In some embodiments, the present invention provides influenza hemagglutinin ("HA") polypeptides, proteins, and protein complexes that comprise a stalk domain that is engineered to facilitate maintenance of its native trimeric conformation, even if the head domain of the HA protein is removed or disrupted. In some embodiments, the present invention provides compositions comprising such polypeptides, proteins, and protein complexes, and methods of use of such proteins and compositions, for example as vaccine immunogens.
INFLUENZA HEMAGGLUTININ PROTEINS AND METHODS OF USE THEREOF

CROSS-REFERENCE TO RELATED APPLICATIONS
[0001] This application claims the benefit of U.S. Provisional Patent Application No. 61/861,989, filed August 3, 2013, the contents of which are hereby incorporated by reference.

SEQUENCE LISTING
[0002] The instant application contains a Sequence Listing which has been submitted electronically in ASCII format and is hereby incorporated by reference in its entirety. Said ASCII copy, created on August 1, 2014, is named Avatar_006_WO1_Sequence_Listing.txt and is 414,314 bytes in size.

COPYRIGHT AND INCORPORATION BY REFERENCE
[0003] A portion of the disclosure of this patent document contains material that is subject to copyright protection. The copyright owner has no objection to the facsimile reproduction by anyone of the patent document or the patent disclosure as it appears in the Patent and Trademark Office patent file or records, but otherwise reserves all copyright rights whatsoever.

[0004] For the purposes of only those jurisdictions that permit incorporation by reference, the text of all documents cited herein is hereby incorporated by reference in its entirety.

BACKGROUND OF THE INVENTION
[0005] The US and world populations continue to be at risk of a pandemic influenza outbreak, analogous to the 1918 Spanish (H1N1) outbreak that killed more than 50 million people. Similarly, weaponized influenza virus remains a major bio-warfare threat. Furthermore, antigenic drift requires individuals seeking protection against influenza to be vaccinated annually, and recent studies have shown that seasonal vaccine products are only weakly efficacious if a mismatch occurs between vaccination strains and circulating strains.

[0006] The development of an effective universal influenza vaccine that provides protection across strains of influenza virus would be of enormous value. Evidence that antibodies specific for the conserved stalk domain of the influenza HA protein can protect
against infection has prompted a concerted effort to identify additional and better monoclonal antibodies, and to develop a protective vaccine to address this significant unmet medical and public health need.

SUMMARY OF THE INVENTION

[0007] Some aspects of the present invention are summarized below. Additional aspects are described in the Detailed Description of the Invention, the Examples, the Figures and the Claims sections of the present patent application.

[0008] The influenza HA protein is known to induce potent neutralizing antibodies that correlate with protection against influenza virus infection. Most existing influenza virus vaccines provide protection based on the generation of antibodies against the highly variable, immunodominant, head domain of the influenza HA protein. However, the head domain is often strain-specific so such vaccines are generally only effective against homologous influenza strains, and do not provide protection against other forms of influenza virus, such as homologous drift variants and heterologous strains. Recently it has been shown that the stalk domain of influenza HA can elicit antibodies that react across influenza virus subtypes, due to the more conserved structure of the stalk domain and the presence of epitopes presented on the conserved stalk. Also, potent neutralizing antibodies (nAbs) have been isolated that specifically bind to the native trimeric conformation of the stalk domain. However, the stalk domain becomes highly unstable and readily transitions to a non-native conformation or disassembles upon removal of the HA head domain - limiting usefulness of the stalk domain on its own (e.g. without the head domain) as a vaccine immunogen. An influenza HA protein having a stalk domain stabilized in its native trimeric conformation could be very valuable – providing a candidate influenza vaccine immunogen capable of providing protection across influenza virus strains. Similarly, such a stabilized influenza HA protein could also be useful for the generation of antibodies, such as diagnostic and therapeutic antibodies.

[0009] Based on an extensive analysis of the structure of the influenza HA protein, the present invention provides a variety of novel design strategies and novel constructs to stabilize or “lock” the stalk domain of the influenza HA protein in its native trimeric conformation. The present invention also provides a variety of engineered influenza HA
polypeptides, proteins, and/or protein complexes, such as those that comprise one or more targeted cross-links (such as di-tyrosine cross-links), one or more to-tyrosine mutations, and/or one or more artificially-introduced protease cleavage sites/motifs. The engineered HA influenza HA polypeptides, proteins, and/or protein complexes of the invention can be made using any suitable influenza HA polypeptide or protein as a starting point. For example, an influenza HA sequence from any influenza type, sub-type, or strain can be used as a starting point for generation of the engineered products described herein. In many of the embodiments described herein, the influenza strain Puerto Rico/8/1934 or “PR8” (which is a strain of the H1N1 influenza subtype of influenza A) was used as the starting point. The amino acid sequence of a wild-type PR8 strain is provided in Figure 9 (SEQ ID NO: 1). However, any other influenza HA sequence from any other influenza type, sub-type, or strain could equally be used. Non-limiting examples of other influenza HA sequences that can be used as the starting point for generating the engineered HA products described herein include, but are not limited to, those illustrated in Figures 55, 56, 57, 58, 59, and 60, and those having the sequences of SEQ ID NO:s 80, 81, 82, 83, 84, 85, 111, 112, 113, 114, and 115. Similarly, codon optimized versions of the nucleotide sequences that encode influenza HA proteins can be used as starting points for the generation of the engineered HA products described herein. Non-limiting examples of codon-optimized HA sequences from the PR8 influenza strain include those having the sequences of SEQ ID NO:s 63, 64, 65, 66, 67, and 68.

[0010] In some embodiments, the present invention provides influenza HA polypeptides, proteins, and/or protein complexes that comprise one or more targeted cross-links in their stalk domain which serve to stabilize or “lock” the stalk domain in its native trimeric conformation. In some embodiments such targeted cross-links are di-tyrosine cross-links. In some embodiments, the present invention provides an influenza HA protein complex comprising a trimeric stalk domain formed by the association of three protomers, wherein the stalk domain comprises one or more targeted cross-links, such as di-tyrosine cross-links, that stabilize the stalk domain in its native trimeric conformation. In some such embodiments, the influenza HA protein complex further comprises one or more cross-links in the influenza HA head domain. In some such embodiments, the influenza HA protein complex does not comprise an intact head domain. In embodiments where di-tyrosine cross-links are used, such cross-links can be made between two tyrosine residues that are naturally
present in an HA polypeptide, protein, and/or protein complex, or between two tyrosine residues that have been introduced by mutation, or between a first tyrosine residue that is naturally present in an HA polypeptide, protein, and/or protein complex and a second tyrosine residue that has been introduced by mutation. In some embodiments, the present invention also provides influenza HA polypeptides, proteins, and/or protein complexes that comprise one or more “to-tyrosine” mutations in the HA stalk domain at locations that have been determined to be desirable locations for the formation of di-tyrosine cross-links to stabilize the stalk domain in its native trimeric conformation. In some embodiments, the influenza HA polypeptides, proteins, and/or protein complexes of the invention (whether containing targeted cross-links (such as di-tyrosine cross-links), or to-tyrosine mutations, or both) are full length HA proteins comprising both the HA stalk domain (with or without the signal peptide) and the HA head domain, and optionally also the HA transmembrane domain. In some embodiments the influenza HA polypeptides, proteins, and/or protein complexes of the invention lack one or more of the HA head domain, the transmembrane domain, and/or the signal peptide. In some embodiments the influenza HA polypeptides, proteins, and/or protein complexes of the invention comprise the HA stalk domain, or at least a portion of the HA stalk domain that is sufficient to assemble into, or form a part of, the normal trimeric stalk conformation. Thus, in some embodiments, it may be possible to remove, add, or substitute certain HA stalk domain amino acids without compromising the ability of the HA polypeptide or protein to assemble into its trimeric conformation.

[0011] In some embodiments the present invention provides influenza HA polypeptides, proteins, and/or protein complexes that comprise to-tyrosine mutations at one or more amino acid positions 403, 406, 411, 422, 429, 432, 433, and 435, where such amino acid numbering is based upon the sequence shown in Figure 9 (SEQ ID NO: 1), or at amino acid positions that correspond to such amino acid positions, for example as determined by alignment of an HA amino acid sequence to SEQ ID NO: 1. Non-limiting examples of influenza HA amino acid sequences that comprise one or more of such to-tyrosine mutations include SEQ ID NOs: 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 86, 87, 88, 89, 90, 91, 92, 93, 96, 97, 98, 99, 100, 101, 102, 103, 104, 105, 106, 107, 108, 109 and 110. In some embodiments the present invention provides influenza HA polypeptides, proteins, and/or protein complexes that comprise di-tyrosine cross-links between one or more pairs of amino acids selected from the following amino acid positions: 308, 403, 406, 411, 422, 429, 432,
433, 435, and 437, where such amino acid numbering is based upon the sequence shown in Figure 9 (SEQ ID NO: 1), or at amino acid positions that correspond to such amino acid positions, for example as determined by alignment of an HA amino acid sequence to SEQ ID NO: 1.

[0012] In some embodiments, the present invention provides influenza HA polypeptides, proteins, and/or protein complexes that comprise one or more artificially-introduced protease cleavage sites that can be used to proteolytically remove the head domain of an HA polypeptide, protein, and/or protein complex. In some embodiments the present invention provides influenza HA polypeptides, proteins, and/or protein complexes that comprise one or more artificially-introduced protease cleavage sites inserted after (e.g. immediately after) amino acid positions 48, 63, 228, 278, 282, 283, 286, and 291, where such amino acid numbering is based upon the sequence shown in SEQ ID NO: 1, or at amino acid positions that correspond to such amino acid positions, for example as determined by alignment of an HA amino acid sequence to sequence ID NO: 1. Non-limiting examples of influenza HA amino acid sequences that comprise one or more of such artificially-introduced protease cleavage sites include SEQ ID NOs: 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, and 30.

[0013] In some embodiments the present invention provides influenza HA polypeptides, proteins, and/or protein complexes that comprise at least one pair of artificially-introduced protease cleavage sites, such that cleavage at both of the pair of cleavage sites will result in removal of the HA head domain. Non-limiting examples of influenza HA amino acid sequences that comprise a pair of such artificially-introduced protease cleavage sites include SEQ ID NOs: 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 26, 27, 28, 29, and 30. In some such embodiments, where there are a pair of artificially-introduced protease cleavage sites, the first such protease cleavage site is inserted after (e.g. immediately after) amino acid position 48 or 63, and the second such protease cleavage site is inserted after (e.g. immediately after) amino acid position 228, 278, 282, 283, 286, or 291, where such amino acid numbering is based upon the sequence shown in SEQ ID NO: 1, or at amino acid positions that correspond to such amino acid positions, for example as determined by alignment of an HA amino acid sequence to sequence ID NO: 1.
In some embodiments, the present invention also provides influenza HA stalk domain polypeptides, proteins, and/or protein complexes that do not comprise an intact HA head domain, such as those generated by proteolytic removal of the influenza HA head domain, for example by cleavage at one or more of the artificially-introduced protease cleavage sites described herein. The stalk domain sequences of influenza HA are discontinuous because the HA protein comprises an N-terminal region comprising stalk domain sequences, followed by a middle region comprising head domain sequences, followed by a C-terminal region comprising additional stalk domain sequences. Accordingly, in some embodiments, proteolytic cleavage/removal of the HA head domain results in the generation of two stalk domain polypeptide fragments – an N-terminal fragment and a C-terminal fragment. In some embodiments the present invention provides such N- and C-terminal stalk domain polypeptides, and/or polypeptides, proteins, or protein complexes that comprise such N- and C-terminal stalk domain polypeptides. In some embodiments such N- and C-terminal stalk domain polypeptides are present in an HA stalk domain protein complex having a native trimeric stalk domain conformation. Non-limiting examples of influenza HA N-terminal stalk domain polypeptides include SEQ ID NOs: 94 and 95. Non-limiting examples of influenza HA C-terminal stalk domain polypeptides include SEQ ID NOs: 96, 97, 98, 99, 100, 101, 102, 103, 104, 105, 106, 107, 108, 109, 110, and 117. A further non-limiting example of an influenza HA N-terminal stalk domain polypeptide is one that consists of, consists essentially of, or comprises, amino acids 1-228 of SEQ ID NO: 117, or amino acids 229 to 519 of SEQ ID NO: 1. In some embodiments the influenza HA N-terminal stalk domain polypeptide comprises one or more to-tyrosine mutations, for example at one or more of positions 403, 406, 411, 422, 429, 432, 433, or 435 of SEQ ID NO: 1, or positions corresponding thereto (for example as determined by alignment to SEQ ID NO: 1) or at one or more of positions 112, 115, 120, 131, 137, 141, 142, or 144 of SEQ ID NO: 117, or positions corresponding thereto (for example as determined by alignment to SEQ ID NO: 117).

In some embodiments, the present invention provides influenza HA polypeptides, proteins, and/or protein complexes that comprise both (a) one or more targeted cross-links, such as di-tyrosine cross-links in their stalk domain which serve to stabilize or “lock” the stalk domain in its native trimeric conformation, and/or one or more “to-tyrosine” mutations in the HA stalk domain at locations that have been determined to be desirable locations for
the formation of di-tyrosine cross-links to stabilize the stalk domain in its native trimeric conformation, for example as described above and elsewhere throughout the present patent specification, and (b) one or more artificially-introduced protease cleavage sites that can be used to proteolytically remove the head domain of the HA polypeptide, protein, and/or protein complex, for example as described above and elsewhere throughout the present patent specification. In some embodiments, the present invention provides an influenza HA polypeptide, protein or protein complex that comprises: (a) a trimeric stalk domain that comprises one or more to-tyrosine mutations, and (b) a head domain that comprises one or more artificially-introduced protease recognition motifs. Non-limiting examples of influenza HA amino acid sequences that comprise both a to-tyrosine mutation and an artificially-introduced protease cleavage site include SEQ ID NOs: 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, and 17. In addition, any of the to-tyrosine mutations and protease cleave site insertions described or illustrated herein can be combined in the same HA polypeptide, protein, or protein complex.

[0016] In some embodiments, the present invention provides an influenza HA protein complex that comprises: (a) a trimeric stalk domain formed by the association of three protomers, wherein the stalk domain comprises one or more artificially-introduced targeted cross links, such as di-tyrosine cross-links (for example, to stabilize the stalk domain in its native trimeric conformation), and (b) a head domain that comprises one or more artificially-introduced protease recognition motifs.

[0017] In some embodiments, the present invention provides a method of making a headless influenza HA polypeptide, protein or protein complex, the method comprising: (a) obtaining or expressing an influenza HA protein comprising (i) a stalk domain and (ii) a head domain containing one or more artificially-introduced protease recognition motifs, (b) allowing the soluble influenza HA protein obtained or expressed in step (a) to fold into its native conformation having a head domain and a trimeric stalk domain comprised of three protomers, (c) introducing one more targeted cross-links, such as di-tyrosine cross-links, into the trimeric stalk domain in order to stabilize the stalk domain in its native trimeric conformation, and (d) subsequently proteolytically cleaving the head domain at the one or more artificially-introduced protease recognition motifs, thereby producing a headless influenza HA protein complex. In some such methods the stalk domain comprises one or
more “to-tyrosine” mutations and step (c) comprises introducing one or more di-tyrosine cross-links into the trimeric stalk domain. In some such methods, the locations of the di-tyrosine cross-links, to-tyrosine mutations, and/or artificially-introduced protease cleavage sites/motifs can be those specified above and/or elsewhere throughout the present patent specification. In some such methods, the influenza HA protein may be expressed in any suitable cell type, including, but not limited to, mammalian cells or insect cells.

[0018] In some embodiments the present invention provides influenza HA polypeptides, proteins, and/or protein complexes that are derived from, comprise, consist essentially of, or consist of any one of the influenza HA amino acid sequences presented herein, or any variants or fragments thereof, that have at least about 40% or 50% or 60% or 65% or 70% or 75% or 80% or 85% or 90% or 95% or 98% or 99% identity with such amino acid sequences presented herein, wherein the influenza HA polypeptides, proteins, and/or protein complexes comprise a a-to-tyrosine mutation at one or more of residues 403, 406, 411, 422, 429, 432, 433, and 435, where such amino acid numbering is based upon the sequence shown in Figure 9 (SEQ ID NO: 1), or at amino acid positions that correspond to such amino acid positions, for example as determined by alignment of an HA amino acid sequence to sequence ID NO: 1.

[0019] In some embodiments the present invention provides influenza HA polypeptides, proteins, and/or protein complexes that are derived from, comprise, consist essentially of, or consist of any one of the influenza HA amino acid sequences presented herein, or any variants or fragments thereof, that have at least about 40% or 50% or 60% or 65% or 70% or 75% or 80% or 85% or 90% or 95% or 98% or 99% identity with such amino acid sequences presented herein, wherein the influenza HA polypeptides, proteins, and/or protein complexes comprise an artificially-introduced protease cleavage site inserted after, for example immediately after, one or more of the following residues: 48, 63, 228, 278, 282, 283, 286 and 291, where such amino acid numbering is based upon the sequence shown in Figure 9 (SEQ ID NO: 1), or at amino acid positions that correspond to such amino acid positions, for example as determined by alignment of an HA amino acid sequence to sequence ID NO: 1.

[0020] In some embodiments the present invention provides influenza HA polypeptides, proteins, and/or protein complexes that are derived from, comprise, consist essentially of, or consist of any one of the influenza HA amino acid sequences presented herein, or any
variants or fragments thereof, that have at least about 40% or 50% or 60% or 65% or 70% or 75% or 80% or 85% or 90% or 95% or 98% or 99% identity with such amino acid sequences presented herein, wherein the influenza HA polypeptides, proteins, and/or protein complexes comprise two artificially-introduced protease cleavage sites, the first such site introduced immediately after residue 48 or 63, and the second such site introduced immediately after residue 228, 278, 282, 283, 286 or 291, where such amino acid numbering is based upon the sequence shown in Figure 9 (SEQ ID NO: 1), or at amino acid positions that correspond to such amino acid positions, for example as determined by alignment of an HA amino acid sequence to sequence ID NO: 1.

[0021] In some embodiments the present invention provides influenza HA polypeptides, proteins, and/or protein complexes that are derived from, comprise, consist essentially of, or consist of any one of the influenza HA amino acid sequences presented herein, or any variants or fragments thereof, that have at least about 40% or 50% or 60% or 65% or 70% or 75% or 80% or 85% or 90% or 95% or 98% or 99% identity with such amino acid sequences presented herein, wherein the influenza HA polypeptides, proteins, and/or protein complexes comprise both (a) a tyrosine residue (whether naturally occurring or arising from a mutation to-tyrosine), at one or more of residues 308, 403, 406, 411, 422, 429, 432, 433, 435, or 437, and (b) an artificially-introduced protease cleavage site inserted immediately after one or more of the following residues: 48, 63, 228, 278, 282, 283, 286 and 291, where such amino acid numbering is based upon the sequence shown in Figure 9 (SEQ ID NO: 1), or at amino acid positions that correspond to such amino acid positions, for example as determined by alignment of an HA amino acid sequence to sequence ID NO: 1.

[0022] In some embodiments the present invention provides influenza HA polypeptides, proteins, and/or protein complexes that are derived from, comprise, consist essentially of, or consist of any one of the influenza HA amino acid sequences presented herein, or any variants or fragments thereof, that have at least about 40% or 50% or 60% or 65% or 70% or 75% or 80% or 85% or 90% or 95% or 98% or 99% identity with such amino acid sequences presented herein, wherein the influenza HA polypeptides, proteins, and/or protein complexes comprise both (a) a tyrosine residue (whether naturally occurring or arising from a mutation to-tyrosine), at one or more of residues 308, 403, 406, 411, 422, 429, 432, 433, 435, or 437, and (b) two artificially-introduced protease cleavage sites - the first such site introduced
immediately after residue 48 or 63, and the second such site introduced immediately after residue 228, 278, 282, 283, 286 or 291, where such amino acid numbering is based upon the sequence shown in Figure 9 (SEQ ID NO: 1), or at amino acid positions that correspond to such amino acid positions, for example as determined by alignment of an HA amino acid sequence to sequence ID NO: 1.

[0023] In some embodiments the present invention provides influenza HA polypeptides, proteins, and/or protein complexes that are derived from, comprise, consist essentially of, or consist of amino acid residues 229 to 519 of SEQ ID NO: 1, or 279 to 519 of SEQ ID NO: 1, or 283 to 519 of SEQ ID NO: 1, or 284 to 519 of SEQ ID NO: 1, or 287 to 519 of SEQ ID NO: 1, or 292 to 519 of SEQ ID NO: 1, or SEQ ID NO: 96, 97, 98, 99, 100, 101, 102, 103, 104, 105, 106, 107, 108, 109, 110, or 117, or amino acid residues 1-228 of SEQ ID NO: 117, or sequences that have at least about 40% or 50% or 60% or 65% or 70% or 75% or 80% or 85% or 90% or 95% or 98% or 99% identity with such amino acid sequences, wherein the influenza HA polypeptides, proteins, and/or protein complexes comprise a tyrosine residue or to-tyrosine mutation at one or more of residues 308, 403, 406, 411, 422, 429, 432, 433, 435, or 437, where such amino acid numbering is based upon the sequence shown in Figure 9 (SEQ ID NO: 1), or at amino acid positions that correspond to such amino acid positions, for example as determined by alignment of an HA amino acid sequence to sequence ID NO: 1, or at one or more of residues 112, 115, 120, 131, 137, 141, 142, or 144, where such amino acid numbering is based upon the sequence shown in Figure 89 (SEQ ID NO: 117), or at amino acid positions that correspond to such amino acid positions, for example as determined by alignment of an HA amino acid sequence to sequence ID NO: 117.

[0024] In some embodiments the present invention provides influenza HA polypeptides, proteins, and/or protein complexes that are derived from, comprise, consist essentially of, or consist of amino acid residues 1 to 47 of SEQ ID NO: 1, or 1 to 62 of SEQ ID NO: 1, or sequences that have at least about 40% or 50% or 60% or 65% or 70% or 75% or 80% or 85% or 90% or 95% or 98% or 99% identity with such amino acid sequences one or more of residues 308, 403, 406, 411, 422, 429, 432, 433, 435, and 437, where such amino acid numbering is based upon the sequence shown in Figure 9 (SEQ ID NO: 1), or at amino acid positions that correspond to such amino acid positions, for example as determined by alignment of an HA amino acid sequence to sequence ID NO: 1.
[0025] In some embodiments the present invention provides compositions and/or influenza HA protein complexes that comprise, consist essentially of, or consist of a first and a second polypeptide, wherein (a) the first (C-terminal) polypeptide comprises, consists essentially of, or consists of amino acid residues 229 to 519 of SEQ ID NO: 1, or 279 to 519 of SEQ ID NO: 1, or 283 to 519 of SEQ ID NO: 1, or 284 to 519 of SEQ ID NO: 1, or 287 to 519 of SEQ ID NO: 1, or 292 to 519 of SEQ ID NO: 1, or SEQ ID NO: 96, 97, 98, 99, 100, 101, 102, 103, 104, 105, 106, 107, 108, 109, 110, or 117, or amino acid residues 1-228 of SEQ ID NO: 117, or sequences that have at least about 40% or 50% or 60% or 65% or 70% or 75% or 80% or 85% or 90% or 95% or 98% or 99% identity with such amino acid sequences, and wherein the influenza HA polypeptides, proteins, and/or protein complexes comprise a tyrosine residue or to-tyrosine mutation, at one or more of residues 308, 403, 406, 411, 422, 429, 432, 433, 435, or 437,, where such amino acid numbering is based upon the sequence shown in Figure 9 (SEQ ID NO: 1), or at amino acid positions that correspond to such amino acid positions, for example as determined by alignment of an HA amino acid sequence to sequence ID NO: 1, or at one or more of residues 112, 115, 120, 131, 137, 141, 142, or 144, where such amino acid numbering is based upon the sequence shown in Figure 89 (SEQ ID NO: 117), or at amino acid positions that correspond to such amino acid positions, for example as determined by alignment of an HA amino acid sequence to sequence ID NO: 117, and wherein (b) the second (N-terminal) polypeptide comprises, consists essentially of, or consists of amino acid residues 1 to 47 of SEQ ID NO: 1, or 1 to 62 of SEQ ID NO: 1, where such amino acid numbering is based upon the sequence shown in Figure 9 (SEQ ID NO: 1), or at amino acid positions that correspond to such amino acid positions, for example as determined by alignment of an HA amino acid sequence to sequence ID NO: 1, or sequences that have at least about 40% or 50% or 60% or 65% or 70% or 75% or 80% or 85% or 90% or 95% or 98% or 99% identity with such amino acid sequences.

[0026] In some embodiments, the present invention provides an influenza hemagglutinin (HA) polypeptide, protein or protein complex comprising, consisting essentially of, or consisting of, an amino acid sequence having at least 50, 55, 60, 65, or 70 % sequence identity to amino acid residues 229 to 519 of SEQ ID NO: 1, wherein the amino acid sequence comprises a point mutation to tyrosine at one or more of amino acid positions 403, 406, 411, 422, 429, 432, 433, and 435, where such amino acid numbering is based upon the
sequence shown in Figure 9 (SEQ ID NO: 1), or at amino acid positions that correspond to such amino acid positions, for example as determined by alignment of an HA amino acid sequence to sequence ID NO: 1. In some embodiments, the present invention provides an influenza hemagglutinin (HA) polypeptide, protein or protein complex comprising, consisting essentially of, or consisting of, an amino acid sequence having at least 50, 55, 60, 65, or 70 % sequence identity to amino acid residues 1 to 228 of SEQ ID NO: 117, wherein the amino acid sequence comprises a point mutation to-tyrosine at one or more of amino acid positions 112, 115, 120, 131, 137, 141, 142, or 144, where such amino acid numbering is based upon the sequence shown in Figure 89 (SEQ ID NO: 117), or at amino acid positions that correspond to such amino acid positions, for example as determined by alignment of an HA amino acid sequence to sequence ID NO: 117. In some such embodiments the influenza HA polypeptide, protein or protein complex forms a part of, and/or is folded into a protein complex having, or capable of forming, a trimeric stalk conformation, and that comprises at least one di-tyrosine cross-link, wherein one or both tyrosines of the at least one di-tyrosine cross-link originate from one of the to-tyrosine mutations. In some such embodiments, the influenza HA polypeptide, protein or protein complex comprises cross-links located between one or more paired tyrosine residues, wherein the paired tyrosine residues are selected from the group consisting of residues 403 and 433; 411 and 422, 403 and 429, 403 and 432, 433 and 435, and 406 and 433, where such amino acid numbering is based upon the sequence shown in Figure 9 (SEQ ID NO: 1), or at amino acid positions that correspond to such amino acid positions, for example as determined by alignment of an HA amino acid sequence to sequence ID NO: 1.

[0027] In some embodiments the HA polypeptides, proteins or protein complexes described herein are capable of folding into a trimeric stalk conformation. In some such embodiments, the influenza HA polypeptides, proteins or protein complexes described herein further comprise one or more point mutations to cysteine. In some embodiments, the influenza HA polypeptides, proteins or protein complexes described herein further comprise a trimerization domain, such as a foldon domain.

[0028] Non-limiting examples of influenza HA polypeptides, proteins and/or protein complexes of the invention include, but are not limited to, those of SEQ ID NOs: 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 86,
87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100, 101, 102, 103, 104, 105, 106, 107, 
108, 109, 110, and 117.

[0029] In some embodiments the influenza HA polypeptides, proteins or protein 
complexes described herein are capable of eliciting production of influenza HA-specific 
antibodies in a subject. In some embodiments, the influenza HA polypeptides, proteins or 
protein complexes described herein are capable of binding to an antibody that recognizes the 
trimeric stalk domain of influenza HA.

[0030] In some embodiments the present invention provides nucleic acid molecules 
encoding the influenza HA polypeptides, proteins or protein complexes described herein.

[0031] In some embodiments, the present invention provides compositions comprising 
the influenza HA polypeptides, proteins or protein complexes described herein, including, 
but not limited to, vaccine compositions. In some such embodiments, such compositions 
may further comprise an adjuvant, a carrier, an immunostimulatory agent, or any 
combination thereof.

[0032] In some embodiments the present invention provides a method of vaccinating a 
subject against influenza, the method comprising administering to a subject a composition 
comprising an effective amount of an influenza HA polypeptide, protein or protein complex 
as described herein.

[0033] These and other embodiments of the present invention are described throughout 
the present patent specification.

BRIEF DESCRIPTIONS OF THE DRAWINGS

[0034] Figure 1. Schematic representation of a headless universal vaccine immunogen 
(PR8) presenting a QNE-bnAb complex (left), and the same bnAb neutralizing HA of (a) 
homologous PR8, (b) drift (NL09), (c) group 1 heterologous (VN04), and (d) group 2 
heterologous (x31) virus on the right.

[0035] Figure 2. Schematic representation of DT-cross-links in headless HA stabilizing 
the stalk trimer. A. DT bonds (top in black) conformationally lock the stalk trimer. B. The 
stalk trimer has fallen apart without conformational locking. The QNE is lost.
Figure 3. Schematic diagram of a top-down view of the stalk, showing HA variant design: two amino acid substitutions per protomer (black and white circles).

Figure 4. (A) DT-specific fluorescence measurement at 405 nm or WT (negative control, left), four HA variants with two amino acid substitutions each, and insulin, as it forms DT bonds with high efficiency (positive control, right). (B) Relative fluorescence of dityrosine mutants. Data represents the average of four replicates with standard deviation indicated by the error bars.

Figure 5. Annotated crystal structure of HA bound to CRC261. The lower circle indicates the targeted area for DT bond formation, the middle circle indicates the targeted area for stalk-proximal proteolytic cleavage, and the upper circle indicates the targeted area for variable loop proteolysis designed to unravel the head to enable stalk-proximal cleavage site access.

Figure 6. Immunofluorescent staining of cells expressing WT and a headless HA protein without cross-linking to stabilize the stalk domain demonstrated that the non-stabilized headless HA protein did not bind one of the most broadly reactive mAbs, C179. A549 cells were transfected with plasmids for the expression of either WT HA or a recombinantly-spliced headless construct without any cross-linking in the stalk domain. 24 hrs post transfection, cells were fixed, permeabilized, and the HA protein was detected with both rabbit polyclonal, pAB (general expression) (upper panels), and mAb C179 anti-stalk (conformational) (lower panels) primary Abs followed by anti-rabbit Alexa 555-conjugated and anti-mouse Alexa 488-conjugated secondary Abs.

Figure 7. DT crosslinks form in the PR8 stalk efficiently, and C179 antigenicity is preserved before & after crosslinking. A. DT-specific fluorescence measurement at ex320/em405nm of WT (neg. control, A), four HA variants with two amino acid substitutions each (to-Tyr substitutions), at residues 403 and 429 (B), 406 and 433 (C), 403 and 433 (D), and 403 and 432 (E) and insulin, which forms DT bonds with high efficiency (positive control, F). B. C179 binding to variants (B-E) before and after DT crosslinking, as measured by sandwich ELISA using goat polyclonal anti-HA antibody for capture (BEI catalog # NR-3148) and the C179 conformational Ab for detection.

Figure 8. 293T cells were untransfected (-) or transfected with WT NA and the
indicated HA plasmids. 72 hours post transfection, VLPs in supernatants and WCEs were analyzed by sandwich ELISA (A, BEI catalog # NR-3148 goat polyclonal anti-HA capture, C179 detection), western blot (B, left panel; PNGase treated WCE), and HA assay (C).

Panel B, right. Cells were transfected as above as indicated with HA and NA. 72 hours post transfection, VLPs were purified over a 30% sucrose-NTE cushion, assayed for total protein, and either mock incubated (WT, 48G) or digested with TEV protease (WT+TEV, 48G+TEV) and PNGase treated. Percent cleavage was determined by western blot.

**Figure 9.** Amino acid sequence (SEQ ID NO. 1) of HA protein from PR8 strain of H1N1 influenza virus. Amino acids 59 through 291 comprise the head domain, which may be proteolytically removed or disrupted in some embodiments. Amino acids 1 through 58 (or 18 to 58 without the signal peptide – which is located at residues 1-17) and 292 through 566 (or 292 through 529 without the transmembrane domain and cytoplasmic tail) comprise the stalk domain. The stalk domain is discontinuous and comprises both an N-terminal and a C-terminal portion of the HA protein. Amino acids 529 through 565 comprise the transmembrane region and cytoplasmic tail. The HA ectodomain (i.e. the outer exposed / non-membrane bound portion) comprises residues 1-528 (or 18 to 528 without the signal peptide).

**Figure 10.** Nucleic acid sequence (SEQ ID NO. 2) of DNA encoding HA protein from PR8 strain of H1N1 influenza virus.

**Figure 11.** Amino acid sequence of a modified PR8 influenza HA protein comprising inserted TEV protease cleavage sites at positions 63 and 278 (underlined), and to-tyrosine mutations at positions 403 (N403Y) and 433 (D433Y) (underlined) (SEQ ID NO:3). The boxed C-terminal sequence comprises the transmembrane region. The amino acid sequence of SEQ ID NO:3 is encoded by the nucleic acid sequence of SEQ ID NO:31 shown in Figure 28.

**Figure 12.** Amino acid sequence of a modified PR8 influenza HA protein comprising inserted TEV protease cleavage sites at positions 63 and 278 (underlined), and to-tyrosine mutations at positions 411 (K411Y) and 422 (N422Y) (underlined) (SEQ ID NO:4). The boxed C-terminal sequence comprises the transmembrane region. The amino
acid sequence of SEQ ID NO:4 is encoded by the nucleic acid sequence of SEQ ID NO:32 shown in Figure 29.

[0046] **Figure 13.** Amino acid sequence of a modified PR8 influenza HA protein comprising inserted TEV protease cleavage sites at positions 63 and 278 (underlined), and to-tyrosine mutations at positions 403 (N403Y), 411 (K411Y), 422 (N422Y), and 433 (D433Y) (underlined) (SEQ ID NO:5). The boxed C-terminal sequence comprises the transmembrane region. The amino acid sequence of SEQ ID NO:5 is encoded by the nucleic acid sequence of SEQ ID NO:33 shown in Figure 30.

[0047] **Figure 14.** Amino acid sequence of a modified PR8 influenza HA protein comprising inserted TEV protease cleavage sites at positions 63 and 282 (underlined), and to-tyrosine mutations at positions 403 (N403Y) and 433 (D433Y) (underlined) (SEQ ID NO:6). The boxed C-terminal sequence comprises the transmembrane region. The amino acid sequence of SEQ ID NO:6 is encoded by the nucleic acid sequence of SEQ ID NO:34 shown in Figure 31.

[0048] **Figure 15.** Amino acid sequence of a modified PR8 influenza HA protein comprising inserted TEV protease cleavage sites at positions 63 and 282 (underlined), and to-tyrosine mutations at positions 411 (K411Y) and 422 (N422Y) (underlined) (SEQ ID NO:7). The boxed C-terminal sequence comprises the transmembrane region. The amino acid sequence of SEQ ID NO:7 is encoded by the nucleic acid sequence of SEQ ID NO:35 shown in Figure 32.

[0049] **Figure 16.** Amino acid sequence of a modified PR8 influenza HA protein comprising inserted TEV protease cleavage sites at positions 63 and 228 (underlined), and to-tyrosine mutations at positions 403 (N403Y), 411 (K411Y), 422 (N422Y), and 433 (D433Y) (underlined) (SEQ ID NO:8). The boxed C-terminal sequence comprises the transmembrane region. The amino acid sequence of SEQ ID NO:8 is encoded by the nucleic acid sequence of SEQ ID NO:36 shown in Figure 33.

[0050] **Figure 17.** Amino acid sequence of a modified PR8 influenza HA protein comprising inserted TEV protease cleavage sites at positions 63 and 283 (underlined), and to-tyrosine mutations at positions 403 (N403Y) and 433 (D433Y) (underlined) (SEQ ID NO:9). The boxed C-terminal sequence comprises the transmembrane region. The amino
acid sequence of SEQ ID NO:9 is encoded by the nucleic acid sequence of SEQ ID NO:37 shown in Figure 34.

[0051] **Figure 18.** Amino acid sequence of a modified PR8 influenza HA protein comprising inserted TEV protease cleavage sites at positions 63 and 283 (underlined), and to-tyrosine mutations at positions 411 (K411Y) and 422 (N422Y) (underlined) (SEQ ID NO:10). The boxed C-terminal sequence comprises the transmembrane region. The amino acid sequence of SEQ ID NO:10 is encoded by the nucleic acid sequence of SEQ ID NO:38 shown in Figure 35.

[0052] **Figure 19.** Amino acid sequence of a modified PR8 influenza HA protein comprising inserted TEV protease cleavage sites at positions 63 and 283 (underlined), and to-tyrosine mutations at positions 403 (N403Y), 411 (K411Y), 422 (N422Y), and 433 (D433Y) (underlined) (SEQ ID NO:11). The boxed C-terminal sequence comprises the transmembrane region. The amino acid sequence of SEQ ID NO:11 is encoded by the nucleic acid sequence of SEQ ID NO:39 shown in Figure 36.

[0053] **Figure 20.** Amino acid sequence of a modified PR8 influenza HA protein comprising inserted TEV protease cleavage sites at positions 48 and 291 (underlined), and to-tyrosine mutations at positions 403 (N403Y) and 433 (D433Y) (underlined) (SEQ ID NO:12). The boxed C-terminal sequence comprises the transmembrane region. The amino acid sequence of SEQ ID NO:12 is encoded by the nucleic acid sequence of SEQ ID NO:43 shown in Figure 40.

[0054] **Figure 21.** Amino acid sequence of a modified PR8 influenza HA protein comprising inserted TEV protease cleavage sites at positions 48 and 291 (underlined), and to-tyrosine mutations at positions 411 (K411Y) and 422 (N422Y) (underlined) (SEQ ID NO:13). The boxed C-terminal sequence comprises the transmembrane region. The amino acid sequence of SEQ ID NO:13 is encoded by the nucleic acid sequence of SEQ ID NO:44 shown in Figure 41.

[0055] **Figure 22.** Amino acid sequence of a modified PR8 influenza HA protein comprising inserted TEV protease cleavage sites at positions 48 and 291 (underlined), and to-tyrosine mutations at positions 403 (N403Y), 411 (K411Y), 422 (N422Y), and 433 (D433Y) (underlined) (SEQ ID NO:14). The boxed C-terminal sequence comprises the
transmembrane region. The amino acid sequence of SEQ ID NO:14 is encoded by the nucleic acid sequence of SEQ ID NO:45 shown in Figure 42.

[0056] Figure 23. Amino acid sequence of a modified PR8 influenza HA protein comprising inserted TEV protease cleavage sites at positions 48 and 291 (underlined), and to-tyrosine mutations at positions 403 (N403Y) and 433 (D433Y) (underlined) (SEQ ID NO:15). The boxed C-terminal sequence comprises the transmembrane region. The amino acid sequence of SEQ ID NO:15 is encoded by the nucleic acid sequence of SEQ ID NO:46 shown in Figure 43.

[0057] Figure 24. Amino acid sequence of a modified PR8 influenza HA protein comprising inserted TEV protease cleavage sites at positions 48 and 291 (underlined), and to-tyrosine mutations at positions 411 (K411Y) and 422 (N422Y) (underlined) (SEQ ID NO:16). The boxed C-terminal sequence comprises the transmembrane region. The amino acid sequence of SEQ ID NO:16 is encoded by the nucleic acid sequence of SEQ ID NO:47 shown in Figure 44.

[0058] Figure 25. Amino acid sequence of a modified PR8 influenza HA protein comprising inserted TEV protease cleavage sites at positions 48 and 291 (underlined), and to-tyrosine mutations at positions 403 (N403Y), 411 (K411Y), 422 (N422Y), and 433 (D433Y) (underlined) (SEQ ID NO:17). The boxed C-terminal sequence comprises the transmembrane region. The amino acid sequence of SEQ ID NO:17 is encoded by the nucleic acid sequence of SEQ ID NO:48 shown in Figure 45.

[0059] Figure 26A-B. Amino acid sequence alignment of modified PR8 influenza HA proteins comprising one inserted protease cleavage site, and the sequence of wild-type PR8 HA from the PR8 strain of influenza virus H1N1 (SEQ ID NO:1 – identified as “PR8HA-WT” in the figure). Underlined amino acid residues indicate protease cleavage sites inserted into the wild-type sequence by substitution and/or replacement of amino acids in SEQ ID NO. 1. Protease cleavage sites are inserted immediately after the following amino acid residues: 291 (SEQ ID NO. 18 and SEQ ID NO. 19), 48 (SEQ ID NO. 20), 286 (SEQ ID NO. 21), 278 (SEQ ID NO. 22), 282, (SEQ ID NO. 23), 63 (SEQ ID NO. 24), or 283 (SEQ ID NO. 25). The inserted protease cleavage sites are TEV protease recognition sequences. The C-terminal sequences shown within the boxed portion of the alignment comprise the
transmembrane regions of the influenza HA proteins. The amino acid sequences of SEQ ID NO. 18, 19, 20, 21, 22, 23, 24 and 25 are encoded by the nucleic acid sequences of SEQ ID NO. 49, 50, 52, 56, 53, 54, 51 and 55, respectively as shown in Figure 46.

[0060] Figure 27A-B. Amino acid sequence alignment of modified PR8 influenza HA proteins comprising two inserted protease cleavage sites, and the sequence of wild-type PR8 HA from the PR8 strain of influenza virus H1N1 (SEQ ID NO:1 – identified as “PR8HA-WT” in the figure). Protease cleavage sites are inserted immediately after the following amino acid residues: 63 and 278 (SEQ ID NO. 26), 63 and 282 (SEQ ID NO. 27), 63 and 283 (SEQ ID NO. 28), 48 and 291 (SEQ ID NO. 29 and 30). The inserted protease cleavage sites are TEV protease recognition sequences. Underlined amino acid residues indicate the sequence located between the protease cleavage sites that would be removed from the HA sequence upon cleavage by a protease (here, TEV protease), for example to facilitate the production of a “headless” HA protein where the head domain is disrupted or removed. The C-terminal sequences shown within the boxed portion of the alignment comprise the transmembrane regions of the influenza HA proteins. Amino acid residues shown in bold (N403, F406, K411, N422, D429, L432, D433 and W435) illustrate positions where tyrosine mutations may be made so as to facilitate the formation of dityrosine bonds in the influenza HA stalk domain, as described herein. The amino acid sequences of SEQ ID NO. 26, 27, 28, 29 and 30 are encoded by the nucleic acid sequences of SEQ ID NO. 57, 58, 62, 60 and 61, respectively as shown in Figure 47.

[0061] Figure 28. Nucleic acid sequence (SEQ ID NO:31) encoding a modified PR8 influenza HA protein comprising inserted nucleic acid residues (shown in lower case) that encode TEV protease cleavage sites at positions 63 and 278 in the protein, and to-tyrosine mutations (shown in lower case) encoded at positions 403 (N403Y) and 433 (D433Y) in the protein. The boxed C-terminal sequence comprises the nucleic acid sequence that encodes the transmembrane region of the protein.

[0062] Figure 29. Nucleic acid sequence (SEQ ID NO:32) encoding a modified PR8 influenza HA protein comprising inserted nucleic acid residues (shown in lower case) that encode TEV protease cleavage sites at positions 63 and 278 in the protein, and to-tyrosine mutations (shown in lower case) encoded at positions 411 (K411Y) and 422 (N422Y) in the
protein. The boxed C-terminal sequence comprises the nucleic acid sequence that encodes the transmembrane region of the protein.

[0063] **Figure 30.** Nucleic acid sequence (SEQ ID NO:33) encoding a modified PR8 influenza HA protein comprising inserted nucleic acid residues (shown in lower case) that encode TEV protease cleavage sites at positions 63 and 278 in the protein, and to-tyrosine mutations (shown in lower case) encoded at positions 403 (N403Y), 411 (K411Y), 422 (N422Y), and 433 (D433Y) in the protein. The boxed C-terminal sequence comprises the nucleic acid sequence that encodes the transmembrane region of the protein.

[0064] **Figure 31.** Nucleic acid sequence (SEQ ID NO:34) encoding a modified PR8 influenza HA protein comprising inserted nucleic acid residues (shown in lower case) that encode TEV protease cleavage sites at positions 63 and 282 in the protein, and to-tyrosine mutations (shown in lower case) encoded at positions 403 (N403Y) and 433 (D433Y) in the protein. The boxed C-terminal sequence comprises the nucleic acid sequence that encodes the transmembrane region of the protein.

[0065] **Figure 32.** Nucleic acid sequence (SEQ ID NO:34) encoding a modified PR8 influenza HA protein comprising inserted nucleic acid residues (shown in lower case) that encode TEV protease cleavage sites at positions 63 and 282 in the protein, and to-tyrosine mutations (shown in lower case) encoded at positions 411 (K411Y) and 422 (N422Y) in the protein. The boxed C-terminal sequence comprises the nucleic acid sequence that encodes the transmembrane region of the protein.

[0066] **Figure 33.** Nucleic acid sequence (SEQ ID NO:36) encoding a modified PR8 influenza HA protein comprising inserted nucleic acid residues (shown in lower case) that encode TEV protease cleavage sites at positions 63 and 282 in the protein, and to-tyrosine mutations (shown in lower case) encoded at positions 403 (N403Y), 411 (K411Y), 422 (N422Y), and 433 (D433Y) in the protein. The boxed C-terminal sequence comprises the nucleic acid sequence that encodes the transmembrane region of the protein.

[0067] **Figure 34.** Nucleic acid sequence (SEQ ID NO:37) encoding a modified PR8 influenza HA protein comprising inserted nucleic acid residues (shown in lower case) that encode TEV protease cleavage sites at positions 63 and 283 in the protein, and to-tyrosine mutations (shown in lower case) encoded at positions 403 (N403Y) and 433 (D433Y) in the
protein. The boxed C-terminal sequence comprises the nucleic acid sequence that encodes the transmembrane region of the protein.

[0068] **Figure 35.** Nucleic acid sequence (SEQ ID NO:38) encoding a modified PR8 influenza HA protein comprising inserted nucleic acid residues (shown in lower case) that encode TEV protease cleavage sites at positions 63 and 283 in the protein, and to-tyrosine mutations (shown in lower case) encoded at positions 411 (K411Y) and 422 (N422Y) in the protein. The boxed C-terminal sequence comprises the nucleic acid sequence that encodes the transmembrane region of the protein.

[0069] **Figure 36.** Nucleic acid sequence (SEQ ID NO:39) encoding a modified PR8 influenza HA protein comprising inserted nucleic acid residues (shown in lower case) that encode TEV protease cleavage sites at positions 63 and 283 in the protein, and to-tyrosine mutations (shown in lower case) encoded at positions 403 (N403Y), 411 (K411Y), 422 (N422Y), and 433 (D433Y) in the protein. The boxed C-terminal sequence comprises the nucleic acid sequence that encodes the transmembrane region of the protein.

[0070] **Figure 37.** Nucleic acid sequence (SEQ ID NO:40) encoding a modified PR8 influenza HA protein comprising inserted nucleic acid residues (shown in lower case) that encode TEV protease cleavage sites at positions 63 and 286 in the protein, and to-tyrosine mutations (shown in lower case) encoded at positions 403 (N403Y) and 433 (D433Y) in the protein. The boxed C-terminal sequence comprises the nucleic acid sequence that encodes the transmembrane region of the protein.

[0071] **Figure 38.** Nucleic acid sequence (SEQ ID NO:41) encoding a modified PR8 influenza HA protein comprising inserted nucleic acid residues (shown in lower case) that encode TEV protease cleavage sites at positions 63 and 286 in the protein, and to-tyrosine mutations (shown in lower case) encoded at positions 411 (K411Y) and 422 (N422Y) in the protein. The boxed C-terminal sequence comprises the nucleic acid sequence that encodes the transmembrane region of the protein.

[0072] **Figure 39.** Nucleic acid sequence (SEQ ID NO:42) encoding a modified PR8 influenza HA protein comprising inserted nucleic acid residues (shown in lower case) that encode TEV protease cleavage sites at positions 63 and 286 in the protein, and to-tyrosine mutations (shown in lower case) encoded at positions 403 (N403Y), 411 (K411Y), 422
(N422Y), and 433 (D433Y) in the protein. The boxed C-terminal sequence comprises the nucleic acid sequence that encodes the transmembrane region of the protein.

**[0073]** Figure 40. Nucleic acid sequence (SEQ ID NO:43) encoding a modified PR8 influenza HA protein comprising inserted nucleic acid residues (shown in lower case) that encode TEV protease cleavage sites at positions 48 and 291 in the protein, and to-tyrosine mutations (shown in lower case) encoded at positions 403 (N403Y) and 433 (D433Y) in the protein. The boxed C-terminal sequence comprises the nucleic acid sequence that encodes the transmembrane region of the protein.

**[0074]** Figure 41. Nucleic acid sequence (SEQ ID NO:44) encoding a modified PR8 influenza HA protein comprising inserted nucleic acid residues (shown in lower case) that encode TEV protease cleavage sites at positions 48 and 291 in the protein, and to-tyrosine mutations (shown in lower case) encoded at positions 411 (K411Y) and 422 (N422Y) in the protein. The boxed C-terminal sequence comprises the nucleic acid sequence that encodes the transmembrane region of the protein.

**[0075]** Figure 42. Nucleic acid sequence (SEQ ID NO:45) encoding a modified PR8 influenza HA protein comprising inserted nucleic acid residues (shown in lower case) that encode TEV protease cleavage sites at positions 48 and 291 in the protein, and to-tyrosine mutations (shown in lower case) encoded at positions 403 (N403Y), 411 (K411Y), 422 (N422Y), and 433 (D433Y) in the protein. The boxed C-terminal sequence comprises the nucleic acid sequence that encodes the transmembrane region of the protein.

**[0076]** Figure 43. Nucleic acid sequence (SEQ ID NO:46) encoding a modified PR8 influenza HA protein comprising inserted nucleic acid residues (shown in lower case) that encode TEV protease cleavage sites at positions 48 and 291 in the protein, and to-tyrosine mutations (shown in lower case) encoded at positions 403 (N403Y) and 433 (D433Y) in the protein. The boxed C-terminal sequence comprises the nucleic acid sequence that encodes the transmembrane region of the protein.

**[0077]** Figure 44. Nucleic acid sequence (SEQ ID NO:47) encoding a modified PR8 influenza HA protein comprising inserted nucleic acid residues (shown in lower case) that encode TEV protease cleavage sites at positions 48 and 291 in the protein, and to-tyrosine mutations (shown in lower case) encoded at positions 411 (K411Y) and 422 (N422Y) in the
protein. The boxed C-terminal sequence comprises the nucleic acid sequence that encodes the transmembrane region of the protein.

