



# Site-Specific Glycosylation of Recombinant Viral Glycoproteins Produced in *Nicotiana benthamiana*

Emmanuel Margolin<sup>1,2,3,4\*†</sup>, Joel D. Allen<sup>5†</sup>, Matthew Verbeek<sup>4</sup>, Michiel van Diepen<sup>1,3</sup>, Phindile Ximba<sup>1,3</sup>, Rosamund Chapman<sup>1,3</sup>, Ann Meyers<sup>4</sup>, Anna-Lise Williamson<sup>1,2,3</sup>, Max Crispin<sup>5\*</sup> and Edward Rybicki<sup>3,4</sup>

<sup>1</sup> Division of Medical Virology, Department of Pathology, Faculty of Health Sciences, University of Cape Town, Cape Town, South Africa, <sup>2</sup> Wellcome Trust Centre for Infectious Disease Research in Africa, University of Cape Town, Cape Town, South Africa, <sup>3</sup> Institute of Infectious Disease and Molecular Medicine, Faculty of Health Sciences, University of Cape Town, Cape Town, South Africa, <sup>4</sup> Biopharming Research Unit, Department of Molecular and Cell Biology, University of Cape Town, Cape Town, South Africa, <sup>5</sup> School of Biological Sciences, University of Southampton, Southampton, United Kingdom

## OPEN ACCESS

### Edited by:

Natacha Soto Pérez,  
Center for Genetic Engineering  
and Biotechnology (CIGB), Cuba

### Reviewed by:

Marina Clemente,  
CONICET Institute of Biotechnological  
Research (IB-INTECH), Argentina  
Somen Nandi,  
University of California, Davis,  
United States

### \*Correspondence:

Emmanuel Margolin  
Emmanuel.margolin@uct.ac.za  
Max Crispin  
max.crispin@soton.ac.uk

† These authors share first authorship

### Specialty section:

This article was submitted to  
Plant Biotechnology,  
a section of the journal  
Frontiers in Plant Science

**Received:** 13 May 2021

**Accepted:** 24 June 2021

**Published:** 22 July 2021

### Citation:

Margolin E, Allen JD, Verbeek M, van Diepen M, Ximba P, Chapman R, Meyers A, Williamson A-L, Crispin M and Rybicki E (2021) Site-Specific Glycosylation of Recombinant Viral Glycoproteins Produced in *Nicotiana benthamiana*.  
*Front. Plant Sci.* 12:709344.  
doi: 10.3389/fpls.2021.709344

There is an urgent need to establish large scale biopharmaceutical manufacturing capacity in Africa where the infrastructure for biologics production is severely limited. Molecular farming, whereby pharmaceuticals are produced in plants, offers a cheaper alternative to mainstream expression platforms, and is amenable to rapid large-scale production. However, there are several differences along the plant protein secretory pathway compared to mammalian systems, which constrain the production of complex pharmaceuticals. Viral envelope glycoproteins are important targets for immunization, yet in some cases they accumulate poorly in plants and may not be properly processed. Whilst the co-expression of human chaperones and furin proteases has shown promise, it is presently unclear how plant-specific differences in glycosylation impact the production of these proteins. In many cases it may be necessary to reproduce features of their native glycosylation to produce immunologically relevant vaccines, given that glycosylation is central to the folding and immunogenicity of these antigens. Building on previous work, we transiently expressed model glycoproteins from HIV and Marburg virus in *Nicotiana benthamiana* and mammalian cells. The proteins were purified and their site-specific glycosylation was determined by mass-spectrometry. Both glycoproteins yielded increased amounts of protein aggregates when produced in plants compared to the equivalent mammalian cell-derived proteins. The glycosylation profiles of the plant-produced glycoproteins were distinct from the mammalian cell produced proteins: they displayed lower levels of glycan occupancy, reduced complex glycans and large amounts of paucimannosidic structures. The elucidation of the site-specific glycosylation of viral glycoproteins produced in *N. benthamiana* is an important step toward producing heterologous viral glycoproteins in plants with authentic human-like glycosylation.

**Keywords:** glycoprotein, glycosylation, occupancy, folding, processing, molecular pharming

## INTRODUCTION

Plant-based expression systems are gaining increased attention for the production of recombinant proteins, including subunit vaccines and therapeutic monoclonal antibodies (Sack et al., 2015; Group et al., 2016). This is largely driven by the potential for more rapid implementation timelines and capacity for rapid production scale-up when compared to established expression systems. The upfront capital investment for a plant-based manufacturing facility is also considerably lower than for mammalian cell production, and the production costs for the raw material are similarly reduced (Nandi et al., 2016). These features are appealing to low- and middle-income countries where biomanufacturing capacity is urgently needed, but is mostly absent (Margolin et al., 2020a). The inherent ease of biomass production also offers the prospect of rapid vaccine production in response to pandemic viral outbreaks (Margolin et al., 2018), as has recently been demonstrated by Medicago Inc., in Canada for a SARS-CoV-2 virus-like particle (VLP) vaccine candidate (Ward et al., 2020a). However, products from plant-based expression systems may not always display desired critical quality attributes: for example, when a biologic requires extensive post-translational modifications, that are not performed in plants as in other cell types, there is consequently a need to further develop plant-based expression technologies to better support the production of complex pharmaceutical products such as glycosylated biologics (Margolin et al., 2020b).

Ideally, a plant-based manufacturing platform should support high yields of recombinant proteins that are appropriately post-translationally processed and biologically active. Encouraging findings have been described for many targets, including the clinical development of influenza and SARS-CoV-2 VLP-based vaccines, and the emergency use of plant-made antibody cocktails for treatment of Ebola virus infection (Landry et al., 2010; Ward et al., 2014; Group et al., 2016; Pillet et al., 2016). However, in many other cases expression yields remain a challenge, and the plant cellular machinery may also not support extensive post-translational modifications for complex biopharmaceutical targets (Margolin et al., 2020c). Virion envelope glycoproteins are amongst the most challenging targets to produce in plants, and the inability to consistently produce native-like glycoproteins from diverse targets constrains the use of the platform for vaccine development against more complex pathogens.

The requirement for “native-like” glycosylation for appropriate immunogenicity is dictated by the antigen of interest rather than being an absolute, and it is worth noting that many targets do not require native glycosylation to elicit appropriate immunogenicity. It is also noteworthy that oligomannose-type glycans are often enriched in viral glycoproteins, such as reported for HIV-1 (Doores et al., 2010; Bonomelli et al., 2011) and SARS-CoV-2 (Watanabe et al., 2020), but are generally less abundant in plants and mammalian systems. Plant-produced influenza hemagglutinin (HA)-based VLP vaccines, for example, have been reported to be efficacious and well tolerated in phase 3 clinical trials despite the presence of typical plant-derived glycans (Ward et al., 2020b). Similarly, an insect-cell produced SARS-CoV-2 spike nanoparticle vaccine has recently been reported to be

protective against COVID-19 (Keech et al., 2020), and the presence of species-specific glycosylation does not appear to negatively impact the folding or immunogenicity of this antigen. In contrast, native-like glycosylation is widely regarded as an essential requirement for an HIV envelope-based vaccine, and subtle differences in glycosylation can dictate the quality of the immune response (Behrens and Crispin, 2017). Additionally, glycosylation is intrinsically linked to protein folding, and aberrant glycosylation and particularly under-glycosylation could compromise the efficient production of viral glycoproteins in a heterologous system (Margolin et al., 2020b).

Remodeling the secretory pathway has been proposed as a strategy to improve the production of complex biologics in plants, by providing a tailor-made environment to support their maturation (Margolin et al., 2020b,e). In the case of viral glycoproteins this would involve eliminating undesired plant-specific glycosylation and expressing the required heterologous cellular machinery to support appropriate glycosylation, folding and processing of the proteins as dictated by the antigen of interest (Meyers et al., 2018; Margolin et al., 2020c,e). Accordingly, the co-expression of human chaperone proteins was reported to improve the yields of several viral glycoproteins in plants and supported the production of antigens that could not previously be produced in the system (Meyers et al., 2018; Margolin et al., 2020c,d). Proteolytic processing of a viral glycoprotein has also been achieved—this time by the co-expression of human furin (Margolin et al., 2020c). However, similar strategies to produce viral glycoproteins with “human-like” glycosylation in plants have not been reported, although considerable progress has been made in this regard for antibodies and other biologics (Montero-Morales and Steinkellner, 2018).

