

COMPOSITIONS AND METHODS RELATED TO EBOLAVIRUS VACCINES**CROSS-REFERENCE TO RELATED APPLICATIONS**

- 5 **[0001]** The subject patent application claims the benefit of priority to U.S. Provisional Patent Application Numbers 63/063,530 (filed August 10, 2020; now pending). The full disclosure of the priority application is incorporated herein by reference in its entirety and for all purposes.

STATEMENT OF GOVERNMENT SUPPORT

- 10 **[0002]** This invention was made with government support under AI129698 and AI123861 awarded by the National Institutes of Health. The government has certain rights in the invention.

BACKGROUND OF THE INVENTION

- 15 **[0003]** Filoviruses such as Zaire Ebola virus (EBOV) and Marburg virus can cause a lethal hemorrhagic fever in humans and nonhuman primates (NHPs). The virus glycoprotein, GP, mediates cell entry by initiating attachment and membrane fusion. Structures of GP-bound neutralizing antibodies (NAbs) from human survivors
- 20 demonstrated that GP harbors all NAb epitopes known to date and is the sole target for vaccine design. Like HIV-1 envelope glycoprotein (Env), filovirus GP is metastable and produces non-functional forms when expressed in various cell lines.
- 25 **[0004]** A lasting campaign has been undertaken to identify and characterize NAbs for EBOV and other filoviruses such as Marburg virus (MARV). As a result, large panels of NAbs were isolated from human survivors, vaccinated humans, and immunized animals. Crystallography and electron microscopy (EM) have revealed multiple sites of vulnerability on EBOV GP. A recent study of 171 monoclonal antibodies (mAbs) defined eight epitope classes, of which six can be recognized by broadly neutralizing antibodies (bNAbs). While recombinant virus-like particles
- 30 (VLPs) can protect against EBOV challenge in animals, the difficulties in manufacturing have hampered their further development as human vaccines.

[0005] Despite recent advances in the development of medical regimen against filoviruses, there is still a strong need in the art for more effective means for treating and preventing human infections and deaths caused by Ebolaviruses. The instant invention is directed to this and other unmet needs in the art.

5

SUMMARY OF THE INVENTION

- [0006] In one aspect, the invention provides engineered or redesigned Ebolavirus glycoprotein (GP) sequences. Relative to a wildtype Ebolavirus GP sequence, the engineered sequences contain (a) a substitution of residue W615 in heptad repeat 2 (HR2) with a smaller hydrophobic residue, and (b) one or more proline substitutions in heptad repeat 1 C segment (HR1_C). The amino acid numbering is based on *Zaire* Ebolavirus GP sequence that has the ectodomain sequence identified by UniProt ID Q05320 (SEQ ID NO:1). In various embodiments, the engineered Ebolavirus GP sequences can be modified or derived from the wildtype GP sequence of a *Zaire* Ebolavirus (EBOV), a *Sudan virus* (SUDV), a *Tai forest virus* (TAFV), a *Bundibugyo virus* (BDBV), or a *Reston Ebolavirus* (RESTV).
- [0007] In some embodiments, residue W615 in the engineered Ebolavirus GP sequences can be replaced with residue L, A, V, I or F. In some embodiments, the proline substitutions in HR1_C include T577P and/or L579P. In some embodiments, the engineered Ebolavirus GP sequences additionally include a truncation at the C-terminus of the membrane proximal external region (MPER). In some of these embodiments, the engineered Ebolavirus GP sequences can further include (a) a deletion of the mucin-like domain (MLD) deleted from the GP1 subunit, and/or (b) a deletion of MPER from the GP2 subunit.
- [0008] In another aspect, the invention provides engineered Ebolavirus GP sequences that contain (a) a truncated soluble GP of an Ebolavirus that has MLD deleted from the GP1 subunit and MPER deleted from the GP2 subunit, and (b) one or more modifications in the HR2 and HR1 regions. In various embodiments, the additional modifications in the HR2 and HR1 regions can be (i) substitution of W615 in heptad repeat 2 (HR2) with a smaller hydrophobic residue, (ii) an extension of HR2 at the C terminus, (iii) one or more proline substitutions in heptad repeat 1 C segment (HR1_C), and (iv) one or more engineered inter-GP disulfide bonds. Again, the amino acid numbering is based on the *Zaire* Ebolavirus GP sequence with UniProt ID Q05320

(SEQ ID NO:1). In various embodiments, these engineered Ebolavirus GP sequences can be derived from *Zaire Ebolaviruses* (EBOV), *Sudan viruses* (SUDV), *Tai forest viruses* (TAFV), *Bundibugyo viruses* (BDBV), or *Reston Ebolaviruses* (RESTV). Some engineered Ebolavirus GP sequences are based on the truncated soluble GP sequence shown in SEQ ID NO:2 or SEQ ID NO:3, or a conservatively modified or substantially identical variant thereof. In some embodiments, the engineered GP sequences can have a N-terminal leader sequence. As exemplification, the N-terminal leader sequence can contain SEQ ID NO:41 or a conservatively modified variant.

[0009] Some engineered soluble Ebolavirus GP sequences contain a substitution of residue W615. In some of these embodiments, residue W615 can be replaced with L, A, V, I or F. Some of these engineered Ebolavirus GP proteins contain the amino acid sequence shown in SEQ ID NOs:4, a conservatively modified variant or a substantially identical sequence thereof. In some embodiments, the engineered soluble Ebolavirus GP proteins can contain an extension of HR2 at the C-terminus. In some of these embodiments, the HR2 extension can be (a) extending HR2 C terminus with a N-terminal fragment of adjacent MPER of the Ebolavirus glycoprotein or (b) replacing a HR2 C-terminal fragment with a longer leucine zipper motif. As exemplification, HR2 extension in the engineered soluble Ebolavirus GP proteins can include (a) extension of the HR2 C terminus from residue 632 to residue 637 in MPER, i.e., addition of residues KTLPD (SEQID NO:32), (b) extension of the HR2 C terminus from residue 632 to residues 643 in MPER, i.e., addition of residues KTLPDQGDNDN (SEQ ID NO:33), or (c) replacement of residues 617-632 with a GCN4 leucine zipper sequence shown in SEQ ID NO:34.

[0010] Some engineered soluble Ebolavirus GP sequences of the invention contain both a HR2 extension noted above and a W615 substitution in HR2. In some of these embodiments, the W615 substitution can be W615L or P612G/W615F double mutation. In some embodiments, the engineered GP sequences contain (a) W615L substitution and (b) extension of the HR2 C terminus from residue 632 to residue 637 in MPER of the Ebolavirus GP. In some other embodiments, the engineered GP sequences contain (a) P612G/W615F double mutation in HR2 and (b) replacement of residues 617-632 with a GCN4 leucine zipper sequence shown in SEQ ID NO:34. Some specific engineered soluble Ebolavirus GP sequences contain an amino acid sequence shown in

any one of SEQ ID NOs:5-8, a conservatively modified variant or a substantially identical sequence thereof.

[0011] In addition to the HR2 extension and/or the W612 substitution, the engineered GP sequences in some embodiments can also contain a C-terminal trimerization motif. In some of these embodiments, the C-terminal trimerization motif can contain SEQ ID NO:29 or SEQ ID NO:30, or a conservatively modified variant or a substantially identical sequence thereof. In addition to the W612 substitution, the engineered GP sequences in some embodiments can also contain one or more proline substitutions in the HR1c segment. Any residue in HR1c can be replaced with a proline residue. In some of these embodiments, the proline substitution comprises T577P or L579P. Some specific examples of engineered Ebolavirus GP sequences having a W612 substitution in HR2 and a proline substitution in HR1c contain an amino acid sequence as set forth in any one of SEQ ID NOs:9-16, a conservatively modified variant or a substantially identical sequence thereof.

[0012] Some engineered GP sequences of the invention contain a HR2 extension, a proline substitution in HR1c, and a substitution at residue W615. In some of these embodiments, the W615 residue can be replaced with L, A, V, I or F. In some of the embodiments, the engineered Ebolavirus GP sequences contain an extension of HR2 C terminus from residue 632 to residue 637 in MPER. Some specific examples of such engineered Ebolavirus GP sequences contain an amino acid sequence as set forth in any one of SEQ ID NOs:17 and 20, a conservatively modified variant or a substantially identical sequence thereof. In some of the embodiments, the engineered Ebolavirus GP sequences can further include a C-terminal trimerization motif. For example, the engineered Ebolavirus GP proteins can include the C-terminal trimerization motif shown in SEQ ID NO:29 or SEQ ID NO:30, or a conservatively modified variant or a substantially identical sequence thereof.

[0013] Some engineered Ebolavirus GP sequences contain (a) W615L substitution, (b) proline substitution T577P or L579P, and (c) HR2 extension from residue 632 to residue 637 in MPER. Some specific examples of these engineered Ebolavirus GP sequences contain amino acid sequence as set forth in any one of SEQ ID NOs:18, 19, 21 and 22, a conservatively modified variant or a substantially identical sequence thereof.

[0014] Some engineered Ebolavirus GP sequences contain one or more engineered disulfide bonds between two neighboring GP protomers. In various embodiments, the engineered disulfide bonds can be created by cysteine substitutions at residue pairs G91/A575 (SS2), F153/Y534 (SS1), T520/A575 (SS3), G157/I532 (SS4), D522/A575 (SS5), and/or K56/G599 (SS6). Some specific examples of these engineered Ebolavirus GP sequences have an amino acid sequence are shown in any one of SEQ ID NOs:23-28, a conservatively modified variant or a substantially identical sequence thereof.

[0015] In some related aspects, the invention provides nucleic acid or polynucleotide sequences that encode an engineered full length or truncated Ebolavirus GP protein sequence described herein. Some engineered Ebolavirus GP sequences also contain a leader peptide encoding sequence at the 5'. Also provided in the invention are pharmaceutical compositions that contain an engineered Ebolavirus GP protein or encoding polynucleotide sequence described herein, and a pharmaceutically acceptable carrier.

[0016] In another aspect, the invention provides vaccine compositions that contain an engineered Ebolavirus GP protein that is displayed on the surface of a self-assembling nanoparticle. In some embodiments, C-terminus of the engineered Ebolavirus GP protein is fused to the N-terminus of subunit of the self-assembling nanoparticle (NP). In some embodiments, C-terminus of the engineered Ebolavirus GP protein is fused to the subunit of the self-assembling nanoparticle via a linker sequence GGGGS (SEQ ID NO:36) or (GGGGS)₂ (SEQ ID NO:42). In some embodiments, the engineered Ebolavirus GP proteins in the NP vaccine compositions contain an amino acid sequence as set forth in any one of SEQ ID NOs:4-28, a conservatively modified variant or a substantially identical sequence thereof. In some embodiments, subunit of the self-assembling nanoparticle contains the polypeptide sequence as shown in SEQ ID NO:36 (E2p), SEQ ID NO:37 (I3-01v9), or SEQ ID NO:38 (ferritin), a conservatively modified variant or a substantially identical sequence thereof.

[0017] Some NP scaffolded vaccine compositions of the invention can additionally contain a locking domain (LD) that is fused to the C terminus of the NP subunit sequence. In some of these embodiments, the employed LD contains the sequence shown in SEQ ID NO:39 or 40, a conservatively modified variant or a substantially identical sequence thereof. Some of the scaffolded vaccines can additionally contain a T-cell epitope that is fused to the C-terminus of the locking domain. In some of these

embodiments, the employed T-cell epitope contains the sequence shown in SEQ ID NO:31, a conservatively modified variant or a substantially identical sequence thereof.

[0018] Some NP vaccine compositions of the invention contain (1) a polypeptide sequence containing from N terminus to C terminus (a) an engineered Ebolavirus GP protein, linker sequence G₄S (SEQ ID NO:43), nanoparticle sequence ferritin, (b) an engineered Ebolavirus GP protein, linker sequence G₄S (SEQ ID NO:43), nanoparticle sequence E2p, or (c) an engineered Ebolavirus GP protein, linker sequence (G₄S)₂ (SEQ ID NO:42), nanoparticle sequence I3-01v9; or (2) a conservatively modified variant of the polypeptide sequence of (1). In some of these embodiments, the engineered Ebolavirus GP sequences contain W615L substitution, proline substitution T577P, and HR2 extension from residue 632 to residue 637 in MPER. These engineered Ebolavirus GP sequences can additionally contain a locking domain and/or a T-cell epitope that is fused to the C-terminus of the nanoparticle subunit sequence. In some embodiments, the vaccine compositions contain (1) a polypeptide sequence containing from N-terminus to C-terminus (a) the engineered Ebolavirus GP protein shown in SEQ ID NO:17, nanoparticle subunit sequence shown in SEQ ID NO:36 (E2p), locking domain shown in SEQ ID NO:39 (LD4), and T cell epitope shown in SEQ ID NO:31 (PADRE) or (b) the engineered Ebolavirus GP protein shown in SEQ ID NO:17, nanoparticle sequence shown in SEQ ID NO:37 (I3-01v9), locking domain shown in SEQ ID NO:40 (LD7), or (2) a conservatively modified variant of the polypeptide sequence.

[0019] A further understanding of the nature and advantages of the present invention may be realized by reference to the remaining portions of the specification and claims.

DESCRIPTION OF THE DRAWINGS

[0020] Figure 1 shows (A) a schematic structure of Ebolavirus glycoproteins (GPs) and (B) the general redesign strategy for generating engineered Ebolavirus GP vaccine immunogen proteins of the invention.

[0021] Figure 2 shows sequence alignment of the H1 (SEQ ID NOs:44-48) and HR2 (SEQ ID NOs:49-53) regions of the glycoproteins (GPs) of randomly selected strains from all 5 Ebolavirus species, Bundibugyo Ebolavirus (AYI50316), Zaire Ebolavirus (AER59712), Reston Ebolavirus (ASU06443), Sudan Ebolavirus (ALH21228) and Tai Forest Ebolavirus (AWK96625).

[0022] Figure 3 shows design, screening, and antigenic characterization of EBOV GP Δ muc trimers with modified stalk and HR1c. (A) Schematic representation of a mucin-deleted GP (GP Δ muc) and three stalk designs (GP Δ muc-2WPZ, GP Δ muc-L, and GP Δ muc-Ext). (B) SEC profiles of four GP Δ muc constructs obtained from a Superdex 200 10/300 column following transient expression in 250 ml HEK293 F cells and mAb114 purification. Left: the SEC curve of GP Δ muc shown in black line; Right: three stalk designs shown for GP Δ muc-2WPZ, -L, and -Ext, respectively. (C) BN-PAGE of GP Δ muc proteins purified by mAb114 (left) and mAb110 (right) columns. Of note, an additional GP Δ muc construct with the 8-aa HR1c replaced by a flexible (GS)₄ loop is included for comparison. (D) ELISA curves of four mAb100/SEC-purified, stalk-modified EBOV GP/GP Δ muc-foldon trimers binding to 10 antibodies. (E) Summary of EC₅₀ values (μ g/ml) of EBOV GP/GP Δ muc-foldon trimers binding to 10 antibodies. (F) ELISA curves of an mAb100/SEC-purified, stalk/HR1c-modified EBOV GP Δ muc-foldon trimer binding to 10 antibodies. (G) Summary of EC₅₀ values (μ g/ml) of a stalk/HR1c-modified EBOV GP Δ muc-foldon trimer binding to 10 antibodies. (H) Plot of EC₅₀ (μ g/ml) values of five GP/GP Δ muc-foldon trimers binding to 10 antibodies. Four pan-Ebolavirus NAbs are indicated. In (E) and (G), EC₅₀ values were calculated for all ELISA data in Prism version 8.4.3.

[0023] Figure 4 shows design and characterization of multilayered EBOV GP Δ muc-presenting NPs. (A) Locking domains (LD) identified from the Protein Data Bank (PDB) for stabilizing the NP-forming interface. Shown in the figure are the PDB IDs of the identified LDs and their full sequences (SEQ ID NOs:54-62). Residues deleted at the N- and/or C-termini of the original sequences for generating actual LDs used in the studies herein are noted in the sequences. (B)-(C) Antigenic evaluation by ELISA. Binding of FR and reengineered E2p (E2p-LD4-PADRE, or E2p-L4P) and I3-01 (I3-01-LD7-PADRE, I3-01-L7P) NPs that present a stabilized GP Δ muc-WL²P² trimer to 10 antibodies with ELISA curves shown in (B) and EC₅₀ values summarized in (C). All GP-NPs were expressed transiently in ExpiCHO cells, purified by a mAb100 column, and further purified by SEC on a Superose 6 10/300 GL column. (D)-(G) Antigenic evaluation by BLI using Octet96. Binding kinetics of GP Δ muc-WL²P²-foldon (D) and GP Δ muc-WL²P² on FR (E), E2p-L4P (F), and I3-01-L7P (G) NPs to 10 antibodies. BLI sensorgrams were obtained from an Octet RED96 instrument using a

series of six concentrations (400-12.5nM by twofold dilution for GPΔmuc; 25-0.78nM and 10-0.31nM by twofold dilution for FR and E2p-L4p/I3-01-L7P, respectively) and quantitation biosensors (see Materials and Methods). (H) SEC profile of GPΔmuc-WL²P²-10GS-I3-01v9-LD7. I3-01v9 is a variant of I3-01 with a redesigned NP-forming interface based on the PDB structure (1VLW). I3-01v9 was reported in our previous study with a construct name “1VLW-V9”.

[0024] Figure 5 shows immunogenicity assessment of EBOV GP/GPΔmuc trimers and GPΔmuc-presenting NPs in BALB/c mice. (A) ELISA binding curves of mouse serum from three trimer groups, in which mice were immunized with WT GP-foldon, GPΔMuc-foldon, and GPΔMuc-WL²P²-fold trimers, and three NP groups, in which mice were immunized with FR, E2p-L4P, and I3-01v9-L7P NPs presenting GPΔMuc-WL²P², to GPΔMuc-WL²P² at w2, w5, w8, and w11. (B) EC₅₀ titers measured for the three trimer groups. (C) EC₅₀ titers measured for the three NP groups. The EC₅₀ titer was measured in the unit of fold of dilution. Of note, serum binding at w2 did not reach the plateau (or saturation) to allow for accurate determination of EC₅₀ titers. Nonetheless, the EC₅₀ values calculated in Prism were used as a quantitative measure of binding antibody titers to facilitate the comparison of different vaccine groups at w2. (D) Ebolavirus-pp (Makona C15) neutralization by 10 Ebolavirus-specific antibodies in TZM-bl cells. (E) MLV-pp neutralization by 10 Ebolavirus-specific antibodies in 293 T cells. (F) MLV-pp neutralization by purified mouse serum IgGs in 293 T cells. Mouse IgGs were collected from all six vaccine groups at w11.

[0025] Figure 6 shows immunogenicity assessment of EBOV GP/GPΔmuc trimers and GPΔmuc-presenting NPs in rabbits. (A) ELISA binding curves of rabbit serum from two trimer groups, in which rabbits were immunized with GPΔMuc-foldon and GPΔMuc-WL²P²-fold trimers, and three NP groups, in which rabbits were immunized with FR, E2p-L4P, and I3-01v9-L7P NPs presenting GPΔMuc-WL²P², to GPΔMuc-WL²P² at w0, w2, w5, w8, w11, and w13. (B) EC₅₀ titers measured for the two trimer groups and three NP groups. The EC₅₀ titer was measured in the unit of fold of dilution. Of note, serum binding at w2 did not reach the plateau (or saturation) but the OD₄₅₀ values were approximately at the same level as those obtained at w5 (fully plateaued). Therefore, the EC₅₀ values calculated in Prism were considered sufficiently accurate to facilitate the comparison of different vaccine groups at w2. Ebolavirus-pp (EBOV-

Makona and BDBV) and MLV-pp neutralization by purified rabbit IgGs from w2, w5, and w8 in 293 T cells is shown in (C), (D), (E) and (F). Data points that do not fit a sigmodal dose response model are shown as dashed lines. The signs of ADE for the two GPΔmuc trimer groups, the FR group, and the two multilayered 60-meric NP groups are w2, w5, and w8, respectively, which are highlighted in dashed rectangles.

[0026] Figure 7 shows unbiased NGS repertoire analysis of bulk-sorted Ebola GP-specific mouse splenic B cells. (A) SEC profile of a biotinylated Avi-tagged Ebola GPΔmuc trimer, termed GPΔmuc-WL²P²-1TD0-Avi-Biot, obtained from a HiLoad Superdex 200 16/600 column. 1TD0 is a trimeric scaffold used to stabilize GPΔmuc in a trimeric state and to deselect B cells directed to foldon in the two trimer groups. (B) Summary of Ebola GPΔmuc-specific bulk sorting of mouse splenic B cells from five vaccine groups. (C) Antibodyomics analysis of NGS data obtained for Ebola GPΔmuc-specific mouse splenic B cells. Distribution of critical B-cell properties are shown for three vaccine groups, in which mice were immunized with (D) GPΔmuc-foldon, (E) GPΔmuc-WL²P²-5GS-FR, and (F) GPΔmuc-WL²P²-10GS-I3-01v9-L7P. For each group, distributions are shown for germline VH/VK gene usage (top), germline VH/VK divergence (bottom left) and CDRH/K3 loop length (bottom right).

DETAILED DESCRIPTION

I Overview

[0027] Ebolaviruses can cause severe hemorrhagic fever. There are five virus species in the Ebolavirus genus, Ebola virus (aka *Zaire Ebolavirus*; EBOV), *Sudan virus* (SUDV), *Tai forest virus* (TAFV), *Bundibugyo virus* (BDBV), and *Reston Ebolavirus* (RESTV). Of these, EBOV, SUDV, TAFV and BDBV all have pandemic potential in humans. For example, EBOV was solely responsible for the largest filovirus outbreak in history during 2013-2016, which spread across nine African countries with 28,600 cases and 11,325 deaths. A previous EBOV outbreak led to 2103 deaths and was declared an international emergency on July 17, 2019 by the World Health Organization (WHO). In recent years, significant progress has been made to counter this deadly virus. Neutralizing antibodies (NAbs) have now been established as effective therapeutics for EBOV infection. Vaccines based on diverse delivery systems have been tested in humans. rVSV-ZEBOV, a replication-competent recombinant

vesicular stomatitis virus (VSV) expressing a *Zaire* EBOV glycoprotein (GP), is the most advanced vaccine candidate that was deployed during the 2018-2019 outbreak. However, GP-specific antibody titers were not noticeably increased seven days after vaccination with rVSV-ZEBOV in humans, contrasting prior findings in nonhuman primates (NHPs). In addition, a recent study reported the neurotropism of rVSV-ZEBOV which caused damage to the eye and brain in neonatal mice. Antibody-dependent enhancement (ADE) of infection was found for EBOV antibodies isolated from human survivors, suggesting that weak or non-NAbs induced by a suboptimal vaccine may cause adverse effects. Currently, no EBOV vaccine has been approved by the U.S. Food and Drug Administration (FDA) for use in humans.

