Development of a Liquid Chromatography High Resolution Mass Spectrometry (LC-HRMS) Method for the Quantitation of Viral Envelope Glycoprotein in Ebola Virus-Like Particle Vaccine Preparations.

Lisa H. Cazares[1,3*], Michael D. Ward[1], Ernie Brueggemann[1], Tara Kenny[1], Paul Demond[2,4], Chris Mahone[1], Karen A.O. Martins[1], Jonathan E. Nuss[1], Trevor Glaros[2], and Sina Bavari[1]

[1]*Molecular and Translational Sciences Division, U.S. Army Medical Research Institute of Infectious Diseases, Frederick, MD 21702*

[2]*BioSciences Division, Biodefense Branch, Edgewood Chemical Biological Center, Gunpowder, MD 21010*

[3] *DoD Biotechnology High Performance Computing Software Applications Institute, Telemedicine and Advanced Technology Research Center, US Army Medical Research and Materiel Command, Fort Detrick, MD 21702*

4 Excet, Inc., 8001 Braddock Road, Suite 105, Springfield, Virginia 22151

*Address correspondence to: Lisa H. Cazares: lisa.h.cazares.ctr@mail.mil
michael.d.ward56.ctr@mail.mil
ernst.e.brueggemann.civ@mail.mil
tara.a.kenny.ctr@mail.mil
paul.s.demond.ctr@mail.mil
christopher.r.mahone.ctr@mail.mil
karen.a.martins2.ctr@mail.mil
jonathan.e.nuss.ctr@mail.mil
trevor.g.glaros.civ@mail.mil
sina.bavari.civ@mail.mil

Abstract:

Background: Ebola virus like particles; (EBOV VLPs, eVLPs), are produced by expressing the viral transmembrane glycoprotein (GP) and the structural matrix protein VP40 in mammalian cells. When expressed, these proteins self-assemble and bud from 'host' cells displaying morphology similar to infectious virions. Several studies have shown that rodents and non-human primates vaccinated with eVLPs are protected from lethal EBOV challenge. The mucin-like domain of envelope glycoprotein (GP_1) serves as the major target for a productive humoral immune response. Therefore GP_1 concentration is a critical quality attribute of EBOV vaccines and accurate measurement of the amount of GP_1 present in eVLP lots is crucial to understanding variability in vaccine efficacy.

Methods: After production, eVLP's are characterized by determining total protein concentration and by western blotting, which only provides semi-quantitative information for GP_1. Therefore, a liquid chromatography high resolution mass spectrometry (LC-HRMS) approach for accurately measuring the concentration of GP_1 in eVLP's was developed. The method employs isotope dilution mass spectrometry using 4 target peptides from 2 regions of the GP_1 protein. Purified recombinant GP_1 was generated to serve as an assay standard. GP_1 quantitation in 5 eVLP lots was performed on an LTQ-Orbitrap Elite MS and the final quantitation was derived by comparing the relative response of the 200 fmol AQUA peptide standards to the analyte response at 4 ppm.

Results: Conditions were optimized to ensure complete tryptic digestion of eVLP, however, persistent missed cleavages were observed in the target peptides. Additionally, N-terminal truncated forms of the GP_1 protein were observed in all eVLP lots, making peptide selection crucial. The LC-HRMS strategy resulted in the quantitation of GP_1 with a lower limit of

quantitation (LLQ) of 1fmol and an average percent coefficient of variation (CV) of 7.4 %. Unlike western blot values, the LC-HRMS quantitation of GP_1 in 5 eVLP vaccine lots correlated positively with survival (after EBOV challenge) in mice.

Conclusions: This method provides a means to rapidly determine eVLP batch quality based upon the quantitation of antigenic GP_1. By monitoring variability based on GP_1 content, the eVLP production process can be optimized, and the total amount of GP_1 needed to confer protection can be accurately determined.

Keywords:
Ebola virus, virus like particles, high resolution mass spectrometry, stable isotope dilution quantitation.

Background:

Ebola is an extremely pathogenic virus that causes hemorrhagic fever and can result in mortality rates of up to 90%. The 2014 Ebola endemic in West Africa brought global attention to a disease that was once only an isolated-regional problem. More than a year later and with a death toll greater than 10,000 people, there is an urgent need for novel therapeutic strategies including treatment and prevention. Virus-like-particles (VLPs) represent a new type of prophylactic vaccine that has had success and is commercialized in products such as, Cervarix (human papillomavirus) [1], and Gardasil (human papillomavirus) [2]. Virus-like particles (VLPs) are generated by exploiting the intrinsic ability of structural viral component proteins, frequently major proteins in the capsid or envelop, to spontaneously self-assemble when expressed in mammalian cells [3]. VLPs are therefore composed of a subset of viral components that mimic the wild-type virus structure but lack viral genetic material, rendering them non-infectious. Unlike recombinant protein vaccines which may elicit a weak immune response due to non-ideal presentation of the viral antigens to the immune system, VLPs are usually antigenically indistinguishable from infectious virus particles [4-6]. These properties make VLPs promising candidates for new efficacious vaccines against many viral pathogens including filoviruses such as Ebola.

Ebola Virus (EBOV) VLPs (eVLPs) are produced by transfection of HEK293 cells with plasmids encoding the genes for viral matrix protein VP40, and envelope glycoprotein (GP) [7-9]. The envelope GP is solely responsible for viral attachment, fusion, and entry of new host cells, and it is therefore a major target of vaccine design efforts. When these proteins are expressed in mammalian cells, they self-assemble and bud from lipid rafts resulting in eVLPs that contain GP, VP40, and other packaged host proteins [10].

Each of the seven genes which comprise the EBOV genome is transcribed into individual messenger RNAs (mRNAs) with the exception of the fourth gene, which encodes for GP. In virus-infected cells, several GP-specific mRNAs are synthesized due to a transcriptional RNA editing phenomenon. Envelope GP is not the primary product of the fourth gene but instead is generated through transcriptional editing, which leads to the insertion by the viral polymerase of an extra adenosine into a stretch of seven other adenosine residues (total 8A) at a specific-editing site near the middle of the coding region [11]. The EBOV polymerase transcribes the unedited GP gene which contains 7 adenosines at the editing site most of the time (>80%), and these transcripts result in the expression of the predominant GP gene product, secretory glycoprotein (sGP) [12]. The addition of 2 adenosine residues at the editing site (total of 9A) codes for a third GP gene product known as second secreted GP (ssGP). Both secreted forms have the same amino-terminal 295 amino acids as GP. Editing (total of 8A) results in the continuation of GP for 381 amino acids beyond the divergence point (see **Figure 1** for sequence alignment) resulting in the pre-processed GP polypeptide (GP0). GP0 is proteolytically cleaved to a large amino-terminal fragment (GP_1) and a smaller carboxy-terminal fragment (GP_2) in the trans-Golgi network by a furin-like enzyme [13]. Mature envelope GP is formed by the re-association of GP_1 and GP_2 through disulfide bonding, and the $GP_{1,2}$ complex is anchored in the membrane by a transmembrane domain near the carboxy terminus of GP_2 [14, 15]. GP_1 contains a highly glycosylated mucin-like domain (MLD) and antibodies that recognize this region of GP_1 have been shown to be protective in mouse models of lethal Ebola virus challenge [16]. In addition, many neutralizing antibodies, including two that comprise part of a promising therapeutic cocktail [17], are directed against the MLD [16, 18, 19].

The GP expression vector used to produce eVLP in HEK293 cells encodes for a transcript containing 8 adenosines and thus should produce only $GP_{1,2}$. Large scale production of eVLPs is performed by contract manufacturing organizations and each lot is characterized by assays that measure total protein and GP1 concentrations (western blotting or single antibody ELISA). Ongoing vaccine studies in our laboratory have shown that eVLPs provide protection against a lethal dose of EBOV in mice and non-human primates when administered with an appropriate adjuvant [20, 21]. Vaccine dosages are administered based on GP_1 protein concentration; however, the effectiveness (based on survival) of each small scale VLP preparation can be highly variable. Therefore improved methods are needed to serve as lot release assays for each eVLP preparation to ensure that only material of sufficient quality is used for *in vivo* evaluation.

This report describes the development of an isotope dilution LC-HRMS method for the absolute quantitation of Ebola GP_1 in eVLP. This protocol resulted in the quantitation of GP_1 with a limit of quantitation of 1fmol and an average percent coefficient of variation of 7.4 %. The optimized MS quantitation of GP_1, in contrast to the western blot quantitation correlated with survival in vaccinated mice after EBOV challenge. This assay provides a means to monitor eVLP batch variability based on GP_1 content, provides a means of monitoring product purity during process development, and will assist in the determination of the dosage needed to confer protection in vaccinated animals.

Materials and Methods:

Generation and Characterization of eVLPs.

eVLPs were produced under a contract with Paragon Bioservices (Baltimore, MD) using a modification of the procedure described by Warfield et al. [22]. In brief, eVLPs were created by transfecting HEK 293 cells with expression vectors containing the genes for GP and VP40 proteins [7, 22-24]. To purify the eVLPs, the clarified cell supernatants were pelleted, separated on a 20–60% continuous sucrose gradient, concentrated by a second centrifugation, and resuspended in endotoxin-free PBS. The gradient fractions containing the eVLPs were determined using western blotting using an anti-GP_1 antibody (6D8). The total protein concentration of each eVLP preparation was determined in the presence of Nonidet P-40 detergent using a detergent-compatible protein assay (Bio-Rad). For these blots unpurified recombinant GP material was used as an assay standard for the generation of a standard curve.

Generation and Characterization of a Recombinant GP_1 standard.

A batch of recombinant Ebola glycoprotein (rGP, carrying an N-terminal poly-histidine tag) was expressed in human HEK293 cells and subsequently purified by immobilized metal affinity chromatography (IMAC). The material was produced under a contract with the Frederick National Laboratory for Cancer Research (Frederick, MD). Analytical scale reverse phase chromatography was used to further fractionate the protein preparation under reducing conditions. Recombinant Ebola glycoprotein material was reduced with 2-mercaptoethanol (final concentration, 0.5 M) during a 30 minute room temperature incubation and then injected (300 µg total protein) onto an apHera C4 column (150 mm x 4.6 mm, 5 µm; Supelco). Mobile phases were as follows: (A) 0.1% trifluoracetic acid (TFA) and (B) acetonitrile/0.1% TFA. The flow rate was set to 0.5mL/min and rGP was separated using the following gradient: 0-3 min: 10% B, 3-5 mins: 10 to 20% B, 5-65 mins: 20 to 45% B, 65-71 mins: 45 to 80% B, and 72-82 mins: 80

to 10% B. During the 20-45% B gradient, nine peaks were collected and dried to completion in a vacuum concentrator. All GP_1 purification experiments were conducted using an Agilent 1200 HPLC system equipped with a UV detector; eluents were continuously monitored at 214 nm.