The expression of glycoproteins in plants leads to the formation of three major glycoforms—namely complex, paucimannosidic and Lewis A structures—which are a feature of the expression system and distinguish plant-produced proteins from those made in conventional expression platforms (Montero-Morales and Steinkellner, 2018). Complex-type glycans containing  $\alpha$ 1,3-fucose and  $\beta$ 1,2-xylose are the most abundant glycan species, and are generated in the Golgi apparatus by  $\alpha$ 1,3-fucosyltransferase and  $\beta$ 1,2-xylosyltransferase, respectively (Wilson et al., 2001). Although these extensions do not occur in mammalian cells and might therefore be expected to appear foreign to mammalian immune systems, they do not seem to impair the immunogenicity of plant-produced influenza vaccines in humans, nor is there any evidence that they pose safety concerns in individuals with pre-existing plant allergies (Ward et al., 2014). However, the observation that these vaccines induced transient IgG and IgE responses to glyco-epitopes cannot be discounted (Ward et al., 2014) and some concerns remain, especially in the context of more heavily glycosylated proteins or where repeated administration may be necessary (Margolin et al., 2020b). Paucimannosidic glycans, comprising truncated structures where the terminal N-acetylglucosamines have been removed from the glycan core, also occur in plants (Liebminger et al., 2011). These are observed in vacuolar and extracellular glycoproteins due to processing by  $\beta$ -hexosaminidases that are localized to these sites (Shin et al., 2017). These modifications

are generally undesirable for human pharmaceuticals as they may promote rapid protein clearance following immunization if they are recognized as foreign (Yang et al., 2015). The final plant glycan structure is the Lewis A epitope which is defined by the presence of  $\beta$ 1,3-galactose and  $\alpha$ 1,4-fucose extensions of the terminal N-acetylglucosamines of the glycan (Fitchette-Laine et al., 1997; Wilson et al., 2001).

Recent studies have also suggested that certain proteins may have lower levels of glycan occupancy when produced in plants, but is presently unclear how widespread this phenomenon is (Jarczowski et al., 2016; Castilho et al., 2018; Göritz et al., 2020; Singh et al., 2020). Encouragingly, a recent study reported similar glycan occupancy for recombinant human cytomegalovirus glycoprotein B when produced in *Nicotiana tabacum* BY-2 and CHO cells. However, under-glycosylation could account for the poor yields of other glycoproteins in the system, given that inadequate glycan occupancy could impair protein folding and result in the degradation of aberrantly folded protein (Wormald and Dwek, 1999; Margolin et al., 2020b). This could also result in suboptimal immunogenicity, or the induction of antibodies against epitopes that are not accessible on the wild type virus if glycosylation occludes these regions (Zhou et al., 2017; Margolin et al., 2020b).

It is presently unclear to what extent these different plant-specific glycan structures decorate plant-produced viral glycoproteins, or if under-glycosylation occurs in this context. Therefore, in this study we aimed to define the site-specific glycan occupancy of cleavage-independent heavily glycosylated envelope glycoproteins from HIV and Marburg virus (MARV) that were produced in *Nicotiana benthamiana* plants. These examples were chosen as challenging test cases for the development of plant expression systems capable of producing humanized mimetics of these extensively post-translationally modified viral glycoproteins. A head-to-head comparison of the site-specific glycan occupancy of the proteins was performed with the equivalent mammalian cell-produced proteins. We also determined the glycosylation of plant-produced Epstein-Barr virus (EBV) gp350 and a cleaved HIV Env SOSIP.664 antigen from a previous study (Margolin et al., 2020c), to determine whether the structures observed were common to heavily glycosylated glycoproteins produced in the system.

## MATERIALS AND METHODS

### Design of Genes for Heterologous Expression in Plants and Mammalian Cells

The design and codon optimization of the synthetic gene sequences for HIV-1 CAP256 SU gp140 NFL, HIV-1 CAP256 SU SOSIP.664, EBV gp350, furin and calreticulin have been previously described (van Diepen et al., 2018; Margolin et al., 2019, 2020c). The MARV glycoprotein antigen (GP) generated in this study was based on the sequence of the Lake Victoria isolate (strain Musoke-80, UniProt accession #P35253). A soluble derivative of the antigen gene was constructed by truncating the ORF after amino acid 648 to remove the transmembrane and

cytoplasmic domains. The gene sequence was further modified by replacing the native leader sequence with the heterologous leader peptide heavy chain (LPH) leader peptide (Margolin et al., 2019) or the tissue plasminogen activator (TPA) leader sequence, and replacing the furin cleavage site with a (GGGS)<sub>2</sub> linker (van Diepen et al., 2018). A synthetic Kozak sequence (CCACC) was included at the 5' end of the gene. The MARV glycoprotein coding sequence was synthesized to reflect the preferred human codon usage.

### Construction of Expression Plasmids for Heterologous Glycoprotein Production

Recombinant protein expression was done using the pEAQ-*HT* and pTHpCapR plasmid systems for glycoprotein expression in plants and mammalian cells, respectively (Sainsbury et al., 2009; Tanzer et al., 2011; van Diepen et al., 2018; Margolin et al., 2019). A stably transfected HEK293 cell line was previously generated for production of the cleavage-independent soluble HIV Env gp140 antigen from the CAP256 SU virus (van Diepen et al., 2018). Similarly, recombinant *Agrobacterium tumefaciens* strains encoding the matched HIV Env antigen, the EBV gp350 antigen and the CAP256 SU SOSIP.664 gp140 were previously generated as part of independent studies (Margolin et al., 2019, 2020c). The MARV GP $\Delta$ TM sequence was assembled in pEAQ-*HT*. The LPH leader was replaced with the TPA leader sequence for expression in mammalian cells and the fusion gene cloned into the pTHpCapR expression plasmid. The recombinant pEAQ-*HT* expression plasmid was electroporated into *A. tumefaciens* AGL1 as described previously (Margolin et al., 2019).

### Recombinant Viral Glycoprotein Production in Plants and Mammalian Cells

All recombinant viral glycoproteins were produced in plants by agroinfiltration of *N. benthamiana* with recombinant *A. tumefaciens* bacteria as described previously (Margolin et al., 2020c). The HIV Env gp140 NFL protein was produced using a stably transfected cell line that was developed as part of an independent study (van Diepen et al., 2019). MARV GP $\Delta$ TM was expressed by transient transfection of HEK293T cells, under serum-free conditions as previously described (van Diepen et al., 2018). Expression of the MARV GP $\Delta$ TM protein was verified by western blotting of crude cell lysates using a polyclonal rabbit antibody raised against a synthetic peptide (CDQIKKDEQKEGTGW) in immunized rabbits. The recombinant proteins were purified from either cell culture media or leaf lysate following homogenization. The HIV and MARV glycoproteins were purified by sequential *Galanthus nivalis* lectin (GNL) affinity chromatography and gel filtration, as previously described, whereas the EBV gp350 $\Delta$ TM protein was purified directly by GNL-affinity chromatography (Margolin et al., 2019; van Diepen et al., 2019). Size exclusion chromatography profiles were normalized by dividing each datapoint by the peak signal and then represented as a proportion of 1. The data was presented as overlaid elution profiles using GraphPad Prism 5 software. The purified proteins were quantified using the Bio-Rad protein DC assay. The recombinant antigens were also



resolved on 4–12% NativePAGE Bis-Tris gels and stained with Bio-Safe™ Coomassie.

## Site-Specific N-Glycan Analysis of Recombinant Viral Glycoproteins

In order for each glycoprotein to be analyzed by liquid-chromatography-mass spectrometry (LC-MS), three aliquots of purified protein were denatured for 1 h in 50 mM Tris/HCl [pH 8.0], containing 6 M of urea and 5 mM of dithiothreitol (DTT). Next, the proteins were reduced and alkylated by adding 20 mM iodoacetic acid (IAA) and incubating the samples for 1 h in the dark. The samples were then incubated with DTT to remove residual IAA. The alkylated glycoproteins were buffer exchanged into 50 mM Tris/HCl [pH 8.0] and digested separately using trypsin, chymotrypsin (Mass Spectrometry Grade, Promega) and alpha-lytic protease (Sigma-Aldrich) at a ratio of 1:30 (w/w). Following overnight digestion, the resulting peptide/glycopeptides were dried and extracted using C18 Zip-tip (Merck Millipore). After elution, the peptide/glycopeptides were dried again, resuspended in 0.1% formic acid and then analyzed by nanoLC-ESI MS with an Easy-nLC 1200 (Thermo Fisher Scientific) system coupled to a Fusion mass spectrometer (Thermo Fisher Scientific) using higher energy collision-induced dissociation (HCD) fragmentation. Peptides were separated using an EasySpray PepMap RSLC C18 column (75  $\mu\text{m} \times 75 \text{ cm}$ ). A trapping column (PepMap 100 C18, 3  $\mu\text{m}$  particle size, 75  $\mu\text{m} \times 2 \text{ cm}$ ) was used in line with the LC prior to separation with the analytical column. The LC conditions were as follows: 275 min linear gradient consisting of 0–32% acetonitrile in 0.1% formic acid over 240 min, followed by 35 min of 80% acetonitrile in 0.1% formic acid. The flow rate was set to 300 nl/min. The spray voltage was set to 2.7 kV and the temperature of the heated capillary was set to 40°C. The ion transfer tube temperature was set to 275°C. The scan range was 400–1600 m/z. The HCD collision energy was set to 50%, appropriate for fragmentation of glycopeptide ions. Precursor and fragment detection were performed using an Orbitrap at a resolution MS1 = 100,000. MS2 = 30,000. The automatic gain control (AGC) target for MS1 = 4e5 and MS2 = 5e4 and injection time: MS1 = 50 ms MS2 = 54 ms.