[0028] The Ebola virus (EBOV) glycoprotein (GP) can be recognized by neutralizing antibodies (NAbs) and is the main target for vaccine design. The present invention is derived in part from the inventors' studies to rationally redesign the Ebola virus glycoprotein and engineer single-component multilayered nanoparticles as vaccine candidates. As detailed herein, the inventors explored the causes of EBOV GP metastability and designed multilayered NP vaccines for in vivo evaluation. To facilitate GP purification, the inventors developed an immunoaffinity column based on mAb100 that is specific for native-like, trimeric GP. The inventors first examined the contribution of two regions in GP2, namely, the HR2 (the "stalk") region and the heptad repeat 1 C (HR1c) segment, to GP metastability in a mucin-deleted *Zaire* EBOV GP construct (GP Δ muc). The inventors extended the HR2 stalk to residue 637 and introduced a W615L mutation based on the comparison of EBOV and MARV GPs. The inventors assessed the ability of proline mutation in HR1c to prevent GP refolding from pre- to post-fusion conformational changes. While both stalk and HR1c-proline mutations increased the trimer yield, the latter appeared to exhibit a complex effect on GP thermostability. Inter-GP molecule disulfide bonds (SS) were also found to increase the trimer stability. Crystal structures were determined for two redesigned GP Δ muc constructs to validate the stalk and HR1c-proline mutations at the atomic level. The inventors then displayed a redesigned GP Δ muc trimer on ferritin, E2p, and I3-01 NPs. Locking domains (LD) and helper T-cell epitopes were incorporated into E2p and I3-01 60-mers to stabilize the NP shell from the inside and to create multilayered NP carriers.

[0029] In mice and rabbits, it was observed that the GP trimers and NPs can induce distinct antibody responses. Next-generation sequencing (NGS) of GP-specific B cells

showed different profiles for NPs presenting large trimeric spikes versus presenting small antigens such as hepatitis C virus (HCV) E2 core. These studies demonstrate the critical factors of EBOV GP metastability, enable two single-component multilayered self-assembling NPs for designing VLP-type vaccines, and provide EBOV vaccine candidates that warrant further evaluation in NHPs and humans.

5 [0030] In accordance with these studies, the invention provides novel engineered Ebolavirus GP sequences that contain the various modifications disclosed herein. Some embodiments of the invention are directed to immunogen polypeptides and polynucleotide vaccines that are derived from the redesigned Ebolavirus GP sequences described herein. Also provided in the invention are Ebolavirus vaccine compositions containing a displaying platform, including a self-assembling nanoparticle, that displays one or more of the engineered Ebolavirus GP immunogens. Therapeutic applications of the engineered Ebolavirus GP immunogen polypeptides and the related nanoparticle vaccine compositions, e.g., treating or preventing Ebolavirus infections, are also provided in the invention.

15 [0031] Unless otherwise specified herein, the vaccine immunogens of the invention, the encoding polynucleotides, expression vectors and host cells, as well as the related therapeutic applications, can all be generated or performed in accordance with the procedures exemplified herein or routinely practiced methods well known in the art. See, e.g., Methods in Enzymology, Volume 289: Solid-Phase Peptide Synthesis, J. N. Abelson, M. I. Simon, G. B. Fields (Editors), Academic Press; 1st edition (1997) (ISBN-13: 978-0121821906); U.S. Pat. Nos. 4,965,343, and 5,849,954; Sambrook et al., Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Press, N.Y., (3rd ed., 2000); Brent et al., Current Protocols in Molecular Biology, John Wiley & Sons, Inc. (ringbou ed., 2003); Davis et al., Basic Methods in Molecular Biology, Elsevier Science Publishing, Inc., New York, USA (1986); or Methods in Enzymology: Guide to Molecular Cloning Techniques Vol. 152, S. L. Berger and A. R. Kimmerl Eds., Academic Press Inc., San Diego, USA (1987); Current Protocols in Protein Science (CPPS) (John E. Coligan, et. al., ed., John Wiley and Sons, Inc.), Current Protocols in Cell Biology (CPCB) (Juan S. Bonifacino et. al. ed., John Wiley and Sons, Inc.), and Culture of Animal Cells: A Manual of Basic Technique by R. Ian Freshney, Publisher: Wiley-Liss; 5th edition (2005), Animal Cell Culture Methods (Methods in Cell Biology, Vol. 57, Jennie P. Mather and David Barnes editors, Academic Press, 1st

edition, 1998). The following sections provide additional guidance for practicing the compositions and methods of the present invention.

II. Definitions

- 5 **[0032]** Unless defined otherwise, all technical and scientific terms used herein have the same meaning as commonly understood by those of ordinary skill in the art to which this invention pertains. The following references provide one of skill with a general definition of many of the terms used in this invention: *Academic Press Dictionary of Science and Technology*, Morris (Ed.), Academic Press (1st ed., 1992);
- 10 *Oxford Dictionary of Biochemistry and Molecular Biology*, Smith et al. (Eds.), Oxford University Press (revised ed., 2000); *Encyclopaedic Dictionary of Chemistry*, Kumar (Ed.), Anmol Publications Pvt. Ltd. (2002); *Dictionary of Microbiology and Molecular Biology*, Singleton et al. (Eds.), John Wiley & Sons (3rd ed., 2002); *Dictionary of Chemistry*, Hunt (Ed.), Routledge (1st ed., 1999); *Dictionary of Pharmaceutical*
- 15 *Medicine*, Nahler (Ed.), Springer-Verlag Telos (1994); *Dictionary of Organic Chemistry*, Kumar and Anandand (Eds.), Anmol Publications Pvt. Ltd. (2002); and *A Dictionary of Biology (Oxford Paperback Reference)*, Martin and Hine (Eds.), Oxford University Press (4th ed., 2000). Further clarifications of some of these terms as they apply specifically to this invention are provided herein.
- 20 **[0033]** As used herein, the singular forms "a," "an," and "the," refer to both the singular as well as plural, unless the context clearly indicates otherwise. For example, "an Env-derived trimer" can refer to both single or plural Env-derived trimer molecules, and can be considered equivalent to the phrase "at least one Env-derived trimer."
- [0034]** As used herein, the terms "antigen" or "immunogen" are used
- 25 interchangeably to refer to a substance, typically a protein, which is capable of inducing an immune response in a subject. The term also refers to proteins that are immunologically active in the sense that once administered to a subject (either directly or by administering to the subject a nucleotide sequence or vector that encodes the protein) is able to evoke an immune response of the humoral and/or cellular type
- 30 directed against that protein. Unless otherwise noted, the term "vaccine immunogen" is used interchangeably with "protein antigen" or "immunogen polypeptide".
- [0035]** The term "conservatively modified variant" applies to both amino acid and nucleic acid sequences. With respect to particular nucleic acid sequences,

- conservatively modified variants refers to those nucleic acids which encode identical or essentially identical amino acid sequences, or where the nucleic acid does not encode an amino acid sequence, to essentially identical sequences. Because of the degeneracy of the genetic code, a large number of functionally identical nucleic acids encode any given protein. For polypeptide sequences, "conservatively modified variants" refer to a variant which has conservative amino acid substitutions, amino acid residues replaced with other amino acid residue having a side chain with a similar charge. Families of amino acid residues having side chains with similar charges have been defined in the art. These families include amino acids with basic side chains (e.g., lysine, arginine, histidine), acidic side chains (e.g., aspartic acid, glutamic acid), uncharged polar side chains (e.g., glycine, asparagine, glutamine, serine, threonine, tyrosine, cysteine), nonpolar side chains (e.g., alanine, valine, leucine, isoleucine, proline, phenylalanine, methionine, tryptophan), beta-branched side chains (e.g., threonine, valine, isoleucine) and aromatic side chains (e.g., tyrosine, phenylalanine, tryptophan, histidine).
- 15 **[0036]** Epitope refers to an antigenic determinant. These are particular chemical groups or peptide sequences on a molecule that are antigenic, such that they elicit a specific immune response, for example, an epitope is the region of an antigen to which B and/or T cells respond. Epitopes can be formed both from contiguous amino acids or noncontiguous amino acids juxtaposed by tertiary folding of a protein.
- 20 **[0037]** Effective amount of a vaccine or other agent that is sufficient to generate a desired response, such as reduce or eliminate a sign or symptom of a condition or disease, such as a viral infection. For instance, this can be the amount necessary to inhibit viral replication or to measurably alter outward symptoms of the viral infection. In general, this amount will be sufficient to measurably inhibit virus (for example, an
- 25 Ebolavirus) replication or infectivity. When administered to a subject, a dosage will generally be used that will achieve a target concentration that has been shown to be sufficient for in vitro inhibition of viral replication. In some embodiments, an "effective amount" is one that treats (including prophylaxis) one or more symptoms and/or underlying causes of any of a disorder or disease, for example to treat Ebolavirus
- 30 infection. In some embodiments, an effective amount is a therapeutically effective amount. In some embodiments, an effective amount is an amount that prevents one or more signs or symptoms of a particular disease or condition from developing, such as one or more signs or symptoms associated with an Ebolavirus infection.

- [0038] As used herein, a fusion protein is a recombinant protein containing amino acid sequence from at least two unrelated proteins that have been joined together, via a peptide bond, to make a single protein. The unrelated amino acid sequences can be joined directly to each other or they can be joined using a linker sequence. As used
- 5 herein, proteins are unrelated, if their amino acid sequences are not normally found joined together via a peptide bond in their natural environment(s) (e.g., inside a cell). For example, the amino acid sequences of bacterial enzymes such as *B. stearothermophilus* dihydrolipoyl acyltransferase (E2p) and the amino acid sequences of Ebola virus GP are not normally found joined together via a peptide bond.
- 10 [0039] The term "Ebola virus", refers to members of the family *Filoviridae*, which are associated with outbreaks of highly lethal hemorrhagic fever in humans and nonhuman primates. Human Ebola virus pathogens include Ebola virus (EBOV), Sudan virus (SUDV), Bundibugyo virus (BDBV), and Taï Forest virus (TAFV). Reston virus (RESTV) is a monkey pathogen and is not currently considered a human pathogen.
- 15 The natural reservoir of the Ebolaviruses is unknown, and there are currently no available vaccines or effective therapeutic treatments for Ebola virus infections. The Ebola virus genome consists of a single strand of negative sense RNA that is approximately 19 kb in length. This RNA contains seven sequentially arranged genes that produce 8 mRNAs upon infection. Ebola virions contain seven proteins: a surface
- 20 glycoprotein (GP), a nucleoprotein (NP), four virion structural proteins (VP40, VP30, VP35, and VP24), and an RNA-dependent RNA polymerase (L) (Feldmann et al. (1992) *Virus Res.* 24, 1-19; Sanchez et al., (1993) *Virus Res.* 29, 215-240; reviewed in Peters et al. (1996) In *Fields Virology*, Third ed. pp. 1161-1176. Fields, B. N., Knipe, D. M., Howley, P. M., et al. eds. Lippincott-Raven Publishers, Philadelphia).
- 25 [0040] The term "Ebola virus glycoprotein (GP)" refers to the only surface antigen of Ebolaviruses that is expressed as a trimer on the viral surface. It is unusual in that it is encoded in two open reading frames. Transcriptional editing is needed to express the transmembrane form that is incorporated into the virion (Sanchez et al. (1996) *Proc. Natl. Acad. Sci. USA* 93, 3602-3607; Volchkov et al. (1995) *Virology* 214, 421-430).
- 30 The unedited form produces a nonstructural secreted glycoprotein (sGP) that is synthesized in large amounts early during the course of infection. The GP of Ebolaviruses is a transmembrane glycoprotein (GP) that is responsible for both receptor binding and membrane fusion. During assembly of the virus, the glycoprotein

undergoes proteolytic cleavage by host proteases such as furin, resulting in the two subunits, GP1 and GP2, which are linked by a disulfide bond. The GP1 subunit (amino acids 33–501) contains the core of the glycoprotein, receptor binding domain (RBD), a glycan cap, and a large mucin-like domain (MLD) which extends around the RBD. The GP2 (amino acids 502–676) subunit contains the internal fusion loop (IFL), heptad repeats 1 and 2 (HR1 and HR2), the membrane-proximal external region (MPER), the transmembrane region (TM), and the cytoplasmic tail (CT). During the transport of Ebolavirus particles to late endosomes, low pH leads to proteolytic processing of GPs by host cysteine proteases such as cathepsins, and the exposed receptor binding site of the proteolytically digested GP is thought to interact with a host receptor, Niemann Pick C1, followed by membrane fusion. Unless otherwise noted, amino acid residue numbering of the GP from Zaire Ebolavirus strain Mayinga-76 is used herein as the reference, which has a complete genome sequence identified by GenBank ID of AF272001. Its GP ectodomain sequence (UniProt ID Q05320; GenBank ID AAG40168.1) is shown in SEQ ID NO:1 herein.

[0041] As used herein, corresponding residues refers to residues that occur at aligned loci. Related or variant polypeptides are aligned by any method known to those of skill in the art. Such methods typically maximize matches, and include methods such as using manual alignments and by using the numerous alignment programs available (for example, BLASTP) and others known to those of skill in the art. By aligning the sequences of polypeptides, one skilled in the art can identify corresponding residues, using conserved and identical amino acid residues as guides. For example, by aligning the sequences of GP polypeptides from different Ebolavirus species or different isolates of the same species, one of skill in the art can identify corresponding residues, using conserved and identical amino acid residues as guides. Corresponding positions also can be based on structural alignments, for example by using computer simulated alignments of protein structure.

[0042] Immunogen as used herein refers to a protein or a portion thereof that is capable of inducing an immune response in a mammal, such as a mammal infected or at risk of infection with a pathogen. Administration of an immunogen can lead to protective immunity and/or proactive immunity against a pathogen of interest.

[0043] Immune response refers to a response of a cell of the immune system, such as a B cell, T cell, or monocyte, to a stimulus. In some embodiment, the response is

specific for a particular antigen (an "antigen-specific response"). In some embodiments, an immune response is a T cell response, such as a CD4⁺ response or a CD8⁺ response. In some other embodiments, the response is a B cell response, and results in the production of specific antibodies.

- 5 **[0044]** Immunogenic composition refers to a composition comprising an immunogenic polypeptide that induces a measurable CTL response against virus expressing the immunogenic polypeptide, or induces a measurable B cell response (such as production of antibodies) against the immunogenic polypeptide.

- 10 **[0045]** Sequence identity or similarity between two or more nucleic acid sequences, or two or more amino acid sequences, is expressed in terms of the identity or similarity between the sequences. Sequence identity can be measured in terms of percentage identity; the higher the percentage, the more identical the sequences are. Two sequences are "substantially identical" if two sequences have a specified percentage of amino acid residues or nucleotides that are the same (*i.e.*, 60% identity, 15 optionally 65%, 70%, 75%, 80%, 85%, 90%, 95%, or 99% identity over a specified region, or, when not specified, over the entire sequence), when compared and aligned for maximum correspondence over a comparison window, or designated region as measured using one of the following sequence comparison algorithms or by manual alignment and visual inspection. Optionally, the identity exists over a region that is at 20 least about 50 nucleotides (or 10 amino acids) in length, or more preferably over a region that is 100 to 500 or 1000 or more nucleotides (or 20, 50, 200 or more amino acids) in length.

- 25 **[0046]** Homologs or orthologs of nucleic acid or amino acid sequences possess a relatively high degree of sequence identity/similarity when aligned using standard methods. Methods of alignment of sequences for comparison are well known in the art. Various programs and alignment algorithms are described in: Smith & Waterman, Adv. Appl. Math. 2:482, 1981; Needleman & Wunsch, J. Mol. Biol. 48:443, 1970; Pearson & Lipman, Proc. Natl. Acad. Sci. USA 85:2444, 1988; Higgins & Sharp, Gene, 73:237-44, 1988; Higgins & Sharp, CABIOS 5:151-3, 1989; Corpet et al., Nuc. Acids Res. 30 16:10881-90, 1988; Huang et al. Computer Appls. in the Biosciences 8, 155-65, 1992; and Pearson et al., Meth. Mol. Bio. 24:307-31, 1994. Altschul et al., J. Mol. Biol. 215:403-10, 1990, presents a detailed consideration of sequence alignment methods and homology calculations.

- [0047] The term "subject" refers to any animal classified as a mammal, e.g., human and non-human mammals. Examples of non-human animals include dogs, cats, cattle, horses, sheep, pigs, goats, rabbits, and etc. Unless otherwise noted, the terms "patient" or "subject" are used herein interchangeably. Preferably, the subject is human.
- 5 [0048] The term "treating" or "alleviating" includes the administration of compounds or agents to a subject to prevent or delay the onset of the symptoms, complications, or biochemical indicia of a disease (e.g., an Ebolavirus infection), alleviating the symptoms or arresting or inhibiting further development of the disease, condition, or disorder. Subjects in need of treatment include those already suffering
- 10 from the disease or disorder as well as those being at risk of developing the disorder. Treatment may be prophylactic (to prevent or delay the onset of the disease, or to prevent the manifestation of clinical or subclinical symptoms thereof) or therapeutic suppression or alleviation of symptoms after the manifestation of the disease.
- [0049] Vaccine refers to a pharmaceutical composition that elicits a prophylactic
- 15 or therapeutic immune response in a subject. In some cases, the immune response is a protective immune response. Typically, a vaccine elicits an antigen-specific immune response to an antigen of a pathogen, for example a viral pathogen, or to a cellular constituent correlated with a pathological condition. A vaccine may include a polynucleotide (such as a nucleic acid encoding an Ebolavirus GP polypeptide
- 20 disclosed herein), a peptide or polypeptide (such as an Ebolavirus GP polypeptide disclosed antigen), a virus, a cell or one or more cellular constituents. In some embodiments of the invention, vaccines or vaccine immunogens or vaccine compositions are expressed from fusion constructs and self-assemble into nanoparticles displaying an immunogen polypeptide or protein on the surface.
- 25 [0050] Virus-like particle (VLP) refers to a non-replicating, viral shell, derived from any of several viruses. VLPs are generally composed of one or more viral proteins, such as, but not limited to, those proteins referred to as capsid, coat, shell, surface and/or envelope proteins, or particle-forming polypeptides derived from these proteins. VLPs can form spontaneously upon recombinant expression of the protein in
- 30 an appropriate expression system. Methods for producing particular VLPs are known in the art. The presence of VLPs following recombinant expression of viral proteins can be detected using conventional techniques known in the art, such as by electron microscopy, biophysical characterization, and the like. See, for example, Baker et al.

(1991) Biophys. J. 60:1445-1456; and Hagensee et al. (1994) J. Virol. 68:4503-4505. For example, VLPs can be isolated by density gradient centrifugation and/or identified by characteristic density banding. Alternatively, cryoelectron microscopy can be performed on vitrified aqueous samples of the VLP preparation in question, and images
5 recorded under appropriate exposure conditions.

[0051] A self-assembling nanoparticle refers to a ball-shape protein shell with a diameter of tens of nanometers and well-defined surface geometry that is formed by identical copies of a non-viral protein capable of automatically assembling into a nanoparticle with a similar appearance to VLPs. Known examples include ferritin
10 (FR), which is conserved across species and forms a 24-mer, as well as *B. stearotherophilus* dihydrolipoyl acyltransferase (E2p), *Aquifex aeolicus* lumazine synthase (LS), and *Thermotoga maritima* encapsulin, which all form 60-mers. Self-assembling nanoparticles can form spontaneously upon recombinant expression of the protein in an appropriate expression system. Methods for nanoparticle production,
15 detection, and characterization can be conducted using the same techniques developed for VLPs.

III. Engineered Ebolavirus GP sequences

[0052] The invention provides engineered or redesigned Ebolavirus glycoprotein
20 (GP) sequences (polypeptides or polynucleotide sequences) for producing Ebolavirus vaccines. Relative to the wildtype GP sequences, these engineered GP polypeptides can contain various modifications, esp. in the GP2 subunit. The GP2 subunit of Ebolavirus GPs contains the N-terminal peptide, the internal fusion loop, two consecutive two heptad repeat regions (HR1 and HR2), the membrane proximal external region
25 (MPER), and the C-terminal transmembrane domain (see Figure 1A). HR1 is in turn structurally divided into 4 segments, HR1_A, HR1_B, HR1_C and HR1_D. The first two segments, HR1_A and HR1_B (residues 554–575), form an α -helix with an approximately 40° kink at Thr565, which delineates HR1_A from HR1_B. HR1_C forms an extended coil linker between HR1_B and the HR1_D segment. HR1_D forms an amphipathic helix, the
30 hydrophobic faces of the three helices in the trimer pack together to form the interface of the peplomer. HR1_C and HR1_D were first defined to correspond to residues 576-582 and residues 583-598, respectively (Lee et al., Future Virol. 2009; 4: 621–635). Based on a higher-resolution and more complete crystal structure of GP reported in Zhao et al.

Nature 535, 169-172, 2016, HR1_C is redefined herein to encompass aa576-583, and HR1_D refers to aa 584-598. HR2 is a largely alpha-helical section of protein, also termed the “HR2 stalk”, that connects the GP core to the viral membrane. In the EBOV sequence, the HR2 stalk encompasses amino acid residues 599-632.

- 5 **[0053]** As detailed herein, the engineered Ebolavirus GP polypeptides contain various mutations or modifications primarily in and around the HR1_C and HR2 motifs (Figure 1B). Unless otherwise noted, the organization and amino acid numbering with regard to various domains or regions of Ebolavirus GP is based on GP sequence of Zaire Ebolavirus strain Mayinga-76, which has ectodomain sequence described by
- 10 GenBank ID AAG40168.1 (SEQ ID NO:1). Due to substantial sequence conservation in the structural motifs among the different Ebolaviruses and different strains of the same Ebolavirus species, corresponding amino acid residues bordering the various domains, regions and loops of an Ebolavirus GP from any other Ebolavirus and strains can be readily determined (e.g., by sequence alignment) or otherwise known in the art.
- 15 As exemplification, a sequence alignment of the HR1 and HR2 regions of EBOV with other Ebolaviruses are shown in Figure 2. Sequences of the whole HR1+HR2 region are about 84.8% identical across the 5 different species of Ebolaviruses, EBOV, SUDV, TAFV, BDBV and RESTV. As indicated in the figure, residues in the HR1_C segment are highly conserved among the different Ebolaviruses. The first few N-terminal
- 20 residues in the MPER motif for HR2 extension as described herein are also highly conserved among the different Ebolavirus species. The W615 residue exemplified herein for substitution is conserved in all Ebolavirus species. Moreover, logo analysis of sequence conservation for HR2 stalk described in the Examples below showed that W615 is 100% conserved across all strains of all six species in the Ebolavirus genus.
- 25 Thus, it is apparent that the Ebolavirus GP sequence modification exemplified with EBOV can be readily applied to other Ebolavirus species and different strains of the same virus species.