[0078] **Figure 45.** Nucleic acid sequence (SEQ ID NO:48) encoding a modified PR8 influenza HA protein comprising inserted nucleic acid residues (shown in lower case) that encode TEV protease cleavage sites at positions 48 and 291 in the protein, and to-tyrosine mutations (shown in lower case) encoded at positions 403 (N403Y), 411 (K411Y), 422 (N422Y), and 433 (D433Y) in the protein. The boxed C-terminal sequence comprises the nucleic acid sequence that encodes the transmembrane region of the protein.

[0079] **Figure 46A-F.** Alignment of nucleic acid sequences encoding modified PR8 influenza HA proteins comprising one inserted protease cleavage site, and the sequence of wild-type HA protein from the PR8 strain of influenza virus H1N1 (SEQ ID NO:2 – identified as “RR8HA-WT” in the figure). Underlined nucleic acid residues encode TEV protease cleavage sites by substitution and/or replacement of nucleic acid residues of SEQ ID NO. 2. The nucleic acid residues are inserted into the nucleic acid sequence such that the encoded protein will have a protease cleavage site immediately after the following amino acid residues: 291 (SEQ ID NO. 49 and SEQ ID NO. 50), 48 (SEQ ID NO. 52), 286 (SEQ ID NO. 56), 278 (SEQ ID NO. 53), 282 (SEQ ID NO. 54), 63 (SEQ ID NO. 51), or 283 (SEQ ID NO. 55). The boxed C-terminal sequences comprise the sequence that encodes the transmembrane region of the protein.

[0080] **Figure 47A-E.** Alignment of nucleic acid sequences encoding PR8 influenza HA proteins comprising two inserted protease cleavage sites, and the sequence of wild-type HA from the PR8 strain of influenza virus H1N1 (SEQ ID NO:2 – identified as “PR8HA-WT” in the figure). Underlined nucleic acid residues encode TEV protease cleavage sites by substitution and/or replacement of nucleic acid residues of SEQ ID NO. 2. The nucleic acid residues are inserted into the nucleic acid sequence such that the encoded HA protein will have protease cleavage sites immediately after the following amino acid residues: 63 and 278 (SEQ ID NO. 57), 63 and 282 (SEQ ID NO. 58), 63 and 286 (SEQ ID NO. 59), 48 and 291 (SEQ ID NO. 60 and 61), and 63 and 283 (SEQ ID NO. 62). The boxed C-terminal sequences comprise the sequence that encodes the transmembrane region of the protein. Boxed nucleic acid residues (corresponding to amino acid positions N403, F406, K411, N422, D429, L432, D433 and W435 in the encoded HA protein) illustrate positions where
To-tyrosine mutations may be made so as to facilitate the formation of dityrosine bonds in the stalk domain of the encoded influenza HA protein, as described herein.

[0081] **Figure 48.** Nucleic acid sequence encoding HA protein of PR8 strain of influenza virus H1N1 with codon optimization for expression of the encoded HA protein in Homo sapiens (SEQ ID NO:63).

[0082] **Figure 49.** Nucleic acid sequence encoding HA protein of PR8 strain of influenza virus H1N1 with codon optimization for expression of the encoded HA protein in Cricetulus griseus (SEQ ID NO:64).

[0083] **Figure 50.** Nucleic acid sequence encoding HA protein of PR8 strain of influenza virus H1N1 with codon optimization for expression of the encoded HA protein in Nicotiana benthamiana (SEQ ID NO:65).

[0084] **Figure 51.** Nucleic acid sequence encoding HA protein of PR8 strain of influenza virus H1N1 with codon optimization for expression of the encoded HA protein in Pichia pastoris (SEQ ID NO:66).

[0085] **Figure 52.** Nucleic acid sequence encoding HA protein of PR8 strain of influenza virus H1N1 with codon optimization for expression of the encoded HA protein in Saccharomyces cerevisiae (SEQ ID NO:67).

[0086] **Figure 53.** Nucleic acid sequence encoding HA protein of PR8 strain of influenza virus H1N1 with codon optimization for expression of the encoded HA protein in Spodoptera frugiperda (SEQ ID NO:68).

[0087] **Figure 54A-C.** Alignment of amino acid sequences of full-length versions of HA proteins from various strains of influenza virus (Udorn 72 (SEQ ID NO:73), Hong Kong 68 (SEQ ID NO:74), Panama 99 (SEQ ID NO:75), Wisconsin 05 (SEQ ID NO:76), Shanghai 13 (SEQ ID NO:77), Singapore 57 (SEQ ID NO:78), Vietnam 04 (SEQ ID NO:79) and PR8 34 (SEQ ID NO:1), USSR 77 (SEQ ID NO:111), Texas 91 (SEQ ID NO:112), WSN 33 (SEQ ID NO:113), South Carolina 1918 (SEQ ID NO:114), and California 09 (SEQ ID NO:115)). Boxed amino acid residues (corresponding to amino acid positions 403, 406, 411, 422, 429, 432, 433 and 435 in the sequence of wild-type HA from the PR8 strain of influenza virus H1N1 (SEQ ID NO:1 – identified as “WT-PR8-34” in the figure) represent
positions where a mutation to a tyrosine residue is contemplated to facilitate the formation of dityrosine bonds in the stalk region of the HA protein. The italicized C-terminal sequences comprise the sequence that encodes the endogenous transmembrane region of the protein, and which can be removed or disrupted so as to generate a soluble version of influenza HA protein (see, for example, Figures 55 – 60).

[0088] Figure 55. Amino acid sequence of a soluble version of HA protein from the PR8 strain of influenza virus (SEQ ID NO:80). Amino acids 520 – 565 of the endogenous transmembrane region (italicized C-terminal sequence of SEQ ID NO:1 in Figure 54) have been replaced by an optional tag (underlined) comprising a thrombin cleavage domain, a T4 foldon trimerization motif, and a 6xHis tag (SEQ ID NO: 118).

[0089] Figure 56. Amino acid sequence of a soluble version of HA protein from the Hong Kong 68 strain of influenza virus (SEQ ID NO:81). Amino acids 521 – 566 of the endogenous transmembrane region (italicized C-terminal sequence of SEQ ID NO:74 in Figure 54) have been replaced by an optional tag (underlined) comprising a thrombin cleavage domain, a T4 foldon trimerization motif, and a 6xHis tag (SEQ ID NO: 118).

[0090] Figure 57. Amino acid sequence of a soluble version of HA protein from the Wisconsin 05 strain of influenza virus (SEQ ID NO:82). Amino acids 521 – 566 of the endogenous transmembrane region (italicized C-terminal sequence of SEQ ID NO:76 in Figure 54) have been replaced by an optional tag (underlined) comprising a thrombin cleavage domain, a T4 foldon trimerization motif, and a 6xHis tag (SEQ ID NO: 118).

[0091] Figure 58. Amino acid sequence of a soluble version of HA protein from the Vietnam 04 strain of influenza virus (SEQ ID NO:83). Amino acids 522 – 568 of the endogenous transmembrane region (italicized C-terminal sequence of SEQ ID NO:79 in Figure 54) have been replaced by an optional tag (underlined) comprising a thrombin cleavage domain, a T4 foldon trimerization motif, and a 6xHis tag (SEQ ID NO: 118).

[0092] Figure 59. Amino acid sequence of a soluble version of HA protein from the Shanghai 13 strain of influenza virus (SEQ ID NO:84). Amino acids 515 – 560 of the endogenous transmembrane region (italicized C-terminal sequence of SEQ ID NO:77 in Figure 54) have been replaced by an optional tag (underlined) comprising a thrombin cleavage domain, a T4 foldon trimerization motif, and a 6xHis tag (SEQ ID NO: 118).
[0093] **Figure 60.** Amino acid sequence of a soluble version of HA protein from the Singapore 57 strain of influenza virus (SEQ ID NO:85). Amino acids 516 – 562 of the endogenous transmembrane region (italicized C-terminal sequence of SEQ ID NO:78 in Figure 54) have been replaced by an optional tag (underlined) comprising a thrombin cleavage domain, a T4 foldon trimerization motif, and a 6xHis tag (SEQ ID NO: 118).

[0094] **Figure 61.** Amino acid sequence of a modified PR8 influenza HA protein comprising to-tyrosine mutations at positions 403 (N403Y) and 429 (D429Y) (underlined) (SEQ ID NO:86).

[0095] **Figure 62.** Amino acid sequence of a modified PR8 influenza HA protein comprising to-tyrosine mutations at positions 403 (N403Y) and 432 (L432Y) (underlined) (SEQ ID NO:87).

[0096] **Figure 63.** Amino acid sequence of a modified PR8 influenza HA protein comprising one to-tyrosine mutations at position 403 (N403Y) (underlined) (SEQ ID NO:88).

[0097] **Figure 64.** Amino acid sequence of a modified PR8 influenza HA protein comprising to-tyrosine mutations at positions 403 (N403Y) and 433 (D433Y) (underlined) (SEQ ID NO:89).

[0098] **Figure 65.** Amino acid sequence of a modified PR8 influenza HA protein comprising to-tyrosine mutations at positions 433 (D433Y) and 435 (W435Y) (underlined) (SEQ ID NO:90).

[0099] **Figure 66.** Amino acid sequence of a modified PR8 influenza HA protein comprising one to-tyrosine mutations at position 435 (W435Y) (underlined) (SEQ ID NO:91).

[0100] **Figure 67.** Amino acid sequence of a modified PR8 influenza HA protein comprising to-tyrosine mutations at positions 406 (F406Y) and 433 (D433Y) (underlined) (SEQ ID NO:92).
[0101] **Figure 68.** Amino acid sequence of a modified PR8 influenza HA protein comprising to-tyrosine mutations at positions 411 (K411Y) and 422 (N422Y) (underlined) (SEQ ID NO:93).

[0102] **Figure 69A-B.** Amino acid sequence alignment of modified PR8 influenza HA proteins comprising one or more to-tyrosine mutations, and the sequence of wild-type PR8 HA from the PR8 strain of influenza virus H1N1 (SEQ ID NO:1 – identified as “PR8HA-WT” in the figure). Dityrosine bonds may be introduced between various combinations of endogenous tyrosine residues (e.g. Y308 and Y437 of SEQ ID NO:1, shown in bold) and residues comprising to-tyrosine mutations (e.g. N403, F406, K411, N422, D429, L432, D433 and W435 of SEQ ID NO:1, shown as underlined), as described herein.

[0103] **Figure 70.** Protein fragments comprising a ‘headless’ influenza HA protein generated after proteolysis at two protease cleavage sites (63G/278S) inserted into the full-length starting sequence (PR8 HA, SEQ ID NO:1). The first fragment (SEQ ID NO:94) is the N-terminal portion of the stalk domain and the second fragment (SEQ ID NO:96) is the C-terminal portion of the stalk domain comprising two to-tyrosine mutations at amino acid positions 120 and 150 (underlined; corresponding to amino acid positions 403 and 433, respectively, in SEQ ID NO:1).

[0104] **Figure 71.** Protein fragments comprising a ‘headless’ influenza HA protein generated after proteolysis at two protease cleavage sites (63G/278S) inserted into the full-length starting sequence (PR8 HA, SEQ ID NO:1). The first fragment (SEQ ID NO:94) is the N-terminal portion of the stalk domain and the second fragment (SEQ ID NO:97) is the C-terminal portion of the stalk domain comprising two to-tyrosine mutations at amino acid positions 128 and 139 (underlined; corresponding to amino acid positions 411 and 422, respectively, in SEQ ID NO:1).

[0105] **Figure 72.** Protein fragments comprising a ‘headless’ influenza HA protein generated after proteolysis at two protease cleavage sites (63G/278S) inserted into the full-length starting sequence (PR8 HA, SEQ ID NO:1). The first fragment (SEQ ID NO:94) is the N-terminal portion of the stalk domain and the second fragment (SEQ ID NO:98) is the C-terminal portion of the stalk domain comprising four to-tyrosine mutations at amino acid
positions 120, 128, 139 and 150 (underlined; corresponding to amino acid positions 403, 411, 422 and 433, respectively, in SEQ ID NO:1).

[0106] Figure 73. Protein fragments comprising a ‘headless’ influenza HA protein generated after proteolysis at two protease cleavage sites (63G/282S) inserted into the full-length starting sequence (PR8 HA, SEQ ID NO:1). The first fragment (SEQ ID NO:94) is the N-terminal portion of the stalk domain and the second fragment (SEQ ID NO:99) is the C-terminal portion of the stalk domain comprising two to-tyrosine mutations at amino acid positions 122 and 152 (underlined; corresponding to amino acid positions 403 and 433, respectively, in SEQ ID NO:1).

[0107] Figure 74. Protein fragments comprising a ‘headless’ influenza HA protein generated after proteolysis at two protease cleavage sites (63G/282S) inserted into the full-length starting sequence (PR8 HA, SEQ ID NO:1). The first fragment (SEQ ID NO:94) is the N-terminal portion of the stalk domain and the second fragment (SEQ ID NO:100) is the C-terminal portion of the stalk domain comprising two to-tyrosine mutations at amino acid positions 130 and 141 (underlined; corresponding to amino acid positions 411 and 422, respectively, in SEQ ID NO:1).

[0108] Figure 75. Protein fragments comprising a ‘headless’ influenza HA protein generated after proteolysis at two protease cleavage sites (63G/282S) inserted into the full-length starting sequence (PR8 HA, SEQ ID NO:1). The first fragment (SEQ ID NO:94) is the N-terminal portion of the stalk domain and the second fragment (SEQ ID NO:101) is the C-terminal portion of the stalk domain comprising four to-tyrosine mutations at amino acid positions 122, 130, 141 and 152 (underlined; corresponding to amino acid positions 403, 411, 422 and 433, respectively, in SEQ ID NO:1).

[0109] Figure 76. Protein fragments comprising a ‘headless’ influenza HA protein generated after proteolysis at two protease cleavage sites (63G/283G) inserted into the full-length starting sequence (PR8 HA, SEQ ID NO:1). The first fragment (SEQ ID NO:94) is the N-terminal portion of the stalk domain and the second fragment (SEQ ID NO:102) is the C-terminal portion of the stalk domain comprising two to-tyrosine mutations at amino acid positions 121 and 151 (underlined; corresponding to amino acid positions 403 and 433, respectively, in SEQ ID NO:1).
[0110] **Figure 77.** Protein fragments comprising a ‘headless’ influenza HA protein generated after proteolysis at two protease cleavage sites (63G/283G) inserted into the full-length starting sequence (PR8 HA, SEQ ID NO:1). The first fragment (SEQ ID NO:94) is the N-terminal portion of the stalk domain and the second fragment (SEQ ID NO:103) is the C-terminal portion of the stalk domain comprising two to-tyrosine mutations at amino acid positions 129 and 140 (underlined; corresponding to amino acid positions 411 and 422, respectively, in SEQ ID NO:1).

[0111] **Figure 78.** Protein fragments comprising a ‘headless’ influenza HA protein generated after proteolysis at two protease cleavage sites (63G/283G) inserted into the full-length starting sequence (PR8 HA, SEQ ID NO:1). The first fragment (SEQ ID NO:94) is the N-terminal portion of the stalk domain and the second fragment (SEQ ID NO:104) is the C-terminal portion of the stalk domain comprising four to-tyrosine mutations at amino acid positions 121, 129, 140 and 151 (underlined; corresponding to amino acid positions 403, 411, 422 and 433, respectively, in SEQ ID NO:1).

[0112] **Figure 79.** Protein fragments comprising a ‘headless’ influenza HA protein generated after proteolysis at two protease cleavage sites (48G/291G) inserted into the full-length starting sequence (PR8 HA, SEQ ID NO:1). The first fragment (SEQ ID NO:95) is the N-terminal portion of the stalk domain and the second fragment (SEQ ID NO:105) is the C-terminal portion of the stalk domain comprising two to-tyrosine mutations at amino acid positions 113 and 143 (underlined; corresponding to amino acid positions 403 and 433, respectively, in SEQ ID NO:1).

[0113] **Figure 80.** Protein fragments comprising a ‘headless’ influenza HA protein generated after proteolysis at two protease cleavage sites (48G/291G) inserted into the full-length starting sequence (PR8 HA, SEQ ID NO:1). The first fragment (SEQ ID NO:95) is the N-terminal portion of the stalk domain and the second fragment (SEQ ID NO:106) is the C-terminal portion of the stalk domain comprising two to-tyrosine mutations at amino acid positions 121 and 132 (underlined; corresponding to amino acid positions 411 and 422, respectively, in SEQ ID NO:1).

[0114] **Figure 81.** Protein fragments comprising a ‘headless’ influenza HA protein generated after proteolysis at two protease cleavage sites (48G/291G) inserted into the full-
length starting sequence (PR8 HA, SEQ ID NO:1). The first fragment (SEQ ID NO:95) is the N-terminal portion of the stalk domain and the second fragment (SEQ ID NO:107) is the C-terminal portion of the stalk domain comprising four to-tyrosine mutations at amino acid positions 113, 121, 132 and 143 (underlined; corresponding to amino acid positions 403, 411, 422 and 433, respectively, in SEQ ID NO:1).

[0115] Figure 82. Protein fragments comprising a ‘headless’ influenza HA protein generated after proteolysis at two protease cleavage sites (48G/291S) inserted into the full-length starting sequence (PR8 HA, SEQ ID NO:1). The first fragment (SEQ ID NO:95) is the N-terminal portion of the stalk domain and the second fragment (SEQ ID NO:108) is the C-terminal portion of the stalk domain comprising two to-tyrosine mutations at amino acid positions 113 and 143 (underlined; corresponding to amino acid positions 403 and 433, respectively, in SEQ ID NO:1).

[0116] Figure 83. Protein fragments comprising a ‘headless’ influenza HA protein generated after proteolysis at two protease cleavage sites (48G/291S) inserted into the full-length starting sequence (PR8 HA, SEQ ID NO:1). The first fragment (SEQ ID NO:95) is the N-terminal portion of the stalk domain and the second fragment (SEQ ID NO:109) is the C-terminal portion of the stalk domain comprising two to-tyrosine mutations at amino acid positions 121 and 132 (underlined; corresponding to amino acid positions 411 and 422, respectively, in SEQ ID NO:1).

[0117] Figure 84. Protein fragments comprising a ‘headless’ influenza HA protein generated after proteolysis at two protease cleavage sites (48G/291S) inserted into the full-length starting sequence (PR8 HA, SEQ ID NO:1). The first fragment (SEQ ID NO:95) is the N-terminal portion of the stalk domain and the second fragment (SEQ ID NO:110) is the C-terminal portion of the stalk domain comprising four to-tyrosine mutations at amino acid positions 113, 121, 132 and 143 (underlined; corresponding to amino acid positions 403, 411, 422 and 433, respectively, in SEQ ID NO:1).

[0118] Figure 85. (A) 293T Cells were transfected with constructs for the expression of the indicated HA dityrosine mutants (403Y, 411Y-422Y, 403Y-433Y and 433Y-435Y) and soluble HA protein (with a C-terminal foldon domain) was purified from supernatants 72 hours post transfection by Ni²⁺ affinity chromatography. Pure HA protein was then
subjected to dityrosine crosslinking conditions in the presence (+) or absence (-) of the required ARP peroxidase enzyme and analyzed by reducing SDS-PAGE followed by Coomassie blue staining. The arrow marks the migration of the monomer and crosslinked trimer, as indicated. (B) To confirm the formation of Dityrosine crosslinks, the purified crosslinked and uncrosslinked samples, obtained as described in A, were analyzed for DT-specific fluorescence: excitation wavelength: 320nm, emission wavelength: 405nm. (C) Binding of the soluble 403Y-433Y HA mutant, before and after crosslinking, to the broadly neutralizing V<sub>H</sub>1-69 stalk-specific mAb 8D4 by direct capture ELISA at 20 μg/ml of 8D4.

[0119] **Figure 86.** (A) 293T cells were transfected with plasmids for the expression of HA (WT and the indicated insertion mutants) and NA. Virus-like particles were analyzed by direct capture ELISA from transfected cell supernatants with a globular head antibody, PY-102. (B) Binding of the single insertion HA mutants (insertion at positions 63, 278, and 286) to a broadly neutralizing V<sub>H</sub>1-69 stalk-specific mAb by direct capture ELISA at 50 μg/ml (normalized for HA presence in supernatants). (C) VLPs obtained as described in A were purified over a 30% Sucrose-NTE cushion. 10 μg of total protein was then incubated in cleavage buffer in the presence (+) or absence (-) of TEV protease (Promega), according to the manufacturer’s instructions. Cleavage efficiency was monitored by Western Blot using an anti-HA2 antibody. Arrows indicate the cleavage product.

[0120] **Figure 87.** (A) 293T cells were transfected with plasmids for the expression of HA (WT and the indicated double-insertion mutants) and NA. Virus-like particles were analyzed by direct capture ELISA from transfected cell supernatants with a globular head antibody, PY-102. (B) Binding of the double-insertion HA mutants (insertions at positions 63+278, and 63+286) to two broadly neutralizing V<sub>H</sub>1-69 stalk-specific mAb by direct capture ELISA at 50 μg/ml (normalized for HA presence in supernatants).

[0121] **Figure 88.** Amino acid sequences of influenza HA C-terminal fragments generated following proteolysis at a protease cleavage site inserted at position 291 of the wild-type PR8 HA amino acid sequence (SEQ ID NO:1). The sequence of SEQ ID NO:108 comprises to-tyrosine mutations at positions 113 and 143 (underlined; corresponding to positions 403 and 433, respectively, of SEQ ID NO:1). The sequence of SEQ ID NO:109 comprises to-tyrosine mutations at positions 121 and 132 (underlined; corresponding to positions 411 and 422, respectively, of SEQ ID NO:1). The sequence of SEQ ID NO:110
comprises to-tyrosine mutations at positions 113, 121, 132 and 143 (underlined; corresponding to positions 403, 411, 422 and 433, respectively, of SEQ ID NO:1). The C-terminal transmembrane region is underlined in each sequence.

[0122] Figure 89. Amino acid sequence of an influenza HA protein C-terminal fragment (SEQ ID NO: 117). This fragment is generated following proteolysis at a protease cleavage site inserted at position 291 of the wild-type PR8 HA amino acid sequence (SEQ ID NO:1). Underlined amino acid residues N112, F115, K120, N131, D137, L141, D142 and W144 illustrate positions in SEQ ID NO: 117 where to-tyrosine mutations can be made to facilitate the formation di-tyrosine bonds. The underlined residues correspond to positions N403, F406, K411, N422, D429, L432, D433 and W435 respectively, of SEQ ID NO:1. The C-terminal transmembrane region is underlined. In some embodiments the transmembrane region is absent (i.e. the fragment does not contain the last 46 amino acid residues (229-274 of SEQ ID NO: 117, but contains residus 1-228 of SEQ ID NO: 117). Tyrosine residues at positions 17 and 146 (shown in bold italic) are endogenous tyrosine residues that may be used in the formation of dityrosine bonds. These endogenous residues correspond to tyrosine residues at positions 308 and 437, respectively, of SEQ ID NO:1.

DETAILED DESCRIPTION OF THE INVENTION

[0123] The present invention provides, in part, influenza HA polypeptides, proteins and/or protein complexes (such as those that comprise a stalk domain having its native conformation and that may or may not comprise an intact head domain), methods of making such polypeptides, proteins and/or protein complexes, compositions (such as pharmaceutical compositions and vaccine compositions) comprising such polypeptides, proteins and/or protein complexes, and methods of use of such polypeptides, proteins and/or protein complexes, for example in vaccination methods, therapeutic methods and other methods. In some embodiments, the influenza HA polypeptides, proteins and/or protein complexes may be useful as immunogens, for example in influenza vaccines.

Definitions and Abbreviations

[0124] As used in the present specification the terms “about” and “approximately,” when used in relation to numerical values, mean within + or – 20% of the stated value.
The abbreviation “HA” as used herein refers to a hemagglutinin protein. The abbreviation “Ab” as used herein refers to antibody. The abbreviation “bnAbs” as used herein refers to broadly neutralizing antibodies. The abbreviation “QNE” as used herein refers to quaternary neutralizing epitopes. The abbreviation “DT” as used herein refers to di-tyrosine. As used herein the phrase “full-length” when used in relation to an influenza HA protein or polypeptide does not require an HA protein or polypeptide that is as long as a wild-type influenza HA protein. Rather the term is used to refer to an influenza HA protein or polypeptide that comprises, at least, both a stalk domain and a head domain. Such stalk and head domains may or may not be as long as those found in a wild-type influenza HA protein or polypeptide. For example, an influenza HA protein or polypeptide that is missing the transmembrane domain found in a wild-type influenza HA protein or polypeptide may still be referred to as a “full-length” HA protein or polypeptide herein if it has a stalk domain and a head domain. In some embodiments, the phrase “full-length,” when used in relation to an influenza HA protein or polypeptide, may refer to an influenza HA protein or polypeptide that, in addition to a stalk and head domain, also comprises a transmembrane domain. As used herein the phrase “soluble” when used in relation to an influenza HA protein or polypeptide refers to an influenza HA protein or polypeptide that does not comprise a transmembrane domain. Such soluble HA proteins or polypeptides may comprise either a stalk domain and a head domain, or stalk domain in the absence of a head domain.

As used herein the terms “protein” and “polypeptide” are used interchangeably, unless otherwise stated. As used herein the term “protein complex” refers to an assembly of two or more proteins or protein subunits, such as two or more monomers. Unless otherwise stated, all description herein that relates to proteins and/or polypeptides applies equally to protein complexes, and vice versa.

As used herein the terms “stabilized” and “locked” are used interchangeably, for example in relation to the effect of cross-linking in stabilizing or locking the stalk domain of an influenza HA protein, polypeptide, or protein complex in its native trimeric conformation. These terms do not require 100% stability. Rather these terms denote a degree of improved or increased stability. For example, in some embodiments, when the term “stabilized” is used in relation to a stalk domain cross-linked in its native trimeric conformation, the term denotes that the native trimeric conformation of the stalk domain has
greater stability than it would have had prior to or without such cross-linking. Stability, and relative stability, may be measured in various ways as described in other sections of this application, for example based on the half-life of the native trimeric conformation of the stalk domain. The improvement or increase in stability may be to any degree that is useful or significant for the intended application. For example, in some embodiments stability may be increased by about 10%, 25%, 50%, 100%, 200% (i.e. 2-fold), 300% (i.e. 3-fold), 400% (i.e. 4-fold), 500% (i.e. 5-fold), 1000% (i.e. 10-fold), or more.

[0128] As used herein the terms “stem” and “stalk” are used interchangeably to refer to a stalk domain, or portion thereof, of an influenza HA protein or polypeptide.

[0129] As used herein the term “engineered” when used in relation to the influenza HA polypeptides, proteins and/or protein complexes of the invention refers generally to influenza HA polypeptides, proteins and/or protein complexes that have been altered in some way as compared to the wild-type versions of those polypeptides, proteins and/or protein complexes, for example, but not limited to, by way of removal or disruption of a particular portion or domain of the wild-type polypeptide, protein and/or protein complex (such as a transmembrane domain or a head domain) or by introduction of one or more point mutations (such as those introduced to facilitate formation of di-tyrosine bonds) or by way of introduction of one or more protease recognition motifs not ordinarily present in the wild-type polypeptide, protein and/or protein complex, or by any other modification of the polypeptide, protein and/or protein complex as compared to its wild-type form.

[0130] Other definitions and abbreviations are found throughout the specification.

Influenza and Influenza Viruses

[0131] Influenza, commonly known as "the flu", is an infectious disease of birds and mammals caused by RNA viruses of the family Orthomyxoviridae, the influenza viruses. Influenza spreads around the world in seasonal epidemics, resulting in about three to five million yearly cases of severe illness and about 250,000 to 500,000 yearly deaths, rising to millions in some pandemic years. In the 20th century three influenza pandemics occurred, each caused by the appearance of a new strain of the virus in humans, and killed tens of millions of people. Often, new influenza strains appear when an existing flu virus spreads to
humans from another animal species, or when an existing human strain picks up new genes from a virus that usually infects birds or pigs.

[0132] There are three different types of influenza virus, type A, type B, and type C, with various subtypes and strains within those types.

[0133] Influenza type A viruses are the most virulent human pathogens among the three influenza types and cause the most severe disease. The influenza A virus can be subdivided into different subtypes or serotypes including, but not limited to H1N1 (which caused Spanish Flu in 1918, and Swine Flu in 2009), H2N2 (which caused Asian Flu in 1957), H3N2 (which caused Hong Kong Flu in 1968), H5N1 (which caused Bird Flu in 2004), H7N7, H1N2 (which is endemic in humans, pigs and birds), H9N2, H7N2, H7N3, H10N7, and H7N9. Wild aquatic birds are the natural hosts for a large variety of influenza A. However, domestic poultry, such as turkeys and chickens, can also become very sick and die from avian influenza, and some avian influenza A viruses also can cause serious disease and death in wild birds.

[0134] Influenza type B almost exclusively infects humans and is less common than influenza A. The only other animals known to be susceptible to influenza B infection are the seal and the ferret. Influenza type B mutates at a rate 2–3 times slower than type A and consequently is less genetically diverse, with only one influenza B serotype known. As a result of this lack of antigenic diversity, a degree of immunity to influenza B is usually acquired at an early age. However, influenza B mutates frequently enough that lasting immunity is not possible.

[0135] Influenza type C virus infects humans, dogs and pigs, sometimes causing both severe illness and local epidemics. However, influenza C is less common than the other types and usually only causes mild disease.

[0136] Influenza viruses A, B and C are very similar in their overall structure. They each comprise a viral envelope containing two main types of glycoproteins, and a central core containing the viral RNA genome and other viral proteins. Hemagglutinin (“HA”) and neuraminidase (“NA”) are the two large envelope glycoproteins. HA is a lectin that mediates binding of the virus to target cells and entry of the viral genome into the target cell. The various influenza A subtypes are classified based on their antibody responses to the HA
and NA proteins. For example, an “H7N2 virus” designates an influenza A subtype that has an HA 7 protein and an NA 2 protein. Similarly an “H5N1” virus has an HA 5 protein and an NA 1 protein. There are currently around 17 known HA subtypes and around 10 known NA subtypes. Many different combinations of HA and NA proteins are possible. Influenza A subtypes H1N1, H1N2, and H3N2 are currently the main types in general circulation in the human population. There are also several prominent subtypes of the avian influenza A viruses that are known to infect both birds and humans – such as H5N1, H7N2, H7N7, H7N3, and H7N7 subtypes.

[0137] Within influenza type A, one can group the various different influenza subtypes in a variety of different ways, if desired. For example, influenza type A subtypes are frequently classified or grouped into different antigenic groups and antigenic subgroups based on their HA protein. Such groupings relate to the antigenicity and degree of HA sequence identity between the different subgroups. Influenza subtypes in the same antigenic group or antigenic subgroup are more similar to each other in terms of antigenicity and HA sequence than those in other antigenic groups. Antigenic group 1 consists of H1, H2, H5, H6, H8, H9, H11, H12, H13, and H16 influenza A subtypes. Antigenic group 2 consists of H3, H4, H14, H7, H10, and H15 influenza A subtypes. Within antigenic group 1, there are three antigenic subgroups, which will be referred to herein as antigenic subgroup 1A, 1B, and 1C. Antigenic subgroup 1A consists of H1, H2, H5 and H6 influenza A subtypes. Antigenic subgroup 1B consists of H11, H13 and H16 influenza A subtypes. Antigenic subgroup 1C consists of H8, H9, and H12 influenza A subtypes.

[0138] In some embodiments herein the HA polypeptides, proteins and protein complexes of the invention are generated from HA sequences from any influenza type – including type A, B, or C. In some embodiments herein the HA polypeptides, proteins and protein complexes of the invention are generated from HA sequences from influenza type A. In some embodiments herein the HA polypeptides, proteins and protein complexes of the invention are generated from HA sequences from influenza type A, antigenic group 1. In some embodiments herein the HA polypeptides, proteins and protein complexes of the invention are generated from HA sequences from influenza type A, antigenic group 1A.

[0139] In some embodiments herein the HA polypeptides, proteins and protein complexes of the invention can be used to vaccinate a subject, and provide protection
against, any influenza type – including type A, B, or C. In some embodiments herein the 
HA polypeptides, proteins and protein complexes of the invention can be used to vaccinate a 
subject, and provide protection against, influenza type A. In some embodiments herein the 
HA polypeptides, proteins and protein complexes of the invention can be used to vaccinate a 
subject, and provide protection against, influenza type A, antigenic group 1. In some 
embodiments herein the HA polypeptides, proteins and protein complexes of the invention 
can be used to vaccinate a subject, and provide protection against, influenza type A, 
antigenic group 1A. In some embodiments herein the HA polypeptides, proteins and protein 
complexes of the invention can be used to vaccinate a subject, and provide protection 
against, influenza subtype H1. In some embodiments herein the HA polypeptides, proteins 
and protein complexes of the invention can be used to vaccinate a subject, and provide 
protection against, influenza subtypes H1 and H2. In some embodiments herein the HA 
polypeptides, proteins and protein complexes of the invention can be used to vaccinate a 
subject, and provide protection against, influenza subtypes H1, H2 and H5. In some 
embodiments herein the HA polypeptides, proteins and protein complexes of the invention 
can be used to vaccinate a subject, and provide protection against, influenza subtypes H1, 
H2, H5 and H6.

[0140] Tables A and B below provide some examples of the sequence identity between 
the HA protein of H1N1 strain PR8, or certain fragments of the HA protein, and 
corresponding proteins or fragments from other influenza subtypes and strains, including 
some from antigenic groups 1 and 2.

[0141] Table A. Percent identity of full-length influenza HA amino acid sequences to the 
PR8 amino acid sequence of SEQ ID NO:1 (as illustrated in Figure 54).

<table>
<thead>
<tr>
<th>Antigenic group 1:</th>
<th>% identity to SEQ ID NO:1</th>
</tr>
</thead>
<tbody>
<tr>
<td>HA-USSR-77.pro (SEQ ID NO:111) (H1 subtype)</td>
<td>90.1</td>
</tr>
<tr>
<td>HA-Texas-91.pro (SEQ ID NO:112) (H1 subtype)</td>
<td>87.8</td>
</tr>
<tr>
<td>HA-WSN-33.pro (SEQ ID NO:113) (H1 subtype)</td>
<td>90.1</td>
</tr>
<tr>
<td>HA-SouthCarolina-1918.pro (SEQ ID NO:114) (H1)</td>
<td>88.2</td>
</tr>
<tr>
<td>Influenza HA protein C-terminal fragment</td>
<td>% identity to C-terminal fragment of SEQ ID NO:1</td>
</tr>
<tr>
<td>----------------------------------------</td>
<td>---------------------------------------------</td>
</tr>
<tr>
<td><strong>Antigenic group 1:</strong></td>
<td></td>
</tr>
<tr>
<td>HA-USSR-77 (H1 subtype)</td>
<td>92.4</td>
</tr>
<tr>
<td>HA-Texas-91 (H1 subtype)</td>
<td>90.2</td>
</tr>
<tr>
<td>HA-WSN-33 (H1 subtype)</td>
<td>89.6</td>
</tr>
<tr>
<td>HA-SouthCarolina-1918 (H1 subtype)</td>
<td>89.6</td>
</tr>
<tr>
<td>HA-California-09 (H1 subtype)</td>
<td>85.4</td>
</tr>
<tr>
<td>HA-Singapore-57 (H2 subtype)</td>
<td>71.6</td>
</tr>
<tr>
<td>HA-Vietnam-04 (H5 subtype)</td>
<td>70.4</td>
</tr>
<tr>
<td><strong>Antigenic group 2:</strong></td>
<td></td>
</tr>
<tr>
<td>HA-Udorn-72 (H3 subtype)</td>
<td>43.9</td>
</tr>
<tr>
<td>HA-HongKong-68 (H3 subtype)</td>
<td>44.5</td>
</tr>
<tr>
<td>HA-Panama-99 (H3 subtype)</td>
<td>43.3</td>
</tr>
<tr>
<td>HA-Wisconsin-05 (H3 subtype)</td>
<td>43.0</td>
</tr>
<tr>
<td>HA-Shanghai-13 (H7 subtype)</td>
<td>44.2</td>
</tr>
</tbody>
</table>

**Table B.** Percent identity of fragment that remains after cleavage of influenza HA protein at cleavage sites 48 and 291.
[0143] In addition to the sequence identities shown in the above table, the percent identity between PR8 (SEQ ID NO: 1) and sequences of H6, H9, H11, and H13 subtypes over a C-terminal fragment remaining after proteolytic cleavage of the influenza HA protein, was found to be 68.2%, 54.7%, 56.2%, and 50.5%, respectively.

**Influenza HA Polypeptides, Proteins and Protein Complexes**

[0144] In some embodiments the present invention provides engineered influenza HA polypeptides, proteins and/or protein complexes, compositions comprising such polypeptides, proteins and/or protein complexes, and methods of use of such polypeptides, proteins and/or protein complexes. Such proteins can be made using any suitable influenza virus HA protein as a starting point. For example, the proteins of the invention can be made using an influenza HA protein from any suitable influenza type (such as A, B, or C), subtype (including, but not limited to, H1N1, H1N2, and H3N2 subtypes) or strain (e.g. the H1N1 A/Puerto Rico/8/1934 ("PR8") strain (SEQ ID NO. 1)) of influenza virus as the starting point. One of the important features of the influenza HA polypeptides, proteins and/or protein complexes described herein is that they comprise the trimeric stalk domain of the HA protein which, unlike the highly variable head domain, is more conserved between influenza types, subtypes and strains. Accordingly, in addition to being useful as vaccine immunogens against homologous types, subtypes, and strains of influenza virus (i.e. against influenza viruses of the same type, subtype and/or strain as used as the starting point for making the influenza HA polypeptides, proteins and/or protein complexes described herein), the HA polypeptides, proteins and/or protein complexes of the invention may also be useful as vaccine immunogens against heterologous types, subtypes, and strains of influenza virus (i.e. against influenza viruses of a different type, subtype and/or strain to that used as the starting point for making the engineered HA polypeptides, proteins and/or protein complexes).

[0145] In some embodiments the present invention provides approaches for stabilizing the stalk domain of an influenza HA protein in its native trimeric conformation, including providing specific locations within the influenza HA protein that can be or should be cross-linked, and providing mutant forms of the HA protein that can facilitate the formation of such cross-links. Such cross-links and mutations can be used alone (e.g. in the context of a wild type HA protein or in the context of an HA protein that does not comprise any man-
made mutations or other man-made modifications), or can be used in combination with one or more other man-made mutations, modifications, cross-links, or stabilization strategies. Thus, for example, the approaches described herein can be used in conjunction with the use of added foldon trimerization domains, stabilizing antibodies (such as 6F12, C179, CR6261, F10, A66 and D8), and/or other partially or potentially stabilizing modifications or mutations.

[0146] The present inventors have performed extensive analysis of the structure of the influenza HA protein and have developed a variety of novel design strategies and novel engineered influenza HA polypeptides, proteins, and/or protein complexes. The present invention also provides methods for making and using such influenza HA polypeptides, proteins, and/or protein complexes. In some embodiments, the present invention provides specific locations within the amino acid sequence of the influenza HA protein at which, or between which, targeted cross-links can be made in order to “lock” the stalk domain of the HA protein in its native trimeric conformation. In some embodiments, the targeted cross-links are di-tyrosine cross-links. Where di-tyrosine cross-links are used, the present invention provides specific amino acid residues (or pairs of amino acid residues) that either comprise a pre-existing tyrosine residue or can be or are mutated to a tyrosine residue such that di-tyrosine cross-links can be made.

[0147] The engineered influenza HA polypeptides, proteins and/or protein complexes described herein can be made based on the sequence of any suitable influenza HA polypeptide, protein and/or protein complex, such as a wild-type (WT) influenza HA protein or polypeptide, or mutant, homolog, derivative, analog, ortholog, or any other derivative of an influenza HA polypeptide, protein and/or protein complex, provided that the HA polypeptide, protein and/or protein complex has a stalk domain, or a portion of a stalk domain, that is capable of folding into, or forming a part of, an stalk domain having a native trimeric conformation and/or is capable of binding to one or more anti-stalk antibodies. Amino acid sequences of suitable influenza HA polypeptides, proteins and/or protein complexes, and nucleic acid sequences that encode such influenza HA polypeptides, proteins and/or protein complexes, are known in the art and any such amino acid or nucleic acid sequence may be used. Furthermore, amino acid sequences of several suitable influenza HA polypeptides, proteins and/or protein complexes, and nucleic acid sequences that encode
such influenza HA polypeptides, proteins and/or protein complexes, are provided herein. While any suitable influenza virus HA protein can be used as a starting point for making the soluble influenza HA polypeptides, proteins and/or protein complexes described herein, such an HA protein should at least comprise a stalk domain, or a portion of a stalk domain, that is capable of folding into a native trimeric conformation and/or that is capable of binding to one or more anti-stalk antibodies, such as neutralizing anti-stalk antibodies. In some embodiments the HA protein used as a starting point is a full-length wild-type HA protein comprising a head domain and a stalk domain, and optionally also a transmembrane domain. In some embodiments the HA protein used lacks a transmembrane domain or lacks a functional or intact transmembrane domain. In some embodiments the HA protein comprises a T4 foldon trimerization motif. In some embodiments the HA proteins that are used as a starting point for making the influenza HA polypeptides, proteins and/or protein complexes described herein: (a) comprise a stalk domain, or a portion of a stalk domain, that is capable of folding into a native trimeric conformation and/or that is capable of binding to one or more neutralizing anti-stalk antibodies, (b) comprise a T4 foldon trimerization motif, and (c) lack a functional or intact transmembrane domain.

[0148] Throughout the present patent specification, when reference is made to specific amino acid residues or specific amino acid regions in the influenza HA protein by referring to their amino residue number or numbers (such as amino acid residues 403 and 422, for example), and unless otherwise stated, the numbering is based on the HA amino acid sequence provided herein, Figure 9 and SEQ ID NO: 1 – which is an amino acid sequence of a wild-type HA protein from influenza strain PR8 (influenza type A - H1N1 subtype). However, it should be noted, and one of skill in the art will understand, that different HA sequences may have different numbering systems, for example, if there are additional amino acid residues added or removed as compared to SEQ ID NO: 1 (for example, as illustrated in Figures 26 and 27 and many of the other Figures and sequences herein). As such, it is to be understood that when specific amino acid residues are referred to by their number, the description is not limited to only amino acids located at precisely that numbered position when counting from the beginning of a given amino acid sequence, but rather that the equivalent/corresponding amino acid residue in any and all HA sequences is intended - even if that residue is not at the same precise numbered position, for example if the HA sequence is shorter or longer than SEQ ID NO. 1, or has insertions or deletions as compared to SEQ
ID NO. 1. One of skill in the art can readily determine what is the corresponding/equivalent amino acid position to any of the specific numbered residues recited herein, for example by aligning a given HA sequence to SEQ ID NO. Thus, in embodiments where specific amino acid residues of the influenza HA protein are referred to, it is to be understood that the invention is not to be limited to sequences having the specified amino acid properties (e.g. presence of a tyrosine residue, a mutation, or an insertion of a protease recognition site, etc.) at only those precise numbered amino acid positions. Rather the specified amino acid properties may be located at any position in any influenza HA protein that is equivalent/corresponding to the numbered positions recited for the PR8 influenza HA protein of SEQ ID NO:1. This description applies equally where references are made to specific nucleic acid residues or specific nucleic acid regions in a nucleotide sequence encoding an influenza HA protein by referring to their nucleic acid residue number or numbers. Thus, unless otherwise stated, the numbering is based on the nucleotide sequence provided herein in Figure 10 and SEQ ID NO. 2.

[0149] In some embodiments, the influenza HA polypeptides, proteins or protein complexes of the present invention can be derived from (or can comprise, consist essentially of, or consist of) the amino acid sequences of any suitable influenza HA polypeptide, protein or protein complex sequence known in the art, including, without limitation: the amino acid sequence of the PR8 strain of H1N1 influenza virus (for example, in a full-length form (SEQ ID NO:1) or a soluble form (SEQ ID NO:80; or amino acid residues 1-519 thereof), the amino acid sequence of the Udorn 72 strain of H3N2 influenza virus (for example, in a full-length form (SEQ ID NO:73) or a soluble form comprising amino acid residues 1-520 thereof), the amino acid sequence of the Hong Kong 68 strain of H3N2 influenza virus (for example, in a full-length form (SEQ ID NO:74) or a soluble form (SEQ ID NO:81; or amino acid residues 1-520 thereof)), the amino acid sequence of the Panama 99 strain of H3N2 influenza virus (for example, in a full-length form (SEQ ID NO:75) or a soluble form comprising amino acid residues 1-520 thereof), the amino acid sequence of the Wisconsin 05 strain of H3N2 influenza virus (for example, in a full-length form (SEQ ID NO:76) or a soluble form (SEQ ID NO:82; or amino acid residues 1-520 thereof)), the amino acid sequence of the Shanghai 13 strain of H7N9 influenza virus (for example, in a full-length form (SEQ ID NO:77) or a soluble form (SEQ ID NO:84; or amino acid residues 1-514 thereof)), the amino acid sequence of the Singapore 57 strain of H2N2 influenza virus (for
example, in a full-length form (SEQ ID NO:78) or a soluble form (SEQ ID NO:85; or amino acid residues 1-515 thereof), the amino acid sequence of the Vietnam 04 strain of H5N1 influenza virus (for example, in a full-length form (SEQ ID NO:79) or a soluble form (SEQ ID NO:83; or amino acid residues 1-521 thereof), the amino acid sequence of the USSR 77 strain of H1N1 influenza virus (for example, in a full-length form (SEQ ID NO:111) or a soluble form comprising amino acid residues 1-519 thereof), the amino acid sequence of the Texas 91 strain of H1N1 influenza virus (for example, in a full-length form (SEQ ID NO:112) or a soluble form comprising amino acid residues 1-519 thereof), the amino acid sequence of the WSN 33 strain of H1N1 influenza virus (for example, in a full-length form (SEQ ID NO:113) or a soluble form comprising amino acid residues 1-518 thereof), the amino acid sequence of the South Carolina 1918 strain of H1N1 influenza virus (for example, in a full-length form (SEQ ID NO:114) or a soluble form comprising amino acids 1-519), the amino acid sequence of the California 09 strain of H1N1 influenza virus (for example, in a full-length form (SEQ ID NO:115) or a soluble form comprising amino acids 1-519), or any fragment thereof. In some embodiments, the influenza HA proteins and polypeptides of the present invention can be derived from (or can comprise, consist essentially of, or consist of) amino acid sequences that have at least about 40%, 50%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 98%, or 99% sequence identity to any known influenza HA sequences or to HA sequences from any known influenza groups, subgroups, families, subfamilies, types, subtypes, genera, species, strains, and/or clades, or any fragment thereof. Furthermore, in addition to the large number of specific amino acid and nucleotide molecules and sequences provided herein (including SEQ ID NO:s 1-110); the present invention also provides and encompasses amino acid and nucleotide molecules and sequences that have at least about 40%, 50%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 98%, or 99% sequence identity to any such molecules and sequences. Thus, for every embodiment herein that refers to a specific sequence or specific SEQ ID NO (such as SEQ ID NO:s 1-110), the present invention also includes variations of such embodiments that include amino acid and nucleotide molecules and sequences that have at least about 40%, 50%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 98%, or 99% sequence identity to such specific sequences or SEQ ID NOs.

[0150] In some embodiments the present invention provides engineered influenza HA polypeptides, proteins and/or protein complexes that comprise a stalk domain (for example,
that has, or is capable of forming, its native trimeric conformation) and that do not comprise a head domain. Such polypeptides, proteins and/or protein complexes can be referred to as “headless” influenza HA polypeptides, proteins and/or protein complexes.

[0151] In some embodiments the present invention provides influenza HA polypeptides, proteins and/or protein complexes that comprise a stalk domain having its native trimeric conformation and a head domain. Such proteins may be referred to as “head-on” influenza HA polypeptides, proteins and/or protein complexes. In some embodiments such polypeptides, proteins and/or protein complexes may also comprise one or more engineered protease recognition motifs that can be used for proteolytic disruption and/or removal of the head domain. In some embodiments such head-on influenza HA polypeptides, proteins and/or protein complexes may be useful as, for example, intermediates in the production of “headless” influenza HA polypeptides, proteins and/or protein complexes - as described herein.

[0152] “Headless” HA variants can be obtained or generated by a variety of methods. For example in some embodiments, headless HA variants can be obtained by removal of all or part of the HA head domain, for example by proteolytic removal of the head domain, or by another other suitable means. In other embodiments headless HA variants can be obtained by expression of a nucleotide sequence encoding only the stalk domain. In some embodiments, “headless” HA variants can be generated by proteolytic cleavage of a full-length influenza HA protein at protease recognition motifs inserted into the protein such that following cleavage, the head domain sequence is cut out and at least two protein fragments comprising the stalk domain remain. Figure 27 illustrates examples of protease cleavage motifs and shows intervening sequences of the head domain that are cut out following protease treatment. Thus, in some embodiments, for example as shown in Figure 27, a “headless” influenza HA variant comprises at least two protein fragments – e.g. an N-terminal fragment and a C-terminal fragment – comprising the stalk domain. In some embodiments one or more fragments of a “headless” influenza protein comprise one or more to-tyrosine mutations, and/or one-or more dityrosine crosslinks. Such mutations and/or crosslinks will typically be present in the C-terminal fragment of the “headless” HA protein. (See for example SEQ ID NOs: 96-110 and 117). Figures 70-84 and 89 illustrate examples of some such HA peptides. In some embodiments such peptides (for example, SEQ ID
NOs: 96-110 and 117) may be comprised within a larger HA molecule comprising a head domain, or they may be present in a “headless” HA protein. In some embodiments, several such peptides may associate to form an HA protein complex that is in, or is capable of forming, a trimeric stalk domain. In some embodiments an influenza HA polypeptide, protein and/or protein complex comprises the amino acid sequence of SEQ ID NO: 108, 109, 110. In some embodiments an influenza HA polypeptide, protein and/or protein complex comprises the amino acid sequence of SEQ ID NO:94 and SEQ ID NO:96, or the amino acid sequence of SEQ ID NO:94 and SEQ ID NO:97, or the amino acid sequence of SEQ ID NO:94 and SEQ ID NO:98, or the amino acid sequence of SEQ ID NO:94 and SEQ ID NO:99, or the amino acid sequence of SEQ ID NO:94 and SEQ ID NO:100, or the amino acid sequence of SEQ ID NO:94 and SEQ ID NO:101, or the amino acid sequence of SEQ ID NO:94 and SEQ ID NO:102, or the amino acid sequence of SEQ ID NO:94 and SEQ ID NO:103, or the amino acid sequence of SEQ ID NO:94 and SEQ ID NO:104, or the amino acid sequence of SEQ ID NO:95 and SEQ ID NO:106, or the amino acid sequence of SEQ ID NO:95 and SEQ ID NO:107, or the amino acid sequence of SEQ ID NO:95 and SEQ ID NO:108, or the amino acid sequence of SEQ ID NO:95 and SEQ ID NO:109, or the amino acid sequence of SEQ ID NO:95 and SEQ ID NO:110. In some embodiments an influenza HA polypeptide, protein and/or protein complex comprises an N-terminal HA peptide comprising, consisting essentially of, or consisting of, SEQ ID NO:94 or SEQ ID NO: 95, and a C-terminal HA peptide comprising, consisting essentially of, or consisting of, SEQ ID NO:96, 97, 98, 99, 100, 101, 102, 103, 104, 105, 106, 107, 108, 109 or 110, or a C-terminal HA peptide comprising, consisting essentially of, or consisting of, amino acid residues 229 to 519 of SEQ ID NO:1 wherein the amino acid sequence comprises a point mutation to tyrosine at one or more of amino acid positions 403, 406, 411, 422, 429, 432, 433, or 435, or a C-terminal HA peptide comprising, consisting essentially of, or consisting of, amino acid residues 1 to 228 of of SEQ ID NO:117 wherein the amino acid sequence comprises a point mutation to tyrosine at one or more of amino acid positions 112, 115, 120, 131, 137, 141, 142, or 144.

