(12) STANDARD PATENT (11) Application No. **AU 2016316723 B2**
(19) AUSTRALIAN PATENT OFFICE

(54) Title
Stabilized viral class I fusion proteins

(51) International Patent Classification(s)
C07K 2/00 (2006.01) ***A61K 39/145*** (2006.01)
A61K 39/12 (2006.01) ***A61K 39/21*** (2006.01)

(21) Application No: **2016316723** (22) Date of Filing: **2016.09.01**

(87) WIPO No: **WO17/037196**

(30) Priority Data

(31) Number	(32) Date	(33) Country
62/213,466	**2015.09.02**	**US**

(43) Publication Date: **2017.03.09**
(44) Accepted Journal Date: **2021.03.25**

(71) Applicant(s)
Janssen Vaccines & Prevention B.V.

(72) Inventor(s)
Langedijk, Johannes Petrus Maria

(74) Agent / Attorney
Shelston IP Pty Ltd., L 9 60 Margaret St, Sydney, NSW, 2000, AU

(56) Related Art
WO 2014174018 A1

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
9 March 2017 (09.03.2017)

WIPO | PCT

(10) International Publication Number
WO 2017/037196 A1

(51) International Patent Classification:
A61K 39/12 (2006.01) *A61K 39/21* (2006.01)
A61K 39/145 (2006.01) *C07K 2/00* (2006.01)

(21) International Application Number:
PCT/EP2016/070654

(22) International Filing Date:
1 September 2016 (01.09.2016)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
62/213,466 2 September 2015 (02.09.2015) US

(71) Applicant: **JANSSEN VACCINES & PREVENTION B.V.** [NL/NL]; Archimedesweg 4, 2333 CN Leiden (NL).

(72) Inventor: **LANGEDIJK, Johannes, Petrus, Maria**; Archimedesweg 4-6, 2333 CN Leiden (NL).

(74) Agents: **MANTEN, Annemieke** et al.; Archimedesweg 4-6, 2333 CN Leiden (NL).

(81) Designated States *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JP, KE, KG, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) Designated States *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

Declarations under Rule 4.17:

— *as to applicant's entitlement to apply for and be granted a patent (Rule 4.17(ii))*

— *as to the applicant's entitlement to claim the priority of the earlier application (Rule 4.17(iii))*

— *of inventorship (Rule 4.17(iv))*

Published:

— *with international search report (Art. 21(3))*

WO 2017/037196 A1

(54) Title: STABILIZED VIRAL CLASS I FUSION PROTEINS

(57) Abstract: The present invention provides stable pre-fusion class I fusion protein in the pre-fusion conformation, comprising one or more mutations in the hinge-loop that is present between the base helix and the RR1.

Stabilized viral class I fusion proteins

The present invention relates to the field of medicine. The invention in particular relates to recombinant pre-fusion class I fusion proteins and uses thereof, e.g. in immunogenic compositions.

Background of the invention

Viral fusion proteins are dynamic fusion machines that drive membrane fusion by irreversible protein refolding from a metastable pre-fusion conformation to a stable post-fusion conformation. The fusogenicity of the protein is important for viral infection.

The fusion proteins of enveloped viruses can be classified in different types based on the general irreversible folding mechanism they display to drive fusion of the virus with the target cell. Fusion proteins from unrelated viruses, such as the fusion protein F from Paramyxoviridae, Ebola GP, Retroviridae envelope protein, Coronaviridae spike, Herpesviridea gB, Orthomyxovirideae Hemagglutinin (HA) and Hemagglutinin Esterase (HE), and others are classified as class I fusion proteins and refold from a labile pre-fusion state to a stable post-fusion state through a similar mechanism although they do not exhibit any significant sequence homologies. Class I fusion proteins thus fuse the viral and host-cell membranes by irreversible protein refolding from the labile pre-fusion conformation to the stable post-fusion conformation. Structures have been determined for a variety of class I fusion proteins in pre-fusion conformation and post-fusion conformation providing insight into the mechanism of this complex fusion machine.

Except for Ebola GP and herpes gB, typically, the inactive mature class I fusion protein (e.g. F_0 for paramyxoviruses, HA for Orthomyxoviruses) requires cleavage during intracellular maturation, often by a furin-like protease that results in an N-terminal part and

an C-terminal part. The cleavage site is near or adjacent to a stretch of 20–25 hydrophobic amino acids (the fusion peptide), followed by a heptad repeat region in the C-terminal part. Since these are class I membrane proteins, the C-terminus contains the transmembrane domain (TM) and after cleavage the membrane bound C-terminal part exposes the N-terminal hydrophobic fusion peptide (FP) (Figure 1). In order to refold from the pre-fusion to the post-fusion conformation, there are two regions that need to refold, which are referred to as the refolding region 1 (RR1) and refolding region 2 (RR2). For all class I fusion proteins, the RR1 includes the FP and heptad repeat A (HRA). After a trigger, the HRA's of all three protomers in the trimer transform from a helix, or from an assembly of loops and secondary structures, to a long continuous trimeric helical coiled coil (Figure 1). The FP, located at the N-terminal segment of RR1, is then able to extend away from the viral membrane and inserts in the proximal membrane of the target cell. Next, the refolding region 2 (RR2), which is located C-terminal to RR2 closer to the TM and often includes the heptad repeat B (HRB), relocates to the other side of the fusion protein and binds the HRA coiled-coil trimer with the HRB domain to form the six-helix bundle (6HB) or with an extended polypeptide chain like Influenza HA. These similarities have been recognized by a nomenclature that places viral fusion proteins with these sequence and structural features into the so-called class I viral fusion protein group (Earp et al. Current topics in microbiology and immunology 185: 26-66, (2005); Jardetzky et al. Current opinion in virology 5: 24-33 (2014)).

When viral fusion proteins are used as a vaccine component the fusogenic function is not important. In fact, only the mimicry of the component is important to induce cross reactive antibodies that can bind the virus. Therefore, for development of robust efficacious vaccine components it is desirable that the meta-stable fusion proteins are maintained in their pre-fusion conformation. A stabilized fusion protein in the pre-fusion conformation can induce an efficacious immune response.

Any discussion of the prior art throughout the specification should in no way be considered as an admission that such prior art is widely known or forms part of common general knowledge in the field.

Unless the context clearly requires otherwise, throughout the description and the claims, the words "comprise", "comprising", and the like are to be construed in an inclusive sense as opposed to an exclusive or exhaustive sense; that is to say, in the sense of "including, but not limited to".

Summary of the invention

The present invention provides stable, recombinant, class I fusion proteins stabilized in the pre-fusion conformation.

In one aspect, the present disclosure provides a stable pre-fusion class I fusion protein, comprising one or more mutations in the hinge-loop that is present between the base helix and the RR1, wherein the class I fusion protein is a filovirus Fusion F protein.

In another aspect, the present disclosure provides a nucleic acid sequence encoding a protein according to the present invention.

In another aspect, the present disclosure provides a composition comprising: a protein according to the present invention; and/or a nucleic acid according to the present invention.

In certain embodiments, the pre-fusion polypeptides are soluble.

The invention also provides nucleic acid molecules encoding the pre-fusion polypeptides according to the invention and vectors comprising such nucleic acid molecules.

The invention also relates to compositions, preferably immunogenic compositions, comprising a class I pre-fusion polypeptide, a nucleic acid molecule and/or a vector, and to the use thereof for inducing an immune response against the class I fusion protein, in particular to the use thereof as a vaccine.

The invention also relates to methods for inducing an anti-virus immune response in a subject, comprising administering to the subject an effective amount of a pre-fusion polypeptide, a nucleic acid molecule encoding said polypeptide, and/or a vector comprising said nucleic acid molecule. Preferably, the induced immune response is characterized by neutralizing antibodies to the virus and/or protective immunity against said virus infection. In particular aspects, the invention relates to a method for inducing neutralizing anti-viral antibodies in a subject, comprising administering to the subject an effective amount of an immunogenic composition comprising a pre-fusion polypeptide, a nucleic acid molecule encoding said polypeptide, and/or a vector comprising said nucleic acid molecule.

Brief description of the Figures

FIG.1: Cartoon of fusion protein trimer in pre-fusion conformation (top left) and in the intermediate state (top right) and schematic representation of the conserved elements of fusion domain of class I fusion proteins (bottom). Fusion peptide (FP), refolding region 1

(RR1), refolding region 2 (RR2) and transmembrane region (TM) are indicated. Molecular structures of fusion proteins are very diverse and may contain a separate receptor binding domain (top left, checkered), however all fusion proteins contain a recognizable fusion domain (diagonal fill, black fill). Generally, the FP is followed by a heptad repeat region in the C-terminal part (black and diagonal striped fill) which both form the RR1. After a trigger, the N-terminal part of RR1 (black) which is connected to the base helix (diagonal fill) via a hinge loop is assembled on the base helix to form a long continuous helical coiled coil. The post-fusion form of the fusion protein is formed when refolding region 2 binds the refolded RR1. Since these are class I membrane proteins, the C-terminus contains the transmembrane domain (TM).

FIG. 2: Cartoons of a fragment of RR1 (black), hinge loop and base helix (diagonal striped fill) for several class I fusion proteins in the pre-fusion (top row) and post-fusion conformation (bottom row). After a trigger, the N-terminal part of RR1 (black fill) which is connected to the base helix (diagonal striped fill) via a hinge loop (top) is assembled on the base helix to form a long continuous helix (bottom). The small white circles in the hinge of RSV F and HIV gp41 represent substitutions to proline that stabilize the pre-fusion conformation. The small black circles are potential positions which are structurally homologous to the white circles and potential positions for stabilizing substitutions.

FIG. 3: Trimer percentage and yield of proline substitutions in HIV-1 env sequence based on a consensus of C clade Envs with a stabilizing disulfide at position 501-605 (ConC_SOS). A) Trimer percentage determined using ELISA with C-tag purified Env (black bars) and AlphaLISA on cell-free supernatant (grey bars). The trimer percentage of the I559P variant obtained by ELISA was used to normalize the AlphaLISA dataset. B) Trimer yield

determined using ELISA with purified Env (black bars) and AlphaLISA on cell-free supernatant (grey bars). The data was normalized to I559P variant. ND: not determined. Proline substitutions of HIV-1 env variants are described in more detail in Table 3.

FIG. 4: A) SEC-MALS profile of Galanthus nivalus lectin-purified HIV-1 ConC_SOS Env variant with I559P substitution (left) and L556P, I559P double substitution (right). Chromatograms show aggregates (at the left of the chromatogram), followed by trimers, followed by two peaks with smaller subunits. B) Percentage of total trimer population remaining after incubation of crude supernatant of HIV-1 env variants for 1 hour at 60°C, using PGT145 binding in AlphaLISA.

FIG. 5: A) NativePAGE analysis of cell free supernatant of ebola GP variants with proline substitutions in hinge-loop. Supernatants of transfected expi293 cells were analyzed on NativePAGE and stained with Coomassie. Lanes show expression and quaternary structure of variants. B) The trimer and monomer bands of the WT, T577P and L579P mutants were determined and their relative percentages calculated.

FIG.6: Analysis of melting temperature (Tm) of wt ebola GP, strain EBOV14 and a variant with the T577P substitution using differential scanning fluorometry (DSF).

FIG 7: A) Trimer percentage based on bands obtained with NativePAGE analysis of soluble GPs of Ebola strains Mayinga and Sudan Gulu and the variant with a substitution in the hinge loop at position 577 to Proline. B) relative trimer yield of the different soluble GPs based on NativePAGE.

FIG. 8: A) Multimer – specific binding of Influenza mini-HA variants with either Proline or Leucine at position 73 in the hinge-loop. B) Overview of expression levels, binding to broadly neutralizing antibodies and trimer binding (as shown in A) of the mini-HA variants. C) Multimer content of mini-HA variant with Tyr or Pro at position 63 shown as percentage of multimer content of 63P vs control with 63Y.

Detailed description of the invention

Virus-cell fusion is the means by which all enveloped viruses, including human pathogens such as for example the influenza virus, the human immunodeficiency virus (HIV) and Ebola virus, enter cells and initiate disease-causing cycles of replication. Virus-cell fusion is executed by one or more viral surface glycoproteins, including one that is generally denoted as the fusion protein. The fusion proteins of enveloped viruses can be classified in different types based on their general irreversible folding mechanism. Thus, fusion proteins from unrelated viruses, such as the fusion protein F from Paramyxoviridae, Ebola GP, Retroviridae envelope protein, Orthomyxoviridea Hemagglutinin (HA) and Hemagglutinin Esterase (HE) are classified as class I fusion proteins. Class I fusion proteins fuse the viral and host-cell membranes by irreversible protein refolding from the labile pre-fusion conformation to the stable post-fusion conformation.

Other known class I fusion proteins are for example the GP protein of Arenaviridae, the E1/E2 protein of Togaviridae, the E protein of Flaviviridae, E(TBEV), E1/E2 (HCV), the GN/GC protein of Bunyaviridae, the G protein of Rhabdoviridae, and the gB, gD, gH protein of Herpesviridae.