- [0054]** Compared to a full length wildtype Ebolavirus GP (see, e.g., Figure 1A), the engineered Ebolavirus GP sequences of the invention typically contain, in addition
- 30 to the mutations or modifications described herein, structural motifs that correspond to the GP ectodomain without the N-terminal leader, MLD and MPER (see, e.g., SEQ ID NO:2). In various embodiments, the engineered Ebolavirus GP sequences of the invention can include additional structural motifs or domains of the full length GP

beyond the structural components present in SEQ ID NO:2. Thus, some engineered Ebolavirus GP sequences of the invention can additionally contain one or more of (1) a leader sequence, (2) part or all of MLD, (3) part or all of MPER, (4) part or all of the transmembrane domain, and (5) part or all of the cytoplasmic tail.

- 5 **[0055]** Some embodiments of the invention are directed to the engineered Ebolavirus GP sequences that correspond to full length Ebolavirus GPs with one or more stabilizing modifications or mutations in the HR2 and HR1 regions described herein. In some embodiments, the engineered Ebolavirus GP sequences can contain substitution at residue W615 in HR2 of a wildtype GP sequence. In some of these
- 10 embodiments, the W615 residue can be replaced with a small hydrophobic residue, e.g., L, A, V, I or F. Additionally or alternatively, the engineered Ebolavirus GP sequences can have one or more proline substitutions in the HR1_C segment. As exemplified herein, any residue in the HR1_C segment, T576, T577, E578, L579, R580, T581, F582 and S583, can be substituted with proline. In some of these embodiments, the
- 15 engineered Ebolavirus GP sequences contain T577P and/or L579P substitutions in HR1_C. In some embodiments, further cysteine substitutions can be introduced into the GP sequence to generate inter-GP disulfide bonds as exemplified herein. In various embodiments, the leader peptide sequence at the N-terminus of the GP sequences can be removed. In some other embodiments directed to polynucleotide sequences or
- 20 vectors that express the engineered Ebolavirus GP proteins, a sequence that encodes the leader peptide (e.g., SEQ ID NO:41) is included at the 5'-end of the engineered Ebolavirus GP polynucleotide sequence.

- [0056]** In some embodiments, the engineered Ebolavirus GP sequences of the invention contain only the ectodomain (i.e., MPER at the C-terminus) or otherwise a
- 25 soluble portion of Ebolavirus GP proteins along with the alterations in the HR1 and HR2 regions described herein. In some of these embodiments, the truncated or altered soluble Ebolavirus GP sequence also has MLD deleted. In some embodiments, the shortened soluble GP sequence additionally has MPER at the C-terminus of the GP ectodomain removed. In various embodiments, the expressed and assembled trimer
- 30 protein also does not contain the leader peptide sequence (SEQ ID NO:41). An example of such a shortened GP soluble sequence based on a Zaire Ebolavirus (EBOV) GP ectodomain sequence (SEQ ID NO:1) is shown in SEQ ID NO:2, which has the leader, MLD and MPER sequences deleted from the wildtype ectodomain sequence. In

some of embodiments, the shortened soluble GP sequence further contains a T42A mutation in the GP1 base motif (GP Δ muc; SEQ ID NO:3).

- [0057]** In addition to the truncation at the C-terminus, engineered soluble Ebolavirus GP sequences of the invention typically contain additional sequence modifications in the HR1 and HR2 regions of a soluble GP sequence (e.g., SEQ ID NO:2 or 3). In various embodiments, the additional sequence modifications include substitution of residue W615 in HR2, extension of HR2 by (1) adding one or more adjacent residues in the MPER or (2) replacing some C-terminal HR2 residues with a longer heterologous sequence, substitutions of one or more residues in the HR1c segment with proline, and introduction of one or more disulfide bonds. As described herein, these additional sequence modifications, alone or in any combination, can promote GP trimer formation, reduce metastability, and stabilize Ebolavirus GP in a native-like trimer conformation.
- [0058]** In some embodiments, the engineered Ebolavirus GP sequences of the invention contain a truncated or shortened soluble GP sequence (e.g., SEQ ID NO:2 or SEQ ID NO:3, or a conservatively modified or substantially identical variant thereof) with additional modifications in HR2. In the exemplified EBOV ectodomain GP sequence (SEQ ID NO:1), the HR2 motif encompasses residues 599-632. In some of these embodiments, the inward-facing residue in the 3-dimensional structure W615 is replaced with an amino acid residue that is smaller and more hydrophobic. In various embodiments, the W615 residue can be replaced with Leu, Phe, Ala, Val, or Ile so as to improve the HR2 stalk packing. These substitutions are intended to stabilize the Ebolavirus GP trimer. In some exemplified embodiments, the sequence substitution is W615L. In some other embodiments, the HR2 can contain a further substitution at residue 612 in addition to W615 substitution. In some of these embodiments, the amino acid substitutions are P612G/W615F. Some specific examples of engineered Ebolavirus GP sequences containing W615 substitution are shown in SEQ ID NOs:4 and 5.
- [0059]** In addition to the substitutions in HR2, the engineered soluble Ebolavirus GP sequences can alternatively or additionally contain an extension of the HR2 motif. In some these embodiments, extension of the HR2 motif (ending at residue 632) involves adding one or more adjacent residues in the MPER motif (starting at residue 633) that are naturally present in the Ebolavirus GP sequence. In various embodiments,

the HR2 extension can be addition of its C-terminal adjacent residues in MPER up to residue 635, 636, 637, 638, 639, 640, 641, 642, 643, 644, 645, 646 or beyond. In some exemplified embodiments, the HR2 extension include extension to the TACE cleavage site (residue 637) in MPER, i.e., having MPER N-terminal residues KTLPD (SEQ ID NO:32). The TACE cleavage site is responsible for GP shedding from the virion surface after the host TACE cleavage. In some other embodiments, the HR2 extension can include 6 additional adjacent residues in the MPER motif, i.e., having MPER N-terminal residues KTLPDQGDNDN (SEQ ID NO:33). In some embodiments, extension of the HR2 motif involves substitution of some of the HR2 C-terminal residues with a longer heterologous sequence. In some of these embodiments, some of the HR2 C-terminal residues can be replaced with a longer GCN4 leucine zipper sequence. For example, HR2 can be extended by replacing residues 617-632 in HR2 with a 31 aa sequence (SEQ ID NO:34) from a GCN4 leucine zipper with PDB ID 2WPZ as exemplified herein. Some specific embodiments of engineered Ebolavirus GP proteins containing an HR2 extension are shown in SEQ ID NOs:5-7. Specific embodiments of engineered Ebolavirus GP proteins containing both an HR2 extension and a residue W615 substitution are shown in SEQ ID NOs:6 and 8.

[0060] Additionally or alternatively to the above-noted modifications in the HR2 region, the engineered Ebolavirus GP proteins of the invention can also contain mutations in the C segment of the HR1 region (HR1_C). Typically, the mutations in HR1_C contain one or more proline substitutions that can stabilize the GP trimer and reduce metastability. In various embodiments, the proline substitution can be present at each position of HR1_C. As exemplified herein with EBOV GP sequence (SEQ ID NO:1), proline substitution can be introduced at each of residues 576-583 (TTELRTFS; SEQ ID NO:35). Thus, in various embodiments, engineered Ebolavirus GP proteins of the invention can contain one or more HR1_C mutations among T576P, T577P, E578P, L579P, R580P, T581P, F582P, and S583P. In some exemplified embodiments, the engineered Ebolavirus GP proteins of the invention contain T577P (P²) or L579P (P⁴) mutations in HR1_C. Specific embodiments of engineered Ebolavirus GP sequences containing one or more proline substitutions in the HR1_C segment are shown in SEQ ID NOs:9-16.

[0061] In some embodiments, the engineered Ebolavirus GP sequences of the invention contain both proline substitution in HR1_C and modifications in HR2 noted

above. For example, the engineered GP proteins can contain a proline substitution, and a W615 substitution and/or a further HR2 extension into MPER. The proline substitution can be at any residue in HR1_C (e.g., at residue 577 or 579). W615 substitution can be any of W615L, W615F, W615A, W615V, or W615I. HR2 extension in these in these engineered GP proteins can be, e.g., extension to residue 637 (i.e., the TACE cleavage site) in MPER. Specific embodiments of engineered Ebolavirus GP sequences containing both a proline substitution in HR1_C and also HR2 modifications are shown in SEQ ID NOs:17-22.

[0062] Other than the modifications in the heptad regions noted above, some engineered Ebolavirus GP sequences of the invention can alternatively or additionally contain one or more engineered cysteine residues for forming inter-GP disulfide bonds. By forming disulfide bonds between neighboring protomers of the GP trimer, these Cys substitutions and resulting engineered disulfide bonds can similarly function to promote trimer formation and to stabilize the GP in a native-like trimer conformation. Depending on the specific GP sequence to be used, one can readily determine the appropriate positions for introducing one or more inter-GP disulfide bonds. This can be performed, e.g., via analyzing potential amino acid pairs in the 3-dimensional structure of the GP and subsequent biochemical and immunological characterizations as described herein. As exemplified with EBOV GP sequence (SEQ ID NO:1) herein, the engineered GP trimer protein can contain one or more engineered inter-protomer SS bonds among G91/A575 (SS2), F153/Y534 (SS1), T520/A575 (SS3), G157/I532 (SS4), D522/A575 (SS5) and K56/G599 (SS6). In some embodiments, the engineered GP trimer proteins of the invention contains engineered disulfide bond at G91/A575. Specific embodiments of engineered Ebolavirus GP proteins containing such engineered disulfide bonds are shown in SEQ ID NOs:23-28.

[0063] In some embodiments, the engineered Ebolavirus GP sequences of the invention can contain a C-terminal trimerization motif. This motif functions to further stabilize the trimer and also to increase the trimer ratio within the total protein yield. Suitable trimerization motifs for the invention include, e.g., T4 fibrin foldon (PDB ID: 4NCV) and viral capsid protein SHP (PDB: 1TD0). T4 fibrin (foldon) is well known in the art, and constitutes the C-terminal 30 amino acid residues of the trimeric protein fibrin from bacteriophage T4, and functions in promoting folding and trimerization of fibrin. See, e.g., Papanikolopoulou et al., J. Biol. Chem. 279: 8991-8998, 2004; and

Guthe et al., J. Mol. Biol. 337: 905-915, 2004. Similarly, the SHP protein and its used as a functional trimerization motifs are also well known in the art. See, e.g., Dreier et al., Proc Natl Acad Sci USA 110: E869–E877, 2013; and Hanzelmann et al., Structure 24: 140–147, 2016. In some exemplified embodiments, the trimerization motif in the engineered GP proteins comprise a foldon sequence shown in SEQ ID NO:29 or the 1TD0 protein sequence shown in SEQ ID NO:30. In some other embodiments, the employed trimerization motif can contain a sequence that is a conservatively modified variant or substantially identical (e.g., at least 90%, 95% or 99% identical) sequence of the exemplified sequence. In some embodiments, the trimerization motif can be inserted with a short GS linker. In various embodiments, the linker can contain 1-6 tandem repeats of GS. In some embodiments, an His6-tag can be added to the C-terminus of the trimerization motif to facilitate protein purification, e.g., by using a Nickel column.

[0064] In addition to the specific examples of GP polypeptides set forth in SEQ ID NOs:4-28, engineered Ebolavirus GP sequences of the invention also encompass sequences having an amino acid sequence that is substantially identical to one of these sequences, including conservatively modified variant sequences. In various embodiments, the engineered Ebolavirus GP proteins of the invention of the invention can have an amino acid sequence that is identical to any of SEQ ID NOs:4-28, except for one or more amino acid residue substitutions of non-conserved residues in HR1 and HR2 among different Ebolavirus species or strains, or substitutions in the non-conserved region or motif of the GP sequence of different Ebolavirus species or strains.

[0065] As exemplified herein with EBOV strain Mayinga-76 (SEQ ID NO:1), GP sequences from different Ebolavirus species and strains can all be readily employed to generate engineered Ebolavirus GP sequences in accordance with the strategy described herein. Ebolaviruses suitable for the invention include any of EBOV, SUDV, TAFV, BDBV and RESTV, as well as any strain of a given Ebolavirus species. In some embodiments of the invention, the engineering Ebolavirus GP proteins are chimeric. These chimeric Ebolavirus GP proteins or immunogen polypeptides can contain a chimeric GP sequence with different subunits or domains derived from multiple Ebolavirus species or from multiple strains of the same Ebolavirus species. For example, the GP1 and GP2 subunit sequences in the chimeric engineered Ebolavirus GP sequences can be derived from two different Ebolavirus species (e.g., GP2 from

EBOV and GP1 from SUDV or TAFV) or from two different strains of the same Ebolavirus species (e.g., two different EBOV strains).

[0066] Many GP sequences of the five different Ebolavirus species and a great number of different strains of a given Ebolavirus species are known in the art. For example, such sequence information can be readily obtained from the Ebola Database maintained by the Los Alamos National Laboratory (LANL) and other Ebolavirus related databases such as the Virus Pathogen Resource (ViPR). See, e.g., Kuiken et al., *Nucleic Acids Res.* 40: D587-92, 2012; Pickett et al., *Nucleic Acids Res.* 40: D593-98, 2012; and Swetha et al., *Adv Bioinformatics.* 2016: 1673284. Admittedly, there is a considerable degree of variability among GP sequences of different Ebolavirus species and different strains of the same Ebolavirus species. However, as noted above, a certain number of conserved residues and motifs are present in HR1 and HR2 in all Ebolavirus species, which are the locations where mutations or modifications are introduced in the engineered GP proteins of the invention. In the engineering Ebolavirus GP proteins of the invention, one can readily determine the appropriate residues for modifications via sequence alignment and also considering conserved Ebolavirus GP motifs and residues known in the art. See, e.g., Leroy et al., *J. General Virol.* 83: 67-73, 2002; Arslan et al., *Bioinformatics.* 32: 151-154, 2017; Wec et al., *Cell* 169: 878-890, 2017; Jun et al., *FEMS Microbiol Rev.* 39: 764-778, 2015; Pappalardo et al., *Sci. Rep.* 6: 23743, 2016; and Ruedas et al., *J. Virol.* 91: e00392-17, 2017. Engineered Ebolavirus GP sequences based on any of the other Ebolavirus species and strains, or a combination of different strains or species, can all be generated using the same strategy exemplified herein for the EBOV GP sequence (SEQ ID NO:1).

[0067] As detailed below, the engineering Ebolavirus GP proteins may be conjugated to a presenting platform (e.g., nanoparticles or VLPs) via various means. Preferably, the conjugation is achieved via covalent linkage, e.g., protein fusions or insertions. In some preferred embodiments, the protein sequence is fused with the presenting platform sequence via a linker sequence. In various embodiments, other modifications can also be made to the engineering Ebolavirus GP proteins or the conjugating partner in order to improve stability or antigenicity.

[0068] The various engineered Ebolavirus GP molecules of the invention can be obtained or generated in accordance with the protocols exemplified herein or methods well known in the art. See, e.g., Sambrook et al., *Molecular Cloning: A Laboratory*

Manual, Cold Spring Harbor Press, N.Y., (3rd ed., 2000); and Brent et al., Current Protocols in Molecular Biology, John Wiley & Sons, Inc. (ringbou ed., 2003). Upon recombinant expression (e.g., in HEK293 F cells as detailed herein), the proteins can be purified by any of the routinely practiced procedures. See, e.g., *Guide to Protein Purification*, Ed. Deutscher, Meth. Enzymol. 185, Academic Press, San Diego, 1990; and Scopes, *Protein Purification: Principles and Practice*, Springer Verlag, New York, 1982. Substantial purification denotes purification from other proteins or cellular components. A substantially purified protein is at least 60%, 70%, 80%, 90%, 95% or 98% pure. Once purified, antigenicity and other properties of the engineered

5 Ebolavirus GP proteins can also be readily examined with standard methods, e.g., antigenic profiling using known bNAbS and non-NAbS, differential scanning calorimetry (DSC), electron microscopy, binding analysis via ELISA, Biolayer Interferometry (BLI), Surface Plasmon Resonance (SPR), and co-crystallography analysis as exemplified herein.

15

IV. Scaffolded Ebolavirus GP vaccine compositions

[0069] The invention provides Ebolavirus GP based vaccine compositions that contain a heterologous scaffold presenting or incorporating an engineered Ebolavirus GP protein described herein. Any heterologous scaffold can be used to present the

20 engineered Ebolavirus GP protein in the construction of the vaccines of the invention. These include nanoparticles, virus-like particles, protein carriers (e.g., immunoglobulin chains or domains such as Fc, KLH, BSA, tetanus toxoid, and diphtheria toxoid), as well as various chemical scaffolds. In some embodiments, a virus-like particle (VLP) such as bacteriophage Q β VLP and nanoparticles can be used. In some preferred

25 embodiments, the heterologous scaffold for presenting or displaying the engineered Ebolavirus GP protein is a nanoparticle. Various nanoparticle platforms can be employed in generating the vaccine compositions of the invention. In general, the nanoparticles employed in the invention need to be formed by multiple copies of a single subunit. The nanoparticles are typically ball-like shaped, and/or have rotational

30 symmetry (e.g., with 3-fold and 5-fold axes), e.g., with an icosahedral structure exemplified herein. Additionally or alternatively, the amino-terminus of the particle subunit has to be exposed and in close proximity to the 3-fold axis, and the spacing of three amino-termini has to closely match the spacing of the carboxyl-termini of the

Ebolavirus GP protein. In some preferred embodiments, the immunogens comprise self-assembling nanoparticles with a diameter of about 20nm or less (usually assembled from 12, 24, or 60 subunits) and 3-fold axes on the particle surface.

[0070] In some preferred embodiments, the engineered Ebolavirus GP protein is presented on self-assembling nanoparticles such as self-assembling nanoparticles derived from E2p, I3-01v9 and ferritin (FR) as exemplified herein. E2p is a redesigned variant of dihydrolipoyl acyltransferase from *Bacillus stearothermophilus* that has been shown to self-assemble into thermostable 60-meric nanoparticle. See, e.g., He et al., *Nat. Commun.* 7:12041, 2016. Similarly, I3-01 is an engineered protein that can self-assemble into hyperstable nanoparticles. See, e.g., Hsia et al., *Nature* 535, 136-139, 2016. A modified version of I3-01, I3-01v9 is used here as exemplification. Ferritin is a globular protein found in all animals, bacteria, and plants. The globular form of ferritin is made up of monomeric subunit proteins (also referred to as monomeric ferritin subunits), which are polypeptides having a molecule weight of approximately 17-20 kDa. Amino acid sequences of E2p, I3-01v9 and ferritin nanoparticle subunits as exemplified herein are shown in SEQ ID NOs:36-38, respectively. Relative to the original sequence, E2p sequence shown in SEQ ID NO:36 contains an Ala substitution at residue 92 as underscored in the sequence below. Sequences of some other suitable nanoparticle sequences are also known in the art. See, e.g., WO2017/192434, WO2019/089817 and WO19/241483. In various embodiments, the Ebolavirus nanoparticle vaccines of the invention can employ any of these known nanoparticles, as well as their conservatively modified variants or variants with substantially identical (e.g., at least 90%, 95% or 99% identical) sequences.

[0071] E2p subunit sequence (SEQ ID NO:36)

AAAKPATTEGEFPETREKMSGIRRAIAKAMVHSKHTAPHVTLMDEADVTKLV
AHRKKFKAIAAEKGKLTFLPYVVKALVSALREYPVLNTAIDDETEEIIQKHYY
NIGIAADTDRLGLLPVIKHADRKPIFALAEQINELAEKARDGKLTPEGMGKASC
TITNIGSAGGQWFTPVINHPEVAILGIGRIAEKPIVRDGEIVAAPMLALSLSFDHR
MIDGATAQKALNHIKRLSDPELLLM

[0072] I3-01v9 subunit sequence (SEQ ID NO:37)

MKMEELFKKHKIVAVLRANSVEEAKMKALAVFVGGVHLIEITFTVPDADTVIK
ELSFLKELGAIIGAGTVTSVEQCRKAVESGAEFIVSPHLDEEISQFCKEKGVFYM
PGVMTPTLVKAMKLGHTILKLFPGEVVGPQFVKAMKGPFPNVKFPVPTGGVN
LDNVCEWFKAGVLAVGVGSALVKGITAEVAAKAAAFVEKIRGCTE

[0073] Ferritin sequence (SEQ ID NO:38)

DIKLLNEQVNKEMNSSNLYMSMSSWCYTHSLDGAGLFLFDHAAEEYEHAKK
 LIIFLNENNVQVQLTSISAPEHKFEGLTQIFQKAYEHEQHISESINNIVDHAIKSKD
 HATFNFLQWYVAEQHHEEVLFKDILDKIELIGNENHGLYLADQYVKGIKSRK

5 [0074] In addition to these exemplified nanoparticle sequences, many other
 nanoparticles or VLPs known in the art may also be used in the practice of the
 invention. These include, e.g., *Aquifex aeolicus* lumazine synthase, *Thermotoga*
Maritima encapsulin, *Myxococcus xanthus* encapsulin, bacteriophage Qbeta virus
 particle, Flock House Virus (FHV) particle, ORSAY virus particle, and infectious
 10 bursal disease virus (IBDV) particle. Other molecules that may be used as the
 presenting platform of the nanoparticle vaccines of the invention include, e.g.,
 molecules with the following PDB IDs: 1JIG (12-mer Dlp-2 from *Bacillus anthracis*),
 1UVH (12-mer DPS from *Mycobacterium smegmatis*), 2YGD (24-mer eye lens
 chaperone α B-crystallin), 3CS0 (24-mer DegP24), 3MH6 and 3MH7 (24-mer HtrA
 15 proteases), 3PV2 (12-mer HtrA homolog DegQ WT), 4A8C (12-mer DegQ from *E.*
Coli.), 4A9G (24-mer DegQ from *E. Coli.*), 4EVE (12-mer HP-NAP from *Helicobacter*
pylori strain YS29), and 4GQU (24-mer HisB from *Mycobacterium tuberculosis*).

[0075] In various embodiments, the Ebolavirus GP protein to be displayed on a
 nanoparticle platform may optionally contain a trimerization motif described above,
 20 e.g., foldon or SHP. Some Ebolavirus nanoparticle vaccine compositions can
 additionally contain other structural components that function to further enhance
 stability and antigenicity of the displayed immunogen. In some embodiments, a
 locking protein domain can be inserted into the nanoparticle construct, e.g., by
 covalently fused to the C-terminus of the nanoparticle subunit. The locking domain can
 25 be any dimeric protein that is capable of forming an interface through specific
 interactions such as hydrophobic (van der Waals) contacts, hydrogen bonds, and/or salt
 bridges. General guidance on selecting locking domains and specific examples are
 described in the art, e.g., PCT2019/036917. In some exemplified embodiments, the
 locking domain used in the Ebolavirus nanoparticle vaccines of the invention can
 30 contain locking domain LD4 or LD7 exemplified herein.