Glycopeptide fragmentation data were extracted from the raw file using Byonic™ and Byologic™ software (Version 3.5; Protein Metrics Inc.). The glycopeptide fragmentation data were evaluated manually for each glycopeptide; the peptide was scored as true-positive when the correct b and y fragment ions were observed along with oxonium ions corresponding to the glycan identified. The MS data was searched using the Protein Metrics' N-glycan libraries: for mammalian expression the 309 N-glycan library was used and for material produced in plants the 52 plant library was used. The relative amounts of glycan at each site, as well as the unoccupied proportion, were determined by comparing the extracted chromatographic areas for different glycotypes with an identical peptide sequence. All charge states for a single glycopeptide were summed. The precursor mass tolerance was set at 4 ppm and 10 ppm for fragments. A 1% false discovery rate (FDR) was applied. The relative amounts of each glycoform at each site, as well as

the unoccupied proportion, were determined by comparing the extracted ion chromatographic areas for different glycopeptides with an identical peptide sequence. Glycans are categorized according to the detected compositions by LC-MS. Compositions containing Hex(10-12) HexNAc(2) were categorized as M9Glc1-3 and Hex(9-5) HexNAc(2) as M9-M5. These are classified as oligomannose-type glycans and are colored green. Hybrid-type glycans contain compositions consisting of Hex(5)HexNAc(3)X or Hex(6)HexNAc(4)X. Remaining glycan compositions were assigned as complex-type. Any composition Hex(3) HexNAc(2) or smaller is classified as truncated. The proportion of unoccupied N-linked glycan sites at each site are colored gray.

## RESULTS

### Viral Glycoprotein Selection and Antigen Design

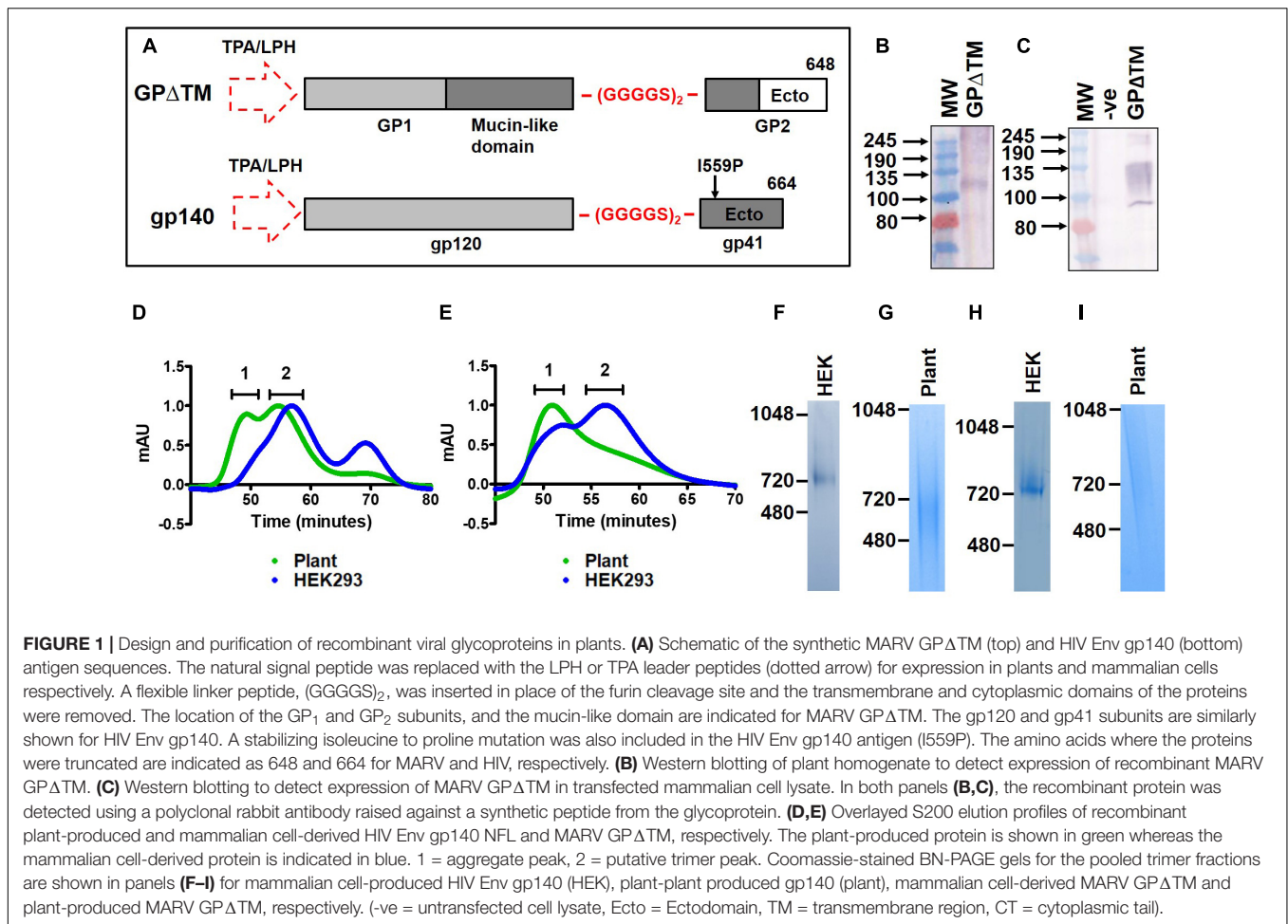
Viral envelope glycoproteins from HIV-1, EBV, and MARV were selected for this study as they are all extensively glycosylated and are therefore ideal models to interrogate the impact of the plant cellular machinery on glycosylation. Both HIV gp140 and EBV gp350 $\Delta\text{TM}$  were successfully expressed in plants as part of a previous study exploring the impact of co-expressing human chaperones on viral glycoprotein production in *N. benthamiana* (Margolin et al., 2020c). MARV is a re-emerging virus with a high probability of causing disease outbreaks in humans, prompting the suitability of plant-based expression to be explored for the development of a vaccine against a prototype emerging filovirus (Letko et al., 2020). MARV and HIV glycoproteins were both expressed as soluble cleavage-independent antigens for ease of production and recovery (Figure 1A) using previously described approaches (Margolin et al., 2019, 2020c; van Diepen et al., 2019).

### Transient Expression of MARV GP $\Delta\text{TM}$ in Plants and Mammalian Cells

Marburg virus GP $\Delta\text{TM}$  was transiently expressed in plants and mammalian cells by agroinfiltration and transfection, respectively. Western blotting of crude extract from agroinfiltrated leaves yielded a product of ~115–125 kDa and some diffuse higher molecular weight products which were not well resolved by SDS-PAGE (Figure 1B). Western blotting of lysate from transfected HEK293 cells produced a similar product, although the protein appeared slightly larger and the signal was fairly diffuse (Figure 1C). An additional product of <100 kDa was observed for the mammalian cell-produced protein which may have arisen from intracellular processing. Expression of the recombinant HIV Env gp140 antigens and EBV g350 $\Delta\text{TM}$  were both previously described (Margolin et al., 2019, 2020c; van Diepen et al., 2019).

### Purification of Soluble MARV and HIV Glycoproteins

The soluble glycoproteins (HIV Env gp140 and MARV GP $\Delta\text{TM}$ ) were captured by GNL affinity chromatography and then fractionated by gel filtration, to recover trimeric glycoproteins