Each fraction of purified rGP_1 was re-dissolved in 100 µl of 8M urea/PBS. The protein concentration of each fraction was estimated by measuring the optical density (OD) at 280 nm in a spectrophotometer and assuming an extinction coefficient at 1% equal to 10 (under this assumption, a 1 mg/ml solution of a protein would have an OD reading of 1.0). Protein from each fraction (500 ng) and 1 µg of the original unfractionated GP material were resolved on a 4-12% BOLT SDS PAGE gel (Life Technologies) and stained with silver (Pierce Silver Stain kit, Fisher Scientific) following the manufacturer's instructions. Following the initial characterization experiment, a larger scale purification experiment was conducted to obtain a sufficient quantity of GP_1. In this iteration, 300 µg of unpurified recombinant GP material was fractionated by reverse phase HPLC and a single peak corresponding to GP_1 was manually collected. The OD at 280 nm was recorded and a preliminary protein concentration was determined for the sample using a theoretical molar extinction coefficient of 45,380 (calculated from the primary sequence of GP_1 using the protein parameter tool on the ExPASy server, http://web.expasy.org/protparam/). The sample was subsequently aliquoted and dried under vacuum centrifugation. SDS PAGE was used to compare the rGP_1 pool to the original unfractionated rGP material. For this experiment, 2.5 µg of rGP_1 and 3.3 µg of unfractionated rGP were resolved on a 4-12% BOLT SDS PAGE gel (Life Technologies) and stained with Coomassie Blue (Imperial protein stain, Fisher Scientific). Lastly, the protein content of the pooled and purified rGP_1 preparation was determined by amino acid analysis (AAA) following

acid catalyzed hydrolysis by Biosynthesis (Lewisville, TX). AAA conducted on triplicate rGP_1 samples determined that on average, each aliquot contains 1.8 µg of protein.

Western Blot Analysis.

Based on total protein concentration, approximately 20-50ng of each eVLP lot was loaded onto a 4-12% SDS PAGE gel and run under reducing conditions. Known amounts of recombinant Ebola GP material (purified GP_1 and unpurified) were also loaded on the gel. Two separate gels were run for the eVLP lots tested and transferred to PVDF membranes. Each blot was blocked overnight with Odyssey blocking buffer in phosphate buffered saline (PBS) (LI-COR Biosciences Lincoln, NE) and then incubated with primary antibody against GP_1 (6D8 or F88.H3D5, 1:1000) for 1 hour at room temperature. After washing 3X with PBS + 0.1% Tween-20 for 5 minutes, secondary antibody (1:5000) goat α-mouse IRDye® 680 labelled (LI-COR) was added and the blots were incubated an additional hour. The blots were again washed 3X with PBST, and then stored in PBS until visualized with an Odyssey infrared imaging system (LI-COR Biosciences Lincoln, NE: model number 9210).

Preparation of eVLP and rGP_1 Standard Proteolytic Digests

Upon receipt of each lot of eVLP from the contractor, stocks were divided into 10 µg aliquots based on the total protein concentration and stored at -80°C until use. For simplicity, each of the 5 lots of eVLP used in this study was designated using alphabetical values (A-E). Sample preparation for MS was performed by first increasing the volume of each aliquot to 50µL with 'Solution **tA'** (25mM Tris-HCl, pH 8.0), reducing with 55 mM DTT at 55°C for 30 minutes, and then alkylating with 68 mM iodoacetamide at room temperature for 45 minutes. Both of these

steps were performed in the presence of 0.05% ProteaseMax™ (Promega Madison, WI). The total volume was then increased to 95 µL with 'Solution **tD'** (25 mM Tris-HCl, pH 8.0, 10% acetonitrile) and 4 µL of a 0.1 µg/µL sequencing grade trypsin/lys-C solution (Promega) and 1 µL of 1% ProteaseMax™ were added followed by incubation at 42°C for 4 hours. Digests were heated to 90°C for 5 minutes, dried completely by speed-vac and stored at -80°C until analyzed. The purified rGP_1 standard was digested using the same protocol as the eVLP's with the exception that the concentration of the trypsin/lys-C was reduced 4-fold.

Quantitation of GP_1 by LC-HRMS

AQUA Ultimate™ peptides (Thermo Fisher Scientific) were synthesized based on the results of extensive survey runs of digested and purified rGP_1 to determine which endogenous peptide sequences had the fewest possible post-translational or artefactual modifications and that also resulted in unambiguous MS^2 spectra for identification, as well as a consistent and chromatographically distinct extracted ion chromatograms (XIC) for quantitative measurement. The following four peptide sequences were selected: *301-IRSEELSFTAVSNR-314*, *303-SEELSFTAVSNR-314*, 65-*SVGLNLEGNGVATDVPSATK*-84, and 65-*SVGLNLEGNGVATDVPSATKR*-85. Each peptide had a C-terminal amino acid modified with ^{13}C and ^{15}N isotopes resulting in a 10 Da and 8 Da mass increase for arginine and lysine respectively. AQUA peptides were supplied by the manufacture in a 40% acetonitrile, 0.1% formic acid solution at 5pmol/µL. A 2X working solution was prepared in 40% acetonitrile, 0.1% formic acid by adding 8 µl of each stock peptide into a total volume of 200 µL (200 fmol/µL). The analyte digest was resuspended in 60µL or 80 µL 40% acetonitrile, 0.1% formic and a 5-point, 2-fold serial dilution performed. AQUA peptides were then spiked into each

analyte dilution at a 1:1 (*v:v)* ratio resulting in a 100 fmol/µL AQUA standard concentration. In addition, a blank was prepared by diluting the AQUA standards 1:1 with 40% acetonitrile, 0.1% formic acid. Samples were resolved on an Acclaim PepMap 100 column (1mm x 100mm) packed with 3um, 100A C18 particles and analyzed in triplicate from lowest to highest concentration by loading 2 µL onto an Ultimate 3000 HPLC (Thermo Fisher Scientific), Mobile phases were as follows: (A) 0.1% formic acid (FA) and (B) acetonitrile/0.1% FA. The flow rate was set to 75 µL/min and peptides were eluted using a 17-minute linear gradient of 1-34% mobile phase B. The column eluent was connected to an Orbitrap Elite mass spectrometer with a HESI-2 ion source (Thermo Fisher Scientific) with a sheath gas pressure of 20 psi and an auxiliary gas flow of 5 units. The electrospray ionization voltage was 5.0 kV with an ion transfer tube temperature of 350 ^{o}C and S-lens RF at 50%. The automatic gain control target was 5.0 x10^4 for Orbitrap in SIM mode and 1.0 x10^4 for linear ion trap in MS/MS mode. The maximum injection time for MS/MS was set to 30 milliseconds. Four consecutive 200 amu SIM scans over the range of m/z 415-1215 at a resolution of 60,000 were used to detect the 20 ions of interest followed by 4 targeted MS/MS low resolution CID scans of the most prominent analyte peptides for sequence verification. For each peptide (heavy and light), both the doubly and triply charged ions were considered and used for quantitation. The average of triplicate extracted ion chromatogram (XIC) counts of each of the 4 standard AQUA peptides, the 4 analyte peptides and deamidated SVG peptides were obtained using XCalibur 2.0 (Thermo Scientific) with automatic integration baseline window set at 10 scans, area noise factor at 5, and peak noise factor set to 20. The XIC counts from each SVG, SVGR, SVG^{deam}, and SVGR^{deam} peptide charge state were first summed in each individual replicate run and then the average for the three technical repeats was determined to represent the contribution of Peptide Set 2 at each dilution.

The SEE and IRSEE (Peptide Set 1) values were obtained similarly. The AQUA standard peptide XIC counts were then used to calculate the ratio of AQUA peptide standard to the 'light' analyte peptide at each dilution using a mass tolerance of 4 ppm. This ratio or relative response was used to generate standard curves which were then used to determine the amount of analyte in fmols injected on-column. These fmol values were then converted to µg to calculate the total GP_1 using a total protein mass of 54,768 Da (UniProt entry Q05320, 33-501).

Limit of Quantitation and Linearity of Analyte Peptides

A previously quantified digest of a Brown Lot eVLP was diluted to 140 fmol/µL GP_1 in 40% acetonitrile, 0.1% formic acid and serially diluted 2-fold down to 0.5 fmol/µL for a total of 9 dilutions. Using a 2 µL injection volume, each dilution was run in triplicate as described above and XIC area standard curves generated for each of the 4 quantitation peptides ranging from 275 fmol to 1.0 fmol. The similar procedure was carried out on the AQUA peptides except the dilution was carried to 0.4 fmol/µL.

Deamidation of AQUA Peptide Standards

A 40 pmol aliquot of AQUA SVG peptide was resuspended in 200 µL 50 mM NH_2HCO_3 pH 8.1 and incubated at 50°C for 3 days then dried to completion by speed-vac. The sample was resuspended in 200 µL 40% acetonitrile, 0.1% formic acid and 2µL was injected using the instrument and chromatographic conditions outlined above. Target masses were aligned by charge state and retention time and XIC values derived as described above using a mass tolerance of 4 ppm.

In-Gel Trypsin Digestion

A 5 µg aliquot of VLP was fractionated by SDS-PAGE onto a 4%-12% gel (BioRad) and the 10 highest intensity bands excised and minced into 1 x 1 mm plugs. Each sample was serially processed in 100 µL solution **tA,** then solution **tB** (25mM Tris-Cl, pH 8.0, 50% Acetonitrile), and finally 100% Acetonitrile before being evaporated to dryness in a SpeedVac. Each gel slice was then reduced and alkylated by incubation in 55 mM DTT at 55°C followed by incubation with 68 mM Iodoacetamide for 45 minutes at room temperature. Bands were dried to completion and 10 µL of a 12.5 ng/µL sequencing grade modified trypsin solution (Promega, Madison, WI) in solution **tD** was added and incubated at room temperature for 30 minutes until trypsin was absorbed. 70 µL solution tD was then added and samples incubated overnight at 37°C. Peptides were then extracted 2X by incubating in 50% Acetonitrile, 0.1% formic acid and the combined digest dried to completion by speedvac.

Animals, Vaccinations, and Viral challenge

Research was conducted under an IACUC approved protocol in compliance with the Animal Welfare Act, PHS Policy, and other Federal statutes and regulations relating to animals and experiments involving animals. The facility where this research was conducted is accredited by the Association for Assessment and Accreditation of Laboratory Animal Care, and adheres to principles stated in the Guide for the Care and Use of Laboratory Animals, National Research Council, 2011. C57BL/6 mice were obtained from NCI Charles River. Female mice between 8-12 weeks of age were vaccinated with 100 µl injections containing 10 µg of GP (as determined by western blot) via the intramuscular (IM) route, in the caudal thigh. Each lot of eVLP was irradiated at 1e6 rad to ensure sterility and contained less than 25 EU/ml endotoxin and less than

10 colony forming units (CFU) of bacteria per vaccination. VLP were diluted in sterile saline and vaccinations were administered two times, with three weeks between vaccinations. Viral challenge occurred four weeks after the second vaccination. A challenge dose of 1,000 pfu of mouse-adapted Ebola virus [25] was administered via the intraperitoneal route (IP). The survival data was pooled from 2-3 studies with n=10 mice each.

Statistical Analysis (differences between lots, animal survival rates.)

Survival studies were evaluated using Fisher's exact test with multiple testing corrections performed by permutation based on the number of comparison's performed. The significance of the deviation from a null hypothesis (p value) was reported for the survival observed in animals vaccinated with each eVLP lot.