[0153] It should be noted that amino acid residues 1 through 58 (or 18 to 58 without the signal peptide – which is located at residues 1-17) and 292 through 566 (or 292 through 529 without the transmembrane domain and cytoplasmic tail) of the PR8 HA amino acid sequence (SEQ ID NO. 1) represent the influenza HA stalk domain sequence. The stalk
domain is discontinuous and comprises both an N-terminal and a C-terminal portion of the HA protein. The amino acid sequences provided here in may comprise additional domains that may be present or partially present or absent in some embodiments but not in others, for example the head domain (e.g. amino acid residues 59-291 of the PR8 HA amino acid sequence (SEQ ID NO. 1)), and/or the transmembrane and cytoplasmic region (e.g. amino acid residues 529 or 530 to 565 of the PR8 HA amino acid sequence (SEQ ID NO. 1)), and/or the signal peptide (e.g. amino acid residues 1-17 of the PR8 HA amino acid sequence (SEQ ID NO:1), and/or one or more optional exogenous (non-HA) sequences such as epitope tags, foldon domains, and the like. For example, in some embodiments an optional foldon trimerization domain, thrombin cleavage site, 6xHis-tag (SEQ ID NO: 118), and/or a strep tag may be present. In some embodiments these additional sequences may be absent, modified, rearranged or replaced. For example, in some embodiments different trimerization domains may be used, or different epitope tags may be used. In some embodiments these additional sequences may be absent, modified, rearranged or replaced, for example with different transmembrane or cytoplasmic domains.

[0154] In some embodiments the present invention provides influenza HA polypeptides, proteins, and/or protein complexes that are derived from, comprise, consist essentially of, or consist of any one of influenza HA amino acid sequences presented herein, or any variants or fragments thereof that have at least about 40% or 50% or 60% or 65% or 70% or 75% or 80% or 85% or 90% or 95% or 98% or 99% identity with such amino acid sequences presented herein, wherein the influenza HA polypeptides, proteins, and/or protein complexes comprise a tyrosine residue (whether naturally occurring or arising from a mutation to-tyrosine), at one or more of residues 308, 403, 406, 437, 411, 422, 429, 432, 433, and 435.

[0155] In some embodiments the present invention provides influenza HA polypeptides, proteins, and/or protein complexes that are derived from, comprise, consist essentially of, or consist of any one of influenza HA amino acid sequences presented herein, or any variants or fragments thereof that have at least about 40% or 50% or 60% or 65% or 70% or 75% or 80% or 85% or 90% or 95% or 98% or 99% identity with such amino acid sequences presented herein, wherein the influenza HA polypeptides, proteins, and/or protein complexes comprise an artificially-introduced protease cleavage site inserted immediately after one or more of the following residues: 48, 63, 228, 278, 282, 283, 286 and 291.
In some embodiments the present invention provides influenza HA polypeptides, proteins, and/or protein complexes that are derived from, comprise, consist essentially of, or consist of any one of influenza HA amino acid sequences presented herein, or any variants or fragments thereof that have at least about 40% or 50% or 60% or 65% or 70% or 75% or 80% or 85% or 90% or 95% or 98% or 99% identity with such amino acid sequences presented herein, wherein the influenza HA polypeptides, proteins, and/or protein complexes comprise two artificially-introduced protease cleavage sites, the first such site introduced immediately after residue 48 or 63, and the second such site introduced immediately after residue 228, 278, 282, 283, 286 or 291.

In some embodiments the present invention provides influenza HA polypeptides, proteins, and/or protein complexes that are derived from, comprise, consist essentially of, or consist of any one of influenza HA amino acid sequences presented herein, or any variants or fragments thereof that have at least about 40% or 50% or 60% or 65% or 70% or 75% or 80% or 85% or 90% or 95% or 98% or 99% identity with such amino acid sequences presented herein, wherein the influenza HA polypeptides, proteins, and/or protein complexes comprise both (a) a tyrosine residue (whether naturally occurring or arising from a mutation to-tyrosine), at one or more of residues 308, 403, 406, 437, 411, 422, 429, 432, 433, and 435, and (b) an artificially-introduced protease cleavage site inserted immediately after one or more of the following residues: 48, 63, 228, 278, 282, 283, 286 and 291.

In some embodiments the present invention provides influenza HA polypeptides, proteins, and/or protein complexes that are derived from, comprise, consist essentially of, or consist of any one of influenza HA amino acid sequences presented herein, or any variants or fragments thereof that have at least about 40% or 50% or 60% or 65% or 70% or 75% or 80% or 85% or 90% or 95% or 98% or 99% identity with such amino acid sequences presented herein, wherein the influenza HA polypeptides, proteins, and/or protein complexes comprise both (a) a tyrosine residue (whether naturally occurring or arising from a mutation to-tyrosine), at one or more of residues 308, 403, 406, 437, 411, 422, 429, 432, 433, and 435, and (b) two artificially-introduced protease cleavage sites - the first such site introduced immediately after residue 48 or 63, and the second such site introduced immediately after residue 228, 278, 282, 283, 286 or 291.
In some embodiments the present invention provides influenza HA polypeptides, proteins, and/or protein complexes that are derived from, comprise, consist essentially of, or consist of amino acid residues 229 to 519 of SEQ ID NO: 1, or 279 to 519 of SEQ ID NO: 1, or 283 to 519 of SEQ ID NO: 1, or 284 to 519 of SEQ ID NO: 1, or 287 to 519 of SEQ ID NO: 1, or 292 to 519 of SEQ ID NO: 1, or sequences that have at least about 40% or 50% or 60% or 65% or 70% or 75% or 80% or 85% or 90% or 95% or 98% or 99% identity with such amino acid sequences, wherein the influenza HA polypeptides, proteins, and/or protein complexes comprise a tyrosine residue (whether naturally occurring or arising from a mutation to-tyrosine), at one or more of residues 308, 403, 406, 437, 411, 422, 429, 432, 433, and 435.

In some embodiments the present invention provides influenza HA polypeptides, proteins, and/or protein complexes that are derived from, comprise, consist essentially of, or consist of amino acid residues 1 to 47 of SEQ ID NO: 1, or 1 to 62 of SEQ ID NO: 1, or sequences that have at least about 40% or 50% or 60% or 65% or 70% or 75% or 80% or 85% or 90% or 95% or 98% or 99% identity with such amino acid sequences.

In some embodiments the present invention provides compositions and/or influenza HA protein complexes that comprise, consist essentially of, or consist of a first and a second peptide, wherein (a) the first peptide comprises, consists essentially of, or consists of amino acid residues 229 to 519 of SEQ ID NO: 1, or 279 to 519 of SEQ ID NO: 1, or 283 to 519 of SEQ ID NO: 1, or 284 to 519 of SEQ ID NO: 1, or 287 to 519 of SEQ ID NO: 1, or 292 to 519 of SEQ ID NO: 1, or sequences that have at least about 40% or 50% or 60% or 65% or 70% or 75% or 80% or 85% or 90% or 95% or 98% or 99% identity with such amino acid sequences, and wherein the influenza HA polypeptides, proteins, and/or protein complexes comprise a tyrosine residue (whether naturally occurring or arising from a mutation to-tyrosine), at one or more of residues 308, 403, 406, 437, 411, 422, 429, 432, 433, and 435, and wherein (b) the second peptide comprises, consists essentially of, or consists of amino acid residues 1 to 47 of SEQ ID NO: 1, or 1 to 62 of SEQ ID NO: 1, or sequences that have at least about 40% or 50% or 60% or 65% or 70% or 75% or 80% or 85% or 90% or 95% or 98% or 99% identity with such amino acid sequences.
links, wherein at least one of the following amino acid residues within the influenza HA polypeptides, proteins, and/or protein complexes is artificially cross-linked to another amino acid residue in the influenza HA protein: Y308, N403, F406, Y437, K411, N422, D429, L432, D433, and W435. In some such embodiments, where the indicated position is not a tyrosine, that residue is mutated to tyrosine. In some such embodiments the cross-link is a di-tyrosine cross-link.

[0163] In some embodiments the present invention provides influenza HA polypeptides, proteins, and/or protein complexes that comprise one or more artificially-introduced cross-links, wherein such artificially introduced cross-links connect two of the following amino acid residues: Y308, N403, F406, K411, Y437, N422, D429, L432, D433, and W435. In some such embodiments, where the indicated position is not a tyrosine, that residue is mutated to tyrosine. In some such embodiment the cross-link is a di-tyrosine cross-link.

[0164] In some embodiments the present invention provides influenza HA polypeptides, proteins, and/or protein complexes in which the amino acid residues in one or more of the following pairs of amino residues are cross-linked to each other by an artificially introduced cross-link: 308/403, 308/435, 403/437, 403/429, 403/432, 403/433, 406/429, 406/433, 411/422, 422/433, 433/435 and 437/435. In some such embodiments, where the indicated position is not a tyrosine, that residue is mutated to tyrosine. In some such embodiments the cross-link is a di-tyrosine cross-link.

[0165] In some embodiments, the present invention contemplates the targeted introduction of one or more cross-links at any suitable position(s) in an influenza HA polypeptide, protein or protein complex, for example, in the stalk domain where the cross-link will or may stabilize the stalk domain in a native trimeric conformation or other conformation capable of binding anti-stalk antibodies, such as neutralizing or broadly neutralizing anti-stalk antibodies. Such stabilization may be achieved, for example, by introducing cross-links that stabilize interactions or folds within a stalk monomer or stalk protomer (intramolecular cross-link), and/or interactions between one or more stalk monomers or stalk protomers (intermolecular cross-link), or any combination of such crosslinks. In some such embodiments the cross-link is a di-tyrosine cross-link. For example, in some embodiments intermolecular di-tyrosine cross-links may be formed between tyrosine residues at positions 403 and 433, 411 and 422, or 433 and 435. Similarly,
in some embodiments intermolecular di-tyrosine cross-links may be formed between a tyrosine at residue 403 and another tyrosine residue, and/or between a tyrosine at residue 433 and another residue, such as, in particular, any of the other residues described herein as potential sites for di-tyrosine cross-links, such as tyrosines (whether natural or mutated) located at residues 308, 403, 406, 411, 422, 429, 432, 433, 435, or 437.

[0166] In some embodiments the present invention provides influenza HA polypeptides, proteins, and/or protein complexes comprising an artificially introduced cross-link between two of the following regions: amino acid residues from about position 298 to about 318, amino acid residues from about position 393 to about position 413, amino acid residues from about position 396 to about position 416, amino acid residues from about position 401 to about position 421, amino acid residues from about position 412 to about position 432, amino acid residues from about position 419 to about position 439, amino acid residues from about position 422 to about position 442, amino acid residues from about position 423 to about position 443, amino acid residues from about position 425 to about position 445 and amino acid residues from about 427 to about 447. In some such embodiments the cross-link is a di-tyrosine cross-link.

[0167] In embodiments where the influenza HA polypeptides, proteins, and/or protein complexes of the invention comprise one or more di-tyrosine cross-links, di-tyrosine cross-links may be introduced between two endogenous tyrosine residues, between two tyrosine residues originating from “to-tyrosine” mutations, or between a tyrosine residue originating from a “to-tyrosine” mutation and an endogenous tyrosine residue. In some embodiments, more than one di-tyrosine cross-link is introduced into an influenza HA protein or polypeptide.

[0168] In embodiments where the influenza HA polypeptides, proteins, and/or protein complexes of the invention comprise one or more di-tyrosine cross-links, non-limiting examples of amino acid positions where a “to-tyrosine” mutation can be introduced include N403, F406, K411, N422, D429, L432, D433, W435, or any combination thereof.

[0169] In embodiments where the influenza HA polypeptides, proteins, and/or protein complexes of the invention comprise one or more di-tyrosine cross-links, non-limiting
examples of preexisting or endogenous tyrosine residues that can be used to form a di-
tyrosine cross-link include Y308 and Y437, or any combination thereof.

[0170] In embodiments where the influenza HA polypeptides, proteins, and/or protein
complexes of the invention comprise one or more di-tyrosine cross-links, non-limiting
examples of residue pairs between which a di-tyrosine cross-link can be introduced include
403/429, 403/432, 403/433, 406/429, 406/433, 411/422 and 433/435, or any combination
thereof.

[0171] In embodiments where the influenza HA polypeptides, proteins, and/or protein
complexes of the invention comprise one or more di-tyrosine cross-links, non-limiting
examples of regions or secondary structures of the influenza HA protein from which amino
acids may be selected for tyrosine substitution and/or di-tyrosine cross-linking include the
stalk domain (e.g. amino acid residues 1 (with the signal peptide) or 18 (without the signal
peptide) to 58, and 292 to 529 (without the transmembrane and cytoplasmic domain(s)) or
566 (with the transmembrane domain). In some embodiments the lower region of the stalk
domain (comprising amino acid residues 18-46, 334-343, 344-390 and 449-503 of SEQ ID
NO:1), and/or the head domain (e.g. amino acid residues 59 to 291 of SEQ ID NO:1), of the
influenza HA polypeptides, proteins, and/or protein complexes of the invention may also
comprise one or more di-tyrosine cross-links and/or one or more to-tyrosine mutations.

[0172] Non-limiting examples of other regions of influenza HA proteins from which one
or more amino acids may be selected for tyrosine substitution and/or cross-linking include
amino acid residues from about position 298 to about position 313, amino acid residues from
about position 393 to about position 413, amino acid residues from about position 396 to
about position 416, amino acid residues from about position 401 to about position 421,
amino acid residues from about position 412 to about position 432, amino acid residues from
about position 419 to about position 439, amino acid residues from about position 422 to
about position 442, amino acid residues from about position 423 to about position 443,
amino acid residues from about position 425 to about position 445, and amino acid residues
from about position 427 to about position 447.

[0173] In some embodiments, the present invention provides influenza HA polypeptides,
proteins, and/or protein complexes that are derived from, comprise, consist essentially of, or
consist of, the amino acid sequence of SEQ ID NO: 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 86, 87, 88, 89, 90, 91, 92, 93, 96, 97, 98, 99, 100, 101, 102, 103, 104, 105, 106, 107, 108, 109 or 110 (each of which are mutants of the influenza HA amino acid sequence that comprise one or more protease recognition sequences to facilitate proteolytic cleavage of the head domain of the HA protein, and/or one or more “to tyrosine” mutations to facilitate di-tyrosine cross-linking and to facilitate “locking” of the stalk domain of the influenza HA protein in a particular conformation, for example, in its native trimeric conformation), or any fragment thereof, such as fragments comprising amino acid the stalk domain of the influenza HA protein, or any other fragments of the influenza HA protein that may be generated proteolytically and/or that may be assembled into or form part of a functional influenza HA protein. In some embodiments, the present invention provides influenza HA polypeptides, proteins, and/or protein complexes that are derived from, comprise, consist essentially of, or consist of, an amino acid sequence having at least about 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 98%, or 99% sequence identity to SEQ ID NO: 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 86, 87, 88, 89, 90, 91, 92, 93, 96, 97, 98, 99, 100, 101, 102, 103, 104, 105, 106, 107, 108, 109 or 110, or any fragment thereof.

[0174] Non-limiting examples of amino acid positions in an influenza HA protein or polypeptide to which di-tyrosine cross-links may be targeted include positions Y308 (pre-existing/endogenous Tyr residue) and N403Y (to-Tyr substitution), the positions Y308 (pre-existing/endogenous Tyr residue) and W435Y (to-Tyr substitution), the positions N403Y (to-Tyr substitution) and Y437 (pre-existing/endogenous Tyr residue), the positions N403Y (to-Tyr substitution) and D429Y (to-Tyr substitution), the positions N403Y (to-Tyr substitution) and L432Y (to-Tyr substitution), the positions N403Y (to-Tyr substitution) and D433Y (to-Tyr substitution), the positions N406Y (to-Tyr substitution) and D429Y (to-Tyr substitution), the positions N406Y (to-Tyr substitution) and D433Y (to-Tyr substitution), the positions D433Y (to-Tyr substitution) and W435Y (to-Tyr substitution), the positions K411Y (to-Tyr substitution) and W422Y (to-Tyr substitution), and the positions Y437 (pre-existing/endogenous Tyr residue) and W435Y (to-Tyr substitution). In some embodiments, the influenza HA polypeptides, proteins and/or protein complexes of the invention comprise one of the above listed di-tyrosine cross-links. In some embodiments,
the influenza HA polypeptides, proteins and/or protein complexes of the invention comprise two of the above listed di-tyrosine cross-links (for example, SEQ ID NO: 5, 8, 11, 14, and 17). In some embodiments, the influenza HA polypeptides, proteins and/or protein complexes of the invention comprise three of the above listed di-tyrosine cross-links. In some embodiments, the influenza HA polypeptides, proteins and/or protein complexes of the invention comprise four of the above listed di-tyrosine cross-links. In some embodiments, the influenza HA polypeptides, proteins and/or protein complexes of the invention comprise five or more of the above listed di-tyrosine cross-links. In some embodiments, the influenza HA polypeptides, proteins and/or protein complexes of the invention comprise any combination or one or more of the above listed di-tyrosine cross-links.

[0175] Non-limiting examples of influenza HA polypeptides, proteins and/or protein complexes designed to have more than one di-tyrosine cross-link include influenza HA proteins with two “to-tyrosine" mutations, where each such tyrosine residue forms a crosslink with different endogenous / preexisting tyrosine residues, or influenza HA proteins with four “to-tyrosine” mutations, e.g. N403Y/K411Y/N422Y/D433Y, as illustrated by SEQ ID NOs: 5, 8, 11, 14 and 17 where the tyrosine at position 403 is designed to pair with the tyrosine at position 411, and the tyrosine at position 422 is designed to pair with the tyrosine at position 433, thus stabilizing the stalk domain of the HA protein by the formation of two di-tyrosine cross-links.

[0176] A bond between a first HA polypeptide and second HA polypeptide within the same protein complex (e.g. monomers that arrange to form a trimer) is an example of an inter-molecular bond. The invention provides exemplary influenza HA proteins and polypeptides comprising cross-links designed to stabilize inter-molecular interactions, as well as influenza HA polypeptides, proteins or protein complexes derived from such sequences and including the specific “to-tyrosine” mutations present in such sequences. For example, one introduced tyrosine in one monomer is designed to pair with the other introduced tyrosine on the adjacent monomer.

[0177] In some embodiments, an HA polypeptide is intra-molecularly cross-linked (e.g., both tyrosines of the cross-link are located within the same HA polypeptide). The invention provides exemplary influenza HA proteins and polypeptides comprising cross-links designed to stabilize intra-molecular interactions, including without limitation, SEQ ID
NO:____, as well as influenza HA polypeptides, proteins or protein complexes derived from such sequences and including the specific “to-tyrosine” mutations present in such sequences.

[0178] In some embodiments (including all of those described above, and those involving influenza HA polypeptides, proteins, and/or protein complexes having any of the specific amino acid sequences recited herein, and those involving variants or fragments of such influenza HA polypeptides, proteins, and/or protein complexes having less than 100% identity to the specific amino acid sequences provided herein), the influenza HA polypeptides, proteins, and/or protein complexes of the invention should have one or more desired properties, such as being capable of (1) forming a native trimeric conformation of the stalk domain, (2) having the stalk domain “locked” in a native trimeric conformation by cross-linking, (3) binding to an influenza HA stalk-specific antibody, (4) binding to a neutralizing antibody, (5) binding to a broadly neutralizing antibody, (6) binding to an antibody selected from the group consisting of 6F12, C179, CR6261, F10, A66, and D8, (7) binding to and/or activating a B cell receptor, (8) eliciting an antibody response in an animal, (9) eliciting a protective antibody response in an animal, (10) eliciting production of neutralizing antibodies in an animal, (11) eliciting production of broadly neutralizing antibodies in an animal, (12) eliciting production of antibodies that recognize quaternary neutralizing epitopes (QNEs) in an animal, and/or (13) eliciting a protective immune response in an animal. In some embodiments the influenza HA polypeptides, proteins, and/or protein complexes described herein are capable of eliciting a protective immune response against one or more influenza virus strains in an animal and/or capable of eliciting a protective immune response against both homologous and heterologous influenza virus strains in an animal.

[0179] Unless otherwise stated, all description herein that relates to specific influenza HA polypeptides, proteins, and protein complexes, relates equally to all homologs, orthologs, analogs, derivatives, mutant forms, fragments, chimeras, fusion proteins etc. thereof, such as those that have certain desired properties or features (for example those that have a stalk domain, or a portion of a stalk domain, that is capable of folding into a native trimeric conformation, or that have desired functional properties, including, but not limited to, being capable of binding to, or eliciting the production of, one or more anti-HA antibodies, such as antibodies that are specific to the influenza HA stalk domain).
Similarly, all description herein that relates to specific polypeptides, proteins, and/or protein complexes polypeptides, proteins, and/or protein complexes (e.g. those having specific amino acid sequences or those from a specific influenza type, subtype, or strain) relates equally to other related forms of such polypeptides, proteins, and/or protein complexes that may exist in nature (for example in different influenza types, subtypes or strains) or that are related to the specific sequences provides herein but have been altered artificially in some way, such as by recombinant means, chemical means, or any other means. The influenza HA polypeptides, proteins, and/or protein complexes described herein can have, or can be derived from, the nucleotide and/or amino acid sequences of any suitable influenza HA polypeptides, proteins, and/or protein complexes known in the art. In some embodiments, the influenza HA polypeptides, proteins, and/or protein complexes of the invention may be, or may be derived from, derivatives and/or analogs of specific influenza HA polypeptides, proteins, and/or protein complexes described herein or known in the art, including proteins that are substantially homologous to any such proteins, or fragments thereof (e.g., in various embodiments, those having at least about 50% or 55% or 60% or 65% or 70% or 75% or 80% or 85% or 90% or 95% or 98% or 99% identity with an amino acid or nucleic acid sequence of any specific influenza HA polypeptides, proteins, and/or protein complexes described herein or known in the art, when aligned using any suitable method known to one of ordinary skill in the art, such as, for example, using a computer homology program known in the art) or whose encoding nucleic acid is capable of hybridizing to a coding nucleic acid sequence of a protein of the invention, under high stringency, moderate stringency, or low stringency conditions.

In some embodiments, the invention provides fragments of the influenza HA polypeptides, proteins, and/or protein complexes described herein, such as those comprising, consisting essentially of, or consisting of, at least about 10 amino acids, 20 amino acids, 50 amino acids, 100 amino acids, 200 amino acids, or 500 amino acids.

In some embodiments one or more amino acid residues within a specific influenza HA polypeptide, protein, or protein complex as described herein, or as known in the art, can be deleted, added, or substituted with another amino acid. In embodiments where such mutations are introduced, the influenza HA polypeptides, proteins, or protein complexes can be micro-sequenced to determine a partial amino acid sequence. In other embodiments the
nucleic acid molecules encoding the influenza HA polypeptides, proteins, and/or protein complexes can be sequenced to identify and/or confirm the introduction of mutations.

[0183] In some embodiments, one or more amino acid residues can be substituted by another amino acid having a similar polarity and that may acts as a functional equivalent, resulting in a silent alteration. In some embodiments substitutions for an amino acid within the sequence may be selected from other members of the class to which the amino acid belongs e.g. to create a conservative substitution. For example, the nonpolar (hydrophobic) amino acids include alanine, leucine, isoleucine, valine, proline, phenylalanine, tryptophan and methionine. The polar neutral amino acids include glycine, serine, threonine, cysteine, tyrosine, asparagine, and glutamine. The positively charged (basic) amino acids include arginine, lysine and histidine. The negatively charged (acidic) amino acids include aspartic acid and glutamic acid. Such substitutions are generally understood to be conservative substitutions.

[0184] In some embodiments artificial, synthetic, or non-classical amino acids or chemical amino acid analogs can be used to make the influenza HA polypeptides, proteins, and/or protein complexes of the invention. Non-classical amino acids include, but are not limited to, the D-isomers of the common amino acids, fluoro-amino acids, and “designer” amino acids such as β-methyl amino acids, Cγ -methyl amino acids, Nγ -methyl amino acids, and amino acid analogs in general. Additional non-limiting examples of non-classical amino acids include, but are not limited to: α- aminocaprylic acid, Acpa; (S)-2-aminoethyl-L-cysteine/HCl, Aecys; aminophenylacetate, Afa; 6-amino hexanoic acid, Alhx; γ-amino isobutyric acid and α-aminoisobutyric acid, Aiba; alloisoleucine, Aile; L-allylglycine, Alg; 2-amino butyric acid, 4-aminobutyric acid, and α -aminobutyric acid, Aba; p-amino phenylalanine, Aphe; b-alanine, Bal; p-bromophenylalane, Brphe; cyclohexylalanine, Cha; citrulline, Cit; β-chloroalanine, Clala; cycloleucine, Cle; p-chlorophenylalanine, Clphe; cysteic acid, Cya; 2,4-diaminobutyric acid, Dab; 3-amino propionic acid and 2,3-diaminopropanic acid, Dap; 3,4-dehydroproline, Dhp; 3,4-dihydroxylphenylalanine, Dhphe; p-fluorophenylalanine, Fphe; D-glucoseaminic acid, Gaa; homoarginine, Hag; δ-hydroxylysine/HCl, Hlys; DL-β-hydroxynorvaline, Hnvl; homoglutamine, Hog; homophenylalanine, Hoph; homoserine, Hos; hydroxyproline, Hpr; p-iodophenylalanine, Iphe; isoserine, Ise; α-methylleucine, Mle; DL-methionine-S-methylsulfoniumchloide,
Msmet; 3-(1-naphthyl) alanine, 1Nala; 3-(2-naphthyl) alanine, 2Nala; norleucine, Nle; N-methylalanine, Nma; Norvaline, Nva; O-benzylserine, Obser; O-benzyltyrosine, Obtyr; O-ethyltyrosine, Oetyr; O-methylserine, Omser; O-methylthreonine, Omthr; O-methyltyrosine, Omtyr; Ornithine, Orn; phenylglycine; penicillamine, Pen; pyroglobutamic acid, Pga; piperolic acid, Pip; sarcosine, Sar; t-butylglycine; t-butyllalanine; 3,3,3-trifluoralanine, Tfa; 6-hydroxydopa, Thphe; L-vinylglycine, Vig; (-)-(2R)-2-amino-3-(2-aminoethylsulfanyl)propanoic acid dihydrochloride, Aaspa; (2S)-2-amino-9-hydroxy-4,7-dioxanonoic acid, Ahdna; (2S)-2-amino-6-hydroxy-4-oxahexanoic acid, Ahoa; (-)-(2R)-2-amino-3-(2-hydroxyethylsulfanyl)propanoic acid, Ahsopa; (-)-(2R)-2-amino-3-(2-hydroxyethylsulfanyl)propanoic acid, Ahspa; (2S)-2-amino-12-hydroxy-4,7,10-trioxadodecanoic acid, Ahtda; (2S)-2,9-diamino-4,7-dioxanonoic acid, Dadna; (2S)-2,12-diamino-4,7,10-trioxadodecanoic acid, Datda; (S)-5,5-difluoronleucine, Dfnl; (S)-4,4-difluoronorvaline, Dfnv; (3R)-1,1-dioxo-[1,4]thiaziane-3-carboxylic acid, Dtea; (S)-4,4,5,5,6,6,6-heptafluoronleucine, Hfnl; (S)-5,5,6,6,6-pentafluoronleucine, Pfnl; (S)-4,4,5,5,5-pentafluoronorvaline, Pfnv; and (3R)-1,4-thiazinane-3-carboxylic acid, Tca.

Furthermore, the amino acid can be D (dextrorotary) or L (levorotary). For a review of classical and non-classical amino acids, see Sandberg et al., 1998 (Sandberg et al., 1998). New chemical descriptors relevant for the design of biologically active peptides. A multivariate characterization of 87 amino acids. J Med Chem 41(14): pp. 2481-91.

**Nucleic Acids**

[0185] In addition to providing certain influenza HA polypeptides, proteins, and/or protein complexes, as described herein, the present invention also provides nucleic acids encoding such influenza HA polypeptides, proteins, and/or protein complexes, and compositions and vectors comprising such nucleic acids. Such nucleic acids can be obtained or made using any suitable method known in the art. For example, nucleic acid molecules encoding influenza HA polypeptides, proteins, and/or protein complexes may be obtained from cloned DNA or made by chemical synthesis. In some embodiments the nucleic acids may be obtained by reverse transcribing RNA prepared by any of the methods known to one of ordinary skill in the art, such as random- or poly A-primed reverse transcription.

Whatever the source, a nucleic acid molecule encoding an influenza HA polypeptide, protein, and/or protein complex of the present invention can be cloned into any suitable vector, such as those to be used for propagation of the nucleic acid molecule or those to be
used for expression of the nucleic acid molecule. The nucleic acid may be cleaved at specific sites using various restriction enzymes, if needed. In embodiments requiring expression, the nucleic acid can be operatively linked to a promoter suitable for directing expression in the desired cell type, such as a mammalian cell or an insect cell, and may be incorporated into any suitable expression vector, such as a mammalian or insect expression vector. A nucleic acid molecule encoding an influenza HA polypeptide, protein, and/or protein complex of the present invention optimized by methods known in the art to improve expression levels of the protein expressed therefrom. For example, codon optimization may be used to minimize or eliminate variations in codon usage between species. In some embodiments an influenza HA polypeptide, protein, and/or protein complex of the present invention is derived from a nucleic acid molecule that has been codon optimized for expression in humans (see, for example, SEQ ID NO.63 and Figure 48), Cricetulus griseus (see, for example, SEQ ID NO.64 and Figure 49), Nicotiana benthamiana (see, for example, SEQ ID NO.65 and Figure 50), Pichia pastoris (see, for example, SEQ ID NO.66 and Figure 51), Saccharomyces cerevisiae (see, for example, SEQ ID NO.67 and Figure 52) or Spodoptera frugiperda (see, for example, SEQ ID NO.68 and Figure 53).

[0186] In some embodiments, the present invention provides nucleic acids that are derived from, comprise, consist essentially of, or consist of, the nucleic acid sequence of SEQ ID NO: 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, or 62 (each of which encodes a mutant of the influenza HA amino acid sequence that comprises one or more protease recognition sequences to facilitate proteolytic cleavage of the head domain of the HA protein, and/or one or more “to tyrosine” mutations to facilitate di-tyrosine cross-linking and to facilitate “locking” of the stalk domain of the influenza HA protein in a particular conformation, for example, in its native trimeric conformation), or any fragment thereof, such as fragments encoding the stalk domain of the influenza HA protein.

[0187] Furthermore, one or skill in the art can readily visualize, or make, nucleic acid molecules that comprise any one or more of the specific “to-tyrosine” mutations described herein, for example, by locating the nucleotide codon that encodes the specific amino acid residue to be mutated, and mutating the nucleotides at that codon as necessary to result in a tyrosine-encoding codon.
Cross-Linking

[0188] In some embodiments the influenza HA polypeptides and/or proteins of the invention are assembled into protein complexes having a desired conformational structure, such as the native trimeric conformation of the stalk domain, and are cross-linked in order to stabilize that conformation. Details of particular regions of the influenza HA protein that can be cross-linked, as well as particular influenza HA mutants designed to facilitate such cross-linking, are described in other sections of this application. In some embodiments the cross-links may be used to stabilize the tertiary and/or quaternary structures of the influenza HA protein. In some embodiments, the cross-linking may be intra- and/or intermolecular cross-linking. In some embodiments, the cross-links that are used are targeted cross-links. In some embodiments, the cross-links that are used are stable under physiological conditions. In some embodiments, the cross-links that are used do not lead to aggregate formation of the influenza HA protein, for example during expression and/or during storage (such as storage of compositions comprising high concentrations of the influenza HA protein). In some embodiments the introduction of such cross-links may enhance the effectiveness of the influenza HA polypeptides, proteins and proteins of the invention as immunogens, such as vaccine immunogens. In some embodiments the introduction of such cross-links may stabilize epitopes within the influenza HA protein, for example, epitopes in the stalk domain, such that the epitopes can be recognized by particular antibodies, elicit production of antibodies, and/or activate B cell receptors upon antibody binding.

[0189] In some embodiments targeted cross-linking can be used. A targeted cross-link is one that can be made to form at a particular position or positions within the influenza HA protein or protein complex. Several strategies may be used to target cross-links to specific locations in an influenza HA protein or polypeptide, such as the specific locations described herein. The present invention provides residue pairs within the influenza HA protein that, when cross-linked, can or may stabilize an influenza HA polypeptide, protein, or protein complex in a conformation that is capable of binding to, or eliciting the production of, neutralizing antibodies, and/or that is capable of generating a neutralizing antibody response in an animal. A targeted cross-link may be introduced at one or more of the locations or positions specified herein by exploiting the physical and/or chemical properties of certain amino acid side chains, for example by making use of enzymatic reactions that recognize specific amino acid sequences or three-dimensional structures, or by incorporating non-
natural amino acids that have the ability to form cross-links in a folded protein or protein complex.

[0190] Cross-links or modifications may be targeted to specific sites in the structure of the influenza HA protein or polypeptide, for example the stalk domain, in order to achieve the desired outcome, e.g. stabilization of the stalk domain in its native trimeric conformation. The present invention contemplates the targeted introduction of one or more cross-links and/or other stabilizing modifications at any suitable position(s) in an influenza HA protein or polypeptide, preferably where the cross-link or modification stabilizes the stalk domain in its native trimeric conformation, or provides enhanced stabilization of the native trimeric conformation of the stalk domain. The invention contemplates that any influenza HA protein amino acid residue, residue pair, secondary structure or other region described herein for di-tyrosine cross-linking may also be used in the formation of other targeted cross-links or bonds or other modifications, including but not limited to amino acid positions Y308, N403, N406, K411, W422, D429, L432, D433, W435, and Y437 or any combination thereof; residue pairs 308/403, 308/435, 403/437, 403/429, 403/432, 403/433, 406/429, 406/433, 411/422, 433/435 and 437/435, or any combination thereof; regions or secondary structures including, for example the HA protein stalk domain or head domain; and other regions of influenza HA protein including the transmembrane domain or the lower region of the stalk domain.

[0191] In some embodiments the influenza HA polypeptides, proteins and protein complexes of the invention comprise cross-links in the stalk domain, such cross-links need not be located only in the stalk domain. In some embodiments cross-links may be located anywhere throughout the influenza HA polypeptide, protein or protein complex, including the head domain in “head-on” polypeptides, proteins and/or protein complexes, as desired. Preferably, an influenza HA polypeptide, protein and/or protein complex comprising cross-links in other regions (e.g. outside of the stalk domain) will retain one or more desired properties such as being capable of (1) forming a native trimeric conformation of the stalk domain, (2) having the stalk domain “locked” in a native trimeric conformation by cross-linking, (3) binding to an influenza HA stalk-specific antibody, (4) binding to a neutralizing antibody, (5) binding to a broadly neutralizing antibody, (6) binding to an antibody selected from the group consisting of 6F12, C179, CR6261, F10, A66, and D8, (7) binding to and/or
activating a B cell receptor, (8) eliciting an antibody response in an animal, (9) eliciting a protective antibody response in an animal, (10) eliciting production of neutralizing antibodies in an animal, (11) eliciting production of broadly neutralizing antibodies in an animal, (12) eliciting production of antibodies that recognize quaternary neutralizing epitopes (QNEs) in an animal, and/or (13) eliciting a protective immune response in an animal.

[0192] A wide variety of methods of cross-linking proteins intra- and inter-molecularly are known in the art, including those having cross-links with varying lengths of spacer arms, and those with and without fluorescent and functional groups for purification. Such methods include, but are not limited to, the use of heterobifunctional cross-linkers (e.g. succinimidoacetylthioacetate (SATA), trans-4-(maleimidylmethyl) cyclohexane-1-carboxylate (SMCC), and succinimido 3-(2-pyridylthio)propionate (SPDP)), homobifunctional cross-linkers (e.g. succinimido 3-(2-pyridylthio)propionate), photoreactive cross-linkers (e.g. 4-azido-2,3,5,6-tetrafluorobenzoic acid, STP ester, sodium salt (ATFB, STP ester), 4-azido-2,3,5,6-tetrafluorobenzoic acid, succinimido ester (ATFB, SE), 4-azido-2,3,5,6-tetrafluorobenzyl amine, hydrochloride, benzophenone-4-isothiocyanate, benzophenone-4-maleimide, 4-benzoxybenzoic acid, succinimido ester, N-((2-pyridylthio)ethyl)-4-azidosalicylamide (PEAS; AET), thiol reactive cross-linkers (e.g. maleimides and iodoacetamides), amine reactive cross-linkers (e.g. glutaraldehyde, bis(imido esters), bis(succinimido esters), diisocyanates and diacid chlorides). Because thiol groups are highly reactive and relatively rare in most proteins by comparison to amine groups, thiol-reactive cross-linking may be used in some embodiments. In cases where thiol groups are missing or not present at appropriate sites in the structures of influenza HA protein, they can be introduced using one of several thiolation methods. For example, Succinimido trans-4-(maleimidylmethyl)cyclohexane-1-carboxylate can be used to introduce thiol-reactive groups at amine sites.

[0193] Several oxidative cross-links are known, such as disulfide bonds (which form spontaneously and are pH and redox sensitive), and di-tyrosine bonds (which are highly stable, and irreversible under physiological conditions).

[0194] In some embodiments the cross-links stabilize the tertiary structure of an influenza HA protein. In some embodiments the cross-links stabilize the quaternary structure of an
influenza HA protein. In some embodiments the cross-links stabilize both the tertiary and quaternary structure of an influenza HA protein.

[0195] In some embodiments an influenza HA polypeptide, protein and/or protein complex of the invention has cross-links that are thermostable. In some embodiments an influenza HA polypeptide, protein and/or protein complex of the invention has cross-links that are not toxic. In some embodiments an influenza HA polypeptide, protein and/or protein complex of the invention has cross-links that are targeted cross-links, or non-targeted cross-links, or reversible cross-links, or irreversible cross-links, or cross-links formed by use of homo-bifunctional cross-linking agents, or cross-links formed by use of hetero-bifunctional cross-linking agents, or cross-links formed by use of reagents that react with amine groups, or cross-links formed by use of reagents that react with thiol groups, or cross-links formed by use of reagents that are photoreactive, or cross-links formed between amino acid residues, or cross-links formed between mutated amino acid residues incorporated into the structure of the proteins or protein complexes, or oxidative cross-links, or di-tyrosine bonds, or glutaraldehyde cross-links, or any combination thereof. In some embodiments an influenza HA polypeptide, protein and/or protein complex of the invention does not have glutaraldehyde cross-links.

[0196] In some embodiments an influenza HA polypeptide, protein and/or protein complex of the invention does not have any artificially-introduced disulfide bonds, or if it does have such disulfide bonds, also has additional artificially-introduced cross-links. In some embodiments an influenza HA polypeptide, protein and/or protein complex of the invention does not have any artificially introduced disulfide bonds, but may have naturally occurring disulfide bonds. Disulfide bonds can be introduced artificially when cysteine side-chains are engineered by point mutation. Disulfide bonds are, however, known to be pH sensitive and to be dissolved under certain redox conditions, and the preventative and/or therapeutic utility of proteins and/or protein complexes engineered with disulfide cross-links, for example to be used as immunogens in vivo, may therefore be compromised. Furthermore, undesired disulfide bonds often form between proteins with free sulphydryl groups that mediate aggregate formation (see, for example, Harris RJ et al. 2004, Commercial manufacturing scale formulation and analytical characterization of therapeutic recombinant antibodies. Drug Dev Res 61 (3): 137 – 154; Costantino & Pikal (Eds.), 2004.

[0197] If the structure and/or immunogenicity of an influenza HA polypeptide, protein and/or protein complex is compromised or altered by a cross-link, maintaining its overall structure and function can be achieved by controlling the availability of amino acid side-chains for the cross-linking reaction or by introducing additional cross-links or other stabilizing modifications. For example, in the case of DT cross-linking, tyrosyl side-chains that are available for reaction, but that lead to the distortion of the structure of the complex, and that compromise the immunogenicity/antigenicity of the influenza HA protein, can be removed by mutating such residues to another amino acid such as, for example, phenylalanine. Furthermore, point mutations may be introduced at positions where the amino acid side-chains will react with cross-linking agents or each other, such that the formation of the bond(s) causes the most beneficial outcome. These positions may also be identified as described herein.

[0198] When at a selected residue a reactive side-chain is not already present, a point mutation may be introduced, for example using molecular biological methods to introduce such a point mutation into the cDNA of a nucleic acid directing its expression, such that a reactive side-chain is present and available for the reaction.

[0199] Cross-links that may be used include, but are not limited to, reversible cross-links resulting from the use of homo- and hetero-bifunctional cross-linking agents that react with amine and/or thiol groups, photoreactive cross-link reagents, any cross-links that may form between non-classical amino acids incorporated into the structure of an influenza HA polypeptide, protein and/or protein complex, any oxidative cross-links, such as, but not limited to, di-tyrosine cross-links / bonds, heterobifunctional cross-linkers (e.g. succinimidyl
acethylthioacetate (SATA), trans-4-(maleimidylmethyl) cyclohexane-1-carboxylate (SMCC), and succinimidyl 3-(2-pyridyldithio)propionate (SPDP)), homobifunctional cross-linkers (e.g. succinimidyl 3-(2-pyridyldithio)propionate), photoreactive cross-linkers (e.g. 4-azido-2,3,5,6-tetrafluorobenzoic acid, STP ester, sodium salt (ATFB, STP ester), 4-azido-2,3,5,6-tetrafluorobenzoic acid, succinimidyl ester (ATFB, SE), 4-azido-2,3,5,6-tetrafluorobenzyl amine, hydrochloride, benzophenone-4-isothiocyanate, benzophenone-4-maleimide, 4-benzoxybenzoic acid, succinimidyl ester, N-[(2-pyridyldithio)ethyl]-4-azidosalicylamide (PEAS; AET), thiol reactive cross-linkers (e.g. maleimides and iodoacetamides), amine reactive cross-linkers (e.g. glutaraldehyde, bis(imido esters), bis(succinimidyl esters), diisocyanates and diacid chlorides).

[0200] The present invention also contemplates the introduction of targeted non-covalent tyrosine-stacking interactions as "cross-links" to stabilize protein-protein interactions and/or desired protein or peptide conformations, such as the native trimeric conformation of a stalk domain of an influenza HA polypeptide, protein and/or protein complex. The cross-link comprises a targeted pi stacking interaction including but not limited to a T-shaped, sandwich, or parallel displaced pi stacking interaction between the aromatic side chains of an introduced/engineered tyrosine and an endogenous tyrosine, phenylalanine, histidine, or tryptophan within the influenza HA polypeptide, protein and/or protein complex, or between the aromatic side chain of an introduced/engineered tyrosine and a second introduced/engineered tyrosine within the influenza HA polypeptide, protein and/or protein complex.

[0201] Irreversible cross-links, as used in the context of this application, include those that are not significantly dissolved under physiologically relevant conditions. It is preferred that the type of cross-links used should not lead to aggregate formation during expression or when the influenza HA polypeptides, proteins and/or protein complexes of the invention are stored at high concentrations. Disulfide bonds are not irreversible cross-links. Rather they are reversible cross-links and may dissolve under physiologically relevant conditions and/or lead to aggregate formation during protein expression and/or production or when stored in high concentrations.

[0202] In some embodiments cross-links may be targeted to the specific regions of influenza HA polypeptides, proteins and/or protein complexes described herein in order to
achieve the desired conformational stabilization and/or the desired immunogenic properties (e.g. the ability to maintain the stalk domain in its native trimeric conformation and/or to bind to broadly neutralizing antibodies). Alternatively, proteins with the cross-links at the locations specified herein may be isolated from a mixture of cross-linked and un-cross-linked proteins with and without desired modifications, for example based on chemical, physical, and/or functional characteristics. Such characteristics may include, for example, trimerization, the presence of a stalk domain having a native trimeric conformation, and/or any desired antigenic, immunogenic, or biochemical characteristics.

[0203] Alternatively, in some embodiments, cross-links may not be targeted, and proteins with the desired cross-links or properties may be isolated from a mixture of modified and unmodified proteins made using a non-targeted cross-linking system.

[0204] In embodiments where influenza HA polypeptides, proteins or protein complexes with the desired cross-links are to be isolated from a mixture of cross-linked and un-cross-linked proteins, such isolation or separation may be performed on the basis of one or more characteristics including, but not limited to, molecular weight, molecular volume, chromatographic properties, mobility in electrophoresis, antigenic and biochemical characteristics, fluorescence characteristics, solubility, binding to antibodies, structural characteristics, immunological characteristics, or any other suitable characteristics.

[0205] In addition to the specific cross-linking positions described herein, additional positions within influenza HA polypeptides, proteins or protein complexes can be identified at which further cross-links can be made, for example where a reactive side-chain would be able to form a bond with a reactive side-chain elsewhere on the influenza HA polypeptide, protein or protein complex. In some embodiments, such additional positions can be selected, for example, to maintain or improve the immunogenicity/antigenicity of the protein, polypeptide or protein complex. In some embodiments, such additional positions to be cross-linked can be selected in pairs.

Di-tyrosine (DT) Cross-Linking

[0206] In some embodiments the present invention provides influenza HA polypeptides, proteins and/or protein complexes that comprise di-tyrosine (DT) cross-links, and methods
of making such DT-cross-linked influenza HA polypeptides, proteins and/or protein
complexes.

[0207] Di-tyrosine cross-linking introduces one or more covalent carbon-carbon bonds
into proteins or protein complexes. This provides a method for stabilizing proteins, protein
complexes, and conformations, by introduction of intra- and/or inter-polypeptide di-tyrosine
bonds while maintaining their structural and functional integrity (See Marshall et al., US
Patent Numbers 7,037,894 & 7,445,912, the contents of which are hereby incorporated by
reference). The minimally altering, and zero-length DT cross-link is not hydrolyzed under
physiological conditions, and has been demonstrated to maintain proteins’ structural
integrity by liquid chromatography / mass spectrometry (LC/MS). Di-tyrosine cross-links
are known to be safe, as they form naturally in vivo, both in the context of proteins evolved
to utilize their specific characteristics (e.g. Elvin CM et al. 2005, Nature 437:999-1002;
Tenovuo J & Paunio K 1979, Arch Oral Biol.;24(8):591-4), and as a consequence of non-
specific protein oxidation (Giulivi et al. 2003, Amino Acids 25(3-4):227-32), and as they are
present in large quantities in some of our most common foods: DT bonds form the structure
of wheat gluten – the quarternary protein structure comprising the glutenin subunits – e.g. in
bread dough during mixing and baking (Tilley et al. 2001, Agric. Food Chem 49, 2627). Di-
tyrosine bonds do not form spontaneously in vitro. Rather, the enzymatic cross-link reaction
is carried out under optimized conditions to preserve protein structure and function.
Therefore, non-specific bonding/aggregation does not occur (unlike with disulfide bonding),
and therefore large-scale manufacturing of a DT stabilized immunogen may be economically
more feasible.

[0208] Tyrosyl side-chains are present in many redox enzymes, and catalysis of the
enzyme-specific reactions often involves tyrosyl radicals that are long-lived and have
comparatively low reactivity. Under optimized conditions radical formation is specific to
tyrosyl side-chains. In close proximity, tyrosyl side chains undergo radical coupling and
form a covalent, carbon-carbon bond. Tyrosyl radicals that do not react revert to non-
radicalized tyrosyl side-chains (Malencik & Anderson, 2003. Di-tyrosine as a product
of oxidative stress and fluorescent probe. Amino Acids 25: 233–247). Therefore, tyrosyl side-
chains must be situated in close proximity to form DT bonds, either within a single folded
polypeptide chain, or on closely interacting protein domains within a complex. Because a
Cα-Cα separation of approximately 5-8 Å is a prerequisite to bond formation (Brown et al., 1998). Determining protein-protein interactions by oxidative cross-linking of a glycine-glycine-histidine fusion protein. Biochemistry 37, 4397-4406; Marshall et al. 2006, US Patent No. 7,037,894), and because no atom is added in the formation of these bonds, the resulting “staple” is “zero length” and non-disruptive to the protein structure.

[0209] Tyrosine residues to be cross-linked may be naturally present in the primary structure of the protein to be cross-linked or may be added by controlled point mutation. To form DT bonds, proteins with tyrosyl side chains can be subjected to reaction conditions that lead to the formation of DT bonds. Such conditions are, or become, oxidative reaction conditions, as the DT bond formation reaction is an oxidative cross-linking reaction. In some embodiments the DT cross-linking reaction conditions yield proteins that are otherwise not, or not detectably, modified. Such conditions may be obtained by use of enzymes that catalyze the formation of H₂O₂, such as peroxidases. DT bond formation may be monitored by spectrophotometry with an excitation wavelength of around 320 nm, and fluorescence measured at a wavelength of around 400 nm (see, for example, Fig. 4A), while loss of tyrosyl fluorescence is monitored also monitored by standard procedures. When loss of tyrosyl fluorescence is no longer stoichiometric with DT bond formation, the reaction may be stopped by any methods known to one skilled in the art, such as, for example, by the addition of a reducing agent and subsequent cooling (on ice) or freezing of the sample. Further details of how to perform DT cross-linking are known in the art and are described in, for example, Marshall et al. 2006, US Patent No. 7,037,894, the contents of which are hereby incorporated by reference.

[0210] The major advantages of di-tyrosine cross-linking in protein engineering include (i) the ability to target specific residues for cross-linking (based on the primary, secondary, tertiary, and/or quaternary structures of proteins and complexes), (ii) minimal structural modification, (iii) specificity of the reaction (tyrosine is the only amino acid known to form cross-links under specific cross-linking conditions); (iv) stability of the linkage, (v) zero length of the cross-link (no atom is added), and (vi) the scalability of the cross-linking chemistry.