[0076] In some embodiments, Ebolavirus nanoparticle vaccines of the invention
 can also contain a T-cell epitope to promote robust T-cell responses and to steer B cell
 development towards bNAbs. The T-cell epitope can be located at any position in
 relation to the other structural components as long as it does not impact presentation of

the immunogen polypeptides on the nanoparticle surface. Any T-cell epitope sequences or peptides known in the art may be employed in the practice of the present invention. They include any polypeptide sequence that contain MHC class-II epitopes and can effectively activate CD4+ and CD8+ T cells upon immunization, e.g., T-helper epitope

5 that activates CD4+ T helper cells. See, e.g., Alexander et al., *Immunity* 1, 751-761, 1994; Ahlers et al., *J. Clin. Invest.* 108:1677-1685, 2001; Fraser et al., *Vaccine* 32, 2896-2903, 2014; De Groot et al., *Immunol. Cell Biol.* 8:255-269, 2002; and *Gene Ther.* 21: 225-232, 2014. In some embodiments, the T cell epitope inserted into the Ebolavirus nanoparticle vaccine construct is a universal pan DR epitope peptide

10 (PADRE). See, e.g., Hung et al., *Mole. Ther.* 15: 1211-19, 2007; Wu et al., *J. Biomed. Sci.* 17: 88, 2010; and Bissati et al., *npj Vaccines* 2: 24, 2017. Other examples of suitable T-cell epitope are also described in the art, e.g., the D and TpD epitope (Fraser et al., *Vaccine* 32, 2896-2903, 2014). In some embodiments, the employed PADRE peptide contains a sequence AKFVAAWTLKAAA (SEQ ID NO:31), a conservatively

15 modified variant or substantially identical (e.g., at least 90%, 95% or 99% identical) sequence thereof. In some of these embodiments, the PADRE epitope is inserted at the C-terminus of a locking domain, e.g., LD4 or LD7 as exemplified herein. In some of these embodiments, a GS restriction site can be added between the LD and the PADRE epitope.

20 [0077] The scaffolded Ebolavirus vaccine compositions of the invention can be constructed in accordance with standard recombinant techniques and the methods described herein (see, e.g., the Examples herein) and/or other methods that have been described in the art, e.g., He et al., *Nat. Comm.* 7, 12041, 2016; Kong et al., *Nat. Comm.* 7, 12040, 2016; He et al., *Sci Adv.* 4(11):eaau6769, 2018; and PCT

25 publications WO2017/192434, WO2019/089817 and WO19/241483. In various embodiments, nanoparticle displaying any of the engineered Ebolavirus GP proteins can be constructed by fusing the Ebolavirus GP polypeptide to the subunit of the nanoparticle (e.g., E2p subunit). Preferably, C-terminus of the Ebolavirus GP polypeptide is fused to the N-terminus of the nanoparticle subunit. In some

30 embodiments, a short peptide spacer can be used to connect the Ebolavirus GP polypeptide and the nanoparticle. For example, the spacer can contain a GS restriction site and/or a longer G4S (SEQID NO:43) or (G4S)₂ (SEQID NO:42) linker as exemplified herein. Some embodiments of the invention are directed to nanoparticles

displaying engineered Ebolavirus GP protein GP Δ muc-WL²P² protein (SEQ ID NO:17) with different combination of nanoparticle subunit sequence, locking domain sequence and/or T-cell epitope. Some specific examples of such nanoparticle vaccine compositions have a sequence structure of GP Δ muc-WL²P²-AS-G₄S-ferritin, GP Δ muc-WL²P²-AS-E2p-LD4-PADRE, or GP Δ muc-WL²P²-AS-(G₄S)₂-I3-01v9-LD7-PADRE.

[0078] Once constructed, the antigenicity and structural integrity of the nanoparticle displayed Ebolavirus GP polypeptides can be readily analyzed via standard assays, e.g., antibody binding assays, biolayer interferometry, and negative-stain electron microscopy (EM). As exemplified herein, the various fusion molecules can all self-assemble into nanoparticles that display immunogenic epitopes of the Ebolavirus GP proteins. By eliciting a robust neutralizing antibody response, these nanoparticles are useful for vaccinating individuals against a broad range of Ebolavirus infections.

V. Polynucleotides and expression constructs

[0079] The engineered Ebolavirus GP proteins and the nanoparticle vaccine compositions of the invention are typically produced by first generating expression constructs (i.e., expression vectors) that contain operably linked coding sequences of the various structural components described herein. Accordingly, in some related aspects, the invention provides polynucleotides (DNA or RNA) that encode the engineered Ebolavirus GP proteins or polypeptides, and that encode the subunit of nanoparticles displayed the engineered Ebolavirus GP polypeptides as described herein, as well as expression vectors that harbor such polynucleotides and host cells for producing the Ebolavirus GP immunogen polypeptides and the vaccine compositions (e.g., HEK293 F cells and ExpiCHO cells exemplified herein). The fusion polypeptides encoded by the polynucleotides or expressed from the vectors are also included in the invention.

[0080] The polynucleotides and related vectors can be readily generated with standard molecular biology techniques or the protocols exemplified herein. For example, general protocols for cloning, transfecting, transient gene expression and obtaining stable transfected cell lines are described in the art, e.g., Sambrook et al., *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Press, N.Y., (3rd ed., 2000); and Brent et al., *Current Protocols in Molecular Biology*, John Wiley & Sons, Inc. (ringbou edition, 2003). Introducing mutations to a polynucleotide sequence by

PCR can be performed as described in, e.g., *PCR Technology: Principles and Applications for DNA Amplification*, H.A. Erlich (Ed.), Freeman Press, NY, NY, 1992; *PCR Protocols: A Guide to Methods and Applications*, Innis *et al.* (Ed.), Academic Press, San Diego, CA, 1990; Mattila *et al.*, *Nucleic Acids Res.* 19:967, 1991; and
5 Eckert *et al.*, *PCR Methods and Applications* 1:17, 1991.

[0081] The selection of a particular vector depends upon the intended use of the fusion polypeptides. For example, the selected vector must be capable of driving expression of the fusion polypeptide in the desired cell type, whether that cell type be prokaryotic or eukaryotic. Many vectors contain sequences allowing both prokaryotic
10 vector replication and eukaryotic expression of operably linked gene sequences. Vectors useful for the invention may be autonomously replicating, that is, the vector exists extrachromosomally and its replication is not necessarily directly linked to the replication of the host cell's genome. Alternatively, the replication of the vector may be linked to the replication of the host's chromosomal DNA, for example, the vector may
15 be integrated into the chromosome of the host cell as achieved by retroviral vectors and in stably transfected cell lines. Both viral-based and nonviral expression vectors can be used to produce the immunogens in a mammalian host cell. Nonviral vectors and systems include plasmids, episomal vectors, typically with an expression cassette for expressing a protein or RNA, and human artificial chromosomes (see, e.g., Harrington
20 *et al.*, *Nat. Genet.* 15:345, 1997). Useful viral vectors include vectors based on lentiviruses or other retroviruses, adenoviruses, adenoassociated viruses, cytomegalovirus, herpes viruses, vectors based on SV40, papilloma virus, HBP Epstein Barr virus, vaccinia virus vectors and Semliki Forest virus (SFV). See, Brent *et al.*, *supra*; Smith, *Annu. Rev. Microbiol.* 49:807, 1995; and Rosenfeld *et al.*, *Cell* 68:143,
25 1992.

[0082] Depending on the specific vector used for expressing the fusion polypeptide, various known cells or cell lines can be employed in the practice of the invention. The host cell can be any cell into which recombinant vectors carrying a fusion of the invention may be introduced and wherein the vectors are permitted to
30 drive the expression of the fusion polypeptide is useful for the invention. It may be prokaryotic, such as any of a number of bacterial strains, or may be eukaryotic, such as yeast or other fungal cells, insect or amphibian cells, or mammalian cells including, for example, rodent, simian or human cells. Cells expressing the fusion polypeptides of the

- invention may be primary cultured cells or may be an established cell line. Thus, in addition to the cell lines exemplified herein (e.g., CHO cells), a number of other host cell lines capable well known in the art may also be used in the practice of the invention. These include, e.g., various Cos cell lines, HeLa cells, HEK293, AtT20, BV2, and N18 cells, myeloma cell lines, transformed B-cells and hybridomas.
- [0083]** The use of mammalian tissue cell culture to express polypeptides is discussed generally in, e.g., Winnacker, *From Genes to Clones*, VCH Publishers, N.Y., N.Y., 1987. The fusion polypeptide-expressing vectors may be introduced to the selected host cells by any of a number of suitable methods known to those skilled in the art. For the introduction of fusion polypeptide-encoding vectors to mammalian cells, the method used will depend upon the form of the vector. For plasmid vectors, DNA encoding the fusion polypeptide sequences may be introduced by any of a number of transfection methods, including, for example, lipid-mediated transfection (“lipofection”), DEAE-dextran-mediated transfection, electroporation or calcium phosphate precipitation. These methods are detailed, for example, in Brent et al., *supra*. Lipofection reagents and methods suitable for transient transfection of a wide variety of transformed and non-transformed or primary cells are widely available, making lipofection an attractive method of introducing constructs to eukaryotic, and particularly mammalian cells in culture. For example, LipofectAMINE™ (Life Technologies) or LipoTaxi™ (Stratagene) kits are available. Other companies offering reagents and methods for lipofection include Bio-Rad Laboratories, Clontech, Glen Research, Life Technologies, JBL Scientific, MBI Fermentas, PanVera, Promega, Quantum Biotechnologies, Sigma-Aldrich, and Wako Chemicals USA.
- [0084]** For long-term, high-yield production of recombinant fusion polypeptides, stable expression is preferred. Rather than using expression vectors which contain viral origins of replication, host cells can be transformed with the fusion polypeptide-encoding sequences controlled by appropriate expression control elements (e.g., promoter, enhancer, sequences, transcription terminators, polyadenylation sites, etc.), and selectable markers. The selectable marker in the recombinant vector confers resistance to the selection and allows cells to stably integrate the vector into their chromosomes. Commonly used selectable markers include neo, which confers resistance to the aminoglycoside G-418 (Colberre-Garapin, et al., *J. Mol. Biol.*, 150:1, 1981); and hygro, which confers resistance to hygromycin (Santerre et al., *Gene*, 30:

147, 1984). Through appropriate selections, the transfected cells can contain integrated copies of the fusion polypeptide encoding sequence.

VI. Pharmaceutical compositions and therapeutic applications

5 [0085] The invention provides pharmaceutical or immunogenic compositions and related methods of using the engineered Ebolavirus GP sequences and nanoparticles displaying the GP proteins as described herein for preventing and treating Ebolavirus infections. In some embodiments, an engineered Ebolavirus GP sequence (a polypeptide or a polynucleotide sequence) or a nanoparticle displaying an engineered
10 protein is included in a pharmaceutical composition. The pharmaceutical composition can be either a therapeutic formulation or a prophylactic formulation. Typically, the composition additionally includes one or more pharmaceutically acceptable vehicles and, optionally, other therapeutic ingredients (for example, antibiotics or antiviral drugs). Various pharmaceutically acceptable additives can also be used in the
15 compositions.

[0086] Some of the pharmaceutical compositions of the invention are vaccines. For vaccine compositions, appropriate adjuvants can be additionally included. Examples of suitable adjuvants include, e.g., aluminum hydroxide, lecithin, Freund's adjuvant, MPL™ and IL-12. In some embodiments, the engineered Ebolavirus GP
20 sequences and related vaccines as disclosed herein can be formulated as a controlled-release or time-release formulation. This can be achieved in a composition that contains a slow release polymer or via a microencapsulated delivery system or bioadhesive gel. The various pharmaceutical compositions can be prepared in accordance with standard procedures well known in the art. See, e.g., Remington's
25 Pharmaceutical Sciences, 19th Ed., Mack Publishing Company, Easton, Pa., 1995; Sustained and Controlled Release Drug Delivery Systems, J. R. Robinson, ed., Marcel Dekker, Inc., New York, 1978); U.S. Pat. Nos. 4,652,441 and 4,917,893; U.S. Pat. Nos. 4,677,191 and 4,728,721; and U.S. Pat. No. 4,675,189.

[0087] Therapeutic methods of the invention involve administering an engineered
30 Ebolavirus GP sequence of the invention or a pharmaceutical composition containing the polypeptide to a subject having or at risk of developing an Ebolavirus infection (e.g., EBOV infection). In some embodiments, a pharmaceutical composition of the invention is employed in therapeutic or prophylactic applications for treating

Ebolavirus infections or eliciting an protective immune response against an Ebolavirus species or strain in a subject. For example, the composition can be administered to a subject to induce an immune response to an Ebolavirus species, e.g., to induce production of broadly neutralizing antibodies to the Ebolavirus species. For subjects at risk of developing an Ebolavirus infection, a vaccine composition of the invention can be administered to provide prophylactic protection against viral infection. Depending on the specific subject and conditions, the pharmaceutical compositions of the invention can be administered to subjects by a variety of administration modes known to the person of ordinary skill in the art, for example, intramuscular, subcutaneous, intravenous, intra-arterial, intra-articular, intraperitoneal, or parenteral routes. In general, the pharmaceutical composition is administered to a subject in need of such treatment for a time and under conditions sufficient to prevent, inhibit, and/or ameliorate a selected disease or condition or one or more symptom(s) thereof. Symptoms of Ebolavirus exposure or infection include, e.g., inflammation of the liver, decreased appetite, fatigue, abdominal pain, jaundice, flu-like symptoms, itching, muscle pain, joint pain, intermittent low-grade fevers, sleep disturbances, nausea, dyspepsia, cognitive changes, depression headaches and mood changes.

[0088] Typically, the immunogenic composition of the invention is administered in an amount sufficient to induce an immune response against an Ebolavirus. For therapeutic applications, the compositions should contain a therapeutically effective amount of an engineered Ebolavirus GP sequence or nanoparticle vaccine composition described herein. For prophylactic applications, the compositions should contain a prophylactically effective amount of the engineered Ebolavirus GP sequence or a nanoparticle displaying a GP protein. The appropriate amount of the polypeptide immunogen or the nanoparticle composition can be determined based on the specific disease or condition to be treated or prevented, severity, age of the subject, and other personal attributes of the specific subject (e.g., the general state of the subject's health and the robustness of the subject's immune system). Determination of effective dosages is additionally guided with animal model studies followed up by human clinical trials and is guided by administration protocols that significantly reduce the occurrence or severity of targeted disease symptoms or conditions in the subject.

[0089] For prophylactic applications, the immunogenic composition is provided in advance of any symptom, for example in advance of infection. The prophylactic

administration of the immunogenic compositions serves to prevent or ameliorate any subsequent infection. Thus, in some embodiments, a subject to be treated is one who has, or is at risk for developing, an Ebolavirus infection, for example because of exposure or the possibility of exposure to the Ebolavirus. Following administration of a therapeutically effective amount of the disclosed therapeutic compositions, the subject can be monitored for Ebolavirus infection, symptoms associated with Ebolavirus infection, or both.

[0090] For therapeutic applications, the immunogenic composition is provided at or after the onset of a symptom of disease or infection, for example after development of a symptom of an Ebolavirus infection, or after diagnosis of an Ebolavirus infection. The immunogenic composition can thus be provided prior to the anticipated exposure to Ebolaviruses in order to attenuate the anticipated severity, duration or extent of an infection and/or associated disease symptoms, after exposure or suspected exposure to the virus, or after the actual initiation of an infection.

[0091] The pharmaceutical composition of the invention can be combined with other agents known in the art for treating or preventing Ebolavirus infections. Administration of the pharmaceutical compositions and the known anti-viral agents can be either concurrently or sequentially.

[0092] Pharmaceutical compositions containing an engineered Ebolavirus GP protein or nanoparticle vaccine of the invention can be provided as components of a kit. Optionally, such a kit includes additional components including packaging, instructions and various other reagents, such as buffers, substrates, antibodies or ligands, such as control antibodies or ligands, and detection reagents. An optional instruction sheet can be additionally provided in the kits.

25

VII. Sequences of some engineered Ebolavirus GP polypeptides or structural motifs

EBOV GP ectodomain sequence (GenBank AAG40168.1) (SEQ ID NO:1):

MGVTGILQLP RDRFKRTSFFLWVILFORTFSI-PLGVIHN STLQVSDVDK
 30 LVCRDKLSST NQLRSVGLNL EGNVATDVP SATKRWGFRS GVPPKVVNVE
 AGEWAENCYN LEIKKPDGSE CLPAAPDGIR GFPRCRYVHK VSGTGPCAGD
 FAFHKEGAFF LYDRLASTVI YRGTTFAEGV VAFLLLPQAK KDFSSHPLR
 EPVNATEDPS SGYYSTTIRY QATGFGTNET EYLFEVDNLT YVQLESRFTP
 35 QFLLQLNETI YTSGKRSNTT GKLIWKVNPE IDTTIGEWAF WETKKNLTRK
 IRSEELSFTV VSNGAKNISG QSPARTSSDP GTNTTTEDHK IMASENSSAM

VQVHSQGREA AVSHLTTLAT ISTSPQSLTT KPGPDNSTHN TPVYKLDISE
 ATQVEQHHRR TDNDSTASDT PSATTAAGPP KAENTNTSKS TDFLDPATTT
 SPQNHSETAG NNNTHHQDTG EESASSGKLG LITNTIAGVA GLITGGRRTR
 REAIVNAQPK CNPNLHYWTT QDEGAAIGLA WIPYFGPAAE GIYIEGLMHN
 5 QDGLICGLRQ LANETTQALQ LFLRATTELRTFSILNRKAI DFLLRWGGT
 CHILGPDCCI EPHDWTKNIT DKIDQIIHDFVD - KTLPDQGD NDNWWTGWRQ
 WIPAGIGVTG VIIAVIALFC ICKFVF

Truncated EBOV GP_{ECTO} sequence (with leader/MLD/MPER deleted) (SEQ ID NO:2):
 10 PLGVIHNSTLQVSDVDKLVCRDKLSSTNQLRSVGLNLEGNGVATDVPSATKRW
 GFRSGVPPKVVNYEAGEWAENCYNLEIKKPDGSECLPAAPDGIRGFPRCRYVH
 KVSOGTGPCAGDFAFHKEGAFFLYDRLASTVIYRGTTFAEGVVAFLILPQAKKDF
 FSSHPLREPVNATEDPSSGYSTTIRYQATGFGTNETEYLFVVDNLTYYVQLESF
 TPQFLLQLNETIYTSRKRSNTTGKLIWKVNPEIDTTIGEWAFWETKKNLTRKIRS
 15 EELSFTTVVSTHHQDTGEESASSGKLGITNTIAGVAGLITGGRRTRREAIVNAQP
 KCNPNLHYWTTQDEGAAIGLAWIPYFGPAAEGLIYIEGLMHNQDGLICGLRQLA
 NETTQALQLFLRATTELRTFSILNRKAIDFLLQRWGGTCHILGPDCCIEPHDWTK
 NITDKIDQIIHDFVD

20 Leaderless GP_{Amuc} (GP_{ECTO} with leader/MLD/MPER deleted +T42A) (SEQ ID NO:3)
 PLGVIHNSALQVSDVDKLVCRDKLSSTNQLRSVGLNLEGNGVATDVPSATKR
 WGRSGVPPKVVNYEAGEWAENCYNLEIKKPDGSECLPAAPDGIRGFPRCRYV
 HKVSGTGPCAGDFAFHKEGAFFLYDRLASTVIYRGTTFAEGVVAFLILPQAKK
 25 DFFSSHPLREPVNATEDPSSGYSTTIRYQATGFGTNETEYLFVVDNLTYYVQLES
 RFTTPQFLLQLNETIYTSRKRSNTTGKLIWKVNPEIDTTIGEWAFWETKKNLTRKI
 RSEELSFTTVVSTHHQDTGEESASSGKLGITNTIAGVAGLITGGRRTRREAIVNA
 QPKCNPNLHYWTTQDEGAAIGLAWIPYFGPAAEGLIYIEGLMHNQDGLICGLRQ
 LANETTQALQLFLRATTELRTFSILNRKAIDFLLQRWGGTCHILGPDCCIEPHDW
 30 TKNITDKIDQIIHDFVD

Leaderless GP_{Amuc}-W615L (SEQ ID NO:4)
 PLGVIHNSALQVSDVDKLVCRDKLSSTNQLRSVGLNLEGNGVATDVPSATKR
 WGRSGVPPKVVNYEAGEWAENCYNLEIKKPDGSECLPAAPDGIRGFPRCRYV
 35 HKVSGTGPCAGDFAFHKEGAFFLYDRLASTVIYRGTTFAEGVVAFLILPQAKK
 DFFSSHPLREPVNATEDPSSGYSTTIRYQATGFGTNETEYLFVVDNLTYYVQLES
 RFTTPQFLLQLNETIYTSRKRSNTTGKLIWKVNPEIDTTIGEWAFWETKKNLTRKI
 RSEELSFTTVVSTHHQDTGEESASSGKLGITNTIAGVAGLITGGRRTRREAIVNA
 QPKCNPNLHYWTTQDEGAAIGLAWIPYFGPAAEGLIYIEGLMHNQDGLICGLRQ
 40 LANETTQALQLFLRATTELRTFSILNRKAIDFLLQRWGGTCHILGPDCCIEPHDL
 TKNITDKIDQIIHDFVD

Leaderless GP_{Amuc}-2WPZ (612/615 double mutation + 2WPZ extension) (SEQ ID NO:5)
 45 PLGVIHNSALQVSDVDKLVCRDKLSSTNQLRSVGLNLEGNGVATDVPSATKR
 WGRSGVPPKVVNYEAGEWAENCYNLEIKKPDGSECLPAAPDGIRGFPRCRYV
 HKVSGTGPCAGDFAFHKEGAFFLYDRLASTVIYRGTTFAEGVVAFLILPQAKK
 DFFSSHPLREPVNATEDPSSGYSTTIRYQATGFGTNETEYLFVVDNLTYYVQLES

RFTPQFLLQLNETIYTS GKRSNTTGKLIWKVNPEIDTTIGEWAFWETKKNLTRKI
 RSEELSFTVVSTHHQDTGEESASSGKLGLITNTIAGVAGLITGGRRTTRREAIVNA
 QPKCNPNLHYWTTQDEGAAIGLAWIPYFGPAAEGIYIEGLMHNQDGLICGLRQ
 LANETTQALQLFLRATTELRTFSILNRKAIDFLLQRWGGTCHILGPDCCIEGHDF
 5 TKQLEDKVEENLSKVYHNENEVARLKKLVGER

Leaderless GPΔmuc-L (C-terminal extension to residue 637) (SEQ ID NO:6)
 PLGVIHNSALQVSDVDKLVCRDKLSSTNQLRSVGLNLEGNGVATDVPSATKR
 WGFSGVPPKVVNYEAGEWAENCYNLEIKKPDGSECLPAAPDGIRGFPRCRYV
 10 HKVSGTGPCAGDFAFHKEGAFFLYDRLASTVIYRGTTFAEGVVAFLILPQAKK
 DFFSSHPLREPVNATEDPSSGYSTTIRYQATGFGTNETEYLFVNDLTYVQLES
 RFTPQFLLQLNETIYTS GKRSNTTGKLIWKVNPEIDTTIGEWAFWETKKNLTRKI
 RSEELSFTVVSTHHQDTGEESASSGKLGLITNTIAGVAGLITGGRRTTRREAIVNA
 QPKCNPNLHYWTTQDEGAAIGLAWIPYFGPAAEGIYIEGLMHNQDGLICGLRQ
 15 LANETTQALQLFLRATTELRTFSILNRKAIDFLLQRWGGTCHILGPDCCIEPHDW
 TKNITDKIDQIIHDFVDKTLDP