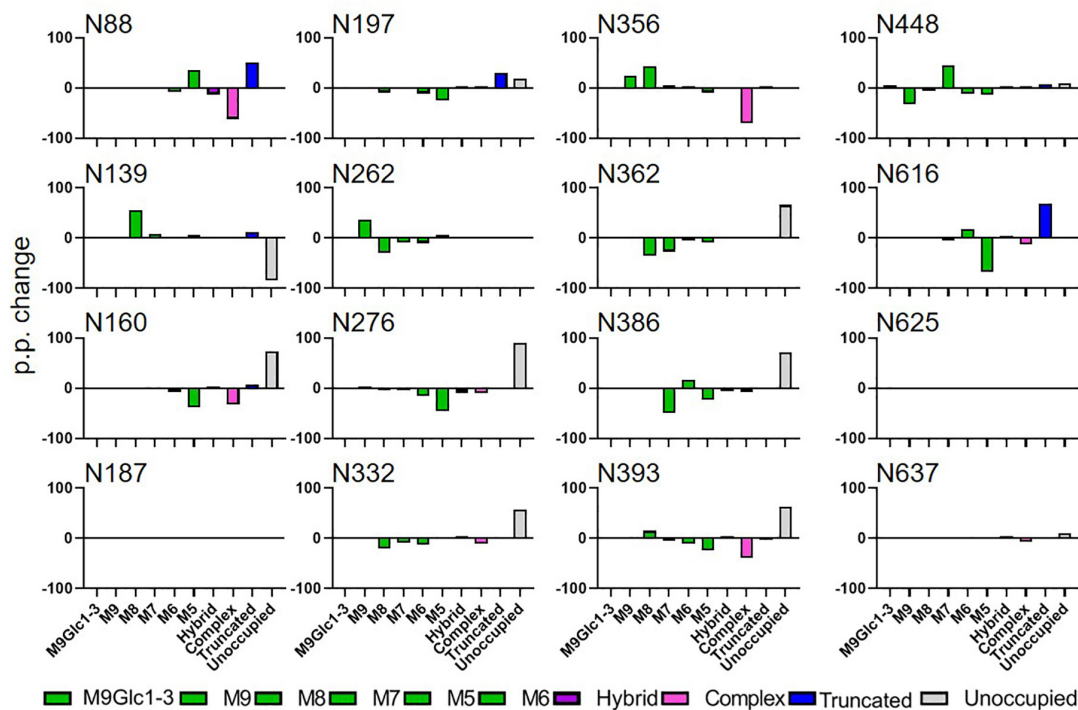


and to exclude non-trimeric protein species (Figures 1D,E). The size exclusion chromatography elution profiles for the plant-produced and mammalian cell-derived glycoproteins were overlaid to compare the relative abundance of the different protein species in both systems. Both glycoproteins exhibited a marked increase in higher molecular weight aggregates when produced in plants as reflected in the profile shifting toward the left (Figures 1D,E). While the plant-produced HIV Env gp140 exhibited a peak of the expected size for trimers (Figure 1D), the plant-produced MARV GP $\Delta$ TM did not yield a defined peak as expected (Figure 1E). Instead, a diffuse shoulder was observed suggesting that the purified protein was highly heterogenous, even after purification.

Coomassie-stained BN-PAGE gels mirrored the observations from size exclusion chromatography. The pooled trimer peak derived from mammalian cells yielded a defined band of ~720 kDa for both HIV Env gp140 (Figure 1F) and MARV GP $\Delta$ TM (Figure 1H), which is consistent with our previous observations for HIV (van Diepen et al., 2019) and published accounts for filovirus glycoprotein trimers (Rutten et al., 2020). In contrast, Coomassie staining of the pooled fractions corresponding to the plant-derived glycoprotein trimers yielded a diffuse signal (Figures 1G,I).

## Aberrant Glycosylation of HIV Env Using *N. benthamiana* as a Production System

To determine the site-specific changes in glycosylation of HIV-1 Env when using *N. benthamiana* as the expression system we applied LC-MS to the purified proteins (Figure 2). The data is displayed as the percentage point change in site-specific glycosylation when the protein is expressed using *N. benthamiana*, so a positive value represents a category which is more abundant in *N. benthamiana* compared to mammalian produced protein. This methodology has been used previously to compare and contrast the glycosylation of the viral glycoprotein from BG505 and other HIV strains (Behrens et al., 2016; Struwe et al., 2018). There are distinct glycan processing states on HIV Env which form key epitopes for broadly neutralizing antibodies. The early glycan processing pathway in both mammals and plants is conserved and involves the attachment and subsequent trimming of oligomannose-type glycans in the ER and early Golgi apparatus (Montero-Morales and Steinkellner, 2018). The density of N-linked glycans on HIV Env perturb this process, and the resultant mature protein contains remnants of these under-processed glycans. These oligomannose-type glycans can form epitopes for broadly neutralizing antibodies that target several regions of the Env spike. Two key regions for antibody



**FIGURE 2 |** Comparison of site-specific glycosylation of HIV Env gp140 NFL when produce in plants or mammalian cells. The data represent the percentage point change in glycosylation when the protein was expressed in *N. benthamiana* compared to HEK293, so a positive value represents a category of glycan that is enriched on material produced in *N. benthamiana*. Glycans are categorized according to the detected compositions by LC-MS. Compositions containing Hex(10-12) HexNAc(2) were categorized as M9Glc1-3 and Hex(9-5) HexNAc(2) as M9-M5. These are classified as oligomannose-type glycans and are colored green. Hybrid-type glycans contain compositions consisting of Hex(5)HexNAc(3)X or Hex(6)HexNAc(4)X. Remaining glycan compositions were assigned as complex-type. Any composition Hex(3) HexNAc(2) or smaller is classified as truncated. The proportion of potential N-linked glycans which are classified as unoccupied are coloured grey.

recognition are the intrinsic mannose patch (IMP), focused around the N332 supersite of vulnerability, and the trimer-associated patch (TAMP), with the N160 glycans of the V1/V2 region forming a key part of several antibody epitopes (Behrens and Crispin, 2017). The TAMP is only present when correctly processed and cleaved Env is analyzed, and forms at the trimer interface and the apex of the protein (Behrens et al., 2017). These two key regions of oligomannose-type glycans are key for immunogen design efforts (Haynes et al., 2019).

When comparing the glycosylation of CAP256 SU gp140 expressed in mammalian and plant cells the distribution and processing of these oligomannose-type glycans is affected. Sites that form the IMP such as N332 and N386 show a decrease in oligomannose-type glycans (40 and 55 percentage points, respectively) when *N. benthamiana* is used as the expression system (Supplementary Table 1). The TAMP is also disrupted with a similar reduction in oligomannose-type glycans at N160 and N197. Oligomannose-type glycans increase in abundance at N356, however this site appears more fully processed in HEK293 cells.

The reason for the decrease in oligomannose-type glycans at key bnAb recognition sites is the near universal increase in potential N-linked glycosylation sites that lack post-translational modifications. Sites N160, N197, N276, N332, N362, N386,

N393, N447, and N637 all show an increase in the proportion of unoccupied glycan sites. This represents an asparagine of an NxS/T sequon which has the capacity for the addition of an N-linked glycan, but where the glycosylation machinery has not attached a glycan. Recombinant expression systems are suboptimal for recapitulating the high occupancy of viral-derived Env, especially on gp41, and this has also been observed in mammalian expression systems, albeit to a lesser extent (Derking et al., 2021).

This can be seen on CAP256 SU gp140 expressed in both HEK293 and *N. benthamiana*, for example at N625 (Supplementary Table 1). The presence of these so-called glycan holes has been shown to induce non-neutralizing antibody responses during immunizations which do not protect from HIV acquisition (Crooks et al., 2015; Zhou et al., 2017). The CAP256 SU gp140 expressed using *N. benthamiana* contains the same glycan holes as with HEK293 cells, however, the presence of additional holes at sites such as N332 will likely generate detrimental antibody responses when used in immunizations. Finally, at three sites on CAP256 SU gp140 from *N. benthamiana*, N88, N197 and N616, a substantial increase in smaller glycan structures were observed (ranging from a single hexosamine to  $\text{Man}_3\text{GlcNAc}_2$ ) which were not observed on mammalian proteins.

**TABLE 1** | Site-specific glycosylation of plant produced CAP256 SU gp140 SOSIP.664 and EBV gp350ΔTM.**CAP256 SU SOSIP.664**

	85	130	139	160	187	197	230	234	241	262	276	289	295	301	332	356	362	386	393	406	411	448	462	611	616	625	637	Average
M9GIC1-3	0	0	0	0	0	0				0	0				0	8	0	0	0			0	33	0	0	0	0	2
M9	42	0	26	0	0	0				35	0				50	2	0	0	8			1	0	0	0	0	9	
M8	12	18	42	16	0	0				44	0				28	61	26	31	12			23	0	0	0	1	17	
M7	26	7	0	2	0	0				14	0				7	12	6	24	10			18	0	0	0	0	7	
M6	0	0	0	0	0	0				2	0				6	2	2	0	3			6	0	0	0	0	1	
M5	2	0	0	2	0	0				2	0				3	6	1	6	1			5	0	0	0	0	2	
Hybrid	0	0	0	0	0	0				0	0				0	0	0	0	0			0	0	0	0	0	0	
Complex	0	0	0	0	0	1				0	0				0	2	1	0	0			1	0	0	0	0	0	
Truncated	0	3	18	3	0	25				0	0				3	0	0	0	1			5	65	0	0	0	7	
Unoccupied	0	71	14	76	100	75				0	99				2	0	60	39	65			42	2	0	99	99	47	

