNTTNEDHK		1676,6	
$(Man)_4 (GlcNAc)_2$	1073,0	2731,6	2729,2
$(Man)_5 (GlcNAc)_2$	1235,1	2893,7	2894,6

Asn 314	Glycan mass	theoretical glycopeptide mass	measured glycopeptide mass
N-linked glycan on peptide 309-326	[M] average	[M+H]+ average	[M+H]+ average
peptide IMASENSSAMVQVHSQGR		1933,2	
$(Fuc)_1 + (Man)_3 (GlcNAc)_2$	1057,0	2972,2	2972,2

Asn 354	Glycan mass	theoretical glycopeptide mass	measured glycopeptide mass
N-linked glycan on peptide 350-363	[M] average	[M+H]+ average	[M+H]+ average
peptide TGPDNSTHNTPVYK		1531,6	
$(Man)_6 (GlcNAc)_2$	1397,3	2910,9	2911,0

Asn 381 1 MC	Glycan mass	theoretical glycopeptide mass	measured glycopeptide mass
N-linked glycan on peptide 378-399	[M] average	[M+H]+ average	[M+H]+ average
peptide RADNDSTASDTPPATTAAGPLK		2158,3	
$(Fuc)_1 + (Man)_3 (GlcNAc)_2$	1057,0	3197,3	3197,3
$(Man)_7 (GlcNAc)_2$	1559,4	3699,7	3697,7
$(Man)_8 (GlcNAc)_2$	1721,5	3861,8	3862,5

Asn 404 #	Glycan mass	theoretical glycopeptide mass	measured glycopeptide mass
N-linked glycan on peptide 400-407	[M] average	[M+H]+ average	[M+H]+ average
peptide AENTNTSK		864,9	
$(Fuc)_1 + (Man)_3 (GlcNAc)_2$	1057,0	1903,9	1901,8

Asn 422/430	Glycan mass	theoretical glycopeptide mass	measured glycopeptide mass
N-linked glycan on peptide 408-446	[M] average	[M+H]+ average	[M+H]+ average
peptide SADSLDLATTTSPQNYSETAGNNNTHHQDTGEES		4025,0	
$2 \times (Fuc)_1 + (Man)_3 (GlcNAc)_2$	2 x 1057	6103,0	6099,0

Asn 531	Glycan mass	theoretical glycopeptide mass	measured glycopeptide mass
N-linked glycan on peptide 528-542	[M] average	[M+H]+ average	[M+H]+ average
peptide QLANETTQALQLFLR		1747,0	
$(Man)_8\ (GlcNAc)_2$	1721,5	3451,4	3451,9
$(Man)_{10}\ (GlcNAc)_2$	2045,8	3775,7	3774,2

Asn 586	Glycan mass	theoretical glycopeptide mass	measured glycopeptide mass
N-linked glycan on peptide 586-590	[M] average	[M+H]+ average	[M+H]+ average
peptide NITDK		590,7	
$(Man)_7\ (GlcNAc)_2$	1559,4	2132,1	2132,5

Table 23: Site specific N-glycan profile assignments of EBOV Makona GP1,2 from Sf9 cells
#: glycosylation sites whose corresponding deglycosylated glycopeptide was not visible in MALDI spectra after PNGase F digestion.

The dominant N-glycan structure in the overall permethylated N-glycan profile is $(Fuc)_1$ + $(Man)_3\ (GlcNAc)_2$. This structure is observed on most but not all of the different N-glycosylation sites. Notably glycosylation sites Asn 531 and Asn 586 appear to be occupied by high mannose type glycans. For the glycopeptide harboring both glycosylation sites Asn422/430 we have assigned an intense ion at m/z 6099 to the presence of two $(Fuc)_1$ + $(Man)_3\ (GlcNAc)_2$ residues bound to the peptide, even if the mass difference is superior to 3 Da. We could not assign any glycan structures bound to peptide 188-215 harboring N-glycosylation sites Asn196/206. This peptide was also not visible in its deglycosylated form. It may be possible that the different forms of this glycopeptide were not detectable in this assay due to the important mass of the expected ions.

The spectra of glycopeptides from Sf9 cell derived EBOV Makona GP1.2 contains several ions that cannot be assigned to occupied N-glycopeptides. Some of these ions correspond to peptides that have been co-purified with the glycopeptide fraction. Others may indicate the presence of O-glycopeptides.

The following Table 24 shows glycan structures bound to each of the theoretical N-glycosylation sites of EBOV Makona GP1.2 from 293T cells. Redundant ion masses that can be assigned to more than one glycan/glycopeptide combination are marked by an asterisk*.

Asn 8	Glycan mass	theoretical glycopeptide mass	measured glycopeptide mass
N-linked glycan on peptide 1-28	[M] average	[M+H]+ average	[M+H]+ average
peptide ILPGVIHNSTLQVSDVDK		1936,2	
$(GlcNAc)_2$ $(Fuc)_1$ + $(Man)_3(GlcNAc)_2$	1463,4	3381,6	3381,3
$(Gal)_1$ $(GlcNAc)_2$ + $(Man)_3(GlcNAc)_2$	1479,4	3397,6	3399,0
$(Gal)_3$ $(GlcNAc)_3$ $(Fuc)_1$ + $(Man)_3(GlcNAc)_2$	2153,0	4071,2	4071,8
$(Gal)_4$ $(GlcNAc)_4$ $(Fuc)_1$ + $(Man)_3(GlcNAc)_2$	2518,3	4436,5	4438,0

Asn172	Glycan mass	theoretical glycopeptide mass	measured glycopeptide mass
N-linked glycan on peptide 169-187	[M] average	[M+H]+ average	[M+H]+ average
peptide EPVNATEDPSSGYYSTTIR		2088,2	
$(GlcNAc)_1$ + $(Man)_3(GlcNAc)_2$	1114,0	3184,2	3181,6
$(GlcNAc)_2$ + $(Man)_3(GlcNAc)_2$	1317,2	3387,4	3387,1
$(GlcNAc)_2$ $(Fuc)_1$ + $(Man)_3(GlcNAc)_2$	1463,4	3533,6	3533,6*
$(Man)_7(GlcNAc)_2$	1559,4	3629,6	3631,5
$(Gal)_1$ $(GlcNAc)_1$ $(Fuc)_1$ + $(Man)_4(GlcNAc)_2$	1584,5	3654,7	3654,7
$(Gal)_1$ $(GlcNAc)_2$ $(Fuc)_1$ + $(Man)_3(GlcNAc)_2$	1625,5	3695,7	3694,8*
$(Gal)_2$ $(GlcNAc)_2$ $(Fuc)_1$ + $(Man)_3(GlcNAc)_2$	1787,7	3857,9	3857,9
$(Gal)_1$ $(GlcNAc)_3$ $(Fuc)_1$ + $(Man)_3(GlcNAc)_2$	1828,7	3898,9	3898,8
$(Gal)_2$ $(GlcNAc)_3$ + $(Man)_3(GlcNAc)_2$	1844,7	3914,9	3913,6
$(Gal)_2$ $GlcNAc)_3$ $(Fuc)_1$ + $(Man)_3(GlcNAc)_2$	1990,8	4061,0	4060,8
$(Gal)_3$ $(GlcNAc)_3$ $(Fuc)_1$ + $(Man)_3(GlcNAc)_2$	2153,0	4223,2	4222,3
$(Gal)_2$ $(GlcNAc)_4$ $(Fuc)_1$ + $(Man)_3(GlcNAc)_2$	2194,0	4264,2	4264,9
$(Gal)_3$ $(GlcNAc)_4$ $(Fuc)_1$ + $(Man)_3(GlcNAc)_2$	2356,2	4426,4	4426,2
$(Gal)_4$ $(GlcNAc)_4$ $(Fuc)_1$ + $(Man)_3(GlcNAc)_2$	2518,3	4588,5	4588,3

Asn196/206 #	Glycan mass	theoretical glycopeptide mass	measured glycopeptide mass
N-linked glycan on peptide 188-215	[M] average	[M+H]+ average	[M+H]+ average
peptide YQATGFGTNETEYLFEVDNLTYVQLESR		3289,5	
no occupied peptides found			

Asn 225 #	Glycan mass	theoretical glycopeptide mass	measured glycopeptide mass
N-linked glycan on peptide 216-233	[M] average	[M+H]+ average	[M+H]+ average
peptide FTPQFLLQLNETIYASGK		2071,4	
$(GlcNAc)_1 + (Man)_3(GlcNAc)_2$	1114,0	3167,4	3167,8*
$(Man)_5(GlcNAc)_2$	1235,1	3288,5	3287,2*
$(GlcNAc)_1$ $(Fuc)_1 + (Man)_3(GlcNAc)_2$	1260,2	3313,6	3313,6*
$(GlcNAc)_2 + (Man)_3(GlcNAc)_2$	1317,2	3370,6	3369,7*
$(Man)_6(GlcNAc)_2$	1397,3	3450,7	3451,2*
$(Gal)_1$ $(GlcNAc)_2 + (Man)_3(GlcNAc)_2$	1479,4	3532,8	3533,6*
$(Man)_7(GlcNAc)_2$	1559,4	3612,8	3613,4*
$(Gal)_1$ $(GlcNAc)_1 + (Man)_5(GlcNAc)_2$	1600,5	3653,9	3653,1
$(Gal)_2$ $(GlcNAc)_2 + (Man)_3(GlcNAc)_2$	1641,5	3694,9	3694,8*
$(Gal)_2$ $(GlcNAc)_3 + (Man)_3(GlcNAc)_2$	1844,7	3898,1	3898,1

Asn 236 1MC	Glycan mass	theoretical glycopeptide mass	measured glycopeptide mass
N-linked glycan on peptide 234-240	[M] average	[M+H]+ average	[M+H]+ average
peptide RSNTTGK		763,8	
$(GlcNAc)_1 + (Man)_4(GlcNAc)_2$	1276,2	2022,0	2020,8
$(Gal)_1$ $(GlcNAc)_1$ $(Fuc)_1 + (Man)_4(GlcNAc)_2$	1584,5	2330,3	2332,5
$(Gal)_1$ $(GlcNAc)_1 + (Man)_5(GlcNAc)_2$	1600,5	2346,3	2343,9
$(GlcNAc)_4$ $(Fuc)_1 + (Man)_3(GlcNAc)_2$	1869,8	2615,6	2613,6

Asn 264 1MC	Glycan mass	theoretical glycopeptide mass	measured glycopeptide mass
N-linked glycan on peptide 263-267/264/268	[M] average	[M+H]+ average	[M+H]+ average
peptide KNLTR/NLTRK		631,8	
$(Man)_6(GlcNAc)_2$	1397,3	2011,1	2013,4
$(Man)_8(GlcNAc)_2$	1721,5	2335,3	2323,5

Asn 285	Glycan mass	theoretical glycopeptide mass	measured glycopeptide mass
N-linked glycan on peptide 285-293	[M] average	[M+H]+ average	[M+H]+ average
peptide NISGQSPAR		930	
$(GlcNAc)_1$ $(Fuc)_1$ + $(Man)_4(GlcNAc)_2$	1422,3	2334,3	2334,1
$(GlcNAc)_3$ $(Fuc)_1$ + $(Man)_3(GlcNAc)_2$	1666,6	2578,6	2577,1
$(Gal)_2$ $(GlcNAc)_2$ $(Fuc)_1$ + $(Man)_3(GlcNAc)_2$	1787,7	2699,7	2699,4
$(Gal)_2$ $GlcNAc)_3$ $(Fuc)_1$ + $(Man)_3(GlcNAc)_2$	1990,8	2902,8	2903,5
$(Gal)_3$ $(GlcNAc)_3$ $(Fuc)_1$ + $(Man)_3(GlcNAc)_2$	2153,0	3065,0	3063,7
$(Gal)_2$ $(GlcNAc)_4$ $(Fuc)_1$ + $(Man)_3(GlcNAc)_2$	2194,0	3106,0	3107,9
$(Gal)_3$ $(GlcNAc)_4$ $(Fuc)_1$ + $(Man)_3(GlcNAc)_2$	2356,2	3268,2	3269,5

Asn 301	Glycan mass	theoretical glycopeptide mass	measured glycopeptide mass
N-linked glycan on peptide 294-308	[M] average	[M+H]+ average	[M+H]+ average
peptide TSSDPETNTTNEDHK		1676,6	
$(Man)_4(GlcNAc)_2$	1073,0	2731,6	2728,8
$(Man)_6(GlcNAc)_2$	1397,3	3055,9	3053,3
$(GlcNAc)_3$ + $(Man)_3(GlcNAc)_2$	1520,4	3179,0	3180,1*
$(Gal)_1$ $(GlcNAc)_2$ $(Fuc)_1$ + $(Man)_3(GlcNAc)_2$	1625,5	3284,1	3284,2
$(Gal)_2$ $GlcNAc)_3$ $(Fuc)_1$ + $(Man)_3(GlcNAc)_2$	1990,8	3649,4	3650,3*
$(Gal)_1$ $(GlcNAc)_4$ $(Fuc)_1$ + $(Man)_3(GlcNAc)_2$	2031,9	3690,5	3691,9
$(Gal)_2$ $(GlcNAc)_4$ $(Fuc)_1$ + $(Man)_3(GlcNAc)_2$	2194,0	3852,6	3850,0
$(Gal)_3$ $(GlcNAc)_4$ $(Fuc)_1$ + $(Man)_3(GlcNAc)_2$	2356,2	4014,8	4013,7

Asn 314	Glycan mass	theoretical glycopeptide mass	measured glycopeptide mass
N-linked glycan on peptide 309-326	[M] average	[M+H]+ average	[M+H]+ average
peptide IMASENSSAMVQVHSQGR		1933,2	
$(GlcNAc)_1$ $(Fuc)_1$ + $(Man)_3(GlcNAc)_2$	1260,2	3175,4	3177,1
$(Man)_6(GlcNAc)_2$	1397,3	3312,5	3311,6
$(Gal)_2$ $(GlcNAc)_2$ $(Fuc)_1$ + $(Man)_3(GlcNAc)_2$	1787,7	3702,9	3701,5
$(Gal)_1$ $(GlcNAc)_3$ $(Fuc)_1$ + $(Man)_3(GlcNAc)_2$	1828,7	3743,9	3742,4
$(Gal)_3$ $(GlcNAc)_3$ $(Fuc)_1$ + $(Man)_3(GlcNAc)_2$	2153,0	4068,2	4065,5
$(Gal)_4$ $(GlcNAc)_4$ $(Fuc)_1$ + $(Man)_3(GlcNAc)_2$	2518,3	4433,5	4434,5