All class I fusion proteins typically comprise the same structural features (see Figure 1). The refolding region 1 (RR1) includes the fusion peptide (FP) and heptad repeat A (HRA). After a trigger, the HRA's and FP's of all three protomers in the trimer transform

from a helix, or form an assembly of loops and secondary structures, to a long continuous trimeric helical coiled coil (Figure 1). The FP, located at the N-terminal segment of RR1, is then able to extend away from the viral membrane and inserts in the proximal membrane of the target cell. Next, the refolding region 2 (RR2), which is located C-terminal to RR2 closer to the TM and often includes the heptad repeat B (HRB), relocates to the other side of the fusion protein and binds the HRA coiled-coil trimer with the HRB domain to form the six-helix bundle or with an extended polypeptide chain like influenza HA. These similarities have been recognized by a nomenclature that places viral fusion proteins with these sequence and structural features into the so-called class I viral fusion protein group (Earp et al. Current topics in microbiology and immunology 185: 26-66, (2005); Jardetzky et al. Current opinion in virology 5: 24-33 (2014).

Although the formation of the RR1 coiled coil and the relocation of RR2 to complete the 6 helix bundle are the fundamental similarities, the divergence in the fusion mechanism is great. The structural studies of several class I fusion proteins showed that they represent very distinct structural subfamilies. The overall folding of the fusion proteins and the trigger for unfolding is different for the family members and the sequence of events may differ for each class I fusion protein. For instance, for many paramyxoviruses it is proposed that the HRB will be exposed and released before the release and formation of the HRA domain.

Apart from the RR1 and RR2, all known class I fusion proteins share another distinctive structural feature, which is referred to herein as the base helix. The base helix is a part of the dynamic fusion protein that does not change conformation during the transformation from the pre to post-fusion conformation. The base helix is at the heart of the fusion protein and forms a trimeric helical base. After the trigger and the initiation of the refolding, the trimeric helical base is extended by assembly of the first helix of RR1. In respiratory syncytial virus fusion protein (RSV F), the base helix is helix alpha 5. After

refolding, helix α4 is the first structural element that assembles on top of the structurally conserved, preformed α5 helix base. The flexible α4-α5 loop is an important structural element because it is the first region in this dynamic protein that needs to hinge in order to achieve the assembly of α4 on top of the α5 (Figure 2).

In the research that led to the present invention, it was discovered that hydrophobic residues, and in particular a proline (Pro) substitution in this hinge loop can stabilize the loop. Stability is obtained by hydrophobic interactions but especially by the rigidity of the proline which limits the backbone conformations of the peptide chain in the hinge region. The proline substitution in this critical hinge region stabilized the pre-fusion conformation of RSV F by restricting the hinge movement and obstructing the movement and assembly of helix 4 on top of helix 5. Surprisingly, also for HIV a stabilizing proline substitution in the hinge loop has been described (Sanders et al. Journal of virology 76, 8875-8889 (2002)).

Recently the crystal structure of the pre-fusion conformation of the fusion domain gp41 of HIV has been solved (Pancera et al. Nature 514, 455-461 (2014)) and by comparing it to the known structure of the gp41 post-fusion 6HB, helix 7 can be identified as the base helix. After refolding of RR1, the hinge loop between α6 and α7 and helix α6 are mounted on top of the base helix α7 to form the long extended coiled coil. Similar to the α4- α5 hinge loop in RSV pre-fusion F, the α6- α7 hinge-loop of pre-fusion HIV gp41 is also the most disordered element of the fusion protein. The gp41 α6- α7 loop is in fact completely disordered and no electron density could be measured in the crystal structure. An important substitution that increased the stability of pre-fusion gp41 is I559P which is located at a structural homologous position as the stabilizing S215P in RSV-F. Although the gp41 I559P is not visible in the structure, the position in the hinge-loop between the base helix and the mounted helix is very similar to position 215 in the hinge-loop of pre-fusion RSV-F (Figure 2). Although the I559P was designed to destabilize the 6HB post-fusion conformation, the

proline at position 559 prevents the hinge movement of the α6- α7 loop in a similar fashion as the proline at position 215 inhibits the hinge movement of the α4- α5 hinge in pre-fusion RSV F (Krarup et al., Nature Communications 8143 (2015), doi:10.1038/ncomms9143). This analogy between two unrelated class I fusion proteins reveals the importance of the hinge as a critical step in refolding of RR1. It also shows that stabilizing the unstable hinge loop is a successful strategy to stabilize the unstable pre-fusion conformation of class I fusion proteins. Introduction of a proline in the hinge loop and preferably relatively close to the base helix thus can be used as a common solution to stabilize the pre-fusion conformation of class I fusion proteins which would make them superior vaccine components.

According to the invention, stabilized hinge loops have been designed for several class I fusion proteins, in particular the class I fusion proteins for which a pre-fusion structure has been elucidated, like influenza HA, retrovirus envelope protein and Ebola GP. For all these examples, it has been shown according to the invention that proline residues in the hinge loop C-terminal to the base helix stabilize the pre-fusion conformation. Therefore, the base helices in the class I fusion proteins were identified and based on the position of the stabilizing proline in RSV F and HIV gp41 hinge loop, the approximate position for a stabilizing proline in the other hinge loops were deduced (Figure 2 and Table 1).

The present invention thus provides recombinant pre-fusion class I fusion proteins that are stabilized in the pre-fusion conformation. The stable class I pre-fusion proteins of the invention are in the pre-fusion conformation, i.e. they comprise (display) epitopes that is specific to the pre-fusion conformation of the fusion protein. An epitope that is specific to the pre-fusion conformation protein is an epitope that is not presented in the post-fusion conformation. Without wishing to be bound by any particular theory, it is believed that the pre-fusion conformation of class I fusion proteins may contain epitopes that are the same as those on the protein expressed on natural virions, and therefore may provide advantages for

eliciting protective neutralizing antibodies. In certain embodiments, the proteins of the invention thus comprise at least one epitope that is recognized by a pre-fusion specific monoclonal antibody.

According to the invention it has been shown that class I fusion proteins can be stabilized in the hinge-loop that is present between the base helix and the RR1 by mutation of one or more specific amino acid residue(s), in particular by mutation of one or more specific hydrophobic amino acid residues into proline (Pro). Table 2 discloses the amino acid sequence of the regions comprising the hinge loop of several class I fusion proteins. The actual hinge loops correspond to the underlined sequences in Table 2.

In certain embodiments, the class I fusion protein is a retroviral envelope protein.

In certain embodiments, the retroviral envelope protein is a HIV-1 envelope protein. In certain embodiments, the hinge loop (underlined) is comprised within the amino acid sequence:

540 QARQLLSGIVQQQNNLLRAIEAQQHLLQLTVWGIKQLQARI 580 (SEQ ID NO: 1) or 539 QARQLLSGIVQQQSNLLRAIEAQQHMLQLTVWGIKQLQTRV 579 (SEQ 2).

In certain embodiments, the stable HIV-1 envelope protein according to the invention comprises a mutation in the hinge loop of the amino acid residue Leu on position 555, a mutation of amino acid residue Leu on position 556, and/or a mutation of the amino acid residue Ala at position 558.

In certain embodiments, the stable pre-fusion HIV-1 envelope protein according to the invention comprises a mutation in the hinge loop of the amino acid residue Leu on position 555 into Pro, a mutation of the amino acid residue Leu on position 556 into Pro, and/or a mutation of the amino acid residue Ala on position 558 into Pro.

In certain embodiments, the stable pre-fusion HIV-1 envelope protein according to the invention comprises an amino acid sequence selected from the group consisting of:

QARQLLSGIVQQQNNPLRAIEAQQHLLQLTVWGIKQLQARI (SEQ ID NO: 3);

QARQLLSGIVQQQNNLPRAIEAQQHLLQLTVWGIKQLQARI (SEQ ID NO: 4);

QARQLLSGIVQQQNNLLRPIEAQQHLLQLTVWGIKQLQARI (SEQ ID NO: 5);

QARQLLSGIVQQQSNPLRAIEAQQHMLQLTVWGIKQLQTRV (SEQ ID NO: 6);

QARQLLSGIVQQQSNLPRAIEAQQHMLQLTVWGIKQLQTRV (SEQ ID NO: 7);

and

QARQLLSGIVQQQSNLLRPIEAQQHMLQLTVWGIKQLQTRV (SEQ ID NO: 8).

It will be understood by the skilled person that the numbering of the amino acid residues relates to the numbering of the amino acid residues in the full-length HIV-1 envelope protein.

In certain embodiments, the present invention thus provides a stable pre-fusion HIV-1 envelope protein comprising an amino acid sequence, wherein the amino acid residue on position 555, 556 and/or 558 is proline.

In certain embodiments, the present invention provides a stable pre-fusion HIV-1 envelope protein comprising the amino acid sequence of SEQ ID NO: 9, wherein the amino acid residue on position 555, 556 and/or 558 is proline.

In certain embodiments, the stable pre-fusion HIV-1 envelope protein further comprises a proline on position 559.

In certain embodiments, the retroviral envelope protein is a HIV-2 envelope protein.

In certain embodiments, the hinge loop (underlined) is comprised within the amino acid sequence:

538 QSRTLLAGIVQQQQLLDAVKRQQELLRLTVWGTKNLQSRV 578 (SEQ ID NO: 30).

In certain embodiments, the stable HIV-2 envelope protein according to the invention comprises a mutation in the hinge loop of the amino acid residue Leu on position 553 and/or a mutation of amino acid residue Leu on position 554, a mutation of amino acid residue Ala at position 556, and/or a mutation of amino acid residue Val at position 557.

In certain embodiments, the stable HIV-2 envelope protein according to the invention comprise a mutation in the hinge loop of the amino acid residue Leu on position 553 into Pro, a mutation of amino acid residue Leu on position 554 into Pro, a mutation of amino acid residue Ala at position 556 into Pro, and/or a mutation of amino acid residue Val at position 557 into Pro.

In certain embodiments, the stable pre-fusion HIV-2 envelope protein according to the invention comprises an amino acid sequence selected from the group consisting of:

LLAGIVQQQQQPLDAVKRQQELLRLTVWG (SEQ ID NO: 31);

LLAGIVQQQQQLPDAVKRQQELLRLTVWG (SEQ ID NO: 32);

LLAGIVQQQQQLLDPVKRQQELLRLTVWG (SEQ ID NO: 33); and

LLAGIVQQQQQLLDAPKRQQELLRLTVWG (SEQ ID NO: 34).

It will be understood by the skilled person that the numbering of the amino acid residues relates to the numbering of the amino acid residues in the full-length HIV-2 envelope protein.

In certain embodiments, the present invention thus provides a stable pre-fusion HIV-2 envelope protein comprising an amino acid sequence, wherein the amino acid residue on position 553, 554, 556 and/or 557 is proline.

In certain embodiments, the present invention provides a stable pre-fusion HIV-2 envelope protein comprising the amino acid sequence of SEQ ID NO: 35, wherein the amino acid residue on position 553, 554, 556 and/or 557 is proline.

In certain embodiments, the class I fusion protein is a filovirus fusion protein (GP).

In certain embodiments, the filovirus GP protein is an Ebola virus GP protein.

In certain embodiments, the hinge loop (underlined) is comprised within the amino acid sequence 553 GLICGLRQLANETTQALQLFLRATTELRTFSILNRKAIDFLLQR 596 (SEQ ID NO: 10).

In certain embodiments, the stable Ebolavirus GP protein according to the invention comprises a mutation in the hinge loop of the amino acid residue Thr at position 577, and/or a mutation of the amino acid residue Leu at position 579. In certain embodiments, the stable Ebolavirus GP protein according to the invention comprises a mutation in the hinge loop of the amino acid residue Thr at position 577 into Pro, and/or a mutation of the amino acid residue Leu at position 579 into Pro.

In certain embodiments, the stable Ebolavirus GP protein according to the invention comprises an amino acid sequence selected from the group consisting of:

GLICGLRQLANETTQALQLFLRATPELRTFSILNRKAIDFLLQR (SEQ ID NO: 11); and

GLICGLRQLANETTQALQLFLRATTEPRTFSILNRKAIDFLLQR (SEQ ID NO: 12).

It will be understood by the skilled person that the numbering of the amino acid residues relates to the numbering of the amino acid residues in the full-length Ebola virus GP protein.

In certain embodiments, the present invention thus provides a stable pre-fusion Ebolavirus GP protein, wherein the amino acid residue on position 577 and/or 579 is proline.

In certain embodiments, the present invention provides a stable pre-fusion Ebolavirus GP protein comprising the amino acid sequence of SEQ ID NO: 13, wherein the amino acid residue on position 577 and/or 579 is proline.

In certain embodiments, the filovirus GP protein is a Marburg virus GP protein.

In certain embodiments, the hinge loop (underlined) is comprised within the amino acid sequence 554 NLVCRLRRLANQTAKSLELLLRVTTEERTFSLINRHAIDFLLAR 597 (SEQ ID NO: 14).

In certain embodiments, the Marburg virus GP protein according to the invention comprise a mutation in the hinge loop of the amino acid residue Thr at position 578, and/or a mutation of the amino acid residue Glu at position 580.

In certain embodiments, the Marburg virus GP protein according to the invention comprise a mutation in the hinge loop of the amino acid residue Thr at position 578 into Pro, and/or a mutation of the amino acid residue Glu at position 580 into Pro.

In certain embodiments, the stable Marburg virus GP protein according to the invention comprises an amino acid sequence selected from the group consisting of:

NLVCRLRRLANQTAKSLELLLRVTPEERTFSLINRHAIDFLLAR (SEQ ID NO: 15); and

NLVCRLRRLANQTAKSLELLLRVTTEPRTFSLINRHAIDFLLAR (SEQ ID NO: 16).