**EBV gp350ΔTM**

	66	106	133	185	188	214	248	196	337	347	364	375	397	405	430	454	462	476	516	538	552	566	587	608	629	643	646	664	675	702	720	754	765	774	799	834	Average
M9GIC1-3	0	0	0	0	0	0	1	0	0	0	0	0	0	0	83	2	6	3		0	0	0	3	2	0		0	0	0	0	0	0	0	1	23	4	
M9	28	6	5	7	0	0	29	13	55	3	3	0	15	9	4	14	13		4	0	0	5	3	0		7	0	0	0	0	0	0	9	48	9		
M8	65	50	31	42	21	2	32	55	29	64	16	0	42	5	19	29	30		27	37	2	13	8	28		32	51	36	43	0	0	0	14	23	27		
M7	7	11	18	11	19	1	9	18	5	0	7	0	13	1	11	10	24		0	15	2	10	8	13		13	49	11	12	0	86	0	10	6	12		
M6	0	1	15	5	2	0	7	4	5	0	2	0	7	1	0	3	7		0	0	0	2	2	0		4	0	0	0	0	1	0	2	0	2		
M5	0	5	18	6	10	0	16	5	6	21	4	0	13	0	1	0	12		0	0	0	3	3	0		5	0	0	0	0	0	0	3	0	4		
Hybrid	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0		0	0	29	0	0	0		0	0	0	0	55	0	0	0	0	3		
Complex	0	3	4	7	5	0	2	1	0	0	2	0	4	0	4	4	11		1	0	48	3	2	13		2	0	13	0	42	0	0	1	0	5		
Truncated	0	14	3	11	42	0	4	2	0	1	30	0	2	0	45	35	0		62	48	1	38	48	46		37	0	40	44	2	8	6	59	0	20		
Unoccupied	0	10	6	12	0	96	1	1	0	11	36	99	4	1	15	0	0		6	0	0	23	24	0		1	0	0	0	0	5	94	0	0	14		

Different tints of blue, from dark through to light, indicate the highest through to the lowest numbers in the table.



## Incomplete Processing of Complex Glycans on MARV GP $\Delta$ TM Expressed in *N. benthamiana* Compared to Mammalian Cells

To investigate whether the observed changes in glycosylation were specific to HIV Env, we performed a comparative analysis of MARV GP $\Delta$ TM produced in HEK293 and *N. benthamiana*. Data of sufficient quality could only be obtained for 8 sites with the full list of sites shown in **Supplementary Table 2**. As with HIV Env, the site-specific glycosylation of MARV GP $\Delta$ TM varied extensively between HEK293 and *N. benthamiana* (**Figure 3**). For the HEK293 cell-produced protein, several sites displayed extensive glycan processing typical of mammalian glycoproteins. Interestingly, a high proportions of hybrid-type glycans were observed, which are named as such because they contain one arm which is fully processed and one arm presenting terminal mannose residues. Importantly, both the hybrid and complex-type glycans observed on HEK293 MARV GP were not present on material produced in *N. benthamiana*. This includes mature glycans modified with xylose which were included in the library used to search the LC-MS data. Instead, there was a global increase of oligomannose-type glycans, for example at N96 in mammalian cells where 41% of the glycans were oligomannose-type. In contrast this increased to 71% when produced in plants, at 30 percentage point increase (**Figure 3** and **Supplementary Table 2**). The presence of elevated amounts of pauciglycan structures was also observed at five sites, with pauciglycans consisting of 20% of the averaged compositions of all sites obtained by LC-MS. Sites N173 and N572 also lacked glycans on 70 and 100% of the sites, whereas for HEK293 these sites were fully occupied. The higher levels of oligomannose-type glycans observed for MARV GP $\Delta$ TM is likely due to the cell line used for protein production. In the case of HIV Env, the presence of oligomannose-type glycans is due to steric clashes inhibiting the ability of glycan processing enzymes whereas the lack of oligomannose-type glycans on MARV GP $\Delta$ TM suggests a differential mechanism. The glycan processing machinery in plants may be unable to process such extensive glycosylation to the same extent as HEK293 cells.

In order to determine if this glycosylation signature was specific to the proteins in question, or rather a feature of producing heavily glycosylated glycoproteins in plants, the site-specific glycosylation of previously produced EBV gp350 and CAP256 SU gp140 SOSIP.664 (Margolin et al., 2020c) were also determined (**Table 1**). Both antigens displayed a similar glycosylation profile comprising of large amounts of oligomannose-type glycans, low levels of complex glycans and considerable amounts of paucimannosidic glycans. Similar to the plant-produced MARV GP $\Delta$ TM and HIV Env gp140 antigens, a large amount of under-glycosylation was observed. Interestingly for EBV gp350 and CAP256 SU gp140 SOSIP produced in plants, sites displayed glycosylated Man<sub>9</sub>GlcNAc<sub>2</sub> (categorized as M9Glc1-3 in **Table 1**). This precursor oligosaccharide is present during the early ER folding stages and the removal of the glucose monosaccharides acts as a checkpoint to signal that the protein is correctly folded. The presence of these glycans

in the resultant purified protein is unusual, given that 0% of the glycans from both CAP256 SU gp140 and MARV GP $\Delta$ TM produced in HEK293 cells contained these moieties. These observations are consistent with the increase in oligomannose-type glycans observed on MARV GP $\Delta$ TM suggesting that immaturely processed glycoproteins are being released from the ER during homogenization of plant material and purified.

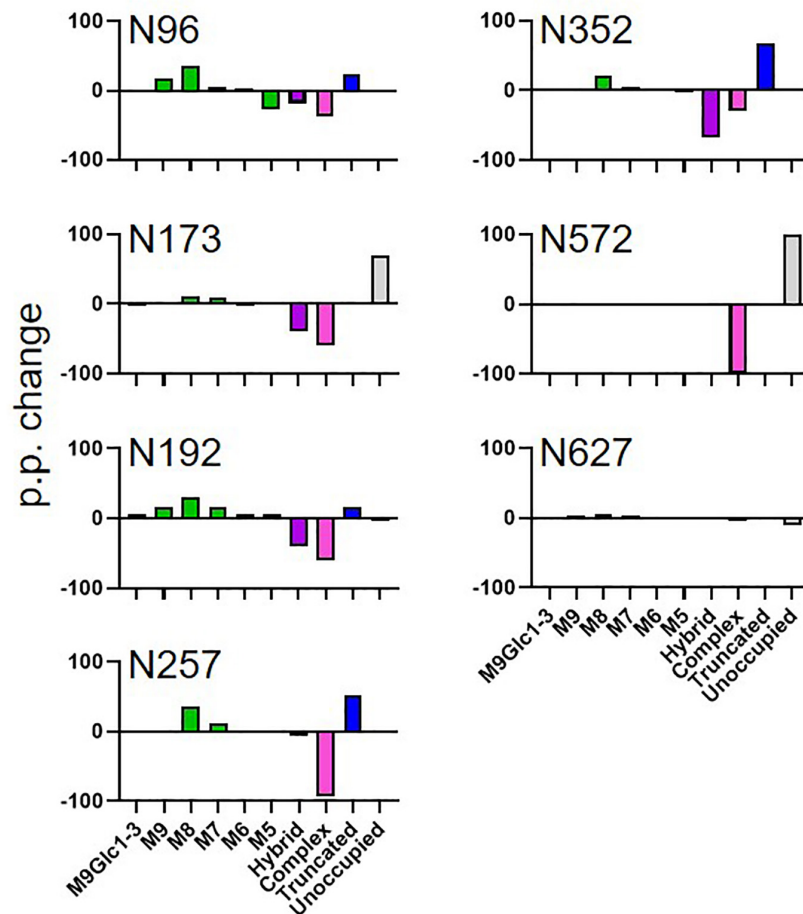
## DISCUSSION

Viral glycoprotein-based subunit vaccines are amongst the most complex recombinant proteins that are being pursued in plant expression systems (Margolin et al., 2018). These proteins have intricate architectures with extensive post-translational modifications, and are heavily reliant on the host cellular machinery to support their maturation. However, some of these maturation events may not occur optimally in plants, necessitating molecular engineering of the host machinery to support the synthesis of these proteins (Margolin et al., 2020b,e). Approaches that have been successfully implemented to increase glycoprotein yields are co-expression of human chaperones, and ectopic expression of the protease furin to support proteolytic processing in plants (Margolin et al., 2020c). Considerable success in this regard has been reported for the production of antibodies and other less complex glycoproteins (Montero-Morales and Steinkellner, 2018), but it is apparent that continued development of plant-based expression systems will be required for the production of complex viral glycoproteins that recapitulate the post-translational processing of mammalian production systems.

Host-derived glycosylation is central to glycoprotein folding and regulates the interaction of the proteins with ER-resident chaperones (Watanabe et al., 2019). In addition, these glycans also determine the immunogenicity of the glycoprotein in natural infection or following vaccination (Dalziel et al., 2014; Watanabe et al., 2019). Therefore, it may be necessary to reproduce the native glycosylation of many viral glycoproteins to elicit appropriate immune responses in the context of vaccination. As a starting point to develop approaches to produce natively glycosylated viral glycoproteins in plants, we interrogated the site-specific glycosylation of recombinant plant-produced MARV and HIV Env glycoproteins and compared them to the equivalent proteins produced in mammalian cells. We also determined the glycosylation of EBV gp350 $\Delta$ TM and HIV Env gp140 SOSIP.664 produced in plants in a previous study (Margolin et al., 2020c).