[0211] In some embodiments, targeted DT cross-links may be introduced at one or more of the specific locations in the influenza HA protein that are recited herein. In other
embodiments, additional positions within influenza HA polypeptides, proteins or protein complexes can be identified at which DT cross-links can be made. In some embodiments, di-tyrosine bonds or cross-links are targeted to specific residue pairs within the structure of an influenza HA polypeptide, protein and/or protein complex where DT bonds will, or are expected to, form due to, for example, their close proximity. In some embodiments tyrosyl side chains are already present at amino acid residues to be cross-linked. In some cases naturally occurring tyrosine residues may constitute either one or both of the paired tyrosine residues necessary for di-tyrosine bond formation. However, in other cases the influenza HA polypeptides, proteins and/or protein complexes of the invention are mutated or engineered to add one or more tyrosine residues, or to substitute one or more non-tyrosine residues for tyrosine residues. Such mutations are referred to herein as “to-tyrosine” mutations, and can be introduced at locations where it is desirable to form di-tyrosine cross-links / bonds. In some embodiments, the present invention provides mutant influenza HA polypeptides, proteins, and/or protein complexes in which tyrosyl side chains are introduced at desired cross-linking positions by introducing point mutations to tyrosine in a nucleic acid sequence encoding the influenza HA polypeptide, protein, or protein complex. Alternatively, in some embodiments influenza HA proteins, polypeptides or protein complexes, or portions thereof, may be synthesized to include tyrosine residues or amino acids having tyrosyl side chains at desired cross-linking positions. Conversely, in some embodiments the present invention provides mutant influenza HA polypeptides, proteins, and/or protein complexes in which tyrosyl side chains are removed at undesirable cross-linking positions by introducing point mutations from tyrosine in a nucleic acid sequence encoding the influenza HA polypeptide, protein, or protein complex, or influenza HA polypeptides, proteins, or protein complexes may be synthesized to exclude tyrosine residues or amino acids having tyrosyl side chains at positions where cross-linking is not desired. For example, at least one of the tyrosyl side chains can be replaced with another side chain, such as a phenylalanine side chain (see, for example, Marshall CP et al., US Patent No. 7,037,894, the contents of which are hereby incorporated by reference). Accordingly, the influenza HA polypeptides, proteins and protein complexes of the invention may comprise point mutations “to tyrosine” or “from tyrosine.” Such mutations can be made by altering the nucleic acid sequences that encode the influenza HA polypeptides, proteins and/or protein complexes of the invention using any suitable mutagenesis methods known in the art. Alternatively, mutant influenza HA polypeptides,
proteins and/or protein complexes may be synthesized, purified, and/or produced by any other suitable methods known in the art.

[0212] In some embodiments, the present invention contemplates the targeted introduction of one or more di-tyrosine cross-link at any suitable position(s) in an influenza HA polypeptide, protein or protein complex, for example, in the stalk domain where the cross-link will or may stabilize the stalk domain in a native trimeric conformation or other conformation capable of binding anti-stalk antibodies, such as neutralizing or broadly neutralizing anti-stalk antibodies. Such stabilization may be achieved, for example, by introducing cross-links that stabilize interactions or folds within a stalk monomer (intra-molecular cross-linking) and/or interactions between one or more stalk monomers that comprise that stalk trimer (inter-molecular cross-linking), or any combination of intra- and/or inter-molecular crosslinks.

Proteolytic Cleavage

[0213] In some embodiments of the invention the influenza HA polypeptides, proteins and/or protein complexes of the invention (and/or intermediates in the synthesis thereof), comprise one or more protease recognition motifs that can be used, for example, to facilitate proteolytic removal of the head domain. Any suitable protease recognition motifs known in the art can be used. Such engineered protease recognition sites can be located at any suitable location in the influenza HA polypeptide, protein, and/or protein complex in which they will be useful for the disruption and/or removal of the head domain but preferably will not disrupt the native trimeric conformation of (and/or conformation of neutralizing epitopes in) the stalk domain. Such locations can be determined using methods known in the art, including, but not limited to, testing the effect of introducing engineered protease recognition sites in functional assays, antibody binding assays, antigenic assays, structural assays, and the like. In some embodiments such engineered protease recognition motifs may be located within a variable loop region – as such regions are known to tolerate variations in amino acid sequence without significantly altering the structure and/or function of the influenza HA protein. The influenza HA proteins of the invention can be engineered to introduce one or more protease recognition sequences by, for example, inserting one or more amino acids that comprise, or comprise part of, a protease recognition site (see for example SEQ ID NO. 18, 19, 21, 23 and 25), or by substituting one or more amino acids from the
influenza HA protein with different amino acids that comprise, or comprise part or, a protease recognition site (see for example SEQ ID NO. 24), or by performing a combination of insertion and substitution of amino acids (see for example SEQ ID NO. 20 and 22) in order to generate a protease recognition sequence within the influenza HA protein sequence. The engineered protease recognition site will typically consist of up to about 20 amino acid residues. In some embodiments the influenza HA polypeptides, proteins and/or protein complexes described herein comprise an engineered protease recognition motif at one or more of the following primary head-removal sites: amino acid residues 53-67, amino acid residues 60-76, amino acid residues 269-277, and amino acid residues 277-290, and may optionally also comprise an engineered protease recognition motif at one or more of the following secondary head-removal sites: amino acid residues 142-146, and amino acid residues 155-164. In some embodiments the influenza HA polypeptides, proteins and/or protein complexes of the present invention comprise a protease recognition sequence that begins at an amino acid residue position within one of the following regions of the influenza HA protein: amino acid residues 40-68, amino acid residues 60-76, amino acid residues 77-114, amino acid residues 120-141, amino acid residues 142-146, amino acid residues 148-178, amino acid residues 182-188, amino acid residues 195-201, amino acid residues 209-242, amino acid residues 250-255, amino acid residues 260-285, amino acid residues 277-290, and amino acid residues 286-320. In some embodiments such protease recognition motifs may allow proteolytic cleavage at one or more of the Sa, Ca, Sb and Cb antigenic sites in the influenza HA head domain. In some embodiments the protease recognition motif is inserted into the influenza HA protein immediately following the amino acid at position 48, 63, 278, 282, 286, or 291. In some embodiments the protease recognition motif is inserted into the influenza HA protein within one or more of the following regions of the influenza HA protein: amino acid residues 38-58, amino acid residue 53-73, amino acid residues 268-288, amino acid residues 272-292, amino acid residues 276-296 and amino acid residues 281-301. In some embodiments the protease recognition motifs may comprise a PreScission Protease recognition sequence (for example, LEVLFQGP (SEQ ID NO. 69)) or TEV recognition sequence, (for example, ENLYFQG (SEQ ID NO. 70) or ENLYFQS (SEQ ID NO. 71)), or any combination thereof. Nucleotide sequences encoding such protease recognition sites can be engineered into the nucleic acids that encode the influenza HA polypeptides, proteins, and/or protein complexes of the invention using standard molecular biology techniques known in the art.
Making and Analyzing Influenza HA Polypeptides, Proteins, and Protein Complexes

[0214] In some embodiments the present invention provides methods for making the influenza HA polypeptides, proteins, and protein complexes of the invention. The influenza HA polypeptides, proteins, and protein complexes of the invention can be made by any suitable means known in the art. In some embodiments the influenza HA polypeptides, proteins, and/or protein complexes of the invention can be made by recombinant means. In some embodiments, the influenza HA polypeptides, proteins, and/or protein complexes of the invention, or any portion thereof, can be made by chemical synthesis means. For example, a peptide corresponding to a portion of a protein or protein complex as described herein can be synthesized by use of a peptide synthesizer.

Recombinant Production Methods

[0215] In embodiments where the influenza HA polypeptides, proteins and protein complexes of the invention are made by recombinant means, nucleic acids encoding the influenza HA polypeptides, proteins and protein complexes of the invention can be expressed in any suitable cell type, including, but not limited to mammalian cells, avian cells (such as EB66 duck cells) and insect cells (such as SF9 or Hi5 cells, using a baculovirus expression system). Methods for expressing polypeptides and proteins from nucleic acid molecules are routine and well known in the art, and any suitable methods, vectors, systems, and cell types known in the art can be used. For example, typically nucleic acid sequences encoding the influenza HA polypeptides, proteins and/or protein complexes of the invention will be placed into a suitable expression construct containing a suitable promoter, which will then be delivered to cells for expression.

Chimeric / Fusion Proteins & Oligomerization Domains

[0216] In some embodiments it may be desirable to add chimeric domains to the influenza HA polypeptides, proteins and/or protein complexes described herein, to produce chimeric proteins /fusion proteins, for example to facilitate the analysis and/or isolation and/or purification of the influenza HA polypeptides, proteins and/or protein complexes described herein. In some embodiments, the influenza HA polypeptides, proteins and protein complexes of the invention may comprise leader sequences, precursor polypeptide sequences, secretion signals, localization signals, epitope tags, protease cleavage sites, and
the like. Epitope tags that can be used include, but are not limited to, FLAG tags, glutathione S-transferase (GST) tags, green fluorescent protein (GFP) tags, hemagglutinin A (HA) tags, histidine (His) tags, luciferase tags, maltose-binding protein (MBP) tags, c-Myc tags, protein A tags, protein G tags, streptavidin (strep) tags, and the like.

[0217] In some embodiments it may be desirable to add oligomerization domains to facilitate the assembly of influenza HA polypeptides, proteins and/or protein complexes as described herein, and/or to facilitate stabilization of stalk domain in a native trimeric conformation, and/or to stabilize other structural features of the influenza HA polypeptides, proteins and/or protein complexes. In some embodiments the oligomerization domains are trimerization motifs, including, but not limited to, the T4 foldon motif. There are a wide variety of trimerization domains in natural proteins that can be used for these purposes including, but not limited to, those described in Habazettl et al., 2009 (Habazettl et al., 2009. NMR Structure of a Monomeric Intermediate on the Evolutionarily Optimized Assembly Pathway of a Small Trimerization Domain. J. Mol. Biol.), Kammerer et al., 2005 (Kammerer et al., 2005. A conserved trimerization motif controls the topology of short coiled coils. Proc Natl Acad Sci USA 102 (39): 13891-13896), Innamorati et al., 2006 (Innamorati et al., 2006. An intracellular role for the C1q-globular domain. Cell signal 18(6): 761-770), and Schelling et al., 2007 (Schelling et al., 2007. The reovirus σ-1 aspartic acid sandwich: A trimerization motif poised for conformational change. Biol Chem 282(15): 11582-11589). Stabilizing trimeric protein complexes can also be accomplished using the GCN4 and T4 fibrininit motifs (Pancera et al., 2005. Soluble Mimetics of Human Immunodeficiency Virus Type 1 Viral Spikes Produced byReplacement of the Native Trimerization Domain with a Heterologous Trimerization Motif: Characterization and Ligand Binding Analysis. J Virol 79(15): 9954-9969; Guthe et al., 2004. Very fast folding and association of a trimerization domain from bacteriophage T4 fibrin. J.Mol.Biol. v337 pp. 905-15; Papanikolopoulou et al., 2008. Creation of hybrid nanorods from sequences of natural trimeric fibrous proteins using the fibrininit trimerization motif. Methods Mol Biol 474:15-33). Heterologous oligomerization motifs may be introduced by any recombinant methods known to one of ordinary skill in the art in order to stabilize the protein-protein interactions of the proteins of the present invention.
In some embodiments it may be desirable to add more than one additional domain and/or tag to the influenza polypeptides, proteins and/or protein complexes described herein, and any combination of suitable chimeric and/or oligomerization domains may be added to makedesired influenza HA polypeptides, proteins and/or protein complexes. In some embodiments, the additional domains are engineered at or in the transmembrane region of an influenza HA protein, for example by insertion and/or substitution of one or more amino acids in the transmembrane region such that all or a portion of the transmembrane region is replaced by the additional domains. In some embodiments the additional domains comprise a thrombin cleavage site, a T4 foldon motif and a histidine tag (e.g. a 6xHis tag (SEQ ID NO: 118)). In some embodiments the additional domains are encoded by a nucleic acid sequence comprising
CGTCTCTGTGGTCCGCTGGTCTCCGGGTTCTGTTACATCCCGGAAGCTCGGC
GTACGCTCGCTACGTACGCCGTAAGACGGTGAAATGGGTCTCTGTCTACCTT
CCTGCACCACCACCACCACCACCACCTGA (SEQ ID NO: 72). In some embodiments the influenza HA polypeptides, proteins and/or protein complexes comprise a tag comprising, consisting of, or consisting essentially of the amino acid sequence
RSLVPRGSPGSGYIPEAPRDGQAYVRKDGEWVLLSTFLHHHHHH (SEQ ID NO: 116).

Chimeric influenza HA polypeptides, proteins and/or protein complexes can be made by any method known to one of ordinary skill in the art, and may comprise, for example, one or several influenza HA polypeptides, proteins and/or protein complexes of the invention, and/or any fragment, derivative, or analog thereof (for example, consisting of at least a domain of a polypeptide, protein, or protein complex of the invention, or at least 6, and preferably at least 10 amino acids of thereof) joined at its amino- or carboxy-terminus via a peptide bond to an amino acid sequence of another protein or other protein domain or motif. In some embodiments such chimeric proteins can be produced by any method known to one of ordinary skill in the art, including, but not limited to, recombinant expression of a nucleic acid encoding a chimeric protein (e.g. comprising a first coding sequence joined in-frame to a second coding sequence); ligating the appropriate nucleic acid sequences encoding the desired amino acid sequences to each other in the proper coding frame, and expressing the chimeric product.

Post-translational Modifications
In some embodiments, the influenza HA polypeptides, proteins and protein complexes described herein may be altered by adding or removing post-translational modifications, by adding or removing chemical modifications or appendices, and/or by introducing any other modifications known to those of ordinary skill in the art. Included within the scope of the invention are influenza HA polypeptides, proteins and protein complexes that are modified during or after translation or synthesis, for example, by glycosylation (or deglycosylation), acetylation (or deacetylation), phosphorylation (or dephosphorylation), amidation (or deamidization), pegylation, derivatization by known protecting/blocking groups, proteolytic cleavage, or by any other means known in the art. For example, in some embodiments the influenza HA polypeptides, proteins and/or protein complexes may be subjected to chemical cleavage by cyanogen bromide, trypsin, chymotrypsin, papain, V8 protease, NaBH4, acetylation, formylation, oxidation, reduction, metabolic synthesis in the presence of tunicamycin, etc. In some embodiments such post-translational modifications can be used to render the influenza HA polypeptides, proteins, and/or protein complexes of the present invention more immunogenic, more stable, and/or more capable of binding to, or eliciting the production of, neutralizing and broadly neutralizing antibodies.

Obtaining Influenza HA Protein in Desired Conformations

In some embodiments the influenza HA polypeptides and/or proteins of the invention are assembled into protein complexes having a desired conformational structure, such as the native trimeric structure of the stalk domain, and are cross-linked in order to stabilize that conformation. As described elsewhere in the present application, the influenza HA protein comprises a trimer formed from three monomers. In some embodiments, prior to and/or during the enzymatic cross-linking reaction, the influenza HA protein may be obtained in (and/or maintained in) the desired conformation, for example while cross-linking is performed. In some embodiments the influenza HA protein may be produced and/or isolated in such a way that most, or substantially all, of the influenza HA molecules have a stalk domain present in a native trimeric conformation. For example, when the HA protein is expressed or obtained in a form that still comprises the head domain, the stalk domain will typically assume its native trimeric stalk confirmation. In some embodiments influenza HA molecules in a desired conformation may be separated from a mixed population of influenza HA protein molecules comprising some that are in the desired conformation (e.g. native
trimeric conformation of the stalk domain) and some that are in other conformations (e.g. stalk domain in a monomeric and/or dimeric conformation). In some embodiments, the influenza HA protein is expressed in cells (for example as its membrane bound or soluble form) and spontaneously assembles into its normal conformation (e.g. having a stalk domain in its native trimeric conformation). In some embodiments no additional stabilization may be necessary to retain stalk domain the influenza HA protein in its native trimeric form. In some embodiments the expressed and assembled/folded influenza HA protein may be kept under particular conditions, or in particular compositions, that favor formation and/or maintenance of the native trimeric conformation of the stalk domain. The influenza HA protein may be obtained and/or isolated and/or maintained in the desired conformation using any suitable method known in the art, including, but not limited to, standard protein purification methods, such as ion exchange chromatography, size exclusion chromatography, and/or affinity chromatography methods. In some embodiments the influenza HA protein may be expressed in the presence of, co-expressed with, or contacted with, molecules that bind to the influenza HA protein and stabilize it in its desired conformation, including, but not limited to, antibodies, small molecules, peptides, and/or peptidomimetics. Non-limiting examples of antibodies that bind to the stalk domain in its native trimeric conformation include 6F12, C179, CR6261, F10, A66, and D8. Other antibodies that can be used to characterize or stabilize the HA polypeptides, proteins and protein complexes of the invention include, but are not limited to, 18A3, 18C11, 18E7, 18E12, 18H9, 16B5, 10A14, 5K24, FI6v3, 6K14, 6J24, 8D4, anti-influenza human antibodies of the V\textsubscript{H}1-69 heavy chain lineage, and anti-influenza human antibodies of the V\textsubscript{H}3-30 heavy chain lineage. In some embodiments, the influenza HA protein may be obtained, isolated, or maintained in its desired conformation by controlling the ionic strength of the media/buffer in which the protein is present (such as by using high or low ionic strength media). In some embodiments the influenza HA protein may be obtained, isolated, or maintained at one or more temperatures that favor preservation of the desired conformation. In some embodiments the influenza HA protein may be obtained, isolated, or maintained over a period of time that diminishes the degree to which the desired conformation lost.

[0222] In some embodiments analysis may be performed to confirm that the desired conformation, such as the native trimeric conformation of the stalk domain, has been formed
and/or maintained in the influenza HA protein. Such analysis may be performed prior to
cross-linking, during the cross-linking process, after the cross-linking process, or at any
combination of such stages. Such analysis may comprise any suitable methods known in the
art for assessing the 3-dimensional structure of a protein or protein complex, including
functional analysis, crystallographic analysis, and the like. In some embodiments such
analysis may include assessing binding of the influenza HA protein to certain antibodies,
such as those that are specific to the native trimeric conformation of the stalk domain and/or
those that are known to bind to antigenic sites in the stalk domain or elsewhere in the
influenza HA protein, as described elsewhere herein, including, but not limited to the 6F12,
C179, CR6261, F10, A66, and D8 antibodies.

Protein Purification

[0223] In some embodiments the methods for making influenza HA polypeptides,
proteins, and protein complexes of the invention may comprise purifying the influenza HA
polypeptides, proteins, or protein complexes before, during, or after, one or more steps in the
manufacturing process. For example, in some embodiments the influenza HA polypeptides,
proteins, and/or protein complexes of the invention may be purified after completion of all
of the manufacturing steps. In some embodiments the influenza HA polypeptides, proteins,
and/or protein complexes of the invention may be purified before commencing the cross-
linking process or after one or more of the intermediate method steps in the process, for
example, after expression of an influenza HA polypeptide or protein, after assembly of a
protein complex, after obtaining the influenza HA protein in a desired conformation, during
or after performing a cross-linking reaction, or after removal of the head domain. The
influenza HA polypeptides, proteins, and/or protein complexes of the invention may be
isolated or purified using any suitable method known in the art. Such methods include, but
are not limited to, chromatography (e.g. ion exchange, affinity, and/or sizing column
chromatography), ammonium sulfate precipitation, centrifugation, differential solubility, or
by any other technique for the purification of proteins known to one of ordinary skill in the
art. In specific embodiments it may be necessary to separate the desirable influenza HA
polypeptides, proteins, and/or protein complexes of the invention from those that were not
sufficiently cross-linked, or those in which the head domain was not sufficiently removed.
This can be done using any suitable system known in the art. For example, influenza HA
proteins having a stalk domain in the native trimeric conformation can be separated from
those that have a stalk domain that is not in the native trimeric conformation using antibody-based separation methods. The influenza HA polypeptides, proteins, and/or protein complexes of the invention may be purified from any source used to produce them. For example, the influenza HA polypeptides, proteins, and/or protein complexes of the invention may be purified from sources including insect, prokaryotic, eukaryotic, mono-cellular, multi-cellular, animal, plant, fungus, vertebrate, mammalian, human, porcine, bovine, feline, equine, canine, avian, or tissue culture cells, or any other source. The degree of purity may vary, but in various embodiments, the purified influenza HA polypeptides, proteins, and/or protein complexes of the invention are provided in a form in which they comprise more than about 10%, 20%, 50%, 70%, 75%, 80%, 85%, 90%, 95%, 98%, 99%, or 99.9% of the total protein in the final composition. In some embodiments the influenza HA polypeptides, proteins, and/or protein complexes of the invention may be isolated and purified from other proteins, or any other undesirable products (such as non-cross-linked products or products where removal of the head domain is insufficient or incomplete), by standard methods including, but not limited to, chromatography, glycerol gradients, affinity chromatography, centrifugation, ion exchange chromatography, size exclusion chromatography, and affinity chromatography, or by any other standard technique for the purification of proteins known in the art. The influenza HA polypeptides, proteins, and/or protein complexes to be isolated may be expressed in high or low ionic media, or isolated in high or low ionic buffers or solutions. The influenza HA polypeptides, proteins, and/or protein complexes of the invention may also be isolated at one or more temperatures that favor preservation of the desired conformation. They may also be isolated over a period of time that diminishes the degree to which a preparation would have lost the desired conformation. The degree to which a preparation of proteins retains one or more desired conformations (such as the native trimeric conformation of the stalk domain and/or conformations that favor binding to neutralizing antibodies, or other desired properties) may be assayed by any suitable method known in the art, including, for example, but not limited to, biochemical, biophysical, immunologic, and virologic analyses. Such assays include, for example, but are not limited to, immunoprecipitation, enzyme-linked immunosorbent assays (ELISAs), or enzyme-linked immunosorbent spot (ELISPOT) assays, crystallographic analysis (including co-crystallization with antibodies), sedimentation, analytical ultracentrifugation, dynamic light scattering (DLS), electron microscopy (EM), cryo-EM tomography, calorimetry, surface plasmon resonance (SPR), fluorescence resonance energy transfer (FRET), circular
dichroism analysis, and small angle x-ray scattering, neutralization assays, antibody-dependent cellular cytotoxicity assays, and/or virologic challenge studies in vivo.

[0224] The yield of the influenza HA polypeptides, proteins, and/or protein complexes of the invention can be determined by any means known in the art, for example, by comparing the amount of the final engineered proteins (such as cross-linked influenza HA proteins) as compared to the amount of the starting material, or as compared to the amount of the materials present in any preceding step of the production methods. Protein concentrations can determined by standard procedures, such as, for example, Bradford or Lowry protein assays. The Bradford assay is compatible with reducing agents and denaturing agents (Bradford, M, 1976. Anal. Biochem. 72: 248). The Lowry assay has better compatibility with detergents and the reaction is more linear with respect to protein concentrations and read-out (Lowry, O J, 1951. Biol. Chem. 193: 265).

Exemplary Production Methods

[0225] In some embodiments the present invention provides methods for making “headless” influenza HA polypeptides, proteins and/or protein complexes as described herein. In some embodiments methods for making “headless” influenza HA polypeptides, proteins and/or protein complexes comprise: (a) expressing an influenza HA protein having (i) both a stalk domain and a head domain, and (ii) one or more engineered protease recognition motifs in or near its head domain, (b) allowing the soluble influenza HA protein expressed in step (a) to fold into its native conformation having a trimeric stalk domain and a head domain, (c) introducing one more cross-links into the trimeric stalk domain, wherein the cross-links stabilize the stalk domain in its native trimeric conformation, and (d) subsequently proteolytically disrupting or removing the head domain, thereby producing a headless influenza HA protein. In some such embodiments the cross-links are targeted cross-links, such as di-tyrosine cross-links. In some embodiments the methods also involve first (at least prior to step (c)) identifying one or more regions in the HA protein in which the introduction of one or more cross-links in step (c) could stabilize the conformation of the stalk in its native trimeric conformation and/or stabilize the stalk in a conformation that allows binding of one or more broadly neutralizing anti-stalk antibodies. In some embodiments methods for making “headless” influenza HA polypeptides, proteins and/or protein complexes comprise: (a) expressing an influenza HA protein having: (i) both a stalk
domain and a head domain, (ii) one or more “to-tyrosine mutations within its stalk domain, and (iii) one or more engineered protease recognition motifs within or close to its head domain, (b) allowing the influenza HA protein to fold into its native conformation having a trimeric stalk domain and a head domain, (c) introducing one or more di-tyrosine cross-links into the trimeric stalk domain, wherein the di-tyrosine cross-links are stable under physiological conditions and stabilize the stalk domain in its native trimeric conformation, and (d) subsequently proteolytically removing the head domain, thereby producing a soluble headless influenza HA protein. In some embodiments the method also involves identifying first (at least prior to step (c)) one or more regions in the HA protein in which the introduction of one or more DT cross-links in step (c) could stabilize the conformation of the stalk in its native trimeric conformation and/or stabilize the stalk in a conformation that allows binding of one or more broadly neutralizing anti-stalk antibodies. In such methods the soluble influenza HA protein will typically comprises one or more protease recognition motifs that can be used to facilitate proteolytic removal of the head domain, as described above and in other sections of this application.

[0226] In some embodiments the methods for making “headless” influenza HA polypeptides, proteins and/or protein complexes described herein may further comprise performing an analysis after commencement or completion of the proteolytic cleavage step(s) to confirm that the head domain of the influenza HA protein has been sufficiently disrupted or removed. In some such embodiments this analysis may comprise, for example, performing an SDS PAGE gel mobility shift assay or using a head-specific antibody.

[0227] In some embodiments the present invention provides methods for making “head-on” influenza HA polypeptides, proteins and/or protein complexes as described herein. In some embodiments methods for making “head-on” influenza HA polypeptides, proteins and/or protein complexes comprise: (a) expressing an influenza HA protein comprising a stalk domain and a head domain, (b) allowing the expressed influenza HA protein to fold into its native conformation having a trimeric stalk domain, and (c) introducing one more physiologically stable cross-links into the HA protein in the trimeric stalk domain and optionally also in the head domain, thereby producing an engineered “head-on” influenza HA protein having a cross-linked stalk domain. In some such embodiments the cross-links are targeted cross-links, such as di-tyrosine cross-links. In some embodiments methods for
making “headless” influenza HA polypeptides, proteins and/or protein complexes comprise: (a) expressing an influenza HA protein having one or more “to-tyrosine” mutations at targeted positions within its stalk domain and optionally also in the head domain, (b) allowing the influenza HA protein to fold into its native conformation having a trimeric stalk domain and a head domain, and (c) performing a DT cross-linking reaction to cross-link tyrosine residues in the stalk domain and optionally also in the head domain, thereby producing an engineered “head-on” influenza HA protein having a DT-cross-linked stalk domain. In such methods the influenza HA protein may comprise one or more protease recognition motifs that could be used, if desired, to facilitate subsequent proteolytic removal of the head domain of the “head-on” protein to generate a “headless” influenza HA protein.

Properties of Influenza HA Polypeptides, Proteins and/or Protein Complexes

[0228] In some embodiments, the influenza HA polypeptides, proteins and/or protein complexes of the invention, including in particular those that are cross-linked as described herein, have certain structural, physical, functional, and/or biological properties. Such properties may include one or more of the following, or any combination of the following: presence or absence of a head domain, existence of the stalk domain in its native trimeric conformation; improved stability of the native trimeric conformation of the stalk domain (as compared to non-cross-linked influenza HA proteins); improved half-life of the influenza HA protein (as compared to non-cross-linked influenza HA proteins); improved thermostability (as compared to non-cross-linked influenza HA proteins); prolonged shelf-life (as compared to non-cross-linked influenza HA proteins); prolonged half-life inside the body of a subject (as compared to non-cross-linked influenza HA proteins); ability to be stored in solution without forming aggregates (including when present at a high concentration in solution); reduced aggregation in solution (as compared to non-cross-linked influenza HA proteins); binding to an antibody; binding to a neutralizing antibody; binding to a broadly neutralizing antibody; binding to a stalk-specific antibody; binding to a conformationally-specific antibody; binding to an antibody that recognizes a stalk domain epitope; binding to an antibody selected from the group consisting of 6F12, C179, CR6261, F10, A66, and D8; binding to a B cell receptor; activation of a B cell receptor; eliciting an antibody response in vivo; eliciting a protective antibody response in vivo; eliciting production of neutralizing antibodies in vivo; eliciting production of broadly neutralizing antibodies in vivo; eliciting production of antibodies that recognize quaternary neutralizing
epitopes (QNEs) in vivo; eliciting a protective immune response in vivo; and/or eliciting a humoral immune response in vivo. In the case of binding to antibody molecules, in some embodiments the influenza HA polypeptides, proteins, and/or protein complexes of the invention bind to the antibodies (such as stalk-specific antibodies, and/or 6F12, C179, CR6261, F10, A66, and D8) with high specificity and/or with high affinity.

Assays for Properties

[0229] In some embodiments the influenza HA polypeptides, proteins, and/or protein complexes of the invention, or any intermediates in their manufacture, may be analyzed to confirm that they have desired properties, such as the desired structural, physical, functional, and/or biological properties - such as those properties listed above or identified elsewhere in this patent specification. For example, in some embodiments in vitro or in vivo assays can be performed to assess the influenza HA protein’s conformational structure, stability (e.g. thermostability), half-life (e.g. inside the body of a subject), aggregation in solution, binding to an antibody (such as a neutralizing antibody, broadly neutralizing antibody; stalk-specific antibody; antibody that recognizes stalk domain epitopes, conformationally-specific antibody, 6F12, C179, CR6261, F10, A66, and Da), binding to a B cell receptor, activation of a B cell receptor, antigenicity, immunogenicity, ability to elicit an antibody response, ability to elicit a protective antibody/immune response, ability to elicit production of neutralizing antibodies, or ability to elicit production of broadly neutralizing antibodies. In embodiments where the influenza HA polypeptides, proteins, and/or protein complexes of the invention are tested in an animal in vivo, the animal may be any suitable animal species, including, but not limited to a mammal (such as a rodent species (e.g. a mouse or rat), a rabbit, a ferret, a porcine species, a bovine species, an equine species, an ovine species, or a primate species (e.g. a human or a non-human primate), or an avian species (such as a chicken)).

[0230] Assays for assessing a protein’s conformational structure are well known in the art and any suitable assay can be used, including, but not limited to, crystallographic analysis (e.g. X-ray crystallography or electron crystallography), sedimentation analysis, analytical ultracentrifugation, electron microscopy (EM), cryo-electron microscopy (cryo-EM), cryo-EM tomography, nuclear magnetic resonance (NMR), small angle x-ray scattering, fluorescence resonance energy transfer (FRET) assays, and the like.
[0231] Assays for assessing a protein’s stability are well known in the art and any suitable assay can be used, including, but not limited to, denaturing and non-denaturing electrophoresis, isothermal titration calorimetry, and time-course experiments in which proteins are incubated and analyzed over time at varying protein concentrations, temperatures, pHs or redox conditions. Proteins may also be analyzed for susceptibility to proteolytic degradation.

[0232] Assays for assessing binding of proteins to antibodies are well known in the art, and any suitable assay can be used, including, but not limited to, immunoprecipitation assays, enzyme-linked immunosorbent assays (ELISAs), enzyme-linked immunosorbent spot assays (ELISPOTs), crystallographic assays (including co-crystallization with antibodies), surface plasmon resonance (SPR) assays, fluorescence resonance energy transfer (FRET) assays, and the like.

[0233] Assays for assessing neutralization activity are well known in the art, and any suitable assay can be used. For example, assays can be performed to determine the neutralizing activity of antibodies or antisera generated by vaccination/immunization of animals with the influenza HA polypeptides, proteins, and/or protein complexes of the invention. Neutralization assays known in the art include, but are not limited to, those described by Dey et al. 2007 (Dey et al., 2007, Characterization of Human Immunodeficiency Virus Type 1 Monomeric and Trimeric gp120 Glycoproteins Stabilized in the CD4-Bound State: Antigenicity, Biophysics, and Immunogenicity. J Virol 81(11): 5579-5593) and Beddows et al., 2006 (Beddows et al., 2007, A comparative immunogenicity study in rabbits of disulfide-stabilized proteolytically cleaved, soluble trimeric human immunodeficiency virus type 1 gp140, trimeric cleavage-defective gp140 and monomeric gp120. Virol 360: 329-340).

[0234] Assays for assessing whether a vaccine immunogen is capable of eliciting an immune response and/or providing protective immunity are well known in the art, and any suitable assay can be used. For example, assays can be performed to determine whether vaccination/immunization of animals with the influenza HA polypeptides, proteins, and/or protein complexes of the invention provide an immune response and/or protective immunity against infection with influenza virus. In some embodiments comparisons may be made
between placebo and test vaccinated groups with regard to their rates of infection or seroconversion or viral loads.

[0235] Assays for assessing a protein’s pharmacokinetics and bio-distribution are also well known in the art, and any suitable assay can be used to assess these properties of the influenza HA polypeptides, proteins, and/or protein complexes of the invention.

Compositions

[0236] In some embodiments the present invention provides compositions comprising any of the influenza HA polypeptides, proteins, and/or protein complexes described herein. In some embodiments such compositions may be immunogenic compositions, vaccine compositions and/or therapeutic compositions. In some embodiments, such compositions may be administered to subjects. In some embodiments the influenza HA polypeptides, proteins, and/or protein complexes described herein may be present in virus-like particles or “VLPs.”

[0237] In some embodiments the influenza HA polypeptides, proteins, and/or protein complexes of the invention may be provided in a composition that comprises one or more additional active components, such as one or more additional vaccine immunogens or therapeutic agents. In some embodiments the influenza HA polypeptides, proteins, and/or protein complexes of the invention may be provided in a composition that comprises one or more other components, including, but not limited to, pharmaceutically acceptable carriers, adjuvants, wetting or emulsifying agents, pH buffering agents, preservatives, and/or any other components suitable for the intended use of the compositions. Such compositions can take the form of solutions, suspensions, emulsions and the like. The term "pharmaceutically acceptable carrier" includes various diluents, excipients and/or vehicles in which, or with which, the influenza HA polypeptides, proteins, and/or protein complexes of the invention can be provided. The term "pharmaceutically acceptable carrier" includes, but is not limited to, carriers known to be safe for delivery to human and/or other animal subjects, and/or approved by a regulatory agency of the Federal or a state government, and/or listed in the U.S. Pharmacopeia, and/or other generally recognized pharmacopeia, and/or receiving specific or individual approval from one or more generally recognized regulatory agencies for use in humans and/or other animals. Such pharmaceutically acceptable carriers, include, but are not limited to, water, aqueous solutions (such as saline solutions, buffers, and the
like), organic solvents (such as certain alcohols and oils, including those of petroleum, animal, vegetable or synthetic origin, such as peanut oil, soybean oil, mineral oil, sesame oil), and the like. In some embodiments the compositions of the invention also comprise one or more adjuvants. Exemplary adjuvants include, but are not limited to, inorganic or organic adjuvants, oil-based adjuvants, virosomes, liposomes, lipopolysaccharide (LPS), molecular cages for antigens (such as immune-stimulating complexes (“ISCOMS”)), Ag-modified saponin/cholesterol micelles that form stable cage-like structures that are transported to the draining lymph nodes), components of bacterial cell walls, endocytosed nucleic acids (such as double-stranded RNA (dsRNA), single-stranded DNA (ssDNA), and unmethylated CpG dinucleotide-containing DNA), AUM, aluminum phosphate, aluminum hydroxide, and Squalene. In some embodiments virosomes are used as the adjuvant.

Additional commercially available adjuvants that can be used in accordance with the present invention include, but are not limited to, the Ribi Adjuvant System (RAS, an oil-in-water emulsion containing detoxified endotoxin (MPL) and mycobacterial cell wall components in 2% squalene (Sigma M6536)), TiterMax (a stable, metabolizable water-in-oil adjuvant (CytRx Corporation 150 Technology Parkway Technology Park/Atlanta Norcross, Georgia 30092)), Syntex Adjuvant Formulation (SAF, an oil-in-water emulsion stabilized by Tween 80 and pluronic polyoxyethlene/polyoxypropylene block copolymer L121 (Chiron Corporation, Emeryville, CA)), Freund’s Complete Adjuvant, Freund’s Incomplete Adjuvant, ALUM - aluminum hydroxide, Al(OH)3 (available as Alhydrogel, Accurate Chemical & Scientific Co, Westbury, NY), SuperCarrier (Syntex Research 3401 Hillview Ave. P.O. Box 10850 Palo Alto, CA 94303), Elvax 40W1,2(an ethylene-vinyl acetate copolymer (DuPont Chemical Co. Wilmington, DE)), L-tyrosine co-precipitated with the antigen (available from numerous chemical companies); Montanide (a manide-oleate, ISA Seppic Fairfield, NJ)), AdjuPrime (a carbohydrate polymer), Nitrocellulose-absorbed protein, Gerbu adjuvant (C-C Biotech, Poway, CA), and the like.

[0238] In some embodiments the compositions of the invention comprise an “effective amount” of an influenza HA polypeptide, protein, and/or protein complex of the invention. An “effective amount” is an amount required to achieve a desired end result. Examples of desired end results include, but are not limited to, the generation of a humoral immune response, the generation of a neutralizing antibody response, the generation of a broadly neutralizing antibody response, and the generation of protective immunity. The amount of
an influenza HA polypeptide, protein, and/or protein complex of the invention that is
effective to achieve the desired end result will depend on variety of factors including, but not
limited to, the type, subtype, and strain of the influenza virus against which protection or
some other therapeutic effect is sought, the species of the intended subject (e.g. whether a
human or some other animal species), the age and/or sex of the intended subject, the planned
route of administration, the planned dosing regimen, the seriousness of any ongoing
influenza infection (e.g. in the case of therapeutic uses), and the like. The effective amount
– which may be a range of effective amounts – can be determined by standard techniques
without any undue experimentation, for example using in vitro assays and/or in vivo assays
in the intended subject species or any suitable animal model species. Suitable assays
include, but are not limited to, those that involve extrapolation from dose-response curves
and/or other data derived from in vitro and/or in vivo model systems. In some embodiments
the effective amount may be determined according to the judgment of a medical or
veterinary practitioner based on the specific circumstances.

Uses of the Influenza HA Polypeptides, Proteins & Protein Complexes of the Invention

[0239] In some embodiments, the influenza HA polypeptides, proteins, and protein
complexes of the invention may be useful as research tools, as diagnostic tools, as
therapeutic agents, as targets for the production of antibody reagents or therapeutic
antibodies, and/or as vaccines or components of vaccine compositions. For example, in
some embodiments the influenza HA polypeptides, proteins, and protein complexes of the
invention are useful as vaccine immunogens in animal subjects, such as mammalian subject,
including humans. These and other uses of the influenza HA polypeptides, proteins, and
protein complexes of the invention are described more fully below. Those of skill in the art
will appreciate that the influenza HA polypeptides, proteins, and protein complexes of the
invention may be useful for a variety of other applications also, and all such applications and
uses are intended to fall within the scope of this invention.

Tools for Studying Influenza HA Antibodies

[0240] In one embodiment, the influenza polypeptides, proteins, and protein complexes
of the invention may be useful as analytes for assaying and/or measuring binding of, and/or
titers of, anti-HA antibodies, for example in ELISA assays, Biacore/SPR binding assays,
and/or any other assays for antibody binding known in the art. For example, the influenza
HA polypeptides, proteins, and protein complexes of the invention could be used to analyze, and/or compare the efficacy of anti-HA antibodies.

Tools for Generation of Antibodies

[0241] The influenza HA polypeptides, proteins, and protein complexes of the invention (including any intermediates and/or variants produced during manufacture of the influenza HA polypeptides, proteins, and protein complexes) may also be useful for the generation of therapeutic antibodies and/or antibodies that can be used as research tools or for any other desired use. For example, the influenza HA polypeptides, proteins, and protein complexes of the invention can be used for immunizations to obtain antibodies to the influenza HA protein for use as research tools and/or as therapeutics. In some embodiments the influenza HA polypeptides, proteins, and protein complexes of the invention can be used to immunize a non-human animal, such as a vertebrate, including, but not limited to, a mouse, rat, guinea pig, rabbit, goat, non-human primate, etc. in order to generate antibodies. Such antibodies, which may be monoclonal or polyclonal, and/or cells that produce such antibodies, can then be obtained from the animal. For example, in some embodiments influenza HA polypeptides, proteins, and protein complexes of the invention may be used to immunize a mouse and to produce and obtain monoclonal antibodies, and/or hybridomas that produce such monoclonal antibodies. Such methods can be carried out using standard methods known in the art for the production of mouse monoclonal antibodies, including standard methods for hybridoma production. In some embodiments influenza HA polypeptides, proteins, and protein complexes of the invention may be used for the production of a chimeric (e.g. part-human), humanized, or fully-human antibody, for example using any of the methods currently known in the art for production of chimeric, humanized and fully human antibodies, including, but not limited to, CDR grafting methods, phage-display methods, transgenic mouse methods (e.g. using a mouse that has been genetically altered to allow for the production of fully human antibodies, such as the Xenomouse) and/or any other suitable method known in the art. Antibodies to the influenza HA polypeptides, proteins, and protein complexes of the invention made using such systems can be characterized antigenically using one or a set of several antigens, preferably including the influenza HA polypeptides, proteins, and protein complexes of the invention themselves. Additional characterization of such antibodies may be carried out by any standard methods known to one of ordinary skill in the art, including, but not limited to,
ELISA-based methods, SPR-based methods, biochemical methods (such as, but not limited to, iso-electric point determination), and methods known in the art for studying biodistribution, safety, and efficacy of antibodies – for example in preclinical and clinical studies.

Administration to Subjects

[0242] In some embodiments, the present invention provides methods that comprise administering the influenza HA polypeptides, proteins and/or protein complexes of the invention (or compositions comprising such influenza HA polypeptides, proteins and/or protein complexes) to subjects. Such methods may comprise methods for treating individuals having influenza virus (i.e. therapeutic methods) and/or methods for protecting individuals against future influenza virus infection (i.e. prophylactic methods).

[0243] Subjects to which the influenza HA polypeptides, proteins and/or protein complexes of the invention, or compositions comprising such influenza HA polypeptides, proteins and/or protein complexes, can be administered (for example in the course of a method of treatment or a method of vaccination) include any and all animal species, including, in particular, those that are susceptible to influenza virus infection or that can provide model animal systems for the study of influenza virus infection. In some embodiments, the subjects are mammalian species. In some embodiments, the subjects are avian species. Mammalian subjects include, but are not limited to, humans, non-human primates, rodents, rabbits, and ferrets. Avian subjects include, but are not limited to chickens, such as those on poultry farms. In some embodiments the subjects to which the influenza HA polypeptides, proteins and/or protein complexes of the invention, or compositions comprising such influenza HA polypeptides, proteins and/or protein complexes are administered, either have influenza, or are at risk of influenza infection, for example due to the subject’s age and/or underlying medical conditions. In some embodiments, the subject is immuno-compromised. In some embodiments, the subject has heart disease, lung disease, diabetes, renal disease, dementia, stroke and/or rheumatologic disease. In some embodiments, the subject is a human of greater than about 50 years in age, greater than about 55 years in age, greater than about 60 years in age, greater than about 65 years in age, greater than about 70 years in age, greater than about 75 years in age, greater than about 80 years in age, greater than about 85 years in age, or greater than about 90 years
in age. In some embodiments, the subject is a human of less than about 1 month in age, less than about 2 months in age, less than about 3 months in age, less than about 4 months in age, less than about 5 months in age, less than about 6 months in age, less than about 7 months in age, less than about 8 months in age, less than about 9 months in age, less than about 10 months in age, less than about 11 months in age, less than about 12 months in age, less than about 13 months in age, less than about 14 months in age, less than about 15 months in age, less than about 16 months in age, less than about 17 months in age, less than about 18 months in age, less than about 19 months in age, less than about 20 months in age, less than about 21 months in age, less than about 22 months in age, less than about 23 months in age, or less than about 24 months in age.

[0244] Various delivery systems are known in the art and any suitable delivery systems can be used to administer the compositions of the present invention to subjects. Such delivery systems include, but are not limited to, intradermal, intramuscular, intraperitoneal, intravenous, subcutaneous, intranasal, epidural, and oral delivery systems. The compositions of the present invention may be administered by any convenient route, for example by infusion or bolus injection, by absorption through epithelial or mucocutaneous linings (e.g., oral mucosa, rectal and intestinal mucosa, etc.) and may be administered together with other biologically active agents. Administration can be systemic or local. Pulmonary administration can also be employed, e.g., by use of an inhaler or nebulizer, and formulation with an aerosolizing agent.

[0245] In some embodiments it may be desirable to administer the pharmaceutical compositions of the invention locally to a tissue in which the influenza HA polypeptide, protein or protein complex may be most effective in generating a desirable outcome. This may be achieved by, for example, local infusion, injection, delivery using a catheter, or by means of an implant, such as a porous, non-porous, or gelatinous implant or an implant comprising one or more membranes (such as sialastic membranes) or fibers from or through which the protein or protein complexes may be released locally. In some embodiments a controlled release system may be used. In some embodiments a pump may be used (see Langer, supra; Sefton, 1987. CRC Crit. Ref. Biomed. Eng. 14: 201; Buchwald et al., 1980. Surgery 88: 507; Saudek et al., 1989. N. Engl. J. Med. 321: 574). In some embodiments polymeric materials may be used to facilitate and/or control release of the influenza HA

[0246] In some embodiments, administration of the influenza HA polypeptide, protein and/or protein complex of the invention can be performed in conjunction with administration of one or more immunostimulatory agents. Non-limiting examples of such immunostimulatory agents include various cytokines, lymphokines and chemokines with immunostimulatory, immunopotentiating, and pro-inflammatory activities, such as interleukins (e.g., IL-1, IL-2, IL-3, IL-4, IL-12, IL-13); growth factors (e.g., granulocyte-macrophage (GM)-colony stimulating factor (CSF)); and other immunostimulatory agents, such as macrophage inflammatory factor, Flt3 ligand, B7.1; B7.2. The immunostimulatory agents can be administered in the same formulation as the influenza HA protein or polypeptide, or can be administered separately.

[0247] In some embodiments, the influenza HA polypeptides, proteins, and/or protein complexes of the invention, or compositions comprising them, can be administered to subjects in a variety of different influenza virus vaccination methods or regimens. In some such embodiments, administration of a single dose is preferred. However, in other embodiments, additional dosages can be administered, by the same or different route, to achieve the desired prophylactic effect. In neonates and infants, for example, multiple administrations may be required to elicit sufficient levels of immunity. Administration can continue at intervals throughout childhood, as necessary to maintain sufficient levels of protection against influenza virus infection. Similarly, adults who are particularly susceptible to influenza virus infection, such as, for example, the elderly and
immunocompromised individuals, may require multiple immunizations to establish and/or maintain protective immune responses. Levels of induced immunity can be monitored, for example, by measuring amounts of neutralizing secretory and serum antibodies, and dosages adjusted or vaccinations repeated as necessary to elicit and maintain desired levels of protection.

[0248] In some embodiments, dosing regimens may comprise a single administration/immunization. In other embodiments, dosing regimens may comprise multiple administrations/immunizations. For example, vaccines may be given as a primary immunization followed by one or more boosters. In some embodiments of the present invention such a “prime-boost” vaccination regimen may be used. For example, in some such prime-boost regimens a composition comprising an influenza HA polypeptide, protein or protein complex as described herein may be administered to an individual on multiple occasions (such as two, three, or even more occasions) separated in time, with the first administration being the “priming” administration and subsequent administrations being “booster” administrations. In other such prime-boost regimens a composition comprising an influenza HA polypeptide, protein or protein complex as described herein may be administered to an individual after first administering to the individual a composition comprising a viral or DNA vector encoding an influenza HA polypeptide, protein or protein complex as a “priming” administration, with one or more subsequent “booster” administrations of a composition comprising an influenza HA polypeptide, protein or protein complex as described herein. Boosters may be delivered via the same and/or different route as the primary immunization. Boosters are generally administered after a time period after the primary immunization or the previously administered booster. For example, a booster can be given about two weeks or more after a primary immunization, and/or a second booster can be given about two weeks or more after the first boosters. Boosters may be given repeatedly at time periods, for example, about two weeks or greater throughout up through the entirety of a subject's life. Boosters may be spaced, for example, about two weeks, about three weeks, about four weeks, about one month, about two months, about three months, about four months, about five months, about six months, about seven months, about eight months, about nine months, about ten months, about eleven months, about one year, about one and a half years, about two years, about two and a half years, about three
years, about three and a half years, about four years, about four and a half years, about five years, or more after a primary immunization or after a previous booster.

Preferred unit dosage formulations are those containing a dose or unit (e.g. an effective amount), or an appropriate fraction thereof, of the influenza HA polypeptides, proteins, and/or protein complexes of the invention. In addition to such ingredients, formulations of the present invention may include other agents commonly used by one of ordinary skill in the art. Pharmaceutical compositions provided by the invention may be conveniently presented in preferred unit dosage formulations prepared using conventional pharmaceutical techniques. Such techniques include the step of bringing into association the active ingredient and the pharmaceutical carrier(s) or excipient(s) or other ingredients. In general, the formulations are prepared by uniformly and intimately bringing into association the active ingredient with liquid carriers. Formulations suitable for parenteral administration include aqueous and non-aqueous sterile injection solutions which may contain antioxidants, buffers, bacteriostats and solutes which render the formulation isotonic with the blood of the intended recipient, and aqueous and non-aqueous sterile suspensions which may include suspending agents and thickening agents. The formulations may be presented in unit-dose or multi-dose containers, for example, sealed ampoules and vials, and may be stored in a freeze-dried (lyophilized) condition requiring only the addition of the sterile liquid carrier, for example, water for injections, immediately prior to use. Extemporaneous injection solutions and suspensions may be prepared from sterile powders, granules and tablets commonly used by one of ordinary skill in the art.

Kits

The present invention further provides kits comprising influenza HA polypeptides, proteins or protein complexes of the invention, or compositions containing such polypeptides, proteins or protein complexes. To facilitate use of the methods and compositions of the invention, any of the components and/or compositions described herein, and additional components useful for experimental or therapeutic or vaccine purposes, can be packaged in the form of a kit. Typically, the kit contains, in addition to the above components, additional materials which can include, e.g., instructions for using the components, packaging material, a container, and/or a delivery device.
[0251] Various embodiments of the present invention may also be further described by the following non-limiting examples:

EXAMPLES

[0252] The numbers in square brackets/parentheses in the Examples section of the present application are citations to the numbered references provided as a reference list herein.