It will be understood by the skilled person that the numbering of the amino acid residues relates to the numbering of the amino acid residues in the full-length Marburg virus GP protein.

In certain embodiments, the present invention thus provides a stable pre-fusion Marburg virus GP protein, wherein the amino acid residue on position 578 and/or 580 is proline.

In certain embodiments, the present invention provides a stable pre-fusion Marburg GP protein comprising the amino acid sequence of SEQ ID NO: 17, wherein the amino acid residue on position 578 and/or 580 is proline.

In certain embodiments, the class I fusion protein is an influenza hemagglutinin (HA) protein of an influenza A virus. In certain embodiments, the influenza hemagglutinin (HA) protein is a HA protein of an influenza A virus of phylogenetic group 1. In certain embodiments, the influenza hemagglutinin (HA) protein is a H1 HA protein.

The influenza HA protein is typically composed of a HA1 domain (comprising about 329 amino acid residues) and a HA2 domain (comprising about 175 amino acid residues).

In certain embodiments, the hinge loop (underlined) is comprised within the amino acid sequence 38 QKSTQNAINGITNKVNSVIEKMNTQFTAVGKEFNKLERRMENLNKKVDDGFID (SEQ ID NO: 18).

In certain embodiments, the HA protein comprises a mutation in the hinge loop (B-loop) of HA2 of the amino acid residue Phe at position 63 (HA2 numbering) and/or a mutation of the amino acid residue Leu at position 73.

In certain embodiments, the HA protein comprises a mutation in the hinge loop of the amino acid residue Phe at position 63 into Pro, and/or a mutation of the amino acid residue Leu at position 73 into Pro.

In certain embodiments, the stable pre-fusion H1 HA protein according to the invention comprises an amino acid sequence selected from the group consisting of:

QKSTQNAINGITNKVNSVIEKMNTQPTAVGKEFNKLERRMENLNKKVDDGFI D (SEQ ID NO: 19); and

QKSTQNAINGITNKVNSVIEKMNTQFTAVGKEFNKPERRMENLNKKVDDGFI D (SEQ ID NO: 20).

It will be understood by the skilled person that the numbering of the amino acid residues relates to the numbering of the amino acid residues in the HA2 domain.

In certain embodiments, the present invention thus provides a stable pre-fusion HA protein, wherein the amino acid residue in the HA2 domain on position 63 and/or 73 is proline.

In certain embodiments, the present invention provides a stable pre-fusion HA protein comprising the amino acid sequence of SEQ ID NO: 21, wherein the amino acid residue in the HA2 domain on position 63 and/or 73 is proline.

In certain embodiments, the influenza hemagglutinin (HA) protein is a HA protein of an influenza A virus of phylogenetic group 2. In certain embodiments, the influenza hemagglutinin (HA) protein is a H3 HA protein.

In certain embodiments, the hinge loop (underlined) is comprised within the amino acid sequence

47 KSTQAAINQINGKLNRLIGKTNEKFHQIEKEFSEVEGRIQDLEKYVEDTKID

(SEQ ID NO: 22).

In certain embodiments, the stable influenza HA protein comprises a mutation in the hinge loop (B-loop) of HA2 of the amino acid residue Phe at position 72 and/or a mutation of the amino acid residue Val at position 82.

In certain embodiments, the stable influenza HA protein comprises a mutation in the hinge loop (B-loop) of HA2 of the amino acid residue Phe at position 72 into Pro and/or a mutation of the amino acid residue Val at position 82 into Pro.

In certain embodiments, the stable pre-fusion H3 HA protein according to the invention comprises an amino acid sequence selected from the group consisting of:

LKSTQAAINQINGKLNRLIGKTNEKPHQIEKEFSEVEGRIQDLEKYVEDTKID (SEQ ID NO: 23); and

LKSTQAAINQINGKLNRLIGKTNEKFHQIEKEFSEPEGRIQDLEKYVEDTKID (SEQ ID NO: 24).

It will be understood by the skilled person that the numbering of the amino acid residues relates to the numbering of the amino acid residues in the HA2 domain.

In certain embodiments, the present invention thus provides a stable pre-fusion HA protein, wherein the amino acid residue in the HA2 domain on position 72 and/or 82 is proline.

In certain embodiments, the present invention provides a stable pre-fusion HA protein comprising the amino acid sequence of SEQ ID NO: 25, wherein the amino acid residue on position 72 and/or 82 is proline.

In certain embodiments, the class I fusion polypeptide is an influenza hemagglutinin (HA) polypeptide of an influenza B virus.

In certain embodiments, the hinge loop is comprised within the amino acid sequence LKSTQEAINKITKNLNSLELEVKNLQRLSGAMDELHNEILELDEKVDDLRAD (SEQ ID NO: 26).

In certain embodiments, the stable HA polypeptides comprise a mutation in the hinge loop (B-loop) of HA2 of the amino acid residue Leu at position 62 and/or a mutation of the amino acid Leu at position 72.

In certain embodiments, the stable HA polypeptides comprise a mutation in the hinge loop (B-loop) of HA2 of the amino acid residue Leu at position 62 into Pro and/or a mutation of the amino acid Leu at position 72 into Pro.

In certain embodiments, the stable pre-fusion B HA protein according to the invention comprises an amino acid sequence selected from the group consisting of:

LKSTQEAINKITKNLNSLELEVKNPQRLSGAMDELHNEILELDEKVDDLRAD (SEQ ID NO: 27); and

LKSTQEAINKITKNLNSLELEVKNLQRLSGAMDEPHNEILELDEKVDDLRAD (SEQ ID NO: 28).

It will be understood by the skilled person that the numbering of the amino acid residues relates to the numbering of the amino acid residues in the full-length influenza B virus HA2 protein.

In certain embodiments, the present invention thus provides a stable pre-fusion HA protein, wherein the amino acid residue on position 62 and/or 72 of HA2 is proline.

In certain embodiments, the present invention provides a stable pre-fusion HA protein comprising the amino acid sequence of SEQ ID NO: 29, wherein the amino acid residue on position 62 and/or 72 of HA2 is proline.

According to the invention, the stable pre-fusion class I proteins thus comprise at least one stabilizing mutation in the hinge loop as compared to the wild-type class I fusion. As used throughout the present application, the amino acid positions are given in reference to the sequence of the full length class I fusion protein (or in reference to the HA2 domain for the influenza HA proteins). Sequence alignments can be done using methods well known in the art, e.g. by CLUSTALW, Bioedit or CLC Workbench.

An amino acid according to the invention can be any of the twenty naturally occurring (or 'standard' amino acids) or variants thereof, such as e.g. D-amino acids (the D-enantiomers of amino acids with a chiral center), or any variants that are not naturally found in proteins, such as e.g. norleucine. The standard amino acids can be divided into several groups based on their properties. Important factors are charge, hydrophilicity or hydrophobicity, size and functional groups. These properties are important for protein structure and protein–protein interactions. Some amino acids have special properties such as cysteine, that can form covalent disulfide bonds (or disulfide bridges) to other cysteine residues, proline that induces turns of the polypeptide backbone, and glycine that is more flexible than other amino acids. Table 1 shows the abbreviations and properties of the standard amino acids.

In certain embodiments, the stable pre-fusion class I proteins are full length.

In certain embodiments, the stable pre-fusion class I proteins are soluble proteins. for example soluble proteins based on the ectodomain or subdomains of the ectodomain.

It will be appreciated by a skilled person that the mutations can be made to the protein by routine molecular biology procedures. The mutations according to the invention preferably result in increased expression levels and/or increased stabilization of the pre-fusion class I polypeptides as compared to the class I fusion polypeptides that do not comprise these mutation(s).

The pre-fusion class I fusion protein polypeptides according to the invention are stable, i.e. do not readily change into the post-fusion conformation upon processing of the polypeptides, such as e.g. purification, freeze-thaw cycles, and/or storage etc.

In certain embodiments, the pre-fusion class I fusion protein polypeptides according to the invention have an increased stability upon storage a 4°C as compared to a class I fusion protein polypeptide without the mutation(s). In certain embodiments, the polypeptides are stable upon storage at 4°C for at least 30 days, preferably at least 60 days, preferably at least 6 months, even more preferably at least 1 year.

In certain embodiments, the class I fusion protein polypeptides according to the invention have an increased stability when subjected to heat, as compared to class I fusion protein polypeptides without said mutation(s). In certain embodiments, the pre-fusion conformation of the class I fusion protein polypeptides are heat stable for at least 30 minutes at a temperature of 55° C, preferably at 58° C, more preferably at 60° C With "heat stable" it is meant that the polypeptides still display the at least one pre-fusion specific epitope after having been subjected for at least 30 minutes to an increased temperature (i.e. a temperature of 55°C or above.

In certain embodiments, the proteins display the at least one pre-fusion specific epitope after being subjected to 1 to 6 freeze-thaw cycles in an appropriate formulation buffer.

As used throughout the present application nucleotide sequences are provided from 5' to 3' direction, and amino acid sequences from N-terminus to C-terminus, as custom in the art.

The present invention further provides nucleic acid molecules encoding the pre-fusion class I proteins according to the invention.

In preferred embodiments, the nucleic acid molecules encoding the proteins according to the invention are codon-optimized for expression in mammalian cells, preferably human cells. Methods of codon-optimization are known and have been described previously (e.g. WO 96/09378). A sequence is considered codon-optimized if at least one non-preferred codon as compared to a wild type sequence is replaced by a codon that is more preferred. Herein, a non-preferred codon is a codon that is used less frequently in an organism than another codon coding for the same amino acid, and a codon that is more preferred is a codon that is used more frequently in an organism than a non-preferred codon. The frequency of codon usage for a specific organism can be found in codon frequency tables, such as in http://www.kazusa.or.jp/codon. Preferably more than one non-preferred codon, preferably most or all non-preferred codons, are replaced by codons that are more preferred. Preferably the most frequently used codons in an organism are used in a codon-optimized sequence. Replacement by preferred codons generally leads to higher expression.

It will be understood by a skilled person that numerous different polynucleotides and nucleic acid molecules can encode the same polypeptide as a result of the degeneracy of the genetic code. It is also understood that skilled persons may, using routine techniques, make nucleotide substitutions that do not affect the polypeptide sequence encoded by the nucleic

acid molecules to reflect the codon usage of any particular host organism in which the polypeptides are to be expressed. Therefore, unless otherwise specified, a "nucleotide sequence encoding an amino acid sequence" includes all nucleotide sequences that are degenerate versions of each other and that encode the same amino acid sequence. Nucleotide sequences that encode proteins and RNA may or may not include introns.

Nucleic acid sequences can be cloned using routine molecular biology techniques, or generated de novo by DNA synthesis, which can be performed using routine procedures by service companies having business in the field of DNA synthesis and/or molecular cloning (e.g. GeneArt, GenScripts, Invitrogen, Eurofins).

The invention also provides vectors comprising a nucleic acid molecule as described above. In certain embodiments, a nucleic acid molecule according to the invention thus is part of a vector. Such vectors can easily be manipulated by methods well known to the person skilled in the art, and can for instance be designed for being capable of replication in prokaryotic and/or eukaryotic cells. In addition, many vectors can be used for transformation of eukaryotic cells and will integrate in whole or in part into the genome of such cells, resulting in stable host cells comprising the desired nucleic acid in their genome. The vector used can be any vector that is suitable for cloning DNA and that can be used for transcription of a nucleic acid of interest. Suitable vectors according to the invention are e.g. adenovectors, such as e.g. Ad26 or Ad35, alphavirus, paramyxovirus, vaccinia virus, herpes virus, retroviral vectors etc. The person skilled in the art is capable of choosing suitable expression vectors, and inserting the nucleic acid sequences of the invention in a functional manner.

Host cells comprising the nucleic acid molecules encoding the pre-fusion class I proteins form also part of the invention. The pre-fusion class I proteins may be produced through recombinant DNA technology involving expression of the molecules in host cells, e.g. Chinese hamster ovary (CHO) cells, tumor cell lines, BHK cells, human cell lines such as

HEK293 cells, PER.C6 cells, or yeast, fungi, insect cells, and the like, or transgenic animals or plants. In certain embodiments, the cells are from a multicellular organism, in certain embodiments they are of vertebrate or invertebrate origin. In certain embodiments, the cells are mammalian cells. In certain embodiments, the cells are human cells. In general, the production of a recombinant proteins, such the pre-fusion class I proteins of the invention, in a host cell comprises the introduction of a heterologous nucleic acid molecule encoding the class I proteins in expressible format into the host cell, culturing the cells under conditions conducive to expression of the nucleic acid molecule and allowing expression of the polypeptide in said cell. The nucleic acid molecule encoding a protein in expressible format may be in the form of an expression cassette, and usually requires sequences capable of bringing about expression of the nucleic acid, such as enhancer(s), promoter, polyadenylation signal, and the like. The person skilled in the art is aware that various promoters can be used to obtain expression of a gene in host cells. Promoters can be constitutive or regulated, and can be obtained from various sources, including viruses, prokaryotic, or eukaryotic sources, or artificially designed.

Cell culture media are available from various vendors, and a suitable medium can be routinely chosen for a host cell to express the protein of interest, here the pre-fusion class I proteins. The suitable medium may or may not contain serum.