Building on previous work to express HIV Env gp140 in plants and mammalian cells (van Diepen et al., 2018, 2019; Margolin et al., 2019, 2020c), we produced a soluble MARV glycoprotein in both systems. This is to our knowledge the first report describing the expression of this filoviral glycoprotein in a plant expression system, and adds to an increasing number of viral glycoproteins that have been expressed in plants. The HIV and MARV glycoproteins were purified by sequential affinity chromatography and gel filtration steps, and the size exclusion elution profiles were overlaid with the equivalent mammalian cell-produced proteins for comparison. Whilst the HIV Env gp140 antigen produced a peak that is consistent with the





**FIGURE 3** | Comparison of MARV GP $\Delta$ TM glycosylation when the protein was expressed in plants and mammalian cells. The data represent the percentage point change in glycosylation when the protein was expressed in *N. benthamiana* compared to HEK293, so a positive value represents a category of glycan that is enriched on material produced in *N. benthamiana*. Glycans are categorized according to the detected compositions by LC-MS. Compositions containing Hex(10-12) HexNAc(2) were categorized as M9Glc1-3 and Hex(9-5) HexNAc(2) as M9-M5. These are classified as oligomannose-type glycans and are colored green. Hybrid-type glycans contain compositions consisting of Hex(5)HexNAc(3)X or Hex(6)HexNAc(4)X. Remaining glycan compositions were assigned as complex-type. Any composition Hex(3) HexNAc(2) or smaller is classified as truncated. The proportion of potential N-linked glycans which are classified as unoccupied are coloured grey.

expected size for trimeric Env, the MARV GP $\Delta$ TM protein yielded a diffuse shoulder, suggesting high levels of heterogeneity for the recombinant protein. In both cases, there was a striking shift of the plant-produced protein toward the left of the profile indicating an increase in size. This corresponds to increased aggregation in the system, which is consistent with previous reports describing unresolved higher molecular weight products following the expression of HIV Env gp140 antigens in plants (Rosenberg et al., 2013; Margolin et al., 2019).

Given the extensive under-glycosylation of the protein observed in this study it is likely that these aggregates comprise of aberrantly folded protein species. This is also supported by the presence of glucosylated mannose species that were observed. Whilst several studies have reported under-glycosylation in a range of plant-produced proteins (Jarczowski et al., 2016; Castilho et al., 2018; Singh et al., 2020), this is the most extensive under-glycosylation that has been reported to date. This

observation may account for the inferior immunogenicity of the plant-produced HIV gp140 antigen compared to the mammalian cell-produced protein that was observed in a previous study, where the trimer elicited high titers of binding antibodies in immunized rabbits but did not induce any appreciable autologous neutralizing antibodies (Margolin et al., 2019). If these antibodies were raised against epitopes that were shielded by glycans on the wild type virus, it would explain why they failed to neutralize. It is well established that antibodies targeting holes in the glycan shield are readily induced against the HIV Env glycoprotein, and therefore reproducing the glycan shield on trimer immunogens is critical to prevent undesired off-target immune responses (Zhou et al., 2017; Crispin et al., 2018; Ringe et al., 2019). Glycans can also constitute important epitopes targeted by neutralizing antibodies, and in the absence of glycosylation these epitopes would not be reproduced (Crispin et al., 2018). The N160 glycan at the apex of the trimer,

for example, is almost completely unoccupied in the plant-produced antigen, and therefore the protein is not expected to exhibit appreciable binding by PG-9 and other related antibodies that recognize this epitope (McLellan et al., 2011). Under-glycosylation was similarly observed at the N332 glycan site which is commonly targeted by broadly neutralizing antibodies during natural infection (Sok et al., 2016). Future work should compare the reactivity of plant-produced and mammalian cell-derived proteins with prototype monoclonal antibodies from natural infection. These antibodies could serve as useful tools to further investigate how closely these plant-produced antigens resemble their mammalian cell-produced counterparts, and will help establish if host engineering approaches improve their folding and antigenicity.

Based on the work described here, and in other recent studies of plant-produced enzymes and antibodies, it seems likely that under-glycosylation of heterologous proteins may be more common than was previously appreciated (Castilho et al., 2018; Göritzer et al., 2020). Given the central role of glycosylation in protein folding, this could be an important contributory factor to inefficient production of certain viral glycoproteins in the system, and could be expected to pose a similar challenge for other similarly glycosylated target proteins (Margolin et al., 2018). The molecular basis for under-glycosylation in the system is not well understood but it has been suggested that this is related to the distinct recognition preferences of the plant oligosaccharyltransferase complex (Margolin et al., 2020e). It is also probably a consequence of the purification process where homogenization of plant tissue will liberate proteins from all stages of the secretory pathway, including those that have not yet been properly glycosylated. Encouragingly, it has recently been shown that the co-expression of *Leishmania major* LmSTT3D increased the glycan occupancy of several model proteins in *N. benthamiana* (Castilho et al., 2018; Göritzer et al., 2020).

The almost complete absence of plant-specific complex glycans in the purified proteins is an unexpected observation, and contrasts with other reports describing plant-produced viral glycoproteins from influenza virus and HIV-1 (Rosenberg et al., 2013; Le Mauff et al., 2015). However, it is difficult to compare these to the current study due to differences in glycoprotein complexity and different antigen design strategies: both glycoproteins produced in this study contain considerably more N-glycan sequons than influenza HA which only contains six N-glycan sites, and therefore the burden on the host cellular machinery is expected to be commensurately greater (Le Mauff et al., 2015). The HIV-1 Env trimer described in this study was also engineered to preserve the structure, whereas the previously described glycoprotein was modified to remove the cleavage site, fusion peptide and immunodominant region of gp41 (Rosenberg et al., 2013). These alterations are expected to have impacted protein folding, and based on subsequent insights into the protein structure, would be likely to have compromised the folding of the glycoprotein (Ringe et al., 2013). It is also plausible that the processing by host mannosidases could be less efficient in plants resulting in poor formation of complex glycans. These enzymes may be less abundant in plants than mammalian cells as the requirement for glycosylation is often considerably less.

Although the levels of truncated and paucimannose-type glycans following expression in plants is surprising, it is not unprecedented. Similar observations have been reported for other heterologous plant-produced proteins, such as  $\alpha$ 1-antitrypsin (Castilho et al., 2014), IgA (Dicker et al., 2016), and bovine follicle stimulation hormone (Dirnberger et al., 2001). These truncated glycans arise from the removal of *N*-acetylglucosamines from the non-reducing end by  $\beta$ -hexosaminidases (HEXOs) (Liebminger et al., 2011). The responsible enzymes are localized to vacuoles and the plasma membrane in *N. benthamiana*, and previous work has shown that their downregulation by RNA interference increased the levels of intact complex-type glycans (Shin et al., 2017). These structures do not naturally occur on viral glycoproteins and therefore this approach will be necessary to eliminate this undesirable processing *in planta*.

In conclusion, we have delineated the site-specific glycosylation of several model human viral glycoproteins that were produced in *N. benthamiana*, and identified key constraints in the host glycosylation machinery. This work adds to a growing body of evidence suggesting that remodeling the secretory pathway may be necessary to support the production of complex pharmaceutical targets in the plant expression system, and provides a rational starting point to produce recombinant glycoproteins with humanized glycosylation in plants (Margolin et al., 2020e). The data presented here has important implications for plant molecular farming of viral glycoproteins and suggests that under-glycosylation may be more widespread than previously realized.

## DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Material**, further inquiries can be directed to the corresponding author/s.

## AUTHOR CONTRIBUTIONS

EM, JA, and MC conceptualized the study. EM and RC designed the gene sequences and conceived the cloning strategy. EM and MV conducted protein expression in plants. EM, MD, and PX conducted protein expression in mammalian cells. JA carried out the site-specific glycosylation. A-LW and RC supervised mammalian cell culture aspects of the project. ER and AM supervised plant-based protein expression component of the project. EM and JA drafted the manuscript. Funding for the project in Cape Town was obtained by EM, A-LW and ER. All authors contributed to data analysis and reviewed the manuscript.