Leaderless GPΔmuc-Ext (C-terminal extension to residue 643) (SEQ ID NO:7)
 PLGVIHNSALQVSDVDKLVCRDKLSSTNQLRSVGLNLEGNGVATDVPSATKR
 20 WGFSGVPPKVVNYEAGEWAENCYNLEIKKPDGSECLPAAPDGIRGFPRCRYV
 HKVSGTGPCAGDFAFHKEGAFFLYDRLASTVIYRGTTFAEGVVAFLILPQAKK
 DFFSSHPLREPVNATEDPSSGYSTTIRYQATGFGTNETEYLFVNDLTYVQLES
 RFTPQFLLQLNETIYTS GKRSNTTGKLIWKVNPEIDTTIGEWAFWETKKNLTRKI
 RSEELSFTVVSTHHQDTGEESASSGKLGLITNTIAGVAGLITGGRRTTRREAIVNA
 25 QPKCNPNLHYWTTQDEGAAIGLAWIPYFGPAAEGIYIEGLMHNQDGLICGLRQ
 LANETTQALQLFLRATTELRTFSILNRKAIDFLLQRWGGTCHILGPDCCIEPHDW
 TKNITDKIDQIIHDFVDKTLDPDQGDNDN

Leaderless GPΔmuc-W615L-L-foldon (or termed GPΔmuc-WL²-foldon) (SEQ ID NO:8)
 PLGVIHNSALQVSDVDKLVCRDKLSSTNQLRSVGLNLEGNGVATDVPSATKR
 WGFSGVPPKVVNYEAGEWAENCYNLEIKKPDGSECLPAAPDGIRGFPRCRYV
 HKVSGTGPCAGDFAFHKEGAFFLYDRLASTVIYRGTTFAEGVVAFLILPQAKK
 DFFSSHPLREPVNATEDPSSGYSTTIRYQATGFGTNETEYLFVNDLTYVQLES
 35 RFTPQFLLQLNETIYTS GKRSNTTGKLIWKVNPEIDTTIGEWAFWETKKNLTRKI
 RSEELSFTVVSTHHQDTGEESASSGKLGLITNTIAGVAGLITGGRRTTRREAIVNA
 QPKCNPNLHYWTTQDEGAAIGLAWIPYFGPAAEGIYIEGLMHNQDGLICGLRQ
 LANETTQALQLFLRATTELRTFSILNRKAIDFLLQRWGGTCHILGPDCCIEPHDL
 TKNITDKIDQIIHDFVDKTLPDASGYIPEAPRDGQAYVRKDGWVLLSTFL
 40

Leaderless GPΔmuc-W615L-P1 (T576P) (SEQ ID NO:9)
 PLGVIHNSALQVSDVDKLVCRDKLSSTNQLRSVGLNLEGNGVATDVPSATKR
 WGFSGVPPKVVNYEAGEWAENCYNLEIKKPDGSECLPAAPDGIRGFPRCRYV
 HKVSGTGPCAGDFAFHKEGAFFLYDRLASTVIYRGTTFAEGVVAFLILPQAKK
 45 DFFSSHPLREPVNATEDPSSGYSTTIRYQATGFGTNETEYLFVNDLTYVQLES
 RFTPQFLLQLNETIYTS GKRSNTTGKLIWKVNPEIDTTIGEWAFWETKKNLTRKI
 RSEELSFTVVSTHHQDTGEESASSGKLGLITNTIAGVAGLITGGRRTTRREAIVNA
 QPKCNPNLHYWTTQDEGAAIGLAWIPYFGPAAEGIYIEGLMHNQDGLICGLRQ

LANETTQALQLFLRAPTELRTFSILNRKAIDFLLQRWGGTCHILGPDCCIEPHDL
TKNITDKIDQIIHDFVD

Leaderless GP Δ muc-W615L-P2 (T577P) (SEQ ID NO:10)

5 PLGVIHNSALQVSDVDKLVCRDKLSSTNQLRSVGLNLEGNGVATDVPSATKR
WGFRSGVPPKVVNYEAGEWAENCYNLEIKKPDGSECLPAAPDGIRGFPRCRYV
HKVSGTGPCAGDFAFHKEGAFFLYDRLASTVIYRGTTFAEGVV AFLILPQAKK
DFFSSHPLREPVNATEDPSSGYSTTIRYQATGFGTNETEYLFVNDLTYVQLES
RFTPQFLLQLNETIYTS GKRSNTTGKLIWKVNPEIDTTIGEWAFWETKKNLTRKI
10 RSEELSFTVVSTHHQDTGEESASSGKLGLITNTIAGVAGLITGGRRTRREAIVNA
QPKCNPNLHYWTTQDEGAAIGLAWIPYFGPAAEGIYIEGLMHNQDGLICGLRQ
LANETTQALQLFLRAPTELRTFSILNRKAIDFLLQRWGGTCHILGPDCCIEPHDL
TKNITDKIDQIIHDFVD

15 Leaderless GP Δ muc-W615L-P3 (E578P) (SEQ ID NO:11)

PLGVIHNSALQVSDVDKLVCRDKLSSTNQLRSVGLNLEGNGVATDVPSATKR
WGFRSGVPPKVVNYEAGEWAENCYNLEIKKPDGSECLPAAPDGIRGFPRCRYV
HKVSGTGPCAGDFAFHKEGAFFLYDRLASTVIYRGTTFAEGVV AFLILPQAKK
DFFSSHPLREPVNATEDPSSGYSTTIRYQATGFGTNETEYLFVNDLTYVQLES
20 RFTPQFLLQLNETIYTS GKRSNTTGKLIWKVNPEIDTTIGEWAFWETKKNLTRKI
RSEELSFTVVSTHHQDTGEESASSGKLGLITNTIAGVAGLITGGRRTRREAIVNA
QPKCNPNLHYWTTQDEGAAIGLAWIPYFGPAAEGIYIEGLMHNQDGLICGLRQ
LANETTQALQLFLRATTPRLTFSILNRKAIDFLLQRWGGTCHILGPDCCIEPHDL
TKNITDKIDQIIHDFVD

25

Leaderless GP Δ muc-W615L-P4 (L579P) (SEQ ID NO:12)

PLGVIHNSALQVSDVDKLVCRDKLSSTNQLRSVGLNLEGNGVATDVPSATKR
WGFRSGVPPKVVNYEAGEWAENCYNLEIKKPDGSECLPAAPDGIRGFPRCRYV
HKVSGTGPCAGDFAFHKEGAFFLYDRLASTVIYRGTTFAEGVV AFLILPQAKK
30 DFFSSHPLREPVNATEDPSSGYSTTIRYQATGFGTNETEYLFVNDLTYVQLES
RFTPQFLLQLNETIYTS GKRSNTTGKLIWKVNPEIDTTIGEWAFWETKKNLTRKI
RSEELSFTVVSTHHQDTGEESASSGKLGLITNTIAGVAGLITGGRRTRREAIVNA
QPKCNPNLHYWTTQDEGAAIGLAWIPYFGPAAEGIYIEGLMHNQDGLICGLRQ
LANETTQALQLFLRATTEPRTFSILNRKAIDFLLQRWGGTCHILGPDCCIEPHDL
35 TKNITDKIDQIIHDFVD

Leaderless GP Δ muc-W615L-P5 (R580P) (SEQ ID NO:13)

PLGVIHNSALQVSDVDKLVCRDKLSSTNQLRSVGLNLEGNGVATDVPSATKR
WGFRSGVPPKVVNYEAGEWAENCYNLEIKKPDGSECLPAAPDGIRGFPRCRYV
40 HKVSGTGPCAGDFAFHKEGAFFLYDRLASTVIYRGTTFAEGVV AFLILPQAKK
DFFSSHPLREPVNATEDPSSGYSTTIRYQATGFGTNETEYLFVNDLTYVQLES
RFTPQFLLQLNETIYTS GKRSNTTGKLIWKVNPEIDTTIGEWAFWETKKNLTRKI
RSEELSFTVVSTHHQDTGEESASSGKLGLITNTIAGVAGLITGGRRTRREAIVNA
QPKCNPNLHYWTTQDEGAAIGLAWIPYFGPAAEGIYIEGLMHNQDGLICGLRQ
45 LANETTQALQLFLRATTELPTFSILNRKAIDFLLQRWGGTCHILGPDCCIEPHDL
TKNITDKIDQIIHDFVD

Leaderless GP Δ muc-W615L-P6 (T581P) (SEQ ID NO:14)

PLGVIHNSALQVSDVDKLVCRDKLSSTNQLRSVGLNLEGNVATDVPSATKR
 WGRSGVPPKVVNYEAGEWAENCYNLEIKKPDGSECLPAAPDGIRGFPRCRYV
 HKVSGTGPCAGDFAFHKEGAFFLYDRLASTVIYRGTTFAEGVVAFLILPQAKK
 DFFSSHPLREPVNATEDPSSGYYSTTIRYQATGFGTNETEYLFVVDNLTYVQLES
 5 RFTPQFLLQLNETIYTS GKRSNTTGKLIWKVNPEIDTTIGEWAFWETKKNLTRKI
 RSEELSFTTVVSTHHQDTGEESASSGKLGLITNTIAGVAGLITGGRRTREAIVNA
 QPKCNPNLHYWTTQDEGAAIGLAWIPYFGPAAEGIYIEGLMHNQDGLICGLRQ
 LANETTQALQLFLRATTELRPFSILNRKAIDFLLQRWGGTCHILGPDCCIEPHDL
 TKNITDKIDQIIHDFVD

10

Leaderless GPAmuc-W615L-P7 (F582P) (SEQ ID NO:15)

PLGVIHNSALQVSDVDKLVCRDKLSSTNQLRSVGLNLEGNVATDVPSATKR
 WGRSGVPPKVVNYEAGEWAENCYNLEIKKPDGSECLPAAPDGIRGFPRCRYV
 HKVSGTGPCAGDFAFHKEGAFFLYDRLASTVIYRGTTFAEGVVAFLILPQAKK
 15 DFFSSHPLREPVNATEDPSSGYYSTTIRYQATGFGTNETEYLFVVDNLTYVQLES
 RFTPQFLLQLNETIYTS GKRSNTTGKLIWKVNPEIDTTIGEWAFWETKKNLTRKI
 RSEELSFTTVVSTHHQDTGEESASSGKLGLITNTIAGVAGLITGGRRTREAIVNA
 QPKCNPNLHYWTTQDEGAAIGLAWIPYFGPAAEGIYIEGLMHNQDGLICGLRQ
 LANETTQALQLFLRATTELRTPSILNRKAIDFLLQRWGGTCHILGPDCCIEPHDL
 20 TKNITDKIDQIIHDFVD

Leaderless GPAmuc-W615L-P8 (S583P) (SEQ ID NO:16)

PLGVIHNSALQVSDVDKLVCRDKLSSTNQLRSVGLNLEGNVATDVPSATKR
 WGRSGVPPKVVNYEAGEWAENCYNLEIKKPDGSECLPAAPDGIRGFPRCRYV
 25 HKVSGTGPCAGDFAFHKEGAFFLYDRLASTVIYRGTTFAEGVVAFLILPQAKK
 DFFSSHPLREPVNATEDPSSGYYSTTIRYQATGFGTNETEYLFVVDNLTYVQLES
 RFTPQFLLQLNETIYTS GKRSNTTGKLIWKVNPEIDTTIGEWAFWETKKNLTRKI
 RSEELSFTTVVSTHHQDTGEESASSGKLGLITNTIAGVAGLITGGRRTREAIVNA
 QPKCNPNLHYWTTQDEGAAIGLAWIPYFGPAAEGIYIEGLMHNQDGLICGLRQ
 30 LANETTQALQLFLRATTELRTFPILNRKAIDFLLQRWGGTCHILGPDCCIEPHDL
 TKNITDKIDQIIHDFVD

Leaderless GPAmuc-W615L-L-P2 (SEQ ID NO:17)

PLGVIHNSALQVSDVDKLVCRDKLSSTNQLRSVGLNLEGNVATDVPSATKR
 35 WGRSGVPPKVVNYEAGEWAENCYNLEIKKPDGSECLPAAPDGIRGFPRCRYV
 HKVSGTGPCAGDFAFHKEGAFFLYDRLASTVIYRGTTFAEGVVAFLILPQAKK
 DFFSSHPLREPVNATEDPSSGYYSTTIRYQATGFGTNETEYLFVVDNLTYVQLES
 RFTPQFLLQLNETIYTS GKRSNTTGKLIWKVNPEIDTTIGEWAFWETKKNLTRKI
 RSEELSFTTVVSTHHQDTGEESASSGKLGLITNTIAGVAGLITGGRRTREAIVNA
 40 QPKCNPNLHYWTTQDEGAAIGLAWIPYFGPAAEGIYIEGLMHNQDGLICGLRQ
 LANETTQALQLFLRATPELRTFSILNRKAIDFLLQRWGGTCHILGPDCCIEPHDL
 TKNITDKIDQIIHDFVDKTLPD

45 Leaderless GPAmuc-W615L-L-P2 + restriction site "AS" and foldon (SEQ ID NO:18)
 PLGVIHNSALQVSDVDKLVCRDKLSSTNQLRSVGLNLEGNVATDVPSATKR
 WGRSGVPPKVVNYEAGEWAENCYNLEIKKPDGSECLPAAPDGIRGFPRCRYV
 HKVSGTGPCAGDFAFHKEGAFFLYDRLASTVIYRGTTFAEGVVAFLILPQAKK
 DFFSSHPLREPVNATEDPSSGYYSTTIRYQATGFGTNETEYLFVVDNLTYVQLES

RFTPQFLLQLNETIYTSGKRSNTTGKLIWKVNPEIDTTIGEWAFWETKKNLTRKI
 RSEELSFTVVSTHHQDTGEESASSGKLGLITNTIAGVAGLITGGRRTTRREAIVNA
 QPKCNPNLHYWTTQDEGAAIGLAWIPYFGPAAEGIYIEGLMHNQDGLICGLRQ
 LANETTQALQLFLRATPELRTFSILNRKAIDFLLQRWGGTCHILGPDCCIEPHDL
 5 TKNITDKIDQIIHDFVDKTLPDASGYIPEAPRDGQAYVRKDGWVLLSTFL

Leaderless GPΔmuc-W615L-L-P2 + restriction site “AS” + G4S linker + 1TD0 (SEQ ID NO:19)
 10 PLGVIHNSALQVSDVDKLVCRDKLSSTNQLRSVGLNLEGNGVATDVPSATKR
 WGFRRSGVPPKVVNYEAGEWAENCYNLEIKKPDGSECLPAAPDGIRGFPRCRYV
 HKVSGTGPCAGDFAFHKEGAFFLYDRLASTVIYRGTTFAEGVVAFLILPQAKK
 DFFSSHPLREPVNATEDPSSGYYSTTIRYQATGFGTNETEYLFVNDLTYVQLES
 RFTPQFLLQLNETIYTSGKRSNTTGKLIWKVNPEIDTTIGEWAFWETKKNLTRKI
 15 RSEELSFTVVSTHHQDTGEESASSGKLGLITNTIAGVAGLITGGRRTTRREAIVNA
 QPKCNPNLHYWTTQDEGAAIGLAWIPYFGPAAEGIYIEGLMHNQDGLICGLRQ
 LANETTQALQLFLRATPELRTFSILNRKAIDFLLQRWGGTCHILGPDCCIEPHDL
 TKNITDKIDQIIHDFVDKTLPDASGGGGSEVRIFAGNDPAHTATGSSGISSPTPAL
 TPLMLDEATGKLVVWDGQKAGSAVGILVPLEGTETALTYKSGTFATEAIHW
 20 PESVDEHKKANAFAGSALSHAA

Leaderless GPΔmuc-W615L-L-P4 (or termed GPΔmuc-WL²P⁴) (SEQ ID NO:20)
 PLGVIHNSALQVSDVDKLVCRDKLSSTNQLRSVGLNLEGNGVATDVPSATKR
 WGFRRSGVPPKVVNYEAGEWAENCYNLEIKKPDGSECLPAAPDGIRGFPRCRYV
 25 HKVSGTGPCAGDFAFHKEGAFFLYDRLASTVIYRGTTFAEGVVAFLILPQAKK
 DFFSSHPLREPVNATEDPSSGYYSTTIRYQATGFGTNETEYLFVNDLTYVQLES
 RFTPQFLLQLNETIYTSGKRSNTTGKLIWKVNPEIDTTIGEWAFWETKKNLTRKI
 RSEELSFTVVSTHHQDTGEESASSGKLGLITNTIAGVAGLITGGRRTTRREAIVNA
 QPKCNPNLHYWTTQDEGAAIGLAWIPYFGPAAEGIYIEGLMHNQDGLICGLRQ
 30 LANETTQALQLFLRATTEPRTFSILNRKAIDFLLQRWGGTCHILGPDCCIEPHDL
 TKNITDKIDQIIHDFVDKTLPD

Leaderless GPΔmuc-W615L-L-P4 + restriction site “AS” and foldon (SEQ ID NO:21)
 PLGVIHNSALQVSDVDKLVCRDKLSSTNQLRSVGLNLEGNGVATDVPSATKR
 35 WGFRRSGVPPKVVNYEAGEWAENCYNLEIKKPDGSECLPAAPDGIRGFPRCRYV
 HKVSGTGPCAGDFAFHKEGAFFLYDRLASTVIYRGTTFAEGVVAFLILPQAKK
 DFFSSHPLREPVNATEDPSSGYYSTTIRYQATGFGTNETEYLFVNDLTYVQLES
 RFTPQFLLQLNETIYTSGKRSNTTGKLIWKVNPEIDTTIGEWAFWETKKNLTRKI
 RSEELSFTVVSTHHQDTGEESASSGKLGLITNTIAGVAGLITGGRRTTRREAIVNA
 40 QPKCNPNLHYWTTQDEGAAIGLAWIPYFGPAAEGIYIEGLMHNQDGLICGLRQ
 LANETTQALQLFLRATTEPRTFSILNRKAIDFLLQRWGGTCHILGPDCCIEPHDL
 TKNITDKIDQIIHDFVDKTLPDASGYIPEAPRDGQAYVRKDGWVLLSTFL

Leaderless GPΔmuc-W615L-L-P4 + restriction site “AS” +G4S linker + 1TD0 (SEQ ID NO:22)
 45 PLGVIHNSALQVSDVDKLVCRDKLSSTNQLRSVGLNLEGNGVATDVPSATKR
 WGFRRSGVPPKVVNYEAGEWAENCYNLEIKKPDGSECLPAAPDGIRGFPRCRYV
 HKVSGTGPCAGDFAFHKEGAFFLYDRLASTVIYRGTTFAEGVVAFLILPQAKK
 DFFSSHPLREPVNATEDPSSGYYSTTIRYQATGFGTNETEYLFVNDLTYVQLES

RFTPQFLLQLNETIYTS GKRSNTTGKLIWKVNPEIDTTIGEWAFWETKKNLTRKI
 RSEELSFTVVSTHHQDTGEESASSGKLGLITNTIAGVAGLITGGRRTTRREAIVNA
 QPKCNPNLHYWTTQDEGAAIGLAWIPYFGPAAEGIYIEGLMHNQDGLICGLRQ
 LANETTQALQLFLRATTEPRTFSILNRKAIDFLLQRWGGTCHILGPDCCIEPHDL
 5 TKNITDKIDQIIHDFVDKTLPDASGGGGSEVRIFAGNDPAHTATGSSGISSPTPAL
 TPLMLDEATGKLVVWDGQKAGSAVGILVLPLEGTETALTYYKSGTFATEAIHW
 PESVDEHKKANAFAGSALSHAA

Leaderless GP Δ muc-SS1 (Y534C/F153C, rC_{β} - C_{β} =3.819A) (SEQ ID NO:23)
 10 PLGVIHNSALQVSDVDKLVCRDKLSSTNQLRSVGLNLEGNGVATDVPSATKR
 WGFRRSGVPPKVVNYEAGEWAENCYNLEIKKPDGSECLPAAPDGIRGFPRCRYV
 HKVSGTGPCAGDFAFHKEGAFFLYDRLASTVIYRGTTFAEGVVAFLILPQAKK
 DFFSSHPLREPVNATEDPSSGYYSTTIRYQATGFGTNETEYLFVEVDNLTYVQLES
 RFTPQFLLQLNETIYTS GKRSNTTGKLIWKVNPEIDTTIGEWAFWETKKNLTRKI
 15 RSEELSFTVVSTHHQDTGEESASSGKLGLITNTIAGVAGLITGGRRTTRREAIVNA
 QPKCNPNLHYWTTQDEGAAIGLAWIPCFGPAAEGIYIEGLMHNQDGLICGLRQ
 LANETTQALQLFLRATTELRTFSILNRKAIDFLLQRWGGTCHILGPDCCIEPHDW
 TKNITDKIDQIIHDFVD

20 Leaderless GP Δ muc-SS2 (A538C/G91C, rC_{β} - C_{α} =3.994A) (SEQ ID NO:24)
 PLGVIHNSALQVSDVDKLVCRDKLSSTNQLRSVGLNLEGNGVATDVPSATKR
 WGFRRSCVPPKVVNYEAGEWAENCYNLEIKKPDGSECLPAAPDGIRGFPRCRYV
 HKVSGTGPCAGDFAFHKEGAFFLYDRLASTVIYRGTTFAEGVVAFLILPQAKK
 DFFSSHPLREPVNATEDPSSGYYSTTIRYQATGFGTNETEYLFVEVDNLTYVQLES
 25 RFTPQFLLQLNETIYTS GKRSNTTGKLIWKVNPEIDTTIGEWAFWETKKNLTRKI
 RSEELSFTVVSTHHQDTGEESASSGKLGLITNTIAGVAGLITGGRRTTRREAIVNA
 QPKCNPNLHYWTTQDEGAAIGLAWIPYFGPCAEGIYIEGLMHNQDGLICGLRQ
 LANETTQALQLFLRATTELRTFSILNRKAIDFLLQRWGGTCHILGPDCCIEPHDW
 TKNITDKIDQIIHDFVD

30 Leaderless GP Δ muc-SS3 (T520C/A575C, rC_{β} - C_{β} =4.008A) (SEQ ID NO:25)
 PLGVIHNSALQVSDVDKLVCRDKLSSTNQLRSVGLNLEGNGVATDVPSATKR
 WGFRRSGVPPKVVNYEAGEWAENCYNLEIKKPDGSECLPAAPDGIRGFPRCRYV
 HKVSGTGPCAGDFAFHKEGAFFLYDRLASTVIYRGTTFAEGVVAFLILPQAKK
 35 DFFSSHPLREPVNATEDPSSGYYSTTIRYQATGFGTNETEYLFVEVDNLTYVQLES
 RFTPQFLLQLNETIYTS GKRSNTTGKLIWKVNPEIDTTIGEWAFWETKKNLTRKI
 RSEELSFTVVSTHHQDTGEESASSGKLGLITNTIAGVAGLITGGRRTTRREAIVNA
 QPKCNPNLHYWTCQDEGAAIGLAWIPYFGPAAEGIYIEGLMHNQDGLICGLRQ
 LANETTQALQLFLRCTTELRTFSILNRKAIDFLLQRWGGTCHILGPDCCIEPHDW
 40 TKNITDKIDQIIHDFVD

