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

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
19 December 2024 (19.12.2024)

WIPO | PCT

(10) International Publication Number
WO 2024/259230 A1

(51) International Patent Classification:
*A61K 39/12* (2006.01) *A61K 39/00* (2006.01)
*A61P 31/22* (2006.01)

(21) International Application Number:
PCT/US2024/034001

(22) International Filing Date:
14 June 2024 (14.06.2024)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
63/508,091 14 June 2023 (14.06.2023) US

(71) Applicant: **CEVA SANTE ANIMALE** [FR/FR]; 10 avenue de la Ballastiere, 33500 Libourne (FR).

(72) Inventor; and
(71) Applicant *(for US only)*: **KAMRUD, Kurt** [US/US]; 8906 Rosehill Road, Lenexa, Kansas 66215 (US).

(72) Inventor: **ROTH, Jason**; 8906 Rosehill Road, Lenexa, Kansas 66215 (US).

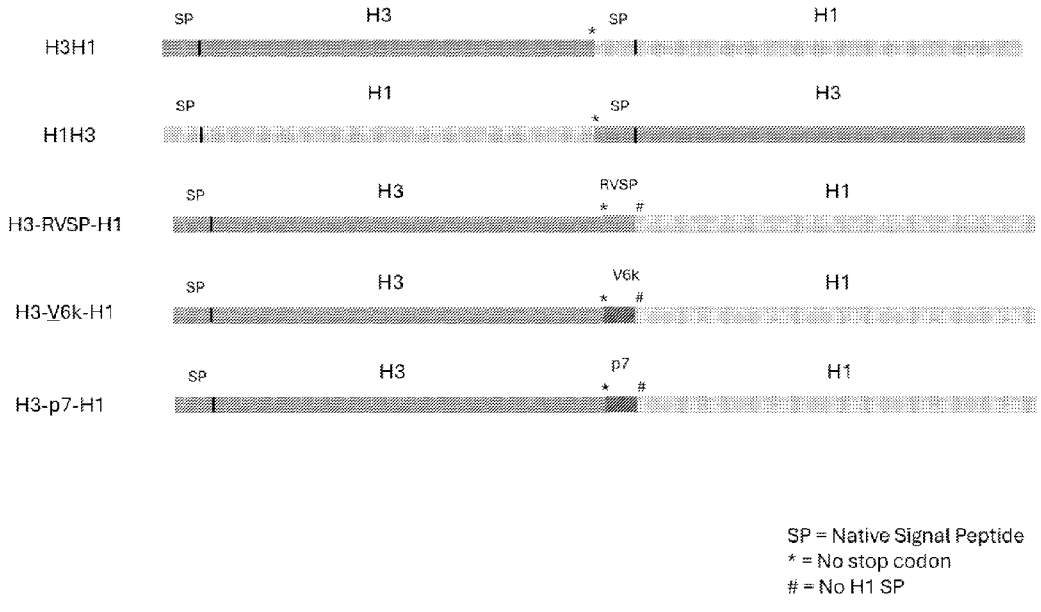
(74) Agent: **OLLIS, Andrew** et al.; Oblon, McClelland, Maier & Neustadt, L.L.P., 1940 Duke Street, Alexandria, Virginia 22314 (US).

(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, CV, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IQ, IR, IS, IT, JM, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, MG, MK, MN, MU, 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, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, WS, ZA, ZM, ZW.

WO 2024/259230 A1

(54) **Title:** NUCLEIC ACID CONSTRUCT COMPRISING SIGNAL PEPTIDE CLEAVAGE SITES EXPRESSING MULTIPLE ANTIGENS

FIG. 1




(57) **Abstract:** A nucleic acid construct encoding multiple polypeptide antigens or immunogens or other polypeptides of interest. The construct expresses a protein that comprises immunogenic or antigenic sequences of two, three or more polypeptides of interest. Once expressed by the nucleic acid construct, the protein, is processed by cleavage of an N-terminal signal peptide and cleavage of 6K, 6K-like, or internal signal peptide cleavage sites to release each polypeptide of interest. Vectors, including replicons and live virus vectors, comprising the construct may be used to deliver it to a subject in need of vaccination. Vaccines and methods of treatment using the nucleic acid construct are also described.

*[Continued on next page]*

**(84) Designated States** *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, CV, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SC, 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, ME, 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).

**Published:**

— *with international search report (Art. 21(3))*
— *before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments (Rule 48.2(h))*
— *with sequence listing part of description (Rule 5.2(a))*

# NUCLEIC ACID CONSTRUCT COMPRISING SIGNAL PEPTIDE CLEAVAGE SITES EXPRESSING MULTIPLE ANTIGENS

## CROSS-REFERENCE TO RELATED APPLICATION

This application claims priority to U.S. Provisional 63/508,091 filed June 14, 2023 and which is incorporated by reference for all purposes.

## REFERENCE TO A SEQUENCE LISTING

In accordance with 37 CFR §1.831-1835 and 37 CFR §1.77(b)(5), the specification makes reference to a Sequence Listing submitted electronically as a .xml file named "545868WO_ST26. This .xml file was generated on June 12, 2024 and is 45,864 bytes in size. The entire contents of the Sequence Listing are hereby incorporated by reference.

## BACKGROUND OF THE INVENTION

*Field of the Invention.* The present invention relates to the fields of molecular biology, biotechnology, and immunology. Specifically, it is directed to a nucleic acid construct encoding multiple polypeptide antigens or immunogens or other polypeptides of interest. The invention provides a novel way using embedded signal protease sites to simultaneously express two or more antigens from a single polynucleotide construct. The construct can be used along with various delivery systems including, but not limited to, plasmids, viral vectors, self-replicating RNA, or live virus vaccines.

*Description of Related Art.* Various systems have been proposed for expression of multiple recombinant proteins from the same vector or nucleic acid construct. Constructs which attempt to express multiple proteins from a single construct such as a plasmid, viral vector, polycistronic construct, or self-replicating RNA construct suffer from a number of problems including from gene interference, where the expression of one gene on the construct interferes with the expression of another gene; from variable gene expression, where different genes are expressed to different extents; from position effects, where the relative positioning of genes upstream or downstream on a construct affect gene expression; from epigenetic modifications which differ among genes on the same construct and can lead to differential expression of the genes; from transcriptional interference, where transcription of one gene interferences with that of another, leading to altered expression levels; from size limitations that restrict the number of non-coding segments on a construct; from gene incompatibility where co-expression of two or more genes leads to

undesirable results; from promoter competition, which leads to different expression of different genes; or from regulatory interactions among genes on the same construct leading to undesirable results.

In view of the above, the inventors sought to engineer a single polynucleotide construct capable of efficiently and reliably expressing multiple recombinant polypeptides of interest. This construct encodes a long protein which, once expressed, is cleaved into individual proteins. For example, it can encode two or more viral structural antigens from an animal virus which are then post-translationally processed into individual viral antigens or immunogens. As each independent antigen is derived from the same long protein, this method can produce stoichiometric amounts of each antigen. As processing of the protein typically occurs in the endoplasmic reticulum (ER), the polypeptides of interest once freed from the protein may be further processed or may further associate with one another, for example, to form tertiary or quaternary epitopes or structures absent from conventionally expressed recombinant proteins.

Maintaining stoichiometry of the amounts of viral or other types of proteins is beneficial as it permits proper assembly of, stability of, and efficient production of, multiprotein complexes, such as viral capsids; can provide for proper post-translational trafficking and processing of the recombinant proteins once expressed and can help form immunogenic complexes of recombinant proteins comprising conformational epitopes comprising chains or residues of different proteins in a quaternary complex.

## BRIEF SUMMARY OF THE INVENTION

One aspect of the invention is directed to nucleic acid constructs comprising polynucleotides coding for a protein or a recombinant protein comprising an N-terminal signal peptide, a first polypeptide of interest, a second polypeptide of interest and an internal signal sequence, such as a 6K or 6K-like sequence, positioned between the first polypeptide of interest and the second polypeptide of interest; wherein the internal signal sequence is cleavable by a protease to yield independent first and second polypeptides. The N-terminal signal peptide can function to translocate the protein from the cytoplasm to the endoplasmic reticulum.

Other aspects of the invention involve vectors that transfect or transform cells, such as host cells, in a non-human animal to be vaccinated and a method for vaccinating or immunizing a non-human animal.

The invention also concerns methods for making and using the polypeptides expressed by the nucleic acid construct and processed into individual polypeptides of interest by cellular proteases.

Specific, non-limiting embodiments include the following.

1. A nucleic acid construct comprising polynucleotides coding for a protein comprising an N-terminal signal peptide, a first polypeptide of interest, a second polypeptide of interest and an internal signal peptide sequence positioned between the first polypeptide and the second polypeptide; wherein the internal signal sequence is cleavable by a protease to yield independent first and second polypeptides; preferably the first and second polypeptides of interest comprise HA antigens and/or the internal signal sequence comprises an RVSP linker. In some embodiments, one or more internal signal peptide sequences are replaced by linker sequences such as RVSP or said internal signal sequences comprise linker sequences.

2. The nucleic acid construct of embodiment 1, wherein the construct further encodes an additional internal signal peptide sequence and a third polypeptide of interest positioned after the second polypeptide of interest.

3. The nucleic acid construct of embodiment 1 or 2, wherein the construct further encodes an additional internal signal peptide sequence and a subsequent polypeptide of interest positioned after the last downstream polypeptide of interest.

4. The nucleic acid construct of embodiment 1-3, wherein the first, the second and subsequent polypeptides of interest differ from each other.

5. The nucleic acid construct of any one of embodiments 1-4 that further comprises a promoter or other regulatory region upstream from the polynucleotide encoding the first polypeptide of interest.

6. The nucleic acid construct of embodiments 1-5, wherein the N-terminal signal peptide directs the protein encoded by the construct into the endoplasmic reticulum.

7. The nucleic acid construct of any one of embodiments 1-6, wherein the encoded N-terminal signal peptide comprises 15 to 150 residues.

8. The nucleic acid construct of any one of embodiments 1-7, wherein the encoded N-terminal signal peptide comprises 20-30 residues.

9. The nucleic acid construct of any one of embodiments 1-8, wherein the encoded internal signal peptide sequence comprises 15 to 75 residues.

10. The nucleic acid construct of any one of embodiments 1-9, wherein the encoded internal signal peptide sequence comprises less than 20 residues.

11. The nucleic acid construct of any one of embodiments 1-10, wherein the encoded N-terminal or internal signal peptide sequences comprise a motif comprising:

an N-region comprising 1 to 5 charged residues including but not limited to Arg and Lys,

an H region spanning 7-15 contiguous hydrophobic residues of which 3-5 residues are leucine residues (including but not limited to Leu and Gly or other residues forming an alpha helix),

a C-region comprising 3-7 uncharged amino acid residues (including but not limited to Ala and Val, or other residues forming a beta sheet), and/or

a Pro-region or negative region comprising 1-6 charged amino acid residues. including but not limited to Glu and Ala.

12. The nucleic acid construct of any one of embodiments 1-11, wherein the encoded N-terminal and/or internal signal peptide regions comprise a cleavage motif of AXA or VXA between a C region and a Pro-region corresponding to a N region.

13. The nucleic acid construct of any one of embodiments 1-12, wherein the encoded N-terminal and/or internal signal peptide sequence comprises a viral internal cleavage site.

14. The nucleic acid construct of any one of embodiments 1-13, wherein the encoded N-terminal and/or internal signal peptide sequence comprises all or portion of a 6K polypeptide from a virus in the Alphavirus genus or other viral 6K-like polypeptide.

15. The nucleic acid construct of any one of embodiments 1-14, wherein the encoded N-terminal and/or internal signal peptide sequence comprises an E2/p7 or p7/NS2 cleavage site of hepatitis C virus.

16. The nucleic acid construct of any one of embodiments 1-15, wherein the encoded N-terminal and/or internal signal peptide sequence comprises a cleavage site from a virus in the order Bunyavirales.

17. The nucleic acid construct of any one of embodiments 1-16, wherein the encoded N-terminal and/or internal signal sequence comprises a WAASA motif.

18. The nucleic acid construct of any one of embodiments 1-17, wherein the first and second encoded polypeptides comprise non-viral or viral polypeptides.

19. The nucleic acid construct of any one of embodiments 1-18, wherein the first and second encoded polypeptides comprise antigenic or immunogenic polypeptides of a virus that infects domesticated animals.

20. The nucleic acid construct of any one of embodiments 1-19, wherein the first and second encoded polypeptides are antigenic or immunogenic polypeptides of Infectious Laryngotracheitis virus (ILTV) polypeptides or polypeptides of another virus that infects avians.

21. The nucleic acid construct of any one of embodiments 1-19, wherein the first and second encoded polypeptides comprise antigenic or immunogenic polypeptides of a virus that infects non-human mammals.

22. A vector comprising the nucleic acid construct of any one of embodiments 1-21 that is engineered so that can be transformed or transfected into a host cell to express the protein.

23. The vector of embodiment 22 that is a modified live virus vector.

24. The vector of embodiment 23 that is a modified live pox, herpesvirus, adenovirus, lentivirus, or retrovirus vector.

25. The vector of embodiment 22 that is DNA or modified DNA, which, optionally, replicates within a host cell.

26. The vector of embodiment 22 that is a DNA launched platform that when introduced into a host cell transcribes and translates the nucleic acid construct into the polyprotein; and that optionally comprises an origin of replication and/or a selectable marker.

27. The vector of embodiment 22 that is RNA or modified RNA, which, optionally, replicates within a host cell.

28. The vector of embodiment 22 that further comprises a 5' cap, a 5' and 3' untranslated region (UTR) and/or a poly A tail.

29. The vector of embodiment 22 that is an RNA launched platform that when introduced into a host cell translates the nucleic acid construct into the polyprotein; and that optionally replicates itself when introduced into the host cell.

30. The vector of embodiment 22 that is self-replicating RNA and that comprises a 5' cap, 5' UTR, a nucleic acid sequence encoding alphavirus non-structural proteins nsP1, nsP2, nsP3 and nsP4, a 26S promoter, the nucleic acid sequence encoding the first polypeptide of interest, the nucleic acid encoding the signal sequence, the nucleic acid encoding the second polypeptide of interest, and a 3'UTR and poly A tail.

31. The vector of embodiment 22 that is a Togavirus vector including but not limited to a SFV, SINV, or VEEV vector.

32. The vector of embodiment 22 that is a Flaviviridae vector including but not limited to a Flavivirus vector or a Pestivivus vector.

33. The vector of embodiment 22 that is an Orthomyxoviridae vector including but not limited to an Influenza virus vector.

34. The vector of embodiment 22 that is an Rhabdoviridae vector including but not limited to a Lyssavirus vector, a Vesiculovirus vector, or a Novirhabdovirus vector.

35. The vector of embodiment 22 that is a Paramyxoviridae vector including but not limited to a Pneumomyxovirinae vector, a Pneumovirus vector, a Paramyxovirinae vector, a Respirovirus vector, an Avulavirus vector, a Rubulavirus vector, a Morbilivirus vector, Picornaviridae family vector, Astroviridae family vector, and Nidovirales order vector.

36. A host cell transformed with, or transfected with, the vector of any one of embodiments 22-35

37. The host cell of embodiment 36 that is a cell of a non-human animal.

38. The host cell of embodiment 36 that is a host cell from a mammal.

39. The host cell of embodiment 36 that is an avian host cell.

40. The host cell of embodiment 36 to which a virus or other microorganism encoding the domain of the first, second, third, or subsequent polypeptide of interest is tropic.

41. The host cell of embodiment 36 that is a bacterial, yeast, insect, mammalian or plant cell or a culture cell derived therefrom.

42. A protein encoded by the nucleic acid construct of any one of embodiments 1-21.

43. A first and second polypeptide of interest produced by cleavage of the protein encoded by the construct of embodiment 1.

44. A first, second, and third polypeptide of interest produced by cleavage of the protein encoded by the construct of embodiment 2.