EXAMPLE 1

[0253] The US and world populations continue to be at risk of a pandemic Influenza outbreak, and weaponized influenza virus remains a major bio-warfare/terrorism threat [23,24]. An influenza virus HA-based vaccine immunogen capable of eliciting Ab responses to the conserved stalk QNEs, instead of the immunodominant head of HA, is expected to give rise to broadly neutralizing antibodies that could protect from homologous (H1N1), as well as homologous drift variant, group 1 heterologous (H5N1), and group 2 heterologous challenge (H3N2). Thus, a single, universal immunogen could elicit protective immune responses against seasonal, pandemic, and weaponized influenza virus. Underscoring the commercial and public health impact that influenza virus has on the population is the fact that life insurance companies in the United States today are required to hold capital against a potential reoccurrence of the 1918 Spanish Flu pandemic (Oliver, Wyman, & Co, 2012 & [25]). The approach described herein has the potential to provide a broadly protective influenza vaccine that could enable stockpiling large amounts of vaccine product and eliminate real threats that derive from the ramp-up times in manufacturing required to address each new threat.

[0254] Current influenza virus vaccines protect mostly against homologous virus strains, requiring new trivalent vaccine cocktails to be matched seasonally to circulating strains. Protection is primarily due to high affinity antibodies to hemagglutinin (HA), and is often strain-specific due to a focusing of the immune response predominantly against the highly variable, immunodominant head domain of the HA protein. The HA stalk, however, is highly conserved across influenza strains, and considerable evidence now suggests that better responses to conserved regions of the stalk would provide broader protection [1 – 4]. Immunization with DNA coding for HA elicits predominantly stalk-specific Ab responses,
and data showing limited heterosubtypic protection by vaccination with HA DNA by electroporation was recently described [5]. Furthermore, vaccination with a ‘headless’ HA protein based immunogen (“Headless HA”, an HA construct from which the variable head domain is removed) results in the induction of Ab responses with significantly enhanced heterosubtypic binding activity [1,6]. A prime-boost combination of a viral or DNA vector encoding HA, followed by a Headless protein boost holds the promise of generating broadly heterospecific responses that yield long-lasting protection. However, good protection against heterologous challenges currently still remains elusive [7 – 10].

[0255] Significant attention has been focused on the identification and characterization of broadly neutralizing antibodies (“bnAbs”) in order to reverse engineer an immunogen capable of eliciting similar antibody responses [9,11]. A number of these bnAbs have been described, and the most potent bind conserved, complex/conformation-specific epitopes that are presented on the conserved stalk of influenza HA trimers, but not on protomers of the same complex [7,12,13]. Isolation of these human Abs proves that a broadly protective vaccine is, in fact, an achievable goal (a “protomer” is a subunit of the trimer, that itself is a HA1/H2A heterodimer). These trimer/complex-specific epitopes are therefore called quaternary neutralizing epitopes (QNEs), and they are believed to represent key sites of vulnerability of influenza viruses since they have the potential to elicit potent quaternary bnAbs. [14,15]. Only the intact trimeric stalk exhibits the broadly protective QNE (see Figure 2). A Headless construct that is locked in its trimeric, native conformation, and that binds the potent and broadly protective quaternary bnAbs could provide a universal influenza immunogen and could elicit potent bnAbs in vaccinated subjects.

[0256] Recently, a headless influenza hemagglutinin (“Headless HA”) immunogen has been shown to elicit antibody (“Ab”) responses focused on the highly conserved stalk region of influenza hemagglutinin (HA) that are broadly cross-reactive. It has also become clear that the most potent and broadly neutralizing/protective Abs (bnAbs) against the stalk region are trimer-specific (i.e. recognize the quaternary structure of the stalk), and that their corresponding quaternary epitopes are not displayed when the head of influenza HA is removed. In the absence of the head domain, the stalk trimer apparently falls apart. The present invention provides a Headless HA immunogen in which the trimeric conformation of the stalk region is stabilized or “conformationally locked” – for example by introduction of
targeted cross-links - before the head is proteolytically removed. This Headless HA immunogen should retain binding to quaternary bnAbs and present quaternary neutralizing epitopes (“QNEs”) as an influenza immunogen. Such a conformationally-locked Headless HA trimer may enable the long-sought goal of broad protection against influenza viruses from a single vaccination regimen.

[0257] Minimally modifying dityrosine (“DT”) stabilization technology enzymatically introduces safe, targeted, zero-length, and irreversible DT bonds to lock proteins and complexes in native conformations. Application of this technology fully preserves protein structure and avoids aggregation because DT bonds do not form spontaneously. Bonds only form between Tyr side-chains in very close structural proximity, and are introduced after the protein has fully folded and is in its native state. Targeted DT crosslinking enables the design of an improved influenza vaccine immunogen by conformationally locking QNEs to maximize broad protection.

[0258] The methods described in the present example involve 3 steps. The first step involves expressing soluble, full-length influenza HA with “to-Tyr” substitutions at targeted positions within the stalk region. The second step involves introduction of stabilizing DT crosslinks. And the third step involves proteolytically removing the head domain of the influenza HA in order to focus the immune responses on the DT-Headless HA QNEs.

[0259] Preliminary studies using a recombinant, soluble HIV Env trimer have demonstrated that DT crosslinking can be used to conformationally-lock the Env immunogen in its native, trimeric conformation, so that it improves binding to the most potent HIV quaternary bnAbs, analogous to the flu quaternary anti-stalk bnAbs, demonstrating the feasibility of this approach. HIV Env and influenza virus Headless are highly analogous in that both are unstable trimers when expressed recombinantly; and in both, key QNEs are only presented in the native trimeric complex. In other preliminary studies targeted DT bonds have been successfully introduced into the influenza HA stalk.

[0260] DT crosslinking of a recombinant PR8 HA construct in its native, trimeric conformation can be performed to confirm binding to key bnAbs, and subsequently the “head” domain can be removed by engineering proteolytic cleavage sites, while maintaining the DT-locked, native antigenic conformation of the stalk trimer. The resulting Headless
HA immunogen can be tested to confirm that it elicits broad protection in a C57BL/6 mouse model. Pre-clinical testing for efficacy can be performed in a highly predictive ferret lethal challenge model. Pre-clinical testing for safety can be performed in rabbits.

Targeted DT Cross-Linking

[0261] By generating native, soluble, and recombinant HA trimers and applying targeted dityrosine (DT) “staples” to covalently cross-link trimerizing interactions in the stalk of the trimer, DT-stabilized HA trimmers will be engineered with fully preserved antigenic profiles. Covalent stabilization of the trimer in the HA stalk region will be engineered to render stable the quaternary structure of the stalk, and this will allow subsequent proteolytic removal of the head while preserving the QNEs of the stalk. DT bonds are introduced to stabilize the complex after the protein/complex is fully folded, and therefore locks the native conformation, while maintaining structural functional integrity of the protein [16 – 18].

These safe, irreversible, and zero-length cross-links form only between Tyr residues in very close structural proximity, and do not distort the structure of the protein. Nor do they cause non-specific aggregate formation, as observed with disulfide bonds [17,19 – 22]. Targeted DT cross-linking technology can be applied to covalently stabilize a soluble HA trimer in its correctly folded conformation, and then one can determine whether it does, in fact, present key QNEs. Subsequently the immunodominant head can be removed by introducing sequence-specific protease cleavage sites - making use of variable loop tolerance for amino acid variation and information gathered from transposon-based mutagenesis analysis of HA. Presentation of QNEs on Headless HA is expected to improve upon the breadth of protection in lethal challenge studies with drift variant and heterologous viruses. The inventors’ prior work in HIV shows that highly glycosylated multimers (e.g. HIV Env) can efficiently be locked together by DT cross-linking at various locations within the cleaved Env trimer – while maintaining the relevant quaternary structure and antigenicity.

Conformationally locking the influenza virus HA trimeric complex

[0262] The HIV envelope spike is trimerized through well characterized interactions at its base as well as interactions at the spike’s apex [33, 34]. In order to stabilize the trimerizing interactions at the apex of the spike, tyrosine substitutions were introduced, and the protein was expressed, purified, and DT cross-linked. By fluorescence, 7 variants were identified that form intermolecular, trimerizing cross-links with an average of 80%+ efficiency prior to
any optimization, as quantified using DT-specific excitation (320nm) and emission (405nm) wavelengths. The ability of these constructs to bind conformational and trimer-specific bnAbs was assayed. DT crosslinking fully preserves binding of the anti-CD4 binding site bnAb b12, which binds both protomers and trimers, and the anti-V2 bnAb PG9, which preferentially binds trimers, but also binds monomers. In addition, conformational locking also significantly reduces binding to non-neutralizing mAbs, such as b6 & b13, in ELISA assays. The position of the DT bonds was confirmed by MS/MS of tryptic fragments of the DT-Env trimer. More importantly, a conformationally locked HIV Env trimer was found to bind significantly better to one of the most extremely broadly neutralizing and potent anti-HIV Env bnAbs, PG16, by comparison to the WT protomer; the PG16 epitope is only presented on the nativefunctional HIV envelope trimer [28]. Improved PG16 binding correlates with a significant reduction in binding to a poorly neutralizing anti-V2 mAb, CH58, that binds an α-helical conformer of an overlapping epitope that PG16 binds as a β-sheet. The next step with this DT-locked, soluble HIV Env trimer will be to test it in animal immunogenicity experiments.

[0263] In influenza HA, the trimeric structure of the HA protein in complex with the CR6261 bnAb was analyzed. Five examples of possible HA variants (N403Y_D429Y; N403Y_L432Y; N403Y_D433Y; N406Y_D429Y; and N406Y_D433Y) were initially identified, each with two point mutations that were predicted to form intermolecular bonds and stabilize the stalk trimer at the membrane distal/head proximal end (see schematic for design in Figure 3) without altering the CR6261 quaternary epitope. Expression vectors encoding four of these variants were generated, and the variants were expressed and subjected to cross-linking conditions. Spectrofluorometry was used to determine whether these variants were forming DT bonds using the DT-specific excitation and emission wavelengths at which DT bonds fluoresce powerfully in direct proportion to their molar concentration. All four variants, but not wild-type HA, formed DT bonds efficiently (Figure 4). Based on comparison to the positive control (insulin) and a DT standard, cross-linking efficiency of >70% is estimated for all four of these constructs prior to any optimization [35].

Removing the HA head from the conformationally locked HA
Proteolytic removal of the HA head domain from the DT-locked HA trimer requires engineering recognition motifs into the HA1 head domain for a substrate-specific protease (e.g. TEV). Using a transposon-based mutagenesis screen, four regions within the PR8 HA1 globular head have been identified that tolerate the insertion of foreign sequences approximately the same size as an engineered TEV protease site. Without further optimization, two of these regions (located at amino acid residues 128 and 223) would allow proteolytic cleavage of 3 of the 4 major antigenic sites in the PR8 globular head – the Sa, Ca, and Sb sites [36]. The remaining Cb site will also be removed. Viruses with insertions at these sites in HA1 remain capable of fusion, and the HA complex thus remains functionally intact. The proteolytic reaction will then be performed.

These data demonstrate that the approach of locking together the HA trimer in the stalk, and subsequently removing the immunodominant head domain, will preserve vaccine-relevant QNEs of headless HA, and will lock the immunogen in an antigenically favorable conformation. This, in turn, suggests that the DT-locked headless trimer(s) described herein are expected to induce broadly protective antibody responses in vivo.

Conformationally locking the influenza virus HA trimeric complex

Experimental Design. Soluble forms (e.g. lacking the transmembrane domain and possessing the T4 foldon trimerization motif) of the WT HA and variants described above will be expressed in SF9 or Hi5 cells as secreted proteins and purified by well-established methods [37 – 38]. The antigenic effect of the to-Tyr substitutions and the DT cross-linking will be determined in ELISAs using a panel of anti-HA stalk broadly neutralizing mAbs (e.g. 6F12, C179, CR6261, F10, A66 and D8), as structural changes caused by to-Tyr substitutions may reduce or enhance binding to some of these antibodies. Methods: Full-curve binding assays will compare WT HA to the uncross-linked and cross-linked HA variants. Changes in binding will be determined using non-linear regression analysis (Graphpad software) of binding curves to calculate and compare EC50 values for each construct with each mAb. Intermolecular bond formation will be confirmed by gel-shift in reducing SDS-PAGE (Western blot/Coomassie; DT bonds are not reduced); DT cross-linking will be quantified by spectrofluorometry, as described above. Such methods can be used to produce HA variants that form intermolecular DT bonds, and that retain binding to
key anti-stalk quaternary bnAbs equal to wild-type PR8 HA after cross-linking the engineered influenza immunogen.

Proteolytically removing the HA head from the conformationally locked HA

[0267] PreScission Protease recognition sequences (LEVLFQGP (SEQ ID NO:69) (cleavage between Q and G residues) and/or TEV recognition sequences (ENLYFQG (SEQ ID NO:70) (cleavage between Q and G residues) and ENLYFQS (SEQ ID NO:71) (cleavage between G and S residues)) can be inserted at defined (e.g. amino acid residues 128 and 223) or additional positions to remove most of the globular head of HA from the baculovirus expressed, purified, fully folded, DT-stabilized, soluble HA precursor. Following antigenic confirmation, amino acid analysis and mass spectrometry can be performed to characterize the cross-linked molecule biochemically.

[0268] Proteolysis of the head domain can be carried out by standard biochemical procedures and assayed by SDS-PAGE electro-mobility shift from a molecular weight corresponding to a complete DT-HA trimer (225kD) to that of a headless trimer (135kDa) (Coomassie stain, Western blot). Removal of the head from the DT cross-linked HA stalk can be confirmed with Head-specific Abs, for example in Western blots and ELISA. The same bnAbs and assays described above can be used to confirm preservation of the most relevant QNEs in DT-Headless HA.

[0269] Amino acid analysis can be performed to assess any non-specific changes to amino acid side chains, and to confirm the presence of DT bonds (the DT moiety itself can be specifically detected). In order to identify the position of the DT bonds in DT-Headless, LC-MS/MS analysis of deglycosylated tryptic digest can be performed, for example on a Thermo Scientific LTQ Mass Spectrometer with a Michrom Paradigm HPLC and Vacuum Spray ionization source.

[0270] Biochemical characterization can be performed to identify variants of DT-stabilized, Headless HAs that retain binding to key anti-stalk quaternary bnAbs equal to the wild-type soluble PR8 HA trimer. If necessary, additional cleavage sites can be engineered in order to first unravel the head, and thereby improve the efficiency of proteolytic cleavage. Similarly, PreScission and/or TEV proteases and their cleavage sites can be used as described above.
Testing protection against challenge with drift and heterologous viruses

A PR8 HA variant can be expressed in mg-quantities, DT crosslinked, proteolyzed, purified, and antigenically characterized. PR8, NL09, and VN04 HALO/PR8_6+2 mutant virus preparations can be made. To establish the LD50 for each of the challenge viruses, for each virus 4 groups of 4 C57BL/6 mice (female, 6- to 8-week-old (Charles River Laboratories) can be inoculated, using 10-fold dilutions of the indicated viruses for each group around the published LD50 for each virus. To establish the optimal dose of purified DT-locked Headless HA trimer immunogen that protects 80%+ of animals from 5 X the LD50 dose of homologous (PR8) challenge, 4 groups of 5 C57BL/6 mice (female, 6- to 8-week-old (Charles River Laboratories) can be immunized with a prime-boost strategy consisting of consecutive injections of varying amounts of the purified DT-Headless HA immunogen with a fixed amount of Poly I/C adjuvant (10μg). Briefly, each group can be immunized with 0μg, 2.5μg, 5μg, and 10μg of DT-locked Headless trimer formulated with Poly I/C as an adjuvant. Three weeks later, the mice can be boosted, each with an equivalent amount of the adjuvanted immunogen. Three weeks after the boost, they can be challenged intranasally with a 5XLD50 dose of homologous (PR8) influenza virus. Mice can be monitored and evaluated for morbidity and mortality for a suitable time, such as 14 days. Mice losing more than 25% of their initial weight can be sacrificed and scored as dead. Survival can be defined as <25% weight loss. To test immunized mice for protection against a drift variant and group 1 heterologous challenges, three groups of C57BL/6 mice can be immunized with 10 μg of Poly I/C adjuvant only (“Adjuvant Only” control groups) and the remaining three groups can be immunized according to the schedule described above with the optimal dose of adjuvanted DT-Headless HA immunogen identified above (“DT-locked Headless Trimer” groups).

Two weeks after the final immunization, one group of Adjuvant Only and one group immunized with the optimized dose of DT-locked Headless trimer each can be challenged intranasally with a lethal dose of homologous virus (PR8 H1N1), the mouse-adapted novel swine pandemic drift variant (NL/09, H1N1), and with heterosubtypic, group 1 influenza virus (VN04 HALO/PR8_6+2 mutant H5N1) (Table 1). Mice can be monitored and evaluated for morbidity and mortality for 10 days and scored as described above.

Table 1: Immunization groups to assess breadth of protection
<table>
<thead>
<tr>
<th>Immunogen</th>
<th>Adjuvant only (negative controls)</th>
<th>DT-locked Headless trimer</th>
</tr>
</thead>
<tbody>
<tr>
<td>Challenge:</td>
<td>Homologous</td>
<td></td>
</tr>
<tr>
<td>homologous</td>
<td>(PR8)</td>
<td>(PR8) (positive control)</td>
</tr>
<tr>
<td>vs.</td>
<td>Drift variant</td>
<td></td>
</tr>
<tr>
<td>drift variant</td>
<td>(NL09)</td>
<td>(NL09)</td>
</tr>
<tr>
<td>vs.</td>
<td>Group 1 Heterosubtypic (VN04 HALO/PR8_6+2 mutant)</td>
<td>F. Group 1 Heterosubtypic (VN04 HALO/PR8_6+2 mutant)</td>
</tr>
</tbody>
</table>

[0274] Statistical Considerations: In view of the fact that both the predictor (adjuvant only vs. adjuvant + DT Headless immunogen) and the outcome (death vs. survival) are dichotomous, the null hypothesis that the vaccine has no effect can be tested with Fisher's Exact Test. To calculate the minimum number of animals per group (equal numbers in all groups) necessary to detect an effect at the 95% confidence level (p<.05), the power can be set to 80% and an assumed effect size of 50% can be used (80% lethality in the control group, 30% lethality in the vaccinated groups). Accordingly, each analyte and control group should use a minimum of 15 animals.

[0275] All methods can be carried out according to standard procedures, for example as described in Steel et al. 2010 [1]. For example, in ELISA assays the antigen (PR8 HA) can be immobilized with an α-foldon mAb (e.g. 74550, Fibrogen Inc.) or an α-stalk mAb to a non-quaternary epitope in order to optimize presentation of its native structure. Antigen-specific Ig in serum can be detected using labeled α-mouse Abs.

[0276] It is expected that DT-Headless will successfully induce protection against drift virus (group D: NL09, H1N1), and/or a homologous strain (group F: H5N1). If need be the immunogen can be reformulated with a different/additional adjuvant and/or the doses tested can be increased, and immunogen dose-calibration testing can be repeated. In addition, if need be the prime-boost regimen can be altered to include a third boost with purified DT-Headless HA antigen. The number of animals used in the final challenge study can be altered / increased to achieve an acceptable confidence level from homologous vs. drift and heterologous challenges.

References for Example 1

2. Pica, N. et al. Hemagglutinin stalk antibodies elicited by the 2009 pandemic influenza


33. Palese, P. Personal Communication. *Professor and Chair of Microbiology, Professor of Medicine, Infectious Disease - Mt. Sinal School of Medicine*
38. Crowe, J. E. Personal Communication. *Director, Vanderbilt Vaccine Center, Vanderbilt University Medical Center*

**EXAMPLE 2**

[0277] Recombinant, soluble protein immunogens represent a significant opportunity in the fight against natural and weaponized pathogens. Broadly neutralizing antibodies (bnAbs) against many pathogens have been described in recent years, many of which bind quaternary structures only displayed by protein complexes – which themselves are often unstable. Therefore, an urgent need exists to “lock” protein-based vaccine immunogens into the same native quaternary conformation as they are presented by the pathogen itself.

[0278] The present example relates to a headless hemagglutinin-based universal flu vaccine made using a system that includes (i) performing site-directed mutagenesis at positions where resulting Tyr residues are predicted to be in close structural proximity, (ii) expressing and purifying the mutant protein, and (iii) subsequently enzymatically crosslinking/locking of the fully-folded protein complex. DT crosslinking is targeted and zero-length, DT bonds are irreversible and do not form spontaneously, and, most
importantly, introduction of DT-bonds preserves protein structure and function, since it occurs once the protein is fully folded.

[0279] At present, a universal influenza vaccine immunogen is not available. Recently, however, headless HA constructs trimerized in the transmembrane domain were described that focus Ab responses on the highly conserved stalk, and that, indeed, elicit broadly protective responses. Soluble headless HA trimerized by a foldon motif, however misfolds, does not present key quaternary neutralizing epitopes (QNEs), and elicits insufficiently protective responses. The present invention provides an alternative system that involves (i) DT-crosslinking a soluble HA construct in its native, trimeric conformation, and (ii) removing the “head” domain by engineering and cutting proteolytic cleavage sites in the variable loops and at the base of the head. The interactions between the subunits of the HA soluble stem (lacking the transmembrane domain) can be locked by dityrosine crosslinking while maintaining the structural integrity of HA trimers. Based on crystal structures, constructs with Tyr side-chains in the stem and in close structural proximity can be made, while avoiding bnAb binding sites. These constructs can be expressed and the resulting proteins purified by His-tag affinity chromatography. Tests can be performed to determine whether the constructs form intermolecular DT crosslinks by screening for DT-specific fluorescence and by gel shift analyses (e.g. Western blots).

[0280] Using of a panel of anti-stem bnAbs, functional preservation of DT crosslinked HA trimer can be measured by ELISA using anti-stem bnAbs. Thermodynamic stabilization can be assayed to confirm the positions of DT bonds and the constructs’ structural integrity after crosslinking biophysically. Constructs can be selected based on favorable antigenic and/or biochemical profiles. It is expected that binding to quaternary anti-stem bnAbs such as 6F12, C179, CR6261, F10, A66 and D8 will be fully maintained.

[0281] Proteolytic cleavage sites can be engineered to unravel and remove the head of native, DT-stabilized HA trimers. Fully folded HA constructs with 4 or more cleavage sites for 1 or 2 proteases can be designed, generated, expressed, and DT crosslinked, and then purified by His-tag affinity chromatography before digesting with protease(s) to remove the head. Antigenic and biochemical and biophysical analyses can be performed to confirm preservation/integrity of QNEs in DT-headless HA after proteolytic digestion and immunogenic analysis can be performed in mice.
Seasonal and pandemic influenza viruses remain a serious threat to human health, due to their ability to evade immune surveillance through rapid genetic drift and re-assortment. In the US alone, influenza causes seasonal epidemics that contribute to hundreds of thousands of hospitalizations and an average of 30,000 deaths annually, while creating a serious economic burden for individuals and the economy as a whole [1-3]. Pandemic outbreaks occur when a virulent strain of virus emerges that infects people with little or no immunity, and rapidly spreads across the globe, representing one of the most serious threats to human health. The 1918 Spanish Flu (H1N1) pandemic caused an estimated 50 million deaths; the 1957 Asian influenza (H2N2) pandemic and the 1968 Hong Kong (H3N2) pandemic each caused several million deaths [6]. Because influenza viruses are readily accessible and are easily transmitted by aerosol, the possibility for genetic engineering represents an enormous threat of weaponization, biowarfare, and bioterrorism [7,8]. Vaccines hold the greatest promise of providing protection in order to control infection.

Although highly effective when matched to circulating strains, current influenza virus vaccines protect mostly against homologous virus strains. Protection is primarily due to high-avidity antibodies against the highly variable, immunodominant head domain of the hemagglutinin (HA) protein, which is specific to each strain of influenza. Therefore, new trivalent vaccine cocktails must be tailored each year to the prevalent influenza strains in circulation. Conventional, egg-based influenza vaccine manufacturing requires that strains be selected 9 months before the start of the season. Unfortunately, predictions of the circulating strains are often inaccurate, resulting in vaccines that are poorly matched, and therefore poorly protective [9-11]. A multitude of development programs are underway to address this problem, many of them in advanced stages, but the approach proposed herein has the potential to move one or more programs beyond the safety and efficacy hurdles, and enable a truly long-term broadly protective vaccine product for both seasonal and pandemic influenza.

The stem of HA is highly conserved across a multitude of influenza strains, and considerable evidence now suggests that vaccination with a ‘headless’ HA consisting primarily of the HA stem results in the induction of antibody responses with significantly enhanced heterosubtypic binding activity and broad protection against lethal challenge [12-15, 16, 17]. As such, headless HA holds significant promise as a universal vaccine.
immunogen capable of protecting against all strains of influenza [16] [17]. Interestingly, immunization with DNA coding for HA has been observed to generate predominantly stalk-specific Ab responses, and data describing heterosubtypic protection by vaccination with HA DNA by electroporation was recently described [50]. A prime-boost combination of an expression vector encoding HA, followed by soluble headless protein boost holds the promise of generating broadly heterospecific responses that yield long-lasting protection.

The present invention provides a soluble ‘headless’ HA trimer covalently stabilized in its correctly folded conformation that presents key quaternary neutralizing epitopes (QNEs). Targeted dityrosine crosslinking technology is used to stabilize a full-length HA trimer, and subsequently the head is removed using sequence/substrate-specific proteases - making use of variable loop tolerance for amino acid variation.

Dityrosine (DT) crosslinking provides a method for stabilizing protein folds, complexes, and conformations by enzymatically introducing zero-length crosslinks, while maintaining structural and functional integrity of the protein [20,21]. Dityrosine bonds provide conformational stability and rigidity to protein structures and have been described in many diverse natural settings. DT crosslinks form naturally in vivo, both in the context of proteins evolved to utilize their specific characteristics [22-24], and as a consequence of protein oxidation [25]. DT bonds form the structure of wheat gluten – the quaternary protein structure comprising the glutenin subunits – and are present in large quantities in some of our most common foods [26]. No other amino acids form crosslinks or are modified when the reaction is carried out under mild conditions, though the tyrosyl side-chains themselves may oxidize if positioned too far apart, thus limiting the efficiency of the reaction, particularly under sub-optimal conditions. DT crosslinks are not hydrolyzed under normal physiological conditions, and do not form spontaneously in vitro. These features of the DT-crosslinking provide important advantages over conventional S-S chemistry; namely spontaneous and/or undesired protein products do not form and non-specific bonding/aggregation does not occur on maturation and processing. Because the reaction can be tightly controlled, development of a large-scale high-yield process can be relatively straight-forward, making the large-scale manufacturing a DT stabilized immunogen more feasible economically.
One of the key features of DT crosslinking is that it is highly dependent on the structural proximity of tyrosyl side-chains, which must therefore be engineered within the structure of a protein or protein complex. Because no carbons are added in the formation of the bond, the resulting “staples” are non-disruptive to the overall protein fold and, critically, specific sites within the protein structure can be targeted with high specificity. The necessary tyrosines may be present in the primary structure of the protein or added by “to tyrosine” point mutations, while Tyr residues that form undesirable DT bonds can be mutated (to Phe, for example) to reduce background.

Protein immunogens are folded chains of amino acid polypeptides, sometimes consisting of several polypeptide subunits. The rate of spontaneous unfolding, conformational transition, and dissociation determines a protein’s functional half-life. Covalent non-peptide bonds between non-adjacent amino acid side chains can dramatically affect the rate of unfolding, and thus the half-life of a protein or protein complex. At least two different chemistries have evolved to accomplish covalent cross-links in proteins in vivo to stabilize their conformations and/or retard unfolding: these are disulfide bonds and dityrosine (DT) bonds.

One major advantage of a directed DT cross-linking approach is that covalent bonds targeted to specific locations can reinforce particular 3-D arrangements of epitopes’ secondary, tertiary, and/or quaternary structures, thereby preventing undesirable conformational transitions, and have the potential to provide a high degree of thermodynamic stabilization and conformational locking without adversely affecting the antigenic properties of protein immunogens.

Disulfide bonds have been found in many eukaryotic proteins of diverse function. Intra-molecular S-S cross-links are often essential in stabilizing protein domains, and inter-molecular S-S bonds provide stability for the quaternary structure of protein complexes. These bonds can form spontaneously, and therefore do not require an additional manufacturing and purification process, but also reduce manufacturing yields due to free sulfhydryl-mediated aggregate formation. Furthermore, because they are formed as the protein is folding in the ER/Golgi apparatus, they can lead to structural distortions that would affect QNE presentation and the breadth of immunogenic protection.
The C-C bond created by DT-crosslinking is stable under virtually any physiological and/or operational conditions that are likely to be used in accordance with the present invention, including those used in the process of immunization and vaccination. DT bonds are “zero length” – i.e. no atom is added. The cross-linking catalyst simply initiates bond formation between two tyrosines and is not incorporated into the product. Thus, no undesirable chemical modification of the protein occurs. DT cross linking is also very specific - no amino acids other than tyrosines have been shown to form cross-links or to be modified when the reaction is carried out under mild conditions. In addition, there is a strict distance requirement between the tyrosine side-chains, with the bond forming only when the two are in very close proximity. Furthermore, DT crosslinks do not form spontaneously, and, as described above, form only between Tyr residues in close proximity. DT crosslinking a protein can therefore lock it in its pre-existing native/functional conformation. In the context of headless HA design, this allows one to (i) engineer headless in an antigenically/immunogenically favorable conformation, e.g. by introducing point mutations, and then (ii) lock it in this preferred conformation by DT crosslinking.

Dityrosine bonds (DT bonds) that have important biological functions have been identified in proteins of several species, presumably in environments where disulfide bonds would be unsuitable. Specific DT bonds have, for example, been described in the cuticlin protein of Caenorhabditis elegans [27], the cell wall proteins of bamboo shoots [28], and parchment collagen [29]. In all of these cases, the proteins have evolved such that specifically placed DT cross-links contribute to the structural rigidity underlying the proteins’ functionality. The importance of such bonds is also evidenced by the fact that in yeast, for example, a metabolic pathway has been described that leads to the formation of DT bonds in specialized proteins [30].

Furthermore, due to the distinct fluorescent properties of DT bonds, in the absence of atomic level structures, their formation can easily be assayed using conventional 96- and 384-well fluorescence plate readers. This also makes optimization of cross-linking conditions simple and efficient.

The present methods involve (a) generating a DT stabilized full-length HA molecule that retains a stalk-specific antigenic profile equivalent to that of WT HA, (b) removing the head domain from the fully folded DT-HA by proteolytic cleavage while
retaining the same ‘stalk-specific’ antigenic profile as WT HA. Immunogenicity may be confirmed in animal studies.

[0295] The present example utilizes HA from the H1N1 A/Puerto Rico/8/1934 (‘PR8’) strain of influenza as the starting point. The majority of influenza virus research in mice employs lab adapted PR8 or the A/WSN/1933 (H1N1) [WSN] influenza viruses. Immunogenicity and challenge studies can be carried out in BALB/c mice with homologous and heterologous H1N1 PR8 and H3N2 X31 challenges. X31 is a reassortant virus carrying the HA and NA genes of A/Hong Kong/1/1968 (H3N2) in the background of PR8 [35].

[0296] To identify HA constructs which allow dityrosine bonds to form and stabilize the HA trimer, the trimeric HA crystal structure is analyzed (pdb file 3GBN) and proximal residues are selected for tyr-substitution away from the binding sites of quaternary neutralizing antibodies (see Figure 5). Once the in silico design of “to-tyr” point mutants (2T-HAs) is complete, cDNA encoding the ectodomain of wild-type HA (PR8) and to-tyr substitution mutants can be generated and cloned into a baculovirus transfer vector (pAcGP67A) using standard molecular biology techniques. WT and 2T-HA proteins can be expressed in SF9 or Hi5 cells and secreted HA can be purified over lectin-based glyco-affinity columns and MonoQ anion-exchange columns. Following purification, secreted HA trimmers can be isolated from monomers and high molecular weight aggregates by size exclusion chromatography (SEC) over a Superdex200 column.

[0297] To evaluate whether or not the designed 2T-HA constructs form intermolecular DT cross-links, the purified proteins can be analyzed before and after exposure to DT crosslinking conditions by gel-shift in reducing SDS-PAGE (Western blot and Coomassie stain) and for DT-specific fluorescence. Constructs capable forming DT cross-links with an efficiency of >50% can be taken forward for further characterization. Based on preliminary studies with HIV env trimers, it is believed that crosslinking efficiencies of greater than 80% are attainable without significant process optimization. Biochemical and biophysical analysis of DT-crosslinked HA trimers (DT-HA) be can be performed to compare their thermostability with that of un-crosslinked HA in normal human serum at 37°C over a time-course of 1-30 days. Trimeric DT-HA and control (uncrosslinked) trimeric HA can be analyzed each day for the presence of retained trimer by Western blot. Likewise, a 60-day, 25°C time course in PBS (pH 7.4) of purified, trimeric DT-HA and control (uncrosslinked)
trimeric HA can be analyzed weekly by SEC. The proportion of total material in the trimeric and monomeric fractions can be quantified using standard peak-integration software and the ratio of trimer to monomer in the DT-HA and control samples can be determined. Given that DT-HA constructs can be identified based on their stability in reducing SDS-PAGE, it is expected that 100% of the DT crosslinked trimer will remain trimeric under the experimental conditions described above, while labile uncrosslinked HA trimers will dissociate into monomeric subunits throughout the duration of the time course.

A central advantage of DT crosslinking technology over other crosslinking methodologies is the ability to form covalent intermolecular crosslinks without disrupting the antigenic profile of vaccine immunogen candidates. The effect of the both the “to-tyr” mutations and the DT crosslinking can be determined by ELISA using a panel of anti-HA stem broadly neutralizing mAbs (e.g. 6F12, C179, CR6261, F10, A66 and D8). Full-curve binding assays can be used to compare WT HA trimers to the 2T-HA mutant trimers (uncrosslinked) and to DT-HA trimers (crosslinked). Changes in binding following the introduction of to-tyr mutations as well as after DT crosslinking can be determined using non-linear regression analysis of binding curves to calculate and compare EC50 values for each construct with each mAb. The position of to-tyr mutations can be distal to and non-overlapping with amino acids involved in binding of the anti-stem bnAbs listed above. It is possible that structural changes caused by tyrosine substitutions may reduce or enhance binding to some of these antibodies. However, preliminary studies using HIV suggest that DT crosslinking fully preserves a protein candidate’s antigenic profile and a similar degree of antigenic preservation is expected following DT crosslinking of influenza HA.

In order to assess non-specific changes to amino acid side chains throughout the entire crosslinked protein, comparative amino acid analysis (AAA) can be performed on uncrosslinked (control) and crosslinked constructs. Amino acid analysis can also be used to confirm the presence of DT bonds since dityrosine crosslinks withstand even the acid-hydrolysis used to prepare samples for AAA and dityrosine itself can be specifically detected in the analysis. In order to directly identify the position of the dityrosine bonds in DT-HA, mass spectrometry analysis of deglycosylated tryptic digests can be used, for example by performing LC-MS/MS on a Thermo Scientific LTQ Mass Spectrometer with a Michrom Paradigm HPLC and Michrom Vacuum Spray ionization source. Collectively
these studies can be used to identify and characterize HA constructs capable of forming trimerizing DT bonds. Such constructs may, even prior to removal of the immune-dominant HA head domain, provide improved HA immunogens stably presenting stalk specific QNEs.

[0300] Previously reported recombinant headless HA constructs do not retain the fully native, quaternary structure of the HA stem and thus, these constructs do not bind known quaternary specific bnAbs. Following baculovirus expression and purification of DT-HA construct(s) as described above the head domain can be removed proteolytically - post-folding and after DT crosslinking - in order to generate a stable headless HA which retains binding to broadly protective, conformation-dependent quaternary antibodies. In order to enable the proteolytic removal of the globular head domain of HA, protease cleavage sites can be introduced into HA1. Head-removal sites can be introduced at, for example, positions 60-76 (N-terminal site) and 277-290 (c-terminal site) through standard molecular biology techniques [19]. Crystal structures of HA indicate that these positions are solvent-exposed and could be made further accessible to proteases by removing the structural constraints that may hamper efficient proteolysis through the introduction of additional cleavage sites into the HA1 variable loop domains (AA positions 142-146 and 155-164) [37]. Unraveling the head can be used to further improve protease substrate access, if required. Introduction of cleavage sites into the HA variable loops is not expected to alter the overall conformation of the HA trimer as these sites are highly tolerant of amino acid substitutions. Indeed, all of these amino acid positions (e.g. 142-146 and 155-164) have changed in infectious virus isolates collected from 1968 through 1999 [38]. Cleaving HA1 in the variable loops can be performed to destabilize the head's globular structure, allowing complete exposure and efficient cleavage at the primary head-removal sites (53-67 and 269-277). PreScission Protease (GE Healthcare Life Sciences) recognition sequences (LEVLFQGP (SEQ ID NO:69)) and TEV (Tobacco Etch Virus protease) recognition sequences ENLYFQG (SEQ ID NO:70) and ENLYFQS (SEQ ID NO:71)) can be used/introduced. TEV cleavage can be carried out at a substrate to enzyme ratio of 1:50-200 w/w in a 25 mM Tris-HCl buffer with 150-500mM NaCl, and 14 mM β-mercaptoethanol at pH 7.0. PreScission Protease cleavage can be performed in a 50mM Tris-HCl buffer, with 150 mM NaCl, 1 mM EDTA and 1 mM dithiothreitol (DTT) at pH 7.0. Removal of the head can be assayed by electro-mobility shift from a molecular weight corresponding to a full-length DT-HA trimer (~225 kDa) to that of a headless trimer (~135 kDa) by SDS-PAGE, followed by coomassie stain and
Western blot. Head-specific detection Abs can be used to confirm removal of the head from the DT crosslinked HA stem by Western blot and ELISA. If HA head-removal is incomplete, the positions of Precission Protease and TEV sites can be swapped, or, only a single type of site can be introduced at all desired cleavage positions.

[0301] In order to test the immunogenicity of the DT-headless constructs, mouse immunogenicity studies can be performed. BALB/c mice (6-8 weeks old) can be anesthetized with isoflurane 3-5% and subsequently immunized in a prime-boost regime/schedule with two intramuscular injections 3 weeks apart, first with DNA comprising 37.5μg of pGag-EGFP and 75μg of pDZ_PR8_HA followed by electroporation pulsing (prime), and subsequently with 25μg of WT HA, foldon/GCN4-stabilized HA trimers, or DT-headless protein (boost). Protein (boost) immunogens can be formulated with Alum (Aluminum phosphate, 300μg/dose). Two weeks following the second injection (boost), serum can be collected and assayed for anti-HA responses relative to pre-immunization serum and adjuvant only controls. Overall anti-HA IgG and IgM titers for each group can be determined by ELISA. Heterosubtypic reactivity of anti-sera to 10 different purified group 1 and group 2 HAs can be determined Western blot and ELISA. Immunogens from each group are expected to elicit anti-HA antibody responses. In order to investigate the heterosubtypic neutralization capacity of anti-serum from each group, the ability of these sera to neutralize a panel of heterologous influenza viruses (HK/68 H3, Bris/07 H3, Neth/03 H7, Cal/09 H1, Sing/57 H2, Vict/04 H5, HK/97 H6, HK/99 H9) can be tested. Anti-serum can be serially diluted 2-fold, mixed with an equal volume of virus, and incubated for 2h at 37°C. Virus-serum mixtures can be added to target cells (MDCK) in serum free media containing trypsin and incubated for 3h prior to replacement of the media. Cells can be monitored for cytopathic effects 3-5 days following exposure to virus-serum mixtures.

[0302] A major objective of this immunogen design and development process is to generate a DT-headless immunogen capable of eliciting bnAbs and protecting against heterologous influenza challenge. To directly investigate the ability of DT-headless to elicit protective responses against influenza infection, 3 groups of 20 BALB/c mice can be immunized with WT HA, foldon/GCN4 headless, or DT-headless, compared to non/pre-immunized and adjuvant-only immunized controls (groups 4 and 5, 20 mice each), and
challenged intranasally with a lethal dose of homologous (PR8) or heterologous (X31) virus – 10 mice each – 2 weeks following the second immunization (boost). Mice can be anesthetized with an intra-peritoneal injection of ketamine (75mg/kg) and xylazine (15mg/kg) prior to challenge, and body weight can be monitored daily. <20% weight loss can be used as a surrogate for survival. It is expected that each immunogen (WT HA, foldon/GCN4 headless, DT-headless) will provide some degree of protection against PR8 challenge. However, it is expected that immunization with a DT-headless immunogen will provide significantly improved protection against heterologous influenza challenge and that this protection will correlate with the titers of bnAbs recognizing conserved QNEs that presented on the native HA stem in its trimeric form.

[0303] The baculovirus expression vector system (BEVS) can be used for manufacturing of recombinant HA antigen as this system is well established and suitable production/purification protocols have been well described and validated [10]. Generally, such protocols involve harvesting infected cells by centrifugation, detergent-mediated protein solubilization, followed by purification involving two chromatographic (IE and HIC columns) steps [10]. Due to the large difference in MW of the trimeric stalk as compared to the monomeric head, and the enzymes used in processing, gel filtration can also be used. IE chromatography can also be used.

[0304] Two enzymes are used in the processes described herein - peroxidases to catalyze the formation of DT bonds and proteases to cleave off the HA head after cross-linking. Both are commercially available.

[0305] Purity of the finished immunogens can be ascertained by conventional gel electrophoresis and HPLC. Cross-linking can be assessed by a combination of gel electrophoresis under denaturing conditions, fluorescence measurements, and amino acid analysis. Immunogenicity can be assessed by profiling against a panel of selected antibodies as described above. HPLC-based assays can be used to identify and measure protein sugar compositions.

[0306] DT-Headless HA can be formulated with an adjuvant selected based on technical specifications and other considerations. Adjuvanted HA formulated with a variety of excipients and stabilizing agents/preservatives can be lyophilized, and following rehydration
tested biophysically (dynamic light scattering) and antigenically. The effect of storage at room temperature, 4°C and -20°C can be tested to determine long-term storage conditions, stability, and potency.

[0307] Animal efficacy studies (e.g. conducted in ferrets) can be performed and acute and long-term animal safety studies can be performed. Ferrets are susceptible to human influenza viruses and develop some of the symptoms of influenza that are seen in humans; furthermore, they are large enough to monitor clinical parameters (e.g. temperature, pulse, and respiratory rate), and relatively large amounts of sera can be obtained for use in serologic and antigenic characterization.

References for Example 2


EXAMPLE 3

[0308] Prior attempts at engineering headless HA have included expressing HA proteins in which the globular head region was recombinantly spliced out. Such prior headless HA constructs generated considerable excitement in the field, because they elicited improved, cross-reactive Ab responses. These Abs, however, were not cross-protective and only protected against homologous challenge. These prior recombinant headless constructs do not bind the current repertoire of cross-protective, conformational stalk Abs, which suggests at least some degree of stalk mis-folding in the absence of the intact globular head. These prior observations were confirmed using one of the most broadly cross-reactive anti-stalk Abs, C179, by immunofluorescence analysis (see Figure 6). Application of DT-based conformational locking will circumvent this shortcoming, by holding together the stalk trimer in its native conformation, prior to proteolytic removal of the head, and thus result in a DT-locked Headless HA immunogen that will focus Ab responses on the critical stalk QNEs.

EXAMPLE 4
[0309] DT crosslinks were introduced into the PR8 HA stalk domain, and the DT-crosslinked HA trimer maintained native antigenicity. Based on the crystal structures of the 1918 H1N1 HA trimer in complex with the cr6261 bnAb (pdb file: 3GBN) and of the PR8 HA (pdb file: 1RU7), tyrosine substitutions were successfully engineered into the HA stalk domain in order to enable DT crosslink formation, which should maintain quaternary antigenicity upon proteolytic removal of the globular head. 293T cells were subsequently transfected with secreted variants of the tyrosine mutants and measured at 405nm fluorescence in transfected cell supernatants, to determine the formation of DT bonds. A large increase in 405nm fluorescence (highly specific for DT bonds) demonstrates robust crosslinking in several tyrosine mutants (Figure 7 A). Based on comparison to the positive control (insulin) and a DT standard, a cross-linking efficiency of >70% was confirmed for four of these constructs, prior to any optimization. As shown in Figure 7B, C179 Ab binding is unchanged before and after the crosslinking reaction. These data show that the PR8 HA stalk can be cross-linked and that the key quaternary stalk epitope bound by one of the most broadly cross-reactive, conformational mAbs, C179 (2), is maintained.

[0310] Targeted protease cleavage sites were also successfully introduced and used to cleave the PR8 HA head domain. Extensive analysis of the PR8 HA structure and transposon-based mutagenesis studies revealed multiple locations within the globular head region that could tolerate insertion of proteolytic cleavage sites. Out of 20+ possible sites identified, two constructs that allow insertion were generated. One site is located at the base of the globular head domain (“48G”), while the other resides closer to the variable loops of the protein (“128S”). Both insertions express well as indicated by Western blot of whole cell extracts (Figure 6B, left) and form virus-like particles (VLPs) in sufficient quantity for detection in transfected cell supernatants by C179 ELISA (Figure 8A). Of the two constructs generated, the 48G insertion is predicted to be least accessible for the protease, yet would most completely remove the Head domain, due to its location close to the base of the head. In order to demonstrate that the 48G site is sufficiently accessible, TEV protease cleavage was performed on the HA 48G protein, using WT HA as a negative control. As shown in Figure 8B, right, TEV protease cleavage of the HA 48G protein results in the removal of the first 48 AAs (6.5kDa) of HA, but no cleavage occurs in the WT HA protein. Furthermore, HA 48G also maintains hemagglutination activity when assayed directly from
transfected cell supernatants, suggesting that it remains folded in its functional conformation (Figure 8C).

**EXAMPLE 5**

[0311] Introduction: In designing a conformationally locked headless HA, the atomic structures of the 1918 HA:cr6261 complex (PDB:3GBN) and PR8 HA (PDB:1RU7) were analyzed to identify positions that 1) enable dityrosine (DT) crosslinking in the stalk at a sufficient distance from the cr6261 epitope to maintain stalk bnAb binding; and 2) enable insertion of protease cleavage sites, that can be used to remove the head.

PR8 HA trimers were successfully locked in their native trimeric conformation using DT crosslinks at several locations in the HA stalk; and these DT-locked HAs maintain native stalk antigenicity.

[0312] Several tyrosine mutations were engineered into the stalk of PR8 HA that enable the trimers to be locked in their native prefusion state, at high efficiency. Fig 85A demonstrates a clear shift to the trimeric state (reducing SDS-PAGE) following DT crosslinking; and Fig 85B confirms that di-tyrosine bonds have formed by specific fluorescence at 405nm. Densitometry of the cross-linked species demonstrates greater than 80% conversion to the trimeric state. Most importantly binding of 8D4, a stalk-specific bnAb, is fully maintained (Fig 85C). Crystallographic analysis has shown that 8D4 binds the same epitope as cr6261. These data confirm that the PR8 HA can be cross-linked in its stalk while maintaining the native conformation of the key V\textsubscript{H}1-69 quaternary stalk epitope.

Multiple C- and N-terminal TEV protease recognition sites have been successfully engineered into the head of PR8 HA, individually and in combination.

[0313] Regions were identified in the head of PR8 HA into which TEV protease cleavage sites can be inserted without disrupting HA’s function. For structure-based design, PR8 HA and TEV protease recognition site structural data were combined, and cleavage site insertions were specifically targeted into regions of HA’s head based on the following criteria: i) proximity to the stalk apex, in order to maximize removal of the immune-dominant head; ii) similarity between the secondary structures of HA and the TEV cleavage site, to minimize structural perturbation; iii) regions identified as tolerant of insertion based
on the transposon mutagenesis screen with data from a transposon-based mutagenesis screen
(Heaton and Palese PNAS Vol. 110, No. 50; pp. 20248-53).

[0314] In total, approximately 50 insertion sites were screened individually, and assayed
for their ability to be incorporated into VLPs by expressing only HA (WT or with insertion)
and NA. This assay encompasses several of HA’s functional attributes, including
expression, cell surface accumulation, membrane microdomain localization, and particle
formation, and was therefore performed with full length HA proteins. This approach
identified several C-terminal (e.g. at positions 278, 282, 283, 286, and 291) and 2 N-terminal
(positions 48 and 63) where TEV insertions maintain the function of WT HA in the VLP
formation assay (Fig. 86 A). Several of these also maintain stalk bnAb binding, and cleave
efficiently (Fig. 86 B & C). Two C-terminal insertions (positions 278 and 286) and 1 N-
terminal insertion (position 63) were prioritized for further analysis and testing in
combination. A second N-terminal insertion (position 48), binds well to certain anti-stalk
V_{H1}-69 bnAbs (e.g. 18A3), but less well to others, and therefore could nonetheless provide a
reasonable alternative to the insertion at position 63.

[0315] Some of the prioritized insertion sites have been tested in various combinations
(e.g. 63-278 and 63-286) and have shown that both of the combinations of insertions also
maintain efficient VLP formation (Fig 87A) and bind well to stalk bnAbs (Fig87B).

[0316] Going forward both components (to-tyrosine mutations and proteolytic cleavage
site insertions) can be introduced into a single HA molecule. DT crosslinking can then be
applied to lock the stalk of HA in its trimeric, prefusion conformation, and following this the
head can be removed proteolytically to generate a fully native, headless HA.

[0317] While the foregoing invention has been described in some detail for purposes of
clarity and understanding, it will be clear to one skilled in the art from a reading of this
disclosure that various changes in form and detail can be made without departing from the
true scope of the invention. The invention may also be further defined in terms of the
following claims.
CLAIMS

1. An influenza hemagglutinin (HA) polypeptide, protein or protein complex comprising an amino acid sequence having at least 65 % sequence identity to amino acid residues 229 to 519 of SEQ ID NO: 1, wherein the amino acid sequence comprises a point mutation to tyrosine at one or more of amino acid positions 403, 406, 411, 422, 429, 432, 433, and 435, or an amino acid residue corresponding thereto.

2. An influenza hemagglutinin (HA) polypeptide, protein or protein complex according to claim 1, wherein the amino acid sequence comprises point mutations to tyrosine at two or more of amino acid positions 403, 406, 411, 422, 429, 432, 433, and 435, or amino acid residues corresponding thereto.

3. An influenza hemagglutinin (HA) polypeptide that consists essentially of an amino acid sequence having at least 65 % sequence identity to amino acid residues 229 to 519 of SEQ ID NO: 1, wherein the amino acid sequence comprises a point mutation to tyrosine at one or more of amino acid positions 403, 406, 411, 422, 429, 432, 433, and 435, or an amino acid residue corresponding thereto.

4. An influenza HA polypeptide, protein or protein complex according to claim 1, wherein the polypeptide, protein or protein complex is folded into a trimeric stalk conformation and comprises at least one di-tyrosine cross-link, wherein one or both tyrosines of the at least one di-tyrosine cross-link originate from a point mutation to tyrosine.

5. An influenza HA polypeptide according to claim 3, wherein the polypeptide is comprised with an influenza HA protein complex that is folded into a trimeric stalk conformation, and wherein the influenza HA protein complex comprises at least one di-tyrosine cross-link, wherein one or both tyrosines of the at least one di-tyrosine cross-link originates from a point mutation to tyrosine.