Results:

Selection and Evaluation of GP_1 Target Peptides for Quantitation by LC-HRMS

In the development of a reproducible protein MS quantitation scheme, the selection of target peptides is a crucial step, especially when the protein of interest is expressed in multiple isoforms, and is highly post-translationally modified. In both the infectious virions and eVLP preparations, GP_1 and GP_2 are proteolytically processed from the GP0 polypeptide and disulfide linked to form the mature $GP_{1,2}$ transmembrane protein complex [2, 15] (see **Figure 1A**). Four peptides were initially identified as target candidates for the quantitation of GP_1 primarily due to their ionization characteristics, lack of post-translational modifications and relative distance within the sequence. During initial LC-HRMS method development we found two of these peptides (173- *GTTFAEGVVAFLILPQAK[^{13}C6, ^{15}N2]*-190) and (479-

LGLITNTIAGVAGLITGGR[^{13}C6, ^{15}N4]-497) failed to show consistent linearity. The GTT peptide and LGL peptide have a Grand Average of Hydropathy score (GRAVY) [26] of 0.933 and 1.08 respectively, indicating a high level of hydrophobicity, which can hinder reliable quantitation. The remaining 2 peptides 65-*SVGLNLEGNGVATDVPSATK[^{13}C6, ^{15}N2]*-84 and 303-*SEELSFTAVSNR[^{13}C6, ^{15}N4]*-314 (designated SVG and SEE, respectively) provided highly reproducible linear standard curves and were selected for use in the assay (see **Figure 1A and 1B**). The selection of these 2 peptides, as we subsequently discovered, also offered a way to distinguish full length GP_1 from truncated versions of the protein, as the SEE peptide sequence is found only in the full length GP_1 molecule. Isotopically labelled AQUA Ultimate™ peptides (Thermo Scientific) of each peptide sequence were synthesized and used for quantitation. Synthetic AQUA (Absolute QUAntitation) peptides are chemically and physically indistinguishable from their endogenous counterparts with respect to retention time, ionization efficiency, and MS/MS fragmentation except they are modified to include 13C and 15N isotopes that increase their relative mass by very precise increments [21]. For this study, AQUA Ultimate™ peptides (Thermo Fisher Scientific) were selected as they have the highest available concentration precision and purity.

During the initial survey runs which were conducted to optimize the digestion of the eVLP for completeness and reproducibility, it was observed that two missed cleavage sites appeared regularly: a C-terminal arginine on the SVG peptide and an N-terminal arginine on the SEE peptide. We rigorously searched for additional missed cleavages as well as non-specific cleavages upstream and downstream of the fully tryptic peptides, and we found no evidence that these species were present (see **Supplemental Table 1A** and **1B**). Given that the ratio of missed cleavage to fully tryptic peptides was highly variable (0.4% - 45%), the 2 peptides representing

these missed cleavages (65-*SVGLNLEGNGVATDVPSATK**R**[^{13}C6, ^{15}N4]*-85 and 301-***IR****SEELSFTAVSNR[^{13}C6, ^{15}N4]*-314*)* were synthesized and evaluated for reproducibility and linearity. These peptides were chromatographically distinct, generated linear standard curves and were therefore suitable for use in the quantitation assay (see **Figure 1B** and **Supplemental Figure 1**). It was also observed that one of the two asparagine residues in the endogenous SVG peptide, but not both, were routinely deamidated. Since all XIC counts from this species need to be combined in order to account for the full stoichiometric contribution of the SVG peptide we evaluated whether the non-deamidated AQUA SVG peptide standard could be used to quantitate the level of deamidated analyte peptide. The standard AQUA SVG and SVGR peptides were fully deamidated by incubating them at 55°C for 2.5 days at pH 8.1 (see **Supplemental Figure 2**). Interestingly, even with this harsh treatment, the doubly deamidated species comprised only 5% of the total SVG peptide compliment, indicating that under normal processing conditions it would be a highly unlikely modification. The XIC response of the deamidated peptide standards were then compared to the non-treated peptide standard of the same concentration. As shown in **Figure 1C**, the response was essentially identical. Therefore, the XIC counts derived from the SVG and SVGR standard AQUA peptides can be used to quantify the additional XIC counts from the endogenous deamidated peptide species without having to synthesize additional labeled deamidated standards. We did not observe deamidation of the single asparagine in the SEE target peptide.

Determination of Optimal Digestion Conditions for GP_1 within the eVLPs.

The proteolytic enzyme of choice is a mass spectrometry grade Trypsin/Lys-C combination (Promega V5073) as it is well characterized, versatile and highly specific. Initial digestion

388 experiments and LC-HRMS analysis of the eVLPs revealed that some regions of GP_1 are very
389 resistant to proteolytic digestion even in the presence of enhancing surfactants such as
390 ProteaseMax™. To ensure as complete a digestion of the eVLP GP_1, we conducted extensive
391 testing using a variety of buffer formulations, reagents, and pre-digestion treatments. These
392 treatments included deglycosylation, sonication and high temperature. Since $GP_{1,2}$ is a heavily
393 glycosylated membrane embedded protein, we performed PNGase deglycosylation prior to
394 digestion in the hope of reducing steric hindrance of the sugars and thereby enhancing trypsin
395 proteolysis. Although we observed a modest improvement in overall peptide count as well as a
396 reduction in frequency of the SVG/SEE missed cleavages, we did not observe any appreciable
397 differences in the ratios of the peptides selected for use in quantitation (data not shown). We
398 concluded the additional deglycosylation procedure would only add to the complexity of the
399 assay. We also tested the cleavable detergent/surfactant, ProteaseMax™ (Promega, Madison,
400 WI), which is designed to enhance the performance of trypsin, and is especially useful for
401 membrane proteins. This reagent dramatically reduced the overall number of missed cleavages
402 and allowed the digest time to be reduced from 16 hours to 4 hours without any loss of digestion
403 efficiency. Despite these efforts, we were unable to completely eliminate the occurrence of the
404 target peptide missed cleavages described above. However, we did not observe any additional
405 upstream or downstream missed cleavage species from either target peptide in survey runs from
406 each eVLP lot tested (**Supplemental Table 1**). Missed cleavage species were observed in 8.2%
407 of the SEE peptide and 31% in the SVG peptide. These values represent the typical level
408 observed in all 5 eVLP lots tested after trypsin digestion. We therefore concluded that the 4
409 peptides selected for the assay would be adequate for quantitation of GP_1 present in eVLP

preparations. The final peptide sequences and charge states used for quantitation of Ebola GP_1 are shown in **Table 1**.

Reverse Phase Purification of GP_1 Standard

In any protein quantitation experiment, the assumption is that unique peptides from different regions within a protein would display a 1:1 molar relationship. However, early quantitation experiments with test lots of recombinant GP material and eVLP revealed a variable target peptide (SVG:SEE) stoichiometric ratio between the lots which was otherwise consistent within each lot. In some cases the disparity between the SVG quantitation and the SEE quantitation was as high as 25%. In order to rule out experimental error as the cause of the discrepancy, we prepared a pure monomeric full length rGP_1 standard from recombinant GP material that could be used to assess the accuracy of the quantitation method. As shown in **Figure 2A**, a reverse phase chromatography procedure was performed that fractionated reduced rGP material into multiple sub-species. Fractions were collected and processed for SDS PAGE analysis and silver stained. As seen in **Figure 2B,** fractions 1-4 and fractions 6-7 constitute GP_1 and GP_2, respectively. Interestingly, fractions 1-4 yielded nearly identical SDS PAGE profiles despite observing multiple shoulder peaks on the reverse phase chromatogram. Ultimately however, the fractionation procedure resulted in a significant enrichment of individual protein species within the rGP preparation. As seen in **Figure 2C**, SDS PAGE analysis confirmed that the fractionated material was highly enriched for GP_1. Collectively, this data indicates that the procedure significantly reduced the amount of heterogeneity in the original sample and produced an enriched version of GP_1 that was suitable for use as an assay standard.

Validation of the Quantitation Method with Purified rGP_1 Standard

Quantitative AAA analysis indicated each aliquot of purified rGP_1 contained an average of 1.8 µg GP_1 protein. In order to evaluate the accuracy of our method, four rGP_1 aliquots were resuspended in either 60 µL or 80 µL 40% acetonitrile, 0.1% formic acid and quantitated using the newly developed HRAM LC-MS/MS. As seen in **Table 2**, after averaging the individual peptide set values, the GP_1 concentration was determined to be 1.50 µg/aliquot for trial 1 and 1.49 µg/aliquot for trial 2 using a dilution volume of 60 µL. These values are within 16.8% and 17.4% of the value obtained with AAA (1.8µg). The SVG/SEE stoichiometric disparity, designated as ΔS, was -7.00% for trial 1 and 9.3% for trial 2. For trials 3 and 4 using a dilution volume of 80 µL, the GP_1 concentration was 1.56 and 1.64 µg/aliquot respectively. These values are within 13.4% for trial 3 and 9.9% for trial 4 of the value determined by AAA analysis. These data indicate that the LC-HRMS method and the combination of these 4 peptides (Set 1 and Set 2) is sufficient to account for the GP_1 protein present with an average accuracy 85.6%.

Development of a High Resolution/Accurate Mass (HR/AM) Quantification of GP in eVLPs:

Since the purified rGP_1 standard returned acceptable LC-HRMS quantitation results with both target peptide sets, we sought to determine the disparity observed in the quantitation of GP_1 in the eVLPs when using Set 1 and Set 2 peptide pairs. While the eVLPs are designed to produce only $GP_{1,2}$ by altering the primary sequence used to transfect the HEK293 cells, the presence of multiple forms of GP was observed by western blotting using two monoclonal antibodies with epitopes located in different portions of the molecule (see **Figure 2D&E**). The mouse monoclonal antibody 6D8 binds at amino acids 389-405 and therefore has affinity for only Ebola

456 GP_1 [16]. This is the antibody routinely employed for the determination of GP content in the
457 eVLP preparations by quantitative western blot or ELISA. Antibody H3D5 is a mouse
458 monoclonal antibody which binds at amino acids 72-109 and therefore has affinity for all forms
459 of GP (both secreted and membrane bound) (see **Figure 1A**) each containing the SVG peptide
460 sequence. This antibody has reactivity with all subtypes of Ebola GP_1, for all subspecies [27]. As
461 shown in **Figure 2D** and **2E**, the predominant band visualized in the unfractionated rGP material,
462 purified $rGP_{1,}$ and 2 lots of eVLPs (lots 'A' and 'E') using both antibodies is fully glycosylated
463 GP_1 (~140kDa), however the H3D5 blot shows the presence of strong distinct bands of a lower
464 molecular weight (~50-100 kDa) present in both eVLP lots and the unpurified rGP material.
465 These bands are much reduced in the rGP_1 purified standard. The additional bands visible in the
466 eVLP western blot using the H3D5 antibody do not correspond to the correct molecular weight
467 for either sGP or ssGP (50 and 47 kDa respectively). In order to verify sequence identity these
468 bands were excised from a gel of one eVLP lot ('A') and stained for total protein with coomassie
469 blue. The 10 most intense bands were excised, trypsin digested, and analyzed with long-gradient
470 CID survey runs as well as targeted LC-HRMS MS to identify any GP protein fragments
471 contributing to the peptide quantitative variability. The results of this sequencing experiment are
472 shown in **Supplemental Figure 3**. All bands excised were confirmed to contain EBOV GP_1 or
473 GP_2 peptides. A gradual loss of C-terminal peptide identifications for GP_1 was observed as the
474 smaller products visible in the gel were sequenced, suggesting the presence of truncated forms of
475 GP_1 in the eVLP. This data indicates that peptides derived from the N-terminus would not be
476 suitable candidates for quantitation of GP_1. We therefore concluded that the SEE target peptide
477 (Set 1) was the only reliable standard for the quantitation of GP_1 in eVLP.