45. A first, second, third and any subsequent polypeptides of interest produced by cleavage of the protein encoded by the construct of embodiment 3.

46. A method for producing two or more polypeptides of interest comprising transforming or transfecting a vector of any one of embodiments 21-34 into a host cell competent to cleave the N-terminal and internal signal peptide domains.

47. The method of embodiment 46 that produces similar or 1:1 ± 20 mole% stoichiometric amounts of the first, second, third or subsequent polypeptides of interest.

The foregoing paragraphs have been provided by way of general introduction and are not intended to limit the scope of the following claims. The described embodiments, together with further advantages, will be best understood by reference to the following detailed description taken in conjunction with the accompanying drawings.

## BRIEF DESCRIPTION OF THE DRAWINGS

A more complete appreciation of the disclosure and many of the attendant advantages thereof will be readily obtained as the same becomes better understood by reference to the following detailed description when considered in connection with the accompanying drawings described below.

FIG. 1 describes five replicons: H3H1, H1H3, H3-RVSP-H1, H3-V6k-H1 and H3-p7-H1. The term "SP" refers to signal peptide sequences. The star indicates the absence of stop codons and the "#" denotes no H1 signal peptide. This figure illustrates each construct of Example 2 *infra* associated with a linker. In some embodiments a construct will contain elements without stop codons as exemplified in this figure or as placed between particular elements as shown in this figures.

FIG. 2 (from Owji, et al., 2018, *infra*) depicts a general signal peptide structure and cleavage site comprising a positively charged N region, an H region that comprises a hydrophobic core, and a C region involved in cleavage. A 6K element may comprise two signal peptides separated by other residues of a viral 6K or 6K-like protein. The H region sequence is described by SEQ ID NO:6.

FIG. 3A (from Löber, et al., 2001, *infra*, incorporated by reference) depicts Hantaan virus glycoprotein precursor and putative signal peptide cleavage site comprising the N, H and C regions. The sequences of the N, H and C regions are given by SEQ ID NOS: 7-9.

FIG. 3B (from Löber, et al., 2001, *infra*, incorporated by reference) depicts putative signal peptide cleavage sites comprising the N, H and C regions from different bunyavirus glycoprotein precursors. The arrows indicate the potential cleavage sites between either the two glycoproteins G1 and G2 (genera *Hantavirus* and *Phlebovirus*) or $NS_m$ and G1 (genus *Bunyavirus*). The N region sequences from top to bottom are given by SEQ ID NOS: 10-19. The H region sequences from top to bottom are given by SEQ ID NOS: 20-29 and the C region sequences from top to bottom are described by SEQ I DNOS: 30-39.

FIG. 4A. Analysis of cleavage site for HCV strain Hp7, see Lin, et al., *Processing in the hepatitis C virus E2-NS2 region: identification of p7 and two distinct E2-specific products with different C termini,* J. VIROL. 1994, 68(8) 5063-5073. See also sequences in Fig. 5 of Lin, et al., *id.* which are incorporated by reference. The combined E2, P7 and NS2 sequences are described by SEQ ID NO: 40. The lower sequence comprises residues 1 to 69 of SEQ ID NO: 40.

FIG. 4B. Analysis of cleavage sites for Bunyamwera. The Buyamwera sequence is described by SEQ ID NO: 41. The signal peptide predictions in FIGS. 4A and 4B were obtained from an online signal peptide predictor.

FIG. 5A describes 6K and 6K-like sequences and their domain structures. The V6k, S6k, BSP, HSP and P7 sequences in FIG. 5A are described by SEQ ID NOS: 42-46 respectively.

FIG. 5B describes 6K and 6K-like sequences and describes the different types of amino acid residues. The V6k, S6k, BSP, HSP, and P7 sequences in FIG. 5B are described by SEQ ID NOS: 47-51, respectively.

FIG. 6 pertains to the SIV-H3-H1 Linker Construct.

## DETAILED DESCRIPTION OF THE INVENTION

As disclosed herein, the invention pertains to a nucleic acid construct engineered to comprise a coding sequence that expresses a protein comprising two, three, four, or more polypeptides of interest separated by 6K, 6K-like, or internal signal peptide cleavage sites. Once expressed the protein encoded by the coding sequence is considered to be transported to the endoplasmic reticulum with the help of the N-terminal signal peptide Proteases, such as signal peptidases, in the ER then cleave the protein at the N-terminal or internal signal peptide cleavage sites releasing the two or more polypeptides of interest which are then free to be separately processed or to associate. The invention also pertains to vectors and host cells comprising the

nucleic acid construct and expressing the protein comprising the first, second or subsequent polypeptides of interest.

As used herein, the term "polynucleotide" or "nucleic acid" refers to a linear polymer whose molecule is composed of many nucleotide units. It may constitute a subsection of a nucleic acid molecule embedded in a longer polynucleotide or, alternatively, constitute a stand-alone nucleic acid with discrete termini such as a terminal 5' phosphate group and a free 3' hydroxy group.

As used herein, the term "polypeptide "or "protein" refers to chain of amino acids linked by peptide bonds, which has a defined N and C terminus, or to a chain of amino acids which forms a domain, region, or fragment of a larger polypeptide, such as a portion of a fusion protein, chimeric protein, or polyprotein. Thus, a polypeptide can be an independent, stand-alone polypeptide or can be a subunit of a larger polypeptide adjacent to or flanked by other sequences.

The term"polyprotein" as defined herein refers to a protein that, after synthesis, is cleaved to produce two or more distinct polypeptides. A polyprotein is typically encoded by a single mRNA coding sequence that is translated into the polyprotein. The single mRNA may be one found in nature or an artificial construct, such as a construct comprising coding sequences obtained from different viral species or strains or a coding sequence that encodes a limited number of antigens from the same virus or microorganism not necessarily in the same order as they appear in the viral or microorganism's chromosome or mRNA. Polyproteins are found in nature. As defined herein, they may also be synthetic or recombinant polypeptides such as cleavable chimeric or fusion proteins comprising immunogenic or antigenic stretches of amino acid residues from two, three, or more different polypeptide antigens.

The term ***"coding sequence"*** refers to a stretch of polynucleotides, which may be RNA or DNA, which directly, or in the case of DNA ultimately, encodes a stretch of amino acids forming a polypeptide. A coding sequence may be one found in nature or may be artificially constructed.

The term ***"region"*** refers to different functional subparts of a polyprotein or chimeric protein encoded by a nucleic acid construct, such as stretches of amino acid that will later be liberated as independent polypeptides or other functional stretches, *e.g.* "a N-terminal signal peptide region". "Region" as well as "fragment" or "segment" or "domain" may be used to refer to subparts of a longer polyprotein, chimeric protein, or fusion protein.

An ***"open reading frame"*** or ORF, for example, in viral polynucleotides typically begins with a start codon and ends with a stop codon. It may encode one or more viral polypeptides.

Embodiments of the invention include but are not limited to the following.

***Nucleic acid constructs.*** One aspect of the invention involves a polynucleotide or nucleic acid construct that encodes a protein comprising an N-terminal signal peptide, a first polypeptide of interest, an internal signal sequence or 6K or 6K-like cleavable by a protease, and a second polypeptide of interest. Upon contact of the protein with a protease recognizing the signal peptide, such as those present in the 6K or 6K-like sequences, the first and second polypeptides are cleaved and then released as independent polypeptides. In some embodiments, a lone signal peptide cleavage sequence will be positioned between an upstream and downstream polypeptide of interest. In other embodiments, multiple signal peptide sequences will be positioned between the polypeptides of interest encoded by the construct. These multiple signal peptide sequences include but are not limited to signal peptide sequences within a 6K or 6K-like polypeptide sequence.

The nucleic acid sequence encoding the protein may be considered to fall within a single open reading frame (ORF) as this entire stretch of polynucleotides has the potential to be translated into protein, *i.e.,* into the proteins or polypeptides disclosed herein.

Advantageously the invention can express two, three, or more proteins that are never associated with each other in nature, for example, the coding sequence can encode influenza HA polypeptide antigens or epitopes and also porcine reproductive and respiratory viral antigens or epitopes. It could also encode a subset of antigens from a particular virus (or microorganism) but not others. Thus, it could omit antigens that inhibit or dominate immune responses. It could also encode antigens from a virus in a different order than that found in nature or in the same order. For example, if viral antigens A, B, C and D are expressed by a virus mRNA in the order ABCD in a viral ORF, then a construct of the invention may express the same antigens in a different order such as BADC, DCBA, etc. based on their order of appearance in the coding sequence of the polynucleotide construct.

***Three or more polypeptides of interest.*** In some embodiments, the nucleic acid construct disclosed herein encodes additional polypeptides of interest. For example, in these embodiments, the protein can further comprise an additional 6K or 6K-like region or internal signal peptide and a third polypeptide of interest positioned downstream from the second polypeptide of interest.

Similarly, the construct may further encode a protein comprising fourth, fifth, or subsequent polypeptide of interest each preceded by a 6K or 6K-like region or an internal signal sequence.

As disclosed herein, a ***polyprotein*** is a large protein molecule that is a single protein chain that contains multiple functional regions or subunits, each of which can be cleaved or processed into a separate, functional protein. These functional proteins can include structural proteins, antigens, immunogens, enzymes or other proteins of interest derived from a longer polyprotein or polypeptide construct. In some embodiments of the invention, a polyprotein comprises two, three or more polypeptide of interest derived from an immunogen or antigen, such a viral immunogen or antigen. In a further embodiment, the two, three or more polypeptides of interest may derive from the same gene of interest or from different genes of interest. In another embodiment, the two, three or more polypeptides of interest are different from each other, for example, are portions of different viral antigens..

The ***coding sequence*** of the construct of the invention encodes a protein or polypeptide comprising two, three, or more antigenic or functional segments separated by cleavable signal peptides, such as 6K or 6K-like polypeptides. Cleavage of the signal peptides releases these separated antigenic regions as independent polypeptides which can be independently trafficked, complexed, or folded. Often the resulting independent polypeptides produced by cleavage are present in stoichiometric amounts based on the number of polypeptide regions of interest encoded by the construct. This avoids problems of expressing known amounts of multiple antigens from different vectors or different RNA transcripts.

While not being bound by any particular theory or explanation, the inventors consider that when a protein encoded by the construct disclosed herein is transported to the endoplasmic reticulum it undergoes a series of events that lead to the cleavage and/or activation of its individual components, such as the first, second, or subsequent polypeptides of interest derived from different polypeptide antigens. The protein is then considered to be transported across the cytoplasmic membrane of the host cell an into the ER lumen, a process that is facilitated by the presence of a signal peptide at the N-terminal of the protein. This N-terminal signal peptide serves as a molecular tag that helps direct the polyprotein into the ER. The signal peptide is subsequently cleaved by a signal peptidase, which is a specialized protease found in the ER lumen. This cleavage releases the mature protein from the N-terminal signal peptide. The mature protein is then cleaved into its individual compositions by proteases found within the ER lumen which

recognize specific cleavage sites within the protein, such as the 6K, 6K-like or signal peptidase cleavage sites engineered into the nucleic acid construct.

***Regulatory elements of the nucleic acid construct.*** In some embodiments the nucleic acid construct disclosed herein further comprises a promoter, enhancer, silencer, or other regulatory region, such as a promoter or ribosome binding site upstream from the protein coding sequences or for RNA constructs or transcripts a 3' UTR, polyA tail, or terminator sequence.

***Other peptide sequences in the polyprotein.*** In some embodiments, the protein may include other peptide sequences, such as a rigid or flexible linker or an affinity tag between the domains disclosed above. A linker may facilitate folding and proper processing the protein when translocated to the endoplasmic reticulum (ER). For applications involving recovery of individual polypeptides of interest, an affinity tag (*e.g.* a His-tag comprising 6-10 His residues or a FLAG-tag: DYKDDDK) can facilitate isolation or purification of a polypeptide of interest produced by processing of the polyprotein. Such additional peptide elements may, or may not, form portions of the regions of the protein comprising the two, three, or more polypeptides of interest.

The protein encoded by the nucleic acid construct disclosed herein, typically comprises an ***N-terminal signal peptide region.*** Preferably, in terms of processing the protein into individual polypeptides, the N-terminal signal peptide region will be one that efficiently translocates a translated protein into the ER. In some embodiments, the N-terminal signal peptide region comprises 15, 20, 30, 40, 50, 60, 70, 80, 90, 100, 110, 120, 130, 140 to 150 residues or any intermediate length within this range. By comparison the internal signal sequence peptides mentioned below often comprise 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70 to 75 residues or in other embodiments less than 15 or 20 residues.

***6K and internal signal peptide cleavage regions.*** The protein encoded by the nucleic acid constructs disclosed herein comprises cleavable sequences such as those of an alphavirus 6K or 6K-like protein or an internal signal peptidase cleavage site of other viruses. Cleavage of these sequences, which are interposed between the polypeptides of interest, releases the two, three or more polypeptides of interest from the polyprotein or polypeptide construct.

***Internal signal peptide sequences.*** The constructs disclosed herein comprise sequences encoding one or more internal signal peptides, 6K, or 6K-like proteins which are positioned adjacent to and between coding sequences for polypeptides of interest, such as coding sequences

for different viral antigens. One or more internal signal peptide sequences may be incorporated into a construct as disclosed herein.

Signal peptides and their cleavage sites may be identified for use within the constructs disclosed herein using a computer program such as ***Signal IP - 6.0*** which is further described and incorporated by reference to <hypertext transfer protocol secure://services.healthtech.dtu.dk/services/SignalP-6.0/> (last accessed May 10, 2023). The SignalP 6.0 server predicts the presence of signal peptides and the location of their cleavage sites in proteins from Archaea, Gram-positive Bacteria, Gram-negative Bacteria and Eukarya. In Bacteria and Archaea, Signal IP 6.0 can discriminate between five types of signal peptides:

Sec/SPI: "standard" secretory signal peptides transported by the Sec translocon and cleaved by Signal Peptidase I (*Lep*),

Sec/SPII: lipoprotein signal peptides transported by the Sec translocon and cleaved by Signal Peptidase II (*Lsp*),

Tat/SPI: Tat signal peptides transported by the Tat translocon and cleaved by Signal Peptidase I (*Lep*),

Tat/SPII: Tat lipoprotein signal peptides transported by the Tat translocon and cleaved by Signal Peptidase II (*Lsp*),

Sec/SPIII: Pilin and pilin-like signal peptides transported by the Sec translocon and cleaved by Signal Peptidase III (*PilD/PibD*).

Additionally, SignalP 6.0 predicts the regions of signal peptides. Depending on the type, the positions of *n-*, *h-* and *c-regions* as well as of other distinctive features are predicted. SignalP 6.0 is based on a transformer protein language model with a conditional random field for structured prediction. Preferred settings for this program include organism: “Eukarya” (predicts Sec/SP1 SPs) or “other”; “Long output” or “Short output” (no figures); and “Fast” or “Slow” model mode. Slow model mode may be used to accurately describe region borders.

***Linker sequences.*** In some embodiments an RVSP protein sequence, for example those derived from Rvs161 or Rvvs167 proteins may be used to internally link antigens expressed by the constructs disclosed herein, see e.g. the construct H1-RVSP-H3 in Fig. 1. Such linkage may provide for proper orientation, spacing and folding of the antigens expressed. Such linkage may substitute for other linkers or signal peptide sequences or may be used in conjunction with them.

***The 6K protein or 6K-like proteins.*** The term "6K" or "6K-like" protein refers to a small membrane protein encoded by alphaviruses and certain other viruses and is named after its approximate molecular weight of 6 kDa.