Asn 354 #	Glycan mass	theoretical glycopeptide mass	measured glycopeptide mass
N-linked glycan on peptide 350-363	[M] average	[M+H]+ average	[M+H]+ average
peptide TGPDNSTHNTPVYK		1531,6	
$(GlcNAc)_2$ $(Fuc)_1$ + $(Man)_3(GlcNAc)_2$	1463,4	2977,0	2977,9
$(GlcNAc)_3$ + $(Man)_3(GlcNAc)_2$	1520,4	3034,0	3036,5
$(GlcNAc)_3$ $(Fuc)_1$ + $(Man)_3(GlcNAc)_2$	1666,6	3180,2	3180,1*
$(Gal)_1$ $(GlcNAc)_1$ $(Fuc)_1$ + $(Man)_5(GlcNAc)_2$	1746,6	3260,2	3262,3
$(Gal)_2$ $(GlcNAc)_2$ $(Fuc)_1$ + $(Man)_3(GlcNAc)_2$	1787,7	3301,3	3302,1*
$(Gal)_1$ $(GlcNAc)_3$ $(Fuc)_1$ + $(Man)_3(GlcNAc)_2$	1828,7	3342,3	3343,6
$(GlcNAc)_4$ $(Fuc)_1$ + $(Man)_3(GlcNAc)_2$	1869,8	3383,4	3385,4
$(NeuAc)_1$ $(Gal)_1$ $(GlcNAc)_3$ $(Fuc)_1$ + $(Man)_3(GlcNAc)_2$	2120,0	3633,6	3633,9

Asn 381 #	Glycan mass	theoretical glycopeptide mass	measured glycopeptide mass
N-linked glycan on peptide 379-399	[M] average	[M+H]+ average	[M+H]+ average
peptide ADNDSTASDTPPATTAAGPLK		2002,1	
$(GlcNAc)_1$ + $(Man)_4(GlcNAc)_2$	1276,2	3260,3	3262,4
$(GlcNAc)_2$ + $(Man)_3(GlcNAc)_2$	1317,2	3301,3	3302,1
$(Gal)_1$ $(GlcNAc)_2$ $(Fuc)_1$ + $(Man)_3(GlcNAc)_2$	1625,5	3609,6	3609,2
$(GlcNAc)_3$ $(Fuc)_1$ + $(Man)_3(GlcNAc)_2$	1666,6	3650,7	3650,3*
$(Gal)_1$ $(GlcNAc)_4$ $(Fuc)_1$ + $(Man)_3(GlcNAc)_2$	2031,9	4016,0	4016,8

Asn 404 #	Glycan mass	theoretical glycopeptide mass	measured glycopeptide mass
N-linked glycan on peptide 400-407	[M] average	[M+H]+ average	[M+H]+ average
peptide AENTNTSK		864,9	
$(GlcNAc)_1$ + $(Man)_5(GlcNAc)_2$	1438,3	2285,2	2285,7
$(GlcNAc)_2$ $(Fuc)_1$ + $(Man)_3(GlcNAc)_2$	1463,4	2310,3	2311,3
$(NeuAc)_1$ $(Gal)_3$ $(GlcNAc)_3$ $(Fuc)_1$ + $(Man)_3(GlcNAc)_2$	2444,3	3291,2	3289,6
$(Gal)_4$ $(GlcNAc)_4$ $(Fuc)_1$ + $(Man)_3(GlcNAc)_2$	2518,3	3365,2	3364,9

Asn 422/430 #	Glycan mass	theoretical glycopeptide mass	measured glycopeptide mass
N-linked glycan on peptide 408-446	[M] average	[M+H]+ average	[M+H]+ average
peptide SADSLDLATTTSPQNYSETAGNNNTHHQDTGEESASSGK		4025,0	
no occupied peptides found			

Asn 531	Glycan mass	theoretical glycopeptide mass	measured glycopeptide mass
N-linked glycan on peptide 528-542	[M] average	[M+H]+ average	[M+H]+ average
peptide QLANETTQALQLFLR		1747	
$(Man)_5(GlcNAc)_2$	1235,1	2964,1	2964,5
$(Man)_6(GlcNAc)_2$	1397,3	3126,3	3126,1
$(GlcNAc)_1 + (Man)_5(GlcNAc)_2$	1438,3	3167,3	3167,8*
$(Man)_7(GlcNAc)_2$	1559,4	3288,4	3287,2*
$(Gal)_1 (GlcNAc)_1 (Fuc)_1 + (Man)_4(GlcNAc)_2$	1584,5	3313,5	3313,6*
$(Gal)_1 (GlcNAc)_1 + (Man)_5(GlcNAc)_2$	1600,5	3329,5	3329,6
$(Gal)_2 (GlcNAc)_2 + (Man)_3(GlcNAc)_2$	1641,5	3370,5	3369,7*
$(Man)_8(GlcNAc)_2$	1721,5	3450,5	3451,2*
$(Gal)_1 (GlcNAc)_4 + (Man)_3(GlcNAc)_2$	1885,8	3614,8	3613,4*
$(NeuAc)_1 (Gal)_1 (GlcNAc)_2 (Fuc)_1 + (Man)_3(GlcNAc)_2$	1919,8	3648,8	3647,9

Asn 586	Glycan mass	theoretical glycopeptide mass	measured glycopeptide mass
N-linked glycan on peptide 586-590	[M] average	[M+H]+ average	[M+H]+ average
peptide 586-590 NITDK		590,7	
$(GlcNAc)_1 + (Man)_3(GlcNAc)_2$	1114,0	1686,7	1683,9
$(Man)_6(GlcNAc)_2$	1397,3	1970,0	1969,2

Table 24: Site specific N-glycan profile assignments of EBOV Makona GP1.2 from 293T cells

#: glycosylation sites whose corresponding deglycosylated glycopeptide was not visible in MALDI spectra after PNGase F digestion.

*: redundant ion mass that can be assigned to more than 1 glycan/glycopeptide combination

Site specific N-glycosylation of EBOV Makona GP1,2 from 293T cells reflects the complexity of glycan structures already described for the overall permethylated N-glycan profile. We were able to identify only few glycopeptides with sialic acid containing glycans bound. This is due to the low abundance of sialic acid in N-glycans already seen in the overall permethylated profile. For glycopeptides harboring two glycosylation sites Asn 196/206 on peptide188-215 and Asn422/430 on peptide 408-446 we could not assign any glycan structures. These peptides were also not visible in their deglycosylated forms. It may be possible that the occupied forms of these glycopeptides were not detectable in this assay due to the important mass of the expected ions. The spectra of glycopeptides from 293T cell derived EBOV Makona GP1,2 contains several ions that cannot be assigned to occupied N-glycopeptides. Some of these ions correspond to peptides that have been co-purified with the glycopeptide fraction. Others may indicate the presence of O-glycopeptides (see below).

Preliminary information on site specific O-glycosylation of EBOV Makona GP1.2

The exact positions of GP1,2 O-glycosylation sites could not be identified. Nevertheless the spectra of the glycopeptide fraction contain a series of ions that could represent occupied O-glycopeptides. For EBOV Makona GP1.2 from 293T cells the ions series m/z 1592, 1795, 1998, 2160, 2323, 2526, 2688, 2891 and 3053 is present in spectrum P2574. The mass differences between the listed ions correspond either to 162 Da representing one hexose residue or to 203 Da corresponding to HexNAc. The spectrum of GP1.2 from Sf9 cells also contains some of these ions. To confirm whether the listed ions actually correspond to O-glycopeptides further studies will be necessary.

Summary of site occupation. As shown in Figure 66, 13 of 17 predicted N-linked sites were occupied in mammalian cell-derived GP1,2, and 15 of 17 sites were occupied in insect cell-derived GP1,2. Each site was occupied by more than one glycan, suggesting heterogenous glycan occupation of each site. The two predicted N-linked sites in GP2 were occupied, confirming the importance of these sites for GP1,2 expression and function. O-glycan site occupation was attempted, but was too complex to assess, due to the large number of predicted sites and the unknown accuracy of the prediction algorithms.

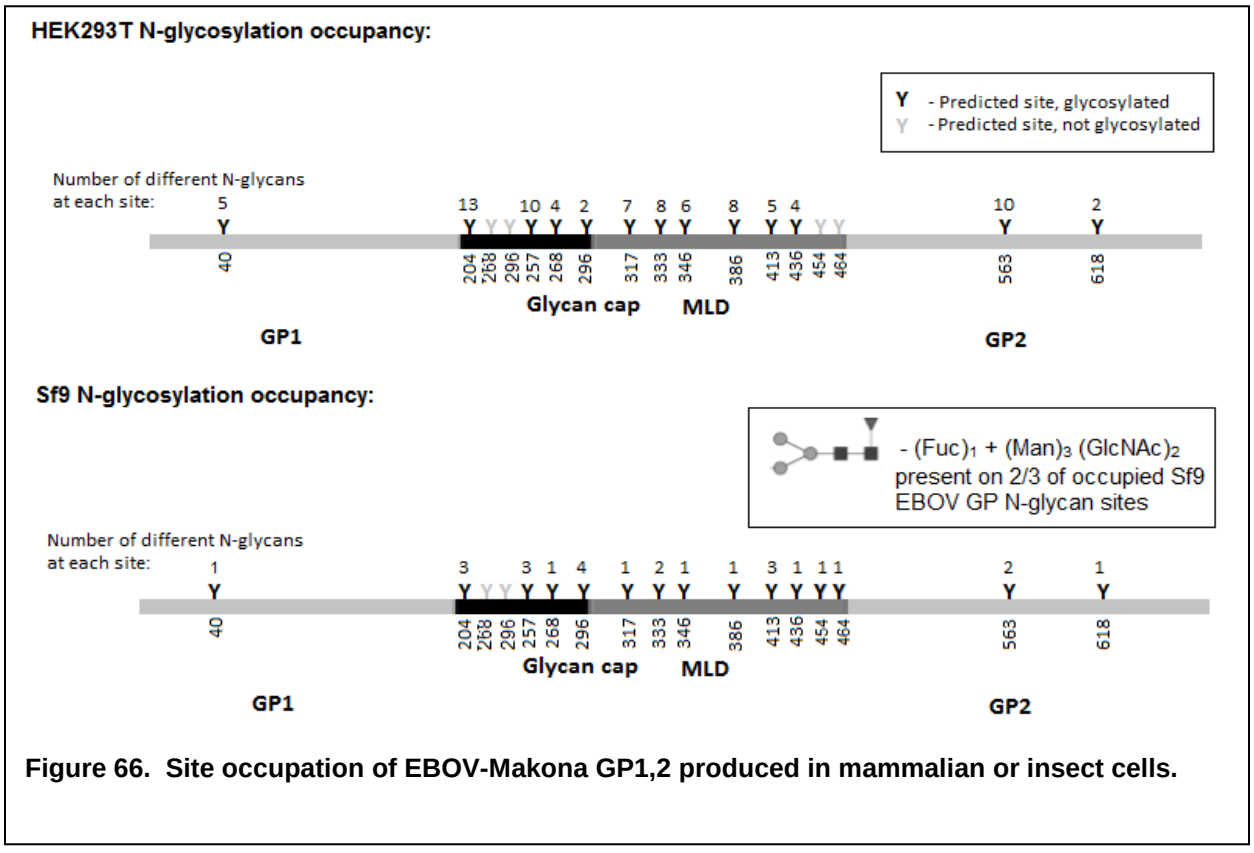


Figure 66. Site occupation of EBOV-Makona GP1,2 produced in mammalian or insect cells.

Key Outcomes and other achievements

a) Global analysis and comparison of GP1,2 N-linked glycosylation patterns on EBOV-Mayinga, EBOV-Makona, SUDV, BDBV, and TAFV produced in mammalian cells. Prior to this work, only EBOV-Mayinga GP1 had been thoroughly investigated for glycosylation patterns. We now report that overall N-linked glycosylation of ebolavirus GP1,2 are similar, with ~50 different glycan structures, largely composed of hybrid glycans. There are some relatively minor differences between N-linked glycan amounts and structures between the ebolaviruses.

b) Global analysis and comparison of GP1,2 O-linked glycosylation patterns on EBOV-Mayinga, EBOV-Makona, SUDV, BDBV, and TAFV produced in mammalian cells. We found substantial O-linked glycan differences between ebolavirus GP1,2. These differences were in both the relative amounts and types of O-glycan structures. O-linked glycan analysis has not previously been reported for ebolavirus GP1,2s.

c) Differences in O-linked glycans between EBOV-Mayinga and EBOV-Makona GP1,2. Although the Mayinga and Makona strains of EBOV are within the same species, we found substantial differences in levels of individual O-linked glycan levels between GP1,2 of these strains. Whether this affects the pathogenesis or immunogenicity of the virus GP1,2s awaits further analysis.

d) Differences in N –and O-linked glycosylation of insect cell- versus mammalian cell-derived GP1,2. Insect cell-derived GP1,2 from EBOV-Mayinga, SUDV, and TAFV contained only 8 N-linked glycans, and only 1 major N-glycan, accounting for ~75% of the all N-glycans, compared to mammalian cell-derived GP1,2 which had a range of ~50 N-glycans, with no species accounting for more that ~11% of the total N-glycans. Surprisingly, we found no O-linked glycans in insect-cell derived GP1,2, even though these cells are known to possess O-linked glycan machinery.

CHAPTER 3. DETERMINING THE EFFECTS OF DIFFERENTIAL GLYCOSYLATION OF FILOVIRUS GP1,2 ON ACTIVATION OF DCS IN VITRO

Having characterized glycosylation patterns in mammalian, insect, and plant-derived filovirus GP1,2, the next step of the project was to determine if these different glycosylation patterns affected immunogenicity. To accomplish this task, we utilized an *in vitro* immunogenicity assay with steady-state dendritic cell subsets and filovirus-like particles generated in mammalian or insect cells. Due to the difficulty in producing plant-derived GP1,2, we chose to concentrate our efforts on mammalian and insect cell-derived GP1,2 in the virus-like particle format.

Generation of virus-like particles expressing filovirus GP1,2.

Choice of VP40. First, we had to decide how to generate virus-like particles (VLPs) that expressed different filovirus GP1,2. We could use a single VP40 gene as the structural protein, or we could match each GP1,2 to its own VP40. We chose to use a single VP40 gene (EBOV-Yambuku) for co-expression with the different filovirus GP1,2s, in order to reduce any variabilities in VP40 expression or composition. The idea behind this is that this would reduce the number of variables (VP40 expression levels, potential differences between different VP40s and their association with GP1,2s, potentially different scaffolding between different VP40s). It is also possible that a homologous VP40 might interact more strongly with its GP1,2. However, we felt that keeping the same VP40 would simplify analysis of *in vitro* and *in vivo* immunogenicity.