478

Testing of the LC-HRMS Quantitation Method for Reproducibility with an eVLP Digest

Three aliquots from a single test lot of eVLP (lot 'A') were used to evaluate the quantitation method for reproducibility. As shown in **Figure 3A**, the workflow was as follows: heavy AQUA standard peptides were added at a fixed concentration of 200 fmols/injection while varying the concentration of the eVLP analyte digest (10 µg based on total protein concentration) over 4 two-fold dilutions. Each complete quantitation set contained a 200 fmol/injection AQUA peptide standard blank and was run in triplicate from which average XIC and percent CV values were calculated. The XIC area contributions from each charge state were summed to provide the total fmols for each peptide species (see **Table 3**). The entire quantitation was performed in triplicate with an analyte resuspension volume of 120µL. Replicates 1, 2, and 3 resulted in a calculated GP_1 concentration (based on the SEE peptide standard set 1 only) of 0.62, 0.53 and 0.61 mg/ml, respectively with an average of 0.59 mg/ml ± 0.025 mg/ml and a percent CV of 7.4%. Therefore, each aliquot of this eVLP contained an average of 1.22µg of GP_1, or 12.2% of the total protein concentration of 10 µg. The average ΔS (SVG/SEE stoichiometric disparity) value was 17.3%.

In order to reduce the possibility of including peptide ion counts from contaminating ion species, our LC-HRMS method included a second stage MS/MS step to fragment each of the analyte peptides to confirm target sequence identity. We were able to confidently identify each of the 4 analyte quantitation peptides in at least the 3 highest dilutions of the test eVLP lot used to determine assay reproducibility. Representative MS/MS spectra of the SVG and SEE target analyst peptides are shown in **Figure 4**. With the exception of the SEE peptide at the highest dilution, peptide assignments from every dilution run were of sufficient quality to obtain non-

ambiguous sequence identifications. These data suggest that the newly developed LC-HRMS method can quantitate the amount of GP_1 in eVLP reproducibly.

Linearity and Limit of Quantitation of Analyte Peptides

To assess the limit of quantitation and sensitivity of the assay, and to ensure the range of protein concentrations tested remains linear relative to our single standard peptide concentration, we performed a linearity and limit of quantitation experiment. While the observed range of concentrations over 5 dilution points spanned from 6 to 250 fmols, it was necessary to show that we could extrapolate to concentrations that fell outside the fixed concentration represented by the AQUA peptide standards. We therefore prepared a dilution of a previously quantified eVLP lot ('A') such that a 9-point 2-fold serial dilution resulted in an on-column GP_1 load of between 275 fmol and 1 fmol. The averaged triplicate XIC values were plotted and CV% values determined (see **Figure 1B**). The quantitation remained linear across the entire concentration range (1-275fmol) with R^2 values for SVG (0.9999), SVGR (0.9972), SEE (0.9986) and IRSEE (0.9979) well within the margin of significance. While CV% values at the highest dilutions were typically less than 5%, the values in the 2 most dilute concentrations spanned a range of 7.6-17.3%. This is within acceptable limits of variability and therefore the quantitative accuracy of the assay is reliable down to 1fmol.

Quantitation of Multiple Lots of eVLP and Comparison with Western Blot Quantitation.

The optimized protocol developed for the digestion and LC-HRMS quantitation of GP_1 was performed on digests of 5 different lots of eVLP's. These lots were produced by an outside contractor (Paragon BioServices, Baltimore, MD) and, at the time of our study, were being used

for a number of in-house animal studies. Using primary aliquots which were stored at -80 and would therefore experience only one freeze-thaw, eVLPs (10ug total protein) were digested and triplicate LC-HRMS quantitation runs were performed. Resuspension volumes for all eVLP digests were 120 µL. As shown in **Figure 3B**, the final quantitation is derived by comparing the relative response of the 200 fmol AQUA standards (SEE and IRSEE: Set 1) to the endogenous analyte response at 4 ppm at each dilution. The average XIC's were then calculated and used in the quantitation to obtain the final concentration of GP_1 protein present in the eVLPs. This quantitation was performed in duplicate and the concentrations of GP_1 for all 5 lots of eVLPs are shown in **Table 4**. The lowest percentage of GP_1 relative to the total protein concentration was found in lot 'E' (1.3%) with a final concentration of 0.10 mg/ml GP_1. The next lowest values are found in the lot 'D' (3.2%, 0.16 mg/ml), while the highest percentage of GP_1 relative to the total protein concentration was found in lot 'A' (15.8%, 0.59 mg/ml). This represents nearly an order of magnitude difference in relative GP_1 concentration between the VLP lots 'A' and 'E'. The ΔS values for each eVLP lot tested ranged from 7.35% for the lot 'B' to 25.5% for lot 'E'.

For each eVLP lot, the GP_1 concentration was also determined after production by the contractor via western blot with the 6D8 antibody and unpurified rGP material as a quantitation standard. As shown in **Table 4**, the range of concentration for GP_1 was 0.71-1.4 mg/ml. Total protein values provided for each lot ranged from 3.8-7.2 mg/ml. Since the western blot quantitation and the MS quantitation results were vastly different we decided to investigate the source of this discrepancy by repeating the western blot on the eVLP lots with the highest and lowest calculated GP_1 (as determined by LC-HRMS) using the 6D8 and H3D5 antibodies (see **Figure 2D and 2E**). As previously mentioned above, both eVLP lots displayed strong signals for GP_1 at ~140kDa using the 6D8 antibody. However, the H3D5 antibody revealed the presence

of truncated products previously seen in the test eVLP lot and the unpurified rGP material. These truncated products are highly abundant in the eVLP lot 'E', which returned the lowest concentration of GP_1 by LC-HRMS quantitation with the largest ΔS value (25.5%), whereas eVLP lot 'A' appears to have fewer detectable GP_1 fragments, and returned a ΔS value of 15.8%.

Correlates of VLP efficacy

In the hopes of using immune correlates as another measure of eVLP quality, the western blot and LC-HRMS quantitation results were compared to survival data in mice for each of these eVLP vaccine preparations. Each lot of eVLP was used to immunize mice (n=20) which were then challenged with a murine adapted Ebola Zaire virus. For each vaccination dose, volumes of eVLP were used which were surmised to contain 10µg of GP (as determined from the western blot quantitation performed by Paragon Biosciences). As shown in **Figure 5**, lot 'E' exhibited the lowest average survival rate after Ebola challenge (40%), and animals vaccinated with the lots 'A' and 'B' exhibited 100% survival. Lot 'E' contained the lowest calculated GP_1 concentration as determined by LC-HRMS whereas the lot 'A' contained the highest. The difference in survival between vaccination with lots 'A' and 'E' was significant (p=0.001). As shown in **Figure 5B**, if we plot the percent survival versus the GP_1 LC-HRMS quantitation in each eVLP lot, expressed as percent total protein, a strong positive correlation is observed (R^2 = 0.9149). A weaker correlation is observed if the absolute quantitation values for GP_1 (expressed as mg/ml) obtained by LC-HRMS are plotted versus survival (p=0.9025). In contrast western blot quantitation values (GP_1 as a percentage of total protein) did not display strong linear correlation with percent survival (R^2= 0.6904) and there was no correlation observed between the western blot GP1 concentrations (expressed as mg/ml) and survival.

The western blot quantitation of lot 'E' returned a value of 1.1mg/ml of GP_1. Therefore a 10µg GP dose would require 9.1 µl of the eVLP preparation for vaccination. However, based on the LC-HRMS quantitation, we can retroactively estimate that the animals were given only 0.9 µg of the 10 µg dose desired, which was adequate to protect only 4/10 animals vaccinated. Conversely, the western blot concentration for lot 'A' (1 mg/ml) is also higher than the LC-HRMS quantitation (0.59 mg/ml), and the 10µl dose thought to contain 10 µg of GP_1 actually contained 5.9 µg which was adequate to protect 100% of the vaccinated animals after Ebola challenge. Therefore the observed differences in eVLP efficacy between eVLP lots 'A' and 'E' are due to vastly different concentrations of antigenic GP_1. From the LC-HRMS quantitation of lot 'B', which also provided 100% survival, we can calculate that a vaccine dose (based on the western blot quantitation) of 10ug GP_1 would actually contain 3.5 µg which appears to be the minimal vaccination dose required to confer 100% survival in mice after Ebola challenge.

Discussion:

Provided that technical pitfalls such as incomplete protein extraction, incomplete proteolysis or protein side-chain modifications are appropriately controlled and considered, protein quantitation by MS using an AQUA strategy can be robust, accurate and reproducible, while achieving low limits of detection [28-30]. Ideally, target peptides should be well separated on a protein of interest to ensure that the entire protein is sufficiently denatured and digested prior to quantitation. Additionally, potential sites of post-translational modifications or residues susceptible to artefactual modifications should be avoided. Despite these considerations, peptide selection is an empirical exercise that balances ideal characteristics with practical limitations. For example, large proteins yield more potential target peptides than small proteins, and sequence

features can greatly constrain peptide selection. As we discovered with the quantitation of Ebola GP_1, a protein of interest may have significant sequence homology with other proteins in a complex mixture, making it difficult to adhere to the peptide selection criteria described above. The quantitation of the Ebola Zaire GP_1 in eVLP preparations was a unique challenge due to the fact that during eVLP production and purification, truncated forms of the GP protein are produced and retained throughout the post-production processing. This prevented the use of target peptides located in the first 90-100 amino acids of the GP sequence. Additionally very few suitable target peptides were available in the C-terminal region of the protein due to the high frequency of glycosylation sites and high hydrophobicity. Therefore we have chosen an unusual strategy of LC-HRMS quantitation in which 2 peptides from overlapping regions of the protein are employed. This necessitated the development of a purified GP_1 standard to provide quality control and assay validation. The average percent accuracy of our method based on quantitation of the rGP1 standard AAA analysis was 85.6%. While the HPLC fractionation we performed resulted in a significant enrichment of GP_1 from GP_2 and truncated products of GP, contaminating protein species may still be contributing to the final concentration based on AAA analysis. Indeed, the H3D5 western blot of the purified rGP_1 revealed immune-reactive species of lower molecular weight, which may be the source of this overestimation.