The term ***6K protein*** refers to a small alphavirus protein that separates the E2 and E1 membrane glycoproteins. During translation signal peptide (SP) sequences found on the N and C termini of the 6K protein are cleaved to relieve the upstream E2 membrane protein and downstream E1 membrane protein from the small 6K protein. Alphavirus 6K proteins are small (58–61 amino acids), hydrophobic, and associate with membranes. Alphaviruses, such as Venezuelan Equine Encephalitis Virus (VEEV), typically encode the 6K protein between their E2 and E1 glycoproteins. Other viruses of the Alphavirus genus that encode a 6K protein include, but not limited to, Sindbis virus (SINV), Chickungunya virus (CHIKV), Semliki Forest Virus (SFV), Ross River Virus (RRV), Sagiyama virus (SAGV), Getah virus (GETV), Middleburg virus (MIDV), Bebaru virus (BEBV), O'nyong nyong virus (ONNV), Ndumu (NDUV), and Barmah Forest virus (BFV). New World alphavirus such as Venezuelan Equine Encephalitis Virus (VEEV), western equine encephalitis virus (WEEV), and eastern equine encephalitis virus (EEEV) also encode a 6K protein. The "6K" protein comprises an internal signal peptide that facilitates the expression of two membrane-bound viral antigens. When a 6K protein is analyzed an internal signal peptide region is found which permits cleavage of the 6K protein away from E1 and E2.

Cleavage of a 6K protein to generate individual structural proteins (*e.g.*, the first and second polypeptides of interest encoded by the construct) occurs via cellular proteases such as signal peptidases. In some embodiments, a nucleic acid construct as disclosed herein may encode an exogenous protease that recognizes and cleaves the 6K domains of the polyprotein it encodes.

The term ***6K-like protein*** refers to a non-alphavirus protein that separates an upstream protein from a downstream protein. That is, it has a similar arrangement to the alphavirus 6K protein described above but is not from an alphavirus. An example of a 6K-like protein is the HCV p7 protein. It separates the E2 and NS2 HCV proteins and also has SP at its N and C terminus that are used to cleave the upstream E2 protein and downstream NS2 protein from the small p7 protein.

A 6K or 6K-like viral protein sequence is not required to be present between upstream and downstream polypeptide sequences encoded by a construct when an alternative internal signal peptide separating them is sufficient to provide a cleavage site. As mentioned above one example

of a 6K-like protein is HCV p7 which is a viral protein found between and in the same open reading frame as E2 and NS2. Other examples of 6K or 6K-like proteins encoded by polynucleotide sequences that are incorporated into the constructs disclosed herein include:

>p7:ALENLVILNAASLAGTHGLVSFLVFFCFAWYLKGRWVPGAVYAFYGMWPL LLLLLALPQRAYA. (SEQ ID NO: 1)

>S6k:ASVAETMAYLWDQNQALFWLEFAAPVACILIITYCLRNVLCCCKSLSFLVL LSLGATARA. (SEQ ID NO: 2)

>V6k:GRPRAETTWESLDHLWNNNQQMFWIQLLIPLAALIVVTRLLRCVCCVVPF LVM AGAAGAGA. (SEQ ID NO: 3)

***Signal peptide sequences*** distinct from those in 6K and 6K-like proteins include a Hantavirus signal peptide that is used to allow viral maturation of two envelope proteins from the same open reading frame. Other examples include Bunyavirus and Hantavirus sequences:

>BSP: SLIISILLSVLILSFVTPIEGT. (SEQ ID NO: 4)

>HSP: RTLNLFRYKSRCYIFTMWIFLLVLESILWAASA. (SEQ ID NO: 5)

***Signal sequence cleavage sites including 6K or 6K-like cleavage sites.*** N-terminal as well as internal signal peptide sequences and signal peptide cleavage sites may be determined using a program such as SignalP 6.0 or can be recognized by the presence of a structural motif such as those disclosed herein.

An N-terminal signal peptide typically directs an immature protein into the endoplasmic reticulum (ER) where the immature protein is processed, for example, by cleavage of the N-terminal signal peptide and internal signal peptides separating regions of the immature protein. For example, an immature protein or polyprotein comprising regions of two, three, or more antigens separated by internal signal peptides may be processed into separate antigens of interest. The cleavage of the N-terminus signal peptide is typically the first step in processing the protein into its mature form (*e.g.*, no N-terminal signal peptide) , Subsequently, the partially processed protein is then cleaved into the two or more antigenic or immunogenic polypeptides of interest via the internal signal peptidase sequence sites to produce independent monomer polypeptides of interest.

The independent monomer polypeptides of interest, processed from the larger immature protein or polyprotein, may then be trafficked for secretion from a cell or for insertion into the

cell membrane. Independent monomer polypeptides that lack a transmembrane domain and are otherwise incapable of binding to the membrane are often secreted from the cell.

In some embodiments, the internal signal peptide may remain fused to an upstream or downstream polypeptide of interest. When this signal peptide also comprises a transmembrane domain, this may result in insertion of the polypeptide of interest into the membrane.

In some alternative embodiments, the proteases may be derived from viruses or be an exogenous protease encoded by the nucleic acid construct itself.

***Cleavage motifs.*** Many signal sequences comprise common motifs; see Owji, H. et al., *A comprehensive review of signal peptides: Structure, roles, and applications,* EUR. J. CELL BIOL. , 2018, 97:422-441 (incorporated by reference) or Löber, et al., *The Hantaan virus glycoprotein precursor is cleaved at the conserved pentapeptide WAASA*, VIROLOGY (2001), 289, 224-229 (incorporated by reference).

The nucleic acid construct disclosed herein may comprise a N-terminal or internal signal sequence that comprises a cleavage motif comprising:

a. an N-region comprising 1 to 5 charged residues (including but not limited to Arg and Lys),

b. an H region spanning 7-15 contiguous hydrophobic residues of which 3-5 residues are leucine residues (including but not limited to Leu and Gly or other residues forming an alpha helix),

c. a C-region comprising 3-7 uncharged amino acid residues (including but not limited to Ala and Val, or other residues forming a beta sheet), and/or

d. a N-region comprising 1-6 charged amino acid residues. (including but not limited to Glu and Ala). See Fig. 2 of Owji, H. et al., *supra.* (incorporated by reference).

In some embodiments, the nucleic acid construct disclosed herein comprise N-terminal and/or internal signal peptide sequences comprising a cleavage motif of AXA or VXA between a C region and a Pro-region corresponding to a N region. Fig. 2 of Owji, H. et al., *supra.* (incorporated by reference).

In other embodiments, the nucleic acid constructs disclosed herein comprise of one or more internal cleavage sites derived from a virus. These may include all or a portion of a 6K sequence from a virus in the Alphavirus genus; other viral 6K-like sequences; an E2/p7 or p7/NS2 cleavage site of hepatitis C virus; an internal signal peptidase cleavage site from a virus in the order

Bunyavirales; or a cleavage site comprising a WAASA motif; see Lin et al., *Processing in the Hepatitis C Viruses E2-NS2 Region: Identification of p7 and Two Distinct E2-Specific Products with Different C Termini,* J. VIROLOGY, 1994, 68(8), 5063-5076; Shi, X, et al., *Visualizing the Replication Cycle of Bunyamwere Orthobunyavirus Expressing Fluorescent Protein-Tagged Gc Glycoprotein,* J. VIROLOGY, 84(7), 8460-8469; Shi, et al., *Bunyamwera orthobunyavirus glycoprotein precursor is processed by cellular signal peptidase and signal peptide peptidase,* PNAS, 2016, <hypertext transfer protocol secure://doi.org/10.1073/pnas.1603364113>; Löeber, C. et al., *The Hantaan Virus Glycoprotein Precursor Is Cleaved at the Conserved Pentatpeptide WAASA.* VIROLOGY, 2001, 289, 224-229. The above references are each incorporated by reference with regard to signal peptide structure and signal peptide cleavage.

Signal peptidase cleavage sites from Hantaan and other viruses including those in FIG. 3A and 3B are described by Löber, et al., VIROLOGY (2001), 289, 224-229 (incorporated by reference).

The specific cleavage motifs for mammalian signal peptides can vary among different proteins and cell types, but they involve short, conserved sequences of amino acids that are recognized by the host proteases. The cleavage of the signal peptides is a regulated and controlled event, allowing the protein to be processed into its mature form and properly transported to its final destination around the cell. There are several publications that are incorporated by reference to describe, signal sequence cleavage motifs, which are the specific amino acid sequences found in signal peptides that are recognized and cleaved by proteases. These include and are incorporated by reference to Owji, H. et al., *A comprehensive review of signal peptides: Structure, roles, and applications,* EUR. J. CELL BIOL. , 2018, 97:422-441; Rapoport, T. A. (2010). *Protein translocation across the eukaryotic endoplasmic reticulum and bacterial plasma membranes.* NATURE, 2007, 450(7170), 663-9; Walter, P. et al., *Signal Sequence Recognition and Protein Targeting to the Endoplasmic Reticulum Membrane*, ANNUAL REVIEW OF CELL BIOLOGY, 1994, 10,87; and Kapp, K. et al., *Post-targeting functions of signal peptides,* PROTEIN TRANSPORT INTO THE ENDOPLASMIC RETICULUM, 2009.

These publications provide detailed information on the different types of signal sequences, their mechanisms of action, and the proteases involved in cleaving them. They also provide information on the different cleavage motifs that have been described for various signal sequences and their specificity for different proteases.

***Signal Peptidase*:** This is a membrane-bound protease that specifically cleaves the signal peptides from newly synthesized proteins, for example, releasing them into the lumen of the endoplasmic reticulum (ER). These include:

*Furin*: This is a cysteine protease that cleaves signal peptides at specific sites, allowing the transport of the mature protein to its final destination.

*Site-1 Protease (S1P):* This is a member of the subtilisin-like protease family and is involved in the maturation and sorting of newly synthesized proteins.

*Site-2 Protease (S2P):* This is another member of the subtilisin-like protease family and is involved in the maturation of newly synthesized proteins.

*PACE4* (Protease-Activated Cysteine Endopeptidase 4): This is a protease involved in the processing of growth factors and cytokines.

In some embodiments, cells expressing a protease described above are used as host cells to express and process the polyprotein encoded by the construct of the invention.

*Signal peptide cleavage regions or sites.* In some embodiments, there can be at least three different cleavage sites. The first one can be the alphavirus 6K, which includes all viruses from the Alphavirus genus. The second site can be 6k-like, one example is the HCV p7. Finally, the third site can be internal signal peptides or ISPs for short. These are often from viruses in the Bunyavirales order, which includes Hantaan, Bunyamwera viruses, and others.

Another common characteristic of most of these cleavage sites is the presence of at least one transmembrane domain. 6k and ISPs have 1 domain and p7 has two domains. A transmembrane domain may comprise 16-30, such as 20-25, amino acid residues that typically are predominantly non-polar residues.

***The polypeptide regions of interest.*** The polynucleotide or nucleic acid construct disclosed herein can be used to express two, three or more polypeptides of interest. There is no particular limitation on the type of polypeptide encoded by the construct. For applications involving immunization or vaccination, the polypeptide sequences of interest are typically derived from polypeptides comprising microbial T or B cell epitopes, such as polypeptides comprising one or more bacterial or viral antigen epitopes. These polypeptides may correspond to an entire microbial antigen or to an immunogenic or antigenic fragment such as a fragment comprising one or more B-cell or T-cell epitopes.

In some embodiments, these microbial polypeptides are antigenic or immunogenic polypeptides of a virus that infects domesticated animals. For example, polypeptides of interest may be those of pathogens causing a specific disease including the following:

poultry diseases: Newcastle disease (ND), avian influenza (AI), infectious bursal disease (IBD), infectious bronchitis (IB), Marek's disease, fowl cholera, infectious laryngotracheitis (ILT), and coccidiosis;

feline diseases: feline leukemia, feline infectious peritonitis;

canine diseases: distemper, adenovirus, parvovirus and leptospirosis; and

porcine diseases: Porcine Circovirus, swine Influenza virus, Rotavirus, Porcine Reproductive and Respiratory Syndrome Virus (PRRSV), Lawsonia, African Swine Fever (ASF), Classical Swine Fever (CSF) and Porcine Epidemic Diarrhea (PED).

A non-limited list of antigens which may be encoded by the constructs include gB, gD, gE and gI antigens from ILTV; H5, H9, H7 , H1N1 or H1N2 from AIV and SIV; protein F for NDV; or VP1 or VP2 for IBV.

***Vectors.*** A suitable vector for incorporating the polynucleotide of nucleic acid construct disclosed herein may be a plasmid vector, viral vector, artificial chromosome, or transposon. Some vectors are attenuated or live virus vectors. Typically, a vector is selected which will efficiently transform or transfect a host cell with the nucleic acid construct disclosed herein under conditions that permit expression of the polyprotein encoded by the construct.

In some embodiments, the vector is a modified live virus vector, such as a modified live pox, herpesvirus, adenovirus, lentivirus, or retrovirus vector.

In some embodiments, the vector comprises DNA or modified DNA, which, optionally, may replicate within a host cell. The vector may be a DNA launched platform that when introduced into a host cell transcribes and translates the nucleic acid construct into RNA and then into the polyprotein; and that optionally comprises an origin of replication and/or a selectable marker.

In other embodiments, the vector may be composed of RNA or modified RNA, which, optionally, replicates within a host cell. Such a vector may further comprise a 5' cap, a 3' UTR and/or a poly A tail. A vector may comprise an RNA launched platform that when introduced into a host cell translates the nucleic acid construct into the polyprotein; and that optionally replicates itself when introduced into the host cell. In some embodiments, the vector comprises self-replicating RNA and that comprises a 5' cap, a nucleic acid sequence encoding alphavirus non-

structural proteins nsP1, nsP2, nsP3 and nsP4, a 26S promoter, the nucleic acid sequence encoding the first polypeptide of interest, the nucleic acid encoding the signal sequence, the nucleic acid encoding the second polypeptide of interest, and a 3'UTR and poly A tail.

Other specific embodiments of vectors include a *Togavirus* vector including but not limited to a SFV, SINV, or VEEV vectors; a *Flaviviridae* vector including but not limited to a *Flavivirus* vector or a *Pestivivus* vector; an *Orthomyxoviridae* vector including but not limited to an *Influenza* virus vector; an *Rhabdoviridae* vector including but not limited to a *Lyssavirus* vector, a *Vesiculovirus* vector, or a *Novirhabdovirus* vector; a *Paramyxoviridae* vector including but not limited to a *Pneumomyxovirinae* vector, a *Pneumovirus* vector, a *Paramyxovirinae* vector, a *Respirovirus* vector, an *Avulavirus* vector, a *Rubulavirus* vector, or a *Morbilivirus* vector; see Mogeler, M.A. & Kamrud, K. I., *RNA-based viral vectors*, EXPERT REV. VACCINES Early online 1-30 (2014)(incorporated by reference).

***Host cells.*** Another aspect of the invention is directed to a host cell transformed with, or transfected with, the nucleic acid constructs or corresponding vectors disclosed herein. Host cells may be those of a non-human animal such as a mammal or avian. In some embodiments, the host cells will be those of, or derived from those of a domesticated animal, such canines, felines, bovines, equines, goats, sheep, pigs, or poultry such as chickens or turkeys. In some embodiments, the host cells may be taken or derived from amphibians, reptiles or fish. In some embodiments, the host cell may be a yeast, fungal, or plant cell or a prokaryotic cell provided these cells are competent to cleave the signal peptide cleavage sites in the polyprotein.

In some embodiments, when the polypeptides of interest are derived from microbial or viral pathogen, the host cell is selected to be one to which the microorganism or virus is tropic. Similarly, the N-terminal and/or the internal signal peptide cleavage sites may be selected from those found in the host cell or in a microbial or viral pathogen.

***Method for simultaneously expressing two, three, or more polypeptides or interest.*** Another aspect of the invention is directed to a method for producing two, three, four, five or more polypeptides of interest comprising transforming or transfecting a vector comprising the nucleic acid construct disclosed herein into a host cell competent to cleave the N-terminal and internal signal peptide domains. The host cell may be one *in vivo*, such is in a tissue or fluid of a subject such as a vaccinated subject, one *ex vivo* taken from a living organisms, or a cell cultivated *in vitro*,

such as a culture cell line such as Chinese Hamster Ovary (CHO) cells, HeLa cells, or cultured *Sporodoptera frugiperda* insect cells.