Generation of mammalian cell- and insect cell-derived VLPs. Mammalian cell-derived VLPs were generated by transfecting HEK-293T cells with EBOV-Mayinga VP40 in a pcDNA3.1 vector (gene synthesis by Genscript) and a filovirus GP1,2 (wild-type sequence, in pcDNA3.1, synthesized by Genscript) using JetPrime transfection reagent. In contrast to transmembrane-

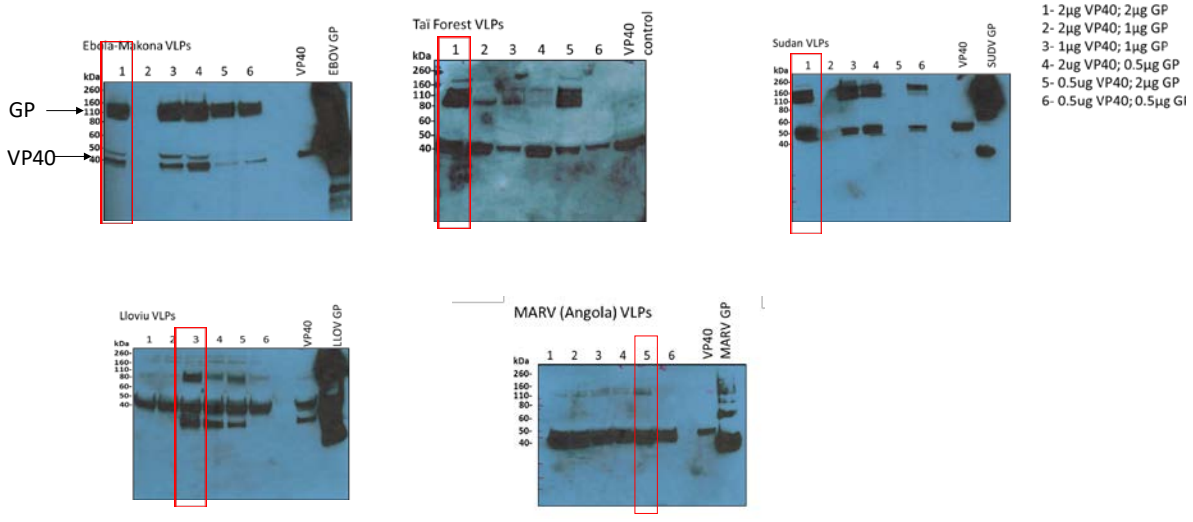


Figure 67. VLP generation and the effects of VP40:GP1,2 ratio. 293T cells were transfected with different ratios of EBOV-Yambuku VP40 and filovirus GP1,2.

deleted GP1,2 MARV and RESTV, we did not have any expression problems with any full-length, unmodified filovirus GP1,2.

We determined optimal VP40:GP1,2 ratios for generation of VLP, and found that different filovirus GP1,2s required different ratios for optimal expression. A VP40:GP1,2 ratio of 1:1 (with 2 ug of each) was optimal for generation of EBOV-Makona, EBOV-Mayinga, TAFV, SUDV, and BDBV VLPs, while a 1:1 ratio with half the plasmid amount was better for LLOV VLPs, and a 1:4 VP40:GP1,2 ratio was best for MARV-Angola VLPs. Therefore, these ratios were used to generate mammalian VLPs (Figure 67). Supernatants from transfected 293T cells were centrifuged at 10,000 rpm, and pellets were ultracentrifuged through a stepwise (60/30/10) sucrose gradient overnight. Bands were collected and ultracentrifuged, and pellets were resuspended in Phosphate buffered saline (PBS). Endotoxin was quantitated, and only preps that had negligible levels of endotoxin were used. Early experiments found that water used in the buffer and sucrose solutions contaminated the VLPs, but this was resolved by using purchased endotoxin-free water.

For insect cell-derived VLPs, the VP40 and GP1,2s were cloned into baculovirus vectors, and clones were transfected into Sf9 cells. The presence of virus carrying VP40 or filovirus GP1,2 was confirmed by western blotting of supernatants from transfected cells. Positive supernatants were used to infect Sf9 cells, and were plaque purified. Baculovirus stocks were generated and

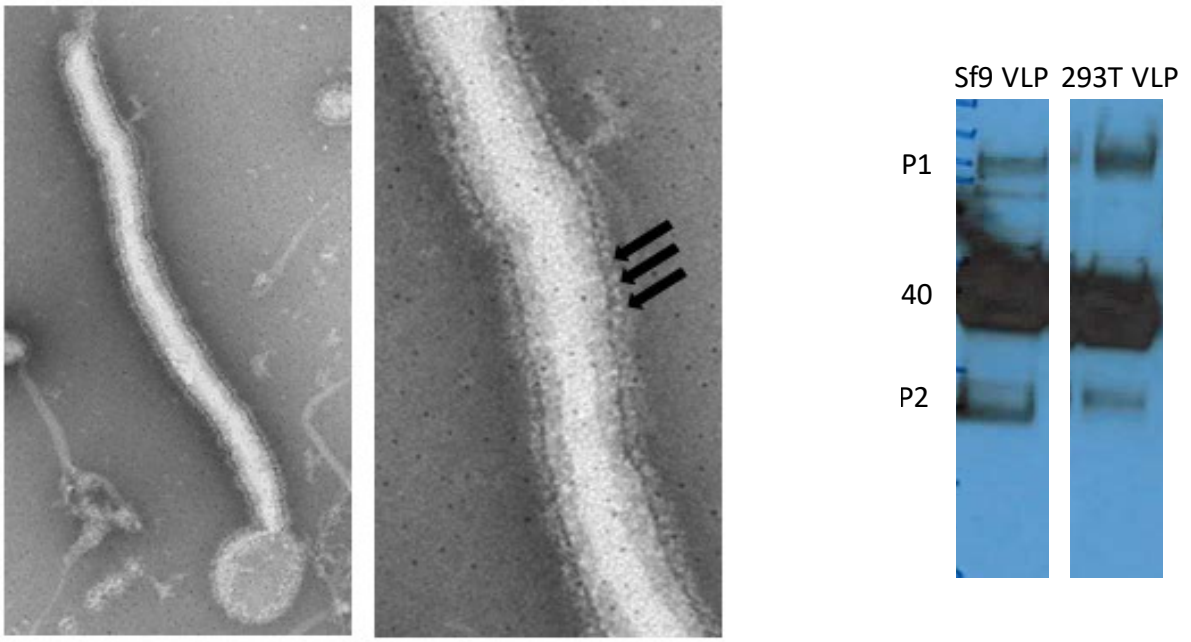


Figure 68. Generation of VLPs. Electron micrograph of mammalian-derived EBOV-Makona (left) and western blots for VP40 and GP1/GP2 for insect- and mammalian cell-derived VLPs (right).

titered. Sf9 cells were infected at a 2:5 VP40:GP1,2 ratio, and supernatants were collected and processed in the same manner as mammalian-derived VLPs.

Therefore, both insect and mammalian cell-derived VLPs were generated (Figure 68).

Effects of differential filovirus GP1,2 glycosylation on activation of dendritic cells *in vitro*.

Generation of myeloid and plasmacytoid dendritic cells. Mouse bone marrow cells were incubated at 1 million cells/mL with 100 ng/mL Flt3L for 8 days. Cells were harvested, and stained with antibodies against CD11c, B220, CD24, MHC II, CD40, and CD86. We found that >80% of cells were CD11c positive. Plasmacytoid DCs were identified as CD11c+/B220+, CD8a+ DCs as CD11c+/B220-/CD24+, and CD8a- cells as CD11c+/B220-/CD24- (Figure 69).

Analysis of VLP levels to induce DC activation. To test appropriate levels of VLPs to use to stimulate DCs *in vitro*, different amounts of VLPs were incubated with DCs (1, 10, and 100 ug/mL) for 24 and 48 hours. As shown in Figure 70, stimulation as analyzed by increased CD86 and MHC II levels was observed at 1 and 10 ug/mL. Interestingly, the 100 ug/mL concentration eliminated the activation seen at 1 and 10 ug/mL. Therefore, further experiments were conducted at 10 ug/mL, to increase the sensitivity of the assay, at 24 hours (since 48 hours did not show any major change in activation—data not shown).

Insect cell-derived versus mammalian cell-derived VLP activation of DCs. DCs were incubated with EBOV-Makona, EBOV-Yambuku, SUDV, BDBV, TAFV, or MARV VLPs generated in mammalian or insect cells at 10 ug/mL for 24 hours. Cells were stained and analyzed by flow cytometry. As shown in Figures 71-75-71, insect cell-derived VLPs were more

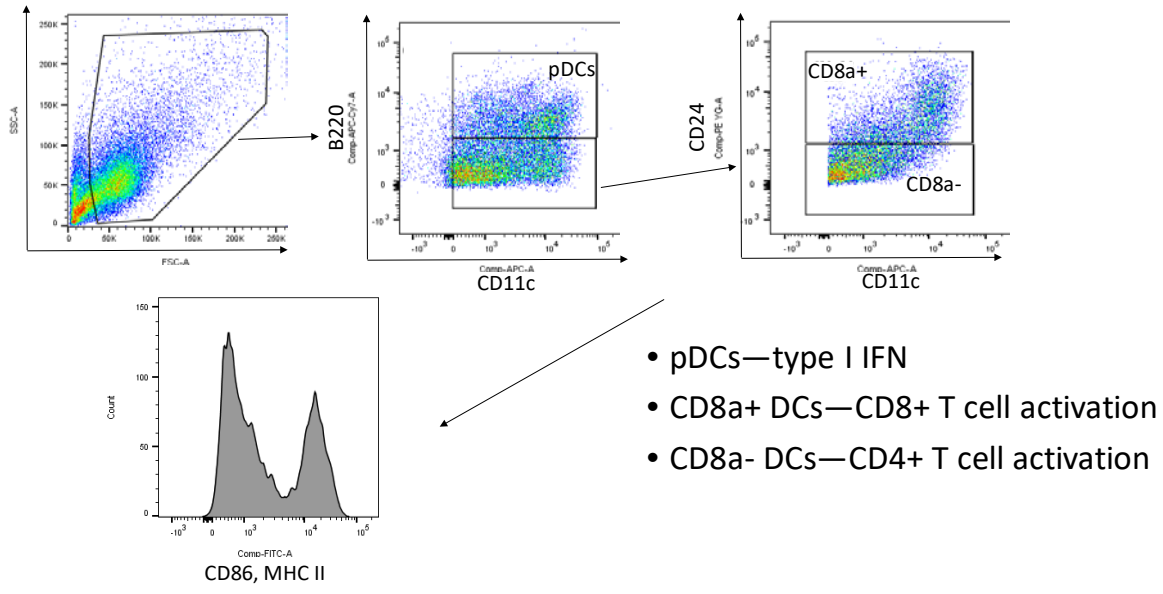


Figure 69. Generation and analysis of mouse dendritic cell subsets *in vitro*.

immunostimulatory CD8a+, CD8a-, and pDCs as measured by the activation markers CD86 and CD40. MHC II levels were less elevated.

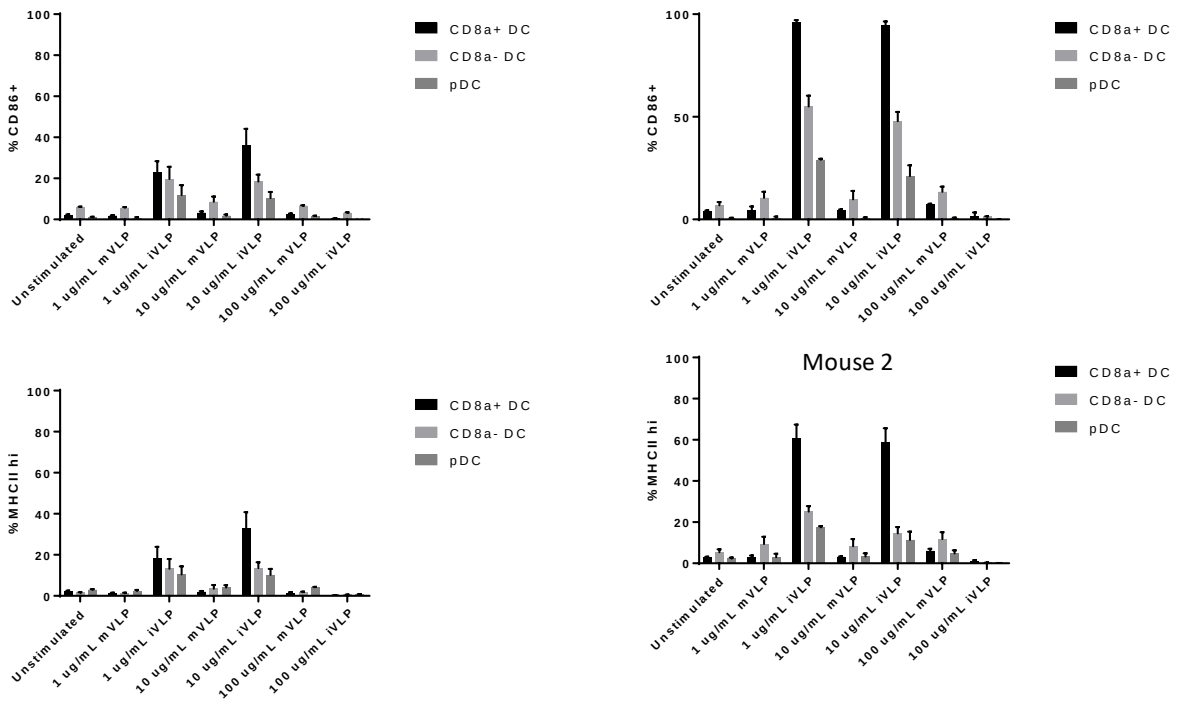


Figure 70. Stimulation of *in vitro*-derived DCs by VLPs. EBOV-Makona VLPs were incubated with DCs at 1, 10, and 100 ug/mL for two days. DCs from two mice are shown. Experiments were performed in triplicate.