## FUNDING

This work is based upon research supported by the South African Medical Research Council with funds received from the South African Department of Science and Technology, the

South African Research Chairs Initiative of the Department of Science and Technology, and the National Research Foundation (Grant number: 64819). Further funding support was provided by core funding from the Wellcome Trust (203135/Z/16/Z). The work in the laboratory of MC was supported by the International AIDS Vaccine Initiative (IAVI) through grant OPP1153692/INV-008352 funded by the Bill & Melinda Gates Foundation.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpls.2021.709344/full#supplementary-material>

## REFERENCES

- Behrens, A. J., and Crispin, M. (2017). Structural principles controlling HIV envelope glycosylation. *Curr. Opin. Struct. Biol.* 44, 125–133. doi: 10.1016/j.sbi.2017.03.008
- Behrens, A. J., Harvey, D. J., Milne, E., Cupo, A., Kumar, A., Zitzmann, N., et al. (2017). Molecular architecture of the cleavage-dependent mannose patch on a soluble HIV-1 envelope glycoprotein trimer. *J. Virol.* 91:e1894–16.
- Behrens, A. J., Vasiljevic, S., Pritchard, L. K., Harvey, D. J., Andev, R. S., Krumm, S. A., et al. (2016). Composition and antigenic effects of individual glycan sites of a trimeric HIV-1 envelope glycoprotein. *Cell Rep.* 14, 2695–2706. doi: 10.1016/j.celrep.2016.02.058
- Bonomelli, C., Doores, K. J., Dunlop, D. C., Thaney, V., Dwek, R. A., Burton, D. R., et al. (2011). The glycan shield of HIV is predominantly oligomannose independently of production system or viral clade. *PLoS One* 6:e23521. doi: 10.1371/journal.pone.0023521
- Castilho, A., Beihammer, G., Pfeiffer, C., Goritzer, K., Montero-Morales, L., Vavra, U., et al. (2018). An oligosaccharyltransferase from *Leishmania major* increases the N-glycan occupancy on recombinant glycoproteins produced in *Nicotiana benthamiana*. *Plant Biotechnol. J.* 16, 1700–1709. doi: 10.1111/pbi.12906
- Castilho, A., Windwarder, M., Gattinger, P., Mach, L., Strasser, R., Altmann, F., et al. (2014). Proteolytic and N-glycan processing of human alpha1-antitrypsin expressed in *Nicotiana benthamiana*. *Plant Physiol.* 166, 1839–1851. doi: 10.1104/pp.114.250720
- Crispin, M., Ward, A. B., and Wilson, I. A. (2018). Structure and immune recognition of the HIV glycan shield. *Annu. Rev. Biophys.* 47, 499–523. doi: 10.1146/annurev-biophys-060414-034156
- Crooks, E. T., Tong, T., Chakrabarti, B., Narayan, K., Georgiev, I. S., Menis, S., et al. (2015). Vaccine-elicited Tier 2 HIV-1 neutralizing antibodies bind to quaternary epitopes involving glycan-deficient patches proximal to the CD4 binding site. *PLoS Pathog.* 11:e1004932. doi: 10.1371/journal.ppat.1004932
- Dalziel, M., Crispin, M., Scanlan, C. N., Zitzmann, N., and Dwek, R. A. (2014). Emerging principles for the therapeutic exploitation of glycosylation. *Science* 343:1235681. doi: 10.1126/science.1235681
- Derking, R., Allen, J. D., Cottrell, C. A., Sliopen, K., Seabright, G. E., Lee, W. H., et al. (2021). Enhancing glycan occupancy of soluble HIV-1 envelope trimers to mimic the native viral spike. *Cell Rep.* 35:108933. doi: 10.1016/j.celrep.2021.108933
- Dicker, M., Tschofen, M., Maresch, D., Konig, J., Juarez, P., Orzaez, D., et al. (2016). Transient glyco-engineering to produce recombinant IgA1 with defined N- and O-glycans in plants. *Front. Plant Sci.* 7:18. doi: 10.3389/fpls.2016.00018
- Dirnberger, D., Steinkellner, H., Abdennebi, L., Remy, J. J., and van de Wiel, D. (2001). Secretion of biologically active glycoforms of bovine follicle stimulating hormone in plants. *Eur. J. Biochem.* 268, 4570–4579. doi: 10.1046/j.1432-1327.2001.02384.x
- Doores, K. J., Bonomelli, C., Harvey, D. J., Vasiljevic, S., Dwek, R. A., Burton, D. R., et al. (2010). Envelope glycans of immunodeficiency virions are almost entirely oligomannose antigens. *Proc. Natl. Acad. Sci. U.S.A.* 107, 13800–13805. doi: 10.1073/pnas.1006498107
- Fitchette-Laine, A. C., Gomord, V., Cabanes, M., Michalski, J. C., Saint Macary, M., Foucher, B., et al. (1997). N-glycans harboring the Lewis a epitope are expressed at the surface of plant cells. *Plant J.* 12, 1411–1417. doi: 10.1046/j.1365-313x.1997.12061411.x
- Göritzer, K., Goet, I., Duric, S., Maresch, D., Altmann, F., Obinger, C., et al. (2020). Efficient N-glycosylation of the heavy chain tailpiece promotes the formation of plant-produced dimeric IgA. *Front. Chem.* 8:346. doi: 10.3389/fchem.2020.00346
- Group, P. I. W., Multi-National, P. I. I. S. T., Davey, R. T. Jr., Dodd, L., Proschan, M. A., Neaton, J., et al. (2016). A randomized, controlled trial of ZMapp for ebola virus infection. *N. Engl. J. Med.* 375, 1448–1456. doi: 10.1056/nejmoa1604330
- Haynes, B. F., Burton, D. R., and Mascola, J. R. (2019). Multiple roles for HIV broadly neutralizing antibodies. *Sci. Transl. Med.* 11:eaa2686. doi: 10.1126/scitranslmed.aaz2686
- Jarczowski, F., Kandzia, R., Thieme, F., Klimyuk, V., and Gleba, Y. (2016). *Methods of Modulating N-Glycosylation site Occupancy of Plant-Produced Glycoproteins and Recombinant Glycoproteins*. US Patent 2016/0115498 A1. Washington, DC: U.S. Patent and Trademark Office.
- Keech, C., Albert, G., Cho, L., Robertson, A., Reed, P., Neal, S., et al. (2020). Phase 1-2 trial of a SARS-CoV-2 recombinant spike protein nanoparticle vaccine. *N. Engl. J. Med.* 383, 2320–2332.
- Landry, N., Ward, B. J., Trepanier, S., Montomoli, E., Dargis, M., Lapini, G., et al. (2010). Preclinical and clinical development of plant-made virus-like particle vaccine against avian H5N1 influenza. *PLoS One* 5:e15559. doi: 10.1371/journal.pone.0015559
- Le Mauff, F., Mercier, G., Chan, P., Burel, C., Vaudry, D., Bardor, M., et al. (2015). Biochemical composition of haemagglutinin-based influenza virus-like particle vaccine produced by transient expression in tobacco plants. *Plant Biotechnol. J.* 13, 717–725. doi: 10.1111/pbi.12301
- Letko, M., Seifert, S. N., Olival, K. J., Plowright, R. K., and Munster, V. J. (2020). Bat-borne virus diversity, spillover and emergence. *Nat. Rev. Microbiol.* 18, 461–471. doi: 10.1038/s41579-020-0394-z
- Liebming, E., Veit, C., Pabst, M., Batoux, M., Zipfel, C., Altmann, F., et al. (2011). Beta-N-acetylhexosaminidases HEXO1 and HEXO3 are responsible for the formation of paucimannosidic N-glycans in *Arabidopsis thaliana*. *J. Biol. Chem.* 286, 10793–10802. doi: 10.1074/jbc.m110.178020
- Margolin, E., Burgers, W. A., Sturrock, Mendelson, M., Chapman, R., Douglass, N., et al. (2020a). Prospects for SARS-CoV-2 diagnostics, therapeutics and vaccines in Africa. *Nat. Rev. Microbiol.* 18, 690–704. doi: 10.1038/s41579-020-00441-3
- Margolin, E., Chapman, R., Meyers, A. E., van Diepen, M. T., Kimba, P., Hermanus, T., et al. (2019). production and immunogenicity of soluble plant-produced HIV-1 Subtype C Envelope gp140 Immunogens. *Front. Plant Sci.* 10:1378. doi: 10.3389/fpls.2019.01378