This method may be used to produce the two or more polypeptides of interest in approximate stoichiometric ratios, such as about 1:1 molar ratio (or 1:1:1, 1:1:1:1, *etc*) which may vary from a predicted stoichiometric ratio such as 1:1 or 1:1:1 by ± 1, 2, 5, 10, 20 mole%. Other stoichiometric ratios of polypeptides of interest may be produced, for example, by encoding the same polypeptide of interest in two positions of a construct expressing three polypeptides of interest. In some embodiments, the co-expressed polypeptides of interest may associate with each other in the ER and form new epitopes or tertiary or quaternary structures.

***Purification of polypeptides of interest.*** Once the polypeptides of interested are released from the polyprotein, they may be further isolated or purified by conventional biochemical techniques including size exclusion, ion exchange, affinity chromatography, hydrophobic interaction, or reversed-phase chromatography.

***Compositions/Vaccines.*** A related aspect of this technology is a composition comprising the vectors disclosed herein and a pharmaceutically acceptable carrier. In one embodiment the composition comprises RNA or DNA comprising the constructs disclosed herein and Lipid InOrganic Nanoparticles (LION). In some embodiments a composition will comprise the nucleic acid construct disclosed herein that encodes a polyprotein encoding two, three or more microbial or viral antigens or immunogenic fragments, and a carrier, adjuvant or excipient in a form suitable for administration to produce a prophylactic or therapeutic response against infection. Such a composition may produce or enhance a humoral or cellular response in an immunized subject against the microbial or viral pathogen from which the polypeptides of interest were derived.

***Therapeutic compositions.*** The nucleic acid constructs, vectors, host cells, and polypeptides of interest may be administered to a subject in need thereof based on the particular subject, the immune response desired, and the type of vaccine or vector. An RNA or DNA construct may be administered in saline, in combination with lipids or polyethylene glycol, aluminum salts or other adjuvants. In one embodiment, the construct or vector is administered in combination with a Lipid InOragnic Nanoparticle.

***LION. Lipid InOrganic Nanoparticle (LION).*** In some embodiments, the platform is administered as a nanoemulsion particle that has a hydrophobic core and comprises a mixture of a liquid oil and one or more inorganic solid nanoparticles. The nanoemulsion particle can also be

referred to herein as Lipid InOrganic Nanoparticles (LIONs). The liquid oil is mixed with the one or more inorganic nanoparticles to form a hydrophobic core. The liquid oil is typically metabolizable. Suitable liquid oil can be a vegetable oil, animal oil, or synthetically prepared oil. In some embodiments, the liquid oil is a fish oil. In some embodiments, the liquid oil is a naturally occurring or synthetic terpenoid. In some embodiments, the liquid oil is squalene, triglyceride (such as capric/caprylic triglyceride or myristic acid triglyceride), vitamin E, lauroyl polyoxylglyceride, monoacylglycerol, soy lecithin, sunflower oil, soybean oil, olive oil, grapeseed oil, or a combination thereof. In one embodiment, the liquid oil is squalene, triglyceride (such as capric/caprylic triglyceride or myristic acid triglyceride), vitamin E, lauroyl polyoxylglyceride, monoacylglycerol, soy lecithin, or a combination thereof. In one embodiment, the liquid oil is squalene, triglyceride (such as capric/caprylic triglyceride or myristic acid triglyceride), sunflower oil, soybean oil, olive oil, grapeseed oil, or a combination thereof. In some embodiments, the liquid oil is squalene (either naturally occurring or synthetic, optionally in combination with any of the above listed liquid oils. The inorganic nanoparticles may be formed from one or more same or different metals (any metals including transition metal), such as from metal salts, metal oxides, metal hydroxides, and metal phosphates. Examples include silicon dioxide ($SiO_2$), iron oxides ($Fe_3O_4$, $Fe_2O_3$, FeO, or combinations thereof), aluminum oxide ($A_{l2}O_3$), aluminum oxyhydroxide (AlO(OH)), aluminum hydroxyphosphate ($Al(OH)_x(PO_4)_y$), calcium phosphate ($Ca_3(PO_4)_2$), calcium hydroxyapatite ($Ca_{10}(PO_4).{}_6(OH)_2$), iron gluconate, or iron sulfate. In some embodiments, the inorganic solid nanoparticle is a metal oxide, such as a transition metal oxide. In one embodiment, the inorganic solid nanoparticle is an iron oxide, for instance, magnetite ($Fe_3O_4$), maghemite (.gamma.-$Fe_2O_3$), wustite (FeO), hematite (α-$Fe_2O_3$), or combinations thereof. In some embodiments, the inorganic solid nanoparticle is a metal hydroxide, such as an aluminum hydroxide or aluminum oxyhydroxide. The inorganic solid nanoparticle may contain a reporter element detectable via imaging methods to allow for imaging and tracking the resulting nanoemulsion particles in the body. For instance, the inorganic solid nanoparticle may contain a reporter element detectable via magnetic resonance imaging (MRI), such as a paramagnetic, superparamagnetic, ferrimagnetic or ferromagnetic compound. Exemplary inorganic solid nanoparticle materials that are MRI-detectable are iron oxides, iron gluconates, and iron sulfates. The inorganic solid nanoparticle typically has an average diameter (number weighted average diameter) ranging from about 3 nm to about 50 nm. For instance, the inorganic solid nanoparticle

can have an average diameter of about 5 nm, about 10 nm, about 15 nm, about 20 nm, about 25 nm, about 30 nm, about 35 nm, about 40 nm, about 45 nm, or about 50 nm. The inorganic solid nanoparticle may be surface modified before mixing with the liquid oil. For instance, if the surface of the inorganic solid nanoparticle is hydrophilic, the inorganic solid nanoparticle may be coated with hydrophobic molecules (or surfactants) to facilitate the miscibility of the inorganic solid nanoparticle with the liquid oil in the "oil" phase of the nanoemulsion particle. Phosphate-terminated lipids (such as phosphatidylated lipids), phosphorous-terminated surfactants, carboxylate-terminated surfactants, sulfate-terminated surfactants, or amine-terminated surfactants can be used for surface modification of the inorganic solid nanoparticle. Typical phosphate-terminated lipids or phosphorous-terminated surfactants are trioctylphosphine oxide (TOPO) or distearyl phosphatidic acid (DSPA). Typical sulfate-terminated surfactants include but not limited to sodium dodecyl sulfate (SDS). Typical carboxylate-terminated surfactants include oleic acid. Typical amine terminated surfactants include oleylamine. In one embodiment, the inorganic solid nanoparticle is a metal oxide such as an iron oxide, and a surfactant, such as oleic acid, oleylamine, SDS, DSPA, or TOPO, is used to coat the inorganic solid nanoparticle, before it is mixed with the liquid oil to form the hydrophobic core. In one embodiment, the inorganic solid nanoparticle is a metal hydroxide, such as an aluminum hydroxide or aluminum oxyhydroxide, and a phosphate-terminated lipid or a surfactant, such as oleic acid, oleylamine, SDS, TOPO or DSPA is used to coat the inorganic solid nanoparticle, before it is mixed with the liquid oil to form the hydrophobic core. The lipids used to form nanoemulsion particles can be cationic lipids, anionic lipids, neutral lipids, or mixtures thereof. In some embodiments, the lipids used are cationic lipids. For example, positively charged lipids that can have favorable interactions with negatively charged bioactive agent (such as DNAs or RNAs) may be used in the nanoemulsion composition. Suitable cationic lipids include 1,2-dioleoyloxy-3-(trimethylammonium)propane (DOTAP); 3.beta.[N(N',N'-dimethylaminoethane)-carbamoyl]cholesterol (DC Cholesterol); dimethyldioctadecylammonium (DDA); 1,2-dimyristoyl-3-trimethylammoniumpropane (DMTAP), dipalmitoyl(C16:0)trimethyl ammonium propane (DPTAP); distearoyltrimethylammonium propane (DSTAP); N-[1-(2,3-dioleyloxy)propyl]-N,N,N-trimethylammonium chloride (DOTMA); N,N-dioleoyl-N,N-dimethylammonium chloride (DODAC); 1,2-dioleoyl-sn-glycero-3-ethylphosphocholine (DOEPC); 1,2-dioleoyl-3-dimethylammonium-propane (DODAP); and 1,2- dilinoleyloxy-3-dimethylaminopropane

(DLinDMA); 1,1'-((2-(4-(2-((2-(bis(2-hydroxydodecyl)amino)ethyl) (2-hydroxydodecyl)amino)ethyl)piperazin-1-yl)ethyl)azanediyl)bis(dodecan-- 2-ol) (C12-200); and combinations thereof. A typical cationic lipid is DOTAP. Other examples for suitable lipids include, but are not limited to, the phosphatidylcholines (PCs), such as distearoylphosphatidylcholine (DSPC), dioleoyl phosphatidylcholine (DOPC), 1-palmitoyl-2-oleoylphosphatidylcholine (POPC), dipalmitoylphosphatidylcholine (DPPC), dipalmitoylphosphatidylcholine (DMPC), *etc.*; phosphatidylethanolamines (PEs), such as 1,2-distearoyl-sn-glycero-3-phosphoethanolamine (DSPE), dioleoylphosphatidylethanolamine (DOPE), etc.; phosphatidylglycerol (PGs); and PEGylated lipids including PEGylated version of any of the above lipids (*e.g.*, DSPE-PEGs). The nanoemulsion particle can further contain one or more surfactants, which can be a hydrophobic surfactant or a hydrophilic surfactant. In some embodiments, the nanoemulsion particle further comprises a hydrophobic surfactant. In some embodiments, the nanoemulsion particle further comprises a hydrophilic surfactant. In one embodiment, the nanoemulsion particle further comprises a hydrophobic surfactant and a hydrophilic surfactant. Suitable hydrophobic surfactants include those having a hydrophilic-lipophilic balance (HLB) value of 10 or less, for instance, 5 or less, from 1 to 5, or from 4 to 5. An exemplary hydrophobic surfactant is a sorbitan ester (such as sorbitan monoester or sorbitan trimester). For instance, the hydrophobic surfactant can be a sorbitan ester having a HLB value from 1 to 5, or from 4 to 5. In some embodiments, the hydrophobic surfactant is a sorbitan monoester or a sorbitan triester. Exemplary sorbitan monoesters include sorbitan monostearate and sorbitan monooleate. Exemplary sorbitan triesters include sorbitan tristearate and sorbitan trioleate. Suitable hydrophilic surfactants include those polyethylene oxide-based surfactants, for instance, a polyoxyethylene sorbitan ester (polysorbate). In some embodiments, the hydrophilic surfactant is a polysorbate. Exemplary polysorbates are polysorbate 80 (polyoxyethylene sorbitan monooleate, or Tween 80), polysorbate 60 (polyoxyethylene sorbitan monostearate, or Tween 60), polysorbate 40 (polyoxyethylene sorbitan monopalmitate, or Tween 40), and polysorbate 20 (polyoxyethylene sorbitan monolaurate, or Tween 20). In one embodiment, the hydrophilic surfactant is polysorbate 80. The nanoemulsion particle can have an oil-to-surfactant molar ratio ranging from about 0.1:1 to about 20:1, from about 0.5:1 to about 12:1, from about 0.5:1 to about 9:1, from about 0.5:1 to about 5:1, from about 0.5:1 to about 3:1, or from about 0.5:1 to about 1:1. The nanoemulsion particle can have a hydrophilic surfactant-to-lipid (*e.g.*, cationic lipid) ratio

ranging from about 0.1:1 to about 2:1, from about 0.2:1 to about 1.5:1, from about 0.3:1 to about 1:1, from about 0.5:1 to about 1:1, or from about 0.6:1 to about 1:1. The nanoemulsion particle can have a hydrophobic surfactant-to-lipid (*e.g.*, cationic lipid) ratio ranging from about 0.1:1 to about 5:1, from about 0.2:1 to about 3:1, from about 0.3:1 to about 2:1, from about 0.5:1 to about 2:1, or from about 1:1 to about 2:1. The nanoemulsion particle can comprise from about 0.2% to about 40% w/v liquid oil, from about 0.001% to about 10% w/v inorganic solid nanoparticle, from about 0.2% to about 10% w/v lipid (*e.g.*, cationic lipid), from about 0.25% to about 5% w/v hydrophobic surfactant (*e.g.*, sorbitan ester), and from about 0.5% to about 10% w/v hydrophilic surfactant. In certain embodiments, the nanoemulsion particle comprises: a hydrophobic core comprising a mixture of one or more inorganic nanoparticles containing at least one metal oxide nanoparticle optionally coated with a phosphate-terminated lipid, a phosphorous-terminated surfactant, a carboxylate-terminated surfactant, a sulfate-terminated surfactant, or an amine-terminated surfactant, and a liquid oil containing naturally occurring or synthetic squalene; a cationic lipid comprising DOTAP; a hydrophobic surfactant comprising a sorbitan ester selected from the group consisting of sorbitan monostearate, sorbitan monooleate, and sorbitan trioleate; and a hydrophilic surfactant comprising a polysorbate. In one embodiment, the nanoemulsion particle comprises: a hydrophobic core comprising a mixture of: one or more inorganic nanoparticles containing iron oxide nanoparticles, and a liquid oil containing naturally occurring or synthetic squalene; the cationic lipid DOTAP; a hydrophobic surfactant comprising sorbitan monostearate; and a hydrophilic surfactant comprising polysorbate 80. In this LION composition, the LION particle can comprise from about 0.2% to about 40% w/v squalene, from about 0.001% to about 10% w/v iron oxide nanoparticles, from about 0.2% to about 10% w/v DOTAP, from about 0.25% to about 5% w/v sorbitan monostearate, and from about 0.5% to about 10% w/v polysorbate 80. In one embodiment, the LION particle comprises from about 2% to about 6% w/v squalene, from about 0.01% to about 1% w/v iron oxide nanoparticles, from about 0.2% to about 1% w/v DOTAP, from about 0.25% to about 1% w/v sorbitan monostearate, and from about 0.5%) to about 5% w/v polysorbate 80. In certain embodiments, the nanoemulsion particle comprises: a hydrophobic core comprising a mixture of: one or more inorganic nanoparticles containing at least one metal hydroxide or oxyhydroxide nanoparticle optionally coated with a phosphate-terminated lipid, a phosphorous-terminated surfactant, a carboxylate-terminated surfactant, a sulfate-terminated surfactant, or an amine-terminated surfactant, and a liquid oil containing naturally

occurring or synthetic squalene; a cationic lipid comprising DOTAP; a hydrophobic surfactant comprising a sorbitan ester selected from the group consisting of sorbitan monostearate, sorbitan monooleate, and sorbitan trioleate; and a hydrophilic surfactant comprising a polysorbate. In one embodiment, the nanoemulsion particle comprises: a hydrophobic core comprising a mixture of one or more inorganic nanoparticles containing aluminum hydroxide or aluminum oxyhydroxide nanoparticles optionally coated with TOPO, and a liquid oil containing naturally occurring or synthetic squalene; the cationic lipid DOTAP; a hydrophobic surfactant comprising sorbitan monostearate; and a hydrophilic surfactant comprising polysorbate 80. In this LION composition, the LION particle can comprise from about 0.2% to about 40% w/v squalene, from about 0.001% to about 10% w/v aluminum hydroxide or aluminum oxyhydroxide nanoparticles, from about 0.2% to about 10% w/v DOTAP, from about 0.25% to about 5% w/v sorbitan monostearate, and from about 0.5% to about 10% w/v polysorbate 80. In one embodiment, the LION particle comprises from about 2% to about 6% w/v squalene, from about 0.01% to about 1% w/v aluminum hydroxide or aluminum oxyhydroxide nanoparticles, from about 0.2% to about 1% w/v DOTAP, from about 0.25% to about 1% w/v sorbitan monostearate, and from about 0.5%) to about 5% w/v polysorbate 80. Nanoparticles and nanoemulsions have been described in the literature and the terms are used herein to refer to those particles having a size less than 1000 nanometers. The nanoemulsion particle (LION) typically has an average diameter (z-average hydrodynamic diameter, measured by dynamic light scattering) ranging from about 20 nm to about 200 nm. In some embodiments, the z-average diameter of the LION particle ranges from about 20 nm to about 150 nm, from about 20 nm to about 100 nm, from about 20 nm to about 80 nm, from about 20 nm to about 60 nm. In some embodiments, the z-average diameter of the LION particle ranges from about 40 nm to about 200 nm, from about 40 nm to about 150 nm, from about 40 nm to about 100 nm, from about 40 nm to about 90 nm, from about 40 nm to about 80 nm, or from about 40 nm to about 60 nm. In one embodiment, the z-average diameter of the LION particle is from about 40 nm to about 80 nm. In one embodiment, the z-average diameter of the LION particle is from about 40 nm to about 60 nm. The average polydispersity index (PDI) of the nanoemulsion particles (LIONs) can range from about 0.1 to about 0.5. For instance, the average PDI of the LION particles can range from about 0.2 to about 0.5, from about 0.1 to about 0.4, from about 0.2 to about 0.4, from about 0.2 to about 0.3, or from about 0.1 to about 0.3.