A "heterologous nucleic acid molecule" (also referred to herein as 'transgene') is a nucleic acid molecule that is not naturally present in the host cell. It is introduced into for instance a vector by standard molecular biology techniques. A transgene is generally operably linked to expression control sequences. This can for instance be done by placing the nucleic acid encoding the transgene(s) under the control of a promoter. Further regulatory sequences may be added. Many promoters can be used for expression of a transgene(s), and are known to the skilled person, e.g. these may comprise viral, mammalian, synthetic

promoters, and the like. A non-limiting example of a suitable promoter for obtaining expression in eukaryotic cells is a CMV-promoter (US 5,385,839), e.g. the CMV immediate early promoter, for instance comprising nt. –735 to +95 from the CMV immediate early gene enhancer/promoter. A polyadenylation signal, for example the bovine growth hormone polyA signal (US 5,122,458), may be present behind the transgene(s). Alternatively, several widely used expression vectors are available in the art and from commercial sources, e.g. the pcDNA and pEF vector series of Invitrogen, pMSCV and pTK-Hyg from BD Sciences, pCMV-Script from Stratagene, etc, which can be used to recombinantly express the protein of interest, or to obtain suitable promoters and/or transcription terminator sequences, polyA sequences, and the like.

The cell culture can be any type of cell culture, including adherent cell culture, e.g. cells attached to the surface of a culture vessel or to microcarriers, as well as suspension culture. Most large-scale suspension cultures are operated as batch or fed-batch processes because they are the most straightforward to operate and scale up. Nowadays, continuous processes based on perfusion principles are becoming more common and are also suitable. Suitable culture media are also well known to the skilled person and can generally be obtained from commercial sources in large quantities, or custom-made according to standard protocols. Culturing can be done for instance in dishes, roller bottles or in bioreactors, using batch, fed-batch, continuous systems and the like. Suitable conditions for culturing cells are known (see e.g. Tissue Culture, Academic Press, Kruse and Paterson, editors (1973), and R.I. Freshney, Culture of animal cells: A manual of basic technique, fourth edition (Wiley-Liss Inc., 2000, ISBN 0-471-34889-9)).

The invention further provides compositions comprising a pre-fusion class I protein and/or a nucleic acid molecule, and/or a vector, as described above. The invention thus provides compositions comprising a pre-fusion class I protein that displays an epitope that is

present in a pre-fusion conformation of the class I protein but is absent in the post-fusion conformation. The invention also provides compositions comprising a nucleic acid molecule and/or a vector, encoding such pre-fusion class I protein. The invention further provides immunogenic compositions comprising a pre-fusion class I protein, and/or a nucleic acid molecule, and/or a vector, as described above. The invention also provides the use of a stabilized pre-fusion class I protein, a nucleic acid molecule, and/or a vector, according to the invention, for inducing an immune response against said class I proteins in a subject. Further provided are methods for inducing an immune response against a class I fusion protein in a subject, comprising administering to the subject a pre-fusion class I fusion protein, and/or a nucleic acid molecule, and/or a vector, according to the invention. Also provided are pre-fusion class I fusion proteins, nucleic acid molecules, and/or vectors, according to the invention for use in inducing an immune response against said class I fusion protein in a subject. Further provided is the use of the pre-fusion class I fusion protein, and/or nucleic acid molecules, and/or vectors according to the invention for the manufacture of a medicament for use in inducing an immune response against said class I fusion protein in a subject. In certain embodiments, the pre-fusion class I proteins according to the invention are for use as a vaccine.

The pre-fusion class I fusion proteins, nucleic acid molecules and/or vectors according to the invention may be used e.g. in stand-alone treatment and/or prophylaxis of a disease or condition caused by the virus comprising said class I fusion protein, or in combination with other prophylactic and/or therapeutic treatments, such as (existing or future) vaccines, antiviral agents and/or monoclonal antibodies.

The invention further provides methods for preventing and/or treating virus infection in a subject utilizing the pre-fusion class I fusion proteins, nucleic acid molecules and/or vectors according to the invention. In a specific embodiment, a method for preventing and/or treating a

virus infection in a subject comprises administering to a subject in need thereof an effective amount of a pre-fusion class I fusion protein, nucleic acid molecule and/or a vector, as described above. A therapeutically effective amount refers to an amount of a class I fusion protein, nucleic acid molecule or vector that is effective for preventing, ameliorating and/or treating a disease or condition resulting from infection by a virus comprising said class I fusion protein. Prevention encompasses inhibiting or reducing the spread of virus or inhibiting or reducing the onset, development or progression of one or more of the symptoms associated with infection by said virus. Amelioration as used in herein may refer to the reduction of visible or perceptible disease symptoms, viremia, or any other measurable manifestation of a viral infection.

For administering to subjects, such as humans, the invention may employ pharmaceutical compositions comprising a pre-fusion class I fusion protein, a nucleic acid molecule and/or a vector as described herein, and a pharmaceutically acceptable carrier or excipient. In the present context, the term "pharmaceutically acceptable" means that the carrier or excipient, at the dosages and concentrations employed, will not cause any unwanted or harmful effects in the subjects to which they are administered. Such pharmaceutically acceptable carriers and excipients are well known in the art (see Remington's Pharmaceutical Sciences, 18th edition, A. R. Gennaro, Ed., Mack Publishing Company [1990]; Pharmaceutical Formulation Development of Peptides and Proteins, S. Frokjaer and L. Hovgaard, Eds., Taylor & Francis [2000]; and Handbook of Pharmaceutical Excipients, 3rd edition, A. Kibbe, Ed., Pharmaceutical Press [2000]). The pre-fusion class I polypeptides, or nucleic acid molecules, preferably are formulated and administered as a sterile solution although it may also be possible to utilize lyophilized preparations. Sterile solutions are prepared by sterile filtration or by other methods known per se in the art. The solutions are then lyophilized or filled into

pharmaceutical dosage containers. The pH of the solution generally is in the range of pH 3.0 to 9.5, e.g. pH 5.0 to 7.5.

In certain embodiments, a composition according to the invention further comprises one or more adjuvants. Adjuvants are known in the art to further increase the immune response to an applied antigenic determinant. The terms “adjuvant” and "immune stimulant" are used interchangeably herein, and are defined as one or more substances that cause stimulation of the immune system. In this context, an adjuvant is used to enhance an immune response to the class I fusion protein of the invention. Examples of suitable adjuvants include aluminium salts such as aluminium hydroxide and/or aluminium phosphate; oil-emulsion compositions (or oil-in-water compositions), including squalene-water emulsions, such as MF59 (see e.g. WO 90/14837); saponin formulations, such as for example QS21 and Immunostimulating Complexes (ISCOMS) (see e.g. US 5,057,540; WO 90/03184, WO 96/11711, WO 2004/004762, WO 2005/002620); bacterial or microbial derivatives, examples of which are monophosphoryl lipid A (MPL), 3-O-deacylated MPL (3dMPL), CpG-motif containing oligonucleotides, ADP-ribosylating bacterial toxins or mutants thereof, such as *E. coli* heat labile enterotoxin LT, cholera toxin CT, and the like; eukaryotic proteins (e.g. antibodies or fragments thereof (e.g. directed against the antigen itself or CD1a, CD3, CD7, CD80) and ligands to receptors (e.g. CD40L, GMCSF, GCSF, etc), which stimulate immune response upon interaction with recipient cells. In certain embodiments the compositions of the invention comprise aluminium as an adjuvant, e.g. in the form of aluminium hydroxide, aluminium phosphate, aluminium potassium phosphate, or combinations thereof, in concentrations of 0.05 – 5 mg, e.g. from 0.075-1.0 mg, of aluminium content per dose.

In other embodiments, the compositions do not comprise adjuvants.

The invention is further illustrated in the following Examples.

Examples

EXAMPLE 1

In order to stabilize the labile pre-fusion conformation of HIV-1 envelope protein, amino acid residues at position 555, 556 and 558 (according to the HXB2 numbering) in the hinge loop were substituted for proline residues. The total gp140 expression, trimeric nature (trimer percentage) and total trimer yield of the envelopes were compared to a variant with the known stabilizing I559P substitution (Sanders et. al., J Virol 2002, supra) and an envelope protein without proline substitutions in the hinge loop. Trimer percentage, yield and stability were determined using ELISA, AlphaLISA and/or SEC-MALS. Part of the wild type hinge loop, the single Pro substitutions, the double Pro substitutions and a triple Pro substitution are shown in Table 3.

Generation of HIV-1 Env consensus C sequence

The HIV-1 Env sequence was based on a consensus sequence based on clade C. Therefore an alignment was downloaded (HIV Sequence Alignments. 3,434 sequences in total) from the Los Alamos database (http://www.hiv.lanl.gov/content/index). Only the C-clade ENVs (1,252 sequences) were used in the consensus maker in the Los Alamos HIV database website.

Recombinant proteins based on consensus C with the stabilizing SOS modification (501C-605C) were expressed in Expi293 cells (Life Technologies), as described below. Env proteins used in ELISA contained a C-terminal C-tag and additional I201C-A433C mutations (Kwon et al. Nat Struct Mol Biol, (2015), 22(7):522-531). Env proteins used in AlphaLISA and SEC-MALS contained a C-terminal SortA-Flag-35GS-His-tag. The cells were transiently transfected using ExpiFectamine293 (Life Technologies) according to the manufacturer's instructions and cultured in a shaking incubator at 37 °C and 8% CO2. The culture

supernatants containing HIV envelope protein (Env) were harvested on day 3 (for AlphaLISA) or day 5 (for ELISA and SEC-MALS) after transfection and sterile-filtered. C-tagged Env was purified using C-tag affinity chromatography and the SortA-Flag-35GS-His-tagged Env was purified using Galanthus nivalus lectin chromatography, followed by SEC-MALS. The recombinant HIV env proteins were purified by a 2-step purification protocol applying a Galantus nivalis-lectin column (Vectorlabs, AL-1243, lot Z0325) for the initial purification and subsequently a superdex200 Increase column (GE) for the polishing step to remove residual contaminants. For the initial lectin step the culture supernatant was diluted with 40 mM Tris, 500 mM NaCl pH7.5 and passed over a 4 ml CV Tricorn 10-50 Lectin Agarose Column at 4 ml per minute. Subsequently the column was washed with 4 column volumes (CV) of 40 mM Tris, 500 mM NaCl pH7.5 and eluted 4 CV of 40 mM Tris, 500 mM NaCl, 1M Mannopyronoside pH7.5 with an upflow of 1.6 mL/min. The eluate was concentrated using a spin concentrator (50K, Amicon Ultra, Millipore) and the protein was further purified using a superdex200 column using 50 mM Tris, 150 mM NaCl pH 7.4 as running buffer. The second peak contained the HIV gp140 trimer. The fractions containing this peak were again pooled and the protein concentration was determined using OD280 and stored a 4°C until use. The identity of the band was verified using Western blotting (not shown) SDS-PAGE analysis and Western blot. Cell culture supernatants or purified protein samples were analyzed on 4-12% (w/v) Bis-Tris NuPAGE gels, 1X MOPS (Life Technologies) under reducing or non-reducing conditions and blotted using the iBlot technology (Life Technologies). All procedures were performed according to manufacturer's instructions. For purity analysis the gels were stained with Krypton Infrared Protein Stain (Thermo Scientific) or SYPRO Rubi protein stain (Bio-Rad). The blots were probed with anti-His-HRP. The gels and the blot membranes were scanned on an Odyssey instrument (Li-Cor) and images analyzed using Odyssey 3.0 software (Li-Cor).

The total Env expression level for input in ELISA was determined by calculation of the area under the curve of the C-tag purification chromatogram. C-tagged Env was captured using His-tagged anti-C-tag VHH. The correct trimeric conformation of the purified protein was confirmed by binding to PGT145 antibody in ELISA. A 100% trimeric Env was taken along as a control to calculate the trimer percentage (Figure 3A). Trimer yield was calculated by multiplying the trimer percentage with the total Env expression level and normalized to I559P, which was set at 1 (Figure 3B). All substitutions show higher trimer percentage and yield compared to the protein without proline substitutions in the hinge-loop. All double substitutions increase trimer percentage and in particular trimer yield compared to I559P variant. Based on these finding a triple mutant L556P, A558P, I559P was constructed and compared to I559P and double mutants L556P, I559P and A558P, I559P. These variants were analyzed in cell-free supernatant using AlphaLISA.

AlphaLISA is a bead-based proximity assay in which singlet oxygen molecules, generated by high energy irradiation of Donor beads, transfer to Acceptor beads, which are within a distance of approximately 200 nm. Subsequently, a cascading series of chemical reactions results in a chemiluminescent signal (Eglen et al. Curr Chem Genomics (2008), 1: 2-10). For the AlphaLISA® assay the constructs were equipped with a Flag-His tag (with a 35GS linker in between). The HIV constructs were expressed in Expi293 cells, which were cultured for 3 days in 96 well plates (200 µl/well). Crude supernatants were diluted 120 times in AlphaLISA buffer (PBS + 0.05% Tween-20 + 0.5 mg/mL BSA). Subsequently 10 µl of these dilutions were transferred to a half-area 96-well plate and mixed with 40µl acceptor beads, mixed with donor beads and PGT145. The sequence of PGT145 was derived from PDB file 3U1S and was expressed like the Envs (no furin added) and purified using MAb select SuRe affinity chromatography. The beads were mixed well before use. After 2 hours of incubation at RT, non-shaking, the signal was measured with Neo (BioTek) The donor beads

were conjugated to ProtA (Cat#: AS102M, Lot#1831829, Perkin Elmer), which could bind to the mAb. The acceptor beads were conjugated to an anti-His antibody (Cat#: AL112R, Lot# 2036860, Perkin Elmer) to detect the His-tag of the protein. For the quantification of the protein yield, a combination of Nickel-conjugated donor beads (Cat#: AS101M, Lot#: 2027498, Perkin Elmer) together with acceptor beads carrying anti-Flag antibody (Cat#: AL112R, Lot#: 2036860, Perkin Elmer) were used. The average mock signal was subtracted from the AlphaLISA counts measured for the different Env proteins. As a reference the ConC_SOSIP backbone was used, the whole data set was divided by ConC_SOSIP signal, to normalize the signals to the backbone. Trimer percentages were obtained by dividing these normalized signals to the normalized signal obtained for the quantification. The backbone trimer percentage was set to 30% according to the results of SEC-MALS.