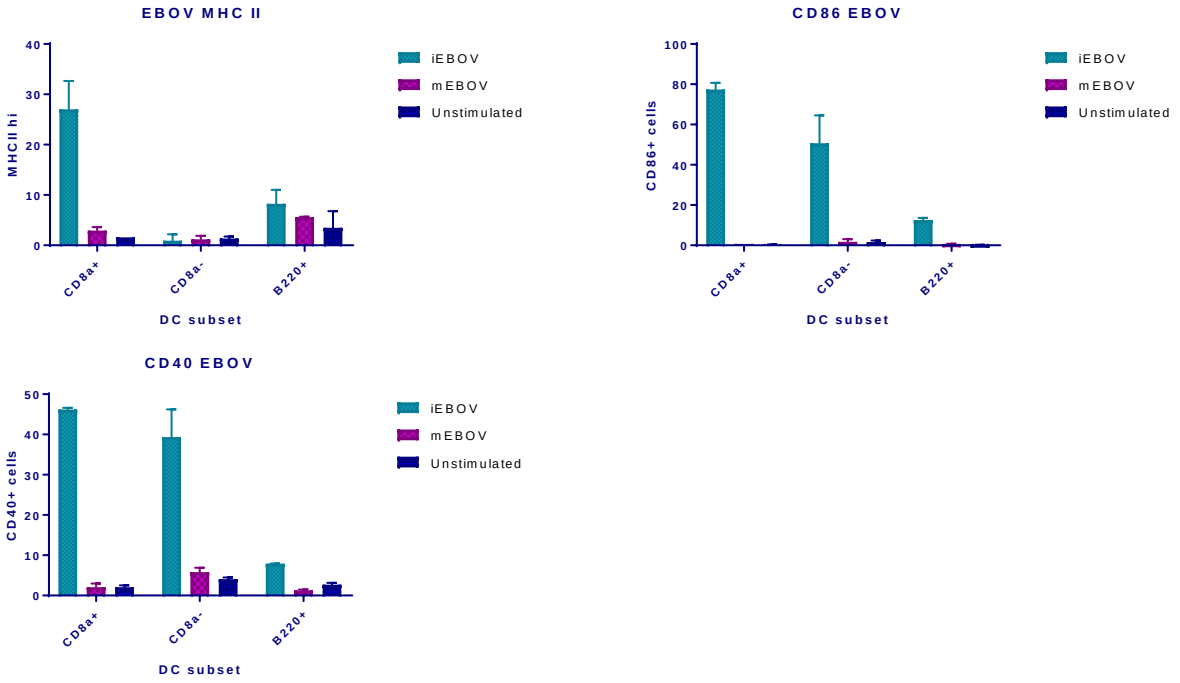


Figure 71. Insect cell-derived EBOV-Makona VLPs are more stimulatory to DC subsets than mammalian cell-derived VLPs.

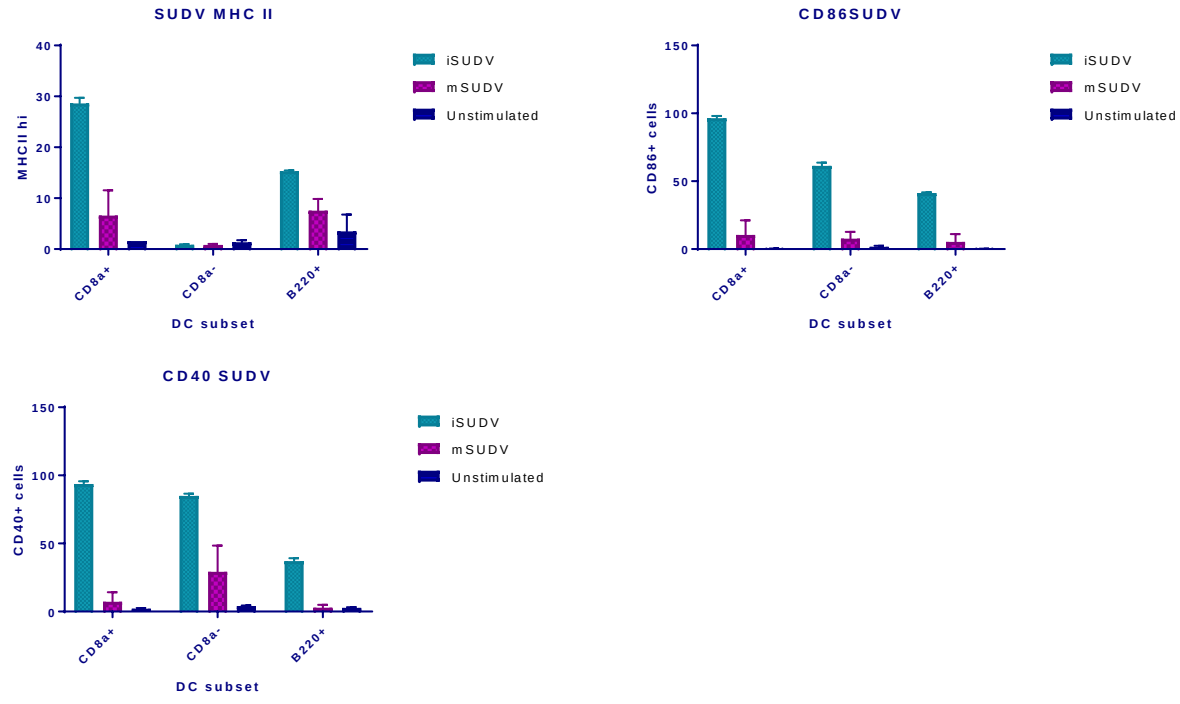


Figure 72. Insect cell-derived SUDV VLPs are more stimulatory to DC subsets than mammalian cell-derived VLPs.

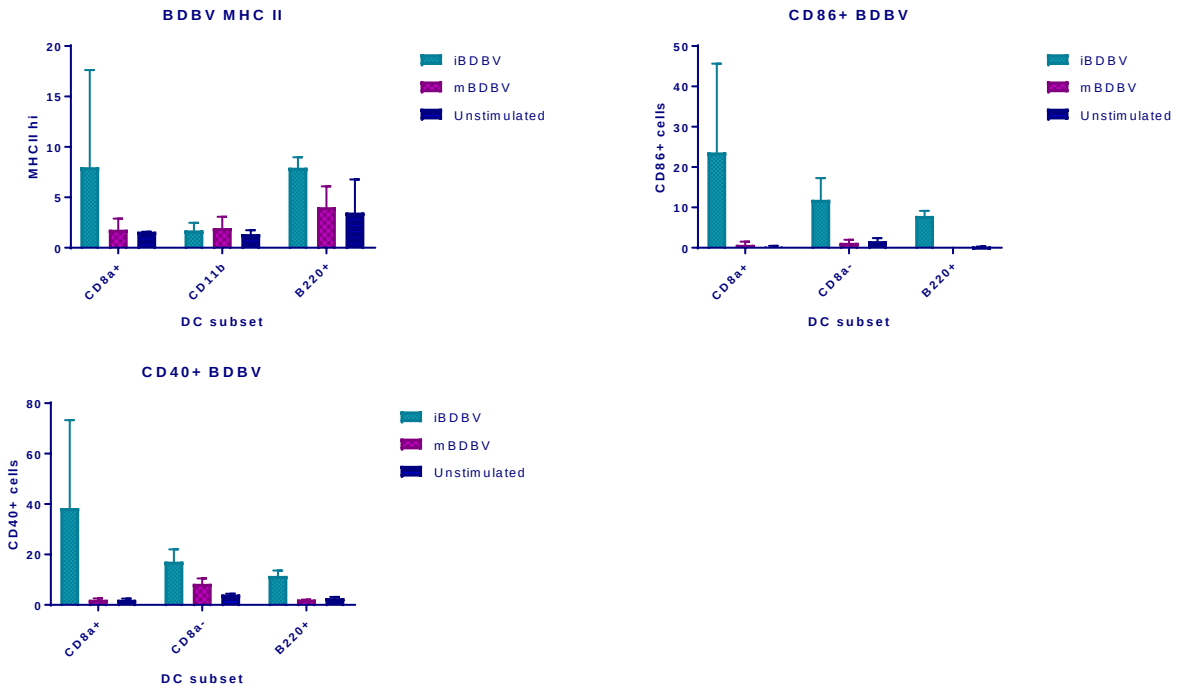


Figure 73. Insect cell-derived BDBV VLPs are more stimulatory to DC subsets than mammalian cell-derived VLPs.

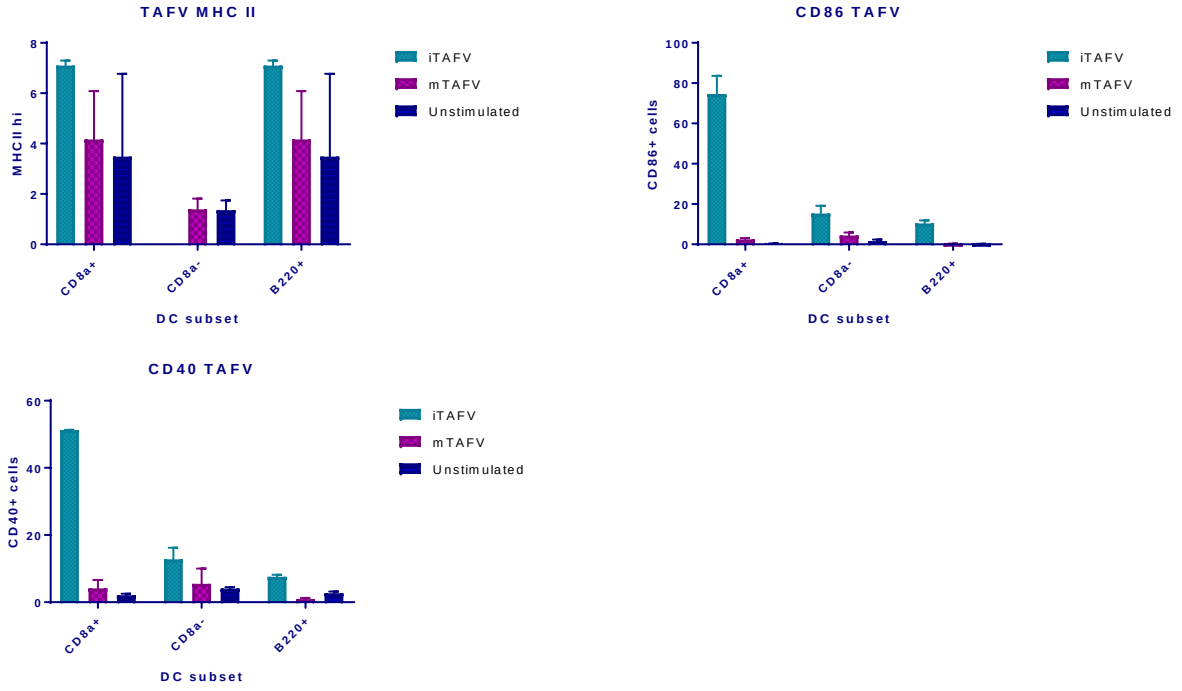


Figure 74. Insect cell-derived TAFV VLPs are more stimulatory to DC subsets than mammalian cell-derived VLPs.

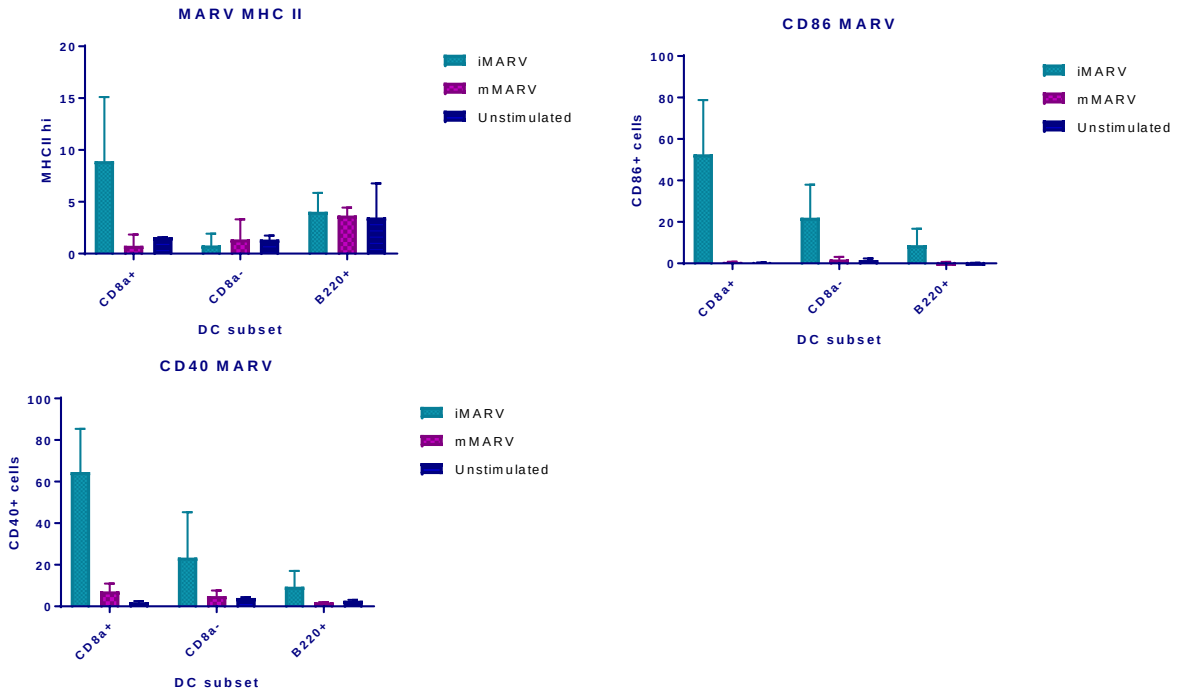
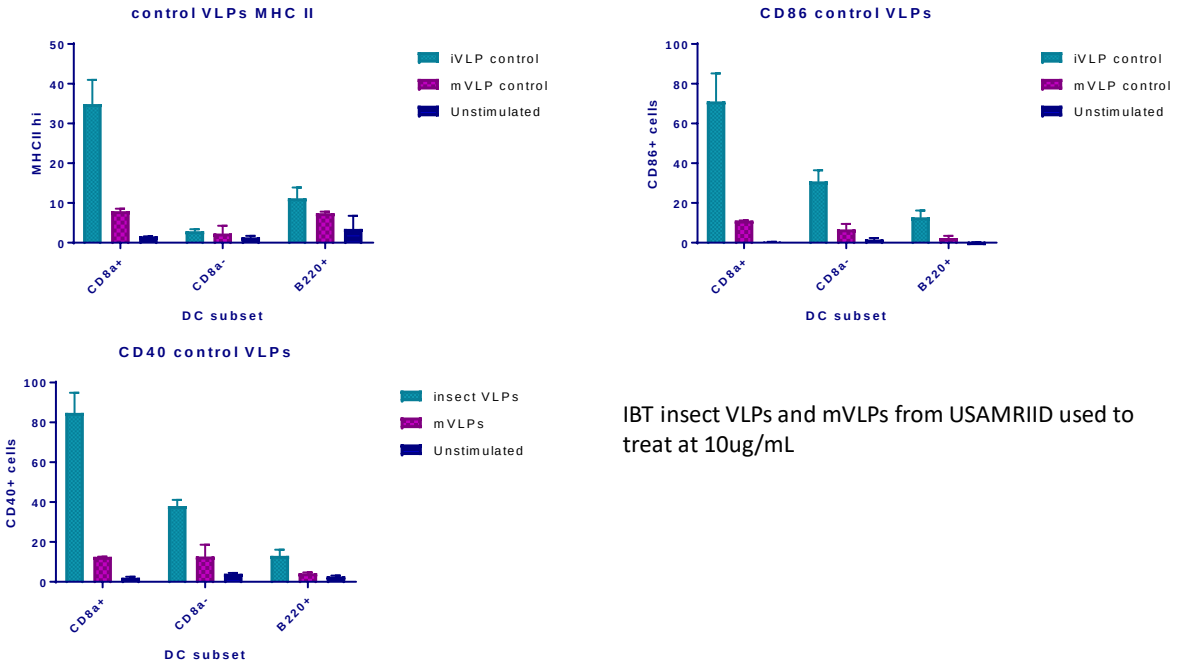


Figure 75. Insect cell-derived MARV-Angola VLPs are more stimulatory to DC subsets than mammalian cell-derived VLPs.