***Modes of administration.*** In some therapeutic embodiments, the nucleic acid constructs disclosed herein, or vectors comprising them, are formulated for *in vivo* delivery. For administration to a non-human animal, the constructs or vectors according to the present application can be given by any enteral or parenteral route. In some embodiments, the construct or vector is administered subcutaneously, intravenously, intramuscularly, intra-articularly, intra-synovially, intrasternally, intrathecally, intrahepatically, intrathymically, into a sex organ, intralesionally, intracranially, intraventricularly, orally, by inhalation spray, topically, rectally, nasally, buccally, vaginally or via an implanted reservoir.

***Treatment.*** Another aspect of this technology is a method for preventing or treating a microbial or viral disease, its symptoms, comorbidities, or sequelae by administering a vector comprising the construct disclosed herein that a encodes and expresses two, three or more microbial or viral antigens, immunogens, or epitopes. Administration of the vaccine containing the nucleic acid construct or vector may induce innate, humoral, and/or cellular immune responses to the encoded microbial or viral antigens (the polypeptides or interest) including inducing neutralizing antibodies or antigen-specific T cells that recognize cells infected with the microbial or viral pathogens.

***Applications for expressed polypeptides of interest.*** In some embodiments, the polypeptides of interest either alone or in combination with other polypeptides of interest expressed by the same nucleic acid construct may be formulated as diagnostic or therapeutic products, for example, polypeptides of interest derived from a virus may be used to detect antibodies to the virus or to induce humoral or cellular immune response to the virus.

## EXAMPLE 1

*Molecular Construction and Generation of Replicons and Formulation of 6 SIV replicons with HA and NA genes linked by 6K or 6K-like elements*

Replicons comprising nucleic acid constructs encoding portions of the HA and NA genes from an H1N1 SIV (swine influenza virus) isolate are produced. The portions of the constructs encoding the HA and NA genes are separated by sequences encoding a 6K or 6K-like linker.

The small 6K or 6K-like linker elements from the following viruses are used, Venezuelan equine encephalitis virus (VEE), called 6K (V6K); Semliki Forest virus (SFV), also called 6K (S6K); hepatitis C virus (HCV), called p7; bunyavirus (Bunya), called signal peptide (SP, BSP); and hantavirus (Hanta), also called SP (HSP). Similar elements from other viruses may also be used. Many viruses code for 6K or 6K-like sequence that are translated as a portion of a larger polyprotein which is processed into independent monomer proteins.

Each of the six constructions which are described in the priority document comprises a 6K or 6K-like protein sequence separating the two viral antigen coding sequences either neuraminidase (NA) or hemagglutinin (HA). Five of the GOI designs have the SIV HA gene in the upstream (5') location relative to the linker element followed by the NA gene in the downstream (3') location (relative to the linker element). The sixth GOI design has the NA gene in the 5' location and the HA gene in the 3' location, relative to the linker element.

The six constructs respectively incorporate V6K (from Venezuelan equine encephalitis virus), V6K (from Venezuelan equine encephalitis virus), S6K (from Semliki Forest virus (SFV), BSP (from bunyavirus), HSP (from Hantavirus), and p7 (from hepatitis C virus)..

The 5' and 3' ends of a GOI are modified as necessary to contain additional nucleotides to aid in cloning into the vector.

On the 5' end, the EcoRV restriction endonuclease site (GATATC) and Kozak sequence (GCCACC) are added upstream of the 5' gene's start codon.

On the 3' end, the GCTGC sequence and PacI restriction endonuclease site (TTAATTAA) are added downstream of the 3' gene's stop codon.

The GOIs are double-stranded DNA fragments.

*Ligation of the Genes of Interest into the VEE backbone.* The pSRT_463_HA_v2 (pSRT) vector, which contains the Venezuelan equine encephalitis virus nsP1-4 replication machinery genes and a 26S promoter that drives the expression of the GOI are used as the replicon expression vector. The pSRT-463_HA-VA_v2 vector is based on a synthetically modified alpha replicon RNA technology (SMARRT). This vector and technology are described by, and incorporated by reference to Maine, C. J., et al., *Self-Replicating RNAs Drive Protective Anti-tumor T Cell Responses to Neoantigen Vaccine Targets in a Combinatorial Approach,* MOL. THER. 2021, 3; 29(3):1186-1198. In some embodiments of the invention a SMARRT type vector or platform is used to deliver the coding sequence according to the invention.

A GOI and pSRT are both digested with EcoRV and PacI restriction endonucleases to prepare the ends for cloning. Each digestion reaction is purified to eliminate the endonucleases and any unwanted DNA fragments.

A GOI is ligated into the pSRT vector using ElectroLigase®. The products are electroporated into TransforMax™ EPI300™ Electropcometent *E. coli* and resulting colonies screened for the presence of the GOI by isolating the plasmid DNA and analyzing via restriction fragment length polymorphism and next generation sequencing.

Clones that have been ligated correctly with no mutations are made into glycerol stocks for long-term storage.

*Amplification of Replicon DNA.* Samples taken from each glycerol stock for each replicon are struck for isolation on animal origin free LB agar plates, containing the kanamycin selection antibiotic, independently and allowed to replicate overnight. Isolated colonies are lifted and used to inoculate animal origin free Terrific broth cultures containing the kanamycin selection antibiotic. The cultures are incubated overnight at 37°C with shaking to allow replication of the replicon plasmids containing the GOI. The bacterial paste is harvested by centrifugation and the NucleoBond Xtra Maxi EF kit is used to isolate purified endotoxin-free plasmid DNA. This DNA is tested for quantity by OD260, quality by measuring % supercoiled, and identity by Sanger or Lumina DNA Sequencing.

*In vitro Transcription of RNA.* The purified DNA for each replicon is linearized by digesting the DNA with the NotI restriction endonuclease to aid in transcription. The linearized DNA are transcribed into RNA overnight at 30°C using the T7 DNA-dependent RNA polymerase and equal molar rNTPs. Subsequently, the transcription reaction products are treated with DNase I to digest the DNA template. The synthetic RNA product will be purified by centrifugal filtration. The 5' end of the uncapped RNA product are enzymatically capped with Guanyltransferase, rGTP, and SAM. The final capped synthetic RNA product is again purified by centrifugal filtration.

*Formulation.* The capped RNA transcripts of each replicon are encapsulated with Lipid Inorganic Nanoparticles (LION) and formulated according to the study protocol.

## EXAMPLE 2

### *H3 and H1 Linker Experiment Summary*

*H3/H1 6K Experiment Materials and Methods*

*Electroporation.*_Synthetic RNA of each construct--H1, H3, H3H1, H1H3, H3-RVSP-H1, H3-V6K-H1, and H3-p7-H1 --as well as a cell control was electroporated into BHK-21 cells using the Lonza 4D-Nucleofector® X Unit and SF Cell Line Nucleofector® X Kit L (Lonza).

BHK-21 cells were harvested using 0.25% trypsin-EDTA and resuspended in SF Buffer at a cell density of $1x10^6$ cells/100 µL electroporation reaction.

Electroporated cells were recovered in Dulbecco's Modified Eagle's Medium containing 2% fetal bovine serum and distributed into four wells of a 96-well tissue culture plate and one well of a 6-well tissue culture plate.

The cell plates were then incubated at 37 ± 2°C and 5% CO2 for 21.5 hours.

After incubation, the cell plates were removed from the incubator and protein expression was evaluated by both indirect immunofluorescence assay (IFA) and Western blotting.

*IFA.* Cell culture media from the 96-well tissue culture plate was decanted and cells washed twice with 1X PBS followed by fixation for 20 minutes at room temperature using Fix and Perm Medium A solution (Invitrogen).

After fixation, cells were washed three times and blocked with a PBS buffer containing 0.1% tween 20 and 3% FBS for 1 hour at 37 ± 2°C. Cells were washed and incubated with hyperimmune antisera against swine influenza virus (SIV) H1 gamma and H3 IV-B for 1 hour.

After washing, cells were then incubated with a goat anti-swine IgG-FITC conjugate (Invitrogen) for 1 hour and then washed again.

Images of each of the wells were obtained using an EVOS™ FL Digital Inverted Fluorescence Microscope (Invitrogen).

*Protein extraction.* Protein from each of the constructs was extracted using the commercial Mem-PER™ Plus Membrane Protein Extraction Kit (Thermo Fisher Scientific) following manufacturer's instructions.

Secreted, cytosolic and membrane protein fractions were collected and frozen at ≤ -50°C.

*Western blotting (H1 and H3 blots).* Western blots were performed using the membrane protein fraction of each construct. Protein concentrations were obtained by measuring UV Absorbance at 280 nm using the Nanodrop™ One (Thermo Fisher Scientific).

Protein from each construct was denatured under reducing conditions at 70°C for 10 minutes using 10X Bolt™ Sample Reducing Agent and 4X Bolt™ LDS Sample Buffer (Thermo Fisher Scientific).

Denatured samples were loaded into wells of a 4-12% Bis-Tris Plus gel along with a MagicMark™ XP Western Protein Standard (Invitrogen). Two gels were loaded in identical orientation with 20 µg protein loaded per well and run at 200 V for 22 minutes.

Transfer of the proteins to nitrocellulose membranes was performed using the iBlot™ 2 (Thermo Fisher Scientific) gel transfer device.

Blotted membranes were then probed separately with swine antisera, one blot receiving antisera against SIV H1 gamma and the other receiving antisera to SIV H3 IV-B, both detected using a goat anti-swine horseradish peroxidase (HRP) secondary antibody (EMD Millipore) using the iBind™ Flex Western Device (Thermo Fisher Scientific).

After incubation for 2.5 hours at room temperature, the membranes were then washed with deionized water and incubated with SuperSignal™ West Pico PLUS Chemiluminescent Substrate (Thermo Fisher Scientific) for five minutes.

Western blot images were obtained using the C-Digit® Blot Scanner (LI-COR).

*Summary of Results.* In both H1 and H3 Western blots, an approximately 80 kDa size band representing glycosylated hemagglutinin (HA0) was observed for the single gene H1 and H3 constructs when probed against matching antisera, indicating proper protein expression; see arrows in FIG. 6.

The cell control in FIG. 6, last column, did not display fluorescence in the immunofluorescent assay (IFA) nor a band in the Western blot.

The linker constructs displayed variable H1 and H3 expression.

Construct H3-RVSP-H1 displayed relatively high and equivalent expression of both H3 and H1 monomers as observed in the Western blot and IFA. Constructs H3-p7-H1 and H3-V6K-H1 also displayed a good expression of both H3 and H1 monomers.

The H3H1 construct displayed expression of both H1 and H3 proteins in both Western blot and IFA.

An HA0 band was not observed for construct H1H3 in the H1 blot indicating that HA0 was not properly expressed by this construct.

*Terminology.* Terminology used herein is for the purpose of describing particular embodiments only and is not intended to be limiting of the invention.

Unless expressly stated, the terms used herein are intended to have the plain and ordinary meaning as understood by those of ordinary skill in the art.

The following definitions are intended to aid the reader in understanding the present disclosure but are not intended to vary or otherwise limit the meaning of such terms unless specifically indicated.

While aspects of the present disclosure have been described in conjunction with the specific embodiments thereof that are proposed as examples, alternatives, modifications, and variations to the examples may be made. The description and specific examples, while indicating embodiments of the technology, are intended for purposes of illustration only and are not intended to limit the scope of the technology. Moreover, recitation of multiple embodiments having stated features is not intended to exclude other embodiments having additional features, or other embodiments incorporating different combinations of the stated features. Specific examples are provided for illustrative purposes of how to make and use the compositions and methods of this technology and, unless explicitly stated otherwise, are not intended to be a representation that given embodiments of this technology have, or have not, been made or tested.

As used herein, the words "preferred" and "preferably" refer to embodiments of the technology that afford certain benefits, under certain circumstances. However, other embodiments may also be preferred, under the same or other circumstances. Furthermore, the recitation of one or more preferred embodiments does not imply that other embodiments are not useful and is not intended to exclude other embodiments from the scope of the technology.

It should be noted that, as used in the specification and the appended claims, the singular forms “a,” “an,” and “the” include plural referents unless the context clearly dictates otherwise.

As used herein, the term “and/or” includes any and all combinations of one or more of the associated listed items and may be abbreviated as “/”. A and/or B includes A, B, and (A + B).

As used herein in the specification, but not in the claims, including as used in the examples and unless otherwise expressly specified, all numbers may be read as if prefaced by the word “substantially”, “about” or “approximately,” even if the term does not expressly appear. The phrase “about” or “approximately” when used in the specification or claims may be used when describing magnitude and/or position to indicate that the value and/or position described is within a reasonable expected range of values and/or positions. A reasonable range of values based on the disclosure and file history should be used when an express numerical range is not recited.

A numeric value may also be expressly defined by a numeric range, such as a value that is +/- 0.1% of the stated value (or range of values), +/- 0.2% of the stated value (or range of values), +/- 0.5% of the stated value (or range of values), +/- 1% of the stated value (or range of values), +/- 2% of the stated value (or range of values), +/- 5% of the stated value (or range of values), +/- 10% of the stated value (or range of values), +/- 15% of the stated value (or range of values), +/- 20% of the stated value (or range of values), *etc.*

Any numerical range recited herein is intended to include all sub-ranges and values subsumed therein. Where a range of values is provided, it is to be understood that each intervening value between an upper and lower limit of the range and any other stated or intervening value in that stated range is encompassed within the disclosure. Where the stated range includes upper and lower limits, ranges excluding either of those limits are also included.

As referred to herein, the terms "can" and "may" and their variants are intended to be non-limiting, such that recitation that an embodiment can or may comprise certain elements or features does not exclude other embodiments of the present invention that do not contain those elements or features.

Unless otherwise specified, accession numbers, trademarked formulations or products, or commercial products are the last versions available as of the filing date of this application.

All publications and patent applications mentioned in this specification are herein incorporated by reference in their entirety to the same extent as if each individual publication or patent application was specifically and individually indicated to be incorporated by reference, especially referenced is disclosure appearing in the same sentence, paragraph, page, or section of the specification in which the incorporation by reference appears.

The citation of references herein does not constitute an admission that those references are prior art or have any relevance to the patentability of the technology disclosed herein. Any discussion of the content of references cited is intended merely to provide a general summary of assertions made by the authors of the references and does not constitute an admission as to the accuracy of the content of such references.