The correct trimeric conformation was confirmed by binding to PGT145 antibody binding and expression levels were quantified by detecting the SortA-Flag-35GS-His-tag. Trimer percentage was calculated by dividing the PGT145 signal by the quantification signal (Figure 3A). The trimer yield was derived from AlphaLISA by normalizing the PGT145 signal to that of I559P, which was set at 1 (Figure 3B) The triple mutant has the highest trimer percentage and yield compared to all other variants. Trimer percentage and yield of Galanthus nivalus lectin-purified I559P and L556P, I559P were confirmed using size exclusion chromatography (SEC) and multi-angle light scattering (MALS) analysis using a high-performance liquid chromatography system (Agilent Technologies) and miniDAWN TREOS (Wyatt) instrument coupled to a Optilab T-rEX Refractive Index Detector (Wyatt). In total, 40 µg of protein was applied to a TSK-Gel G3000SWxl column (Tosoh Bioscience) equilibrated in running buffer (150 mM Sodium Phosphate buffer, 50 mM NaCl, pH 7.0) at 1 mL/min. The data were analyzed by the Astra 6 software package and molecular weight calculations were derived from the refractive index signal. The trimer content from SEC-

MALS chromatograms were determined by calculating the area under the curve, yielding comparable results as ELISA and AlphaLISA (Figure 4A). Trimer stability was determined using PGT145 binding in AlphaLISA of crude supernatant that was incubated for 1 hour at 60°C. For the thermostability assay 20 µl crude supernatant was heated at 60°C for 1 hour, in a PCR block. The plates were centrifuged 5 min. at maximum speed, to remove aggregates. The non-treated and heat treated supernatants were diluted 40x. Then the AlphaLISA assay was performed as described above, using PGT145 as trimer specific antibody and Nickel/Flag for the quantification.

Percentage of total trimer population remaining after incubation was calculated by dividing the PGT145 signal before incubation and the signal after incubation. All double and triple mutants show higher trimer stability than the I559P variant (Figure 4B). These results show that hinge stabilization is successful for the class I fusion protein of the retrovirus HIV-1.

EXAMPLE 2

In order to stabilize the pre-fusion conformation of a filovirus fusion protein GP, residues in the hinge loop at position 575, 576, 577, 579, 581 and 583 were substituted to Proline and the expression level, multimeric nature of the Ebola GP and the stability was accessed using NativePAGE and differential scanning fluorometry (DSF), respectively. Figure 5 shows the NativePAGE analysis of supernatants from ebola GP variants with proline substitutions at position 575, 576, 577, 579, 581 and 583. As shown in Figure 5 A, only variants with substitutions at positions 577 and 579 showed high expression and had higher trimer content compared to the wt sequence. The trimer and monomer bands of the WT, T577P and L579P mutants were determined (Fig 5B) and their relative percentages calculated (Fig 5C). NativePAGE Bis-Tris gels system (Life technologies) was used to analyse the

supernatants from transiently transfected cells. Subsequently the gels were stained by Coomassie. The Native PAGE gel was scanned using Biorad ChemiDoc MP. The quantification was done using Biorads Image Lab software, using the "Lane and band" analysis tool. Therefore the protein lanes are highlighted and the protein bands of interest are selected. The software calculates the intensity of the highlighted bands and their relative amounts were calculated.

Next, HIS-tag purified GP variants were tested for temperature stability by DSF. Protein was mixed with SYPRO orange fluorescent dye (Life Technologies S6650) in a 96-well optical qPCR plate. The optimal dye and protein concentration was determined experimentally. All protein dilutions were performed in PBS, and a negative control sample containing the dye only was used as a reference subtraction. The measurement was performed in a qPCR instrument (Applied Biosystems ViiA 7) using the following parameters: a temperature ramp from 25–85 °C with a rate of 0.015 °C per second. Data was collected continuously. The negative first derivative of the Sypro Orange signal measured at several intervals during a temperature ramp up to 85 degrees. Raw data from the Sypro Orange assay was analyzed using an R script which parses ViiA7 output (machine used for the Sypro assay) and merges the sample information with the melt curve data. Spotfire software was used to subtract the controls from the data (buffer with Sypro Orange). Then it determines top and bottom of the melt curves and normalizes the curves for values between 0 and 100. Duplicates are averaged from duplicate runs and Tm determined at half maximal value. Melting temperatures (point where the signal reaches 50% of the maximum value) could be determined after fitting the data using Spotfire for the wt protein (61.5°C) and the T577P variant (64°C). The fitted data are shown in Figure 6. The variant with the T577P substitution showed a higher temperature stability than the wt protein.

To study the universality of the stabilizing effect of T577P, the 577P substitution was also evaluated in the GP of ebola strain Mayinga (NCBI Reference Sequence: NP_066246.1), ebola strain Sudan Gulu (NCBI Reference Sequence: YP_138523.1) As shown in Figure 7, the proline substitution increased expression levels and trimer content of filovirus GP. These results show that hinge stabilization is also successful for the class I fusion protein of ebola virus.

EXAMPLE 3

The fusion domain of Influenza is sequestered in the Hemagglutinin protein that has a head domain (mostly HA1) that is responsible for binding to sialic acid (hemagglutination functionality) and a stem domain (mostly HA2) that contains the fusion domain. In order to study the stabilization of the Influenza Hemagglutinin fusion domain, point mutations were made in a so-called mini-HA (#4650) that corresponds to a semi-stable stem region and contains the HA2 ectodomain and a fragment of HA1 (Impagliazzo et. al., Science 24 August, 2015). In order to stabilize the semi-stable pre-fusion conformation of Influenza mini-Hemagglutinin 4650, HA2 Ser73 was substituted to Leu or Pro to remove the glycan at position 71 and study the effect of the hydrophobic rigidifying Proline residue versus the hydrophobic Leucine residue in the hinge loop (B-loop). Because the mini-Hemagglutinin did not contain the head domain, a variant was made in which Lys68 was substituted to Gln68 because the Lysine is originally involved in a conserved salt-bridge between the B-loop and the head domain. Because the head was deleted, the salt-bridge was gone and Lysine could be changed to the neutral Glutamine. Recombinant proteins were expressed in 293 Freestyle cells (as described below). Stability of the mini-hemagglutinin variants was accessed by an ELISA in which the multimer content of the protein is measured. Supernatants of transfected cell were tested in multimer ELISA. Plates were coated with the stem-specific broadly

neutralizing MAb CR9114 (Dreyfus et. al., Science (2012), 337(6100):1343-8). Supernatants were titrated and incubated with the coated plates. After washing, the captured multimers were incubated with biotinylated CR9114. Next the wells were incubated with HRP-conjugated streptavidin followed by addition of HRP substrate (Impagliazzo et. al., Science 2015). Figure 8 shows that variants with P73 in the hinge loop (B-loop) showed higher multimer expression than variants with L73. Additionally, a substitution of Tyr to Pro at position 63 of mini-HA #4650 also resulted in an almost 2-fold higher multimer expression (Fig 8C). These results show that hinge stabilization is also successful for the class I fusion protein of the orthomyxovirus Influenza HA.

EXAMPLE 4

Expression of protein constructs

The constructs were synthesized and codon-optimized at GenScript (Piscataway, NJ 08854) or at Gene Art (Life Technologies, Carlsbad, CA). The constructs were cloned into pCDNA2004 or generated by standard methods involving site-directed mutagenesis and PCR, and sequenced. HEK-Expi293 cells or HEK293F cells were transiently transfected with pCDNA2004 plasmid with the protein insert (and in case of HIV Env 90% env and 10% Furin-pCDNA2004), according to the manufacturer's instructions and cultured for 5 days at 37°C and 10% CO2. The culture supernatant was harvested and spun for 5 minutes at 300 g to remove cells and cellular debris. The spun supernatant was subsequently sterile filtered using a 0.22 μm vacuum filter and stored at 4°C until use.

Table 1. Standard amino acids, abbreviations and properties

Amino Acid	3-Letter	1-Letter	Side chain polarity	Side chain charge (pH 7.4)
alanine	Ala	A	non-polar	Neutral
arginine	Arg	R	polar	Positive
asparagine	Asn	N	polar	Neutral
aspartic acid	Asp	D	polar	Negative
cysteine	Cys	C	non-polar	Neutral
glutamic acid	Glu	E	polar	Negative
glutamine	Gln	Q	polar	Neutral
glycine	Gly	G	non-polar	Neutral
histidine	His	H	polar	positive(10%) neutral(90%)
isoleucine	Ile	I	non-polar	Neutral
leucine	Leu	L	non-polar	Neutral
lysine	Lys	K	polar	Positive
methionine	Met	M	non-polar	Neutral
phenylalanine	Phe	F	non-polar	Neutral
proline	Pro	P	non-polar	Neutral
serine	Ser	S	polar	Neutral
threonine	Thr	T	polar	Neutral
tryptophan	Trp	W	non-polar	Neutral
tyrosine	Tyr	Y	polar	Neutral
valine	Val	V	non-polar	Neutral

Table 2

HIV-1, HXB2 540 QARQLLSGIVQQQNNLLRAIEAQQHLLQLTVWGIKQLQARI 580 (SEQ ID NO:1)
HIV-1, ConC 539 QARQLLSGIVQQQSNLLRAIEAQQHMLQLTVWGIKQLQTRV 579 (SEQ ID NO:2)
HIV-2, BAH97710.1 538 QSRTLLAGIVQQQQQLLDAVKRQQELLRLTVWGTKNLQSRV 578 (SEQ ID NO:30)

Filo: alpha 1 - hingeloop - alpha2
Ebola, mayinga 553 GLICGLRQLANETTQALQLFLRATTELRTFSILNRKAIDFLLQR 596 (SEQ ID NO:10)
Marburg 554 NLVCRLRRLANQTAKSLELLLRVTTEERTFSLINRHAIDFLLAR 597 (SEQ ID NO:14)

Influenza A, H1 38 QKSTQNAINGITNKVNSVIEKMNTQFTAVGKEFNKLERRMENLNKKVDDGFID (SEQ ID NO:18)
Influenza A, H3 47 LKSTQAAINQINGKLNRLIGKTNEKFHQIEKEFSEVEGRIQDLEKYVEDTKID (SEQ ID NO:22)
Influenza B, B/Yamanashi/166/1998
38 LKSTQEAINKITKNLNS-LELEVKNLQRLSGAMDELHNEILELDEKVDDLRAD (SEQ ID NO:26)

Table 3. Hinge loop of HIV-1 ConcC, residues 547- 571 (with Env numbering following standard HXB2 convention) and variants with Proline substitutions

Code	substitution	sequence
HIV150400:	ConC_SOS	LLSGIVQQQSNLLRAIEAQQHMLQLTVWG
HIV150399:	$I_{559}P$	LLSGIVQQQSNLLRA**P**EAQQHMLQLTVWG
HIV150401:	$A_{558}P$	LLSGIVQQQSNLLR**P**IEAQQHMLQLTVWG
HIV150402:	$L_{556}P$	LL**S**GIVQQQSNL**P**RAIEAQQHMLQLTVWG
HIV150403:	$L_{555}P$	LLSGIVQQQSN**P**LRAIEAQQHMLQLTVWG
HIV150404:	$_{558}P_{559}P$	LLSGIVQQQSNLLR**PP**EAQQHMLQLTVWG
HIV150405:	$_{556}P_{559}P$	LLSGIVQQQSNL**P**RA**P**EAQQHMLQLTVWG
HIV160114:	$_{556}P_{558}P_{559}P$	LLSGIVQQQSNL**P**R**PP**EAQQHMLQLTVWG

Sequences

HIV-1 HXB2 gp160 with signal peptide (SEQ ID NO: 9).