IBT insect VLPs and mVLPs from USAMRIID used to treat at 10ug/mL

Figure 76. VLPs from other sources exhibit similar DC activation profiles compared to those made in our laboratory.

Since the VLPs generated for use in Figures 71-75 were made in our lab, we wanted to ensure that the results we found were not due to lab-specific VLP generation methods. Therefore, we used insect cell-derived EBOV-Yambuku VLPs from a commercial source (IBT) and mammalian-derived EBOV-Yambuku VLPs from United States Army Medical Research Institute of Infectious Diseases (USAMRIID), and repeated the assay. As shown in Figure 76, the data were similar, with insect VLPs exhibiting increased activation of DC subsets.

To analyze whether activation of different DC subsets was due to binding affinity, we collaborated with Dr. Huh Davies to conjugate mammalian-derived EBOV GP1,2 to fluorescent beads. After incubation of EBOV-GP1,2 beads to DCs, we found that all three DC subsets bound to the beads (Figure 77). Future experiments will compare binding of mammalian-GP1,2-coated versus insect GP1,2-coated beads to DC subsets.

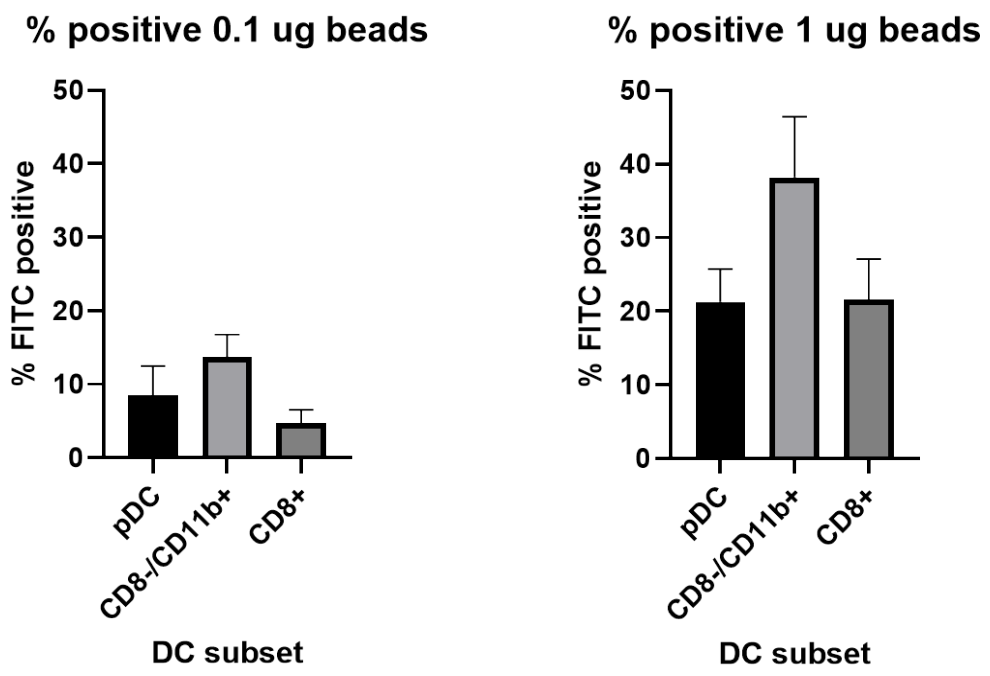


Figure 77. Binding of EBOV GP1,2-coated fluorescent beads to DC subsets.

Key Outcomes

a) Assessment of *in vitro* immunogenicity of insect- versus mammalian cell-derived VLPs expressing filovirus GP1,2. We found that for all filovirus GP1,2s tested (EBOV-Yambuku, EBOV-Makona, SUDV, BDBV, TAFV, and MARV), insect cell-derived VLPs were much more

immunostimulatory for dendritic cells, as assessed by upregulation of multiple cell surface activation markers (CD86, CD40, MHC II).

b) Differential activation of dendritic cells subsets by filovirus VLPs. We found that filovirus VLPs preferentially activate CD8a+ dendritic cells, but also activate CD8a- and plasmacytoid DCs at lower levels. This could have distinct effects on adaptive immunity, as CD8a+ T cells are potent stimulators of CD8+ T cells.

c) Confirmation of our findings with our laboratory-made VLPs and VLPs from outside sources. We found that data obtained with VLPs made by industry (IBT) or other labs (USAMRIID) corroborated with the data obtained by the VLPs made in our laboratory.

CHAPTER 4. ASSESSMENT OF ANTIBODY AND T CELL RESPONSES IN MICE VACCINATED WITH VLPS CONTAINING DIFFERENTIALLY GLYCOSYLATED GP1,2

All current filovirus vaccine platforms utilize the highly glycosylated filovirus attachment protein (glycoprotein, or GP1,2) as an antigen. However, the different vaccine platforms are expressed in various cell types that endow differential glycosylation on GP1,2s, and the importance of the various types of glycosylation in promoting immunogenicity of filovirus GPs is not well known. The first step, in Chapter 2, was to analyze N- and O-linked glycosylation on filovirus GP1,2s expressed in mammalian and insect cells. The second step, in Chapter 3, was to analyze glycosylation in plant-derived GPs, perform *in silico* prediction of filovirus GP glycosylation, and begin to determine effects of differential glycosylation of filovirus GP on

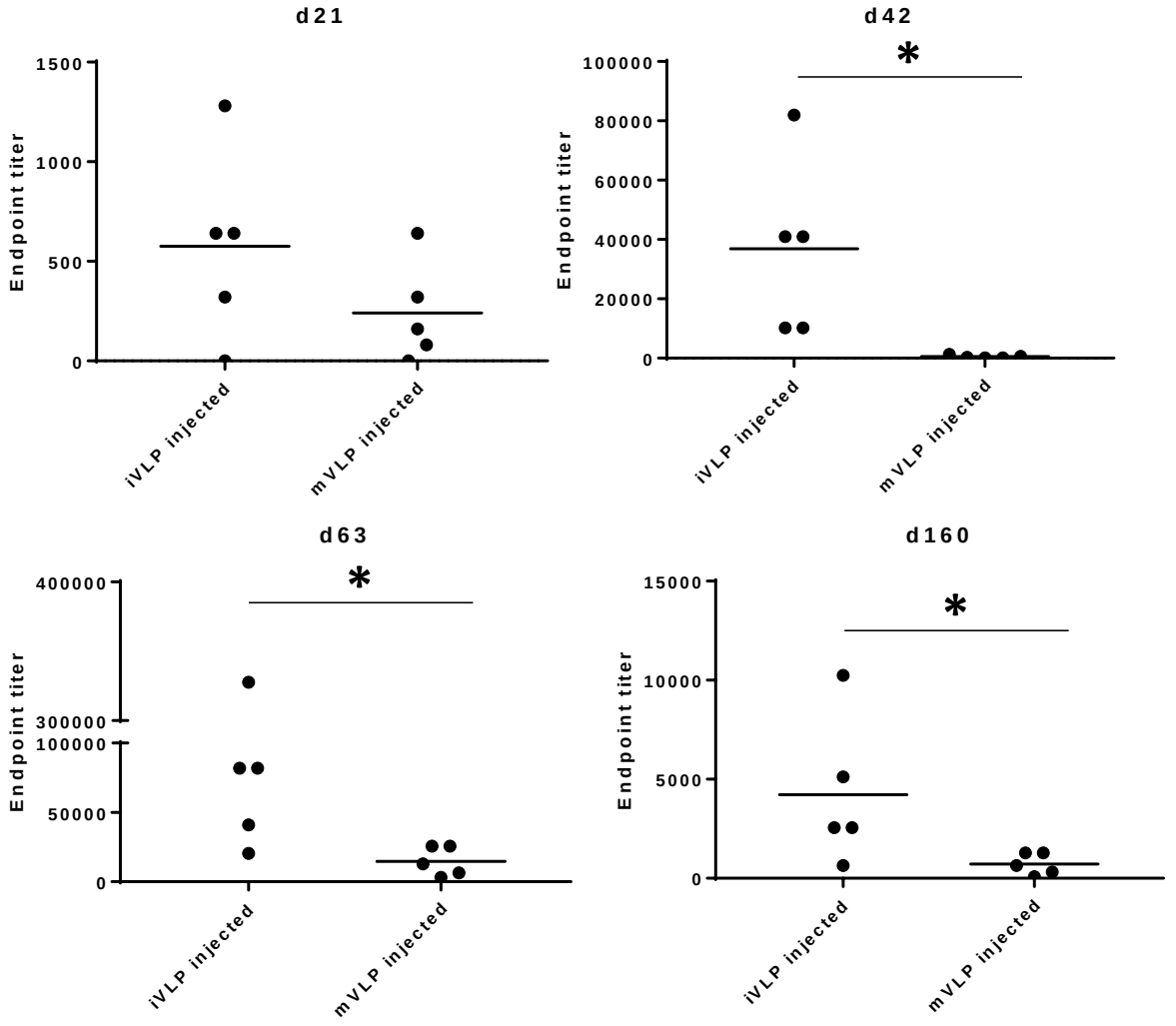


Figure 78. Insect cell-derived VLPs induce higher IgG ELISA titers compared to mammalian cell-derived VLPs. Mice were injected i.p. with 10 ug of VLP on days 0, 21, and 42. Mice were bled and sera was analyzed for anti-GP1,2 binding.

dendritic cell *in vitro.* In this Chapter, we assessed the effects of GP1,2 glycosylation on antibody and T cell responses *in vivo*.

Assessment of differential glycosylation on antibody titers *in vivo* after VLP vaccination. To determine whether glycosylation differences affect immunogenicity *in vivo*, we vaccinated mice with 4.6 ug of EBOV-Makona VLPs made in mammalian cells, versus 1.1 ug VLP made in insect cells, without adjuvant, on days 0, 21, and 42. ELISAs were conducted using purified GP1,2 as an antigen. As shown in Figure 78, much higher IgG antibodies against GP1,2 were generated in insect-derived VLPs than mammalian-derived VLPs on days 42, 63, and 160.

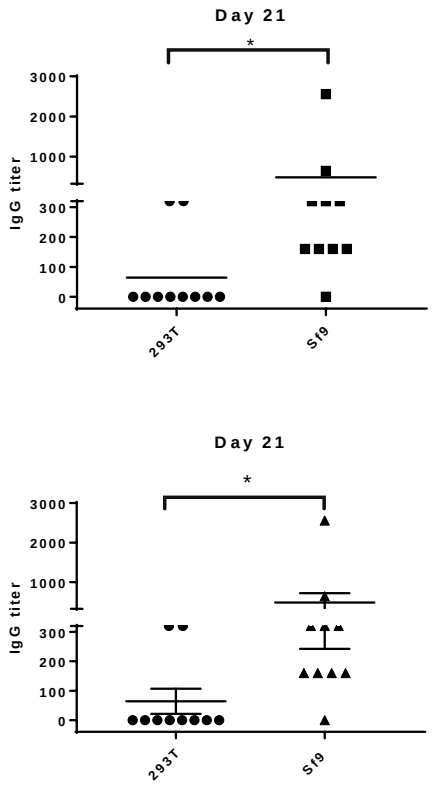


Figure 79. Insect-derived EBOV-Makona GP1,2 induces higher levels of IgG antibody in vaccinated mice. Mice were vaccinated intraperitoneally with 1 ug of mammalian (293T) cell-derived, insect (Sf9) cell-derived, or *in vitro* deglycosylated GP1,2 on day 0 and sera was analyzed for GP1,2 antibodies by ELISA on day 21 using mammalian-derived GP1,2 as the target antigen (top) or insect-derived GP1,2 as the target antigen (bottom).

Determination of total antibody titers induced by differential glycosylation of EBOV-Makona GP1,2

As shown in Chapter 3, using a VLP-based system, insect-derived GP1,2 (from all filovirus species tested) induced higher dendritic cell activation *in vitro* than mammalian-derived GP1,2. Similarly, as shown in Figure 79, EBOV VLPs made in insect cells induced higher antibody titers after vaccination in mice compared to mammalian-derived VLPs. We hypothesized that the lack of O-glycosylation or the more limited, basic N-glycosylation of insect-derived GP1,2 caused the increase in *in vitro* and *in vivo* immunogenicity.

However, we could not rule out the possibility that other factors (such as insect membrane proteins on the membrane of VLPs) could have affected the results. Therefore, we chose to analyze *in vivo* immunogenicity

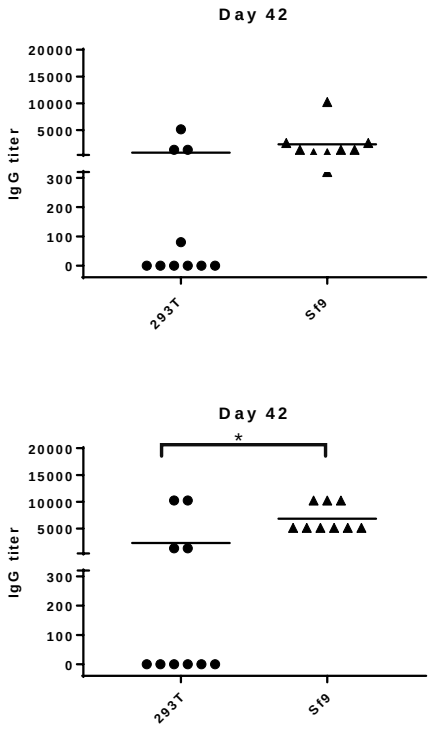


Figure 80. Insect-derived EBOV-Makona GP1,2 induces higher levels of IgG antibody in vaccinated mice. Mice were vaccinated intraperitoneally with 1 ug of mammalian (293T) cell-derived, insect (Sf9) cell-derived, or *in vitro* deglycosylated GP1,2 on day 0 and day 21 and sera was analyzed for GP1,2 antibodies by ELISA on day 42 using mammalian-derived GP1,2 as the target antigen (top) or insect-derived GP1,2 as the target antigen (bottom).

each case, insect-derived GP1,2 induced higher levels of IgG antibodies against GP1,2 after a single vaccination. After two rounds of vaccination, similar results were found, where mice receiving insect-derived GP1,2 had higher levels of IgG antibody (Figure 80).