IN THE CLAIMS

1. A nucleic acid construct comprising polynucleotides coding for a protein comprising an N-terminal signal peptide, a first polypeptide of interest, a second polypeptide of interest and an internal signal peptide sequence positioned between the first polypeptide and the second polypeptide; wherein the internal signal sequence is cleavable by a protease to yield independent first and second polypeptides; preferably the first and second polypeptides of interest comprise HA antigens and/or the internal signal sequence comprises V6k.

2. The nucleic acid construct of claim 1, wherein the construct further encodes an additional internal signal peptide sequence and a third polypeptide of interest positioned after the second polypeptide of interest.

3. The nucleic acid construct of claim 1 or 2, wherein the construct further encodes an additional internal signal peptide sequence and a subsequent polypeptide of interest positioned after the last downstream polypeptide of interest.

4. The nucleic acid construct of claim 1-3, wherein the first, the second and subsequent polypeptides of interest differ from each other.

5. The nucleic acid construct of any one of claims 1-4 that further comprises a promoter or other regulatory region upstream from the polynucleotide encoding the first polypeptide of interest.

6. The nucleic acid construct of claims 1-5, wherein the N-terminal signal peptide directs the protein encoded by the construct into the endoplasmic reticulum.

7. The nucleic acid construct of any one of claims 1-6, wherein the encoded N-terminal signal peptide comprises 15 to 150 residues.

8. The nucleic acid construct of any one of claims 1-7, wherein the encoded N-terminal signal peptide comprises 20-30 residues.

9. The nucleic acid construct of any one of claims 1-8, wherein the encoded internal signal peptide sequence comprises 15 to 75 residues.

10. The nucleic acid construct of any one of claims 1-9, wherein the encoded internal signal peptide sequence comprises less than 20 residues.

11. The nucleic acid construct of any one of claims 1-10, wherein the encoded N-terminal or internal signal peptide sequences comprise a motif comprising:

an N-region comprising 1 to 5 charged residues including but not limited to Arg and Lys,

an H region spanning 7-15 contiguous hydrophobic residues of which 3-5 residues are leucine residues (including but not limited to Leu and Gly or other residues forming an alpha helix),

a C-region comprising 3-7 uncharged amino acid residues (including but not limited to Ala and Val, or other residues forming a beta sheet), and/or

a Pro-region or negative region comprising 1-6 charged amino acid residues. including but not limited to Glu and Ala.

12. The nucleic acid construct of any one of claims 1-11, wherein the encoded N-terminal and/or internal signal peptide regions comprise a cleavage motif of AXA or VXA between a C region and a Pro-region corresponding to a N region.

13. The nucleic acid construct of any one of claims 1-12, wherein the encoded N-terminal and/or internal signal peptide sequence comprises a viral internal cleavage site.

14. The nucleic acid construct of any one of claims 1-13, wherein the encoded N-terminal and/or internal signal peptide sequence comprises all or portion of a 6K polypeptide from a virus in the Alphavirus genus or other viral 6K-like polypeptide.

15. The nucleic acid construct of any one of claims 1-14, wherein the encoded N-terminal and/or internal signal peptide sequence comprises an E2/p7 or p7/NS2 cleavage site of hepatitis C virus.

16. The nucleic acid construct of any one of claims 1-15, wherein the encoded N-terminal and/or internal signal peptide sequence comprises a cleavage site from a virus in the order Bunyavirales.

17. The nucleic acid construct of any one of claims 1-16, wherein the encoded N-terminal and/or internal signal sequence comprises a WAASA motif.

18. The nucleic acid construct of any one of claims 1-17, wherein the first and second encoded polypeptides comprise non-viral or viral polypeptides.

19. The nucleic acid construct of any one of claims 1-18, wherein the first and second encoded polypeptides comprise antigenic or immunogenic polypeptides of a virus that infects domesticated animals.

20. The nucleic acid construct of any one of claims 1-19, wherein the first and second encoded polypeptides are antigenic or immunogenic polypeptides of Infectious Laryngotracheitis virus (ILTV) polypeptides or polypeptides of another virus that infects avians.

21. The nucleic acid construct of any one of claims 1-19, wherein the first and second encoded polypeptides comprise antigenic or immunogenic polypeptides of a virus that infects non-human mammals.

22. A vector comprising the nucleic acid construct of any one of claims 1-21 that is engineered so that can be transformed or transfected into a host cell to express the protein.

23. The vector of claim 22 that is a modified live virus vector.

24. The vector of claim 23 that is a modified live pox, herpesvirus, adenovirus, lentivirus, or retrovirus vector.

25. The vector of claim 22 that is DNA or modified DNA, which, optionally, replicates within a host cell.

26. The vector of claim 22 that is a DNA launched platform that when introduced into a host cell transcribes and translates the nucleic acid construct into the polyprotein; and that optionally comprises an origin of replication and/or a selectable marker.

27. The vector of claim 22 that is RNA or modified RNA, which, optionally, replicates within a host cell.

28. The vector of claim 22 that further comprises a 5' cap, a 5' and 3' untranslated region (UTR) and/or a poly A tail.

29. The vector of claim 22 that is an RNA launched platform that when introduced into a host cell translates the nucleic acid construct into the polyprotein; and that optionally replicates itself when introduced into the host cell.

30. The vector of claim 22 that is self-replicating RNA and that comprises a 5' cap, 5' UTR, a nucleic acid sequence encoding alphavirus non-structural proteins nsP1, nsP2, nsP3 and nsP4, a 26S promoter, the nucleic acid sequence encoding the first polypeptide of interest, the nucleic acid encoding the signal sequence, the nucleic acid encoding the second polypeptide of interest, and a 3'UTR and poly A tail.

31. The vector of claim 22 that is a Togavirus vector including but not limited to a SFV, SINV, or VEEV vector.

32. The vector of claim 22 that is a Flaviviridae vector including but not limited to a Flavivirus vector or a Pestivivus vector.

33. The vector of claim 22 that is an Orthomyxoviridae vector including but not limited to an Influenza virus vector.

34. The vector of claim 22 that is an Rhabdoviridae vector including but not limited to a Lyssavirus vector, a Vesiculovirus vector, or a Novirhabdovirus vector.

35. The vector of claim 22 that is a Paramyxoviridae vector including but not limited to a Pneumomyxovirinae vector, a Pneumovirus vector, a Paramyxovirinae vector, a

Respirovirus vector, an Avulavirus vector, a Rubulavirus vector, a Morbilivirus vector, Picornaviridae family vector, Astroviridae family vector, and Nidovirales order vector.

36. A host cell transformed with, or transfected with, the vector of any one of claims 22-35

37. The host cell of claim 36 that is a cell of a non-human animal.

38. The host cell of claim 36 that is a host cell from a mammal.

39. The host cell of claim 36 that is an avian host cell.

40. The host cell of claim 36 to which a virus or other microorganism encoding the domain of the first, second, third, or subsequent polypeptide of interest is tropic.

41. The host cell of claim 36 that is a bacterial, yeast, insect, mammalian or plant cell or a culture cell derived therefrom.

42. A protein encoded by the nucleic acid construct of any one of claims 1-21.

43. A first and second polypeptide of interest produced by cleavage of the protein encoded by the construct of claim 1.

44. A first, second, and third polypeptide of interest produced by cleavage of the protein encoded by the construct of claim 2.

45. A first, second, third and any subsequent polypeptides of interest produced by cleavage of the protein encoded by the construct of claim 3.

46. A method for producing two or more polypeptides of interest comprising transforming or transfecting a vector of any one of claims 21-34 into a host cell competent to cleave the N-terminal and internal signal peptide domains.

47. The method of claim 46 that produces similar or 1:1 ± 20 mole% stoichiometric amounts of the first, second, third or subsequent polypeptides of interest.

FIG. 1

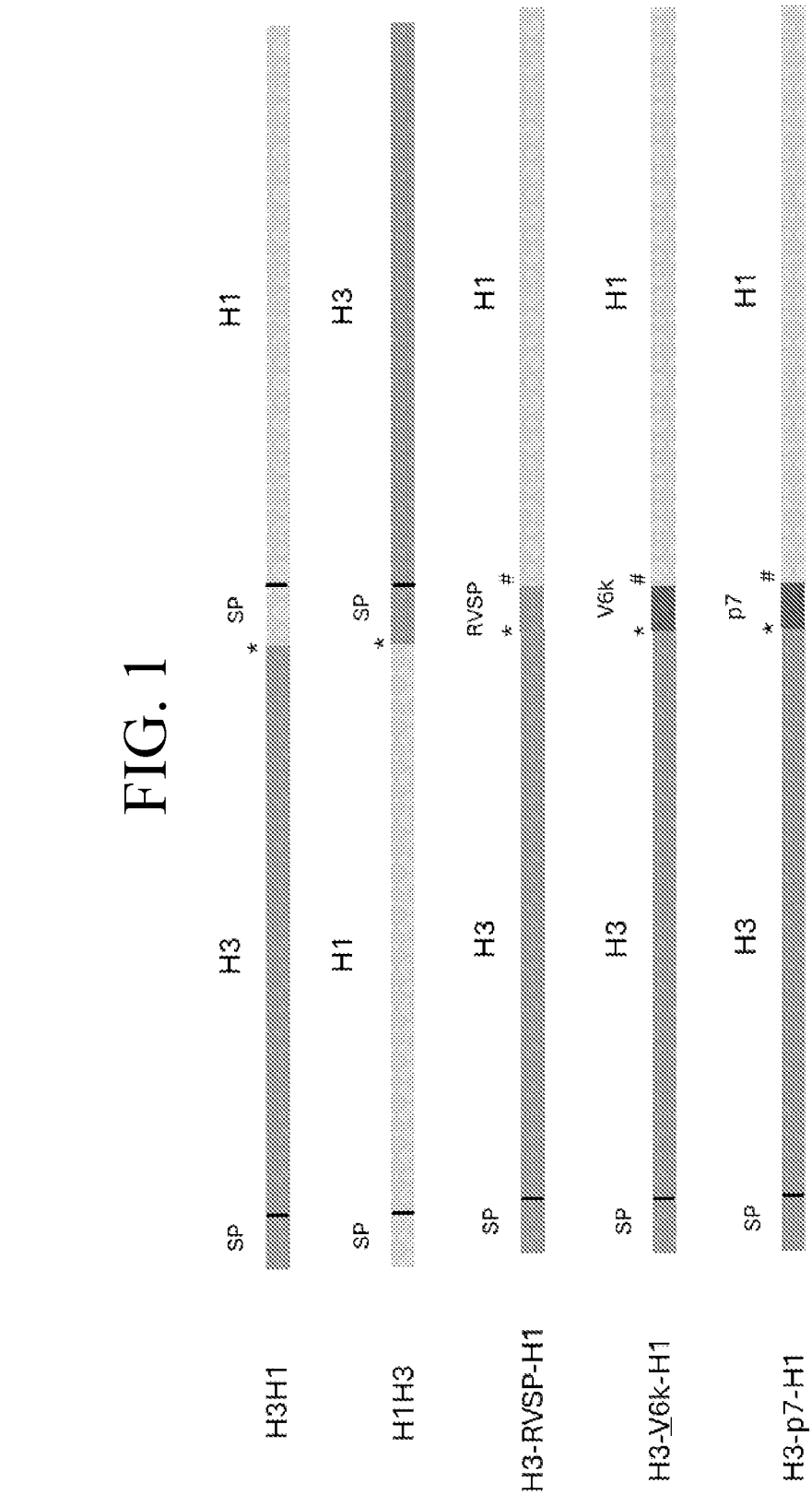

H3H1
SP
H3
*
SP
H1
H1H3
SP
H1
*
SP
H3
H3-RVSP-H1
SP
H3
*
#
RVSP
H1
H3-V6k-H1
SP
H3
*
#
V6k
H1
H3-p7-H1
SP
H3
*
#
p7
H1
SP = Native Signal Peptide
* = No stop codon
# = No H1 SP

FIG. 2

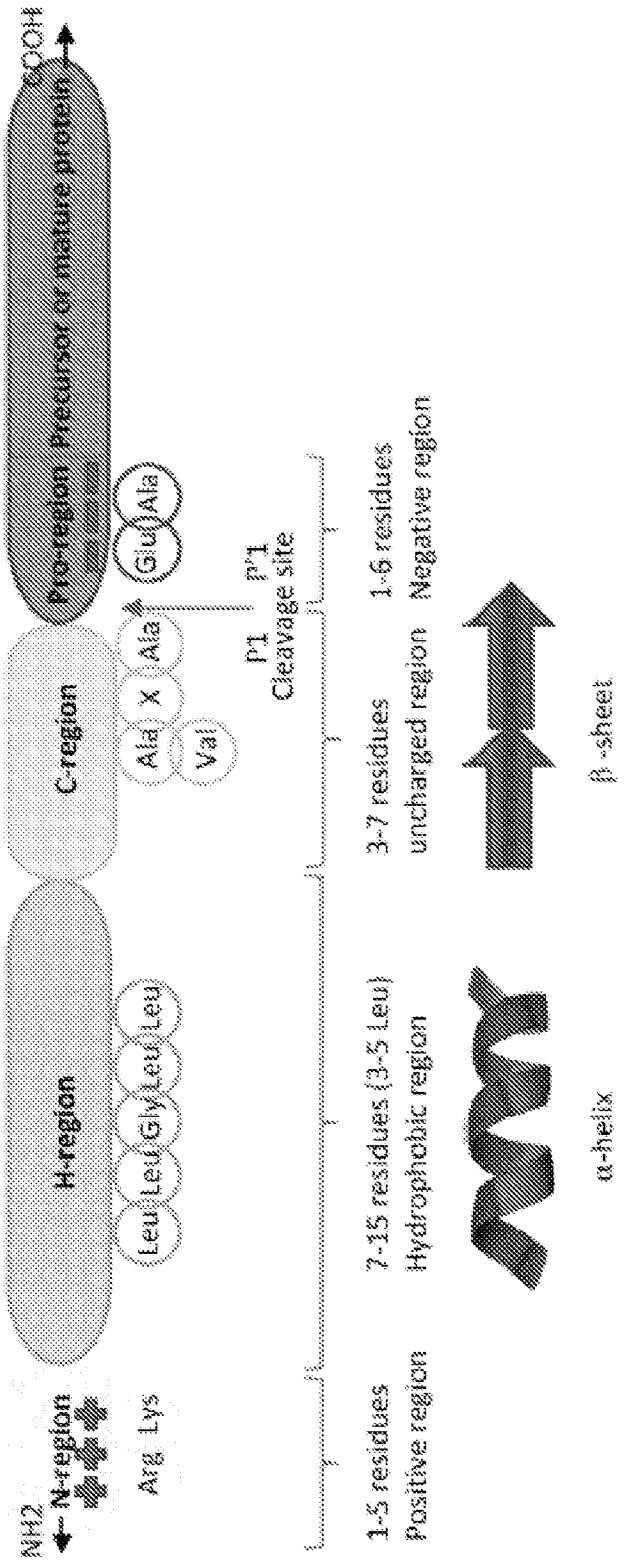
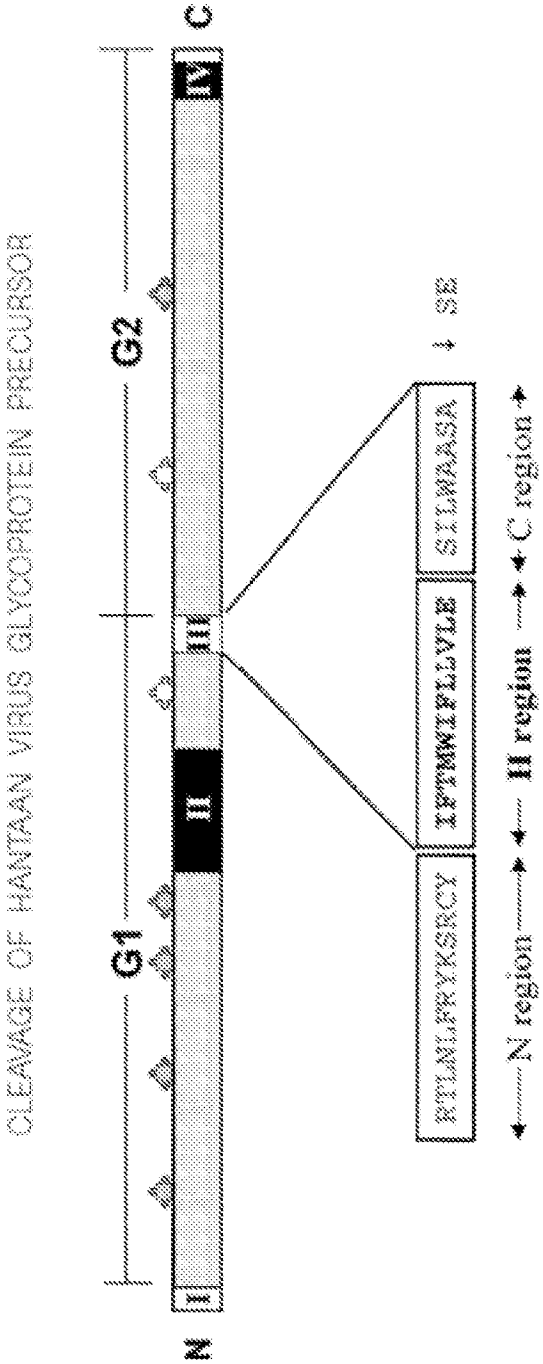