MRVKEKYQHLWRWGWRWGTMLLGMLMICSATEKLWVTVYYGVPVWKEATTTLFCASDAKAYDTEVHNV
WATHACVPTDPNPQEVVLVNVTENFNMWKNDMVEQMHEDIISLWDQSLKPCVKLTPLCVSLKCTDLKN
DTNTNSSSGRMIMEKGEIKNCSFNISTSIRGKVQKEYAFFYKLDIIPIDNDTTSYKLTSCNTSVITQA
CPKVSFEPIPIHYCAPAGFAILKCNNKTFNGTGPCTNVSTVQCTHGIRPVVSTQLLLNGSLAEEEVVI
RSVNFTDNAKTIIVQLNTSVEINCTRPNNNTRKRIRIQRGPGRAFVTIGKIGNMRQAHCNISRAKWNN
TLKQIASKLREQFGNNKTIIFKQSSGGDPEIVTHSFNCGGEFFYCNSTQLFNSTWFNSTWSTEGSNNT
EGSDTITLPCRIKQIINMWQKVGKAMYAPPISGQIRCSSNITGLLLTRDGGNSNNESEIFRPGGGDMR
DNWRSELYKYKVVKIEPLGVAPTKAKRRVVQREKRAVGIGALFLGFLGAAGSTMGAASMTLTVQARQL
LSGIVQQQNNLLRAIEAQQHLLQLTVWGIKQLQARILAVERYLKDQQLLGIWGCSGKLICTTAVPWNA
SWSNKSLEQIWNHTTWMEWDREINNYTSLIHSLIEESQNQQEKNEQELLELDKWASLWNWFNITNWLW
YIKLFIMIVGGLVGLRIVFAVLSIVNRVRQGYSPLSFQTHLPTPRGPDRPEGIEEEGGERDRDRSIRL
VNGSLALIWDDLRSLCLFSYHRLRDLLLIVTRIVELLGRRGWEALKYWWNLLQYWSQELKNSAVSLLN
ATAIAVAEGTDRVIEVVQGACRAIRHIPRRIRQGLERILL

HIV-2 Envelope protein (SEQ ID NO: 35).

MAHINSHLLISLLLISVYGCMCKQYVTVFYGIPAWRNATVPLFCATANRDTWGTVQCLPD
SGDYTEISVNITEAFDAWNNTVTEQAVDDVWNLFETSLKPCVKLTPLCVAMNCSTNNTRT
NNTTASTTTGNSTTPIVVNEAIPCVKANNCSGIGLEDVVNCTFNMTGLRQDERKQYNDTW
YRRDLECEGTRCYMRTCNTSVIQESCDKHYWDSLRFRYCAPPGYAILRCNDTNYSGFMHN
CSKVVVSSCTRMMETQTSTWFGFNGTRAENRTYMYWHGRDNRTIISLNRYYNLTMHCRRP
GNKTVLPITIMSGRRFHSRPVINERPRQAWCWFEGNWTEAMREVKETVMKHPRYTGIKNI
TKINLVGPSAGSDPEARYMWTNCRGEFFYCNMTWFLNWVEGKNGTKRNYVPCHIRQIVNT
WHKVGKYVYLPPREGLLSCNSTVTSIIANIEWIDSNETNITMSAEVGELYRLELGDYKLV
EITPIGFAPTNIKRYSSATPRNRRGVMVLGFLGFLATAGSAMGAASLTLSAQSRTLLAGI
VQQQQQLLDAVKRQQELLRLTVWGTKNLQSRVTAIEKYLKDQALLNSWGCAFRQVCHTTV
PWPNESLTPNWTDMTWQQWEEKVHYLDANITQLLEEAQIQQEKNMYELQKLNHWDVFSNW
FDFTSWMAYIRLGLYVVVGLIVLRIVIYVIQMLARLRKGYRPVFSSPPSYTQQIPIHKHR
GQPANEETEDEGGREGDYRSWPWQIEYAHFLIRQLRNLLIWLYNGCRNLLLRTSQILQPA
LQPLRLSLAYLQYGISWLQEAIQAATRAAGETLANAGRALWEALRRTAGAIIAVPRRIRQ
GLELALL

EBOLA GP (SEQ ID NO: 13) **(STRAIN MAYINGA-76)**

MGVTGILQLPRDRFKRTSFFLWVIILFQRTFSIPLGVIHNSTLQVSDVDKLVCRDKLSSTNQLRSVGL
NLEGNGVATDVPSATKRWGFRSGVPPKVVNYEAGEWAENCYNLEIKKPDGSECLPAAPDGIRGFPRCR
YVHKVSGTGPCAGDFAFHKEGAFFLYDRLASTVIYRGTTFAEGVVAFLILPQAKKDFFSSHPLREPVN

ATEDPSSGYYSTTIRYQATGFGTNETEYLFEVDNLTYVQLESRFTPQFLLQLNETIYTSGKRSNTTGK
LIWKVNPEIDTTIGEWAFWETKKNLTRKIRSEELSFTVVSNGAKNISGQSPARTSSDPGTNTTTEDHK
IMASENSSAMVQVHSQGREAAVSHLTTLATISTSPQSLTTKPGPDNSTHNTPVYKLDISEATQVEQHH
RRTDNDSTASDTPSATTAAGPPKAENTNTSKSTDFLDPATTTSPQNHSETAGNNNTHHQDTGEESASS
GKLGLITNTIAGVAGLITGGRRTRREAIVNAQPKCNPNLHYWTTQDEGAAIGLAWIPYFGPAAEGIYI
EGLMHNQDGLICGLRQLANETTQALQLFLRATTELRTFSILNRKAIDFLLQRWGGTCHILGPDCCIEP
HDWTKNITDKIDQIIHDFVDKTLPDQGDNDNWWTGWRQWIPAGIGVTGVIIAVIALFCICKFVF

MARBURG GP (SEQ ID NO: 17)

MKTTCLLISLILIQGVKTLPILEIASNIQPQNVDSVCSGTLQKTEDVHLMGFTLSGQKVADSPLEASK
RWAFRAGVPPKNVEYTEGEEAKTCYNISVTDPSGKSLLLDPPTNIRDYPKCKTIHHIQGQNPHAQGIA
LHLWGAFFLYDRIASTTMYRGKVFTEGNIAAMIVNKTVHKMIFSRQGQGYRHMNLTSTNKYWTSSNGT
QTNDTGCFGTLQEYNSTKNQTCAPSKKPLPLPTAHPEVKLTSTSTDATKLNTTDPNSDDEDLTTSGSG
SGEQEPYTTSDAATKQGLSSTMPPTPSPQPSTPQQGGNNTNHSQGVVTEPGKTNTTAQPSMPPHNTTT
ISTNNTSKHNLSTPSVPIQNATNYNTQSTAPENEQTSAPSKTTLLPTENPTTAKSTNSTKSPTTTVPN
TTNKYSTSPSPTPNSTAQHLVYFRRKRNILWREGDMFPFLDGLINAPIDFDPVPNTKTIFDESSSSGA
SAEEDQHASPNISLTLSYFPKVNENTAHSGENENDCDAELRIWSVQEDDLAAGLSWIPFFGPGIEGLY
TAGLIKNQNNLVCRLRRLANQTAKSLELLLRVTTEERTFSLINRHAIDFLLARWGGTCKVLGPDCCIG
IEDLSRNISEQIDQIKKDEQKEGTGWGLGGKWWTSDWGVLTNLGILLLLSIAVLIALSCICRIFTKYI
G

Hemagglutinin Influenza A virus (group 1)(A/Brisbane/59/2007(H1N1) (SEQ ID NO: 21).

MKVKLLVLLCTFTATYADTICIGYHANNSTDTVDTVLEKNVTVTHSVNLLENSHNGKLCL
LKGIAPLQLGNCSVAGWILGNPECELLISKESWSYIVEKPNPENGTCYPGHFADYEELRE
QLSSVSSFERFEIFPKESSWPNHTVTGVSASCSHNGESSFYRNLLWLTGKNGLYPNLSKS
YANNKEKEVLVLWGVHHPPNIGNQKALYHTENAYVSVVSSHYSRKFTPEIAKRPKVRDQE
GRINYYWTLLEPGDTIIFEANGNLIAPRYAFALSRGFGSGIINSNAPMDKCDAKCQTPQG
AINSSLPFQNVHPVTIGECPKYVRSAKLRMVTGLRNIPSIQSRGLFGAIAGFIEGGWTGM
VDGWYGYHHQNEQGSGYAADQKSTQNAINGITNKVNSVIEKMNTQFTAVGKEFNKLERRM
ENLNKKVDDGFIDIWTYNAELLVLLENERTLDFHDSNVKNLYEKVKSQLKNNAKEIGNGC
FEFYHKCNDECMESVKNGTYDYPKYSEESKLNREKIDGVKLESMGVYQILAIYSTVASSL
VLLVSLGAISFWMCSNGSLQCRICI

Hemagglutinin Influenza A virus (group 2)K7N5L2_9INFA (SEQ ID NO: 25).

QKLPGNDNST ATLCLGHHAV PNGTIVKTIT NDQIEVTNAT ELVQSSSTGG
ICDSPHQILD GENCTLIDAL LGDPQCDGFQ NKKWDLFVER SKAYSNCYPY
DVPDYASLRS LVASSGTLEF NNESFNWTGV TQNGTSSACK RRSNNSFFSR
LNWLTHLKFK YPALNVTMPN NEKFDKLYIW GVHHPGTDND QISLYAQASG
RITVSTKRSQ QTVIPNIGSR PRVRDIPSRI SIYWTIVKPG DILLINSTGN
LIAPRGYFKI RSGKSSIMRS DAPIGKCNSE CITPNGSIPN DKPFQNVNRI
TYGACPRYVK QNTLKLATGM RNVPEKQTQG IFGAIAGFIE NGWEGMVDGW
YGFRHQNSEG IGQAADLKST QAAINQINGK LNRLIGKTNE KFHQIEKEFS
EVEGRIQDLE KYVEDTKIDL WSYNAELLVA LENQHTIDLT DSEMNKLFER
TKKQLRENAE DMGNGCFKIY HKCDNACIGS IRNGTYDHDV YRDEALNNRF
QIK

Hemagglutinin Influenza B virus (B/Yamanashi/166/1998) (SEQ ID NO: 29)

MKAIIVLLMVVTSNADRICTGITSSNSPHVVKTATQGEVNVTGVIPLTTTPTKSHFANLK
GTKTRGKLCPTCLNCTDLDVALGRPMCVGVTPSAKASILHEVRPVTSGCFPIMHDRTKIR
QLPNLLRGYEKIRLSTQNVINAEKAPGGPYRLGTSGSCPNATSRSGFFATMAWAVPKDNN
KTATNPLTVEVPHICTKEEDQITVWGFHSDDKTQMKNLYGDSNPQKFTSSANGVTTHYVS
QIGGFPDQTEDGGLPQSGRIVVDYMVQKPGKTGTIVYQRGILLPQKVWCASGRSKVIKGS
LPLIGEADCLHEKYGGLNKSKPYYTGEHAKAIGNCPIWVKTPLKLANGTKYRPPAKLLKE
RGFFGAIAGFLEGGWEGMIAGWHGYTSHGAHGVAVAADLKSTQEAINKITKNLNSLSELE
VKNLQRLSGAMDELHNEILELDEKVDDLRADTISSQIELAVLLSNEGIINSEDEHLLALE
RKLKKMLGPSAVDIGNGCFETKHKCNQTCLDRIAAGTFNAGEFSLPTFDSLNITAASLND
DGLDNHTILLYYSTAASSLAVTLMIAIFIVYMISRDNVSCSICL

2016316723 01 Mar 2021

Claims

1. A stable pre-fusion class I fusion protein, comprising one or more mutations in the hinge-loop that is present between the base helix and the RR1, wherein the class I fusion protein is a filovirus Fusion F protein.

2. The protein according to claim 1, wherein the filovirus F protein is an Ebola virus F protein.

3. The protein according to claim 2, comprising a mutation in the hinge loop of the amino acid residue Thr at position 577 and/or a mutation of the amino acid residue Leu at position 579.

4. The protein according to any one of claims 1-3, wherein the class I fusion protein is an Ebola virus F protein comprising an amino acid sequence selected from the group consisting of:
GLICGLRQLANETTQALQLFLRATPELRTFSILNRKAIDFLLQR (SEQ ID NO: 11); and
GLICGLRQLANETTQALQLFLRATTEPRTFSILNRKAIDFLLQR (SEQ ID NO: 12).

5. The protein according to claim 1, wherein the filovirus F protein is a Marburg virus F protein.

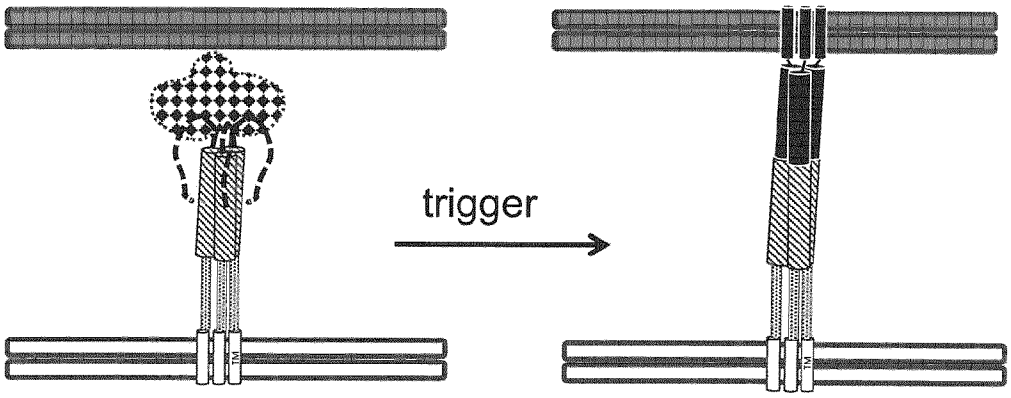
6. The protein according to claim 5, comprising a mutation in the hinge loop of the amino acid residue Thr at position 578, and/or a mutation of the amino acid residue Glu at position 580.