Leaderless GP Δ muc-SS4 (I532C/G157C, rC_{β} - C_{α} =4.202A) (SEQ ID NO:26)
 PLGVIHNSALQVSDVDKLVCRDKLSSTNQLRSVGLNLEGNGVATDVPSATKR
 WGFRRSGVPPKVVNYEAGEWAENCYNLEIKKPDGSECLPAAPDGIRGFPRCRYV
 45 HKVSGTGPCAGDFAFHKECAFFLYDRLASTVIYRGTTFAEGVVAFLILPQAKKD
 FFFSSHPLREPVNATEDPSSGYYSTTIRYQATGFGTNETEYLFVEVDNLTYVQLESR
 FTPQFLLQLNETIYTS GKRSNTTGKLIWKVNPEIDTTIGEWAFWETKKNLTRKIR
 SEELSFTVVSTHHQDTGEESASSGKLGLITNTIAGVAGLITGGRRTTRREAIVNAQ
 PKCNPNLHYWTTQDEGAAIGLAWCPYFGPAAEGIYIEGLMHNQDGLICGLRQL

ANETTQALQLFLRATTELRTFSILNRKAIDFLLQRWGGTCHILGPDCCIEPHDWT
KNITDKIDQIIHDFVD

- Leaderless GP Δ muc-SS5 (D522C/A575C, rC_{β} - C_{β} =4.410Å) (SEQ ID NO:27)
- 5 PLGVIHNSALQVSDVDKLVCRDKLSSTNQLRSVGLNLEGNGVATDVPSATKR
WGFRSGVPPKVVNYEAGEWAENCYNLEIKKPDGSECLPAAPDGIRGFPRCRYV
HKVSGTGPCAGDFAFHKEGAFFLYDRLASTVIYRGTTFAEGVV AFLILPQAKK
DFSSHPLREPVNATEDPSSGYYSTTIRYQATGFGTNETEYLFVNDLTYVQLES
RFTPQFLLQLNETIYTS GKRSNTTGKLIWKVNPEIDTTIGEWAFWETKKNLTKI
- 10 RSEELSFTVVSTHHQDTGEESASSGKLGLITNTIAGVAGLITGGRRTRREAIVNA
QPKCNPNLHYWTTQCEGAAIGLAWIPYFGPAAEGIYIEGLMHNQDGLICGLRQ
LANETTQALQLFLRCTTELRTFSILNRKAIDFLLQRWGGTCHILGPDCCIEPHDW
TKNITDKIDQIIHDFVD

- 15 Leaderless GP Δ muc-SS6 (K56C/G599C, rC_{β} - C_{α} =4.597Å) (SEQ ID NO:28)
PLGVIHNSALQVSDVDCLVCRDKLSSTNQLRSVGLNLEGNGVATDVPSATKRW
GFRSGVPPKVVNYEAGEWAENCYNLEIKKPDGSECLPAAPDGIRGFPRCRYVH
KVS GTGPCAGDFAFHKEGAFFLYDRLASTVIYRGTTFAEGVV AFLILPQAKKDF
FSSHPLREPVNATEDPSSGYYSTTIRYQATGFGTNETEYLFVNDLTYVQLESRF
- 20 TPQFLLQLNETIYTS GKRSNTTGKLIWKVNPEIDTTIGEWAFWETKKNLTKIRS
EELSFTVVSTHHQDTGEESASSGKLGLITNTIAGVAGLITGGRRTRREAIVNAQP
KCNPNLHYWTTQDEGAAIGLAWIPYFGPAAEGIYIEGLMHNQDGLICGLRQLA
NETTQALQLFLRATTELRTFSILNRKAIDFLLQRWGCTCHILGPDCCIEPHDWT
NITDKIDQIIHDFVD

- 25 Foldon (SEQ ID NO:29)
GYIPEAPRDGQAYVRKDG EWVLLSTFL:

- 1TD0 (SEQ ID NO:30)
- 30 EVRIFAGNDPAHTATGSSGISSPTPALTPLMLDEATGKLVVWDGQKAGSAVGIL
VLPLEGTETALTYKSGTFATEAIHWPE SVDEHKKANAFAGSALSHAA:

- GCN leucine zipper motif (SEQ ID NO:34)
KQLED KVEEN LSKVY HNENE VARLK KLVGE R
- 35

Locking domain LD4 (SEQ ID NO:39):
FSEEQKKALDLAFYFDRRLTPEWRRYLSQRLGLN EEQIERWFRRKEQQIGWSH
PQFEK

- 40 Locking domain LD7 (SEQ ID NO:40):
SPAVDIGDRLDELEKALEALSAEDGHDDVGQRLESLLRRWNSRRAD

Leader sequence of EBOV GP (SEQ ID NO:41)
MGVTGILQLPRDRFKRTSFFLWVILFQRTFSI

45

EXAMPLES

[0093] The following examples are offered to illustrate, but not to limit the present invention.

5 Example 1 Tag-free immunoaffinity purification of EBOV GP trimers

[0094] EBOV GP contains a heavily glycosylated mucin-like domain (MLD), which shields the glycan cap and neutralizing epitopes in GP1 and GP2. A soluble, mucin-deleted form of *Zaire* EBOV GP (GP Δ muc) that produced a high-resolution crystal structures (Zhao et al., Nature 535: 169-172, 2016) was used as a basic construct
 10 to investigate GP metastability. In HIV-1 vaccine research, immunoaffinity chromatography (IAC) columns based on bNAbs 2G12 and PGT145 (67, 68) have been widely used for the purification of native-like Env trimers. While 2G12 targets a glycan patch on a single gp120, PGT145 binds the trimer apex and can separate closed trimers from partially open and misfolded Envs. These two bNAb columns have also been used
 15 to purify HIV-1 gp140-presenting NPs. Such GP-specific antibody columns have not been reported in EBOV vaccine research. Recently, Corti et al. identified two potent NAbs, mAb114 and mAb100, from a human survivor (*Science* 351, 1339-1342, 2016). Misasi et al. elucidated the structural basis for neutralization by mAb114, which targets the receptor binding site (RBS), and mAb100, which interacts with the GP1/GP2
 20 interface and internal fusion loop (IFL) of two GP subunits (*Science* 351, 1343-1346, 2016).

[0095] Here, we examined the utility of mAb114 and mAb100 as IAC columns. The GP Δ muc constructs with and without a C-terminal trimerization motif, foldon, were transiently expressed in 250 ml HEK293 F cells and purified on an antibody
 25 column prior to size-exclusion chromatography (SEC) using a Superdex 200 10/300 GL column and blue native polyacrylamide gel electrophoresis (BN-PAGE). With mAb114, both GP Δ muc samples showed aggregate (~9ml), dimer (~12ml), and monomer (~13.5-14ml) peaks in the SEC profiles, but only GP Δ muc-foldon showed a visible trimer peak (~10.5-11ml) in SEC and a faint band of slightly less than 440 kD
 30 on the BN gel. Following mAb100 purification, GP Δ muc showed overall low yield, whereas GP Δ muc-foldon demonstrated high trimer purity with no dimer and monomer peaks. Consistently, GP Δ muc-foldon registered a single trimer band on the BN gel.

Taken together, while both mAb114 and mAb100 columns can be used for GP purification, mAb100 offers an effective IAC method for purifying native-like GP trimers due to its recognition of a quaternary GP epitope. The stark difference in trimer yield between the two soluble GP Δ muc constructs suggests a strong tendency for trimer dissociation.

Example 2 Effect of the HR2 stalk on EBOV GP metastability

[0096] EBOV GP possesses a long, extended HR2 stalk. Even in the high-resolution GP Δ muc structures, the HR2 stalk still contains less helical content than most coiled-coils in the database, ~15 aa vs ~30 aa, and becomes unwound towards the C terminus, suggesting an inherent instability in HR2. Recently, King et al. solved a 3.17Å-resolution structure for the MARV GP Δ muc trimer in complex with a therapeutic human mAb, MR191 (*Cell Host Microbe* 23, 101-109.e104, 2018). Surprisingly, the MARV HR2 stalk adopted a well-formed coiled-coil with canonical sidechain packing along the three-fold axis. To identify the cause of this difference, we obtained EBOV and MARV GP sequences from the Virus Pathogen Database and Analysis Resource (ViPR). A total of 274 EBOV GPs and 87 MARV GPs were used for sequence conservation analysis of the region spanning the CX₆CC motif and the HR2 stalk, aa 601-632 for EBOV and aa 602-633 for MARV, respectively. Most inward-facing amino acids were conserved except for W615 in EBOV, or L616 in MARV. Indeed, structural analysis revealed a critical difference at this position: three W615s in EBOV GP (PDB: 5JQ3) formed a wide triangle at the neck of the coiled-coil with a C α distance of 11.1 Å and a C β distance of 9.0 Å; in contrast, with a smaller and more hydrophobic L616, a C α distance of 10.5 Å and a C β distance of 8.3 Å were observed in MARV GP (PDB: 6BP2).

[0097] Based on this finding, we hypothesized that a W615L mutation may stabilize the EBOV GP trimer. To further examine the effect of the stalk, we created three GP Δ muc constructs by replacing aa 617-632 with a GCN4 leucine zipper (PDB: 2WPZ, aa 3-33) and by extending the C terminus to 637 and 643 to include a newly identified bNAb epitope (King et al., *Nat Commun* 10, 1788, 2019) that spans HR2 and the membrane-proximal external region (MPER), termed “L” and “Ext”, respectively (Fig. 3A). Double mutation P612G/W615F was introduced to the GP Δ muc-

2WPZ construct between the CX₆CC motif and the coiled-coil to reduce any structural strain in this region. These constructs were characterized by SEC and BN-PAGE following transient expression in 250-ml 293 F cells and mAb114 purification. Indeed, all three designs increased trimer yield with GPΔmuc-2WPZ showing the most visible trimer peak in SEC (Fig. 3B). Consistently, trimer bands were observed for all three constructs on the BN gel, albeit with less intensity for GPΔmuc-L and -Ext (Fig. 3C, left). Upon mAb100 purification, all three GPΔmuc variants showed more visible trimer bands than wildtype GPΔmuc (Fig. 3C, right), supporting the notion that the HR2 stalk is critical to GP trimer stability.

10 [0098] We next combined the W615L mutation and the “L” extension in a single construct named GPΔmuc-W615L-L-foldon, or simply GPΔmuc-WL²-foldon. This construct, along with GPΔmuc-foldon, was expressed transiently in 1-liter 293 F cells and purified using an mAb100 column prior to SEC on a HiLoad Superdex 200 16/600 GL column. In three production runs, GPΔmuc-WL²-foldon consistently outperformed

15 the wildtype construct, showing a two-fold higher trimer peak in the SEC profile and an ~2.6-fold greater trimer yield after SEC (1.3 mg vs 0.5 mg). Thermostability was assessed by differential scanning calorimetry (DSC) for two purified GP trimers. The thermal denaturation midpoint (T_m) value of the stalk-stabilized trimer was 3 °C higher than that of the wildtype trimer (67 °C vs 64 °C). Consistently, stalk stabilization also

20 increased the onset temperature (T_{on}) from 52.4 °C to 62.5 °C, with a narrower half *width* of the peak ($\Delta T_{1/2}$) than the wildtype trimer (3.8 °C vs. 5.1 °C). Antigenicity was assessed for four mAb100/SEC-purified EBOV GP trimers in enzyme-linked immunosorbent assay (ELISA) (Fig. 3D-E). A panel of 10 antibodies was used, including three NAbs targeting the base (KZ52, c2G4 and c4G7), two human NAbs –

25 mAb100 (IFL) and mAb114 (RBS), a non-NAb directed to the glycan cap (c13C6), and four pan-Ebolavirus bNAbs targeting the HR2-MPER epitope (BDBV223) and IFL (ADI-15878, ADI-15946, and CA45). The GPΔmuc trimer showed notably improved antibody binding with respect to the GP_{ECTO} trimer with an up to 7.6-fold difference in half maximal effective concentration (EC₅₀), indicating that MLD can effectively shield

30 GP from antibody recognition. The two HR2 stalk modifications led to further enhanced recognition of RBS by mAb114 and IFL by CA45, respectively, although the EC₅₀ values only showed moderate changes. A ~40% reduction in EC₅₀ was observed

for GP Δ muc-WL²-foldon binding to BDBV223. Taken together, two HR2 stalk mutations can effectively improve trimer yield, thermostability, and antigenicity for EBOV GP.

5 Example 3 Effect of the HR1c bend on EBOV GP metastability

[0099] We hypothesized that HR1c is essential to EBOV GP metastability. Since HR1c in wildtype EBOV GP is equivalent in length (8 aa) to a truncated HR1_N in the prefusion-optimized HIV-1 Env, metastability in EBOV GP may not be sensitive to the HR1c length and likely requires a different solution. We thus hypothesized that a
10 proline mutation in HR1c, termed P¹⁻⁸, may rigidify the HR1c bend and improve the EBOV GP trimer stability.

[00100] To examine this possibility, eight GP Δ muc-W615L variants, each bearing a proline mutation in HR1c but without the L extension and foldon at the C terminus, were validated experimentally. All constructs were transiently expressed in 250-ml 293
15 F cells and purified using an mAb114 column, which captures all GP species. The proline mutation at most positions in HR1c showed little effect on the composition of GP species except for T577P (P²) and L579P (P⁴), which displayed notable trimer peaks at ~11ml in the SEC profiles. In a separate experiment, all eight constructs were transiently expressed in 250-ml 293 F cells and purified using an mAb100 column.
20 Only P² and P⁴ showed any measurable trimer yield, with a notably high SEC peak observed for P⁴ that corresponds to well-formed trimers. The mAb100-purified GP was also analyzed by BN-PAGE, which showed a trimer band for P² and P⁴. Overall, the T577P mutation, P², can substantially increase trimer yield, while the L579P mutation, P⁴, exhibited a less pronounced effect.

25 [00101] Next, the T577P mutation (P²) was incorporated into the GP Δ muc-WL²-foldon construct, resulting in a construct named GP Δ muc-WL²P²-foldon. This construct was expressed transiently in 1-liter 293 F cells and purified using an mAb100 column for SEC characterization on a HiLoad Superdex 200 16/600 GL column. In three production runs, GP Δ muc-WL²P²-foldon generated a trimer peak that was two- and
30 four-fold higher than GP Δ muc-WL²-foldon and wildtype GP Δ muc-foldon, respectively, with an average yield of 2.6 mg after SEC. Protein collected in the SEC range of 55.5-62.0 ml was analyzed by BN-PAGE, which displayed a trimer band across all fractions

without any hint of impurity. The thermostability of GP Δ muc-WL²P²-foldon was determined by DSC after mAb100 and SEC purification.

[00102] Unexpectedly, two transition peaks were observed in the thermogram, one registered at a lower T_m of 61.6 °C and the other at a higher T_m of 68.2 °C. To this end, a second construct bearing the L579P mutation (P⁴), termed GP Δ muc-WL²P⁴-foldon, was also assessed by DSC. Although only one peak was observed in the thermogram with a T_m of 67.0 °C, a slight widening at the onset of the peak suggested a similar unfolding behavior upon heating. DSC thus revealed the complexity associated with a proline-rigidified HR1_C bend, which may increase the trimer yield at the cost of reducing GP thermostability. The antigenicity of GP Δ muc-WL²P²-foldon was assessed using the same panel of 10 antibodies by ELISA (Fig. 3F-G) and bio-layer interferometry (BLI). The T577P mutation (P²) appeared to improve GP binding to most antibodies with respect to GP Δ muc-WL²-foldon (Fig. 3G), with a 40% reduction in EC₅₀ observed for bNAb BDBV223, which targets HR2-MPER. Although BLI profiles were almost indistinguishable between wildtype and redesigned GP Δ muc-foldon trimers – all with fast on-rates and flattened dissociation curves, a two-fold higher signal at the lowest concentration (12.5 nM) was observed for GP Δ muc-WL²P²-foldon binding to bNAb BDBV223, consistent with the ELISA data.

[00103] Our results thus demonstrated the importance of HR1_C to EBOV GP metastability and an unexpected sensitivity of HR1_C to proline mutation. Recently, Rutten et al. tested proline mutations in HR1_C along with a K588F mutation to stabilize filovirus GP trimers (*Cell Rep.* 30, 4540–50, 2020). While a similar pattern of increased trimer yield was observed for the T577P mutant, the reported thermostability data appeared to be inconsistent with our DSC measurement. Further investigation is warranted to fully understand the role of HR1_C in filovirus-cell fusion and its impact on GP stability.

Example 4 GP stabilization with disulfide bond mutations

[00104] Since EBOV GP already contains an endogenous SS bond linking GP1 and GP2 (C53-C609), we examined whether inter-GP SS bonds can be used to promote trimer formation and to lock GP in a “closed” trimer. Based on a high-resolution EBOV GP Δ muc structure (PDB: 5JQ3), we identified inter-GP amino acid pairs whose C_α-C_α

distances are within a cutoff of 4.72 Å. A total of nine pairs were identified. After visual inspection, three were removed as they may interfere with an existing SS bond or a hydrophobic cluster. The remaining six were divided into three groups: the IFL-head group (SS1/2/4), the IFL-NHR group (SS/5), and the HR2 group (SS6). Six GPΔmuc-SS constructs were designed and then characterized by SEC following transient expression in 250-ml 293 F cells and mAb114 purification. Diverse SEC profiles were observed, with SS2 showing a substantial trimer peak, consistent with a band of slightly below 440 kD on the BN gel. The mAb100-purified materials were also analyzed by BN-PAGE, with trimer bands observed for SS2, SS3, and SS5. Antigenicity was assessed for the three SS mutants in ELISA using six antibodies. While all three SS mutants outperformed wildtype GPΔmuc, SS2 showed higher affinity for NAbs targeting the base and IFL. Taken together, a well-positioned inter-GP SS bond can effectively stabilize EBOV GP in a native-like trimer conformation.

15 Example 5 Crystallographic analysis of redesigned EBOV GPΔmuc trimers

[00105] To understand how the stalk and HR1c mutations affect EBOV GP, we solved crystal structures for an unliganded GPΔmuc-foldon trimer with WL² and WL²P² at 2.3 Å and 3.2 Å, respectively. Both proteins showed a three-lobed, *chalice*-like structure, with Cα root-mean-square deviations (r.m.s.d.) of 0.92-1.14 Å to a high-resolution structure (PDB: 5JQ3) at a single subunit level. WL²P² yielded a more complete structure than WL² at the glycan cap (R302-V310) and the HR2 stalk (I627-D637). In the WL²P² structure, the glycan cap covers the RBS with glycan moieties visible for N238/N257/N268 in GP1 and for N563 in GP2, while in the WL² structure for N238/N257 in GP1 and N563/N618 in GP2. GP1 consists mainly of β-strands which form a broad semi-circular groove that clamps the α3 helix and the β19-β20 strands in GP2. The T577P mutation appeared to have a minimum effect on the conformation of the HR1c bend, as indicated by a Cα r.m.s.d. of 0.19 Å for this 8-aa segment.

[00106] In the WL²P² structure, the backbone carbonyl (CO) groups of R574, A575, and T576 in one subunit formed moderate-to-strong hydrogen bonds with the head group of R164 side chain in an adjacent subunit, whereas only one CO-NH distance was within the 3.5 Å cutoff in wildtype GPΔmuc. The W615L mutation exerted a

visible structural effect on the HR2 stalk. We have speculated that a bulky, inward-facing W615 at the top of the coiled-coil destabilizes the HR2 stalk and a W615L mutation would improve its packing. Indeed, the $\text{Ca-C}\alpha/\text{C}\beta\text{-C}\beta$ distances between two W615s of adjacent subunits in wildtype GP Δ muc, 11.1/9.0 Å, were reduced to 10.1/8.0 Å and 10.6/8.2 Å in WL² and WL²P², respectively. As a result, the coiled-coil region in EBOV HR2 stalk added one more helical turn, thus resembling MARV HR2 stalk. The “L” extension in the WL²P² structure could be fully modeled to D637 as a well-ordered loop anchored to the C-terminal foldon motif, rendering a complete HR2 stalk and partial MPER. Superposition of HR2 stalks yielded Ca r.m.s.d. values of 1.46 Å, 2.05 Å, and 1.85 Å with respect to EBOV-Mayinga (PDB: 5JQ3), SUDV (PDB: 3S88), and BDBV (PDB: 6EA5) GPs, respectively, suggesting some degree of structural variability in this region.

[010107] The WL²P² structure was then compared to a recently reported *Makona* GP Δ muc structure (PDB: 6VKM) containing the T577P/K588F mutation. In total, 353 of 398 residues in the WL²P² structure matched with the double mutant with a Ca r.m.s.d. of 0.89 Å. A more complete cathepsin cleavage loop was modeled in WL²P² than previous structures (aa 197-210 vs. aa 193-213), suggesting that this loop bridges over the IFL and interacts with IFL-directed NAbs such as mAb100 (42). In addition, we observed more electron densities for the β 18 loop in the glycan cap (aa 294-310) and the HR2 stalk than the double mutant. For the HR1_C bend, WL²P² showed more favorable hydrogen bonding patterns with a Ca r.m.s.d of 0.26 Å. A Ca r.m.s.d of 1.7 Å was obtained for the IFL region between the two structures. Lastly, the WL²P² structure was docked into a panel of known GP/antibody complexes. Overall, WL²P² has preserved all critical GP-antibody contacts. The mAb100/GP complex is of most interest as mAb100 has been used for GP purification with substantial purity. Cryo-EM has revealed additional electron density near the mAb100 light chain that likely corresponds to portions of the β 13– β 14 loop (aa 190-210) (Misari et al., Science 351: 1343-46, 2016). However, this density was not observed in a 6.7 Å-resolution crystal structure of the same complex (PDB: 5FHC). In the WL²P² structure, density could be seen for up to H197, which would be in the proximity of mAb100 light chain in the docked WL²P²/mAb100 model. Collectively, our structures validated the rationally designed mutations and provide atomic details for regions that are unavailable in

previous structures. The WL²P² structure also provides a potential explanation for the increased trimer yield, although the cause of the two-peak thermogram remains unclear.