To confirm that changes in glycosylation was responsible for the of purified subunit GP1,2, derived from either insect (Sf9) or mammalian (293T) cells. In this manner, the glycosylation would be the only difference between the GP1,2s.

To that end, we purified EBOV-Makona GP1,2 from insect or mammalian cells, resulting in vast differences in glycosylation (mammalian-derived GP1,2 has extensive N- and O-glycosylation compared to no O-glycosylation and basic N-glycosylation of insect-derived GP1,2, as we have published. Mice were vaccinated intraperitoneally with 1 ug of GP1,2 mixed with a 1:1 ratio of alum on Days 0, 21, and 42. Blood was collected on Days 0, 21, 42, and 63 and analyzed for antibody binding to purified EBOV-Makona GP1,2 in ELISAs. On day 21 after a single vaccination, Insect-derived GP1,2 induced significantly higher IgG antibody titers than mammalian-derived GP1,2 (Figure 79). To verify that the anti-GP1,2 antibodies induced by vaccination bound fully glycosylated GP1,2, ELISAs were performed using both mammalian-derived (Figure 79, top) or insect-derived GP1,2 (Figure 79, bottom). In

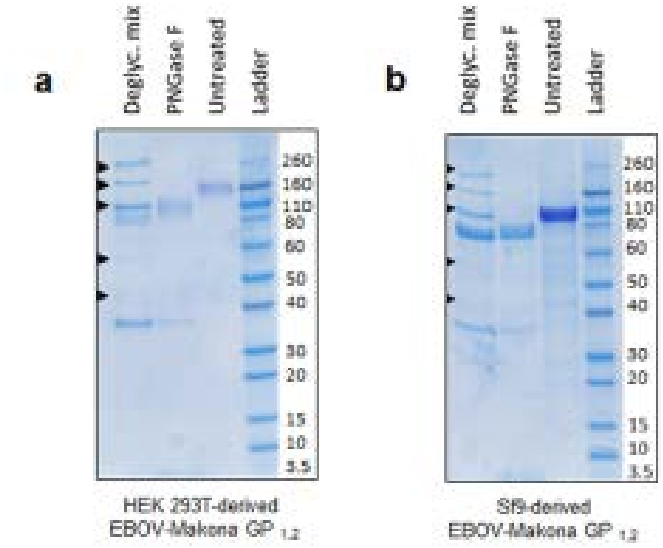


Figure 81. Deglycosylation of N- and O-glycans on EBOV-Makona GP1,2 made in insect or mammalian cells. GP1,2 from mammalian (a) or insect (b) cells was incubated with enzymes that cleave N-glycans (PNGaseF) or both N- and O-glycans (mix) and molecular weight was compared with uncleaved protein by SDS-PAGE and colloidal blue staining. The extra bands in "mix" lanes are the glycosidases.

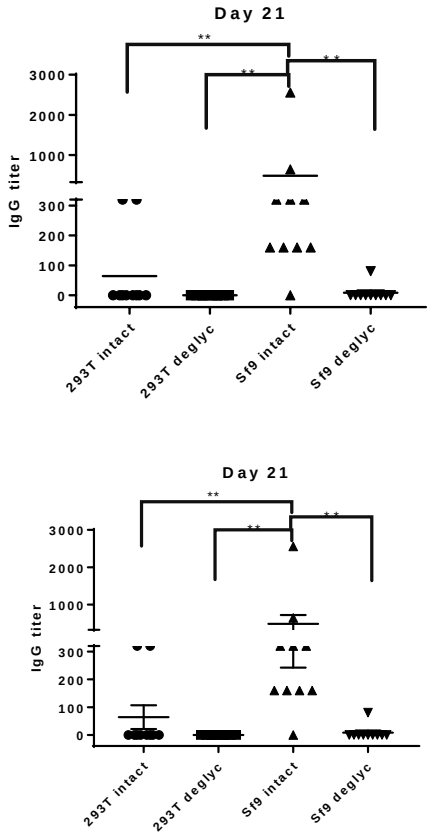


Figure 82. Effects of deglycosylation on GP1,2 immunogenicity. Mice were vaccinated intraperitoneally with 1 ug of mammalian (293T) cell-derived, insect (Sf9) cell-derived, or *in vitro* deglycosylated GP1,2 on days 0 and 21 and sera was analyzed for GP1,2 antibodies by ELISA using mammalian-derived GP1,2 as the target antigen (top) and insect-derived GP1,2 as the target antigen (bottom).

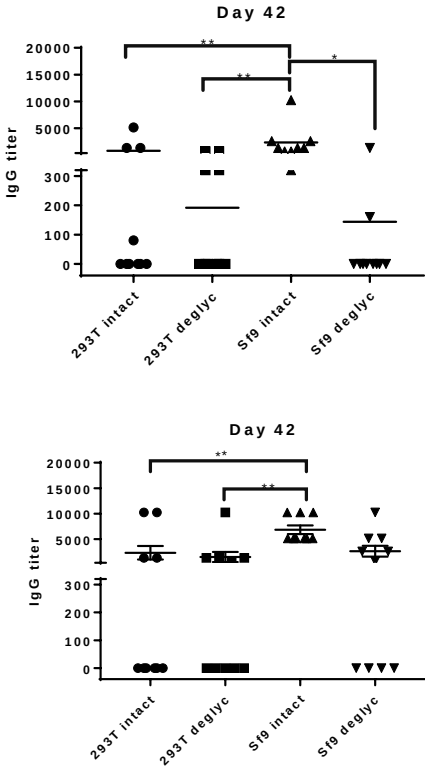


Figure 83. Effects of deglycosylation on GP1,2 immunogenicity. Mice were vaccinated intraperitoneally with 1 ug of mammalian (293T) cell-derived, insect (Sf9) cell-derived, or *in vitro* deglycosylated GP1,2 on days 0 and 21 and sera was analyzed for GP1,2 antibodies by ELISA using mammalian-derived GP1,2 as the target antigen (top) or insect-derived GP1,2 as the target antigen (bottom).

difference in GP1,2 immunogenicity in these experiments, we used a mix of N- and O-glycosidases to deglycosylate insect- or mammalian-derived GP1,2. This resulted in near-complete deglycosylation (Figure 81), as the fully deglycosylated bands ran at the expected size for GP1,2 that lacks post-translational modification (~75 kDa). These deglycosylated GP1,2 were re-purified to remove the glycosidase enzymes, and used to vaccinate mice.

Figure 82 shows the results from mice 21 days post-vaccination with the non-deglycosylated insect and mammalian-derived GP1,2 (labeled “intact”) compared to mice vaccinated with the

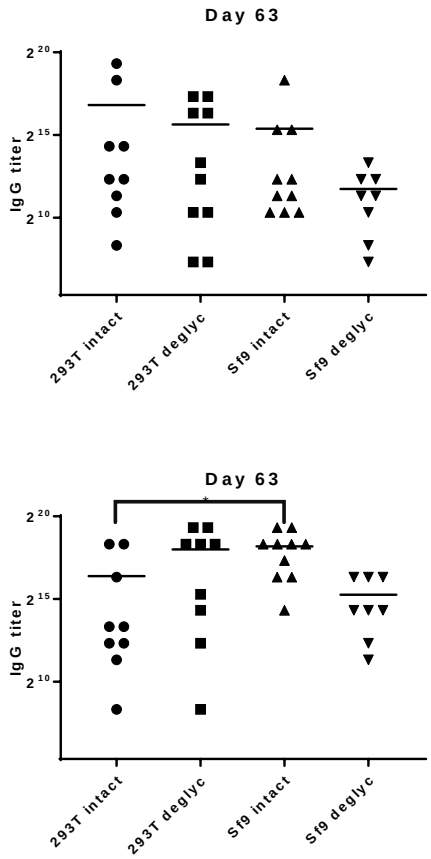


Figure 84. IgG titers after 3 rounds of vaccination. Mice were vaccinated intraperitoneally with 1 ug of mammalian (293T) cell-derived, insect (Sf9) cell-derived, or *in vitro* deglycosylated GP1,2 on day 0, day 21, and day 42 and sera was analyzed for GP1,2 antibodies by ELISA on day 63 using mammalian-derived GP1,2 as the target antigen (top) or insect-derived GP1,2 as the target antigen (bottom).

no difference seen between deglycosylated and non-deglycosylated GP1,2 (Figure 84). However, there was an increase in insect-derived GP1,2 IgG levels compared to mammalian-derived GP1,2 when using insect, but not deglycosylated GP1,2 (labeled "deglyc"). Mice were vaccinated with 1 ug of GP1,2 mixed with alum at a 1:1 ratio. Twenty-one days after the first vaccination, mice that had been vaccinated with deglycosylated insect-derived GP1,2 had decreased specific IgG compared to mice that had been vaccinated with non-deglycosylated insect-derived GP1,2, regardless of the cellular origin of the target antigen. For mammalian-derived GP1,2 it appeared there was a similar decrease in deglycosylated-GP1,2-vaccinated mice, but it was not statistically significant since only 2 mice had detectable antibody titers in the control GP1,2 group.

On day 42, 21 days after the second vaccination, there was a significant decrease in IgG titers from deglycosylated insect-derived GP1,2 compared to non-deglycosylated GP1,2 when using mammalian-derived GP1,2 as the antigen target in ELISAs (Figure 83, top). Although there was also a decrease in studies using insect-derived GP1,2 as an antigen target, this was not statistically significant (Figure 83, bottom). We did not detect significant differences in titers in mice vaccinated with deglycosylated or non-deglycosylated mammalian-derived GP1,2 (Figure 83, top and bottom).

Twenty-one days after the third vaccination, there was

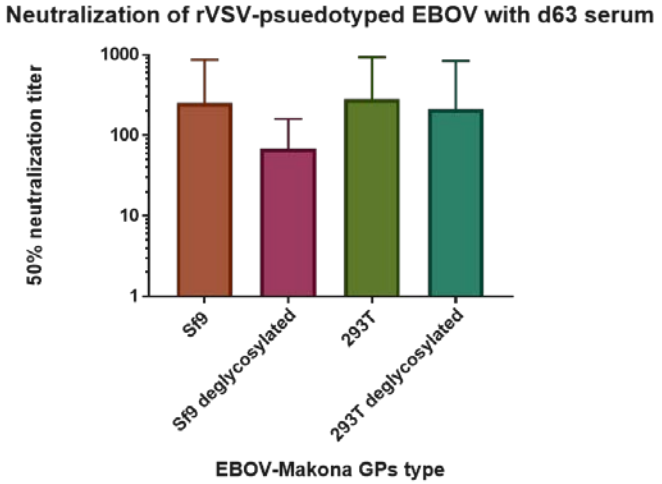


Figure 85. Neutralization assays. Sera from day 63 (21 days after the 3rd vaccination, as shown in Figure 6) were tested for neutralization in intact or deglycosylated insect-derived and mammalian-derived EBOV GP1,2-vaccinated mice.

mammalian, GP1,2 as a target antigen (Figure 84).

Therefore, after 1 or 2 rounds of vaccination, insect-derived GP1,2 induces higher IgG antibody titers that mammalian-derived GP1,2. Deglycosylation of insect-derived GP1,2 eliminates immunogenicity after 1 round of vaccination and possibly after 2 rounds of vaccination. After 3 rounds of vaccination, these differences disappear, possibly due to a "catching up" effect of multiple vaccinations.

To analyze functional antibody activity, we utilized replication incompetent recombinant Vesicular Stomatitis Virus (VSV) expressing luciferase and pseudotyped with EBOV GP1,2. Sera from day 63 (21 days after the 3rd vaccination, as shown in Figure 6) were tested for neutralization in intact or deglycosylated insect-derived and mammalian-derived EBOV GP1,2-vaccinated mice. As shown in Figure 85, there was no significant difference between the groups, matching the ELISA data in Figure 84. However, it is important to note that all samples had significant neutralizing titers, showing that the sera contained functional antibody.

Assessment of T cell activation by intracellular cytokine staining.

Spleens from day 63 mice (21 days after the 3rd vaccination with EBOV-Makona GP1,2, were tested for epitope-specific T cells. Pooled splenocytes from each group were incubated with 15mers overlapping by 10 amino acids spanning the GP1,2 sequence for 4 hours in the presence of brefeldin A, and cells were stained for CD3, CD8, CD4, IFN-gamma, IL-2, and TNFa antibodies. Flow cytometric analysis revealed CD8+ T cells responding to epitopes in vaccinated but not unvaccinated control mice (Figure 86).

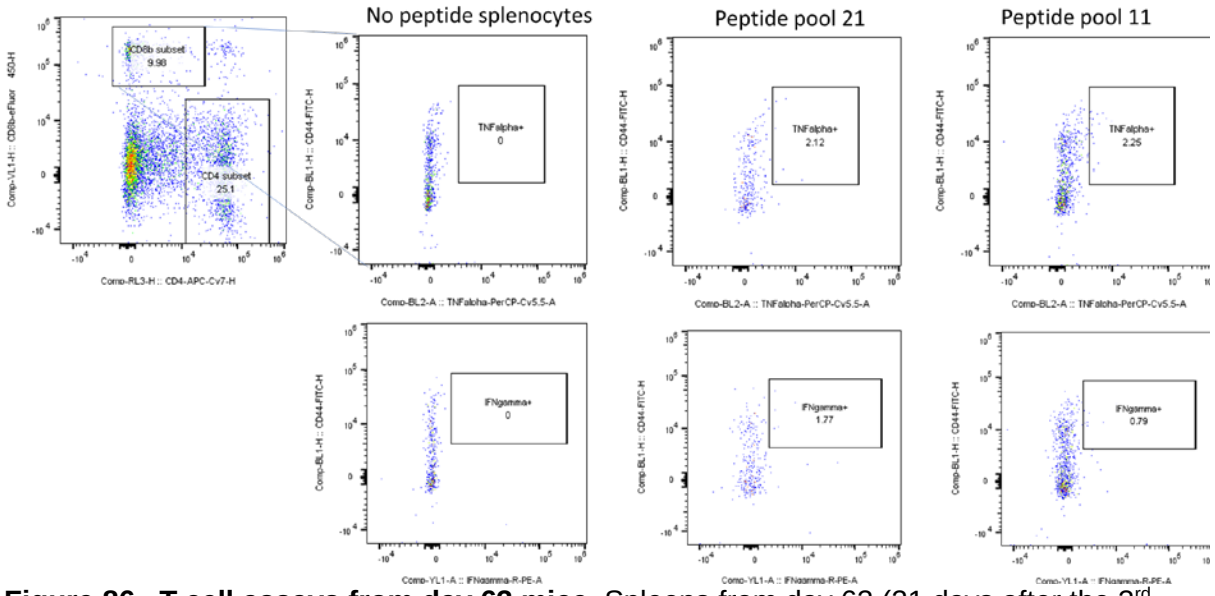


Figure 86. T cell assays from day 63 mice. Spleens from day 63 (21 days after the 3rd vaccination) were tested for epitope-specific T cells. Shown are representative negative controls and positive samples.