7. The protein according to claim 5 or 6, wherein the class I fusion protein is a Marburg virus F protein comprising an amino acid sequence selected from the group consisting of:
NLVCRLRRLANQTAKSLELLLRVTPEERTFSLINRHAIDFLLAR (SEQ ID NO: 15); and

NLVCRLRRLANQTAKSLELLLRVTTEPRTFSLINRHAIDFLLAR (SEQ ID NO: 16).

8. A nucleic acid sequence encoding a protein according to any one of the preceding claims.

9. A composition comprising: a protein according to any one of claims 1 to 7; and/or a nucleic acid according to claim 8.



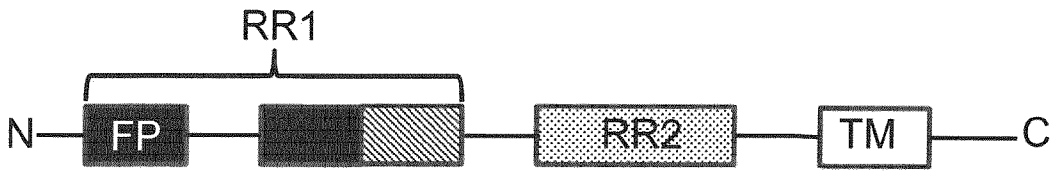


Fig. 1

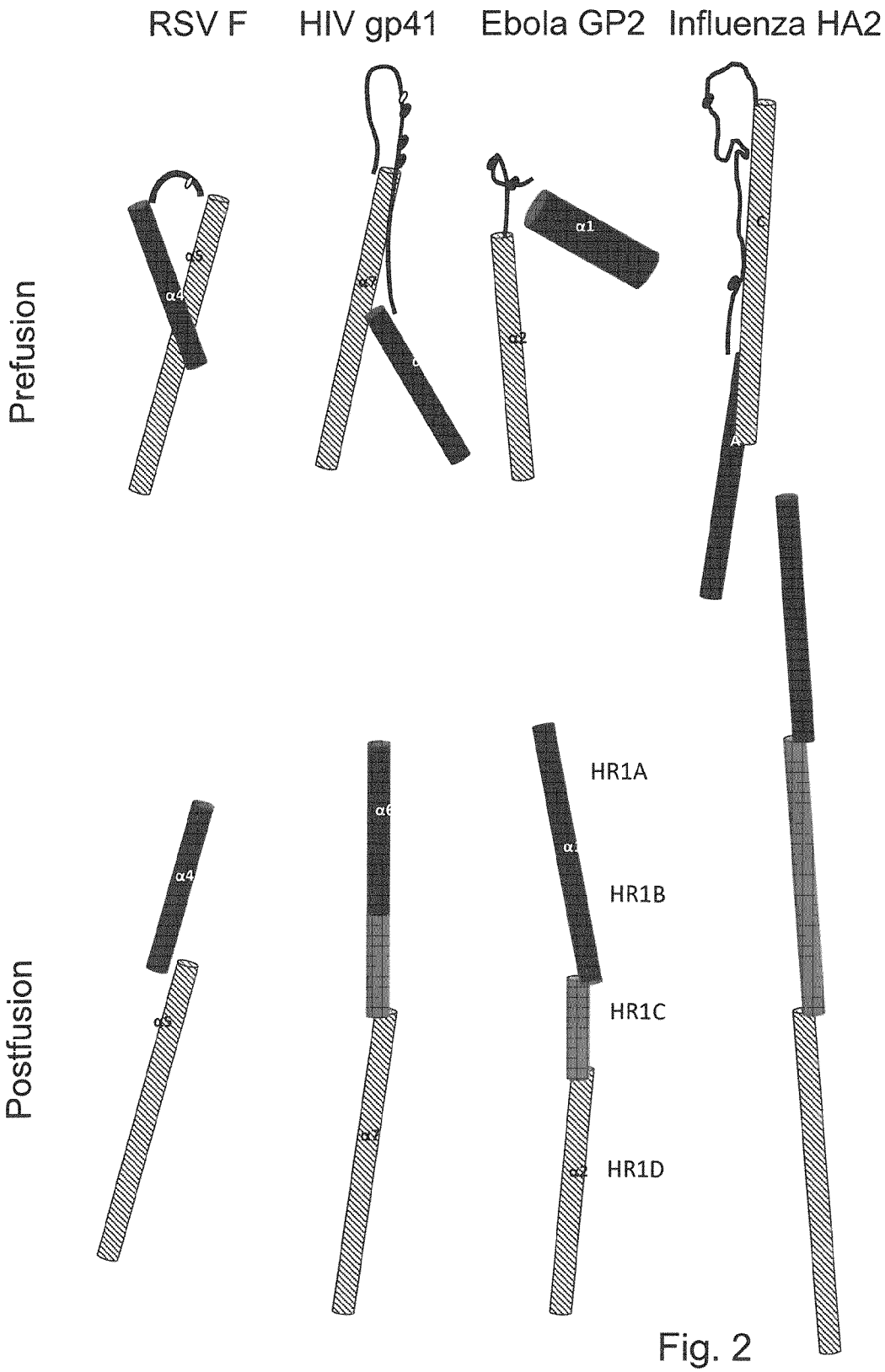

RSV F
HIV gp41
Ebola GP2
Influenza HA2
Prefusion
Postfusion
α5
α4
α3
α1
α2
C
N
HR1A
HR1B
HR1C
HR1D

Fig. 2

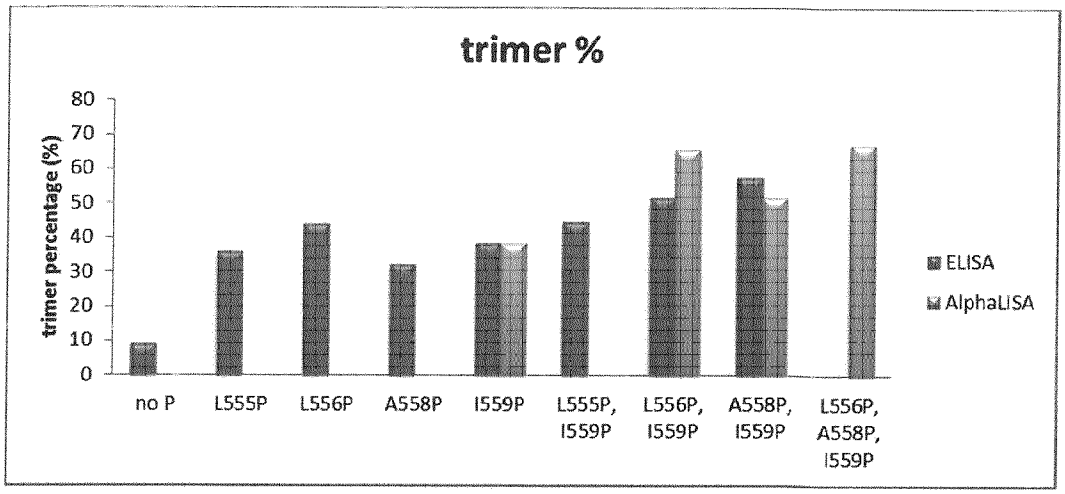

trimer %
trimer percentage (%)
80
70
60
50
40
30
20
10
0
no P
L555P
L556P
A558P
I559P
L555P, I559P
L556P, I559P
A558P, I559P
L556P, A558P, I559P
ELISA
AlphaLISA

Fig. 3A

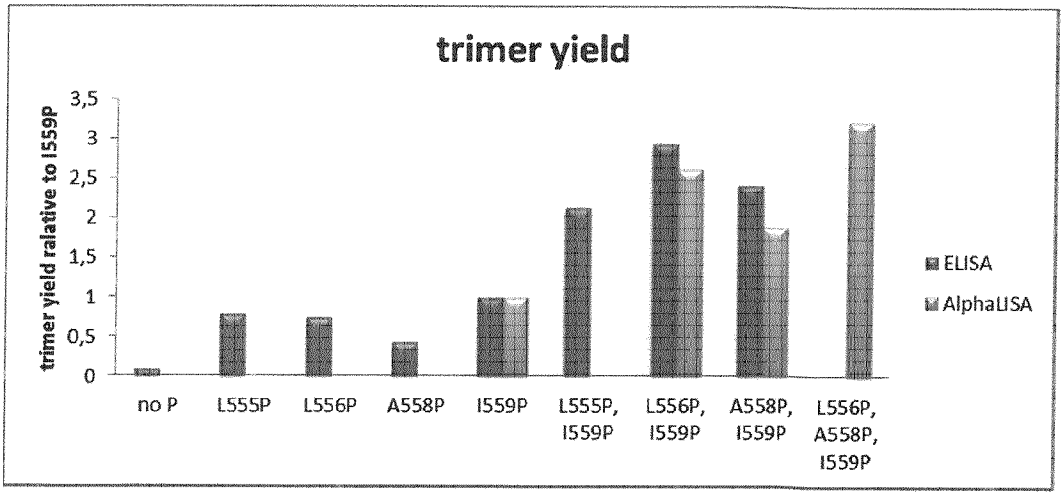

trimer yield
trimer yield ralative to I559P
3,5
3
2,5
2
1,5
1
0,5
0
no P
L555P
L556P
A558P
I559P
L555P, I559P
L556P, I559P
A558P, I559P
L556P, A558P, I559P
ELISA
AlphaLISA

Fig. 3B

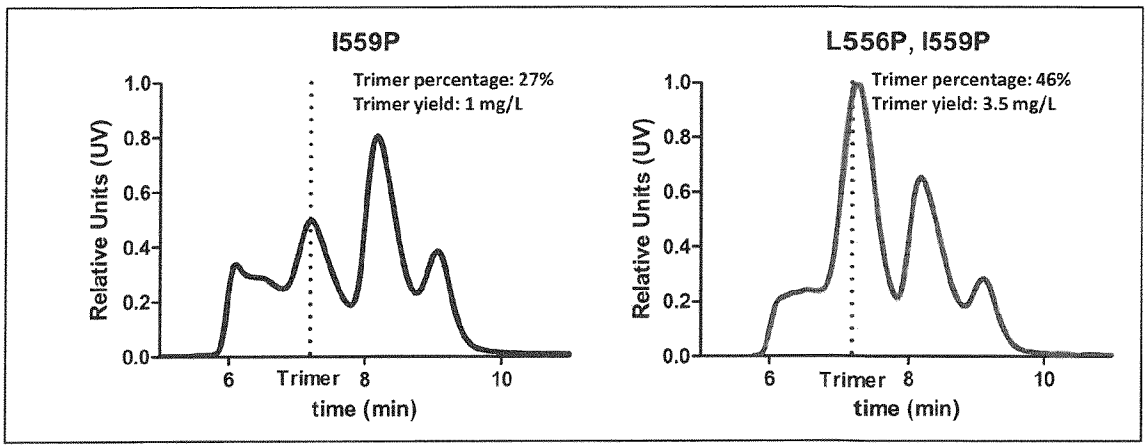

I559P
Trimer percentage: 27%
Trimer yield: 1 mg/L
Relative Units (UV)
1.0
0.8
0.6
0.4
0.2
0.0
6
Trimer
8
10
time (min)
L556P, I559P
Trimer percentage: 46%
Trimer yield: 3.5 mg/L
Relative Units (UV)
1.0
0.8
0.6
0.4
0.2
0.0
6
Trimer
8
10
time (min)

Fig. 4A

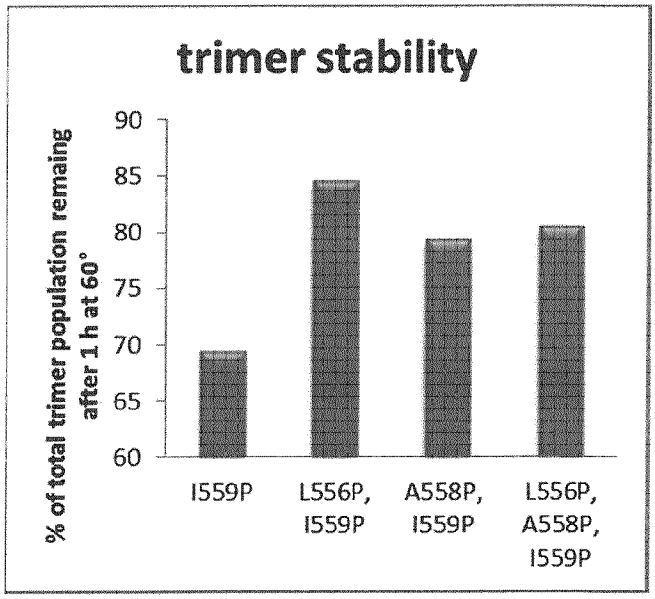

trimer stability
% of total trimer population remaing after 1 h at 60°
90
85
80
75
70
65
60
I559P
L556P, I559P
A558P, I559P
L556P, A558P, I559P