Example 6 Display of EBOV GPΔmuc trimers on multilayered hyperstable

5 nanoparticles

- [00108]** Self-assembling NPs provide an alternative to recombinant VLPs for vaccine development. We tested several protein NPs as “multivalent carriers” to display the redesigned GPΔmuc trimers for in vivo assessment. Specifically, we modeled the GPΔmuc trimer structure on ferritin (FR), E2p, and I3-01, resulting in GP-presenting
- 10 NPs of 34.5 nm, 45.9 nm, and 49.2 nm, respectively. Superposition of the GPΔmuc C termini onto the FR and E2p N termini yielded Cα r.m.s.d. values of 7.0 Å and 5.5 Å, suggesting that GPΔmuc can be fused to FR with a short G4S linker and to E2p without linker, respectively. However, the large spacing between the N termini of I3-01 subunits (~50.5 Å) requires a long linker to connect with the C termini of a GPΔmuc
- 15 trimer, which form a long, narrow stalk. Based on computational modeling, a 10-aa (G4S)₂ linker would suffice and result in a Cα r.m.s.d. of 0.8 Å. Here, we first displayed two GPΔmuc trimers, wildtype and WL²P², on FR, E2p, and I3-01 with a 5-aa linker, no linker, and a 10-aa linker, respectively. All six GP-NP constructs were transiently expressed in 100-ml ExpiCHO cells followed by mAb100 purification and SEC on a
- 20 Superose 6 10/300 GL column. Overall, WL²P² outperformed wildtype GPΔmuc with greater NP yield and purity. Based on molecular weight (m.w.), the SEC peaks centered at 15 ml could be unassembled GP-NP species, suggesting an inherent instability for wildtype E2p and I3-01. The mAb100-purified GPΔmuc-WL²P²-presenting NP samples were further analyzed by negative stain EM, showing NPs mixed with impurity.
- 25 **[00109]** Previously, we demonstrated the use of a pan-reactive T-cell epitope as both a linker and as a built-in T-cell help in an HIV-1 NP construct based on I3-01 (55), suggesting that additional structural and functional components can be incorporated into such large 60-mers. Here, we sought to stabilize the E2p and I3-01 NPs via engineering. To this end, we fused a dimeric locking domain (LD) to the C terminus of
- 30 an NP subunit, and then a T-cell epitope to the C terminus of a LD. We hypothesized that LD can stabilize the non-covalent NP-forming interface and the T-cell epitopes can form a cluster at the NP core to elicit a strong T-cell response upon vaccination. To test

this hypothesis, we selected nine LDs from 815 homodimers in the protein database (PDB) (Fig. 4A). Based on structural modeling, LDs 1-7 were tested for E2p and five LDs (4-5 and 7-9) were tested for I3-01, all displaying GP Δ muc-WL²P². Following transient expression in 100-ml ExpiCHO cells and mAb100 purification, 12 LD-
 5 containing NP samples were characterized by SEC. Notably, LD4 and LD7 increased the NP peak (UV₂₈₀ value) by 5- and 2.5-fold for E2p and I3-01, respectively, with substantially improved NP purity. Further incorporation of a T-cell epitope, PADRE, didn't alter or slightly improved the NP yield and purity.

[00110] A total of seven GP-NP samples, with three variants for each 60-mer, were
 10 further analyzed by BN-PAGE. FR and two E2p variants displayed a single high-m.w. band corresponding to well-formed NPs, whereas wildtype E2p and all three I3-01 samples showed additional low-m.w. bands at 232-440 kD on the gel, indicative of unassembled GP-NP species. The mAb100/SEC-purified GP Δ muc-WL²P²-presenting FR, E2p-LD4-PADRE (E2p-L4P), and I3-01-LD7-PADRE (I3-01-L7P) samples were
 15 analyzed by negative stain EM. In addition to high-purity NPs for all three samples, an array of well-formed GP Δ muc spikes could be readily seen on the surface of FR and E2p-L4P NPs, consistent with SEC and BN-PAGE. Antigenicity was assessed for these three purified NP samples by ELISA against the same antibody panel (Fig. 4B-C). Compared to the individual WL²P² trimer, three NPs exhibited an epitope-specific
 20 binding pattern. Overall, NP display appeared to improve antibody binding to the RBS and glycan cap in GP1 and reduce antibody binding for bNAbs targeting the base and IFL at the GP1/GP2 interface and the GP2 stalk. This finding raised concerns that these conserved bNAb epitopes at the GP base and IFL on the NP-displayed trimers may not be as accessible as on the soluble trimers. To this end, BLI was performed to further
 25 probe antibody recognition of various GP epitopes (Fig. 4D-G). Using comparable molar concentrations for GP, three NPs showed considerably higher binding signals than the soluble trimer, suggesting that GP-presenting NPs may be more effective in promoting BCR clustering due to the avidity effect. Nonetheless, three NPs based on FR, E2p-L4P, and I3-01v9-L7P (Fig. 4H), in which I3-01v9 is a previously reported I3-
 30 01 variant, were selected for in vivo evaluation and comparison with trimers in mice and rabbits.

Example 7 Immunogenicity of EBOV GP trimers and NPs in BALB/c mice

[00111] We immunized BALB/c mice four times with three-week intervals to obtain an initial readout of immunogenicity for three EBOV GP/GPΔmuc trimers and three NPs. A soluble GP_{ECTO} trimer, GP-foldon, was included to represent the wildtype GP (with MLD) used by EBOV vaccines in clinical testing. For the I3-01v9-L7P group, mice were immunized with 20μg mAb100-purified instead of 50μg mAb100/SEC-purified protein due to the low yield of this NP. We first assessed the GP-specific antibody response in mouse sera using GPΔmuc-WL²P²-1TD0 as a probe, in which 1TD0 (PDB: 1TD0) is a trimerization motif (Fig. 5A-C). Both GPΔmuc groups significantly outperformed the GP_{ECTO} group throughout the immunization with P values < 0.0064, suggesting that MLD can shield GP from antibody recognition. In contrast, little difference was found between the two GPΔmuc groups, with WL²P² showing a slightly higher average EC₅₀ titer at w2 and w5 that was reversed at later time points. Compared to the GPΔmuc-WL²P² trimer group, all NP groups showed lower EC₅₀ titers except for the E2p-L4P NP group at w2, which yielded a P value of 0.0381.

[00112] This finding was somewhat unexpected, as significant differences were found between E2 core and NP groups at both w2 and w5 in a recent HCV vaccine study (He et al., Sci. Adv. 6: eaaz6225, 2020), suggesting that antibody titers induced by antigen-presenting NPs may be greatly influenced by antigen size, structure, and epitope distribution. NP display may occlude antibody access to the base and stalk epitopes, which are the targets of many bNAbs. This result may also be attributed to other factors such as dosage. As the NP carrier accounts for 21-33% of the total mass of an NP vaccine and the same dose (50 μg protein) has been used for all vaccine groups except for the I3-01v9-L7P NP group, mice in the NP groups would receive significantly less GP antigen than mice in the trimer group.

[00113] We then validated the Ebolavirus pseudoparticle (Ebolavirus-pp) neutralization assay using a panel of 10 antibodies against EBOV-Makona and a BDBV strain in 293 T cells. As expected, early EBOV NAb KZ52, c2G4, and c4G7 neutralized EBOV but not BDBV. NAb mAb100 and mAb114 neutralized both Ebolavirus species with different potencies. Four bNAbs, three directed to IFL and one targeting HR2-MPER, cross-neutralized EBOV and BDBV. A non-NAb, c13C6, which binds the glycan cap and is part of a therapeutic antibody cocktail, appeared to enhance

viral infection, suggesting a potential ADE effect. The Ebolavirus-pp assay was also performed in TZM-bl cells (Fig. 5D), which were used to screen small-molecule inhibitors. While all the tested NAb/bNAbs blocked pseudovirus infection, c13C6 exhibited a similar pattern of ADE. When tested against pseudoparticles bearing the murine leukemia virus (MLV) Env, MLV-pps, c13C6 showed enhancement of infection while NAb and bNAbs remained non-reactive (Fig. 5E).

[00114] We next assessed the immunized mouse samples against EBOV and BDBV in the Ebolavirus-pp assay. Immunoglobulin G (IgG) was purified from the mouse serum at the last time point (w11) to eliminate non-specific, anti-viral activities.

Distinct patterns of antibody response were observed, suggesting the elicitation of NAb and c13C6-like ADE-causing non-NAbs. Among the three trimer groups, while GP_{ECTO} showed a moderate NAb response with ADE observed for 2-4 mice, a substantial increase in both NAb and c13C6-like responses was observed for GP Δ muc, suggesting that the removal of MLD can equally expose NAb epitopes and the glycan cap, a main target for ADE-causing antibodies in natural infection. The WL²P² mutation appeared to have largely reversed the adversary effect caused by the removal of MLD. Among the three NP groups, E2p-L4P was the best performer and showed primarily NAb response with little hint of ADE. However, a c13C6-like antibody response was observed for 2-3 mice in the FR group and for all mice in the I3-01v9-L7P group. Since c13C6 but not any of the NAb/bNAbs reacted with MLV-pps (Fig. 5E), we sought to use the MLV-pp assay to “gauge” c13C6-like antibody responses induced by different vaccines (Fig. 5F). Indeed, we observed enhanced MLV-pp infection in excellent agreement with the ADE effect observed in the Ebolavirus-pp assay. The MLV-pp assay also indicated that the E2p-L4P NP induced a minimum c13C6-like response comparable to GP_{ECTO} with the MLD shielding. The high level of ADE effect observed for the GP Δ muc and I3-01v9-L7P groups appeared to be associated with the presence of open trimers and unassembled GP-NP species, respectively.

[00115] The mouse data offered critical insights into the effect of various GP forms and NP carriers on vaccine-induced antibody response. In brief, a multilayered E2p NP with 20 closed GP Δ muc trimers on the surface provides a promising vaccine candidate. The benefit of NP display may not be fully reflected by binding antibody titers and should be judged by the type of antibodies elicited. The c13C6-like antibodies that

target the glycan cap and cross-react with small secreted GP (ssGP) may cause adverse effects in vaccination.

Example 8 Immunogenicity of EBOV GP trimers and NPs in rabbits

- 5 [00116] Two GPΔmuc trimers, wildtype and WL²P², and three NPs presenting the WL²P² trimer were also assessed in rabbits with four injections over three-week intervals. Rabbit sera collected at six timepoints during immunization were analyzed by ELISA using the same trimer probe (Fig. 6A-B). Of note, rabbits were immunized with 20μg of mAb100/SEC-purified I3-01v9-L7P NP to reduce the c13C6-like responses.
- 10 Between the GPΔmuc groups, the WL²P² group showed higher EC₅₀ titers for all six time points except for w11, with a significant P value of 0.0229 obtained for w5 (two weeks after the second injection). Among the three NP groups, a consistent pattern was observed, with the I3-01v9 group showing the highest EC₅₀ tier and the FR group the lowest throughout the immunization. A significant difference was found between the
- 15 I3-01v9-L7P group and the E2p-L4P group for w8, w11, and w13 (four weeks after the last injection), with a P value of 0.0021-0.0053.
- [00117] Compared to the GPΔmuc-WL²P² trimer group, the I3-01v9-L7P NP group showed higher EC₅₀ tiers for all six time points, with significant P values obtained for w8, w11, and w13. In contrast, the FR and E2p-L4P groups yielded lower EC₅₀ titers
- 20 than their respective trimer group at w2 and w5, but this pattern was reversed at w8 and w11 with significant P values (0.0421 and 0.0492 for FR and E2p, respectively) at w8. However, the advantage of these two NP groups diminished at w11, showing lower EC₅₀ tiers than the trimer group at the last time point, w13. Purified IgGs from time points Pre, w2, w5, w8, and w11 were analyzed in Ebolavirus-pp and MLV-pp assays
- 25 (Fig. 6C-F). At the last time point, all vaccine groups showed NAb responses to both Ebolavirus species with no sign of ADE, suggesting a different pattern compared to the mixed NAb/c13C6-like responses in mice. Of note, the I3-01v9-L7P NP group yielded higher average IC₅₀ (50% inhibitory concentration) titers, at 211.3 μg/ml and 11.72 μg/ml for EBOV and BDBV, respectively, than other groups, confirming that the
- 30 unassembled GP-NP species – not the NP itself – were the cause of ADE in mouse immunization.

[00118] Consistently, all vaccine groups at w11 showed no enhancement of MLV-pp infection, in contrast to the mouse data (Fig. 5F). Therefore, c13C6-like antibodies appeared to be absent in serum toward the end of rabbit immunization. However, rabbit IgGs obtained from the earlier time points demonstrated an increasing NAb response accompanied by a declining c13C6-like antibody response (Fig. 6C-F). In Ebolavirus-pp neutralization assays against EBOV-Makona, ADE was first observed for the two trimer groups, the FR group, and the two multilayered NP groups at w2, w5, and w8, but then disappeared at w5, w8, and w11, respectively. Our analysis thus revealed a unique pattern for vaccine-induced c13C6-like antibodies in rabbits, which might change epitope specificity through the mechanism of “gene conversion” upon repeated antigen stimulation.

Example 9 B cell response profiles associated with EBOV GP trimers and NPs

[00119] Previously, we combined antigen-specific B cell sorting and NGS to obtain a quantitative readout of B cell responses induced by HCV E2 core (E2mc3) and E2mc3-E2p NP (He et al., Sci. Adv. 6: eaaz6225, 2020). Diverse heavy-chain variable gene (V_H) usage, a higher degree of V_H mutations, and a broader range of heavy chain complementarity determining region 3 (HCDR3) length were observed for E2p, alluding to a common B cell mechanism for effective NP vaccines. Can this mechanism be generalized to EBOV NP vaccines? Here, we applied the same strategy to obtain GP-specific B cell response profiles for different EBOV vaccine platforms. We first created an Avi-tagged GP Δ muc-WL²P²-1TD0 trimer probe to sort GP-specific mouse splenic B cells (Fig. 7A). Sorted B cells from 25 mice (Fig. 7B), five per group, were subjected to NGS on Ion GeneStudio S5. The NGS data were analyzed by a mouse Antibodyomics pipeline (Fig. 7C). Detailed B cell profiles were derived to facilitate the comparison of different vaccine platforms (Fig. 7D-F).

[00120] We mainly focused on the GP Δ muc-WL²P²-foldon group and the multilayered E2p group to compare B cell responses induced by GP Δ muc in its soluble form versus NP-displayed form. In terms of germline gene usage, similar patterns were observed for V_H and V_K genes. Namely, the stabilized GP Δ muc trimer activated more V_H/V_L genes (9.4/9.4) than its NP form (6/7), with significant P values of 0.0163 and 0.0076 for V_H and V_L , respectively. In contrast, the E2p NP decorated with 60 HCV E2 cores activated more V_H – but not V_L – genes than the E2 core. In terms of the degree of

somatic hypermutation (SHM), no significant difference was found between the two groups. Nonetheless, the NP group showed a visible shift in the SHM distribution, with higher germline V_H/V_K divergence, on average 6.4%/2.9%, than the trimer group, on average 5.3%/2.6%. In the HCDR3 analysis, two metrics – the average loop length and the r.m.s. fluctuation (r.m.s.f) of loop length – were calculated for comparison. Unlike in the HCV vaccine study where HCDR3 r.m.s.f. yielded a P value of <0.0001 between the E2 core and NP groups, no significant difference was found between the EBOV trimer and NP groups, although the E2p-L4P NP induced antibodies with longer HCDR3 loops. Overall, EBOV and HCV NPs demonstrated differential B cell patterns with respect to individual antigens. There were no apparent correlations between B cell profiles and vaccine-induced NAb/c13C6-like responses. Our results suggest that antigen size, structure, glycosylation, and epitope distribution, other than the multivalent NP display, may also be critical to shaping the B cell response.

Example 9 Some exemplified materials and methods

[00121] Expression and purification of EBOV GPΔmuc and GPΔmuc-presenting NPs: Wildtype and redesigned GPΔmuc constructs were transiently expressed in HEK293 F cells (Thermo Fisher) for biochemical, biophysical, and antigenic analyses. Briefly, 293 F cells were thawed and incubated with FreeStyle™ 293 Expression Medium (Life Technologies, CA) in a shaker incubator at 37°C, 135 rpm and 8% CO₂. When the cells reached a density of 2.0×10^6 /ml, expression medium was added to reduce cell density to 1.0×10^6 ml⁻¹ for transfection with polyethyleneimine (PEI) (Polysciences, Inc). Next, 900 μg of plasmid in 25 ml of Opti-MEM transfection medium (Life Technologies, CA) was mixed with 5 ml of PEI-MAX (1.0 mg/ml) in 25 ml of Opti-MEM. After 30-min incubation, the DNA-PEI-MAX complex was added to 1L 293 F cells. Culture supernatants were harvested five days after transfection, clarified by centrifugation at 2200 rpm for 22 min, and filtered using a 0.45 μm filter (Thermo Scientific). GPΔmuc proteins were extracted from the supernatants using an mAb114 antibody column or an mAb100 antibody column. Bound proteins were eluted three times, each with 5ml of 0.2 M Glycine (pH=2.2) and neutralized with 0.5ml of Tris-Base (pH=9.0). The proteins were further purified by size exclusion chromatography (SEC) on a Superdex 200 Increase 10/300 GL column or a HiLoad Superdex 200 16/600 column (GE Healthcare). GPΔmuc-presenting nanoparticles were

produced in ExpiCHO cells (Thermo Fisher). Briefly, ExpiCHO cells were thawed and incubated with ExpiCHO™ Expression Medium (Thermo Fisher) in a shaker incubator at 37 °C, 135 rpm and 8% CO₂. When the cells reached a density of 10×10⁶ ml⁻¹, ExpiCHO™ Expression Medium was added to reduce cell density to 6×10⁶ ml⁻¹ for transfection. The ExpiFectamine™ CHO/plasmid DNA complexes were prepared for 100-ml transfection in ExpiCHO cells following the manufacturer's instructions. For these nanoparticle constructs, 100 µg of plasmid and 320 µl of ExpiFectamine™ CHO reagent were mixed in 7.7 ml of cold OptiPRO™ medium (Thermo Fisher). After the first feed on day one, ExpiCHO cells were cultured in a shaker incubator at 33 °C, 115 rpm and 8% CO₂ following the Max Titer protocol with an additional feed on day five (Thermo Fisher). Culture supernatants were harvested 13 to 14 days after transfection, clarified by centrifugation at 4000 rpm for 25 min, and filtered using a 0.45 µm filter (Thermo Fisher). The mAb100 antibody column was used to extract nanoparticles from the supernatants, which was followed by SEC on a Superose 6 10/300 GL column. For GPDmuc and GPDmuc-presenting nanoparticles, protein concentration was determined using UV₂₈₀ absorbance with theoretical extinction coefficients.

[00122] *Blue Native Polyacrylamide Gel Electrophoresis (BN-PAGE)*: EBOV GPDmuc and GPDmuc-presenting nanoparticles were analyzed by *blue native* polyacrylamide gel electrophoresis (*BN-PAGE*) and stained with Coomassie blue. The proteins were mixed with sample buffer and G250 loading dye and added to a 4-12% Bis-Tris NativePAGE™ gel (Life Technologies). BN-PAGE gels were run for 2 to 2.5 hours at 150 V using the NativePAGE™ running buffer (Life Technologies) according to the manufacturer's instructions.

[00123] *Enzyme-Linked Immunosorbent Assay (ELISA)*: Each well of a Costar™ 96-well assay plate (Corning) was first coated with 50 µl PBS containing 0.2 µg of the appropriate antigens. The plates were incubated overnight at 4 °C, and then washed five times with wash buffer containing PBS and 0.05% (v/v) Tween 20. Each well was then coated with 150 µl of a blocking buffer consisting of PBS, 40 mg ml⁻¹ blotting-grade blocker (Bio-Rad), and 5% (v/v) FBS. The plates were incubated with the blocking buffer for 1 hour at room temperature, and then washed five times with wash buffer. For antigen binding, antibodies were diluted in the blocking buffer to a maximum concentration of 10 µg ml⁻¹ followed by a 10-fold dilution series. For each antibody

dilution, a total of 50 μ l volume was added to the appropriate wells. For animal sample analysis, serum or plasma was diluted by 10-fold for mouse and 50-fold for rabbit in the blocking buffer and subjected to a 10-fold dilution series. For each sample dilution, a total of 50 μ l volume was added to the wells. Each plate was incubated for 1 hour at room temperature, and then washed 5 times with PBS containing 0.05% Tween 20. For antibody binding, a 1:5000 dilution of goat anti-human IgG antibody (Jackson ImmunoResearch Laboratories, Inc), or for animal sample analysis, a 1:2000 dilution of horseradish peroxidase (HRP)-labeled goat anti-mouse or anti-rabbit IgG antibody (Jackson ImmunoResearch Laboratories), was then made in the wash buffer (PBS containing 0.05% Tween 20), with 50 μ l of this diluted secondary antibody added to each well. The plates were incubated with the secondary antibody for 1 hour at room temperature, and then washed 5 times with PBS containing 0.05% Tween 20. Finally, the wells were developed with 50 μ l of TMB (Life Sciences) for 3-5 min before stopping the reaction with 50 μ l of 2 N sulfuric acid. The resulting plate readouts were measured at a wavelength of 450 nm. Of note, the week 2 serum binding did not reach the plateau (or saturation) to allow for accurate determination of EC₅₀ titers. Nonetheless, the EC₅₀ values calculated in Prism were used as a quantitative measure of antibody titers to facilitate the comparison of different vaccine groups at week 2.