6. An influenza HA polypeptide, protein or protein complex according to claim 4, wherein the cross-links are located between one or more paired tyrosine residues, wherein the paired tyrosine residues are selected from the group consisting of residues 403 and 433; 411 and 422, 403 and 429, 403 and 432, 433 and 435, and 406 and 433.
7. An influenza HA protein complex according to claim 5, wherein the cross-links are located between one or more paired tyrosine residues, wherein the paired tyrosine residues are selected from the group consisting of residues 403 and 433; 411 and 422, 403 and 429, 403 and 432, 433 and 435, and 406 and 433.

8. An influenza HA polypeptide, protein or protein complex according to claim 1, wherein the influenza HA polypeptide, protein or protein complex has the amino acid sequence of SEQ ID NO: 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 86, 87, 88, 89, 90, 91, 92, 93, 96, 97, 98, 99, 100, 101, 102, 103, 104, 105, 106, 107, 108, 109, 110, or 117, or an amino acid sequence having 65% or greater sequence identity to any of such sequences.

9. An influenza HA polypeptide, protein or protein complex according to claim 1, wherein the polypeptide, protein or protein complex is capable of binding to an HA-stalk-specific antibody.

10. An influenza HA polypeptide, protein or protein complex according to claim 9, wherein the polypeptide, protein or protein complex is capable of binding to antibody C179.

11. An influenza HA polypeptide, protein or protein complex according to claim 1, wherein the polypeptide, protein or protein complex is capable of folding into a trimeric stalk conformation.

12. An influenza HA polypeptide, protein or protein complex according to claim 1, wherein the polypeptide, protein or protein complex further comprises a trimerization domain.

13. An influenza HA polypeptide, protein or protein complex according to claim 12, wherein the trimerization domain is a foldon domain.

14. A nucleic acid molecule encoding an influenza HA polypeptide, protein or protein complex according to claim 1.

15. A composition comprising an influenza HA polypeptide, protein or protein complex according to claim 1.
16. The composition of claim 16, wherein the composition is a vaccine composition.

17. The composition of claim 17, wherein the composition further comprises an adjuvant, a carrier, an immunostimulatory agent, or any combination thereof.

18. A composition comprising an influenza HA polypeptide according to claim 2.

19. A composition comprising an influenza HA polypeptide according to claim 2 and an influenza HA polypeptide consisting essentially of an amino sequence having at least 65 % sequence identity to SEQ ID NO: 94 or SEQ ID NO: 95.
Figure 1
Figure 4 A

DT Specific Fluorescence (ex:320 em:405)

- WT PRB
- WT PRB-DT
- 403-420
- 403-429-DT
- 403-430
- 403-432-DT
- 403-432
- 403-433
- 406-433
- 408-433-DT
- Insulin
- Insulin+DT
Figure 4 B
Figure 7

Relative DT Fluorescence

OD 450nm

No crosslink
DT crosslink
**Figure 8**

A

C179 ELISA Supernatants

<table>
<thead>
<tr>
<th></th>
<th>WT</th>
<th>48G</th>
<th>128S</th>
<th>(-)</th>
</tr>
</thead>
<tbody>
<tr>
<td>OD_{450nm}</td>
<td>0.4</td>
<td>0.3</td>
<td>0.2</td>
<td>0.1</td>
</tr>
</tbody>
</table>

B

<table>
<thead>
<tr>
<th></th>
<th>(-)</th>
<th>WT</th>
<th>48G</th>
<th>128S</th>
</tr>
</thead>
<tbody>
<tr>
<td>WCEs</td>
<td>107A</td>
<td>64-</td>
<td>51-</td>
<td>39-</td>
</tr>
</tbody>
</table>

Purified VLPs

- WT
- WT+TEV
- 48G
- 48G+TEV

**Anti-HA**

C

<table>
<thead>
<tr>
<th></th>
<th>(-)</th>
<th>WT HA</th>
<th>48G</th>
</tr>
</thead>
</table>

HA assay
Figure 9

(SEQ ID NO.1)

MKANLLVLLSALAAADATIGYHANNSTDGVVTVLTVTVSVMNLDSHGKLCRLKGIAPQLLG
KCNIAGWLLGNPECPLPVRSSWVSYIVETPNSENGICYPGDFIDYEEILREQLSVSFFFERFEIPKESSW
PNHNVTNGVTACSHKEGKSFFYRNLWLTEKGSYPYFKLKNSYVNKKGKEVLLWGIHHPPNSKEQNIYQN
ENAVYSVVTSNYNRRFTEIAERPQVRDQAGRMYWYWTLLKKPGDITFEANGLAPIAMYAFALSQFGSG
IITSNASMHCNTKQCTPLGAINSSLFYQNIHPVTIGECPKYVRSKLRMVTGLRNTPISQSRGLFGAIA
GFIEGGWTDIGWBGYHHQEQGSGYAAODQKSTQNAINGITNKVNTVIEKMNQFTAVGKEFNKLEKRM
ENLNKKEDDGFQDIWTYNAELLVLLENERTLDFHDSNKNLYEKVQLRNNAEIGNGCFEFYHKCDINE
CMESVRNGTYDPKYGSSKLNRKEDVGKLESNGYQILAIYSTASSSVLVLVLSGAI5FWMCNKLQ direkti

gi|21693169|gb|AAM75158.1|AF389118_1 hemagglutinin HA [Influenza A virus (A/Puerto Rico/8/34/Mount Sinai(H1N1))]

CRICI
Influenza A virus (A/Puerto Rico/8/34/Mount Sinai(H1N1))
Figure 12
(SEQ ID NO.4)
Figure 13
(SEQ ID NO.5)

>PR8HA-63G278S-403Y411Y422Y433Y.pro

MKANLLVLLSALAAADALTICIGYHANNSTDVTVDTVLEKNTVTTHSVNITLLEDSHNKLCRLKGELFLYPOQKCHNIAGKLGL
NFECDPPLLVPVSNWYFIVETPNSENGICYPQDFIDYEELREQLSSVSSFERFEIFPKESSWPNHNTINGVTAACSHEGKSSF
YRNLALTEKEGSYPKLKSYVVIKKGKEVLVLQGIHHPFNSKEQHIIYQNENAYSVTVTSHYRRFTPIABFKVRDQA
GRNYYWTLKPGDTIIIEAINGHLIAVAFALSRGFQRNLQFQSNASHHECNYTKCQPTLGAINGSLOPYQNIIHFVTIGEC
FKYVRSAKLMMVTGSIPQSRGLFGAIAFIEGGWTGIDGWYGYYHQQEQGSYAADQKSTQHAININGITNKVNTVI
EMYIQFTAVGNYFNPKEKRMELYIKKVDGFLYIWTYNAELLVLLEHERTLDHDSNKVNYEKVSQLKNNAKEIING
CFEYHKCDHCMESVNGTYDYPKYSEEKLNREKVDGVKLESNGIYQLAIYSTVASSVLVFLGLAISFGMCMNGSL
QCRIICH
>PR8HA-63G282S-403Y433Y.pro
MKANLVLVLSALAAADALTICIGYHANNSTDVTVDTVLEKNTTTHSVMHLEDHNGKLCRLKGEHLYFGKCNIAGWLG
NFECDDLVPVRSGYIVETPISNGICYPGDFIDYEELREQLSSSVSSFPERFEIFFKESSPNHINTNGVTAACSHEGKSSF
YNHLLLALTEKGGYFKLKNLYVKKGVKVIVLWGIIHHPNKFSEQQHITYQHENTAYVSVVTVHRRFTP81BRPKVQDQA
GRMNYWTLLKPGDIIIFEAIGNHLIAFMYAFALSRGFGGSIIEENLYFQSTSHASMHECNTKQCTFPLGAINESSLIFYQNIHP
VTIGECFPYVRSAKLRTGLRTPSIQSRLFQAIAGFEQGWTGMIDGWYTHQNEQGGSYAADQK3TQNAINGITNH
KVTVEKMIQFTAVGKEFNPKENHNLNKVDDGFLYINTYJAELLVLLLENERTLDPHDSGVKINLYKSVQSLKINNA
KEIGNGCPEFYHKDNCEMESVNTYDYFKYSEESKNREKVGKLESNGYOILAISTVAGSLVLYVSGIAISEYXMN

C5H8GSLOCRICK
>PR8HA-63G282S-411Y422Y.pro
MKANLLVLLSALAAADADTCIGYHANNSTDTRTVLKEKVTVTHSVNHLLEDSHNGKLRCRLKGKGEILYFQGKCNIAKWL
NPFCDPPLLVPVRSWGYIVETPSNGICYPGDFIDYEELREQLSSVSSFERFEIIPKESSWPNHINTNGVTACSHGBKSS
YRNLLALTEKEGSYPKLNSTYVNNKGEVNLW#GIHHPPNSKEQHNYQHENAYVSVVTSHYRHRFTPQAIABRFKVRDQA
GRMNYYWTLLKPGDTIIFEAINGHLIAHPYAFAL5RGFGSGIIIEHLILYFQGSTSHASMHECNTKQCTPLGAINSSLFMPQNH
VTIGECFKVYRSAKLRMVTGLRMTPSQTQSRLLFGAIAGFIEGGW7TMIDGWSGYHYHNEQGSGYAADQKTQNAINGITN
KVNTvieKvHnQPTAVGYEFNKLKvtkRMVNLKVKVDDGFLDIWYNAEELLVLLHETFCLDFHDSNVHMKLYEVKSQKLNN
KEIGHGCFEFYHCDNICECMESVRNGTYDYPKYEESKLREKVGDQKHLSMGYOILAYSTVASLTVLLVSCGAISEY
QSNGSLQCRIKX
>PR8HA-63G283G-403Y433Y.pro
MKANLLVLSALAAADATCIGYHHNSTDIDTVLVTLEKNTVTTHSVIDLLEDSHIGKLCRLGKEHLYFGQGNCIAGNWLLG
NPECDDPLLVPVRSWSYIVETPNSENGICYPFDFIDYEELREQLSSVSFSEFERFEIPKESSWPNHITNIGVTAACSHEGKSSF
YRHLLALTEKEGSYPKLNYVNIKGGEKVLVLKIHHPNSKEEQHNYIHQNAYSVVTVTSHYNRRFTPEIAERPKVRDQA
GRNNYYWTLLKPDITIEAINGHLAPLYAFALSRGFSGIIITAELYFQQSINSMECNHTCQFTPLGAINSSLFYQNIHP
VTIGECFYVRSAKLRLMRTLSIQSRGLFAIAGFIEGGWTQMDGWYGNYHQNEQ鲜QSSGAADQKSTQNAINGITN
KVNTVIEKMYIQFTAOGKRFKNLMKVLKDDGFLYIWTNYAEILVLEHERTLDFHDSIVKHLYEYVKSQLKIN
KEINGCGCFEFYHKCNECMESVRNGTYDYPKSEESKLNREKVDGWFLESNHQIOIALYSTGASSLVLVLSIGAGFENK
CSNGSLACR ICT
>PR8HA-63G283G-411Y422Y.pro
MKANLLVLLSALAAADADTICIGYHANNSTDTVDVLEKNTVTHSVMHLEDSHNGKLCRLKGEHLYPQGKCIAGWLLG
NFEDPPLLPRSGYIVETFHSGNCYPGDFIDYELSEQLSSVSFFERFPEIFKESWPHNTINGVTAACSHEGKSSF
YRLLALTEKEGSYPKLKNYSNYVKKGEVVLGVIHPFNSKEQQHIYQHEAYVSVTVTSYHRRTPEIABRFKVRDQA
GRMNYYWTLLKPGDTIFEAINGKLIAMYAFALSERGFSGIGITENLYPQSSHASMRCNTKCQFTPLGAINSSLFYQNIHP
VTIGECFXYRSAKLRMVTGLRNTPSIQSRGLFGIAIFEGGGTMIDGWYGYHQQEQGSGYAADQKSTQNAINGITN
KVNTVIKMNIVQPTAVYEFPKLEKRMEVLHNVDDGFLDINTYNAELLVLLENERTLDFHSNVHKLYEKVSKQLKIN
KEIGNGCFEFYHKCDNIECMESVRNTYDYFKESEKLNREKVDDGKLESMGYOILAIYSTVASSLVLILSICAIFFNN
[CSHGLQCRICK]
>PR8HA-63G283G-403Y411Y422Y433Y.pro
MKANLLVLLSALAAADADTICIGYHANNISTDTVDTVLEKNTVTSHSVHNLLEDSSHNGKLCRLKGEHLYPOQKCNIAAGWLLG
NPECDEPLLFPVRSWYSIVETFPSSENGICYPGDFIDYEELBQLSSVSFSFERRFHEPPKESWPNMNININGTAACSHECGSSF
YRHLLALTEKEGSYPKLKSYSVNHKGEVLVLWGIIHPFNSKEQOHYYQHENAYVSVTVTSHYHRRTP81ABRKFQVRDQA
GRMNYYWTLKKPDTIIFEAINGHLIAFMAL5RGFSGSITENHLYPQQSNIASMHBNCNTKCTQTPFLGAIIISSLFYQNIHP
VTIGECFKYVRSKLRMVTGLRNTPSIQSRLFGLAIAGFIEGGWTGMIDGWYGHYHQQNEQGSGYAADOKSTQAINGITHN
KVNTVIEKMYIQPTAVGVEFQNKLERKMFLNKVDGFLYINTYAELLQLLENERTLDFHGUVKNLYEKVSQLINKNIA
KEIGNGCFFEFYHKCDNEMESVNGTYDYPKYSEESKLNREKVDGKLESIGYOIILAISTVAGSSLVLVSICGASEFN
CSNHSLQCRIC
>PRA8A-48G291G-403Y433Y.pro
MKANLLVLLSAALADADTICIGYHANNSDTVDVLEKTVTTVTHSVEILYFQGNLLEDASHNGKLRLIGIAFLQLGLGCN
IAGWLGNFCDPLLFRSWSYVFETPNSEGNGICYPDFIDYEELREQLSSVSSFERRFEIIPKESSWPHHTYNGSTAACS
HEKSSFYRHLWLTEKESYPKLNSYVHKGKEVLVLWGLIHHPPNSKEKIQNIYQHNYAYVSVVTSHYURRFTPEIAER
FKVRDQAGRMNYYTLKFPDTIIEFANGLHIAHMAFALSRGFGSIITSMASHMECEHLVYFQHNTKCPFTPLGAINESSL
FYQNIHFVTIGECFVTSAKLRTVMGTPISQSRGFGAIAGFEIGGWTGMDGWGNYHHQNEQSGYADQKSTQNG
AINGITVKVHTVIKEMYIQPTAVGKEFNNLEKRMENLNNKVDGFLYWTYNAELVLENLNEERTQDHPDSNVMLYEKKV
SQLKNNAKEIGNGCFEFYHKCDNENCSVRNGTYDFKSYEEKLNKEKVDGGLKLSMGYCALAXYTVASSLYLLVL

**Figure 20**

(SEQ ID NO.12)
Figure 21

(SEQ ID NO.13)

>PR8HA-48G291G-411Y422Y.pro
MKANLLVLGLALAAADADTICIGYHANSTDTVDTVEKNTVTTHGVENLYFQGNILLED5HNGKLRCRLGIAFLQLKGCN
IAAGKLLGNPECOPLLFVRSWSYIVETPENSEMIGICYPGDFIDYEELREQLSSVSSFERFEIFPKESWPHHINTNGVTAACS
HEGKSSFYRNLWLTEKEGSGYPDKNSYVHKKGKEVLMLGIGIHPPJSKEQNIYQHNYAYVSVTSHYNRRTPEIAER
PKVRDQAGRNIYYTWLKKPGDTIIFEEANGHIAFMIALGSGGILTHASMHIECNLYFQGNTKQQTPLGAINSSL
FYQNIHPVTIGECPKYRSAKLMVTLRNPISQSRLFGAIAGFIEGGWGMIDGWGNYHQPNEQGSGYAADQKSTQN
AINIGITNKVNHTVIKMNQFTAVGYEFPNLKEKRMEDYNKKVDGFLDIWTYHAELLVLLEHERTLDPHDSNVKINLYEKK
SQLNNAKEIINGCFEFYHKCDNECNSVRHGTYDPKYSEESKLLNREKVGDVSLSSXGIXQI1ALAIKTQVASSLVLVLVL
GAISFXMSNGSLOCRICI
Figure 22

(SEQ ID NO.14)

>PR8HA-48G291G-403Y411Y422Y433Y.pro
MKANLVLSSALAAAAADDTICIGYHANIADTDTVDTVLKNIKTVTHSVERLYPQGNLLEDSDHNGKLCRLKGIAPLQLGKC
IAGWLLGFECQPLLFRSWSYIVETPNSENGICYPGFDIDYEELREQLSSVSSFERFEIFPKESSWPHSNHTINGVTAA
CSREKGSSFYRHNLLWTEKGSYPKLNSYVNHKGKEVLWGLHPPNISKEQQNIYQNEAXYSVTSHYVRRFTPEIAER
FKVRIQAGRMIIYYWTLKPGDTIIIFEANGLLIAENYAFALSARGFSGMITHSMAHHECENLYPQHNTKQTPGLAINSSL
FYQNIHFVTIGECPTVKVSAKLRTVGLANHTPSIQRGLFGIAAGFIEGGWTGMIDGWYGYHHQNEQGSYAADQKSTQN
AINGITNKVHTVIKMYIQPTAVGYEFNKLEKRMEYNKVKDDLFLYINTYNAELVLLENERTLDPHSNWKLYEYKV
SQLKNAKEIGNCFCFEYHKCDNECNESVRNGTYDFKYS3EESKLHREKVDGVALEGMMGIYO11AIDXVTAVSSLLVLLVSL
GAISPXMCNHSGL6CRICII
>PR8HA-48G291S-411Y422Y.pro
MKANLLVLLSALAAADADTCIGYHANNSTDTVDTVLEKNTVTTHSVEINLYFQGHLLEDSHNGKLCRLGIGALQLGKCN
IAGWLLGFECQPLLPRWSYSYVETPNSENGICYPGDFIDYEELREQLSSVSSFERFEIFPKESSWPHNHTNGVTAA
CRREGKSSFYRHLWLTEKEGSYPKLNSYVHNGKEVLLLWGIHHPNISKEQQQNYQNEHAYVSVVTSHYUNRRTFPEIAER
FKVRDQAGRMNYYWTLFPGTDIIIFEANGNLIPHYAFALSRGFGSGIITSNASHNESCENLYFQSNTKCQTPLLAINSSL
PYQNIHFVTIGECFKTVRSAKLRRMTGILTPISIQSERFGFIAAGFIEEElGWTGMDGWYGHHQNEQGSGYAADQKSTQN
AINGITNKVTIVIEKNIQPTAVGYEFNKLEKRMELYLNNKVDGDLDINTYNAELVLLENEERTLDFDSNKLYEYKV
SQLKINNAKEIGNCCEFEPHYKCDNECMESVRHGTYDFKYSSESKLHREKGDVYKLSMGIYQILAIYSTVASSLVLVLUSL
GAISPMCSHGSLOCRIGI
Figure 27 B

FR8HA-63G278S SIQRGLFGIAIGFIEEGWGTMDGWYGHQHEQGSGYADQKSTQAINGITNKTVNT 399
FR8HA-63G282S SIQRGLFGIAIGFIEEGWGTMDGWYGHQHEQGSGYADQKSTQAINGITNKTVNT 405
FR8HA-63G283G SIQRGLFGIAIGFIEEGWGTMDGWYGHQHEQGSGYADQKSTQAINGITNKTVNT 405
FR8HA-48G291G SIQRGLFGIAIGFIEEGWGTMDGWYGHQHEQGSGYADQKSTQAINGITNKTVNT 412
FR8HA-48G291S SIQRGLFGIAIGFIEEGWGTMDGWYGHQHEQGSGYADQKSTQAINGITNKTVNT 412
FR8HA-WT SIQRGLFGIAIGFIEEGWGTMDGWYGHQHEQGSGYADQKSTQAINGITNKTVNT 398

FR8HA-63G278S IKRNQFTAVGKEFKLKKRMENLNKVDGFLIDWYNAELLVLLEELERTLDHPSNVI 459
FR8HA-63G282S IKRNQFTAVGKEFKLKKRMENLNKVDGFLIDWYNAELLVLLEELERTLDHPSNVI 465
FR8HA-63G283G IKRNQFTAVGKEFKLKKRMENLNKVDGFLIDWYNAELLVLLEELERTLDHPSNVI 465
FR8HA-48G291G IKRNQFTAVGKEFKLKKRMENLNKVDGFLIDWYNAELLVLLEELERTLDHPSNVI 472
FR8HA-48G291S IKRNQFTAVGKEFKLKKRMENLNKVDGFLIDWYNAELLVLLEELERTLDHPSNVI 472
FR8HA-WT IKRNQFTAVGKEFKLKKRMENLNKVDGFLIDWYNAELLVLLEELERTLDHPSNVI 458

FR8HA-63G278S KYLYEKVSQLKINNAKEIIGNGCFEFYHKCDHECMESVRNGTYDFKYSEESKLJNREKVDG 519
FR8HA-63G282S KYLYEKVSQLKINNAKEIIGNGCFEFYHKCDHECMESVRNGTYDFKYSEESKLJNREKVDG 525
FR8HA-63G283G KYLYEKVSQLKINNAKEIIGNGCFEFYHKCDHECMESVRNGTYDFKYSEESKLJNREKVDG 525
FR8HA-48G291G KYLYEKVSQLKINNAKEIIGNGCFEFYHKCDHECMESVRNGTYDFKYSEESKLJNREKVDG 532
FR8HA-48G291S KYLYEKVSQLKINNAKEIIGNGCFEFYHKCDHECMESVRNGTYDFKYSEESKLJNREKVDG 532
FR8HA-WT KYLYEKVSQLKINNAKEIIGNGCFEFYHKCDHECMESVRNGTYDFKYSEESKLJNREKVDG 518

FR8HA-63G278S VKLESGIYQIALYTVASSVLLLLVGALSIFFWNNCSHSGLQCRICIQ 566 (SEQ ID NO. 26)
FR8HA-63G282S VKLESGIYQIALYTVASSVLLLLVGALSIFFWNNCSHSGLQCRICIQ 572 (SEQ ID NO. 27)
FR8HA-63G283G VKLESGIYQIALYTVASSVLLLLVGALSIFFWNNCSHSGLQCRICIQ 572 (SEQ ID NO. 28)
FR8HA-48G291G VKLESGIYQIALYTVASSVLLLLVGALSIFFWNNCSHSGLQCRICIQ 579 (SEQ ID NO. 29)
FR8HA-48G291S VKLESGIYQIALYTVASSVLLLLVGALSIFFWNNCSHSGLQCRICIQ 579 (SEQ ID NO. 30)
FR8HA-WT VKLESGIYQIALYTVASSVLLLLVGALSIFFWNNCSHSGLQCRICIQ 565 (SEQ ID NO. 1)
Figure 28

(SEQ ID NO.31)

>PR8HA-63G278S-403Y433Y.seq

ATGAAGGCAAAACCTACTGTCCTGTATAAGTGCACCTGTGCAGCTGACAGATCAACAAATATGTATAGGCCTAACATGCAGAA

CAATTCACCAGACACTGTGACACAGTAACCTCAGAAGAAATGACAGATGACACACTCTGTAACTCTGCTGCTGAGACAGGC

ACAACGGAAAAACTATGATGATAAAGAAGgagaaatctgtacttccagggaaAAATGTAACATCGCCGAGATGGCTTTGGGA

AACCCAGAATTGGCAACCCACTGCTTCCAGTGAGATCATTGCTCTCATTAGAAACCAAACTCTCGAGAATGGAATATG

TTATCCACAGATTTCATCGACTATGAGGAGCTGAGGGAAGCAATTGAGCTCAAGTTGACATCACTATCGAAAGATTCGAATAAT

TTCCCAAAAGAAGAAATTTGCTATGGTACCGCAACCACACACAAAAGCGGAATACGGCCAGCATGCTCCCAGTAGGGGAAAGCAGTTTT

TACAGAAAATTTGCTATGGTACCGAAGGAGGGGCTCATACCCAAAGCCTGAAAAATCTTTATGTTAACAACAAAAAGGGGA

AGAAGTCTCTTGATCTGTGGGTATTCATCCACCCGGCTTACAGTTAGGAACACAGAATATCCTATGAGAATGAATGCTT

ATGTTCTCTGATGTGACTCAAAATTAACAGGAAGTTTACCAGGAAACTACAGATAAGAGTCAAGC

GGGAGGATGAACATTCTACTCGAACCTTGGCTAAACCCCGGAGACACAAATAAATTTGAGGGAAATGGGAAATCAATAGCACC

AATGTATGCTTTCGCACTGTGAGTAAGGGGTTTGGgagaatctgtacttccagTCAAACGCATCAGTGAAGGTGAAC

CGAAGTGTGACACACCGCTGAGCTTAAACACAGCAGTCTCTACTTACGAAATATAACACAGTCTCAGAACATTAGGAAGATGC

CCCAAAATACCCGAGGATGTCGAACATGTTAGGGTGTTACAGAACTAGGAAACACTCCGTCCATTCAATTCCAGAGGTCTATT

TGAGGACCATATGGCGGTGGTTATGAGGGGAGTGGAGTTAGTGATAGTTGTTATCATCAGAAGTAAAC

AGGATCAGGCTATGCAAGGGATCAAAAAAGCACAACAAATGCCATTACGGGATTACAAACAGTTGAAACTGTTATC

GAGAAATGATCATTGCAATCCACTACAGCTGTGGTTAAGAGTTTACAAACAAATGAAAAAAGATGGAAAATTAAAATAAA

AGTTGAGTAGGTGATTCTCTGTCATACATTGGACATATAATAGCAGAATGTGATTGTACTGCTGAAAAATGGAAGCAGTTGTCT

TCATGACTCCTAAATGAGAAGATCTGATGAGAAGATGAAAAAAAGCATAATTTAGCAGAAAAGAAATCAGGAATGGA

TGTTTTGAGTTCTCAACCAAGATGGTGAACATTGAGTGAAATGGAAGCTATGATCTATCCAAATATTTC

AGAAGGCTCAGATGGCAACGGAGGAAAAGGTATAGGGAGTTGAAAATCTGGAATCAATGGGGAATTGCTATCGAGATTCTGCGACT

ACTCAACTGTCGCCAGTTCACGTGTGCTTTTGGCTCTCCTGGGGGCAATCAAGTTTCTGGAATGTGGTCTTCAATGGATCTTTG

CAGTGCAGAATATGCGATCTGGA
Figure 29

(SEQ ID NO.32)

>PR8HA-63G278S-41LY22Y.seq

ATGAAGGCAAAACCTACTGTCCTGTTAAGTGACATTGCACTGCGACTGCAAGATGCAAGACACCAATATGTATAGGCTACCATGC

CAATTCACGGAACACTGTTGACACAGTACCTGAGAGAAATGTGACAGTGAACACTCTGCTTAACCTGCTGCGAAGACAGCC

ACAACCGAAAAACTATGTAGATAAAAGAAGagaaatctgtacttccqagggaAAATGTAACATGCGCGATGTGCTTTGAGA

ACCCAGAATTGCGACCACACTGCTTTACGTGAACTATGTGCTCTATAGAAACACACAAACTCTGGAAGATGGAATATG

TTATTCCCAAGAGATTTTCTAGACTATGAGGAGCTAGGGAGCAATTTAGGCAGCTAGTCTCATTACGGAAAGATTCGAATAT

TTCCCAAGAAAGATTTTCTAGACTATGAGGAGCTAGGGAGCAATTTAGGCAGCTAGTCTCATTACGGAAAGATTCGAATAT

TACAGAAATTGGCTATGGCTAGGGAAGGGGCGAGTCAATCCCCATGAGGGAAGGCAGTTTT

AGAAGTCTCTTGTACTTGGGCTCTATCCACGTAGTTAGGAACACAGAATATGCTATAGATGACAGTGCT

GGGAGGATGAACTATTACTCTTAGCAGTCATATATAGCTAAAAACCGGAGACACAAATATATATGGCGGCAAATGGAATCTAAATAGACC

AATGTATGCTTTTTCACGACCTGAGGGGCTTTGGGagaaatctgtacttccqagggaAAATGTAACATGCGCGATGTGCTTTGAGA

CGAAGTGTCACACACGCTCAGGACATATATACCGAGTAGCCGCTACATAGAAGATGGAATATG

CCAAATACTCAGGAGTCGCAAAATAGGGAAGTTACAGACTAGAAACCTCGCATCAGCCATTAGGGAGTTT

TGGAGCCATTTGCGGCTTTTTTTAGAGGGGATGGAGTTAGTATGGGGATGGGTTATCATCATACGAGAATGGAAC

AGGGAATCAGGTATGCAAGCGATCAAAAAAAGCACAACAAATGCATTTACCAAGGGATTACAAAAACAGGACACACTGTTATC

GAGAAATGAACATTCATTCACAGCTGTTGGTAcGAATTCACAAATTTAGAAAAAGGATGGAATATCCTAAATAAAA

AGTGGATTGATTCTTTGCGACATATATAGCGAATTGTATTAGCTATCGTGGGAAATGGAAAGGCACCTGAGGATT

TCCATGACTCAATATGTTAGAAATGATGAAAGATAAAAGCCAAATTTAAAGAATATGCAAAAAATCGGAAATGGA

TGTTTTGAGTCTTTCACTACCAAGTGGGACAGAATGCAATGCAAGGAAGAATTGGGATTATAGATGATTACACTATCTGAGCAGATC

AGAAGAGTCACACACGCTCAGGAGTAAATGGGAATCAATGGGGATCTTCAATATGACAGTCTCGGCGATTT

ACTCAACTGTGCGCAGTTCTACTGTTGCTTTTGCTTCTCTCGCGCGGCAATCGTTCTGAGTTGTTCTAATGGAATCTTGG

CAGTGCAGAAATATGGCATCTGGA
>PR8HA-63G278S-403Y411Y222Y433Y.seq
ATGAGGCAACACTCTGCTCTTTAATAGCAGCAGCTGCAAGAGCATGAAATATGTAAGCTACATTGGAACCA
CAATTCGCAACACTCTGCTCTTTAATAGCAGCAGCTGCAAGAGCATGAAATATGTAAGCTACATTGGAACCA
ACAACGAAACTATCTGATGATAATGAAAGgagaacctgtctacacctgacatggaacca
AACCAGAATTGCGGACCCACTGCTCTCAGATGAAACTCTATGTAAGAATACACCATCAAGTGAATCTGAA
TATCCACGACTATCTAGGAGGCTGAGGGAACCATTGAGGCTACATTGCAATGGAACCA
TCCCCACAAAGAAGTACAGGTAAGCAGCAGCTGCAAGAGCATGAAATATGTAAGCTACATTGGAACCA
TACAGAAAATTTGCTATGGGCTACGGGAGAAGGGAAAGCAGTTTT
AGAAGTCTTGTATGTTGGGTTATCTACATCCCTGCTATACCTTTAATGGGATGGAACCA
ATGCTCTGTATGCTACCTAATATACACGGGAGTTATCCACGCTACATTGCAATGGAACCA
GGGAGGATGAAACTATTACTCAGGACCTTTGCAAAAACCGGGGAGACCAAAATATTTTTGAGGCAGGAAATTTAACTGCAAC
ATGTAATGCTTTCGCACTGATAGGGGCTTTTGGgagaacctgtctacacctgacatggaacca
CAGGATCTACAGCCTGCAAGATCAATGCAAAACTCTATGTAAGAATACACCATCAAGTGAATCTGAA
CACAATACCTGAGGATCGATAATGGAACCA
CTACGAAATTGCAAGGATAATGTAAGCTACATTGGAACCA
GAGAAATGAAATTCTAAATCAGGCTTACGTCTTGTACAGGCTGAGGAGTACATGTAAGAATACACCATCAAGTGAATCTGAA
AGTTGAGGATAATGGAACCA
TGCTAATGCTGATGAATGGAACCA
AGAAGTCTTGTATGTTGGGTTATCTACATCCCTGCTATACCTTTAATGGGATGGAACCA
ACTCAAAGCAGGATCAGGCTGAGGGAACCATTGAGGCTACATTCTAGGAGGCTGAGGGAACCATTGAGGCTACATTGCAATGGAACCA
CAGTGCAGAAATATGGCATCTGAA
>PR8HA-63G282S-403Y433S.seq
ATGAAGGCAAAACCTACTGGCTCTGTTAAGTGCGATCGTGCAGATCGAGACACAAATATGTATAGGCTACCATGCGAA
CAATTCAACCGACACTGTGGACACAGTGACTTCGAGAAGAATGTGACAGTACACACTCTGTTCATTACGTCGAGAACAGCC
ACAACGGAAAACACTATGTAGATAAAAGGagaagatactgtacttcagagggaAAATGAACATCGCGGAGATGCTCTTGGGA
AACCCAGAATTGGGACCACACTGCTTCCAGTGGATGCTAGTCCTACATTGTAGAAGAACACCAACTCTAGAAGATGGAAATATG
TTATCCAGACTTATTTGGACCTGATGAGGAGTGGAGCACTCTGAGCATTAGCTACCAGAAATATCGAAATAT
TTCCCAAGAAGATCAGAGTACGCCCCAACAACACAACAAACCGGATACGCAGTCGCTCCATGAGGGAGAAAAGCAGTTTTT
TACAGAAATTTGTCTAGTGCGTGACGGGAAGGGGCGTCATAACCAAAGCTGAAAAATTCTTATGTGAAACAAAAAGGGAA
AGAAGTCCTTGTAATGTGGGATTTACTCATCACCACGGCTCTAAACAGTAGGAACACAGAATATCTATCAGAATGGAAATGCTT
AGTCTCTGTAGTGAACCTAATATAACAGGAGATTATTACCCCGGAATAAGCAGAAAGACCAAATATAGAGACGCAGT
GGAGGATGAAACATTTACTGACCCTTGCTAAACCCGAGGAGACCAATATAATTTGGAGGAATGGCAATGAAATCTCAATAGCC
AATGTATGCTTTGCACTGAGTAGGCTTGGGCTCCGACATCGCgagaatctgtacttcagagcACCTCAACGCAT
CAATGCACTAGTGTAACACAGTACGCCTCAACACCgCtCgGAGCTATAAACAGCAGTCTCCCTTCACGAGAATATACACCCA
GTCACAATAGGAGATGGCCCAAATATGCAGTGCGGCAATTGAGGATGTTACAGGACTAGGGACACGTCCGTCCAT
TCAAATCCAGAGGTCTATTTGGGACCTGCGGCTTTATTGAAAGGGGATGGATGGAATGATAGATGATGATGGTATGGTT
ATCATGATCGAGAACGAGGACTCGATGTGACGCGATCAAAAGGACGACACAAAAATGCCATTAACGGAATCAAAAC
AAGGTGAACACTGTTATCGAGAAAATGtACATCTCAATTCACAGCTGTGGTAAAGAATTCAACAAATTGAAAAAGGAT
GGAAAAATAAAATAAAAGGTGTAGTGGATGTTCTGtACATTTGGCAATATAATGCAGAATTGTGAATCTCAGGGAAA
ATGAAAGGAGATCTTCGATTCTCCATGACTCAAAATGTGGAAGAATCTTGCAAGGAAAAAGGAAATGAAAAATGCC
AAAGAAATTCCGAAATATGGGATTTGTTGACTTACACCAAGTGTGAGCAATGGAATGTGAAAGATGAGATGACTTAA
TGATTATCCCAAAATATCCAGAGAGTCGAATGGTGAACAGGGAAAAGGTTAGATTGGTGAATATGGGAATATGCGAATCT
ATCGAGATTTCGGGCGATCTACTCAACTGCTGGCCGAGTTTCAGTGGTTTTTGGCTCCCTGGGGGCGCAATCAGTTTCCTGGGATG
TGGTCTATAGGATCTTTGGCAGAATATGCGCATCTGA
Figure 32

(SEQ ID NO.35)

>PR8HA-63G282S-411Y422Y.seq

ATGAAGGCAAAACCTACTGGCTCTGGTTAAGTGCGACCTGCGTGCAGATCGACGACAAATATGTATAGCTACCATGCGAAG

CAATTCAGACAGACTGTTGACACAGACTGCGAAAGAATAGTGAACAGTGAACACTCTGCTGTATACCTGCTGATGACAGAGCC

ACAACCGAAAACCTATGAGTTAAAGGAAGGgaatctgtaacttcagggagAAATGTAACATCGGCCGATGGCTCTTGAGGA

AACCAGATGGGACCCACTGCTGTCACTGAGATCAGTGTGGTCTACATTGCCGAGATGGAATGAAATG

TTATCCAGAGTATTCATCAGCTATGAGGAGCTGAGGAGCAACTTAGGAATCTGCTGATCATTCCGAAAGATTGCAAAATAT

TTCCCAAAAGAAAGATCCTAGGGCCCAACCACCAACCAACCAACCGGAGTAACGCGCATGCTCCATGAGGGGGAAGAAGAAGT

TACAGAAAATTTGCTATGCTGATGCGGAGAAGGGGGTGTCACTACCCCAAGCTGAAAAATTCTTATGTAACAAAAAGAGGA

AGAATCTCTTGTACTGAGGAGGTTATCATTACTCACCCCGCTAATACAGTAAAGGAACACAGAAATATCTATACGAAATGAAAATGCTT

ATGTCCTCTGTAGTGACTTAAATATACAGGAGATTTACTCCCGGAAAATAGCAGAAAGACCAAAGATAAGGATACACGT

GGGAGGATGAAACTTTACTGCAAGCCTTAAACCCCGGAGACACAAATATTATTGAGGGCAATAAGGAAATCTTAGAACACC

AATGTAATGCTTTCCGACTGAGTAGGAGCTTTGGGTCCCGGTCATCTCgagaatctgtaacttcagggagaCCACAAACGAT

CAATGCATAGTGTAAACAGACAGTCAAAACCACCgCTcGCGACTTTAAAACAGCAGTCCTCCCTATCCGAGAATATACACCCA

GTCACAAATAGGAAGTGGCCCAAAATACGTACGGAGGTGCGCCTAATGAGGGATGTTACAGGACTAAAGGACACTCCGGCTCAT

TCAATCCAGAGGCTCTATTGGAGCGCATGCGGCTATTATTTATGAAAGGGGAGTAGAAGTGAATGGAATTGATGATTGGTT

ATCATCAGAGAAATGACAGGAGGTCAACATGCAAGCGATTTAAAAGGCAAAGACAAATGCAATACAGGAGGATCAAACAC

AAGGTGAAACACTGGTTATCGAGAAAAATGAAACATTTCAACTCACGCTGTGTTATcGAATTCACAAACTATAGAAAAAGGAT

GGAaAAtcTTAAAATAAAAAGGTGAGTGTAGGATTTCTGGACATTATATAATGCAGAATTGTGATCTACTGGAAG

ATGAAAAGGACTCTGGATTTCCATGACTCAAAATTGTAAGAATCTGTAATGGAAGATAAAAAGGCAATTTAAAAGGAATATGCC

AAAGAAATCCGCAAAATGCGTTTGGAGTCTTACACCCAACTGGTACAAATGAGAAGGAGGGAAAGTGAGATGGGATGAGATG

TGGATTATCCAAAATATCGGAGAAGTCAAATGGAACAGGGAAAGATGATAGTGAGTGAATTTGGGAAATCTCAAGTGGGATCT

ATCAGATTTGTCGGGAGTACTCAACTGCGCCAGTTCTACTGGTGTATTGGTCTTCCCTGGGGCGAACACTGATTTCTGGAATG

TGTTCTAAATGGAATCTTTTCAGTGCAAGAAATATGCACTCTGA
Figure 33

(SEQ ID NO.36)

>PR8HA-63G282S-403Y11Y22Y33Y.seq

ATGAAGGCAAACTACTTGCTGTTAAGTGCACTTGCACTGCGATGCAAGACACAAATATGTTATAGGCTACCATGCGAA
CAATCCACGACACTGGTGGCACACACTTGCGAGGAAATGAGACGTGACACTGCTGTCAGACACACACTCTGTGTTAACCCTGCTGAGACAGCC
ACAAAGGAAAACATGTAGATAAAGAGGagaatctgtacttccagggaaAAATGTAAACATCGCGGATAGGGCTTGGGGA
AACCCAGATGCGACCCACCTGCTGCATTACGATGCTGCTACATTTGATAGAAAACCACAACTCTGAGAATGGGAAATATG
TTATCCCGAGATTTTCTACGGACTATGAGGAGCTAGGAAGGCTAAGCTCGTACATGCTACATCATGGCAGAGTGTTTTT
TTCCCAAAGAAAGATCGTCGCAACCCCAACCACACACACACACGGAGAAACCGCAGCATGCTCCCATGAGGGGAAACGAGTCTT
TACAGAAATTTGCATGGCTGACGGGAAAGGGGCTACATACCCAAAGCTGAAAAATCTTTATGAGACAAAAAGGGAA
AGAAGTCTGGTGCTGGGTTATTCATCACCACGCCCTAATACGTAAGGAAACACAGAAATATCTATGCAAGAATGGATAATTTT
ATGTCCTCTGTAGTGACACAAAAATTATAACAGGAAGATTTTACCCCCGAAAAATAGCAAGAAGACAAAAATAGAGATCAAGC
GGAGGATGAACTATTACCTGACCTTGAACACCAATTTATTGGACGGCAAATGGGAAATCTAATAGCACCC
AATGATATGCCTTTCGACATCGAGTGAGGCTTTGGGTCGCGCATCGAAGaactgtaacttcagacgACCTCAACGCAAT
CAATGCGATAGTGGCTAACACAGTGCTCAGCAGCCTATACCTGCTCTCCCTTACCAAGATATACACCCA
GTCACAATAGGAGCTGCCAAAATACCTGACGAGTGCCCAATTTGAGGATTTACAGGACTAAAGAGCACTCCGTCCAT
TCAATCCAGAGCTCTATTTGGAGCCATGTCGCGGTTTTATTTGAAAGGGGATGGACTGGAATGAGATAGTGGATGTGATGTGT
ATCATCATCGAAATGAAAGGATCTCAGCGATCAAAAGAAGCACAACAAATGCCTATTACCGGATTACCAAA
AAGGTGAACACTGGTTATCGAGAAAATGctGATTCACATTCCACTACACAGCTGGTGTAcGAATTCACAAATAATGAAAGGAAT
GGAtAcTAAAATAAAAAGTTTGAATGAGGATTCTCTGTAcATTTCTGCATATAATGCAAGATTTGATTGCATCTGCGAAAA
ATGAAAGGACTCTGGATTTCCATTACGACTAATAATGTAAGAAAGTAAACAAGCATTAAAGAGAATATGCC
AAAAAATCCGCAAATGAGATTTTGCTGCTACCAACAGTGACAAATGGAAGTGGAAAGAAGTATAGAAAGTTGGGACTTA
TGATTATCCTTTGCGACCTACTCAACTGTCGTCAGTTCATGACTGGTCTTTTGCTTCCCTGGGGACAAATCAAGTCTGTGGATG
TGTTCTAATGGATCTTTGCGATGCAAGATATGCATCTGA
>PR8HA-63G283G-403Y433Y.seq
ATGAAAGCTAACCTGCTGCTGACGCCCTGCTGATCCAGTACCCATCTGATCGCTATCCACAGCAAGCC
CAAGCACCAGACACCTGACGGATTACCTGCTGGAAAGGAACGTTACGCACTGCTGGAAGTACG
ACAGGGCGAAGCTTTACAGCTGACTAGGAAACTGCGGGACAGCTGACCGTGCAGTCCAGCTAG
TTCCAAAGAGAGACAGCTGGGCCCACCAACGACCAACAGCAGCTGAGCACCCGGCCCTGAGC
ACCCGGGAGAAGAGGCAAGCAGTACGTTAAGAAGCAAGCCAA
AGAGGTGTGCTGGTGGGTCGTTGCACCACCCACCCACCCACCCACCCACTTAAAAAGGAGCACAG
ACAGCAAGAGGTGCCCCAAATACGTGACAGGGTTTACAGGTAGTTTACAAGGAGAAGACACTGCG
TCCATTCCAGAGGTCTATTTGGCAGCCATTTGCGGTTTTTTATATTGAGGGGATGAGATAGATGAGATTGATTGTT
ATCATTCACAGATTGAAACAGGGAATCAAGGCTATATGCAAGCGATCCCCAATAAGGACAAGAATGCGATTATACAAC
AAGGTTAAGCCACCTTTATCGAGAAATGGTCACATCCTAATTCCAGCTGCTGGTAAAGCATTAACAAATAGGAAAGGAT
GGAAAAAATAAAAAAGGGTATGAGGATTTCTTGACATAATTGACAAATTGTTAGTCTACTCTGGAAA
ATGAAAGGACCTTTGTATTTTCTAGTTAAATTGAGAAGATACCTTTGAAGAATGAAAAGCCAAATCAAAGAAATTAGCC
AAGAAATACTCGGAAATGGATTTTTGATCTTACACCAAGATGTAGCAAGAAATGAAAATTGGACTTATA
TGATTATCCAAAATTTCGAGAGATGCAAAGGACAGGGGAAAGAGTGAATGAGGTGAAGATTGAAATATCATGCTGGAT
ACTAGTTTCTGGCCATACAAAACTGATGCACCAGTTTACTGTTGCCAGTCTGCTGTCCTTGGGATG
TTTCTAAATGGAATTTTGCGAAGATAATGAGATCTCGTA

Figure 34
(SEQ ID NO.37)
(SEQ ID NO.39)

>PR8HA-63G283G-403Y411Y422Y433Y.seq
ATGAAGGCTAACCTGCTGCTGCTGAGCGCCCTGGCTGACGTGATGCCGATACCCTCTGACATCGCTACCGCAACGCAAA
CAACAGCACCAGACCTGAGTACCCGTGCTGGAAGAAACGTCGTTGACACGCGACCGGCCTTGATGCTAGATGCGAC
ACAACCAGAAGCTGCTGCACCCAAGGGGAGAATCTCAAGGAGTCTACATCGCTGCACATGACATCGCTACCGCAACGC
CTACCCCGGCAGCTTCACTACAGGAAACTGCGAGCAGCTGACGACGCTGCTTCAGTTGAGATTCGAGATCT
TCCCCAAAGAGAGACACCTGCGCCCAACACCAACACCAACAGGCGTGACAGCGCCCTGTAGCCACAGGGAAGGCAAGAGCAGGCTC
TACAGAACACTGCTGCTGCTGAGCCGAGAAAGGAGGAGCAGCTACCCAAAGCTGAAGAAGACAGCTACGTGACACAAGAAGGCAAA
AGAGTGCTGCTGCTGCTGAGCTGGCCATCCACCAACCCCAACTCTTAAAGAGCAGCAAAACTTACCAGAAAGATCGAAACGCC
ACGTGTCCTGCTGAGACCAGAAGCTCCGGCTTCCACCCCGAGATGCGGAGAGCCTAAAGTGGGGAGATCGGCGCC
G6CAGAATGAACTACTACCTGAGCGCCCTGGCTGAGCCGAGCAGCCATACATCTTTGAGGCAAGGCAACCTGATCGGCC
TATGTACCGCTCTGCCCTGGAGCAGGTCGCTCGCAGCCATACACGgaacaacctgtactctcaaggtaccAAGCAGC
GCATGCAAGTGGCAAACCAAGTGCAGCCGACCCCGCTCCGAGCTATATAAAGCAGCAGTCTCCCTTACCCAGAATATACACCCA
GTCACAAATAGGAGATGCCCCAAATACCTGACAGGATGCGGCAAAATTGAGATGGTTACAGGACTAAAGGAAACTCCGCTCAT
TCAATCCAGAGTCTATTTGAGCCCATTCGCGGTTTTATTGAGGGGAGTGACTGGGATATGATAGTGATGATGATTT
ATCATTACAGAAATGAAAGAGGATCAAGGCTATGCAACGCGATCAAIAAAGAGCACAACAAATGCAATTAAAGGAT
AAGGTGAACACTGTATCGAGAAATGTCATCTCAATTCAACAGCTGTTGGTcGAATTCACAAACAAATAGAAAGAGAT
GGATAcTAAATAAAAAATGTTGATGATTGATTTTCTGTACATTTTGCAATATAATGCGAAATTGTATGGTCTACTGAGA
ATGAAAGGACTCCTGGATTCCCATGACTCTAATAATGGAAGAATCTGTGTAGAAAGTATGAAAAAGCCAAATATAGAATTGCC
AAAAGAATTCGAAATATTTTTGTGACTCCACCAAGTGTGCAAACTGAGAATGCGGTAAAGTGAAGAATTTGGAATCTT
TGGATATCCCATAATTCCGAGAGAGTCAAAATGTGAAACAGGGAAAGAGGTGAGATGGATGAAATTTGGGAAATCCAT
ATCGAGATTTCTGCGCAGTACTCACTACGTGCCAGTTTACGTGGCTTTTGCTTTGCTTCCCTGGGGAGAAATCGAGT
TGGTCTAATGGAGTCTTTGCGAGAATATGCGCATGCTGA
Figure 37

(SEQ ID NO.40)

>PR8HA-63G286S-403Y433Y.seq

ATGAGGCCAACCTACTTGTGGTATAATGACACTGCTGAGACTGAGACCACACAATTATGAGCCATCGACATCG
CAATTCAACCGGACACTGTGTTGACACAGTGACACTGCAAGAAATGAGACAGACACTGCTGCTGAGACGAC
ACAACGGAAAACATATGGATATAAAGAGagatacttgatacttcagggagaaATGTAACACTCGCGGATGTG
ACGCTCTTGGGA
AACCCAGAATGGGACCCACTGCGCTCTTAGAGATATGCTGCTACTACATTGATAAAGACCCAAAACACTCGA
AGAAATGGAATTTCTCATCGAATAGAGGAGCGTGAAGGGAACACTGCTATGCTATCTACATGGACAAAGAT
TTCCAAAGAAGAGAAGGCTATGGCGGCAACCAACACACACAAAGCGGATAACGCAGCTGCTCCATGAGGGA
GGGAAGAGCACTTGTGTTGGGTCTTTACATCCACGGCCCTAAGAGACAGAATGAAATGAGACGCTACGGAG
AGAAGTTCTTAGCTATGCGGTCTACCTATCGAATAGGGAACACAGAATTTATATATGCAATGAAAGATGCT
ATATGCTCTGGTTAATACATGCTCATTCTCCGGTAAGGAAAGACACGCAAATGACAGACTACG
GGGAGGCGATGAAACTATATGCTGTACACATTTTTTTCTATTGGGTACCATTTTTTGGCAATGAGGACACAC
AAAGAGCAGATGATGCTGCTGAGAATACGCTACATGAGAATGAGTAGGTGCTGCTGAGAATGAGAGGAGA
CAATGAGCAGATGATGCTGCTGAGAATACGCTACATGAGAATGAGAGGAGA
GTCACAATTACGGACATGCGCCAAATACGTAGAGGTGACGATCAGTTCAAGGATGTTACAGGACTAAGGACAC
ACTCGAGTCA
TCATCCAGAGCTAGCTGGAGGACCTGGGCTGTTTTATTGAAGGGGATGGAGTGATAGATGATGGTAGTGTT
ATCATACGAAATGAAAAGGAGTAGCTATGACGCGACTCAGCTATCAGCTATCAGCTATCAGCTATCAGCTAT
AAGGGAACTCTGTTATCGAGAAAAATGAACATTTACATCGACATCTGCTGAGGTAAAGAACTTACAAATTGA
GAAAATTAAAAAAGGTGATGAGGGGCTTCGGATATCTGAATATGCAGATGTTGATCTACTGTGAG
AGGAAAGGAGCTGCTGTAGGATTTCCCATGACTCAGAAAATGTAAGAAGAATGACTGATGAGAAGAGATTAAA
GCAAATTAAGAAATAGC
AAAAGAAATCGGAAAAATGAGGTTTGGATCCTACACACAGCTGAGAATGAAATGCAAGTGAAGGAAATGGGAG
ACTTA
TGATATCCCAAATATTACCAAGAGACTCAGACTGCGGAGTCCACTGGGTTTGGGTCTTCCCTGGGCGGAAATCAG
TGGGATGGTTCTTGGCAATGAGGACACAC
GTTCTACTGCTGCGAAATATGCGACATCG
Figure 38