Fig. 4B

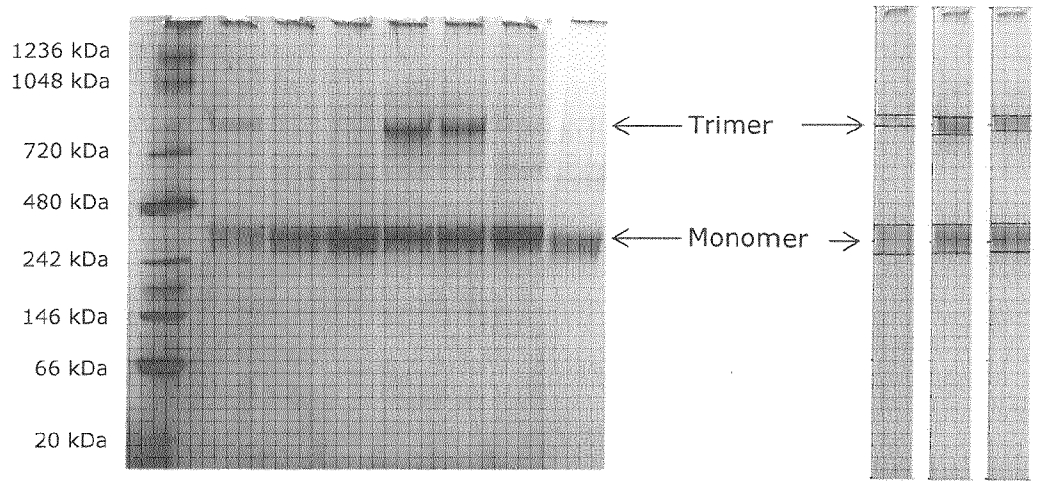
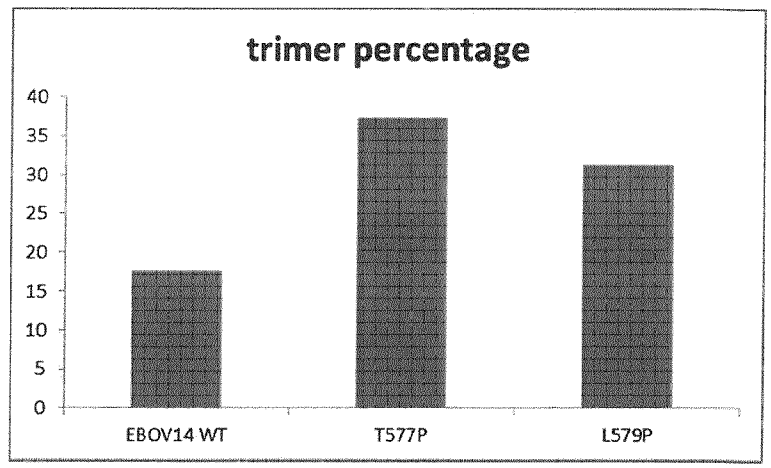

1236 kDa
1048 kDa
720 kDa
480 kDa
242 kDa
146 kDa
66 kDa
20 kDa
Trimer
Monomer

Fig. 5A


trimer percentage
40
35
30
25
20
15
10
5
0
EBOV14 WT
T577P
L579P

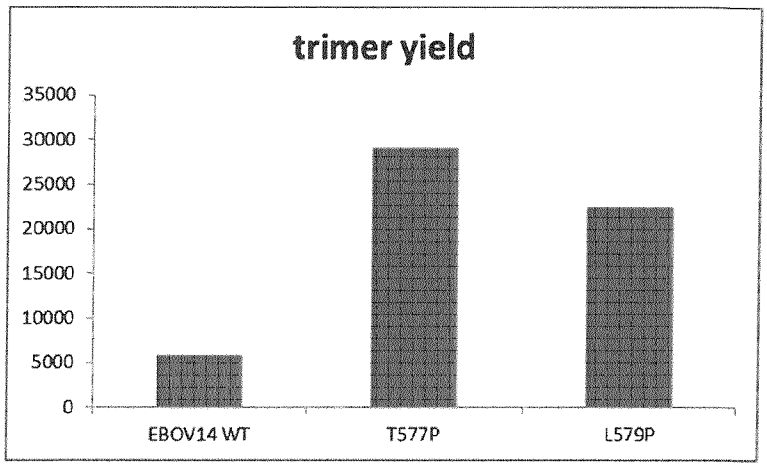

trimer yield
35000
30000
25000
20000
15000
10000
5000
0
EBOV14 WT
T577P
L579P

Fig. 5B

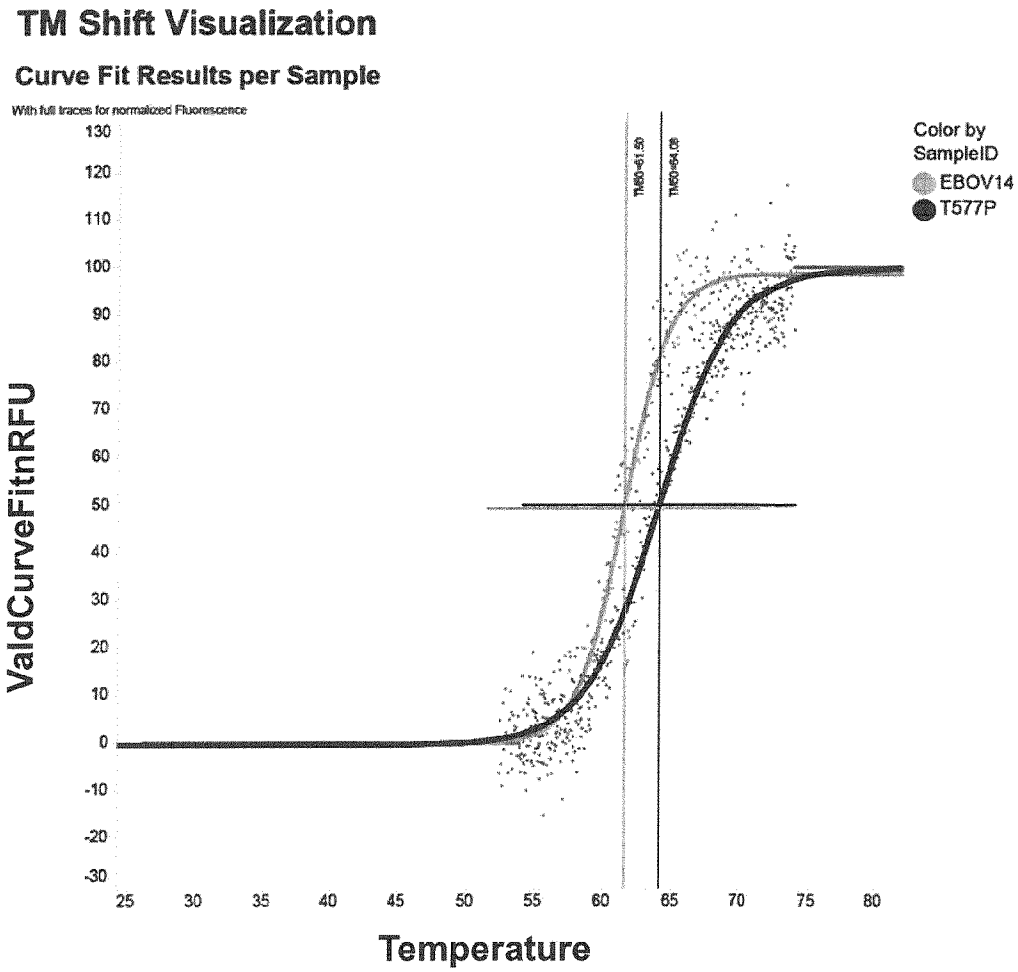
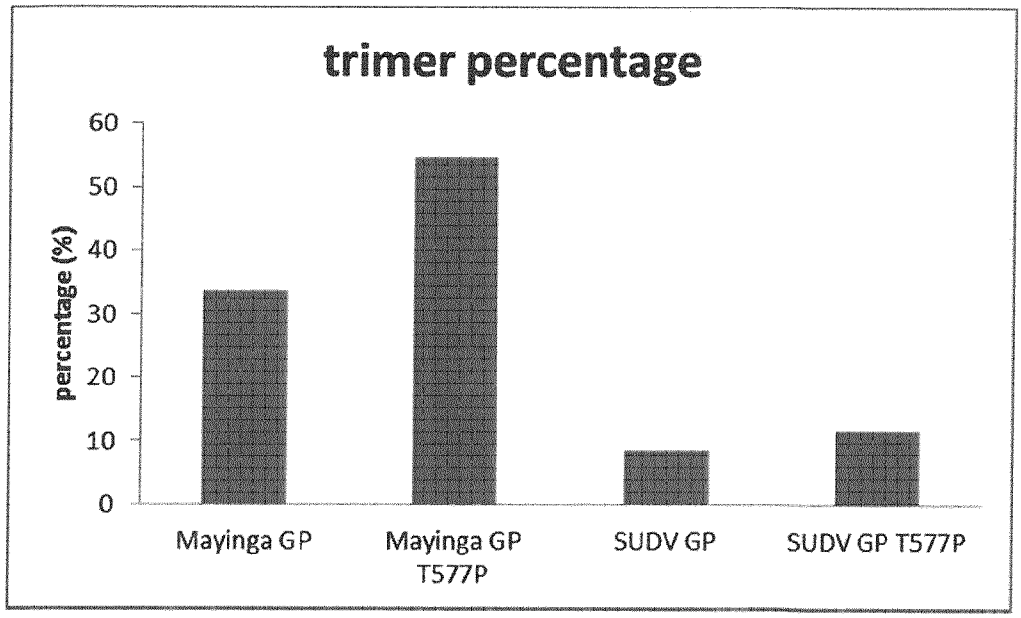

TM Shift Visualization
Curve Fit Results per Sample
With full traces for normalized Fluorescence
ValdCurveFitnRFU
Temperature
Color by SampleID
EBOV14
T577P
TM50=61.50
TM50=64.08
130
120
110
100
90
80
70
60
50
40
30
20
10
0
-10
-20
-30
25
30
35
40
45
50
55
60
65
70
75
80

Fig. 6


trimer percentage
percentage (%)
60
50
40
30
20
10
0
Mayinga GP
Mayinga GP T577P
SUDV GP
SUDV GP T577P

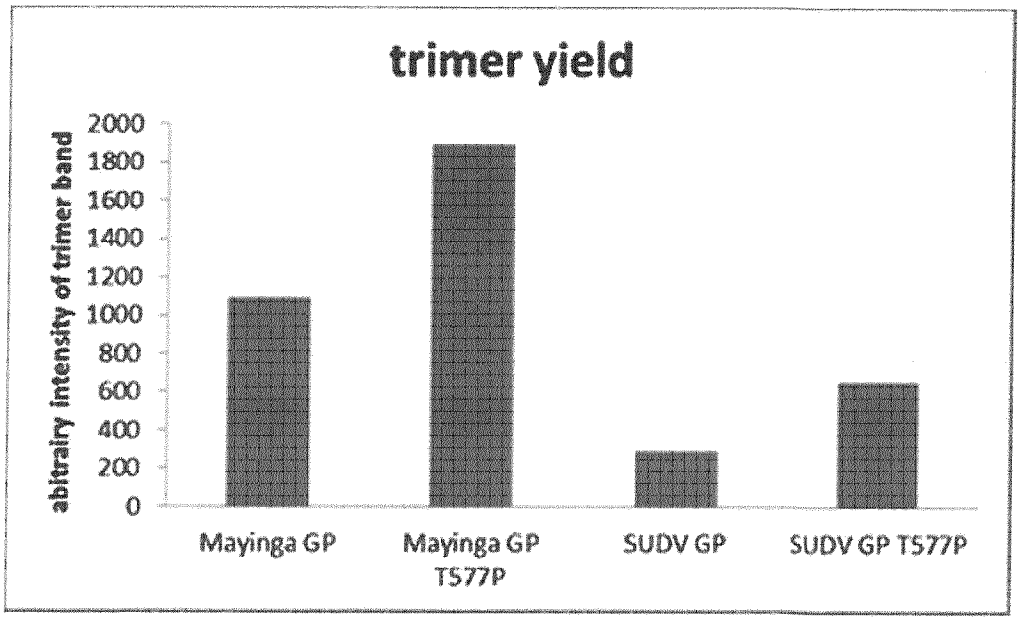

trimer yield
abitrairy intensity of trimer band
2000
1800
1600
1400
1200
1000
800
600
400
200
0
Mayinga GP
Mayinga GP T577P
SUDV GP
SUDV GP T577P

Fig. 7

Fig. 8A

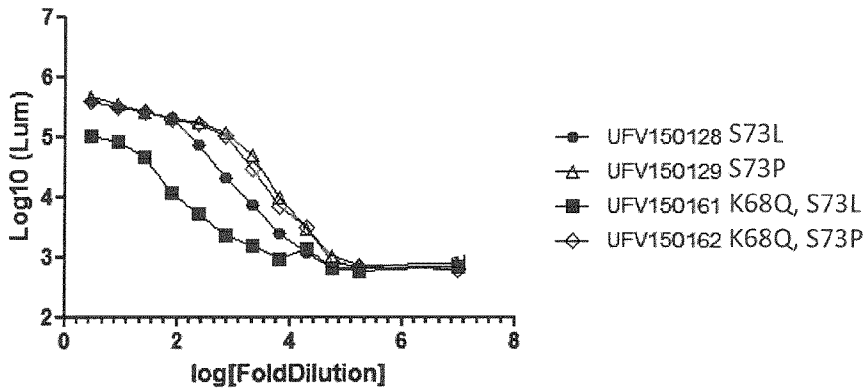


Fig. 8B

Construct	Description	Expression	Binding CR9114	Binding CR6261	Binding CR14045	Multimer ELISA
UFV150128	S73L	-	-	-	-	3.02
UFV150129	S73P	+	+	+	+	3.71
UFV150161	K68Q, S73L	-	-	-	-	2.00
UFV150162	K68Q, S73P	++	+	+	+	3.64

Fig. 8C