Figure 87 shows the peptide pool array used to determine epitopes. Green shading indicates

1	2	3	4	5	6	7	8	9	10	11	12	Peptide pool
1	13	25	37	49	61	73	85	97	109	121	133	13
2	14	26	38	50	62	74	86	98	110	122	134	14
3	15	27	39	51	63	75	87	99	111	123		15
4	16	28	40	52	64	76	88	100	112	124		16
5	17	29	41	53	65	77	89	101	113	125		17
6	18	30	42	54	66	78	90	102	114	126		18
7	19	31	43	55	67	79	91	103	115	127		19
8	20	32	44	56	68	80	92	104	116	128		20
9	21	33	45	57	69	81	93	105	117	129		21
10	22	34	46	58	70	82	94	106	118	130		22
11	23	35	47	59	71	83	95	107	119	131		23
12	24	36	48	60	72	84	96	108	120	132		24

Figure 87. Peptide pools used in T cell assays from day 63 mice. Spleens from day 63 (21 days after the 3rd vaccination) were incubated with the above pools of peptides. Green-shaded cells show "hits" where 15mers were found to induce cytokine responses in T cells. Peptides 106 and 107 (in black box) contain the WIPYFGPAE epitope, which is known to be a T cell epitope in EBOV GP1,2 and was recognized in this assay.

"hits" of epitopes that induced cytokine responses in different groups. Importantly, the epitope WIPYFGPAE (contained in peptide 106 and 107 in Figure 87) was found to induce cytokine responses in the vaccinated mice, serving as an internal control and validating this assay. Individual groups of vaccinated mice (insect-derived GP1,2 mammalian-derived GP1,2, deglycosylated insect-derived GP1,2 and deglycosylated mammalian-derived GP1,2) showed a different array of peptide pools that induced positive responses in T cells (Figure 88). This suggests potential differences in T cell responses based on differential GP1,2 glycosylation.

293T intact	293T deglyc	Sf9 intact	Sf9 deglyc
5	5	1	4
9	6	8	10
10	9	21	11
11	10	23	17
16	11		21
18	21		22
21	22		
22			
23			

Figure 88. Peptide pools identified in different vaccinated groups from day 63 mice. Mice vaccinated with GP1,2 from deglycosylated or non-deglycosylated GP1,2 were tested for T cell cytokine responses from spleens on day 63 (21 days after the 3rd vaccination). The data above show pools of peptides where positive responses were found.

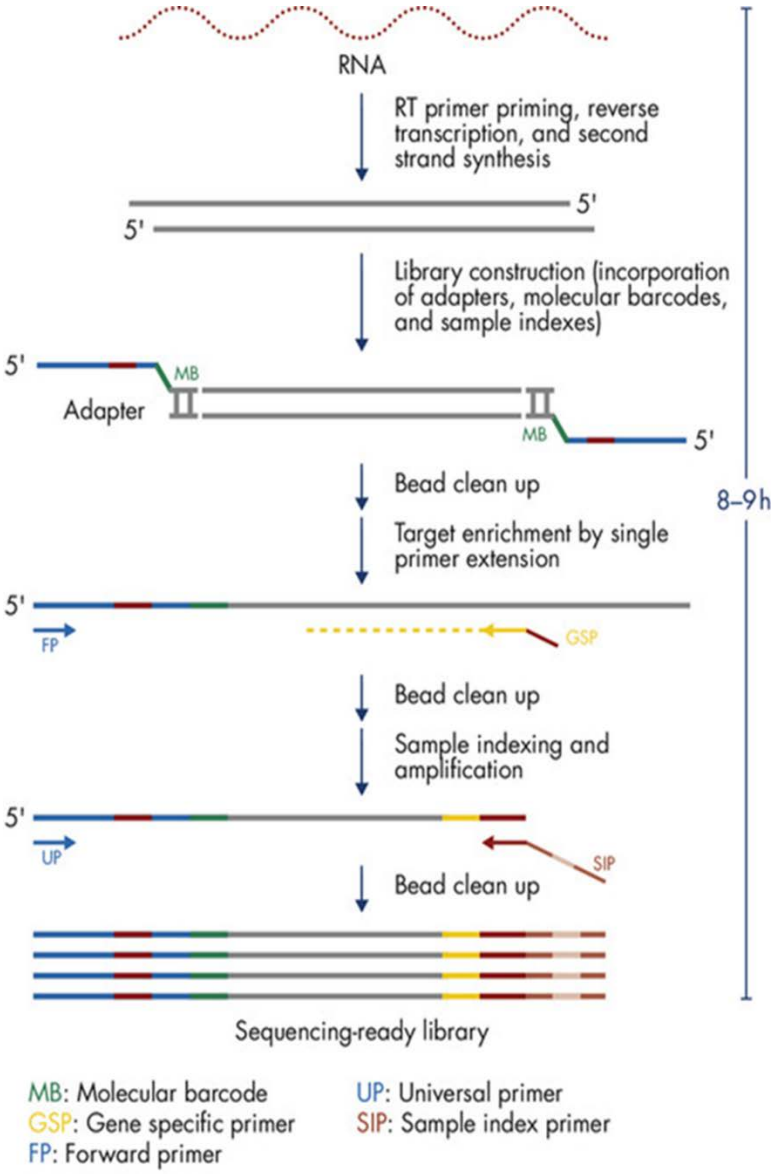


Figure 89. Workflow for T cell receptor sequencing from mRNA.

Assessment of T cell activation by intracellular cytokine staining.

T-cell receptors are encoded by 4 different genes: alpha (TRA), beta (TRB), gamma (TRG) and delta (TRD) and T cells are either alpha-beta or delta-gamma types. One region of the T-cell receptors, the CDR3, is hypervariable due to VDJ and can be sequenced to determine the number of clonotypes present in a sample. One approach, TCR sequencing, uses total RNA extracted from the cell population, and then uses a pool of primers for gene-specific reverse transcription of alpha (TRA), beta (TRB), gamma (TRG) and delta (TRD) TCR mRNAs, followed by gene-specific primers for specific enrichment (Figure 89). To determine abundance of each clonotype, the transcripts have unique molecular identifiers attached, so that when sequenced the number of transcripts in the samples can be determined. Once sequencing has occurred, diversity of the T cells in the population can be assessed using the Shannon-Wiener index. It is computed by the

formula $H = -\sum pi\ R\ i{=}1\ \log pi$, where pi is the proportion of unique molecular identifiers from the ith clonotype and R is the total number of clonotypes. The Shannon-Wiener index was originally proposed in ecology to evaluate the diversity of species but can be applied to the diversity of clonotypes. This allows for an analysis of how well a vaccine induces broad T cell responses as measured by diversity of TCR sequences compared to a control group.

To further analyze the effects of filovirus GP1,2 glycosylation on immunogenicity, we analyzed TCR diversity in mice vaccinated with insect, mammalian, insect deglycosylated, mammalian deglycosylated, or mock-vaccinated groups. Figure 90 below shows the clonotype diversity of sequenced TCRs from mouse spleens. There are significant differences in the clonotype diversity between the vaccine groups; this preliminary analysis suggests that vaccination with insect GP1,2 resulted in higher TCR diversity in both alpha and beta chains than mammalian-derived GP1,2, consistent with our other data showing stronger immune responses in insect GP1,2 vaccinated mice.

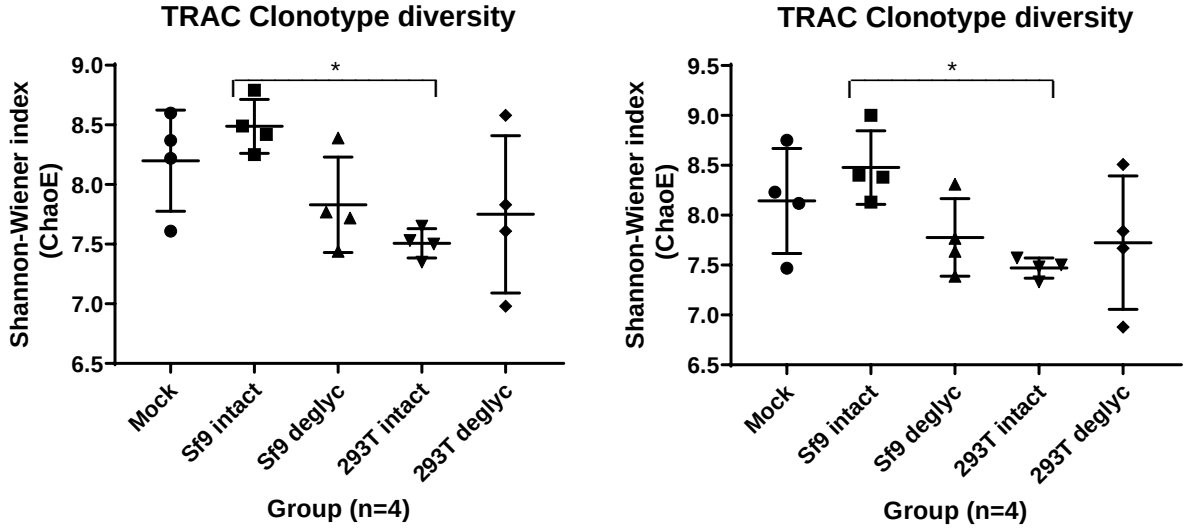


Figure 90: Preliminary data from Alpha and Beta chain TCR clonotype diversity as measured by the Shannon-Wiener formula from sequenced splenocytes from vaccinated mice. Four mice in each group from each of the vaccinated groups (insect GP intact, insect GP deglycosylated, mammalian GP intact, mammalian GP deglycosylated and mock (PBS only) vaccinated) were subjected to TCR sequencing.

Live virus challenge of vaccinated mice. We have received approval for live virus challenge of vaccinated mice. The Integrated Research Facility (IRF) at the National Institutes of Health (NIH) has agreed to challenge mice vaccinated with purified MARV GP1,2 and purified EBOV GP1,2 with mouse-adapted MARV and mouse adapted EBOV, respectively. Mice will be vaccinated with mammalian (293T)-derived, insect (Sf9)-derived, enzymatically deglycosylated 293T-derived, or enzymatically deglycosylated Sf9-derived GP1,2. These experiments were

slated for early 2020 but the coronavirus pandemic has delayed this project. However, we will still pursue these experiments even though the grant has ended.

CHAPTER 5. CONCLUSIONS AND FUTURE DIRECTIONS

The goal of this project was to 1) determine glycosylation patterns on filovirus GP1,2 expressed in mammalian, insect, and plant cells; 2) analyze *in vitro* immunogenicity of differentially glycosylated filovirus GP1,2; and 3) assess *in vivo* immunogenicity of differentially glycosylated filovirus GP1,2. We found that mammalian cell-derived GP1,2 had extensive and varied N- and O-linked glycosylation, insect cell-derived GP1,2 had simple N-glycosylation, and plant cell-derived GP1,2 had high mannose N-glycosylation. We found that insect-derived GP1,2 was more immunogenic to dendritic cells *in vitro* than mammalian-derived GP1,2 in a VLP format. Our in vivo studies indicated that insect-derived GP1,2 was more immunogenic in generation of antibody and T cell responses that mammalian-derived or deglycosylated GP1,2. Together, these data show that changing glycosylation on viral glycoproteins can have significant effects on immunogenicity, which is an important concept for vaccine development and production strategies. This grant has been used to train over 20 scientists at different levels of training and was included in several publications and many more posters and oral presentations.

We were unable to complete the *in vivo* challenge experiments with mice vaccinated with EBOV or MARV mammalian, insect, or deglycosylated GP1,2 due to the SARS-CoV-2 pandemic. We did obtain permission to have these experiments conducted with the NIH, and we plan to pursue this experiment even though the grant has ended.

Therefore, our future experiments will include live virus challenge of mice vaccinated with differentially glycosylated GP1,2.Other experiments will revolve around more extensive testing of TCR diversity and clonotype magnitude in these vaccinated mice.

We appreciate the support of DTRA for this project.

REFERENCES

1. Clarke EC, Collar AL, Ye C, Cai Y, Anaya E, Rinaldi D, et al. Production and Purification of Filovirus Glycoproteins in Insect and Mammalian Cell Lines. Scientific reports. 2017;7(1):15091. doi: 10.1038/s41598-017-15416-3. PubMed PMID: 29118454; PubMed Central PMCID: PMCPMC5678155.
2. Yang X-L, Tan CW, Anderson DE, Jiang R-D, Li B, Zhang W, et al. Characterization of a filovirus (Měnglà virus) from Rousettus bats in China. Nature Microbiology. 2019;4(3):390-5. doi: 10.1038/s41564-018-0328-y.
3. Shi M, Lin XD, Chen X, Tian JH, Chen LJ, Li K, et al. The evolutionary history of vertebrate RNA viruses. Nature. 2018;556(7700):197-202. Epub 2018/04/06. doi: 10.1038/s41586-018-0012-7. PubMed PMID: 29618816.
4. Bradfute SB, Bavari S. Correlates of immunity to filovirus infection. Viruses. 2011;3(7):982-1000. doi: 10.3390/v3070982. PubMed PMID: 21994766; PubMed Central PMCID: PMC3185794.
5. Sachs JL. Resolving the first steps to multicellularity. Trends Ecol Evol (Amst). 2008;23(5):245-8. doi: 10.1016/j.tree.2008.02.003.
6. Moremen KW, Tiemeyer M, Nairn AV. Vertebrate protein glycosylation: diversity, synthesis and function. Nat Rev Mol Cell Biol. 2012;13(7):448-62. doi: 10.1038/nrm3383.
7. Hebert DN, Lamriben L, Powers ET, Kelly JW. The intrinsic and extrinsic effects of N-linked glycans on glycoproteostasis. Nature chemical biology. 2014;10(11):902-10. doi: 10.1038/nchembio.1651.
8. Eichler R, Lenz O, Garten W, Strecker T. The role of single N-glycans in proteolytic processing and cell surface transport of the Lassa virus glycoprotein GP-C. Virology journal. 2006;3:41. doi: 10.1186/1743-422X-3-41. PubMed PMID: 16737539; PubMed Central PMCID: PMCPMC1524727.
9. McGinnes LW, Morrison TG. The Role of Individual Oligosaccharide Chains in the Activities of the HN Glycoprotein of Newcastle Disease Virus. Virology. 1995;212(2):398-410. doi: 10.1006/viro.1995.1497.
10. Panda A, Elankumaran S, Krishnamurthy S, Huang Z, Samal SK. Loss of N-Linked Glycosylation from the Hemagglutinin- Neuraminidase Protein Alters Virulence of Newcastle Disease Virus. Journal of virology. 2004;78(10):4965-75. doi: 10.1128/JVI.78.10.4965-4975.2004.
11. Shi X, Elliott RM. Analysis of N-Linked Glycosylation of Hantaan Virus Glycoproteins and the Role of Oligosaccharide Side Chains in Protein Folding and Intracellular Trafficking. Journal of virology. 2004;78(10):5414-22. doi: 10.1128/JVI.78.10.5414-5422.2004.
12. Doores KJ. The HIV glycan shield as a target for broadly neutralizing antibodies. The FEBS Journal. 2015;282(24):4679-91. doi: 10.1111/febs.13530.
13. Reed DJ, Lin X, Thomas TD, Birks CW, Tang J, Rother RP. Alteration of glycosylation renders HIV sensitive to inactivation by normal human serum. The Journal of Immunology. 1997;159(9):4356-61.
14. Quiñones-Kochs MI, Buonocore L, Rose JK. Role of N-Linked Glycans in a Human Immunodeficiency Virus Envelope Glycoprotein: Effects on Protein Function and the Neutralizing Antibody Response. Journal of virology. 2002;76(9):4199-211. doi: 10.1128/JVI.76.9.4199-4211.2002.
15. Crispin M, Ward AB, Wilson IA. Structure and Immune Recognition of the HIV Glycan Shield. Annual Review of Biophysics. 2018;47(1):499-523. doi: 10.1146/annurev-biophys-060414-034156.
16. Mårdberg K, Nyström K, Tarp MA, Trybala E, Clausen H, Bergström T, et al. Basic amino acids as modulators of an O-linked glycosylation signal of the herpes simplex virus type 1 glycoprotein gC: functional roles in viral infectivity. Glycobiology. 2004;14(7):571-81. doi: 10.1093/glycob/cwh075.