The presence of truncated GP products in the eVLP preparations is the likely source of variation between the quantitation of GP_1 with the two standard peptide sets. This hypothesis is supported by the data obtained during the rGP_1 standard assay development and testing. The LC-HRMS quantitation of purified, rGP_1 resulted in an average ΔS of 6.6% as compared to the eVLP method validation and reproducibility trials, which showed an average ΔS of 16.5%. Furthermore, we have never observed a higher quantitation result for GP_1 from peptide Set 1

(SEE/IRSEE) located in the middle of the molecule as compared to results obtained with peptide Set 2 (SVG/SVG.R) located near the N-terminus in any of the eVLP preparations. This supports our hypothesis that N-terminal sequence fragments in the SVG region of the protein are indeed part of the GP protein compliment of eVLP and that the moderate variability shown in our data is the result of experimental variation only.

Targeted MS approaches, in particular selected reaction monitoring (SRM), employing triple quadrupole mass spectrometers, have become the standard technique for quantitatively analyzing tens to hundreds of peptides and/or small molecules across a large number of samples. A limitation of triple quadrupole instruments however, is the relatively low resolution of precursor *m*/*z* measurements, which can allow interference from nominally isobaric background contaminants in complex mixtures. Newer instrumentation has facilitated the use of high-resolution accurate MS for quantitative analysis. This approach is often referred to as LC-HRMS and provides both qualitative and quantitative information during analysis by providing full-scan accurate mass data for the entire chromatographic run [31, 32]. Qualitative and quantitative information are obtained post-data acquisition by extracting ion chromatograms with signals centered around the m/z of the analyte(s) of interest and a predefined mass extraction window. This comprehensive detailed data obtained for each sample after LC-HRMS analysis was crucial for the development of a successful quantitation strategy for GP_1. For example, common modifications such as deamidation cause isotopic interferences, particularly when SRM-based methods are employed using low-resolution MS [31]. Conversely, full-scan HRMS data allowed the unequivocal confirmation of deaminated endogenous target peptide species which improved the accuracy our quantitation method. Furthermore, high resolution MS/MS survey scans proved to be essential for the optimization and assessment of digestion efficiency. As revealed in this

work, the detailed data provided by LC-HRMS was essential to overcome the bioanalytical challenge of GP_1 quantitation in eVLP, and allowed us to address potential issues prior to development of a more streamlined quantitation scheme.

This study also highlights the superiority of mass spectrometry methods such as SRM and LC-HRMS for protein quantitation and characterization over western blotting and other immuno-affinity methods, which has been the topic of discussion in recent review articles [33, 34]. A western blot assay depends on the specificity of a single antibody, and quantitative information often relies on a protein standard that may be poorly characterized, especially if evaluation is also based on reactivity to a single antibody. This can lead to quantitative inconsistencies such as those which we observed in the GP_1 western blot quantitation performed after eVLP production. However, there is definite value in validation by orthogonal immune-affinity approaches, and the use of the H3D5 antibody allowed us to confirm the presence of shorter versions of the GP1protein in eVLP preparations and unpurified recombinant GP material.

The use of crudely purified GP standard in the western blot quantitation unintentionally led to an overestimation of the final GP_1 concentration in eVLP, since the total protein concentration for the unpurified recombinant GP also included GP_2 as well as truncated protein species. Data obtained from the eVLP mouse vaccination study revealed that the amount of GP_1 in each eVLP lot as determined by LC-HRMS, unlike the quantitative western blot, correlated with survival after Ebola challenge. The highest observed correlation with animal survival was obtained using the percent of GP_1 in relation to the total protein in the VLP. This would suggest that the “density” of GP_1 in relation to other proteins (both viral and host derived) present in the eVLP particle is directly related to the efficacy of that particular VLP preparation.

The impact of the truncated products on eVLP quality or suitability for vaccination has not been determined. However, the LC-HRMS data has revealed that the eVLP lots which exhibited the lowest percent survival (lots 'D' and 'E') also contained the lowest amount of GP_1 and the highest ΔS values (indicating an abundance of truncated GP1 products). These truncated products may be the result of a frame-shift anomaly, ribosomal slippage or simply general protein degradation. Due to the fact that sGP is produced in greater abundance than $GP_{1,2}$, and since the proteins share a common N-terminus, it is speculated that sGP functions as a decoy molecule for EBOV-specific neutralizing and non-neutralizing antibodies [35]. Additionally, recent studies have shown that sGP actively subverts the host immune response to induce cross-reactivity with epitopes it shares with membrane-bound $GP_{1,2}$ [36]. Therefore, truncated versions of the GP_1 protein may indeed compromise the quality of eVLP vaccines.

Conclusions:

A LC-HRMS approach resulted in the successful quantitation of GP_1 in eVLP vaccine preparations. The use of this newly developed assay will allow us to monitor variability based on GP_1 content, providing quality control information to further optimize and refine the eVLP production process. Finally, using this quantitative LC-HRMS approach, the total amount of GP_1 necessary to confer protection can be accurately determined; a crucial factor in successful vaccine development.

List of abbreviations:

EBOV (Ebola virus)

VLP (Virus-like particles)

LC-HRMS (liquid chromatography high resolution mass spectrometry)

GP (Glycoprotein)

Competing interests:

None of the authors have competing financial interests.

Author's contributions:

LHC participated in the experimental design, coordinated the efforts, analyzed results, and wrote the draft of the manuscript, MDW conceived the experimental design, performed sample preparation, data analysis and presentation , EB optimized instrument methods and performed all of the instrument runs, TK participated in sample processing and edited the manuscript, PD edited the manuscript, CM assisted with sample processing and assisted with the purification of the GP_1 standard, KOM performed the vaccinations and EBOV challenge studies in mice, and edited the manuscript. JEN conceived of the method for the purification of the GP1 standard and performed the purification procedure and edited the manuscript, TG participated in experimental design and edited the manuscript, SB contributed to study design and coordinated the research efforts. All authors have read and approved the manuscript.

Acknowledgments:

This research was funded by MCS/JVAP under USAMRIID project number 1325119. Opinions, interpretations, conclusions and recommendations are those of the authors and are not necessarily endorsed by the US Army. This work was sponsored by the U.S. Army Medical Research and Materiel Command. The U.S. Army Medical Research Acquisition Activity (Fort

Detrick, MD) is the awarding and administering acquisition office. This document was prepared as an account of work sponsored by an agency of the U.S. Government. Neither the U.S. Government nor any of their employees makes any warranty, expressed or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the U.S. Government The views and opinions of authors expressed herein do not necessarily state or reflect those of the U.S. Government and shall not be used for advertising or product endorsement purposes.

The author's would like to acknowledge Christopher Cooper Ph.D for helpful conversations during the preparation of this manuscript.

References:

1. Szarewski, A., et al., *Efficacy of the human papillomavirus (HPV)-16/18 AS04-adjuvanted vaccine in women aged 15-25 years with and without serological evidence of previous exposure to HPV-16/18.* Int J Cancer, 2012. **131**(1): p. 106-16.
2. Group, F.I.S., *Quadrivalent vaccine against human papillomavirus to prevent high-grade cervical lesions.* N Engl J Med, 2007. **356**(19): p. 1915-27.
3. Garcea, R.L. and L. Gissmann, *Virus-like particles as vaccines and vessels for the delivery of small molecules.* Curr Opin Biotechnol, 2004. **15**(6): p. 513-7.
4. Bosio, C.M., et al., *Ebola and Marburg virus-like particles activate human myeloid dendritic cells.* Virology, 2004. **326**(2): p. 280-7.
5. Moron, V.G., et al., *In vivo, dendritic cells can cross-present virus-like particles using an endosome-to-cytosol pathway.* J Immunol, 2003. **171**(5): p. 2242-50.
6. Noad, R. and P. Roy, *Virus-like particles as immunogens.* Trends Microbiol, 2003. **11**(9): p. 438-44.
7. Bavari, S., et al., *Lipid raft microdomains: a gateway for compartmentalized trafficking of Ebola and Marburg viruses.* J Exp Med, 2002. **195**(5): p. 593-602.
8. Swenson, D.L., et al., *Generation of Marburg virus-like particles by co-expression of glycoprotein and matrix protein.* FEMS Immunol Med Microbiol, 2004. **40**(1): p. 27-31.

9. Noda, T., et al., *Ebola virus VP40 drives the formation of virus-like filamentous particles along with GP*. J Virol, 2002. **76**(10): p. 4855-65.
10. Licata, J.M., et al., *Overlapping motifs (PTAP and PPEY) within the Ebola virus VP40 protein function independently as late budding domains: involvement of host proteins TSG101 and VPS-4.* J Virol, 2003. **77**(3): p. 1812-9.
11. Sanchez, A., et al., *The virion glycoproteins of Ebola viruses are encoded in two reading frames and are expressed through transcriptional editing.* Proc Natl Acad Sci U S A, 1996. **93**(8): p. 3602-7.
12. Mehedi, M., et al., *A new Ebola virus nonstructural glycoprotein expressed through RNA editing.* J Virol, 2011. **85**(11): p. 5406-14.
13. Volchkov, V.E., et al., *Processing of the Ebola virus glycoprotein by the proprotein convertase furin.* Proc Natl Acad Sci U S A, 1998. **95**(10): p. 5762-7.
14. Barrientos, L.G., et al., *Disulfide bond assignment of the Ebola virus secreted glycoprotein SGP*. Biochem Biophys Res Commun, 2004. **323**(2): p. 696-702.
15. Ito, H., et al., *Ebola virus glycoprotein: proteolytic processing, acylation, cell tropism, and detection of neutralizing antibodies.* J Virol, 2001. **75**(3): p. 1576-80.
16. Wilson, J.A., et al., *Epitopes involved in antibody-mediated protection from Ebola virus.* Science, 2000. **287**(5458): p. 1664-6.
17. Olinger, G.G., Jr., et al., *Delayed treatment of Ebola virus infection with plant-derived monoclonal antibodies provides protection in rhesus macaques.* Proc Natl Acad Sci U S A, 2012. **109**(44): p. 18030-5.
18. Martinez, O., et al., *Impact of Ebola mucin-like domain on antiglycoprotein antibody responses induced by Ebola virus-like particles.* J Infect Dis, 2011. **204 Suppl 3**: p. S825-32.
19. Dowling, W., et al., *Influences of glycosylation on antigenicity, immunogenicity, and protective efficacy of ebola virus GP DNA vaccines.* J Virol, 2007. **81**(4): p. 1821-37.
20. Martins, K.A., T.K. Warren, and S. Bavari, *Characterization of a putative filovirus vaccine: virus-like particles.* Virol Sin, 2013. **28**(2): p. 65-70.
21. Warfield, K.L., et al., *Ebola virus-like particle-based vaccine protects nonhuman primates against lethal Ebola virus challenge.* J Infect Dis, 2007. **196 Suppl 2**: p. S430-7.
22. Warfield, K.L., et al., *Ebola virus-like particles protect from lethal Ebola virus infection.* Proc Natl Acad Sci U S A, 2003. **100**(26): p. 15889-94.
23. Warfield, K.L., et al., *Marburg virus-like particles protect guinea pigs from lethal Marburg virus infection.* Vaccine, 2004. **22**(25-26): p. 3495-502.
24. Swenson, D.L., et al., *Virus-like particles exhibit potential as a pan-filovirus vaccine for both Ebola and Marburg viral infections.* Vaccine, 2005. **23**(23): p. 3033-42.
25. Bray, M., et al., *A mouse model for evaluation of prophylaxis and therapy of Ebola hemorrhagic fever.* J Infect Dis, 1999. **179 Suppl 1**: p. S248-58.
26. Kyte, J. and R.F. Doolittle, *A simple method for displaying the hydropathic character of a protein.* J Mol Biol, 1982. **157**(1): p. 105-32.
27. Ou, W., et al., *Development and characterization of rabbit and mouse antibodies against ebolavirus envelope glycoproteins.* J Virol Methods, 2011. **174**(1-2): p. 99-109.
28. Kline, K.G. and M.R. Sussman, *Protein quantitation using isotope-assisted mass spectrometry.* Annu Rev Biophys, 2010. **39**: p. 291-308.