NH2
N-region
Arg Lys
H-region
Leu Leu Gly Leu Leu
C-region
Ala X Ala
Val
Glu Ala
P1
Cleavage site
P'1
Pro-region Precursor or mature protein
COOH
1-5 residues
Positive region
7-15 residues (3-5 Leu)
Hydrophobic region
3-7 residues
uncharged region
1-6 residues
Negative region
α-helix
β -sheet

FIG. 3A


CLEAVAGE OF HANTAAN VIRUS GLYCOPROTEIN PRECURSOR
N
C
G1
G2
I
II
III
IV
RTLNLFRYKSRCY
IPTMWTPLSVLE
STLWAASA
SS
N region
II region
C region

# FIG. 3B

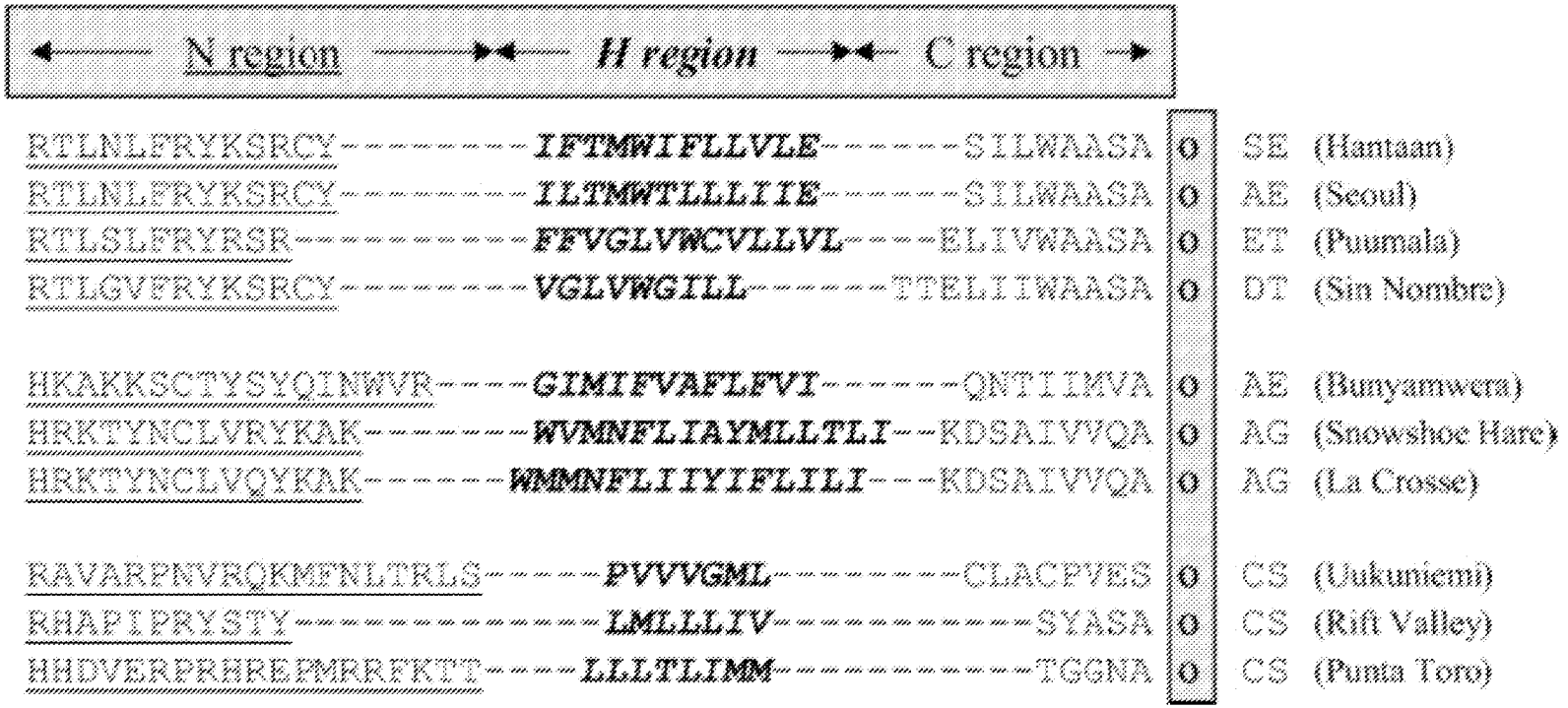
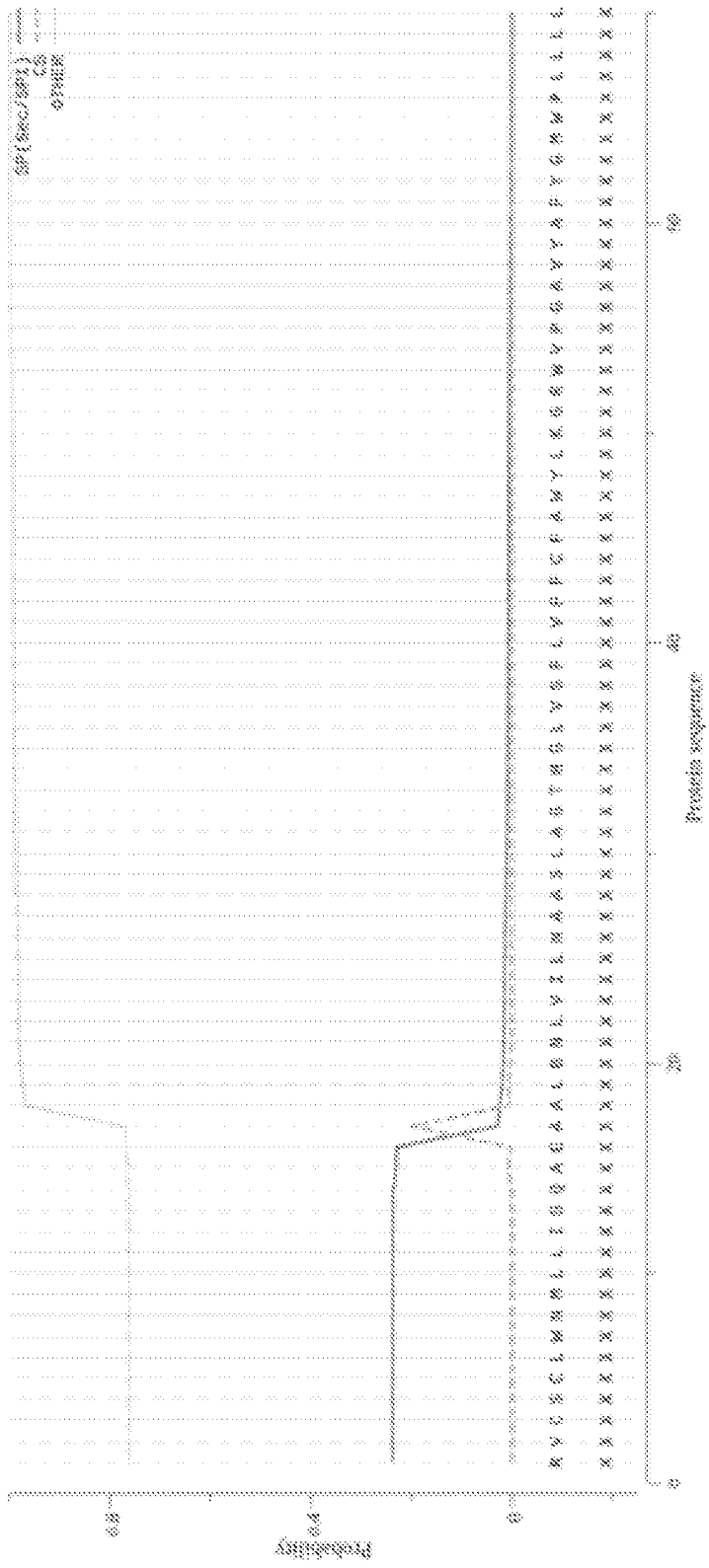

N region
H region
C region
RTLNLFRYKSRCY---------IFTMWIFLLVLE-------SILWAASA 0 SE (Hantaan)
RTLNLFRYKSRCY---------ILTMWTLLLIIE-------SILWAASA 0 AE (Seoul)
RTLSLFRYRSR-----------FFVGLVWCVLLVL----ELIVWAASA 0 ET (Puumala)
RTLGVFRYKSRCY---------VGLVWGILL--------TTELIIWAASA 0 DT (Sin Nombre)
HKAKKSCTYSYQINWVR-----GIMIFVAFLFVI-------QNTIIMVA 0 AE (Bunyamwera)
HRKTYNCLVRYKAK---------WVMNFLIAYMLLTLI--KDSAIVVQA 0 AG (Snowshoe Hare)
HRKTYNCLVQYKAK--------WMMNFLIIYIFLILI---KDSAIVVQA 0 AG (La Crosse)
RAVARPNVRQKMFNLTRLS-----PVVVGML---------CLACPVES 0 CS (Uukuniemi)
RHAPIPRYSTY-------------LMLLLIV-----------SYASA 0 CS (Rift Valley)
HHDVERPRHREPMRRFKTT-----LLLTLIMM-----------TGGNA 0 CS (Punta Toro)