[00124] Bio-Layer interferometry (BLI): The kinetics of GP Δ muc and GP Δ muc-presenting nanoparticle binding to a panel of 10 antibodies was measured using an Octet Red96 instrument (fortéBio, Pall Life Sciences). All assays were performed with agitation set to 1000 rpm in fortéBio 1 $\times$ kinetic buffer. The final volume for all the solutions was 200 μ l per well. Assays were performed at 30 °C in solid black 96-well plates (Geiger Bio-One). 5 μ g ml⁻¹ of antibody in 1 $\times$ kinetic buffer was loaded onto the surface of anti-human Fc Capture Biosensors (AHC) for GP Δ muc and of anti-human Fc Quantitation Biosensors (AHQ) for nanoparticles for 300 s. A 60 s biosensor baseline step was applied prior to the analysis of the association of the antibody on the biosensor to the antigen in solution for 200 s. A two-fold concentration gradient of antigen, starting at 400 nM for GP Δ muc trimers, 25 nM for FR NP, and 10 for E2p/I3-01v9 NPs was used in a titration series of six. The dissociation of the interaction was followed for 300 s. Correction of baseline drift was performed by subtracting the mean value of shifts recorded for a sensor loaded with antibody but not incubated with

- antigen and for a sensor without antibody but incubated with antigen. Octet data were processed by FortéBio's data acquisition software v.8.1. Experimental data were fitted with the binding equations describing a 2:1 interaction to achieve optimal fitting. Of note, GPΔmuc trimer binding was also measured using AHQ to facilitate the comparison of antibody binding with nanoparticles.
- [00125] Differential scanning calorimetry (DSC): Thermal melting curves of WT and redesigned GPΔmuc trimers were obtained with a MicroCal VP-Capillary calorimeter (Malvern). The purified GPΔmuc produced from 293F cells were buffer exchanged into 1×PBS and concentrated to 27–50 μM before analysis by the instrument. Melting was probed at a scan rate of 90 °C·h⁻¹ from 25 °C to 110 °C. Data processing, including buffer correction, normalization, and baseline subtraction, was conducted using the standardized protocol from the Origin 7.0 software.
- [00126] Protein production, crystallization and data collection: Two Zaire EBOV GPΔmuc-foldon constructs, one with the W615L mutation and the L extension (to aa 637) and the other with an additional T577P mutation, were expressed in HEK293 S cells. The expressed GP was purified using an mAB100 antibody column followed by size-exclusion chromatography (SEC) on a HiLoad Superdex 200 16/600 column (GE Healthcare). The freshly purified samples of GP were used for crystallization. The crystallization experiments were carried out using the sitting drop vapor diffusion method on an automated CrystalMation™ robotic system (Rigaku) at both 4 °C and 20 °C at The Scripps research Institute (TSRI) (Elslinger *et al.*, *Acta Crystallogr. Sect. F Struct. Biol. Cryst. Commun.* 66, 1137-1142, 2010). EBOV GP was concentrated to ~10 mg/ml in 50 mM Tris-HCl pH 8.0. The reservoir solution contained 12% (w/v) PEG 6000 and 0.1 M Sodium citrate, pH 4.5. The diffractable crystals were obtained after two weeks at 20 °C. The crystals of EBOV GP were cryoprotected with 25% glycerol, mounted in a nylon loop and flash frozen in liquid nitrogen. The data were collected for two crystals of GPΔmuc-WL²-foldon and GPΔmuc-WL²P²-foldon at Advanced Photon Source (APS) beamline 23IDB, which were diffracted to 2.3 Å and 3.2 Å resolution, respectively. The diffraction data sets were processed with HKL-2000. Crystals belonged to rhombohedral H32 and tetragonal P321 space groups with cell dimensions of GPΔmuc-WL²-foldon $a = 114.58 \text{ Å}$, $b = 114.58 \text{ Å}$ and $c = 312.38 \text{ Å}$ and GPΔmuc-

WL²P²-foldon $a = 114.06 \text{ \AA}$, $b = 114.06 \text{ \AA}$ and $c = 136.22 \text{ \AA}$, respectively. The overall completeness of the two data sets was 95.77% and 99%, respectively.

[00127] Structure determination and refinement: The structures of EBOV GP were determined by molecular replacement (MR) using Phaser from the CCP4i suite with the coordinates of *Zaire* Ebola GP (PDB: 5JQ3) and the program MOLREP. The polypeptide chains were manually adjusted into electron density using Coot, with structure validation carried out using MolProbity. The final R_{cryst} and R_{free} values for the refined structures are 19.2% and 22.9% and 25% and 29.3 for GP Δ muc-WL²-foldon and GP Δ muc-WL²P²-foldon, respectively. The data processing and refinement parameters are listed in table S1.

[00128] Electron microscopy (EM) assessment of nanoparticle constructs: The initial EM assessment of EBOV GP Δ Muc nanoparticles was conducted at the Scripps Core Microscopy Facility. Briefly, nanoparticle samples were prepared at the concentration of 0.01 mg/ml. Carbon-coated copper grids (400 mesh) were glow-discharged and 8 μ L of each sample was adsorbed for 2 minutes. Excess sample was wicked away and grids were negatively stained with 2% uranyl formate for 2 minutes. Excess stain was wicked away and the grids were allowed to dry. Samples were analyzed at 80kV with a Talos L120C transmission electron microscope (Thermo Fisher) and images were acquired with a CETA 16M CMOS camera.

[00129] Mouse immunization and sample collection: The Institutional Animal Care and Use Committee (IACUC) guidelines were followed with animal subjects tested in the immunization study. Eight-week-old BALB/c mice were purchased from The Jackson Laboratory. Mice were housed in ventilated cages in environmentally controlled rooms at Scripps Research, in compliance with an approved IACUC protocol and AAALAC guidelines. Mice were immunized at weeks 0, 3, 6 and 9 for a total of four times. Each immunization consisted of 200 μ L of antigen/adjuvant mix containing 50 μ g of vaccine antigen (or 20 μ g for the I3-01v9 NP) and 100 μ L of adjuvant, AddaVax or Adju-Phos (InvivoGen), via the intraperitoneal (i.p.) route. Blood was collected two weeks after each immunization. All bleeds were performed through the facial vein (submandibular bleeding) using lancets (Goldenrod). While intermediate bleeds were collected without anticoagulant, terminal bleeds were collected using EDTA-coated tubes. Serum and plasma were heat inactivated at 56 °C for 30 min, spun at 1000 RPM for 10 min, and sterile filtered. The cells were washed once in PBS and

then resuspended in 1 ml of ACK Red Blood Cell lysis buffer (Lonza). After two rounds of washing with PBS, peripheral blood mononuclear cells (PBMCs) were resuspended in 2 ml of Bambanker Freezing Media (Lymphotec). In addition, spleens were also harvested and grounded against a 70- μ m cell strainer (BD Falcon) to release the splenocytes into a cell suspension. Splenocytes were centrifuged, washed in PBS, treated with 5 ml of Red Blood Cell Lysis Buffer Hybri-Max (Sigma-Aldrich), and frozen with 10% of DMSO in FBS. While serum and plasma were used in EBOV neutralization assays, 80% of the plasma from individual mice at week 11 was purified using a 0.2-ml protein G spin kit (Thermo Scientific) following the manufacturer's instructions. Purified mouse IgGs at week 11 (w11) were assessed in pseudovirus neutralization assays. Rabbit immunization and blood sampling were carried out under a subcontract at ProSci (San Diego, CA). Five groups of female New Zealand White rabbits, four rabbits per group, were immunized intramuscularly (i.m.) with 50 μ g (20 μ g for the I3-01v9 NP) of vaccine antigen formulated in 250 μ l of adjuvant, AddaVax or Adju-Phos (InvivoGen), with a total volume of 500 μ l, at weeks 0, 3, 6, and 9. Blood samples, 20 ml each time, were collected from the auricular artery at day 0 (Pre), weeks 2, 5, 8, and 11. For the last time point (week 13), more than 100 ml of blood was taken via cardiac puncture. Serum was separated from blood and heat inactivated for ELISA binding assays. Purified rabbit IgGs were assessed in pseudovirus neutralization assays.

[00130] Pseudovirus neutralization assay: Ebolavirus pseudoviral particle (Ebolavirus-pp) neutralization assays were utilized to assess the neutralizing activity of previously reported mAbs and vaccine-induced antibody responses in mice and rabbits. Ebolavirus-pps were generated by co-transfection of HEK293 T cells with the pNL4-3.lucR-E- plasmid (NIH AIDS reagent program) and the expression plasmid encoding the GP gene of an EBOV Makona strain (GenBank Accession number: KJ660346) or a BDBV Uganda strain (GenBank Accession number: KR063673) at a 4:1 ratio by lipofectamine 3000 (Thermo Fisher Scientific). After 48 to 72 h, Ebolavirus-pps were collected from the supernatant by centrifugation at 4000 rpm for 10 min, aliquoted, and stored at -80 °C before use. The mAbs at a starting concentration of 10 μ g/ml, or purified IgGs at a starting concentration 300 μ g/ml for mouse and 1000 μ g/ml for rabbit, were mixed with the supernatant containing Ebolavirus-pps and incubated for 1 h at 37°C in white solid-bottom 96-well plate (Corning). Based on recent studies on EBOV infectivity in various cell lines, HEK293 T cells or TZM-bl cells were used for

Ebolavirus-pp neutralization assays. Briefly, HEK293 T cells or TZM-bl cells at 1×10^4 were added to each well and the plate was incubated at 37°C for 48 h. After incubation, overlying media was removed, and cells were lysed. The firefly luciferase signal from infected cells was determined using Bright-Glo Luciferase Assay System (Promega) according to the manufacturer's instructions. Data were retrieved from a BioTek microplate reader with Gen 5 software, the average background luminescence from a series of uninfected wells was subtracted from each well, and neutralization curves were generated using GraphPad Prism 8.4.3, in which values from wells were compared against a well containing Ebolavirus-pp only. Lentiviral vectors pseudotyped with the murine leukemia virus (MLV) Env gene, termed MLV-pps, were produced in HEK293 T cells and included in the neutralization assays as a negative control. As non-NAb c13C6 exhibited enhanced MLV-pp infection, the MLV-pp assay was also used to detect the c13C6-like, ADE-causing antibody response in immunized animal samples.

[00131] Bulk sorting of EBOV GPΔmuc-specific mouse B cells: Spleens were harvested from immunized mice 15 days after the last immunization and cell suspension was prepared. Cells were stained as follows: dead cells were excluded by staining with Fixable Aqua Dead Cell Stain kit (Thermo Fisher L34957). Receptors FcγIII (CD16) and FcγII (CD32) were blocked by adding 20 μl of 2.4G2 mAb (BD Pharmingen N553142). Cells were then incubated with 10 μg/ml of biotinylated GPΔmuc, WT or UFOg². Briefly, GPΔmuc was generated by biotinylation of the individual Avi-tagged EBOV GPΔmuc using biotin ligase BirA according to the manufacturer's instructions (Avidity LLC). Biotin excess was removed by SEC on a HiLoad Superdex 200 16/600 column (GE Healthcare). In the SEC profile, the Avi-tagged GPΔmuc peak is centered at 14.5 ml, while a broader peak of biotin ligase can be found at 65-70 ml (WT) or 60-65 ml (UFOg²). Cells and biotinylated proteins were incubated for 5 min at 4 °C, followed by the addition of 2.5 μl of anti-mouse IgG fluorescently labeled with FITC (Jackson ImmunoResearch 115-095-071) and incubated for 15 min at 4 °C. Finally, 5 μl of premium-grade allophycocyanin (APC)-labeled streptavidin were added to the cells and incubated for 15 min at 4 °C. In each step, cells were washed with DPBS and the sorting buffer was 0.5 ml FACS buffer. FITC⁺ APC⁺ GPΔmuc-specific B cells were sorted using BD FACSAria II into Eppendorf tube with 500 μl of FACS buffer.

[00132] Next-generation sequencing (NGS) and Bioinformatics Analysis of Mouse B cells: A 5'-rapid amplification of cDNA ends (RACE) protocol has been reported for unbiased sequencing of mouse B cell repertoires. Here, this protocol was applied to bulk-sorted, E2-specific mouse splenic B cells. Briefly, 5'-RACE cDNA was obtained from bulk-sorted splenic B cells of each mouse with SMART-Seq v4 Ultra Low Input RNA Kit for Sequencing (TaKaRa). The immunoglobulin PCRs were set up with Platinum *Taq* High-Fidelity DNA Polymerase (Life Technologies) in a total volume of 50 μ l, with 5 μ l of cDNA as template, 1 μ l of 5'-RACE primer, and 1 μ l of 10 μ M reverse primer. The 5'-RACE primer contained a PGM/S5 P1 adaptor, while the reverse primer contained a PGM/S5 A adaptor. We adapted the mouse 3'-C γ 1-3/3'-C μ inner primers and 3'-mC κ outer primer as reverse primers for 5'-RACE PCR processing of heavy and light (κ) chains. A total of 25 cycles of PCR was performed and the expected PCR products (500-600 bp) were gel purified (Qiagen). NGS was performed on the Ion S5 GeneStudio system. Briefly, heavy and light (κ) chain libraries from the same mouse were quantitated using Qubit® 2.0 Fluorometer with Qubit® dsDNA HS Assay Kit, and then mixed using a ratio of 3:1 before being pooled with antibody libraries of other mice at an equal ratio for sequencing. Template preparation and (Ion 530) chip loading were performed on Ion Chef using the Ion 520/530 Ext Kit, followed by sequencing on the Ion S5 system with default settings. The mouse *Antibodyomics* pipeline was used to process the raw data and determine distributions for germline gene usage, somatic hypermutation (SHM), germline divergence, and H/KCDR3 loop length.

[00133] The invention thus has been disclosed broadly and illustrated in reference to representative embodiments described above. It is understood that various modifications can be made to the present invention without departing from the spirit and scope thereof.

[00134] It is further noted that all publications, sequence accession numbers, patents and patent applications cited herein are hereby expressly incorporated by reference in their entirety and for all purposes as if each is individually so denoted. Definitions that are contained in text incorporated by reference are excluded to the extent that they contradict definitions in this disclosure.

WHAT IS CLAIMED IS:

1. An engineered Ebolavirus glycoprotein (GP) protein, comprising (a) substitution of residue W615 in heptad repeat 2 (HR2) with a smaller hydrophobic residue, and (b) one or more proline substitutions in heptad repeat 1 C segment (HR1C); wherein the amino acid numbering is based on *Zaire* Ebolavirus GP sequence with UniProt ID Q05320 (SEQ ID NO:1).
2. The engineered Ebolavirus GP protein of claim 1, wherein the Ebolavirus is *Zaire Ebolavirus* (EBOV), *Sudan virus* (SUDV), *Tai forest virus* (TAFV), *Bundibugyo virus* (BDBV), or *Reston Ebolavirus* (RESTV).
3. The engineered Ebolavirus GP protein of claim 1, wherein residue W615 is replaced with L, A, V, I or F.
4. The engineered Ebolavirus GP protein of claim 1, wherein the proline substitutions comprise T577P or L579P.
5. The engineered Ebolavirus GP protein of claim 1, further comprising a truncation at the C-terminus of the membrane proximal external region (MPER).
6. The engineered Ebolavirus GP protein of claim 5, further comprising (a) a deletion of the mucin-like domain (MLD) deleted from the GP1 subunit, and/or (b) a deletion of MPER from the GP2 subunit.
7. An engineered Ebolavirus glycoprotein (GP) protein, comprising:
 - (a) a truncated soluble GP of an Ebolavirus that has the mucin-like domain (MLD) deleted from the GP1 subunit and the membrane proximal external region (MPER) deleted from the GP2 subunit, and
 - (b) substitution of residue W615 in heptad repeat 2 (HR2) with a smaller hydrophobic residue; wherein the amino acid numbering is based on *Zaire* Ebolavirus GP sequence corresponding to UniProt ID Q05320 (SEQ ID NO:1).
8. The engineered Ebolavirus GP protein of claim 7, further comprising one or more modifications selected from (i) an extension of HR2 at the C terminus, (ii) one or more proline substitutions in heptad repeat 1 C segment (HR1C), and (iii) one or more engineered inter-GP disulfide bonds.

9. The engineered Ebolavirus GP protein of claim 7, wherein the Ebolavirus is *Zaire Ebolavirus*; (EBOV), *Sudan virus* (SUDV), *Tai forest virus* (TAFV), *Bundibugyo virus* (BDBV), or *Reston Ebolavirus* (RESTV).

10. The engineered Ebolavirus GP protein of claim 7, wherein the truncated soluble GP comprises SEQ ID NO:2 or SEQ ID NO:3, or a conservatively modified or substantially identical variant thereof.

11. The engineered Ebolavirus GP protein of claim 7, further comprising a N-terminal leader sequence.

12. The engineered Ebolavirus GP protein of claim 11, wherein the leader sequence comprises SEQ ID NO:41 or a conservatively modified variant.

13. The engineered Ebolavirus GP protein of claim 7, wherein W615 is replaced with L, A, V, I or F.

14. The engineered Ebolavirus GP protein of claim 7, comprising the sequence SEQ ID NO:4, a conservatively modified variant or a substantially identical sequence thereof.

15. The engineered Ebolavirus GP protein of claim 7, further comprising an extension of HR2 at the C-terminus.

16. The engineered Ebolavirus GP protein of claim 15, wherein the HR2 extension comprises (a) extending HR2 C terminus with a N-terminal fragment of the adjacent membrane proximal external region (MPER) of the Ebolavirus glycoprotein or (b) replacing a HR2 C-terminal fragment with a longer leucine zipper motif.

17. The engineered Ebolavirus GP protein of claim 15, wherein the HR2 extension comprises (a) extending the HR2 C terminus from residue 632 to residue 637 in MPER (addition of residues KTLPD (SEQID NO:32)), (b) extending the HR2 C terminus from residue 632 to residues 643 in MPER (addition of residues KTLPDQGDNDN (SEQ ID NO:33)), or (c) replacing residues 617-632 with a GCN4 leucine zipper sequence shown in SEQ ID NO:34).

18. The engineered Ebolavirus GP protein of claim 15, wherein the W615 substitution comprises W615L substitution or P612G/W615F double mutation.

19. The engineered Ebolavirus GP protein of claim 15, comprising (a) W615L substitution and (b) extension of the HR2 C terminus from residue 632 to residue 637 in MPER of the Ebolavirus GP.

20. The engineered Ebolavirus GP protein of claim 18, comprising (a) P612G/W615F double mutation in HR2 and (b) replacement of residues 617-632 with a GCN4 leucine zipper sequence shown in SEQ ID NO:34.

21. The engineered Ebolavirus GP protein of claim 15, comprising an amino acid sequence as set forth in any one of SEQ ID NOs:5-8, a conservatively modified variant or a substantially identical sequence thereof.

22. The engineered Ebolavirus GP protein of claim 15, further comprising a C-terminal trimerization motif.

23. The engineered Ebolavirus GP protein of claim 22, wherein the C-terminal trimerization motif comprises SEQ ID NO:29 or SEQ ID NO:30, or a conservatively modified variant or a substantially identical sequence thereof.

24. The engineered Ebolavirus GP protein of claim 7, further comprising one or more proline substitutions in the HR1C segment.

25. The engineered Ebolavirus GP protein of claim 24, wherein the proline substitution comprises T577P or L579P.

26. The engineered Ebolavirus GP protein claim 24, wherein W615 is replaced with L, A, V, I or F.

27. The engineered Ebolavirus GP protein of claim 24, comprising an amino acid sequence as set forth in any one of SEQ ID NOs:9-16, a conservatively modified variant or a substantially identical sequence thereof.

28. The engineered Ebolavirus GP protein of claim 24, further comprising an extension of the HR2 C terminus.

29. The engineered Ebolavirus GP protein of claim 28, wherein the HR2 C terminus extension comprises extension from residue 632 to residue 637 in MPER.

30. The engineered Ebolavirus GP protein of claim 28, comprising an amino acid sequence as set forth in any one of SEQ ID NOs:17 and 20, a conservatively modified variant or a substantially identical sequence thereof.

31. The engineered Ebolavirus GP protein of claim 28, further comprising
5 a C-terminal trimerization motif.

32. The engineered Ebolavirus GP protein of claim 31, wherein the C-terminal trimerization motif comprises SEQ ID NO:29 or SEQ ID NO:30, or a conservatively modified variant or a substantially identical sequence thereof.

33. The engineered Ebolavirus GP protein of claim 32, comprising (a)
10 W615L substitution, (b) proline substitution T577P or L579P, and (c) HR2 extension from residue 632 to residue 637 in MPER.

34. The engineered Ebolavirus GP protein of claim 31, comprising an amino acid sequence as set forth in any one of SEQ ID NOs:18, 19, 21 and 22, a conservatively modified variant or a substantially identical sequence thereof.

35. The engineered Ebolavirus GP protein of claim 7, further comprising
15 an engineered disulfide bond between two neighboring GP protomers.

36. The engineered Ebolavirus GP protein of claim 35, wherein the engineered disulfide bond is engineered between residues G91/A575 (SS2), F153/Y534 (SS1), T520/A575 (SS3), G157/I532 (SS4), D522/A575 (SS5) or K56/G599 (SS6).

37. The engineered Ebolavirus GP protein of claim 35, comprising an amino acid sequence as set forth in any one of SEQ ID NOs:23-28, a conservatively modified variant or a substantially identical sequence thereof.

38. A polynucleotide encoding the engineered Ebolavirus GP protein of claim 1 or claim 7.

39. A polynucleotide of claim 38, wherein the engineered Ebolavirus GP protein further comprises at the N-terminus a leader peptide sequence.

40. A pharmaceutical composition, comprising the engineered Ebolavirus GP protein of claim 7, and a pharmaceutically acceptable carrier.

41. A vaccine composition, comprising an engineered Ebolavirus GP protein of claim 7 that is displayed on the surface of a self-assembling nanoparticle.

42. The vaccine composition of claim 41, wherein C-terminus of the engineered Ebolavirus GP protein of claim 7 is fused to the N-terminus of subunit of the self-assembling nanoparticle.

43. The vaccine composition of claim 42, wherein C-terminus of the engineered Ebolavirus GP protein is fused to the subunit of the self-assembling nanoparticle via a linker sequence GGGGS (SEQ ID NO:41) or (GGGGS)₂ (SEQ ID NO:42).

44. The vaccine composition of claim 41, wherein the engineered Ebolavirus GP protein of claim 7 comprises an amino acid sequence as set forth in any one of SEQ ID NOs:4-28, a conservatively modified variant or a substantially identical sequence thereof.

45. The vaccine composition of claim 41, wherein subunit of the self-assembling nanoparticle comprises the polypeptide as shown in SEQ ID NO:36 (E2p), SEQ ID NO:37 (I3-01v9), or SEQ ID NO:38 (ferritin), a conservatively modified variant or a substantially identical sequence thereof.

46. The vaccine composition of claim 41, further comprising a locking domain (LD) that is fused to the C terminus of the NP subunit.

47. The vaccine composition of claim 46, wherein the LD comprises SEQ ID NO:39 or 40, a conservatively modified variant or a substantially identical sequence thereof.

48. The vaccine composition of claim 46, further comprising a T-cell epitope that is fused to the C-terminus of the locking domain.

49. The vaccine composition of claim 48, wherein the T-cell epitope comprises SEQ ID NO:31, a conservatively modified variant or a substantially identical sequence thereof.

50. The vaccine composition of claim 41, comprising (1) a polypeptide sequence containing from N terminus to C terminus (a) the engineered Ebolavirus GP

protein, linker sequence G₄S (SEQ ID NO:43), nanoparticle sequence ferritin, (b) the engineered Ebolavirus GP protein, linker sequence G₄S (SEQ ID NO:43), nanoparticle sequence E2p, or (c) the engineered Ebolavirus GP protein, linker sequence (G₄S)₂ (SEQ ID NO:42), nanoparticle sequence I3-01v9; or (2) a conservatively modified variant of the polypeptide sequence; wherein the engineered Ebolavirus GP protein comprises W615L substitution, proline substitution T577P, and HR2 extension from residue 632 to residue 637 in MPER.

51. The vaccine composition of claim 50, further comprising a locking domain and/or a T-cell epitope that is fused to the C-terminus of the nanoparticle subunit sequence.

52. The vaccine composition of claim 51, comprising (1) a polypeptide sequence containing from N-terminus to C-terminus (a) the engineered Ebolavirus GP protein shown in SEQ ID NO:17, nanoparticle subunit sequence shown in SEQ ID NO:36 (E2p), locking domain shown in SEQ ID NO:39 (LD4), and T cell epitope shown in SEQ ID NO:31 (PADRE) or (b) the engineered Ebolavirus GP protein shown in SEQ ID NO:17, nanoparticle sequence shown in SEQ ID NO:37 (I3-01v9), locking domain shown in SEQ ID NO:40 (LD7), or (2) a conservatively modified variant of the polypeptide sequence.

20