29. Chahrour, O., D. Cobice, and J. Malone, *Stable isotope labelling methods in mass spectrometry-based quantitative proteomics.* J Pharm Biomed Anal, 2015. **113**: p. 2-20.
30. Liebler, D.C. and L.J. Zimmerman, *Targeted quantitation of proteins by mass spectrometry.* Biochemistry, 2013. **52**(22): p. 3797-806.
31. Rochat, B., *Quantitative/qualitative analysis using LC-HRMS: the fundamental step forward for clinical laboratories and clinical practice.* Bioanalysis, 2012. **4**(14): p. 1709-11.
32. Huang, M.Q., Z.J. Lin, and N. Weng, *Applications of high-resolution MS in bioanalysis.* Bioanalysis, 2013. **5**(10): p. 1269-76.
33. Aebersold, R., A.L. Burlingame, and R.A. Bradshaw, *Western blots versus selected reaction monitoring assays: time to turn the tables?* Mol Cell Proteomics, 2013. **12**(9): p. 2381-2.
34. Taylor, A.E., B. Keevil, and I.T. Huhtaniemi, *Mass spectrometry and immunoassay: how to measure steroid hormones today and tomorrow.* Eur J Endocrinol, 2015. **173**(2): p. D1-12.
35. de La Vega, M.A., et al., *The multiple roles of sGP in Ebola pathogenesis.* Viral Immunol, 2015. **28**(1): p. 3-9.
36. Li, W., et al., *Characterization of Immune Responses Induced by Ebola Virus Glycoprotein (GP) and Truncated GP Isoform DNA Vaccines and Protection Against Lethal Ebola Virus Challenge in Mice.* J Infect Dis, 2015. **212 Suppl 2**: p. S398-403.

Figure Legends:

Figure 1. Target Peptide Selection and Characterization.

Panel A Top) Sequence alignment of the 3 proteins (GP_1, sGP and ssGP) derived from the Ebola GP transcript showing the locations of peptide candidates for use in the quantification of Ebola GP_1 (red dotted boxes) as well as the location of peptides rejected for the final assay (black boxes). All three proteins share sequence homology in the first 295 amino acids. Peptides identified in survey runs were evaluated for absence of post translational modifications, ionization efficiency and protein location.. **Panel A Bottom)** Schematic of fully processed $GP_{1,2}$ transmembrane protein, showing the location of the receptor binding site (RBS) and mucin-like domain (MLD) of $GP_{1,}$ as well as the extracellular domain (ECD), transmembrane region (TM) and cytoplasmic tail (CT) of GP_2. GP_1 and GP_2 are disulfide linked to form the mature $GP_{1,2}$ complex. **Panel B**. Standard curves for each target analyte peptide over a 9 point dilution

showing linearity from 275 fmols to 1 fmol total GP_1. An aliquot of the previously quantified eVLP lot 'A' (200 fmols/µL SEE at 120 µL dilution) was resuspended 83 µL 40% acetonitrile, 0.1% Formic (137.5 fmols/µL) and serially diluted. A 2 µL injection utilizing the described instrument method was run in triplicate for each dilution. R^2 values for all four peptides are well within the margin of significance for linearity. Also shown in tabular form are the %CV values for each triplicate XIC measurement for each peptide at each dilution. These data indicate linearity down to 1 fmol with the largest CV% (SVGR – 17.3%) in dilution number '8' of the serially diluted series. **Panel C**. AQUA-SVG peptide signal response for non-deamidated (circle) and deamidated (triangle) peptide. AQUA-SVG peptide was deamidated by incubating 40 pmols at 50°C/pH 8.0 for 2.5 days while a matching 40 pmol aliquot was stored at -20°C. A 5-point, 2-fold serial dilution was performed resulting in a 250 to 15.6 fmol/µL concentration range for each sample. LC-HRMS was run in triplicate on each dilution and the average counts plotted.

Figure 2. **Purification and characterization of a recombinant GP_1 standard**.

Panel A. Representative chromatogram of preparative C4 reverse phase HPLC of 300µg reduced recombinant GP material indicating fraction collection points. **Panel B**. SDS PAGE followed by silver-staining of fractions 1-7 showing the separation of GP_1 (top arrow) and GP_2 (bottom arrow). Material from fraction 1 was divided into 1.8µg aliquots and used for quantitation standard. **Panel C.** Silver stained SDS PAGE performed under reducing conditions comparing the rGP starting material and the purified rGP_1 standard. GP1 (top arrow) and GP2 (bottom arrow). **Panel D**. Western blot of eVLP Lot 'A', eVLP Lot 'E', unpurified rGP and the purified rGP_1 standard using the monoclonal antibody 6D8 showing the detection of fully

842 glycosylated GP_1 (arrow). **Panel E.** Western blot of eVLP Lot 'A', eVLP Lot 'E', unpurified rGP
843 and the purified rGP_1 standard using the monoclonal antibody H3D5 showing the detection of
844 fully glycosylated GP_1 (arrow) and GP protein fragments.

845

846 **Figure 3**. **Illustration of the eVLP quantitation method workflow, calculations and eVLP**
847 **GP_1 result table**.

848 **Panel A.** VLP digests are resuspended in 120 µL and five 2-fold serial dilutions performed.
849 Each dilution is mixed 1:1 with a solution containing 200 fmol/µL of each of the four
850 isotopically labeled AQUA peptides and run in triplicate using 2 µL injections. **Panel B.**
851 Method used for calculating the GP_1 concentration in the rGP_1 standard at each dilution.
852 Average XIC area counts from the 2+ and 3+ charge states are first summed for each AQUA
853 ($\mathbf{P_{aq}}$) and analyte ($\mathbf{P_v}$) peptide. The values from the SEE and IRSEE are summed to provide the
854 total counts for Peptide 'Set 1' and the SVG and SVGR values are summed to provide total
855 counts for Peptide 'Set 2'. The final quantitation is derived by comparing the relative response
856 of the 200 fmol AQUA standard to the endogenous analyte response at 4 ppm and averaging the
857 response between the 2 peptide sets. For absolute eVLP rGP_1 quantitation, only the values
858 derived from peptide set 1 (SEE/IRSEE) were used.

859

860 **Figure 4**. **Average MS/MS fragmentation spectra of the SVG and SEE peptide.**

861 Replicate CID fragmentation spectra and 300 ppm theoretical ion tables of the SVG (**A,B**) and
862 SEE (**C,D**) analyte peptides derived from eVLP lot 'A'. Panels A and B represent the y-series
863 assignments at the 1:2 (1) and 1:32 (5) dilution samples respectively (see dilution scheme in

Figure 3A). Prominent y-series sequence ions are indicated. The SVG series contains 10 consecutive y-series ions resulting in a MASCOT Ions Score of 69 and an Expect score of $6.7x10^{-6}$ at dilution '1' with the '5' dilution Ions Score at 57 with an Expect Score of $8.8x10^{-5}$. The SEE peptide contains 7 y-series ions in both the '1' and '5' dilutions with Ions Scores of 55 and 41 with Expect Scores of $8.2x10^{-5}$ and 0.0021 respectively. With the exception of the SEE '5' dilution, peptide assignments from every dilution run were of sufficient quality to obtain non-ambiguous sequence identifications.

Figure 5. **Survival data in mice for quantified eVLP lots and correlation with LC-HRMS and western blot quantitation**. **A)** Average percent survival (grey bars) after two vaccinations in mice (n=10 for each eVLP lot) with 10µg of GP_1 (as calculated from 6D8 WB results) of the indicated lot. Vaccinations were three weeks apart, with four weeks between the final vaccination and the challenge. The percent GP_1 to total protein concentration (right axis) obtained using LC-HRMS is represented by the white bars. Survival in all vaccinated groups was significant ($p<.005$) when compared to saline controls. Fisher's Exact test was used to compare survival between the group vaccinated with eVLP lot 'E' and the other lots (* indicates $p < 0.01$, ** indicates $p < .001$) **B)** LC-HRMS and western plot values for GP_1 content in each eVLP lot tested (based on % total protein or mg/ml) were plotted against percent survival in mice after EBOV challenge (since Lots A and B both gave 100% the lower lot B value was used resulting in a total of 4 datapoints). The strongest linear correlation ($R^2=0.9149$) was obtained with the LC-HRMS GP_1 values based on % total protein followed by the LC-HRMS values for GP_1 in mg/ml ($R^2=0.9025$).

887 **Supplemental Figure 1**. **XIC profiles of each peptide charge state used in the quantitation**
888 **showing the retention time alignment at 4 ppm**. Data was taken from eVLP lot 'B' dilution
889 '2' replicate. The 4 profiles from each peptide are ordered from AQUA 2+ (H 2+), analyte 2+ (L
890 2+), AQUA 3+ (H 3+) and analyte 3+ (L 3+). **Panel A** represents the SVG and SVGR peptide
891 Set 2, **Panel B** is the SEE and IRSEE peptide Set 1 and **Panel C** shows the XIC profiles of the
892 deamidated SVG and SVGR analyte peptides. XIC values were acquired with at least 7 data
893 points sampled across the elution profile.

894

895 **Supplemental Figure 2**. **The binary behavior of the 2 asparagines (N) within the SVG**
896 **peptide**. In order to assess whether the AQUA SVG peptide counts would show a similar
897 response to the deamidated analyte SVG peptide a 40 pmol aliquot of the AQUA peptide SVG
898 standard was incubated at 55°C for 2.5 days at pH 8.1 resulting in complete deamidation likely at
899 asparagine 9. Target m/z values were 969.4969 **(A)**, 646.6684 **(B)**, 969.0069 **(C)**, 646.3404 **(D)**,
900 969.9909 **(E)** and 646.9964 **(F)**. The most abundant species by far is the singly deamidated 2+
901 ion followed by the singly deamidated 3+ ion. The doubly deamidated ions comprise 5% of the
902 total counts. *****Note the complete absence of signal in the non-deamidated mass ranges.