- Margolin, E., Chapman, R., Williamson, A. L., Rybicki, E. P., and Meyers, A. E. (2018). Production of complex viral glycoproteins in plants as vaccine immunogens. *Plant Biotechnol. J.* 19, 1531–1545. doi: 10.1111/pbi.12963
- Margolin, E., Crispin, M., Meyers, A., Chapman, R., and Rybicki, E. P. (2020b). A roadmap for the molecular farming of viral glycoprotein vaccines: engineering glycosylation and glycosylation-directed folding. *Front. Plant Sci.* 11:609207. doi: 10.3389/fpls.2020.609207
- Margolin, E., Oh, Y. J., Verbeek, M., Naude, J., Ponnendorf, D., Meshcheriakova, Y. A., et al. (2020c). Co-expression of human calreticulin significantly improves the production of HIV gp140 and other viral glycoproteins in plants. *Plant Biotechnol. J.* 18, 2109–2117. doi: 10.1111/pbi.13369
- Margolin, E., Verbeek, M., Meyers, A., Chapman, R., Williamson, A. L., and Rybicki, E. P. (2020d). Calreticulin co-expression supports high level production of a recombinant SARS-CoV-2 spike mimetic in *Nicotiana benthamiana*. *bioRxiv [Preprint]* doi: 10.1101/2020.06.14.150458:2020.06.14.150458
- Margolin, E. A., Strasser, R., Chapman, R., Williamson, A. L., Rybicki, E. P., and Meyers, A. E. (2020e). Engineering the plant secretory pathway for the production of next-generation pharmaceuticals. *Trends Biotechnol.* 38, 1034–1044. doi: 10.1016/j.tibtech.2020.03.004
- McLellan, J. S., Pancera, M., Carrico, C., Gorman, J., Julien, J. P., Khayat, R., et al. (2011). Structure of HIV-1 gp120 V1/V2 domain with broadly neutralizing antibody PG9. *Nature* 480, 336–343.
- Meyers, A. E., Peter, R. E., and Margolin, E. A. (2018). *Co-Expression of Human Chaperone Proteins in Plants for Increased Expression of Heterologous Polypeptides*.
- Montero-Morales, L., and Steinkellner, H. (2018). Advanced plant-based glycan engineering. *Front. Bioeng. Biotechnol.* 6:81. doi: 10.3389/fbioe.2018.00081
- Nandi, S., Kwong, A. T., Holtz, B. R., Erwin, R. L., Marcel, S., and McDonald, K. A. (2016). Techno-economic analysis of a transient plant-based platform for monoclonal antibody production. *MAbs* 8, 1456–1466. doi: 10.1080/19420862.2016.1227901
- Pillet, S., Aubin, E., Trepanier, S., Bussiere, D., Dargis, M., Poulin, J. F., et al. (2016). A plant-derived quadrivalent virus like particle influenza vaccine induces cross-reactive antibody and T cell response in healthy adults. *Clin. Immunol.* 168, 72–87. doi: 10.1016/j.clim.2016.03.008
- Ringe, R. P., Pugach, P., Cottrell, C. A., LaBranche, C. C., Seabright, G. E., Ketas, T. J., et al. (2019). Closing and opening holes in the glycan shield of HIV-1 envelope glycoprotein SOSIP trimers can redirect the neutralizing antibody response to the newly unmasked epitopes. *J. Virol.* 93:e1656-18.
- Ringe, R. P., Sanders, R. W., Yasmeen, A., Kim, H. J., Lee, J. H., Cupo, A., et al. (2013). Cleavage strongly influences whether soluble HIV-1 envelope glycoprotein trimers adopt a native-like conformation. *Proc. Natl. Acad. Sci. U.S.A.* 110, 18256–18261. doi: 10.1073/pnas.1314351110
- Rosenberg, Y., Sack, M., Montefiori, D., Forthal, D., Mao, L., Hernandez-Abanto, S., et al. (2013). Rapid high-level production of functional HIV broadly neutralizing monoclonal antibodies in transient plant expression systems. *PLoS One* 8:e58724. doi: 10.1371/journal.pone.0058724
- Rutten, L., Gilman, M. S. A., Blokland, S., Juraszek, J., McLellan, J. S., and Langedijk, J. P. M. (2020). Structure-based design of prefusion-stabilized filovirus glycoprotein trimers. *Cell Rep.* 30, 4540.e3–4550.e3.
- Sack, M., Hofbauer, A., Fischer, R., and Stoger, E. (2015). The increasing value of plant-made proteins. *Curr. Opin. Biotechnol.* 32, 163–170. doi: 10.1016/j.copbio.2014.12.008
- Sainsbury, F., Thuenemann, E. C., and Lomonosoff, G. P. (2009). pEAQ: versatile expression vectors for easy and quick transient expression of heterologous proteins in plants. *Plant Biotechnol. J.* 7, 682–693. doi: 10.1111/j.1467-7652.2009.00434.x
- Shin, Y. J., Castilho, A., Dicker, M., Sadio, F., Vavra, U., Grunwald-Gruber, C., et al. (2017). Reduced paucimannosidic N-glycan formation by suppression of a specific beta-hexosaminidase from *Nicotiana benthamiana*. *Plant Biotechnol. J.* 15, 197–206. doi: 10.1111/pbi.12602
- Singh, A. A., Poole, O., Kwezi, L., Lotter-Stark, T., Stoychev, S. H., Alexandra, K., et al. (2020). Plant-based production of highly potent anti-HIV antibodies with engineered posttranslational modifications. *Sci. Rep.* 10:6201.
- Sok, D., Pauthner, M., Briney, B., Lee, J. H., Saye-Francisco, K. L., Hsueh, J., et al. (2016). A prominent site of antibody vulnerability on HIV Envelope incorporates a motif associated with CCR5 binding and its camouflaging glycans. *Immunity* 45, 31–45. doi: 10.1016/j.immuni.2016.06.026
- Struwe, W. B., Chertova, E., Allen, J. D., Seabright, G. E., Watanabe, Y., Harvey, D. J., et al. (2018). Site-specific glycosylation of virion-derived HIV-1 Env is mimicked by a soluble trimeric immunogen. *Cell Rep.* 24, 1958.e5–1966.e5.
- Tanzer, F. L., Shephard, E. G., Palmer, K. E., Burger, M., Williamson, A. L., and Rybicki, E. P. (2011). The porcine circovirus type 1 capsid gene promoter improves antigen expression and immunogenicity in a HIV-1 plasmid vaccine. *Virol. J.* 8:51. doi: 10.1186/1743-422x-8-51
- van Diepen, M. T., Chapman, R., Douglass, N., Galant, S., Moore, P. L., Margolin, E., et al. (2019). Prime-boost immunizations with DNA, modified vaccinia virus ankara, and protein-based vaccines elicit robust HIV-1 Tier 2 neutralizing antibodies against the CAP256 superinfecting virus. *J. Virol.* 93:e2155-18.
- van Diepen, M. T., Chapman, R., Moore, P. L., Margolin, E., Hermanus, T., Morris, L., et al. (2018). The adjuvant AlhydroGel elicits higher antibody titres than AddaVax when combined with HIV-1 subtype C gp140 from CAP256. *PLoS One* 13:e0208310. doi: 10.1371/journal.pone.0208310
- Ward, B. J., Gobeil, P., Séguin, A., Atkins, J., Boulay, I., Charbonneau, P.-Y., et al. (2020a). Phase 1 trial of a candidate recombinant virus-like particle vaccine for Covid-19 disease produced in plants. *medRxiv [Preprint]* doi: 10.1101/2020.11.04.20226282
- Ward, B. J., Landry, N., Trepanier, S., Mercier, G., Dargis, M., Couture, M., et al. (2014). Human antibody response to N-glycans present on plant-made influenza virus-like particle (VLP) vaccines. *Vaccine* 32, 6098–6106. doi: 10.1016/j.vaccine.2014.08.079
- Ward, B. J., Makarkov, A., Séguin, A., Pillet, S., Trepanier, S., Dhaliwall, J., et al. (2020b). Efficacy, immunogenicity, and safety of a plant-derived, quadrivalent, virus-like particle influenza vaccine in adults (18–64 years) and older adults (>=65 years): two multicentre, randomised phase 3 trials. *Lancet* 396, 1491–1503. doi: 10.1016/S0140-6736(20)32014-6
- Watanabe, Y., Allen, J. D., Wrapp, D., McLellan, J. S., and Crispin, M. (2020). Site-specific glycan analysis of the SARS-CoV-2 spike. *Science* 369, 330–333. doi: 10.1126/science.abb9983
- Watanabe, Y., Bowden, T. A., Wilson, I. A., and Crispin, M. (2019). Exploitation of glycosylation in enveloped virus pathobiology. *Biochim. Biophys. Acta* 1863, 1480–1497. doi: 10.1016/j.bbagen.2019.05.012
- Wilson, I. B., Zeleny, R., Kolarich, D., Staudacher, E., Stroop, C. J., Kamerling, J. P., et al. (2001). Analysis of Asn-linked glycans from vegetable foodstuffs: widespread occurrence of Lewis a, core alpha1,3-linked fucose and xylose substitutions. *Glycobiology* 11, 261–274. doi: 10.1093/glycob/11.4.261
- Wormald, M. R., and Dwek, R. A. (1999). Glycoproteins: glycan presentation and protein-fold stability. *Structure* 7, R155–R160.
- Yang, W. H., Aziz, P. V., Heithoff, D. M., Mahan, M. J., Smith, J. W., and Marth, J. D. (2015). An intrinsic mechanism of secreted protein aging and turnover. *Proc. Natl. Acad. Sci. U.S.A.* 112, 13657–13662. doi: 10.1073/pnas.1515464112
- Zhou, T., Doria-Rose, N. A., Cheng, C., Stewart-Jones, G. B. E., Chuang, G. Y., Chambers, M., et al. (2017). Quantification of the impact of the HIV-1-glycan shield on antibody elicitation. *Cell Rep.* 19, 719–732.

**Conflict of Interest:** EM, RC, AM, AL-W, and ER have filed patent applications describing the development of approaches to support production of glycoproteins in plants including US 2019/0337994 A1, WO 2018 220595 A1, PA174002\_PCT, and PA2106659.4.

The remaining author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2021 Margolin, Allen, Verbeek, van Diepen, Ximba, Chapman, Meyers, Williamson, Crispin and Rybicki. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.