17. Bagdonaite I, Nordén R, Joshi HJ, King SL, Vakhrushev SY, Olofsson S, et al. Global Mapping of O-Glycosylation of Varicella Zoster Virus, Human Cytomegalovirus, and Epstein-Barr Virus. J Biol Chem. 2016;291(23):12014-28. doi: 10.1074/jbc.M116.721746.
18. Carette JE, Raaben M, Wong AC, Herbert AS, Obernosterer G, Mulherkar N, et al. Ebola virus entry requires the cholesterol transporter Niemann–Pick C1. Nature. 2011;477(7364):340-3. doi: 10.1038/nature10348.
19. Cote M, Misasi J, Ren T, Bruchez A, Lee K, Filone CM, et al. Small molecule inhibitors reveal Niemann-Pick C1 is essential for Ebola virus infection. Nature. 2011;477(7364):344-8. doi: 10.1038/nature10380. PubMed PMID: 21866101; PubMed Central PMCID: PMC3230319.
20. de La Vega M-A, Wong G, Kobinger GP, Qiu X. The Multiple Roles of sGP in Ebola Pathogenesis. Viral immunology. 2014;28(1):3-9. doi: 10.1089/vim.2014.0068.
21. Lennemann NJ, Rhein BA, Ndungo E, Chandran K, Qiu X, Maury W. Comprehensive functional analysis of N-linked glycans on Ebola virus GP1. mBio. 2014;5(1):e00862-13. doi: 10.1128/mBio.00862-13. PubMed PMID: 24473128; PubMed Central PMCID: PMCPMC3950510.
22. Lennemann NJ, Walkner M, Berkebile AR, Patel N, Maury W. The Role of Conserved N-Linked Glycans on Ebola Virus Glycoprotein 2. The Journal of infectious diseases. 2015;212(suppl_2):S204-S9. doi: 10.1093/infdis/jiv201.
23. Baribaud F, Pöhlmann S, Leslie G, Mortari F, Doms RW. Quantitative Expression and Virus Transmission Analysis of DC-SIGN on Monocyte-Derived Dendritic Cells. Journal of virology. 2002;76(18):9135-42. doi: 10.1128/JVI.76.18.9135-9142.2002.
24. Lin G, Simmons G, Pohlmann S, Baribaud F, Ni H, Leslie GJ, et al. Differential N-linked glycosylation of human immunodeficiency virus and Ebola virus envelope glycoproteins modulates interactions with DC-SIGN and DC-SIGNR. Journal of virology. 2003;77(2):1337-46. PubMed PMID: 12502850; PubMed Central PMCID: PMC140807.
25. Simmons G, Reeves JD, Grogan CC, Vandenberghe LH, Baribaud F, Whitbeck JC, et al. DC-SIGN and DC-SIGNR bind ebola glycoproteins and enhance infection of macrophages and endothelial cells. Virology. 2003;305(1):115-23. PubMed PMID: 12504546.
26. Alvarez CP, Lasala F, Carrillo J, Muñiz O, Corbí AL, Delgado R. C-Type Lectins DC-SIGN and L-SIGN Mediate Cellular Entry by Ebola Virus in cis and in trans. Journal of virology. 2002;76(13):6841-4. doi: 10.1128/JVI.76.13.6841-6844.2002.
27. Takada A, Fujioka K, Tsuiji M, Morikawa A, Higashi N, Ebihara H, et al. Human Macrophage C-Type Lectin Specific for Galactose and N-Acetylgalactosamine Promotes Filovirus Entry. Journal of virology. 2004;78(6):2943-7. doi: 10.1128/JVI.78.6.2943-2947.2004.
28. Becker S, Spiess M, Klenk H-D. The asialoglycoprotein receptor is a potential liver-specific receptor for Marburg virus. Journal of General Virology. 1995;76(2):393-9. doi: 10.1099/0022-1317-76-2-393.
29. Geyer H, Will C, Feldmann H, Klenk HD, Geyer R. Carbohydrate structure of Marburg virus glycoprotein. Glycobiology. 1992;2(4):299-312. PubMed PMID: 1421752.
30. Noyori O, Matsuno K, Kajihara M, Nakayama E, Igarashi M, Kuroda M, et al. Differential potential for envelope glycoprotein-mediated steric shielding of host cell surface proteins among filoviruses. Virology. 2013;446(1):152-61. doi: 10.1016/j.virol.2013.07.029.
31. Francica JR, Varela-Rohena A, Medvec A, Plesa G, Riley JL, Bates P. Steric Shielding of Surface Epitopes and Impaired Immune Recognition Induced by the Ebola Virus Glycoprotein. PLoS pathogens. 2010;6(9). doi: 10.1371/journal.ppat.1001098.
32. Ritchie G, Harvey DJ, Stroeher U, Feldmann F, Feldmann H, Wahl-Jensen V, et al. Identification of N-glycans from Ebola virus glycoproteins by matrix-assisted laser desorption/ionisation time-of-flight and negative ion electrospray tandem mass spectrometry. Rapid communications in mass spectrometry

: RCM. 2010;24(5):571-85. doi: 10.1002/rcm.4410. PubMed PMID: 20131323; PubMed Central PMCID: PMC3399782.
33. Feldmann H, Will C, Schikore M, Slenczka W, Klenk HD. Glycosylation and oligomerization of the spike protein of Marburg virus. Virology. 1991;182(1):353-6. PubMed PMID: 2024471.
34. Feldmann H, Nichol ST, Klenk HD, Peters CJ, Sanchez A. Characterization of filoviruses based on differences in structure and antigenicity of the virion glycoprotein. Virology. 1994;199(2):469-73. doi: 10.1006/viro.1994.1147. PubMed PMID: 8122375.
35. Bradfute SB, Dye JM, Jr., Bavari S. Filovirus vaccines. Human vaccines. 2011;7(6):701-11. PubMed PMID: 21519188; PubMed Central PMCID: PMC3219077.
36. Friedrich BM, Trefry JC, Biggins JE, Hensley LE, Honko AN, Smith DR, et al. Potential vaccines and post-exposure treatments for filovirus infections. Viruses. 2012;4(9):1619-50. doi: 10.3390/v4091619. PubMed PMID: 23170176; PubMed Central PMCID: PMC3499823.
37. Blaney JE, Marzi A, Willet M, Papaneri AB, Wirblich C, Feldmann F, et al. Antibody quality and protection from lethal Ebola virus challenge in nonhuman primates immunized with rabies virus based bivalent vaccine. PLoS pathogens. 2013;9(5):e1003389. doi: 10.1371/journal.ppat.1003389. PubMed PMID: 23737747; PubMed Central PMCID: PMC3667758.
38. Wilson JA, Hevey M, Bakken R, Guest S, Bray M, Schmaljohn AL, et al. Epitopes involved in antibody-mediated protection from Ebola virus. Science. 2000;287(5458):1664-6. PubMed PMID: 10698744.
39. Olinger GG, Bailey MA, Dye JM, Bakken R, Kuehne A, Kondig J, et al. Protective cytotoxic T-cell responses induced by venezuelan equine encephalitis virus replicons expressing Ebola virus proteins. Journal of virology. 2005;79(22):14189-96. doi: 10.1128/JVI.79.22.14189-14196.2005. PubMed PMID: 16254354; PubMed Central PMCID: PMC1280180.
40. Sullivan NJ, Hensley L, Asiedu C, Geisbert TW, Stanley D, Johnson J, et al. CD8+ cellular immunity mediates rAd5 vaccine protection against Ebola virus infection of nonhuman primates. Nature medicine. 2011;17(9):1128-31. doi: 10.1038/nm.2447. PubMed PMID: 21857654.
41. Bukreyev A, Rollin PE, Tate MK, Yang L, Zaki SR, Shieh WJ, et al. Successful topical respiratory tract immunization of primates against Ebola virus. Journal of virology. 2007;81(12):6379-88. Epub 2007/04/13. doi: 10.1128/JVI.00105-07. PubMed PMID: 17428868; PubMed Central PMCID: PMC1900097.
42. Ledgerwood JE, Costner P, Desai N, Holman L, Enama ME, Yamshchikov G, et al. A replication defective recombinant Ad5 vaccine expressing Ebola virus GP is safe and immunogenic in healthy adults. Vaccine. 2010;29(2):304-13. doi: 10.1016/j.vaccine.2010.10.037. PubMed PMID: 21034824.
43. Martin JE, Sullivan NJ, Enama ME, Gordon IJ, Roederer M, Koup RA, et al. A DNA vaccine for Ebola virus is safe and immunogenic in a phase I clinical trial. Clinical and vaccine immunology : CVI. 2006;13(11):1267-77. doi: 10.1128/CVI.00162-06. PubMed PMID: 16988008; PubMed Central PMCID: PMC1656552.
44. Herbert AS, Kuehne AI, Barth JF, Ortiz RA, Nichols DK, Zak SE, et al. Venezuelan equine encephalitis virus replicon particle vaccine protects nonhuman primates from intramuscular and aerosol challenge with ebolavirus. Journal of virology. 2013;87(9):4952-64. doi: 10.1128/JVI.03361-12. PubMed PMID: 23408633; PubMed Central PMCID: PMC3624300.
45. Garbutt M, Liebscher R, Wahl-Jensen V, Jones S, Moller P, Wagner R, et al. Properties of replication-competent vesicular stomatitis virus vectors expressing glycoproteins of filoviruses and arenaviruses. Journal of virology. 2004;78(10):5458-65. PubMed PMID: 15113924; PubMed Central PMCID: PMCPMC400370.

46. Warfield KL, Posten NA, Swenson DL, Olinger GG, Esposito D, Gillette WK, et al. Filovirus-like particles produced in insect cells: immunogenicity and protection in rodents. The Journal of infectious diseases. 2007;196 Suppl 2:S421-9. doi: 10.1086/520612. PubMed PMID: 17940979.
47. Mellquist-Riemenschneider JL, Garrison AR, Geisbert JB, Saikh KU, Heidebrink KD, Jahrling PB, et al. Comparison of the protective efficacy of DNA and baculovirus-derived protein vaccines for EBOLA virus in guinea pigs. Virus research. 2003;92(2):187-93. PubMed PMID: 12686428.
48. Sun Y, Carrion R, Jr., Ye L, Wen Z, Ro YT, Brasky K, et al. Protection against lethal challenge by Ebola virus-like particles produced in insect cells. Virology. 2009;383(1):12-21. doi: 10.1016/j.virol.2008.09.020. PubMed PMID: 18986663; PubMed Central PMCID: PMC2657000.
49. Phoolcharoen W, Bhoo SH, Lai H, Ma J, Arntzen CJ, Chen Q, et al. Expression of an immunogenic Ebola immune complex in Nicotiana benthamiana. Plant biotechnology journal. 2011;9(7):807-16. doi: 10.1111/j.1467-7652.2011.00593.x. PubMed PMID: 21281425; PubMed Central PMCID: PMCPMC4022790.
50. Kushnir N, Streatfield SJ, Yusibov V. Virus-like particles as a highly efficient vaccine platform: diversity of targets and production systems and advances in clinical development. Vaccine. 2012;31(1):58-83. doi: 10.1016/j.vaccine.2012.10.083. PubMed PMID: 23142589.
51. Okumura A, Pitha PM, Yoshimura A, Harty RN. Interaction between Ebola virus glycoprotein and host toll-like receptor 4 leads to induction of proinflammatory cytokines and SOCS1. Journal of virology. 2010;84(1):27-33. doi: 10.1128/JVI.01462-09. PubMed PMID: 19846529; PubMed Central PMCID: PMC2798428.
52. Martinez O, Valmas C, Basler CF. Ebola virus-like particle-induced activation of NF-kappaB and Erk signaling in human dendritic cells requires the glycoprotein mucin domain. Virology. 2007;364(2):342-54. doi: 10.1016/j.virol.2007.03.020. PubMed PMID: 17434557; PubMed Central PMCID: PMC2034500.
53. Ye L, Lin J, Sun Y, Bennouna S, Lo M, Wu Q, et al. Ebola virus-like particles produced in insect cells exhibit dendritic cell stimulating activity and induce neutralizing antibodies. Virology. 2006;351(2):260-70. doi: 10.1016/j.virol.2006.03.021. PubMed PMID: 16678231.
54. Ou W, Delisle J, Jacques J, Shih J, Price G, Kuhn JH, et al. Induction of ebolavirus cross-species immunity using retrovirus-like particles bearing the Ebola virus glycoprotein lacking the mucin-like domain. Virology journal. 2012;9:32. doi: 10.1186/1743-422X-9-32. PubMed PMID: 22273269; PubMed Central PMCID: PMC3284443.
55. Hevey M, Negley D, Geisbert J, Jahrling P, Schmaljohn A. Antigenicity and vaccine potential of Marburg virus glycoprotein expressed by baculovirus recombinants. Virology. 1997;239(1):206-16. Epub 1998/01/14. doi: S0042-6822(97)98883-8 [pii]

10.1006/viro.1997.8883. PubMed PMID: 9426460.