903

904 **Supplemental Figure 3**. **LC-MS/MS protein identification of coomassie blue stained bands**
905 **(A)** derived from a 5 µg aliquot of the eVLP lot 'B' after SDS-PAGE analysis. The 10 highest
906 intensity bands from a coomassie stained 4-12% NuPAGE Novex Bis-Tris gel were excised and
907 digested with Trypsin/Lys-C. Each sample was analyzed on a 60 minute LC-MS/MS survey MS
908 run and searched with MASCOT (v. 2.4) against an SProt (2014_02) database with human and
909 *Ebola zaire* taxonomy specified. Locations of the 5 most prominent peptides are indicated as

well as the total number of GP_1 peptides observed for each sample (GP_1 peptide counts). A minimum Expect score of 0.005 with an FDR of 1% was used for peptide validation. No N-terminal peptides downstream of AA 192 were observed in gel bands 7 thru 10. These data indicate that while the majority of the GP_1 protein appears to be full-length, a significant number of N-terminal GP_1 fragments are present. **(B)**An identical sample was analyzed via western blot using the H3D5 antibody.

Table 1. Masses of analyte and AQUA standard peptides used for quantitation of GP_1 in eVLP

		Analyte		AQUA Standard	
Set	Sequence	2(+) m/z	3(+) m/z	2(+) m/z	3(+) m/z
1	SEELSFTAVSNR	670.3281	447.2211	675.3322	450.5572
1	IRSEELSFTAVSNR	804.9206	536.9495	809.9248	540.2856
2	SVGLNLEGNGVATDVPSATK	964.9998	643.669	969.0069	646.3404
2	SVGLNLEGNGVATDVPSATKR	1043.0498	695.7027	1048.0549	699.0388
2	SVGLNLEGNGVATDVPSATK-deam	965.4918	643.997	N/A	N/A
2	SVGLNLEGNGVATDVPSATKR-deam	1043.5424	696.0307	N/A	N/A

Table 2. HR/AM-MS method validation using purified recombinant GP_1 standard.

Trial	Dilution	Peptide	On-Column fmoles	On-Column ng	Total In Sample (µg)	Ave. (µg)	ΔS	Accuracy (%)	Precision (% CV)
1	60	Set 1	472.1	25.9	1.55	1.50	-7.00%	83.2	0.35
		Set 2	439.0	24.0	1.44				
2	60	Set 1	432.3	32.2	1.42	1.49	9.30%	82.6	
		Set 2	472.4	35.2	1.55				
3	80	Set 1	347.7	25.9	1.52	1.56	4.60%	86.6	3.2
		Set 2	363.7	27.1	1.59				
4	80	Set 1	363.3	27.0	1.59	1.64	5.50%	90.1	
		Set 2	383.4	28.5	1.68				

921

Table 3. GP_1 quantitation in three replicates of a single eVLP lot

Sample	Digest (µg)	fmoles Pep Set 1	fmoles Pep Set 2	GP_1 Aliquot (µg)	GP_1 Concentration (mg/ml)	GP_1 Pct. Total Protein	GP_1 ΔS (%)	GP_1 CV%
Lot A (1)	10	199.1	231.1	1.31	0.62	13.1	16.1	7.40%
Lot A (2)	10	172.3	207.1	1.13	0.53	11.3	20.2	
Lot A (3)	10	183.8	212.2	1.21	0.61	12.1	15.5	

922

923

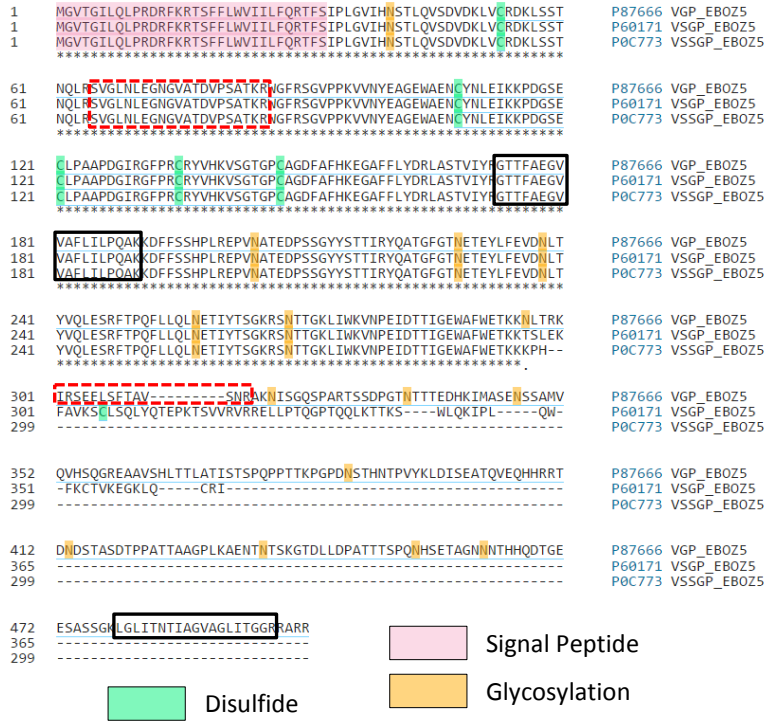
Table 4. LC-HRMS results of GP_1 quantitation in 5 lots of eVLP and comparison with quantitative western blot values

Lot	Concentration (mg/ml) Protein	rGP WB	MS	MS (CV%)	Average ΔS (%)	GP1 % TP
A	4.7	1.0	0.59	5.6%	15.8	12.6
B	3.8	1.14	0.35	10.3%	7.35	9.2
C	4.8	1.4	0.31	11.6%	12.7	6.5
D	4.9	0.7	0.16	8.8%	24.1	3.2
E	7.2	1.1	0.10	5.4%	25.5	1.3