(SEQ ID NO.41)

>FR88A-63G286S-411Y422Y.seq

ATGAAGGCAAAACCTACTGGCTCTGGTTAAGTGACCTTGCGAGCTGACAGCAGACACAAATATGTATAGCTACCAGTACGAA

CAATTCAACGCACACTGTTGACACAGCTACTGAGAAGAATGATGACAGACACACTCTGTTAACCTGCTGAGACACGCC

ACAACGGAAAACATATGAGATTAAAGAAGagaactctgtactttccaggggaaATGTAACATCGCGGATGCTTGGGGA

AACCCAGAATTGGCGACCACACTGCTCTCCAGTGAGATCATGTCCTACATTGTAAGAACCAGAATCTGGAAGATGGAATATG

TTATCCGAGATTTCTCATCGACTATGAGGAGGCTGAGGAGAATGAGCTGACTATGATCATTGGAAGATTTGGAAATATAT

TTCCCAAAAGGAAGAGATCTGTCGCGCCCAACCACACACACAAACCGGATACACCGGCACAGTCCTCCATGGAGGGGAAGACAGTTTT

TACAGAAAAATTGCTATGGCTGACGGGAAAAGGGGCTCATACCACAAAGCTGAAATATCTTATGTGCAACAAAGAGGGAA

AGAAGTCCTCTGTACTGTGGGGTTATCTCATTACCAGGAAACACAGAATATCTATCCAGAATGAAAATGGCTT

ATGTCTCTGTAATGCAATTCACATACAGGATTTTCTCCCGGAAAATAACGAAAAGACCCAAAATGAGAGATACAGCT

GGGAGGATGAACTATATGCTCGAGTCTAAACCAGGAGACACAAATATAATTGAGGGCAATGGGAAATCTAATAGCACC

AATGTATGCTTTGCGACTGAGTATGGGTTTTGGGTGCGGCTATCATCATTCAACACCGCagagaactctgtactttccagagcT

CAATGCAACACAGTGAAGCTGGAACACACACACACAGCTGCTCATGTAATGAGCACTCCCTTACCCAAAGAATACCCCTAC

GTCACAACTGAGGAGTGCCCCAATATCAGTACAGGGATGCTGCAAAATTGAGGATGTATACAGGAACTAAGAAGACACTCGGTCAAT

TCATGCCAGAGTCATATGAGGAGCCTGCGTTTTTTATTGAAGGGGAGTAGATGATGAGATTGATGATGTT

ATCATCATACGAGATTGAAACAGGAGACGGCTATGCAAGCGGATCAAAAAAAGCAACACAAATGGCCATTAAAGGATTCAAAAC

AAGGTGAAACACTGTTATCGAGAAATGAAACATTCAATTCACAGCTGTTGGAATcGAATTCACAAATATAGAAAAAGGAT

GGAAaAcTTAAATAAAAAAGGTGATGATTGGAATTTCTGGACATATAATATGCAAGATATTGTGATCTACTGGGAA

ATGAAAAGGCATCTGGGATTTTCCATGACTCAATAAGTGAAGAATCTGTAAGAAGATGAAAAAGGACATTAAGAATACGCC

AAGGAAAATCCCGGAAATGGATTTTGGATCTACACAAAGTGTGACAAATGAAATGCAATGGGAAAGTGTAAAGAATGGGACTTTA

TGTTATCCAAATATCCAGAAGACTCAATTGGAACAGGGAAAAGGATGAGATGGGATGAATGGAATATCAGATGCTGTGGGATAG

TGTTCTATGGAATTTTGCGACAGAATATGCGATCTCGA
Figure 40

(SEQ ID NO.43)
Figure 41

(SEQ ID NO.44)

>PRSHA-48G291G-411Y422Y.seq
ATGAAGGCTAACTGCTGTGGCTGGCTGATGGGCTGCCATGTAACCCTCTGGCTACCCACGCAA
CAACAGCAGACACGCAGTGGATTACCCGGCTGGGAAAGAAAGACCTGAGTGACCCCATCAGA
GCAATCTGCTGGAGGATAGACCCCCAACGGCAAGCTGCGGCTGACGGAATTGCCTGCCC
ATCCGCGCTGCTGGCTGGGCCACCCCGAGTGATCTCTGCTGCTGCGAGCTGTCTACATGC
CAAGCTGACAGCCGTAAGCGATCTCTACCCCCGCACTCTGACTACAGGAAACGCGACGCT
GCTTGAGAAGAAGGAAAGAGCTGCTGGCTGGGCCATCCACCCCCCACCACACTCTGAAAG
CTAGTGAACAAAGAAGGAAAGAGCTGCTGGCTGGGCCATCCACCCCCCACCACACTCTGAA
CTAAGCGGACTCCTCGGTCCATCTCAAGTACGTACGTTTGGAAGCGGAGTTTGCTCCACCC
TGATAGATGGATGTATGGTTATCATCACATCAGAAATGACGAGGTATCGCTAGCGATCAAA
GCCATTAACGGGATTACAAAACAGGTGAAACCTGTATTGAGAAATGACCTACATCAGCTG
CAACAAATTTTAGAAAAAGGTGGAAACATTTAAAATAGTTGATGTGATTTCTGGAATTTTG
AACATATAATGTTAGTTTCTACGACGAAATGAGGACCTCTGAGATTTGCTGCAATATGG
AGCATTGAAGTAAATATCGAAAGAAGAAATCGGAAAATGTGATTTGTGATCTACACAAATG
AGTGTAAGGAAATGGGACTTTATGATTTATCACCAATGGAACATAGCTGGGTACATAGGA
AGTGTAAGGAAATGGGACTTTATGATTTATCACCAATGGAACATAGCTGGGTACATAGGA
AGTGTAAGGAAATGGGACTTTATGATTTATCACCAATGGAACATAGCTGGGTACATAGGA
AGTGTAAGGAAATGGGACTTTATGATTTATCACCAATGGAACATAGCTGGGTACATAGGA
>PR8HA-48G291S-403Y333Y.seq
ATGAAGGCTAACCTGCTGGCTGCTGAGGCCCTTGCTGCGTGATGCGGATACCATCTGCTAGCCCGTACCCATTCGCTAGCCACGCAA
CAACAGCACCAGCACCAGTTGATACCGCTGGAAAGAACAGTCGACCAGCACCACAGCTGGGAAATCTATTTGCAAG
GCAATCTGCTGGAGATAGCAACACAGCAGCTGCGCTGAGAGGGAACTGGCGCTGAGCTGGGCAAGCTGGAAT
ATCCCGCTGGCTGGCTGGGACCCCGTATCTTGCTGTGCTGGAGCTTGCTGCTGGGACCTCACATCGTGAAACC
CAACAGCAGGAGCTAGCTGATCTCCCGGAGCTCCATGACTAGAGAGCTGCGAGCGCTGAGACGCGCTGAG
GCTTCGaaAGAATTGGCATCTTCCCAAAAAGAGAAGAGCAGCTGCGCCAACCAACCAACCAACCAATCACGCGCTGAGC
CACGAGGGCAAGGCGAAGCTTCTAAGGAACTACCCGACGAGGAGGACAGCTGACACGGCCACCTGCTTCC
CTACGGAGACAGCCTAGCTACGTTCCGTTGAGACAGCAACACTACAAACGGGCGTTTCAACCCCCGAGATCGCCGAGAGG
CCTAACAGTCCGAGATCGCCGAGCAATGAACTACATACGCTGAGCCCGCAGCCCGCAGCCCAACGACCACTCAGTCGCCGCC
CAACGGCAACCTGATCCCGCTTACGTCAGCTGGAGCTGGACAGCGACGCTGAGCAGCCACGAGCTGAG
GCTACGACGAGCTGAGCGAGCTGACTCCCAAAAGCAGCAGCCGAGGACGGCAGCTCCGAGCTCAGTAAGAGACGACAGTCTC
CTTACGAGAATATACACCGACGATCATACGAGAGAAGTGGCGCCAAAATACGTCAGAGTGCCCACAAAATGGAGATGTTACAGG
ACTAACGGCAAACACTCCGGTCATACCCATACGAGGAGCTCAGTTGGAGCCATGCGGCTTTTATCAGAGGGGATGAGCTGGAA
TGATAGTGGAAGTGATATGTATCATCATCAGAAATGACGAGGATCGATCGGCTATGAGGGAACAAATGAAGACAGCTGAGCTGGAA
GCCATTACGGAGTTACAACAAACAGGTCACGTGTTATCGAGAAATGCAATTACATACGACGCTGAGCTGGTAAAGAATT
CAACAAAAATTAGAAAAAAGAGCTTAATAATTAAAAAGGTTGAGTTGAGATTCTGGTCACATTGGCCAAATATATATGGTGAA
AATTGGTAACTACGTTCCGAGATTTCTGAGTACCTCAAAATGTGAAAGAATCTGTAATGAGAAAATA
AGCGAATTAAAGAATATGCCAAGAAAGATCTGGAATTTTGAGTTGCTACACAAAGATGTGCAATGATGATGCTAG
AGGTAAAGAATAAGATGGCGCTTTAGTTATATCCAAATATTCGAGAGTGAAACGTTAGCTGAGTGAGAATGCTAG
AATTGGAAAATGTTGGGAGTATACCTGAGGTTCAGTTCCACGAGGTTCGTCGCGGGCTGAGTTCTTCGCTCCCTG
GGGCAATCATTTCTGGATGTTCTTCTAATGGGATCTTTTGGCATCGAGAATATGCAACCTTA

Figure 43
(SEQ ID NO.46)
Figure 44

(SEQ ID NO.47)

>PRSA-48G291S-411Y422Y.seq
ATGAAGGCCTAACCTGGTGGCTGCTGGACGCCCTTGCTGGACTGGACACCTATTCTTCTGCGGCGAGCTACGCTATCGGCTACACCACGCAAA
CAACAGAGGACCCGCCTGAGATGCCTGCTGGGAAGAGAAACCTGAACGCTGGACCCACAGCAGCTGAAACCTTCTATTTCAAG
GCAATCTGGTGGAAAGATAGACACCCCAAGGGCAAGCCTGCTGGCTGAAAGGCTGAGAAGCCTGCTGGCTGAGCTTCTCCACTGCTGGAAACCC
CAACAGCGAAGGCGACATCTGTACCTGCCGCCCTGACGCTACTAGAGGAACTGCGGAGCAGCTGAGAACGCTGGTCC
GCTTCgaaAGATTGCGAGATCTTCACCAAGAGACGCACTGCTGGCCTACCAACCGCAACGAGCAGCAGCTGAGAACGCTGGTCC
CAGGAGGGCAAGAACGAGCTTCTACCGGAAACCTTCTGGCTGAGCTGGCAGAAAGGCGAGCAGCTGCTGGTCC
CTACGTGAACCAAGAAAGGCAAGAGGAGGCGAGTGGTCTCAACCACCCCTCAACTCTAAAAGAGAGCGAGACAA
CTTACACCAGAGAAGAACCGCTGGCTGGGCTGAGCCACGAACACTACACCCGGGCGATCCACCGGAGAG
CTTAAAGTGCGGGAACAGGCGGAGAATGAACTACTACTTGGACCCCTGCTGAAGCCCCGGGCAACATGATCTACTTGAGG
CAACGGCAGCCTGATGCCCTCTGCTGTAGAGGCAGCTGCTGGCTGAGCTTCTCCACTGCTGGTCC
GCATGCAAGCTGGCGAGAACCTGTACTTCCAAAAGCAACACCAAAGTGCCAGACCCCGCTGGAGCTATAAACAGCAGTCTC
CCCTACAGAGAATACACACACCAGTCACATAGAGGAGTGGCCTAAAAATACGTCAAGGATGCTGCTACAGG
ACTAAGGAAACACTCCGTCCATTCATCACAAGGCTATTTGAGCCATTTGCCGTGGTTTTTATGAAAGGGAATGAGCTGGTCC
TGATAAGATGGATTGTTATCCTACTACAAAGAAATGGAACACGGTACAGCTGCTGAAAGGCTGAGAAGCCTGCTGGTCC
GCCATTAACGGGATTACACACAAAGGTGAACACTGTATATCGAAATAGACATTCAATCCACAGCCTGGGCTLaGAAT
CAACAAAATTTAGAAAAAGGTGAAAATACCTAATAAAAAAGTTGATGATGGATTTCTGGGACTTTTGGGCAATATTTAGCAG
AATTTGATGTTACTCAGGAAAATGGAAGACCTGCTGGATTTCCATGACTCAAAATGTGAAAGAATCGTGGATGAGAAATGAAA
AGCCATTTAAAGAAATATGCAGAAAGAAATCGGAAAGGTTGATGTTTGTAGTCTACCAGAACAGTGTGCAACTGTGACATGGA
AGGTTAGAAATGGGACTTATGATTACTCCAAATATTCGAGAAAGTCAAGGAGGAAAAGTGAATGGAGATGGA
AATTTGAAATCAATGCGGAGATCTACGATATCCTTGCCGAGCTACTCAACTCTGGGCTGCTGGCTGTTTTGTCCTCCCTTG
GGGCGCAATTCTTTCTGGATGTGTCTTCTAAGTAGGATCTTTTGCGACTATGCAATCTGAA
Figure 45

(SEQ ID NO.48)

>PR8HA-48G291S-403Y411Y422Y433Y.seq

ATGAAGGCTAACCTGCTGTGCTGTGCTGAGCGGCTTGGCTGAGGATGCCAATCAGCATCAGCTGCAACACCAGCAAA
CAACGCAACGGACACGGTATACCCTGCTGAGAAAGACAGTAGGCACCGGTTGAGGGGAAACCTGGATTTTTCAAG
GCAATCTGCTGAGAGATAGACCAACCCGACGCTGTCTAGGCAAGGGAAATGCGGCTGAGCAGCTGGAAGTCTAA
ATCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT
CAACAGCGGAAACGGCGACTCTCTACCCCGGAGAATCTACGAGAAGAGCTCGAAGACGCGACGCTGAGACCGCCG
GCTTCCggaaAGATTCAAGATCTTCCCAAAAGAGAGAGACCTGCGCCCAACCAACCAACACACACACACACAC
CACGGCCAGAGACGACTCTAGACCCGAAATCTGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG
CTACGTGAAAGAAAGAGGACAAAGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG
CCTAAAGGTGCGGAGATCAGGCGGCAAGATGAAACTAATCTCTGGACCTCTGCTGAGGCCGCGACCAATCAGCTCTCTG
CAACGGAACCTGATCCGGCCTAACGTTACGAGGACGCTGACCCGCTGACAGCAGAGCTGACGCGACGCTGACG
GCATGCAGGAGTCAGAAGACCTGATCTCTCAAAAGGAAACCAAGATCCGGTACCCGCTGACGCGACGCTGACG
CCTACAGTATACATACCCAGTCGCAATAGGAGATGCCTCCCCAAAATACGTTCAGGAGTTGCTGACAGGACAGGG
ACATGATTAGGAGATGTTATCCTCATCATCAAGTGAAGGAGACGCTAGCTAGGCTATGCAAAAAGCACAACAAAAT
GCCATTACGGGATTACAAAACACAGGGTAAACCTGTTATCGGAAGAATGACATTTCAATTTACGCTGTTGCTG
CAACAAATTGAAAGAAAAGGATGGAATCCTTTAAATGCCGTATGAGATTCTAGTCTCTGGATTTCTGACATTTTG
AATTGGTTAGTCTACTCTGGAAATGAGGACATCTGGAGTTTCCTACGAATCTCTGAAGAATCTGTGATAGGAAAGT
AAGCGATTAAAGATAATGCCAAAGAAATGCGGAAATTTGAGATTTTGGATGCTCATCACAAGATGTCAGAACATGG
AAGGTGTAAGAAATGGGACTTATGTTATACCTCAATTTCCGAAATATATAGGAAGATGCGAAGTACAGGAAAGAAG
AATTGGAAATCATAGGGGATCTACTGAGATCTTGCCGAGATCTACTCAACTGGTCGCACTTCTGCTGCTGCTGCTG
GGGGCAATCAGTTTCCTGGAGTGTTCTTAAATGGAATCTTGGCAGTACGAAATATGCAATCTGA
Figure 46 A

FR8HA-291G.seq
ATGAAGGCACAACCTACTGTCTGTTAAGTGAGACTTGACATCTGAGATGGCACACATATA
60
FR8HA-291S.seq
ATGAAGGCACAACCTACTGTCTGTTAAGTGAGACTTGACATCTGAGATGGCACACATATA
60
FR8HA-63G.seq
ATGAAGGCACAACCTACTGTCTGTTAAGTGAGACTTGACATCTGAGATGGCACACATATA
60
FR8HA-WT.seq
ATGAAGGCACAACCTACTGTCTGTTAAGTGAGACTTGACATCTGAGATGGCACACATATA
60
FR8HA-48G.seq
ATGAAGGCACAACCTACTGTCTGTTAAGTGAGACTTGACATCTGAGATGGCACACATATA
60
FR8HA-278S.seq
ATGAAGGCACAACCTACTGTCTGTTAAGTGAGACTTGACATCTGAGATGGCACACATATA
60
FR8HA-282S.seq
ATGAAGGCACAACCTACTGTCTGTTAAGTGAGACTTGACATCTGAGATGGCACACATATA
60
FR8HA-283G.seq
ATGAAGGCACAACCTACTGTCTGTTAAGTGAGACTTGACATCTGAGATGGCACACATATA
60
FR8HA-286S.seq
ATGAAGGCACAACCTACTGTCTGTTAAGTGAGACTTGACATCTGAGATGGCACACATATA
60

FR8HA-291G.seq
TGTAATAGCTACATGCAACTCAGACAAATCGAAGACAACACTGGTACGAGAGGAGTAG
120
FR8HA-291S.seq
TGTAATAGCTACATGCAACTCAGACAAATCGAAGACAACACTGGTACGAGAGGAGTAG
120
FR8HA-63G.seq
TGTAATAGCTACATGCAACTCAGACAAATCGAAGACAACACTGGTACGAGAGGAGTAG
120
FR8HA-WT.seq
TGTAATAGCTACATGCAACTCAGACAAATCGAAGACAACACTGGTACGAGAGGAGTAG
120
FR8HA-48G.seq
TGTAATAGCTACATGCAACTCAGACAAATCGAAGACAACACTGGTACGAGAGGAGTAG
120
FR8HA-278S.seq
TGTAATAGCTACATGCAACTCAGACAAATCGAAGACAACACTGGTACGAGAGGAGTAG
120
FR8HA-282S.seq
TGTAATAGCTACATGCAACTCAGACAAATCGAAGACAACACTGGTACGAGAGGAGTAG
120
FR8HA-283G.seq
TGTAATAGCTACATGCAACTCAGACAAATCGAAGACAACACTGGTACGAGAGGAGTAG
120
FR8HA-286S.seq
TGTAATAGCTACATGCAACTCAGACAAATCGAAGACAACACTGGTACGAGAGGAGTAG
120

FR8HA-291G.seq
GTGACAGTGACGACACTCCTGTT--------------------------AACCTGCTGACGAGACG
159
FR8HA-291S.seq
GTGACAGTGACGACACTCCTGTT--------------------------AACCTGCTGACGAGACG
159
FR8HA-63G.seq
GTGACAGTGACGACACTCCTGTT--------------------------AACCTGCTGACGAGACG
159
FR8HA-WT.seq
GTGACAGTGACGACACTCCTGTT--------------------------AACCTGCTGACGAGACG
159
FR8HA-48G.seq
GTGACAGTGACGACACTCCTGTT--------------------------AACCTGCTGACGAGACG
159
FR8HA-278S.seq
GTGACAGTGACGACACTCCTGTT--------------------------AACCTGCTGACGAGACG
159
FR8HA-282S.seq
GTGACAGTGACGACACTCCTGTT--------------------------AACCTGCTGACGAGACG
159
FR8HA-283G.seq
GTGACAGTGACGACACTCCTGTT--------------------------AACCTGCTGACGAGACG
159
FR8HA-286S.seq
GTGACAGTGACGACACTCCTGTT--------------------------AACCTGCTGACGAGACG
159

FR8HA-291G.seq
CACACAGGAAACTACTAGTATGGATTAAAAAGGA-ATAAGCCC-CACTACAATTTGGAAGAAT
217
FR8HA-291S.seq
CACACAGGAAACTACTAGTATGGATTAAAAAGGA-ATAAGCCC-CACTACAATTTGGAAGAAT
217
FR8HA-63G.seq
CACACAGGAAACTACTAGTATGGATTAAAAAGGA-ATAAGCCC-CACTACAATTTGGAAGAAT
217
FR8HA-WT.seq
CACACAGGAAACTACTAGTATGGATTAAAAAGGA-ATAAGCCC-CACTACAATTTGGAAGAAT
217
FR8HA-48G.seq
CACACAGGAAACTACTAGTATGGATTAAAAAGGA-ATAAGCCC-CACTACAATTTGGAAGAAT
217
FR8HA-278S.seq
CACACAGGAAACTACTAGTATGGATTAAAAAGGA-ATAAGCCC-CACTACAATTTGGAAGAAT
217
FR8HA-282S.seq
CACACAGGAAACTACTAGTATGGATTAAAAAGGA-ATAAGCCC-CACTACAATTTGGAAGAAT
217
FR8HA-283G.seq
CACACAGGAAACTACTAGTATGGATTAAAAAGGA-ATAAGCCC-CACTACAATTTGGAAGAAT
217
FR8HA-286S.seq
CACACAGGAAACTACTAGTATGGATTAAAAAGGA-ATAAGCCC-CACTACAATTTGGAAGAAT
217

FR8HA-291G.seq
ACATCGCCGAATAGCTCTTGAGGAAGCAGAACGAGACTGGTCCATGACATCAT
277
FR8HA-291S.seq
ACATCGCCGAATAGCTCTTGAGGAAGCAGAACGAGACTGGTCCATGACATCAT
277
FR8HA-63G.seq
ACATCGCCGAATAGCTCTTGAGGAAGCAGAACGAGACTGGTCCATGACATCAT
277
FR8HA-WT.seq
ACATCGCCGAATAGCTCTTGAGGAAGCAGAACGAGACTGGTCCATGACATCAT
277
FR8HA-48G.seq
ACATCGCCGAATAGCTCTTGAGGAAGCAGAACGAGACTGGTCCATGACATCAT
277
FR8HA-278S.seq
ACATCGCCGAATAGCTCTTGAGGAAGCAGAACGAGACTGGTCCATGACATCAT
277
FR8HA-282S.seq
ACATCGCCGAATAGCTCTTGAGGAAGCAGAACGAGACTGGTCCATGACATCAT
277
FR8HA-283G.seq
ACATCGCCGAATAGCTCTTGAGGAAGCAGAACGAGACTGGTCCATGACATCAT
277
FR8HA-286S.seq
ACATCGCCGAATAGCTCTTGAGGAAGCAGAACGAGACTGGTCCATGACATCAT
277
Figure 46B

**FR8HA-291G.seq**
GGTCCTACATGGAGAAAACCAAAACTCGAGAATGGGATATGTGATATCCAGGAGTTTTCA 337

**FR8HA-291S.seq**
GGTCCTACATGGAGAACAACCAAAACTCGAGAATGGGATATGTGATATCCAGGAGTTTTCA 337

**FR8HA-63G.seq**
GGTCCTACACATGGAGAACAACCAAAACTCGAGAATGGGATATGTGATATCCAGGAGTTTTCA 337

**FR8HA-WT.seq**
GGTCCTACACATGGAGAACAACCAAAACTCGAGAATGGGATATGTGATATCCAGGAGTTTTCA 337

**FR8HA-48G.seq**
GGTCCTACACATGGAGAACAACCAAAACTCGAGAATGGGATATGTGATATCCAGGAGTTTTCA 358

**FR8HA-278S.seq**
GGTCCTACACATGGAGAACAACCAAAACTCGAGAATGGGATATGTGATATCCAGGAGTTTTCA 337

**FR8HA-282S.seq**
GGTCCTACACATGGAGAACAACCAAAACTCGAGAATGGGATATGTGATATCCAGGAGTTTTCA 337

**FR8HA-283G.seq**
GGTCCTAACTATGGAGAACAACCAAAACTCGAGAATGGGATATGTGATATCCAGGAGTTTTCA 337

**FR8HA-286S.seq**
GGTCCTAACTATGGAGAACAACCAAAACTCGAGAATGGGATATGTGATATCCAGGAGTTTTCA 337

-------------------------------------------------------------------

**FR8HA-291G.seq**
TCGACTATGAGGAGCCTGGAGAAGCTGAAATGGAGATTTGGTCTCAGTGTGCTGTCACTTGAAGGATTTGCAA 397

**FR8HA-291S.seq**
TCGACTATGAGGAGCCTGGAGAAGCTGAAATGGAGATTTGGTCTCAGTGTGCTGTCACTTGAAGGATTTGCAA 397

**FR8HA-63G.seq**
TCGACTATGAGGAGCCTGGAGAAGCTGAAATGGAGATTTGGTCTCAGTGTGCTGTCACTTGAAGGATTTGCAA 397

**FR8HA-WT.seq**
TCGACTATGAGGAGCCTGGAGAAGCTGAAATGGAGATTTGGTCTCAGTGTGCTGTCACTTGAAGGATTTGCAA 397

**FR8HA-48G.seq**
TCGACTATGAGGAGCCTGGAGAAGCTGAAATGGAGATTTGGTCTCAGTGTGCTGTCACTTGAAGGATTTGCAA 418

**FR8HA-278S.seq**
TCGACTATGAGGAGCCTGGAGAAGCTGAAATGGAGATTTGGTCTCAGTGTGCTGTCACTTGAAGGATTTGCAA 397

**FR8HA-282S.seq**
TCGACTATGAGGAGCCTGGAGAAGCTGAAATGGAGATTTGGTCTCAGTGTGCTGTCACTTGAAGGATTTGCAA 397

**FR8HA-283G.seq**
TCGACTATGAGGAGCCTGGAGAAGCTGAAATGGAGATTTGGTCTCAGTGTGCTGTCACTTGAAGGATTTGCAA 397

**FR8HA-286S.seq**
TCGACTATGAGGAGCCTGGAGAAGCTGAAATGGAGATTTGGTCTCAGTGTGCTGTCACTTGAAGGATTTGCAA 397

-------------------------------------------------------------------

**FR8HA-291G.seq**
TATTCTCCAAGACACCTCTGCCCACCACACACAAACACGAAGTACTGCTGACATGTGCTGTCACTTTTAAGGATTTGCAA 457

**FR8HA-291S.seq**
TATTCTCCAAGACACCTCTGCCCACCACACACAAACACGAAGTACTGCTGACATGTGCTGTCACTTTTAAGGATTTGCAA 457

**FR8HA-63G.seq**
TATTCTCCAAGACACCTCTGCCCACCACACACAAACACGAAGTACTGCTGACATGTGCTGTCACTTTTAAGGATTTGCAA 457

**FR8HA-WT.seq**
TATTCTCCAAGACACCTCTGCCCACCACACACAAACACGAAGTACTGCTGACATGTGCTGTCACTTTTAAGGATTTGCAA 457

**FR8HA-48G.seq**
TATTCTCCAAGACACCTCTGCCCACCACACACAAACACGAAGTACTGCTGACATGTGCTGTCACTTTTAAGGATTTGCAA 478

**FR8HA-278S.seq**
TATTCTCCAAGACACCTCTGCCCACCACACACAAACACGAAGTACTGCTGACATGTGCTGTCACTTTTAAGGATTTGCAA 457

**FR8HA-282S.seq**
TATTCTCCAAGACACCTCTGCCCACCACACACAAACACGAAGTACTGCTGACATGTGCTGTCACTTTTAAGGATTTGCAA 457

**FR8HA-283G.seq**
TATTCTCCAAGACACCTCTGCCCACCACACACAAACACGAAGTACTGCTGACATGTGCTGTCACTTTTAAGGATTTGCAA 457

**FR8HA-286S.seq**
TATTCTCCAAGACACCTCTGCCCACCACACACAAACACGAAGTACTGCTGACATGTGCTGTCACTTTTAAGGATTTGCAA 457

-------------------------------------------------------------------

**FR8HA-291G.seq**
CCCATGAGGAGAAGAACGTGTTTTTACAGAAAATTTTCTTACGAGAATGGGATATTGCTGACAGGAAAGGATTTGCAA 517

**FR8HA-291S.seq**
CCCATGAGGAGAAGAACGTGTTTTTACAGAAAATTTTCTTACGAGAATGGGATATTGCTGACAGGAAAGGATTTGCAA 517

**FR8HA-63G.seq**
CCCATGAGGAGAAGAACGTGTTTTTACAGAAAATTTTCTTACGAGAATGGGATATTGCTGACAGGAAAGGATTTGCAA 517

**FR8HA-WT.seq**
CCCATGAGGAGAAGAACGTGTTTTTACAGAAAATTTTCTTACGAGAATGGGATATTGCTGACAGGAAAGGATTTGCAA 517

**FR8HA-48G.seq**
CCCATGAGGAGAAGAACGTGTTTTTACAGAAAATTTTCTTACGAGAATGGGATATTGCTGACAGGAAAGGATTTGCAA 538

**FR8HA-278S.seq**
CCCATGAGGAGAAGAACGTGTTTTTACAGAAAATTTTCTTACGAGAATGGGATATTGCTGACAGGAAAGGATTTGCAA 517

**FR8HA-282S.seq**
CCCATGAGGAGAAGAACGTGTTTTTACAGAAAATTTTCTTACGAGAATGGGATATTGCTGACAGGAAAGGATTTGCAA 517

**FR8HA-283G.seq**
CCCATGAGGAGAAGAACGTGTTTTTACAGAAAATTTTCTTACGAGAATGGGATATTGCTGACAGGAAAGGATTTGCAA 517

**FR8HA-286S.seq**
CCCATGAGGAGAAGAACGTGTTTTTACAGAAAATTTTCTTACGAGAATGGGATATTGCTGACAGGAAAGGATTTGCAA 517

-------------------------------------------------------------------

**FR8HA-291G.seq**
CATACCAGAACGCTGAAAAATTCTTATTTGACAAAGAAAAAGGGGAAAGGTGCTCGTTGACTGGAAGGATTTGCAA 577

**FR8HA-291S.seq**
CATACCAGAACGCTGAAAAATTCTTATTTGACAAAGAAAAAGGGGAAAGGTGCTCGTTGACTGGAAGGATTTGCAA 577

**FR8HA-63G.seq**
CATACCAGAACGCTGAAAAATTCTTATTTGACAAAGAAAAAGGGGAAAGGTGCTCGTTGACTGGAAGGATTTGCAA 577

**FR8HA-WT.seq**
CATACCAGAACGCTGAAAAATTCTTATTTGACAAAGAAAAAGGGGAAAGGTGCTCGTTGACTGGAAGGATTTGCAA 577

**FR8HA-48G.seq**
CATACCAGAACGCTGAAAAATTCTTATTTGACAAAGAAAAAGGGGAAAGGTGCTCGTTGACTGGAAGGATTTGCAA 598

**FR8HA-278S.seq**
CATACCAGAACGCTGAAAAATTCTTATTTGACAAAGAAAAAGGGGAAAGGTGCTCGTTGACTGGAAGGATTTGCAA 577

**FR8HA-282S.seq**
CATACCAGAACGCTGAAAAATTCTTATTTGACAAAGAAAAAGGGGAAAGGTGCTCGTTGACTGGAAGGATTTGCAA 577

**FR8HA-283G.seq**
CATACCAGAACGCTGAAAAATTCTTATTTGACAAAGAAAAAGGGGAAAGGTGCTCGTTGACTGGAAGGATTTGCAA 577

**FR8HA-286S.seq**
CATACCAGAACGCTGAAAAATTCTTATTTGACAAAGAAAAAGGGGAAAGGTGCTCGTTGACTGGAAGGATTTGCAA 577

******************************************************************************
Figure 46 C

FR8HA-291G.seq  GGGTTATTTCTACCCCCCTACAGTAAGGAAACACAGAAATATCTATCGGATGAAAGAAATG  637
FR8HA-291S.seq  GGGTTATTTCTACCCCCCTACAGTAAGGAAACACAGAAATATCTATCGGATGAAAGAAATG  637
FR8HA-63G.seq  GGGTTATTTCTACCCCCCTACAGTAAGGAAACACAGAAATATCTATCGGATGAAAGAAATG  637
FR8HA-WT.seq  GGGTTATTTCTACCCCCCTACAGTAAGGAAACACAGAAATATCTATCGGATGAAAGAAATG  637
FR8HA-48G.seq  GGGTTATTTCTACCCCCCTACAGTAAGGAAACACAGAAATATCTATCGGATGAAAGAAATG  658
FR8HA-278S.seq  GGGTTATTTCTACCCCCCTACAGTAAGGAAACACAGAAATATCTATCGGATGAAAGAAATG  637
FR8HA-282S.seq  GGGTTATTTCTACCCCCCTACAGTAAGGAAACACAGAAATATCTATCGGATGAAAGAAATG  637
FR8HA-283G.seq  GGGTTATTTCTACCCCCCTACAGTAAGGAAACACAGAAATATCTATCGGATGAAAGAAATG  637
FR8HA-286S.seq  GGGTTATTTCTACCCCCCTACAGTAAGGAAACACAGAAATATCTATCGGATGAAAGAAATG  637

FR8HA-291G.seq  CTATTATGTCTGATGTGATCTCAATATATAATTACAGAGATTTACCCGGAAATGAGAAATG  697
FR8HA-291S.seq  CTATTATGTCTGATGTGATCTCAATATATAATTACAGAGATTTACCCGGAAATGAGAAATG  697
FR8HA-63G.seq  CTATTATGTCTGATGTGATCTCAATATATAATTACAGAGATTTACCCGGAAATGAGAAATG  697
FR8HA-WT.seq  CTATTATGTCTGATGTGATCTCAATATATAATTACAGAGATTTACCCGGAAATGAGAAATG  697
FR8HA-48G.seq  CTATTATGTCTGATGTGATCTCAATATATAATTACAGAGATTTACCCGGAAATGAGAAATG  697
FR8HA-278S.seq  CTATTATGTCTGATGTGATCTCAATATATAATTACAGAGATTTACCCGGAAATGAGAAATG  718
FR8HA-282S.seq  CTATTATGTCTGATGTGATCTCAATATATAATTACAGAGATTTACCCGGAAATGAGAAATG  697
FR8HA-283G.seq  CTATTATGTCTGATGTGATCTCAATATATAATTACAGAGATTTACCCGGAAATGAGAAATG  697
FR8HA-286S.seq  CTATTATGTCTGATGTGATCTCAATATATAATTACAGAGATTTACCCGGAAATGAGAAATG  697

FR8HA-291G.seq  GACCCAAAGTAAAGATCAAGCTGCGGGGAGATAGAATACTATGTGGAACCCTCTGCAAAACCGG  757
FR8HA-291S.seq  GACCCAAAGTAAAGATCAAGCTGCGGGGAGATAGAATACTATGTGGAACCCTCTGCAAAACCGG  757
FR8HA-63G.seq  GACCCAAAGTAAAGATCAAGCTGCGGGGAGATAGAATACTATGTGGAACCCTCTGCAAAACCGG  757
FR8HA-WT.seq  GACCCAAAGTAAAGATCAAGCTGCGGGGAGATAGAATACTATGTGGAACCCTCTGCAAAACCGG  757
FR8HA-48G.seq  GACCCAAAGTAAAGATCAAGCTGCGGGGAGATAGAATACTATGTGGAACCCTCTGCAAAACCGG  778
FR8HA-278S.seq  GACCCAAAGTAAAGATCAAGCTGCGGGGAGATAGAATACTATGTGGAACCCTCTGCAAAACCGG  757
FR8HA-282S.seq  GACCCAAAGTAAAGATCAAGCTGCGGGGAGATAGAATACTATGTGGAACCCTCTGCAAAACCGG  757
FR8HA-283G.seq  GACCCAAAGTAAAGATCAAGCTGCGGGGAGATAGAATACTATGTGGAACCCTCTGCAAAACCGG  757
FR8HA-286S.seq  GACCCAAAGTAAAGATCAAGCTGCGGGGAGATAGAATACTATGTGGAACCCTCTGCAAAACCGG  757

FR8HA-291G.seq  GAGACCAATATATTTTGGCGGAAAATGGAATATAGCTATAGCAGCAAAATGAAGTCTTGGCACAC  817
FR8HA-291S.seq  GAGACCAATATATTTTGGCGGAAAATGGAATATAGCTATAGCAGCAAAATGAAGTCTTGGCACAC  817
FR8HA-63G.seq  GAGACCAATATATTTTGGCGGAAAATGGAATATAGCTATAGCAGCAAAATGAAGTCTTGGCACAC  817
FR8HA-WT.seq  GAGACCAATATATTTTGGCGGAAAATGGAATATAGCTATAGCAGCAAAATGAAGTCTTGGCACAC  817
FR8HA-48G.seq  GAGACCAATATATTTTGGCGGAAAATGGAATATAGCTATAGCAGCAAAATGAAGTCTTGGCACAC  817
FR8HA-278S.seq  GAGACCAATATATTTTGGCGGAAAATGGAATATAGCTATAGCAGCAAAATGAAGTCTTGGCACAC  838
FR8HA-282S.seq  GAGACCAATATATTTTGGCGGAAAATGGAATATAGCTATAGCAGCAAAATGAAGTCTTGGCACAC  817
FR8HA-283G.seq  GAGACCAATATATTTTGGCGGAAAATGGAATATAGCTATAGCAGCAAAATGAAGTCTTGGCACAC  817
FR8HA-286S.seq  GAGACCAATATATTTTGGCGGAAAATGGAATATAGCTATAGCAGCAAAATGAAGTCTTGGCACAC  817

FR8HA-291G.seq  TGGATGAGGAGCTTTGGCTCGGCATACGACTTATGACGTTCCAGCTTTCTGCAAC  851
FR8HA-291S.seq  TGGATGAGGAGCTTTGGCTCGGCATACGACTTATGACGTTCCAGCTTTCTGCAAC  851
FR8HA-63G.seq  TGGATGAGGAGCTTTGGCTCGGCATACGACTTATGACGTTCCAGCTTTCTGCAAC  851
FR8HA-WT.seq  TGGATGAGGAGCTTTGGCTCGGCATACGACTTATGACGTTCCAGCTTTCTGCAAC  851
FR8HA-48G.seq  TGGATGAGGAGCTTTGGCTCGGCATACGACTTATGACGTTCCAGCTTTCTGCAAC  851
FR8HA-278S.seq  TGGATGAGGAGCTTTGGCTCGGCATACGACTTATGACGTTCCAGCTTTCTGCAAC  851
FR8HA-282S.seq  TGGATGAGGAGCTTTGGCTCGGCATACGACTTATGACGTTCCAGCTTTCTGCAAC  851
FR8HA-283G.seq  TGGATGAGGAGCTTTGGCTCGGCATACGACTTATGACGTTCCAGCTTTCTGCAAC  851
FR8HA-286S.seq  TGGATGAGGAGCTTTGGCTCGGCATACGACTTATGACGTTCCAGCTTTCTGCAAC  851

* * *
Figure 46 D

FR8HA-291G.seq
---------AAACGCATCAATGCAAGTGGAAATCTGACTCTCCAGGAAACAGGAGT 904
FR8HA-291S.seq
---------AAACGCATCAATGCAAGTGGAAATCTGACTCTCCAGGAAACAGGAGT 904
FR8HA-63G.seq
---------AAACGCATCAATGCAAGTGGAAATCTGACTCTCCAGGAAACAGGAGT 883
FR8HA-WT.seq
---------AACACGACTCAATGCAAGTGGAAATCTGACTCTCCAGGAAACAGGAGT 883
FR8HA-48G.seq
---------AACACGACTCAATGCAAGTGGAAATCTGACTCTCCAGGAAACAGGAGT 904
FR8HA-278S.seq
---------TCAACGACTCAATGCAAGTGGAAATCTGACTCTCCAGGAAACAGGAGT 886
FR8HA-282S.seq
---------TCAACGACTCAATGCAAGTGGAAATCTGACTCTCCAGGAAACAGGAGT 904
FR8HA-283G.seq
---------TCAACGACTCAATGCAAGTGGAAATCTGACTCTCCAGGAAACAGGAGT 904
FR8HA-286S.seq
---------TCAACGACTCAATGCAAGTGGAAATCTGACTCTCCAGGAAACAGGAGT 904

* * * ********* 964

-----------------------------------------------

FR8HA-291G.seq
GTCAACACCGCTCCGAGATTTACATCCATATTGCAGCTGACTCTCCAGGAAACAGGAGT 964
FR8HA-291S.seq
GTCAACACCGCTCCGAGATTTACATCCATATTGCAGCTGACTCTCCAGGAAACAGGAGT 964
FR8HA-63G.seq
GTCAACACCGCTCCGAGATTTACATCCATATTGCAGCTGACTCTCCAGGAAACAGGAGT 964
FR8HA-WT.seq
GTCAACACCGCTCCGAGATTTACATCCATATTGCAGCTGACTCTCCAGGAAACAGGAGT 964
FR8HA-48G.seq
GTCAACACCGCTCCGAGATTTACATCCATATTGCAGCTGACTCTCCAGGAAACAGGAGT 964
FR8HA-278S.seq
GTCAACACCGCTCCGAGATTTACATCCATATTGCAGCTGACTCTCCAGGAAACAGGAGT 964
FR8HA-282S.seq
GTCAACACCGCTCCGAGATTTACATCCATATTGCAGCTGACTCTCCAGGAAACAGGAGT 964
FR8HA-283G.seq
GTCAACACCGCTCCGAGATTTACATCCATATTGCAGCTGACTCTCCAGGAAACAGGAGT 964
FR8HA-286S.seq
GTCAACACCGCTCCGAGATTTACATCCATATTGCAGCTGACTCTCCAGGAAACAGGAGT 964

* * * ********* 964

-----------------------------------------------

FR8HA-291G.seq
CAATAGGGAGGGTCGCACATTGCAAGTGCTGATTACAGTTTGCAGCTGACTCTCCAGGAAACAGGAGT 1024
FR8HA-291S.seq
CAATAGGGAGGGTCGCACATTGCAAGTGCTGATTACAGTTTGCAGCTGACTCTCCAGGAAACAGGAGT 1024
FR8HA-63G.seq
CAATAGGGAGGGTCGCACATTGCAAGTGCTGATTACAGTTTGCAGCTGACTCTCCAGGAAACAGGAGT 1024
FR8HA-WT.seq
CAATAGGGAGGGTCGCACATTGCAAGTGCTGATTACAGTTTGCAGCTGACTCTCCAGGAAACAGGAGT 1024
FR8HA-48G.seq
CAATAGGGAGGGTCGCACATTGCAAGTGCTGATTACAGTTTGCAGCTGACTCTCCAGGAAACAGGAGT 1024
FR8HA-278S.seq
CAATAGGGAGGGTCGCACATTGCAAGTGCTGATTACAGTTTGCAGCTGACTCTCCAGGAAACAGGAGT 1024
FR8HA-282S.seq
CAATAGGGAGGGTCGCACATTGCAAGTGCTGATTACAGTTTGCAGCTGACTCTCCAGGAAACAGGAGT 1024
FR8HA-283G.seq
CAATAGGGAGGGTCGCACATTGCAAGTGCTGATTACAGTTTGCAGCTGACTCTCCAGGAAACAGGAGT 1024
FR8HA-286S.seq
CAATAGGGAGGGTCGCACATTGCAAGTGCTGATTACAGTTTGCAGCTGACTCTCCAGGAAACAGGAGT 1024

* * * ********* 1024

-----------------------------------------------

FR8HA-291G.seq
GGGATGGAGGTCGCACATTGCAAGTGCTGATTACAGTTTGCAGCTGACTCTCCAGGAAACAGGAGT 1084
FR8HA-291S.seq
GGGATGGAGGTCGCACATTGCAAGTGCTGATTACAGTTTGCAGCTGACTCTCCAGGAAACAGGAGT 1084
FR8HA-63G.seq
GGGATGGAGGTCGCACATTGCAAGTGCTGATTACAGTTTGCAGCTGACTCTCCAGGAAACAGGAGT 1084
FR8HA-WT.seq
GGGATGGAGGTCGCACATTGCAAGTGCTGATTACAGTTTGCAGCTGACTCTCCAGGAAACAGGAGT 1084
FR8HA-48G.seq
GGGATGGAGGTCGCACATTGCAAGTGCTGATTACAGTTTGCAGCTGACTCTCCAGGAAACAGGAGT 1084
FR8HA-278S.seq
GGGATGGAGGTCGCACATTGCAAGTGCTGATTACAGTTTGCAGCTGACTCTCCAGGAAACAGGAGT 1084
FR8HA-282S.seq
GGGATGGAGGTCGCACATTGCAAGTGCTGATTACAGTTTGCAGCTGACTCTCCAGGAAACAGGAGT 1084
FR8HA-283G.seq
GGGATGGAGGTCGCACATTGCAAGTGCTGATTACAGTTTGCAGCTGACTCTCCAGGAAACAGGAGT 1084
FR8HA-286S.seq
GGGATGGAGGTCGCACATTGCAAGTGCTGATTACAGTTTGCAGCTGACTCTCCAGGAAACAGGAGT 1084

* * * ********* 1084

-----------------------------------------------
>PR8HA-WT-Homo sapiens
ATGAAGGCAACCTGCTGTTGCTGCTGCAGCTGCCCGTGGCTGAGCACAGGATCGCCGAGATCCACTGCTGTATCGGGTT
ACCAGCAGCCAAACACACGGACACCGCTGAGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT
GAACCTGCTGGAAGATAGCAGACCAGCCACAGTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG
AAGTGCAATATCGCCGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG
CCTACATCGTTAAAAACCCCAAAAGCAAGGTGAAACGCTATCGTTAAAAACCCCAAAAGCAAGGTGAAACGCTATCGTTAAAAAC
ACTGCGCGAGCAGCTGACGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG
CCAACCCACAAAACCCGCGCGCAGGAGCAGGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG
TGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG
CTTGGAGGCCAGCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG
ATTACGACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG
ATCATATGGCGCGCCGCGCAGGACTGGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG
GCAAGCCTGCCATGTACGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG
CTTGGAGGCCAGCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG
ATTACGACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG
GCAAGCCTGCCATGTACGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG
CTTGGAGGCCAGCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG
ATTACGACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG
GCAAGCCTGCCATGTACGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG
CTTGGAGGCCAGCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG
ATTACGACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG
GCAAGCCTGCCATGTACGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG
CTTGGAGGCCAGCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG
ATTACGACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG
GCAAGCCTGCCATGTACGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG
CTTGGAGGCCAGCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG
ATTACGACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG
GCAAGCCTGCCATGTACGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG
CTTGGAGGCCAGCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG
ATTACGACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG
GCAAGCCTGCCATGTACGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG
CTTGGAGGCCAGCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG
ATTACGACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG
GCAAGCCTGCCATGTACGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG
CTTGGAGGCCAGCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG
ATTACGACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG
GCAAGCCTGCCATGTACGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG
CTTGGAGGCCAGCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG
>PR8HA-WT-Criculus-griseus
ATGAAGGCCAACCCTCTTGCTGCTGCTGCTGCTGCTCTGGCTGCCGCTGATGGCTGACACCATCTGATGCTGCT
ACCAGCACAACCCTCCACCGGACACCCTGCTGGGATACCCGGCTGCTGCTGCGCTGACACCTGGAGCTGAG
GAATCTGCTGCGAAATATCCACACGCGCAAGCTGGTCCGGCTGCTGAGGGAATCGCGCCCCCCCTTCTAGCTG
GAAGTGAATATCGCTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT
ACCTACATCGTGGAAACCCCAATCCCAGAAGCAGGCTCTGCTTACCCCGGGCGAGACGCTACTGACGAG
ACPTGCGGAGCAGCTGCTCCTCCGGCTGCTGACCTCCGGCCCTTCTCACAGGGGAAAGGCAAGCTTACGCC
CCCAACCAACACACACCGCGGTTGAGCCGCCGCTTCTCTCAGGGGAAAGGCAAGCTTACGCC
CTCAATCTCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT
ATCACATCCTACACACCGCCCTCCATGACAGTGAACACCAAGTGCCAGACCCCCTGGGCGCCATCAACA
GCTCCCGCTTACAGGAACATCCACCCCCTGCTGACATCCGGGAGTGGCCTCCCCGAAATATGGCGGAGCGCC
GCTCGGGATTGTGCACCACCGCTGAGAACACCCCTTCCATCCAGCTAGGGCGTCCGCGTATCGCC
GGCTTTATCAGGAGGCGCCTGGACCCCGGATGATCAGGAGCGGTTGCTACGGCTACCAACAGTACAGCAG
GCTCCGGCTACCCCGCGCAAGAATCTACCAGAGACATACCCAGCGCAAGGCTCTGAGGAAACAGTACAGCAG
CGTGATCGAAGATGAACATCCAGTCCACCGCGGCTGGGACAAGAGCTTTCAACAAACAGTGGAAACACGGATG
GAAACCTGGAACAAAGGTTGGGAGCAAGCTCTGCTGGCACTCTGACATCCACGCGAGCTGGCTGCTGGC
TGCTGGAAACGAGCGGACCCTTGGACTCTCCACGGACTCCAAACGTGGAAACACTGTACGGAAGAAGTGAGACT
CCAGCTGAAAACACCGCAGAACAGATCGCGAAGCTGCTTCCGTTAACAAGACAGTGCAACACGGAG
TGCAAGTAATCCTGGGACAAGGGCACCCTACGAGACTCTCCACGGAGAAAGACGAAGCTGAAACCCGG
AAAAAGTGGAACCGCTGGTAACTCCACGGGATCTATGCTGCTGCGGACCTACTCCACCGTGCC
CTCCAGCCTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT
TGCCGGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT

Figure 49
(SEQ ID NO.64)
>PR8HA-WT-Nicotiana-benthamiana
ATGAGGCTAACCTTCGCTGCTCTGCTGCTGCTGCTGATGCTGATCACATCTGCATTTGTT
ACCAGGCTAACACAGACCCAGATTCTGTTGAGATTACGCTconstGACCACGACCTCTCCTGT
GAACTGCTTGGAGATCTACCAACGCTGAAGCTGTGAGGCTCTTAAGGTATATGCTCCACTGCAGCTTGTT
AAGTGCAATATCGCTGTGTTGCTCTGCTGCTGCTGCTGCTGATGCTGATCACATCTGCATTTGTT
CCCTACATTGTTGAAAAGCTCTACTACAGCAGGAACCCGATATGCTGCTACTCCCTGCTGCTGATATCGATTG
ACPAGAGAGAGCTCCAGCTGCTTGCAGTCTATTACCATCTTTGGAGAGGTTCGAGATCTTCTCCCTCAAGAGTCCTCTTTG
CCTACTACCCAGACTAAAGGTTTGCCGTTAGGAGTAAAGCTCAGCTGACCTCTTGGAGAGGACC
CTTAGCTAGACGGTCTGCCTTAACCTGATCGCTCTCTATATGCTCTTCTGCTCTGTAGGGTTTCCGTTCTGTATT
ATCATTACCAAGCAACGCTCTTCTATGCAAGGTGGAATTCATACAAAGGTTGACGAGATCCCTTGGTTGCTATCAAAC
GCATCTCTCCCTACGCAAGCACTTACCCCTGCTGAGATTCGTTGGTGAGTCCTAAGCTAATGTGGTTAGGTCTGCTAAG
GTTAGGATGTTGACCCGCTCTTGAGGAACACCCCTTCTATCCAGTCGGGACGTTCCTCGAGCTATCGCT
GGATTATTGAGGAGGTGGTGTGAAACCTGGTTACGATGCTGATGTTGATAGCTTTACTCAGTACGAGAAGCAGG
GGTTCTGGATACCTGCTGCTGCTAAGAGCTACCCAGAAAGCTATACCAACCGACTAAAGGTTGAAGTGGAAC
CGTGCAGAAGATGAACATCCACTGTTACCCGCTGCTGCTGCTGCTGCTGATGCTGATCACATCTGCATTTGTT
GAAAACCTATACAAAGAAGGTTTATGTTGTTTCTCGGATATCGCTCAACAGCTGCTGCTGCTGCTGCT
TTCTTGAGAAGAGAGGCCTCGTCTGGATTTTCCAGCAGTAAGCAACGCTGAGAACCCTGTAGCAGAGT
CCAGCTTAAAGAACAACGCTTAAGAGACTGCTGCTGCTGCTGCTGCTGATGCTGATCACATCTGCATTTGTT
TGCAAGGATCTGCGATCTCAG
Figure 51

(SEQ ID NO.66)

>PR8HA-WT-Pichia-pastoris
ATGAGGCTACTTGTTTCTCGCTTGTCTGCTGCTGATGCTGACATGATCTGATATGTGTT
ACCAGCTAAACACTCCACTGACACTGTGGAGACAGCATTTTGGAGAGAAACCTGTTCTGCTGACTTCTGGT
TAATTTTTTGGAGGACACTCCACAAACGCTTATACAGTTGAGTATGATAGAGGAGTTATCGCTTCATATGGCTGGAG
AACTGTAACATTGCTGTTGTTGTTGGAATACCAGAGTGTAGTCTCTTGTGGCAGTTAGACTCTGGT
CCTACATCGTGGAGACTCACAATCTCGAGAAGCAGCTTCTGCTGTTACCTCCACAGGAGGAAAGTCCTCCATCTACTCAAGAAC
TGTGTGTTGACTGAGAAAGAGGTTCTCTACACAAAAAGTGAAGAAGATTCTCTCGAGAAGAATCTCTG
GAGAAGTTACGTGGTTTCTGTTGCTATCTCCACACTACAACGAGATTTCAACTCGACAGATCGAGAAAGAC
CAAGGTTAGAGATCAGGGTTGCTGTAAGATGAAGACTAAGCTGAGCTGTGAGTGGTCTAAGTACCTGAGATACGCTAA
CTTCCAGCTACTACCGTATCTGTGCTCCACCTAACGAGCTTGCTCTTGGCAGAGGTTCCGCTCGGT
ATCATCCTCTACACGGCTTCCATGCAGGTAACACTAAGCTGACATTCAATGCTGTTGGTCTAATCTAAGAC
CCCCTCCTCCAGCTACCCACACACACCCCAGTTCTGATCCTGGAGTGCTACGGTGCTATCGACGCTGAGGAAAGAC
GTGGAGATGTTACTGTTTCTTGGAGAAGACCTCCATACCTAGCTAGGGTGCTACCGCTACACACAAG
GTTCCTGGTATGCTGGGTTCAGAGGTCCACACTACCCAGAGTTAAGTATACCCAGATCGAGGCTGTTCGCGT
GCTATGTGCTACGCTGCTACTAGATGTTAAGTATACCCAGATCGAGGCTGTTCGCGT
GTTTTTATGGAAGTGGTTTGGACTGCTGTGATCGACGGTGTTGCTACCGCTACACACAAG
GTGTTGTTACGCTGCTGACACAAATGCAAAGGTAGATACCTAACGAGTATCAACAGGGAGTTACAC
TGCTATCGAGAAGATGACACTCACGTCTGCTGTTGGAATAGGTCACAGAAGGTGAAGAGAAAGAGATG
GAAAATCTTGGAAACAGAGAAGTTGACGGCTTCTTGGGAGACATCTGAGGCTTCAAACGCGTAGGTGTGTTCTGT
GTGGAGAAGACGAGAATTTTGGGACTTCCACAGCTCCACACGTTAAGAACTTGGTACGAGAGGTTAACAC
CCATTAGGAAGCAAAGCTTACAGGAGATCGTTAGCTGGTCTACTGCACAAAGTGTGACAACAG
TGAGTTGGAATCCGTTGAAAGGCTGCTTACGCTACCCCTAATGACTCAGAGACAGTTAAAGCAG
AAAAGGGTGTAGGGTTAAGTTTGGGAAATCCATGGGTATCTACTAACAATCTGGCTAC
TTCCCTCTGGTTTGTGGTTTCTTGGTTCTGCTGTTGTGCTACCTTCTGAGTTGTTTCCACAGGTTTCCGCA
GTAGAATCTCTATCTAA
>PR8HA-WT-Saccharomyces-cerevisiae
ATGAAGGCCAACACCTTTGTTTTTGGTCGCCTGTGCTGCTGATGATCATAATATGTATTTGATT
ACCATGCCAACACTTCACTCCGATACTTTGGAATTGTGCACTAATGGGATGTTCTTGG
CAATTTGGAGTCAAGGATTCCCTAGAAGATTGCGATTTGGTACCAATTTGCATTGAATTTG
AATGTAATATTTGGCTTGGTTGTTGTTGGTGTTAACCAGAATGATCCTATATTGATCCAGTTAATCTTG
CTTACACATCGTGGAAAATCCCAATCCGAGAAATGTTATTTTGGGATCCACACCTGATTGAGA
ATGGAGAGAACAAATTTGCTCCTCCGTTTCTCATTCAAGGAATTTGCAAATCTCCTCTTGG
CACAACCATACAATATGAGTTACTGCTGCTGTTGTTCCATAGAGTAAAAGTTCTTCTTACTAGAATCT
TTTGTTGGGTGGGACGGAAAAGAGGTCTTCTTACCCAAAGTGGAGAAACTCTCCTACGTGTAAC
AGAAGTTTGTGTTCTTGGGGAATTTCACTCTCACCACAAACTCAAGAACAACAACACTACACAAAC
GAAAGGCTCTGCTTCTTGGATATGGCCAATGTATGTCTGTTTCGGTTCTAGAGGTATTTGTTGGTCCC
ATTAGTTAGCTTCCAAACGCTTCTATGCAAGATCACAACACTAATGGGCTATCATAATCT
CATCTTACGCGACACAAAAATCAGCCACAGTTACTTATGGTGAATGTCCTAATAGTATGTTAGATCCCAA
ATGGAAGATTTGTTACCGGTTGAGAAAACCCCCATCATTCATATCAAGAGGTGTTTTGGTGTGATATTGCC
GTTTTATTGGAAATGTTGGTGAGACTTGTATGATTGTATGTTGGTATGTTACCATCAACAAAGGACAG
GTTCTGGTACGCTGCTGTAACAAAGATCTACCTAAAAATGGCCATTAACGTTTATTACAAACAAGTTAAC
CGTCATCGAAAAGATGAACATTCAATTGCAGGCCGCAGGCAGAGATCTCAGAAGATGGAGAGATG
GAAAATCTGGAACAAAAAGGGTTCGAGCTGTTTTCTTGGAATTGTGAAATTCAAAACAGTGTGATAAAC
TGCATCGAAATCTCCTAAGAAACGCTTACATGTTACCTCAGATTTACCCAAACACGTGAAATGCAAGG
AAAAGGTGGACGGTCTAAGGTGAATCTAAGCCTATTTGCTGTTGCTGTTGGGCCATCTACTCTACTGTG
CTCTTTAATAGGTTTGTGTTAGTTGTTGGTTGGTTGGTTGGTTGGTTGGTTGGTTGGTTGGTTGGTTGGTTG
AAGTGAATCTGTATATTGGA
Figure 54 A

HA-USSR-77.pro
-----MKALKLVLCLALSDAT---ADTGICGHANHHSTDTSVDTSVEKU41
HA-Texas-91.pro
-----MKALKLVLCLACTATY---ADTGICGHANHHSTDTSVDTSVEKU41
WT-PR8-34.pro
-----MKALKLVLCSLAGAAAD---ADTGICGHANHHSTDTSVDTSVEKU41
HA-WSN-33.pro
-----MKALKLVLVLYFVATD---ADTGICGHANHHSTDTSVDTSVEKU41
HA-SouthCarolina-1983.pro
-----MEARLVLCLACAATN---ADTGICGHANHHSTDTSVDTSVEKU41
HA-California-09.pro
-----MKALVLVLLYFTAFAT---ADTGICGHANHHSTDTSVDTSVEKU41
HA-Singapore-57.pro
-----MAIIYLILLTAVR---GDICGICGHANNNSTKEDTVLTERH43
HA-Vietnam-04.pro
-----MEKTIVLLFAVSVLK---SDICGICGHANNSQVTQDMEKU41
HA-Udorn-72.pro
MTKIALSYLIFCIVLQQDFPDNGNATMICLGHVHNVGLTVKTISDQ50
HA-HongKong-68.pro
MTKIALSYLIFCICLQLDQLDNGNATMICLGHVHNVGLTVKTISDQ50
HA-Panama-99.pro
MTKIALSYLILVLPQK LGPNDNGNATMICLGHVH NSW melTVKTISDQ50
HA-Wisconsin-05.pro
MTKIALSYLILVLPQK LGPNDNGNATMICLGHVH NSW melTVKTISDQ50
HA-Shanghai-13.pro
MI----TQTLVFALAIIFTH--ADCGICGHANHHSTS GTWNLTERG42

HA-USSR-77.pro
TVTHSVNLLEDSHNGKICRLKGIALPLQGKCSIAWGILHNCHPCEE SLVSK91
HA-Texas-91.pro
TVTHSVNLLEDSHNGKICRLKGIALPLQGKCSIAWGILHNCHPCEE SLFSE91
WT-PR8-34.pro
TVTHSVNLLEDSHNGKICRLKGIALPLQGKCHNIAWGILHNCHPCEEPLPVR91
HA-WSN-33.pro
AVTSHSVNLLEDSHNGKICRLKGIALPLQGKCHNIAWGILHNCHPCECSLIF91
HA-SouthCarolina-1983.pro
TVTHSVNLLEDSHNGKICRLKGIALPLQGKCHNIAWGILHNCHPCEEDLLTAS91
HA-California-09.pro
TVTHSVNLLEDSHNGKICRLKGIALPLQGKCHNIAWGILHNCHPCEE SLTA91
HA-Singapore-57.pro
TVTHADILEKTHNKICLGNPFFLELGDSIAWGILHNCEPCDLRSSYP89
HA-Vietnam-04.pro
TVTHAQDLILEKTHNKICLCDLGKFLRLDSCAVWGILHNCEPCDEFINV90
HA-Udorn-72.pro
EVNATLVEGSTGKCHI--NPHRILDGDCTLDDLADDGDGCDFQNEH98
HA-HongKong-68.pro
EVNATLVEGSTGKCHI--NPHRILDGDCTLDDLADDGDGCDFQNEH98
HA-Panama-99.pro
EVNATLVEGSTGKCHI--NPHRILDGDCTLDDLADDGDGCDFQNEH98
HA-Wisconsin-05.pro
EVNATLVEGSTGKCHI--NPHRILDGDCTLDDLADDGDGCDFQNEH98
HA-Shanghai-13.pro
EVNATeterBTHuIFRICS--KGHRVTVLDGQCGLLGLTGFQPCDFQLEF90

HA-USSR-77.pro
SKSYIAETPNSGTCTGYFADYEBEELQSLSSVESFPERFIEFPKESWP141
HA-Texas-91.pro
SKSYIAETPNSGTCTGYFADYEBEELQSLSSVESFPERFIEFPKESWP141
WT-PR8-34.pro
SKSYIVETPSNGTCTGYPDIFYDEELQSLSSVESFPERFIEFPKESWP141
HA-WSN-33.pro
SKSYIVETPSNGTCTGYPDIFYDEELQSLSSVESFPERFIEFPKESWP141
HA-SouthCarolina-1983.pro
SKSYIVETPSNGTCTGYPDIFYDEELQSLSSVESFPERFIEFPKESWP141
HA-California-09.pro
SKSYIVETPSNGTCTGYPDIFYDEELQSLSSVESFPERFIEFPKESWP141
HA-Singapore-57.pro
EKSIVEMKMRPGDCGLYFSGNDFELKLHSLVSFGHKEEKLLELPFWR138
HA-Vietnam-04.pro
EKSIVEMKMRPGDCGLYFSGNDFELKLHSLVSFGHKEEKLLELPFWR138
HA-Udorn-72.pro
TWDLFVERKAFS-HCYPVDVPSLASLVAASG--TEFLSEPFTTW144
HA-HongKong-68.pro
TWDLFVERKAFS-HCYPVDVPSLASLVAASG--TEFLSEPFTTW144
HA-Panama-99.pro
KGDLFVERKAFS-HCYPVDVPSLASLVAASG--TEFLSEPFTTW144
HA-Wisconsin-05.pro
KGDLFVERKAFS-HCYPVDVPSLASLVAASG--TEFLSEPFTTW144
HA-Shanghai-13.pro
ADL11ERGERSD-VCYPKFVHEAQLIRGREG--DILEMAEGFTY136

HA-USR-77.pro
KHNVTRGTASCSHGKGSFSSYYRLLWLTE--KNGSYPHLKSYYVNYKEKEE189
HA-Texas-91.pro
NHTVTKGTVTSSCHHGKSGSFYRNLWLTE--KNGSYPHLKSYYVNYKEKEE189
WT-PR8-34.pro
NNH-TNVTGTVTSSCHHGKSGSFYRNLWLTE--KGSYPHLKSYYVNYKEKEE189
HA-WSN-33.pro
NHTVTKGTVTSSCHHGKSGSFYRNLWLTE--KGSYPHLKSYYVNYKEKEE189
HA-SouthCarolina-1983.pro
NHTVTKGTVTSSCHHGKSGSFYRNLWLTE--KGSYPHLKSYYVNYKEKEE189
HA-California-09.pro
NHTVTKGTVTSSCHHGKSGSFYRNLWLTE--KGSYPHLKSYYVNYKEKEE189
HA-Singapore-57.pro
QHTTGG--SRACAVGSHFSFRNHMTL--KGGYVFPAKDGSNHNTSGSEGQ187
HA-Vietnam-04.pro
SHEASLGVSSACPYQKSSFFHRVVLK--KGSTYPTIKRSHYHNNQED187
HA-Udorn-72.pro
GVT-QINGSSNACKRQGFDGFSRLNWL--SAGTYPVLFVTPMMNNDF191
HA-HongKong-68.pro
GVT-QINGSSNACKRQGFDGFSRLNWL--SAGTYPVLFVTPMMNNDF191
HA-Panama-99.pro
GVT-QINGSSNACKRQGFDGFSRLNWL--SAGTYPVLFVTPMMNNDF191
HA-Wisconsin-05.pro
GVT-QINGSSNACKRQGFDGFSRLNWL--SAGTYPVLFVTPMMNNDF191
HA-Shanghai-13.pro
GIR-TNGATSACCR-SSSSFIAEMKVLSSNIDNAAFPMQMTKSYHTRGK184
>PR8HA-soluble
MKANLLVLLSAAAADADTICIGYAHNNSTDTVTDVKVTIINNLLEDSHNGKLCLRKLGIAPLQLGKCNIA
NGLLNPECDDLPPVRWSYIVETPNSBNGICYPGDFIDYELREQLSSSVSSEFERIEFPKES
KIPNFHNTINGVTAAACSHEGKSSFYRNLLWETKEGSYPKLNISYVKKKKEVLVLKCGIHPNPNSKEQ
QHIYQMNAYVSVSVTVSTYHRRFTEIAERPVKRDQA
GRNNTYWTLKPGDITIEANGLMPIFAFGALSРGFSGIITSNASMHECNTKCQFQLGAHSVSLP
YQNIHPVTIGECP
KTVSAKLRMVTGRUTPSIQSRLFGAIAFGIEGWTGMIDGWYGYHHQNEQGSGYAADQKSTQA
INGITNKVNTVIE
KMNHQFTAVGKEFKLEKRMENLNNKVDGFLDITYNAEELVLGRENERTLDFHDSNKNLYEKVKSL
QKNKNAEKIGNGC
FEFYHCDNEMCEMSVINGTYDTPKYESKLNREKVDGVRSLVPGRGSPSGGYIPEAIPDGQAAYVRKDGE
KWLLSTFLHHH
HHH
>HA-HongKong-68-soluble.pro

MKTIALSYIFCLALGQDLFGNMTATLCLGHHAVPNGTLVKTIITDDQIEVTNATELVQSSSTGKICNPHRILDGIDC
TLIDALGDHPDCFQNETWDLFFERSKAFSNCPYDFPDYASLRLVASSGSTLFEFITEGFTVTQNGGNAACKRPGF
NGFFSRLLMLTKGSTYPVLNVTMPNNDNFDKLYIWGVHHPSTQNEQTSLYVQESGRVTSTRSQQSIIPNIGSRPWVR
GQSSRISIYWTVKPGVDSLVINSINGLIALPRGYFKMRTGKSSIMSSSEPDCITCIEFSITPHGSPIDKFPQNVKITYGA
CFKYVKQNTKLATGMNSVFEKQTRGLGAIAGFENGWEGMDQNYGFRBQNSGTQADLKTQSAIDQINGKLRNV
IEKTNKPHQIEKEFSEVEGRIDQLEKVEVTIKLWYNAELLVALENQHTIDLTDSEMNLKEFTQVRENAEMGN
GCFKIIYHKCDNACIESIRNGTYHDVYRDALNNRPQIKGVRLVPGRSPGSYIFPEAPROGQAYVRRDGEWLLSTFLH
HHHHH
>HA-Wisconsin-05-soluble.pro

MKTIIALSYILCLVFAQKLFPGNDNSTATCLGHHAVPNIGHTVIKTITNDDIQEVNTNATELVQGSSSTGGICDSPHQILDENC
TLIDALLGPQCDGFQKIKWDLFVERSAYNCPYDFVDPHLRSLVASSGTLFEHNDESFNTGVTQNGTSSSCKRRSH
NSFFSRLNWLTHLKFYPAINTMPNNEKFDKLYIGWVHHPVTDNDIQILFYAAASGRTVSTKRSQQTVIPNIGSRPRIR
NIPSRIISYWTIVKPGDILLNSTGHLLAFRGYFKIRSGKSSIMRSRAPIGKCHSECITPHGSIIPNDKPFQNVNRITYGA
CFRYVKQNTKLATGMHRVPEKQTRGIFGAIAEGFIENGWEGMDGNYGFHRQNFLSEIGQAGDLKSTQAATINIQINGKLNRL
IGKTENEKPHQIKIEKEFSEVEGRIQDEKLEYVETDKIDLWSYNAELLVALQENQHTIDLDSEM9KLFERTKQCLRENAEDMGN
GCFKIYHKCDNACIGSIRNGTVDHDYVYRDEALINRPQIKGVRLVFPGSYPGYIFPEAPROQAYVRKDQWEVLSTFLHH
HHHHH
Figure 58

(SEQ ID NO.83)

>HA-Vietnam-04-soluble.pro
MEKIVLLFAIVSLVK3DQICIGYHANNTEQVDIEMKIVVTVAQDILEEKHKNGKLCDDLGVKFLILRDCSVAGWLLGN
FMCDEFINVPEWSVIVEKANFVNDCYFGDFNDYEELKLHLGHRINHEKIQIIFKSSSSHEASLGVSSACPYQKGSFF
RNVWLIKKSHTYPIKRSYNTNDQEDLVLWGIHHFNDAAEQTKLYQIPTYTISVGSTLNLQRLVPRIAATRSKVNGQSG
RMEFFWTILKPDAINIFSINEHFIAPYEAYIKIVKKGDSITMKSELEYGNCNTKQTFMIAGINSFHFHHIPLTIGCPK
YVKSHRLVLATGLRNPQRERRKRRKRGFLGAIAGFIEGGWQGMVDGWYGYHSHNEQGGYAADKESTQKAIIDGVTNKVNS
IIDKMTQFEAVGREFINLERRIELHKKMDGFVLVWNYANELLVMEHERTDLFDHDSVWKLYDKVRLQLRDIAKELG
NGCFEFYHKCDNEMSVNGTYPQYSEARLKRKEISGVRSLVPRGSYGIPAEAPRGQAAYVRKDG EWVLLSTFL
HHHHHH
>HA-Shanghai-13-soluble.pro

MNTQILVFALIAIITNADKICLGHHAVSNGTKVNTLEQGVEVVHATETVERTNIPRCKGKRTVDLGQCGLGTITG
FPQCDQFLFSDADLIEEREGSDVCYPGKFVNEALRQILRESGIDKEAMGFTYSGRTNGATSACRRSGSSFYAEMKW
LLSNTDIAAFFQMTKSYYKTRKGFALIVWGHHVSTABQTKLYSGNHKLVTVGSNYQSFVPSFQXPAQFVYNGLSGRID
FHWLMLNPNDTFTSFNGAFAIPDRASFLRGRKSMQISGSGVQVDANCEGDCHSGCTISLHLPFQIDSRAVGTKCPYVKQ
RSLLLATGMKIVPEIFKGRGLFGAIAGFENGWELDGIQGRHONAQQGETAADDYKSTQASICQITGTKNRLIEKTNQ
QFELIDNFEVEKQIHNVINGNTRDSITEWVSNAELVAMENQHTIDLADSEMDLKLYERVKQLENAEDGTSCFEIF
HKDDDCMASIRNUTHDKRYEEAMQNRIQIDFVRLVPRGSFGGYIPEAPRDGQAYVRKDGWVLLSTFLHHHHHH
Figure 60

(SEQ ID NO.85)

>HA-Singapore-57-soluble.pro

MAIIYLILLFTAVRGDQICIGYHANNSTEKVDTILERNVTTHAKDILEKTHNGKLCKLNGIPPLELGDCSIAGWLLGNP
ECDRLLSVPEWSYMEKENPRDGLCYPSFNDYEELFLHLLSSVKKHFEKVKILPKDRWTQHTTSGRSACAVSGNPSSFRNI
MVWLTKKSHYVPVAKGSYNNTSGEQMLIIDWVHFNDETEQRTLYQINVGYVSVMGSTLKRSTFPDIAERPKEKGSLRM
EFWSWLLDMWTIHFESTGHLHAPYGFKISKRGSSGIMKTEGTECETKCQTPGAGNTTLFPHHVHFLTGECFKYV
KSEKLVLATGLNVPQIESRGLFGAIAGFIEGQGQHMVDGWGYHHSNDQSGSYAADKESTQKAFDIGITNKNOVIEKMN
TQFEAVGKEFANLILLERLEINKKMEDGDIFLDVNTYMAELLVLMHEERTLDPHDSNVMNLVDKVRMLEQRLNVKELGNCGPEF
YHKCDDECMDSGVKHGTYDFKYEEESKLHREIKGVRSLVPRGSPGSGYIFEAFRDGQAAYVRKDEGWLLSTFLLLLHHHHH
>PR8HA-403Y429Y

MKANLLVLLSALAAADADTICGYHANNSTDTVDTVLEKNVTVTHSVNLLLEDNHGKLCRLKGIAPQLQGKCNIAGW
LLGNPECDPLLVPVRSWSYIVETPNSENGICYPGDFIDYEELREQLSVSSFERFEIFPKES8WFENHNNGVTAACSH
EGKSSFYRNLLWLTEKEGSGYFKLNSYVNKKGKEVLVLWGLIHHPNSKEQQ8IQHNNAYV5VVTSSYNNRFTPRTIA
ERPKVRDQAGRMNYYWTLKPGDTIIFEANGNLIAPMYAFALSRGFGSGLTSMASONMHCN1CTQCTPLGAINSSLPY
QNIHPV'TIGECPKYVRSAKLMVTLRNTFSIQSRGLFAGAIFEGGWGWTGMIDGWYHYHQUEQGSGYAADQKSTQ
NAINGITNKNVTVIEKMYIQTAVGKEFKRMENLNNKVDYQFLDIWTVYAESLVLLENERTLDPHDSNVKNLY
EKVPSQLKNNAKEIGNGCEFELYKCDNESCVRNGTYDPYKESKLNREKVDGBKLESMIYQILAIYSTVASS
LVLLVSLGAISFWMCN5GLQ5CRIC1
>PR8HA-403Y432Y

MKANLLVLLSALAAADADTICIGYHANNSTDTVDTVLEKNVTVTHSVNLLEDHSHNGKLCRLKGIAPLQLGKCNIAGW
LGNPECDCPLLPPVRSWSYIVETPNASENGICYPGDFIDYEELREQLSSVSSFERFEIPKESSWPNHNTNGVTAACSH
EGKSSFYRNLPLLWTELKEGSYPKLN5VNYVKKGKEVLVWGHPPNSIKEQQINIYQnenAYYSVVTSSNYNRRFTPMT
ERPKVRDQAGRMNYYYWTLKPGDIIIFEAENGNLIIAPMYAFALSRGFG5GIITSNASMHCNTKQPQLGAINSSLPY
QNIHVPITGECPKYVRSAKLVMVTGLRTPSIQ5SRGLFGAAIFEGGWGTGMDGWYHYHQNEQGSGYAADQKSTQ
NAINGITNKVNTVIEKMYIQFTAVGKEFNKLEKRMLNKKVDDGFYDIWTYNAELLVLENETLDFHDSNVKNLY
EKVKSQILKNNAKEIGNCCEFJFYHCDNECHESVRNGTDYPKSEESKLNREKVDGVKLESMGYQILAYSTVASS
LVLVSLGAILSFWMCSNGSLQCRICI
>PR8HA-403Y433Y.pro

MKAANLVLSSLAAADADTICIGYHANNSTDTVDVTLEKNNVTTHSVNLLEDSHNGKLRLKGIAPLQLGKCNIAGW
LLOGPCEPDPILPERSWYSYIVETPNSENGICYPGDIDYEELREQLSSVSSFERFEIFPKESSWPHNHTNGVTAAACSH
EGKSSFYRNLWLTEKEGSYPKLKNYSYVNNKGEVLLVLWGIHPFNPSEQNIIYQ NCAAAY VS VS VTSV SY NRRFTEIA
ERPKVRDQAGRMNNYYWTLKPGDITTFEBANGNLIAPMYAFALSQIFGSIGAITSNASMHECNTKCTQPLGAINSSLPY
QNIHPVITIGECPKYVRSAKLRMVTGLRNTPISQSRGLFGAIAGFIEGWTGMDGWYGYHQQNEQGSGYADQKSTQ
NAINGITNKVNTVIEKMYIQFTAVGKEFNIKLRMENLNKVVDDGFILWTYNAELLVLLENETLDPHDSNVKNLY
EKVKSQLKNNAKEIGNGCQERYHKCDNCEMSVRNGTYDPKYSEESKLREKVGDVKLESMGYIQILAIYSTVASS
LVLLVSLGAIISPWMCNAGSLQCRICL
>PR8HA-433Y435Y.pro

MKANLLVLLSALAAADADTICIGYHANNSDTDTVTVLEKNTVTVSHSVNLEDSDHNGKLCRLKGIAPLQLGKCIAGW
LLGNPECDPLLPPRSWSYIVETPNSENGICYPGDYIDYEELREQLSSVSSFERFEIFPKESSWNHNITNGVTAACS
EGKSSFYRNLLWLTEKEGYSYPKLKNYVNNKDGKVLVILWGGIHHPNSKEQQNIYQENAYVSVVTSYNNRFTPEIA
ERPQVRDQAGRMNYYWTLLKPGDTTTIEANGNLIAPMYAFALSRRGFGSGITSNASMHECNTKCQTPLGAINESSLPY
QNIHPVTIGECPKYVRSAKLRMTVGLRTPSIQSRGLFIAIGFIEGGWTGMIDGWYGYHQQNEQGSGYAADQKSTQ
NAINGITNKVNTVIEKMNQFTAVGKEFKRMLNKKVDDGFLYTYTYNAELLVLENERTLDPHDSNVKNY
EKVSQLKNNAKIEIGNCCEFHYHKCDNECMESVRNTYDYPKYSEESKLNREKVDGVKLESMSQYILAIYSTVASS
LVLLVSLG AISPAGMA PWSMCSNSLQCRICI
80 / 107

Figure 66

(SEQ ID NO.91)

>PR8HA-435Y.pro

MKANLLVLLSALAAADADTICICYHANNSTDTVDVTLEKNTVTSHSVNLLEDSHNGKLCRLKGAIPQLQGKCIAGW
LLGNECDPLLPVRSWSYIVETPNSENGICYPDFIDYEELREQLSSVSFERFEIFEIPKESSWPNHINTGVTACSH
EGKSSFYRNLLLWTLEKGSYPKLNKSVNKKGKEVLVLWGIHHPFNSKEEQQNIQNYNENAYVSVTNSYNNRFTPTEIA
ERPKVRDQAGRBNYTYWLKFGDDTIEEANGNLIAPMYAFALSRGFGSGLITTTNAMHCSNHTKCTQFLGAINSSLFY
QNIHPVTIGECPKYVRSAKLRMTGLRNPISQRSGLFGAIAGFIEGGWTGMDGYYHYHLQEQGSGYAADQKSTQ
NAINGITNKVNTVEKMNQFTAVGKRFKLEKRMENLNNKVDGFLHITYYNANELLVLENERTILDPHDSNVKNLY
EKVKSQKNNAEIGNGCDFEFYHKCDNCEMESVRNGTDYFKYSEESKLNYKVDGVKLESMGIYQILAIYSTVASS
LVLLVSLGAIISPWMCSSNLQCRICL
>PR8HA-406Y433Y

MKANLVLSSLADAADACIYHANNSTDTVTVLEKNTVTHSVNLLEDSHNGKLRCRLKGIAPLQLGKCNIAGW
LLGNPECDDLLPQVRSWSYIVETPNSENGICYPGDFIDYEELREQLSSVSSFERFEIFPKESSWPNHNTNGVTACSH
EGKSSFYRNLLWLTKEGYSYPKLNSYVNKKGKEVLVGLWGIHHPNSKEEQQNYQENAYYSVVTSNYNRFTPEIA
ERPQDQAGRMISSYWTTKKPGDIIFEANGNIAMPSYGFRGFSGIRITSNASMHECTKQTPILGAINSSLPY
QNIHPVTIGECPYVSAKLRMVTLRNTPSIQSRGLFGAIAGFIEGWTGMIDGWYGYHHNEQGSGYAADQKSTQ
NAINGITNKVNTVIEMNIIYTAVGKEFKNKLEKRMENLNVKDDGFYIWTYNAELLVLENERTLDHDSNVKNLY
EKVSKQLKNACEIGNCFEFYHKCDNCHESVRRGTYDFPYSEESKLREKVDGVKLESMGICYQILAYSTVASS
LVLLVSLGAIFWMCSNGSLQCRICI
>PR8HA-411Y422Y.pro

MKANLLVLLSALAAADATCIGYHANNSTDVTVDVTEKNTVTHSVNLLEDSHNGKLCRLKGIAPLQLGKCNIAGWL
LGNPEDPLLPVRSWSYIVETPNSENGICYPDFIDYEELREQLSSVSSFERFEIFPKESWWPNHNTNGVTAACSH
EGKSSFYRNLLWLTEKEGSYFKLKNYSYNKKGKEVLVLWGIHFPNSKEQQNIYQENAYVSVVTSYNRRFTPEIA
ERPKVRODQAGRMNYYWTLLKPGDTIIEANGNLIAPMYAFALSRGFGSGITSNASMHECNTKCTPLGAINSSSLPY
QNIHFVTIGECPKYVRSAKLRMTGLRTPSIQSRGLFGAIAGFIEGGWTGMIDGWWYGHQNEQGSGYAADQKSTQ
NAINGITNKNVTVEKNIQFTAVGVEENKLERMEYLNKKVDGDGLDIWNYAELLVLENERTLDPHDSNKNLY
EKVKSQLKNNAKEIGNCFEFYHKCDNCEMESVRNGTYDPKYSEESKLNEKVDGBKLESQGYIQLAIYSTVASS
LVLLVSLGASFWMCNSGSQCRICI
>PR8HA-63G278S-N.pro

MKANLLVLLSALAAADTICIGYHANNSTDVDVTLEKNTVTTHSNLLEDSHNGKLCRLKGENLYFQ (SEQ ID NO:94)

>PR8HA-63G278S-C-120Y150Y.pro

SNASMHECNTKQTPLGAINSSLPYQNIHPTIECPKYVRSAKLRMVTLGRLNTPISQSRGLFGAILAGFIEGWTGMDGWYGYHAYQNEQSGYAADQKTQNAINGTNKVINTVIEKMYIQFATAVGKEFNLKEMLENLKVKDDGFLYIWTYNAELLVLLENERTLDHFHSNVKNLYEKVKSQKLNNAKEIGNGCFEFYHKCDNECMESVRNGTYDYPKYSHEEKLNRKVDGKLESMDIYQILAIYSTVASSLVLVLSLGAISFWMCNNGSLQCRICI (SEQ ID NO:96)
>PR8HA-63G278S-N.pro

MKANLLVLLSALAAADADTICIGYHANNSTDVTVDLVLEKNTVTVTHSVNLLEDSHNGKLCRLKGENLYFQ (SEQ ID NO: 94)

>PR8HA-63G278S-C-128Y139Y.pro

SNASMHECNTKCQTPLGAINSLLPYQNIHPVTIGECPKYVRSAKLMVTGLRNTPSILQSRGLFGAIAAGPIEGGWGTGMIDGWYGYHQQNEQGSGYADQKSTQNAINGITNKVNTVIEKMIQFTAVGYEFNKLEKRMELNKKVDDGFLDIWTVNAAELLVLLENERTLDFHSNVKLYEKVKSSLKNAKEIGNGCFEPYHKCDNECMESVRNGTYDPKYSEESKLNREKVGDVQLESMGYQILAIYSTVASSVLLVSLGAISFWMCNQGLQCRIC (SEQ ID NO: 97)
MKANLLVLLSALAAADADTICIGYHNASTDTOVDVLEKNTVTHSVNLLEDSHNGKLCRLKGENLYFQ (SEQ ID NO: 94)

SNASMHECNTKCQTPSLGAINSSLPYQNHIHPVTIGECPKYVRSAKLITGLRNTPSIQSRGLFGAIAAGPIEGWGTGMTIDGWYGHQNEQSGSYAADQKSTQANAINGITNKVTIEKMYIQFTAQVYGCFNLEKRMELNKKVDGFLYIWTYNAELLVILERNERTLDPHDSNVALYKVEKSKQanineKEIGNGCFEPYHKDNECMESVRNGYTDYPKYSEESKLNREEKVDGVKLESMDYQILAIYSTVASSLLVVLSGAISFWMCNSGLQCRICI (SEQ ID NO: 98)
MKAQLLLVLLSALAAADADTCIGYHANSTDTVDVLEKNVTVENTSVPNLEDNNGKLCRLKGENLTYFQ (SEQ ID NO: 94)

>PR8HA-63G282S-C-122Y152Y.pro

STSNASHNHECNTKQTPLGAINSSLPYQNHVTIGECFKYVRSALMRMTGLRNTPSIQSRGLFGAIAGFIEGGWT
GMDGWYHYHQQEQGSGYAADQKSTQAINGITNKVNTVIERMYIQFTAVGKEFNKLEKMNQKNNVDDGFLYW
TYNAELLVLENERTLDFHSDNVKNLYEKVSQLKNNAKEIGNCGCEFYHKCDNEMESVRNTYDYPKIESKLN
REKVDVKLESMSGIYQILAISTVASSLVLIVSLGAIJFWMCSNGSLQCRICI (SEQ ID NO: 99)
>PR8HA-63G282S-N.pro
MKANLLVLLSALAAADATCIGYHANNSTDRTVDTVLEKNTVTTHSINLLEDSHNGKLCRLKGENLTFQ (SEQ ID NO: 94)

>PR8HA-63G282S-C-130Y141Y.pro
STSNASMGHECNTKQTPGLAINSSLPYQNIHPVTIGECFVYRASKLRMVTGLRNTPSIQSGFLGAIAGFIEGGW
GMIDGWGYHHQNEQSGYAADQKSTQNAINGITNKNTVIEKMNQFTAQYGFEKLEXMEILNKKVDGFLD
TYNAELLVLENERTLDHDSVKNLYEKVSKQLNNAKEIGNGCFFYHKCDNCEMEVRNGTYDFYK'REESKLN
REKVDGVKLESMGYQLAIYSTVASSLVLIVSVSLGAISFWMCNLSLQCRICI (SEQ ID NO: 100)
>PR8HA-63G282S-N.pro

MKANLLVLLSALAAADACTICIGYHANNSTDVTVDTEKNTVTTHSVNLLLEDSHNGKLCRLKGENTLYFQ (SEQ ID NO: 94)

>PR8HA-63G282S-C-122Y130Y141Y152Y.pro

STSNASMHECNKTCQTPLGAINSSLPYQNIHPVTIGECPKYVRSAKLRMTGLRNTPSIQSRGLFGAIAGFIEGWT
GMDGWYGYYHQQNEQSGYADQKSTQNAINGITNKVNTVIERMYIQFTAUGYEFNKLEKRMEYLNNKKVDGFLITE
TYNAELLVLLENERTLFHDSNVKLYEKVSKQLKNAKEIGNGCFEFYHKCDNECMESVRNTYDYPKYESKLN
REKVDGVKLESMGYQILAIYSTVASSLVLVSLGAISFWCMSNGSLQCRICI (SEQ ID NO: 101)
>PR8HA-63G283G-N.pro

MKANLLVLLSALAAADADTICIGYHANNSTDVTDVLEKNTVTHSVNLLEDSHGKLKLCRLKGENLYFQ (SEQ ID NO: 94)

>PR8HA-63G283G-C-121Y151Y.pro

GSNASMHECNCTQCTPLGAINSSLPYQNIIIHPVTIGECPKYVRSARLMVTGLRNTPSIQSRGLFGAIAGFIEGGWTG
MIDGWWYGYYHNOEQGSGYAADQSTQNAINGITNKVNTVEKNYIQTAVGKEFNNKLEKRMMENLNIKKLDGGFLYIWT
YNAELLVLLENERTLDHDSNVKRYEVKSQLKNKEIGNGCFFEFYHKCDNECMSVRNGTYDYPKSEESKLNR
EKVDGVKLESMGIYQILAIYSTVASSLLVLLVSLGAISFWMCSNGSLQCRCI (SEQ ID NO: 102)
>PR8HA-63G283G-N.pro
MKAANLVLSSALAAADADTCIGYHANNSTDTVTVLEKNVTVTSHSNNLLEDNHGKLCRLKGENLYFQ (SEQ ID NO:94)

>PR8HA-63G283G-C-129YI40Y.pro
GSNASMHECNTKCOPIPLGAINTSSLPYQNIHPVTIGECPKYVRSAKLRMVTGLRNTPSIQSRGLFGAIAFIEGGWTG
MIDGWYGYHHQNEQGSQYADKSTQNAINGINHTKVTIEKMNQFTAVGYEFNKLEKRMEYLNNKVDGFDIWT
YNAEILLVLLLENERTLDHFHDNSVKNLYEKVKSQLKNNAKEIGNGCFEFYHKCDNECMESVRNGTYPYKSEESKLNR
EKVDGVKLESMGYIQILAIYSTVASSLVLVSLGAISFWMCSNGSLQCRICI (SEQ ID NO:103)
MKANLVLSSALAADAADTICIGYHANNSTDTDVTLEKNTVTSHSVNLEDSHGKLCKRLGNYLQ (SEQ ID NO: 94)

>PR8HA-63G283G-C-121Y129Y140Y151Y.pro

GSNASMHECNTKQQPTPLGAASLQVQNYHPVTIGECPKYVRSAKLMVTGLRVPSIQSRIQLFGAIAGFIEGNGWIGMDGVQYHHQNEQGSGYADQKSTQNAINGITNKVTIEKMYIQFTAVGYEFNKLEKRMEYLNKVDGFLYIWTYNAELVLLLENERTLDHFDSNVKNLYEKVSQQLNNAKEIGNGCCEFYHKCDNCEMESVRNGTYPKYSSEKSLNR

EKVDGVKLESMSGJQILAIYSTVASSLVLVSLGAISFWMCSNGLQCRIC (SEQ ID NO: 104)
>PR8HA-48G291G-N.pro

MKANLLVLLSALAAADADTICIGYHANNSTDTDVTVLEKNTVTSHVENVLYFQ (SEQ ID NO:95)

>PR8HA-48G291G-C-113Y143Y.pro

GNTKCQPPLGAINSSLPYQNIHPVTIGECPKYVRSAKLMVTGLRNTPSIQSRGFLGAIAGFIEGGWTGMDGGYGY
HHQNEQGSGYAADQKSTQNAINGITNKVTVEKMYIQFTAVGKEFNKLEMENLNKVDGDLYIWTYNAELLVL
LENERTLDFHDSNVKLYEKSVQLKNNAKEIGNCGEFYHKCDNEMVESVNGTYDYPKYESKILNREKVDGVKL
ESMGIYQILAIYSTVASSLVLVLLVSLGAISFWMCSNGLQCRIC (SEQ ID NO:105)
MKANLLVLLSALAAADADTICIGYHANNSTDTVDTVLEKNVTVTSHVENVLYFQ (SEQ ID NO:95)

>PR8HA-48G291G-C-121Y132Y.pro

GNTKCQTPLGAISLPLYQNIHPVTIGECPKYVRSAKLRMVTLNTPSIQSRLFGAIAFEGGKTMIDGMYGY
HHQNEQGSGYADQKSTQAINGITNKVTVEKMNIOFTAVGYEFNKLEKRMELYNNKVDGDFTDWTYNAELVL
LENERTLDHDSNVKNLYEVKSQLKNNAEIGNCCEFHFHCNMECESVRNCTYDYPKYSEESKLNREKVDGVLK
ESMGYQILAIYSTVASSLVLLVSLGAISFWMCNMGSLQUCRICI (SEQ ID NO:106)
>PR8HA-48G291G-N.pro

MKANLLVLLSALAAADADTICIGYHANNSSTDVTVDVELEKNVTVTSHSVENLYFQ (SEQ ID NO:95)

>PR8HA-48G291G-C-113Y121Y132Y143Y.pro

GNTKCQTPLGAINSSLPYQNIHPVTIGECPKYVRSAKLRMTGLRNTPSIQSRGLFGAIAAGFIEGGWGTGMDGWGY
HHQNEQGSGYAADQKSTQANINGITNKVTVEKMYIQFTAvgYEFNKLEKRMEYLNKKVDDGFLXIWTYNANELLVL
LENERTLDHFHDSNVKLYEKLQKNAKEIGNCCEFHFYHKCDNECMESVRNGTYDYPYSEESKLRNEKVDGVKL
ESMGIYQILAIYSTVASSLVLVLVSLGAIASFWSMCNSGLQCRICI (SEQ ID NO:107)
MKANLVLSSALSAAADADTICIGYHANNSTDTVTLEKNVTVTHSVENLYFQ (SEQ ID NO:95)

SNTKCGTQLGAINSSLPYQNIHPTVIGECPKYVRSAKLRVTGLRNPITSQPRLFGAIAGFIEGGWTGIYGDGY
HHQNEQGSGYAADQKSTQNAINGITNKVTIEKMYIQFTAUGKEFKNKLEMNLKVKDDGFLYWTNYAEILLVL
LENERTLDHFSDNVKNLYEKVMLQKNAKEIGNGFHYHKCDNEMESVRNGTYDPKYSEESKIILREKVDGVL
ESMGYQILAIYSTVASSLVVLSLGAISFWMCNSGSLQCRICI (SEQ ID NO:108)
MKANLLVLLSALAAADADTICIGYHANNSTDVDGLVLEKNVTVTHSVENLYFQ (SEQ ID NO:95)

>PR8HA-48G291S-C-121Y132YY.pro

SNTKCIQTPLGAINSSLPYQNIHPVTIGECPKYVRSAKLRMVTLRNPSTQSRGLFGAIAGFIEGGWTGMDGWYGY
HHQNEQGSGYADQKSTQNAINGITNKNVTVIEKMINQFTAVGYEFNKLEKRMELYNKKVDGGFLDⅡWTYNAELLVL
LENERTLDFHDSNVKLYECKQLKNAAKEIGNGCEFHYHKCDNECMESVNGTYDYPKYSEESKLNREKVDGVKL
ESMGIYQLAIYSTVASSLVLVLVSLGAISFWMCNSGSLQCRICI (SEQ ID NO:109)
>PR8HA-48G291S-N.pro

MKANLLVLLSALAAADADTIICYHYANNSTDTVTDTVLEKNVTVTHSVNLYFQ (SEQ ID NO:95)

>PR8HA-48G291S-C-113Y121Y132Y143Y.pro

SNTKCQTPLGAINESSLPYQNIHPVTIGECPKYVRSAKLRMVTGLNTPSIQRGFLGAIAGFIEGGWTGMIDGWYG
HHQNEGSGYADQKSTQAINGITNKVNTVEKMYIOFTAVGYESFNKLERMKYLNKVDGDFLKYWTNAELLLVL
LENERTDFHDDSNVKNYEVKSQLKNNAKEIGNCFEFYHKCDNECMESVRNGTYDPKYSEESKLNRKVVDGVKL
ESMGYQILAIYSTVASSLVLLVSLGAISSFWMCSNGSLQCRICI (SEQ ID NO:110)
Figure 85 A

N403Y   K411Y-N422Y
-DT  +DT  -DT  +DT

N403Y-D433Y   D433Y-W435Y
-DT  +DT  -DT  +DT

Trimer

Monomer
Figure 85 B-C

B

DT-specific Fluorescence

405nm fluorescence

WT  403Y-433Y  433Y-435Y

No DT  DT

C

Stalk bnAb binding

mAb-8D4 OD_{450}

WT (-dt)  WT (+dt)  403Y+433Y (-dt)  403Y+433Y (+dt)
Figure 86 C

48G - + 278S - +

63G 278S + - - + TEV

[Image of gel electrophoresis results]
Ab Binding of Combined Insertions

- AvVH1-69-1
- 18A3

WT  63+27863+286

V_{H1-69} bnAb binding
>PR8HA-C-terminalfragment-core1.pro

SNTKCQTLGAINSSLPYQNIHPVTIGECPKYVRSAKLGMVTGRLNTPS1Q5RGLFAGAIGFIEGGGWIGMDGWYG
HHQNEQSGSAADQKSTQNAINGITNKVTIEKMYIQFTAVGKEFKNLKEKRMENLKNNKVDGFLYIWYTYNAELLVL
LENERTLDHFSDNVKNLYEIKVQLNNAEIGNGCFEFYHKCDNECMESVRNGTYDYPKYSEESKLNRKVDGVKL
ESMGIQYQILAIYSTVASSLVLLVSLGAISFWMSCNSGSLQCRICI (SEQ ID NO:108)

>PR8HA-C-terminalfragment-core2.pro

SNTKCQTLGAINSSLPYQNIHPVTIGECPKYVRSAKLGMVTGRLNTPS1Q5RGLFAGAIGFIEGGGWIGMDGWYG
HHQNEQSGSAADQKSTQNAINGITNKVTIEKMNQFTAVGYE FNKLEKRMELYLNKVDGFLDIWTYNAELLVL
LENERTLDHFSDNVKNLYEIKVQLNNAEIGNGCFEFYHKCDNECMESVRNGTYDYPKYSEESKLNRKVDGVKL
ESMGIQYQILAIYSTVASSLVLLVSLGAISFWMSCNSGSLQCRICI (SEQ ID NO:109)

>PR8HA-C-terminalfragment-core3.pro

SNTKCQTLGAINSSLPYQNIHPVTIGECPKYVRSAKLGMVTGRLNTPS1Q5RGLFAGAIGFIEGGGWIGMDGWYG
HHQNEQSGSAADQKSTQNAINGITNKVTIEKMYIQFTAVGKEFKNLKEKRMENLKNNKVDGFLYIWYTYNAELLVL
LENERTLDHFSDNVKNLYEIKVQLNNAEIGNGCFEFYHKCDNECMESVRNGTYDYPKYSEESKLNRKVDGVKL
ESMGIQYQILAIYSTVASSLVLLVSLGAISFWMSCNSGSLQCRICI (SEQ ID NO:110)
>PR8HA-C-terminalfragment-corel.pro
NTKCQTPLGAInSSLPYQNHPVTIGECPhKVYRSMRMTGLNTPSIQSRLFGAIAFIEGggWTgMIDGwygh
HQNEQGSGYAADQKSTQAINGITNKVNTVIEKMNQETAvGKFENKLECRMENLNKKVDDGFLNTYuYNaELLVLL
ENERTLDHFHDSNVKNLVEKVSQILKNNAKEISGNCCFEFHyHCDNECMESVRNGTYDYPYSEESKLNNREKVDGVKLE
SMGIYQILAIYISTVASSLLVVLVSLGAIISPWMCNSGSLQCRICI