FIG. 4A

HCV strain H p7

E2 p7 NS2

RVCSCLWMMLLISQAEA AIENLVILNAASLAGTHGLVSFLVFFCFAWYLKGRWVPGAVYAFYGMWPLLLLLALPQRAYA LDTEV




From Lin et al 1994

FIG. 4B

Bunyamwera

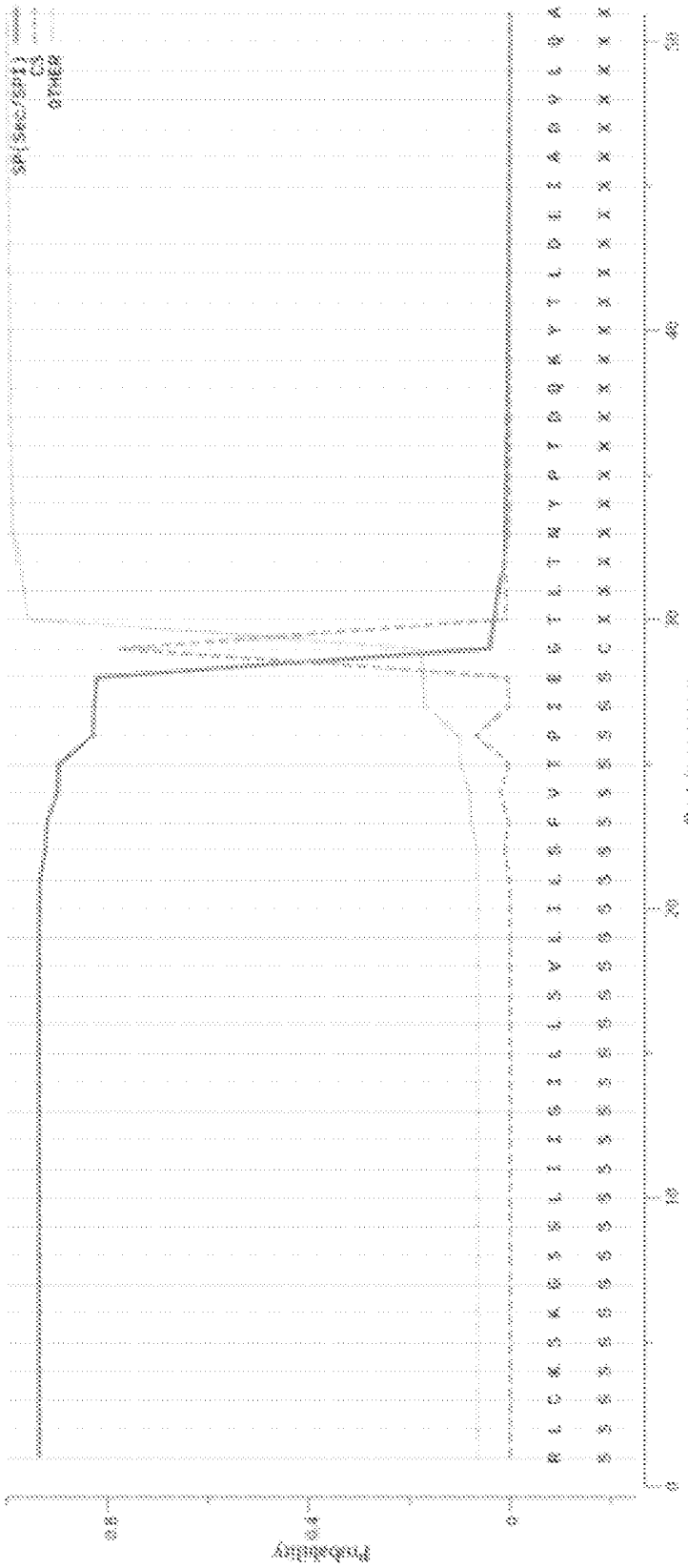


FIG. 5A

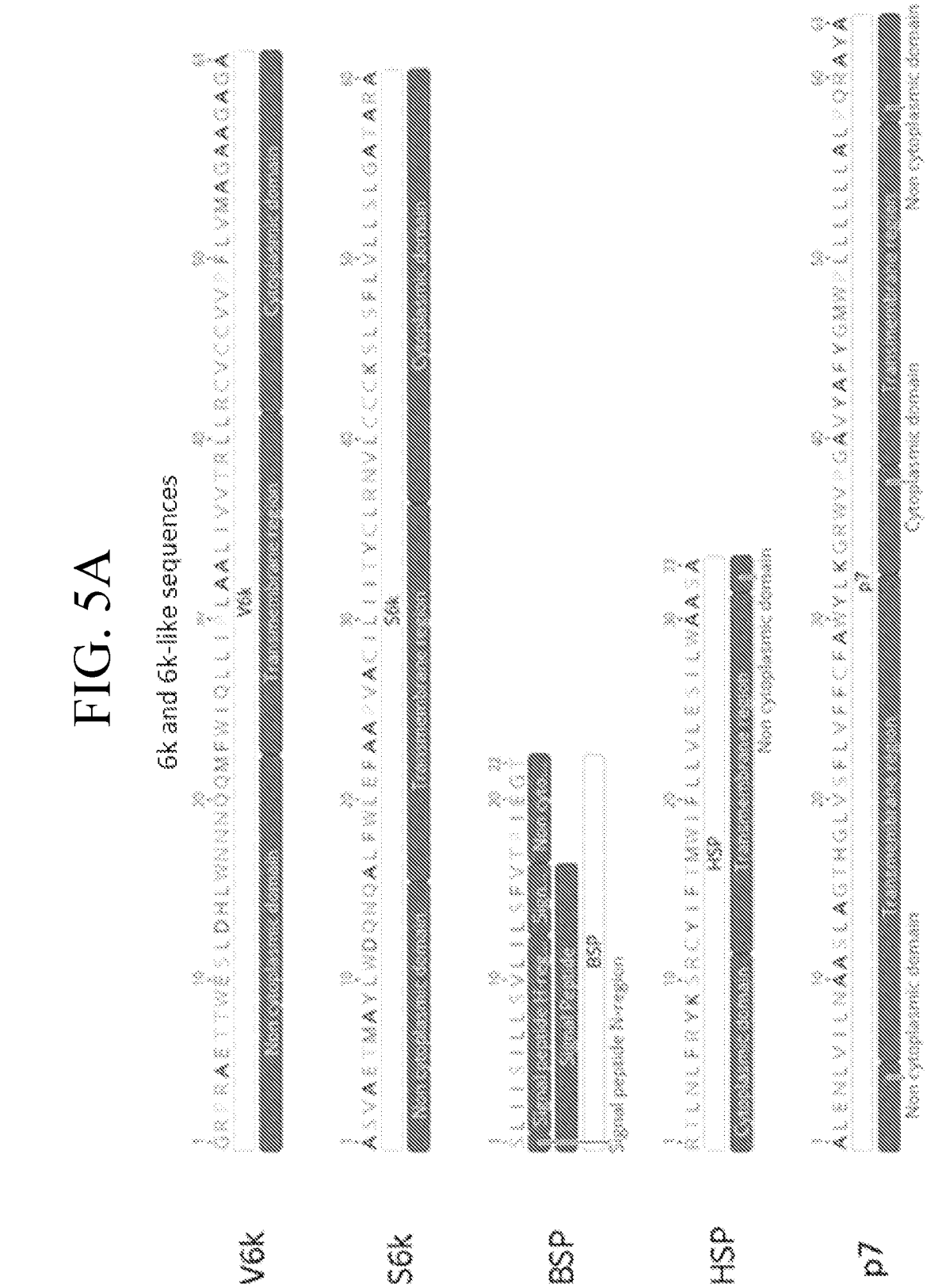

6k and 6k-like sequences
V6k
S6k
BSP
HSP
p7
Signal peptide hydrophobic
Non cytoplasmic domain
Cytoplasmic domain

FIG. 5B

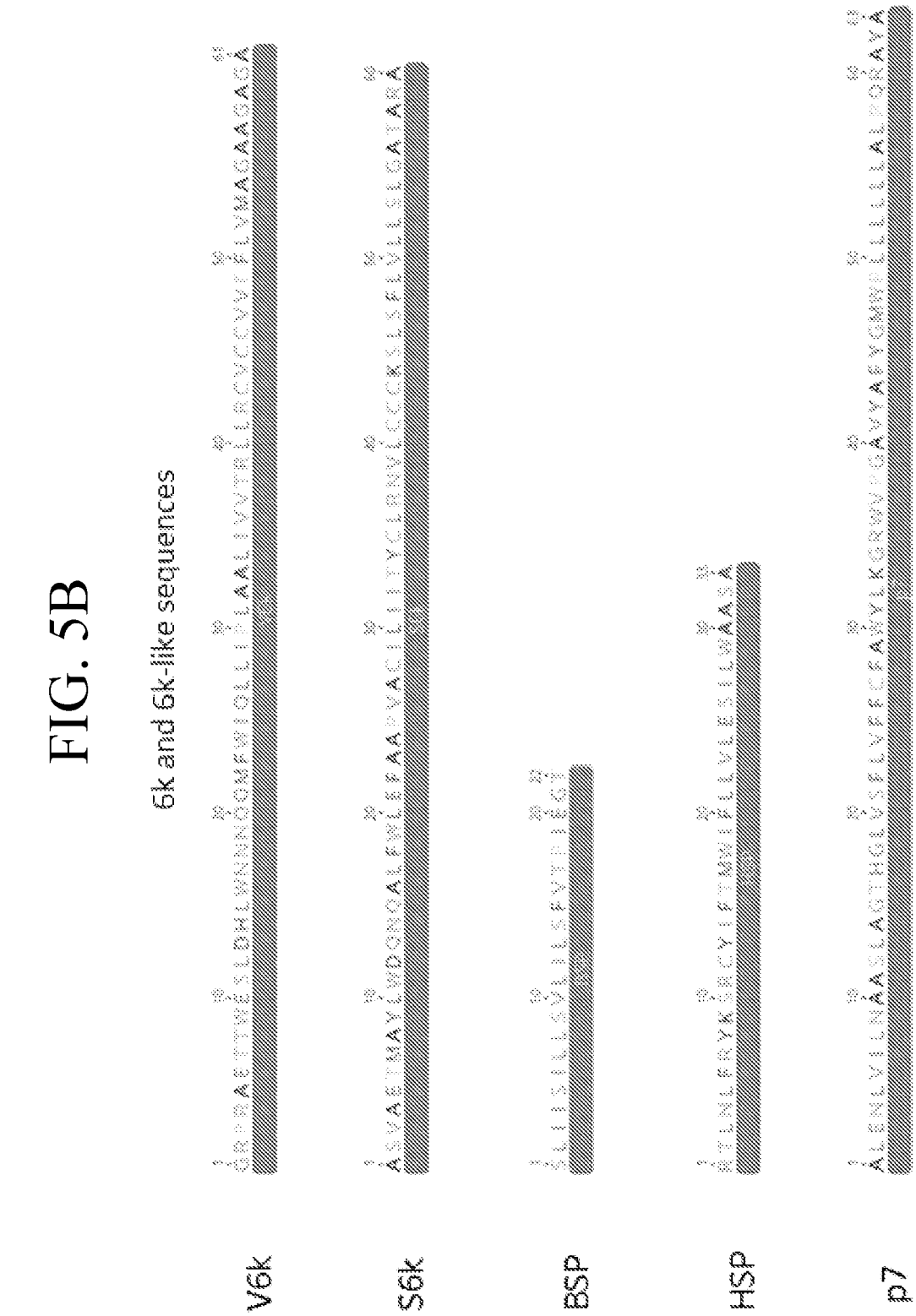
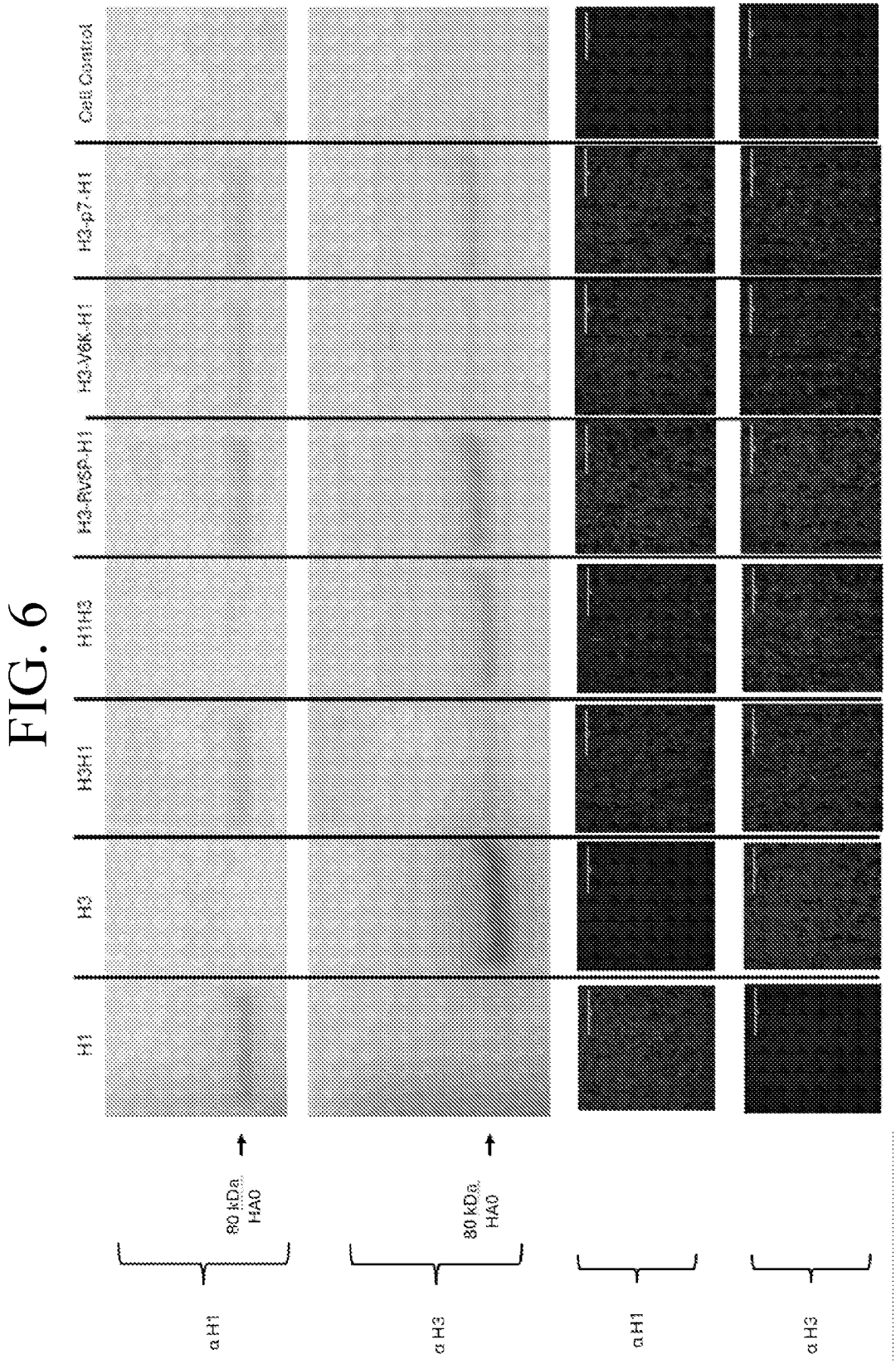


FIG. 6


H1
H3
H3H1
H1H3
H3-PVSP-H1
H3-V6K-H1
H3-p7-H1
Cell Control
80 kDa HA0
80 kDa HA0
α H1
α H3
α H1
α H3

# INTERNATIONAL SEARCH REPORT

International application No
PCT/US2024/034001

**A. CLASSIFICATION OF SUBJECT MATTER**

INV. A61K39/12 A61P31/22
ADD. A61K39/00

According to International Patent Classification (IPC) or to both national classification and IPC

**B. FIELDS SEARCHED**

Minimum documentation searched (classification system followed by classification symbols)

A61K A61P

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

EPO-Internal, FSTA, CHEM ABS Data, BIOSIS, EMBASE, WPI Data

**C. DOCUMENTS CONSIDERED TO BE RELEVANT**

| Category* | Citation of document, with indication, where appropriate, of the relevant passages | Relevant to claim No. |
|---|---|---|
| X | SARASATE EIAMTANASATE ET AL: "C-terminal hydrophobic region leads PRSV P3 protein to endoplasmic reticulum", VIRUS GENES, KLUWER ACADEMIC PUBLISHERS, BO, vol. 35, no. 3, 13 June 2007 (2007-06-13), pages 611-617, XP019557726, ISSN: 1572-994X, DOI: 10.1007/S11262-007-0114-Z figure 1 page 612, left-hand column, paragraph 1 ----- -/-- | 1-47 |

[x] Further documents are listed in the continuation of Box C.

[x] See patent family annex.

* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance;; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance;; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

| Date of the actual completion of the international search | Date of mailing of the international search report |
|---|---|
| 10 October 2024 | 11/11/2024 |

| Name and mailing address of the ISA/ | Authorized officer |
|---|---|
| European Patent Office, P.B. 5818 Patentlaan 2<br>NL - 2280 HV Rijswijk<br>Tel. (+31-70) 340-2040,<br>Fax: (+31-70) 340-3016 | Schmitz, Till |

1

Form PCT/ISA/210 (second sheet) (April 2005)

# INTERNATIONAL SEARCH REPORT

International application No
PCT/US2024/034001

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

| Category* | Citation of document, with indication, where appropriate, of the relevant passages | Relevant to claim No. |
|---|---|---|
| X | TATINENI S ET AL: "Efficient and stable expression of GFP through Wheat streak mosaic virus-based vectors in cereal hosts using a range of cleavage sites: Formation of dense fluorescent aggregates for sensitive virus tracking", VIROLOGY, ELSEVIER, AMSTERDAM, NL, vol. 410, no. 1, 5 February 2011 (2011-02-05), pages 268-281, XP027582695, ISSN: 0042-6822 [retrieved on 2011-01-04] figure 2A(b) | 1-47 |
| X | YAO J S ET AL: "Interactions between PE2, E1 and 6K required for assembly of alphaviruses studied with chimeric viruses", JOURNAL OF VIROLOGY, THE AMERICAN SOCIETY FOR MICROBIOLOGY, US, vol. 70, no. 11, 1 November 1996 (1996-11-01), pages 7910-7920, XP002964147, ISSN: 0022-538X figure 1 | 1-47 |
| A | US 2018/311338 A1 (HENDERSON DANIEL R [US]) 1 November 2018 (2018-11-01) figure 1 | 1-47 |
| A | MORIETTE CORALIE ET AL: "RECOVERY OF A RECOMBINANT SALMONID ALPHAVIRUS FULLY ATTENUATED AND PROTECTIVE FOR RAINBOW TROUT", JOURNAL OF VIROLOGY, THE AMERICAN SOCIETY FOR MICROBIOLOGY, US, vol. 80, no. 8, 1 April 2006 (2006-04-01), pages 4089-4098, XP009076629, ISSN: 0022-538X, DOI: 10.1128/JVI.80.8.4088-4098.2006 figure 6 | 1-47 |
| A | KENNETH LUNDSTROM: "Self-Replicating RNA Viruses for RNA Therapeutics", MOLECULES, vol. 23, no. 12, 13 December 2018 (2018-12-13), page 3310, XP055682865, DE ISSN: 1433-1373, DOI: 10.3390/molecules23123310 cited in the application the whole document section 2.1, Fig. 1, Tables 1-3, p. 5-6 | 1-47 |

1

# INTERNATIONAL SEARCH REPORT

International application No.
PCT/US2024/034001

**Box No. I Nucleotide and/or amino acid sequence(s) (Continuation of item 1.c of the first sheet)**

1. With regard to any nucleotide and/or amino acid sequence disclosed in the international application, the international search was carried out on the basis of a sequence listing:

   a. [X] forming part of the international application as filed.

   b. [ ] furnished subsequent to the international filing date for the purposes of international search (Rule 13*ter*.1(a)).

      [ ] accompanied by a statement to the effect that the sequence listing does not go beyond the disclosure in the international application as filed.

2. [ ] With regard to any nucleotide and/or amino acid sequence disclosed in the international application, this report has been established to the extent that a meaningful search could be carried out without a WIPO Standard ST.26 compliant sequence listing.

3. Additional comments:

# INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No
PCT/US2024/034001

| Patent document cited in search report | Publication date | Patent family member(s) | Publication date |
|---|---|---|---|
| US 2018311338 A1 | 01-11-2018 | AU 2017207744 A1 | 26-07-2018 |
| | | AU 2022201834 A1 | 07-04-2022 |
| | | BR 112018014109 A2 | 11-12-2018 |
| | | CA 3010974 A1 | 20-07-2017 |
| | | CN 108778365 A | 09-11-2018 |
| | | CN 113181106 A | 30-07-2021 |
| | | EP 3402549 A1 | 21-11-2018 |
| | | JP 2019511255 A | 25-04-2019 |
| | | US 2017196966 A1 | 13-07-2017 |
| | | US 2018311338 A1 | 01-11-2018 |
| | | WO 2017123652 A1 | 20-07-2017 |