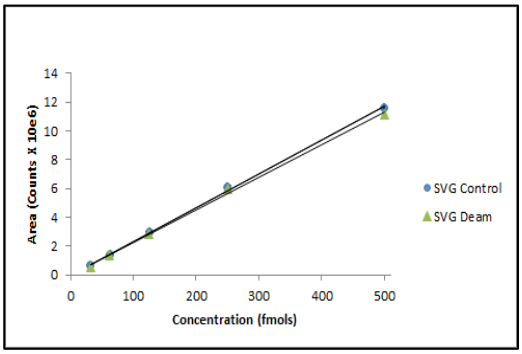
924

Figure 1

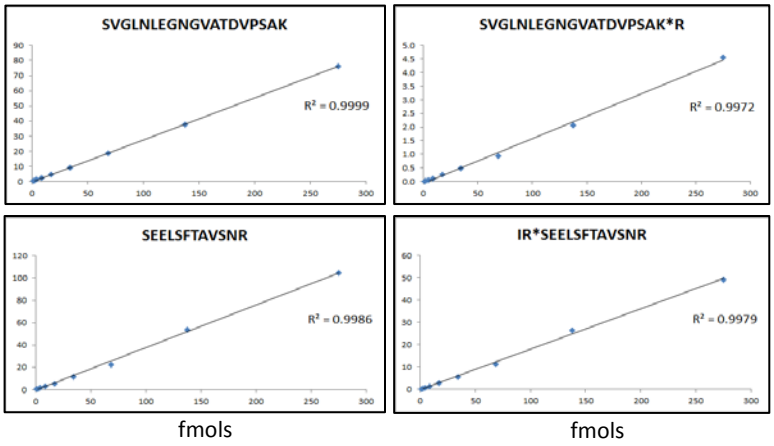
A



C

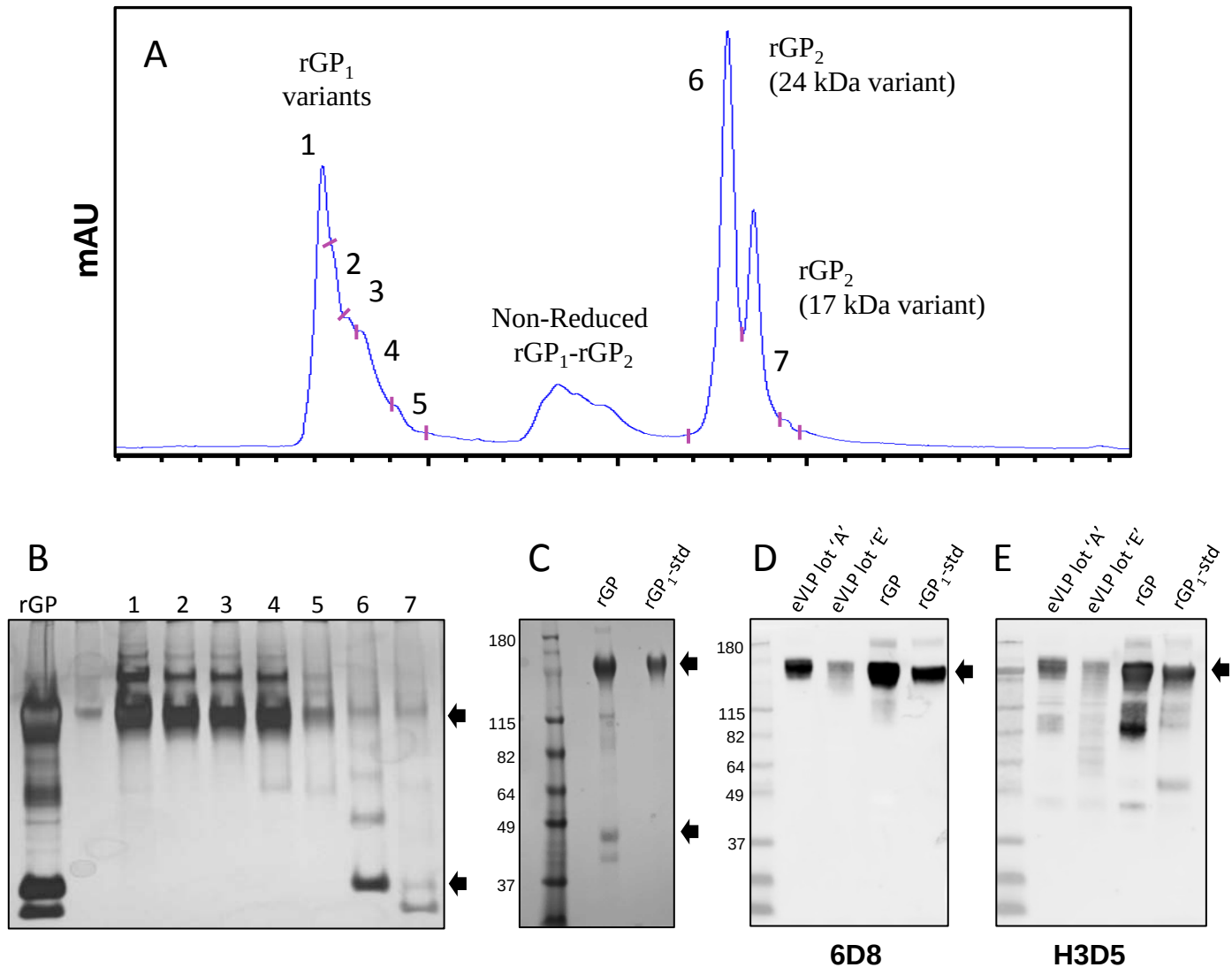


B

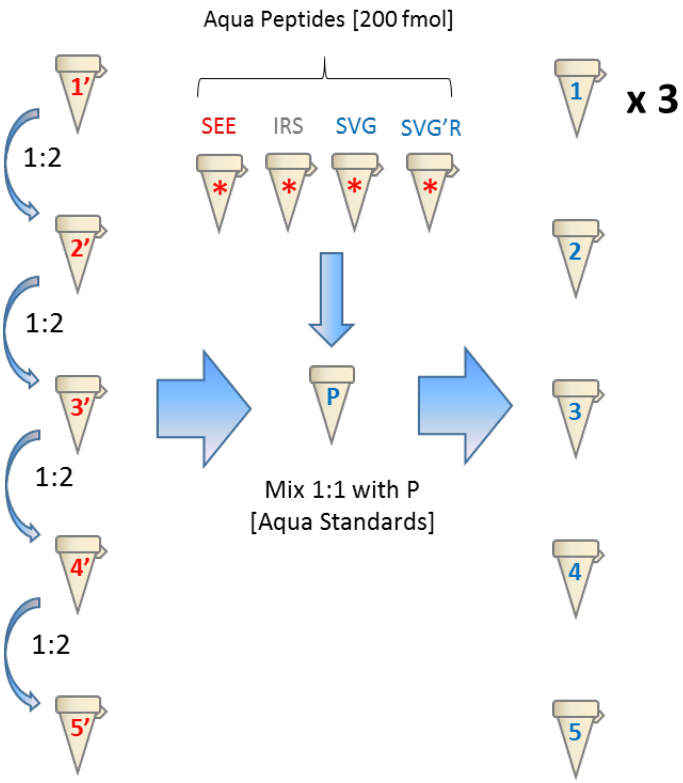


Peptide	Dilution (CV%)								
	1	2	3	4	5	6	7	8	9
SVG	3.9	2.3	5.1	3.0	2.8	12.9	2.8	10.8	11.9
SVGR	4.3	5.1	7.4	5.1	8.7	7.6	1.0	17.3	16.8
SEE	5.6	4.9	8.1	4.2	4.8	11.2	5.9	7.6	14.9
IRSEE	3.7	5.6	1.8	6.1	6.1	5.4	11.4	8.4	16.2

Figure 2



A Workflow


Aqua Peptides [200 fmol]
SEE
IRS
SVG
SVG'R
1'
2'
3'
4'
5'
1:2
1:2
1:2
1:2
P
Mix 1:1 with P
[Aqua Standards]
1
2
3
4
5
x 3

B

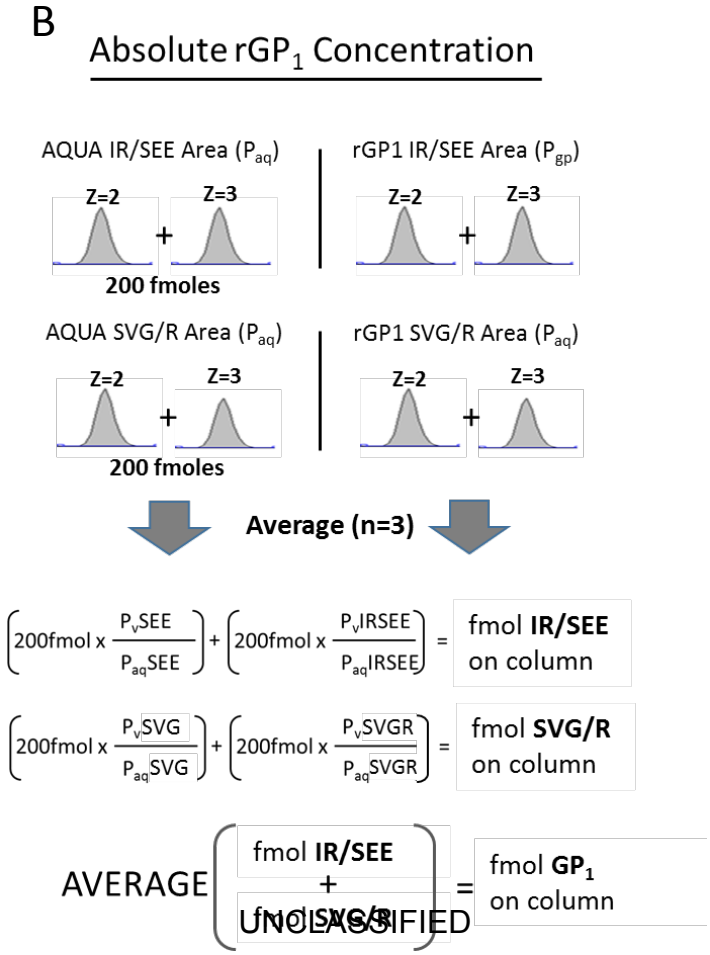
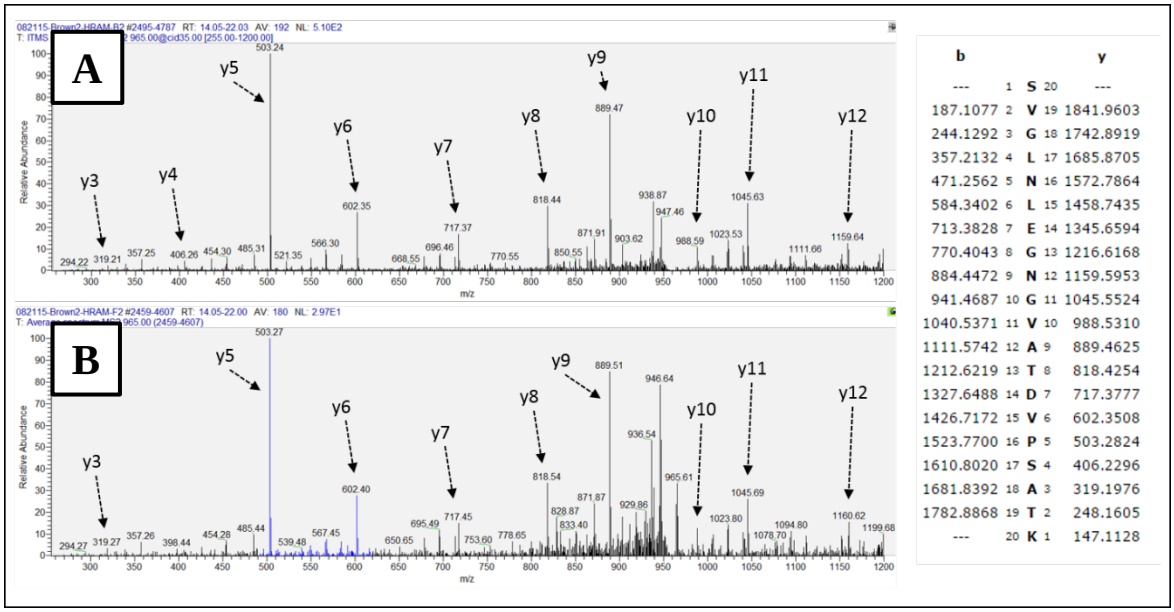

Absolute rGP1 Concentration
AQUA IR/SEE Area (Paq)
rGP1 IR/SEE Area (Pgp)
Z=2
Z=3
200 fmoles
AQUA SVG/R Area (Paq)
rGP1 SVG/R Area (Paq)
Average (n=3)
200fmol x PvSEE / PaqSEE + 200fmol x PvIRSEE / PaqIRSEE = fmol IR/SEE on column
200fmol x PvSVG / PaqSVG + 200fmol x PvSVGR / PaqSVGR = fmol SVG/R on column
AVERAGE (fmol IR/SEE + fmol SVG/R) = fmol GP1 on column
UNCLASSIFIED

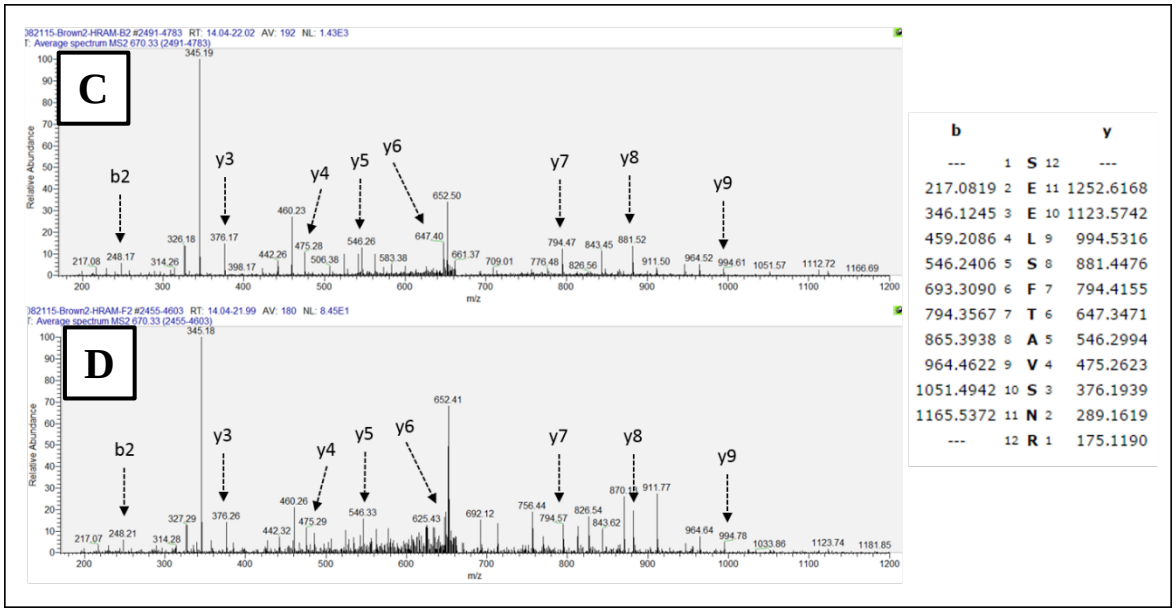
Figure 4

SVGLNLEGNGVATDVPSATK



b			y
---	1	S 20	---
187.1077	2	V 19	1841.9603
244.1292	3	G 18	1742.8919
357.2132	4	L 17	1685.8705
471.2562	5	N 16	1572.7864
584.3402	6	L 15	1458.7435
713.3828	7	E 14	1345.6594
770.4043	8	G 13	1216.6168
884.4472	9	N 12	1159.5953
941.4687	10	G 11	1045.5524
1040.5371	11	V 10	988.5310
1111.5742	12	A 9	889.4625
1212.6219	13	T 8	818.4254
1327.6488	14	D 7	717.3777
1426.7172	15	V 6	602.3508
1523.7700	16	P 5	503.2824
1610.8020	17	S 4	406.2296
1681.8392	18	A 3	319.1976
1782.8868	19	T 2	248.1605
---	20	K 1	147.1128

SEELSFTAVSNR



b			y
---	1	S 12	---
217.0819	2	E 11	1252.6168
346.1245	3	E 10	1123.5742
459.2086	4	L 9	994.5316
546.2406	5	S 8	881.4476
693.3090	6	F 7	794.4155
794.3567	7	T 6	647.3471
865.3938	8	A 5	546.2994
964.4622	9	V 4	475.2623
1051.4942	10	S 3	376.1939
1165.5372	11	N 2	289.1619
---	12	R 1	175.1190

Figure 5